Development and Application of a GNSS-R Error Model for Hurricane Winds

Rajeswari Balasubramaniam, *Member, IEEE*, and Christopher S. Ruf, *Life Fellow, IEEE*

***Abstract*—A parametric error model is developed to represent the uncertainty in retrieval of hurricane force wind speed by a spaceborne GNSS-R instrument. The functional form of the model is constructed based on a bottom–up consideration of the primary contributing sources of uncertainty. Scaling parameters in the model are tuned in a top–down manner using a large population of wind speed retrievals by the CYGNSS satellite, which are colocated in space and time with HWRF reanalysis hurricane winds in the North Atlantic during 2018–2022. The root-mean-squared difference between CYGNSS and HWRF winds is found to depend on a number of variables, two of which are wind speed and receive antenna gain. The parametrized error model represents these dependencies. The model can be used as a design tool to predict expected performance as a function of instrument design. In particular, the model predicts the antenna gain required to achieve a particular level of wind speed uncertainty at a particular wind speed. A case study is considered in which a receive antenna gain of at least 20 dBi is found to be required to reliably distinguish between a Category 4 and Category 5 hurricane. This has implications for the optimal design of a future GNSS-R instrument intended for hurricane observations.**

***Index Terms*—CYGNSS, GNSS-R, hurricanes, science antenna, wind sensitivity.**

I. Introduction

OVER the last two decades, research utilizing reflected GPS signals for remote sensing has demonstrated the ability of GNSS-R to provide useful geophysical information. Several low Earth orbiting missions have shown successful retrieval of ocean roughness, soil moisture, sea ice extent, etc., [1], [2], [3], [4], [5]. The interest in GNSS-R has consistently risen over the years primarily due to its low-cost, low-powered, low-mass systems that could lead to affordable spaceborne constellation of such receivers that can dramatically improve the spatio-temporal sampling of the Earth's surface. This capability is especially useful for ocean remote sensing applications in fast evolving phenomena such as hurricanes. Furthermore, because GNSS-R uses L-band GNSS signals, the observations are less affected by the atmospheric attenuation caused by extreme weather conditions. Therefore, they yield more robust wind estimates in hurricane systems that are characterized by heavy rain conditions.

Manuscript received 7 July 2023; revised 22 October 2023 and 7 December 2023; accepted 14 December 2023. Date of publication 18 December 2023; date of current version 4 January 2024. This work was supported in part by NASA Science Mission Directorate under Contract NNL13AQ00C with the University of Michigan. *(Corresponding author: Christopher S. Ruf.)*

The authors are with the Department of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: rajibala@umich.edu; cruf@umich.edu).

Digital Object Identifier 10.1109/JSTARS.2023.3344371

To test the viability of GNSS-R from space, the U.K. Disaster Monitoring Constellation (U.K.-DMC) platform was equipped with a GNSS-R payload and launched into a 680 km sun-synchronous orbit in September 2003 [6]. This was the first successful dedicated GNSS-R demonstration from space and, over the years, several spaceborne platforms have successfully made GNSS-R observations from space. The complete list of all past and upcoming GNSS-R satellite missions is given in Table I.

To improve the utilization of spaceborne GNSS-R systems for Earth observation and operational weather monitoring, it is helpful to characterize the sensitivity of GNSS-R signals to geophysical variables as a function of instrument parameters. Sensitivity studies connect geophysical retrieval characteristics to the design parameters of the instruments. This aids our understanding of the performance of existing instruments and it can provide guidance with the design of future instruments.

In this work, we assess the sensitivity of GNSS-R measurements to hurricane wind speeds specifically as a function of the gain of the receive antenna that is used to collect the reflected GNSS signals. It is found that the antenna gain sets a limit on the sensitivity of GNSS-R observations to wind speed changes in high wind conditions. This results because the forward scattered signal strength significantly reduces with an increase in wind speed. Table I includes the peak antenna gain of the receive antenna used by each of the past and upcoming spaceborne GNSS-R missions. It can be seen that they all have roughly similar values in the range from 11 to 15 dBi. We examine what the implications of these gain values are on the maximum wind speed that can be measured reliably and consider what (higher) gain is required to reliably measure higher wind speed values.

A semiempirical error model for ocean wind speed retrieval by the CYGNSS satellites is developed. CYGNSS is a NASA mission which uses GNSS-R to determine near-surface ocean wind speed [7], [8], [9]. The functional form of the error model is constructed to represent contributions by the primary sources of error that produce the overall uncertainty in wind speed estimates. The model includes scaling parameters to allow for adjustments to the relative strength of the individual contributions by each error source. The values of the scaling parameters are determined using a large population of colocated wind speed retrievals by the CYGNSS satellite and modeled wind speed by Hurricane Weather Research and Forecasting (HWRF) for North Atlantic storms during 2018–2022. The root-mean-squared difference (RMSD) between CYGNSS and HWRF winds and, in particular, the dependence of the RMSD on a number of

© 2023 The Authors. This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

TABLE I
LIST OF SPACE-BORNE GNSS-R MISSIONS

Satellite missions	Launch year	Altitude and inclination	Operating Frequency	Peak antenna gain
UK-DMC	2003	680 km, 98.2°	GPS L1 C/A	11.8 dBi
TDS-1	2014	635 km, 98.8°	GPS L1 C/A	13 dBi
CYGNSS	**2016**	**510 km, 35°**	**GPS L1 C/A**	**14.6 dBi**
BuFeng-1	2019	579 km, 45°	GPS L1 and BeiDou B1	14 dBi
SPIRE GNSS-R	2019	571 km, 37°	L1 GPS, Galileo, QZSS, Beidou	~11 dBi
FSSCat	2020	530 km, 97.5°	GPS L1 C/A Galileo E1	12.22 dBi
FengYun-3E	2021	808 km, 98.8°	GPS L1/L2 and BeiDou B1, Galileo E1	12 dBi
Triton	Expected in 2023	LEO, near polar	L1/L2 dual pol	14.5 dBi
HydroGNSS	Expected in 2024	LEO, near polar	GPS L1/L5 Galileo E1/E5 dual pol	14 dBi

engineering and geophysical variables are used as constraints on the parametric error model. These variables include instrument calibration, sensitivity of the wind speed inversion algorithm to measurement noise, dependence of measurement noise on antenna gain, the wind speed at the ocean surface, and other geophysical parameters than wind speed (e.g., long wave ocean swell and rain). The model is effectively tuned so that it can reproduce the dependencies that are observed. This error model can then be used as a simulator to choose an optimal gain setting in order to achieve a desired error limit at a given wind speed.

The rest of this article is organized as follows. Section II gives a description of the datasets used in this analysis and the observations derived from the CYGNSS-HWRF matchup analysis. Section III explains the theoretical basis for the error model. Section IV describes the tuning of the parametric error model: uncertainty analysis of the model and a demonstration of the use of this error model in estimating the required peak gain for a specific application. Finally, Section V provides a discussion on possible design considerations and its strengths and weaknesses. Finally, Section VI concludes this article.

