Li *et al. Satellite Navigation* (2023) 4:4
https://doi.org/10.1186/s43020-022-00093-z

Satellite Navigation
https://satellite-navigation.springeropen.com/

ORIGINAL ARTICLE Open Access

Wind speed retrieval using GNSS-R technique with geographic partitioning

Zheng Li, Fei Guo*, Fade Chen, Zhiyu Zhang and Xiaohong Zhang

Abstract

In this paper, the effect of geographical location on Cyclone Global Navigation Satellite System (CYGNSS) observables is demonstrated for the first time. It is found that the observables corresponding to the same wind speed vary with geographic location regularly. Although latitude and longitude information is included in the conventional method, it cannot effectively reduce the errors caused by geographic differences due to the non-monotonic changes of observables with respect to latitude and longitude. Thus, an improved method for Global Navigation Satellite System Reflectometry (GNSS-R) wind speed retrieval that takes geographical differences into account is proposed. The sea surface is divided into different areas for independent wind speed retrieval, and the training set is resampled by considering high wind speed. To balance between the retrieval accuracies of high and low wind speeds, the results with the random training samples and the resampling samples are fused. Compared with the conventional method, in the range of 0–20 m/s, the improved method reduces the Root Mean Square Error (RMSE) of retrieved wind speeds from 1.52 to 1.34 m/s, and enhances the correlation coefficient from 0.86 to 0.90; while in the range of 20–30 m/s, the RMSE decreases from 8.07 to 4.06 m/s, and the correlation coefficient increases from 0.04 to 0.45. Interestingly, the SNR observations are moderately correlated with marine gravities, showing correlation coefficients of 0.5–0.6, which may provide a useful reference for marine gravity retrieval using GNSS-R in the future.

Keywords: CYGNSS, Geographical differences, Ocean wind speed, GNSS reflectometry, Marine gravity

Introduction

The change of sea surface wind speed is one of the important factors affecting the marine environment. Timely and accurate monitoring of sea surface wind speed is of significance for maritime navigation safety and contributes to improving our understanding of the marine climate environment. Traditional wind speed monitoring mainly uses ground-based stations, buoys, and meteorological remote sensing satellites like scatterometers and altimeters, etc. However, these methods suffer from limited measurement range and coverage, high cost, and high power consumption.

*Correspondence:
Fei Guo
fguo@whu.edu.cn
School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China

With a high temporal resolution and wide coverage, Global Navigation Satellite System Reflectometry (GNSS-R) has developed as a valid remote sensing technique for retrieving earth surface geophysical parameters over the past few decades (Garrison and Katzberg, 2000). Moreover, the implementation of Low Earth Orbit (LEO) satellite missions for GNSS-R, e.g., the United Kingdom-Disaster Monitoring Constellation (UK-DMC) (Gleason, 2006), TechDemoSat-1 (TDS-1) (Foti et al., 2015), and Cyclone Global Navigation Satellite System (CYGNSS), has further enhanced the advantages of temporal resolution and coverage (Ruf et al., 2013, 2016). These missions have demonstrated the potential for monitoring ocean and land parameters, e.g., Significant Wave Height (SWH) (Roggenbuck et al., 2019), sea ice extent (Yan and Huang 2016), ocean wind speed (Hammond et al., 2020), soil moisture (Arroyo et al., 2014; Pan et al., 2020), aboveground biomass (Carreno-Luengo et al., 2020; Chen

et al., 2021), etc. The focus of this paper is on wind speed retrieval.

In 1993, Martin-Neira first proposed the concept of sea surface altimetry using Global Positioning System (GPS) direct and reflected signals (Martin-Neira, 1993). In 2000, a theoretical framework was developed to establish the relationship between the properties of reflected GPS signal and the sea surface roughness (Zavorotny & Voronovich, 2000). In 2015, an ocean wind speed retrieval algorithm based on Technology Demonstration Satellite (TDS-1) data was proposed by Foti et al., which demonstrated the performance of spaceborne GNSS-R for low wind speed retrieval (Foti et al., 2015). In 2017, Foti et al. showed that the spaceborne GNSS-R data derived from TDS-1 had the capability of monitoring hurricanes. The results confirmed that GNSS-R signals can detect the ocean condition changes due to very strong and near-surface ocean wind associated with hurricanes (Foti et al., 2017).

In 2014, a Minimum Variance (MV) wind speed estimator based on five observables derived from Delay-Doppler Map (DDM) was developed (Clarizia et al., 2014). Due to its good performance, this wind speed estimator has been used in the Level 2 ocean surface wind speed data product of the CYGNSS mission (Clarizia & Ruf, 2016). In 2021, an improved CYGNSS wind speed retrieval method that combines the retrievals from DDM observables using Particle Swarm Optimization (PSO) algorithm is proposed (Guo et al., 2021). Moreover, several studies have also shown that inputting multiple observables into the Artificial Neural Network (ANN) can improve the accuracy of CYGNSS wind speed retrieval (Reynolds et al., 2020; Li et al., 2021; Asgarimehr et al., 2022). Although these observables are very sensitive to wind speed, they also are affected by many other factors, e.g., topography, swell and degree of wave development, long gravity waves, etc. (Chen-Zhang et al., 2016; Gleason et al., 2020). These factors make the relationship between CYGNSS observables and ocean wind speed complex and exhibit interactions. To obtain accurate wind speed retrieval, it is necessary to use a large amount of data to find out as many factors as possible affecting wind speed and establish the relationship between the observables and wind speed.

However, due to the complex and changeable marine environment, it's difficult to find out all the factors. To establish the accurate relationship between the CYGNSS observables and ocean wind speed, the unknown factors that affect observables are considered as a whole in this paper. Their combined influence on the observables is used for further analysis.

In this paper, an investigation on improving CYGNSS wind speed retrieval by considering geographical differences is performed. The results show that when the wind speed is the same, the observables in different sea areas are not the same, and the same observables show different wind speeds in different sea areas. This geographical difference is the result of a combination of many influencing factors. As mentioned above, it is difficult to analyze all factors individually and exclude their influence on observables. However, simply training all observables and wind speeds without eliminating the influence of other factors will inevitably lead to large retrieval errors. To solve this problem, the sea surface is divided into different areas for independent wind speed retrieval. Unlike the previous method, in each area the effect of these factors on the observables is basically the same, which can effectively reduce the retrieval error caused by these factors. Furthermore, in order to correct the error caused by the low occurrence of the high wind samples, the random training samples and the resampled samples are used for wind speed retrieval, respectively. Although this resampling method can improve the accuracy of high wind speed retrieval, the accuracy of low wind speed retrieval is reduced. To balance between the retrieval accuracies of high and low wind speeds, the results with the random training samples and the resampling samples are fused. The rest of this paper is arranged as follows. "Datasets and data filtering" section is the introduction of the experimental datasets and the data preprocessing. "The geographical difference analysis of CYGNSS observables" section analyzes the geographic differences of the observables and their relationship with marine gravity. "Method" section describes the proposed wind speed retrieval method, and its evaluation is given in "Evaluation section. Finally, the concluding remarks and suggestions for future work are presented in "Conclusions" section.

Datasets and data filtering

Datasets

The CYGNSS is designed to retrieve ocean surface wind speed with a constellation of eight small satellites. In addition to the advantages of high resolution and wide coverage with mean revisit time of 7.2 h over $0.25^\circ \times 0.25^\circ$ latitude–longitude grids (Morris and Ruf, 2017), CYGNSS can also estimate sea surface wind speed under all precipitating conditions and over the full dynamic range of wind speeds experienced in a Tropical Cyclone (TC) (Ruf et al., 2016). In this paper, the CYGNSS L1 level 3.1 version data (https://podaac.jpl.nasa.gov/dataset/CYGNSS_L1_V3.1) in the 3 months from July to September 2019 is used for wind speed retrieval.

The reference wind speed is derived from the fifth generation ECMWF (European Center for Medium-range Weather Forecasts) Atmospheric Reanalysis, (ERA5)

wind speed data provided by ECMWF (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysisera5-single-levels). ERA5 data is the fifth-generation reanalysis data of ECMWF, which provides grid wind speed data hourly with a global spatial resolution of 0.25°. The final reference wind speeds are obtained from the eastward component and the northward component of the wind speed 10 m above the surface in ERA5. After data preprocessing, the ECWMF data is matched with the CYGNSS dataset by applying the following criteria: the distance is less than 20 km, and the time difference is less than 30 min.

