Deep Learning-based GNSS-R Global Vegetation Water Content: Dataset, Estimation, and Uncertainty

Daixin Zhao, *Student Member, IEEE*, Milad Asgarimehr, Konrad Heidler, *Student Member, IEEE*, Jens Wickert, Xiao Xiang Zhu, *Fellow, IEEE*, Lichao Mou*

***Abstract*—Vegetation water content (VWC) is a crucial parameter for understanding vegetation dynamics and hydrological cycle on Earth. With rapid climate changes in recent years, monitoring VWC with high spatiotemporal coverage on a global scale is of paramount importance. Yet, traditional in situ measurements are constrained in remote and densely vegetated regions. Additionally, existing spaceborne remote sensing methods face challenges due to poor cloud penetration capabilities, soil moisture interference, and inadequate temporal resolution. Spaceborne global navigation satellite system reflectometry (GNSS-R) has demonstrated promising potential to overcome these limitations in vegetation monitoring. In this study, we propose a scheme for deep learning-based GNSS-R VWC assessment, leveraging a rapidly growing amount of GNSS-R data with an unprecedented sampling rate. We introduce a triplet dataset, which consists of measurements from the cyclone GNSS (CYGNSS), global land data assimilation system (GLDAS), and soil moisture active passive (SMAP), spanning over three years. Validation is performed using several benchmark models with the proposed dataset. Furthermore, the models' predictive uncertainty is quantified with Monte Carlo (MC) dropout technique to provide a trustworthy representation of estimations. Experimental evaluation of the models demonstrates good consistency between the estimated VWC and ground truth, with a minimum root mean square deviation (RMSD) of 1.0988 kg/m^2 and a bias of 0.002 kg/m^2 over a twelve-month test period. Moreover, a daily global VWC estimation is achieved through the proposed pipeline, filling the gaps of current products and enabling rapid measurements with enhanced temporal availability. We will make the proposed dataset publicly available.**

***Index Terms*—CYGNSS, Dataset, Deep learning, Earth observation, GNSS reflectometry (GNSS-R), Uncertainty, Vegetation water content**

∗ Corresponding author: Lichao Mou

This work is jointly supported by the Hermann von Helmholtz-Gemeinschaft Deutscher Forschungszentren e.V. project "Artificial Intelligence for GNSS Reflectometry: Novel Remote Sensing of Ocean and Atmosphere (AI4GNSSR)" [grant number: ZT-I-PF-5-091], by the German Federal Ministry of Education and Research (BMBF) in the framework of the international future AI lab "AI4EO – Artificial Intelligence for Earth Observation: Reasoning, Uncertainties, Ethics and Beyond" [grant number: 01DD20001], and by the German Federal Ministry of Economics and Technology in the framework of the "national center of excellence ML4Earth" [grant number: 50EE2201C]. We would like to thank the scientific teams associated with the CYGNSS, GLDAS, and SMAP missions at NASA and the University of Michigan. Also, the authors are grateful to Caroline Arnold, Tianqi Xiao, Anja Rösel, and Andrés Camero for their valuable discussions.

Daixin Zhao is with the Chair of Data Science in Earth Observation, Technical University of Munich, 80333 Munich, Germany, and also with the German Aerospace Center (DLR) and GFZ Helmholtz Centre for Geosciences (e-mail: daixin.zhao@tum.de).

Milad Asgarimehr is with Section 1.1 Space Geodetic Techniques, GFZ Helmholtz Centre for Geosciences, 14473 Potsdam, Germany (e-mail: milad@gfz.de).

Konrad Heidler is with the Chair of Data Science in Earth Observation, Technical University of Munich, 80333 Munich, Germany (e-mail: k.heidler@tum.de).

Jens Wickert is with Section 1.1 Space Geodetic Techniques, GFZ Helmholtz Centre for Geosciences, 14473 Potsdam, Germany, and also with the Institute of Geodesy and Geoinformation Science, Technische Unversität Berlin, 10623 Berlin, Germany (e-mail: wickert@gfz.de).

Xiao Xiang Zhu is with the Chair of Data Science in Earth Observation, Technical University of Munich, 80333 Munich, Germany, and also with the Munich Center for Machine Learning, 80333 Munich, Germany (e-mail: xiaoxiang.zhu@tum.de).

Lichao Mou is with the Chair of Data Science in Earth Observation, Technical University of Munich, 80333 Munich, Germany (e-mail: lichao.mou@tum.de).

I. Introduction

As a critical component of Earth's ecosystems, vegetation contributes to hydrological and carbon cycles, supports biodiversity conservation, and maintains ecosystem stability. It aids in mitigating the greenhouse effect and global warming while also providing essential functions such as soil stabilization and water purification. However, recent climate changes and extremes raise significant concerns regarding their adverse impacts on vegetation dynamics, resulting in heightened vulnerability of ecosystems [1]. Hence, monitoring dynamics and variations in terrestrial vegetation on a global scale is of great interest to researchers and stakeholders, as it offers valuable insights into ecological processes and informs adaptive management strategies for sustainable development.

By summarizing the total amount of liquid water contained within the foliage and stems of vegetation, vegetation water content (VWC) is one of the most important parameters for understanding vegetation properties. Due to its strong correlation with transpiration and vegetation drought conditions, VWC can be utilized to infer water stress, detect wildfires, and assess drought events [2], [3]. Considering its significance for vegetation dynamics studies, rapid and accurate estimations of VWC with extended spatial coverage and temporal availability are therefore of paramount importance.

As traditional in situ measurements are limited in remote and densely vegetated regions, numerous efforts have been made to monitor vegetation variables related to VWC from different spaceborne remote sensing platforms. With optical remote sensing, normalized difference water index (NDWI), leaf area index (LAI), and normalized difference vegetation index (NDVI) derived from Landsat and moderate resolution imaging spectroradiometer (MODIS) reflectance data are used to quantify the water content of the canopy [4], [5]. Radar backscatter observations from microwave remote sensing are further employed to estimate VWC and vegetation optical depth (VOD) variations in different regions [6], [7], [8], [9].

Moreover, LiDAR data collected from global ecosystem dynamics investigation (GEDI) and ICESat-2 have been utilized to produce vegetation variables [10], [11]. Nonetheless, these methods are often limited by poor cloud penetration capabilities, soil moisture interference, and low temporal resolution.

Emerging as a novel remote sensing technique, global navigation satellite system reflectometry (GNSS-R) opens up possibilities for estimating VWC that can address the aforementioned challenges. The retrieval of VWC from GNSS-R measurements is based on analyzing the effective L-band reflectivity of reflected GNSS signals. This reflectivity over land is influenced by surface roughness, vegetation scattering effects, and Fresnel reflection coefficients. As GNSS signals traverse vegetation layers, their energy is attenuated both during transmission to and reflection from the ground. The extent of this attenuation depends on vegetation density and its dielectric properties, which are controlled by the vegetation's moisture level. While the physical mechanisms underpinning this process are available [12], few studies have explored VWC retrieval using GNSS-R measurements, with most research focusing on vegetation detection rather than mapping vegetation moisture levels.

Over the past decade, the GNSS-R community has placed a growing interest in vegetation-related parameter retrieval [13], [14]. Several studies start from ground-based GNSS-R measurements with theoretical methods. In [15], vegetation effects on reflected GNSS signals are evaluated utilizing signal-to-noise ratio (SNR) observations. Experiments show that NDVI is inversely correlated with signal magnitude. With up-looking and down-looking vertical polarized antennas, coherently added reflected signals are used to obtain vegetation height retrievals in a field campaign [16]. Another walnut-tree field investigation calculates attenuation differences between GNSS signals received from the open sky and under vegetation [17]. VWC is then derived as a function of differential attenuation and LAI over an 11-month campaign period. Moving towards airborne GNSS-R vegetation remote sensing experiments, researchers demonstrate the usage of polarimetric measurements for plant water inversion in agricultural applications [18]. Carreno-Luengo *et al.* [19] perform empirical studies with GNSS-R receivers onboard 17 stratospheric balloons over boreal forests. Experimental results hint that reflected signal intensity is roughly independent of platform height for a high coherent integration time ($20\,\mathrm{ms}$), and the coherent scattering component is attributed to canopy and soil. In addition, theoretical simulations of the scattering coefficient, polarization scattering mechanism, and cross-polarized reflectivity of vegetation are investigated in [20], [21], and [22], respectively. These pioneering efforts have established a theoretical foundation for vegetation monitoring using GNSS-R observations.

Further advancements have been achieved with spaceborne studies, given the success in ground-based and airborne GNSS-R vegetation monitoring. Utilizing L-band signals, GNSS-R satellites can operate under all weather conditions around the clock, which is particularly beneficial in regions with dense cloud coverage. With low-mass, cost-effective, and power-efficient small satellites, GNSS-R constellations can achieve sub-daily revisit time over different locations on Earth and provide timely observations [23]. Camps *et al.* [24] first employ spaceborne GNSS-R data from TechDemoSat-1 (TDS-1) to analyze signals' sensitivity with respect to NDVI. Based on the impact of vegetation attenuation, this study reveals the potential of vegetation monitoring from space despite footprint heterogeneity. Similarly, above-ground biomass (AGB) shows a pronounced effect in GNSS-R signatures by tuning soil moisture active passive (SMAP) satellite dual-polarization radar receiver to GPS L2 frequency ($1227.6\,\mathrm{MHz}$) [25]. Subsequently, the successful launch of the cyclone GNSS (CYGNSS) constellation [26] has enabled monitoring vegetation parameters with improved spatiotemporal coverage. In the development of a novel CYGNSS-based daily soil moisture inversion method, Kim *et al.* [27] reveal that VWC, especially in densely vegetated regions, can significantly affect soil moisture estimation due to signal scattering and attenuation. Carreno-Luengo *et al.* initially explore the relationship between CYGNSS GNSS-R bistatic reflectivity and vegetation opacity [28], followed by deriving a polynomial function of trailing edge for AGB retrieval over tropical forests [29]. Furthermore, Yueh *et al.* [12] propose a semiempirical model to map the interaction between soil moisture, VWC, and surface roughness with CYGNSS reflectometry measurements.

More recently, deep learning has been garnering increasing attention in various remote sensing applications, owing to its powerful ability to learn underlying mappings between different geophysical parameters [30], [31]. With current existing constellations (e.g., CYGNSS and PRETTY (Passive REflecTomeTry and dosimetrY) [32]) and an increasing amount of future missions (e.g., HydroGNSS [33]), millions of GNSS-R observations are made available each day. Thus, there is a growing trend among researchers to utilize data-driven approaches to fully exploit the potential of Earth's surface monitoring with GNSS-R data. Several studies have leveraged this synergy for applications including wind speed estimation [34], [35], [36], soil moisture retrieval [37], [38], and sea ice detection [39], [40]. In terms of vegetation monitoring with GNSS-R measurements, a few studies have employed artificial neural networks (ANNs) to retrieve vegetation parameters with GNSS-R measurements. By combining equivalent reflectivity, SNR, incidence angle, and geolocation in a feature set, the authors utilize an ANN model to estimate AGB and tree height over five forest locations [41]. Chen *et al.* [42] further enrich the feature set of ANN models with SMAP soil moisture for AGB and canopy height estimation. To harness the full potential of deep neural networks, recent research efforts tend to either integrate full delay-Doppler maps (DDMs) to preserve valuable features or incorporate additional sources of information to enhance vegetation monitoring based on GNSS-R observables [43], [44], [45]. Recently, Asgarimehr *et al.* [46] investigate diurnal VWC cycles in the Amazon and their response to water stress. Using high-frequency observations from the CYGNSS constellation, the research reveals significant diurnal fluctuations in VWC, with higher values in the morning compared to the evening, reflecting the processes of water uptake and transpiration. As a result, GNSS-R-derived VWC is proposed as an indicator of water stress, providing

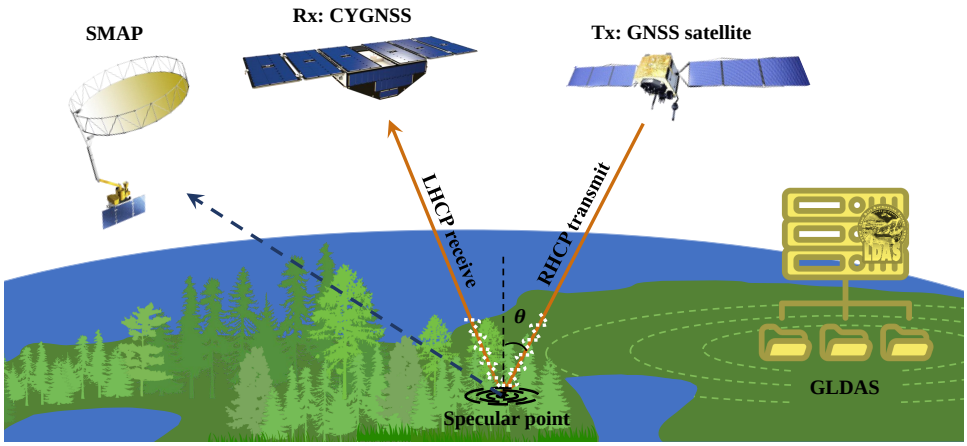


Fig. 1. Schematic of measurements incorporated to build CGS dataset: cyclone GNSS (CYGNSS), global land data assimilation system (GLDAS), and soil moisture active passive (SMAP). A GNSS satellite acts as a transmitter (Tx), transmitting L-band signals from a right-hand circular polarization (RHCP) antenna towards the land surface. Partially reflected signals-of-opportunity (left-hand circular polarization (LHCP) waves) at the specular point are perturbed by vegetation attenuation and are received by a receiver (Rx), i.e., a CYGNSS satellite. The symbol θ denotes the signal incidence angle. SMAP radiometer relies on passively sensed brightness temperatures, whereas GLDAS assimilates satellite observations with land surface models.

better insights into forest responses to increasing aridity in a changing climate. However, the timely estimation of global VWC variations based on GNSS-R measurements has yet to be thoroughly investigated. Besides, given the highly ill-posed nature of GNSS-R-based vegetation monitoring due to complex signal scattering, the absence of a comprehensive and readily accessible large-scale dataset hinders the exploration of potential interactions among different factors and parameters. While most existing studies apply GNSS-R data in combination with ancillary datasets for deep learning-based retrieval tasks, these efforts are often ad hoc and lack standardized quality control and spatiotemporal alignment.