II. DATA DESCRIPTION

A. CYGNSS Level 2 Wind Speed Products

NASA's CYGNSS constellation of GNSS-R satellites has been specifically designed to measure ocean surface winds in fast evolving phenomena, such as hurricanes. It has been observed that the behavior of CYGNSS measurement observables varies significantly under hurricane-like conditions when compared to a fully developed sea. Hence, two GMFs were developed, one for fully developed seas and the other for a young, underdeveloped sea (e.g., in hurricanes) [10]. The Ocean Surface Winds Team of the Center for Satellite Application and Research division of NOAA has also developed a CYGNSS wind speed data product that uses a unified GMF for both fully developed and young seas [11]. Apart from this, there is also a trackwise corrected wind speed product and a storm centric gridded wind speed product available to the science community [12], [13]. The list of all the available CYGNSS hurricane wind products is given in Table II. In this work, we will be using the CYGNSS Level 2 Science Data Record v3.1 (SDR v3.1) wind speed estimates [14] as they are found to be more sensitive to hurricane-force winds than the other products [15].

B. CYGNSS-HWRF Matchup Dataset

For this analysis, the SDR v3.1 release of CYGNSS Level 2 Young Seas Limited Fetch wind speed measurements over 56 major hurricanes from 2018, 2020, 2021, and 2022 are used. The Weather Research and Forecast (WRF) system for hurricane prediction (HWRF) is an operational model developed by the National Centers for Environmental Prediction. HWRF provides three domains (one parent and two nested) and is based on the initial position of the storm and on the National Hurricane Center forecast of the 72-h storm position. The two nested domains move along the storm with a coverage of 24° × 24° and 7° × 7° for the middle nest and the inner nest, respectively [16]. For our purposes, we use the inner nest gridding that offers the finest resolution of about 0.015° (approximately 2 km). A previous study compared HWRF reanalysis winds with several satellite observations and found a bias of less than 1 m/s and a standard deviation of less than 4 m/s for winds in the range of 10–60 m/s [21]. We, therefore, consider HWRF as a reliable and high-quality ground truth reference for purposes of our study.

The CYGNSS wind speed estimates are matched to the HWRF inner nest grid, which has a spacing of 2 km. The HWRF winds are resampled to CYGNSS resolution and are colocated to CYGNSS wind estimates with a maximum temporal separation of 60 min and a maximum spatial separation of 0.25° lat and lon. Fig. 1 shows a density scatter plot of the matched-up CYGNSS and HWRF winds. It can be observed that the density distribution is roughly symmetric with respect to the one-to-one line (shown by dashed lines in Fig. 1) for wind speeds as high as ~35 m/s.

The CYGNSS level 2 winds are filtered by several quality measures for this analysis. Only observations with the overall quality flag set to best quality are used. This results in a total dataset consisting of approximately1.28 million observations in hurricanes with 115 637 samples within the 34 knot wind radii.

TABLE II
CYGNSS HURRICANE WIND SPEED PRODUCTS

CYGNSS product	Start/End date	Wind product	Availability on PO.DAAC
CYGNSS Level 2 Science Data Record v3.1	2018-August to Present	Young Seas Limited Fetch (YSLF) wind	yes
CYGNSS Level 2 Climate Data Record v1.2	2018-August to Present	Young Seas Limited Fetch (YSLF) wind	yes
NOAA CYGNSS Level 2 v1.1	2017-May to 2022-May	Single wind algorithm, all winds	yes
CYGNSS Level 3 Storm Centric Grid	2018-August to Present	Storm centric, hurricane winds	yes
CYGNSS Level 3 Merged Data v0.0	2018-August to Present	Storm centric, all winds	no
NOAA CYGNSS Level 2 v1.2	2017-May to Present	Single wind algorithm, all winds	yes

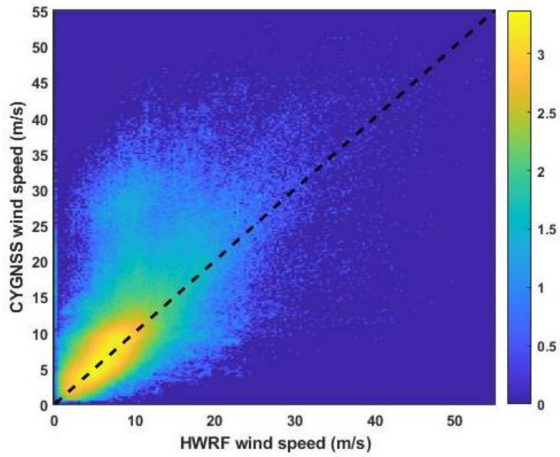

Fig. 1. Density scatter plot of CYGNSS-HWRF matchup winds for 56 major hurricanes from 2018, 2020, 2021, and 2022.

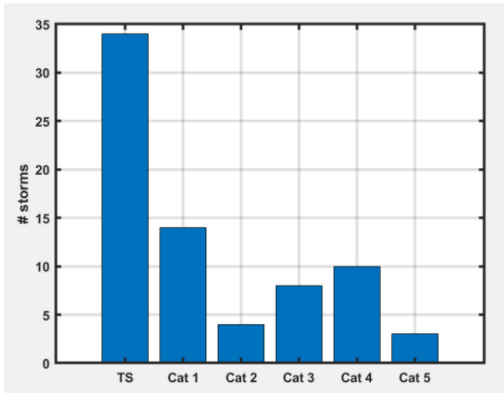

Fig. 2. Distribution of storm categories in the dataset.

TABLE III
CYGNSS-HWRF MATCHUP DATASET

Year	# samples within R34	# samples with ws>=40 m/s	Max wind (m/s)
2018	43 100	1817	78.79
2020	32 394	275	68.46
2021	21 459	362	63.60
2022	18 684	297	52.20
Total dataset	**115 637**	**2 751**	**78.79**

Table III lists the yearwise sample distribution within hurricanes and Fig. 2 shows the distribution of storm categories in this dataset.

In this work, we use the complete dataset for developing the parametric error model and further divide the dataset into ten subsets each with 50% of the samples picked through uniform random sampling. The subsets are used to derive the uncertainty in the error model and in engineering design estimates based on it. More details on this will be discussed in the subsequent sections.

Fig. 3 shows the RMSD between CYGNSS retrieved wind speeds and HWRF reference winds as a function of wind speed and the range corrected gain (RCG). RCG is defined as the receiver antenna gain, multiplied by the range losses, at the specular point. The RCG represents a more complete definition of signal gain than the antenna gain alone since it takes into account both the effect of the receiver antenna and the effect of the range losses. The RCG is used as a proxy for signal-to-noise ratio (SNR), assuming that the noise floor is constant and the variations in signal strength are due to changes in the signal gain. The expression for translating physical antenna gain to RCG is

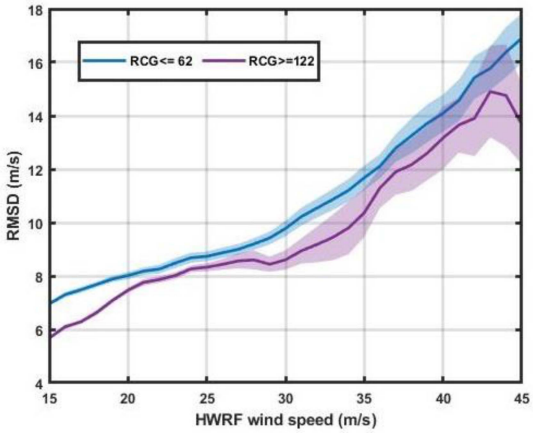

Fig. 3. RMSD as a function of wind speed and RCG. The shaded region represents the standard deviation in RMSD at different wind speeds.