Data filtering

To ensure the quality of the CYGNSS data, the observables must be checked for any anomaly before ocean wind speed retrieval. The initial quality control is mainly based on the following criteria (Asgarimehr et al., 2022; Li et al., 2021; Ruf & Balasubramaniam, 2019):

(1) The receiver antenna gain in the direction of the specular point (sp_rx_gain) is larger than 0 dBi;
(2) The Normalized Bistatic Radar Cross Section (NBRCS) (ddm_nbrcs) is larger than 0;
(3) The Leading Edge Slope (LES) (ddm_les) is larger than 0;
(4) The Signal Noise Ratio (SNR) (ddm_snr) is larger than 0 dB;
(5) The incident angle of all observables is less than 35°;
(6) Data with good overall quality, indicated by the quality_flags.

The geographical difference analysis of CYGNSS observables

Previous studies show that the change of sea surface roughness is mainly caused by the sea surface wind (Foti et al., 2017; Zavorotny & Voronovich, 2000). On the other hand, the sea surface roughness can also reflect the change of wind speed. However, the sea surface roughness is affected not only by the wind speed, but also by other factors. Due to the influence of other factors, the relationship between the CYGNSS observables and the wind speed is complex and not an ideal one-to-one correspondence. Figure 1 shows the relationship between these CYGNSS observables and the wind speeds. In order to reduce the influence of other potential factors on roughness as much as possible, these effects are treated as a whole in geographical analysis.

To analyze the effect of geographical location on the CYGNSS observables, the observables are compared in the following criteria: the ocean is divided into the areas with 5° equal intervals in longitude and latitude, respectively; the average value of the observables corresponding to the same wind speed on each area. Figure 2 shows the correspondence between the wind speed and the three types of CYGNSS observables in different regions, including NBRCS, LES, and SNR. Top-left: The curve of NBRCS along longitude. Top-right: The curve of NBRCS along latitude. Middle-left: The curve of LES along longitude. Middle-right: The curve of LES along latitude. Bottom-left: The curve of SNR along longitude. Bottom-right: The curve of SNR along latitude. It can be seen from Fig. 2 that the corresponding observables of the same wind speed at different geographic areas are different, and the changing trend of the observables corresponding to the different wind speeds is very similar. It should be noted that the observables of different wind speeds in the SNR analysis chart have multiple intersections, while less for LES and the least for NBRCS. This is mainly caused by the different sensitivity of the three types of observables to wind speed. The sensitivity of the three types of observables to wind speed is NBRCS > LES > SNR, which is consistent with the previous finding (Li et al., 2021). In addition, Fig. 2 also tells that the higher the wind speed, the more intersection points will be. The reason

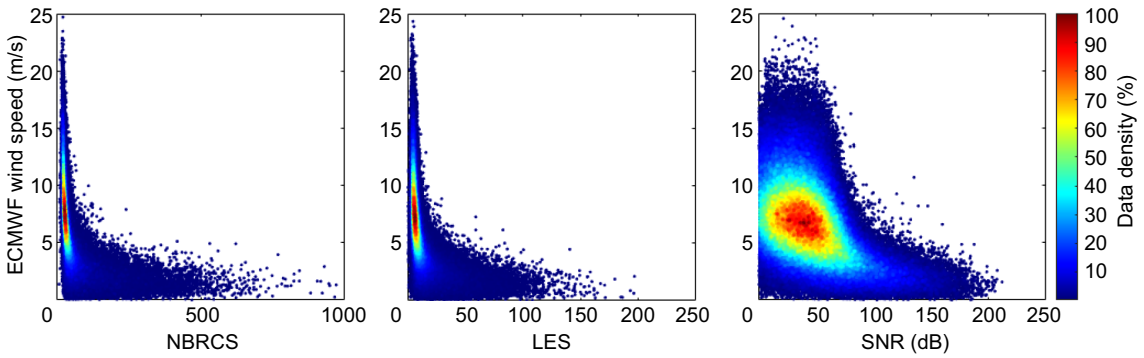

Fig. 1 The density map of the CYGNSS observables and ECMWF wind speed

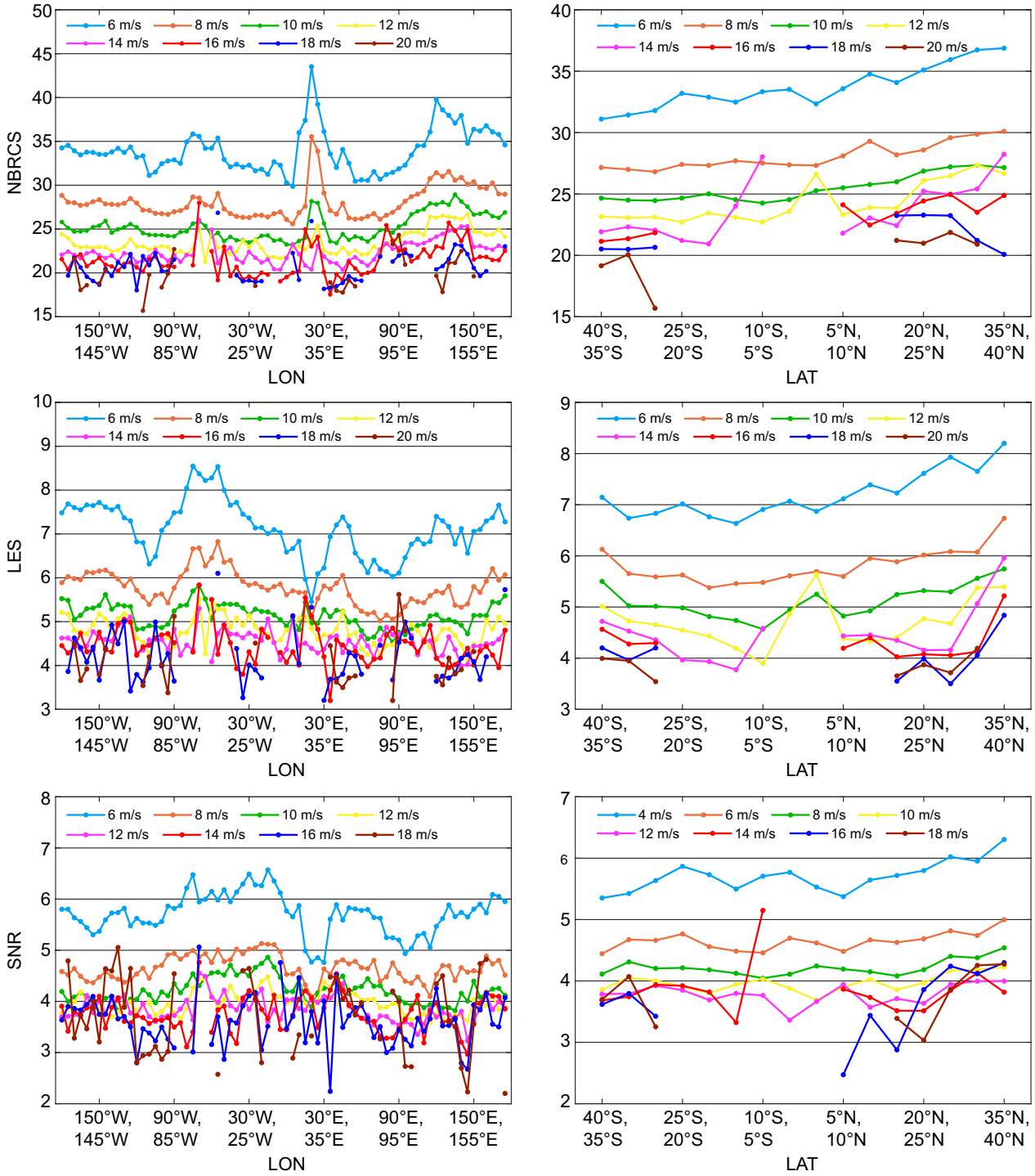

Fig. 2 The variations of observables along longitude and latitude

may be the sensitivity of observables is reduced at high wind speeds. As the wind speed increases, the difference between the observables of different wind speeds gradually decreases. This indicates that the wind speed range corresponding to the same observable will increase, which leads to an increase in retrieval error.

