A Forward Model for Data Assimilation of GNSS Ocean Reflectometry Delay-Doppler Maps

Feixiong Huang, *Student Member, IEEE,* James L. Garrison, *Senior Member, IEEE,* S. Mark Leidner,
Bachir Annane, Ross N. Hoffman, Giuseppe Grieco and Ad Stoffelen, *Senior Member, IEEE*

Abstract—Delay-Doppler maps (DDMs) are generally the lowest-level of calibrated observables produced from Global Navigation Satellite System Reflectometry (GNSS-R). A forward model is presented to relate the DDM, in units of absolute power at the receiver, to the ocean surface wind field. This model and the related Jacobian are designed for use in assimilating DDM observables into weather forecast models. Given that the forward model represents a full set of DDM measurements, direct assimilation of this lower-level data product is expected to be more effective than using individual specular-point wind speed retrievals. The forward model is assessed by comparing DDMs computed from Hurricane Weather Research and Forecasting (HWRF) model winds against measured DDMs from the Cyclone Global Navigation Satellite System (CYGNSS) Level 1a data. Quality controls are proposed as a result of observed discrepancies due to the effect of swell, power calibration bias, inaccurate specular point position and model representativeness error. DDM assimilation is demonstrated using a Variational Analysis Method (VAM) applied to three cases from June 2017, specifically selected due to the large deviation between scatterometer winds and European Centre for Medium-Range Weather Forecasts (ECMWF) predictions. DDM assimilation reduced the root mean square error (RMSE) by 15, 28 and 48%, respectively, in each of the three examples.

Index Terms—GNSS-R, GPS, ocean wind, CYGNSS, data assimilation, DDM, forward model

I. INTRODUCTION

OCEAN remote sensing with Global Navigation Satellite System Reflectometry (GNSS-R) has been developed over the last two decades for use on both airborne and spaceborne platforms. A sequence of satellite technology demonstrations, starting with the Disaster Monitoring Constellation (UK-DMC) [1] and later TechDemoSat-1 (TDS-1) [2] have validated the capability of ocean wind speed retrieval from GNSS-R signals. On December 15, 2016, NASA launched the Cyclone Global Navigation Satellite System (CYGNSS), an 8-satellite constellation intended to use ocean wind retrievals from GNSS-R to improve track and intensity forecasts of

Feixiong Huang and James L. Garrison are with the Department of Aeronautics and Astronautics, Purdue University, West Lafayette, IN, USA (e-mail: huang712@purdue.edu).

S. Mark Leidner is with Atmospheric and Environmental Research, Lexington, MA, USA

Bachir Annane is with Cooperative Institute for Marine and Atmospheric Studies, Miami, FL, USA.

Ross N. Hoffman is with Cooperative Institute for Satellite Earth System Studies, University of Maryland, at NOAA/NESDIS/Center for Satellite Applications and Research, College Park, MD, USA

Giuseppe Grieco and Ad Stoffelen are with Royal Netherlands Meteorological Institute, De Bilt, The Netherlands.

Manuscript received March XX, 2020; revised XXXX XX, 2020.

tropical cyclones (TCs). GNSS transmissions lie in L-band (1-2 GHz), which has a lower rain attenuation than the higher frequencies typically used by scatterometers. This can increase the available observations within the inner core of a TC. Additionally, a constellation of small satellites (enabled by the order-of-magnitude lower size, weight and power of a reflectometry instrument vs. that of an active radar) provides better spatial and temporal sampling with up to 32 simultaneous measurements per second and an average revisit time of 4 hours. This high rate of spatial and temporal sampling enables improved observation of the rapid intensification stage in TC development [3].

The fundamental GNSS-R observable is the delay-Doppler map (DDM), the cross-correlation between a reflected GNSS signal and a locally generated model of the transmitted signal (code and carrier) over a range of delays and Doppler frequencies [4]. The Delay Doppler Mapping Instrument (DDMI) on each CYGNSS satellite can nominally track up to four (4) GNSS reflections, providing one DDM per second at each reflection over a range of 17 delays and 11 Dopplers, sampled at 0.25 chip and 500 Hz, respectively [5]. The CYGNSS Science Operation Center (SOC) generates several levels of data products from these DDMs. At Level 1a (L1a), the DDM is calibrated to absolute power in units of watts. The Level 1b data product contains the bistatic radar cross-section (BRCS), generated from the DDM by removing the effects of satellite geometries, attitudes, direct power and antenna gain [6]. Level 2a data consists of the surface wind speed retrieval at the specular point with a resolution of 25 km. This retrieval uses the average DDM (DDMA) and the leading edge slope (LES) observables computed over a 3×5 window aligned with the specular point delay. Geophysical model functions (GMFs) were derived for these observables by fitting a model to training data [7]. Two GMFs were developed, one for fully developed seas (FDS) and another one for young seas/limited fetch (YSLF) conditions, resulting in two versions of the wind speed retrieval [8]. A minimum variance estimator is used to optimally combine the DDMA and LES retrievals [9]. The mean-square slope (MSS) of the ocean surface is also provided as a Level 2b data product, obtained directly from the BRCS [10]. For all of these retrievals, however, the 25 km resolution requirement limits the usable region of the DDM to only 15 samples around the specular point, discarding more than 90% of the available data. The relationship between the wind field away from the specular point, and the DDM observables at longer delays and larger Doppler frequencies is complicated and generally not invertible due to the delay-

Doppler ambiguity. In this paper, we propose the use of data assimilation (DA) to extract information from a much larger region of each sequential DDM to improve model representation of the wind field over a wider area.

DA has been shown to be an effective method for incorporating observations into numerical models to improve weather forecasts [11]. In the typical DA approach, a state is defined from a discretized geophysical field and observations are processed iteratively to update estimates of this state under physically-based constraints. An initial estimate of the state, known as the background, is provided and the updated state is known as the analysis field [12]. Assimilation can be based on stochastic methods by directly computing the optimal estimate (e.g. the Kalman filter [13], [14]) or variational methods by minimizing a cost function (e.g. adjoint method, 3D-Var [15], [16], 4D-Var [17]). When assimilating indirect observations, a linear observation operator that computes the response of observable to the physical state is required [13]–[16]. For example, a bending angle forward model has been used to assimilate GNSS radio occultation measurements [18], [19] and a geophysical model function has been used to relate scatterometer radar cross-sections to ocean surface wind fields [20]. A Jacobian or tangent linear operator is also usually required for the optimization process.

Recent studies have shown the impact of assimilating simulated or real CYGNSS Level 2 winds using the Gridpoint Statistical Interpolation (GSI)-based 3DVAR system [21]–[23] and the variational analysis method (VAM) [24], [25]. As stated previously, DDM power (a Level 1a product) is a more fundamental GNSS-R observable and incorporates information from a larger wind field than only that at the specular point. Direct assimilation of DDM power is, therefore, expected to better utilize more of the information in the GNSS-R measurement. In this paper, an analytical GNSS-R forward model is developed to link the DDM to a gridded wind speed and enable assimilation of a calibrated 17×11 DDM (CYGNSS Level 1a data) into weather forecast models, where the observation errors are considered to be uncorrelated. Taking in a gridded wind field and GNSS-R metadata, the forward model produces a simulated DDM which will be evaluated through comparison with measured CYGNSS DDMs at a defined delay-Doppler coordinate. An analytical model for the Jacobian matrix, which describes the sensitivity of DDM power in each delay-Doppler bin to the gridded wind field, is also derived. These models meet all of the requirements for optimal assimilation of Level 1a DDM data. Furthermore, all of the code is written as a single callable function in the C programming language so that it can be easily integrated into any data assimilation system.

As a demonstration, the forward model and Jacobian are integrated into the VAM and used to process CYGNSS observations from a few specific examples in which the European Centre for Medium-Range Weather Forecasts (ECMWF) model shows significant differences with scatterometer (AS-CAT, OSCAT) wind fields. In each of these examples, the VAM is shown to improve the agreement between ECMWF and scatterometer winds.

A description of the forward model and Jacobian are pre-sented in Section II. Section III presents an analysis of the model agreement with actual CYGNSS measurements and a comparison of the closed-form Jacobian with finite difference calculations. Section IV presents an example application of the forward model in the VAM. Section V summarizes the results and our conclusion.

II. MODEL DESCRIPTION

Inputs to the forward model can separated into four data types: *Geometry*, including the positions and velocities of the transmitter, receiver and specular point; *Metadata*, including sample time, GPS PRN code, and specular bin index; *Power*, including the Effective Isotropic Radiated Power (EIRP) of the GPS transmitter and antenna parameters of the receiver; *Wind Field*, a gridded wind field in latitude/longitude coordinates. *Geometry, Metadata* and *Power* (except the receiver antenna patterns) are all from CYGNSS Level 1 data. The gridded wind field has a resolution of 0.125 degree in geodetic coordinates. The forward model has two outputs: a modeled DDM produced by the forward operator and a Jacobian matrix computed by analytically differentiating the forward operator. Specific variables names from the CYGNSS Level 1 data are listed in the Appendix. Fundamental code from the CYGNSS End-to-End Simulator (E2ES) [26] has been reused.

A. Forward Operator

In a GNSS-R receiver, the reflected signal, $u_r(t)$, is first cross-correlated with a local copy of transmitted signal over a range of delays, τ and Doppler frequencies, f

$$Y_k(\tau,f) = \frac{1}{T_i}\int_{t_k-T_i}^{t_k} u_r(t)a(t+\tau)e^{2\pi j(f_0+f)t}dt. \quad (1)$$

$Y_k(\tau,f)$ is the complex correlation result at time t_k, $a(t)$ is the baseband signal model (generated from a pseudorandom noise (PRN) code using Binary Phase Shift Keyed (BPSK) modulation) and T_i is the integration time. N sequential complex results are then incoherently averaged.

$$Z(\tau,f) = \frac{1}{N}\sum_{k=1}^{N}|Y_k(\tau,f)|^2 \quad (2)$$

For CYGNSS, T_i = 1 ms and $N = 1000$. Generation of the Level 1a data products requires subtracting the noise floor, P_N, and calibration of the receiver gain, G_R, to produce a measurement (in units of Watts) of the DDM power at the front-end of the receiver [6], [27],

$$P_{DDM}(\tau,f) = \frac{Z(\tau,f)}{G_R} - P_N. \quad (3)$$

Assume that $P_{DDM}(\tau,f)$ is an unbiased measurement,

$$P_{DDM}(\tau,f) = h(\tau,f,\boldsymbol{x}) + \nu(\tau,f) \quad (4)$$

in which $h(\tau,f,\boldsymbol{x})$ is a model for the DDM. The vector $\boldsymbol{x}$ contains a set of wind speed values, ordered in geodetic latitude and longitude coordinates, which define the wind field on the ocean surface. The spatial resolution of the wind field used in this forward model is 0.125 degree in latitude and

longitude. ν is a zero-mean random variable that represents the error of the measurement.

With knowledge of the wind field, satellite geometries, and power parameters, a widely used model for the DDM, derived from the Kirchoff Approximation to Geometric Optics (KA-GO), can be written as a surface integral [4],

$$h(\tau, f, \boldsymbol{x}) = \frac{\lambda^2 P_t G_t G_0}{(4\pi)^3} \iint \left(\frac{G_r(\vec{\rho})}{R_t^2(\vec{\rho}) R_r^2(\vec{\rho})} \sigma^0(\vec{\rho}, x(\vec{\rho})) \right. \\ \left. \times \chi^2(\tau - \tau_g(\vec{\rho}), f - f_g(\vec{\rho})) d\vec{\rho}. \quad (5)$$

The variable of integration, $\vec{\rho}$, is a two-dimensional position vector defining points on the ocean surface. λ is the wavelength of the GPS carrier. $P_t G_t$ is the GPS EIRP which is assumed constant over the entire scattering area. G_r is the receiver antenna gain. $R_t(\vec{\rho})$ and $R_r(\vec{\rho})$ are path distances from the mean ocean surface at $\vec{\rho}$ to the transmitter and receiver, respectively. $x(\vec{\rho})$ is the wind speed at $\vec{\rho}$. $\chi^2(\tau - \tau_g(\vec{\rho}), f - f_g(\vec{\rho}))$ is the ambiguity function in which $\tau_g(\vec{\rho})$ is the "geometric" path delay through a point on the mean ocean surface at $\vec{\rho}$ and $f_g(\vec{\rho})$ is the corresponding Doppler frequency. An additional coefficient, the excess power gain, G_0, is included to account for uncalibrated biases. $G_0 = 1$ by default, but it can be adjusted to best fit the observed data.