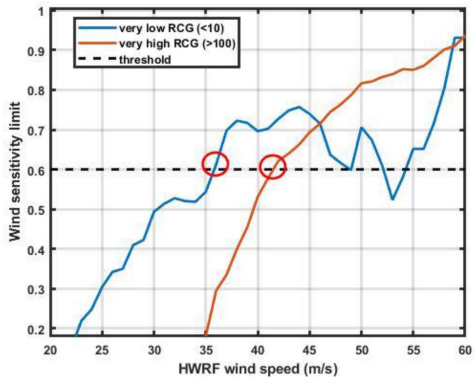

Fig. 4. Maximum wind speed limit for CYGNSS. Maximum wind speed value transitions from 36 to 42 m/s for a shift from very low to very high RCG.

given in [9]. For the sake of convenience of representation, we have dropped the term 10^{-27} in the RCG values. Therefore, an RCG of 50 would represent the number 50×10^{-27}.

Another important point to note is that our analysis of the performance of the wind speed assessment will be w.r.t. RCG for the CYGNSS measurement geometry. Applying these results to other orbits would require a rescaling of the antenna gain to compensate for the change in range. The range of RCGs available in the matchup dataset is partitioned into ten subregions in such a way that each of the ten RCG bins has almost a similar number of samples. This is done to avoid any inconsistency in the results due to sample size. For the sake of visualization here, we are depicting two RCG ranges in Fig. 3. The wind speed range for the analysis is set between 15 and 45 m/s. This range of wind speed is specifically chosen to characterize the performance for high wind conditions. Below 15 m/s, the HWRF reanalysis wind errors become significant, and above 45 m/s, the number of matchup samples is insufficient for developing a statistically significant error model. The figure shows the variation of RMSD with respect to both wind speed and RCG. The standard deviation in the RMSD as a function of wind speed is depicted by the shaded region in the figure. The RMSD metric is given by

$$\text{RMSD}\ (u, \text{RCG}) = \sqrt{\left\langle \left(\boldsymbol{u}_{\text{CYGNSS}} - \boldsymbol{u}_{\text{HWRF}}\right)^2 \right\rangle} \qquad (1)$$

where u is the wind speed bin and RCG is the gain bin over which the error is evaluated, $\boldsymbol{u}_{\text{CYGNSS}}$ is the set of CYGNSS wind samples in the given wind and gain bins, and $\boldsymbol{u}_{\text{HWRF}}$ is the set of corresponding HWRF measurements. The focus of this work is on developing an error model that can best represent this 2-D behavior of the RMSD as a function of wind speed and RCG.

C. Wind Speed Sensitivity

As an aid in understanding the dependence of GNSS-R signals on antenna gain, we consider the signal sensitivity to winds as a function of RCG. This sensitivity assessment is shown in Fig. 4.

We define the wind sensitivity limit as the maximum wind above which CYGNSS measurements lose sensitivity to wind speed changes. This can be represented using the following statistical metric:

$$\text{Wind sensitivity limit } = \text{abs}\left(\frac{\text{bias}}{\text{RMSD}}\right)\Bigg|_{w}. \qquad (2)$$

We use this metric as a qualitative figure of merit to highlight the influence of bias on RMSD. The metric identifies how much of the overall error is driven by bias versus by the retrieval variance. At a particular wind speed, if the bias is the major contributor to the RMSD, we attribute this to a diminished sensitivity of the observed $\sigma_0$to changes in wind speed. This occurs when the ratio approaches unity. We selected two limiting RCG ranges (very low and very high) to highlight the dependence on sensitivity to wind speed. A threshold ratio of 0.6 or lower is assumed to correspond to "significant" wind speed sensitivity by the CYGNSS measurements. It can be observed in Fig. 4 that the wind sensitivity limit improves from 36 m/s at very low RCG values to 42 m/s at very high RCG values.

Another metric to evaluate the wind sensitivity limit is the ratio of RMSD to wind speed as a function of wind speed, as shown in Fig. 5. Two examples using data with different ranges of RCG are shown with their standard deviation plotted as the shaded region around them. It can be observed that the ratio monotonically drops with increasing wind speed up to values of 30–35 m/s, indicating the upper bound of good wind sensitivity. Above this, the ratio begins to increase, suggesting a diminished wind speed sensitivity. Note that the increase begins at a higher wind speed for the data with higher RCG values.

III. Functional Form of Error Model

In Section II, we observed that RMSD varies as a function of wind speed and antenna gain. Use of RMSD can be considered a top–down approach to characterizing retrieval uncertainty in that it considers the overall, end-to-end performance of the retrieved geophysical parameter, wind speed in this case, with respect to an independent estimate (HWRF reanalysis wind) of the same parameter. An error model can be constructed by decomposing

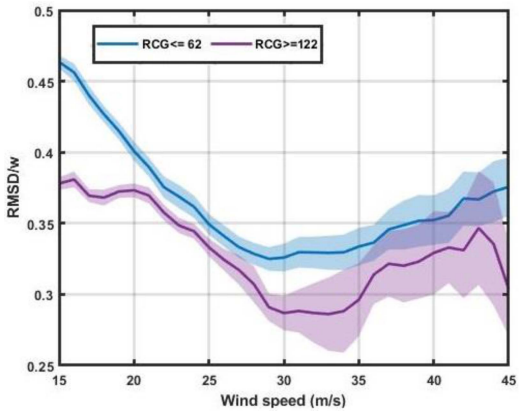

Fig. 5. Observing wind sensitivity limit through the ratio of RMSD to wind speed as a function of wind speed. The shaded region represents the standard deviation in the ratio at different wind speeds.

the RMSD into its individual contributors. If the components are assumed to be independent uncorrelated error sources that can be characterized as zero mean Gaussian distributions, then the total error can be represented as the sum of the squares of the individual error sources [17], or

$$\epsilon_u^2 = \sum_{i=1}^{x} [\epsilon_{qi}]^2 + \epsilon_{\text{bias}}^2 \tag{3}$$

where ϵ_u represents the overall RMSD error and x represents the total number of independent error terms. ϵ_{qi} are the individual errors contributed by imperfect engineering calibration, retrieval algorithm implementation, geophysical sensitivities, and other effects, and ϵ_{bias} represents the error due to biases (i.e., nonzero mean values of the individual errors). This method of error decomposition into its engineering, geophysical retrieval, and other sources is referred to as the bottom–up approach [18]. The individual error terms can be estimated by taking the partial derivatives of the appropriate inversion model with respect to the parameter of interest

$$\epsilon_{qi} = \left| \frac{\partial I}{\partial qi} \right| \Delta qi \tag{4}$$

where I represents the inversion model that maps a particular error source to retrieved wind speed, qi are the respective sources of error, and Δqi are the magnitudes of the errors in each source.

A. Identifying Individual Sources of Error

As the first stage in identifying individual error sources, we split the RMSD into its two major components, namely, error in the normalized bistatic radar cross section (σ_0) and error due to nonwind effects. The error in σ_0 folds into it all the errors associated with unwrapping of the forward model and the error in the empirical geophysical model function (GMF) used in the wind retrieval algorithm to translate σ_0 into equivalent near surface wind speed. Specifically, the forward model that maps wind speed to σ_0 is given by

$$\sigma_0 = \text{GMF}(\text{u}) \tag{5}$$

and the wind speed retrieval algorithm is essentially an inversion of the GMF given measurements of σ_0.