Overall, Fig. 2 demonstrates that the observables corresponding to the same wind speed vary with

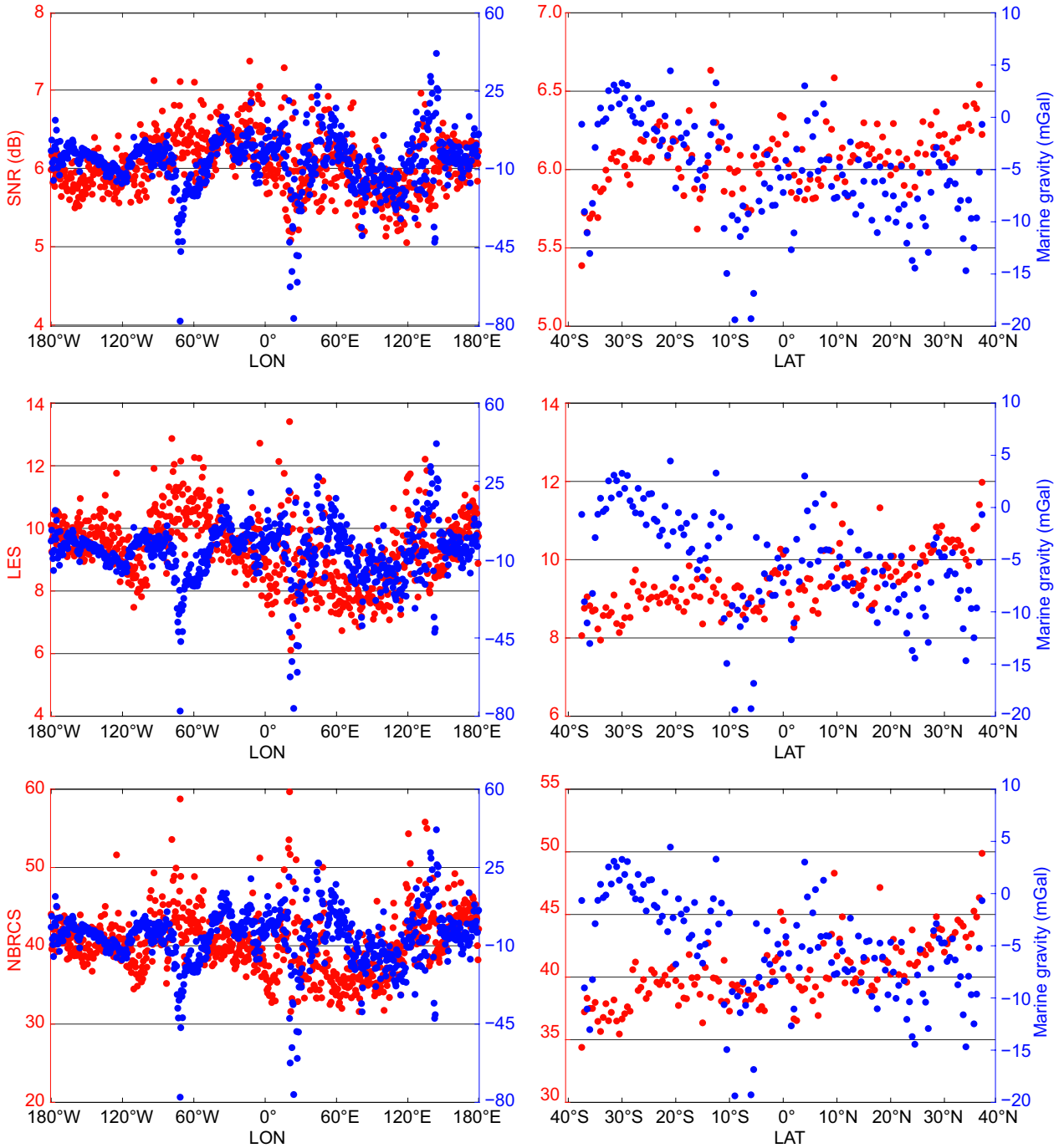

Fig. 3 The comparison between marine gravity and CYGNSS observables in different locations

geographic locations. This phenomenon is likely related to the influencing factors of sea surface roughness, including external and internal factors. Wind speed is the main external factor. The waves on the ocean surface become rougher as the sea surface wind increases. Ideally, different wind speeds would cause different roughness, thus keeping the CYGNSS observables in one-to-one correspondence with wind speed. However, the sea surface roughness may also be influenced by internal factors, such as marine gravity, which will counteract some of the wind effect on the wave. Therefore, when the wind speed is constant, the

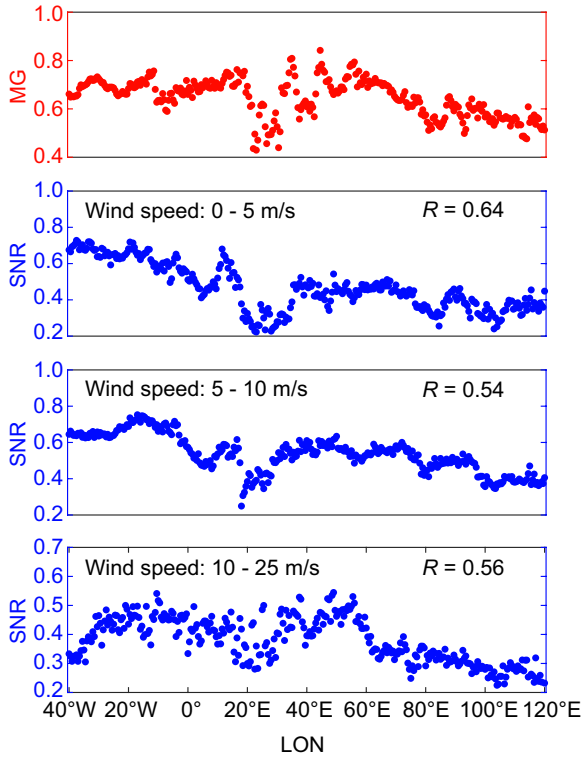

Fig. 4 The correlation between SNR and marine gravity

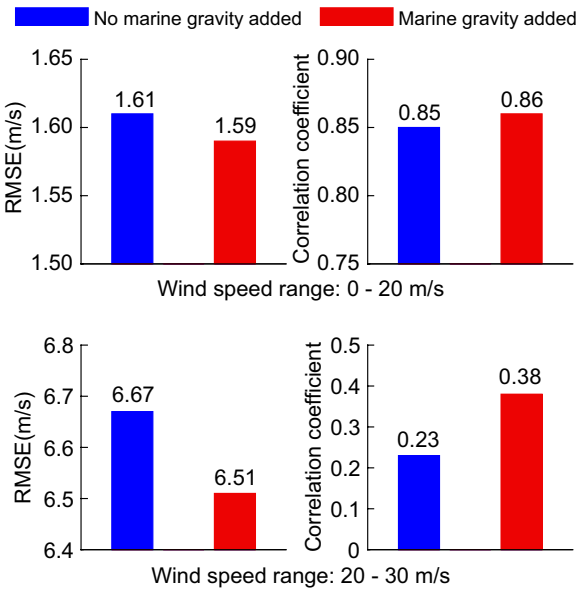

Fig. 5 The CYGNSS wind speed retrieval performance before and after the addition of marine gravity

sea surface in different gravity zones may produce different roughness.

To understand the possible impacts of marine gravity on the geographical differences, the comparison of CYGNSS observables with marine gravity is presented. The marine gravity is derived from CryoSat-2 and Jason-1 (https://topex.ucsd.edu/pub/global_grav_1min/ grav 1 min/) (Garcia et al., 2014). The following criteria are used to match the marine gravity data with the CYGNSS dataset: distance less than 20 km, and the time difference is not considered due to the marine gravity remains stable over time. Figure 3 shows the correlation between CYGNSS observables and marine gravity along longitude and latitude when the wind speed is 5 m/s, respectively. All data are divided into intervals of 0.5° along longitude and latitude, and the average values of various types of data are calculated in each interval. The left panel shows the variation of CYGNSS observables and marine gravity along the longitude direction. The right panel shows the variation of CYGNSS observables and marine gravity along latitude. It can be seen that the correlation between CYGNSS observables and marine gravity along longitude is significantly higher than that along latitude, and the variations range of marine gravity is about -15×10^{-3} to 10×10^{-3} cm/s^2 in the latitude direction and -80×10^{-3} to 60×10^{-3} cm/s^2 in the longitude direction, respectively. The results show that CYGNSS may respond to large-scale changes in ocean gravity to a certain extent.

It is noted that in Fig. 3, SNR has the highest correlation with marine gravity, but the trend of SNR with marine gravity appears anomalous in the range of longitude 80°W–40°W and 130°E–150°E. The cause for these anomalous regions needs further study. The correlation coefficients (R) between SNR and marine gravity except for the anomalous region are given in Fig. 4, where the red dots are the marine gravity and the blue dots are SNR at different wind speeds. Since the amount of data decreases rapidly with increasing wind speed, the data are divided into three ranges according to wind speed (0–5 m/s, 5–10 m/s, 10–25 m/s), and the normalized correlation coefficients are given separately. It can be seen from these plots that the correlation coefficient decreases with increasing wind speed. This can be due to a reduction in data quality and quantity at high wind speeds.