σ^0 is the bistatic radar cross section (BRCS). In the KA-GO model, σ^0 depends on the probability density function of surface slopes, $p(\vec{s}, x)$,

$$\sigma^0(\vec{\rho}, x(\vec{\rho})) = \pi |\Re(\vec{\rho})|^2 \frac{q^4(\vec{\rho})}{q_z^4(\vec{\rho})} p(\vec{s}(\vec{\rho}), x(\vec{\rho})) \quad (6)$$

where $\Re$ is the Fresnel reflection coefficient. $\vec{q}(\vec{\rho})$ is the bisector,

$$\vec{q}(\vec{\rho}) = \frac{2\pi}{\lambda} \left(\frac{\vec{R}_t(\vec{\rho})}{R_t(\vec{\rho})} - \frac{\vec{R}_r(\vec{\rho})}{R_r(\vec{\rho})} \right) (\quad (7)$$

which can be decomposed into two components: q_z perpendicular to the surface and $\vec{q}_\perp$ parallel to the local tangent plane. The slope $\vec{s}$ can then be expressed as $\vec{s} = -\vec{q}_\perp / q_z$. Assuming an isotropic normal distribution for the surface slopes, $p(\vec{s}, x)$ is defined by a single parameter, the omni-directional MSS, m,

$$p(\vec{s}, x) = \frac{1}{2\pi m(x)} exp \left(-\frac{|\vec{s}|^2}{2m(x)} \right) (\quad (8)$$

A common empirical model [28] is used to provide a monotonic relationship between MSS and wind speed.

$$m(x) = 0.225(0.003 + 0.00508f) \quad (9)$$

$$f = \begin{cases} x & \text{if } 0 < x < 3.49 \\ 6 \cdot ln(x) - 4 & \text{if } 3.49 < x < 46 \\ 0.411 \cdot x & \text{if } x > 46 \end{cases} \quad (10)$$

where the numerical MSS value representing the isotropic case, was obtained by averaging the up-wind and cross-wind slope variances, $m = (\sigma_c^2 + \sigma_u^2)/2$.

To accelerate computation, fast Fourier transform (FFT) methods can be applied to (5) formluated as a 2D convolution [29], [30]

$$h(\tau, f, \boldsymbol{x}) = Q(\tau, f, \boldsymbol{x}) ** \chi^2(\tau, f), \quad (11)$$

between the ambiguity function $\chi^2(\tau, f)$ and

$$Q(\tau, f, \boldsymbol{x}) = \iint \left(B(\vec{\rho}) \sigma^0(\vec{\rho}, x(\vec{\rho})) \right. \\ \left. \times \delta(\tau - \tau_g(\vec{\rho})) \delta(f - f_g(\vec{\rho})) d\vec{\rho}. \quad (12)$$

$\delta()$ is the Dirac delta function and

$$B(\vec{\rho}) = \frac{\lambda^2 P_t G_t G_0}{(4\pi)^3} \frac{G_r(\vec{\rho})}{R_t^2(\vec{\rho}) R_r^2(\vec{\rho})}. \quad (13)$$

The delay-Doppler space needs to be discretized for computing (11). For application to data assimilation, the modeled DDM is also required to use the same delay-Doppler coordinates as the observed DDM, so as to enable calculation of difference between the observation and the forward model. Specular point position estimates generated by the DDMI onboard CYGNSS are very coarse and fixed to discrete values. Therefore, the true specular point is not generally located in the center of a single DDM bin [31]. Fractional precise specular bin indices, n_τ^s, n_f^s are generated in post-process and provided as part of the metadata listed in Table I [27]. n_τ^s, n_f^s are used to shift the delay-Doppler coordinates of the modeled DDM to align with those of the observed one.

Equation (11) is discretized and computed using a 2D FFT

$$h[n_\tau, n_f, \boldsymbol{x}] = Q[n_\tau, n_f, \boldsymbol{x}] ** \Xi[n_\tau, n_f] \quad (14)$$

in which

$$\Xi[n_\tau, n_f] = \chi^2(n_\tau \Delta\tau, n_f \Delta f). \quad (15)$$

$n_\tau = 1, 2, ..., N_\tau$ and $n_f = 1, 2, ..., N_f$ are indices in the discrete delay-Doppler space. $\Delta\tau$ and Δf are the delay and Doppler increments. For CYGNSS, $N_\tau = 17$, $N_f = 11$, $\Delta\tau = 0.25 T_c$ where T_c is the period of a GPS C/A code chip in seconds and $\Delta f = 1/(2T_i) = 500$ Hz.

$Q[n_\tau, n_f, \boldsymbol{x}]$ is the discrete form of the surface integral (12). When the Earth-centered Earth-fixed (ECEF) coordinates of the transmitter, receiver and specular point are known, a curvilinear system, identified as the "surface frame" (SURF), is created and used to discretize the surface near the specular point into a $r \times v$ grid at a resolution of 1 km [32]. In this paper, the total area covered by this grid is 120 km × 120 km (vs. 90 km × 90 km in [32]) in order to encompass the expected glistening zone for all scenarios.

A bilinear interpolation is used to interpolate $\boldsymbol{x}$, provided at fixed points in geodetic coordinates, to 1 km grid points in the SURF frame, $\boldsymbol{x}'$.

$$\boldsymbol{x}' = \boldsymbol{M}\boldsymbol{x} \quad (16)$$

$\boldsymbol{x}'$ is a N' × 1 vector, $\boldsymbol{x}$ is a N × 1 vector and $\boldsymbol{M}$ is a N' × N transformation matrix. N depends on the geometry, $N' = 120 \times 120 = 14400$, and the computation of $\boldsymbol{M}$ can be referenced in [32]. $Q[n_\tau, n_f, \boldsymbol{x}]$ can then be computed numerically as

$$Q[n_\tau, n_f, \boldsymbol{x}] = \Delta A \sum_{i=1}^{N'} \left(B(\vec{\rho_i}) \sigma^0(\vec{\rho_i}, x_i') D[n_\tau, n_f, \vec{\rho_i}] \right) \quad (17)$$

in which $\Delta A = 1000^2$ is the discrete increment in surface area and

$$D[n_\tau, n_f, \vec{\rho_i}] = \delta \left[\left\lfloor n_\tau - n_\tau^s - \frac{\tau_g(\vec{\rho_i}) - \tau_g^s}{\Delta \tau} \right\rfloor \right]$$
$$\times \delta \left[\left\lfloor n_f - n_f^s - \frac{f_g(\vec{\rho_i}) - f_g^s}{\Delta f} \right\rfloor \right] \quad (18)$$

$\delta[\,]$ is the Kronecker delta function and $\lfloor\,\rfloor$ stands for rounding to the nearest integer. τ_g^s and f_g^s are path delay and Doppler frequency at the specular point. τ_g and τ_g^s are in units of seconds and f_g and f_g^s are in units of Hz.

B. Jacobian

The Jacobian matrix is a partial derivative matrix of the forward operator, with respect to each wind speed in the input vector, $\boldsymbol{x}$.

$$\boldsymbol{H} = \frac{\partial \boldsymbol{h}(\boldsymbol{x})}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial h[1,1,\boldsymbol{x}]}{\partial x_1} & \cdots & \frac{\partial h[1,1,\boldsymbol{x}]}{\partial x_N} \\ \vdots & \cdots & \vdots \\ \frac{\partial h[1,N_f,\boldsymbol{x}]}{\partial x_1} & \cdots & \frac{\partial h[1,N_f,\boldsymbol{x}]}{\partial x_N} \\ \frac{\partial h[2,1,\boldsymbol{x}]}{\partial x_1} & \cdots & \frac{\partial h[2,1,\boldsymbol{x}]}{\partial x_N} \\ \vdots & \cdots & \vdots \\ \frac{\partial h[N_\tau,N_f,\boldsymbol{x}]}{\partial x_1} & \cdots & \frac{\partial h[N_\tau,N_f,\boldsymbol{x}]}{\partial x_N} \end{bmatrix}. \quad (19)$$

To compute (19), equation (14) has to be linearized with respect to $\boldsymbol{x}'$. However, calculating the derivative with $\boldsymbol{x}'$, at the 1 km resolution, would be computationally expensive. To reduce the computational cost, the SURF frame is resampled at a 10 km resolution by selecting $K = N'/100 = 144$ points out of the N' points. The resulting matrix of surface wind speeds is then "unrolled" into a $K \times 1$ vector, $\boldsymbol{X}$. Fig. 1 shows the selection of the $N'/100$ points in $\boldsymbol{X}$ and their spatial relation to $\boldsymbol{x}'$ and $\boldsymbol{x}$. Equation (14) then becomes

$$h[n_\tau, n_f, \boldsymbol{X}] = Q'[n_\tau, n_f, \boldsymbol{X}] ** \Xi[n_\tau, n_f] \quad (20)$$

in which

$$Q'[n_\tau, n_f, \boldsymbol{X}] = 100\Delta A \sum_{i=1}^{K} \left(B(\vec{\rho_i})\sigma^0(\vec{\rho_i}, X_i)D[n_\tau, n_f, \vec{\rho_i}] \right) \quad (21)$$

and the discrete increment in surface area is $100\Delta A$.

Taking the derivative of (20) with respect to each X_i, we obtain

$$\frac{\partial h[n_\tau, n_f, \boldsymbol{X}]}{\partial X_i} = W_i[n_\tau, n_f, \boldsymbol{X}] ** \Xi[n_\tau, n_f] \quad (22)$$

in which

$$W_i[n_\tau, n_f, \boldsymbol{X}] = 100\Delta A \Big(B(\vec{\rho_i})\pi \, |\Re(\vec{\rho_i})|^2 \frac{q^4(\vec{\rho_i})}{q_z^4(\vec{\rho_i})}$$
$$\times \frac{\partial P(\vec{s}(\vec{\rho_i}), X_i)}{\partial X_i} D[n_\tau, n_f, \vec{\rho_i}] \Big) \quad (23)$$

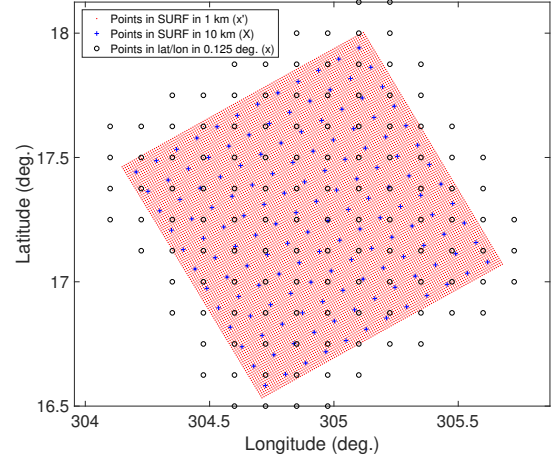Only the slope-PDF is depenent on the wind speed, so this is the only derivative which must be computed. Assuming a

