Wind Speed Retrieval Algorithm for the Cyclone Global Navigation Satellite System (CYGNSS) Mission

Maria Paola Clarizia, *Member, IEEE*, and Christopher S. Ruf, *Fellow, IEEE*

***Abstract*—A retrieval algorithm is presented for the Level 2 ocean surface wind speed data product of the Cyclone Global Navigation Satellite System (CYGNSS) mission. The algorithm is based on the approach described by Clarizia *et al*., 2014. The approach is applied to the specific orbital measurement geometry, antenna, and receiver hardware characteristics of the CYGNSS mission. Several additional processing steps have also been added to improve the performance. A best weighted estimator is used to optimally combine two different partially correlated estimates of the winds by taking their weighted average. The optimal weighting dynamically adjusts for variations in the signal-to-noise ratio of the observations that result from changes in the measurement geometry. Variations in the incidence angle of the measurements are accounted for by the use of a 2-D geophysical model function that depends on both wind speed and incidence angle. Variations in the propagation time and signal Doppler shift at different measurement geometries affect the instantaneous spatial resolution of the measurements, and these effects are compensated by a variable temporal integration of the data. In addition to a detailed description of the algorithm itself, the root-mean-square wind speed retrieval error is characterized as a function of the measurement geometry and the wind speed using a detailed mission end-to-end simulator.**

***Index Terms*—Delay–Doppler (DD) map (DDM), geophysical model function (GMF), GNSS reflectometry, tropical cyclones (TCs), wind speed.**

I. Introduction

THE National Aeronautics and Space Administration's (NASA) Earth Venture Cyclone Global Navigation Satellite System (CYGNSS) is a spaceborne mission focused on tropical cyclone (TC) inner core process studies [1]. CYGNSS is designed to measure ocean surface wind speed in all precipitating conditions (including those experienced in the TC eyewall) and with sufficient frequency to resolve the genesis and rapid intensification phases of the TC life cycle. CYGNSS exploits the ability of the L1 GPS signal at 1575 MHz to penetrate through rain and the fact that the bistatic scattering cross section at L-band is sensitive to changes in wind speed over the full dynamic range of wind speeds experienced in TCs up to 70 m/s [2].

Manuscript received December 19, 2014; revised July 1, 2015, October 20, 2015, and January 18, 2016; accepted February 28, 2016. Date of publication April 6, 2016; date of current version June 1, 2016.

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

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2016.2541343

The GPS reflectometry (GPS-R) receiver on board CYGNSS produces a 2-D image of the diffuse scattered power from the surface in the vicinity of the specular point (SP), as a function of time delay and Doppler frequency shift, known as the delay–Doppler (DD) map (DDM) [3]–[5], [7], [8], [11], [13], [17]. Delays are usually expressed in units of chips, where 1 chip is about 1 μs and represents the duration of the single rectangular pulse of the GPS coarse/acquisition (C/A) code [30]. Each DDM pixel is obtained through a cross-correlation of the received scattered GPS signal with a locally generated replica of the C/A code of the transmitted signal, for the DD coordinate corresponding to that pixel. Such coherent cross-correlation is usually performed over 1 ms [8], [11] and hence is heavily affected by speckle noise, so a number of incoherent accumulations of consecutive cross-correlation values are necessary to reduce the noise level. A typical incoherent accumulation time for DDM measured spaceborne is 1 s. The peak of a DDM usually represents the power coming from the SP [see Fig. 1(a)]. The typical "horseshoe shape" of DDMs represents the power coming from an area around the SP referred to as the glistening zone (GZ), whose extent depends on the sea surface roughness. The horseshoe shape is related to the space-to-DD coordinate transformation that is imposed by the cross-correlation for different pairs of delay and Doppler frequency values in the GNSS-R signal processing chain. The power scattered by the sea surface is mapped in the DD domain through lines of constant delay (isodelay ellipses) and constant Doppler shift (iso-Doppler parabolas) that can be traced on the Earth surface, as shown in Fig. 1(a). However, two important aspects of such transformation have to be considered. One, known as space ambiguity, refers to the fact that a single DD coordinate [τ_1, f_1 in Fig. 1(a)] typically corresponds to two different points on the surface [P and Q in Fig. 1(a)].[1] The other effect is related to the Woodward ambiguity function (WAF), for which a given DDM pixel contains the weighted combination of the scattering contribution from that pixel plus the neighbor pixels in space. This is shown in Fig. 1(b), where the two intersections between an isodelay annulus centered at 10 chips, and of width equal to 2 chips, and an iso-Doppler band centered at 2 kHz, and of width equal to 2 kHz, are highlighted in the spatial

[1]There is an exception, represented by the so-called ambiguity-free line. This is the transverse axis of the parabolic iso-Doppler lines, where the intersection between an iso-Doppler line and an isodelay line is only given by one point.

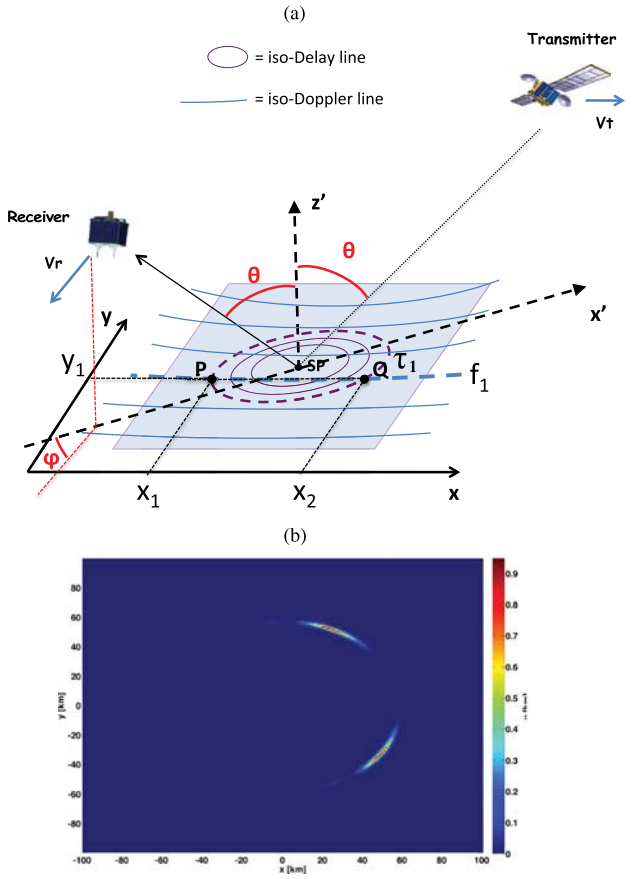

Fig. 1. (a) Example of SP on the sea surface. The incidence angle θ where the specular reflection occurs is shown in red along with the azimuth angle φ, defined as the angle between the receiver velocity vector and the x'-axis of the specular $(x'-z')$ plane that contains the transmitter, the receiver, and the SP. The configuration of (purple) isodelay ellipses and (blue) iso-Doppler parabolas in space is also illustrated, with the dashed isodelay line and dashed iso-Doppler line corresponding to DD coordinates (τ_1, f_1) and to spatial coordinates P and Q. (b) Effect of WAF illustrated for two areas in space from which the GPS reflections are combined into a single pixel in the DDM. The effect of WAF is illustrated as weights (colors) applied to the two areas.

domain. The widths chosen for the isodelay annulus and the iso-Doppler band correspond to the width of the main lobe of the WAF, along the delay and Doppler dimensions. The colors in Fig. 1(b) represent the weights applied by the WAF to each pixel within the two intersections, prior to their combination into a single DD pixel.

To date, two main approaches have been used for estimating sea surface winds from DDMs. The first approach relies on fitting, usually in a least square sense, the measured DDMs to those simulated with a forward scattering model (e.g., 3), to retrieve either a sea surface roughness parameter [e.g., mean-square slope (MSS)] [6]–[8] or surface wind speed [2], [9]–[11].

The second approach uses a so-called DDM observable, a quantity derived from the DDM which depends on the underlying wind and roughness. Examples of DDM observables are the DDM volume [14], the average scattered signal around the SP [DDM average (DDMA)] [12], the bistatic radar cross section (RCS) measured at the SP [13], the volume of DD pixels of normalized DDMs around the SP [14], [15], the leading edge slope (LES) or trailing edge slope (TES) of integrated delay waveforms (IDW) [9], [12], and different measures of the extent of diffuse scattering within the DDM [16]. The approach is to calculate these quantities from the measured DDMs, regress them against the match-up winds measured locally by some other sources or instruments, quantify the relationship between the observable and the wind via an empirical lookup table (LUT), and incorporate the LUT into a retrieval algorithm [12]. In this case, there is no intermediate step of MSS calculation since the wind estimate is derived directly from the measured observable.

The wind speed retrieval algorithm presented here uses the second approach, and it is an adaptation of an algorithm that was developed for use by the United Kingdom Disaster Monitoring Constellation (UK-DMC) mission [8], [11], [13], [17]. This type of wind retrieval algorithm is often used in scatterometry, where empirical geophysical model functions (GMFs) are derived from a large collocation study between observed measurements and *in situ* buoy and/or numerical weather prediction (NWP) model data [21]–[29]. Similarly, CYGNSS will rely on a large data set of collocated wind speed information from different sources (including buoys, model outputs, aircraft measurements, and satellite crossovers) that will be used to construct a robust empirical GMF for each observable. Improvements to the algorithm, relative to [12], are made to optimize its performance for the CYGNSS mission. The algorithm is implemented and tested using a large and diverse population of simulated DDMs, generated using the CYGNSS end-to-end simulator (E2ES), whose details can be found in [19]. These DDMs represent the expected characteristics of CYGNSS on-orbit measurements with respect to wind speed dynamic range, measurement geometry, received signal strength, and thermal and fading noise statistics. The retrieval algorithm uses as its input calibrated DDMs of the radar scattering cross section, which are referred to as the CYGNSS Level 1b (L1b) data product. They are, in turn, derived from Level 1a (L1a) calibrated DDMs of received power. The CYGNSS L1a and L1b calibration algorithms are described in [18]. The ocean surface wind speed produced by the algorithm presented here is the CYGNSS Level 2 (L2) data product. The retrieval algorithm is characterized with respect to spatial resolution and retrieval uncertainty and is shown to meet the CYGNSS mission baseline requirements in both respects.

