Investigating the Sensitivity of Spaceborne GNSS-R Measurements to Ocean Surface Winds and Rain

by

Rajeswari Balasubramaniam

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Climate and Space Sciences and Engineering)
in the University of Michigan
2020

Doctoral Committee:

Professor Christopher Ruf, Chair
Associate Research Scientist Darren McKague
Professor Aaron J. Ridley
Professor Fawwaz T. Ulaby

A falling water droplet.

Rajeswari Balasubramaniam
rajibala@umich.edu
ORCID iD: 0000-0002-9937-1201

To MahaMaya.

ACKNOWLEDGEMENTS

I am extremely grateful to my advisor Prof. Ruf for mentoring me throughout this Ph.D program in various aspects. His valuable and critical analysis of work has helped shape this thesis. I have also learnt many subtle lessons ranging from how to give a research talk to as simple as composing a professional email from him. Next, I would like to acknowledge Dr. Darren McKague for his invaluable guidance and help with understanding and handling different kinds of data. I would also like to share my regards to the other committee members, Prof. Ulaby and Prof. Ridley for their valuable suggestions through my dissertation phase.

I would like to extend my gratitude to the entire CYGNSS science team for their help and ideas at various stages of this research, especially to Dr. Valery Zavarotny and Dr. Maria Paola Clarizia for their critical inputs and guidance. I am thankful to my lab mates Mary Morris, Tianlin Wang, David Meyers, Mohammed Mousavi and Maryam Salim for making this journey a joyful learning experience. I would also like to thank the ClaSP IT staff, Bryan White and Faye Ogasawara for helping me have a smooth sail with my IT needs.

I cannot describe in words how much my family - Tata, Pati, Amma, Appa and Padu have impacted and shaped me into who I am today. I am also ever grateful to my friends and extended family for their support and encouragement.

Last and the most important is my humble pranams to my Paramacharya for always being there and guiding me.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

ABSTRACT

Earth remote sensing using reflected Global Navigation Satellite System (GNSS) signals is an emerging trend, especially for ocean surface wind measurements. GNSS-Reflectometry (GNSS-R) measurements of ocean surface scattering cross section are directly related to the surface roughness at scale sizes ranging from small capillary waves to long gravity waves. These roughness scales are predominantly due to swell, surface winds and other meteorological phenomena such as rain. In this study we are interested in understanding and characterizing the impact of these phenomena on GNSS-R signals in order to develop a better understanding of the geophysical parameters retrieved from these measurements.

In the first part of this work, we look at GNSS-R measurements made by the NASA Cyclone Global Navigation Satellite System (CYGNSS) for developing an effective wind retrieval model function for GNSS-R measurements. In a fully developed sea state, the wind field has a constant speed and direction. In this case, a single Fully Developed Seas (FDS) Geophysical Model Function (GMF) is constructed which relates the scattering cross section to the near surface wind speed. However, the sea age and fetch length conditions inside a hurricane are in general not consistent with a fully developed sea state. Therefore, a separate empirical Young Sea Limited Fetch (YSLF) GMF is developed to represent the conditions inside a hurricane. Also, the degree of under development of the seas is not constant inside hurricanes and conditions vary significantly with azimuthal location within the hurricane due to changes in the relative alignment of the storms forward motion and its cyclonic rotation. The azimuthal dependence of the scattering cross-section is modelled and a modified

azimuthal YSLF GMF is constructed using measurements by CYGNSS over 19 hurricanes in 2017 and 2018.

Next, we study the impact of rain on CYGNSS measurements. At L-band rain has a negligible impact on the transmitted signal in terms of path attenuation. However, there are other effects due to rain, such as changes in surface roughness and rain induced local winds, which can significantly alter the measurements. In this part of the study we propose a 3-fold rain model for GNSS-R signals which accounts for: 1) attenuation; 2) surface effects of rain; and 3) rain induced local winds. The attenuation model suggests a total of 96% or greater transmissivity at L-Band up to 30mm/hr of rain. A perturbation model is used to characterize the other two rain effects. It suggests that rain is accompanied by an overall reduction in the scattering cross-section of the ocean surface and, most importantly, this effect is observed only up to surface wind speeds of $\sim$ 15 m/s, beyond which the gravity capillary waves dominate the scattering in the quasi-specular direction. This work binds together several rain-related phenomena and enhances our overall understanding of rain effects on GNSS-R measurements.

Finally, one of the important objectives for the CYGNSS mission is to provide high quality global scale GNSS-R measurements that can reliably be used for ocean science applications. In this part of the work we develop a Neural Network based quality control filter for automated outlier detection for CYGNSS retrieved winds. The primary merit of the proposed Machine Learning (ML) filter is its ability to better account for interactions between the individual engineering, instrument and measurement conditions than can separate thresholded quality flags for each one.

CHAPTER I

Introduction

Ocean remote sensing plays an integral role in understanding the dynamics of the Earth and the inter-relationship between various Earth systems. Ocean waves span a wide range of scale sizes, from the centimeter scales of small surface waves to planetary level meso-scale structures ranging to several thousands of kilometers (*Maul*, 2012). Fig.1.1 gives a broad picture of the range of scales of ocean waves. At every scale several vital physical, chemical and biological processes occur between the ocean and other components of the Earth system, which determine the Earth's climate patterns. In addition to climate studies, accurate remote sensing of ocean waves is important for operational weather forecasting, port and ship operations and other engineering projects. Ocean surface wind speeds at 10m and 19m have been of particular interest to the oceanography community as they help evaluate the rate of momentum transfer between the sea and atmosphere using boundary layer models, and therefore inform our understanding of global energy circulation.

The advent of ocean remote sensing can be traced back to World War II when the first quantitative application of remote sensing was made in oceanography using aerial hydrographic surveys (*Robinson*, 2010). Since then, the instrumentation, technology and science have rapidly evolved and have provided us with new and exciting approaches to study our planet's oceans. Table 1.1 lists different ocean properties and

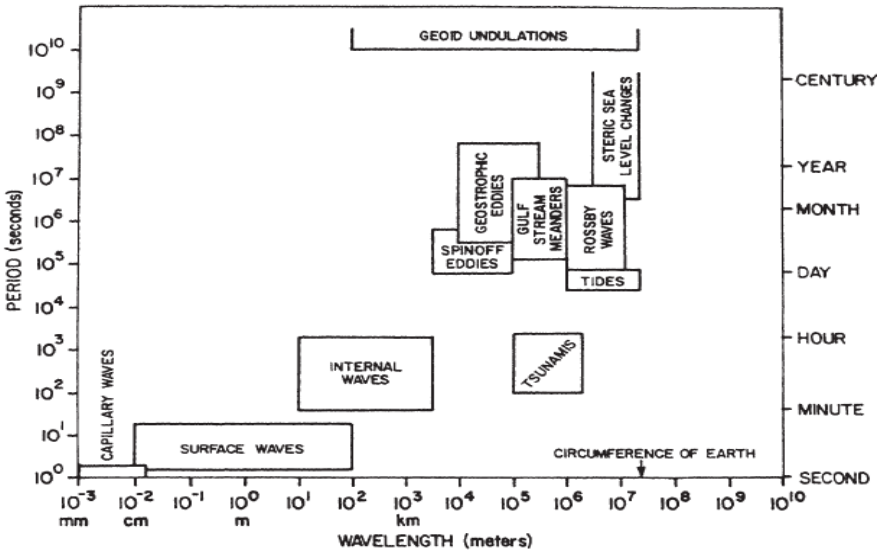


Figure 1.1: Spatial and temporal scales of a few kinds of ocean waves (*Maul*, 2012)

the remote sensing techniques available to measure them. Our specific interest in this study is the accurate measurement of ocean surface winds using microwave remote sensing capabilities.

As with any emerging technology, new techniques require new interpretative ideas. Therefore, it is important to characterize ocean surface waves in terms of measurable parameters using different techniques. The following is a description of important wave parameters that are used to define an ocean surface wave.

In the absence of wind forcing, a perturbation on the sea surface can be assumed to behave like a well-defined wave using simplified linear wave theory (*Young*, 1999). Thus an ocean surface wave can be characterized by a period T, a wavelength λ and a dispersion relation that connects both. The dispersion relation for a sinusoidal surface wave is given by (*Young*, 1999)

$$\omega^2 = gk, \tag{1.1}$$

where $\omega = \frac{2\pi}{T}$ is the wave frequency, $k = \frac{2\pi}{\lambda}$ is the wavenumber and g is the acceler-

Ocean property	**Technique**
Temperature	Infrared radiometry
	Microwave radiometry
Salinity	Microwave Radiometry
	Laser Scatterometry
Currents	Microwave Radar
	HF Radar
	Doppler Laser
Surface winds	Microwave radar/radiometry
Sea State	Microwave radar/radiometry
	Radar altimetry
	Visible radiometry
Tides	Microwave Radar
Ice	Visible radiometry
	Microwave radar

Table 1.1: A non-exhaustive list of ocean properties and the corresponding remote sensing technique used to measure it. (*Maul*, 2012).

ation due to gravity. These types of waves are dispersive in nature i.e, longer waves travel faster than shorter waves. In general, Eq.1.1 characterizes the swell waves generated by a storm in the ocean.

However, wind forcing cannot be neglected as wind is continually creating local waves in the oceans. To completely describe a wave field, the height of the surface needs to be known at every point, which is not possible in practice. Hence sea surface usually is defined using statistical properties of surface heights or slopes, which is called the Wave Spectrum, S. Another important parameter for defining ocean waves is the Significant Wave Height, H_s, which is a statistical measure of the height between the crest and the trough of a sinusoidal wave.

The simplifying assumptions of constant wave amplitude and period associated with finite amplitude wave theory are not representative of the actual sea surface, which in reality is often a combination of different wave heights, periods and wavelengths. Thus it is helpful to represent a sea surface as a linear superposition of sinusoids given by

$$\eta(t) = \sum_{i=1}^{N} a_i sin(\omega_i t + \phi_i), \tag{1.2}$$

where a_i, ω_i and ϕ_i are the amplitude, frequency and phase, respectively, of the ith component. Each of the sinusoidal components satisfies linear wave theory. Thus the variance of a wave profile is given by

$$\sigma^2 = \frac{1}{2N} \sum_{i=1}^{N} a_i^2, \tag{1.3}$$

In the limit as $N \to \infty$ and the individual wave components become differentially small, this becomes a Fourier transform in the continuous frequency domain given by

$$\sigma^2 = \int_0^\infty S(\omega) d\omega \tag{1.4}$$

Here $S(\omega)$ is called the omnidirectional frequency spectrum (*Elfouhaily et al.*, 1997). The advantage of using this variance spectrum is its ability to model a complex water surface using simplified linear wave theory. Let us define $\psi(k_x, k_y)$ as the directional spectrum in Cartesian coordinates, where k_x is the wavenumber in the upwind direction and k_y is the wavenumber in the cross-wind direction. Then we can define the upwind and crosswind Mean Square Slope (MSS) as

$$\begin{aligned} mss_x &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} k_x^2 \psi(k_x, k_y) dk_x dk_y \\ mss_y &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} k_y^2 \psi(k_x, k_y) dk_x dk_y \end{aligned} \tag{1.5}$$

The total MSS is then given by

$$mss = mss_x + mss_y = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(k_x^2 + k_y^2)\psi(k_x, k_y)dk_x dk_y = \int_{0}^{\infty} k^2 S(k)dk \tag{1.6}$$

The term in the integral ($k^2S(k)$) is called the omnidirectional slope spectrum. GNSSR forward scatter is quasi specular incoherent scatter in most conditions. This scattering mechanism is dependent on the long, tilting waves. Therefore an appropriate MSS for surface roughness relevant to GNSS-R is the low pass omnidirectional MSS given by

$$mss = \int_{k_{min}}^{k_{max}} k^2 S(k)dk \tag{1.7}$$

Here k_{max} is called the low-pass wavenumber or the scale dividing parameter and sets an upper limit on the wavenumbers to which GNSS-R measurements are sensitive. An empirical value for the low-pass wavenumber corresponding to 3λ was proposed (*Brown*, 1990). This was later extended to account for incidence angle dependence (*Garrison et al.*, 2002). Thus, the approximate low-pass wavenumber is given by

$$k_{max} = \frac{2\pi}{3\lambda} cos(\theta), \tag{1.8}$$

An appropriate value for the low-pass wavenumber is required in order to derive an accurate MSS estimate for a given wave spectrum. Thus, the wave frequency, wavenumber, wave spectrum and significant wave height form a sufficient set of parameters to define an ocean surface for purposes of predicting and interpreting GNSS-R measurements.

1.1 Microwave remote sensing and instrumentation

Measurement of the ocean at frequencies of 1-40 GHz is commonly referred to as microwave ocean remote sensing. Above 40 GHz, atmospheric scattering and absorption dominate the signal while observations below 1 GHz suffer from lower resolution of surface properties as well as radio frequency interference (*Robinson*, 2010). The measurement of ocean surface winds at microwave frequencies is predominantly sensitive to the surface roughness in the case of active systems and to the natural radio emission from foam formed over the ocean in the case of passive systems. Based on the principle of operation and the nature of the measurements, current spaceborne ocean remote sensing techniques can be classified into four types as described below.

1.1.1 Microwave radiometers

Radiometers are passive sensing systems that observe the natural emission from the surface. At microwave frequencies, the emission from the ocean depends on the water temperature, the dielectric property of the surface and the orientation and shape of the sea surface. As wind blows over the ocean, foam forms on the ocean surface which has a distinct dielectric property unlike that of the sea water underneath it (*Droppleman*, 1970; *Wentz*, 1975). Consequently, wind speed becomes a measurable parameter for a multi-frequency microwave radiometer. Measurable sensitivity to wind speed is found from 6-37 GHz (*Ulaby et al.*, 2014). Thus some of the commonly known radiometers such as the Special Sensor Microwave Imager (SSMI) flown since 1987 (*Hollinger and Lo*, 1983), the microwave imager TMI flown on the Tropical Rainfall Measuring Mission (TRMM) flown since 1997 (*Kummerow et al.*, 1998) and the Advanced Microwave Scanning Radiometer (AMSR-E) (*Kawanishi et al.*, 2003) use this technique to retrieve ocean surface wind speed. The first polarimetric microwave radiometer, WindSat, which was launched in 2003 demonstrated the retrieval of both

wind speed and direction by measuring partially polarized emission from the ocean surface (*Gaiser et al.*, 2004).

1.1.2 Scatterometers

Scatterometers are active microwave instruments that are designed to measure the backscatter of the ocean surface. Most scatterometers operate at C or Ku bands, as these frequencies have the most atmospheric transmissivity (*Ulaby et al.*, 2014). To date, scatterometers are the most effective sensors for mapping the distribution of wind speed and direction. A scatterometer transmits radar pulses to the ocean and measures the energy scattered back from the surface. The backscatter geometry has the transmitter and the receiver at the same location. The back scattered power is then used to derive the normalized radar cross-section at the surface, which contains information about the roughness of the surface. The normalized radar cross-section, σ_0 is evaluated from the radar equation given by

$$P_r = \frac{P_t G^2 \lambda^2 A}{(4\pi)^3 R^4} \sigma_0, \tag{1.9}$$

where P_t and P_r are the transmit and receive powers, G is the gain of the antenna, λ is the wavelength of the transmit signal, A is the effective illuminated area and R is the distance between the transmitter and the surface. The measurements in general are made in the incidence angle range of 20-60 degrees and the dominant mechanism that contributes to the backscattering at these incidence angles is Bragg scattering by surface capillary and short gravity waves that resonate with the wavelength of the signal. The condition for Bragg scattering is illustrated in Fig.1.2 and given by

$$\frac{\lambda_s}{\lambda_r} sin(\theta) = \frac{n}{2}, n = 1, 2, 3..., \tag{1.10}$$

Typical incidence angles	C-band (5.25 GHz)	Ku-Band (13.4 GHz)
20 deg	9.76mm	3.8 mm
30 deg	1.42 cm	5.5 mm
40 deg	1.8 cm	7.16 mm
50 deg	2.18 cm	8.54 mm
60 deg	2.47 cm	9.65 mm

Table 1.2: Fundamental Bragg resonant wavelength for typical incidence angles.

where λ_s is the wavelength of the surface wave, θ is the incidence angle and λ_r is the wavelength of the transmit signal. The Sea Winds instrument on QuikScat is an example of the Ku-band Scatterometer that provided 90% global coverage per day with a spatial resolution of 25 km and a swath of 1400 km (*Graf et al.*, 1998; *Ebuchi et al.*, 2002). The ASCAT instrument on the MetOp is an example of a C-band Scatterometer with a swath width of 500 km and spatial resolution of 50 km and provides almost global coverage per day (*Figa-Saldaña et al.*, 2002).

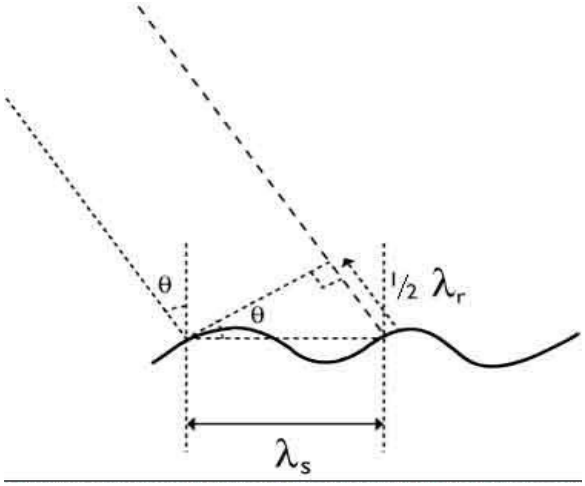


Figure 1.2: Bragg scattering mechanism

The dominant scattering mechanism for scatterometers operating at moderate incidence-angle ranges is Bragg scattering. Under Bragg scattering, the return signal is proportional to the roughness of the ocean surface on the scale of the radar wavelength. As wind speed increases, surface roughness increases and this increases the σ_0 in the backscatter direction. However, it is to be noted that σ_0 also is dependent on

the relative azimuth angle between the radar look direction and the wind direction. Therefore, σ_0 is a function of the radar parameters, the observing geometry and the height profile of the surface. Table 1.2 describes the fundamental Bragg resonant wavelengths for different incidence angles. It can be noted that these resonant wavelengths match the wavelength of capillary waves, which is on the order of a few centimeters.

1.1.3 Synthetic Aperture Radars

Synthetic Aperture Radars (SAR) are a class of active microwave imaging radars that produces line-by-line images by capturing the backscattered radar pulses. However, unlike Scatterometry, SAR does not image the surface from different view angles. Hence, the wind speed can be estimated only with prior knowledge of wind direction at the surface (*Maul*, 2012). It is important to note that SAR produces much finer resolution information of the roughness which allows for studying the microscale changes in wind speed. However, its narrower swath-width results in much longer global revisit times, thus making it less useful for operational weather forecasting. This makes SAR less suitable for global wind speed measurement as compared to scatterometers.

1.1.4 Radar altimeters

A radar altimeter is an active microwave instrument that transmits short radar pulses through its nadir pointing antenna and retrieves sea surface height information from the shape of the return pulse. This technique can be extended to interpret the surface wind speed as well. The altimeter, being a nadir-viewing (zero incidence angle) radar, it responds differently to wind than an oblique-viewing scatterometer or SAR. The maximum signal is returned when the sea is calm, so increased scattering in high

winds reduces the magnitude of the radar echo. When a transmit pulse hits the sea surface, it is reflected by the highest wave crests first and then followed by reflections from the lowest wave troughs. The return wave pulse ramp has a lower slope for larger significant wave height and higher slope for lower significant wave height (*Robinson*, 2010). Thus it is the slope of the leading edge of the return pulse that is correlated with the sea state and the amplitude of the return pulse that is correlated with the surface wind speed. It is to be noted that altimeters are point sampling instruments and therefore they have a relatively sparse sampling than their scatterometer and SAR counterparts.

1.2 Ocean sensing using bistatically reflected GNSS signals

Associated with the three approaches described in the previous section are three important limitations namely, insufficient temporal sampling, gaps between swaths and the high cost of building and maintaining such systems. These limitations have created the need for the development of a low-cost complimentary system that can fill coverage gaps, offers increased temporal sampling and easy maintenance for long term operations. Bistatic radar remote sensing using navigation signals of opportunity is an emerging trend that satisfies these needs.

Global Navigation Satellite Systems (GNSS) are satellite constellations that provide global coverage with timing and positioning information to users located on or near the surface of the Earth. Though the main purpose of the GNSS system is for civilian and military geolocation and navigation applications, it has already had considerable impact as a viable remote sensing instrument. The standard operating frequencies for GNSS systems include L1 (1.575 GHz/19.05 cm), L2 (1.227 GHz/24.45 cm) and L5 (1.176 GHz/ 25.48 cm) bands. A highly successful application of GNSS is the Radio Occultation (RO) technique for atmospheric sensing (*Wickert et al.*, 2004; *Luntama*

GNSS system	Country	Operating bands
GPS	USA	L1, L2, L5
GALILEO	Europe	L1, E6, L5
COMPASS	China	L1, E6, L5
QZSS	Japan	L1, E6, L2, L5
IRNSS	India	L5

Table 1.3: Operating frequency bands for existing GNSS systems .

et al., 2008). RO is a process of measuring the Doppler shift due to atmospheric bending in a transmitted signal between one satellite and another. This Doppler shift due to the atmosphere is useful in deriving the temperature, pressure and density of the atmosphere as a function of altitude.

The use of reflected GNSS signals to study the Earth's surface is also becoming widely popular and many theoretical bases and technology demonstrations are being established (*Lowe et al.*, 2002; *Garrison et al.*, 2002; *Cardellach et al.*, 2003; *Komjathy et al.*, 2004; *Gleason et al.*, 2005). As of today there are many GNSS constellations in place such as the Global Positioning System (GPS) by the United States, The Russian GLONASS system, the European GALILEO, the Chinese BeiDou (COMPASS) system and other regional systems such as the Indian IRNSS and the Japanese QZSS. All the GNSS systems operate at L-Band which has inherent advantages for ocean remote sensing purposes. Also, L-band signals are sensitive to longer waves which are not directly forced by winds. The two important differences between conventional Scatterometers and GNSS-R are the longer wavelength of operation for GNSS-R systems and the fact that GNSS-R measurements use a bistatic geometry, so the forward scattering cross section is the measurement, as opposed to the backscattering cross section for scatterometers. While the backscatter is dominated by Bragg scattering at a narrow range of capillary wavelengths, there is no Bragg resonance with forward scatter and the scattering cross section is determined by a much wider range of wavelengths. Table 1.3 presents the various bands of operation for the current GNSS systems.

The technique of observing the reflections of GNSS signals that contain information about the properties and the characteristics of the scattering surface is called GNSS-Reflectometry (GNSS-R) (*Gleason et al.*, 2009). The technique of GNSS-R can be used to study a variety of ocean geophysical parameters with a primary focus on surface roughness as the measurable parameter. There are various advantages to using GNSS-R for ocean surface wind sensing. Firstly, the L-band operation allows penetration through precipitation and offers an all-day/ all-weather capability. This is especially advantageous for making measurements inside hurricanes. It is also important to note that the bistatic scattering geometry of a GNSS-R system ensures a strong signal return in the specular direction relative to the weak return for the monostatic Scatterometry systems. Finally, the global coverage of GNSS signals imply a possibility of global coverage for the GNSS-R systems. The details about the nature of GNSS-R measurements, the geometry and the existing airborne and spaceborne missions will be discussed in Chapter 2.

1.3 Looking into hurricanes

Hurricanes are destructive to life and property on Earth. In 2018, 22 major hurricanes formed in the Northern Hemisphere within 3 months, creating a record for the most active hurricane season in recorded history and 2017 had serious devastation by the Atlantic hurricanes. Hurricanes are intense low-pressure storm systems that are formed in the tropical latitudes over large bodies of warm waters and then move towards land. They are characterized by thunderstorms, large waves and powerful winds that rotate around their low pressure centers. Measuring the strength of surface winds in the inner-core of the storm is vital to advancing our understanding of hurricanes and predicting their life cycle. Fig.1.3 displays the tracks and intensities of hurricanes, tropical cyclones and typhoons on a global scale from 1850 to 2017

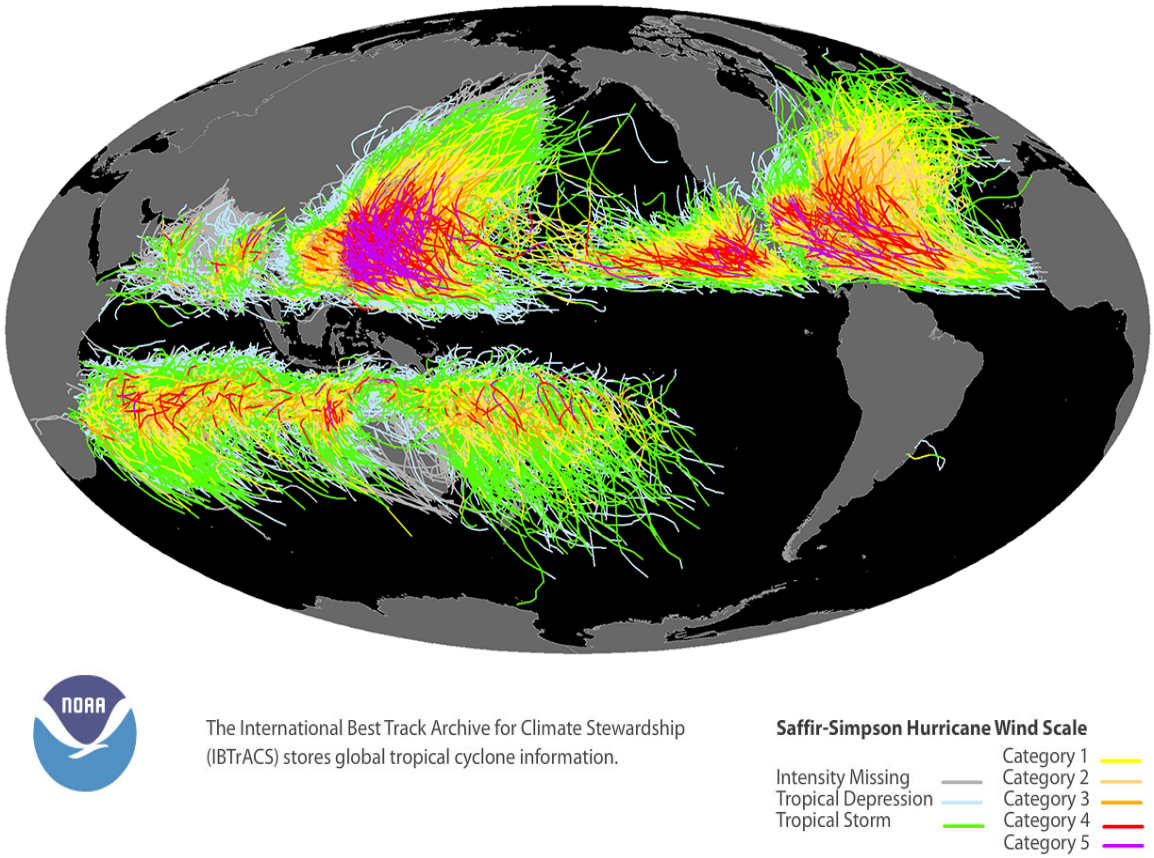


Figure 1.3: Global inventory of hurricane tracks and intensities (From 1850 to 2017). (https://www.ncei.noaa.gov/news/inventory-tropical-cyclone-tracks).

Numerical weather prediction (NWP) models are the primary tool used to predict the track and intensity of hurricanes. These models have traditionally been global in extent, and comprise solutions of the atmospheric momentum, mass, and energy and water vapor conservation equations. These NWPs use data assimilation of short term predictions and large volumes of satellite data for improved estimates of initial conditions. Current research capabilities can predict a hurricane track with a lead time of 5 days with very high accuracy; however, the ability to forecast hurricane intensification remains a challenge. One dominant reason for this is lack of frequent and accurate observations of surface winds in the inner core (*Ruf et al.*, 2016). The dedicated systems for measuring ocean surface winds discussed in Section 1.1 are unable

to provide accurate measurements in the regions of heavy precipitation and therefore there are fewer wind measurements within the inner core of hurricanes. Also, current remote sensing systems that can penetrate heavy rain have very long re-visit periods thereby failing to capture the hurricane during its rapid intensification phase. For instance, the polarimetric radiometer onboard Windsat has a capability to measure wind speeds between 3 and 25 m/s under rain rates up to 25 mm/hr but has a re-visit time of 8 days (*Gaiser et al.*, 2004). The QuikScat Sea-Winds instrument has a capability to measure from 3-30 m/s but a revisit period of 1-2 days. Further, the Sea-Winds instrument can measure wind speed, but with a wind speed error as large as 5m/s for rain rates as low as 2 mm/hr (*Tournadre and Quilfen*, 2003). Similarly, ASCAT retrievals becomes unreliable for rain rates $>$ 6 mm/hr (*Portabella et al.*, 2012). The AMI-wind C- band scatterometer also begins to have noticeable variation in NRCS for rain rates $>$ 5 mm/hr (*Tournadre and Quilfen*, 2003).

The unique role of a GNSS-R system in hurricane wind measurements is its ability to address these two limitations. The current spaceborne GNSS-R mission, Cyclone Global Navigation Satellite System (CYGNSS), is a perfect candidate. CYGNSS uses the all-weather performance of the GPS bistatic radar with the spatial and temporal sampling properties of a constellation of 8 LEO observatories to achieve surface wind measurements in the inner core of hurricanes with sufficient frequency to resolve its genesis and rapid intensification phases (*Ruf et al.*, 2016).

CHAPTER II

GPS Reflectometry Over Oceans

2.1 Nature of GNSS-R measurements

The GNSS-R measurement technique is discussed in this chapter, including the bistatic scattering geometry, a review of previous GNSS-R airborne and spaceborne missions, and finally a detailed discussion about the CYGNSS mission.

2.1.1 Bistatic geometry and Delay Doppler Maps

Among the different GNSS systems, the Global Positioning System (GPS) was the first and is the most widely used. The GPS constellation consists of 32 satellites in 6 equally spaced orbit planes. Each satellite circles the Earth twice a day and at least 4 satellites are visible at any one time from any point on the Earth. GNSS-R measures signals transmitted by a GPS satellite, scattered by the Earth surface, and then received by a ground-based, airborne or spaceborne GPS receiver.

When the GPS signal encounters the Earth surface, it is scattered by the individual facets and some of the reflected components arrive at the receiver. Of all the scattered signals received, the bistatic geometry is dominated by the specular reflection, which is the point of shortest distance between transmitter-Earth-receiver, called the Specular Point (SP). Irrespective of the signal interaction with the surface, the signal time delay

and Doppler frequency shift can be mapped across the surface accurately (*Gleason et al.*, 2009). The scattered signal is a superposition of components scattered from various points on the sea surface. The signal time delay is caused by the path followed by the signal to reach the receiver and the Doppler shift is caused by to the relative motion between the GPS satellites, the SP on the surface and the receiver system.

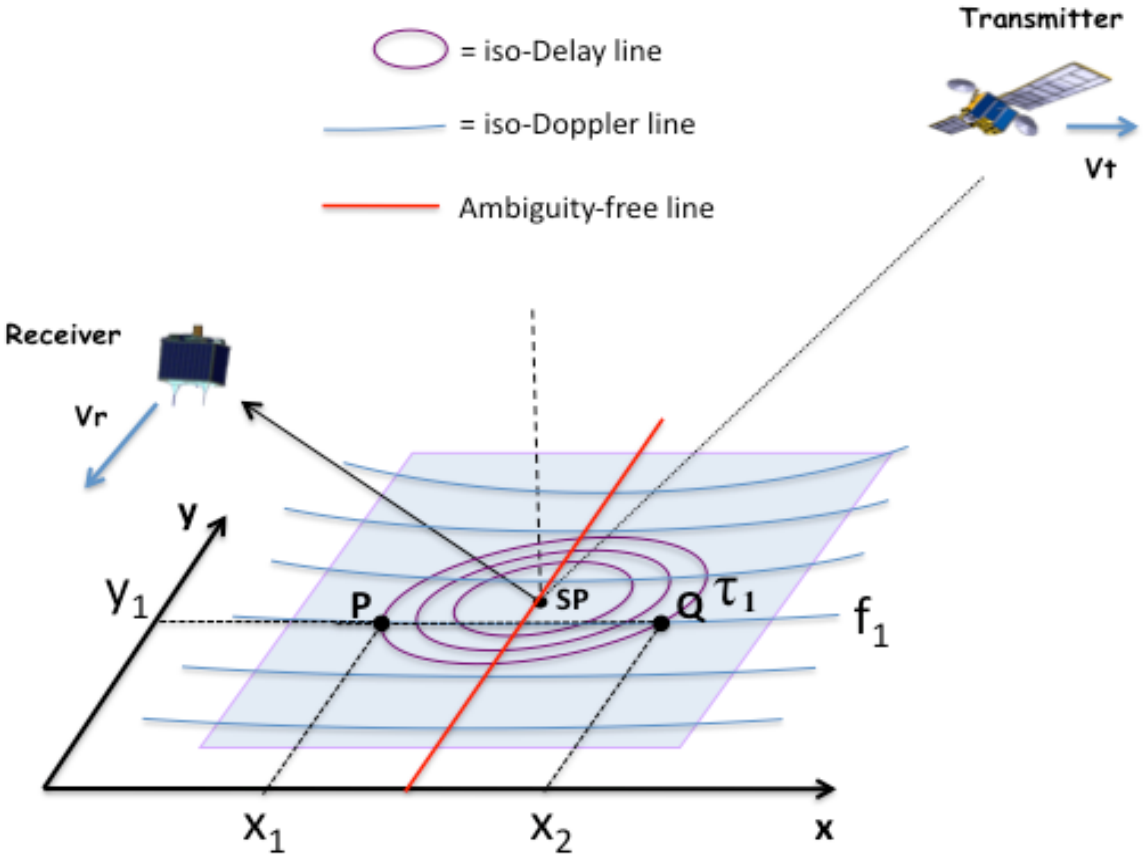


Figure 2.1: Bistatic geometry for GNSS-R systems (left), Delay-Doppler Maps (right) (*Clarizia and Ruf*, 2016b).

The locus of points of constant delay forms an ellipse and is called an iso-range ellipse. The first iso-range ellipse defines the surface ellipse where the path delay is exactly one chip ($\sim$ 300 m for GPS L1 C/A code) longer than the SP. Similarly, lines of common Doppler shift in frequency form the iso-doppler hyperbolas. The glistening zone is the region on the surface where the reflecting facets are likely to scatter the signal towards the receiver. Each point in the glistening zone can be mapped to a

pair of Delay-Doppler values.

The Earth surface viewed by a GNSS-R receiver is a mapping of spatial points on to the Delay-Doppler domain called the Delay Doppler Map (DDM). However, it is to be noted that different points on the sea surface can map to the same Delay-Doppler point as the intersection of Delay and Doppler lines happen at 2 different points on the surface. There is also an ambiguity-free line that has a one-to-one mapping between spatial and delay-doppler coordinates. The ambiguity-free line is a line passing through the specular point and is perpendicular to the Doppler hyperbolas. Away from this line, two different points in space will map to the same DD location, as depicted in Fig.2.1. Though points P and Q are spatially separated, because the same iso-range ellipse and the iso-doppler hyperbola intersect at these two points, both are mapped to the same delay-doppler bin at the receiver.