Method

As mentioned above, given a wind speed, its corresponding CYGNSS observables show regular regional differences. Therefore, the elimination of the geographical differences can contribute to expressing the relationship between the wind speed and CYGNSS observables. In addition, existing research has shown that the

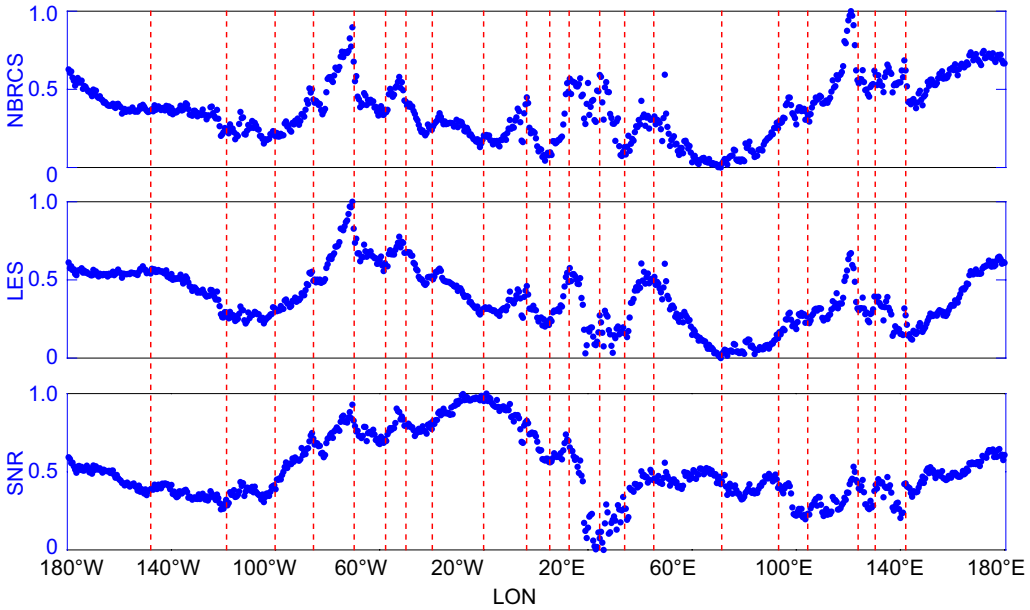

Fig. 6 Geographic partitioning based on the change of CYGNSS observables

low occurrence of the high wind samples will lead to the underestimation of high wind speeds (Guo et al., 2022; Li et al., 2021). Focusing on these two problems, a CYGNSS wind speed retrieval method considering the geographic differences and the low occurrence of the high wind samples is presented in this section.

Theory of GNSS-R wind speed retrieval

The scattered GPS signal power from the ocean can be described as a function of geometric parameters and sea surface roughness, which can be expressed as (Voronovich and Zavorotny 2017):

$$\left|Y\left(\tau,f\right)\right|^2 = \frac{P_t G_t \lambda^2 T_i^2}{(4\pi)^3} A \frac{G_r \Lambda^2(\tau) S^2\left(f\right)}{R_t^2 R_r^2} \sigma_0 \mathrm{d}A \qquad (1)$$

where $\left|Y\left(\tau,f\right)\right|^2$ is the function of the time delay τ and the frequency offset f, representing the Global Navigation Satellite System (GNSS) power; P_t is the GNSS transmitter power; G_t is the GNSS antenna gain; λ is the carrier wavelength; Ti is the coherent integration time; R_t is the distance from the transmitter to the specular reflection point; R_r is the distance from the receiver to the specular reflection point; Λ^2 and S^2 are the components of the Woodward Ambiguity Function (WAF) in delay and delay Doppler frequency, respectively; G_r is the receiver antenna gain; dA is the surface element of

Table 1 The specific information of each area

Area	Longitude range	Wind speed range (m/s)
Area-1	180°–152°W	20.97
Area-2	152°E–123°W	22.80
Area-3	123°W–107°W	24.73
Area-4	107°W–86°W	22.65
Area-5	86°W–71°W	29.69
Area-6	71°W–59°W	28.48
Area-7	59°W–52°W	19.36
Area-8	52°W–42°W	32.58
Area-9	42°W–22°W	30.22
Area-10	22°W–4°W	24.31
Area-11	4°W–4°E	18.51
Area-12	4°E–12°E	18.32
Area-13	12°E–23°E	21.15
Area-14	23°E–33°E	20.24
Area-15	33°E–44°E	21.23
Area-16	44°E–68°E	29.10
Area-17	68°E–89°E	22.65
Area-18	89°E–104°E	23.62
Area-19	104°E–121°E	19.43
Area-20	121°E–129°E	30.65
Area-21	129°E–145°E	24.81
Area-22	145°E–180°	23.08

the scattering area A; σ_0 symbolizes the NBRCS, which is related to the roughness of the glistening zone. Here, σ_0 can be used to retrieve ocean wind speed.

Selection of observables

In addition to NBRCS, the observables derived from DDM commonly used for wind speed retrieval are the LES of Integrated Delay Waveforms (IDW) and SNR (Garrison et al., 2002). The DDM_SNR derived from CYGNSS data is computed by:

$$d_{\mathrm{SNR}} = 10\log_{10}\left(\frac{S_{\max}}{N_{\mathrm{avg}}}\right) \tag{2}$$

where $S_{\max}$ is the maximum value (in raw counts) in a single DDM bin and N_{avg} is the average per-bin raw noise counts. In order to improve the sensitivity of SNR to wind speed, the bottom noise in DDM is removed and given as:

$$d_{\mathrm{SNR}'} = 10\log_{10}\left(\frac{S_{\max}-N_{\mathrm{avg}}}{N_{\mathrm{avg}}}\right) = 10\log_{10}\left(\frac{S_{\max}}{N_{\mathrm{avg}}}-1\right) \tag{3}$$

According to the instruments and geometric effects provided in the bistatic radar equation, the SNR in CYGNSS can be corrected to:

$$\begin{aligned} d_{\mathrm{SNRC}} = {} & d_{\mathrm{SNR}'} + 10\log_{10}\left(R_t^2 R_r^2\right) + 10\log_{10}\left(\cos^2\theta\right) \\ & - 10\log_{10} D_{\mathrm{EIRP}} - 10\log_{10} G_r \end{aligned} \tag{4}$$

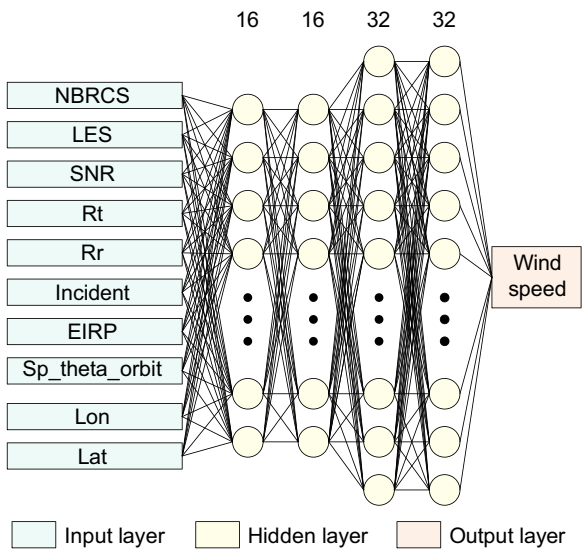

Fig. 7 Simplified visualization of the multilayer perceptron

where θ is the incident angle, D_{EIRP} is the GPS effective isotropic radiated power, and d_{SNRC} represents the corrected SNR.

Although SNR, LES, and NBRCS have high sensitivity to wind speed, Eq. (4) tells that the inclusion of other observables can improve the wind speed retrieval accuracy. Therefore, this study selects NBRCS, LES, SNR, Rt, Rr, incident, EIRP, and Sp_theta_orbit to retrieve the wind speed.

Partitioning strategy

The conventional method attenuates the effect of geographical differences by adding latitude and longitude information. However, the improvement in wind speed retrieval performance is limited due to the non-monotonic geographic variation of observations along latitude and longitude. As described in Sect. 3, after controlling for wind speed, the CYGNSS observables show a moderate correlation with marine gravity, so an attempt is made to add marine gravity to the dataset to improve retrieval performance. Figure 5 shows the comparison of the CYGNSS wind speed retrieval performance before and after the addition of marine gravity, except for the anomalous region. After the addition of marine gravity, in the range of 0–20 m/s, the Root Mean Square Error (RMSE) of retrieval wind speeds decreases from 1.61 to 1.59 m/s, and the correlation coefficient increases from 0.85 to 0.86; while in the range of 20–30 m/s, the RMSE decreases from 6.67 to 6.51, and the correlation coefficient increases from 0.23 to 0.38. The marginal effect can be attributed to the fact that additional parameters, such as temperature, salinity, and seawater density, may also affect sea surface roughness in addition to marine gravity (Liu et al., 2021). These factors are potential contributors to the multiple values of the CYGNSS observables with constant wind speed. Since there are too many possible factors which are interconnected, it is quite challenging to remove each factor's influence separately. Fortunately, the combined effect of these factors is relatively stable in terms of geographic distribution. Thus, it is possible to improve the performance of CYGNSS wind speed retrieval by geographic partitioning.