This paper is organized as follows. Section II describes the overall wind speed retrieval algorithm, summarizing the approach presented in [12] and describing the algorithm steps in detail. Section III illustrates the application of the retrieval algorithm to a data set obtained from a simulation of the full life cycle of a TC. Section IV contains an analysis of the performances of the retrieval algorithm using the simulated data set, and Section V discusses the results and summarizes the main conclusions.

II. Wind Speed Retrieval Algorithm

The main objective of the retrieval algorithm presented here is to meet the baseline requirements imposed for the CYGNSS

mission [1]. These establish the spatial resolution of the retrieved wind speed to be 25 km $\times$ 25 km and the wind speed retrieval uncertainty to be larger than 2 m/s or 10% of the measured wind over the wind speed dynamic range of 2–70 m/s.

The steps of the L2 wind speed retrieval algorithm are given in detail in the subsequent paragraphs and are summarized here.

1) Two DDM "observables," the DDM average (DDMA) and the LES, are derived from L1b DDMs of RCS and DDMs of scattering area, which are generated as explained in [18].
2) Wind speeds are estimated from both DDMA and LES observables using a GMF, empirically constructed for each observable, that maps each observable value into a wind speed value.
3) Wind estimates from DDMA and LES are linearly combined using a best weighted estimator. The coefficients of the linear combination are computed using an adaptive covariance (AC) method.

Three main differences exist between the algorithm presented in [12] and the one presented here. The first is that only two out of five observables (DDMA and LES) are used here for the retrieval algorithm. The inability to obtain DDMs with CYGNSS with an incoherent accumulation time lower than 1 s rules out the possibility to use observables based on the variance of the scattered signal (i.e., DDMV and ADDMV), while the TES observables cannot be adopted here because of spatial resolution constraints. The second difference is that the approach shown in [12] computes a unique set of coefficients for the linear combination of retrievals, derived from a single covariance matrix of the retrieval errors. An approach like this is well suited when the data set is small and its performance statistics are stationary, as is the case for UK-DMC. In the case of CYGNSS, the approach described in [12] is applied in an adaptive way, by assigning different coefficients for the linear combinations for different groups of samples, characterized by different degrees of correlation. This alternative and innovative implementation of the best weighted estimator exploits in a better way the strong dependence between the performance improvement and existing correlation between retrieval errors.

The third difference is that the algorithm is applied here to a much larger DDM data set spanning the full dynamic range of wind speeds, as opposed to the very small UK-DMC data set used in [12], covering winds ranging only from 2 to 12 m/s. Hence, this paper also represents a convincing validation of the retrieval algorithm, particularly at high wind speeds, where the performances of GPS-R are known to be more critical due to the reduction in scattered power.

A. DDMA and LES Observables

The two observables used by the retrieval algorithm are the DDMA and the LES, which have been introduced in [12]. The DDMA is the average of the L1b DDM of RCS over a given delay/Doppler range window around the SP. The DDMA is calculated from the DDM after the additive noise floor has been removed. The noise floor is estimated from the DDM in a region where there is no signal (at negative delay values prior to that of the SP) and is subtracted from each DDM pixel.

The LES is the slope of the leading edge of the IDW. One-dimensional IDWs are derived from the 2-D DDMs by integration along their Doppler dimension over a specified range of Doppler frequencies [12], [32].

One important aspect of the computation of DDMA and LES observables from DDMs is the range of delay and Doppler values over which the observables are computed. The DDMA and LES observables are computed using the same delay/Doppler window, and the final choice of the delay and Doppler ranges has been established as [-0.25 0.25] chips in delay and [-1000 1000] Hz in Doppler, corresponding to 3 $\times$ 5 bins of the DDM. This choice involves a tradeoff between the improvement in signal-to-noise ratio (SNR) that results from averaging across more of the diffuse scattered signal, versus the accompanying degradation in spatial resolution that results from a larger region of integration. A detailed discussion of this tradeoff is given in Section III-B.

DDMA and LES were computed in [12] from the L1a DDMs of scattered power, while here, the observables are computed from L1b DDMs of RCS and DDMs of scattering area [18]. These are L1a DDMs to which the dependence on the transmitter and receiver antenna gain, range losses, and GPS transmit antenna power has been removed, leaving only the RCS term inside the DDM equation. The stringent resolution requirements, which, for the CYGNSS mission, is to retrieve winds with a spatial resolution of 25 km $\times$ 25 km, allow to approximate the receive antenna gain, range loss terms, and GPS transmit antenna power across the DDM subset used to compute the observables with constant values calculated at the specular reflection point [18]. DDMA and LES observables are computed from L1b DDMs of RCS and then normalized by the scattering area corresponding to the selected DD range over which they are computed. This normalization removes the dependence of the observables on the scattering area which also varies with geometry, to finally generate the L1b observables of NRCS (see [18] for details). Another type of correction, called Level 2a (L2a) correction, is subsequently estimated and used to generate a 2-D GMF function of both wind speed and incidence angle. This correction is explained in Section III-D1.

B. Wind Speed Retrieval From DDMA and LES

The winds are estimated from the observables using a GMF that represents the basis for the mapping from the DDM observable to the estimated wind speed. In practice, the GMF is precomputed from a subset of all the data called training data set (see Section III-D). The derivation of the GMF is presented in Section III-D. For each sample of either the DDMA or LES, a wind speed is estimated using the GMF function for the appropriate incidence angle of the sample. Following the approach adopted in scatterometry, where the GMF is a function of both geometry and wind speed, the GMF function can be thought of in this case as a family of functions that relate the observable to the wind speed, there being one GMF for each incidence angle. Once the proper GMF is selected, the estimation is done through the following: 1) linear interpolation, when the

observable value falls within the range of values spanned by the GMF, using the two GMF points before and after the observable value, or 2) extrapolation, when it is outside of that range, using the first or last three points of the GMF to compute the extrapolation slope. The first three points of the GMF have wind speeds of 0.2, 0.6, and 1 m/s, whereas the last 3 points have wind speeds of 45, 46, and 47 m/s.

It is worth noting that, in some cases, the values of the observables are negative, i.e., the noise floor, subtracted from the DDM, is larger than the signal in the DDM itself. These negative values are flagged as bad samples.

C. AC Estimator

The wind speed estimates from DDMA and LES can be combined together to produce an optimal estimate. This best weighted estimator exploits the degree of decorrelation present between the errors in the individual DDMA and LES estimates to minimize the rms error in the wind speed estimate. As shown in [12], the estimator is a linear combination of the individual estimates, with weighting obtained by requiring that the estimator be unbiased (i.e., the expected value of its retrieval is equal to the true quantity to be estimated) and by minimizing its variance. The derivation of the vector of coefficients $\boldsymbol{m}$ for the estimator is described in [12] and given by

$$\boldsymbol{m} = \left(\sum_{i=1}^{N} \sum_{j=1}^{N} c_{i.j}^{-1} \right)^{-1} \boldsymbol{C}^{-1} \mathbf{1} \tag{1}$$

where $\mathbf{1}$ is a vector of ones, $\boldsymbol{C}^{-1}$ is the inverse of the covariance matrix between the individual retrieval errors, and $c_{i,j}^{-1}$ are its elements. In practice, the correlation between DDMA and LES retrieval errors is found to vary as a function of the SNR of the measurements. For this reason, the covariance matrix assumed by the estimator is varied accordingly, in a manner referred to as AC. This approach consists of estimating a different covariance matrix and, therefore, a different pair of coefficients for the linear combination, for different ranges of the SNR. The SNR is estimated using a CYGNSS measurement parameter referred to as the range-corrected gain (RCG). The RCG is the receiver antenna gain at the SP divided by the range losses at the SP or

$$\Gamma_{\mathrm{SP}} = \frac{G_{\mathrm{SP}}^{R}}{\left(R_{\mathrm{SP}}^{\mathrm{Total}}\right)^{2}} \tag{2}$$

where Γ_{SP} represents the RCG, G_{SP}^{R} is the receiver gain at the SP, and $R_{\mathrm{SP}}^{\mathrm{Total}}$ is the total range loss from the transmitter to the surface and the surface to the receiver at the SP. A map of the RCG as a function of incidence and azimuth angles [defined in Fig. 1(a)] is shown in Fig. 2.

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 SNR, assuming that the noise floor is constant and the variations in signal strength are due to changes in the signal gain. Typical low, medium, and high values of RCG for CYGNSS are respectively $3 * 10^{-27}$, $10 * 10^{-27}$, and $20 * 10^{-27}$ m^{-4}. For brevity, we will often refer to RCG values of 3, 10, and 20, omitting the 10^{-27} term.

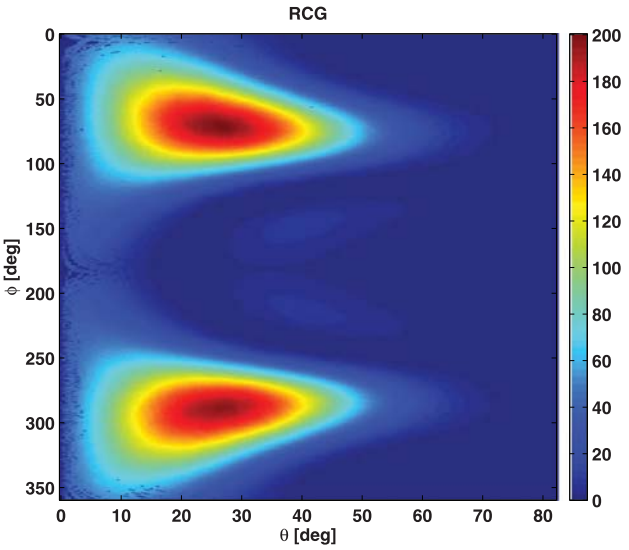

Fig. 2. Map of RCG normalized by a factor of 10^{-27}, as a function of incidence angle and azimuth angle.

Compared to [12, eq. (4.7)], the estimator with the AC approach included becomes

$$u_{\mathrm{MV}}^{i} = \boldsymbol{m}^{i} \cdot \boldsymbol{u}^{i} \; \Gamma_{\mathrm{low}}^{i} < \Gamma < \Gamma_{\mathrm{high}}^{i} \tag{3}$$

where i refers to each of the RCG intervals used to compute the coefficients in the linear combination, $\boldsymbol{m}$ is the vector of coefficients defined in (1), and $\boldsymbol{u}$ is the vector of wind speed estimates from DDMA and LES.