One other important aspect of the glistening zone in the Delay-Doppler coordinate system is its dependence on the geometry of transmitter and receiver and the surface roughness. It has been shown in (*Clarizia and Ruf*, 2016b) that with higher incidence angles the iso-delay ellipse tend to stretch out and become wider. This makes sense physically because larger incidence angles imply longer propagation paths and therefore larger delays.

2.1.2 The Bistatic Radar Cross-Section

The scattering of GPS signals by the Earth surface is largely dependent on surface roughness. The bistatic radar equation proposed by Zavarotny and Voronovich (2000) is the widely used forward scattering model for GNSS-R applications. This model uses the Geometric Optics (GO) limit of the Kirchhoff Approximation (KA) for a rough surface. The approach discretizes the surface into scattering facets and the final received signal is the superposition of the returns from a large number of scatterers. The net received power for a given delay-doppler bin is given by

$$P_G(\tau, f) = T_i^2 \frac{P_T \lambda^2}{(4\pi)^3} \int\int \frac{G_T(\vec{\rho}) G_R(\vec{\rho}) \Lambda^2(\tau, \vec{\rho}) |S(f, \vec{\rho})|^2 \sigma_0(\vec{\rho})}{R_0^2(\vec{\rho}) R^2(\vec{\rho})} d^2\rho, \tag{2.1}$$

where $P_G(\tau, f)$ is the scattered signal power at delay τ and Doppler frequency f measured by the receiver over the coherent integration time T_i. P_T, λ and G_T are the GPS transmit power, carrier wavelength and antenna gain, respectively. R_0 and R are the transmitter to surface and surface to receiver ranges, respectively. G_R is the receiver antenna gain and σ_0 is the Normalized Bistatic Radar Cross Section (NBRCS). Λ and S represent the GPS spreading function and the Doppler zone function of the GPS, respectively. From this bistatic radar equation the value of σ_0 can be derived from the measurement of P_G as the other parameters are known, assumed or measured.

This model only considers the incoherent component of scattering, which means the surface must be "sufficiently" rough to ensure a large Rayleigh parameter and incidence angles much smaller than grazing angles. If the surface is calm or mirror- like, then the coherent component becomes significant which breaks down the roughness assumption and the derived Normalized Bistatic Radar Cross-Section will tend to infinity.

At the GPS signal wavelength, the ocean can be assumed rough under most conditions. For the GO model, the total power received at any Delay-Doppler bin depends on the probability of occurrence of wave slopes at those locations on the surface. Thus the model distribution of ocean surface wave slopes plays an important role in determining the net received power at the receiver. Several wave spectra models are available to emulate a real ocean surface (*Pierson Jr and Moskowitz*, 1964; *Hasselmann et al.*, 1973; *Fung and Lee*, 1982; *Apel*, 1994; *Elfouhaily et al.*, 1997). Of these, the Elfouhaily omnidirectional wave spectrum is believed to offer the best representation of ocean waves for GNSS-R. A detailed discussion of surface roughness as seen by a bistatic system was discussed in Chapter 1. It is important to note here that the

bistatic cross-section is sensitive to a broader range of roughness scales (discussion of the variation in cut-off wavenumber at lower wind speeds is available in detail in Chapter 4) as compared to the monostatic configuration where the cross-section is sensitive only to a narrow portion of the roughness spectrum due to Bragg resonance.

2.2 Previous reflectometry experiments

The idea of sensing the ocean surface using GNSS signals can be traced back to 1988 (*Hall and Cordey*, 1988) when a multistatic- scatterometry system was proposed. The first aircraft measurements of sea roughness (*Garrison et al.*, 1998) were performed in 1997 using a nadir pointing LHCP (Left Hand Circularly Polarised) antenna. The GPS transmit signal is RHCP, therefore it is justified to use a LHCP antenna on the receive end, as reflection inverts the polarization. At steep to moderate elevation angles, the co-pol scattered signal is negligible but the signal on RHCP (Right Hand Circularly Polarised) and LHCP begin to converge at low grazing angles. Garrison et al. (1998) showed that the cross-correlation between the reflected signal and a reference PRN (Pseudo Random Number) code that was recorded as a function of relative time delay demonstrated sensitivity to surface roughness. The first space-based detection of ocean-reflected GPS signals was reported by Lowe et al. (2002) and their approach was proposed for performing altimetry using GPS reflections.

Aircraft experiments for wind measurements and the development of retrieval algorithms were performed (*Garrison et al.*, 2002) to show that the most significant information is contained in the slope of the trailing edge of the reflected waveform. The first effort to study the high wind regimes found in hurricanes was accomplished in 1998 with flights into the outer bands of Hurricane Bonnie as it made landfall near Topsail Beach, NC. With the cooperation of NOAA, a GPS Delay Mapping Receiver was installed on one of the Hurricane Hunter aircraft in 2000 and acquired the first

GPS-reflected data from inside a hurricane (*Katzberg et al.*, 2001). Gleason et al., (2005) also demonstrated the feasibility of global wind measurements from spacecraft altitudes using the experiment onboard the UK-DMC satellite. A second GPS bistatic radar payload was flown on the TDS-1 satellite in 2014 (*Jales and Unwin*, 2015). The payload consisted of a nadir pointing antenna with a peak gain of 13.3 dB for capturing GPS reflections and a receiver called the SGR-ReSI which generated the DDMs. Early analysis of TDS-1 measurements demonstrated capability of retrieving winds between 3-18 m/s with a Root Mean Square Error (RMSE) of 2.2 m/s (*Foti et al.*, 2015).

Other flight campaigns to study sensitivity to soil moisture and sea ice sensing were performed (*Komjathy et al.*, 2004; *Masters*, 2004). Currently the CYGNSS constellation that was launched in 2016 is making global measurements with unprecedented coverage of ocean surface wind speeds over a broad range of precipitation and complex meteorological phenomena such as hurricanes (*Ruf et al.*, 2018a).

2.3 The CYGNSS mission

The Cyclone Global Navigation Satellite System (CYGNSS) is the first dedicated GNSS-R constellation mission selected by NASAs Earth venture program. The goal of the CYGNSS mission is to study and model the inner core of tropical cyclones (TCs) to accurately forecast their intensification. CYGNSS works at the all-weather GPS L1 frequency to study the wind speed near the eye wall. It is a well-known fact that L band reflections are sensitive to wind speed and the GPS signals are less affected by heavy precipitation. CYGNSS has effectively improved its temporal frequency by utilizing 8 micro satellites that are roughly equally spaced around a 520 km circular orbit inclined at 35 degrees to the equator thereby providing a median re-visit period of around 3 hours and a mean re-visit period of 7 hours. Fig.2.2 shows

an example coverage map of the CYGNSS constellation in 24 hrs.

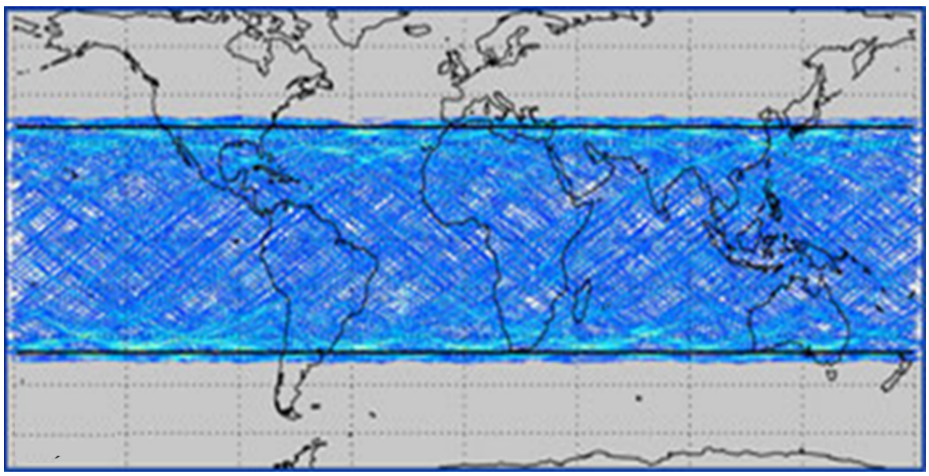


Figure 2.2: CYGNSS spatial coverage tracks over a 24 hr period (*Ruf et al.*, 2016).

Each CYGNSS observatory carries a GNSS-R radar receiver, tuned to the GPS L1 signal at 1.575 GHz, which continuously generates DDMs of surface reflections from specular points within the footprints of its two downward pointing receive antenna beams. The receivers are a third generation version of the Space GPS Receiver Remote Sensing Instrument (SG-ReSI) product line developed by Surrey Satellite Technology Ltd. over the past ∼15 years. The two CYGNSS receive antenna beams are pointed cross track to the direction of orbital motion. Each antenna is a 2x3 element phased array with a fan beam antenna pattern. The patterns cover incidence angles of approximately 5-65 deg and azimuth angles of 75-105 deg and 255-285 deg on the port and starboard sides of the sub-satellite point, respectively. The Delay Doppler Mapping Instrument (DDMI) maps the received GPS signal onto the Delay-Doppler space. It also receives the direct signal using its zenith antenna that helps to determine the location of the observatory. The signals are measured every second and each receiver has 4 channels thereby enabling up to 32 sea surface measurements per second across the constellation.

The CYGNSS mission provides 3 levels of data products. The received DDMs undergo the Level-1 calibration (*Gleason et al.*, 2016) to derive the Bistatic Radar Cross

Section (BRCS) and the normalized areas, which is subsequently used in the retrieval algorithm to extract wind speed information in the viewed region. The inputs to the Level 2 wind retrieval algorithm are the observables extracted from the BRCS (*Gleason*, 2013) and normalized scattering area derived from Level 1 calibration algorithm. An observable must be a best representative of the information contained in the input. The observables chosen for this purpose are the Delay Doppler Map Average (DDMA) and the Leading Edge Slope (LES). A regression based Geophysical Model Function (GMF) is statistically inverted over these observables to retrieve the wind speed (*Clarizia et al.*, 2014). The Level 3 data product is a gridded wind speed map with a spatial grid of 0.2 deg (lat, lon) and a temporal spacing of 60 minutes.

CHAPTER III

CYGNSS Geophysical Model Development

As described in Chapter 1, ocean surface wind speed can be estimated from spaceborne observations made by microwave radiometers and radars. Radiometers measure brightness temperature, from which a surface emissivity observable is derived. The sensitivity to wind results primarily from the generation of foam on the ocean surface, which has a significantly different emissivity than that of water alone (*Wilheit*, 1979; *Chang and Li*, 1998; *Gaiser et al.*, 2004). The estimation of wind speed from the observed emissivity typically relies on a geophysical model function (GMF) that relates the two. GMFs can either be developed from a "first principles" analysis of electromagnetics or, which is more common in practice, they can be constructed empirically by relating a large population of observations to near-coincident measurements of the wind speed made by some other means. Radars measure scattered transmitted signals, from which observables related to the scattering properties of the surface are derived. The sensitivity to wind results primarily from the roughening of the ocean surface, which alters both the radar scattering cross section and the shape of the scattered radar waveform relative to that for a smooth water surface. A GMF approach is also often used with radar wind sensors. A GMF relating wind speed to the scattering cross section observable has been used by ocean scatterometers (*Hersbach et al.*, 2007; *Ricciardulli and Wentz*, 2015). A correction to such a GMF has been considered for

ocean altimeters using additional information about the significant wave height of the ocean contained in the shape of the leading edge of the radar return waveform (*Chelton and McCabe*, 1985; *Glazman and Greysukh*, 1993).

Both scatterometers and altimeters are examples of monostatic radars, which carry a co-located transmitter and receiver and measure the backscatter signal. In contrast, in a bistatic radar configuration, the transmitter and receiver are in different locations. Global Navigation Satellite System-Reflectometry (GNSS-R) instruments, which make use of Global Positioning System (GPS) or other navigation signals as their transmitters, are examples of bistatic radars. Their preferred bistatic measurement geometry is quasi-specular forward scattering because the scattering cross section tends to be largest in that direction. Examples of spaceborne GNSS-R instruments which have successfully measured ocean surface winds are the UK Disaster Monitoring Constellation (UK-DMC) (*Clarizia et al.*, 2009, 2014) and TechDemoSat-1 (TDS-1) (*Foti et al.*, 2015; *Soisuvarn et al.*, 2016; *Foti et al.*, 2017b). These missions have reported their wind speed retrieval performance in the literature, but have not provided the details of the GMFs used by their retrieval algorithms. In addition, both UK-DMC and TDS-1 make measurements at low (near nadir) incidence angles only, due to the orientation of their receive antenna beams, so their GMFs are restricted to those angles. Sensitivity to hurricane force winds has also been demonstrated by TDS-1 (*Foti et al.*, 2017a), but a corresponding high wind GMF has not yet been reported. GNSS-R instruments have also flown on aircraft to measure ocean surface winds (*Garrison et al.*, 2002; *Katzberg et al.*, 2001).

GNSS-R radar receivers measure the scattering cross section of the ocean surface at and in a region surrounding the specular reflection point. The measurements are localized on the Earth surface using a combination of time delay and Doppler frequency filters to form a Delay Doppler Map (DDM) of the surface (*Zavorotny and Voronovich*, 2000). The time delay filter acts similarly as the range gating function

in conventional monostatic radars to isolate regions on the surface with a particular time-of-flight from the transmitter to the receiver. The Doppler filter similarly isolates regions on the surface at which the GPS signal experiences a particular Doppler shift. Numerous observables can be derived from a DDM for use in wind speed retrievals. For airborne applications, wind speed GMFs have been developed based on the relative strength of the scattering at and away from the specular point (*Katzberg et al.*, 2006). This makes use of the fact that specular scattering tends to decrease while diffuse scattering along directions away from the specular point tends to increase as wind speed and surface roughness increase. This approach has the advantage that the observable is a ratio between two received signal strengths, so ancillary measurement parameters such as the strength of the GPS transmitted signal and the gain of the receive antenna, largely cancel out. For spaceborne applications, it is problematic to rely on this ratio-based approach because differences in time delay and Doppler shift correspond to much greater differences in distance from the specular point at satellite altitudes and orbital velocities. As a result, use of such a ratio approach could result in an effective spatial resolution of 100s of km for the derived wind speed, rather than a few km in the case of aircraft (*Clarizia and Ruf*, 2016a). In order to maintain a spatial resolution of 10s of km, spaceborne wind speed retrievals tend to rely only on DDM measurements near the specular point (*Clarizia and Ruf*, 2016b).

DDMs are processed into two observables both of which are used for wind speed retrieval. The first observable is the normalized scattering cross section (σ_0) averaged over an area roughly 25 km in diameter centered on the specular point. The other observable is the leading edge of the slope of the radar return waveform (LES). Both observables are defined in Clarizia (2014), which describes the wind speed retrieval algorithm used by the CYGNSS mission and develops an initial pre-launch GMF based on simulated measurements.

An empirical GMF is developed here for use by the CYGNSS wind speed retrieval

algorithm. It uses near-coincident matchups between the CYGNSS observables and independent estimates of the ocean surface wind speed referenced to a 10 m height (u10). The independent wind speeds are provided by Numerical Weather Prediction (NWP) models at low to moderate wind speeds and by instruments on the NOAA P-3 hurricane hunter aircraft at higher wind speeds in tropical cyclones (*Uhlhorn et al.*, 2007). The GMF developed here is distinct from those reported previously in two primary respects. With previous airborne applications, the observables were either uncalibrated Signal-to-Noise Ratio (SNR) measurements of the received GPS signal, or normalized measures of the extent of diffuse scattering away from the specular point. The GMF developed here instead uses as its observables (a) the absolutely calibrated σ_0 of the surface in the vicinity of the specular direction and (b) the leading edge slope (LES) of the radar return waveform. With previous spaceborne applications, coincident matchup measurements were not available at hurricane force wind speeds from the NOAA P-3 hurricane hunter aircraft. Their availability allows for the extension of the GMF to significantly higher wind speeds than has been reported previously. The GMF development is presented in two stages. First, a Fully Developed Seas (FDS) version is developed based on matchups with NWP model outputs at low to moderate wind speeds. Then a Young Seas/Limited Fetch (YSLF) version is presented based on matchups with measurements by hurricane hunter aircraft during flights through several 2017 Atlantic hurricanes.

3.1 Fully developed seas GMF

The FDS GMF is based on an empirical pairing of CYGNSS Level 1 (L1) observations of σ_0 and LES with the 10 meter referenced ocean surface wind speed (u10), as determined by Numerical Weather Prediction (NWP) models. A large population of these pairings is partitioned into "bins" with respect to u10 and the incidence angle

(θ_{inc}) of the observation. The average values of u10 and the L1 observable within a bin are paired together as one discrete sample of the GMF for that observable. An algebraic parametric model is then fit to the discrete GMF samples to produce a continuously varying GMF so as to be used by the Level 2 (L2) wind speed retrieval algorithm.

3.1.1 Description of training dataset: ECMWF and GDAS matchups

Matchup NWP data used to train the empirical GMF are the 10 meter referenced ocean surface wind speeds provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) and the Global Data Assimilation System (GDAS). ECMWF is an independent intergovernmental organization supported by 34 countries to produce current weather forecasts and climate reanalysis products and to perform related research activities (*Andersson et al.*, 2015). The reanalysis products are used here. GDAS is a system operated by the NOAA National Centers for Environmental Information to organize a variety of surface, balloon, aircraft and spaceborne observations into a gridded model space for use by NOAA's global forecast system. ECMWF and GDAS use a 0.25 deg and 1.0 deg reporting interval, respectively. Bilinear interpolation in space and linear interpolation in time of the reported NWP products are used to estimate u10 at the times and locations of the CYGNSS specular point observations.

A merged "ground truth" u10 product is used to combine model outputs by both ECMWF and GDAS. For wind speeds below 20 m/s, the ECMWF value for u10 is used alone. For wind speeds between 20 and 25 m/s, the arithmetic average of ECMWF and GDAS wind speeds is used. For wind speeds above 25 m/s, the GDAS value alone is used. This merged-product approach is used to accentuate the better accuracy of ECMWF at lower wind speeds and of GDAS at higher wind speeds. In addition, matchups are only used if the u10 values for ECMWF and GDAS differ by

no more than 3 m/s. Examples of the training data set are shown in Fig.3.1 for the σ_0 observable at three values of θ_{inc}. Shown are logarithmic density scatterplots of the merged ground truth u10 values vs. the observable. The general trend is for the scattering cross section to decrease as wind speed increases, as expected for bistatic forward scattering from a wind-roughened ocean surface.

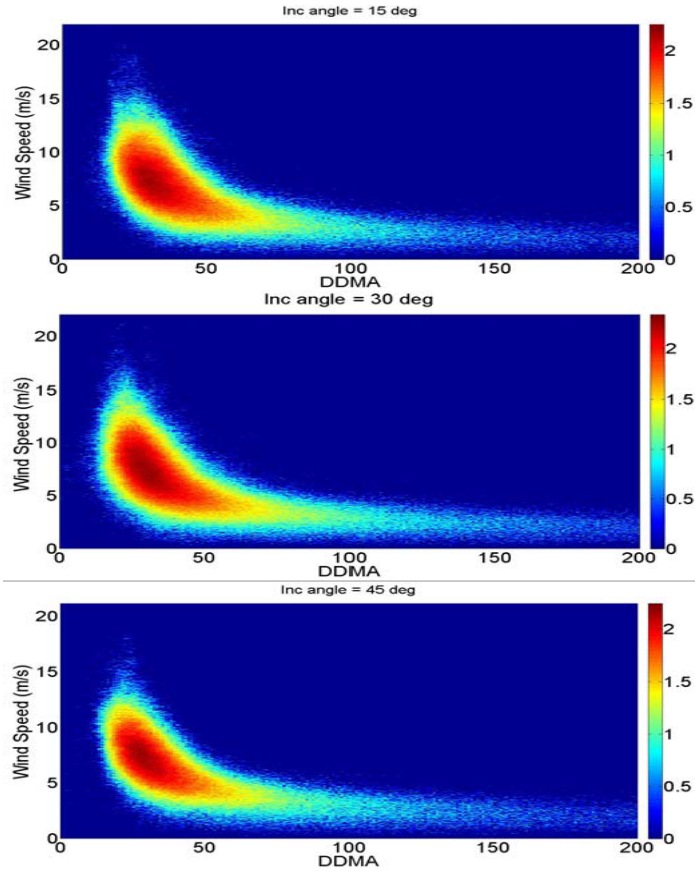


Figure 3.1: Log(density) scatterplots of σ_0 measured by CYGNSS vs. ground truth u10 at incidence angles of 15 deg (top), 30 deg (mid) and 45 deg (bot). The color scale is the log10 of the number density of points.

3.1.2 Binning of matchups for discrete empirical GMF

The CYGNSS L1 observables are filtered prior to use as part of the training used to derive the empirical GMF. The filters are used for reasons of quality control. Specifically:

1. The Doppler coordinate of the specular point in the DDM is required to be greater than the lowest possible value in the map and less than the highest possible value. This discards cases where it is at the edge of the map and the computed Doppler coordinate may be incorrect. In practice, this happens less than 0.1% of the time.

2. The delay coordinate of the specular point in the DDM is required to be greater than the lowest possible value in the map and less than the highest possible value. This discards cases where it is at the edge of the map and the computed delay coordinate may be incorrect. In practice, this happens less than 0.1% of the time.

3. All non-numeric values of the observables are discarded. This eliminates samples for which noise in the calibration data can produce non-physical calibrated L1 data. In practice, this occurs less than 0.1% of the time.

4. The observables are required to be non-negative. This eliminates samples for which noise in the calibration data has produced non-physical calibrated L1 data, as well as measurements that are very close to the measurement noise floor. In practice, this occurs less than 0.1% of the time.

5. All measurements are discarded for which the spacecraft star tracker is not tracking due to solar contamination. Some reported spacecraft attitude data during sun outages are known to be erroneous (with inaccuracies greater than the error allocation in the L1 calibration algorithm for attitude knowledge). This only occurs when the outage is especially long, but all sun outage data are flagged and removed as a precaution. In practice, this occurs less than 1% of the time.

6. All data with a CYGNSS Range Corrected Gain (RCG) of less than 10 are discarded. RCG is a composite measure of receive signal strength that combines

the receive antenna gain in the direction of the specular point with the R-2 propagation range loss from the GPS transmitter to the specular point and from the specular point to the CYGNSS receiver. In practice, data with an $RCG > 1$ can typically produce useful wind speed retrievals, but only data with higher SNR values are used to train the empirical GMF.

7. All observations resulting from transmissions by the GPS Block Type II-F satellites are discarded. Block II-F is the newest family of GPS satellites, and the one for which the CYGNSS team has the least information about its transmitter antenna gain pattern. There are currently 8 II- F satellites in the constellation, out of 31 total in operation.

The behavior of the empirical GMF as a function of u10 and θ_{inc} is smoothed by allowing sequential bins in either dimension to overlap. In the incidence angle dimension, the bin center is incremented every 1 deg from 1 to 70 deg and all samples are included within $+/-2.0$ deg of the center. In the wind speed dimension, the bin center is incremented every 0.1 m/s from 0.05 to 34.95 m/s and all samples are included within a variable bin width that varies according to the population density of samples as a function of wind speed. Specifically, the bin widths used are:

- $+/-0.4$ m/s ($u10 < 2m/s$)
- $+/-0.3$ m/s ($2 < u10 < 5m/s$)
- $+/-0.2$ m/s ($5 < u10 < 9m/s$)
- $+/-0.4$ m/s ($9 < u10 < 11m/s$)
- $+/-0.6$ m/s ($11 < u10 < 14m/s$)
- $+/-0.8$ m/s ($14 < u10 < 17m/s$)
- $+/-1.0$ m/s ($17 < u10$)

A weighted average of all samples within twice of these bin width ranges is performed. Samples within +/− one bin width of the bin center are given twice as much weight as those between 1 x binwidth and 2 x binwidth from the bin center. This tapered weighting approach reduces the introduction of artificial higher frequency components into the GMF than are present in the original samples.

The GMF is also forced to be monotonic as a function of wind speed. The GMF value at 7.05 m/s is computed first (since this is generally the most probable wind speed and so it has the largest population of samples in its near vicinity). GMF values are then sequentially computed in steps of 0.1 m/s above and below this value using the averaging scheme described above. However, values are allowed to either decrease or stay the same with increasing wind speed, and increase or stay the same with decreasing wind speed. This limits the introduction of non-physical variations into the GMF due to undersampling of certain parts of the (wind speed, incidence angle) state space. In practice, this monotonicity algorithm is only enforced at the highest and lowest wind speeds in the population, where the sampling density tends to be lowest.

Examples of the empirical GMFs for both observables (σ_0 and LES) at $\theta_{inc} = 30$ deg, overlaid on the training data from which they were derived, are shown in Fig.3.4. Over the central range of wind speeds where most of the samples occur, the GMF agrees with the highest density part of the scatter plot. At the highest and lowest wind speeds, the smaller size of the population makes the behavior of the GMF more susceptible to errors. The flowchart for GMF development is shown in Fig.3.2.

Examples of the empirical GMFs for both L1 observables across a range of incidence angles are shown in Fig.3.3. The general dependence of observable on wind speed is consistent across all incidence angles. The dependence on incidence angle at a given wind speed is also consistent. Note that the slope of the GMF ($dObs/du10$) is highest at low wind speeds, indicating that wind speed retrievals will, in general,

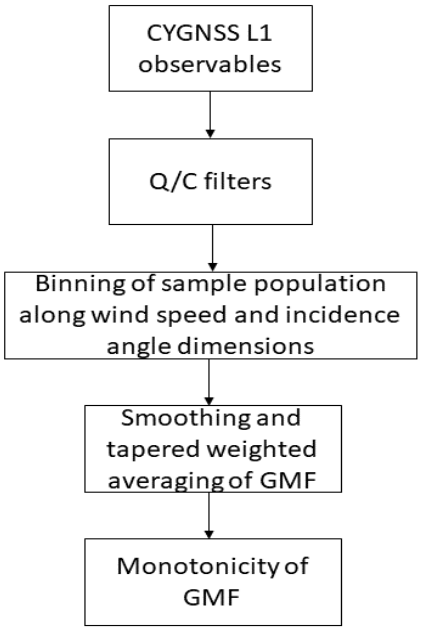


Figure 3.2: Process flowchart for deriving empirical GMFs for the two Level 1 observables, σ_0 and LES.

perform better and be less susceptible to measurement noise and calibration errors at the lower wind speeds. The general behavior of the empirical σ_0 GMF, both as a function of wind speed and of incidence angle, is consistent with scattering model predictions based on the first order small slope approximation method (*Zavorotny and Voronovich*, 2014; *Ruf et al.*, 2016).

An algebraic parametric model is fit in a least-squares sense to the empirical GMF in order to populate the lookup tables used by the CYGNSS Level 2 wind speed retrieval algorithm. This process smooths across some of the behavior in the empirical GMF that is related to measurement noise and insufficient number of samples in the training set. It also interpolates across portions of (wind speed, incidence angle) state space that are not sufficiently populated by the training set. The parametric model assumed for the GMF is divided into two portions based on the observed behavior of the empirical GMF as a function of wind speed. At low wind speeds, a model is

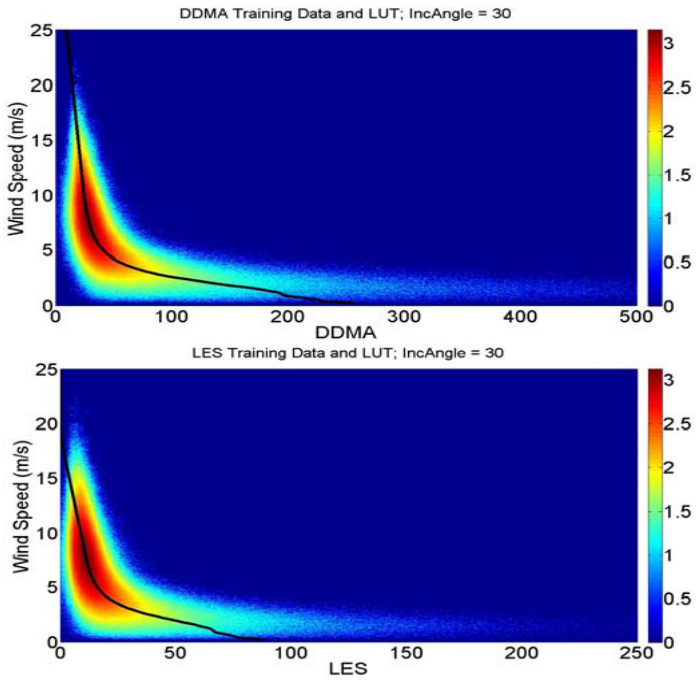


Figure 3.3: Empirical GMFs for the two Level 1 observables, σ_0 or DDMA (top) and LES (bot), at $\theta_{inc} = 30$ deg, overlaid on log(density) scatter plot of the training data from which they were derived.

assumed of the form

$$Obs = a_0 + a_1 u^{-1} + a_2 u^{-2} \tag{3.1}$$

where Obs is the Level 1 CYGNSS observable (either σ_0 or LES), u is the ground truth u10 wind speed, and a_i are the dependent parameters of the model. At high wind speeds, a model is assumed of the form

$$Obs = b_0 + b_1 u + b_2 u^2 \tag{3.2}$$

where b_i are the dependent parameters of the model. The population of samples used to train these two models is different for each observable. For σ_0, samples at wind speeds below 15 m/s are used to determine a_{0-2} and samples above 15 m/s

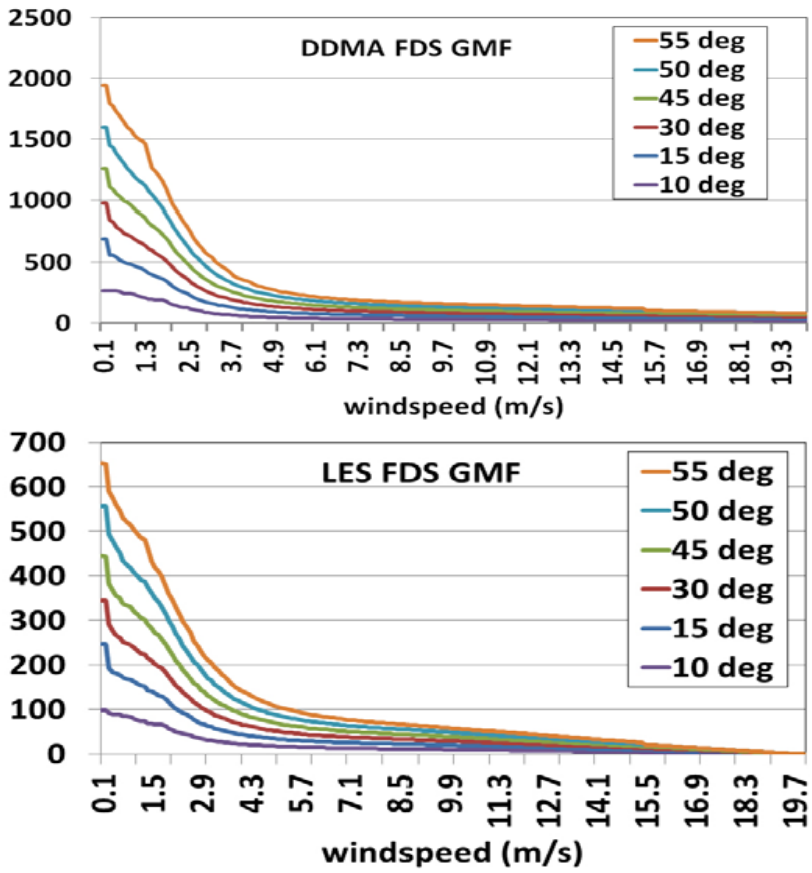


Figure 3.4: Empirical GMFs for σ_0 (top) and LES (bot) at θ_{inc}=10 15,...,55 deg.

are used to determine b_{0-2}. For the LES observable, samples below and above 10 m/s are used. The actual transition from one parametric model to the other in the GMF occurs near, but not exactly at, these wind speed values. A transition point is chosen where the first derivatives of the two models are equal (i.e. a spline fit). The model parameters and the spline fit transition point are chosen independently at each incidence angle.

Examples of the parametric model GMF, together with the empirical GMF from which they are obtained, are shown in Fig.3.5. At the lowest wind speeds (below ~ 2 m/s), the empirical and parameteric models tend to deviate (more so for the LES observable). The sensitivity of the LES observable to wind speed drops to zero at wind speeds above ~ 18 m/s, whereas the σ_0 observable retains its sensitivity up to wind speeds of ~ 30 m/s.

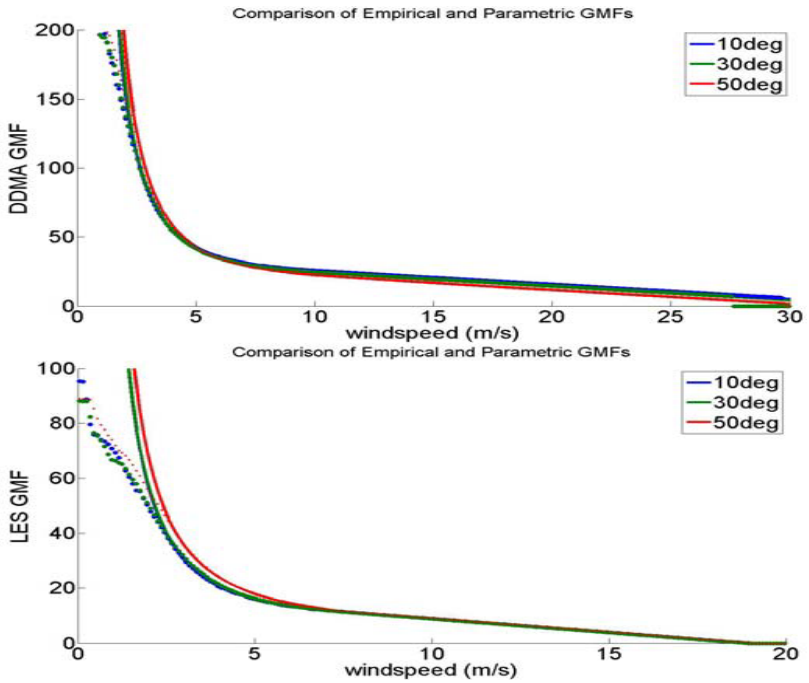


Figure 3.5: Discrete empirical GMFs (symbols) and continuous parametric models (solid lines) derived from them, for the two L1 observables σ_0 (top) and LES (bot), and for θ_{inc} = 10, 30, 50 deg.

Examples of the parametric model GMFs for both observables and over a range of incidence angles are shown in Fig.3.6. The truncation of the GMF at high wind speeds results from the limitations in the dynamic range of wind speeds included in the training data set.