As can be seen from Fig. 2, the observations fluctuate greatly along longitude but remain much stable along latitude. Therefore, the sea surface is divided along longitude according to the change of CYGNSS observables. The observables in a divided area generally change monotonously with the longitude information, and the longitude and latitude information can better attenuate the effect of geographical differences. The partitioning strategy is shown in Fig. 6, where the blue dots are the normalized CYGNSS observables at different wind

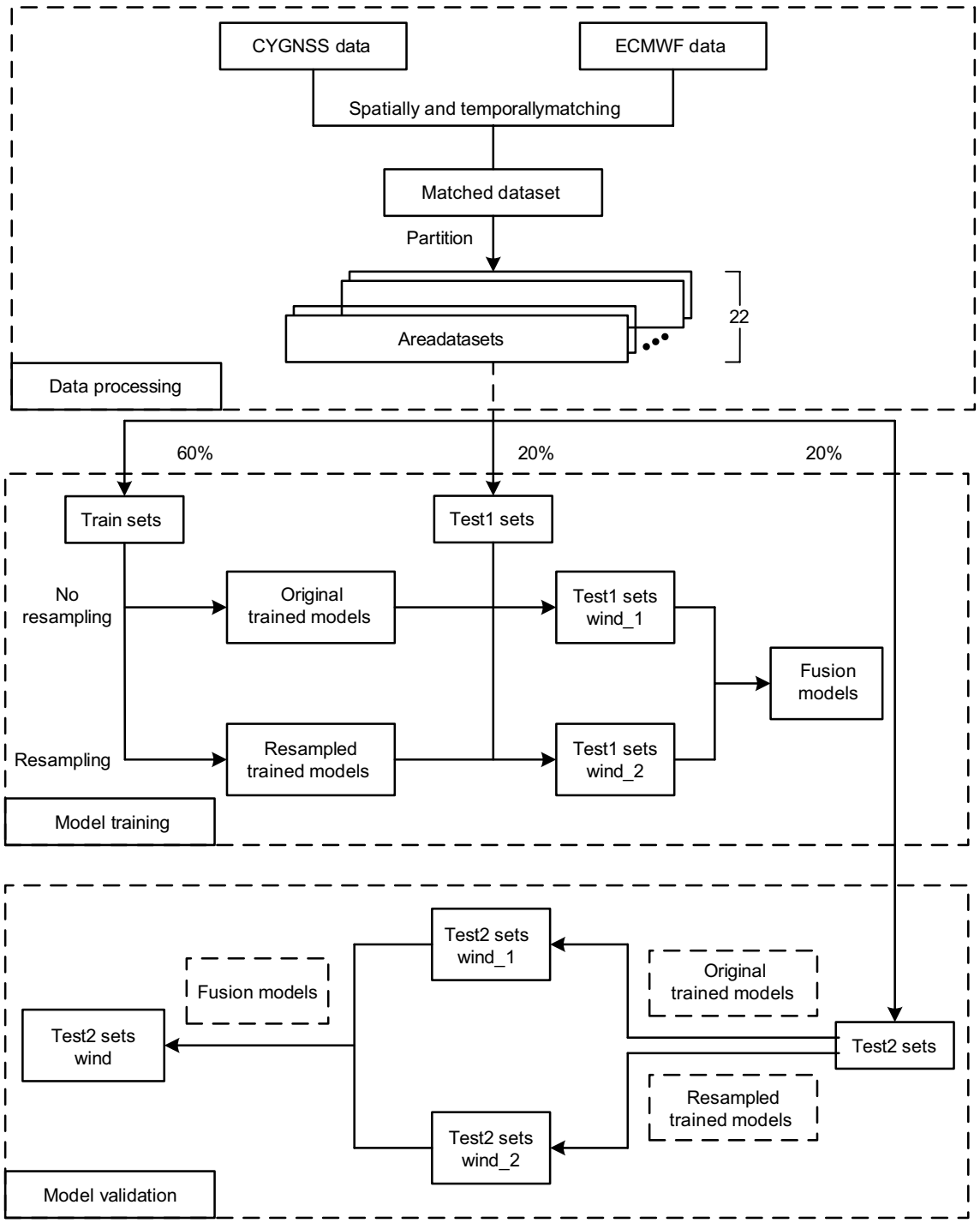

Fig. 8 The experimental structure and process

speeds, and the red dotted lines represent the regional boundaries. Finally, the sea surface is divided into 22 independent areas, and the specific information of each area is shown in Table 1.

Multilayer perceptron configuration

Recently, due to the capacity of artificial neural networks in learning complex relationships and generalizing the results from the training data, it has become a popular application tool. Previous studies have