To this end, we propose the first large-scale triplet dataset, termed CGS dataset, which integrates measurements from the CYGNSS, global land data assimilation system (GLDAS) [47], and SMAP [48] missions for deep learning-based GNSS-R VWC estimation. The first letters of the CYGNSS, GLDAS, and SMAP missions collectively constitute the "CGS" acronym. Fig. 1 depicts a schematic of the different data sources in the CGS dataset and the CYGNSS measurement principle. As demonstrated in our experiments, several deep learning models prove the feasibility of utilizing the proposed dataset to retrieve VWC globally. More specifically, the contributions of this work are threefold:

1) We build the first large-scale dataset to facilitate deep learning-based GNSS-R vegetation monitoring. With a timespan of over three years, observations from the CYGNSS, GLDAS, and SMAP missions are aggregated, filtered, and collocated with standardized quality control and spatiotemporal alignment. The CGS dataset includes variables that describe reflected signal characteristics, surface attributes, and hydrological parameters to support reproducibility and enable further analyses.
2) We benchmark the task of VWC estimation using several widely applied deep learning methods within the GNSS-R community. By employing the CGS dataset, we validate the effectiveness of deep learning-based GNSS-R VWC estimation through comprehensive studies.
3) We quantify the model uncertainty for all benchmarked methods to provide a trustworthy representation of the estimated VWC. In situations where ground truth VWC data are unavailable, uncertainty measurements provide a confidence level of the model predictions.

II. CGS: A TRIPLET DATASET FOR GNSS-R VWC ESTIMATION

To address the lack of large-scale datasets and foster algorithmic innovation, we propose the triplet CGS dataset for deep learning-based GNSS-R VWC estimation. To the best of our knowledge, this represents a pioneering effort in creating a large-scale benchmark dataset for deep learning-based GNSS-R vegetation monitoring. Fig. 2 provides an overview of the data extraction, filtering, and collocation process. The construction of the proposed dataset is detailed as follows.

A. Data sources

1) CYGNSS: Launched in late 2016, NASA's CYGNSS mission represents the first small satellite constellation dedicated to GNSS-R applications. The primary purpose of the mission is to monitor tropical cyclones by analyzing reflected GPS signals as signals-of-opportunity. Beyond its baseline science objectives, recent studies demonstrate the capability of CYGNSS for various terrestrial applications. Although the temporal revisit frequency over land is generally longer than the mean revisit time over ocean surfaces ($\sim$7 h), CYGNSS allows more frequent vegetation monitoring compared to other spaceborne remote sensing platforms with signatures over the tropics and subtropics.

As a fundamental component for the proposed CGS dataset, we choose CYGNSS L1 science data record version 3.1 from July 2019 to December 2022 [49]. Since July 2019, the CYGNSS sampling rate has been increased from once per second to twice per second, effectively reducing along-track beam smearing and enabling a theoretical maximum of 64

TABLE I
VARIABLE SELECTION OF THE PROPOSED CGS DATASET FROM DIFFERENT SOURCES: CYGNSS, GLDAS, AND SMAP. THE NAMING CONVENTIONS AND UNITS OF THE VARIABLES FOLLOW THE ORIGINAL SOURCE CONFIGURATIONS.

Sources	Variables
CYGNSS	brcs, power_analog, raw_counts, eff_scatter, sp_rx_gain, gps_eirp, ddm_snr, gps_ant_gain_db, gps_tx_power_db, tx_to_sp_range, rx_to_sp_range, sp_inc_angle, sp_lat, sp_lon, date, prn_code, spacecraft_num
GLDAS	SoilMoist_S_tavg, lat, lon, date, AvgSurfT_tavg*, CanopInt_tavg*, ECanop_tavg*, ESoil_tavg*, Evap_tavg*, TVeg_tavg*
SMAP	bulk_density, clay_fraction, radar_water_body_fraction, roughness_coefficient, latitude, longitude, date, vegetation_water_content▲, soil_moisture*, surface_temperature*, vegetation_opacity*

▲ Target variable.
* Unused variables for VWC estimation with the CGS dataset, made available for further scientific analyses.

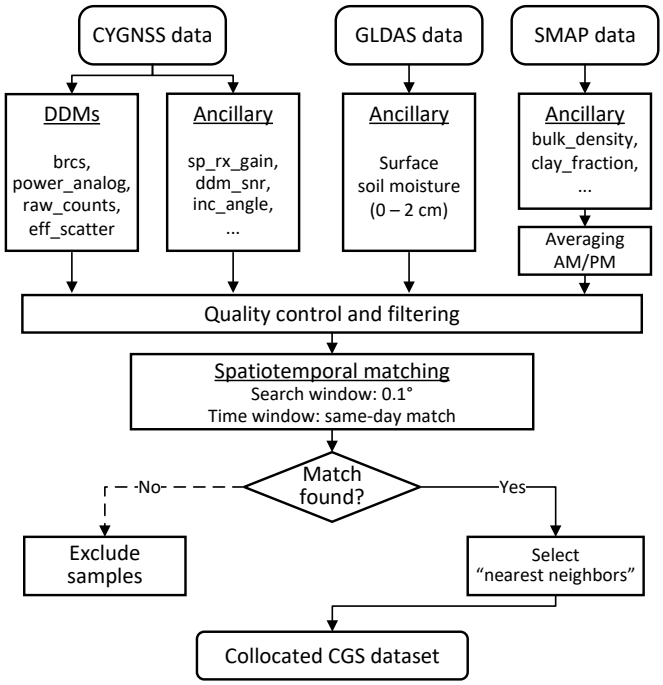


Fig. 2. Workflow for data extraction, quality control, and collocation. Delay-Doppler maps (DDMs) and ancillary parameters from CYGNSS, GLDAS, and SMAP are first extracted and filtered. Spatiotemporal matching is then performed using a $0.1°$ search window within the same date. For each CYGNSS sample, the nearest neighboring samples from GLDAS and SMAP are selected to curate the CGS dataset.

observations per second across the entire constellation [50]. Moreover, improved calibration and adjustments have been applied to version 3.1 data, e.g., normalized bistatic radar cross section (NBRCS) calibration, science antenna pattern adjustment, and coarse quantization effects correction.

DDM is one of the most important GNSS-R observables. Four types of CYGNSS DDMs (namely, DDM BRCS, the corresponding effective scattering area, analog power, and raw counts) are selected and concatenated depth-wise to create three-dimensional (3D) volumetric data for each specular point, where the abscissa and ordinate represent the Doppler and delay dimensions of DDMs, respectively. In terms of per-DDM variables that specify spacecraft configurations, geometries, and characteristics of incident and reflected signal intensities, we choose thirteen different variables available from the L1 science data record. These variables are selected based on theoretical and empirical studies to constitute an ancillary parameter set. For instance, we include variables crucial for deriving the total power (P_r) of GNSS signals scattered near the specular direction, which is determined by the sum of two terms as follows:

$$P_r = \left(\frac{\lambda}{4\pi}\right)^2 \frac{P_t G_t G_r}{(R_{ts}+R_{sr})^2}\Gamma(\theta_i) + \frac{\lambda^2}{(4\pi)^3}\frac{P_t G_t G_r}{R_{ts}^2 R_{sr}^2}\sigma, \quad (1)$$

where the former term represents coherent reflection for bare and slightly rough surfaces, and the latter denotes incoherent reflection by rough or densely vegetated surfaces [19], [51]. The presence of vegetation attenuates and scatters the GNSS signal both before it reaches the ground and after it is reflected back to the receiver [52]. Thus, the scattering and level of attenuation of the reflected signals can be used to estimate VWC. λ is the wavelength of the GPS L1 signal ($19.05\,\mathrm{cm}$); $P_t G_t$ is GPS effective isotropic radiated power (EIRP); G_r represents antenna gain toward the specular point; R denotes the distance, and the t, r, s subscripts represent the GPS transmitter, CYGNSS receiver, and the specular point, respectively. Also, the coherent reflectivity Γ and vegetation attenuation are dependent on incidence angle (θ_i) [12]. σ denotes the BRCS of rough surfaces and vegetation in m^2. Temporal and spatial information of the specular point is further included to provide seasonal and regional behaviors. Additionally, CYGNSS spacecraft number and GPS pseudorandom noise (PRN) code are considered to alleviate inter-satellite variances, GPS block type differences, and intentional changes (i.e., "flex power") based on time and location. The selected DDMs and variables enable a joint representation of signals' characteristics to foster correlation studies for VWC estimation. A detailed summary of these observations is provided in the corresponding row of Table I.

2) GLDAS: Previous studies demonstrate that the CYGNSS effective reflectivity strongly depends on surface soil moisture over land [27]. As L-band signals penetrate into the soil and reflect off the Earth's surface, the intensity of reflected signals varies depending on the soil's dielectric constant, which is a function of soil moisture. Consequently, soil moisture interference inevitably affects the retrieval of vegetation parameters using reflected signals [53]. In this study, we consider soil moisture as a primary measurement for enhancing the ability

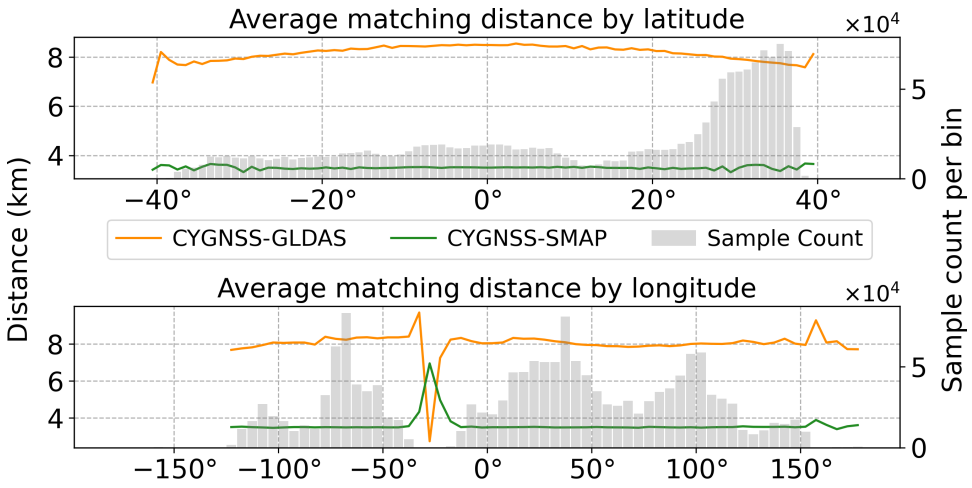


Fig. 3. Average collocation distance between CYGNSS samples and the nearest GLDAS/SMAP neighbors. Histograms show the number of samples per bin as a function of latitude (1° bins) and longitude (5° bins).

TABLE II
SUMMARY OF AVERAGE AND STANDARD DEVIATION OF COLLOCATION DISTANCES BETWEEN CYGNSS AND GLDAS/SMAP DATA. DISTANCES ARE CALCULATED AFTER SPATIOTEMPORAL MATCHING USING A 0.1° SEARCH WINDOW.

Collocation statistics	Absolute latitude range				Absolute longitude range		
	0°–10°	10°–20°	20°–30°	30°–40°	0°–60°	60°–120°	120°–180°
CYGNSS–GLDAS mean (km)	8.47	8.34	8.07	7.80	8.16	8.05	8.08
CYGNSS–GLDAS std (km)	3.17	3.11	3.03	2.95	3.08	3.04	3.06
CYGNSS–SMAP mean (km)	3.52	3.50	3.45	3.53	3.50	3.50	3.52
CYGNSS–SMAP std (km)	1.36	1.35	1.31	1.31	1.33	1.32	1.36

of deep learning models to learn the underlying mapping between the input data and VWC labels.

To achieve a higher temporal resolution of VWC estimation, daily surface soil moisture observations are extracted from GLDAS-2.2 L4 data product with a 0.25×0.25 degrees grid resolution [54]. GLDAS integrates extensive satellite and ground-based data with data assimilation methods to derive land surface models globally. The data is available from the NASA Goddard Earth Sciences Data and Information Services Center (GES DISC). For building the CGS dataset, we incorporate soil moisture in surface (0–2 cm) and several supplementary variables to facilitate further studies on vegetation moisture fluctuations and hydrological cycle dynamics, e.g., average surface skin temperature in K, transpiration and canopy water evaporation in $\mathrm{kg}/(\mathrm{m}^2\,\mathrm{s})$ (see Table I). Temporal distribution of the selected GLDAS samples aligns with the CYGNSS data timespan. Moreover, longitude and latitude data are retained to enable collocation.

3) SMAP: The SMAP mission aims to improve the understanding of Earth's hydrological and carbon cycles by tracking surface soil moisture and freeze-thaw state. It provides global coverage every 2–3 days with its primary measurements and other scientific parameters. As an essential ancillary variable for the SMAP mission, VWC is calculated by combining NDVI foliage water content estimation and annual NDVI extremes [55], given by:

$$\mathrm{VWC} = (1.9134 \times \mathrm{NDVI}^2 - 0.3215 \times \mathrm{NDVI}) + \gamma_s \times \frac{\mathrm{NDVI}_{\max} - \mathrm{NDVI}_{\min}}{1 - \mathrm{NDVI}_{\min}}, \quad (2)$$

where γ_s is the stem factor. Here, MODIS measurements are employed to derive the NDVI values and stem factors associated with different land cover types. Furthermore, since the relative amplitudes of coherent and incoherent scattering in CYGNSS reflected signals are affected by surface dielectric properties and surface roughness, we also include variables such as bulk density, clay fraction, and roughness coefficient. For bulk density and clay fraction, SMAP utilizes global soil property datasets such as harmonized world soil database (HWSD), whereas surface roughness coefficient is estimated using the tau-omega model to account for surface scattering effects [56]. These variables are extracted from SMAP enhanced L3 $9\,\mathrm{km}$ equal-area scalable earth grid (EASE-Grid) version 5 product from July 2019 to December 2022 [57]. Considering the existing multipath effects near water bodies, the fraction of water pixels within each EASE-Grid cell is considered to mitigate potential errors. It is important to note that SMAP daily measurements contain ascending and descending overpasses. For a given grid with multiple measurements, daily averaged values of selected variables are assigned to the corresponding locations. In addition, we include further variables such as surface temperature and vegetation opacity from the mission.

B. Quality control

With millions of data samples available daily, an automated quality control workflow is critical to alleviate potential transmission issues, calibration errors, and atmospheric effects. As deep learning methods are data-driven, high-quality measurements help the models to learn a precise representation of the data and enhance the ability to identify correlations among different variables. Therefore, several filtering criteria based on quality flags and empirical studies are applied to remove unreliable observations.

For CYGNSS measurements, we follow quality control steps of previous efforts for GNSS-R inland application [37], but with more rigorous conditions. Firstly, only measurements with nano star tracker attitude status as "OK" are retained. Quality flags provided by L1 products are further used to filter the remaining data, including "black-body DDM", "DDM in test pattern", "low confidence in GPS EIRP estimation", "S-band transmitter powered up", "radio frequency interference (RFI) detected", and "incorrect ddmi antenna selection" [58], [59]. Afterward, observations are filtered out with empirical thresholds. To minimize geometric effects on GNSS-R signal reflectivity, many previous studies apply an upper threshold of 60° for incidence angle. Here, we retain observations with incidence angles up to 70° to preserve more daily samples and enable the analyses of model behavior near boundary conditions. This allows for broader data coverage and supports further research on the impact of extreme geometry on retrieval quality. We remove the data with a receive antenna gain in the direction of the specular point and a direct SNR that are smaller than $0\,\mathrm{dB}$. Moreover, we discard measurements with DDM peak values falling outside a range of delay bins between 5 and 11.