III. Wind Estimation From Simulated NR TC

The wind speed retrieval algorithm described in Section II is applied to a DDM data set simulated using the CYGNSS E2ES [19] and using a wind field that reproduces the full life cycle (genesis, development, and rapid intensification) of a TC, as input to the E2ES. The sections that follow first provide an overview of the DDM data set used in this paper, followed by an illustration of how the GMF is derived. A detailed discussion about how to choose the DD window for computing the observables is presented, and time averaging (TA) of the observables is introduced as a means of improving the performance when the native (snapshot) spatial resolution is better than that required for CYGNSS.

A. Description of Simulated DDMs

The population of simulated DDMs for this paper has been generated using the CYGNSS E2ES. The E2ES simulates DDMs using an advanced implementation of the GPS bistatic scattering model described in [3], which includes realistic thermal and speckle noise and whose simulated bistatic NRCS accounts for the nonuniformity of the wind field across the

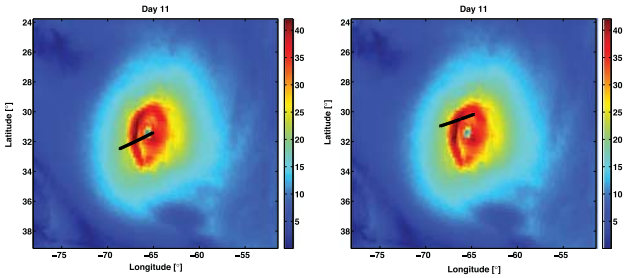

Fig. 3. Snapshots of wind fields of storm (in meters per second), simulated using the NR model, for day 11 of the simulation, with the CYGNSS specular track with the highest RCG values for its SPs, and as close as possible to the hurricane eye. The wind fields are represented here with a resolution of $\sim 0.04^\circ$ in both latitude and longitude.

spatial area that is used to generate the DDM. A detailed description of the E2ES characteristics and its implementation are given in [19]. DDMs have been simulated following the exact specifications for CYGNSS for DD range and resolution (see [18]). A very large variety of different but realistic CYGNSS geometries are generated through orbital simulations and are used to simulate the DDMs, which are then characterized by different incidence angles and different antenna gains at the SP.

The simulated DDM data set is collected from all eight CYGNSS observatories over a time frame of 13 days, during which a full life cycle of a TC has been simulated using the ECMWF/WRF Nature Run (NR) model [31]. Full details of these simulations can be found in [19]. A snapshot of two wind fields near the storm center, after rapid intensification, is shown in Fig. 3 for day 11 of the NR, along with CYGNSS specular tracks crossing the storm directly through the eye (left) and through the eyewall (right).

A DDM was generated for each second of acquired data and for those SPs acquired during an overpass of the satellites in the vicinity of the TC, resulting in more than 2 million DDMs over 13 days.

Each DDM is associated with a ground truth wind, computed as the 25 km $\times$ 25 km spatial average of the winds simulated from the NR model, centered at the SP. Fig. 4 shows the characteristics of the simulation geometries through the probability density function (pdf) of antenna gain (left), range losses (middle), and RCG (right). The pdfs of wind speeds and incidence angles are instead shown in Fig. 5. The percentage of wind speeds higher than 20 m/s is 0.8%.

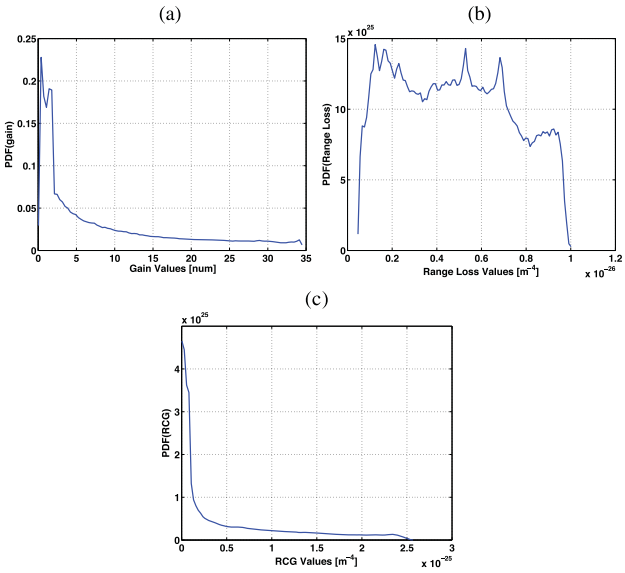

Fig. 4. PDFs of (a) antenna gain of receiver, (b) propagation range losses, and (c) RCG for the full 13 days of DDMs generated using the NR model.

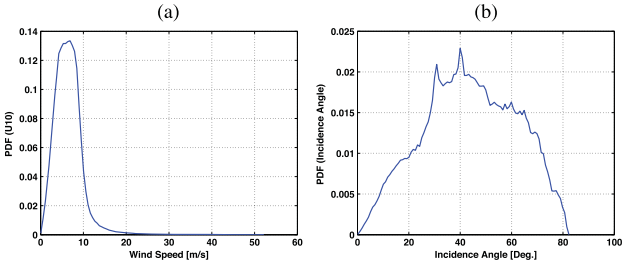

Fig. 5. PDFs of (left) wind speed and (right) incidence angle for the full 13 days of DDMs generated using the NR model.

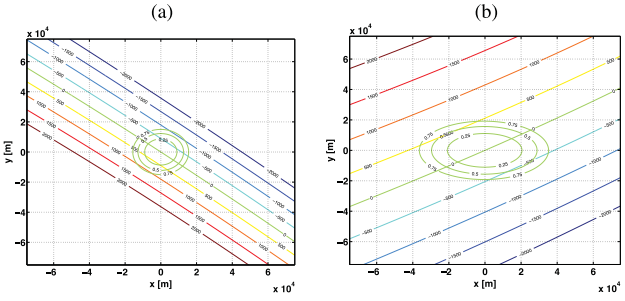

Fig. 6. (a) (Green) Isodelay ellipses (shown for a 0.25-chip step) and (colored lines) iso-Doppler parabolas (shown for a 500-Hz step), illustrated for an incidence angle of 16.5° and an azimuth angle of 229.7°. (b) (Green) Isodelay ellipses and (colored lines) iso-Doppler parabolas, illustrated for an incidence angle of 57.6° and an azimuth angle of 146.0°. The plane of incidence is in this case represented by the y-z plane.

B. Delay and Doppler Range Selection

The choice of delay and Doppler range for the computation of the observables depends on the requirement on spatial resolution and on the configuration of the isodelay and iso-Doppler lines across the GZ. The behavior of these lines depends on the scattering geometry of the measurement.

Fig. 6 shows two examples of different configuration of the isodelay and iso-Doppler lines for two different incidence and azimuth angles. The isodelay ellipses [see Fig. 6(a)] widen and stretch out with increasing incidence angles [i.e., going from Fig. 6(a) to (b)] without changing their orientation, so the geometrical parameter that mostly influences the configuration of the isodelay lines is the incidence angle. In the case of iso-Doppler lines instead, the spacing between them once again increases with increasing incidence angle, but they are also affected by the orientation of the receiver velocity vector

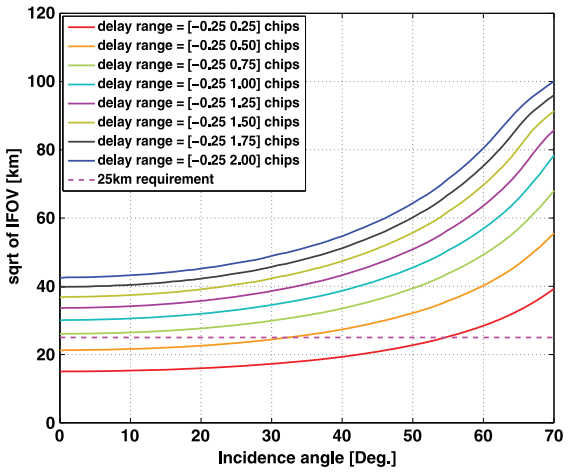

Fig. 7. Illustration of curves of square root of IFOV versus incidence angle, for a variety of delay ranges. The dashed magenta line shows the 25-km requirement.

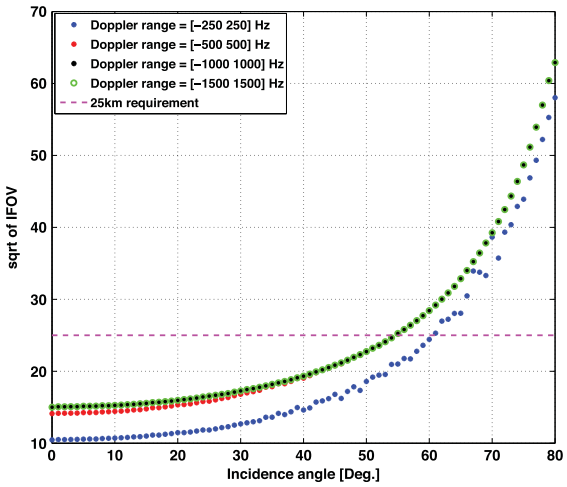

Fig. 8. Illustration of curves of square root of IFOV versus incidence angle, for a fixed delay range of [−0.25 0.25] chips and different Doppler ranges. The dashed magenta line shows the 25-km requirement. Note that the black dots for the [−1000 1000] Hz Doppler case overlap the green circles for the [−1500 1500] Hz Doppler case and that the [−0.25 0.25] chip line, illustrated in red in Fig. 7, is identical to the [−1500 1500] Hz case.

(i.e., by the azimuth angle), which, in turn, influences their orientation. Therefore, the configuration of the iso-Doppler lines is influenced by both the incidence and azimuth angles.