3.1.3 Validation and performance characterization

As a means of assessing whether the derived GMFs properly represent the response of the Level 1 observables to changes in ocean surface wind speed, they are used as the basis for a wind speed retrieval algorithm. The algorithm inverts the GMF to estimate wind speed given the measured observable (*Clarizia and Ruf*, 2016b). The error in this retrieval algorithm (groundtruth retrieval) is considered as a function of the two coordinates of the GMF, incidence angle and wind speed. Note that this com-

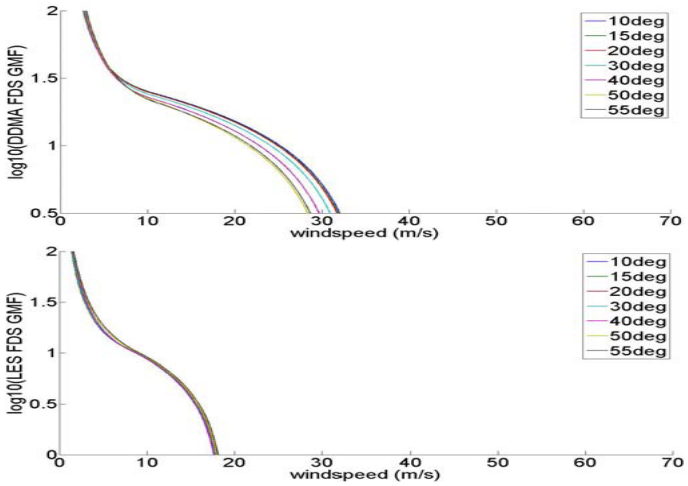


Figure 3.6: Parametric model FDS GMF for the Level 1 observables σ_0 (top) and LES (bot) at incidence angles of 10, 15,..., 55 deg. The dependence on incidence angle is more pronounced in the case of the σ_0 (DDMA) observable. The maximum wind speed at which the observable is sensitive to changes in wind speed is also higher for σ_0 than for the LES observable.

parison uses the same population of data as was used to train the GMFs and so should be considered a test of internal consistency in the generation of the GMFs. A more independent assessment of retrieval performance is presented in Ruf et al.(2018).

The dependence of retrieval error on incidence angle is shown in Fig.3.7. The σ_0-based retrieval on the left shows a positive statistical retrieval bias (retrieved values are larger than groundtruth more often than they are smaller). The LES- based retrieval on the right shows a more unbiased distribution of retrieval errors. Notably, in terms of incidence angle dependence, the highest density of retrieval errors occurs near an error of zero and this is true at all incidence angles and for both L1 observables.

The dependence of retrieval error on the ground truth wind speed is shown in Fig.3.8. Significant positive retrieval biases (retrieved values are larger than groundtruth more often than they are smaller) can be seen at ground truth wind speeds of 5- 15 m/s for σ_0 and at 3-10 m/s for LES. A possible cause for this behavior, and a corresponding mitigation strategy, are considered next.

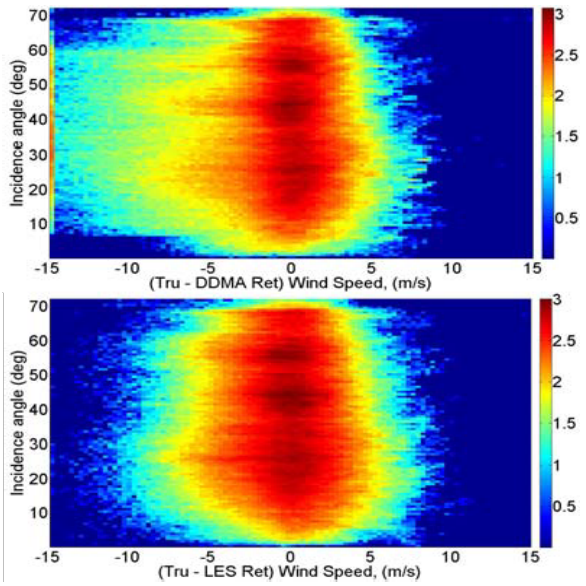


Figure 3.7: Log(density) scatter plot of wind speed retrieval error vs. incidence angle for retrievals using the σ_0 (top) and LES (bot) Level 1 observable. There is no significant dependence on incidence angle.

Fig.3.9 shows the dependence of retrieval error on the difference between the wind speeds retrieved using the two L1 observables. Larger retrieval errors tend to be highly correlated with the difference between the two retrievals. The large positive bias in retrievals based on the σ_0 observable tends to coincide with cases where the σ_0 retrieval is much larger than that from the LES observable. Alternately, samples for which the σ_0 retrieval is much smaller than the LES retrieval tend to coincide with large positive biases in the LES retrieval. The root cause of this behavior may be related to the fact that the two observables respond to different aspects of the sea state, only part of which is forced by local wind speed. If, for example, they respond in different ways to the longer swell portion of the surface wave spectrum, this could explain their different dependence on the retrieval error. One hypothesis is that young sea conditions (not fully developed) may coincide with instances where the two retrievals have significant differences.

In terms of mitigation of this behavior, and improvement in the overall wind speed retrieval performance, the dependence of retrieval error on the difference between the

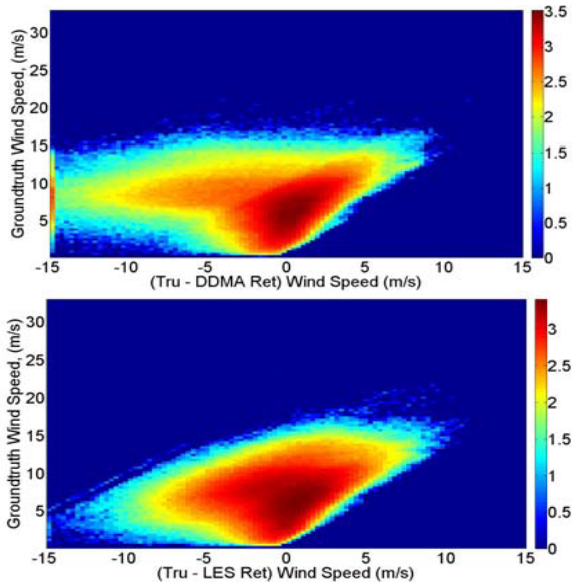


Figure 3.8: Log(density) scatter plot of wind speed retrieval error vs. ground truth wind speed for retrievals using the σ_0 (top) and LES (bot) L1 observable.

σ_0 and LES retrieved wind speeds can be used as a quality control (Q/C) filter. This is illustrated in Fig.3.10, which shows the dependence of the retrieved Minimum Variance (MV) wind speed on the difference. The MV retrieval is a weighted average of the σ_0 and LES retrievals, weighted by the inverse variance of the error in wind speed retrieved by each of the individual observables (*Clarizia and Ruf*, 2016b). Large errors in the MV retrieval can be seen in Fig.3.10 to correlate with large differences between the σ_0 and LES retrievals. A simple Q/C filter could, for example, exclude all retrievals for which the difference is greater than 6 m/s. This filter threshold is illustrated in Fig.3.10. In practice, this Q/C filter discards approximately 4% of the samples.

The effectiveness of the Q/C filter, and the overall quality of the MV retrieval algorithm, is illustrated in Fig.3.11, which compares the ground truth and retrieved wind speeds as a scatter plot and by their RMS and mean difference. The large retrieval biases evident in Fig.3.8 have been largely removed by this Q/C filter. Fig.3.8 also illustrates the performance of the wind speed retrieval below 20 m/s. The RMS

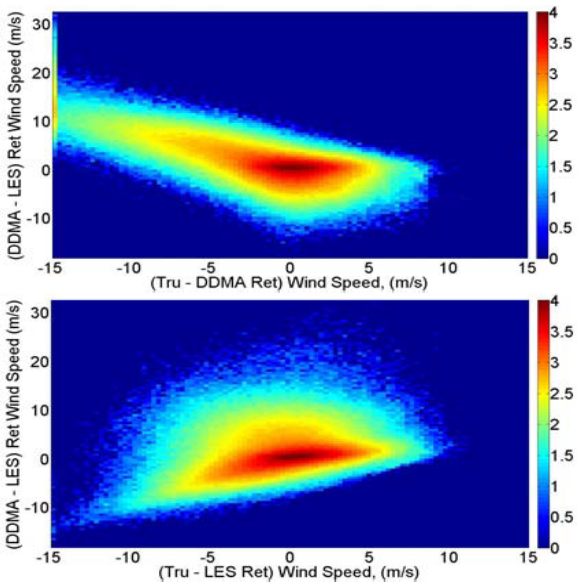


Figure 3.9: Log(density) scatter plot of wind speed retrieval error using the σ_0 (top) and LES (bot) L1 observable vs. the difference between the two retrieved wind speeds.

difference is $\sim$ 2 m/s at low wind speeds and grows to $\sim$ 4 m/s at 20 m/s.

3.2 Young seas/limited fetch GMF

The YSLF GMF is based on matchups between measurements by CYGNSS made during overpasses of 2017 Atlantic hurricanes and near-coincident ocean surface wind speed measurements made by the Stepped Frequency Microwave Radiometer (SFMR) on NOAA P-3 hurricane hunter aircraft (*Uhlhorn et al.*, 2007). These matchups demonstrate a fairly consistent sensitivity of the CYGNSS L1 observables to changes in wind speed at high (30-60 m/s) levels.

The mean high wind sensitivities ($d\sigma_0/du10$ and $dLES/du10$) are used to define a YSLF GMF that is consistent with the fully developed seas GMF at low wind speeds but whose high wind behavior is replaced by the YSLF sensitivities derived from the SFMR matchups over hurricanes.

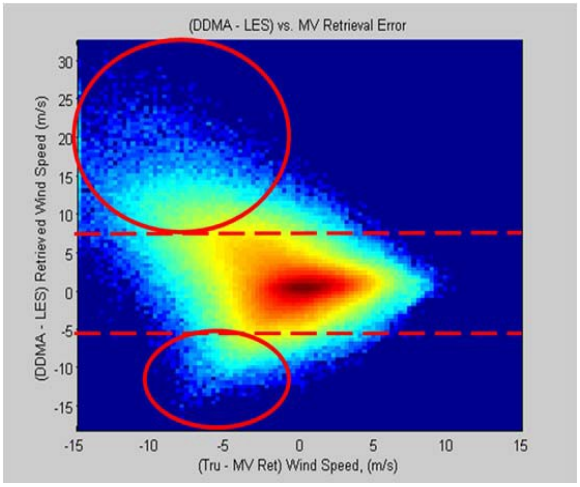


Figure 3.10: Log(density) scatter plot of minimum variance wind speed retrieval error vs. the difference between the wind speeds retrieved by the two individual L1 observables. A Q/C filter that discards retrievals with differences greater than 6 m/s (shown by the two dashed red lines) will eliminate the circled regions with large retrieval errors.

3.2.1 Description of training dataset: NOAA P-3 SFMR matchups

Twenty-five (25) coincident overpasses of hurricanes between CYGNSS and NOAA P-3 aircraft occurred during the 2017 Atlantic hurricane season. Coincidence was defined by locating the aircraft ground track during one of its eyewall penetrations that was closest to a CYGNSS specular point track for that overpass and requiring that they occurred within 60 min of one another. The 25 cases identified in this way include overpasses of Hurricanes Harvey, Irma and Maria. The maximum SFMR wind speed recorded across all cases was 73 m/s and the range of CYGNSS incidence angles covered 13-67 deg. Histograms of the SFMR u10, CYGNSS θ_{inc}, and CYGNSS σ_0 measured across all 25 overpasses are shown in Fig.3.12.

An example of one of the coincident overpasses, occurring over Hurricane Maria on 24 Sep 2017, is shown in Fig.3.13. The CYGNSS measurement of σ_0 can be seen to decrease roughly monotonically as the wind speed measured by SFMR increases.

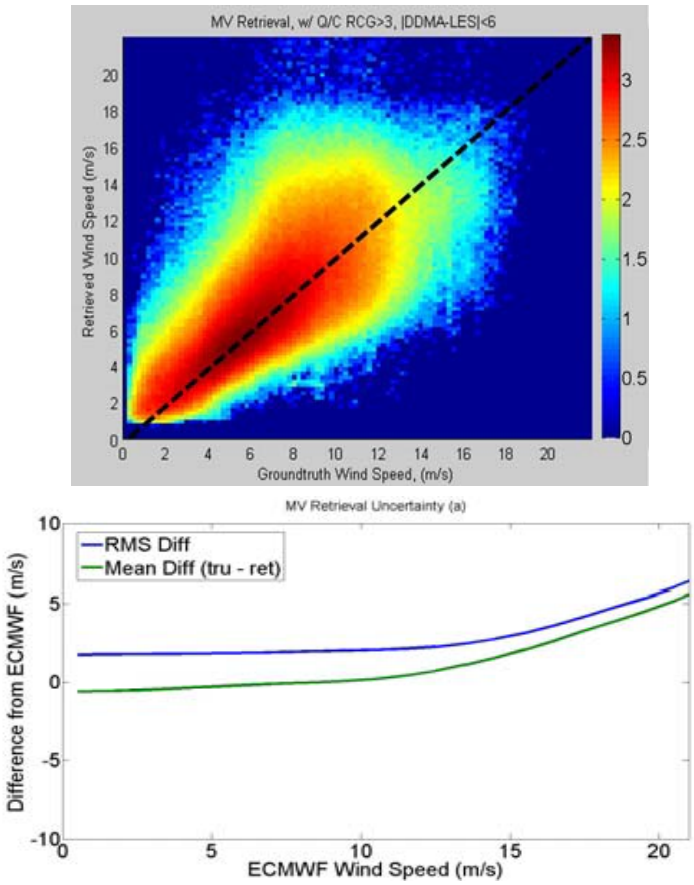


Figure 3.11: Log(density) scatter plot of groundtruth vs. MV retrieved wind speed (top) with black-dashed line of 1:1 agreement, and of RMS and mean retrieval error vs. groundtruth wind speed (bot).

3.2.2 Regression of coincident overpasses to determine GMF

For each of the 25 coincident hurricane overpasses, the L1 observables are related to the coincident SFMR wind speeds by linear regression. One example of this is shown in Fig.3.14, for the case illustrated in Fig.3.13. The slope of the linear regression is taken as the sensitivity of the observable to changes in wind speed.

Fig.3.14 shows the slope of the linear regression determined for each of the 25 cases and for both L1 observables. The set of all 25 regression slopes is averaged together to determine the sensitivity of the YSLF GMF. The resulting sensitivity factors are

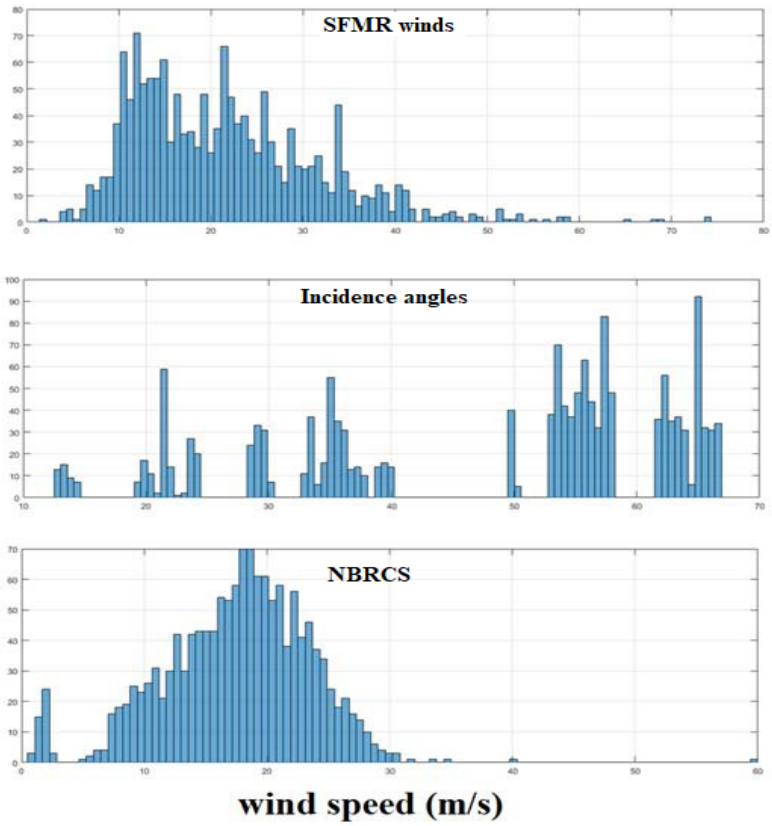


Figure 3.12: Histograms of the SFMR u10, CYGNSS θ_{inc}, and CYGNSS σ_0 measured across all 25 coincident hurricane overpasses that are used to determine the high wind portion of the limited fetch GMF.

$$\begin{aligned} d\sigma_0/du_{10} &= -0.1880(m/s)^{-1} \\ dLES/du_{10} &= -0.0929(m/s)^{-1} \end{aligned} \tag{3.3}$$

3.2.3 Parametric model with hi/lo wind partitions

As with the fully developed seas GMF, an algebraic parametric model is assumed for the YSLF GMF. The parametric model is again divided into low and high wind speed portions. At low wind speeds, a similar model is assumed as for the FDS GMF

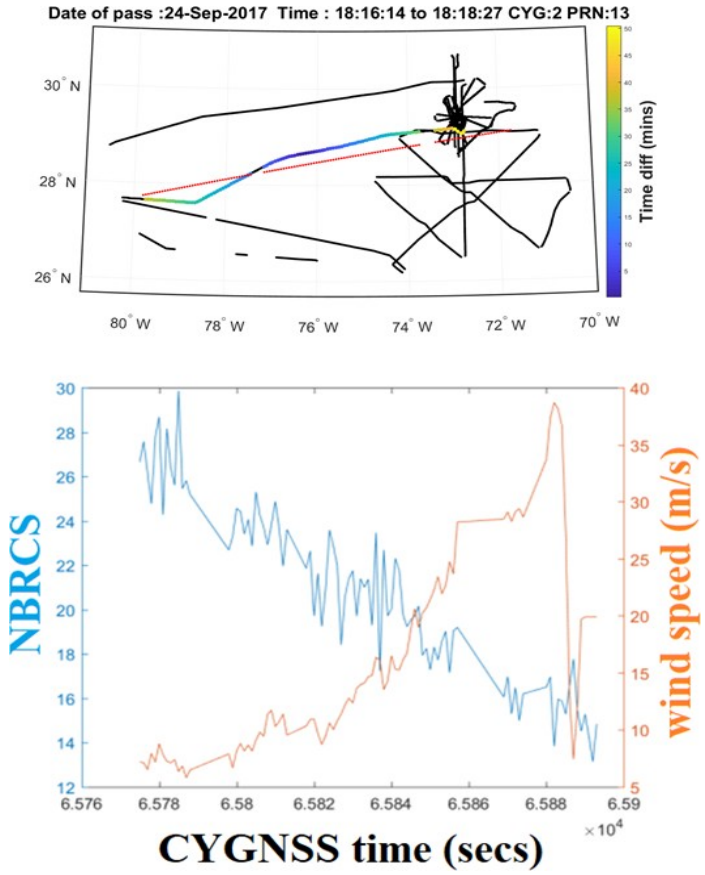


Figure 3.13: Coincident CYGNSS/P-3 overpass of Hurricane Maria on 24 Sep 2017 at 1817 UTC. (top) The P-3 ground track is shown in black. The CYGNSS specular point track is shown in red. The colored portion of the P-3 track is color coded by the time difference. (bot) CYGNSS L1 σ_0 (labeled NBRCS) and SFMR wind speed measured along the coincident track.

$$Obs = a_0 + a_1 u^{-1} + a_2 u^{-2}, \tag{3.4}$$

and the a_{0-2} coefficients are again determined using the fully developed seas training set. At high wind speeds, a linear model is assumed

$$Obs = c_0 + c_1 u, \tag{3.5}$$

with slope coefficient (c1) given by eqn.3.3 as determined from the linear regression

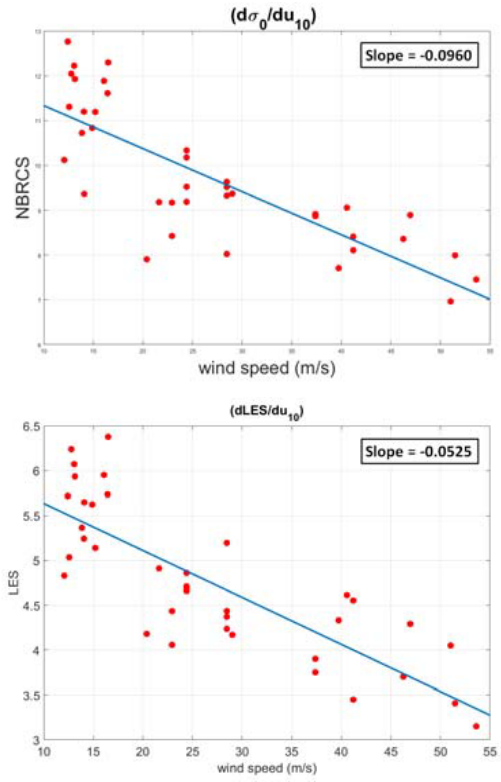


Figure 3.14: Linear regression of CYGNSS L1 σ_0 and LES observables against SFMR wind speed for Hurricane Maria overpass on 24 Sep 2017 at 1817 UTC. The slope of the linear regression is noted in each plot.

of hurricane overpass matchups. The transition between low and high wind speed segments is again selected as the wind speed where the first derivatives of the two models are equal. The low wind speed model parameters and the spline fit transition point are chosen independently at each incidence angle. A common high wind speed sensitivity is assumed for all incidence angles, since there was no clear dependence of sensitivity on incidence angle evident in the coincident hurricane overpass data.

Examples of the YSLF GMF are shown in Fig.3.16, together with the FDS GMFs at the same incidence angles. Several characteristics are noteworthy. The models agree at low wind speeds (by design). At wind speeds in the range 15-25 m/s, the sensitivity ($dObs/du10$) is markedly stronger in the FDS case. The limiting wind speed, above which the value of the observable is zero, is much higher in the YSLF

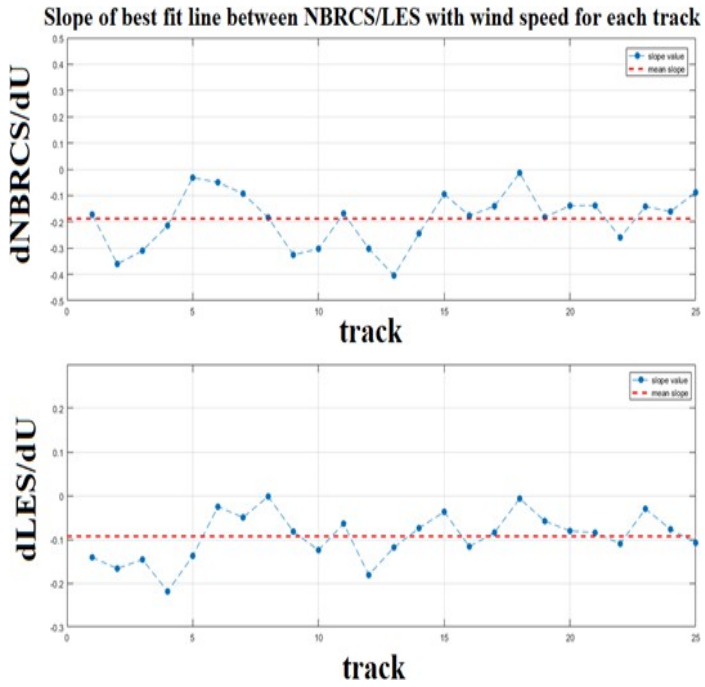


Figure 3.15: Slope of the linear regression determined from each of the 25 coincident hurricane overpasses for both the σ_0 (top) and LES (bot) L1 observable.

case. And finally, there is a range of values of both observables over which two different wind speeds correspond to the same measurement. This represents an ill-posed, multi-valued inversion problem. In practice, it may be necessary to have some *a priori* knowledge about the fetch or sea age of the conditions under observation in order to uniquely convert L1 observable measurements to wind speed.

The multi-valued nature of the mapping from L1 observable to wind speed is illustrated in Fig.3.17. The FDS and YSLF GMFs agree below $\sim$ 12m/s. Above 12 m/s, the behavior of the FDS GMF is derived from matchups with ECMWF and GDAS away from major storms. The behavior of the YSLF GMF above 12 m/s is derived from matchups with P-3 SFMR wind speed measurements in major storms. The YSLF GMFs for both L1 observables are shown in Fig.3.18.

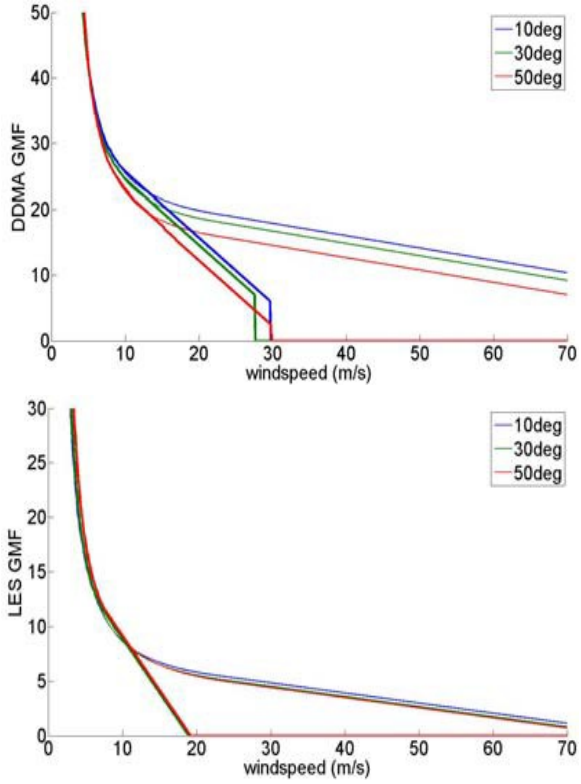


Figure 3.16: Young Seas/Limited Fetch GMF (thin lines) for σ_0 (top) and LES (bottom), together with the Fully Developed Seas GMF (thick lines), at θ_{inc} = 10, 30, 50 deg. The models agree at low wind speeds but have very different high wind speed dependencies.

3.2.4 Validation and performance characterization: Repeatability of limited fetch conditions

The adequacy of the YSLF GMFs to represent the response of the Level 1 observables to changes in ocean surface wind speed in hurricanes is assessed by using it in a wind speed retrieval algorithm. The algorithm inverts the GMF to estimate wind speed given the measured observable, in the same manner as the previous assessment for the FDS GMF. In this case, given the limited number of coincident hurricane overpasses that are available, individual case studies are considered rather than overall performance statistics. The wind speed retrieval performance for each hurricane overpass is evaluated using CYGNSS/ P-3 matchups similar to the one shown in Fig.3.13. The

retrieval results for four overpasses are shown in Fig.3.19, presented as a time series of the CYGNSS retrieved wind speed overlaid with the nearest SFMR and merged ECMWF & GDAS wind speeds.

In general, the lower wind speed portions of the CYGNSS tracks agree well with the ECMWF & GDAS wind speeds and the high wind speed portions near the storm center agree with the SFMR wind speeds. Note that the CYGNSS retrievals at the highest wind speed levels tend to be noisier, as can be expected given the lower slope of the GMF at high wind speeds. The wind speed retrieved by CYGNSS using the FDS GMF is also included in Fig.3.19 for comparison. The CYGNSS YSLF and FDS winds agree closely at low wind speed speeds, which is consistent with their very similar GMFs at low winds. At high wind speeds, the YSLF retrievals are higher than the FDS ones, which is also consistent with the higher values for σ_0 in the YSLF than the FDS GMF at the same wind speed (see Fig.3.17). Occasional drop-outs in the reported FDS winds are evident in Fig.3.19 near the storm center. They result from quality control filters which flag the retrievals as unreliable when the individual FDS DDMA and LES retrievals differ by more than 6 m/s.

3.3 Discussion

The dependence of CYGNSS measurements on the local wind speed at the location of the measurement is multi-valued in the sense that different wind speeds can result in the same values for σ_0 and LES. The relationship appears to be strongly dependent on sea age, with fully developed seas generally exhibiting a repeatable, single-valued mapping. This is demonstrated by the RMS difference between CYGNSS retrieved winds and coincident NWP matchups. As shown in Fig.3.11, the RMS difference is between 2 and 3 m/s for NWP wind speeds below 15 m/s, then begins to rise in conditions that are more likely to include partially developed seas. For the young seas

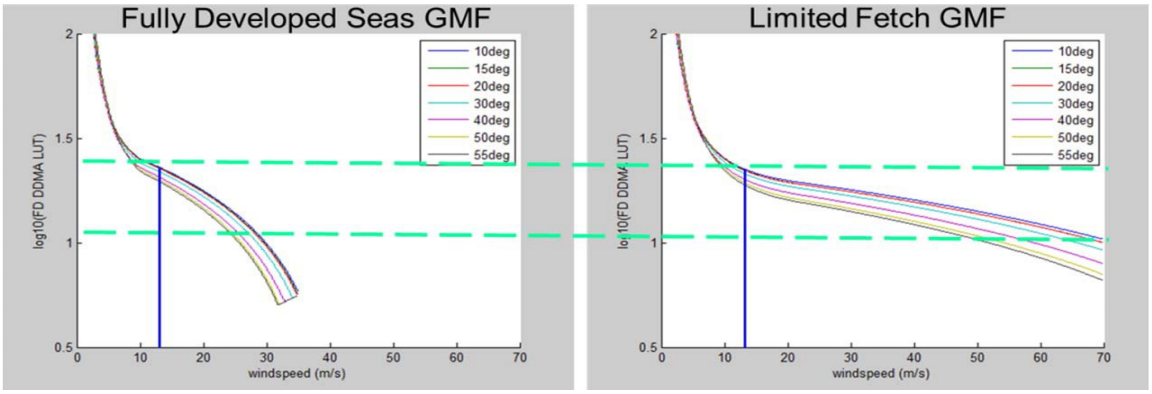


Figure 3.17: Side-by-side comparison of GMFs for σ_0 that are appropriate for Fully Developed Seas (left) and Young Seas/Limited Fetch conditions in hurricanes (right). The region of the L1 observable bounded by the two dashed green lines maps onto two distinct wind speeds from the same value of σ_0. The wind speed marked by the vertical blue line (12 m/s) represents the highest wind speed with a common mapping by both GMFs.

with limited fetch that are more typical of conditions in and near tropical cyclones, coincident matchups with airborne SFMR measurements indicate significantly higher values for the σ_0 and LES measurements than are observed in the fully developed seas cases, given the same wind speed. This general trend continues for matchups near the inner core of the tropical cyclones, at wind speed values which are above those reported by NWP models.

A detailed assessment and characterization of the performance of a CYGNSS wind speed retrieval algorithm based on the GMF is performed(*Ruf et al.*, 2018b). To evaluate performance below 20 m/s, a large ($\sim$ 30 million) population of retrieved winds using the FDS GMF is compared to near-coincident winds reported by ECMWF. The RMS difference between them is found to be 2.0 m/s and the component of that difference due to uncertainty in the CYGNSS wind speed retrieval is estimated to be 1.4 m/s.

Above 20 m/s, performance is evaluated by comparisons between winds retrieved using the YSLF GMF and near- coincident winds measured by the SFMR instrument

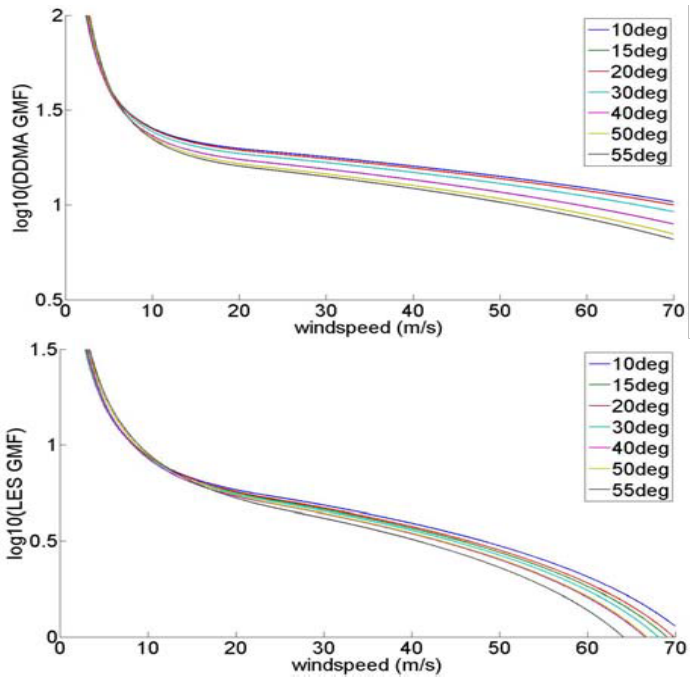


Figure 3.18: Representative YSLF GMFs for the Level 1 observables σ_0 (top) and LES (bot) at incidence angles of 10, 15,..., 55 deg. The dependence on incidence angle is more pronounced at higher wind speeds with both observables.

on the P- 3 hurricane hunter aircraft during eyewall penetrations. In this case, the population of intercomparisons contains 674 samples. The RMS difference between samples is found to be 6.5 m/s and the component of the difference ascribed to uncertainty in the CYGNSS retrievals is 5.0 m/s. The significantly larger uncertainty in CYGNSS retrievals at high wind speeds is attributed to result from two primary causes. One cause is the decrease in sensitivity of the L1 observables to changes in wind speed as the winds increase. This is illustrated in Fig.3.17 by the decrease in slope of both the FDS and YSLF GMFs as wind speed increases. A second cause for the increased retrieval uncertainty at high winds is the sensitivity of the L1 observables to sea state conditions not directly related to wind speed; in particular, to sea age or fetch length in and near tropical cyclones.

The multi-valued dependence on wind speed can be explained by considering the general relationship between GNSS-R measurements and sea state, and the relationship

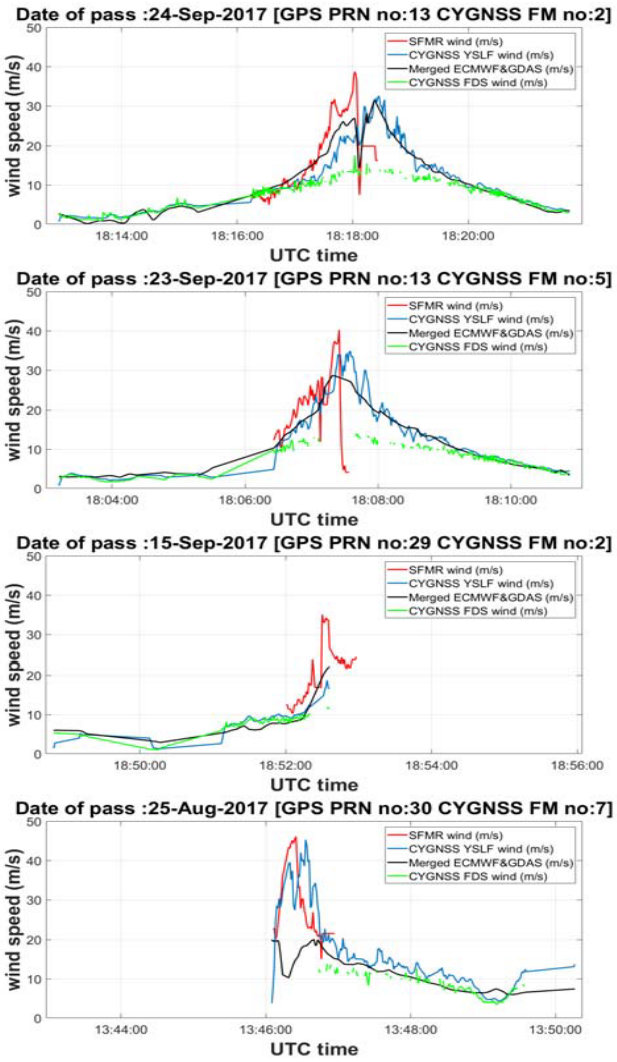


Figure 3.19: Coincident hurricane overpasses by CYGNSS and underflights by the NOAA P-3 hurricane hunter aircraft.

between the sea state at a particular location and the local wind speed there. In general, local winds tend to generate surface roughness nearly instantaneously at the smaller, capillary, wavelength end of the surface height spectrum. The influence of winds on the longer wavelength swell portion of the spectrum takes longer to develop, both in time and in fetch length. For this reason, young seas in limited fetch conditions will tend to have a smaller longwave portion of their spectrum. In fully developed seas, the wind speed has sufficient time to influence the full roughness spectrum and the relationship between capillary and swell waves is more consistent. This

general behavior is consistent with the fact that the derived empirical GMFs map values for σ_0 and LES to higher wind speeds in young sea/limited fetch conditions than in fully developed seas.