In terms of GLDAS variables, missing values are removed along with their corresponding longitude, latitude, and temporal information. For SMAP observations, we apply quality flags "soil moisture retrieval was not successful" and "freeze/thaw state retrieval was not successful" to omit the corresponding data samples. Entries with missing data for the selected variables are further excluded to ensure data integrity when training deep learning models.

C. Collocation

With the variables of interest defined, spatiotemporal alignment is the next step in building the CGS dataset. For each CYGNSS sample, the corresponding GLDAS and SMAP variables need to accommodate differences in spatial resolution and timescales to select the "closest" measurement. Given that CYGNSS satellites operate within a latitude range of $\pm$ 38°, to reduce computational costs, we retain only the data within $\pm$ 40° latitude from GLDAS and SMAP prior to the collocation process. For temporal alignment, both GLDAS and SMAP employ a daily average; therefore, measurements from the same date are selected for matching. The CYGNSS data are then collocated with GLDAS and SMAP measurements using a search window of 0.1 degrees in both longitude and latitude. Concretely, for each preprocessed CYGNSS sample, the specular point coordinates are used to locate a corresponding GLDAS or SMAP grid center within the defined search window. If no matching grid center is found for either GLDAS or SMAP, the corresponding CYGNSS sample is excluded. In cases where multiple candidate grid centers meet the search criteria, the CYGNSS sample is collocated with the nearest neighboring samples from the other missions. The collocation yields an average of over two million measurements annually, resulting in a total sample size of approximately 9.9×10^6.

To quantitatively evaluate the collocation quality and its associated uncertainty, Fig. 3 shows the average collocation distance between CYGNSS samples and the nearest GLDAS/SMAP neighbors as a function of latitude and longitude. Histograms indicate the number of samples per bin, which hints at the global distribution of collocated samples. Overall, the spatial distribution of the samples aligns well with the global terrain and shows higher density within the latitude range of 20°–40°. The GLDAS collocation distance increases toward the equator due to the 0.25° grid resolution results in larger distances between grid cells at lower latitudes. In contrast, the SMAP collocation distance remains relatively uniform across latitudes and longitudes, as the SMAP product uses the EASE-Grid to preserve consistent grid spacing globally. A notable spike in longitude around 25°W is attributed to the lack of collocated data, as also indicated by the reduced sample counts in the histogram. Table II details the collocation statistics across different latitude and longitude intervals for the proposed dataset. The mean collocation distance ranges from approximately $7.80\,\mathrm{km}$ to $8.47\,\mathrm{km}$ for CYGNSS–GLDAS and remains more stable between $3.45\,\mathrm{km}$ and $3.53\,\mathrm{km}$ for CYGNSS–SMAP. Standard deviations further illustrate the variability introduced by the differences in spatial resolution and data availability during the collocation process. A more detailed discussion of data uncertainty is provided in Section VI.

III. Methodology

Following advances in applying deep learning methods to various GNSS-R applications, we use the proposed CGS dataset to train multiple models for estimating global VWC. Several benchmark methods are selected based on their demonstrated success in GNSS-R-based retrieval tasks. With the assumption that deep learning models can effectively explore the underlying mapping from input variables to VWC labels, this section details the preprocessing steps of the CGS dataset and introduces the benchmark models. Furthermore, we thoroughly describe the method employed for uncertainty quantification. An overview of the data preparation, benchmark training, and assessment in this study is depicted in Fig. 4.

A. Data preparation

Before training various benchmark models, several preprocessing steps are required to transform the variables from the CGS dataset. To evaluate models' generality and robustness with future samples beyond the training period, we temporally

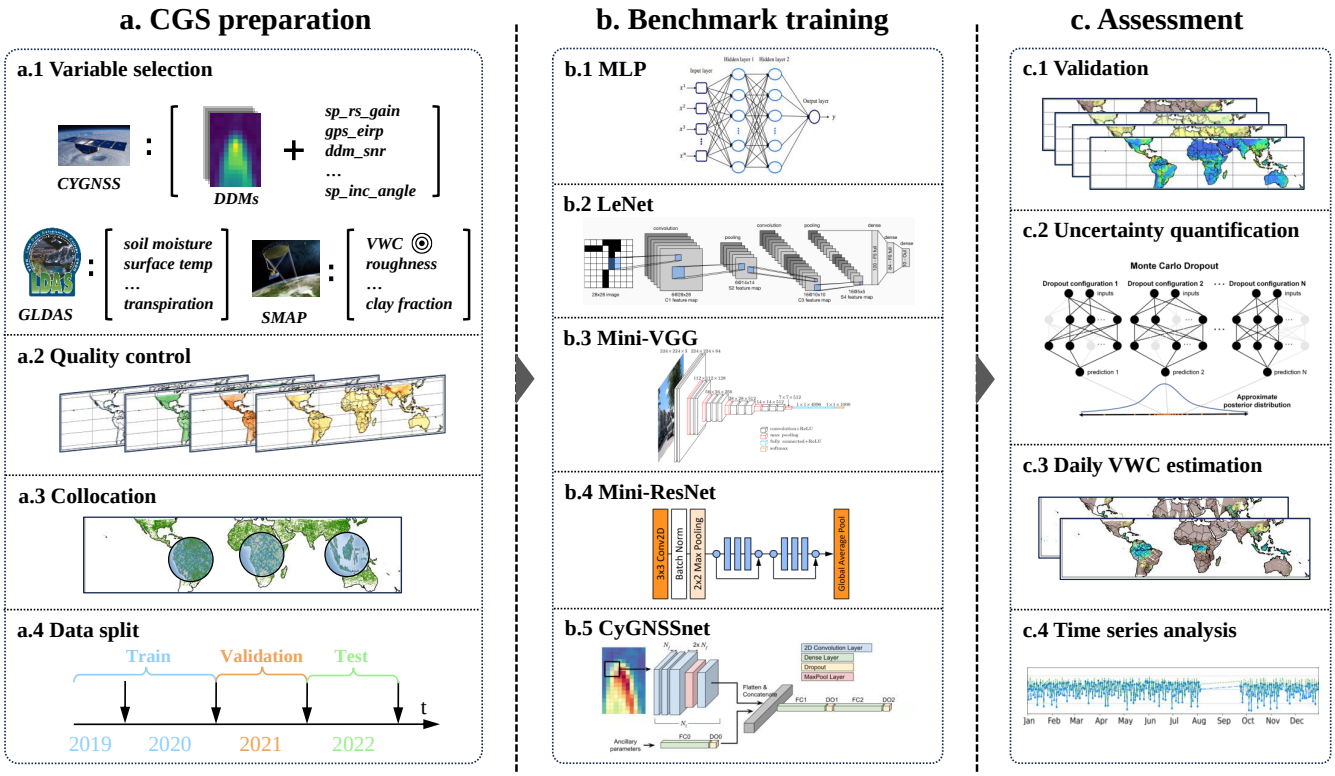


Fig. 4. The proposed scheme for deep learning-based GNSS-R vegetation water content (VWC) assessment. See the text for acronyms and descriptions for each process step.

clustered the data into training (July 2019 to December 2020), validation (in 2021), and test sets (in 2022).

With a wide range of numerical values in selected variables, normalization is a critical step to ensure that all input features contribute equally to the models' predictions. Scaling the input DDMs and other variables to have zero mean and unit variance facilitates model convergence and stabilizes gradients across different features. We, therefore, follow the best practices of normalization to prevent information leakage during the training process. Initially, each variable's mean and standard deviation are calculated for the training set and applied separately. These calculated values are then used to normalize the validation and test sets.

B. Benchmark algorithms

1) MLP: A multilayer perceptron (MLP) is a fundamental type of neural network. The main architecture of an MLP typically consists of at least three layers of neurons: an input layer, one to several hidden layers, and an output layer. With a nonlinear activation function at each neuron, MLP can approximate continuous functions and solve problems that are not linearly separable. The use of MLPs has been demonstrated to be effective for retrieving Earth's surface parameters from GNSS-R data, with applications including the estimation of ocean wind speed and soil moisture [60], [61]. In our approach, CYGNSS DDMs are flattened into 1D input vectors. These vectors are then concatenated with ancillary parameters as a single vector to create an input set for the VWC regression task.

2) LeNet: LeNet represents a pioneering convolutional neural network (CNN) [62]. Despite its simplicity by recent neural network standards, LeNet is notable for using convolutional layers to extract spatial context from image inputs. Its architecture typically consists of convolutional layers and pooling layers, followed by fully connected layers. The pooling layers help to reduce both the spatial dimensions of the input images and the intermediate feature maps, which can reduce computational demands and mitigate overfitting. In adapting LeNet for VWC estimation using the CGS dataset, differing from the MLP setup, DDMs are processed as images to preserve fine-grained delay-Doppler correlations. After processing through the convolutional branch, the output features are flattened and merged with those from an ancillary data branch. Additionally, we replace LeNet's softmax classification layer with a regression head to enable continuous VWC prediction.

3) Mini-VGG: VGG network is a deep CNN known for its simplicity and depth [63]. Based on an analysis of how to increase the depth of CNN architectures, VGG emphasizes the use of small convolutional kernels of size 3×3 to construct a deep network structure with up to 19 layers. The Mini-VGG used in this study is a simplified variant of the VGG model, which makes it more suitable for extracting features from DDMs with a pixel dimension of 17×11. While retaining the core design principles of repeated convolutional blocks followed by max-pooling layers, Mini-VGG includes two convolutional blocks with 64 and 128 filters to reduce model complexity and computational cost. Dropout [64] is applied after each hidden layer during both training and inference.

Similar to adaptations made for LeNet, an additional branch for integrating ancillary parameters is attached.

4) Mini-ResNet: As one of the most widely applied CNNs until now, residual network (ResNet) enables the training of deeper neural networks while mitigating degradation problem [65]. The main idea of ResNet is that every additional layer should more easily contain the identity function as one of its elements, enabled by the use of identity shortcut connections. The hypothesis that fits a residual mapping is easier than the desired underlying mapping with stacked layers, which helps address the vanishing gradient problem. Mini-ResNet refers to a scaled-down version of ResNet architecture with two residual blocks. Compared to deeper ResNet variants (e.g., ResNet-50), the number of blocks and filters is reduced (128 and 256 filters, respectively) to limit model size and avoid overfitting on input DDMs. Previous studies have demonstrated the effectiveness of residual learning in GNSS-R remote sensing, such as ocean wind speed retrieval [66], sea surface height measurement [67], [68], and sea ice detection [40].

5) CyGNSSnet: Introduced by Asgarimehr *et al.* [69], CyGNSSnet represents a domain-specific deep learning model designed to derive global ocean wind speed from CYGNSS measurements. Utilizing a set of convolutional layers for feature extraction from BRCS DDMs and fully connected layers for processing ancillary data, CyGNSSnet significantly improves the accuracy of wind speed measurements compared to an operational retrieval algorithm. Given that the model design is well-suited to process DDMs and ancillary parameters, we employ the CyGNSSnet architecture with minor modifications. Specifically, we adapt the model to accept four types of DDMs and a broader set of variables as inputs while maintaining the hyperparameters of intermediate layers unchanged.

C. Uncertainty quantification

Uncertainties of GNSS-R remote sensing typically arise from measurement ambiguities due to instrumental noise, ionosphere effects, and RFI. Quantifying uncertainties in GNSS-R observations is a non-trivial aspect of assessing the measurement reliability and identifying potential ways of calibration. Previous studies have proposed several methods for evaluating uncertainties in ground-based and spaceborne GNSS-R measurements [70], [71].

In the context of deep learning-based Earth's surface parameter retrieval, model uncertainty, also known as epistemic uncertainty, is an essential factor in describing the confidence of estimations. By definition, this type of uncertainty arises from a lack of knowledge or information during the modeling process [72], [73]. Since data-driven methods are essentially learning from a limited set of GNSS-R measurements, their predictions are inherently connected with ambiguities due to potentially incorrect model assumptions. With increased model complexity and diverse data modalities, models trained on a finite set of data might exhibit overconfidence in their predictions, resulting in higher estimation errors. Furthermore, when inferring geolocations without ground truth labels, model uncertainty provides a trustworthy representation of the estimated VWC values. To account for these aspects, we quantify uncertainty in this study using Monte Carlo (MC) dropout [74].

Typical deterministic neural networks are not able to capture estimation uncertainty. For the task of VWC regression, models generate fixed estimations with a given set of input variables. To obtain predictive distributions, MC dropout allows a neural network trained with dropout layers to be interpreted as a Bayesian approximation of a probabilistic deep Gaussian process. Concretely, dropout is applied not only during model training but also at inference phase to sample an approximated posterior distribution. As shown in [75], MC dropout can be well applied to quantify the estimation uncertainty for GNSS-R-based soil moisture retrieval.

Assume an MLP model with $\mathbf{X} = \{\mathbf{x}_k\}_{k=1}^{N}$ as input variables, $\mathbf{Y} = \{\mathbf{y}_k\}_{k=1}^{N}$ as output VWC values, and N denotes the sample counts. Let $\mathbf{w} = \{\mathbf{W}_i\}_{i=1}^{L}$ represent weight matrices for L hidden layers, the predictive probability can be specified as:

$$p(\mathbf{y}|\mathbf{x}, \mathbf{X}, \mathbf{Y}) = \int p(\mathbf{y}|\mathbf{x}, \mathbf{w}) p(\mathbf{w}|\mathbf{X}, \mathbf{Y}) \mathrm{d}\mathbf{w}. \tag{3}$$

Given that the posterior distribution $p(\mathbf{w}|\mathbf{X}, \mathbf{Y})$ is intractable, a matrices distribution $q(\mathbf{w})$ can approximate the intractable posterior by randomly setting its columns to zero [76], written as:

$$\mathbf{W}_i = \mathbf{M}_i \cdot \mathrm{diag}\left([\mathbf{z}_{i,j}]_{j=1}^{K_i} \right), \tag{4}$$

$$\mathbf{z}_{i,j} \sim \mathrm{Bernoulli}(p_i) \text{ for } i = 1, \ldots, L,\ j = 1, \ldots, K_{i-1}, \tag{5}$$

in which each $\mathbf{z}_{i,j}$ is a Bernoulli random variable that determines whether the neuron j in layer $i-1$ as an input to layer i should be retained or dropped, with p_i representing the dropout rate. Here, matrices $\mathbf{M}_i$ denote variational parameters and $K_i \times K_{i-1}$ represents the dimension of weight matrix $\mathbf{W}_i$.