We define the instantaneous field of view (IFOV) as the physical area included in the intersection between the isorange ellipse and the two iso-Doppler lines corresponding to a given maximum delay and Doppler range. Fig. 7 shows curves of the square root of the IFOV versus the incidence angle, for different delay ranges and considering an infinite Doppler range. In this case, IFOV is the area included in the isodelay ellipse of the maximum delay considered. All of the delay ranges considered for the calculation of the observables begin at −0.25 chips to improve the SNR, but this obviously does not have an effect on the calculation of the IFOV since the first physically existing delay is 0 chip (i.e., the one at the SP). It is clear from Fig. 7 that, for a 25 km × 25 km requirement on the spatial resolution (shown as a magenta dashed line), the best suitable choice is a delay range from −0.25 to 0.25 chips. The range of scattering angles compliant with the spatial resolution requirement for a delay range from −0.25 to 0.50 chips is too narrow, whereas all the other ranges never satisfy the spatial resolution required.

Having selected the delay range, the Doppler range is chosen such that the iso-Doppler lines will not truncate the isodelay ellipse at 0.25 chips, which would cause some useful portion of the scattered signal to be left out, and also will not lie too far from the isodelay ellipse, which would introduce additional noise without adding a signal. Hence, the Doppler range is chosen to try to satisfy the following two conditions.

1) The iso-Doppler lines are the closest possible to the isodelay line selected.
2) The iso-Doppler lines are always outside the isodelay line selected.

The Doppler range that satisfies the aforementioned conditions for a given isodelay line is not unique and depends on the scattering geometry. However, for the chosen delay range, the iso-Doppler range that best satisfies the aforementioned condition is from −1000 to 1000 Hz. Fig. 8 shows again the curves of the square root of the IFOV versus the incidence angle, but this time for a fixed delay range from −0.25 to 0.25 chips and different Doppler ranges.

It is interesting to observe that the first case of Doppler range of [−250 250] Hz is different from the others and noisier since the very small Doppler range chops off a variable part of the area within the isodelay ellipse at 0.25 chips. Instead, small differences are observed for the other cases, and mostly at lower incidence angles. This happens because, at higher incidence angles, the iso-Doppler lines stretch out more rapidly than the isodelay lines and tend to fall quickly outside the 0.25 isodelay ellipse. Thus, the IFOV for higher incidence angles is entirely determined by the delay range.

The final choice for a Doppler range of [−1000 1000] Hz has then been confirmed by applying the full L2 retrieval algorithm to DDMA observables computed using the three different Doppler ranges shown in Fig. 8 and by verifying that the Doppler range of [−1000 1000] Hz is the one providing the lowest rms wind speed retrieval error.

Note that only the physical area included in the isodelay and iso-Doppler lines is considered. Several additional factors should be taken into account in order to fully relate the delay and Doppler ranges to a spatial region. These include the following: 1) The spatial boundaries defined by lines of constant delay and Doppler not conforming to a line of constant distance from the SP. An effective spatial resolution, based, for example, on equal area coverage, needs to be defined. 2) The effect of the WAF needs to be taken into account by a suitable definition of the effective spatial resolution since, at a given point in space, it defines a weighted response function for the contribution of adjacent regions of the delay/Doppler domain to the measurements [20]. The smearing caused by the WAF is expected to slightly increase the physical area, but this increase should be small. Finally, 3) as a consequence of

TA, illustrated in the next section, spatial smearing is produced in the direction of motion of the SP. Even though this spatial smearing is taken into account through the definition of the effective field of view (EFOV), the EFOV is here calculated assuming some simplifications, which are not always valid, as explained in the next section.

Each of these factors will be included in a more complete definition of the spatial resolution, which is currently under development.

C. TA

One extra processing step that improves the performances of the retrieval algorithm is the TA of the collected data. Looking again at Fig. 8, it can be seen that, up to an incidence angle of around 55°, the IFOV of the collected samples is below the spatial resolution requirement of 25 km. In these cases, it is possible to average a number of consecutive samples in time to reduce the noise in the observables, and achieve the spatial resolution limit of 25 km. The maximum number of samples that can be averaged can be estimated as follows. Since the SP is moving on the surface at approximately 6 km/s, an EFOV can be defined for each sample as

$$F_{\text{eff}} = F_{\text{ist}} + \sqrt{F_{\text{ist}}} \cdot (n-1) \cdot l \tag{4}$$

where F_{eff} is the EFOV, F_{ist} is the IFOV, n is the total number of samples to be averaged, and l indicates the SP displacement on the surface per second, which is about 6 km in the case of CYGNSS. Imposing an EFOV equal to 25 km and solving for n, we obtain

$$n = \frac{F_{\text{eff}}}{6\sqrt{F_{\text{ist}}}} - \frac{\sqrt{F_{\text{ist}}}}{6} + 1. \tag{5}$$

A plot of n versus incidence angle for two different delay ranges, and using an EFOV limit of 25 km, is shown in Fig. 9(a). The value n has been rounded down to the next lowest integer value, in order to make sure that the 25-km requirement is never violated. A graphical example of TA is also illustrated in Fig. 9(b), where the different SP tracks for a single CYGNSS observatory are shown as blue crosses, the sample considered for TA is shown in red, and the consecutive samples to be averaged together with the red one are highlighted with black circles.

The function shown in Fig. 9(a) can be then used to know, for each incidence angle, the number of samples that can be averaged if a 25-km requirement on the spatial resolution must be met. The calculation of n here uses simplifying assumptions and does not take into account several aspects, like the exact trajectory of the SP on the surface or the real shape of the IFOV, which is ideally assumed to be a square in (4) but which depends instead on the configuration of the isodelay and iso-Doppler lines at the ranges selected for the computation of the observables. As mentioned in Section III-B, these aspects are currently under investigation, but they are expected to have ultimately a small impact on n, making the simplifications used here valid enough to still allow a reliable estimation of n. In

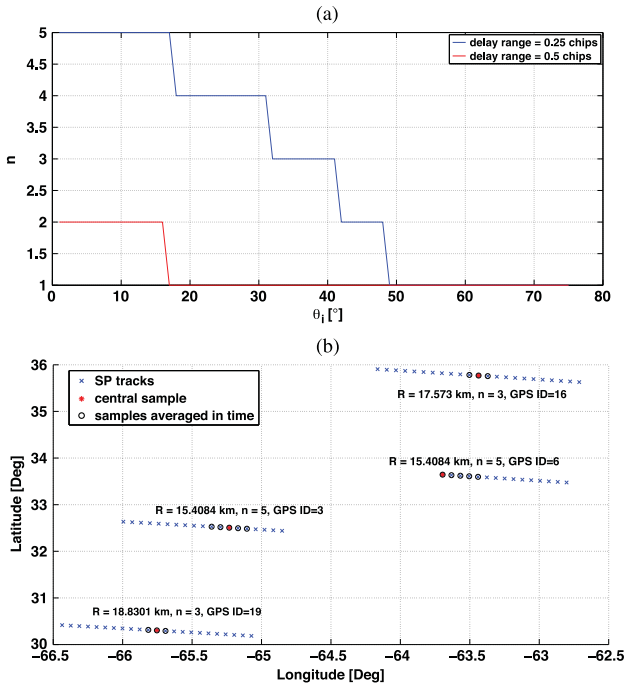

Fig. 9. (a) Number of samples to average, as a function of the incidence angle θi, for two different delay ranges. The blue curve for a delay range of [−0.25 0.25] chips, and corresponding essentially to the red curve in Fig. 7, is the one of interest. The red curve for a delay range of [−0.25 0.50] chips corresponds to the orange curve in Fig. 7. (b) Graphical illustration of how the TA algorithm works, for four simultaneous SP tracks acquired by a single CYGNSS observatory.

principle, two different approaches to TA are possible, and these are the following:

1) average the values of the observables and then retrieve the wind;
2) average the values of the retrieved wind.

These two approaches are not equivalent since the function (i.e., the LUT) relating one to the other is not linear.

We have tested both approaches by computing the total rms retrieval error for each case and for all wind speeds and found that the best performance is achieved by averaging the observables rather than the retrieved wind. Note also from Fig. 8 that observables at incidence angles higher than a threshold of 54.5° do not meet the resolution requirements, and hence, TA is only applicable to those observables at incidence angles lower than this threshold.

D. Derivation of GMF

The training data for the derivation of the GMF are formed as the L1b observables with odd minutes in their time stamp, computed over the optimal delay range of [−0.25 0.25] chips and Doppler range of [−1000 1000] Hz, as found in Section III-B, to which TA has been applied whenever possible. Only data with high enough RCG are selected for GMF generation. We use an RCG lower bound threshold of 20, which corresponds to roughly 22% of the data and covers the full range of azimuth angles and a range of incidence angles between 0° and 60°.

A full 2-D GMF, which is a function of both wind speed and incidence angle, is generated using the following steps.

1) The dependence of the observables on the incidence angle is estimated using observables computed from noise-free DDMs (the so-called L2a correction).
2) A unique GMF is derived for all the samples, to map the selected observables against the true wind speed from the NR, spatially averaged over an area of 25 km $\times$ 25 km and centered at the SP.
3) The GMF obtained in step 2 is scaled according to the L2a correction developed in step 1 in small increments of the incidence angle, to obtain a unique GMF for each incidence angle interval. This represents our final 2-D GMF.

This method constitutes a simplified approach to the derivation of a full 2-D GMF. It will be replaced by a more direct approach that filters the training set by both wind speed and incidence angle once a sufficiently large training population is available.

Greater detail about each of the processing steps is given in the following sections.