The nonwind error component encompasses all others source of errors, such as those related to the fact that there is not a one-to-one correspondence between wind speed and σ_0 (such nonwind geophysical effects include long wave ocean swell and rain), errors in the HWRF reference measurements, errors in spatio-temporal colocation of CYGNSS measurements with the reference, and other possible sources of errors. This first stage of the error model can be represented by

$$\epsilon_u^2 = \epsilon_{u,\text{others}}^2 + \epsilon_{u,\sigma_0}^2 \tag{6}$$

where $\epsilon_{u,\text{others}}^2$ represents the bias error [ϵ_{bias} in (3)] plus any nonwind sources of error, and ϵ_{u,σ_0}^2 represents errors associated with the measurement variable σ_0. As the next stage in this error decomposition process, we break down the error in the measurement variable into two components using the same partial derivative approach

$$\left.\frac{\partial \sigma_0}{\partial u}\right|_u = \left.\frac{\partial \text{GMF}}{\partial u}\right|_u \tag{7}$$

$$\epsilon_{u,\sigma_0}^2 = \left| \left.\frac{\partial \text{GMF}}{\partial u}\right|_u \right|^{-2} \epsilon_{\sigma_0}^2. \tag{8}$$

The error in wind speed due to σ_0 is a composite of error in the GMF used in the wind retrieval algorithm and the engineering calibration equations. The error associated with the GMF can be evaluated by taking the partial derivative of the GMF with respect to wind speed. $\epsilon_{\sigma_0}^2$ represents error in instrument calibration and other potential unknown error sources. As the final stage in the error decomposition process, we break down the errors into two major components: the error associated with antenna gain and all other errors rolled up into another component, as given by

$$\epsilon_{\sigma_0}^2 = \epsilon_{\sigma_0,\text{others}}^2 + \epsilon_{\sigma_0,G}^2 \tag{9}$$

where $\epsilon_{\sigma_0,G}^2$ represents the error in σ_0 associated with antenna gain and $\epsilon_{\sigma_0,\text{others}}^2$ represents all other rolled up errors in σ_0. The purpose of isolating the error associated with the magnitude of the antenna gain is specifically for modeling the errors as a function of antenna gain, which is the primary objective of this work.

B. Tuning Bottom–Up Error Models With Top–Down RMSD

The overall wind speed uncertainty can be represented as the squared sum of the individual sources of error. The RMSD is the representation of the overall performance and is the primary component of the top–down error. The individual error sources shown in (6)–(9) are the components of the bottom–up error analysis approach. The top–down approach is useful as an end-to-end assessment of retrieval performance and, when combined with the bottom–up assessment, can serve to empirically tune the error model on which it is based. Combining the two approaches, we formulate the error model as

$$\epsilon_u^2 = \left| \left.\frac{\partial \text{GMF}}{\partial u}\right|_u \right|^{-2} \left(\epsilon_{\sigma_0,\text{others}}^2 + \epsilon_{\sigma_0,G}^2\right) + \epsilon_{u,\text{others}}^2. \tag{10}$$

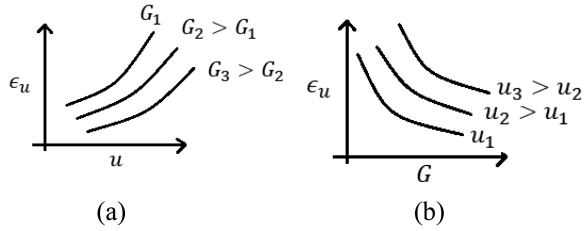

Fig. 6. Conceptual illustration of the dependence of RMSD on (a) wind speed (u) for different values of RCG (G), and (b) on G for different values of u.

Equation (10) is the theoretical error model developed for CYGNSS measurements of ocean surface wind speed. In the next section, the various components of this error model are parametrically tuned to the empirical behavior of the observed RMSD.

IV. Parametric Error Model

As shown in (10), the theoretical error model has five major components. ϵ_u is the RMSD error between CYGNSS and HWRF winds, which is a known parameter; $\frac{\partial \text{GMF}}{\partial u}|_u$ is the slope of the GMF at any given wind speed u, which is a known parameter; $\epsilon_{\sigma_0,\text{others}}$ is the error in σ_0 due instrument calibration uncertainties, uncertainties in unwrapping the forward model and other rolled up sources of unknown engineering errors, which is an unknown parameter; $\epsilon_{\sigma_0,G}$ is the error in σ_0 due to the error in antenna gain estimation, which is also an unknown quantity; finally, $\epsilon_{\text{u, others}}$ accounts for nonwind geophysical effects, error in reference winds, error in spatio-temporal co-location of the matchup, and other rolled up unknown sources of error, which is also an unknown parameter in the error equation. The purpose of the parametric model is to estimate each of these major error components using the available RMSD observations. The following sections derive empirical expressions for each of these error components.

As wind speed increases, the strength of forward scattered power decreases, resulting in reduced GNSS-R sensitivity to changes in wind speed and increased sensitivity to noise and errors in calibration of σ_0. This is consistent with the observed increase in RMSD with increasing wind speed. RMSD also varies inversely with RCG. RCG directly affects the strength of the reflected signal received from the Earth surface, whereas the level of additive noise generated by a radar receiver is independent of the reflected signal strength. As a result, as RCG decreases, SNR decreases and RMSD increases. These two fundamental dependencies are illustrated conceptually in Fig. 6, which qualitatively describes the expected variation in RMSD as a function of wind speed and RCG.

In order to quantify the expected dependence of RMSD on wind speed and RCG in the measurements, we first segment the measurements into distinct RCG and wind speed bins. In total, ten different RCG bins are chosen such that each bin contains approximately 10% of the samples. This is done in order to maintain statistical significance of the error model over the operating range of RCG and wind speed values. Fig. 7 is

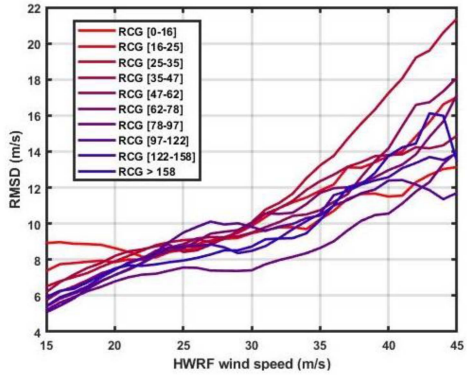

Fig. 7. RMSD as a function of wind speed for different partitions of RCG.

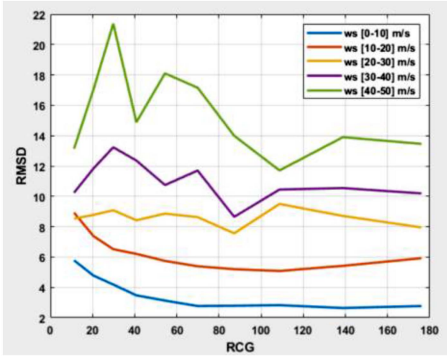

Fig. 8. RMSD as a function of RCG for different partitions of wind speed.

a plot of RMSD on the y-axis and HWRF wind speed on the x-axis and the different colors represent the ten different RCG bins with reddish hue associated with the lowest RCG ranges and bluish hue associated with the highest RCG ranges. It can be observed from this plot that the RMSD behavior is consistent with the expected dependence shown in Fig. 6(a). Next, the RMSD is plotted as a function of RCG in Fig. 8, with the different curves representing different wind speed ranges. The behavior of RMSD is again consistent with the expected dependence plotted in Fig. 6(b).