During the training process, input variables $\mathbf{X}$ and output VWC values $\mathbf{Y}$ are fed into the MLP model to optimize $\mathbf{w}$ with dropout activated through backpropagation. During the inference phase, in contrast to deterministic models, dropout remains activated and is equivalent to sample T times stochastic forward passes from the Bernoulli distribution. A predictive mean VWC using MC dropout is given by:

$$\mathbb{E}(\widehat{\mathbf{Y}}) \approx \frac{1}{T} \sum_{t=1}^{T} \widehat{\mathbf{Y}}(\mathbf{X}, \mathbf{W}_1^t, \mathbf{W}_2^t, \ldots, \mathbf{W}_L^t), \tag{6}$$

where $\widehat{\mathbf{Y}}$ is estimated VWC value and $\mathbf{W}_L^t$ is the network weight at t-th inference time. It is worth noting that evaluation measurements such as root mean square deviation (RMSD) are not applicable when performing inference without ground truth labels. On the contrary, MC dropout allows for the quantification of the model's estimation uncertainty with an uncertainty score. This capability provides valuable insights that are particularly relevant for time-sensitive applications, where rapid and reliable measurements are critical.

IV. Experimental setup

A. Implementation details

All the benchmark models are implemented on the Tensorflow platform [77] and trained on a single NVIDIA RTX

TABLE III
EVALUATION STATISTICS FOR DIFFERENT BENCHMARK MODELS OVER A TWELVE-MONTH TEST PERIOD OF THE CGS DATASET. THE TOTAL NUMBER OF PARAMETERS AND INFERENCE TIME OVER A THOUSAND SAMPLES ARE INDICATED FOR EACH MODEL.

Models	Params	Inf. time	RMSD ($\mathrm{kg/m^2}$)	ubRMSD ($\mathrm{kg/m^2}$)	Bias ($\mathrm{kg/m^2}$)	R^2 score	Uncertainty score
MLP	580K	3 ms	1.1075	1.1073	0.012	0.968	0.583
LeNet	390K	2 ms	1.0988	1.0987	0.002	0.969	0.434
Mini-VGG	450K	5 ms	1.1264	1.1263	−0.016	0.967	0.705
Mini-ResNet	1.2M	17 ms	1.1577	1.1575	0.015	0.966	0.689
CyGNSSnet	970K	7 ms	1.1094	1.1092	0.017	0.968	0.713

3090 GPU. We use Adam optimizer [78] with a batch size of 2048 samples and a learning rate magnitude of 10^{-3} for each model individually. The commonly chosen mean squared error (MSE) is used as the loss function. The maximum number of training epochs is set as 500 with an early-stopping patience of 6 epochs. For uncertainty quantification, we apply MC dropout with a dropout rate of 10% to all benchmark models. After training, we sample ten times for each model over the entire test period of the CGS dataset. This allows us to assess the estimation performance and confidence levels in a joint way. Moreover, we select a day outside the temporal scope of the CGS dataset to produce a daily global VWC coverage.

B. Evaluation metrics

Evaluation is performed over a test period from January 2022 to December 2022 on CGS dataset. We assess all benchmark models both qualitatively and quantitatively using RMSD and bias. Additionally, unbiased RMSD (ubRMSD) is computed by removing the mean from both the estimates and ground truth labels. The coefficient of determination (R^2 score) is also calculated across the entire set of test samples to provide further insights into model performance.

To enable uncertainty quantification in the absence of reference VWC values, we draw ten predictions with dropout enabled and compute the predictive variance as an uncertainty score for each model:

$$\text{Uncertainty score} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{t}\sum_{l=1}^{t}(\hat{v}_{il}-\bar{\hat{v}}_i)^2\right), \quad (7)$$

where $\hat{v}_{il}$ is the l-th prediction for sample i, and $\bar{\hat{v}}_i$ is the mean of the t predictions for the i-th sample, calculated as $\bar{\hat{v}}_i = \frac{1}{t}\sum_{l=1}^{t}\hat{v}_{il}$.

V. RESULTS

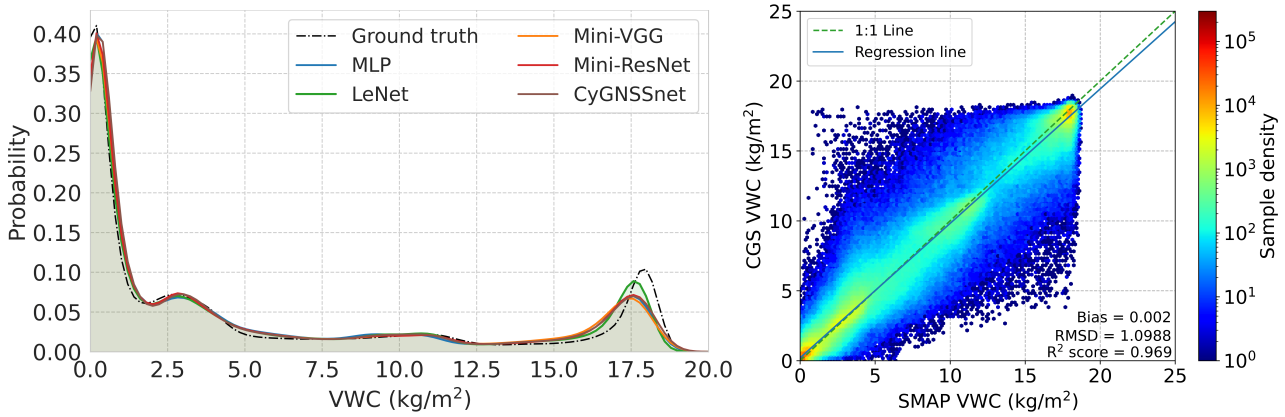
To demonstrate the feasibility of deep learning-based GNSS-R VWC estimation, in this section, we utilize the holdout test set of our curated CGS dataset to evaluate model performance. We provide geographical visualizations of the estimation results to further analyze the spatial distribution of RMSD, bias, and uncertainty scores. Lastly, a global daily VWC estimation is presented compared to ground truth measurements.

A. Benchmark results

Estimations of VWC from five selected benchmark models are compared against SMAP-derived reference labels. The performance of each model over the entire test period of 2022 is quantified with RMSD, ubRMSD, bias, R^2 score, and uncertainty score, as detailed in Table III. In general, by employing the proposed CGS dataset with selected input variables, all models demonstrate commendable performance in VWC estimation when compared to SMAP VWC products. An average RMSD of $1.1199\,\mathrm{kg/m^2}$ and an average bias of $0.006\,\mathrm{kg/m^2}$ across all models highlight the feasibility of accurate VWC retrieval through the incorporation of deep learning algorithms and the proposed dataset. Additionally, an average R^2 score of 0.968 further evidences the effectiveness of this approach for global VWC estimation, given approximately two million samples collected throughout 2022.

Among all the evaluated models, LeNet yields superior performance, evidenced by an RMSD of $1.0988\,\mathrm{kg/m^2}$, an ubRMSD of $1.0987\,\mathrm{kg/m^2}$, and a minimal bias of $0.002\,\mathrm{kg/m^2}$. Enhanced performance is further evidenced by its improved R^2 score. Notably, the uncertainty measurement associated with LeNet surpasses those of competing methods, with the MLP achieving the second lowest uncertainty score, followed by other models. Given the tendency of deep learning methods to favor the majority of data samples, sophisticated models may risk overfitting to local optima during optimization. This can lead to increased estimation errors and potentially higher uncertainty scores. Nonetheless, the comparison between MLP and LeNet emphasizes the advantages of employing the full DDM as an image input. Corroborated by prior studies [79], [80], this approach facilitates the extraction of intricate delay-Doppler correlations, whereas converting the DDM into a 1D feature vector leads to the loss of valuable cross-correlations.

The probability density functions (PDFs) of the ground truth and VWC estimates retrieved by benchmark models are depicted in Fig. 5(a), which helps to assess model behavior across different VWC intervals. As indicated by a black dashed line, reference VWC samples align with an expected natural distribution within the $\pm\ 40^\circ$ latitude range. Unevenly distributed ground truth displays two distinct peaks around the VWC intervals of 0–2 $\mathrm{kg/m^2}$ and 16–18 $\mathrm{kg/m^2}$, corresponding to arid or barren landscapes and densely vegetated rainforests, respectively. Despite the challenges raised by such a heterogeneous and unbalanced data distribution, deep learning algorithms provide satisfactory estimations that agree



(a) (b)

Fig. 5. (a) Probability density functions (PDFs) of SMAP-derived ground truth and VWC values estimated by different benchmark models. (b) Density plots in log-scale for a best-performing model (LeNet) against reference VWC. The color bar measures sample density. Bias, RMSD, and R^2 score are calculated over twelve months of test data. A 1:1 diagonal and a regression line are shown as a green dotted line and a blue line, respectively.

well with reference VWC distribution. Consistent with the numerical results, VWC values predicted by LeNet (depicted in green) exhibit a closer match to the ground truth. Despite slight deviations in peak intervals, MLP (blue), Mini-VGG (orange), Mini-ResNet (red), and CyGNSSnet (brown) also demonstrate a relatively high degree of correspondence. The general consistency in PDFs across all models indicates their ability to capture essential patterns necessary for accurate VWC estimation, enabled by the carefully curated dataset and thorough selection of variables.

Fig. 5(b) presents a log-density plot comparing VWC estimates from the best-performing model (LeNet) with reference labels. A regression line of predicted VWC with respect to SMAP-derived observations is shown in blue. Generally, scatter points are mainly clustered along the 1:1 diagonal line, suggesting overall agreement between the estimated and reference VWC values. Align with the VWC distribution shown in Fig. 5(a), high-density regions corresponding to the two peak intervals are highlighted with yellow to red color, indicating a greater concentration of paired points. Notably, several points (on the order of 10^0–10^2) deviate from the diagonal line. This deviation primarily results from the uneven distribution of VWC labels in real-world scenarios. Since deep learning algorithms are data-driven, less frequent samples in the dataset tend to be underrepresented in the predictions. Although discrepancies are observed within the VWC interval of 5–16 $\mathrm{kg/m^2}$, which accounts for less than one-fifth of the entire test set, the overall strong correlation between predicted VWC and the ground truth remains evident.

B. Geographical analysis

To improve the understanding of the global distribution of VWC estimations obtained using deep learning algorithms on the CGS dataset, we conduct a series of analyses with geographical visualizations. Fig. 6 demonstrates the average ground truth VWC alongside the estimated VWC from the best-performing model over the course of 2022. With an average grid resolution of 1 degree, the qualitative comparison of

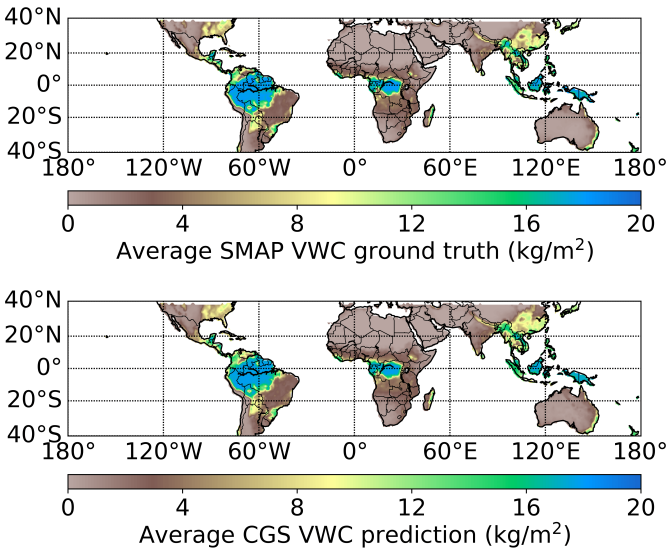


Fig. 6. Geographical distribution of average reference VWC as well as retrieved average VWC of the LeNet model on the CGS dataset during the test period.

these global VWC maps indicates that CGS predictions align well with reference values. In particular, accurate estimations are evident in equatorial regions, e.g., the Amazon region, Congo basin, and Southeast Asia, which are characterized by dense rainforests and high VWC levels. Likewise, the model demonstrates consistent performance in arid regions such as the Sahara, Arabian, and Australian deserts.

The global distributions of RMSD, bias, and uncertainty scores for CGS VWC estimation are presented in Fig. 7. Similarly, a triptych of maps is derived with a grid resolution of $1^\circ \times 1^\circ$. The topmost map reveals a clear geographical pattern where transition regions from densely to moderately vegetated areas exhibit higher RMSD values, indicated by orange to red colors (RMSD of 2–3 $\mathrm{kg/m^2}$). Additionally, regions with frequent agricultural activities, such as central China and the eastern United States, experience relatively larger fluctuations

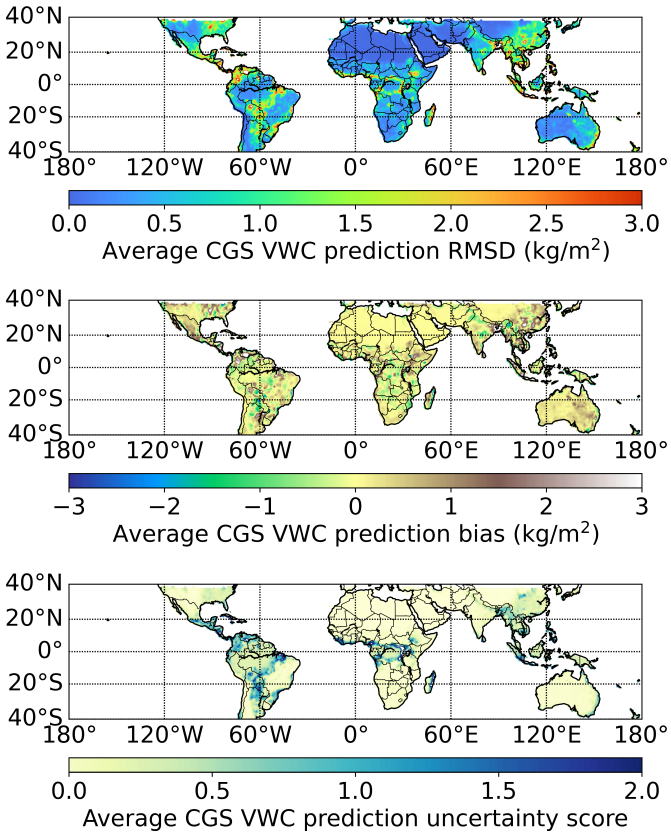


Fig. 7. Average RMSD (top), bias (middle), and uncertainty score (bottom) spatial distributions of the LeNet model compared with reference VWC with a $1^\circ \times 1^\circ$ grid resolution on the test set.