Figure 1. Spatial relation of the grid points in three different resolutions: (1) grid points in SURF in 1km resolution ($\boldsymbol{x}'$); (2) grid points in SURF in 10km resolution ($\boldsymbol{X}$). (3) grid points in latitude/longitude coordinates in 0.125 degree resolution ($\boldsymbol{x}$). $\boldsymbol{X}$ are evenly distributed among $\boldsymbol{x}'$.

normal distribution allows a closed-form expression for this derivative,

$$\frac{\partial P(\vec{s}(\vec{\rho}), X)}{\partial X} = \frac{1}{2\pi m^2(X)} \left(\frac{|\vec{s}(\vec{\rho})|^2}{2m(X)} - 1 \right)$$
$$\times exp \left(-\frac{|\vec{s}(\vec{\rho})|^2}{2m(X)} \right) \frac{dm(X)}{dX} \quad (24)$$

in which

$$\frac{dm(X)}{dX} = \begin{cases} 1.143 \times 10^{-3} & \text{if } 0 < X < 3.49 \\ \frac{6.858 \times 10^{-3}}{X} & \text{if } 3.49 < X < 46 \\ 4.69773 \times 10^{-4} & \text{if } X > 46. \end{cases}$$

A matrix, $\boldsymbol{H_0}$, is formed from $N_\tau N_f \times K$ (187×144 in the particular application to CYGNSS) elements of (22). The last step is to transform this matrix to geodetic coooordinates. With the coordinate transformation in [32], the latitude and longitude of each point in $\boldsymbol{X}$ can be calculated. The transformation matrix $\boldsymbol{T}$ from $\boldsymbol{x}$ (at 1 km) to $\boldsymbol{X}$ (at 10 km)

$$\boldsymbol{X} = \boldsymbol{Tx} \quad (25)$$

can be obtained from $\boldsymbol{M}$ averaging blocks of 100 terms.

$$T_{ij} = \frac{1}{100} \sum_{i' \in I_i} M_{i'j} \quad (26)$$

I_i is the set of indices of the 100 points in $\boldsymbol{x}'$ that are neareast to $\boldsymbol{X}_i$. $M_{i'j}$ is the i'-th row, j-th column element in matrix $\boldsymbol{M}$, $i = 1, 2, ..., K$, $i' = 1, 2, ..., N'$, $j = 1, 2, ..., N$, and $\boldsymbol{T}$ is a matrix in dimension of $K \times N$. Applying the chain rule

$$\frac{\partial h[n_\tau, n_f, \boldsymbol{x}]}{\partial x_i} = \frac{\partial h[n_\tau, n_f, \boldsymbol{X}]}{\partial X_1} \frac{dX_1}{dx_i} + \frac{\partial h[n_\tau, n_f, \boldsymbol{X}]}{\partial X_2} \frac{dX_2}{dx_i}$$
$$+ ... + \frac{\partial h[n_\tau, n_f, \boldsymbol{X}]}{\partial X_K} \frac{dX_K}{dx_i}, \quad (27)$$

the Jacobian matrix (19) can be obtained as

$$\boldsymbol{H} = \boldsymbol{H_0 T}. \quad (28)$$

III. MODEL ASSESSMENT

To assess the forward model and identify possible difficulties in assimilating DDMs, the model is applied to Hurricane Weather Research and Forecasting (HWRF) synoptic-scale (0.125 degree resolution) wind fields [33], [34] and the resulting DDMs are then compared against CYGNSS Level 1a data (version 2.1). CYGNSS measurements collected during overpasses of Hurricane Maria on September 23, 2017 and Cyclone Gita on February 12, 2018 are chosen for this comparison. One track in Maria covers the time from 18:03 to 18:08 UTC and the corresponding HWRF wind field is from the analysis at 18:00 UTC. For Cyclone Gita, the track covers the time from 14:03 to 14:05 UTC and is compared with winds from the HWRF 2-hour forecast at 14:00 UTC. The picked two tracks are the best tracks with good data quality that overpass the center of the tropical cyclones during the time. Since the time difference between the model and measurement is less than 10 minutes in this experiment, change of the wind field in such a short time period can be ignored. The CYGNSS specular tracks and wind fields for both cases are shown in Fig. 2. Measurements from GPS Block IIF satellites are removed from consideration because of the inaccurate estimation of the fluctuating transmitter power. The analysis wind field is reported with a resolution of 0.125 deg.

A. Assessment of forward operator

Data for each variable in Table I were obtained from the CYGNSS L1a product at a 1 Hz rate. The receiver antenna patterns estimated by pre-launch measurements and on-orbit corrections were provided by the CYGNSS SOC [27]. At each time of observation, the modeled DDM produced by the forward operator was compared with that obtained from the L1a data. Fig. 3 shows examples of the forward operator for three different cases inside Cyclone Gita exhibiting average wind speeds in the glistening zone of 6 m/s (a), 15 m/s (b) and 30 m/s (c). It can be observed that for all cases, the modeled and observed DDMs have similar power values and asymmetric shapes. Two metrics were used to quantify the agreement between the forward operator and observations. To compare the absolute power between observed and modeled DDMs, the average relative difference of effective bins was computed,

$$\epsilon_k = \frac{1}{N} \sum_{i,j} \frac{y_k(\tau_i, f_j) - h_k(\tau_i, f_j)}{y_k(\tau_i, f_j)}, \qquad (29)$$

where k is the index of the DDM, y_k and h_k are observed and modeled DDM, i and j are indices of effective bins and N is the number of effective bins. Effective bins are defined as DDM samples having a power larger than $1/10$ of the DDM peak. Fig. 4 shows the selection of DDM bins using this method for two different cases. To compare the shapes of the two DDMs, the correlation coefficient, ρ_k, between the effective bins of the two DDMs is computed from

$$\rho_k = \frac{\sum_{i,j}(y_k(\tau_i, f_j) - \bar{y})(h_k(\tau_i, f_j) - \bar{h}_k)}{\sqrt{(\sum_{i,j}(y_k(\tau_i, f_j) - \bar{y})^2)(\sum_{i,j}(h_k(\tau_i, f_j) - \bar{h}_k)^2)}}$$
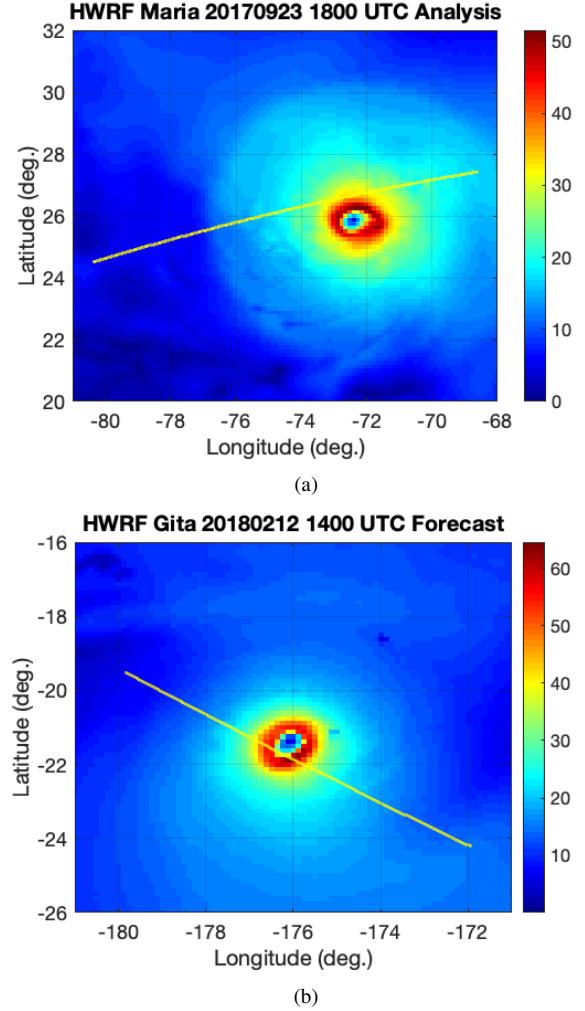$$(30)$$

(a)

(b)

Figure 2. HWRF wind field and overpass of CYGNSS specular points (yellow lines): (a) Hurricane Maria on September 23, 2017 at 18:00 UTC and CYGNSS data on same day at around 18:06 UTC with CYGNSS SV 5 and GPS PRN 13. (b) Hurricane Gita on February 12, 2018 at 14:00 UTC and CYGNSS data on same day at around 14:04 UTC with CYGNSS SV 1 and GPS PRN 23.

where $\bar{y}$ and $\bar{h}_k$ are mean values of all effective bins of the observed and modeled DDM [31]. When the shapes of the two DDMs are more similar, the coefficient should be closer to 1.

Fig. 5 shows the average relative differences, ϵ_k, for the two tracks in Hurricane Maria and Cyclone Gita. Fig. 6 shows the correlation coefficients, ρ_k, for these same tracks. Wind speeds at the specular point along these tracks are also plotted on these figures. In Fig. 5, it can be seen that the relative difference mostly remains below 30%, except for some low wind speed cases for Hurricane Maria. The correlation coefficient plotted in Fig. 6 is mostly larger than 0.9, while there is a decrease in high wind speed cases. Both figures show that the DDMs calculated by the forward operator are generally in agreement with CYGNSS measured ones.

Under the ideal case, the only cause of the difference between observation and model is the difference between the model wind field and the real winds. Adjustment of the model winds, to minimize this difference, is the essential principle of data assimilation. However, there are several other uncer-
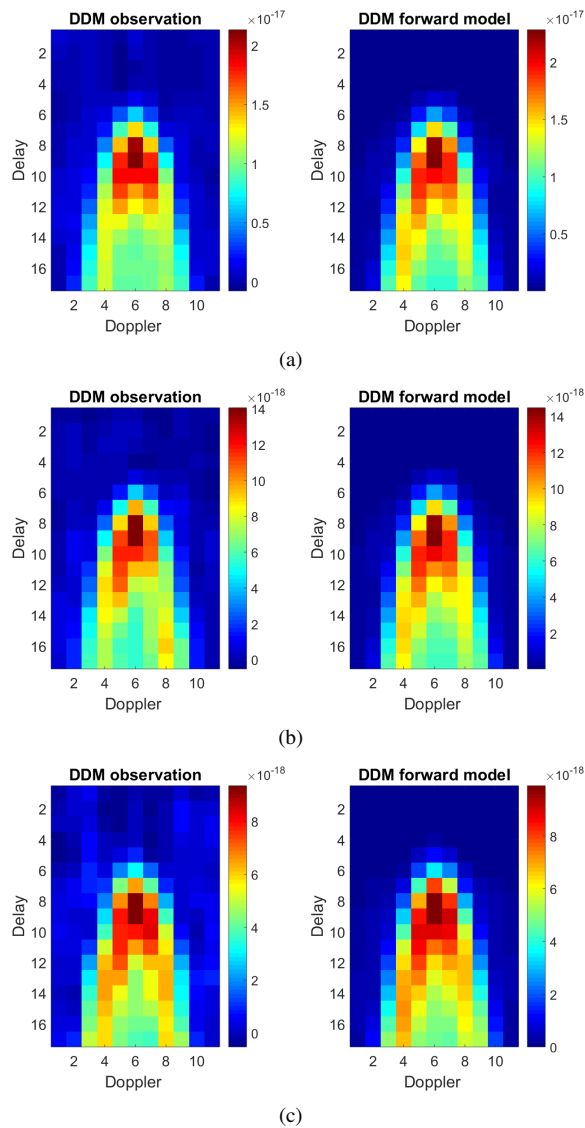
Figure 3. Comparisons between observed DDM and modeled DDM at three different cases for Cyclone Gita on February 12, 2018: wind speed average over the glistening zone = 6 m/s (a), 15 m/s (b), and 35 m/s (c). DDM values are in the unit of Watt.
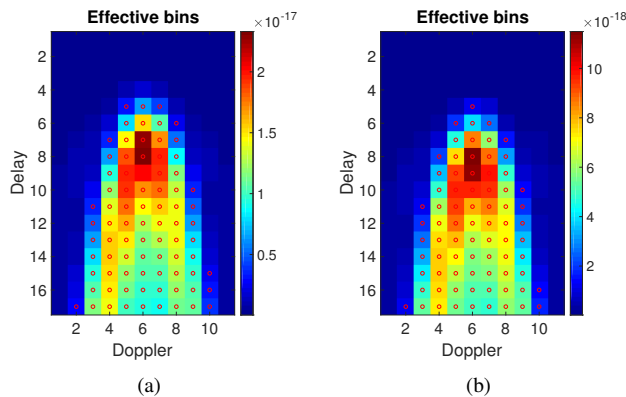