1) L2a Correction: The L2a correction represents the correction for the dependence on the incidence angle $\theta_{\rm SP}$ at the SP, which is not eliminated by the L1b calibration. After the L1b calibration, a dependence on θ_i and θ_r of the DDM of scattered power remains in the equation for the normalized RCS, σ_0. The angles θ_i and θ_r are the incidence and scattering angles formed, respectively by the transmitter range and a generic point on the surface, referred to as $\vec{\rho}$, and the receiver range and the same point $\vec{\rho}$ on the surface. We can then write the expression for the NRCS on $\vec{\rho}$ as [3]

$$\sigma_0(\vec{\rho},\theta_i,\theta_r) = \frac{\pi\,|\Re(\theta_i)|^2\,q^4(\vec{\rho},\theta_i,\theta_r)}{q_z^4(\vec{\rho},\theta_i,\theta_r)} P\left(-\frac{\vec{q}_\perp(\vec{\rho},\theta_i,\theta_r)}{q_z(\vec{\rho},\theta_i,\theta_r)}\right). \tag{6}$$

The scattering cross section depends on the Fresnel reflection coefficient $\Re$ and on the pdf $P(\ldots)$ of the sea surface slopes, where the slopes are defined as $\vec{s} = -\vec{q}_\perp/q_z = [-q_x/q_z, -q_y/q_z]$. The vector $\boldsymbol{q} = [q_x, q_y, q_z]$ is known as the scattering vector, namely, the bisector of the angle formed by the transmitter point on surface and receiver point on surface ranges. Assuming a Gaussian pdf for the sea surface slopes, it can be expressed as in [7], [20]

$$P(\vec{s}) = \frac{1}{2\pi\det(\boldsymbol{M})}\exp\left[-\left(\frac{1}{2}\vec{q}_\perp^{\,T}\boldsymbol{M}^{-1}\vec{q}_\perp\right)\right]$$
$$\boldsymbol{M} = \begin{bmatrix}\cos(\phi) & -\sin(\phi)\\ \sin(\phi) & \cos(\phi)\end{bmatrix}\begin{bmatrix}\sigma_{\rm up}^2 & 0\\ 0 & \sigma_{\rm cross}^2\end{bmatrix}\begin{bmatrix}\cos(\phi) & \sin(\phi)\\ -\sin(\phi) & \cos(\phi)\end{bmatrix} \tag{7}$$

where $\sigma_{\rm up}^2$ and $\sigma_{\rm down}^2$ are the variances of the slope (also known as MSSs) along the wave direction and perpendicular to it and ϕ is the wave direction with respect to the specular frame. If we assume to be close enough to the SP, where incidence and scattering angles are the same ($\theta_i = \theta_r = \theta_{\rm SP}$), then it can be easily verified that the x- and y-components of the scattering vector are null, and (6) becomes

$$\sigma_0(\vec{\rho},\theta_{\rm SP}) = \frac{|\Re(\theta_{\rm SP})|^2}{2\det(\boldsymbol{M})}. \tag{8}$$

In other words, under this approximation, the dependence on the incidence angle is contained only in the Fresnel reflection coefficient (see [3, eqs. (36)–(39)]), and thus, it is sufficient to correct for $\Re$ to eliminate such dependence.

Unfortunately, compensating for the effect of $\Re$ is not sufficient to fully eliminate the dependence of σ_0 on $\theta_{\rm SP}$. This is shown in Fig. 10, where a set of L1b DDMAs of RCS from noise-free simulated DDMs, and normalized by their scattering area, has been generated for a number of different realistic CYGNSS geometries, corresponding to different incidence angles, and for different wind speed values and a constant wave direction. The L1b DDMAs are corrected for the dependence on the Fresnel reflection coefficient (a correction obtained by simply dividing the L1b DDMA value by the squared module of $\Re$ for the specific incidence angle).

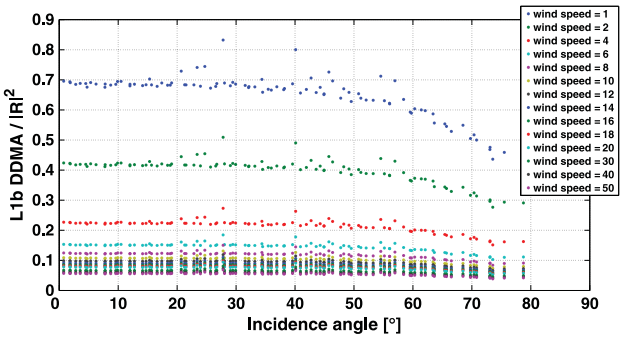

Fig. 10. Plot of noise-free L1b DDMA, versus incidence angle, obtained from noise-free DDMs simulated using the E2ES, and corrected for the dependence on the Fresnel reflection coefficient. The different colors refer to different wind speeds.

Fig. 10 shows that a dependence on incidence angle remains after having removed the dependence on $\Re$. This suggests that the assumption of $\theta_i = \theta_r = \theta_{\rm SP}$ is not sufficiently accurate and that the observables that we are considering cover an area large enough that the incidence and scattering angles over this area cannot be assumed equal to the incidence angle at the SP. From the L1b DDMA, we have therefore developed an empirical correction. Having verified that the dependence of the L1b DDMA on incidence angle does not vary with wind speed and having normalized each of the DDMA curves (i.e., the data sets corresponding to a single wind speed value) by their approximate value at $\theta_i = 0°$, we have found the coefficients a, b, and c of a power law relationship of the form $y(\theta) = a\theta^b + c$, which best fits the normalized DDMA data points in that curve. The normalized data points and the best fit power curve are shown in Fig. 11. This function is used to model the dependence of the L1b DDMA on incidence angle. The same procedure has been applied to L1b LES data, and the final coefficients for the power fit have been found to be the same as those for the DDMA.

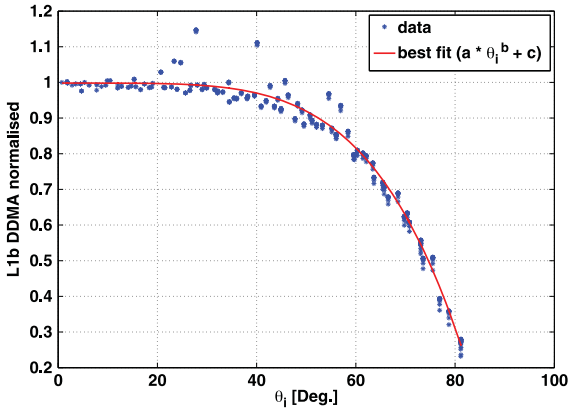

Fig. 11. (Blue) Normalized DDMA samples for all geometries and all wind speeds versus incidence angle, with (red) superimposed best fit power function. The coefficients are $a = -1.14 * 10^{-9}$, $b = 4.61$, and $c = 1.00$.

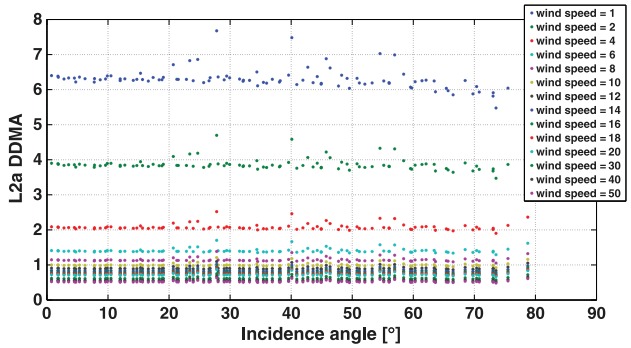

Fig. 12. Plot of noise-free L2a (bottom) DDMA, versus incidence angle, obtained from noise-free DDMs simulated using the E2ES. The different colors refer to different wind speeds.

The final expression to derive the L2a observables is therefore

$$O_{L2a} = \frac{O(\theta)_{L1b}}{y(\theta)} \tag{9}$$

where O stands for the observable and $y(\theta)$ is the power fit shown in Fig. 11. Fig. 12 shows the L2a DDMA, where the L2a correction has been now implemented through the use of the power fit mentioned earlier and where the dependence on the incidence angle has been eliminated.

2) Derivation of GMF From All Samples: A GMF is derived which maps the observables to wind speed, which, in our case, is taken as the true wind speed from the NR, spatially averaged over an area of 25 km $\times$ 25 km and centered at the SP. The GMF is computed in the form of a LUT from DDMA and LES values in the training data set, and their corresponding wind speeds using samples with RCG > 20. The DDMA and LES observables used to derive the GMF are the L2a-corrected ones according to (9). The LUT used by the retrieval algorithm is derived using a tapered approach for the bin widths, with smaller bin widths at low and high winds, where there are fewer available samples, and larger bin widths at medium winds, where the number of samples is large. The LUT function is constructed by taking, for each wind bin, an average with a triangle weighting of all the true winds falling within that bin as the y-value, and the similarly averaged corresponding L2a DDMA values at the x-value. A smoothing filter is then applied to the LUT, to make sure it is a physically reasonable monotonic function. The training data and LUT for both DDMA and LES samples are shown in Fig. 13.

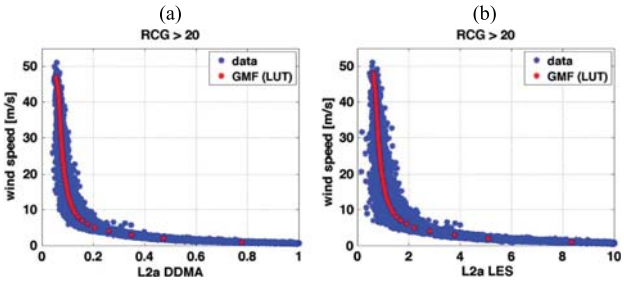

Fig. 13. Ground-truth wind speeds versus (a) L2a DDMA and (b) L2a LES training data, shown as blue points and selected with RCG > 20. The LUT derived from the data is shown as red points.

3) GMF Scaling Using L2a Correction: The dependence of the GMF on incidence angle is explicitly accounted for in the retrieval algorithm by adding a second dimension to the LUT that is indexed by incidence angle. Instead of applying the L2a correction to the individual observables, we scale the GMF obtained in Section III-D2 by the L2a correction to generate a different GMF at each incidence angle interval. This approach is taken for two main practical reasons. First, it allows for a simple software approach that is well suited for operational implementation in a spaceflight program. Second, it is amenable to simplified postlaunch adjustments, via tuning of the associated LUT entries, to account for possible discrepancies in the observed incidence angle dependence of the scattered signal. In fact, the current understanding of the bistatic scattering cross section of the ocean surface and its dependence on incidence angle and wind speed is not sufficient, particularly at high wind speeds, to rule out the possibility that adjustments to these prelaunch model-based GMFs will be required after launch.

We divide the total incidence angle range (from 0.1° to 80°) into 800 intervals of 0.1° each, and we derive a GMF $g(\theta, u)$ for each interval by applying inversely the L2a correction, as given by

$$g(\theta, u) = g(u) \cdot y(\theta) \tag{10}$$

where $g(u)$ is the total GMF derived in Section III-D2 and $y(\theta)$ is the L2a correction derived in Section III-D1. The $g(\theta, u)$ resulting from (10) is shown for five incidence angle intervals 10° wide and one incidence angle interval 5° wide, together with the samples belonging to that interval in Fig. 14. The six incidence angle intervals overall span the range of 0°–55°. The agreement between samples (blue) and GMFs (red) is overall good in all cases. Note, however, that the final 2D-GMF for wind retrieval uses 800 incidence angle intervals, in which the scatter is much smaller than what is seen in the six shown here.