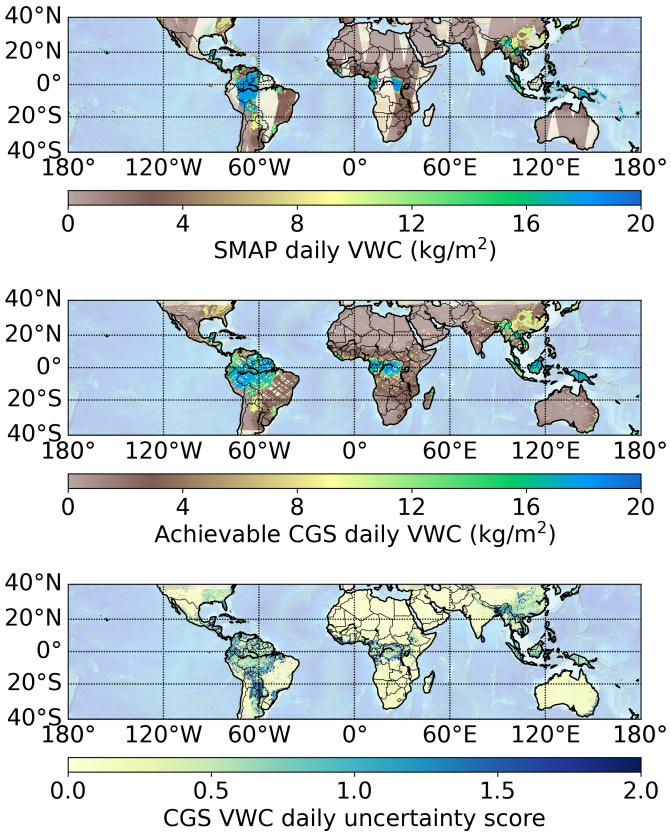


Fig. 8. Geographical distribution of daily VWC coverage for SMAP (top), CGS estimation with LeNet model (middle), and the associated daily uncertainty score for CGS estimations (bottom).

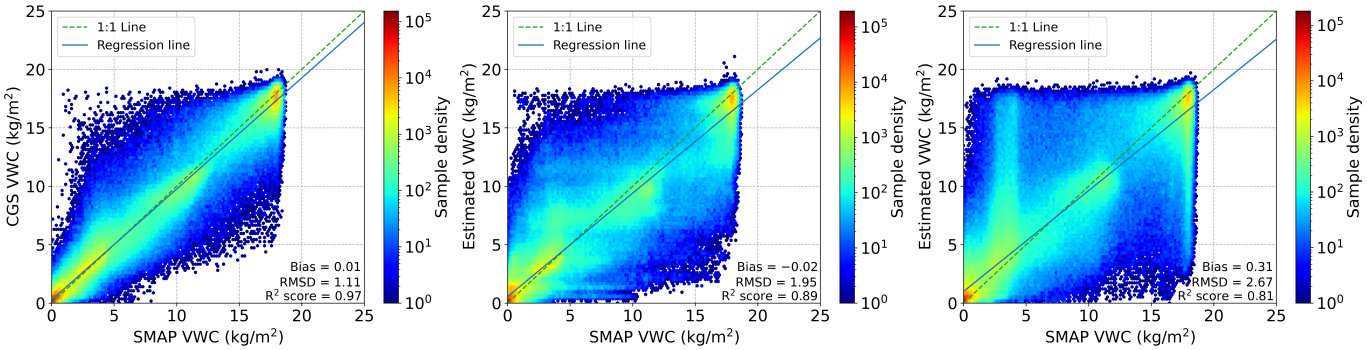
in RMSD. On the contrary, rainforests and sparsely vegetated regions at mid-to-high latitudes generally show lower RMSD values, ranging from 0–0.5 $\mathrm{kg/m^2}$, indicating more reliable predictions in these zones. Furthermore, the middle map of the Fig. 7 demonstrates that bias in VWC predictions is mostly neutral across vast terrestrial expanses. Although some patchy areas exhibit slight over- or underestimations, the overall negligible bias hints the estimated VWC is generally reliable, with only a few regions showing a bias exceeding $\pm\, 2\,\mathrm{kg/m^2}$.

To specifically evaluate predictive confidence, the bottom map in Fig. 7 depicts the global distribution of uncertainty scores associated with the predictions, where lower scores indicate higher confidence. Although this map shares similarities with the RMSD distribution, it is important to note that it reflects the model's stability across multiple inferences rather than predictive errors. Regions exhibiting higher estimation uncertainty can be attributed to several factors. One potential driver is the large footprint (i.e., glistening zone) of CYGNSS observations, which cover areas with a mixture of dense vegetation and bare or sparsely vegetated surfaces. Additionally, complex topography and surface roughness significantly influence the reflected signals, leading to increased estimation uncertainty in mountainous regions, including the Himalayas, the east coast of Australia, and eastern Madagascar. From a data-centric perspective, models trained on real-world datasets inevitably experience uneven data distributions. As a result, predictions for less frequently observed VWC intervals tend to have lower confidence.

With a substantial amount of daily measurements available from the CYGNSS constellation, a daily global VWC estimation is achieved using the proposed CGS dataset pipeline and trained model. Fig. 8 displays the geographical distribution of CGS daily global VWC coverage alongside its SMAP-derived counterpart. This approach fills the gaps in SMAP's daily overpasses, allowing rapid measurements and improving temporal resolution. Regarding spatial resolution, which is strongly influenced by CYGNSS reflection types, it can range from $3.5 \times 0.5\,\mathrm{km}$ to $25 \times 25\,\mathrm{km}$ considering signal integration time and incidence angles [51]. Still, spatial resolution can vary significantly due to surface heterogeneity, which leads to changes in the effective footprint. Since the entire DDMs are used as inputs to train models for predicting VWC, the predicted values are assumed to correspond to the spatial resolution of the label source, as the models learn to map features extracted from the inputs to the provided annotations. It should be noted that the current workflow for generating daily VWC maps still results in gaps when CYGNSS data are not available. However, given the unprecedented spatiotemporal coverage of CYGNSS and the availability of upcoming GNSS-R missions, such products can be leveraged to perform spatial interpolation and address these limitations. Furthermore, the distribution of daily uncertainty scores provides insights into the model's confidence under ambiguous conditions. As shown in the bottom panel of Fig. 8, areas with relatively higher

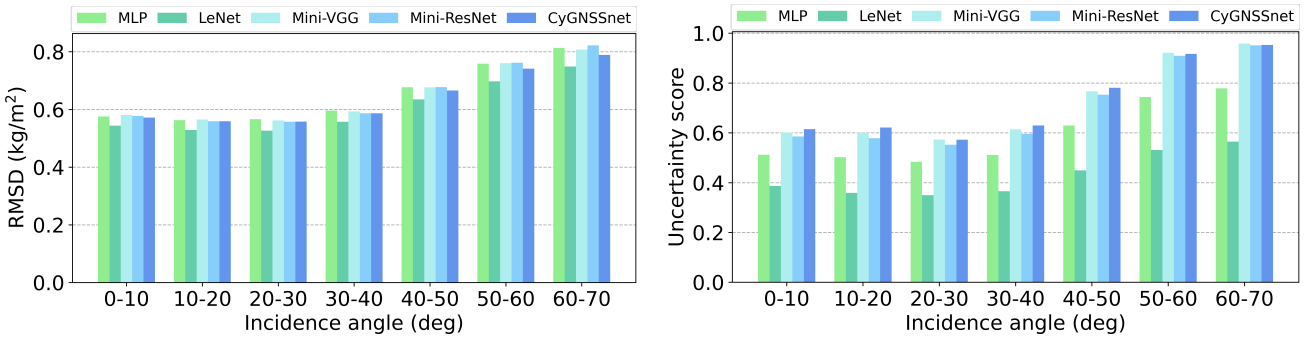
TABLE IV
EVALUATION STATISTICS OF MULTILAYER PERCEPTRONS (MLPS) TRAINED WITH DIFFERENT INPUT CONFIGURATIONS: WITH THE FULL CGS DATASET, SPATIOTEMPORAL VARIABLES ONLY, AND CGS WITHOUT SPATIOTEMPORAL VARIABLES.

Variables	RMSD ($\mathrm{kg/m^2}$)	RMSD ($\mathrm{kg/m^2}$)	Bias ($\mathrm{kg/m^2}$)	R^2 score	Uncertainty score
CGS (full)	1.1075	1.1073	0.012	0.968	0.583
Spatiotemporal only	1.9505	1.9507	−0.022	0.893	0.936
CGS w/o spatiotemporal	2.6722	2.6544	0.310	0.807	0.738



(a) (b) (c)

Fig. 9. Density plots in log-scale for models against SMAP-derived VWC values over a twelve-month test period: (a) full CGS dataset, (b) spatiotemporal variables only, and (c) CGS dataset excluding spatiotemporal variables.



(a) (b)

Fig. 10. Correlation of RMSD (a) and uncertainty score (b) with CYGNSS incidence angle for different benchmark models.

uncertainty are marked with a mosaic pattern, particularly near transitions between forests and bare lands, as well as in mountainous regions. This information is crucial for assessing the reliability of model estimations during periods when reference VWC measurements are unavailable.

VI. DISCUSSION

A. Impact of spatiotemporal variables

We investigate the potential influence of spatiotemporal features on model performance by comparing models trained on the full CGS dataset against models trained without spatiotemporal inputs and those trained on spatiotemporal inputs only. All models use MLPs to ensure a fair comparison, as the spatiotemporal-only model does not incorporate input DDMs during training. Numerical comparisons and density plots are summarized in Table IV and Fig. 9. Not surprisingly, the model trained on the full CGS dataset demonstrates clear advantages across all evaluation metrics. While the spatiotemporal-only model captures coarse distribution patterns, it is insufficient for accurate VWC retrieval on its own. The density plot in Fig. 9(b) illustrates that the spatiotemporal-only model exhibits a broader scatter, notably overestimations for VWC in the 3–5 $\mathrm{kg/m^2}$ range and underestimations above $15\,\mathrm{kg/m^2}$.

To further evaluate whether the models rely on spatiotemporal correlations rather than true causal relationships, we conduct an additional experiment excluding spatiotemporal features entirely from the input set. As shown in Fig. 9(c), the performance is acceptable at low VWC values corresponding to coherent scattering over arid land surfaces and deserts. However, the model exhibits strong oscillations and scatter in the estimation distribution as VWC values increase. One key factor contributing to the degraded estimation accuracy is topographic heterogeneity. While ancillary input features provide partial descriptions of the surface state, they cannot fully capture the variability within the glistening zone, introducing ambiguities that are difficult to resolve without spatial

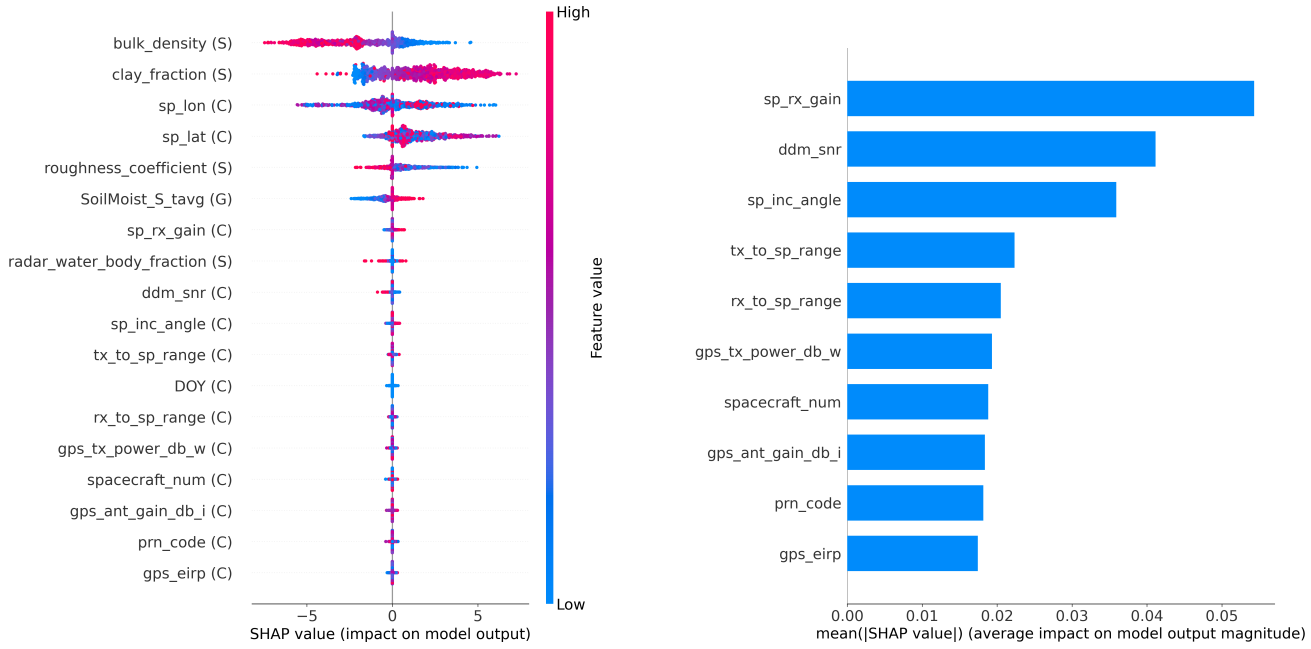


Fig. 11. SHAP (SHapley Additive exPlanations) analysis for feature importance in LeNet-based VWC estimation models. (a) SHAP summary plot for ancillary variables, with dot colors indicating original feature values. (b) Average absolute SHAP values for CYGNSS non-spatiotemporal features.

and temporal context. Furthermore, the mixture of coherent and incoherent scattering components in GNSS-R scattered signals presents a fundamental challenge. Disentangling these components and understanding their interactions in terrestrial retrieval tasks remains an open research question. This unresolved complexity introduces additional uncertainty into VWC estimation. Therefore, including spatiotemporal information in a data-driven setup is essential for resolving these ambiguities and achieving accurate, timely VWC estimation.

B. Correlation with incidence angle

Previous studies demonstrate that the incidence angle is an important variable for GNSS-R vegetation monitoring [12], [81], [82]. The correlation between the models' RMSD and uncertainty scores with the CYGNSS incidence angle is presented in Fig. 10(a) and Fig. 10(b), respectively. It can be observed that both metrics generally increase as the incidence angle rises. For samples with incidence angles within the $0°$–$40°$ interval, the average RMSD for all models remains below $0.6\,\mathrm{kg/m^2}$. Moreover, the lowest values of RMSD and uncertainty score, particularly for the LeNet estimations, are found around incidence angles of $20°$–$30°$. This finding is consistent with [22], where the peak magnitude of simulated NBRCS also falls within the same incidence angle range, while reduced magnitudes are observed at incident angles of $\sim 10°$ and $\sim 70°$. These results suggest that specular measurements with right-hand circular polarization transmit and left-hand circular polarization receive are suitable for vegetation monitoring at grazing incidence angles. Variations in the uncertainty score further reveal the impact of vegetation attenuation at high incidence angles where coherent reflection is enhanced.