Figure 4. Selection of effective DDM bins (red circles), using 1/10 of the peak power in two cases for Cyclone Gita on February 12, 2018: wind speed = 5 m/s (left) and 30 m/s (right). DDM values are in the unit of Watt.
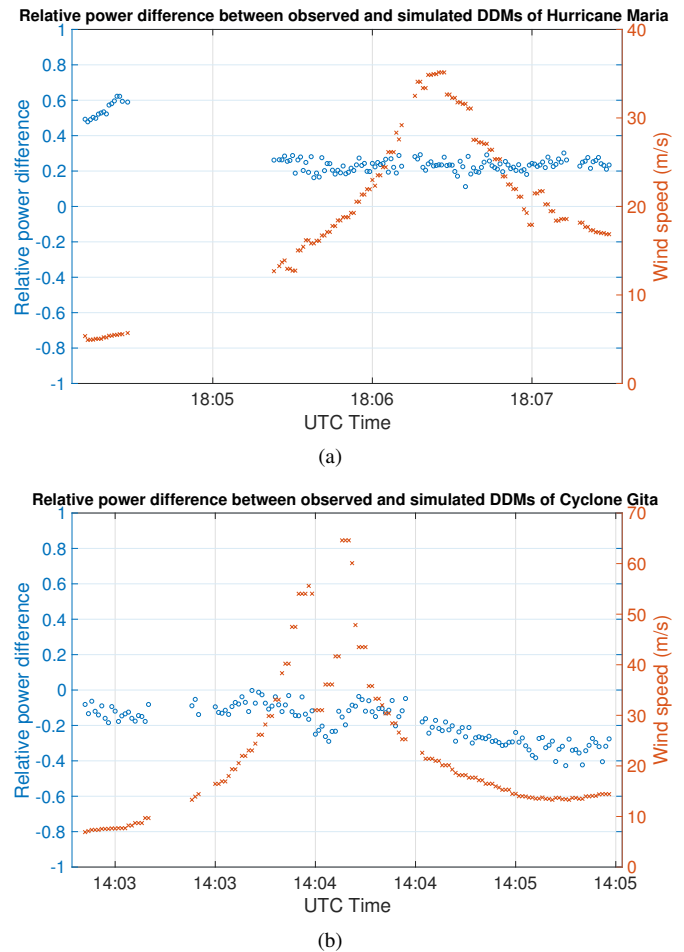

Figure 5. Relative power differences between modeled DDMs and observed DDMs of the two tracks in (a) Hurricane Maria and (b) Cyclone Gita.

tainties that influence the performance of the forward model. These must be understood and accounted for in any practical data assimilation scheme. In our comparison, we selected cases in which the time difference between observation and model (less than 10 minutes) are small enough such that temporal change of wind field can be ignored. The HWRF analysis (Maria) and 2-hour short forecast (Gita) were found to be highly consistent with the best track data according to a verification report [35]. Using these examples, we evaluated four (4) sources of uncertainty:

1) Low wind speed: swell and specular reflection: Ocean surface roughness is not only sensitive to short waves driven by the local wind but also to long waves forced by non-local swell [36], [37]. At medium or high wind speed cases, the local wind plays the dominant role and the effect of swell can be ignored. However, at very low wind speed cases ($\lesssim$ 5 m/s), the local wind is weak and the swell may be evident. In the forward model, MSS is computed as an empirical function of only the local wind speed [28], incorporating no information about swell. Ocean swell can be characterized by the significant wave height (SWH) or MSS from a wave model (i.e. WAVEWATCH III [38] or WAM [39]).
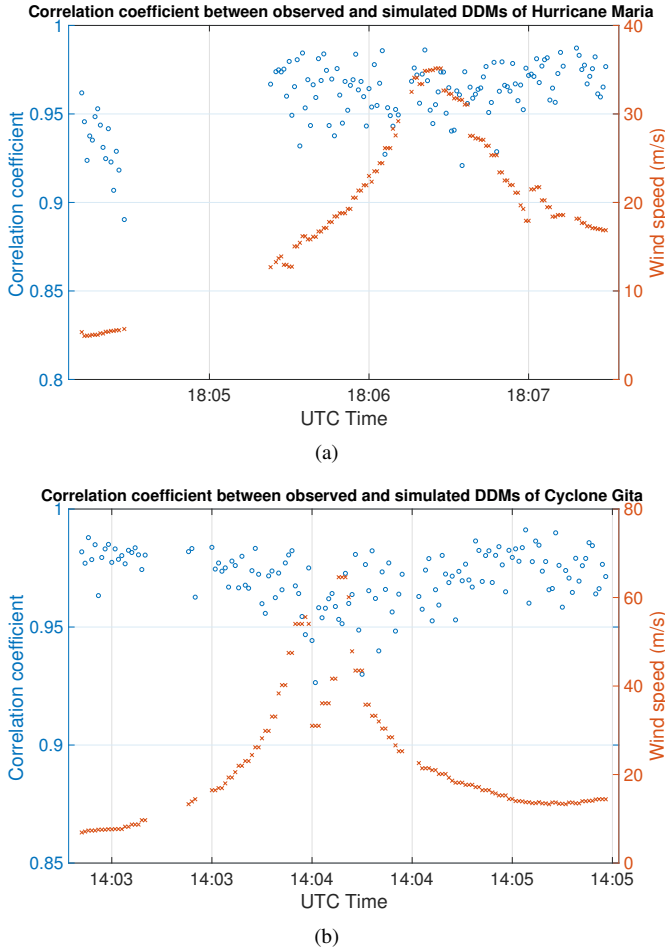
Figure 6. Correlation coefficients between modeled DDMs and observed DDMs of the two tracks in (a) Hurricane Maria and (b) Cyclone Gita.
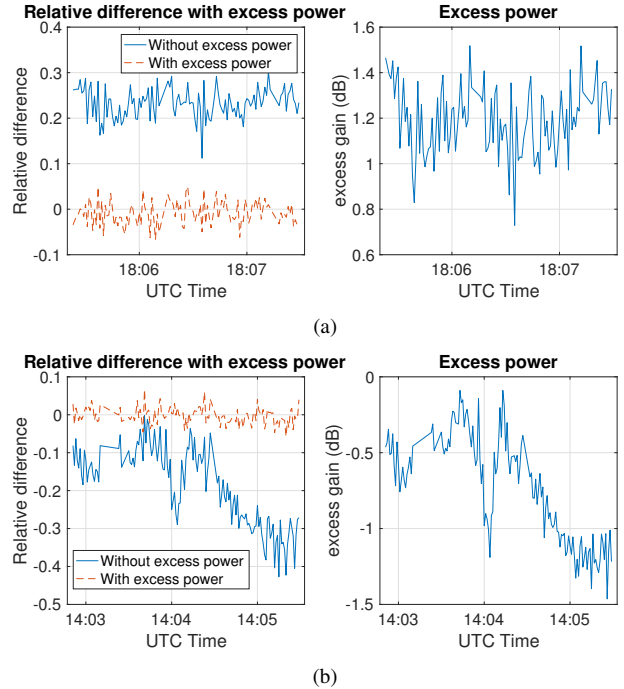

Figure 7. Relative power differences with and without adding the excess power in each modeled DDM and the excess power along the track: (a) Maria (b) Gita.

Bayesian estimation has also been applied to correct for the effect of long waves using SWH [40]. Those methods need auxiliary data of the sea state in addition to the wind speed, which could complicate the forward model. At the present time, data assimilation is parameterized in terms of wind speed (or wind vector) only. Furthermore, at low wind speed, the ocean surface may be smooth enough such that coherent scattering exists and the KA-GO model would no longer be applicable [41]. These un-modeled effects are possible causes of the high relative differences observed for low wind speeds as shown in Fig. 5. To avoid the effect of swell and coherent scattering, quality controls on the background wind speed (i.e., a threshold on the minimum wind speed) or the relative difference in equation (29) can be used to filter out the cases which may exhibit those effects.

2) DDM power bias: A bias is observed in both cases in Fig. 5, in which the relative power differences are all above zero for Maria and all below zero for Gita. One of the major sources for this bias is the inaccurate estimated GPS transmitter EIRP. For CYGNSS, the GPS EIRP is estimated using measurements from the zenith navigation antenna on each CYGNSS satellite and a ground-

based GPS power monitor [42], [43]. Accuracy of the EIRP estimate can be influenced by variation in the transmitter power and calibration of the zenith antenna pattern. To evaluate this influence, the excess power gain G_0 in equation (14) was estimated independently. For each modeled DDM, the excess power is adjusted by fitting all effective bins in the modeled DDM against the observed DDM using a least square method. Fig. 7 shows the relative power differences with and without adjusting G_0 and the excess power along the track for Hurricane Maria and Cyclone Gita. Samples for wind speed at the specular point under 5 m/s have already been removed from this comparison, due to aforementioned concerns about swell and coherent scattering. The relative differences become much closer to zero and the bias is removed after adding the excess power for both cases. It shows that the excess power parameter in the forward model can be used to effectively correct the bias. Errors in the receiver antenna pattern, resulting from changes from exposure to the space environment, for example, would also introduce different power error of each DDM bin. These can be assumed to remain constant through the mission and can thus their effect on the bias can also be calibrated. In the initial studies in this paper, a maximum threshold on the relative power difference in equation (29) will be set for quality control.

3) Inaccurate specular point location: The specular point delay and Doppler, reported by the DDMI, are very coarse, due to the simplified geometry model used for on-board open-loop tracking and the limited precision of discrete steps (0.25 chip, 500 Hz). Specular point
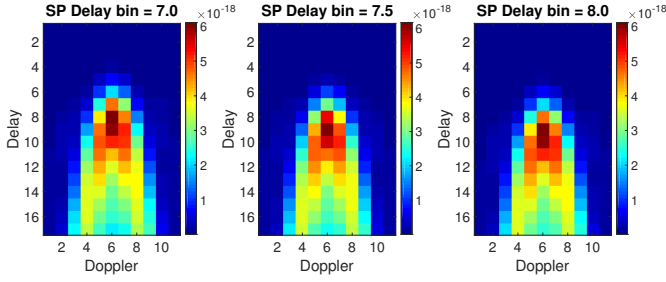
Figure 8. DDMs computed by the forward model using specular delay bins of 7.0, 7.5, 8.0 (from left to right). One example of the Hurricnae Maria case was selected with the specular point at 26.5°N, 72.8°W.
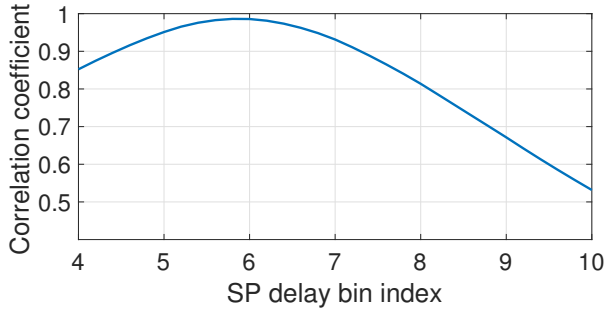

Figure 9. Correlation coefficients between the observed DDM and modeled DDM using different specular delay bin index with a delay step of 0.2 bin. One example of the Hurricnae Maria case was selected with the specular point at 26.5°N, 72.8°W.