The existence of multiple GMFs presents implementation challenges for a wind speed retrieval algorithm in terms of deciding which one to use and under what conditions. Use of the FDS version appears to perform well in most cases away from major storms. Likewise, the YSLF version performs well with most coincident hurricane overpass cases, but not all. The use of a single non-FDS GMF should be considered an approximation to the true dependence of the GMF on sea age or fetch length. It reflects an effective average of the relationship between L1 observables and sea state across the young seas/limited fetch conditions that were present in the 2017 Atlantic hurricanes from which the GMF was derived. A more accurate accounting for the departure from a fully developed sea state might, for example, use a fetch-dependent parametrization of the YSLF GMF, or it might modify the L1 observables based on sea age or fetch length in order to estimate an effective FDS values. These are possible improvements that are under consideration for future development of the GMF.

Rather than attempting to make corrections to the wind speed retrieval algorithm, to account for the sensitivity of the L1 measurements to other aspects of the sea state than those directly forced by local wind speed, an alternative approach might be to directly assimilate the L1 measurements into a coupled wind/wave model that is able to predict GNSS-R measurements given a known sea state using an appropriate rough surface scattering model. Direct data assimilation has proven useful in other situations in which measurements are not uniquely determined by a single geophysical parameter, and this may be the case here as well.

CHAPTER IV

Impact of Rain on GNSS-R Measurements

For several decades, scatterometers have provided valuable ocean surface wind information by measuring the backscattering properties of the surface. The sensitivity to wind in such measurements is due to Bragg scattering from the wind-driven centimeter scale capillary waves. The presence of rain, perturbing the measured radar cross-section, affects about 7% of global scatterometer measurements (*Weissman et al.*, 2012). Other meteorological phenomena often accompany the presence of rain, so understanding its influence on the measurements is critical. Multiple studies have been conducted to examine the effects of rain contamination on scatterometer-derived wind information (*Tang et al.*, 2013, 2015).

When raindrops strike a water surface, the disturbance creates a central column of water referred to as a stalk and a ring of waves that propagate radially outward from the point of impact (*Craeye et al.*, 1997). Studies on the effect of rain on the water surface suggest that at microwave frequencies the dominant surface features, when viewed at oblique angles of incidence, are the ring waves generated by raindrops hitting the water surface, and the stalks are the relevant features at grazing incidence (*Bliven et al.*, 1997; *Craeye et al.*, 1997; *Wetzel*, 1990; *Bass et al.*, 1968). Craeye et al. (1997) have provided a ring-wave spectrum for rain simulated using in-situ experiments over a broad range of rain rates from 5 to 200 mm/h. This ring-wave spectrum

has been widely used to account for the effects of rain in scatterometric measurements (*Xu et al.*, 2014; *Zhang et al.*, 2016). There have also been many physics-based and empirical models of rain effects from C to Ku band (*Nielsen and Long*, 2008; *Stiles and Yueh*, 2002; *Draper and Long*, 2004; *Contreras and Plant*, 2006).

Global Navigation Satellite System-Reectometry (GNSS-R) offers an alternative approach to measure ocean surface wind speed (*Clarizia et al.*, 2009; *Ruf et al.*, 2013; *Foti et al.*, 2015; *Soisuvarn et al.*, 2016). GNSS-R measures reflected GPS signals to extract useful remote sensing information about the reflecting surface. GNSS-R sensors operate in bistatic mode, with the transmitter on the Global Positioning System (GPS) satellite or other navigation platforms. GNSS-R ocean wind measurements are similar to scatterometers in that they are also sensitive to the ocean surface roughness caused by winds, except in the quasi-specular forward scattering direction (*Zavorotny and Voronovich*, 2000).

There have been some initial studies of the effects of rain on GNSS-R measurements for altimetry and ocean wind measurements (*Asgarimehr et al.*, 2018; *Ghavidel and Camps*, 2016). For instance, the study by Asgarimehr et al. (2018) reports around 0.7 dB drop in the magnitude of the Normalized Bistatic Radar Cross-Section from TDS-1 measurements (*Foti et al.*, 2015). However, due to the limited availability of data, there have so far been no clear conclusions drawn about the different types of impacts of rain on the GNSS-R measurements. There have also been theoretical simulations of rain eects on the ocean surface to account for the splash effect (*Zavorotny and Voronovich*, 2018) and to account for attenuation effects (*Asgarimehr et al.*, 2019). Zavorotny and Voronovich (2018) show that the splash effect is visible only at very low wind speeds ($< 5m/s$) where specular scattering can be observed and Asgarimehr et al. (2019) have shown that the total attenuation effect on retrieved wind speed is very small. With the onset of the Cyclone Global Navigation Satellite System (CYGNSS) mission, a large number of GNSS-R measurement samples is now

available on a global scale which encompasses a wide range of wind speeds and rain rates. This has enabled the possibility to analyze and understand the individual effects of rain on GNSS-R measurements more comprehensively.

CYGNSS operates at the all-day all-weather GPS L1 frequency (1.575 GHz) to measure the Normalized Bistatic Radar Cross-Section (σ_0) of the ocean surface. Though the L-band signals can penetrate through clouds and rain, under very heavy precipitation there are other surface effects of rain that can affect the measurements (*Balasubramaniam and Ruf*, 2018; *Wan et al.*, 2019).

This chapter is broadly divided into the following sections: Section 4.1 describes the models proposed to account for different effects of rain. Section 4.2 uses the CYGNSS observations to empirically characterize the perturbations in σ_0, followed by Section 4.3, where a comparison of model and observation-based inferences are made. Finally, the chapter concludes with discussion and conclusions, where key observations are reiterated, and their physical implications discussed.

4.1 Theoretical model

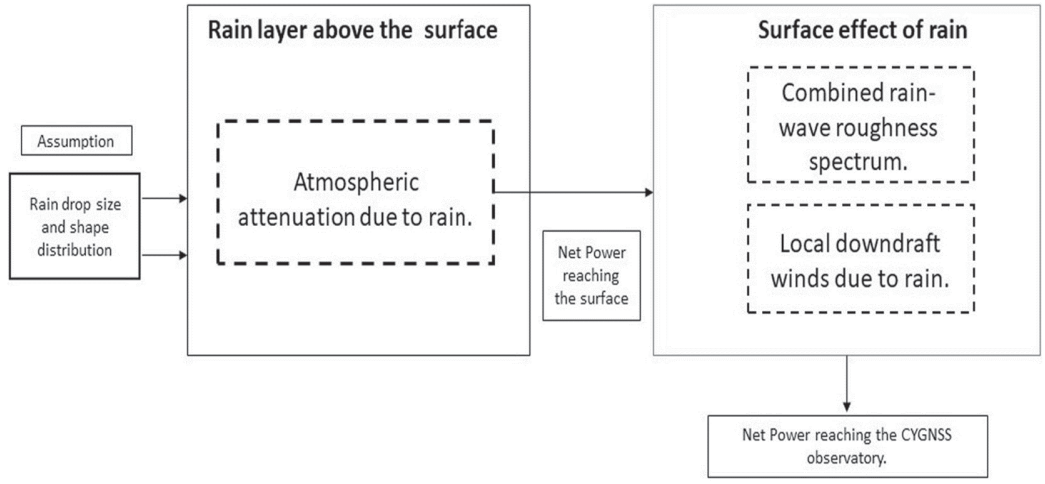


Figure 4.1: Model for the effects of rain on GNSS-R measurements.

The effects of rain on the GNSS-R forward scatter signal can be broadly divided into

3 major components. The first component is the path attenuation of the propagating signal. The second component is related to the effects of raindrops splashing on the water surface, modifying the ocean surface roughness. The third component is related to how precipitation induces locally generated winds by pushing down the air column. In this work, we try to incorporate physics-based principles and direct observations to model each of these effects. Fig.4.1 is the proposed rain model used to account for these effects. GPS satellites transmit L1 C/A navigation signals (1.575 GHz) from a Medium Earth Orbit (MEO) at approximately 20,200 km altitude. These signals are scattered from the ocean surface and received by the CYGNSS observatories in Low Earth Orbit (LEO) at an altitude of approximately 520 km. In the event of rain, attenuation occurs along two paths, first as the GPS signal propagates from the satellite to the ocean surface. The signal is then scattered from the surface of the ocean, with its scattering cross-section dependent on the roughness of the surface. The surface is perturbed by the falling rain drops in two ways. At very low wind speeds (near specular scattering regime), a rain drop hitting the surface can cause modifications to the roughness spectrum. This effect becomes negligible when the surface roughness is dominated by wind generated capillary waves (diffused scattering regime). At higher rain rates, the rain can also induce localized winds by pushing down the air column (*Weissman et al.*, 2012). The reflected L1 signal is then attenuated a second time along its return path towards the CYGNSS observatory. Fig.4.2 provides a pictorial representation of this sequence of interactions.

4.1.1 Radiative transfer model for propagation loss

Absorption of the GPS signal along its propagation path can be modelled using the radiative transfer equation. Let L_a represent the transmissivity of the medium between the transmitter and the surface. For a plane wave, the transmissivity is the negative exponential of the optical depth. The optical depth τ_p can be realized as

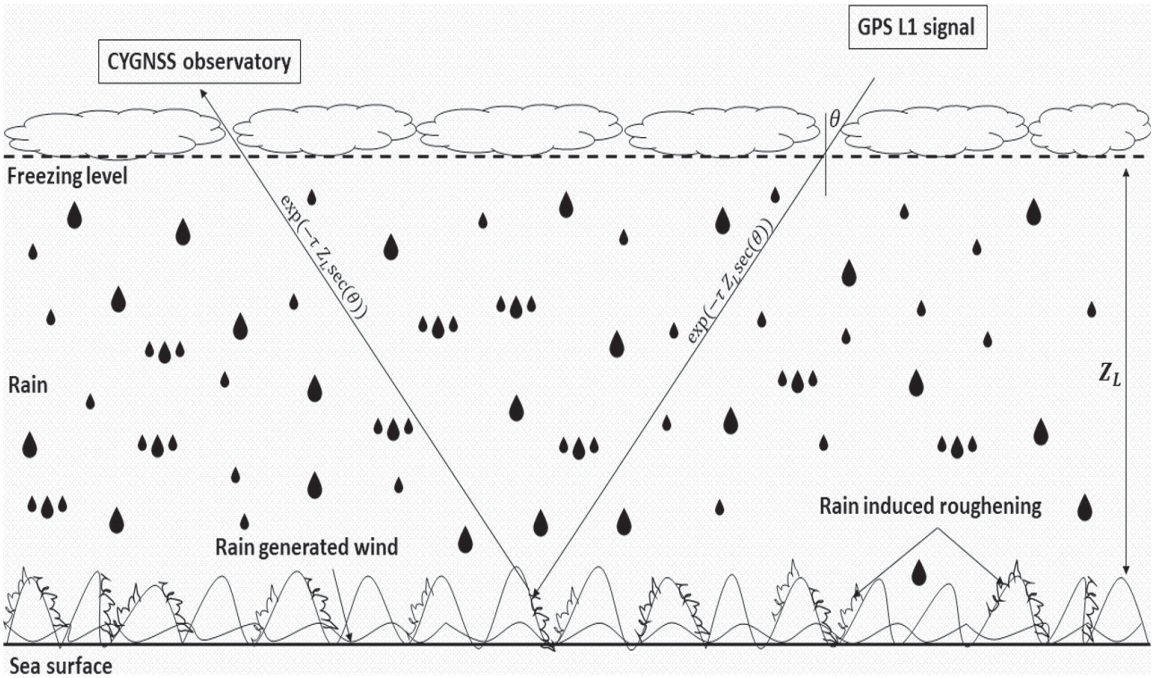


Figure 4.2: Modifications of the received reflected signals through rain.

the slant path integration of specific absorption from the surface up to the freezing level. The integration height used here is the average height of the freezing level in the tropics which is approximately 4800 m (*Harris Jr et al.*, 2000). The model for absorption due to hydrometeors is derived using a regression analysis from Ulaby et al. (2014). The equations governing propagation path loss are given by Eq. (4.1), Eq. (4.2), Eq. (4.3),

$$L_a = exp(-\tau_p) \tag{4.1}$$

$$\tau_p = \int_0^{Z_L sec(\theta)} \alpha_p(z) dz \tag{4.2}$$

$$\alpha_p = \alpha_1(f) R^{b(f)} \tag{4.3}$$

where f is the frequency in GHz, R is the rain rate in mm/hr and $\alpha_1(f)$ and $b(f)$ are regression-based model coefficients given by (*Ulaby et al.*, 2014),

$$\alpha_1(f) = 6.39 * 10^{-5} f^{2.03} \tag{4.4}$$

$$b(f) = 0.851 f^{0.158} \tag{4.5}$$

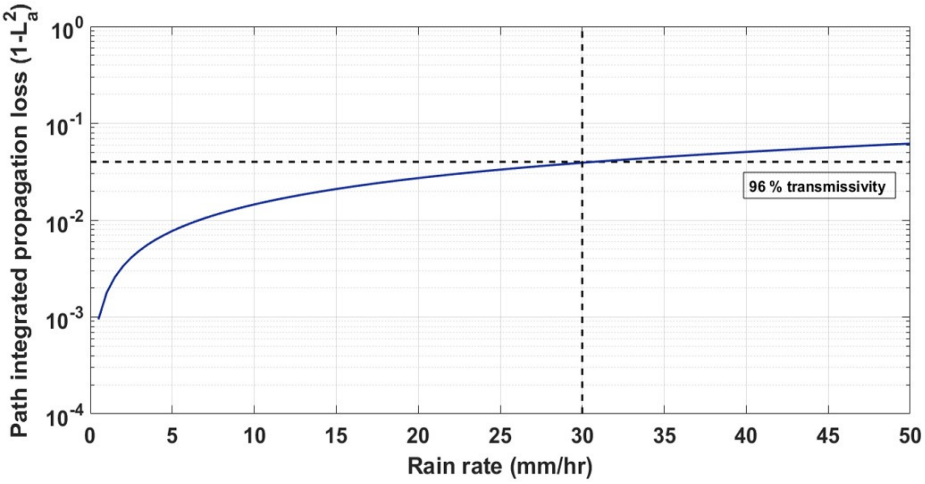


Figure 4.3: Path integrated propagation loss.

These expressions are valid for frequencies up to 2.9 GHz and therefore can be used at the GPS L1 frequency of 1.575 GHz. The double-path integrated loss is shown as a function of rain rate in Fig.4.3. The model indicates that there is a total transmissivity of greater than 96% up to $30mm/hr$. This model re-attests to the assumption that rain has negligible attenuation up to $30mm/hr$ for GPS signals. The modified bistatic scattering forward model, including the double-path attenuation due to rain, is given by Eqn.4.6.

$$P_G(t,f) = T_i^2 \frac{P_T \lambda^2}{(4\pi)^3} Ł_a^2 \int\int \frac{G_T(\vec{\rho}) G_R(\vec{\rho}) F(\vec{\rho}) \Lambda^2(t,\vec{\rho}) |S(\vec{\rho})|^2 \sigma_0^{obs}(\vec{\rho})}{r_0^2(\vec{\rho}) r^2(\vec{\rho})} d^2\vec{\rho} \tag{4.6}$$

where L_a^2 is the transmissivity accounting for rain attenuation. $P_G(\tau, f)$ is the coherently processed scattered signal power at time delay τ and doppler shift f over the coherent integration time T_i. P_T, λ and G_T are the GPS transmit power, carrier wavelength and antenna gain, respectively. r_0 and r are the transmitter-to-surface and surface-to-receiver ranges respectively. G_R is the CYGNSS receiver antenna gain and σ_0 is the Normalized Bistatic Radar Cross Section. Λ and S represent the GPS spreading function and the Doppler zone function of the GPS, respectively. The attenuation-corrected normalised bistatic radar cross-section, $\sigma_0^{obs}|_R$ in linear scale can be derived by inverting the above model:

$$\sigma_0^{obs}|_R = \frac{P_G(4\pi)^3 r_0^2 r^2}{P_T \lambda^2 G_T G_R} \frac{1}{L_a^2} \tag{4.7}$$

Path attenuation due to rain at microwave frequencies plays a minor role in the change in radar cross section. However, properly correcting for its effect will help to accurately characterize the other, more dominant, components such as the modification to surface roughness and local downdraft winds created by the rain.

4.1.2 Modelling surface effects of rain

The global distribution of wind speeds over the ocean has a mean in the range of $6 - 8m/s$. Based on the wind speed, a fully developed ocean wind-wave spectrum can be broadly broken down into 2 regimes of electromagnetic scattering. The first is the near specular regime, for which wind speeds are generally less than 5 m/s, which has a surface Rayleigh parameter much smaller than 1. The second is the diffuse scattering regime, with wind speeds >= 5 m/s, in which very large surface Rayleigh parameters hold and where the geometric optics approximation is valid (*Voronovich and Zavorotny*, 2017; *Zavorotny and Voronovich*, 2018). In the diffuse scattering

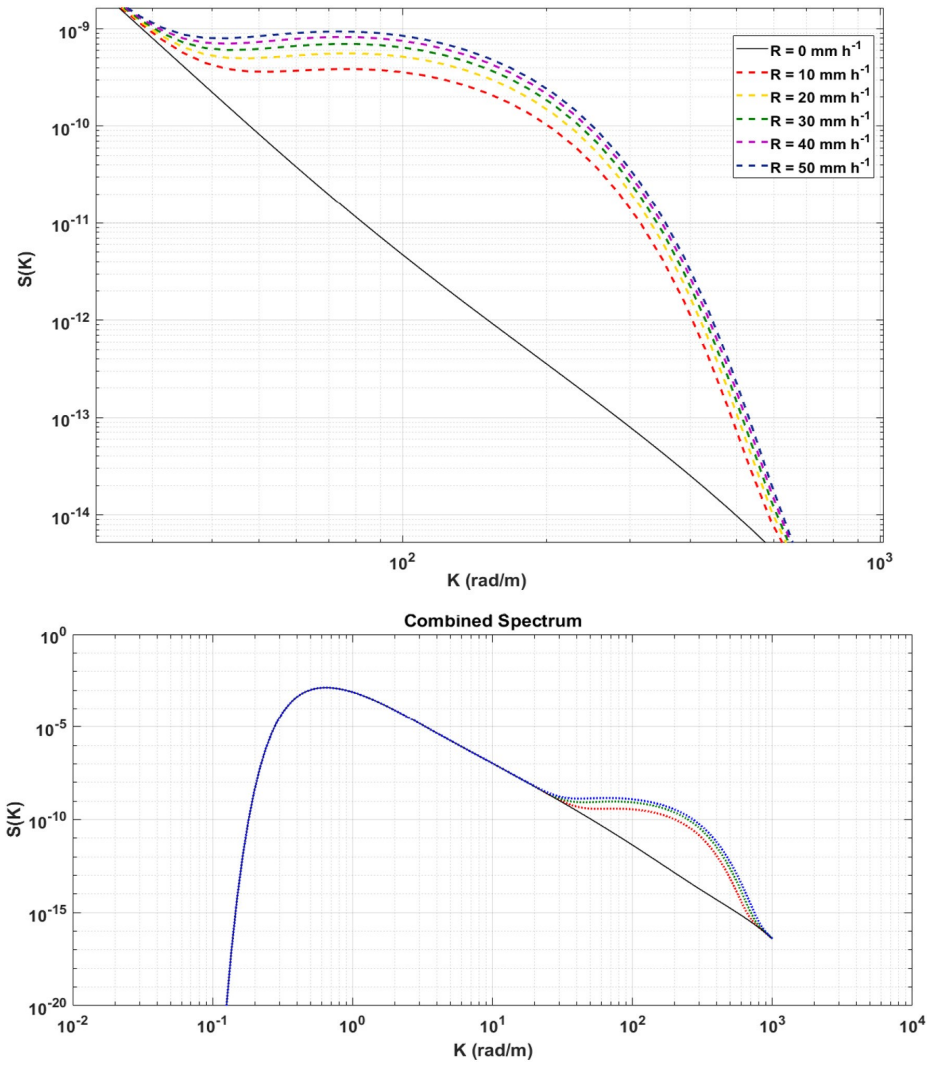


Figure 4.4: Superposition of Elfouhaily and Ring wave spectra to generate the combined rain-wind roughness spectrum for 5 different rain rates and a wind speed = 5 m/s. (top) Closeup of portion of the spectrum affected by rain. (bottom) Full spectrum including the dominant lower wind roughened portion.

regime, surface roughness due to wind-wave interactions dominates over rain induced roughness. The induced surface roughness due to a rain drop hitting the water surface is dominant only at low wind speeds (< 5 m/s) (*Zavorotny and Voronovich*, 2018).

A rain drop hitting the water surface causes 3 major types of surface deformations: crater, crown and stalk. First, a cavity on the water surface is formed along with a crown which then changes into a vertical stalk of water that finally subsides to generate rings of gravity capillary waves (*Bliven et al.*, 1997). The dominant features in this process with respect to microwave frequencies are the stalks and the ring waves. At oblique incidence angles the cm-wavelength ring waves play the dominant role while the stalks play a dominant role only at grazing angles (*Craeye et al.*, 1997). The approach used to model the impact of raindrops on the water surface follows the widely used idea (*Craeye et al.*, 1997) of first order superposition of the rain generated ring wave spectrum and the wind-forced Elfouhaily elevation spectrum (*Elfouhaily et al.*, 1997). The ring wave spectrum is a log-gaussian model of the form:

$$S(k) = \frac{1}{2\pi} v_{gr}(k) S_{peak} exp(-\pi \frac{ln(\frac{f(k)}{f_p})}{\frac{\Delta f}{f_p}}) \tag{4.8}$$

where v_{gr} is the group velocity in m/s, S_{peak} is the power law spectral model given by $6 * 10^{-4} R^{0.53}$, $f(k)$ can be derived from the dispersion relation $w^2 = gk + hk^3$. The remaining quantities are $\Delta f = 4.42 + 0.0028RHz, f_p = 5.772 - 0.0018RHz$. These model coefficients were estimated in (*Craeye et al.*, 1997) for ring waves of wavelengths of 1.94 to $20.8cm$.

The resulting combined rain-wind spectrum for an example wind speed of $5m/s$ and different rain rates is shown in Fig.4.4. The wave spectrum describes the distribution of energy at different wavenumbers. The Elfouhaily wind-wave spectrum has its peak between 10^{-1} and 1 rad/m waves and the ring-wave spectrum has its peak between

$100 - 300$ rad/m. Larger wavenumbers represent shorter waves thus the ring waves introduce additional short-scale roughness over the long scale roughness induced by wind. The GNSS-R σ_0 observations are inversely proportional to the mean square slope of the surface. The mean square slope is the integral of the omnidirectional slope spectrum (*Elfouhaily et al.*, 1997) up to a cut-off wave number which typically is determined using the geometric optics approximation for electromagnetic scattering. At the GNSS-R L-band wavelength, the cutoff wavenumber is approximately $10 - 12$ rad/m. Note that the portion of the roughness spectrum that is perturbed by rain is predominantly above this cut-off. The implications of this relationship are addressed in detail in the Results section.

4.1.3 Rain induced local winds

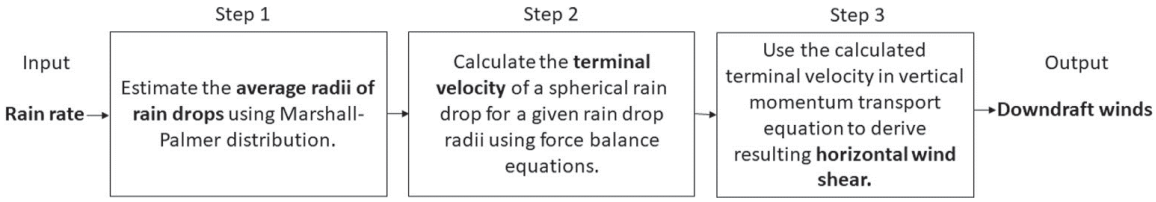


Figure 4.5: Algorithm to calculate rain induced local winds.

In addition to the signal attenuation effect of rain and the rain induced surface roughness due to the drop-splash effect, rain induced local winds can be generated by the downdraft of the air column below a rain cell. This downdraft can alter the surface wave spectrum by raising the near-surface wind speed due to outflow (*Weissman et al.*, 2012). Accurate understanding of this effect entails the consideration of various physical processes together. The end goal of this subsection is to derive a physics-based relationship between surface rain rate and its induced downdraft winds. The proposed algorithm to calculate the downdraft winds is represented in Fig.4.5.

We first derive the drop size distribution of the raindrops using the widely used Marshall-Palmer exponential rain drop distribution (*Marshall and Mc. K. Palmer*,

1948). This distribution provides an empirical relationship between rain rate and rain drop size distribution as given by:

$$N_D = N_0 exp(-\lambda D)$$
$$\lambda = 41 R^{-0.21} cm^{-1}, \tag{4.9}$$

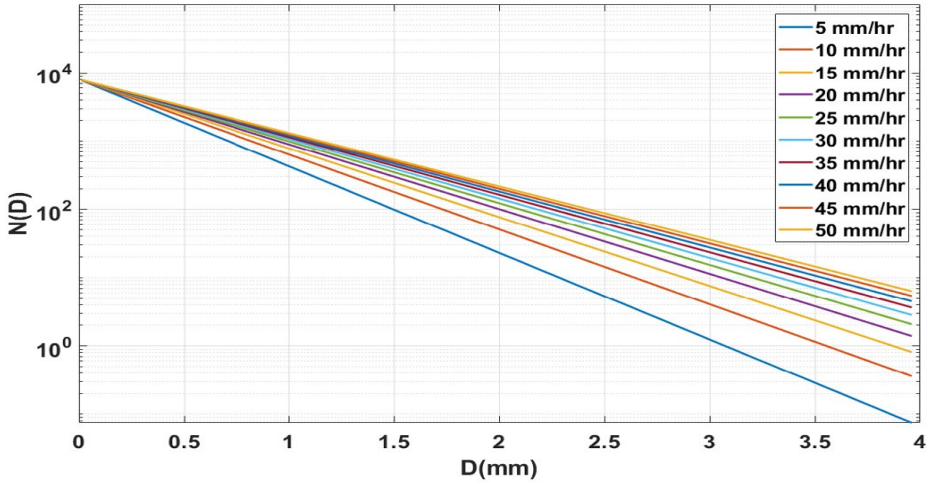


Figure 4.6: Marshall-Palmer distribution based weighed average estimate of rain drop radii.

where N_D is the drop number density for a given diameter of rain drop, D and rain rate R. The number density distribution derived from the above expression is shown in Fig.4.6. The size of the rain drops varies between $0.01mm$ and $4mm$, above which surface tension weakens compared to gravity and the drop splits into smaller sizes. An important assumption made while using the Marshall-Palmer distribution is that the variation in drop size distribution with cloud type and geographic region introduce only second order variability relative to the other major effects described here. Using this density distribution, the half-order moment of the radius of the rain drop, which is required to compute the terminal velocity of the falling rain drops, can be derived as a function of rain rate.

The terminal velocity of the falling raindrops is computed using an empirical power

law model (*Atlas and Ulbrich*, 1977)(*Doviak et al.*, 2006), as given by:

$$v(D) = 386.6D^{0.67} \tag{4.10}$$

where D is the rain drop diameter in meters and v is the terminal velocity. Terminal velocity can be computed from the above equation by using twice the half order moment of the radius derived from the Marshall-Palmer distribution for the diameter in the above equation.

The derived terminal velocity can now be used in the momentum conservation equations (*Holleman*, 2001) to calculate the rain induced downdraft wind speed (derivation provided in Appendix A). The model relating the downdraft wind speed to the surface rain rate derived using the above method is depicted in Fig.4.7.

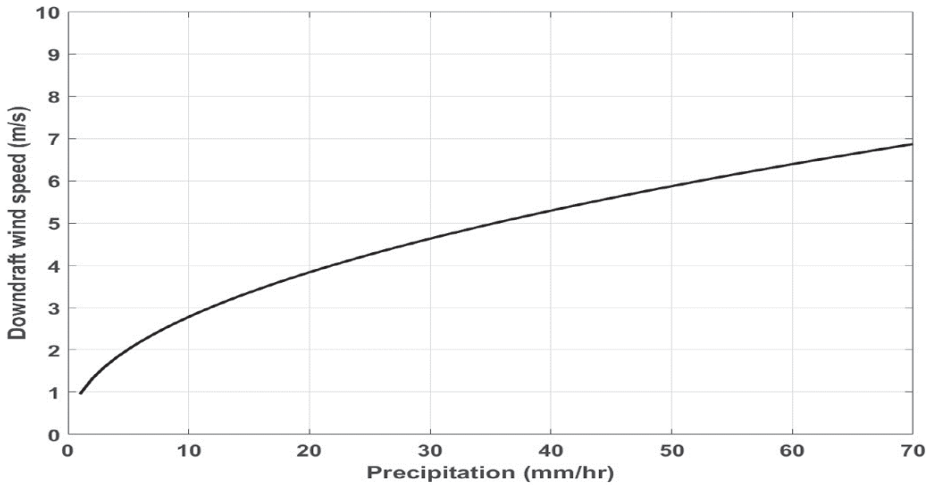


Figure 4.7: Downdraft windspeed model.

It can be observed from this model that downdraft winds increase monotonically with rain rate, as expected. This model provides an accounting for the excess wind generated due to rain, and thus for the change in radar cross-section due to rain-induced downdrafts. In later sections we evaluate how this physically based model compares to CYGNSS radar cross-section measurements.

CYGNSS data products	Description
Level 1	Normalized Bistatic Radar Cross-Section and Leading Edge Slope at the specular point with 25 km baseline resolution
Level 2	Surface wind speed in m/s at the specular point with 25 km baseline resolution
Level 3	Gridded surface wind speed for 0.2 x 0.2 (lat, lon) and temporal resolution of 60 min.

Table 4.1: Description of pertinent CYGNSS data products.

4.2 Observations

4.2.1 Data description

We consider various statistical analyses of CYGNSS data to identify and characterize the rain signature in the measurements. The data used are the $v2.1$ release of CYGNSS Level 1 σ_0 measurements and Level 3 wind retrieval from all 8 CYGNSS observatories over a period of 200 days between DOY 77 and DOY 276 of 2017. The data are matched to the European Numerical Weather Prediction Model, ECMWF 10 m referenced re-analysis ocean surface wind speeds at the specular point for ground truth wind speed. ECMWF wind speed information is matched to CYGNSS specular point locations through bilinear interpolation in space and linear interpolation in time (*Ruf and Balasubramaniam*, 2018).

The data source used for precipitation information is the Integrated Multi-satellitE Retrieval for GPM (IMERG) $V05B$ half hourly final run gridded data product with a spatial resolution of (0.1×0.1 deg) and temporal resolution of 30 mins (*Huffman et al.*, 2015). The pertinent CYGNSS data products are provided in Table.4.1 (*Ruf et al.*, 2016). The GPM measurements are re-gridded to match the Level 3 CYGNSS gridded

data product which has a spatial resolution of 0.2×0.2 deg, a temporal resolution of 60 min, and are collocated to the nearest CYGNSS specular point DDM in space and time. The precipitation information is also matched to the Level 3 CYGNSS gridded wind speed product to observe the global probability density function (pdf) of CYGNSS winds under low and high precipitation conditions.

The CYGNSS Level 1 σ_0 data are filtered by several quality measures for this analysis. Only data with a very high antenna gain ($> 10dB$) are used. The overall quality flag of the Level 1 data is set to best quality with a filter to exclude GPS block type IIF measurements to ensure best quality CYGNSS measurements. The sample population for this analysis contains approximately 10^8 measurements.

4.2.2 Controlled CYGNSS-GPM dataset

In addition to the large CYGNSS-GPM dataset described above, a second control dataset containing 110 individual CYGNSS tracks was generated to test and understand the behavior of σ_0 in the presence of rain. In addition to being passed through the aforementioned quality tests, these CYGNSS tracks are subjected to additional control over various surface properties and satellite geometries (all the tracks are provided in Appendix B). Each CYGNSS track is an overpass by a single CYGNSS observatory over a contiguous strong rain event followed by a rain-free portion. Here strong rain events are identified as events with peak rain rates > 30 mm/hr. This level of rain rate is chosen to correspond to situations where the attenuation effect by rain begins to become significant. It is important to have both strong rain and no-rain conditions present in the same track in order to study variations in σ_0 under similar measurement conditions.

The control parameters for the surface properties are the Significant Wave Height (SWH) and wind conditions along the track between the peak-rain and no-rain re-

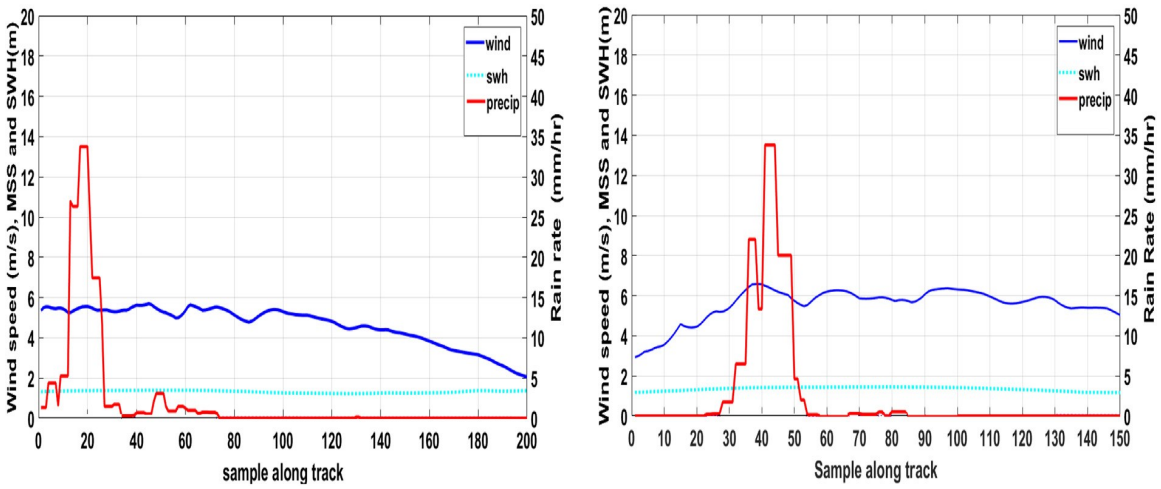


Figure 4.8: Sample CYGNSS tracks in the controlled dataset. Along track precipitation in red and wind speed in blue and SWH in cyan.

gions. SWH controls for the presence of swell effects. For SWH, WaveWatch III IFREMER SWH information is used and, for wind speed, ECMWF $10m$ surface wind speed information is used. The control parameters for satellite geometry are the receiver antenna gain and the incidence angle at the specular reflection point. For each member of the control dataset, all 4 of these control parameters are matched between the regions of strong-rain event and no-rain event. The no-rain event is a measurement sample on a given track where the precipitation is zero and peak rain event is the measurement along the same track but at the location of maximum precipitation. Thus the pairs are matched on the same track between the strongest rain region and the nearest no-rain region, resulting in a total of 110 pairs of matchups. Fig.4.8 depicts some of the tracks in the dataset. Note that the wind speed is fairly uniform across the entire track and the track contains both high precipitation and precipitation-free regions.