The error model shown in (10) can be expressed as a parametric version given by

$$\epsilon_u^2 = a + \left| \frac{\partial \text{GMF}}{\partial u} \right|_u \Bigg|^{-2} (b + F(G)) \tag{11}$$

where the parameter a is a constant bias that represents $\epsilon^2_{u,\text{others}}$. It is referred to as a constant bias in the sense that it is not a function of either wind speed or gain. Next, the parameter b is also a constant bias that represents the error component $\epsilon^2_{\sigma_0,\text{others}}$. The third and the final component of the parametric error model is $F(G)$, which represents the error component due to antenna gain, $\epsilon^2_{\sigma_0,G}$. Values for the three components, a, b, and $F(G)$ are empirically derived from observed RMSD behavior, as described in the upcoming sections.

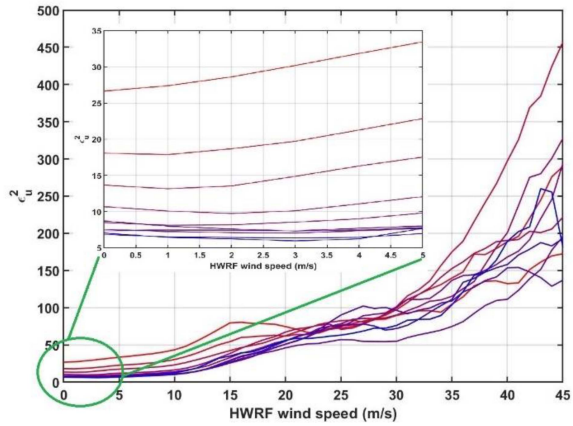

Fig. 9. Square of RMSD due to nonwind effects at very low wind speeds and high RCG values. Each curve corresponds to a different range of RCG, covering the range from 0–16(most red) to 158–200 (most blue).

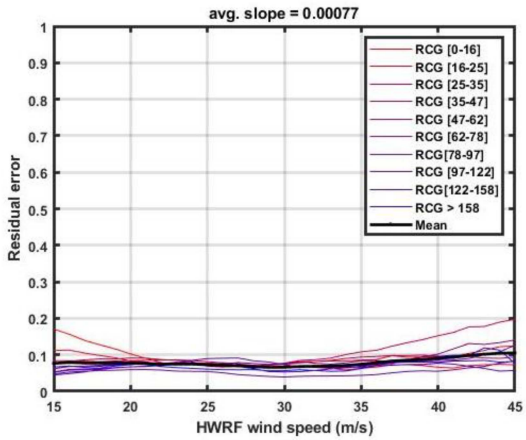

Fig. 10. Residual error after removing the effects of bias and GMF.

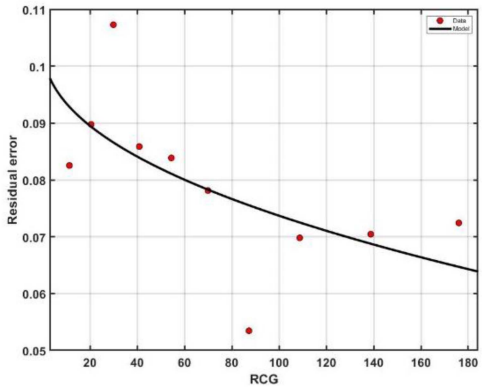

Fig. 11. Estimating the error due to antenna gain.

A. Error Component Due to Nonwind Geophysical Effects

The first error component that needs to be deduced from the observed RMSD is the parameter "a" in (11). This component accounts for errors due to other geophysical effects, such as swell or rain, error in the reference winds, and errors associated with inexact colocation between the measured and reference wind speeds. The parameter "a" also accounts for the effect of biases in the other error components. Since all these error sources are assumed independent of wind speed or RCG, this parameter is considered a constant term in the error model.

We estimate the "a" term in the model by considering only RMSD values in very low wind conditions and at very high RCG values. Since RMSD increases with wind speed and decreases with RCG, this combination of conditions will tend to reduce the dependence of RMSD on either wind speed or RCG and the residual RMSD that still remains is assigned to "a." Specifically, we examine the RMSD value using only samples with wind speed < 5 m/s and considering only the highest RCG values. Fig. 9 shows the behavior of the RMSD versus wind speed partitioned by RCG. The deep blue curve represents the highest RCG values and they converge to $\sim 6.7\ \mathrm{m/s^2}$ at the lowest wind speed values. This is the value assigned to the "a" term in the error model.

B. Residual Error Estimation

Once the bias error due to nonwind/nongain effects is estimated, the other components of the parametric error model can be determined by considering the residual. The residual error is defined by rewriting (11) as

$$\frac{\epsilon_u^2 \ - \ \mathrm{a}}{\left|\frac{\partial \mathrm{GMF}}{\partial u}\right|_u\Big|^{-2}} = (b + F(G)) \tag{12}$$

where the left-hand side of the expression consists of ϵ_u, a, and $|\frac{\partial \mathrm{GMF}}{\partial u}|_u|$, which are all known. This is referred to as the residual error and is independent of the effects of wind and other geophysical effects. It is illustrated by plotting the residual error with respect to wind speed, as shown in Fig. 10. It can be seen that the residual error curves have a near-zero slope across the entire wind speed range indicating that the geophysical components of error have been properly accounted for in the error model.

This residual error now represents only the engineering calibration errors and the error due to antenna gain, as given by the right-hand side of (12). An empirical power series model is fit to the residual error as a function of RCG. The power series fit has the form

$$\text{Residual error } = (b + F(G)) \ = \ b + p_1 (G)^{p_2}. \tag{13}$$

The model parameters b, p_1, and p_2 can be estimated from the measurements by least squares regression. A plot of the residual error as a function of RCG and the corresponding model fit is shown in Fig. 11. Based on this model, the parameter b is found to be 0.1039 and the error component due to antenna gain is given by

$$F(G) = (-0.003613)\,(G)^{0.4609}. \tag{14}$$

This completes the construction of all components of the error model. The final, full parametric model is given by

$$\epsilon_u^2 = 6.7 + \left|\frac{\partial GMF}{\partial u}\right|_u\Big|^{-2} \left(0.1039 + (-0.003613)\,(G)^{0.4609}\right). \tag{15}$$

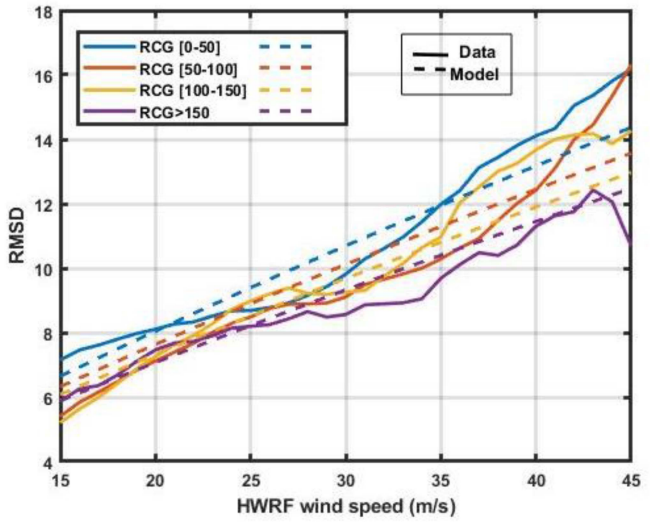

Fig. 12. Error model versus observation.