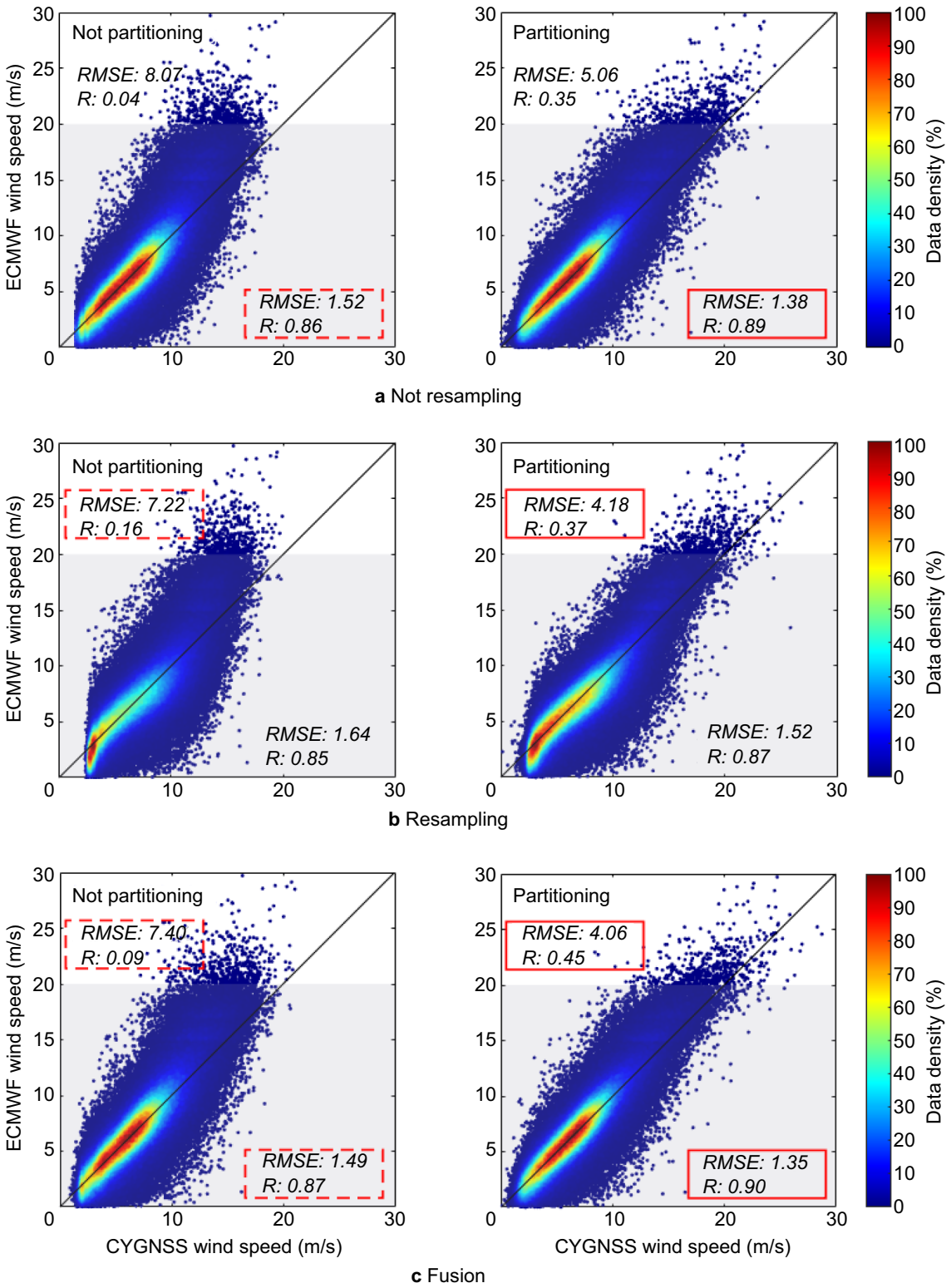

Fig. 9 Density scatterplots of ECMWF wind speed versus retrieved wind speed

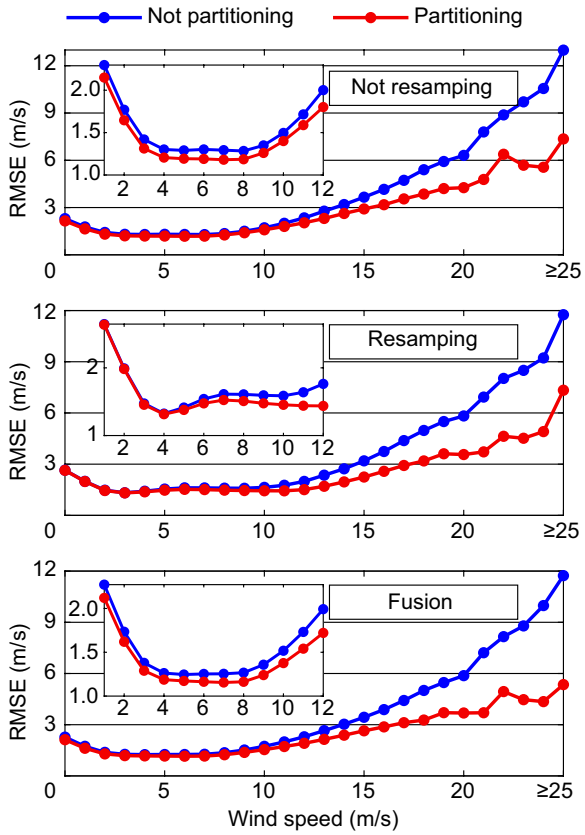

Fig. 10 Wind speed retrieval error curves of different methods

successfully used artificial neural networks for GNSS-R wind speed retrieval (Reynolds et al., 2020; Li et al., 2021). Because Multi-Layer Perceptron (MLP) is one of the most common artificial neural networks, it is used for wind speed retrieval in this paper.

MLP can automatically extract or explore some "reasonable solution rules" hidden in the data by learning from a large amount of sample data. MLP is a gradient descent procedure that computes the value of the derivative in a very efficient way and modifies the weights according to a parameter known as "learning rate" (Mart et al., 2006). Structurally, MLP is usually composed of three layers: input layer, hidden layer, and output layer. There are nodes in each layer, and the nodes in adjacent layers are connected by weights, but the nodes in each layer are independent of each other. The principle of the forward propagation of multilayer perceptron can be summarized as:

$$d_{\text{net}}^{j} = \sum_{i=0}^{N} W_{ij} X_i \tag{5}$$

where X_0 and W_{0j} are the bias ($X_0 = 1$) and its weight, respectively. N represents the number of input nodes. Each hidden node input (d_{net}^{j}) is then transformed through the non-linear transfer function to produce a hidden node output, Y_j (Rani et al., 2014). In this study, the transfer function tanh is used and expressed as follows:

$$Y_j = f\left(d_{\text{net}}^{j}\right) = \frac{e^{d_{\text{net}}^{j}} - e^{-d_{\text{net}}^{j}}}{e^{d_{\text{net}}^{j}} + e^{-d_{\text{net}}^{j}}} \tag{6}$$

The structure of the multilayer perceptron is shown in Fig. 7. Generally, the MLP wind speed retrieval performance of multi- hidden layer is always better than that of a single hidden layer, but adding more hidden MLP layers does not necessarily improve model performance. On the contrary, too many hidden layers will greatly increase the training time and even cause over-fitting and reduce model performance. After testing, the trained neural network structure with four hidden layers (the first and second layers each contain 16 neurons, and the third and fourth layers each contain 32 neurons) can achieve the best performance.

Implementation process

The experimental structure and process are displayed in Fig. 8. The CYGNSS observables data is first spatiotemporally matched with the ECMWF wind. Then, the sea surface is divided into 22 areas according to the partitioning strategy. In each area, 60% of the data is used as the training set, and the rest is for the Test1 and the Test2, 20% each. In order to weaken the underestimation of high wind speed data due to low incidence, the training set is resampled for each area to increase the proportion of high wind speed data in the training set. Training sets before and after resampling are trained separately to obtain the original trained model and the resampled trained model. The two models are used to retrieve the test sets, respectively. The results show that the original model has excellent performance at low wind speeds, while the resampled model has better performance at medium and high wind speeds. To achieve the best performance of the retrieval results in the whole wind speed range, the two retrieval wind speeds are fused. For the Test1 set, the wind speeds obtained by the two models are fed into the multilayer perceptron for training to obtain a fusion model that can accommodate both high and low wind speeds, and the model is validated using the Test2 set. It should be noted that when the trained models remain unchanged, the performance characteristic of the retrieved wind speeds is also basically unchanged. Thus, the obtained fusion model will be universal. On the other hand, if the trained model changes, the fusion model needs to be retrained. Finally, the obtained model is evaluated. For the independent Test2 set, the original

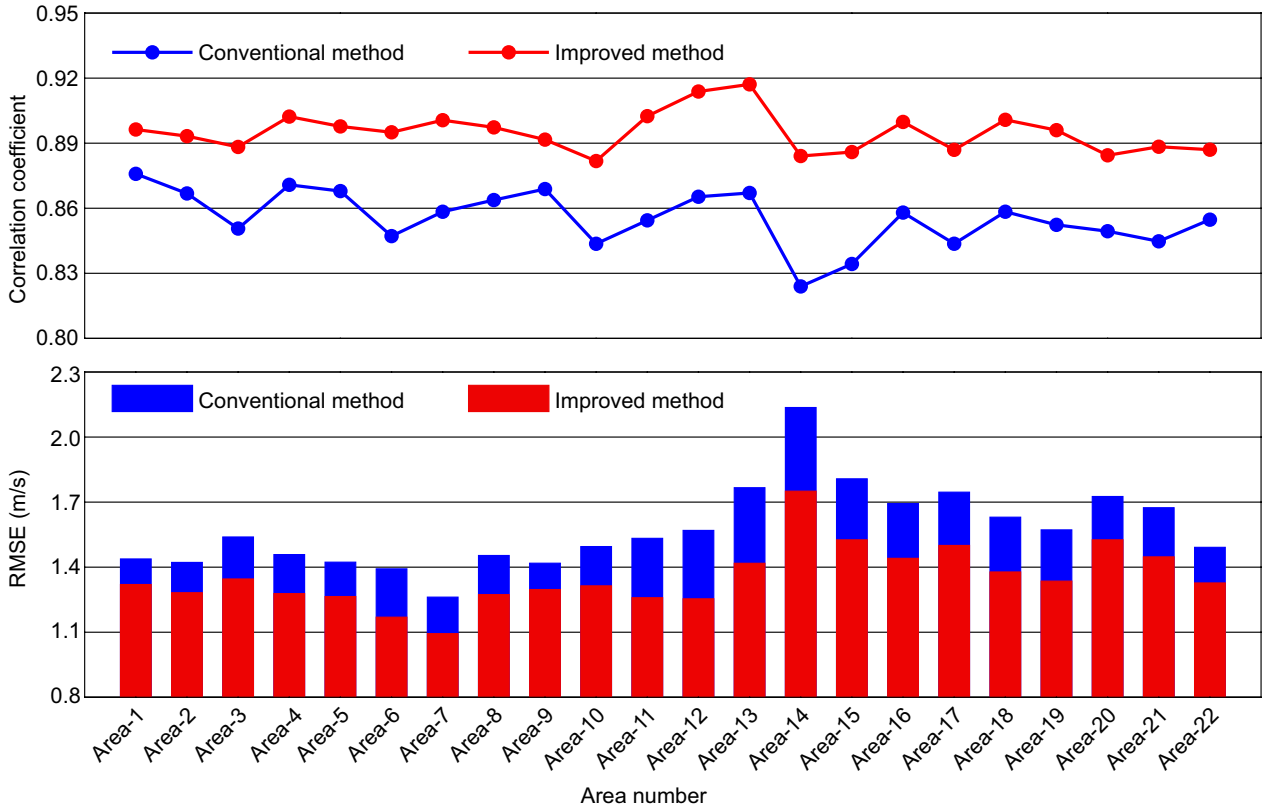

Fig. 11 The retrieval performance comparison between the conventional method and the improved method for each area

trained model and resampled trained model are used to obtain two retrieval wind speeds, and the fusion model is used to fuse the two retrieval wind speeds.