C. Ablation study on feature importance

Understanding feature importance is a non-trivial aspect of data-driven approaches, especially given the highly ill-posed nature of GNSS-R vegetation monitoring. While deep learning models can effectively map both linear and non-linear dependencies in the task, interpretability is essential to gain valuable insights into the underlying physical correlations, identify potential biases, and better understand the sources of uncertainty. A systematic review by Nguyen *et al.* highlights the importance of incorporating interpretability into deep learning to provide insights on feature importance for terrestrial applications [53]. To investigate the decision-making process of deep learning models for VWC retrieval, we conduct ablation studies to asses the individual and collective contributions of two categories of model inputs: DDMs and ancillary parameters.

Table V summarizes the results of ablation experiments for the best-performing model (LeNet) when trained with either the full set of four DDMs or each DDM individually. When using the full set of four DDMs as input, the model achieves the best overall accuracy, while models trained with individual DDMs exhibit slightly degraded performance. Although certain DDM types share similar observational features (e.g., before or after calibration), the ablation results demonstrate that the integration of all DDMs provides complementary information for VWC retrieval. Notably, the model trained with all DDMs shows slightly higher uncertainty, which is expected because mapping multiple DDM types to the target variable introduces additional complexity. Variations in one DDM type can influence the entire decision-making process of the models. Nonetheless, the performance gain highlights the importance of incorporating all available DDM types in the workflow.

TABLE V
EVALUATION STATISTICS OF THE LENET MODEL TRAINED ON DIFFERENT SELECTIONS OF DDM TYPES FROM THE CGS DATASET.

DDM types	RMSD ($\mathrm{kg/m^2}$)	ubRMSD ($\mathrm{kg/m^2}$)	Bias ($\mathrm{kg/m^2}$)	R^2 score	Uncertainty score
CGS (full)	1.0988	1.0987	0.002	0.969	0.434
brcs	1.1044	1.1045	−0.002	0.968	0.427
power_analog	1.1017	1.1018	0.003	0.969	0.426
raw_counts	1.1187	1.1188	−0.002	0.968	0.433
eff_scatter	1.1035	1.1037	−0.004	0.969	0.429

TABLE VI
COMPARISON OF RMSD, UBRMSD, BIAS, AND UNCERTAINTY SCORE FOR DIFFERENT BENCHMARK MODELS ACROSS THREE CASE STUDY REGIONS OVER A TWELVE-MONTH TEST PERIOD OF THE CGS DATASET.

Regions	Models	RMSD ($\mathrm{kg/m^2}$)	ubRMSD ($\mathrm{kg/m^2}$)	Bias ($\mathrm{kg/m^2}$)	Uncertainty score
Amazon	MLP	0.9828	0.9587	−0.216	1.301
	LeNet	0.9091	0.8977	−0.143	0.621
	Mini-VGG	1.0234	0.9929	−0.248	1.831
	Mini-ResNet	1.0135	0.9987	−0.173	1.798
	CyGNSSnet	0.9940	0.9771	−0.183	1.903
Congo	MLP	1.3102	1.3086	−0.060	1.433
	LeNet	1.2337	1.2277	−0.120	0.876
	Mini-VGG	1.3717	1.3430	−0.278	1.978
	Mini-ResNet	1.5065	1.4939	−0.193	1.944
	CyGNSSnet	1.3628	1.3390	−0.253	2.011
SEA	MLP	2.0126	2.0117	−0.056	1.280
	LeNet	1.9271	1.9267	0.029	0.846
	Mini-VGG	2.0140	2.0135	−0.038	1.757
	Mini-ResNet	2.0207	2.0035	0.262	1.830
	CyGNSSnet	1.9610	1.9668	−0.127	1.848

Corroborated by previous studies [83], [84], [85], SHapley Additive exPlanations (SHAP) [86] provides a model-agnostic tool for "unboxing" deep learning models by estimating the marginal contribution of each feature to the predictions. Fig. 11(a) shows the SHAP summary plot for all ancillary parameters, with each dot representing an individual SHAP value for a feature and a sample, and the color indicating the original feature value. Features are sorted by their average impact on the model output. The letters in parentheses after each feature indicate its source from the CGS dataset. Unlike previous work on ocean wind speed retrieval, where features such as NBRCS (σ_0) and incidence angle prove to be dominant, we observe that variables related to soil conditions, surface properties, and geolocation play important roles in vegetation monitoring. This is consistent with theoretical expectations that surface roughness, topographic heterogeneity, and soil water content interference affect the propagation and attenuation of GNSS-R signals, thereby influencing the performance of VWC retrieval. The findings in Table IV reaffirm the importance of including spatiotemporal information is critical for establishing a solid baseline for estimation, while non-spatiotemporal features help capture localized or seasonal variations to generate accurate and timely predictions.

Fig. 11(b) depicts the SHAP values for all CYGNSS non-spatiotemporal features averaged over absolute contributions. Among these features, receiver antenna gain at the specular point emerges as the most influential, followed by SNR of the DDM and incidence angle. Along with the results in Fig. 10(a), these findings emphasize the strong influence of signal quality and measurement geometry on GNSS-R-based vegetation monitoring. These ablation studies demonstrate the value of integrating explainable AI techniques to interpret the complex relationships learned by deep learning models, which is crucial for understanding model behavior and guiding future improvements in model design.

D. Case studies with time series analysis

In order to further assess the reliability of the proposed approach in estimating VWC variations, we focus on several regions of interest (ROIs) for case studies. Three locations with dense vegetation coverage, namely the Amazon rainforest, Congo basin, and Southeast Asia, are selected to evaluate the estimation performance of the benchmark models. Table VII provides details on the latitude and longitude ranges of the selected regions, along with the total sample counts and daily average sample counts over the entire test period.

A detailed comparison of the performance metrics for different models across selected regions is provided in Ta-

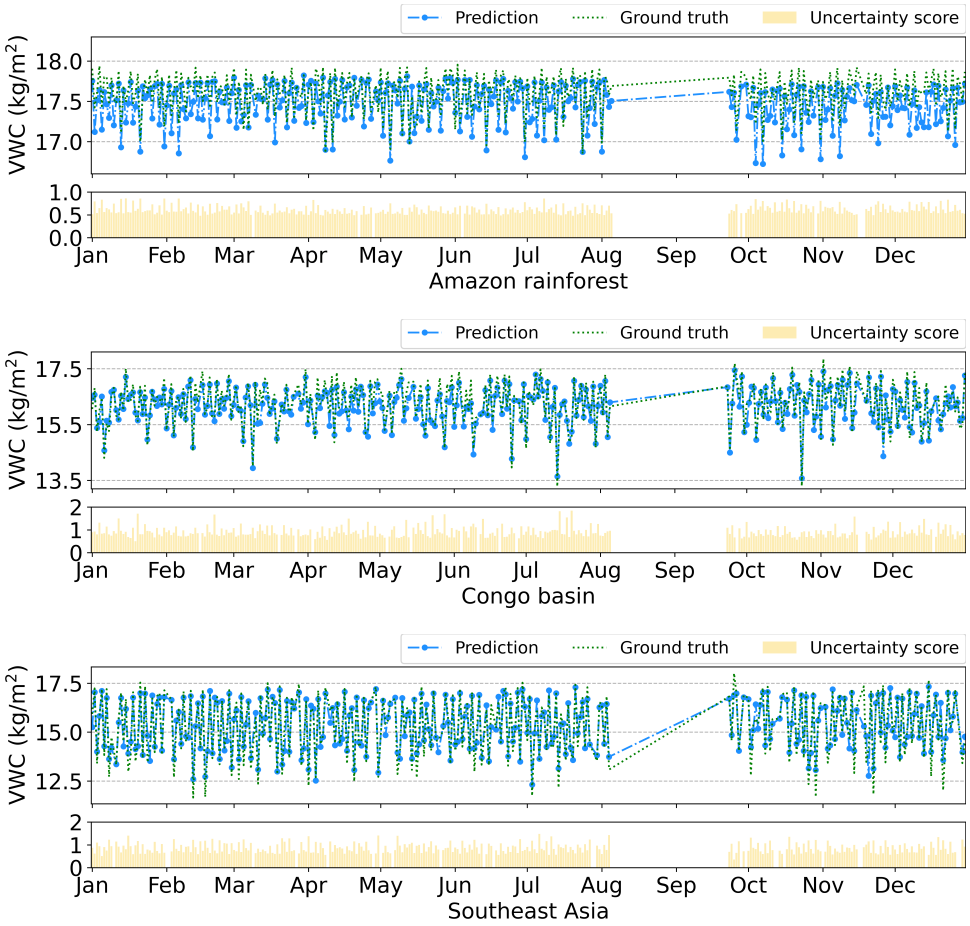


Fig. 12. VWC time series comparing ground truth and LeNet model predictions for 2022 across three case study regions. Histograms of uncertainty scores are included to illustrate the confidence levels of the predictions.

TABLE VII
LATITUDE AND LONGITUDE RANGES, TOTAL SAMPLE COUNTS, AND DAILY AVERAGE SAMPLE COUNTS FOR THREE CASE STUDY REGIONS: AMAZON RAINFOREST, CONGO BASIN, AND SOUTHEAST ASIA.

Region	Latitude / Longitude	Total Samples	Daily Average
Amazon	[10°S, 3°N] / [50°W, 76°W]	8.6×10^4	280
Congo	[5°S, 4°N] / [10°E, 29°E]	5.1×10^4	169
SEA	[6°S, 14°N] / [93°E, 118°E]	2.5×10^4	87

ble VI. In general, the RMSD and ubRMSD values for the Amazon rainforest and Congo basin are relatively low across all benchmark models, whereas the Southeast Asia region exhibits higher values. These differences may stem from variations in footprint heterogeneity: the Amazon and Congo regions primarily consist of dense vegetative basins, while Southeast Asia features more complex topography, including islands and archipelagos. Measurements in coastal regions are essentially influenced by both terrestrial and oceanic surfaces with increased variability. Regarding model biases, densely vegetated regions tend to show slight underestimations on average. Since deep learning models are trained on the global distribution of VWC data rather than localized ROIs, they are inevitably influenced by the majority of VWC intervals. Further, the consistently lower ubRMSD values compared to RMSD suggest that applying bias corrections can help mitigate VWC estimation errors. Among the evaluated models, LeNet demonstrates lower estimation errors and uncertainty scores, indicating that a careful selection of model architecture enhances performance and maintains generalizability across underrepresented intervals. These results highlight that the accuracy and reliability of VWC estimations are affected by regional characteristics and the spatial variability of vegetative cover.

Fig. 12 demonstrates time series plots of VWC estimations across the selected ROIs. Reference VWC and predictions of the best-performing model show strong agreement, especially for the Amazon rainforest and Congo basin. Consistent with the findings in Table VI, minor biases toward underestimation are observed in these regions. In addition, for each VWC time series, a corresponding uncertainty score histogram is depicted at the bottom. It is evident that the uncertainty scores of the Amazon rainforest are comparatively low compared to the other two regions. Spikes in uncertainty may be attributed to land cover changes or varied climatic conditions, which introduce additional challenges for precise VWC measure-

ments with a high degree of confidence. Future efforts could focus on utilizing the proposed CGS dataset to develop region-specific models for localized applications or to create more robust models with improved generalizability and enhanced estimation precision.

E. Challenges and limitations

While our study demonstrates the potential of deep learning models for estimating global VWC using the proposed dataset, several challenges and limitations must be acknowledged. First, the uncertainty quantification framework used in this study primarily captures epistemic uncertainty due to model limitations. However, a portion of the predictive uncertainty may also reflect aleatoric uncertainty, which arises from intrinsic noise in the input data and label sources. For instance, although the SMAP-derived VWC product is the most widely adopted reference for VWC studies, it is an estimated product with its own uncertainty and limitations, particularly in regions with mixed land cover. One potential alternative is to derive VWC from the dual-channel algorithm (DCA) VOD using linear relationships that depend on the foliage type of the vegetation. This approach could provide an independent degree of freedom relative to SMAP VWC.

Second, despite careful collocation, spatial mismatches remain due to differing spatial resolutions (i.e., CYGNSS with variable footprints, GLDAS at $\sim 0.25°$, and SMAP at $\sim 9\,\mathrm{km}$). These mismatches, combined with temporal discrepancies, cause inconsistencies and potential biases during the learning process. Additionally, although the current workflow includes only static variables from SMAP and is therefore not affected by differences between ascending and descending overpass times, the potential influence of diurnal variability in SMAP ancillary parameters should be carefully considered in further correlation analyses. Such temporal variations may propagate into model predictions and introduce ambiguities in downstream applications.

Finally, the CYGNSS measurements are subject to inherent uncertainties. The GNSS-R scattered signals are influenced by noise, variations in scattering components, and differing surface properties, which result in significant variability in the effective footprint size. The spatial heterogeneity complicates the interpretation of GNSS-R signatures, as the footprints often include a mixture of surface types within the glistening zone. Future work could explore spatial interpolation and dynamic gridding strategies to mitigate these issues and optimize the trade-off between spatiotemporal coverage and correlated uncertainties in the retrievals.

VII. Conclusion

In this paper, we investigated deep learning-based GNSS-R VWC estimation by incorporating measurements from CYGNSS, GLDAS, and SMAP missions. Using the proposed CGS dataset, our study delved into the synergy between large-scale datasets and deep learning models to achieve accurate VWC predictions. The presented scheme for global daily VWC estimation could contribute to a deeper understanding of vegetation dynamics and hydrological cycle. By applying MC dropout technique, we further quantified estimation uncertainty, providing a measure of reliability for daily estimations in the absence of reference data.

With the increasing availability of constellations and data, we hope our approach can inspire further research in combining GNSS-R vegetation monitoring with deep learning algorithms. By integrating different sources of GNSS-R measurements and leveraging their unprecedented temporal resolution, it is anticipated that diurnal or near real-time VWC estimation could become feasible. Furthermore, advancements in deep learning models, such as physics-informed neural networks, hold promise for enhancing the physical consistency and generalizability of VWC retrievals and could help disentangle the mixed scattering contributions inherent in GNSS-R observations. In addition, spatial interpolation and dynamic gridding approaches could further improve the spatial completeness and temporal resolution of the estimation product. These developments could help fully exploit the potential of GNSS-R for vegetation and hydrological monitoring at a global scale.

References

[1] G. A. Afuye, A. M. Kalumba, and I. R. Orimoloye, "Characterisation of Vegetation Response to Climate Change: A Review," *Sustainability*, vol. 13, no. 13, p. 7265, 2021.