The control parameters in this dataset are shown in Fig.4.9. The four control parameters are compared at the sample of peak precipitation and zero precipitation sample

along the track. Note that all 4 parameters lie close to the 1 : 1 line, suggesting similar surface conditions and measurement geometry between the rain and no-rain regions. This reduces other potential sources of variation in the σ_0 and suggests that the observed differences in measured σ_0 between rain and no-rain conditions are most likely due to one or more of the 3 possible effects of rain.

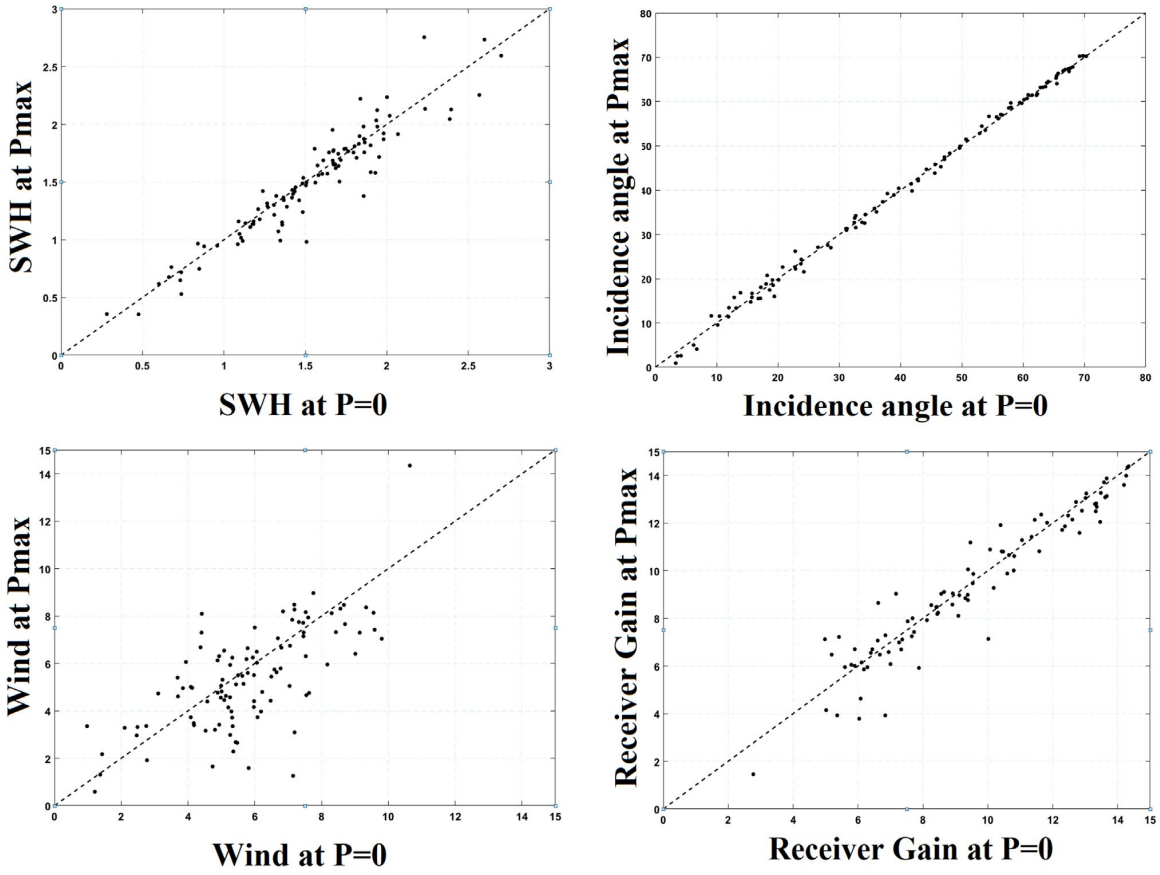


Figure 4.9: Control parameters for controlled CYGNSS-GPM dataset.

4.2.3 Rain signature in CYGNSS wind retrieval

To extract rain signatures in the CYGNSS-GPM dataset, the pdf of all CYGNSS Level 3 wind retrievals globally is constructed for 2 different rain rate conditions. Fig.4.10 shows the pdf for samples with a rain rate of less than $5mm/hr$ and for samples with a rain rate of greater than $10mm/hr$. We further randomly divide the

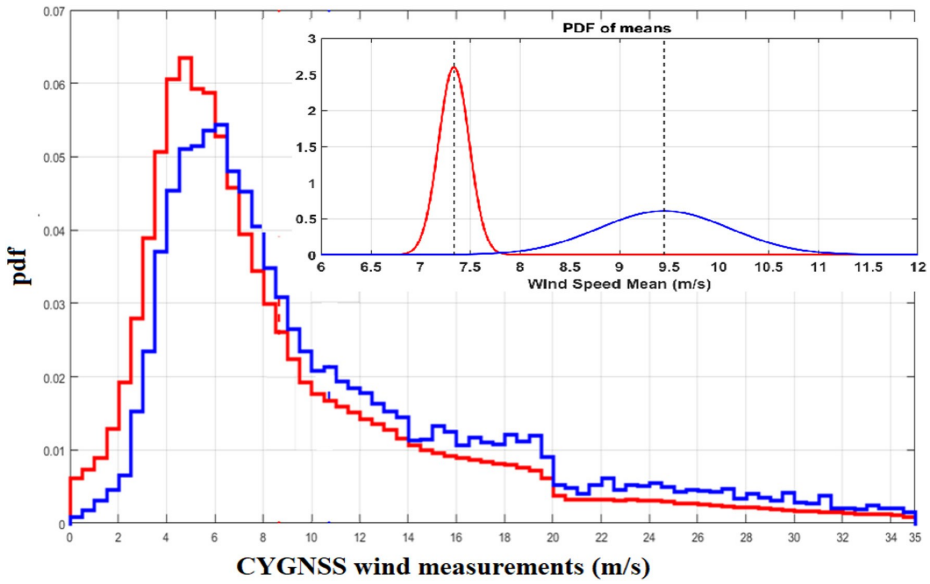


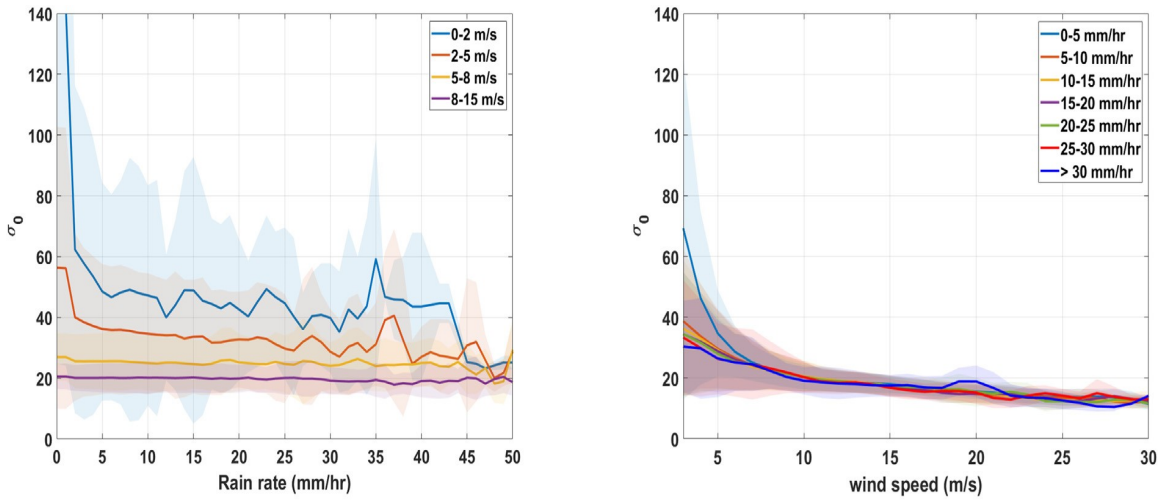
Figure 4.10: PDF of CYGNSS wind retrievals for rain rates $< 5mm/hr$ (red) and $> 10mm/hr$ (blue). Top right: PDF of means of wind retrieval using 10 subsets of data for rain rates $< 5mm/hr$ (red) and $> 10mm/hr$ (blue).

sample population into 10 independent partitions of the dataset, with approximately 10^7 samples in each partition, and consider the distribution of their mean values. A significant difference in the mean of the distributions is evident between the two rain rate conditions. For samples with $< 5mm/hr$ rain rate, the mean is $\sim 7.4m/s$, which is consistent with global average wind speed. For the samples with $> 10mm/hr$ rain rates, the mean is $\sim 9.5m/s$, which indicates a significant shift in the measurements in the presence of rain.

4.2.4 Rain signature in σ_0 measurements

The increase observed in retrieved wind speed in the presence of rain might be due to actual increases in the wind that are associated with the rain or to changes in the surface roughness caused by rain that are misinterpreted as an increase in wind speed. To observe the impact of rain on the measurements, variations in σ_0 are examined.

Fig.4.11 shows scatter plots from the complete matchup population amongst the three



(a) (b)

Figure 4.11: (a) Variation in σ_0 with respect to rain for different ranges of wind speed. (b)Variation in σ_0 with respect to wind speed for different ranges of rain rate.

variables σ_0, wind speed and rain rate, with the abscissa being rain rate (left) or wind speed (right), the ordinate being σ_0 in each case, and the third variable sorted by color. In the right plot, a clear reduction in the magnitude of σ_0 is observed with increase in rain rates. Recall that attenuation of the signal due to rain is negligible below rain rates of $\sim 30mm/hr$. This suggests that one or both of the other two phenomena, surface roughening due to rain and induced winds due to rain, are also associated with a decrease in σ_0. This behavior is examined in detail in the later sections. Note that in both the plots, the dependence of σ_0 on rain rate becomes negligible above the range of $8-15$ m/s. This suggests that the surface roughness due to wind dominates the effects of rain at wind speeds beyond the upper limit of $15m/s$. In the left plot, we observe the variation of σ_0 with rain rate for narrow wind speed bins. We can observe that σ_0 has a strong dependence on rain rate for wind speeds up to 5 m/s, a weaker dependence is visible up to 15 m/s, and above that, no dependence is observed.

4.2.5 Rain signature in direct DDM measurements

A double difference approach is taken to analyze the impact of rain on the CYGNSS DDMs. DDMs are modeled using the mission's End-to-End Simulator, which generates model DDMs for a given specified measurement geometry and an assumed wind speed (*OBrien et al.*, 2014). In this case, the wind speed used is the ECMWF matchup. The simulator assumes rain-free conditions. The difference between the observed and simulated DDMs is the first of two differences. The second difference is performed by considering a DDM at the point of maximum precipitation and a nearby one at zero precipitation along the same track. The double difference examines the difference between the $(obs - sim)$ DDM with rain and without rain. The first difference accounts for possible offsets in the simulator and the second difference highlights the effects of rain. Note that the cases considered here feature only very small variations in either wind speed or measurement geometry along the track, in order to minimize the effects of all other surface and geometric parameters.

An example track is shown in Fig.4.12. The image on the top left is a land mask with overlays showing the IMERG precipitation product by the color scale and the CYGNSS track in black. The red-x marks the CYGNSS sample at maximum precipitation and the green-x indicates the CYGNSS sample at zero precipitation. The double difference DDM method is performed at the 2 sample locations. The plot on the top right is the wind speed and precipitation information along the track. Note the wind speed drop at peak precipitation. We hypothesize that when the rain drops strike the surface, they suppress surface roughening associated with the wind. But the falling rain also pushes down the air column, creating a downdraft which then flows radially outwards, creating higher winds outside the rain region. The bottom left image is the double difference DDM between peak-rain and no-rain events and the bottom right image is the double difference DDM for two adjacent no-rain samples.

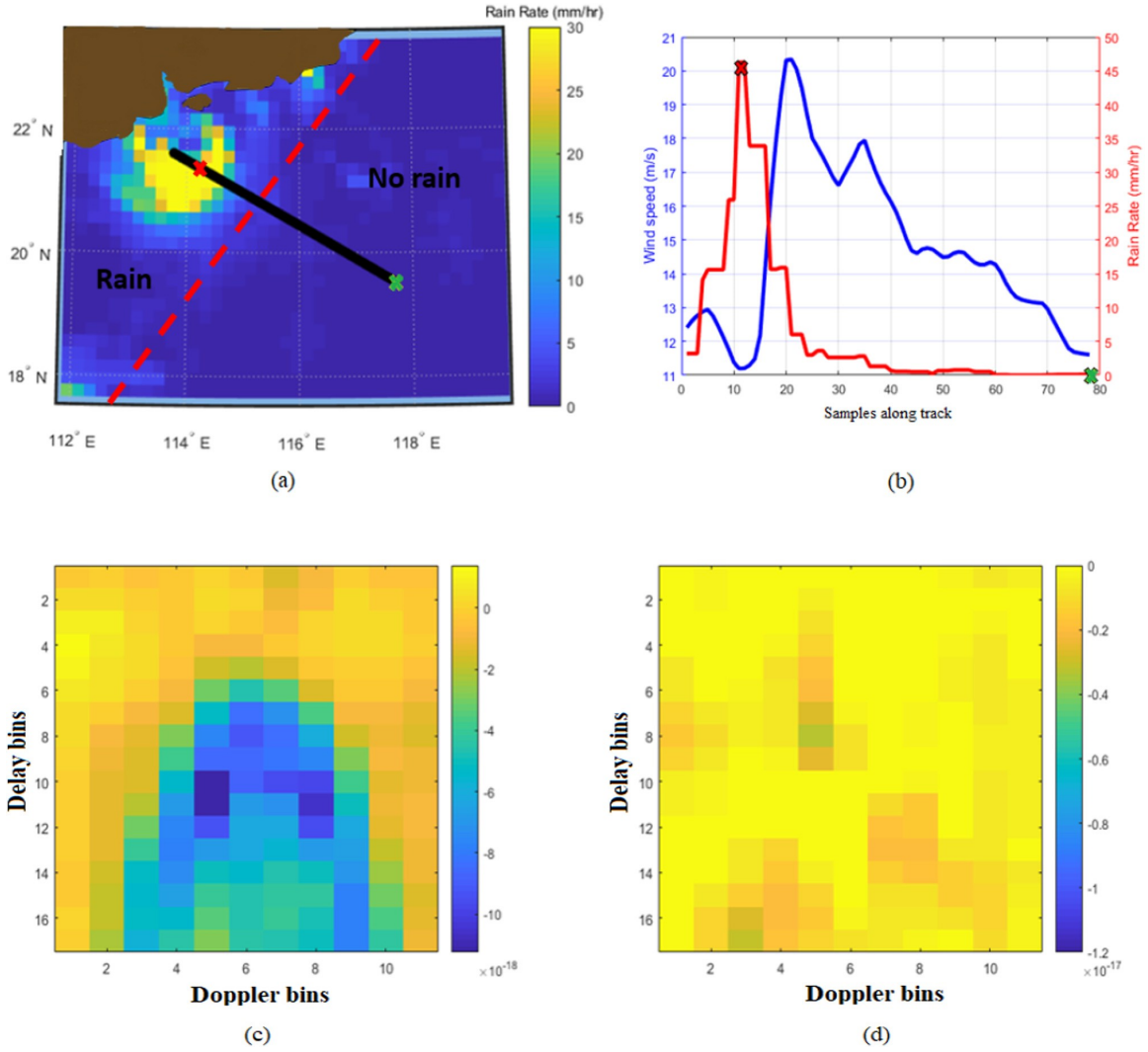


Figure 4.12: Double difference DDM method for observing rain signature.(a) CYGNSS track overpass of a large convective rain cell (DOY 235, FM3, PRN 7) (b) Groundtruth ECMWF wind speed in m/s and GPM IMERG precipitation in mm/hr along the track, (c) Double difference DDM of power in watts between a peak rain and a zero precipitation region, (d) Double difference DDM of power in watts between 2 different zero precipitation regions.

The DDMs shown are the DDM of power difference and their x-axis is the 11 doppler bins and the y-axis is the 17 delay bins centered on the specular point.

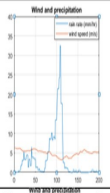
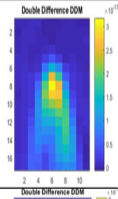
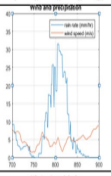
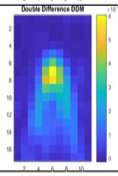
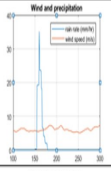
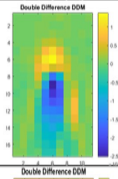
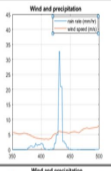
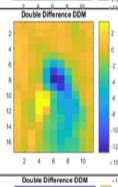
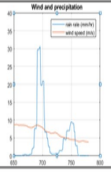
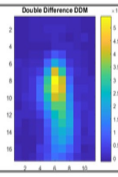
Precipitation	Wind speed	CYGNSS track	Double Difference DDM
Max P = 32.47 mm/hr Min P = 0 mm/hr	W at max P=3.39 m/s W at min P= 4.17 m/s		
Max P = 31.5 mm/hr Min P = 0 mm/hr	W at max P= 5.13 m/s W at min P= 5.66 m/s		
Max P = 35.13 mm/hr Min P = 0 mm/hr	W at max P= 5.59 m/s W at min P= 5. 78 m/s		
Max P = 32.63 mm/hr Min P = 0 mm/hr	W at max P= 6.0285 m/s W at min P= 6.0643 m/s		
Max P = 30.5 mm/hr Min P = 0 mm/hr	W at max P= 8.30 m/s W at min P= 8.56 m/s		

Figure 4.13: Double difference DDM of power for 5 different CYGNSS tracks with near uniform wind along the track. Column 1: Precipitation information (P) at the peak and no precipitation sample points, Column 2: Wind information (W) at the peak and no precipitation sample points, Column 3: CYGNSS track with wind and precipitation information along the track and Column 4: Double difference DDM of power between the peak and no precipitation regions along the track.

The double difference DDM between peak-rain and no-rain events shows most of the difference in scattered power bins to be localized in the horseshoe region, with the highest difference near the specular point, suggesting that the rain information is concentrated near the specular point. The magnitude of this difference is comparable to the magnitude of the power received, which suggests that the effect of rain is not

negligible. We see in the double difference between 2 no-rain samples (bottom right image) that the double-difference DDM has values close to zero everywhere. Fig.4.13 shows the double difference DDM of power for 5 different CYGNSS tracks that has near uniform wind speed along the track. In all these cases, all the surface properties and geometries are very similar between peak rain and no rain region. This validates the assumption that the difference observed in the peak-rain vs no-rain case is caused by the presence of rain and most of the information is contained in the horse shoe region of the DDM.

4.3 Results

In the following subsections, we compare the CYGNSS observations to the proposed rain effect model to understand how well the model represents the underlying rain effects on ocean surface scattering of GPS signals.

4.3.1 Path integrated attenuation through rain

The integrated attenuation along the signal path is modelled in Eqn.4.7. We apply a loss correction to the measured value of σ_0 to account for the drop in received power. The corrected σ_0 is denoted by $\sigma_0^{obs}|_R$. A perturbation model for the surface σ_0 can be written as:

$$\sigma_0^{obs}|_R = \sigma_0^{obs}|_{R=0} - \delta\sigma_0(R), \tag{4.11}$$

where $\sigma_0^{obs}|_{R=0}$ is the σ_0 (in linear scale) if there was no rain and $\delta\sigma_0(R)$ is the perturbation in σ_0 due to effects of rain other than attenuation which is assumed to be a function of rain rate.

The perturbation term is modeled empirically as follows: First, the CYGNSS-GPM matchups are obtained for a narrow range of wind speeds for the rain-free condition. Next, the attenuation corrected $\sigma_0^{obs}|_R$ is calculated based on the proposed model in Eqn.4.11 for different rain-rates. An empirical model for perturbation in σ_0 is then given by:

$$\delta\sigma_0(R) =< \sigma_0^{obs}|_{R=0} > + < \sigma_0^{obs}|_R >, \tag{4.12}$$

where <> denotes the expectation over a large population of CYGNSS Level 1 σ_0 values collected over 200 days in 2017. Thus the rain induced surface roughening and increased surface winds due to rain are modelled as a difference in expectation of the σ_0 distribution for rain-free condition and the distribution of σ_0 at different rain rates.

The sample population for different rain rates is developed by filtering the CYGNSS-GPM matchup measurements by GPM IMERG rain rates. Fig.4.14 shows the sample population density of CYGNSS samples at different rain rates.

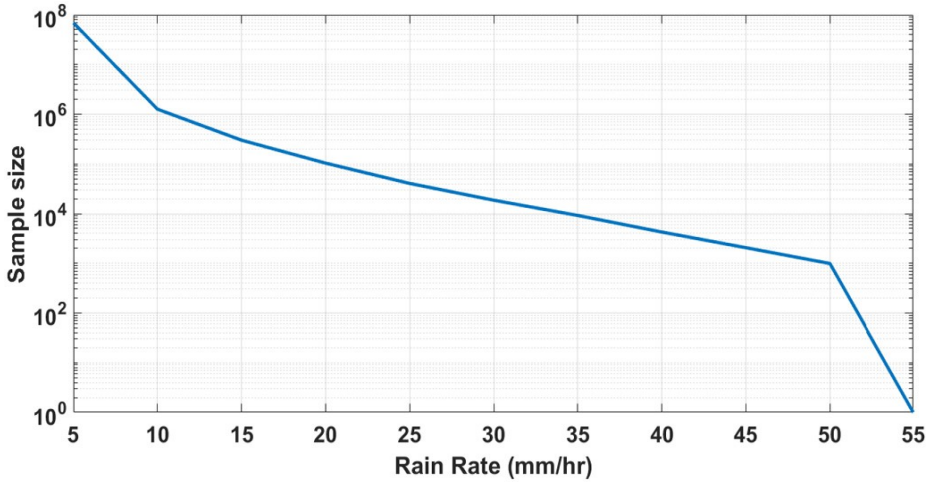


Figure 4.14: CYGNSS samples for different rain rate bins.

To maintain the quality of estimation we fix the threshold of number of samples per bin to be 10,000. The model developed for the σ_0 perturbation is of the form:

$$\delta\sigma_0(R) = a * R^b + c, \tag{4.13}$$

where, a,b and c are empirically derived coefficients and R is the rain rate in mm/hr. The model developed for the perturbation in σ_0 is shown in Fig.4.15.

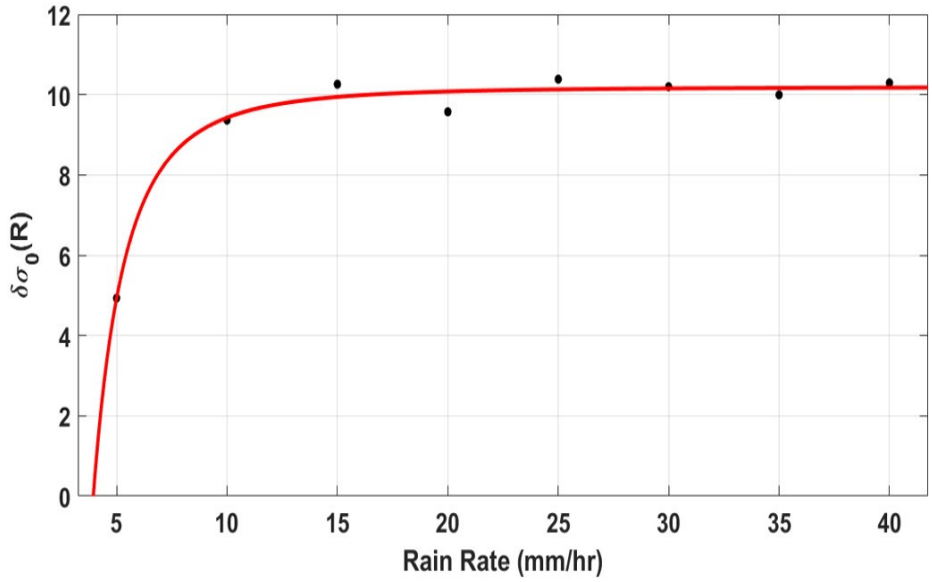


Figure 4.15: Model for rain induced perturbation.

In the figure, note that the perturbation rises sharply for changes in rain rate from 0 to 15 mm/hr and saturates at higher rain rates. This result is consistent with the observation in Fig.4.11 (left), which shows that changes in the observed σ_0 are largest at rain rates below 15 mm/hr and reduce at higher rain rates.

4.3.2 Ring wave impact on roughness spectrum

A composite surface roughness spectrum is constructed using the Elfouhaily omni-directional wind-wave spectrum together with the dominant ring-wave spectrum due to rain described above. The Mean Square Slope (MSS) at the surface is calculated from the composite spectrum by integration, as described by (*Elfouhaily et al.*, 1997):

$$mss = \int_{k_{min}}^{k_{max}} k^2 S(k) dk, \tag{4.14}$$

where k is the wavenumber, k_{min} and k_{max} represent the limits of integration given by $k_{min} = \frac{g}{10}(\frac{\Omega}{u_{10}})^2$ and $k_{max} = \frac{2\pi}{3\lambda} cos(\theta)$, g is the acceleration due to gravity in m/s^2, Ω is the inverse wave age, u_{10} is the 10m surface wind speed, λ is the wavelength of the transmit signal, θ is the incidence angle of the signal at the surface, and $S(k)$ is the composite roughness spectrum.

The expression for the upper limit of the integration, k_{max}, is determined using the geometric optics limit and hence is only an approximate value. It is in the range of $10 - 12$ rad/m at L-Band for typical angles of incidence. We consider here the possibility that the geometric optics approximation is not appropriate at sufficiently low wind speeds and levels of surface roughness, and that a different (higher) value for k_{max} may be more appropriate. The possibility is addressed empirically by comparing observed values of MSS to those predicted by the integral expression given by Eq.4.14, while varying the assumed value of k_{max}.

To illustrate our approach, we consider an example surface with wind speed of 3 m/s and then we evaluate the MSS of this surface with no rain present and when additionally roughened by a rain rate $R = 100$ mm/hr. Fig.4.16 shows the composite spectrum for a rain rate of 100 mm/hr with the integration limits shown in red and blue. Note that the ring wave portion of the spectrum is dominant only for $k > 50$ rad/m, but the low pass wavenumber (k_{max}) is ~ 12 rad/m. The resulting MSS from Eq.4.14. is 0.0111807484 for no rain (R= 0 mm/hr) and 0.01118119895 for a rain rate of 100 mm/hr, or an increase of 0.004%. This suggests that the increase in MSS would be negligible at very high rain rates and low wind speeds, a model prediction which is inconsistent with the behavior of the CYGNSS measurements. This motivates our consideration that the geometric optics limit for k_{max} may be underestimating the

true lowpass wavenumber.

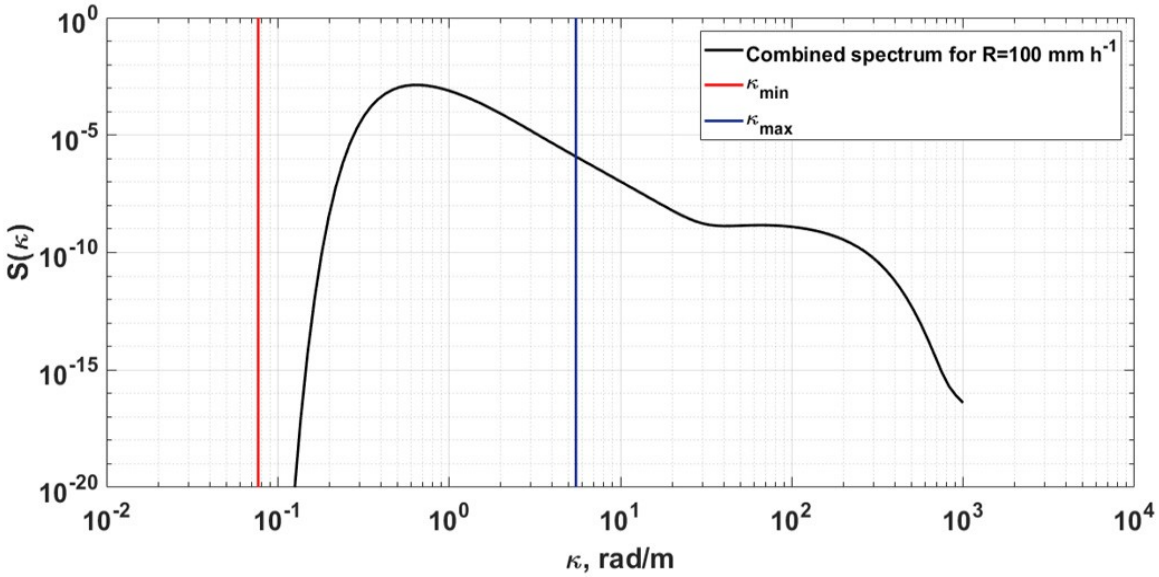


Figure 4.16: Sample ocean surface roughness spectrum for wind speed =3 m/s and rain rate =100mm/hr. A rain-free roughness spectrum would decrease monotonically above $k \sim 0.6$ rad/m. The impact of rain causes a noticeable increase in roughness at $k = 50 - 500$ rad/m.

Our empirical examination of k_{max} uses the controlled dataset of 110 CYGNSS tracks described above. An optimal lowpass wavenumber is estimated by minimizing the absolute difference between the theoretical model for MSS and the MSS observed at the sample location of peak precipitation in the dataset. The optimal wavenumber is given by:

$$k_{opt} = \underset{k_{max}}{min} |mss_{model} - mss_{obs}| \tag{4.15}$$

An example is shown in Fig.4.17 of the absolute difference as a function of k_{max}. A distinct, global minimum in the absolute difference is evident at k_{opt}. This behavior of the absolute difference vs. k_{max} is consistent across all 110 tracks. The optimal lowpass wavenumber tends to be slightly higher than the limit defined by the geometric optics approximation. k_{opt} is derived in this way for all of the 110 tracks and

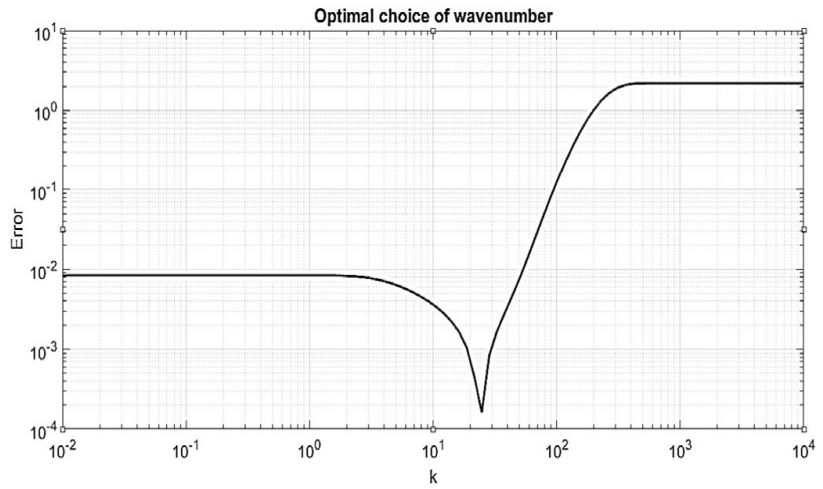


Figure 4.17: Sample error minimisation along track for optimal k.

is compared in Fig.4.18 with the theoretical value for k_{max} in the geometric optics limit. Fig.4.18a shows the theoretical value of k_{max} for each of the tracks. Averaging across all 110 cases, the mean lowpass wavenumber from geometric optics is found to be 7.9486. Fig.4.18b shows the value of k_{opt} for each of the 110 tracks.

In Fig.4.18b, the optimal wavenumber is seen to vary as a function of wind speed, with its highest values close to 50 at very low wind speeds, and its value decreasing with increasing wind speed until if levels off at the mean geometric optics limit above $\sim$ 5m/s. A parametric model for k_{opt} as a function of wind speed, fit by least squares minimization to the 110 empirical values, is also shown in Fig.4.18b. It is given by:

$$k_{maxnew} = k_0 + k_s * exp(\frac{-u}{u_s}), \tag{4.16}$$

where k_{maxnew} is our new lowpass wavenumber, k_0 is the geometric-optics lowpass wavenumber, k_s=104.378 and u_s=0.7116 are the empirical coefficients from the least squares minimization and u is the wind speed. We observe here that the new model for maximum wavenumber converges to the geometric optics limit above 5 m/s and retains its cosine dependence on incidence angle. However, this model suggests that

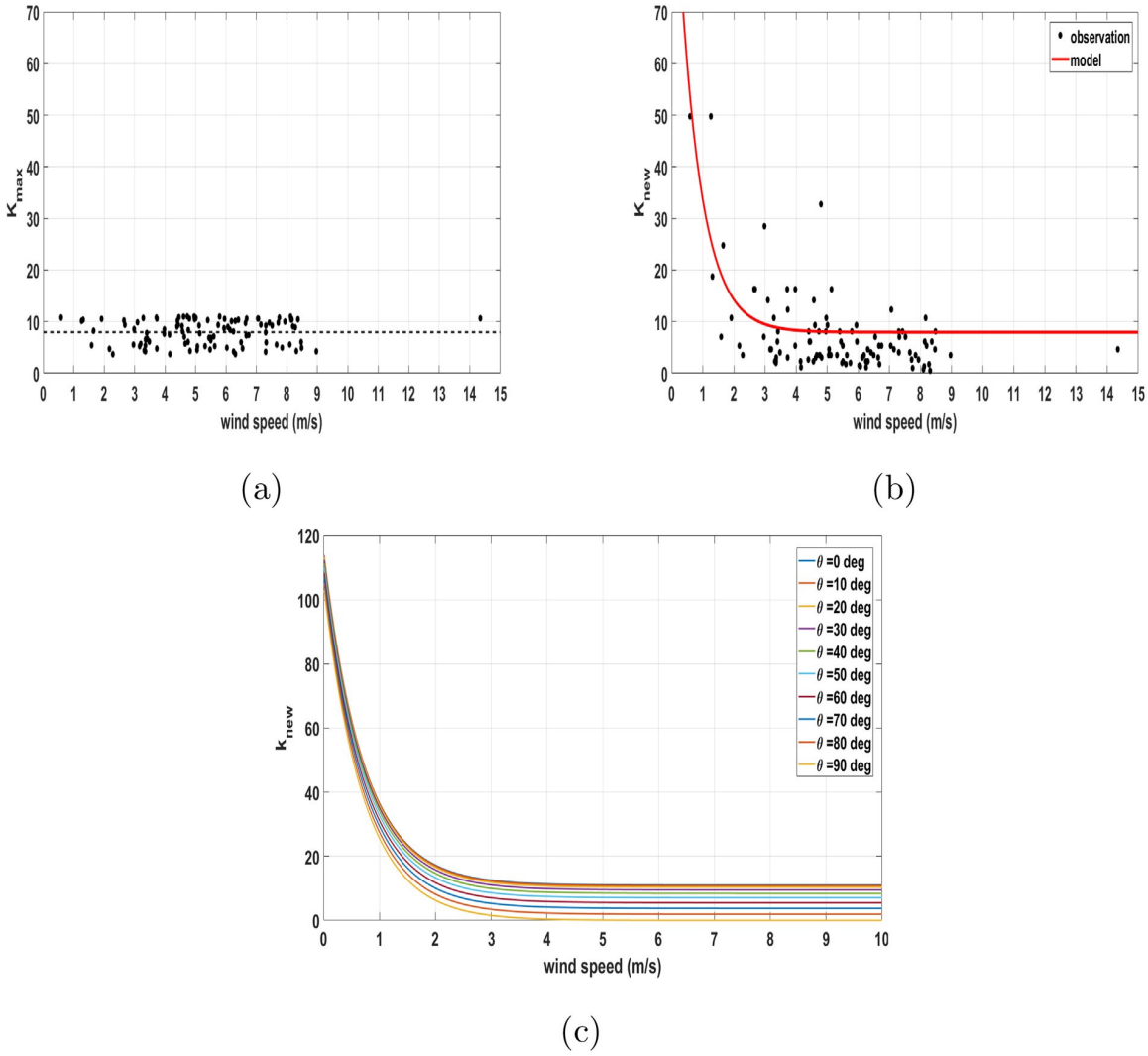


Figure 4.18: (a) theoretical maximum wavenumber(derived from geometric optics approximation), (b) optimal (empirically derived) maximum wavenumber of the surface roughness spectrum that significantly affects the mean squared slope and bistatic radar scattering cross section,(c) A new maximum wavenumber model with wind speed and incidence angle dependence.

the geometric optics limit breaks down at very low wind speeds, and it is under those conditions that the effects of surface roughening by rain become significant and measurable due to an increase in the upper bound of wavenumbers that affect the MSS. These results are consistent with the simulations conducted (*Voronovich and Zavorotny*, 2017) which show that the strong diffuse scattering approach to solving the bistatic radar equation inaccurately predicts the DDM for wind speeds $< 5m/s$.