Evaluation

Figure 9 shows the wind speed inversion results with different methods, where (a), (b), and (c) are the non-resampling method, resampling method, and fusion method, respectively. The results using the methods without partitioning are shown on the left, and with partitioning on the right. It can be seen from the figure that the performances of the methods with partitioning are all better than that of the methods without partitioning. Compared with the unresampled method, the resampled method has better performance in the range of 20–30 m/s, with the RMSE reducing from 5.06 to 4.18 m/s and the R increasing from 0.35 to 0.37. However, in the range of 0–20 m/s, the RMSE increases from 1.38 to 1.52 m/s and the R decreases from 0.89 to 0.87. In contrast, the fused method has better performance at both high and low wind speeds. When the wind speed is less than 20 m/s, the RMSE is 1.34 m/s and the R is 0.90. When the wind speed is greater than 20 m/s, the RMSE is 4.06 m/s and the R is 0.45.

Figure 10 shows the error curves of the retrieval results obtained with different methods. The retrieval results for the wind speeds above 25 m/s are merged together due to very few samples. It can be seen from the figure that the retrieval errors of the methods with partitioning are smaller than those of the methods without partitioning. In the range of 0–10 m/s, the retrieval errors of the unresampled methods are smaller than those of the resampled. When the wind speed is larger than 10 m/s, the retrieval errors of the resampled are smaller than those of the unresampled. After fusion, the retrieval errors are at a low level in the whole wind speed range.

Figure 11 shows the retrieval performances in different areas using the conventional method and the improved method. The conventional method represents the unresampled method without partitioning, and the improved method represents the fusion method with partitioning. Compared with the conventional method, the improved method shows better performance in all areas, with the RMSEs decreasing from (1.26–2.14 m/s) to (1.09–1.75 m/s) and Rs increasing from (0.82–0.88) to (0.88–0.92). The specific retrieval performance in each area is given in Table 2.

Table 2 The wind speed retrieval performance in each area

Area	Conventional method		Improved method	
	RMSE (m/s)	Correlation coefficient	RMSE (m/s)	Correlation coefficient
Area-1	1.44	0.88	1.32	0.90
Area-2	1.42	0.87	1.28	0.89
Area-3	1.54	0.85	1.35	0.89
Area-4	1.46	0.87	1.28	0.90
Area-5	1.42	0.87	1.26	0.90
Area-6	1.39	0.85	1.17	0.90
Area-7	1.26	0.86	1.09	0.90
Area-8	1.46	0.86	1.27	0.90
Area-9	1.42	0.87	1.30	0.89
Area-10	1.50	0.84	1.31	0.88
Area-11	1.53	0.85	1.26	0.90
Area-12	1.57	0.87	1.25	0.91
Area-13	1.77	0.87	1.42	0.92
Area-14	2.14	0.82	1.75	0.88
Area-15	1.81	0.83	1.53	0.89
Area-16	1.69	0.86	1.44	0.90
Area-17	1.75	0.84	1.50	0.89
Area-18	1.63	0.86	1.38	0.90
Area-19	1.57	0.85	1.34	0.90
Area-20	1.73	0.85	1.53	0.88
Area-21	1.67	0.84	1.45	0.89
Area-22	1.49	0.85	1.33	0.89

Since the ANN model was trained and validated with the wind speeds from ECMWF, it is necessary to evaluate the wind speed retrieval performance with an independent source of wind speed measurements. Therefore, the wind speed measurements collected by the National Data Buoy Center (NDBC) buoys were also used as a reference. In order to reduce the impact of reflected signals from land contamination, the sites within 25 km from the coastline were excluded, and five buoy sites at different distances from the coastline (34 km, 50 km, 75 km, 200 km and 380 km) were selected for comparison experiments. The wind speed retrieval of CYGNSS observables matched with buoy data is performed using the conventional model and the improved model, respectively. The results show that the wind speed retrieval performance of the improved model is better than that of the conventional model, with the RMSEs decreasing from (0.81–1.35 m/s) to (0.76–1.31 m/s) and Rs increasing from (0.67–0.90) to (0.68–0.91). Brief information on these buoys and their specific wind speed retrieval performance are given in Table 3.

Conclusions

In this paper, the impact of the geographical differences on CYGNSS observables is investigated. When the wind speeds are the same, it is found that the observables vary with the location. The factors that cause this difference are diverse and complex. Although latitude and longitude information is included in the conventional method, it cannot effectively reduce the errors caused by geographic differences due to the non-monotonic changes of observables with respect to the location along latitude and longitude. Fortunately, the geographical differences of CYGNSS observables are regular and stable. In small areas, the fluctuations of CYGNSS observables become smooth and small. Therefore, an improved GNSS-R wind speed retrieval method is proposed, which divides the sea surface into several independent areas according to the geographical differences of CYGNSS observables. In each area, CYGNSS observables vary monotonically with longitude, and wind speeds are retrieved independently. In addition, to correct the error caused by the low occurrence of the high wind samples, the random training samples and resampled training samples are used for wind speed retrieval in each area, respectively. Although this resampling method can improve the accuracy of high wind speeds retrieval, the accuracy of low wind speeds retrieval is reduced. To balance between the retrieval accuracies for high and low wind speeds, the results with the random training samples and the resampling samples are fused.

Table 3 Brief information about buoys and wind speed retrieval performance

Buoy ID	Lon (°E)	Lat (°N)	DIST (km)	Number of matchups	Conventional model		Improved model	
					RMSE (m/s)	*R*	RMSE (m/s)	R
41009	− 81.18	28.50	34	986	1.35	0.82	1.31	0.83
41013	− 77.76	33.44	50	1602	1.35	0.90	1.27	0.91
42040	− 88.23	29.20	75	657	1.21	0.67	1.20	0.68
41010	− 78.48	28.87	200	422	0.81	0.89	0.76	0.90
51000	− 153.79	23.52	380	705	1.18	0.88	1.10	0.87

Compared with the conventional method, the improved method shows better performance at both low wind speeds and high wind speeds. In the range of 0–20 m/s, the RMSEs and Rs of the conventional method and the improved method are 1.52 m/s, 1.34 m/s and 0.86, 0.90, respectively In the range of 20–30 m/s, the RMSEs and Rs of the conventional method and the improved method are 8.07 m/s, 4.06 m/s and 0.04, 0.45, respectively. In all independent areas, the improved method shows better retrieval performance, with RMSEs decreasing from (1.26–2.14 m/s) to (1.09–1.75 m/s) and Rs increasing from (0.82–0.88) to (0.88–0.92). In conclusion, the results demonstrate that the improved method has a better performance for wind speed retrieval.

As we observed, CYGNSS observables are moderately correlated with marine gravity in some regions. However, many unknown factors such as the physical properties of seawater and the performance of the instrument may also affect the observables as shown in previous studies (Chen et al., 2022; Wang et al., 2021). Exploring these unknown factors can not only improve the performance of ocean wind speed retrieval but also provide a broader application prospect for GNSS-R. In addition, although the results obtained above show that the observables can only reflect the large-scale gravity variation in some regions, it can provide the possibility of using GNSS-R for marine gravity retrieval in the future. Of course, to obtain a better marine gravity retrieval using GNSS-R, it is necessary to exclude the influence of other factors. It is quite challenging to use GNSS-R for marine gravity retrieval and requires further research both experimentally and theoretically.

Acknowledgements
We thank NASA for providing the CYGNSS observation data and ECMWF for providing the wind speed products.

Author contributions
Conceptualization and Methodology: FG and ZL; Writing original draft: ZL and FG; Editing: FG, ZL, FC, ZZ, and XZ; Review: FG, ZL, FC, ZZ, and XZ; All authors read and approved the final manuscript.

Funding
This work was supported by the National Natural Science Foundation of China (Grant No. 42074029), the Fund for Creative Research Groups of China (Grant No. 41721003), and the Natural Science Foundation of Hubei Province for Distinguished Young Scholars (Grant No. 2021CFA039).

Availability of data and materials
The variational CYGNSS wind retrievals are available from the authors.

Declarations

Competing interests
The authors declare that they have no competing interests.