The behavior of the parametric error model as a function of wind speed and RCG is compared with that of the observed RMSD in Fig. 12. The solid lines represent the measurements and the dotted lines represent the model. From Fig. 12, it can be seen that the model does a fair job of representing the measurements in the wind speed range 15–45 m/s and also follows the expected trend of dependence on wind speed and RCG, as discussed in Fig. 6. We note that there is some deviation of the error model from the observations at the low wind speed (<15 m/s) and high wind speed (>40 m/s) ranges. This may be due to other error components not properly accounted for in this model.

C. Uncertainty Analysis

The residual uncertainty in the derived error model given by (15) is assessed here. Model uncertainty can be deconstructed into two primary source: intrinsic and extrinsic. Intrinsic uncertainties are associated with uncertainties in the estimated empirical parameters of the model. These uncertainties are evaluated by performing a simplified N-subset analysis. The total dataset used to derive the parameters is divided into N-subsets (N = 10 in this case). Each subset consists of 50% of the samples from the original dataset selected using uniform random sampling. Each of these subsets is then used to estimate the empirical parameters of the model. The intrinsic uncertainty in each estimated parameter is defined as the standard deviation of the ten estimated parameter values. These uncertainties are listed in Table IV. The ten simulation runs of the model are then used to determine the overall intrinsic uncertainty of the model. This is shown in Fig. 13. From the error surface in Fig. 13, it is observed that the maximum intrinsic error is about 0.45 m/s and the mean intrinsic error is about 0.12 m/s in the wind speed range of 15–45 m/s. The extrinsic uncertainty is determined by considering the absolute difference between the model and the observed RMSD at different wind and RCG values, as given by

$$\text{Extrn. Uncert.} = abs\left(\text{RMSD}_{\text{obs}} - \text{RMSD}_{\text{model}}\right) \quad (16)$$

TABLE IV
ESTIMATED MODEL PARAMETER VALUES

Model parameters	Best fit estimates	Uncertainty (+/-)
a (nonwind geophysical effects)	6.7747	0.0864
b (engineering calibration bias)	0.1039	0.0832
p1 (Gain parameter)	-0.003613	0.0858
p2 (Gain parameter)	0.4609	0.2412

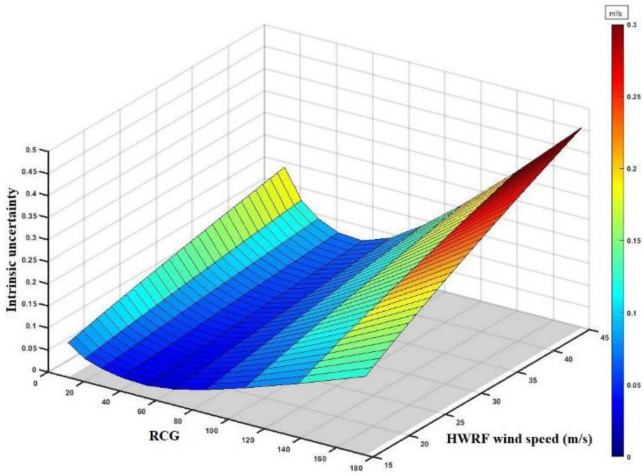

Fig. 13. Intrinsic uncertainty in the parametric error model as a function of wind speed and RCG.

where RMSD_{obs} is the empirical RMSD between CYGNSS and HWRF, and $\text{RMSD}_{\text{model}}$ is the RMSD estimated using the parametric model given by (15). The extrinsic uncertainty error surface is plotted in Fig. 14 over a range of values of wind speed and RCG. The average extrinsic uncertainty in the model, spanning the entire wind speed and RCG range, is ∼0.9 m/s and the median uncertainty is ∼0.7 m/s. The extrinsic uncertainty is highest at low RCG and high wind speed values. This may be due to erroneous assumptions about the dominant sources of error in this operating regime. Notably, however, the error model performs well with high RCG values and high wind speeds. This is the operating regime that is most relevant to the question of how much the high retrieval performance would be improved by using a bigger (higher gain) antenna.

In summary, intrinsic uncertainty characterizes how well we know the model and extrinsic uncertainty characterizes how well the model fits the observations. These two uncertainties are combined in a sum of squares fashion to derive the total uncertainty in the model. The total uncertainty is shown in Fig. 15. Note that the total uncertainty does not have a strong

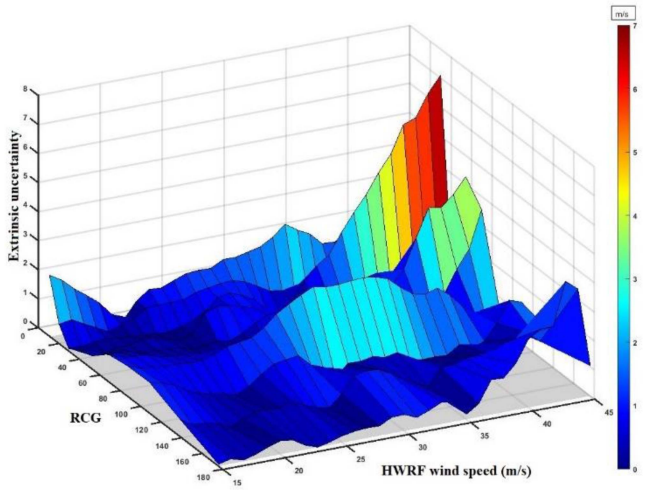

Fig. 14. Extrinsic uncertainty in the parametric error model as a function of wind speed and RCG.

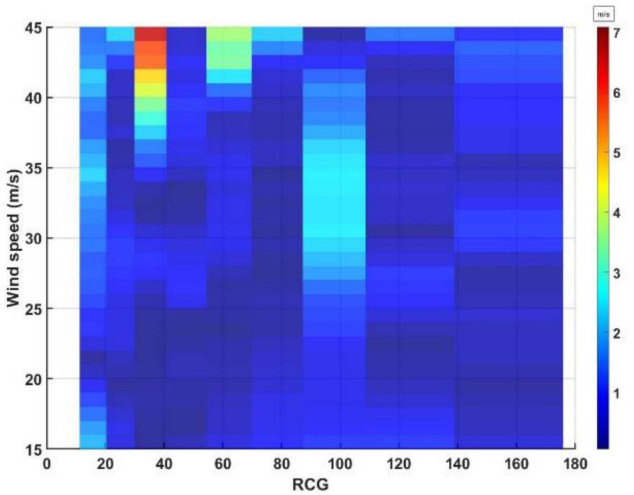

Fig. 15. Total uncertainty in the parametric error model as a function of wind speed and RCG.

dependence on either wind speed or RCG. It is, however, dominated by extrinsic uncertainties as their magnitudes are an order of magnitude greater than the intrinsic uncertainties. The mean total uncertainty is ∼0.96 m/s and the median is 0.67 m/s. The highest value of the uncertainty occurs at the highest wind speed and lowest RCG values.