[2] C. J. Tucker, "Remote sensing of leaf water content in the near infrared," *Remote Sensing of Environment*, vol. 10, no. 1, pp. 23–32, 1980.

[3] J. Peñuelas, I. Filella, C. Biel, L. Serrano, and R. Savé, "The reflectance at the 950–970 nm region as an indicator of plant water status," *International Journal of Remote Sensing*, vol. 14, no. 10, pp. 1887–1905, 1993.

[4] P. Zarco-Tejada, C. Rueda, and S. Ustin, "Water content estimation in vegetation with MODIS reflectance data and model inversion methods," *Remote Sensing of Environment*, vol. 85, no. 1, pp. 109–124, 2003.

[5] T. J. Jackson, D. Chen, M. Cosh, F. Li, M. Anderson, C. Walthall, P. Doriaswamy, and E. Hunt, "Vegetation water content mapping using Landsat data derived normalized difference water index for corn and soybeans," *Remote Sensing of Environment*, vol. 92, no. 4, pp. 475–482, 2004.

[6] M. O. Jones, L. A. Jones, J. S. Kimball, and K. C. McDonald, "Satellite passive microwave remote sensing for monitoring global land surface phenology," *Remote Sensing of Environment*, vol. 115, no. 4, pp. 1102–1114, 2011.

[7] E. R. Hunt, L. Li, M. T. Yilmaz, and T. J. Jackson, "Comparison of vegetation water contents derived from shortwave-infrared and passive-microwave sensors over central Iowa," *Remote Sensing of Environment*, vol. 115, no. 9, pp. 2376–2383, 2011.

[8] D. Chaparro, G. Duveiller, M. Piles, A. Cescatti, M. Vall-llossera, A. Camps, and D. Entekhabi, "Sensitivity of L-band vegetation optical depth to carbon stocks in tropical forests: a comparison to higher frequencies and optical indices," *Remote Sensing of Environment*, vol. 232, p. 111303, 2019.

[9] H. Wang, J.-P. Wigneron, P. Ciais, Y. Yao, L. Fan, X. Liu, X. Li, J. K. Green, F. Tian, S. Tao, W. Li, F. Frappart, C. Albergel, M. Wang, and S. Li, "Seasonal variations in vegetation water content retrieved from microwave remote sensing over Amazon intact forests," *Remote Sensing of Environment*, vol. 285, p. 113409, 2023.

[10] L. Duncanson, A. Neuenschwander, S. Hancock, N. Thomas, T. Fatoyinbo, M. Simard, C. A. Silva, J. Armston, S. B. Luthcke, M. Hofton, J. R. Kellner, and R. Dubayah, "Biomass estimation from simulated GEDI, ICESat-2 and NISAR across environmental gradients in Sonoma County, California," *Remote Sensing of Environment*, vol. 242, p. 111779, 2020.

[11] P. Potapov, X. Li, A. Hernandez-Serna, A. Tyukavina, M. C. Hansen, A. Kommareddy, A. Pickens, S. Turubanova, H. Tang, C. E. Silva, J. Armston, R. Dubayah, J. B. Blair, and M. Hofton, "Mapping global forest canopy height through integration of GEDI and Landsat data," *Remote Sensing of Environment*, vol. 253, p. 112165, 2021.

[12] 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 Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–17, 2022.

[13] X. Wu, P. Guo, Y. Sun, H. Liang, X. Zhang, and W. Bai, "Recent Progress on Vegetation Remote Sensing Using Spaceborne GNSS-Reflectometry," *Remote Sensing*, vol. 13, no. 21, p. 4244, 2021.

[14] N. Pierdicca, D. Comite, A. Camps, H. Carreno-Luengo, L. Cenci, M. P. Clarizia, F. Costantini, L. Dente, L. Guerriero, A. Mollfulleda, S. Paloscia, H. Park, E. Santi, M. Zribi, and N. Floury, "The Potential of Spaceborne GNSS Reflectometry for Soil Moisture, Biomass, and Freeze–Thaw Monitoring: Summary of a European Space Agency-funded study," *IEEE Geoscience and Remote Sensing Magazine*, vol. 10, no. 2, pp. 8–38, 2022.

[15] E. E. Small, K. M. Larson, and J. J. Braun, "Sensing vegetation growth with reflected GPS signals," *Geophysical Research Letters*, vol. 37, no. 12, p. L12401, 2010.

[16] N. Rodriguez-Alvarez, A. Camps, M. Vall-llossera, X. Bosch-Lluis, A. Monerris, I. Ramos-Perez, E. Valencia, J. F. Marchan-Hernandez, J. Martinez-Fernandez, G. Baroncini-Turricchia, C. Pérez-Gutiérrez, and N. Sánchez, "Land Geophysical Parameters Retrieval Using the Interference Pattern GNSS-R Technique," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 1, pp. 71–84, 2011.

[17] N. Rodriguez-Alvarez, X. Bosch-Lluis, A. Camps, I. Ramos-Perez, E. Valencia, H. Park, and M. Vall-llossera, "Vegetation Water Content Estimation Using GNSS Measurements," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 2, pp. 282–286, 2012.

[18] A. Egido, M. Caparrini, G. Ruffini, S. Paloscia, E. Santi, L. Guerriero, N. Pierdicca, and N. Floury, "Global Navigation Satellite Systems Reflectometry as a Remote Sensing Tool for Agriculture," *Remote Sensing*, vol. 4, no. 8, pp. 2356–2372, 2012.

[19] H. Carreno-Luengo, A. Camps, J. Querol, and G. Forte, "First Results of a GNSS-R Experiment From a Stratospheric Balloon Over Boreal Forests," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 5, pp. 2652–2663, 2016.

[20] P. Ferrazzoli, L. Guerriero, N. Pierdicca, and R. Rahmoune, "Forest biomass monitoring with GNSS-R: Theoretical simulations," *Advances in Space Research*, vol. 47, no. 10, pp. 1823–1832, 2011.

[21] X. Wu and S. Jin, "GNSS-Reflectometry: Forest canopies polarization scattering properties and modeling," *Advances in Space Research*, vol. 54, no. 5, pp. 863–870, 2014.

[22] O. Eroglu, M. Kurum, and J. Ball, "Response of GNSS-R on Dynamic Vegetated Terrain Conditions," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 5, pp. 1599–1611, 2019.

[23] C. S. Ruf, S. Gleason, Z. Jelenak, S. Katzberg, A. Ridley, R. Rose, J. Scherrer, and V. Zavorotny, "The CYGNSS nanosatellite constellation hurricane mission," in *2012 IEEE International Geoscience and Remote Sensing Symposium*, pp. 214–216, 2012.

[24] A. Camps, H. Park, M. Pablos, G. Foti, C. P. Gommenginger, P.-W. Liu, and J. Judge, "Sensitivity of GNSS-R Spaceborne Observations to Soil Moisture and Vegetation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 10, pp. 4730–4742, 2016.

[25] H. Carreno-Luengo, S. Lowe, C. Zuffada, S. Esterhuizen, and S. Oveisgharan, "Spaceborne GNSS-R from the SMAP Mission: First Assessment of Polarimetric Scatterometry over Land and Cryosphere," *Remote Sensing*, vol. 9, no. 4, p. 362, 2017.

[26] C. Ruf, S. Gleason, A. Ridley, R. Rose, and J. Scherrer, "The nasa cygnss mission: Overview and status update," in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 2641–2643, 2017.

[27] H. Kim and V. Lakshmi, "Use of Cyclone Global Navigation Satellite System (CyGNSS) Observations for Estimation of Soil Moisture," *Geophysical Research Letters*, vol. 45, no. 16, pp. 8272–8282, 2018.

[28] H. Carreno-Luengo, G. Luzi, and M. Crosetto, "Sensitivity of CyGNSS Bistatic Reflectivity and SMAP Microwave Radiometry Brightness Temperature to Geophysical Parameters Over Land Surfaces," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 1, pp. 107–122, 2019.

[29] H. Carreno-Luengo, G. Luzi, and M. Crosetto, "Above-Ground Biomass Retrieval over Tropical Forests: A Novel GNSS-R Approach with CyGNSS," *Remote Sensing*, vol. 12, no. 9, p. 1368, 2020.

[30] X. X. Zhu, D. Tuia, L. Mou, G. S. Xia, L. Zhang, F. Xu, and F. Fraundorfer, "Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, pp. 8–36, dec 2017.

[31] Q. Yuan, H. Shen, T. Li, Z. Li, S. Li, Y. Jiang, H. Xu, W. Tan, Q. Yang, J. Wang, J. Gao, and L. Zhang, "Deep learning in environmental remote sensing: Achievements and challenges," *Remote Sensing of Environment*, vol. 241, p. 111716, 2020.

[32] A. Dielacher, H. Fragner, and O. Koudelka, "PRETTY - passive GNSS-Reflectometry for CubeSats," *e+i Elektrotechnik und Informationstechnik*, vol. 139, p. 25–32, feb 2022.

[33] M. J. Unwin, N. Pierdicca, E. Cardellach, K. Rautiainen, G. Foti, P. Blunt, L. Guerriero, E. Santi, and M. Tossaint, "An Introduction to the HydroGNSS GNSS Reflectometry Remote Sensing Mission," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 6987–6999, 2021.

[34] M. Asgarimehr, I. Zhelavskaya, G. Foti, S. Reich, and J. Wickert, "A GNSS-R Geophysical Model Function: Machine Learning for Wind Speed Retrievals," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, pp. 1333–1337, aug 2020.

[35] X. Liu, W. Bai, J. Xia, F. Huang, C. Yin, Y. Sun, Q. Du, X. Meng, C. Liu, P. Hu, and G. Tan, "FA-RDN: A Hybrid Neural Network on GNSS-R Sea Surface Wind Speed Retrieval," *Remote Sensing*, vol. 13, no. 23, p. 4820, 2021.

[36] T. Xiao, C. Arnold, D. Zhao, L. Mou, J. Wickert, and M. Asgarimehr, "Deep Learning in Spaceborne GNSS Reflectometry: Correcting Precipitation Effects on Wind Speed Products," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 17860–17875, 2024.

[37] F. Lei, V. Senyurek, M. Kurum, A. C. Gurbuz, D. Boyd, R. Moorhead, W. T. Crow, and O. Eroglu, "Quasi-global machine learning-based soil moisture estimates at high spatio-temporal scales using CYGNSS and SMAP observations," *Remote Sensing of Environment*, vol. 276, p. 113041, 2022.

[38] M. M. Nabi, V. Senyurek, F. Lei, M. Kurum, and A. C. Gurbuz, "Quasi-Global Assessment of Deep Learning-Based CYGNSS Soil Moisture Retrieval," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 5629–5644, 2023.

[39] Q. Yan, W. Huang, and C. Moloney, "Neural Networks Based Sea Ice Detection and Concentration Retrieval From GNSS-R Delay-Doppler Maps," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 8, pp. 3789–3798, 2017.

[40] Y. Hu, X. Hua, W. Liu, and J. Wickert, "Sea Ice Detection from GNSS-R Data Based on Residual Network," *Remote Sensing*, vol. 15, no. 18, p. 4477, 2023.

[41] E. Santi, S. Paloscia, S. Pettinato, G. Fontanelli, M. P. Clarizia, D. Comite, L. Dente, L. Guerriero, N. Pierdicca, and N. Floury, "Remote Sensing of Forest Biomass Using GNSS Reflectometry," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 2351–2368, 2020.

[42] F. Chen, F. Guo, L. Liu, and Y. Nan, "An Improved Method for Pan-Tropical Above-Ground Biomass and Canopy Height Retrieval Using CYGNSS," *Remote Sensing*, vol. 13, no. 13, p. 2491, 2021.

[43] G. Pilikos, M. P. Clarizia, and N. Floury, "Biomass Estimation with GNSS Reflectometry Using a Deep Learning Retrieval Model," *Remote Sensing*, vol. 16, no. 7, p. 1125, 2024.

[44] F. Chen, L. Liu, F. Guo, and L. Huang, "A New Vegetation Observable Derived from Spaceborne GNSS-R and Its Application to Vegetation Water Content Retrieval," *Remote Sensing*, vol. 16, no. 5, p. 931, 2024.

[45] Y. Zhang, J. Bu, X. Zuo, K. Yu, Q. Wang, and W. Huang, "Vegetation Water Content Retrieval from Spaceborne GNSS-R and Multi-Source Remote Sensing Data Using Ensemble Machine Learning Methods," *Remote Sensing*, vol. 16, no. 15, 2024.

[46] M. Asgarimehr, D. Entekhabi, and A. Camps, "Diurnal Vegetation Moisture Cycle in the Amazon and Response to Water Stress," *Geophysical Research Letters*, vol. 51, no. 19, p. e2024GL111462, 2024.

[47] M. Rodell, P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J. K. Entin, J. P. Walker, D. Lohmann, and D. Toll, "The Global Land Data Assimilation System," *Bulletin of the American Meteorological Society*, vol. 85, no. 3, pp. 381–394, 2004.

[48] D. Entekhabi, E. G. Njoku, P. E. O'Neill, K. H. Kellogg, W. T. Crow, W. N. Edelstein, J. K. Entin, S. D. Goodman, T. J. Jackson, J. Johnson, J. Kimball, J. R. Piepmeier, R. D. Koster, N. Martin, K. C. McDonald, M. Moghaddam, S. Moran, R. Reichle, J. C. Shi, M. W. Spencer, S. W. Thurman, L. Tsang, and J. Van Zyl, "The Soil Moisture Active Passive (SMAP) Mission," *Proceedings of the IEEE*, vol. 98, no. 5, pp. 704–716, 2010.

[49] CYGNSS. 2021, "CYGNSS Level 1 Science Data Record Version 3.1." Ver. 3.1. PO.DAAC, CA, USA. Dataset accessed [2023-06-01].

[50] C. Ruf, D. Mckague, D. Posselt, S. Gleason, M. P. Clarizia, V. Zavorotny, T. Butler, J. Redfern, W. Wells, M. Morris, J. Crespo, C. Chew, E. Small, D. Pasqual, T. Wang, A. Warnock, D. Mayers, M. Al-Khaldi, and A. J. O'Brien, *CYGNSS Handbook*. Ann Arbor, MI: Michigan Publishing Services, 2022.

[51] C. Chew and E. Small, "Estimating inundation extent using CYGNSS data: A conceptual modeling study," *Remote Sensing of Environment*, vol. 246, p. 111869, 2020.

[52] A. Egido, S. Paloscia, E. Motte, L. Guerriero, N. Pierdicca, M. Caparrini, E. Santi, G. Fontanelli, and N. Floury, "Airborne gnss-r polarimetric measurements for soil moisture and above-ground biomass estimation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 5, pp. 1522–1532, 2014.