4.3.3 Comparison of downdraft winds with measurements

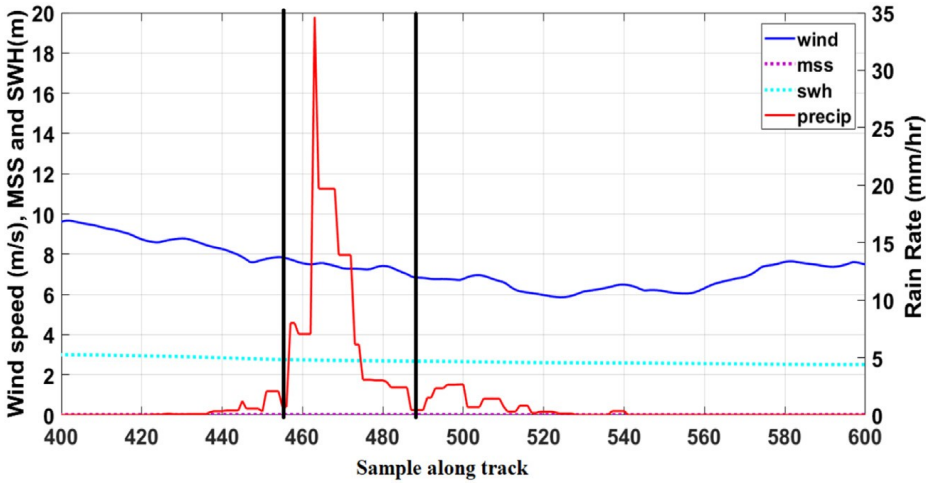


Figure 4.19: Sample CYGNSS track measurement.

A model for the downdraft wind generated by rain was described above. We now compare this model to observations by considering the surface wind speed in a rain-free area immediately adjacent to a rain event. One example CYGNSS track is shown in Fig.4.19 in which the regions of measurements used to calculate the downdraft wind speed are highlighted. The x-axis represents the sample number along the track and the y-axis represents the parameters: wind speed, MSS, SWH and rain rate. The two black vertical lines show the region of rain-free measurement of wind speed on either side of the rain event, where precipitation is close to or identically zero. The observation of downdraft winds is defined as the average of wind measurements at

the closest regions around the peak precipitation region where the rain rate is close to zero. This method of calculating the radial downdraft winds is equivalent to measuring radially outward-moving winds at 2 different locations along the track. These measurements are compared to the downdraft model-based winds which are evaluated by taking the rain rate as an input and solving the vertical momentum transfer equation to derive the radial downdraft winds as shown in Fig.4.5. This procedure is followed for all of the 110 tracks in the controlled CYGNSS dataset and the wind measurements are compared to the downdraft wind model shown in Fig.4.7. A typical convective rain event is usually much smaller than the GPM IMERG grid cell size and the rain rate value provided by its level 3 product per grid is likely not an accurate representation of the real situation. This is because the rain rate is not uniform within the grid cell. If a rain event occupies only 1/4th of the grid size it would lead to a lesser average rain rate in the grid than if the rain event occupies half or more in the grid cell. The IMERG rain product has a similar resolution as the CYGNSS σ_0 so this effect will be most significant if the relationship between the downdraft velocity and rain rate is non-linear, which is the case at lower rain rates as shown in Fig.4.20.

However the downdraft velocity at higher rain rates has a linear dependence on the rain rate (which is our region of interest) and hence the averaging effect of having non-uniform rain rates in the sub-grid space will have less of an impact. The comparison of model and measurements is shown in Fig.4.20. All the 110 tracks in the controlled CYGNSS dataset have a peak precipitation > 30 mm/hr. We observe that the model wind speed passes through the cluster of measurement points. There is considerable scatter between the model wind and the observations, which could be attributed to the simplified and approximate method used to calculate the downdraft wind from the rain rate. However, a general trend of increasing rain rate being accompanied by higher downdraft wind speed can be observed.

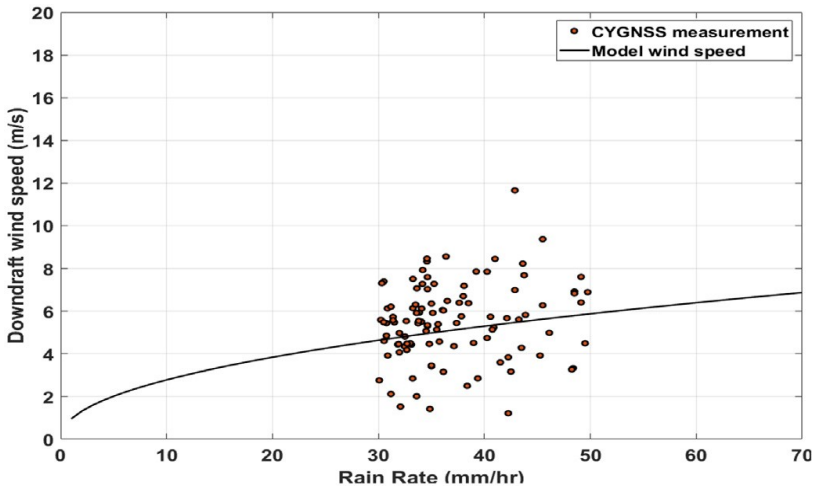


Figure 4.20: Measurements over controlled dataset and model.

4.4 Discussion and conclusions

GNSS-R missions work on the principle of L-Band bistatic scatterometry. For accurate wind retrievals, it is important to understand the dominant rain effects at L-Band and how the σ_0 measurements are affected by them. In this work we have identified 3 major effects of rain and have proposed a model for each. The observations used are from the CYGNSS constellation mission which is currently operational and has been collecting global ocean surface information since 2016.

The high-quality observation dataset encompasses a range of rain events with sufficient samples ($> 10,000$) up to a rain rate of 50 mm/hr. Using the CYGNSS measurements, a detailed empirical analysis was made to differentiate the effects caused by rain. The analysis suggests that: 1) the rain produces a reduction in magnitude of the measured σ_0 with increase in rain rate; and 2) the impact of rain can only be observed for wind speeds up to ~ 15 m/s.

The rain effect model proposed in this work has a 3-fold structure-path attenuation, modified surface roughness and downdraft winds. From the path attenuation model, we observe that the attenuation at L-band due to raindrops is not significant, with transmissivity above 96% for rain rates below 30 mm/hr. After removing the effect

of path attenuation, the empirical model for perturbations in σ_0 sharply rises for a change in rain rate from 0 to 15 mm/hr and saturates at higher rain rates. The model also shows that the perturbation in σ_0 is always negative; thus, rain produces a drop in the magnitude of σ_0.

The dynamic range of wind speeds is split into 2 regimes based on the electromagnetic scattering nature of the surface. The near specular regime is the region of wind speed $<$ 5 m/s and above that is the diffuse scattering regime. In the near specular regime, we observe that the modified rain-wind roughness spectrum with a geometric optics approximation of the lowpass wavenumber integration limit does not adequately account for the observed effects of rain. We empirically estimate the lowpass wavenumber limit and determine that it is a function of wind speed. The limit of integration increases at very low wind speeds and saturates to the theoretical geometric optics limit for larger wind speeds ($>=$ 5 m/s). However, the net effect on a GNSS-R measurement at such low wind speeds is the combined effect of variation in the cutoff wave number with wind speed as well as the coherent component in the forward scattering direction as shown by Voronovich and Zavorotny (2017).

The last of the 3 effects is local downdraft wind due to rain. Of the 3 different effects of rain studied in this work, we have found that the impact of attenuation is negligible and hence the total impact of rain on the measurements is dominated by the splash effect and the downdraft effect. Of these two, the splash effect is restricted to very low wind speeds ($<$ 5 m/s) but the overall impact of rain can be observed up to surface wind speeds of 15 m/s. This suggests that the dominant effect on the moderate wind speed regime of 5-15 m/s is due to the downdraft effect. This model is compared to the simplified method of measuring the downdraft winds from the controlled dataset of 110 tracks. We find that the observation and the model are generally consistent in the estimation of downdraft winds and the model suggests that a maximum downdraft wind speed of 7 m/s should be present at very high rain rates.

CHAPTER V

Azimuthal Dependence of GNSS-R Cross-Section Inside Hurricanes

The measurement of hurricane wind fields has a long history, ranging from airborne measurements (*Jones et al.*, 1981; *Uhlhorn et al.*, 2007; *Wright et al.*, 2001) to spaceborne observations made by microwave radiometers and radars (*Ebuchi et al.*, 2002; *Figa-Saldaña et al.*, 2002; *Gaiser et al.*, 2004). The key challenges for mapping the complex hurricane wind fields are the need for adequate spatial and temporal sampling of such fast evolving phenomena and the ability to penetrate through strong rain bands to measure the surface winds. Global Navigation Satellite SystemReflectometry (GNSS-R) is a relatively new field of remote sensing that uses the existing GNSS signals to study the surface. It greatly improves the sampling and revisit capability by utilizing the existing GPS transmitter constellation, and its L-band measurements are less affected by the heavy precipitation in the rain bands.

A GNSS-R radar measures the scattering cross-section of the surface around the region of specular reflection. The reflected GPS signal observed by a GNSS-R receiver is mapped into delaydoppler space for different time delays, and doppler shifts are observed. This forms the DelayDoppler Map (DDM) of the surface (*Gleason et al.*, 2009).

It is important to note that the scattering crosssection measured by a GNSS-R re-

ceiver is directly related to the surface roughness rather than the surface wind itself. GNSS-R forward scatter is quasi specular incoherent scatter in most conditions. Thus an appropriate Mean Squared Slope (MSS) of the surface as sensed by GNSS-R measurements is an integration of the wave spectrum over a range of wavelengths ranging from several meters up to a few tens of cms. The surface MSS is inversely related to the measured normalized bistatic radar crosssection (σ_0). The Geophysical Model Function (GMF) maps this σ_0 to the ocean surface wind speed empirically to retrieve the near-surface wind speed from the measurements. Some studies have also estimated the MSS of the surface directly from wind speed measurements inside hurricanes using empirically derived relationships (*Gleason et al.*, 2018b). In a hurricane environment, with complex temporal and spatial distribution of wind and wave fields, it is a challenging task to accurately retrieve wind speed from GNSSR measurements. The scattering cross section depends on surface roughness scales spanning a wide range from small capillary waves to long gravity waves. In fully developed seas, with essentially infinite sea age and fetch length, the relative magnitude of the surface roughness at different scale sizes reaches an equilibrium state due to energy cascade and dissipation mechanisms. Inside a hurricane, however, the sea age and fetch length conditions can vary significantly with azimuthal location due to its rotational and translational motion. This can perturb the balance between the roughness at different scales and alter the measured scattering cross section. Despite its complex nature, several simulations (*Fan et al.*, 2009; *Moon et al.*, 2003; *Young*, 2017), directional buoy measurements (*Young*, 2006), and airborne missions (*Uhlhorn et al.*, 2007; *Walsh et al.*, 1985; *Wright et al.*, 2001) have, over the years, helped develop directional wave spectra for hurricanes. The directional wave spectra acquired from several hurricane reconnaissance missions suggest that local wind and wave directions vary sinusoidally with the azimuth angle referenced to the hurricane heading and have a weak radial dependence (*Hwang et al.*, 2017). These results suggest that

remote sensing techniques such as GNSS-R, which depend on surface wave scattering, should consider azimuthal wind-wave response functions for accurate modeling and subsequent wind retrieval.

A GNSS-R GMF describes the relationship between measured scattering crosssection and the 10m reference wind speed. Previous empirical GMFs for hurricane winds have been developed without allowing for possible dependence on azimuthal location within the storm (*Clarizia et al.*, 2014; *Ruf and Balasubramaniam*, 2018). As a result, actual azimuthal dependencies are essentially averaged out, and wind speed retrieval errors will be correlated with azimuth location. An improved, azimuthally dependent, empirical GMF is developed here to better account for the azimuthal variation of the wind and wave directions. A large data set of observations from the CYGNSS mission is used.

In this chapter, the impact of azimuthal variation on the measured scattering crosssection is assessed using CYGNSS data over 19 major hurricanes across different basins during 2017 and 2018. For this analysis, HWRF reanalysis hurricane winds are used as a reference. The remainder of this chapter is structured as follows: Section 5.1 gives a description of the data sets used and the observations from the CYGNSS-HWRF matchup analysis, Section 5.2 describes the empirical GMF developed as a function of azimuth angle; Section 5.3 assesses the performance of the proposed model, and Section 5.4 provides the conclusions of the study.

5.1 Theory and observations

For this analysis, the v2.1 release of CYGNSS Level 1 σ_0 measurements over 19 major hurricanes from 2017 and 2018 is used (*PO.DAAC*, 2018). The data are matched to Hurricane Weather Research and Forecasting (HWRF) reanalysis winds of the inner nest grid spacing of 2 km. The HWRF winds are resampled to CYGNSS resolution

West Pacic	East Pacic	Atlantic	Indian
Jebi	Aletta	Florence	Mekunu
Jelawat	Otis	Harvey	Titli
Mangkhut	Willa	Irma	
Maria		Jose	
Trami		Maria	
Walaka		Michael	
Yutu		Oscar	

Table 5.1: Hurricanes from different basins in the CYGNSS-HWRF matchup dataset.

and are empirically paired to CYGNSS σ_0 observations with a maximum temporal separation of 60 min and a maximum spatial separation of 0.25 deg latitude and longitude. The matchups are then translated into a storm-centric direction of motion-based coordinate system for the purpose of understanding the azimuthal variation of measurements relative to the storm heading. The Weather Research and Forecast (WRF) system for hurricane prediction (HWRF) is an operational model developed by the National Centers for Environmental Prediction (NCEP). HWRF provides three domains (one parent and two nested) and is based on the initial position of the storm and on the National Hurricane Center (NHC) forecast of the 72 hr storm position. The two nested domains move along the storm with a coverage of 24 deg x 24 deg and 7 deg x 7 deg for the middle nest and the inner nest, respectively (*Tallapragada et al.*, 2014). For our purposes, we use the inner nest gridding because it offers the finest resolution of about 0.015 deg (approximately 2 km). The CYGNSS level 1 σ_0 are also filtered by several quality measures for this analysis. Only observations with high antenna gain ($> 5dB$) and the overall quality flag set to best quality are used. This has allowed a total data set consisting of ∼187,000 observations in hurricanes by CYGNSS. Table 5.1 lists the different hurricanes contained in this data set.

The CYGNSS wind retrieval algorithm uses the two measured observables, namely, the normalized bistatic radar scattering crosssection (σ_0) and the slope of the leading edge of the radar return pulse scattered by the ocean surface (LES) (*Clarizia*

and Ruf, 2016b). With these observables, GMFs are empirically derived by pairing near coincident independent estimates of 10m referenced ocean surface wind. For a fully developed sea (FDS), which constitutes the majority of the measurements, the observables are matched to the ground truth reference which is the combination of European Centre for MediumRange Weather Forecasts (ECMWF) and Global Data Assimilation System (GDAS) reanalysis wind speed products. This results in a FDS GMF. For young sea conditions with limited fetch (YSLF), as observed in hurricanes, a YSLF GMF is generated using matchups with near coincident NOAA P3 hurricane hunter passes over the major Atlantic storms in 2017 (*Ruf and Balasubramaniam*, 2018).

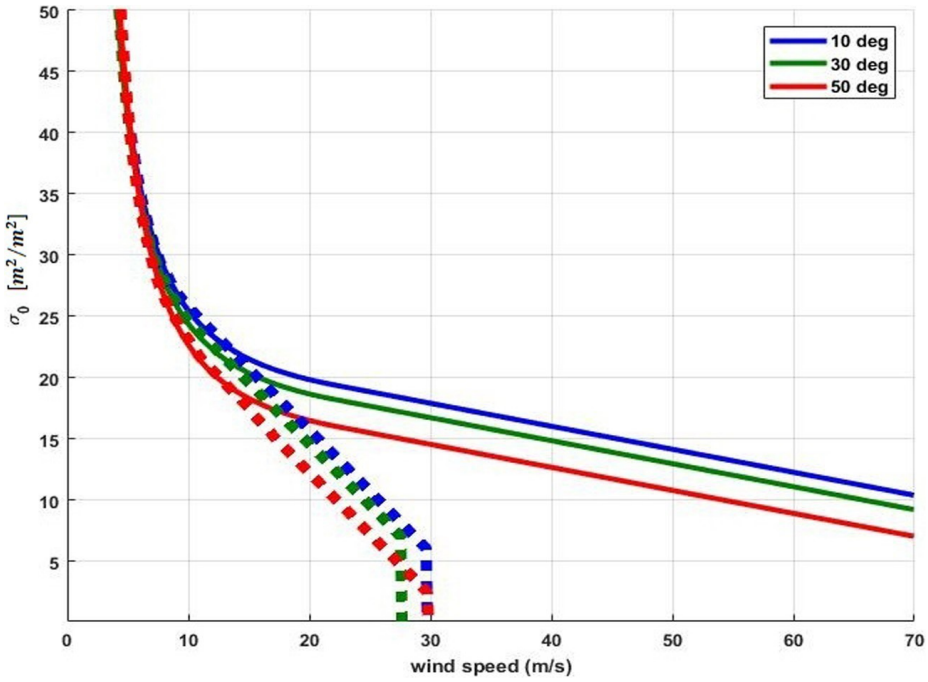


Figure 5.1: CYGNSS GMF for FDS winds shown by dotted lines and YSLF winds shown with solid lines. Incidence angles of 10, 30, and 50 deg are shown.

Examples of the FDS and YSLF GMFs are shown in Fig.5.1 for observations at incidence angles of 10, 30, and 50 deg. Above wind speeds of ∼15 m/s, the two GMFs diverge due to the underdeveloped state of seas near tropical cyclones, which tends to lower the roughness and increase the scattering cross-section. One important

feature to note is the difference in the slope of the two GMFs at higher wind speeds. The YSLF GMF at high wind speeds has a higher value and a shallower slope (lower $|d\sigma_0/du_{10}|$) for all incidence angles. In general, the high wind slope of the GMF can be used as a proxy for sea state development, with lower magnitudes being associated with younger seas.

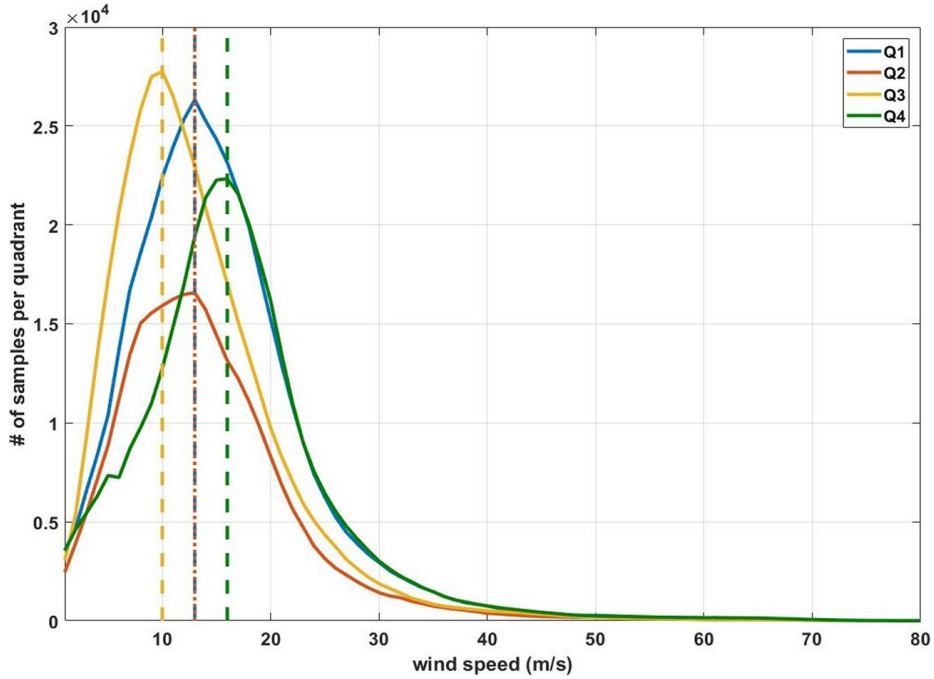


Figure 5.2: Distribution of winds in different quadrants in a storm relative to its heading.

Fig.5.2 shows the distribution of winds in different quadrants of a hurricane using the CYGNSS-HWRF matchup data set. The wind distribution inside the hurricane varies strongly in the azimuthal direction. In theory, the first quadrant is the generation region (shown in Fig.5.3a) and has the maximum energy. Also, Quadrant 1 has the largest wind speeds relative to the surface because the winds generated by the storm in this region are added to the storm motion. The wind generates a spectrum of waves with different group velocities. Waves that have their group velocity equal to the velocity of the forward motion of the storm remain in the intense wind region and receive maximum energy from the wind. Waves with group velocities greater than

the storm's forward motion velocity will outrun the storm and propagate ahead as swell waves and those with a lower group velocities than the storm will be outrun by the storm and be left behind (*Young*, 1999). The swell radiating out ahead of the storm will often interact with the local calm sea in phase quadrature, resulting in a confused sea condition. This is generally observed ahead of the storm and to the rear of the storm, resulting in a younger sea in each of those regions.

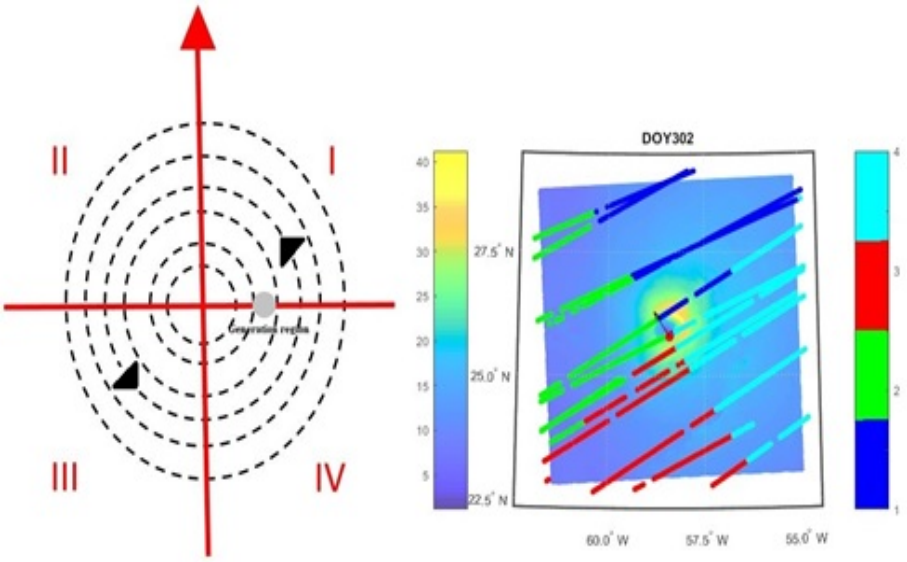


Figure 5.3: (a) Description of stormcentricdirection of motionbased coordinate system used in this work. (b) Sample CYGNSS tracks overlaid on HWRF wind field. HWRF wind intensity is shown by the colorbar on the left, and the quadrant of the CYGNSS sample is shown by the colorbar on the right.

The adequacy of a single slope for the GMF is assessed by determining the slope separately in each quadrant of a storm, relative to the storm heading. The storm quadrants are defined based on the Cartesian representation of quadrants with increasing azimuth angle in the anticlockwise direction (see Fig.5.3a). An example track of CYGNSS overlaid on HWRF, partitioned by storm quadrant (color bar on the right) is shown in Fig.5.3b, the storm heading is represented by the black arrow at the storm center (identified by the red circle).

For the purpose of analysis, the range of HWRF wind speeds from 20 to 70 m/s is divided into bins. The center of each wind speed bin is stepped in 1 m/s increments from 20 to 70 m/s. Within a bin, all corresponding CYGNSS scattering cross-section measurements are averaged together. The wind speeds are binned in this way to reduce biases in the estimation process due to variations in sample size at different wind speeds.

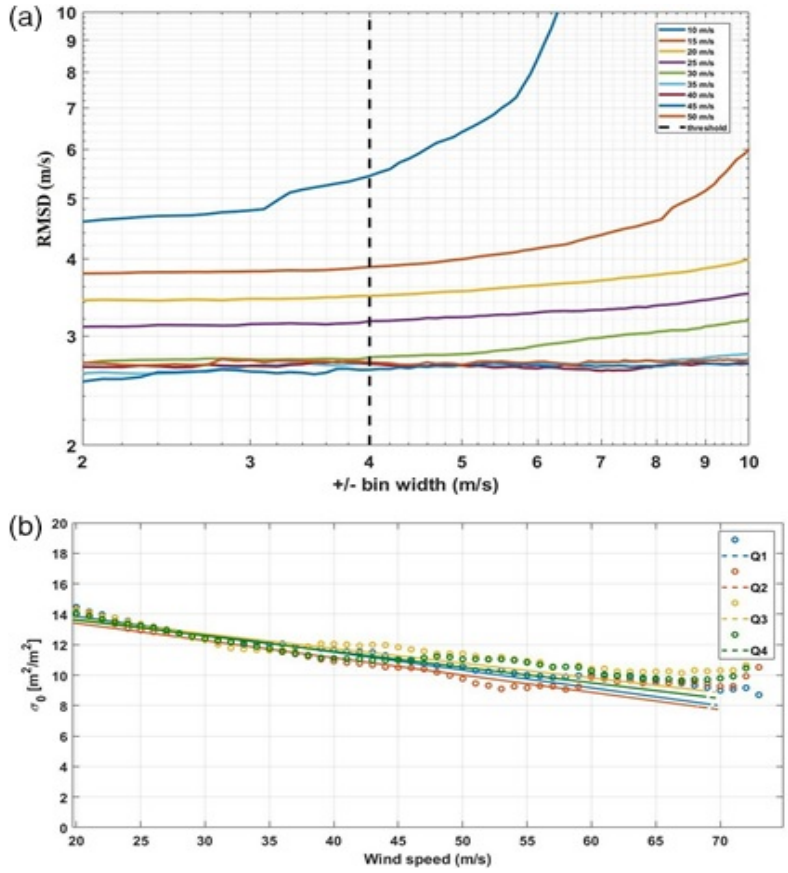


Figure 5.4: (a) Choice of wind speed averaging bin width based on RMS error. (b) GMF for different quadrants.

The width of each wind speed bin is chosen based on the RMS difference plot shown in Fig.5.3a. RMS difference is evaluated by

$$RMSD(w_i) = sqrt < (\boldsymbol{\sigma_{0i}} - \bar{\sigma_{0i}})^2 >, \tag{5.1}$$

Here $\boldsymbol{\sigma_{0i}}$ is the set of radar cross-section measurements in the ith wind bin w_i and $\bar{\sigma_{0i}}$ is the mean value of the radar cross-section in the given bin. The bin width is set as a variable parameter, and the error for different wind speed ranges and bin widths is plotted in Fig.5.4a. The RMS difference is found to be roughly constant for bin widths up to $+/-5$ m/s. Hence, for optimal performance, we choose a width of $+/-4$ m/s for the analysis throughout this chapter. Fig.5.4b shows the GMF for different storm quadrants. The slope is derived by linear regression over the binned radar crosssection, as described above. These GMFs at high winds can be seen to vary with quadrants. This is consistent with the azimuthal variation of the local wind wave directions in published directional wavenumber spectrum data sets (B24, I09, I12, and I14) (*Hwang et al.*, 2018). Quadrant 3 has the highest GMF, indicating a younger sea condition. Quadrant 2 has the lowest GMF, indicating an extended fetch and duration, therefore a longer sea age.

5.2 Harmonic model function

An empirical GMF is developed here which includes first- and second-order harmonic dependence on azimuthal location within the storm. This approach is based on the idea that any azimuthally varying function can be modeled as a linear combination of sinusoids. This technique is commonly used to represent the azimuthal dependence of radar and radiometer observations of ocean surface winds (*Meissner and Wentz*, 2002; *Wentz and Smith*, 1999). The functional form of the model is given by

$$\sigma_0 = f(w, \phi) = a_0(w) + b_1(w)sin(\phi) + b2(w)sin(2\phi), \tag{5.2}$$

where w is the wind speed, ϕ is the azimuth angle, and (a_0, b_1, b_2) are model parameters that depend on wind speed. Note that the slope of the GMF above ~ 20 m/s

is the same for different incidence angles. Apart from the wind speed dependence of σ_0, it also has a dependence on incidence angle as shown in Fig.5.1. However, due to the limited size of the sample population, the dependence on incidence and azimuth angles cannot be separated. One way to address this issue is by maintaining a similar incidence angle distribution at all azimuth angles. This will mitigate the effect and the resultant three-parameter harmonic model can be considered to represent the azimuthal dependence averaged across all incidence angles. It should be noted that the strength of the azimuthal dependence may vary with incidence angle. Additionally, note that if $b_1 = b_2 = 0$, the new GMF essentially defaults to the earlier azimuth-free version.

The σ_0 observations are averaged over wind speed bins that are $+/-4m/s$ wide and the parametric model described above is fit to the observations for wind speed $> 15m/s$ by non-linear least-squares minimization. The three model parameters (a_0, b_1, b_2) are shown vs. wind speed in Fig.5.5 (a-c). The estimated parameters are shown in blue and the black-dashed lines represent the 95% confidence intervals on these estimates. Examples of the full GMF vs. azimuth angle at 20, 40 and 60 m/s are shown in Fig.5.5 (d). Several important features of the GMF behavior can be observed. First, the magnitude of σ_0 decreases with increasing wind speed, similar to the previous GMF behavior. Secondly, the azimuthal dependence of the GMF increases with increasing wind speed. Thirdly, the maximum value of σ_0 occurs in the third quadrant for each wind speed, consistent with the GMF magnitude noted in Fig.5.4 (b). The minimum lies in the second quadrant, and therefore, has a higher sea age and fetch conditions.

These effects are further illustrated in Fig.5.6, which shows the peak-to-peak azimuthal variation in σ_0 vs. wind speed. The azimuthal variation rises steadily between ~ 30 and 55 m/s. The drop-off in azimuthal variation above 55 m/s may be a result of the small number of samples available and the lower sensitivity to changes

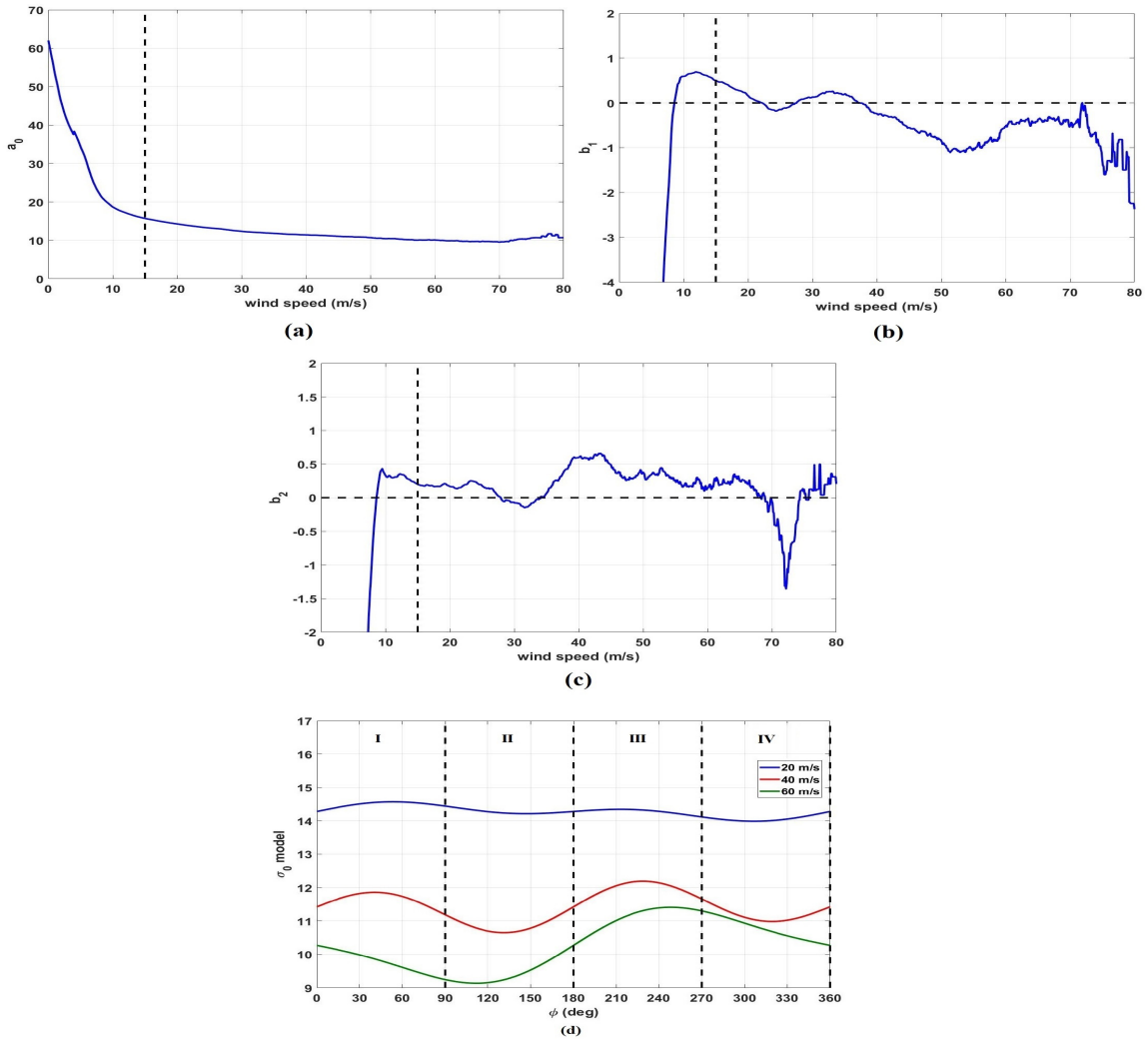


Figure 5.5: (a)-(c) Model parameters for azimuth GMF. (d) Azimuthal model for scattering crosssection shown for three different wind speeds.

in wind speed.