Figure 10. An example of the Jacobian matrix computed from the Hurricane Maria case when the specular point is near the hurricane eyewall. The specular point is at 26.5°N, 72.8°W.

delay and Doppler are re-computed on the ground by the CYGNSS SOC using a more accurate method of specular point calculation and a higher fidelity mean sea surface model [27]. These higher precision specular bin indices, n_τ^s, n_f^s, are obtained from the CYGNSS L1a metadata and used in the forward model (14). However, some errors still remain as a result of errors in the sea surface height model, local height variation, and C/A code tracking error, resulting in some misalignment between the observed and modeled DDMs. Fig. 8 illustrates this effect and shows change in the modeled DDMs with displacement of the specular point delay in the step of 0.5 bin. Fig. 9 shows the correlation coefficients between the observed DDM and modeled DDM using different specular delay bin index. Correlation coefficient between the observed and modeled DDM can be used as quality control to filter out cases with large error in the specular point position, as this metric represents the shape similarity between the two.

4) High wind speed: The empirical relationship between wind speed and MSS in equation (9) was derived by fitting a large collection of airborne GNSS-R observations from high-elevation satellites against wind speed measurements [28]. In the very high wind speed region near the hurricane eyewall, the model should be used cautiously because of the complicated sea state, breaking of the waves or unavailability of the sea surface slope information [44]. The decrease in correlation coefficient in Fig. 6 when the wind speed is high could be possibly
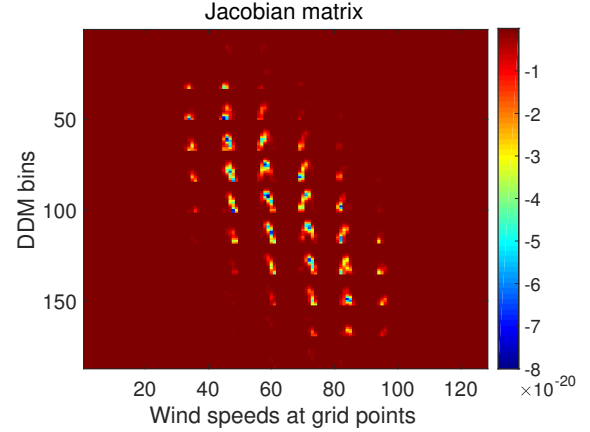
caused by this model representativeness error. Discarding observations with background wind speed exceeding a threshold ($\gtrsim$ 35 m/s) can be used as an additional quality control filter.

The discussion above shows that, while the forward model accurately represents the relationship between the scattered and surface conditions over the full extent of the DDM in most cases, there are some important discrepancies which must be understood and accounted for. We propose that these effects can be mitigated through limiting the useable range of wind speeds, estimating the excess power gain, G_0, and defining quality control tests on the relative power difference and correlation coefficient between the observed and modeled DDM.

B. Validation of the Jacobian

The Jacobian matrix represents the sensitivity of each sample of the DDM to the wind speed at each surface grid point. An example of the matrix, shown in Fig. 10, is computed for Hurricane Maria at a time when the specular point is near the hurricane eyewall. The two dimensions of this matrix are the DDM values (organized with delay and Doppler unrolled into a vector) and surface wind speeds (organized with latitude and longitude unrolled into a vector). Some rows and columns in the matrix are zero because they correspond to delays earlier than the specular point or later than the largest iso-delay ellipse. The scattered power in some is so small that they are dominated by noise and thus the Jacobian is effectively zero. High absolute values in Fig. 10 matrix represent the sensitivity of the DDM bins to wind speeds at specific locations falling within the corresponding delay and Doppler range.

To visually present the matrix, a given column representing the sensitivity of all DDM samples to the surface wind speed at a specific surface point, can be expanded into a matrix in delay and Doppler coordinates. Two examples are shown in Fig. 11. This shows that information about one surface point is present in multiple DDM samples with some degree of independence. Similarly, a given row representing the sensitivity of a specific sample of the DDM to a region of
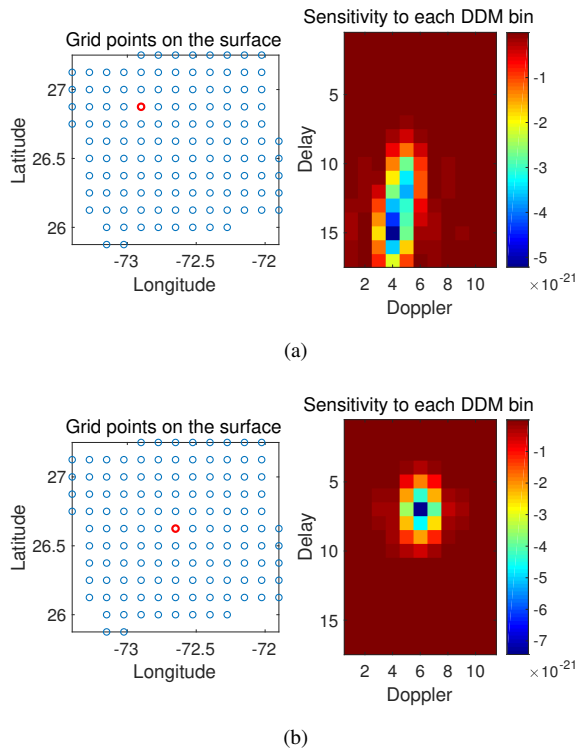
(a)

(b)

Figure 11. Two cases for sensitivity of wind speed in a grid point to all DDM bins computed from the Jacobian matrix in Fig. 10. In each case, a grid point (red circle) of the wind field is selected in the left plot. The sensitivity of the power in all DDM bins with respect to wind speed of the selected grid point is shown in the right plot.
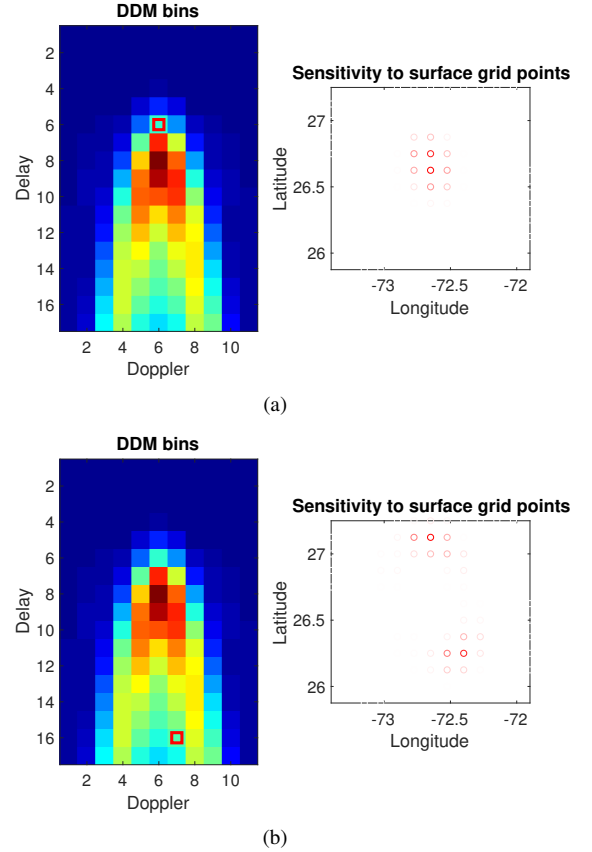

(a)

(b)

Figure 12. Two cases for sensitivity of a DDM bin power to wind speeds on all grid points computed from the Jacobian matrix in Fig. 10. In each case, a DDM bin (red square) is selected in the left plot; The senstivity of wind speed in the wind field grid with respect to the power of the selected DDM bin is shown in the right plot, where the saturation of the red color represents the absolute value of the sensitivity.

wind speeds on the surface grid is shown in Fig. 12. Note that the delay-Doppler ambiguity can be observed in Fig. 12(b) in which two geographically separated regions on the surface both contribute to the observation made at one delay and Doppler.

The Jacobian matrix computed by equation (28) was derived in a closed form by differentiating the forward operator, assuming all geometries, power and wind speed are homogenous within each 10 km cell, rather than the 1 km grid used in the forward operator. This linearization and the 10-km approximation may induce errors in the computation of the Jacobian matrix. To validate the Jacobian matrix, the matrix is computed by the finite difference of the forward model (at 1 km resolution) using a step of 0.0001 m/s. Fig. 13 shows a comparison of Jacobian matrices by finite difference vs. that from the analytical form. It can be seen that they have similar patterns. The relative difference is shown in Fig. 13(c). After discarding the meaningless zero values in the matrix, the relative error is about 0.19 in average and the correlation coefficient between the two matrices is 0.92. In the same way, Jacobian matrices of the track of 100 DDMs for Hurricane Maria were validated by the finite difference with an average relative error is 0.20 and average correlation coefficient is 0.90 for the track. Although the analytical form may exhibit some error due to the larger grid size, it is much efficient with approximately one-tenth of the computation time of the finite difference method.
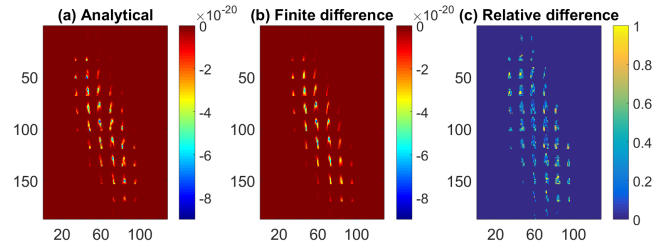

Figure 13. Comparison of the Jacobian matrices computed by finite difference and analytical form for the Jacobian matrix in Fig. 10. The three plots are the analytical version (a), finite difference version (b) and the relative difference between the two (c).

IV. MODEL APPLICATION

In this section, we will demonstrate application of the forward model in data assimilation using a Variational Analysis Method (VAM) [20], [45]. Our intent is to define the algorithm and provide examples demonstrating that observations from the full DDM (CYGNSS Level 1a data products) can be successfully used with the VAM to improve estimates of the ocean wind field. With this goal in mind, we have selected a few specific test cases in which there are large deviations (Root Mean Square Error (RMSE) ≥ 2.0 m/s) between ECMWF forecast model winds and scatterometer observations, and then

show the reduction in this difference after processing the observed DDMs through the VAM. These results should be interpreted as a demonstration of this approach and not as a general assessment of its performance over a large, globally distributed, set of winds as would be necessary to make a quantitative and statistically significant assessment. Such large scale studies, however, are presently in progress.