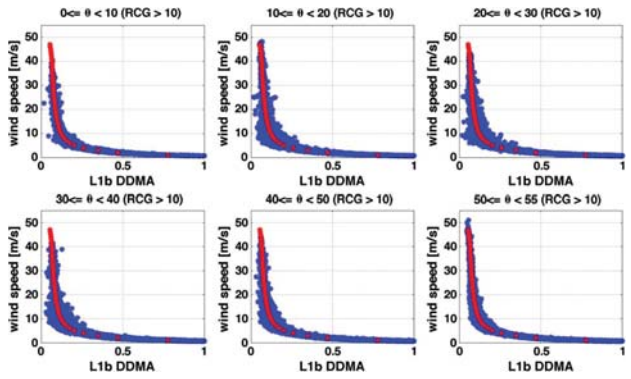

Fig. 14. (Blue) Scatterplot of L1b DDMA samples versus wind speed for six different incidence angle intervals (left to right and top to bottom): [0°–10°], [10*–20°], [20°–30°], [30°–40°], [40°–50°], and [50°–55°]. The red stars show the GMF corresponding to the 0.1° incidence angle interval at the center of the interval considered (e.g., the GMF for the [4.95°–5.05°] interval is shown for the samples belonging to the [0°–10°] interval, and so on).

4) Best Weighted Estimator: Following the implementation of the estimator presented in [12], the DDMA and LES wind estimates are linearly combined into a best weighted estimator, using coefficients calculated according to (1), for the following intervals of RCG:

1) $3 \leq \text{RCG} < 5$;
2) $5 \leq \text{RCG} < 10$;
3) $10 \leq \text{RCG} < 20$;
4) $\text{RCG} \geq 20$.

The covariance matrices used to compute the best weighted estimate coefficients are derived empirically from the population of retrievals since neither the individual pdfs of DDMA and LES wind retrieval errors nor the joint pdf between the two retrieval errors is available in analytical form.

The bias between true and retrieved winds is first computed for each observable and for each RCG interval and is removed, so as to compute the coefficients from unbiased DDMA and LES wind estimates. The dependence of bias on wind speed was found not to be significant and therefore was not included. The coefficients $\boldsymbol{m}$, as well as the bias, are calculated from the training data set (50% of data), where the true wind speeds are known, and then used in the test data set, where the wind speeds are not known, to assess the algorithm performance. This means that the coefficients may be suboptimal when applied to the test data set, but we expect the performance to be asymptotically optimal for increasing number of observations, assuming that the statistical properties of the training data set are the same as those of the test data set. Table I shows the correlation coefficients between DDMA and LES wind estimates, for the four RCG intervals, highlighting that the correlation is lowest at the low RCG values, due to the stronger noise affecting the correlation statistics, and then, it gradually increases for increasing RCG values. In this sense, the separation of the RCG intervals is useful since the best weighted estimator helps improve the performances when the correlation is low, i.e., for the noisiest data.

TABLE I
CORRELATION COEFFICIENT BETWEEN DDMA AND LES WIND ESTIMATES SHOWN FOR THE FOUR RCG INTERVALS UNDER ANALYSIS

	3≤ RCG ≤5	5≤ RCG ≤10	10≤ RCG ≤20	RCG ≥20
Corr. Coeff.	**0.09**	**0.11**	**0.12**	**0.35**

IV. WIND RETRIEVAL ALGORITHM PERFORMANCE ANALYSIS

The performance of the wind speed retrieval algorithm is evaluated using the test data set set aside from the NR simulations. Algorithm performance and error characterization are done in terms of different lower bound thresholds used to determine the CYGNSS spatial coverage statistics.

A. Performance Figure of Merit

The performance of the retrieval algorithm is characterized using two different figures of merit. The first is the absolute wind speed rms error in meters per second, as a function of the true wind speed. This is calculated as

$$\sigma\left(u_{(c)}^{k}\right) = \sqrt{\frac{1}{N_k}\sum_{i=1}^{N_k}\left[\hat{u}_{(ub)}^{i} - u^{i}\right]^2} \tag{11}$$

where the index k represents different intervals of true wind speeds, the index i represents the different wind samples within the kth interval, $u_{(c)}^{k}$ is the wind value at the center of the kth interval, N_k is the number of samples in the kth interval, $\hat{u}_{(ub)}^{i}$ is the ith estimated unbiased wind speed, and u^i is the ith true wind speed. These curves are useful to assess how the error changes and depends on the true wind speed. The second figure of merit is the average relative rms error, expressed in percentage, for wind speeds above a given threshold. The average relative rms error is calculated as

$$\sigma_R = \int_{v>V'} \overline{\sigma}(v) f_v(v|H) dv \tag{12.a}$$

$$\overline{\sigma}\left(u_{(c)}^{k}\right) = \sigma\left(u_{(c)}^{k}\right) / u_{(c)}^{k} \tag{12.b}$$

where σ_R is the average relative rms error, V' is the wind threshold, $u_{(c)}^{k}$ is the wind at the center of the kth interval, and $f_v(v|H)$ is a conditional pdf of true wind speeds, derived from the pdf shown in Fig. 5, with H representing the condition of wind higher than the given threshold. This expresses the rms error as a single number, for wind speeds above a given threshold. The performance of the algorithm is also evaluated after an EFOV filter is applied to filter out all the wind estimates with an incidence angle greater than 54.5° since those samples have an EFOV greater than 25 km × 25 km and are not compliant with the CYGNSS mission resolution requirements (see Fig. 8).

An example of the rms error plots computed according to (11) is shown in Fig. 15, calculated using wind speed intervals of width ±10 m/s. Fig. 15 highlights the progressive reduction in error when various components of the retrieval algorithm are

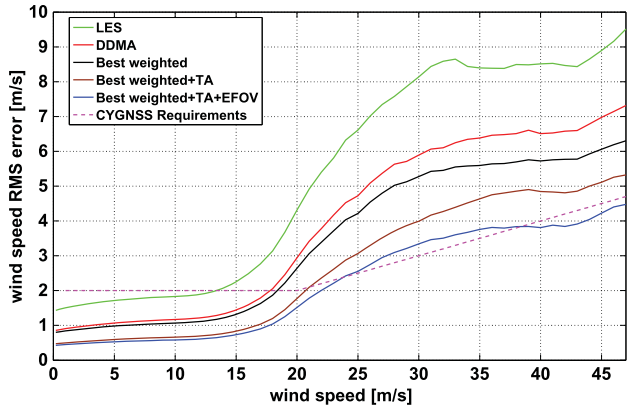

Fig. 15. RMS error curves for (green) LES and (red) DDMA, and the progressive improvements in the performances (black) when the best weighted approach is applied, (brown) when TA is applied to the observables, and (blue) when the EFOV filter is also applied, allowing to meet the (dashed magenta) CYGNSS requirements. The curves are computed from samples with RCG > 10.

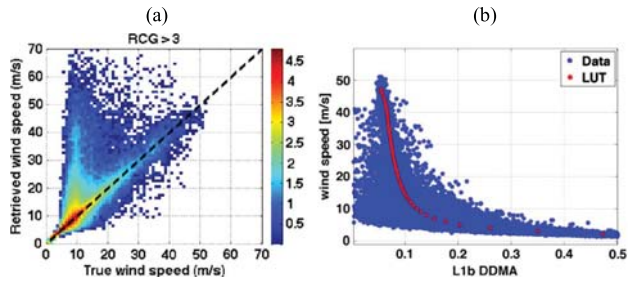

Fig. 16. (a) Density scatterplots of true versus retrieved winds, from best weighted estimator with AC approach, TA, and EFOV filter, for RCG > 3. (b) Scatterplot of L1b DDMA data (blue) versus true wind speed, for RCG > 3, with (red) overlapped GMF.

incorporated one by one. The best weighted estimator (black) improves the performances compared to DDMA (red) and LES (green) alone, particularly for the high wind speed part. The inclusion of TA (brown) and the EFOV filter (blue) further improves performance and ultimately allows the requirements for the wind speed rms error to be met (magenta). Note that the EFOV filter has a net effect of improving the final performances of the algorithm since it filters out a portion of the RCG map characterized by low RCG values (see Fig. 2).

The performance and error characterization shown in the subsequent paragraphs are all derived using the combination of the best weighted estimator, TA, and EFOV filter.

B. Performance and Error Characterization of Retrieval Algorithm

An illustration of the density scatterplot of the retrieved versus the true wind speed is shown in Fig. 16(a), for RCG > 3, where the colorbar indicates the log-density of samples. The correlation coefficient between estimated and true wind speeds is 0.83.

The density plot shows a "two-branch" behavior whose explanation can be found in Fig. 16(b), where the DDMA observables for RCG > 3 are scatter plotted against the true wind speed, with the GMF overlaid. The large number of low

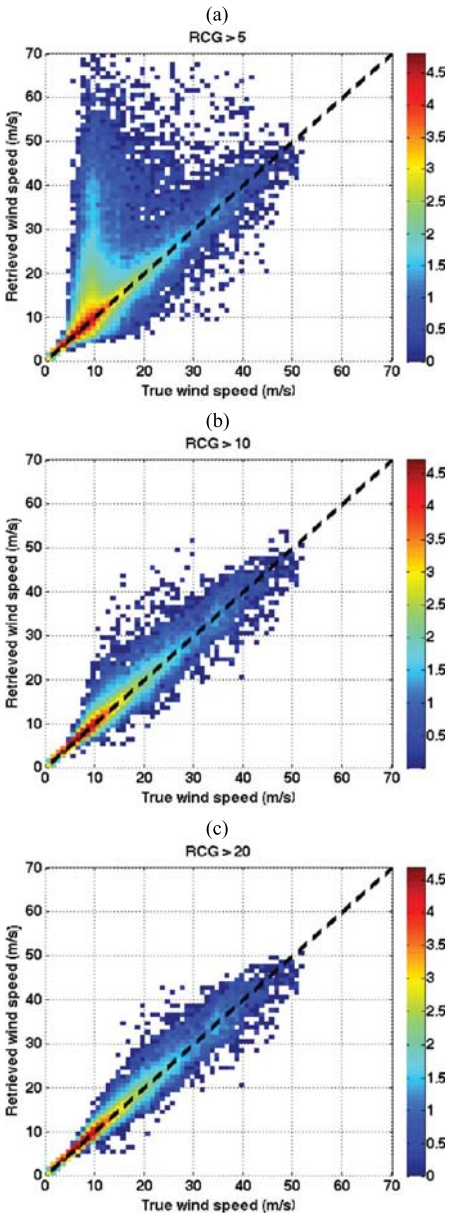

Fig. 17. Density scatterplots of true versus retrieved winds, from best weighted estimator with AC approach, TA, and EFOV filter, for (a) RCG > 5, (b) RCG > 10, and (c) RCG > 20. The color refers to the density value in logarithmic scale.