D. Antenna Gain Estimation for Hurricane Force Winds

Parametric error models can provide guidance in the design of future missions. In the specific example developed here, the model can be used to determine the minimum science antenna gain needed to achieve a particular RMSD for wind speed retrieval.

As an example, for a future mission design study, we use the error model to determine the minimum antenna gain needed to distinguish a Category 4 hurricane from a Category 5. According

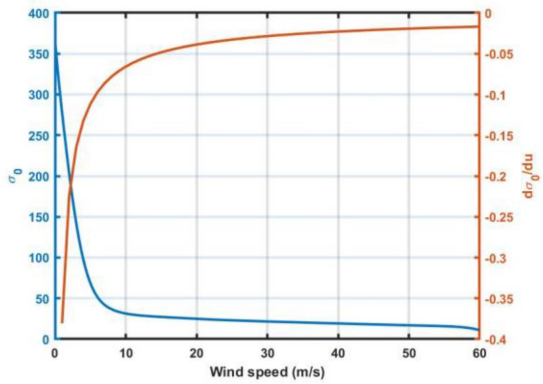

Fig. 16. Wind speed GMF and its slope ($d\sigma_0/du$) as a function of wind speed for a typical 30° angle of incidence.

to the Saffir–Simpson scale, a Category 4 hurricane is defined by a maximum sustained wind speed (Vmax) in the range 58.1–69.7 m/s and a Category 5 hurricane is defined when Vmax is greater than 70 m/s. Therefore, a simplistic approach to differentiate a measurement of Cat 4 wind from a Cat 5 wind is to have a maximum RMSD in the retrieval of ∼12 m/s. Evaluation of (15) requires the knowledge of the slope of the GMF with respect to wind speed. Both the GMF and its slope are shown in Fig. 16 as a function of wind speed for a typical observing angle of incidence of 30° [19].

Let the wind speed of interest be 64 m/s (the midpoint of the Category 4 range). From Fig. 16, the slope of the GMF at this wind speed is $\sim$0.016 $\mathrm{m/s^{-1}}$. Solving (15) for the value of antenna gain at which the RMSD equals 12 m/s at a wind speed of 64 m/s results in a gain value of 20.5 dBi. The calculation is done for a typical CYGNSS incidence angle of 30°, and for this incidence angle, the GPS satellite range is approximately 2.3×10^4 km and the CYGNSS range is approximately 588.9 km. The uncertainty in the RMSD at this value of RCG and windspeed is $+/-$ 1 m/s. This introduces an uncertainty into the determination of the required antenna gain. The uncertainty of $+/-1$ m/s in the error model corresponds to a range for the required antenna gain of 19.6 to 21.3 dBi. The effective area required to realize this gain value is approximately 0.6 m $\times$ 0.6 m. The current CYGNSS science antenna is a 3×2 patch array with an approximate effective area of 0.3 m $\times$ 0.3 m. This implies that an antenna 4 times the current size would be needed to achieve this goal (equivalently a 6×4 patch array).

V. Discussions

The motivation for developing a parametric error model for CYGNSS is to characterize the limitations of the existing system and to develop the necessary tools for translating the geophysical retrieval performance to engineering parameter requirements. This work has specifically focused on hurricane winds with an interest in identifying a suitable science antenna requirement that can capture such high winds. Although we have tried to model as many primary sources of errors as possible, due to the limited number of observations in the hurricanes, several assumptions

have been made. For instance, assuming a constant bias to represent the nonwind geophysical effects, such as rain and swell, might not be a good representation as each of these is known to have a wind speed dependence. Although the CYGNSS Level 2 retrievals are corrected for the effect of swell and are assumed to be free of rain contaminations, such effects can have a second order influence on the observations. These could possibly be a reason for the increase in residual error in the error model at low and high wind speeds (as observed in Fig. 10). Nonetheless, the residual error slope is close to zero through most of the wind speed range of our analysis.

Another interesting consideration is the influence of the magnitude of the antenna gain on the antenna gain uncertainty. These two factors are highly anticorrelated. A high gain magnitude will have lower gain uncertainty and vice versa. These two effects cannot be separated apart from the available information and, hence, get rolled into a single error component.

Many assumptions have been made while developing a parametric error model to best represent the observations. The error model developed in this work is intended to be used as a guide for future GNSS-R mission planning. Based on this error model, the minimum antenna gain required to distinguish a Cat 4 magnitude wind from a Cat 5 wind is found to be at 20.5+/− 0.9 dBi. As seen in the previous section, this would mean a single aperture antenna that is 4 times the size of the current CYGNSS antenna. A single higher gain antenna would have a narrower beamwidth and a narrower field of view for specular point reflections, resulting in a possible reduction in the number of reflections captured. There are several antenna design choices to mitigate this. One is the use of an active beam-forming antenna. This introduces a tradeoff between fewer reflections and a more complex, higher risk antenna with higher a power requirement. A second mitigation approach is the use of fixed multibeam antennas.

Past, current, and planned future missions use antennas with peak gain in the range of 11–15 dBi. This study shows that this range is insufficient for reliably distinguishing hurricane force winds in major storms. Although this study has focused on tuning the error model to the antenna gain, such error models can be developed from observations, for tuning different engineering parameters to suit the geophysical variable of interest measured from GNSS-R systems. This kind of sensitivity study of retrieval characteristics with respect to engineering parameters can be a useful tool for planning of future instrument requirements to meet new geophysical measurement objectives.

VI. Conclusion

A parametric error model for the CYGNSS retrieval of hurricane force winds is presented. This model is constructed based on near coincident matchups with the HWRF reanalysis winds from 56 hurricanes over a period of four years. The model is developed by combining a top–down empirical parameterization with a bottom–up error model functional form that includes an explicit dependence on receive antenna gain. The RMSD of CYGNSS retrieved winds relative to HWRF reference winds is due to the compounding effects of several error components such as those due to nonwind geophysical effects (swell, rain, etc.), error in colocation between CYGNSS and HWRF, error in instrument calibration, and other engineering calibration errors. The different error components have been parametrically modeled based on observations. This error model fairly replicates the dependence of actual observations on wind speed and RCG in the wind speed range 15–45 m/s. The average uncertainty in the model is ∼0.9 m/s over the entire range and the median uncertainty is ∼0.7 m/s. There are some regions in the wind speed/RCG space where the error model performs significantly more poorly, specifically at combinations of high wind speeds and low RCG values. This could be attributed to erroneous assumptions about the dominant sources of error in this measurement regime. It may also be due to the sparsity of observations in this regime. A simple example of estimation of minimum antenna gain required to distinguish Cat 4 hurricane winds from Cat 5 hurricane winds was evaluated using the developed model. It was found that at least 20.5 +/−0.9 dBi of gain is required for this application, which translates to an effective area that is 4 times that of the current science antenna on the CYGNSS observatories. Most/all the current and the upcoming spaceborne GNSS-R missions have antennas much lower than this expected value indicating that they might not have the ability to capture such high winds.