5.3 Performance assessment

To assess the ability of the proposed model to capture the azimuthal variation in σ_0, several statistical measures of performance are considered. The robustness of the model is evaluated by breaking the total dataset into 3 subsets using every third element. The model is then trained on one subset ($1/3rd$ of the data) and the relative

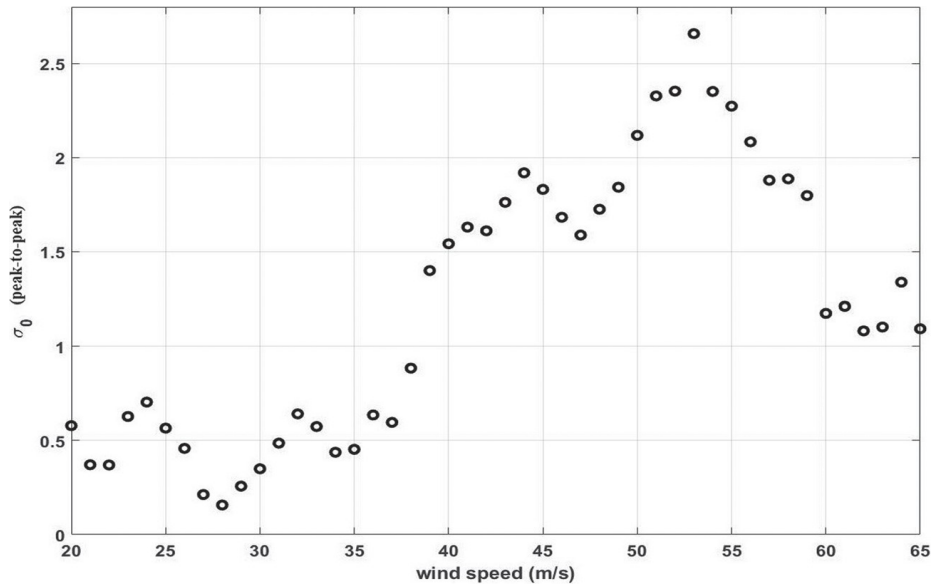


Figure 5.6: Peak-to-peak azimuthal variation of scattering crosssection for different wind speeds.

RMSE is evaluated on another dataset. This relative error formulation is given by

$$RealtiveError(\%) = \frac{\sqrt{<(\sigma_0 - \tilde{\sigma}_0)^2>}}{<\sigma_0>} * 100 \tag{5.3}$$

Here σ_0 is the measurement sample at a given wind speed bin and $\tilde{\sigma}_0$ is the model estimate of scattering cross-section for a given wind speed and azimuth information. Fig.5.7 (a) shows the behavior of relative error vs. wind speed for the different combinations of training and testing subsets. The dashed line represents use of the total dataset and hence is a test of internal consistency in the generation of the GMF. The large relative error at the lower wind speeds ($< 15m/s$) is due to the fact that the model has been trained only for higher wind speeds.

The relative error is consistent over the 3 different datasets (D1, D2 and D3) as well as over the total dataset, hence indicating the robustness of the developed model. Also, for wind speeds $>$ 60 m/s the relative error becomes noisy and this is attributed to the sparse observations at such high winds. Next, the dataset is divided into 2

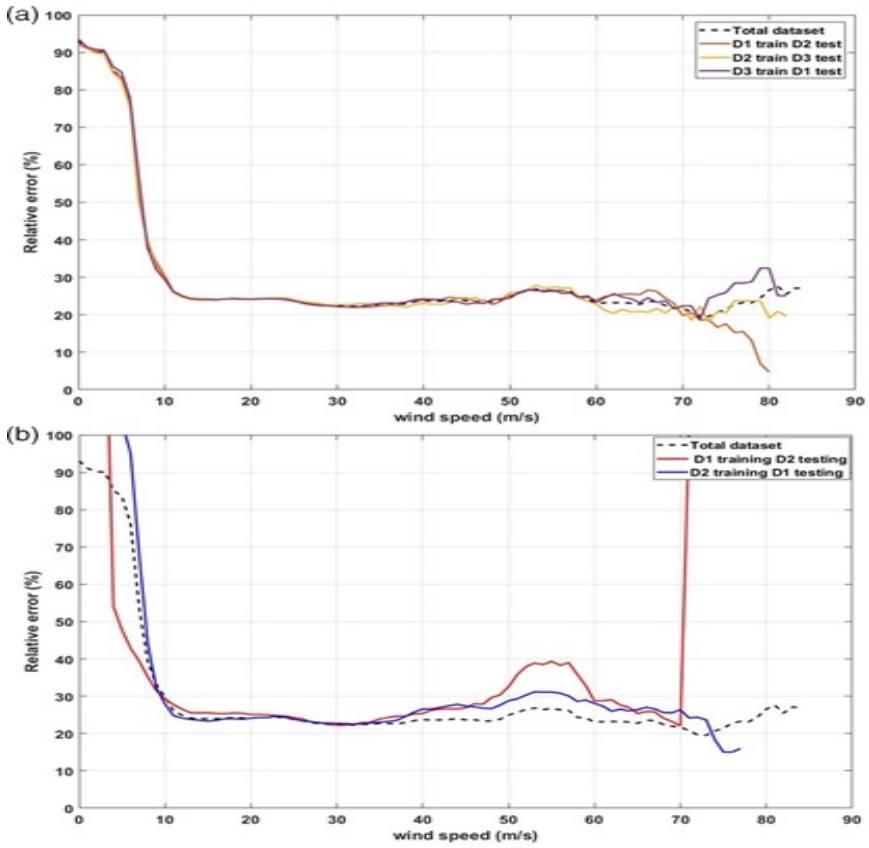


Figure 5.7: (a) RMSE for training and testing on 3 subsets (D1, D2 and D3). (b)RMSE for training and testing of 2 independent datasets (D1 and D2).

independent subsets (D1 and D2) with storms well mixed from different basins and years and the analysis is repeated. The result is shown in Fig.5.7 (b). The relative error is consistent over a wide range of wind speeds from $20 - 60$ m/s re-attesting to the robustness of the model.

The next performance metric is a statistical measure of the percentage of azimuthal variation captured by the model over different wind speed ranges. The metric used for this purpose is given by the following expression and the physical interpretation of its limiting values are described in the equation.

$$1-\frac{\mathbf{var}(\tilde{\sigma}_0)}{\mathbf{var}(\sigma_0)}*100 = \begin{cases} 1, & \text{if model captures azimuth information completely.} \\ 0, & \text{if model does not capture any azimuth information.} \end{cases} \tag{5.4}$$

Here **var(.)** refers to the variance of the sample population, σ_0 is the measurement and $\tilde{\sigma}_0$ is the model. The assumption behind this statistical metric is that the total variance in the observations at a given wind speed is associated with multiple factors, one of which is azimuthal variation. Since the model explicitly accounts for azimuthal variation, any residual variance in the model should be due to factors other than azimuthal variation. If the metric is 1, it indicates that the model has captured all of the azimuth information and, if the metric is 0, no azimuth information is captured. Fig.5.8 shows the behavior of this metric tested for different wind speeds. The metric suggests that the model is consistent and is able to capture greater than 90 % of the azimuthal variation over a broad range of wind speeds between 20 and 60 m/s. Lower wind speeds are shown here for completeness and, again, at higher winds the performance drops due to sparse observations in the region.

Finally, we evaluate the total error in the observed scattering cross-section due to the azimuthal variation. The measure is defined by:

$$Error(\%) = abs\left(1 - \frac{\mathbf{MSD}(\sigma_0, \tilde{\sigma}_0)}{\mathbf{var}(\sigma_0)}\right) * 100. \tag{5.5}$$

Here, **MSD(.)** is the mean squared difference between the observations and the model estimate for a given wind speed while **var(.)** is the total variance in the observation for a given wind speed. If the MSD between the observation and the model is the same as the variance in the observation, it suggests that no systematic azimuthal variation observed and if the MSD is negligible compared to the total

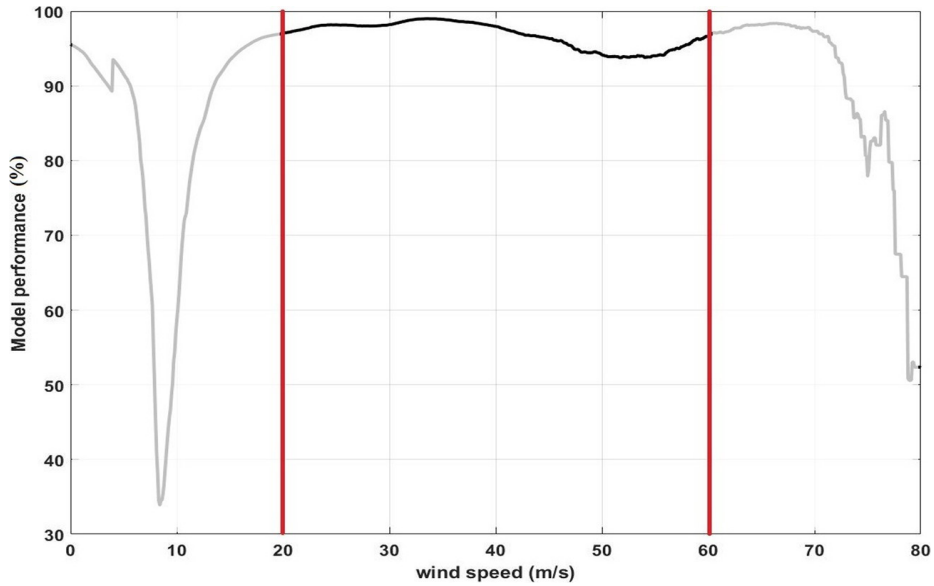


Figure 5.8: Evaluation of azimuthal information captured by the model.

variance in the observation, then most of the error in the σ_0 can be attributed to azimuthal variation. Fig.5.9 shows the % error in the scattering cross-section caused by the azimuthal variation. The model suggests a 2-8 % error in the scattering cross section for the wind speed range 20-60 m/s. The error is close to 2% at 20 m/s and gradually increases to 8 % around 53 m/s, then begins to reduce above that. While the overall error due to azimuthal variation is negligible, understanding the effect of this variation gives important insight into the wave properties inside a hurricane and its impending impact on the scattering cross-section. The increase in azimuthal variation with increase in wind speed also suggests that the GNSS-R scattering cross-section could be sensitive to the directional properties of wind at higher wind speed ranges.

5.4 Discussion and conclusion

The specular bistatic scattering cross section of the ocean surface in tropical cyclones, as measured by GNSS-R radar receivers on the CYGNSS spacecraft, is found to de-

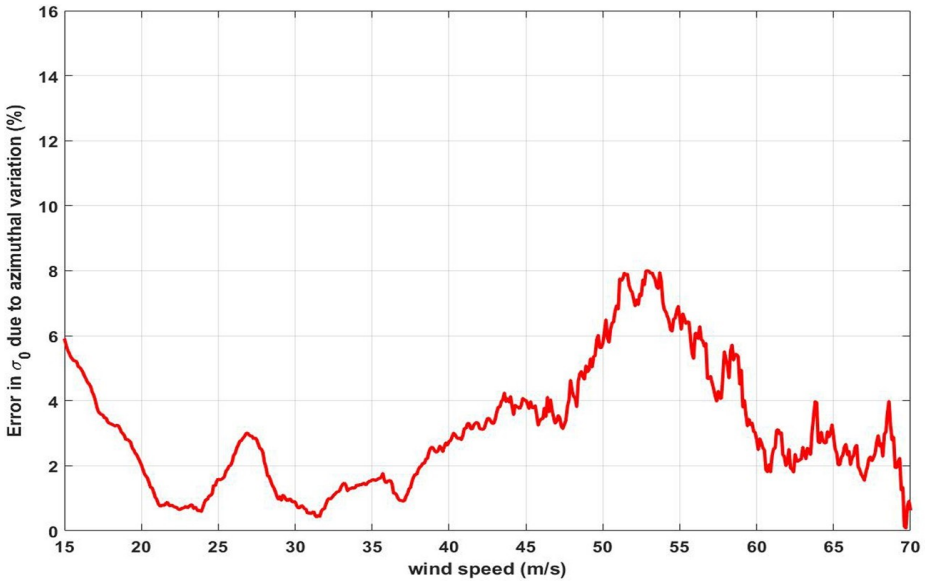


Figure 5.9: Error associated with scattering cross-section due to azimuthal variation.

pend on azimuthal location relative to the direction of storm motion. The dependence is caused by variations in the sea age and fetch length with storm quadrant, which affects the balance between surface roughness at short (capillary) and long (gravity) wavelengths. The roughness spectrum, in turn, affects the scattering cross section. A modified Geophysical Model Function (GMF) is developed using a second order harmonic expansion to represent the azimuthal dependence. The zeroth order term in the GMF is consistent with previous models that have not included an azimuthal dependence. The first and second order terms together explain between 2 and 8% of the total variance in the scattering cross section, with higher explained variance being associated with higher wind speeds. The azimuthal corrections to the GMF are found to be significant between $\sim$ 20 m/s and 60 m/s. Above $\sim 60 m/s$, the results are inconclusive owing to the scarcity of samples.

It is worthwhile to note that the current GMF used by the CYGNSS project does not include an azimuthal dependence in organized storms. Another difference from the GMF developed here is the source of reference winds. The v2.1 GMF was developed using matchups with near-surface wind measurements made by the Stepped

Frequency Microwave Radiometer on NOAA hurricane hunter aircraft (*Ruf and Balasubramaniam*, 2018), whereas the GMF developed here uses matchups with HWRF model winds. Differences in overall magnitude between the two GMFs are likely a result of these differences.

The azimuthally dependent GMF has utility in two regards. It can serve as an indicator of sea state development in the inner core of tropical cyclones, for use in process studies into air-sea and wind-wave interactions. It could also be used to improve wind speed retrieval algorithms in tropical cyclones that are based on GNSS-R observations. Retrieval algorithms essentially invert the GMF to estimate wind speed given the scattering cross section, and a more physically representative forward GMF will allow for a more accurate inversion. This type of study can also be extended to other non-TC weather systems with younger seas or scenarios with limited fetch conditions, e.g, limited fetch on the lee side of major islands that may result in a different relationship between wind speed and MSS or σ_0, and therefore require a modified GMF for accurate wind speed retrieval. The next steps in this work will be to implement the proposed azimuthal GMF for CYGNSS wind retrieval and evaluate its performance.

CHAPTER VI

Machine Learning Based Quality Control of CYGNSS Retrieved Winds

The near surface ocean wind speed retrieval by CYGNSS uses empirical Geophysical Model Functions (GMFs) developed for 2 measurement observables, namely the Normalised Bistatic Radar Cross-Section (NBRCS) and the Leading Edge Slope (LES), derived from the DDMs (*Ruf et al.*, 2016). Wind speed is then estimated by inverse mapping of these observables to reference winds (NWP models and aircraft measurements) using the GMFs and optimally combining the two estimates using a minimum variance estimate of the wind speed (*Clarizia and Ruf*, 2016b).

The error in the retrieved wind speed can come from different levels of measurement processing and it is important to identify and eliminate erroneous measurements in order to provide high quality observations for scientific applications. While previous spaceborne GNSS-R missions were primarily focused on technology demonstrations, CYGNSS attempts to provide an operational service to meet its science goals. Hence high data quality is a top priority.

At the engineering level, the major sources of error can be attributed to errors in estimation of the GPS transmit power, GPS antenna patterns, spacecraft pointing knowledge, and star tracker/science antenna boresight misalignment (*Gleason et al.*, 2016). At the instrument level, possible errors include science antenna gain error, instrument

noise power estimation error, calibration load temperature error, and Digital-Analog quantization error (*Gleason et al.*, 2018a). At the measurement level, errors can occur due to specular point geo-location errors, dependence of observables on other geophysical parameters such as wave age, swell etc., non-unique mapping from wind speed to observable, and error in interpolation in space and time for co-location with reference winds, assumption of wind speed uniformity over $25km^2$ surface etc. (*Ruf et al.*, 2018b).

A number of quality control flags are already in place at different levels of the data processing in order for the science community to pick and choose data with quality requirements specific to their application. At Level 1, some of the major quality flags are star tracker attitude status and one-hertz status flags indicating if Milky Way or the Sun is in the zenith antenna field of view (FOV). The Level 1 quality flags also look indicate spacecraft attitude errors, radio frequency interference (RFI), DDM noise floor errors, rapid rate of change of receiver temperature, and telemetry errors (*Gleason*, 2014). At Level 2, the quality flags indicate retrieval ambiguity, low Range Corrected Gain (RCG) and other data processing errors (*Clarizia et al.*). Despite stringent quality flags filtering the data, there remain occasional outlier samples with large discrepancies between the CYGNSS retrieved wind speed and reference validation winds (shown in Fig.6.1) To improve the data quality of CYGNSS, another layer of quality control is needed which can effectively identify and eliminate these outliers. This is the primary objective of this work.

In this work we develop a Neural Network based quality control filter for CYGNSS Level 2 winds that can effectively identify and remove outliers. We also consider the performance of the CYGNSS retrieved winds before and after this filter is applied to assess its efficacy. The reminder of the chapter is structured as follows: Section 6.1 describes the datasets used, Section 6.2 explains the details of the proposed quality control filter and in Section 6.3 the performance of the proposed algorithm is an-

alyzed and the CYGNSS level 2 data performance is assessed before and after the filter. Section 6.4 discusses the trade-offs in performance and offers conclusions of this study.

6.1 Data description

The Level 2 CYGNSS winds are minimum variance estimated winds from two observables, namely NBRCS and LES for low-moderate wind speed ranges (0-25 m/s). The CYGNSS retrieved winds are matched to near-coincident independent estimates of the ocean surface wind speed referenced to a 10m height (u_{10}) from the Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2) (*Gelaro et al.*, 2017). MERRA-2 is a reanalysis product provided by NASAs Global Modelling and Assimilation Office (GMAO). The reference matchup MERRA-2 gridded data product has a spatial resolution of 0.5 deg x 0.625 deg (lat,lon) and an hourly instantaneous assimilation (*Bosilovich et al.*).

Fig.6.1 shows the density scatter plot of CYGNSS retrieved winds with respect to the MERRA-2 winds. In the figure, the dashed line represents the 1:1 line and the solid line represents the mean retrieved wind speed line, which essentially is the GMF. This plot is generated by dividing the 2-D space into 500 bins or regions. And the matchup winds are assimilated into the nearest bin and finally the log to the base 10 of the number density is taken for better visualization of the density differences. There are several important observations from this plot. Firstly, most of the observations fall along the 1:1 line at lower wind speeds, indicating good retrieval quality. However, a cluster of very high CYGNSS retrieved winds (15-35 m/s) is noticeable at low MERRA-2 winds (5-10 m/s). The improved filtering method developed here targets the removal of these outliers. Secondly, the GMF line and the 1:1 line are very similar up to a MERRA-2 and CYGNSS wind speed of $\sim 10m/s$. Above this range,

the GMF line begins to deviate away from the 1:1 line. This inherent bias in the GMF complicates the identification of outliers by the filter algorithm. The purpose of a quality filter is to remove outliers only and not correct for biases in the retrieval. This is another consideration to be accounted for while designing the filter. Finally, the density of samples at high MERRA-2 wind speeds ($> 20m/s$) is very small relative to the lower wind-speed ranges. Therefore trade-off studies must be performed for filter design to balance between efficiency of outlier removal and retaining as many high wind samples as possible. All of the above objectives will be addressed in the course of developing the filter.

Over and above the existing quality flags, 13 diagnostic variables are used to distinguish outlier samples from good samples. These diagnostic variables are listed in Table 6.1. The choice of diagnostic variables is based on previous calibration experience with GNSS-R data and error analyses (*Gleason et al.*, 2016, 2018a; *Ruf et al.*, 2018b). The diagnostic variables can be categorized into 3 major types: instrument related attributes, measurement geometry related attributes and surface related attributes. In Section 6.3 these diagnostic variables will be assessed for their individual significance in enabling the filter to distinguish between outliers and good samples.

6.2 Proposed quality control method

Outlier/anomaly detection is an active research field spanning a wide range of applications from manufacturing quality control to astronomical detections. Machine learning techniques are widely used for outlier detection and automation of quality control processes (*Mehrotra et al.*, 2017). Despite an emerging trend in the use of machine learning methods for Earth Observation applications, the calibration and validation of satellite measurements are most often handled manually by instrument specialists. Utilizing the capabilities of machine learning tools for calibration and

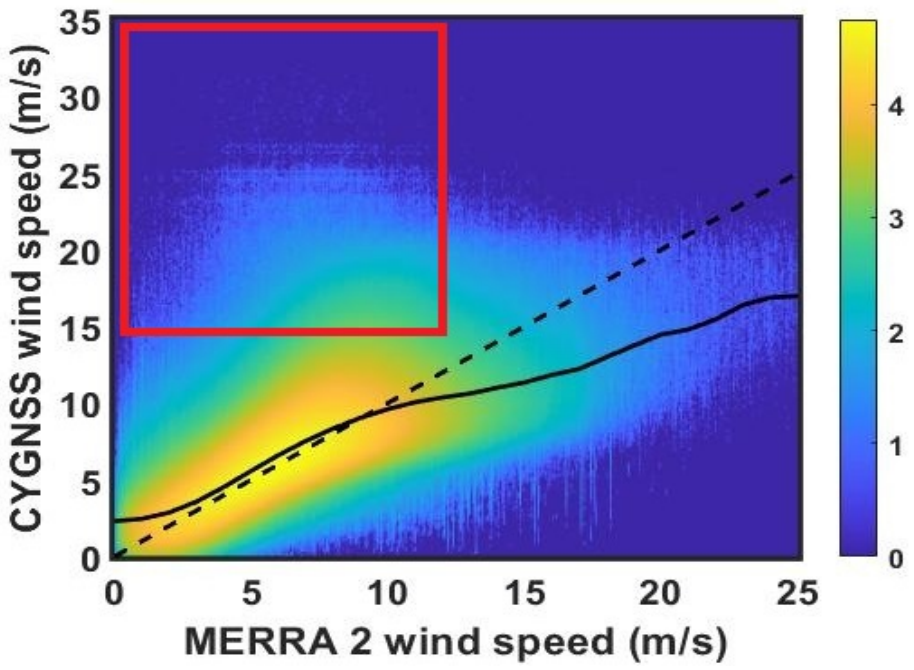


Figure 6.1: Log density plot of CYGNSS Level 2 retrieved winds matched to MERRA-2 reference winds for the wind speed range of 0-25 m/s. The dashed line represents 1:1 agreement between the two winds. The solid line is the average CYGNSS retrieved wind at each MERRA-2 wind speed. A clustering of outliers can be seen near MERRA-2 wind speeds of 0-10 m/s and CYGNSS wind speeds $> 15 m/s$ (shown with a red box). One primary objective of the new filter is removal of this cluster.

validation activities can help to better understand the behavior of the data. Outliers can be defined as sample measurements that have a distinct deviation in their properties when compared to the major proportion of the data (*Hawkins*, 1980). Visually, these measurements are regions of low density in sample space, i.e., they have a low number of neighboring points within a threshold distance compared to the rest of the sample space. Machine learning tools that are widely used for outlier analysis include the supervised classification techniques such as Neural Networks (NN), K-Nearest Neighbors (K-NN), Decision trees, Support Vector Machines (SVM) etc. (*Aggarwal and Sathe*, 2017). As noted in Fig.6.1, there are distinct regions away from the GMF line and the 1:1 line that should be detected and removed. For the CYGNSS quality control filter, supervised training of a Neural Network is used for outlier detection and

Diagnostic variable	Description	Type	Symbol/ abbreviation
prn_code	GPS PRN code	Instrument attribute	PRN
antenna	CYGNSS science antenna	Instrument attribute	ant
nst_att_status	Star tracker attitude status	Instrument attribute	nst
sc_roll	Spacecraft roll	Instrument attribute	Roll
zenith_ant_gain	Zenith antenna Gain	Instrument attribute	Z_{gain}
zenith_power	Zenith antenna power	Instrument attribute	Z_{power}
incidence_angle	Incidence angle at SP[1]	Geometry attribute	θ
azimuth	Azimuth angle of SP	Geometry attribute	ϕ
range_corr_gain	Range Corrected Gain	Geometry attribute	RCG
sp_rx_gain	CYGNSS antenna Gain at SP	Surface attribute	G
ddm_snr	SNR at SP	Surface attribute	SNR
ddm_nbrcs	NBRCS at SP	Surface attribute	DDMA
ddm_les	LES at SP	Surface attribute	LES

Table 6.1: List of diagnostic variables used.

removal. The details of the quality control filter design are explained in this section.

6.2.1 Population definitions

The CYGNSS Level 2 v3.0 data with MERRA-2 wind speed matchups from the year 2018 have a total of $\sim$ 153 million samples. The sample space is divided into 2 regions. The low wind region consists of all samples with CYGNSS retrieved winds, U_{CYG}, less than or equal to 10 m/s. The high wind region consists of samples with U_{CYG} greater than 10 m/s. In both regions MERRA-2 wind speed, U_M is required to be less than 25 m/s. This division of sample space is due to the behavior of the GMF line relative to the 1:1 line. Below 10 m/s, the GMF line is very similar to the 1:1 line (see Fig.6.1) and above this wind speed the GMF line begins to underestimate. Therefore, it is appropriate to have two different training datasets, one for each region.

In the low wind region, a good sample satisfies $|U_{CYG} - < U_{CYG} >|_{(U_M)}| <= 1m/s$. and an outlier satisfies $|U_{CYG} - < U_{CYG} >|_{(U_M)}| > 4m/s$, where $< U_{CYG} >|_{(U_M)}$ refers to the mean value of the wind speed retrieved by CYGNSS for a given value of the MERRA-2 wind speed. This relationship is described by the solid line in Fig.6.1. In the high wind region, a good sample is defined by $(U_{CYG} - U_M) <= 2m/s$ and an outlier as a sample with $(U_{CYG} - U_M) > 3m/s$. The difference in training population definitions at low and high wind speeds is due to the inherent bias in the GMF which can be observed as the deviation of the retrieval mean (solid line) from the 1 : 1 agreement (dashed line) above 10 m/s in Fig.6.1. As the wind retrieval is based on the GMF, a bias in the GMF can lead to under/over-estimation of winds despite being a good measurement. To mitigate the effect of GMF-induced bias on the outlier detection capability of the filter, the filter is trained with respect to the GMF. However, as the filter is reliant on the Level 1 diagnostic variables, which are independent of GMF, the samples lying near the 1:1 line are also good samples and therefore the modified definition of training data is used at high winds. The training datasets for the 2 different sample spaces are shown in Fig.6.2. Such conservative training definitions are used to improve the outlier detection capability of the algorithm. Further analysis of the definition of a good or an outlier sample is discussed later in this section. For training, we use $\sim$ 4 million samples for each wind speed region and $\sim$ 8 million samples for validation. The performance metrics used to evaluate the outlier detection capability of an algorithm are the Probability of Detection (PD) and False Alarm Rate (FAR). For these metrics, the definition of good and outlier samples are different from the training definitions. The validation definitions are based on the NASA mission requirements on wind retrieval error. Thus, the wind speed differences for a good sample shall be less than 2 m/s from the mean and an outlier is defined as those samples having a difference greater than

[1]SP = Specular Point

5 m/s. Finally, the filter is tested over the total population and its performance is assessed.

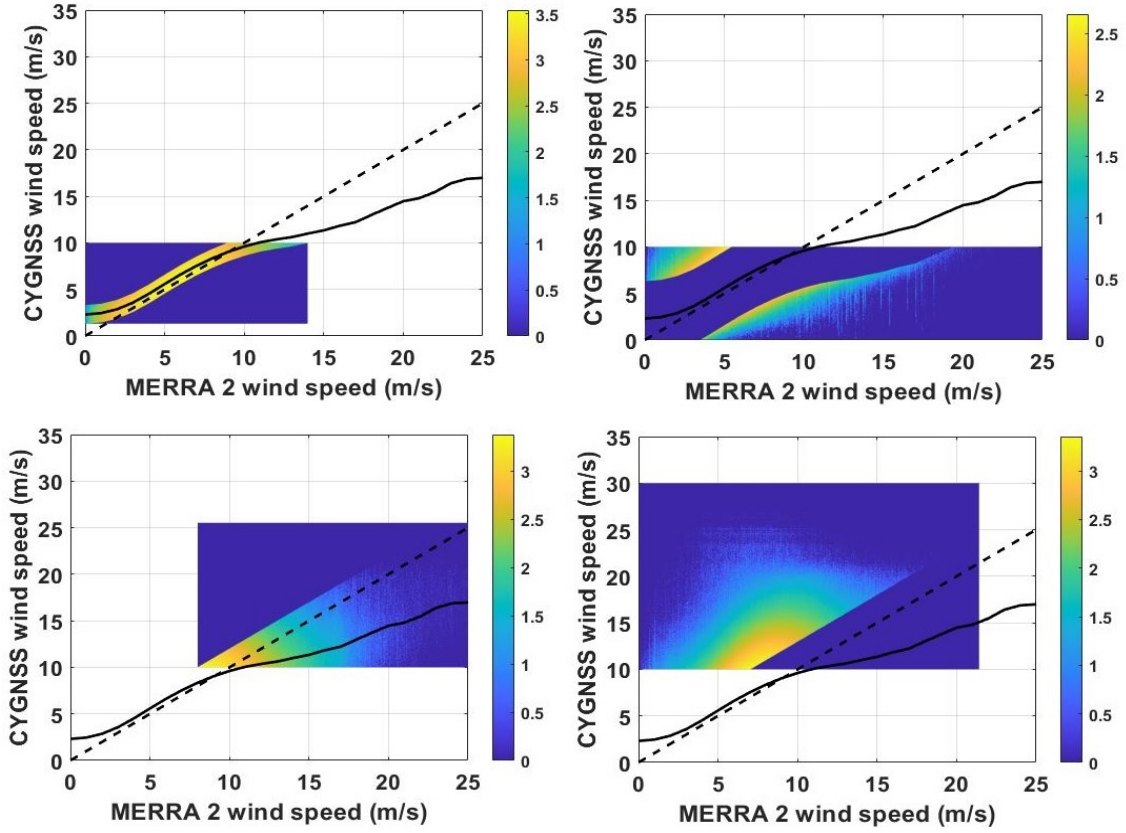


Figure 6.2: Density plot of CYGNSS Level 2 retrieved winds matched to MERRA-2 reference winds used for training. Top row represents the good (left) and outlier (right) training population for ($0 < U_{CYG} <= 10m/s$). The bottom row represents the good (left) and outlier (right) training population for ($U_{CYG} > 10m/s$).

6.2.2 Quality control process design

A block diagram representation of the quality control design process is shown in Fig.6.3. The first stage of the algorithm is feature extraction. The input to this stage is the Fully Developed Seas (FDS) winds over a reference wind speed region of 0-25 m/s. The CYGNSS Level 2 wind retrievals are of two kinds: the FDS and Young Seas Limited Fetch (YSLF) winds. The FDS winds are low to moderate winds (up

to 25 m/s) over fully developed waves in the ocean. This forms the major proportion of the total measurements. The YSLF winds are hurricane force winds measured over the tropical cyclones that have varying wave age and fetch conditions. The filter proposed in this work is developed specifically for FDS winds as this dataset encompasses the majority of the measurements and has a well behaved nature relative to its counterpart. The feature extraction stage extracts the different diagnostic variables listed in Table 6.1 for every sample point. Next is the training stage. One Neural Network (NN) classifier is trained for each of the wind speed intervals over the individual training datasets described above. The last stage is the validation and testing stage where the skill of the filter is assessed. The performance assessment of this Neural Network filter is discussed in detail in Section 3.

Apart from the Neural Network filter, other standard supervised outlier detection techniques such as Logistic Regression, Decision Trees and K-NN are also considered and their PD and FAR performances are listed in Table 6.2. It can be observed from Table 6.2 that K-NN has very good PD when compared to the other 2 algorithms but has a high FAR as well. In general, NN is preferred over K-NN because of the heavy computational memory requirement of K-NN as compared to the memory requirement for training the NN coefficients.

Apart from the Neural Network filter, other standard supervised outlier detection techniques such as Logistic Regression, Decision Trees, Nave Bayes and K-NN are also considered and their confusion matrices are listed in Fig.6.4. In the confusion matrix, the rows represent the true classes, the columns represent predicted classes, and the percentage of samples are given in each of the boxes. Outliers are represented as class 0 and good samples are represented as class 1. Among the various classifiers experimented with, the K-NN and the NN have a similar performance. In general, NN is preferred over K-NN because of the heavy computational memory requirement of K-NN as compared to the memory requirement for training the NN coefficients.

Classifier	Training time (secs)
Logistic Regression	118.43
Decision tree	78.87
Nave Bayes	56.39
K-NN	49,608 (13.78 hrs)
NN	6475 (1.8 hrs)

Table 6.2: Training time required for individual classifiers.

This can be seen in terms of the time requirement for training each of the classifier, shown in Table 6.2. It can be seen that K-NN requires the most time, followed by the NN.

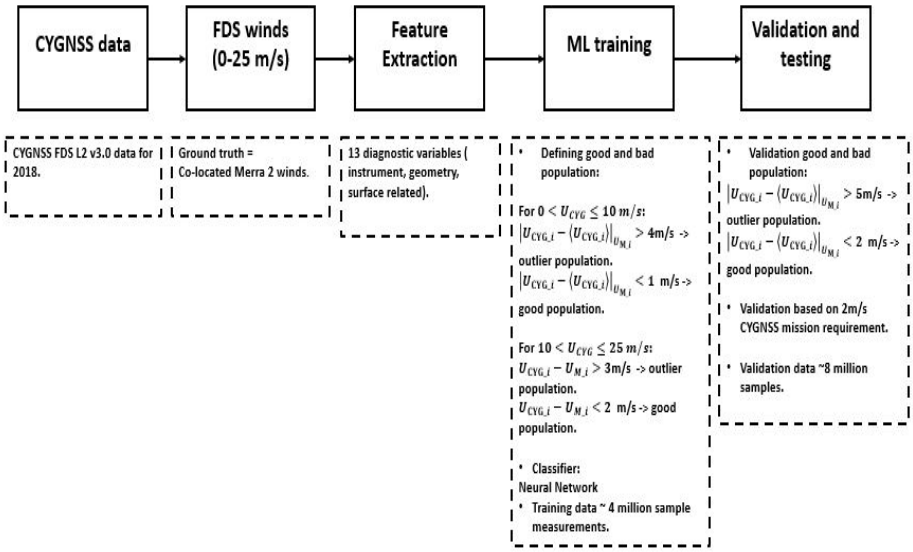


Figure 6.3: Steps involved in the new quality control algorithm for CYGNSS data. The algorithm has 3 major stages Feature Extraction, ML training and Validation/testing.