DDMA observables, occurring around 10-m/s true wind speed, produces wind retrievals much higher than 10 m/s when the GMF is used to map these DDMA values into wind speeds, hence generating the second branch of high winds near the 10-m/s true wind in Fig. 16(a).

The increase in RCG thresholds reduces the noise in the samples and improves the correspondence between retrieved and true winds; in particular, it contributes to the gradual elimination of all those samples characterized by low observable values at low true winds and the disappearance of the "two-branch" behavior. This is shown in Fig. 17, where the density plot of retrieved versus true winds is shown for RCG thresholds

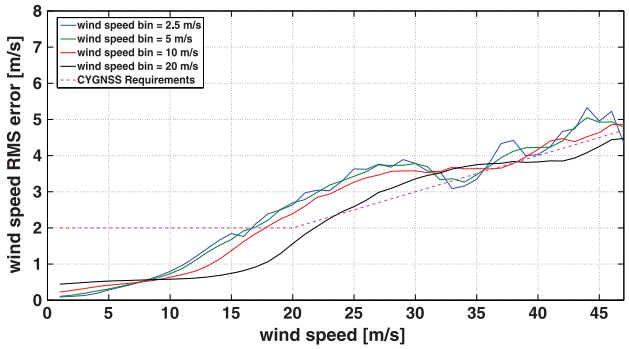

Fig. 18. RMS error curves versus true wind speed, for a fixed RCG lower bound threshold of 10 and for different wind speed bin sizes.

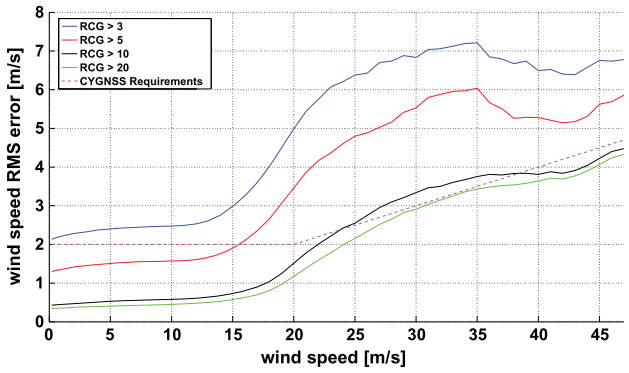

Fig. 19. RMS error curves versus true wind speed, for four different RCG thresholds.

of 5, 10, and 20, with correlation coefficients between estimated and true winds of 0.91, 0.98, and 0.99, respectively.

Fig. 18 shows the wind speed rms error curves for an RCG lower bound threshold of 10, using different widths of the wind speed interval for the calculation of the error.

Fig. 18 demonstrates that the rms error is generally higher when small wind speed intervals are used and decreases with wider intervals. For the performance analysis that follows, we select a wind speed interval of 20 m/s (i.e., ± 10 m/s) since this is consistent with the NASA mission requirement that the wind speed retrieval performance be evaluated over two wind speed intervals, those below 20 m/s and those above it. Wind speed rms error curves as a function of the true wind speed, for different RCG lower bound thresholds, are shown in Fig. 19 along with the baseline mission requirement on the rms error.

From Fig. 19, we see that, while requirements are met for wind speeds lower than 15 m/s using an RCG higher than 5, an RCG higher than 10 is instead the minimum threshold that meets the requirement at wind speeds higher than 15 m/s. The reduction in rms error for wind speeds greater than 35 m/s shown in the figure is a consequence of the reduced number of samples available at higher wind speeds. Above 35 m/s, the number density of available samples drops significantly below that at lower wind speeds, and the rms error, being estimated from a population of samples over a range of wind speeds, becomes overly weighted toward the performance at the lower wind speeds. In fact, the rms error is expected to grow monotonically with wind speed.

TABLE II
SUMMARY OF PERFORMANCES FOR WINDS LOWER THAN 20 m/s, FOR DIFFERENT RCG THRESHOLDS

	RCG ≥ 3	RCG ≥ 5	RCG ≥ 10
RMS Err [m/s]	**2.48**	**1.57**	**0.58**

TABLE III
SUMMARY OF PERFORMANCES FOR WINDS HIGHER THAN 20 m/s, FOR DIFFERENT RCG THRESHOLDS

	RCG ≥ 5	RCG ≥ 10	RCG ≥ 20
Relative RMS Err [m/s]	**17.8 %**	**9.8%**	**8.2%**

TABLE IV
PERCENTAGE OF RETAINED SAMPLES FOR DIFFERENT LOWER BOUNDS OF RCG, WITH EFOV ON (FIRST ROW) AND OFF (SECOND ROW)

EFOV FILTER	RCG ≥ 3	RCG ≥ 5	RCG ≥ 10	RCG ≥ 20
ON	**66%**	**62%**	**50%**	**45%**
OFF	**81%**	**73%**	**57%**	**48%**

It is also useful to characterize the algorithm performance in terms of overall rms error for low-to-medium winds (i.e., lower than 20 m/s) and for high winds (i.e., higher than 20 m/s). Table II illustrates the rms error for all cases with true wind speeds lower than 20 m/s and for different RCG thresholds. The relative rms error is calculated here according to (11), using only samples with true wind below 20 m/s. The mission requirement of an rms error below 2 m/s is met for an RCG lower bound threshold of 5.

Table III shows the relative rms error for true wind speeds higher than 20 m/s and again for different RCG thresholds. The relative rms error is calculated here according to (12). Table III highlights that the minimum RCG threshold required to meet the baseline mission requirement for winds higher than 20 m/s is 10. Such RCG lower bounds inevitably impose some cutoff on the number of samples used to compute the performances. The percentages of retained samples for the different RCG lower bounds are reported in Table IV, both when the EFOV filter is applied and not. Since the EFOV filter primarily eliminates samples with low RCG (see Fig. 2), it preferentially affects the percentage of samples at the low RCG lower bounds.

C. Some Examples of Retrieved Versus True Winds

In this last section, we present some example transects of true and retrieved winds versus time, for cases of storm crossing with high RCG and for the wind field plots shown in Fig. 3. The SP tracks are made of 60 s of data acquisition. The corresponding plots of the true and retrieved wind speeds versus elapsed time, for the 60 s considered, are shown in Fig. 20. We

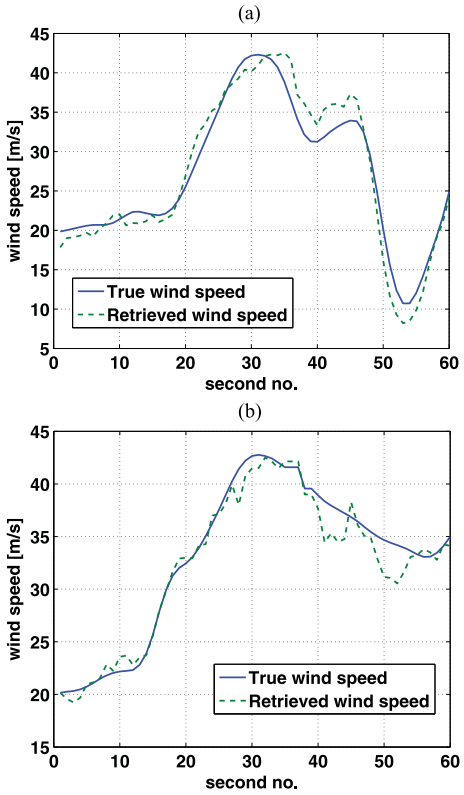

Fig. 20. Illustration of true and retrieved wind speeds, versus second number, for (a) case 1 and (b) case 2 of storm crossing of specular tracks, shown in Fig. 3. The RCG value of the SP at the center of the transect is respectively 161 for case 1 and 190 for case 2.

observe that the retrievals match the true winds well, and this is generally true for high RCG values.

V. Summary and Conclusion

The baseline CYGNSS L2 wind speed retrieval algorithm has been presented. The algorithm uses the wind speed retrieval approach described in [12] and introduces a number of improvements specifically developed for the CYGNSS mission. The algorithm has been tested with simulated DDMs that represent the expected characteristics of CYGNSS in-orbit measurements [19]. The main differences and improvements that have been incorporated into the present algorithm, compared to [12], are as follows:

1) observable DD range selected based on spatial resolution;
2) use of L1b DDMs of RCS instead of L1a DDMs of received power;
3) use of realistic 13-day NR DDMs simulating the full life cycle of a TC, with a 50/50 split of the population for training and evaluation purposes;
4) development of a 2-D GMF function with respect to both wind speed and incidence angle, through the L2a correction;
5) wind speed estimator using adaptive (i.e., RCG-dependent covariance matrix) approach;
6) implementation of TA.

The general performances of the retrieval algorithm for the CYGNSS baseline spatial resolution requirement of 25 km $\times$ 25 km have been presented in Tables II and III and characterized using RCG thresholds for the two distinct cases of wind speed lower than 20 m/s and higher than 20 m/s. The main conclusions that can be drawn from the present analysis are that an RCG higher than 5 is sufficient to meet the requirements for wind speeds lower than 20 m/s while an RCG higher than 10 is necessary to meet the requirements at wind speeds higher than 20 m/s. Some examples of SP transects passing through or very near the eye of the storm have been shown, highlighting that the retrieved winds are in good agreement with the true winds when the samples are acquired with sufficiently high RCG.