A. Method

VAM is a tool to assimilate multi-platform measurements into a vector wind field [20]. Wind vectors are usually more valuable than wind speeds in data assimilation as they implicitly provide additional dynamical information (divergence, vorticity) of the wind field. VAM finds the minimum of the objective cost function J,

$$J(\boldsymbol{\beta}) = J_b(\boldsymbol{\beta}) + J_{ddm}(\boldsymbol{\beta}) + J_c(\boldsymbol{\beta}), \tag{31}$$

where $\boldsymbol{\beta} = [u_1, u_2, ..., u_N, v_1, v_2, ..., v_N]$ is a vector containing the analysis wind field in u (East) and v (North) components. N is the number of grid points in the wind field. $J_b(\boldsymbol{\beta})$ is the wind background term presenting the misfit between the analysis and the backgound,

$$J_b(\boldsymbol{\beta}) = \lambda_b \frac{1}{\sigma_\beta^2} \left(\boldsymbol{\beta} - \boldsymbol{\beta}^b\right)^T \left(\boldsymbol{\beta} - \boldsymbol{\beta}^b\right) \tag{32}$$

where $\boldsymbol{\beta}^b$ is the background wind field vector and σ_β^2 is the variance of each wind vector component in the background. $J_{ddm}(\boldsymbol{\beta})$ is the DDM observation term presenting the misfit between the analysis and the observation,

$$J_{ddm}(\boldsymbol{\beta}) = \lambda_{ddm}(\boldsymbol{h_\beta}(\boldsymbol{\beta}) - \boldsymbol{y})^T \boldsymbol{R}^{-1}(\boldsymbol{h_\beta}(\boldsymbol{\beta}) - \boldsymbol{y}) \tag{33}$$

in which $\boldsymbol{y}$ is the DDM observation organized into a vector and $\boldsymbol{R}$ is the DDM observation error covariance matrix. $\boldsymbol{h_\beta}(\boldsymbol{\beta}) = \boldsymbol{h}(\boldsymbol{x})$ is the vector of DDM samples computed by the forward operator in section II-A using the analysis wind speed field,

$$\boldsymbol{x} = \left[\sqrt{u_1^2 + v_1^2}, \sqrt{u_2^2 + v_2^2}, ..., \sqrt{u_N^2 + v_N^2}\right]^T. \tag{34}$$

$J_c(\boldsymbol{\beta})$ is the smoothness and constraint term including Laplacian, divergence, and vorticity of the difference between the background and analysis,

$$J_c(\boldsymbol{\beta}) = \lambda_{lap}J_{lap}(\boldsymbol{\beta}) + \lambda_{div}J_{div}(\boldsymbol{\beta}) + \lambda_{vor}J_{vor}(\boldsymbol{\beta}) \tag{35}$$

where $J_{lap}(\boldsymbol{\beta})$, $J_{div}(\boldsymbol{\beta})$, and $J_{vor}(\boldsymbol{\beta})$ are computed as described in [20]. λ_b, λ_{ddm}, λ_{lap}, λ_{div} and λ_{vor} are weighting factors of each term. The J_{div} and J_{vor} terms are derived from the Navier-Stokes equations for viscous fluid motion and therefore implicitly constrain the VAM to find solutions that in part satisfy those equations. The combination of terms measuring departures from the background, J_b and J_c, in the VAM define the background error covariance as in more traditional 3D-VAR or 4D-VAR analysis systems. Because the VAM is tailored for 2 dimensions and only for the surface wind components (u, v), it can be viewed as a specialized version of 3D-VAR.

To minimize the cost function $J(\boldsymbol{\beta})$, the gradient has to be calculated,

$$\frac{\partial J(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \frac{\partial J_b(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} + \frac{\partial J_{ddm}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} + \frac{\partial J_c(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \tag{36}$$

specially,

$$\frac{\partial J_{ddm}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = 2\lambda_{ddm}\boldsymbol{H}_\beta^T \boldsymbol{R}^{-1}(\boldsymbol{h_\beta}(\boldsymbol{\beta}) - \boldsymbol{y}) \tag{37}$$

where $\boldsymbol{H}_\beta$ is the partial derivative matrix respect to u, v components and can be related to the Jacobian matrix $\boldsymbol{H}$ computed in section II-B by

$$\boldsymbol{H}_\beta = \begin{bmatrix} \boldsymbol{H}_{1,*} & L_u & \boldsymbol{H}_{1,*} & L_v \\ \boldsymbol{H}_{2,*} & L_u & \boldsymbol{H}_{2,*} & L_v \\ \vdots & & \vdots & \\ \boldsymbol{H}_{P,*} & L_u & \boldsymbol{H}_{P,*} & L_v \end{bmatrix} \tag{38}$$

$\boldsymbol{H}_{i,*}$ is the i-th row of the $\boldsymbol{H}$ matrix. Row vectors

$$L_u = \left[\frac{u_1}{x_1}, \frac{u_2}{x_2}, ..., \frac{u_N}{x_N}\right] \tag{39}$$

and

$$L_v = \left[\frac{v_1}{x_1}, \frac{v_2}{x_2}, ..., \frac{v_N}{x_N}\right] \tag{40}$$

are partial derivatives of $\boldsymbol{x}$ respect to the u and v components. The operator represents the element-wise multiplication between two vectors and $P = N_\tau N_f$ is the number of DDM bins.

The initial value of the analysis $\boldsymbol{\beta}$ is set to equal to the background. The minimization problem is solved by a Quasi-Newton algorithm [46].

B. Experiment description

10-m ocean surface wind vectors provided by ECMWF forecasts [47] were used as the background and DDMs from the CYGNSS v2.1 Level 1 data were used for the observation. Scatterometer (SCAT) winds from ASCAT and OSCAT 25 km wind products were used for validation. The SCAT wind components have been validated to have a bias less than 0.01 m/s and error standard deviation of less than 1 m/s using a triple collocation [48], [49]. Three 20-minute periods of data from June 2017 are chosen for the experiment. Collocations have been made between CYGNSS specular points and SCAT wind vector cell (WVC) with a criterion of 25 km in distance and 40 minutes in time. All CYGNSS data with quality flags not equal to zero and those from GPS block-IIF transmitters (with larger power estimation error) have been filtered out. As mentioned in Section III, additional quality control checks were applied to remove DDMs for which the forward model can not be expected to be valid. These checks were applied prior to processing each DDM: 1) Background wind speed at the specular point between 2 m/s and 35 m/s; 2) Maximum relative power difference $< 100\%$; 3) Correlation coefficient > 0.9; 4) Signal-to-noise ratio (SNR) ≥ 3 dB and 5) Incidence angle $\leq 60°$. About 28% of the data were discarded by these quality control checks.

The original ECMWF forecasts were quadratically interpolated into the center time of every 20-minute period and
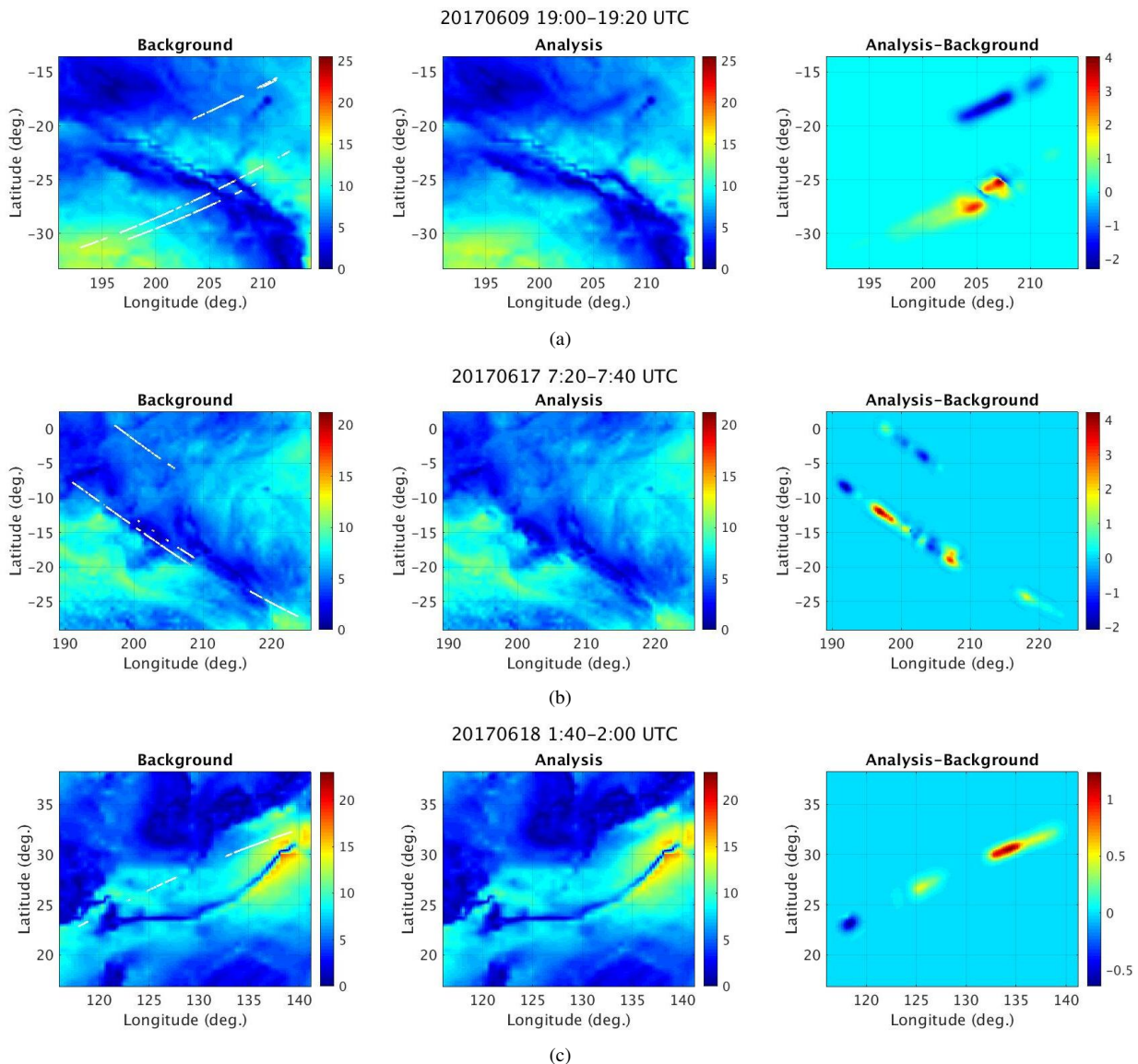
Figure 14. Wind field maps of the ECMWF background, analysis from DDM assimilation and increment of three cases at (a) 19:10 UTC on June 9, 2017, (b) 7:30 UTC on June 17, 2017 and (c) 1:50 UTC on June 18, 2017. The CYGNSS specular point tracks are shown as the white circles in the background maps.

bilinearly interpolated onto the 0.125 degree spatial grid. CYGNSS DDMs were then assimilated into the background using the VAM, as described in IV-A, to produce the analysis. Wind speed at the CYGNSS specular points from both the background and analysis are then compared with SCAT winds. Three example cases, all with an RMSE between background wind field and SCAT data larger than 2 m/s, were selected to serve as a demonstration of possible improvements with data assimilation. Center times of the three cases were 19:10 UTC at June 9 (a), 7:30 UTC at June 17 (b) and 1:50 UTC at June 18 (c).

In the VAM cost function (31), the standard deviation of the ECMWF background wind component σ_β is set to 1 m/s. The DDM error covariance matrix is assumed to be diagonal with relative error in the DDM observation of 10%.

$$\boldsymbol{R} = \mathrm{diag}[(0.1\boldsymbol{y})^2] \qquad (41)$$

diag() indicates a 187×187 diagonal matrix generated from a vector. The weighting factors are set through trial and error, based upon the resulting improvement in the analysis wind field. Specific numbers used in these results are: $(\lambda_b, \lambda_{ddm}, \lambda_{lap}, \lambda_{div}, \lambda_{vor}) = (4, 1/4, 200, 400, 100)$.