[53] H. H. Nguyen, H. Kim, W. Crow, S. Yueh, W. Wagner, F. Lei, J.-P. Wigneron, A. Colliander, and F. Frappart, "From theory to hydrological practice: Leveraging CYGNSS data over seven years for advanced soil moisture monitoring," *Remote Sensing of Environment*, vol. 316, p. 114509, 2025.

[54] B. Li, M. Rodell, S. Kumar, H. K. Beaudoing, A. Getirana, B. F. Zaitchik, L. G. de Goncalves, C. Cossetin, S. Bhanja, A. Mukherjee, S. Tian, N. Tangdamrongsub, D. Long, J. Nanteza, J. Lee, F. Policelli, I. B. Goni, D. Daira, M. Bila, G. de Lannoy, D. Mocko, S. C. Steele-Dunne, H. Save, and S. Bettadpur, "Global GRACE Data Assimilation for Groundwater and Drought Monitoring: Advances and Challenges," *Water Resources Research*, vol. 55, no. 9, pp. 7564–7586, 2019.

[55] S. Chan, R. Bindlish, R. Hunt, T. Jackson, and J. Kimball, *Ancillary data report: Vegetation water content*. SMAP Science Document no. 047: California Institute of Technology, Pasadena, CA, USA, 2013.

[56] P. O'Neill, R. Bindlish, S. Chan, J. Chaubell, E. Njoku, and T. Jackson, "Soil Moisture Active Passive (SMAP) Algorithm Theoretical Basis Document: Level 2 & 3 Soil Moisture (Passive) Data Products," Technical Report JPL D-66480, Jet Propulsion Laboratory, California Institute of Technology, August 2020.

[57] O'Neill, P. E., S. Chan, E. G. Njoku, T. Jackson, R. Bindlish, J. Chaubell, and A. Colliander, *SMAP Enhanced L3 Radiometer Global and Polar Grid Daily 9 km EASE-Grid Soil Moisture, Version 5. Date Accessed 06-01-2023*. NASA National Snow and Ice Data Center Distributed Active Archive Center, 2021.

[58] C. C. Chew and E. E. Small, "Soil Moisture Sensing Using Spaceborne GNSS Reflections: Comparison of CYGNSS Reflectivity to SMAP Soil Moisture," *Geophysical Research Letters*, vol. 45, no. 9, pp. 4049–4057, 2018.

[59] N. Rodriguez-Alvarez, E. Podest, K. Jensen, and K. C. McDonald, "Classifying Inundation in a Tropical Wetlands Complex with GNSS-R," *Remote Sensing*, vol. 11, no. 9, p. 1053, 2019.

[60] J. Reynolds, M. P. Clarizia, and E. Santi, "Wind Speed Estimation From CYGNSS Using Artificial Neural Networks," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 708–716, 2020.

[61] J. F. Munoz-Martin, R. Onrubia, D. Pascual, H. Park, M. Pablos, A. Camps, C. Rüdiger, J. Walker, and A. Monerris, "Single-Pass Soil Moisture Retrieval Using GNSS-R at L1 and L5 Bands: Results from Airborne Experiment," *Remote Sensing*, vol. 13, no. 4, p. 797, 2021.

[62] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[63] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in *International Conference on Learning Representations*, 2015.

[64] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *Journal of Machine Learning Research*, vol. 15, no. 56, pp. 1929–1958, 2014.

[65] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, June 2016.

[66] H. Liu, Y. Hou, S. Jiang, M. Huang, and H. Qu, "GNSS-R Ocean Wind Speed Retrieval Algorithm Based on Fusing Frequency-Domain Information," *IEEE Geoscience and Remote Sensing Letters*, vol. 21, pp. 1–5, 2024.

[67] Q. Wang, W. Zheng, F. Wu, H. Zhu, A. Xu, Y. Shen, and Y. Zhao, "Information Fusion for Spaceborne GNSS-R Sea Surface Height Retrieval Using Modified Residual Multimodal Deep Learning Method," *Remote Sensing*, vol. 15, no. 6, p. 1481, 2023.

[68] J. Xing, D. Yang, Z. Zhang, P. Yang, and F. Wang, "Development of F-ResNet for Spaceborne GNSS-R Sea Surface Height Measurement From CYGNSS," *IEEE Communications Letters*, vol. 27, no. 10, pp. 2712–2716, 2023.

[69] M. Asgarimehr, C. Arnold, T. Weigel, C. Ruf, and J. Wickert, "GNSS reflectometry global ocean wind speed using deep learning: Development and assessment of CyGNSSnet," *Remote Sensing of Environment*, vol. 269, p. 112801, feb 2022.

[70] C. S. Ruf, S. Gleason, and D. S. McKague, "Assessment of CYGNSS Wind Speed Retrieval Uncertainty," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 1, pp. 87–97, 2019.

[71] D. Purnell, N. Gomez, N. H. Chan, J. Strandberg, D. M. Holland, and T. Hobiger, "Quantifying the Uncertainty in Ground-Based GNSS-Reflectometry Sea Level Measurements," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 4419–4428, 2020.

[72] J. Jakeman, M. Eldred, and D. Xiu, "Numerical approach for quantification of epistemic uncertainty," *Journal of Computational Physics*, vol. 229, no. 12, pp. 4648–4663, 2010.

[73] M. Abdar, F. Pourpanah, S. Hussain, D. Rezazadegan, L. Liu, M. Ghavamzadeh, P. Fieguth, X. Cao, A. Khosravi, U. R. Acharya, V. Makarenkov, and S. Nahavandi, "A review of uncertainty quantification in deep learning: Techniques, applications and challenges," *Information Fusion*, vol. 76, pp. 243–297, 2021.

[74] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning," in *Proceedings of The 33rd International Conference on Machine Learning*, pp. 1050–1059, 2016.

[75] Y. Li, S. Yan, and J. Gong, "Quantifying uncertainty in soil moisture retrieval using a Bayesian neural network framework," *Computers and Electronics in Agriculture*, vol. 215, p. 108414, 2023.

[76] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian Approximation: Appendix," *arXiv preprint*, 2016.

[77] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, " TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems," 2015. Software available from tensorflow.org.

[78] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in *3rd International Conference on Learning Representations*, 2015.

[79] D. Zhao, K. Heidler, M. Asgarimehr, C. Arnold, T. Xiao, J. Wickert, X. X. Zhu, and L. Mou, "DDM-Former: Transformer networks for GNSS reflectometry global ocean wind speed estimation," *Remote Sensing of Environment*, vol. 294, p. 113629, 2023.

[80] T. M. Roberts, I. Colwell, C. Chew, S. Lowe, and R. Shah, "A Deep-Learning Approach to Soil Moisture Estimation with GNSS-R," *Remote Sensing*, vol. 14, no. 14, p. 3299, 2022.

[81] J. F. Munoz-Martin, N. Rodriguez-Alvarez, X. Bosch-Lluis, and K. Oudrhiri, "Analysis of polarimetric GNSS-R Stokes parameters of the Earth's land surface," *Remote Sensing of Environment*, vol. 287, p. 113491, 2023.

[82] D. Bai, C. S. Ruf, and D. Moller, "Calibration of the Polarimetric GNSS-R Sensor in the Rongowai Mission," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 63, p. 1–18, 2025.

[83] M. K. Rohil and S. Mathur, "CYGNSS-derived soil moisture: Status, challenges and future," *Ecological Informatics*, vol. 69, p. 101621, 2022.

[84] T. Xiao, M. Asgarimehr, J. Wickert, D. Zhao, L. Mou, and C. Arnold, "Evaluating Feature Impact on Ocean Wind Speed Predictions: An Application of Explainable AI to GNSS Reflectometry Data," in *IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium*, pp. 1854–1858, 2024.

[85] H. Du, W. Li, E. Cardellach, S. Ribó, A. Rius, and Y. Nan, "Deep residual fully connected network for GNSS-R wind speed retrieval and its interpretation," *Remote Sensing of Environment*, vol. 313, p. 114375, 2024.

[86] S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," in *Advances in Neural Information Processing Systems*, vol. 30, Curran Associates, Inc., 2017.

Daixin Zhao (Student Member, IEEE) received the B.E. degree in remote sensing science and technology from Wuhan University, Wuhan, China, and the M.Sc. degree in geomatics engineering in University of Stuttgart, Stuttgart, Germany, in 2014 and 2018, respectively.

He was with the Institute of Photogrammetry and Geoinformation, Leibniz University Hannover, Hannover, Germany, in 2018. He is currently pursuing the Ph.D. degree in deep learning and Earth observation at the Technical University of Munich, Munich, Germany, and the German Aerospace Center (DLR), Weßling, Germany. He is also a research scientist at GFZ Helmholtz Centre for Geosciences, Potsdam, Germany. His research interests include remote sensing and Earth observation, deep learning, and GNSS reflectometry.

Daixin Zhao was a recipient of the Outstanding Student Presentation Award (OSPA) at AGU 2024 Meeting, Washington, D.C.

Milad Asgarimehr is the Principal Investigator of the AI for GNSS-R (AI4GNSSR) project at the GFZ Helmholtz Centre for Geosciences. He began his research at GFZ and Technische Universität Berlin in 2017 after being awarded a Ph.D. Fellowship by Geo.X, the geoscientific competence network in Berlin and Potsdam. His Ph.D. research focused on spaceborne GNSS Reflectometry and remote sensing. His work earned him the Bernd Rendel Prize from the German Research Foundation in 2020, and he successfully defended his Ph.D. that same year with the highest distinction. He was a visiting Researcher at Universitat Politècnica de Catalunya in Barcelona and collaborated remotely with the Massachusetts Institute of Technology as a DAAD PRIME (Postdoctoral Researchers International Mobility Experience) Fellow in 2024. He is currently a research scientist at the GFZ Helmholtz Centre for Geosciences. His research interests include integrating GNSS remote sensing with artificial intelligence for Earth system modeling by combining data from various sources and sensors.

Konrad Heidler (Student Member, IEEE) received the bachelor's degree (B.Sc.) in mathematics, the master's degree (M.Sc.) in mathematics in data science, and the Doctorate in Engineering degree (Dr.-Ing.) from the Technical University of Munich (TUM), Munich, Germany, in 2017, 2020, and 2024, respectively.

He is currently a Post-Doctoral Researcher with TUM, where he is leading the working group for visual learning and reasoning at the Chair for Data Science in Earth Observation. His research work focuses on the application of deep learning for remote sensing in polar regions, solving reasoning tasks with deep learning, and applications of self- and semi-supervised learning in Earth observation.

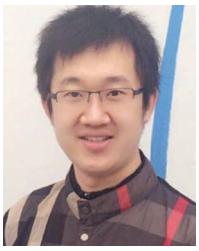
Jens Wickert received the Diploma degree in physics from Technische Universität Dresden, Germany, in 1989 and the Ph.D. degree in geophysics/meteorology from Karl-Franzens University Graz, Austria, in 2002.

He holds a joint Professorship on Global Navigation Satellite System Remote Sensing, Navigation, and Positioning of the GFZ Helmholtz Centre for Geosciences and Technische Universitaet Berlin. He is also GFZ Section Head Space Geodetic Techniques and the GFZ Director on "Atmosphere on Global Change" research. He was Principal Investigator of the pioneering GPS Radio Occultation experiment aboard the German Challenging Mini-satellite Payload (CHAMP) satellite and also the Chair of the Science Advisory Group of the GNSS REflectometry, Radio Occultation, and Scatterometry experiment onboard the International Space Station for GNSS reflectometry and radio occultation (GEROS-ISS). He has authored/coauthored more than 300 ISI listed publications on GNSS Earth observation and received several research awards.

Xiao Xiang Zhu (Fellow, IEEE) received the master (M.Sc.), doctor of engineering (Dr.-Ing.), and "Habilitation" degrees in signal processing from the Technical University of Munich (TUM), Munich, Germany, in 2008, 2011, and 2013, respectively.

She is the Chair Professor for Data Science in Earth Observation with TUM and was the founding Head of the Department "EO Data Science," the Remote Sensing Technology Institute, German Aerospace Center (DLR). Since May 2020, she has been the PI and Director of the international future AI lab "AI4EO – Artificial Intelligence for Earth Observation: Reasoning, Uncertainties, Ethics and Beyond," Munich, Germany. Since October 2020, she has also been serving as a Director of the Munich Data Science Institute (MDSI), TUM. From 2019 to 2022, she has been a co-coordinator of the Munich Data Science Research School and the head of the Helmholtz Artificial Intelligence – Research Field "Aeronautics, Space and Transport." She was a Guest Scientist or Visiting Professor with the Italian National Research Council (CNR-IREA), Naples, Italy, Fudan University, Shanghai, China, the University of Tokyo, Tokyo, Japan, and University of California, Los Angeles, CA, USA, in 2009, 2014, 2015, and 2016, respectively. She is currently a Visiting AI Professor with ESA's Phi-lab, Frascati, Italy. Her main research interests include remote sensing and Earth observation, signal processing, machine learning, and data science, with their applications in tackling societal grand challenges, e.g., Global Urbanization, UN's SDGs, and Climate Change.

Dr. Zhu has been a Member of young academy (Junge Akademie/Junges Kolleg) with the Berlin-Brandenburg Academy of Sciences and Humanities and the German National Academy of Sciences Leopoldina and the Bavarian Academy of Sciences and Humanities. She is a Fellow of the Academia Europaea (the Academy of Europe). She serves in the scientific advisory board in several research organizations, among others the German Research Center for Geosciences (GFZ, 2020–2023) and Potsdam Institute for Climate Impact Research (PIK). She is an Associate Editor of IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, Pattern Recognition and served as the Area Editor responsible for special issues of IEEE Signal Processing Magazine (2021–2023). She is a Fellow of AAIA and ELLIS.



Lichao Mou received the bachelor's degree in automation from Xi'an University of Posts and Telecommunications, Xi'an, China, in 2012, the master's degree in signal and information processing from the University of Chinese Academy of Sciences, Beijing, China, in 2015, and the Dr.-Ing. degree in remote sensing from the Technical University of Munich, Munich, Germany, in 2020.

Since 2019, he has been a Research Scientist with the Remote Sensing Technology Institute, German Aerospace Center, Weßling, Germany, and an AI Consultant at the Helmholtz Artificial Intelligence Cooperation Unit. In 2015, he spent six months with the Computer Vision Group, University of Freiburg, Freiburg, Germany. In 2019, he was a Visiting Researcher with Cambridge Image Analysis Group (CIA), University of Cambridge, Cambridge, U.K. He is currently a Guest Professor with Munich AI Future Laboratory AI4EO, TUM.

Dr. Mou was a recipient of the First Prize in the 2016 IEEE GRSS Data Fusion Contest and a finalist for the Best Student Paper Award at the 2017 Joint Urban Remote Sensing Event and the 2019 Joint Urban Remote Sensing Event.