Received: 12 September 2022 Accepted: 27 December 2022
Published online: 06 February 2023

References

Arroyo, A. A., Camps, A., Aguasca, A., Forte, G. F., & Onrubia, R. (2014). Dual-polarization GNSS-R interference pattern technique for soil moisture mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7*(5), 1533–1544. https://doi.org/10.1109/JSTARS.2014.2320792

Asgarimehr, M., Arnold, C., Weigel, T., Ruf, C., & Wickert, J. (2022). GNSS reflectometry global ocean wind speed using deep learning: Development and assessment of CYGNSSnet. *Remote Sensing of Environment, 269*, 112801. https://doi.org/10.1016/j.rse.2021.112801

Carrenoluengo, H., Luzi, G., & Crosetto, M. (2020). Above-ground biomass retrieval over tropical forests: A novel GNSS-R approach with CYGNSS. *Remote Sensing, 12*(9), 1368. https://doi.org/10.3390/rs12091368

Chen, F., Guo, F., Liu, L., & Nan, Y. (2021). An improved method for pan-tropical above-ground biomass and canopy height retrieval using CYGNSS. *Remote Sensing, 13*(13), 2491. https://doi.org/10.3390/rs13132491

Chen, F., Zhang, X., Guo, F., Zheng, J., Nan, Y., & Freeshahd, M. (2022). TDS-1 GNSS reflectometry wind geophysical model function response to GPS block types. *Geo-Spatial Information Science*. https://doi.org/10.1080/10095020.2021.1997076

Chen-Zhang, D. D., Ruf, C. S., Ardhuin, F., & Park, J. (2016). GNSS-R nonlocal sea state dependencies: Model and empirical verification. *Journal of Geophysical Research Oceans, 12*(11), 8379–8394. https://doi.org/10.1002/2016JC012308

Clarizia, M. P., & Ruf, C. S. (2016). Wind speed retrieval algorithm for the cyclone global navigation satellite system (CYGNSS) mission. *IEEE Transactions on Geoscience and Remote Sensing, 54*(8), 4419–4432. https://doi.org/10.1109/TGRS.2016.2541343

Clarizia, M. P., Ruf, C. S., Jales, P., & Gommenginger, C. (2014). Spaceborne GNSS-R minimum variance wind speed estimator. *IEEE Transactions on Geoscience and Remote Sensing, 52*(11), 6829–6843. https://doi.org/10.1109/TGRS.2014.2303831

Foti, G., Gommenginger, C., Jales, P., Unwin, M., & Rosello, J. (2015). Spaceborne GNSS reflectometry for ocean winds: first results from the UK Techdemosat-1 mission. *Geophysical Research Letters*. https://doi.org/10.1002/2015GL064204

Foti, G., Gommenginger, C., & Srokosz, M. (2017). First spaceborne GNSS-Reflectometry observations of hurricanes from the UK Techdemosat-1 mission. *Geophysical Research Letters, 44*(12), 12358–12366. https://doi.org/10.1002/2017GL076166

Garcia, E. S., Sandwell, D. T., & Smith, W. (2014). Retracking CryoSat-2, Envisat and Jason-1 radar altimetry waveforms for improved gravity field recovery. *Geophysical Journal International, 196*(3), 1402–1422. https://doi.org/10.1093/gji/ggt469

Garrison, J. L., & Katzberg, S. J. (2000). The application of reflected GPS signals to ocean remote sensing. *Remote Sensing of Environment, 73*(2), 175–187. https://doi.org/10.1016/S0034-4257(00)00092-4

Garrison, J. L., Komjathy, A., Zavorotny, V. U., & Katzberg, S. J. (2002). Wind speed measurement using forward scattered GPS signals. *IEEE Transactions on Geoscience and Remote Sensing, 40*(1), 50–65. https://doi.org/10.1109/36.981349

Gleason, S. (2006). *Remote sensing of ocean, ice and land surfaces using bistatically scattered GNSS signals from low earth orbit*. Ph.D. Dissertation, University of Surrey, Guildford, UK.

Gleason, S., Johnson, J., Ruf, C., & Bussy-Virat, C. (2020). Characterizing background signals and noise in spaceborne GNSS reflection ocean observations. *IEEE Geoscience and Remote Sensing Letters, 17*(4), 587–591. https://doi.org/10.1109/LGRS.2019.2926695

Guo, W., Du, H., Cheong, J. W., Southwell, B. J., & Dempster, A. G. (2021). GNSS-R wind speed retrieval of sea surface based on particle swarm optimization algorithm. *IEEE Transactions on Geoscience and Remote Sensing, 99*, 1–14. https://doi.org/10.1109/TGRS.2021.3082916

Guo, W., Du, H., Guo, C., Southwell, B. C., Cheong, J. W., & Dempster, A. W. (2022). Information fusion for GNSS-R wind speed retrieval using statistically modified convolutional neural network. *Remote Sensing of Environment*. https://doi.org/10.1016/j.rse.2022.112934

Hammond, M. L., Foti, G., Gommenginger, C., & Srokosz, M. (2020). Temporal variability of GNSS-Reflectometry ocean wind speed retrieval performance during the UK Techdemosat-1 mission. *Remote Sensing of Environment, 242*, 111744. https://doi.org/10.1016/j.rse.2020.111744

Li, X., Yang, D., Yang, J., Zheng, G., & Li, W. (2021). Analysis of coastal wind speed retrieval from CYGNSS mission using artificial neural network. *Remote Sensing of Environment, 260*, 112454. https://doi.org/10.1016/j.rse.2021.112454

Liu, B., et al. (2021). First assessment of CYGNSS-incorporated SMAP sea surface salinity retrieval over pan-Tropical Ocean. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14*, 12163–12173. https://doi.org/10.1109/JSTARS.2021.3128553

Mart, R., El-Fallahi, A., & Lasdon, L. (2006). Path relinking and GRG for artificial neural networks. *European Journal of Operational Research, 169*(2), 508–519. https://doi.org/10.1016/j.ejor.2004.08.012

Martin-Neira, M. (1993). A passive reflectometry and interferometry system (PARIS): Application to ocean altimetry. *ESA Journal, 17*(4), 331–355.

Morris, M., & Ruf, C. S. (2017). Determining tropical cyclone surface wind speed structure and intensity with the CYGNSS satellite constellation. *Journal of Applied Meteorology and Climatology, 56*(7), 1847–1865. https://doi.org/10.1175/JAMC-D-16-0375.1

Pan, Y., Ren, C., Liang, Y., Zhang, Z., & Shi, Y. (2020). Inversion of surface vegetation water content based on GNSS-IR and MODIS data fusion. *Satellite Navigation, 1*(1), 21. https://doi.org/10.1186/s43020-020-00021-z

Rani, B., Srinivas, K., & Govardhan, A. (2014). Rainfall prediction with TLBO optimized ANN. *Journal of Scientific and Industrial Research, 73*, 643–647.

Reynolds, J., Clarizia, M.P., & Santi, E. (2020). Wind speed estimation from CYGNSS using artificial neuralnetworks. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13*, 708–716. https://doi.org/10.1109/JSTARS.2020.2968156

Roggenbuck, O., Reinking, J., & Lambertus, T. (2019). Determination of significant wave heights using damping coefficients of attenuated GNSS SNR data from static and kinematic observations. *Remote Sensing, 11*(4), 409. https://doi.org/10.3390/rs11040409

Ruf, C., Posselt, D., Majumdar, S., Gleason, S., & Morris, M. (2016). CYGNSS handbook. Michigan Publishing Services.

Ruf, C. S., et al. (2016). New ocean winds satellite mission to probe hurricanes and tropical convection. *Bulletin of the American Meteorological Society*. https://doi.org/10.1175/BAMS-D-14-00218.1

Ruf, C., & Balasubramaniam, R. (2019). Development of the CYGNSS geophysical model function for wind speed. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12*(1), 66–77. https://doi.org/10.1109/JSTARS.2018.2833075

Ruf, C., Gleason, S., Jelenak, Z., Katzberg, S., Ridley, A., Rose, R., et al. (2013). The NASA EV-2 Cyclone Global Navigation Satellite System (CYGNSS) mission. *IEEE Aerospace Conference*. https://doi.org/10.1109/AERO.2013.6497202

Voronovich, A. G., & Zavorotny, V. U. (2017). Bistatic radar equation for signals of opportunity revisited. *IEEE Transactions on Geoscience and Remote Sensing, 56*(4), 1959–1968. https://doi.org/10.1109/TGRS.2017.2771253

Wang, T., Ruf, C. S., Gleason, S., O'Brien, A. J., & Russel, A. (2021). Dynamic calibration of GPS effective isotropic radiated power for GNSS-reflectometry earth remote sensing. *IEEE Transactions on Geoscience and Remote Sensing, 99*, 1–12. https://doi.org/10.1109/TGRS.2021.3070238

Yan, Q., & Huang, W. (2016). Spaceborne GNSS-R sea ice detection using delay-doppler maps: First results from the UK Techdemosat-1 mission. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9*(10), 4795–4801. https://doi.org/10.1109/JSTARS.2016.2582690

Zavorotny, V. U., & Voronovich, A. G. (2000). Scattering of GPS signals from the ocean with wind remote sensing application. *IEEE Transactions on Geoscience and Remote Sensing, 38*(2), 951–964. https://doi.org/10.1109/36.841977

Publisher's Note