C. Results

In the experiment, 579 DDMs in case (a), 593 DDMs in case (b) and 175 DDMs in case (c) were assimilated. The wind field maps of the background, analysis and increment in each case are shown in Fig. 14. The CYGNSS specular point tracks are also shown in the background maps in each case. It can be observed that the assimilation of a track of DDMs can impact a swath of wind field with a width of around 200 km. The comparison between the background (ECMWF) winds and analysis (ECMWF-CYGNSS) winds to the SCAT
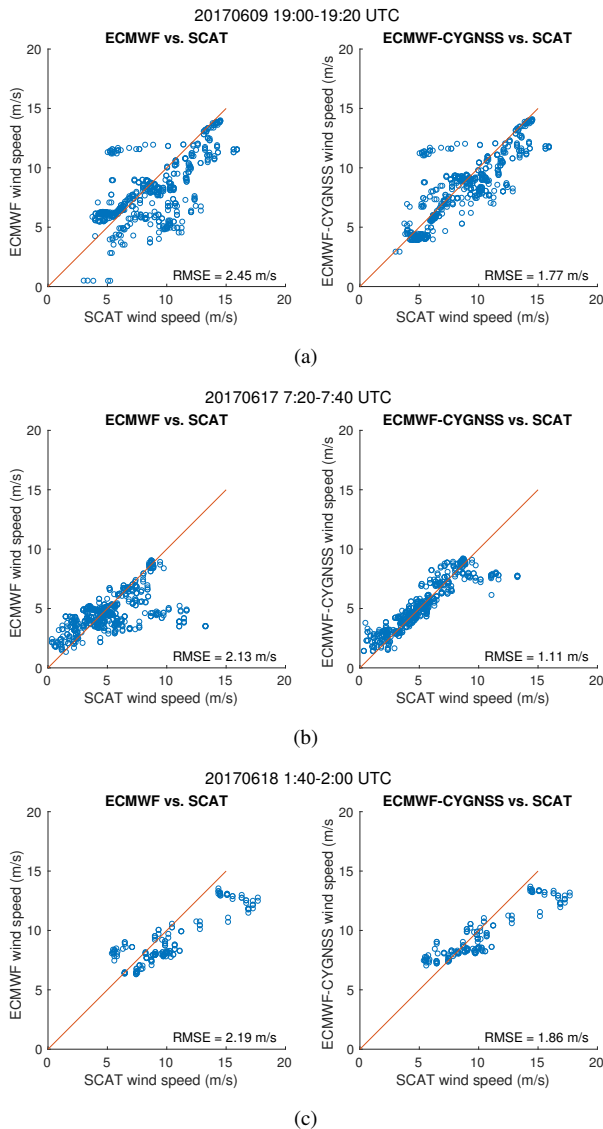
Figure 15. Scatter plots of the background (ECMWF) winds and analysis (ECMWF-CYGNSS) winds versus the scatterometer (SCAT) winds ("truth") in the three cases at (a) 19:10 UTC on June 9, 2017, (b) 7:30 UTC on June 17, 2017 and (c) 1:50 UTC on June 18, 2017. The RMSE is listed at the bottom-right angle in each plot.

winds ("truth") at the specular points is shown in Fig. 15. It shows that the wind speeds in the analysis are more correlated with the "truth" than are the background winds. The RMSE decreases from 2.45, 2.13, and 2.19 m/s to 1.77, 1.11, and 1.86 m/s after the DDM assimilation in the three cases. Although three cases are not sufficient to make a general conclusion about the benefit of DDM assimilation to numerical weather prediction (NWP), they are presented here to demonstrate that this approach is possible and that it can show improvements in certain cases with large initial errors.

V. CONCLUSION

A generalized forward model for assimilating GNSS-R DDMs into numerical weather models has been presented. Assimilation of full DDMs could incorporate more information into forecast models than what is provided by baseline CYGNSS Level 2 wind products. The forward model takes in a small set of satellite parameters (e.g. CYGNSS Level 1a products and antenna patterns) and a gridded wind field. Its output is the modeled DDM and Jacobian matrix.

Performance of the forward model was assessed on two tracks of CYGNSS data from Hurricane Maria and Cyclone Gita by comparing observed and computed DDMs. Results of this comparison show agreement with relative differences under 30% and correlation coefficients larger than 90%. Limitations of the forward model have been presented, including effects of swell, observation bias, specular positioning error and limitations of the KA-GO theory especially at high wind speed cases. Quality controls on wind speed, relative power difference and correlation coefficient between the observed and modeled DDM were defined based upon these deficiencies. An analytical Jacobian matrix was validated by comparison with one computed by finite differences. The average relative error between the analytical version and a Jacobian computed by finite difference is 20%. The average correlation coefficient between the two is 0.90.

A VAM was introduced to demonstrate assimilation of DDMs using the forward model. Three examples exhibiting large differences bewteen the ECMWF background and scatterometer wind fields were selected for this demonstration. All three results shows improvement in agreement with scatterometer data, reducing the RMSE by 15, 28 and 48%. These examples demonstrate the mechanism and show the potential benefit of assimilating DDMs directly into NWP models. A larger study of this improvement, over a globally-distributed set of observations, will be necessary to provide a statistically significant evaluation.

Although the model was developed for CYGNSS, it can be applied to any other spaceborne GNSS-R missions which provide DDM measurements. As long as the satellite geometries, power parameters and metadata listed in section II are also provided, the DDMs can be assimilated into NWP models by the forward model.

APPENDIX

The variables in the CYGNSS Level 1 v2.1 data used by the forward model in section II are listed in Table I.

Table I
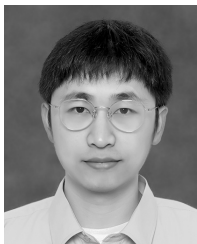PARAMETERS OF CYGNSS LEVEL 1 DATA USED IN THE FORWARD MODEL

Data Type	Variable Name	Type	Units	Description	Symbol
Geometry	tx_pos_x	int	meter	Transmitter posistion X	(none)
Geometry	tx_pos_y	int	meter	Transmitter posistion Y	(none)
Geometry	tx_pos_z	int	meter	Transmitter posistion Z	(none)
Geometry	tx_vel_x	int	m/s	Transmitter velocity X	(none)
Geometry	tx_vel_y	int	m/s	Transmitter velocity Y	(none)
Geometry	tx_vel_z	int	m/s	Transmitter velocity Z	(none)
Geometry	sc_pos_x	int	meter	Receiver posistion X	(none)
Geometry	sc_pos_y	int	meter	Receiver posistion Y	(none)
Geometry	sc_pos_z	int	meter	Receiver posistion Z	(none)
Geometry	sc_vel_x	int	m/s	Receiver velocity X	(none)
Geometry	sc_vel_y	int	m/s	Receiver velocity Y	(none)
Geometry	sc_vel_z	int	m/s	Receiver velocity Z	(none)
Geometry	sp_pos_x	int	meter	Specular point posistion X	(none)
Geometry	sp_pos_y	int	meter	Specular point posistion Y	(none)
Geometry	sp_pos_z	int	meter	Specular point posistion Z	(none)
Power	gps_eirp_watt	float	Watts	GPS EIRP	$P_t G_t$
Power	ddm_ant	int	(none)	The antenna that received signal	(none)
Metadata	prn_code	int	(none)	GPS PRN code	(none)
Metadata	ddm_ant	int	(none)	The antenna that received signal	(none)
Metadata	ddm_sp_delay_row	float	(none)	Specular delay bin index	n_τ^s
Metadata	ddm_sp_dopp_col	float	(none)	Specular Doppler bin index	n_f^s
Metadata	ddm_timestamp_utc	int	(none)	DDM sample time	(none)
Metadata	quality_flags	int	(none)	DDM quality flags	(none)

REFERENCES

[1] M. Clarizia, C. Gommenginger, S. Gleason, M. Srokosz, C. Galdi, and M. Di Bisceglie, "Analysis of GNSS-R delay-doppler maps from the UK-DMC satellite over the ocean," *Geophysical Research Letters*, vol. 36, no. 2, 2009.

[2] G. Foti, C. Gommenginger, P. Jales, M. Unwin, A. Shaw, C. Robertson, and J. Rosello, "Spaceborne GNSS reflectometry for ocean winds: First results from the UK TechDemoSat-1 mission," *Geophysical Research Letters*, vol. 42, no. 13, pp. 5435–5441, 2015.

[3] C. Ruf, M. Unwin, J. Dickinson, R. Rose, D. Rose, M. Vincent, and A. Lyons, "CYGNSS: enabling the future of hurricane prediction [remote sensing satellites]," *IEEE Geoscience and Remote Sensing Magazine*, vol. 1, no. 2, pp. 52–67, 2013.

[4] V. U. Zavorotny and A. G. Voronovich, "Scattering of GPS signals from the ocean with wind remote sensing application," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 2, pp. 951–964, 2000.

[5] M. Unwin, P. Jales, P. Blunt, S. Duncan, M. Brummitt, and C. Ruf, "The SGR-ReSI and its application for GNSS reflectometry on the NASA EV-2 CYGNSS mission," in *Aerospace Conference, 2013 IEEE*. IEEE, 2013, pp. 1–6.

[6] S. Gleason, C. S. Ruf, M. P. Clarizia, and A. J. O'Brien, "Calibration and unwrapping of the normalized scattering cross section for the cyclone global navigation satellite system," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 5, pp. 2495–2509, 2016.

[7] M. P. Clarizia and C. S. Ruf, "Wind speed retrieval algorithm for the cyclone global navigation satellite system (CYGNSS) mission," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 8, pp. 4419–4432, 2016.

[8] C. S. Ruf and R. Balasubramaniam, "Development of the CYGNSS geophysical model function for wind speed," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2018.

[9] M. P. Clarizia, C. S. Ruf, P. Jales, and C. Gommenginger, "Spaceborne GNSS-R minimum variance wind speed estimator," *IEEE transactions on geoscience and remote sensing*, vol. 52, no. 11, pp. 6829–6843, 2014.

[10] C. Ruf, P. Chang, M. Clarizia, S. Gleason, Z. Jelenak, J. Murray, M. Morris, S. Musko, D. Posselt, D. Provost *et al.*, "CYGNSS handbook," *Ann Arbor, MI, USA: Michigan Publishing*, 2016.

[11] O. Talagrand, *Data Assimilation in Meteorology And Oceanography*, ser. International Geophysics Series. Elsevier Science Serials, 2002. [Online]. Available: https://books.google.com/books?id=f57gAAAACAAJ

[12] E. Blayo, E. Cosme, M. Nodet, and A. Vidart, *Introduction to data assimilation*, 2011.

[13] G. Evensen, *Data assimilation: the ensemble Kalman filter*. Springer Science & Business Media, 2009.

[14] D. T. Pham, J. Verron, and M. C. Roubaud, "A singular evolutive extended Kalman filter for data assimilation in oceanography," *Journal of Marine systems*, vol. 16, no. 3-4, pp. 323–340, 1998.

[15] D. F. Parrish and J. C. Derber, "The national meteorological center's spectral statistical-interpolation analysis system," *Monthly Weather Review*, vol. 120, no. 8, pp. 1747–1763, 1992.

[16] D. M. Barker, W. Huang, Y.-R. Guo, A. Bourgeois, and Q. Xiao, "A three-dimensional variational data assimilation system for MM5: Implementation and initial results," *Monthly Weather Review*, vol. 132, no. 4, pp. 897–914, 2004.

[17] P. Courtier, J.-N. Thépaut, and A. Hollingsworth, "A strategy for operational implementation of 4D-Var, using an incremental approach," *Quarterly Journal of the Royal Meteorological Society*, vol. 120, no. 519, pp. 1367–1387, 1994.

[18] L. Cucurull, J. Derber, and R. Purser, "A bending angle forward operator for global positioning system radio occultation measurements," *Journal of Geophysical Research: Atmospheres*, vol. 118, no. 1, pp. 14–28, 2013.

[19] Y.-H. Kuo, S. Sokolovskiy, R. Anthes, and F. Vandenberghe, "Assimilation of GPS radio occultation data for numerical weather prediction," pp. 157–186, 2001.

[20] R. Hoffman, S. Leidner, J. Henderson, R. Atlas, J. Ardizzone, and S. Bloom, "A two-dimensional variational analysis method for NSCAT ambiguity removal: Methodology, sensitivity, and tuning," *Journal of Atmospheric and Oceanic Technology*, vol. 20, no. 5, pp. 585–605, 2003.