The present algorithm has only been applied to two specific observables (DDMA and LES) due to accumulation time and resolution constraints, as pointed out in Section II. However, the best weighted approach can be generally used on any number of different observables, and usually, the larger the number of observables, the better the results, provided that the observables chosen meet the requirements on spatial resolution and have retrieval errors that are sufficiently decorrelated from one another.

Acknowledgment

The authors would like to thank the reviewers for their in-depth reviews and helpful suggestions that have largely contributed to improve this paper.

References

[1] C. Ruf *et al.*, "The NASA EV-2 Cyclone Global Navigation Satellite System (CYGNSS) mission," in *Proc. IEEE Aerosp. Conf.*, 2013, pp. 1–7, DOI: 10.1109/AERO.2013.6497202.
[2] S. J. Katzberg, O. Torres, and G. Ganoe, "Calibration of reflected GPS for tropical storm wind speed retrievals," *Geophys. Res. Lett.*, vol. 33, no. 17, Sep. 2006, Art. no. L18602, DOI:10.1029/2006GL026825.
[3] V. Zavorotny and A. Voronovich, "Scattering of GPS signals from the ocean with wind remote sensing applications," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 2, pp. 951–964, Mar. 2000.
[4] M. Armatys, "Estimation of sea surface winds using reflected GPS signals," Ph.D. dissertation, Dept. Aerosp. Eng. Sci., Univ. Colorado, Boulder, CO, USA, 2001.
[5] E. Cardellach, "Sea surface determination using GNSS reflected signals," Ph.D. dissertation, Dept. Signal Theor., Univ. Politecnica Catalunya, Catalunya, Spain, 2002.
[6] E. Cardellach *et al.*, "Mediterranean balloon experiment: Ocean wind speed sensing from the stratosphere using GPS reflections," *Remote Sens. Environ.*, vol. 88, no. 3, pp. 351–362, Dec. 2003.
[7] O. Germain *et al.*, "The eddy experiment: GNSS-R speculometry for directional sea-roughness retrieval from low altitude aircraft," *Geophys. Res. Lett.*, vol. 31, no. 21, Nov. 2004, Art. no. L12306.
[8] M. Clarizia *et al.*, "Analysis of GNSS-R delay-Doppler maps from the UK-DMC satellite over the ocean," *Geophys. Res. Lett.*, vol. 36, no. 2, Jan. 2009, Art. no. L02608.
[9] J. Garrison, A. Komjathy, V. Zavorotny, and S. J. Katzberg, "Wind speed measurements using forward scattered GPS signals," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 1, pp. 50–65, Jan. 2002.
[10] A. Komjathy *et al.*, "Retrieval of ocean surface wind speed and wind direction using reflected GPS signals," *J. Atmos. Ocean. Technol.*, vol. 21, no. 3, pp. 515–526, Mar. 2004, DOI: http://dx.doi.org/10.1175/1520-0426(2004)021<0515:ROOSWS>2.0.CO;2.
[11] S. Gleason *et al.*, "Detection and processing of bistatically reflected GPS signals from low-earth orbit, for the purpose of ocean remote sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 1229–1241, Jun. 2005.

[12] M. P. Clarizia, C. Ruf, C. Gommenginger, and P. Jales, "Spaceborne GNSS-R minimum variance wind speed estimator," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 11, pp. 6829–6843, Nov. 2014.
[13] S. Gleason, "Space-based GNSS scatterometry: Ocean wind sensing using an empirically calibrated model," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 9, pp. 4853–4863, Sep. 2013, DOI: 10.1109/TGRS.2012.2230401.
[14] J. Marchan-Hernandez *et al.*, "Correction of the sea state impact in the L-band brightness temperature by means of delay-Doppler maps of global navigation satellite signals reflected over the sea surface," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 10, pp. 2914–2923, Oct. 2008.
[15] J. Marchan-Hernandez *et al.*, "Sea-state determination using GNSS-R data," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 621–625, Oct. 2010.
[16] N. Rodriguez-Alvarez *et al.*, "Airborne GNSS-R wind retrievals using delay-Doppler maps," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 626–641, Jan. 2013.
[17] S. T. Gleason, "Remote sensing of ocean, ice and land remote sensing using bistatically scattered GNSS signals from low earth orbit," Ph.D. dissertation, Univ. Surrey, Surrey, U.K., 2006. [Online]. Available: http://clasp-research.engin.umich.edu/missions/cygnss/reference/gnss-overview/Gleason_Thesis_GNSS.pdf
[18] S. Gleason, C. Ruf, M. P. Clarizia, and A. O'Brien, "Calibration and unwrapping of the normalized scattering cross section for the cyclone global navigation satellite system ," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 5, pp. 2495–2509, May 2016, DOI: 10.1109/TGRS.2015.2502245.
[19] A. O'Brien, "CYGNSS End-to-End Simulator Technical Memo," 2014, *available upon request*.
[20] M. P. Clarizia, "Investigating the effect of ocean waves on ocean-reflected GPS measurements," Ph.D. dissertation, School Ocean Earth Sci., Univ. Southampton, Southampton, U.K., 2012. [Online]. Available: http://eprints.soton.ac.uk/359049/.
[21] E. P. W. Attema 1986, "An experimental campaign for the determination of the radar signature of the ocean at C-band," in *Proc. 3rd Int. Colloq. Spectral Sign. Obj. Remote Sens.*, Les Arcs, France, 1986, pp. 791–799.
[22] A. Cavanie and D. Ofler, "ERS-1 wind scatterometer: Wind extraction and ambiguity removal," in *Proc. IGARSS Todays Sol. Tomorrows Inf. Needs*, Zurich, Switzerland, 1986, pp. 395–398.
[23] A. C. M. Stoffelen and D. L. T. Anderson, "Scatterometer data interpretation: Measurement space and inversion," *J. Atmos. Ocean. Technol.*, vol. 14, no. 6, pp. 1298–1313, Dec. 1997.
[24] A. C. M. Stoffelen and D. L. T. Anderson, "Scatterometer data interpretation: Derivation of the transfer function CMOD4," *J. Geophys. Res.*, vol. 102, no. C3, pp. 5767–5780, Mar. 1997.
[25] H. Hersbach, A. Stoffelen, and S. de Haan, "An improved C-band scatterometer ocean geophysical model function: CMOD5," *J. Geophys. Res.*, vol. 112, no. C3, Mar. 2007, Art. no. C03006, DOI:10.1029/2006JC003743.
[26] J. A. Verspeek *et al.*, "Validation and calibration of ASCAT using CMOD5.n," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 1, pp. 386–395, Jan. 2010, doi:10.1109/TGRS.2009.2027896.
[27] J. Verspeek, A. Stoffelen, A. Verhoef, and M. Portabella, "Improved ASCAT wind retrieval using NWP ocean calibration," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 7, pp. 2488–2494, Jul. 2012, DOI:10.1109/TGRS.2011.2180730.
[28] B. S. Gohil, R. Sikhakolli, and R. K. Gangwar, "Development of geophysical model functions for Oceansat-2 scatterometer," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 2, pp. 377–380, Mar. 2013, DOI: 10.1109/LGRS.2012.2207369.
[29] S. Soisuvarn, Z. Jelenak, P. S. Chang, S. O. Alsweiss, and Q. Zhu, "CMOD5.H—A high wind geophysical model function for c-band vertically polarized satellite scatterometer measurements," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 6, pp. 3744–3760, Jun. 2013.
[30] P. Misra and P. Enge, *Global Positioning System: Signals, Measurements, and Performance*, Revised 2nd ed. Lincoln, MA, USA: Ganga-Jamuna Press, 2011.
[31] D. Nolan, R. Atlas, K. T. Bhatia, and L. Bucci, "Development and validation of a hurricane nature run using the joint OSSE nature run and the WRF model," *J. Adv. Model. Earth Syst.*, vol. 5, no. 2, pp. 382–405, Jun. 2013.
[32] T. Elfouhaily, D. R. Thompson, and L. Linstrom, "Delay-Doppler analysis of bistatically reflected signals from the ocean surface: Theory and application," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 3, pp. 560–573, Mar. 2002, DOI: 10.1109/TGRS.2002.1000316.

Maria Paola Clarizia (M'14) received the M.S. degree in telecommunications engineering (*summa cum laude*) from the University of Sannio, Benevento, Italy, in 2007 and the Ph.D. degree in global navigation satellite system (GNSS) reflectometry from the School of Ocean and Earth Science, University of Southampton, Southampton, U.K.

She was a Research Engineer in the field of synthetic aperture radar altimetry and GNSS reflectometry with Starlab Barcelona SL and a Research Scientist in the field of GNSS reflectometry with the Natural Environment Research Council, National Oceanography Centre, Southampton. In April 2013, she joined the University of Michigan, Ann Arbor, MI, USA, as a Postdoctoral Research Fellow, developing wind speed retrieval algorithms for the NASA Cyclone Global Navigation Satellite System mission. Her research is focused on GNSS reflectometry, altimetry, GNSS radio occultation, sea surface simulations and electromagnetic scattering models, data analysis, and statistical processing.

Dr. Clarizia is a member of the European Geophysical Union and the American Meteorological Society.

Christopher S. Ruf (S'85–M'87–SM'92–F'01) received the B.A. degree in physics from Reed College, Portland, OR, USA, and the Ph.D. degree in electrical and computer engineering from the University of Massachusetts at Amherst, Amherst, MA, USA.

He was previously with Intel Corporation; Hughes Space and Communication; the NASA Jet Propulsion Laboratory; and Pennsylvania State University, State College, PA, USA. He is currently a Professor in atmospheric science and space engineering with the University of Michigan, Ann Arbor, MI, USA, and a Principal Investigator of the Cyclone Global Navigation Satellite System NASA Earth Venture mission. His research interests include Global Navigation Satellite System reflectometry 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 Commission F of the Union Radio Scientifique Internationale. He was a former Editor-in-Chief of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (TGRS), and he has served on the editorial boards of *Radio Science* and the *Journal of Atmospheric and Oceanic Technology*. He has been the recipient of four NASA Certificates of Recognition and seven NASA Group Achievement Awards, as well as the 1997 TGRS Best Paper Award, the 1999 IEEE Resnik Technical Field Award, the 2006 IGARSS Best Paper Award, and the 2014 IEEE GRSS-S Outstanding Service Award.