Acknowledgment

The authors thank Dr. B. Annane, Senior Research Associate III, with the University of Miami for the HWRF reference datasets.

References

[1] V. U. Zavorotny and A. G. Voronovich, "Scattering of GPS signals from the ocean with wind remote sensing application," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 2, pp. 951–964, Mar. 2000, doi: 10.1109/36.841977.

[2] S. J. Katzberg, R. A. Walker, J. H. Roles, T. Lynch, and Peter G. Black, "First GPS signals reflected from the interior of a tropical storm: Preliminary results from Hurricane Michael," *Geophysical Res. Lett.*, vol. 28, no. 10, pp. 1981–1984, May 2001, doi: 10.1029/200GL012823.

[3] 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 J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 1, pp. 107–122, Jan. 2019, doi: 10.1109/JSTARS.2018.2856588.

[4] Q. Yan and W. Huang, "Sea ice Remote sensing using GNSS-R: A review," *Remote Sens.*, vol. 11, no. 21, Nov. 2019, Art. no. 2565, doi: 10.3390/rs11212565.

[5] C. Chew and E. Small, "Description of the UCAR/CU Soil moisture product," *Remote Sens.*, vol. 12, no. 10, May 2020, Art. no. 1558, doi: 10.3390/rs12101558.

[6] M. P. Clarizia, C. Gommenginger, S. Gleason, C. Galdi, and M. Unwin, "Global navigation satellite system-reflectometry (GNSS-R) from the U.K.-DMC satellite for remote sensing of the ocean surface," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2008, pp. I-276–I-279, doi: 10.1109/IGARSS.2008.4778847.

[7] C.S. Ruf et al., "A new paradigm in earth environmental monitoring with the CYGNSS small satellite Constellation," *Sci. Rep.*, vol. 8, 2018, Art. no. 8782, doi: 10.1038/s41598-018-27127-4.

[8] *CYGNSS Handbook*. Ann Arbor, MI, USA: Michigan Publishing Services, 2022, doi: 10.3998/mpub.12741920.

[9] M. P. Clarizia and C. S. Ruf, "Wind speed retrieval algorithm for the cyclone global navigation satellite system (CYGNSS) mission," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4419–4432, Aug. 2016, doi: 10.1109/TGRS.2016.2541343.

[10] C. S. Ruf and R. Balasubramaniam, "Development of the CYGNSS geophysical model function for wind speed," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 1, pp. 66–77, Jan. 2019, doi: 10.1109/JSTARS.2018.2833075.
[11] F. Saïd, Z. Jelenak, J. Park, and P. S. Chang, "The NOAA track-wise wind retrieval algorithm and product assessment for CyGNSS," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 4202524, doi: 10.1109/TGRS.2021.3087426.
[12] CYGNSS, Aug. 5, 2018, "CYGNSS level 3 storm centric grid science data record version 1.0," Cyclone Global Navigation Satellite System, doi: 10.5067/CYGNS-L3S10.
[13] CYGNSS, 2022, "CYGNSS level 3 Climate Data record version 1.2," Cyclone Global Navigation Satellite System, doi: 10.5067/CYGNS-L3C12.
[14] M. P. Clarizia, V. Zavorotny, D. Mc Kague, and C. Ruf, "Level 2 wind speed retrieval algorithm theoretical basis document," CYGNSS Project, UM Doc. 148-0138, Rev 6, Univ. of Michigan, Ann Arbor, MI, USA, 2020.
[15] R. Balasubramaniam and C. S. Ruf, "Assessing the sensitivity of GNSS-R signals to ocean surface winds near hurricanes using CYGNSS measurements," in *Proc. 103rd AMS Annu. Meeting*, Denver, CO, USA , 2023, Paper 2747.
[16] S. Gopalakrishnan et al., "Hurricane Weather Research and Forecasting (HWRF) model: 2011 scientific documentation," L. Bernardet, ed., 2011.
[17] S. Gleason, C. S. Ruf, A. J. O'Brien, and D. S. McKague, "The CYGNSS level 1 calibration algorithm and error analysis based on on-orbit measurements," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 1, pp. 37–49, Jan. 2019, doi: 10.1109/JSTARS.2018.2832981.
[18] C. S. Ruf, S. Gleason, and D. S. McKague, "Assessment of CYGNSS wind speed retrieval uncertainty," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 1, pp. 87–97, Jan. 2019, doi: 10.1109/JSTARS.2018.2825948.
[19] M Asgarimehr, J Wickert, and S. Reich, "Evaluating impact of rain attenuation on space-borne GNSS reflectometry wind speeds," *Remote Sens.*, vol. 11, no. 9, 2019, Art. no. 1048, doi: 10.3390/rs11091048.
[20] R. Balasubramaniam and C. Ruf, "Characterization of rain impact on L-Band GNSS-R ocean surface measurements," *Remote Sens. Environ.*, vol. 239, 2020, Art. no. 111607.
[21] A Manaster, L Ricciardulli, and T. Meissner, "Tropical cyclone winds from WindSat, AMSR2, and SMAP: Comparison with the HWRF model," *Remote Sens.*, vol. 13, no. 12, 2021, Art. no. 2347, doi: 10.3390/rs13122347.

Rajeswari Balasubramaniam (Member, IEEE) received the B.E. degree in electrical and electronics engineering from Anna University, Chennai, India, in 2014, the M.E. degree in geo-informatics with a specialization in remote sensing and image processing from the Indian Institute of Space Science and Technology, Thiruvananthapuram, India, in 2016, the M.S. degree in electrical engineering and computer science and the Ph.D. degree in remote sensing from the University of Michigan, Ann Arbor, MI, USA, in 2018 and 2020, respectively.

She is currently a Research Scholar with the Department of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor. Her current work focuses on the development of hurricane wind retrieval algorithms and improving various aspects of calibration and validation for the NASA Cyclone Global Navigation Satellite System mission. Her research interests include GNSS-reflectometry, statistical signal processing, remote sensing inversion theory, data analysis, and microwave engineering.

Christopher S. Ruf (Life Fellow, IEEE) received the B.A. degree in physics from Reed College, Portland, OR, USA, in 1982, and the Ph.D. degree in electrical and computer engineering from the University of Massachusetts at Amherst, Amherst, MA, USA, in 1987.

He is currently a Professor of atmospheric science and space engineering with the University of Michigan, Ann Arbor, MI, USA, and a Principal Investigator of the NASA Cyclone Global Navigation Satellite System mission. He has worked previously with Intel Corporation, Hughes Space and Communication, the NASA Jet Propulsion Laboratory, and Penn State University. His research interests include GNSS-R remote sensing, microwave radiometry, atmosphere and ocean geophysical retrieval algorithm development, and sensor technology development.

Dr. Ruf is a Member of the American Geophysical Union, the American Meteorological Society, and the Commission F of the Union Radio Scientifique Internationale. He is a former Editor-in-Chief of IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING and has served on the editorial boards of *Radio Science* and the *Journal of Atmospheric and Oceanic Technology*. He was the recipient of four NASA Certificates of Recognition and seven NASA Group Achievement Awards, 1997 TGRS Best Paper Award, 1999 IEEE Resnik Technical Field Award, 2006 IGARSS Best Paper Award, and 2014 IEEE GRSS Outstanding Service Award.