[21] S. Zhang and Z. Pu, "Assessing the impact of assimilating CYGNSS ocean surface winds on tropical cyclone analyses and forecasts with regional OSSEs," *32nd Conference on Hurricanes and Tropical Meteorology*, 2016.

[22] S. Zhang, Z. Pu, D. J. Posselt, and R. Atlas, "Impact of CYGNSS ocean surface wind speeds on numerical simulations of a hurricane in observing system simulation experiments," *Journal of Atmospheric and Oceanic Technology*, vol. 34, no. 2, pp. 375–383, 2017.

[23] B. McNoldy, B. Annane, S. Majumdar, J. Delgado, L. Bucci, and R. Atlas, "Impact of assimilating CYGNSS data on tropical cyclone analyses and forecasts in a regional OSSE framework," *Marine Technology Society Journal*, vol. 51, no. 1, pp. 7–15, 2017.

[24] B. Annane, B. McNoldy, S. M. Leidner, R. Hoffman, R. Atlas, and S. J. Majumdar, "A study of the HWRF analysis and forecast impact of realistically simulated CYGNSS observations assimilated as scalar wind speeds and as VAM wind vectors," *Monthly Weather Review*, vol. 146, no. 7, pp. 2221–2236, 2018.

[25] S. M. Leidner, B. Annane, B. McNoldy, R. Hoffman, and R. Atlas, "Variational analysis of simulated ocean surface winds from the Cyclone Global Navigation Satellite System (CYGNSS) and evaluation using a regional OSSE," *Journal of Atmospheric and Oceanic Technology*, no. 2018, 2018.

[26] A. O'Brien, "End-to-End Simulator Technical Memo," *CYGNSS official documentation, Ohio State University*, no. 148-0123, 2014.

[27] S. Gleason, C. S. Ruf, A. J. OBrien, and D. S. McKague, "The CYGNSS level 1 calibration algorithm and error analysis based on on-orbit measurements," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 1, pp. 37–49, 2018.

[28] S. J. Katzberg, O. Torres, and G. Ganoe, "Calibration of reflected GPS for tropical storm wind speed retrievals," *Geophysical Research Letters*, vol. 33, no. 18, 2006.

[29] T. Elfouhaily, D. R. Thompson, and L. Linstrom, "Delay-doppler analysis of bistatically reflected signals from the ocean surface: theory and application," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 3, pp. 560–573, 2002.

[30] J. F. Marchán-Hernández, A. Camps, N. Rodríguez-Álvarez, E. Valencia, X. Bosch-Lluis, and I. Ramos-Pérez, "An efficient algorithm to the simulation of delay–doppler maps of reflected global navigation satellite system signals," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 8, pp. 2733–2740, 2009.

[31] G. Grieco, A. Stoffelen, M. Portabella, M. B. Rivas, W. Lin, and F. Fabra, "Quality control of delay-doppler maps for stare processing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 5, pp. 2990–3000, 2018.

[32] F. Huang, J. L. Garrison, N. Rodriguez-Alvarez, A. J. O'Brien, K. M. Schoenfeldt, S. C. Ho, and H. Zhang, "Sequential processing of GNSS-R delay-doppler maps to estimate the ocean surface wind field," *IEEE Transactions on Geoscience and Remote Sensing*, 2019.

[33] V. Tallapragada, L. Bernardet, M. K. Biswas, S. Gopalakrishnan, Y. Kwon, Q. Liu, T. Marchok, D. Sheinin, M. Tong, S. Trahan *et al.*, "Hurricane Weather Research and Forecasting (HWRF) model: 2015 scientific documentation," *HWRF Development Testbed Center Tech. Rep*, vol. 99, 2014.

[34] R. M. Yablonsky, I. Ginis, B. Thomas, V. Tallapragada, D. Sheinin, and L. Bernardet, "Description and analysis of the ocean component of NOAAs operational Hurricane Weather Research and Forecasting Model (HWRF)," *Journal of Atmospheric and Oceanic Technology*, vol. 32, no. 1, pp. 144–163, 2015.

[35] J. P. Cangialosi, "National hurricane center forecast verification report, 2017 hurricane season," *NOAA National Hurricane Center Tech. Rep*, 2017.

[36] J. Miranda, M. Vall-Ilossera, A. Camps, and N. Duffo, "Sea surface emissivity at l-band: Swell effects," in *Geoscience and Remote Sensing Symposium, 2002. IGARSS'02. 2002 IEEE International*, vol. 5. IEEE, 2002, pp. 2623–2625.

[37] M. P. Clarizia, "Investigating the effect of ocean waves on GNSS-R microwave remote sensing measurements," Ph.D. dissertation, University of Southampton, 2012.

[38] H. L. Tolman *et al.*, "User manual and system documentation of WAVE-WATCH III TM version 3.14," *Technical note, MMAB Contribution*, vol. 276, p. 220, 2009.

[39] T. W. Group, "The WAM model-A third generation ocean wave prediction model," *Journal of Physical Oceanography*, vol. 18, no. 12, pp. 1775–1810, 1988.

[40] M. P. Clarizia and C. S. Ruf, "Bayesian wind speed estimation conditioned on significant wave height for GNSS-R ocean observations," *Journal of Atmospheric and Oceanic Technology*, vol. 34, no. 6, pp. 1193–1202, 2017.

[41] A. G. Voronovich and V. U. Zavorotny, "The transition from weak to strong diffuse radar bistatic scattering from rough ocean surface," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 11, pp. 6029–6034, 2017.

[42] C. S. Ruf, S. Gleason, and D. S. McKague, "Assessment of CYGNSS Wind Speed Retrieval Uncertainty," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2018.

[43] T. Wang, C. S. Ruf, B. Block, D. S. McKague, and S. Gleason, "Design and performance of a GPS constellation power monitor system for improved CYGNSS l1b calibration," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 1, pp. 26–36, 2018.

[44] V. U. Zavorotny and A. G. Voronovich, "Recent progress on forward scattering modeling for GNSS reflectometry," in *Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International*. IEEE, 2014, pp. 3814–3817.

[45] R. Atlas, R. N. Hoffman, J. Ardizzone, S. M. Leidner, J. C. Jusem, D. K. Smith, and D. Gombos, "A cross-calibrated, multiplatform ocean surface wind velocity product for meteorological and oceanographic applications," *Bulletin of the American Meteorological Society*, vol. 92, no. 2, pp. 157–174, 2011.

[46] J.-F. Bonnans, J. C. Gilbert, C. Lemaréchal, and C. A. Sagastizábal, *Numerical optimization: theoretical and practical aspects*. Springer Science & Business Media, 2006.

[47] E. Andersson, A. Persson, and I. Tsonevsky, "User guide to ECMWF forecast products," *ECMWF, v2*, vol. 1, p. 121, 2015.

[48] A. Stoffelen, J. A. Verspeek, J. Vogelzang, and A. Verhoef, "The CMOD7 geophysical model function for ASCAT and ERS wind retrievals," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 5, pp. 2123–2134, 2017.

[49] J. V. Anton Verhoef and A. Stoffelen, *ScatSat-1 wind validation report*, OSI SAF, 2018. [Online]. Available: http://projects.knmi.nl/scatterometer/publications/pdf/osisaf_cdop3_ss3_valrep_scatsat1_winds.pdf

James L. Garrison (SM'14) is a Professor in the School of Aeronautics and Astronautics at Purdue University, West Lafayette, IN, USA with a courtesy appointment in the School of Electrical and Computer Engineering. He is also an affiliate with the Ecological Sciences and Engineering (ESE) interdisciplinary program and serves on the Executive Committee of the Purdue Climate Change Research Center (PCCRC). His research interests include Earth remote sensing using Global Navigation Satellite Systems (GNSS) and signals of opportunity. He is the Principal Investigator for SNOOPI, a NASA mission to demonstrate remote sensing with P-band signals of opportunity. Prior to his academic position, Prof. Garrison was with the National Aeronautics and Space Administration (NASA), Greenbelt, MD, USA. He earned a Ph.D. from the University of Colorado Boulder, Boulder, CO, USA in 1997 and also holds a B.S. from the Rensselaer Polytechnic Institute, Troy, NY, USA and a M.S. from Stanford University, Stanford, CA, USA. He is a fellow of the Institute of Navigation (ION) and Editor in Chief of the IEEE Geoscience and Remote Sensing Magazine.

S. Mark Leidner is a Senior Staff Scientist at Atmospheric and Environmental Research (AER, a Verisk Analytics Company) and a Certified Consulting Meteorologist (CCM, Certificate #681 from the American Meteorological Society, January 2012). He has a BS in Metallurgical Engineering and Materials Science from Carnegie Mellon (1989) and a MS in Meteorology from Penn State (1996). He spent fifteen months at the European Center for Medium Range Weather Forecasts working as a consultant on surface winds from the NASA scatterometer instruments, NSCAT and QuikSCAT. At ECMWF, he introduced the use of US scatterometer data in the Centre's 4D-variational analysis assimilation system. He was also project lead on an Emerging Risk Research Program (ERRP), run by Verisk Insurance Solutions, a division of Verisk Analytics, and sponsored by US and Canadian insurers to examine shifts in exposure due to changing patterns of severe convective weather.

Bachir Annane graduated with a "Diplme d'enseignement Suprieur en Mathmatique pure" (DES in pure mathematics) from the University of Algiers (USTHB), Algeria. In the US, he has a M.S. degree in Applied Mathematics from the University of Central Florida. He also has a M.S. and a Ph.D. in Meteorology from Florida State University. He has been working as a researcher with the Cooperative Institute of Marine and Atmospheric Science (CIMAS/UM) since 2005.

Feixiong Huang (S'18) received the B.S. degree in geodesy and geomatics engineering from Wuhan University, Wuhan, China, in 2014, and the M.S. degree in geomatics from the Lyles School of Civil Engineering, Purdue University, West Lafayette, IN, USA, in 2015, where he is currently pursuing the Ph.D. degree with the School of Aeronautics and Astronautics.
His research interests include the development of a GNSS-R forward model and assimilation of GNSS-R Delay-Doppler Maps into numerical weather models under CYGNSS mission project.

Ross N. Hoffman is the past vice-president of the R & D division of Atmospheric Environmental Research (AER 1999 - 2013). He is now a part-time scientist at the Cooperative Institute for Satellite Earth System Studies (CISESS), University of Maryland, supporting the NOAA/NESDIS/Center for Satellite Applications and Research (STAR), College Park, Maryland. His principal areas of interest are objective analysis and assimilation methods, impact experiments, atmospheric dynamics, climate theory, and atmospheric radiation. He was a long-term member of the NASA scatterometer science teams and has served on some committees of the AMS and NAS. His Ph.D. advisor was Prof. E. N. Lorenz and post-doc advisor was Dr. E. Kalnay.

Giuseppe Grieco received his M. Sc. in Marine Sciences from the University Parthenope of Napoli in 2001 and his Ph. D. in Environmental Engineering from the Universit della Basilicata in 2008. Currently, he is with the Institute of Marine Sciences, Barcelona, Spain, working on satellite remote sensing, with a particular focus on active microwave instruments.

Ad Stoffelen (M'13-SM'14) received the M.Sc. degree in physics from the Technical University of Eindhoven, Eindhoven, The Netherlands, in 1987, and the Ph.D. degree in meteorology on scatterometry from the University of Utrecht, Utrecht, The Netherlands, in 1998. He currently leads a Group on active satellite sensing with the Royal Netherlands Meteorological Institute, De Bilt, The Netherlands, where he is involved in future satellite missions, research and development of information retrieval, numerical weather prediction (NWP) mesoscale wind data assimilation, 24/7 operations, user training, and services. He is furthermore involved in the European Space Agency Doppler Wind Lidar Mission-Aeoluson atmospheric dynamics. His research interests include establishing an international scatterometer virtual constellation.