6.2.3 Neural Network filter design

The NN used for this application consists of a single hidden layer with 10 neurons. The input layer consists of 13 neurons, each for one diagnostic variable and the output layer has one neuron that classifies an input sample as an outlier or a good sample. Only one hidden layer is used as it is a sufficient condition to form any bounded/unbounded

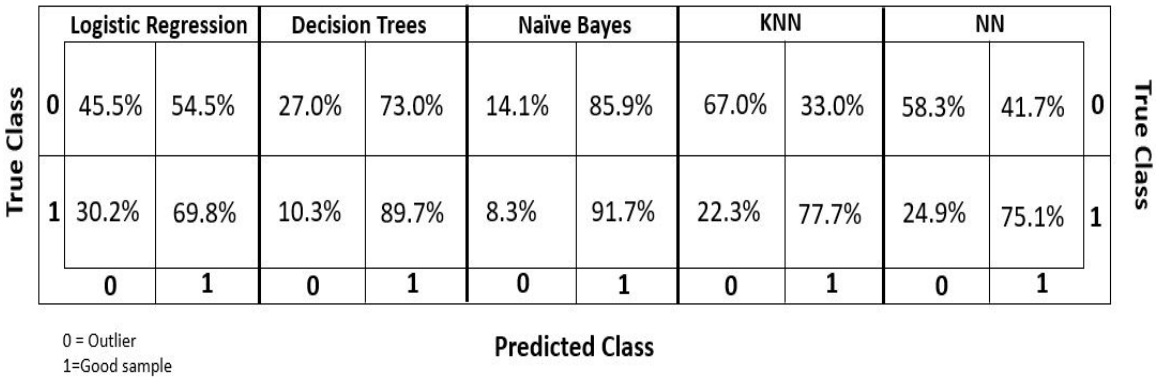


Figure 6.4: Confusion matrices for different classifiers.

convex region in the space spanned by the input (*Lippmann*, 1987). The choice of the number of neurons in the hidden layer is decided by experimentation. In general, a feedforward network can have any shape but the commonly used structure is a pyramidal structure with a decrease in the number of neurons at each layer away from the input. There is practically no upper limit on the number of neurons to be used in this case as the training population is very large ($\sim$ 4 million). Hence, we explored three different neuron counts, namely 5, 10, and 15, and the performance plot in terms of PD and FAR at different wind speeds is plotted to make a choice on the hidden layer size. The performance plot is shown in Fig.6.5.

The blue curves represent PD and the red curves represent FAR. We observe that the three network sizes exhibit a very similar performance in terms of PD and FAR over the entire range of wind speed. For this reason, other performance metrics, such as computation time, network complexity and % samples removed as outliers should be prioritized when choosing the optimal structure. In terms of computational time, NN size = 15 is the shortest, followed by NN size =10, and the longest is NN size =5. This is an expected trend, as simpler networks can take a longer time for error convergence. Next, in terms of network complexity, NN size = 5 has the least number of tunable parameters, followed by NN size=10 and the largest being NN size =15. The % of samples removed as outliers by NN size =5 is $\sim$ 23% of the total data, by

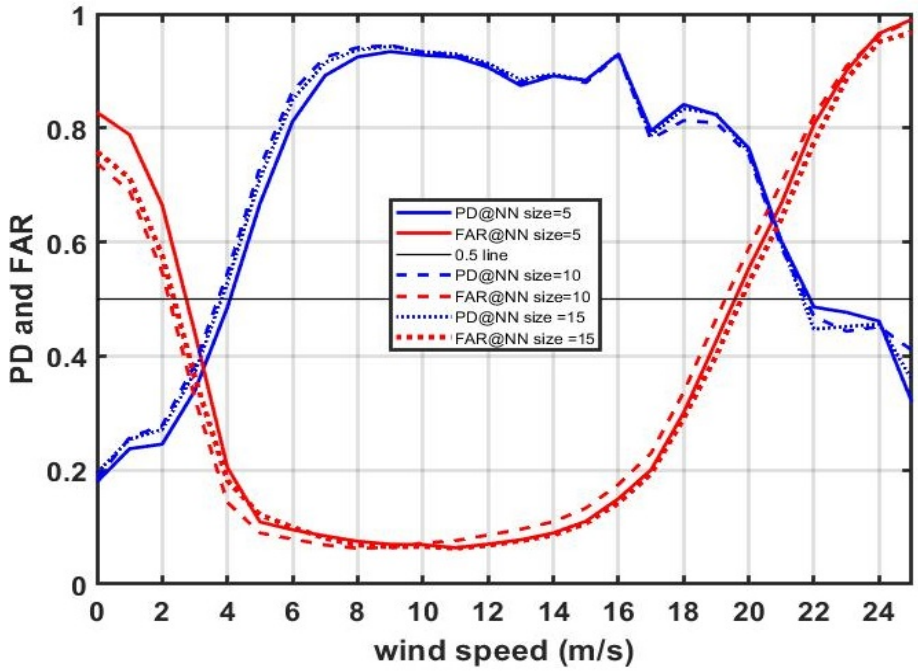


Figure 6.5: PD and FAR curves for 3 different network sizes (5, 10 and 15).

NN size =10 is $\sim$ 20 % and by NN size = 15 is $\sim$ 22%. Therefore, based on all of these considerations, NN size=10 was chosen as the optimal network design for this application.

Thus for purposes of quality control application for CYGNSS, the QC filter has two NNs (NN1 and NN2), each trained for a specific wind speed range (0-10 m/s and 10-25 m/s). The NNs are identical in architecture and contain 2 layers with 10 neurons in the hidden layer, as discussed above. The hidden layer is trained with a sigmoid transfer function and a linear transfer function is used at the output. The optimization algorithm used for this is the widely used Levenberg-Marquardt algorithm.

To evaluate the design space of this filter, the definitions of good and outlier samples are varied and the performance metrics are plotted. Understanding the behavior of the filter for different sample definitions can help users understand how the network handles the outliers and choose an optimum definition based on the application requirements. The family of PD and FAR curves are plotted in Fig.6.6. The blue curves represent PD and red curves represent FAR. The PD metric is affected by

the density of outlier samples and the FAR metric is affected by the density of good samples. Changing the wind speed difference thresholds for good and outlier samples will affect the overall performance of this QC filter. For this study, the wind speed difference from the GMF line for a good sample is varied from 1 m/s to 4 m/s and for an outlier is varied from 3m/s to 7 m/s.

There are many interesting features in Fig.6.6. Firstly, the FAR curves do not vary much with changes in the definition of the good population but there is a significant jump in PD with changes in the definition of the outlier population. This is due to the relatively small percentage of outliers when compared to the total sample population. Next, the FAR metric has the best performance when the good sample definition is set to 1 m/s and gradually degrades with increase in the difference. However, above a wind speed of $\sim$18 m/s, the trend reverses. This is due to the fact that, at higher wind speeds there is a greater degree of scatter in the data (as seen in Fig.6.1) resulting in poorer performance in terms of FARs at very stringent definitions of a good sample. Next, as mentioned earlier, the PD metric seems to have a strong jump with change in outlier definition; with the highest PD performance for a wind speed difference of > 7 m/s from the GMF line. Again, the trend flips in nature at higher wind speed (> 21 m/s), which again is attributed to increased scatter in the data. Finally, it is important to note that the general performance of the filter is not optimal at very low wind speeds (< 3 m/s) for any definition of good and outlier sample. Thus, the ideal operating range for this filter is ~ 5 m/s to 18 m/s; this is where most of the samples lie. The choice of the definitions is dependent on the application. For instance, applications that require very high quality control like monitoring long term variations in wind speed data must go for highest PD performance. Applications at higher wind speeds, which are needed to retain as many higher wind speed samples as possible, must go for lower PD performance. In this work on the assessment of wind retrieval performance, the definition of a good sample is a wind speed difference

<= 2m/s and outliers are defined as > 5 m/s.

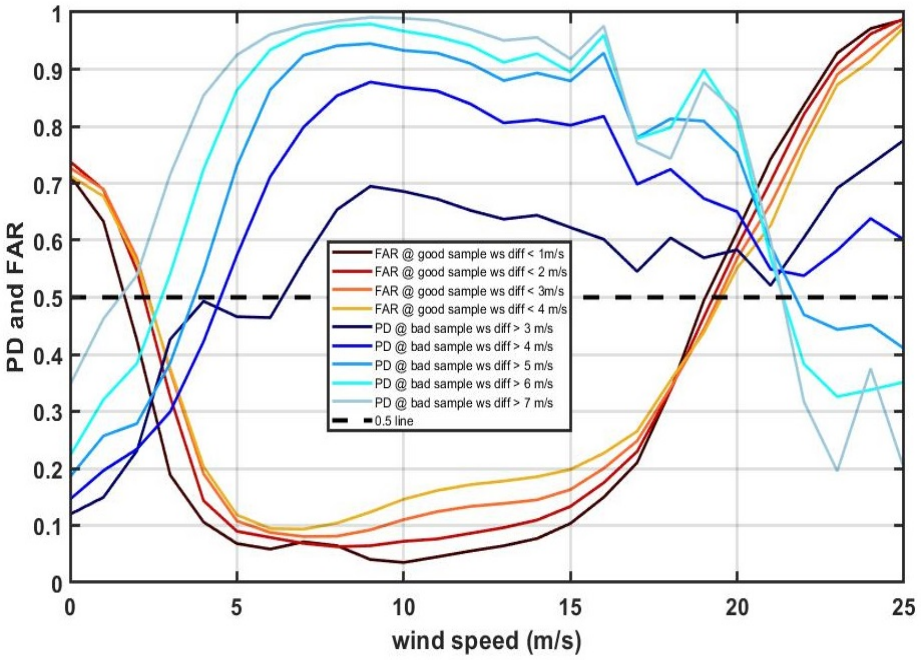


Figure 6.6: Family of PD and FAR curves for different definitions of good and outlier samples. The dark-light blue curves represent PD and the orange-red curves represent FAR.

6.3 Results

In this section the performance of the quality controlled CYGNSS wind speed data set is assessed. Two identical Neural Networks, one for each wind speed region discussed in Section 6.2 are trained. The first NN is applied to CYGNSS winds between 0-8 m/s and the second NN is applied to CYGNSS winds $> 8m/s$. This slight shift between the training and testing wind speed regimes is to improve the net performance of the filter, as the first NN will be biased towards lower winds where the highest density of samples occur and the second NN will again be biased towards the lower winds in the range (10-35 m/s). The resulting quality controlled CYGNSS wind speed dataset is shown in Fig.6.7.

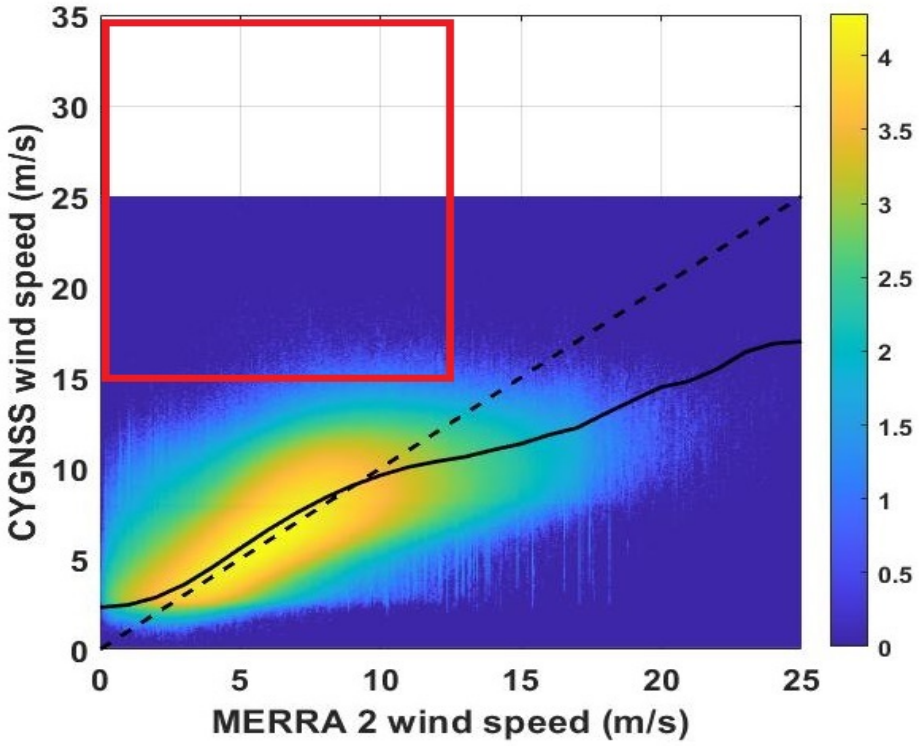


Figure 6.7: CYGNSS retrieved wind dataset after quality control. The outliers in the red box have been mostly eliminated here relative to Fig.6.1

Comparing Fig.6.1 and Fig.6.7 demonstrates the effectiveness of the filter. The large cluster of high CYGNSS winds at low MERRA-2 winds has been removed by this filter. Also, the CYGNSS samples are now evenly distributed along the GMF line (solid black line), in contrast with the original dataset. Finally, a significant reduction of scatter in the dataset can be observed. The performance of this proposed QC filter is assessed in the following subsections based on the error statistics Mean Difference (MD), Root Mean Squared Difference (RMSD) and variance of data. The test dataset consists of all the sample points (∼153 million).

6.3.1 Algorithm performance analysis

To assess the skill of the quality control algorithm, first the validation metrics, PD and FAR, are examined in Fig.6.8. These metrics are based on the design parameters

discussed in the previous section. The optimal range of operation for this filter is $\sim$ 5 m/s to 17 m/s. In this range the FAR for good samples is consistently $< 20\%$ and the PD for outliers is $> 75\%$. The peak performance is between 6-14 m/s where FAR $< 10\%$ and PD is $> 80\%$. This is also the region of maximum data density as the wind speed distribution has a peak near 7 m/s.

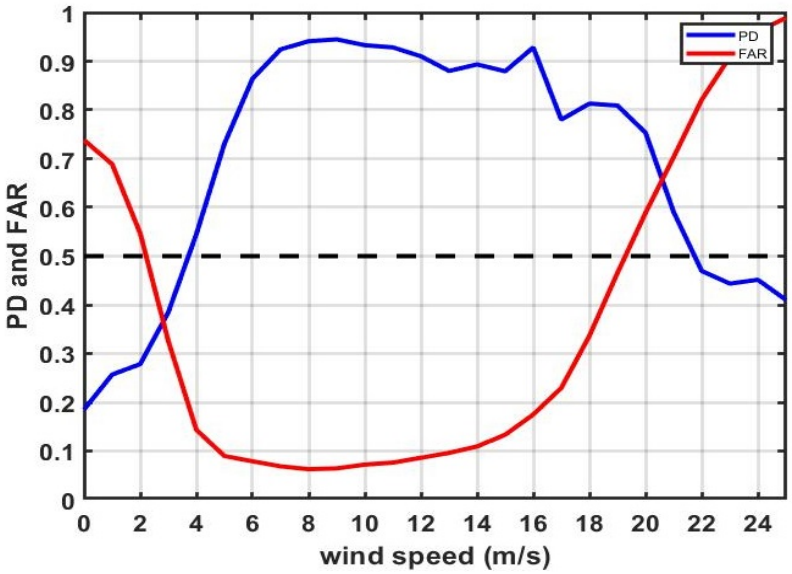


Figure 6.8: PD and FAR metrics for the CYGNSS test dataset.

Next, the skill of the filter is assessed by looking at the ratio of number of outliers identified by the filter to the total number of outliers for a range of wind speed differences. This is shown in Fig.6.9. The x-axis is the difference between CYGNSS wind speed and the GMF line. As per our validation criteria, we have defined a sample as an outlier if the difference is greater than 5 m/s. The 5 m/s threshold is shown in red. It can be observed that $\sim 70\%$ of the outliers are correctly identified for wind speed difference $\sim$ 5 m/s and the filter eliminates close to $\sim 100\%$ of outliers with wind differences $>$ 10m/s.

To understand Fig.6.9 better, we look at the distribution of outliers (wind speed difference >= 5m/s) at different MERRA-2 wind speed bins before and after applying the filter. This data distribution is shown in Fig.6.10. The red distribution shows

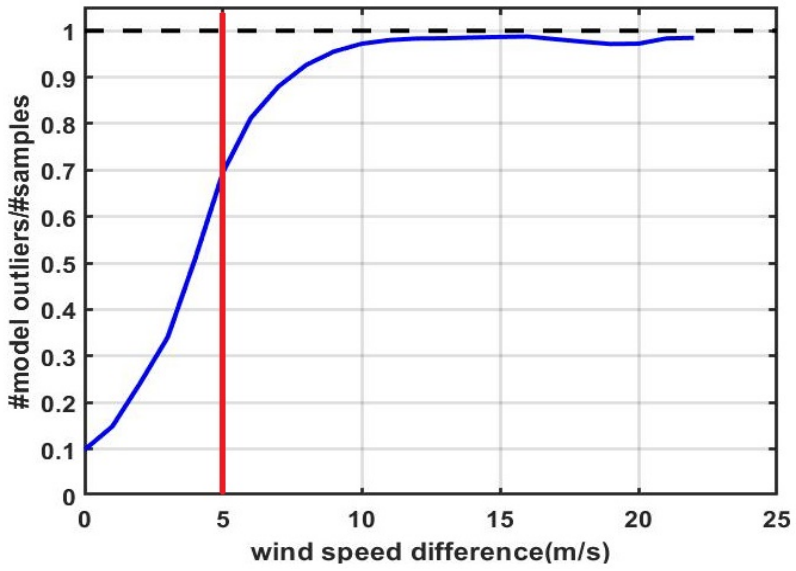


Figure 6.9: Ratio of outliers rightly identified by the filter to the actual # of outliers Vs wind speed difference.

the density of outliers in the original dataset and the blue shows the distribution of outliers after applying the filter. Firstly, a very significant decrease in the outlier population can be observed after filtering. The filtered dataset has approximately 4 times fewer outliers. In the original dataset, most of the outliers are present between 5-10 m/s which is also the peak region for wind speed distribution. In this region the filter has been able to remove a large proportion of the outliers. Next, in the filter design section the low PD and high FAR at high winds region was discussed. Though at first, it may appear as if the filter cannot operate in this wind speed region, the distribution of outliers in this region (plot on the top right) shows that the number of outliers is almost an order of magnitude smaller after the filtering process, indicating that the filter can operate efficiently in this region but the low sample density in the region does not reflect this capability of the filter in the PD and FAR metrics.

Finally, the total wind speed distribution of the dataset before and after applying the filter is plotted in Fig.6.11. After applying the filter, $\sim 20.5\%$ of the data have been removed by the filter as outliers. From Fig.6.11 it can be observed that the largest

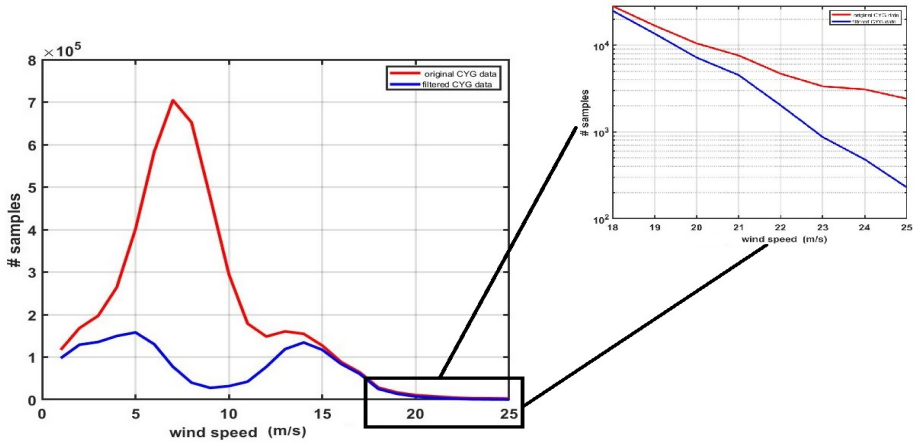


Figure 6.10: Distribution of outliers at different MERRA-2 wind speed bins before and after QC filter.

difference in density occurs at high wind speeds ($> 18m/s$). This is partly due to the high FAR of the filter in this region and partly due to large scatter in the data in this region. A substantial difference in density can also be observed at very low wind speed regions ($< 3m/s$), again owing to the high FAR of the filter in this region.

6.3.2 Identifying dominant feature vectors

In this section the importance of each of the diagnostic variables is assessed using the minimum redundancy maximum relevance algorithm. The algorithm minimizes the redundancy of the feature set and maximizes the set with respect to the training data. Pairwise mutual information of the diagnostic variables is used to quantify its redundancy and relevance (*Darbellay and Vajda*, 1999). Fig.6.12 shows the score for each of the variables based on its importance in distinguishing outliers from good samples.

The most dominant feature is the DDMA (NBRCS). This is as expected because the wind retrieval by CYGNSS is directly related to the two observables NBRCS and LES. The other dominant features are predominantly instrument related such as

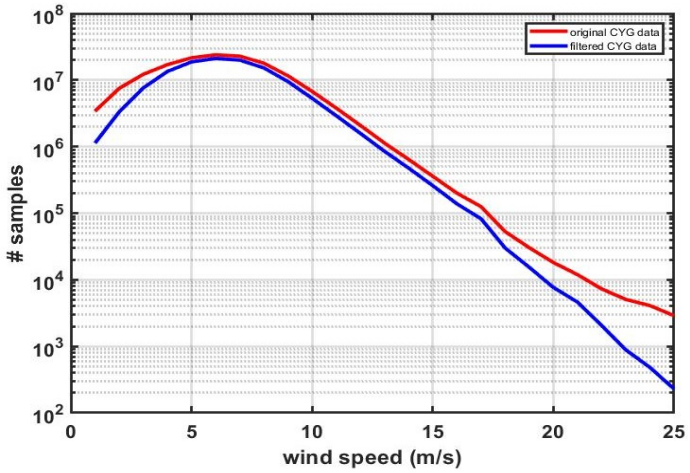


Figure 6.11: Distribution of CYGNSS retrieved winds before and after QC filter.

azimuth angle, PRN, star tracker attitude status and satellite roll. This suggests that most of the outliers are caused by improper instrument calibration.

6.3.3 Wind retrieval performance

The CYGNSS wind retrieval is evaluated based on 3 error statistics, namely, the Mean Difference (MD), RMS difference (RMSD) and variance in the data. The MD and RMSD are evaluated with respect to the 1:1 line which represents a superposition of both variance in the data and the intrinsic bias in the GMF, whereas the variance is a measure of only the degree of scatter in the data. The error statistics are presented in Fig.6.13 and Fig.6.14. In Fig.6.13 the MD and RMSD of the original dataset is shown by solid lines and the filtered dataset is shown by dashed lines. An increase in bias can be observed in the filtered dataset as compared to the original dataset; this is because, after filtering, the samples that are identified as good by the filter are aligned closer to the GMF line rather than the 1:1 line. The increase in bias is more dominant above 10 m/s as the GMF line begins to deviate away from the 1:1 line above this wind speed. Fig.6.14 shows the variance in the data at different wind speed bins. Variance represents the degree of scatter in the data and after applying

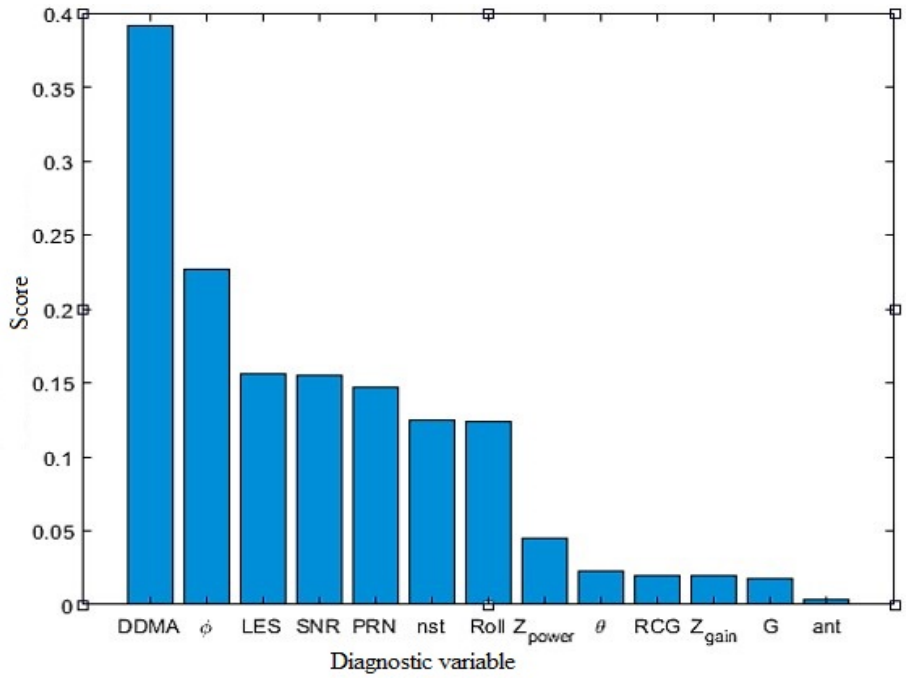


Figure 6.12: Dominant diagnostic variables in identifying outliers. Variable definitions are provided in Table 6.1.

the filter there is a sharp drop in the scatter. The standard deviation in the filtered dataset is <= 2m/s for a wide range of wind speeds. These error statistics show a significant improvement in the nature of retrieval after the QC filter.

6.4 Discussion

The CYGNSS retrieved winds are currently being used for various ocean science applications such as ocean circulation studies, regional and global analysis of ocean winds (*Leidner et al.*, 2018), tropical cyclone studies (*McNoldy et al.*, 2017; *Annane et al.*, 2018; *Li et al.*, 2020; *Mayers and Ruf*, 2019), and assimilation into Numerical Weather Prediction (NWP) models. Data reliability plays an important role in aiding such scientific studies. The CYGNSS wind speed data products are of two kinds - the Fully Developed Seas (FDS) wind retrievals and Young Sea Limited Fetch (YSLF) retrievals. Of these two, the FDS winds form the major proportion of the measurements

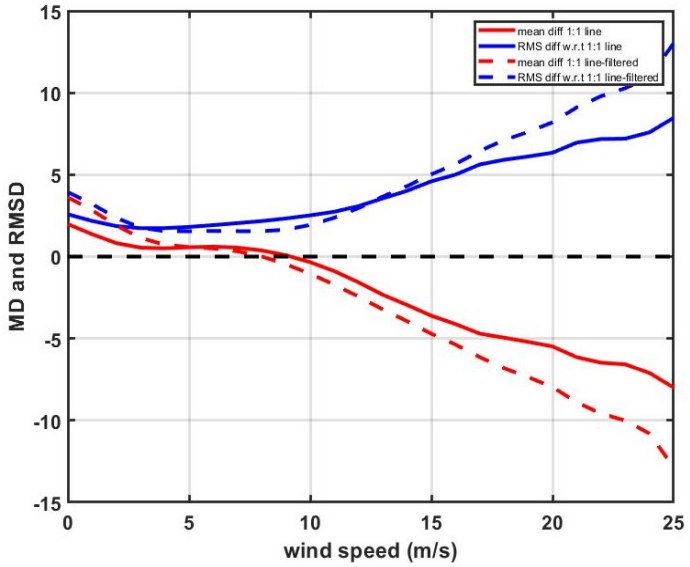


Figure 6.13: Mean difference and RMS difference statistic on CYGNSS retrieved winds before and after QC filter.

and are therefore used for many scientific applications, especially for assimilation into NWP models. The YSLF data product is for hurricane force winds measured over individual storms and therefore is a substantially smaller set of measurements. The QC mechanism developed in this work is for the CYGNSS FDS winds, in order to reduce errors (in particular, outliers) in the retrieval due to various engineering and measurement related errors.

The primary merit of the proposed ML filter is its ability to better account for interactions between the individual engineering, instrument and measurement conditions than can separate thresholded flags for each one. The current approach upon which we are improving uses individual flags and, despite these existing QC filters, there remains considerable scatter in the data hinting that individual and independent thresholds is not an effective way of removing the outliers.

The filter proposed here utilizes the capability of ML tools to learn inherent patterns from the training dataset and quickly come up with any convex boundaries separating the outliers from good data. One other advantage of such filters is that, because the

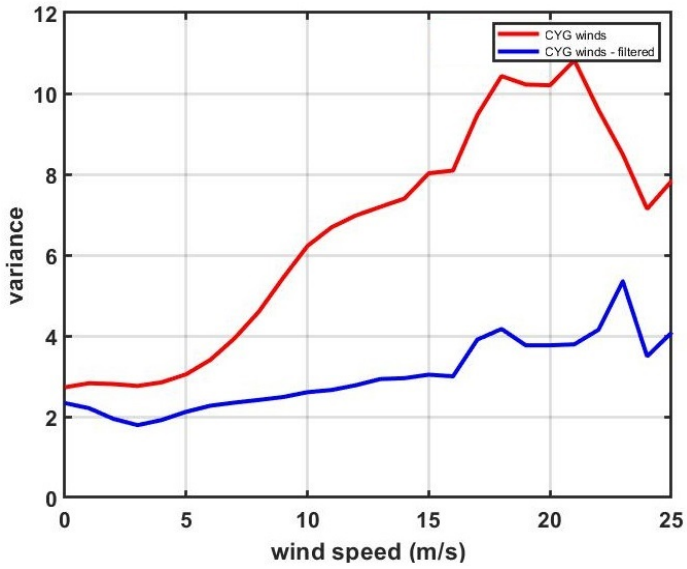


Figure 6.14: Variance in CYGNSS retrieved winds before and after QC filter.

system itself is aging with time, and as shown in this work most of the outliers are due to calibration errors, the new ML-based QC thresholds can be reassessed periodically. In such situations, the ML filters come in handy as their parameters can be tuned easily to respond to any changes.

Assimilating the CYGNSS near surface wind retrievals into NWP models for better forecasting is one of its important uses. In general, NWP models give a weight to meteorological satellite observations based on their error statistics. Thus, reducing errors in the retrieval will help assimilate CYGNSS winds better. Using this filter, the standard deviation of the retrieval is reduced from 2.6 m/s to 1.7 m/s over the wind speed range 0-25 m/s.

At higher wind speed ranges, this filter is too aggressive and removes some valuable high wind measurements. This is due to the fact that high wind data density ($> 20m/s$) is very sparse, hence insufficient for the Neural Network to be able to learn significant patterns from it. To address this situation, one possibility is to assimilate more of CYGNSS high wind data in future years, to better train the Neural Network in this region. However, it is also important to consider here that the CYGNSS FDS

winds are reliable only up to 25 m/s as they have been developed using NOAA/GDAS ocean surface winds as their reference (*Ruf and Balasubramaniam*, 2018).

The direction of focus of future work will be to develop automated machine learning based QC that can effectively remove outliers at all wind speed ranges. Currently this filter is operable only between 5-18 m/s. The lower sample density at high and very low winds, prevent the QC filter from operating in these regions. One possible solution, as mentioned above, is to wait for more CYGNSS measurements in these wind speed regions before the QC is applied. Using ML based QC for YSLF winds can be complicated by the rapidly varying sea state inside hurricanes. In such cases, a physics based definition of an outlier might be needed. One approach to apply quality control for such data is to observe trends along overlapping tracks within a given spatial boundary around the hurricane.

6.5 Conclusion

In this work a Neural Network based Quality Control filter for CYGNSS wind retrieval is developed. The inputs to this filter are the 13 diagnostic variables that broadly represent instrument related, measurement geometry related and surface related attributes. Of these diagnostic tools, the surface related attributes (NBRCS, LES, and SNR) and instrument related attributes (azimuth angle, star tracker status, PRN, satellite roll) play a dominant role in distinguishing outliers from good sample population. The Neural Network is trained over two different training datasets at two different CYGNSS wind regimes based on the behavior of the GMF. The operating range of the filter is between 5-18 m/s. Within this range the probability of outlier detection is $> 75\%$ and the false alarm rates is $< 20\%$. In total $\sim 20.5\%$ of the data is removed as outliers by this filter. At least 75% of the outliers with wind speed difference of at least 5 m/s are removed while $\sim 100\%$ of the outliers with wind speed

difference of at least 10 m/s are removed. This filter has significantly reduced the scatter in the data. The quality filtered dataset has a std. deviation of <= 2 m/s over a wind range of wind speeds. The design space for this filter is also analyzed in this work to identify trade-offs between PD and FAR. The choice of PD and FAR will depend on the application. For example, a low FAR may be especially important for applications in which good spatial and temporal sampling are very important (e.g. to image rapidly changing weather systems) whereas a high PD may be especially important for applications in which the lowest possible uncertainty in wind speed is important (e.g. to detect small trends over long time intervals, such as are associated with global change).

As the next steps in this work, the higher wind regime will be the focus of interest. Strategies to improve the performance of this filter at higher winds while retaining as many samples as possible will be considered. Currently, this filter is developed only for fully developed seas, in the future the feasibility of extending to young seas will also be studied.

CHAPTER VII

Summary and Future Work

The general objective of this thesis is to understand how variations in sea surface roughness affect GNSS-R measurements. This work began with calibrating the CYGNSS system for accurate wind retrieval and development of a Geophysical Model Function (GMF) to retrieve near surface winds under 2 different sea states, namely fully developed seas and young seas with limited fetch. Next, the impact of rain on the GNSS-R signal was characterized to address questions such as, what are the possible effects of rain on the reflected signal and which components dominate in the measurements. As a part of this study we also addressed the impact of the geometric optics based fixed wavenumber cut-off on the derived MSS at very low wind speed (near specular regime). Next, we looked at how the signal modulates with azimuthal location inside hurricanes and we identified strong patterns in the variation which correlated with the degree of sea state development and with its motion. Finally, we developed intelligent quality control filters to improve the performance of CYGNSS retrieved winds under fully developed sea condition. While these research topics might seem to be independent of each other at first look, their findings have helped identify how the GNSS-R measurements behave under individual aspects of a hurricane, which include high winds, heavy precipitation, and varying sea state. They also form the foundational blocks on which future GNSS-R systems can be designed with

a better understanding of what to expect from the reflected signals under different meteorological conditions.

In the following subsections, the novel contributions of this thesis in each of the subtopics discussed above will be summarized; also, a brief discussion of different aspects where this work can be improved further is presented. Finally, a list of publications generated from this work is mentioned at the end of the chapter.

7.1 GMF as a function of sea state

In Chapter 3, Geophysical Model Functions (GMFs) are developed which map the Level 1 observables made by the Cyclone Global Navigation Satellite System (CYGNSS) radar receivers to ocean surface wind speed. The observables are: 1) the normalized bistatic radar cross section (σ_0) of the ocean surface; and 2) the slope of the leading edge of the radar return pulse scattered by the ocean surface. GMFs are empirically derived from measurements by CYGNSS which are nearly coincident with independent estimates of the 10 meter referenced ocean surface wind speed (u10). Two different sources of "ground truth" wind speed are considered - numerical weather prediction model outputs and measurements by the NOAA P-3 hurricane hunter during eyewall penetrations of major hurricanes. The GMFs derived in each case have significant differences that are believed to result from differences in the state of development of the long wave portion of the ocean surface height spectrum that result from characteristic differences in wave age and fetch length near vs. far from the center of a hurricane.

The proposed GMFs current implementation challenges for a wind speed retrieval algorithm in terms of deciding which one to use and under what conditions. A more accurate and a generic model for a GMF could possibly use a fetch-dependent parametrization of the YSLF GMF, or it might modify the L1 observables based on

sea age or fetch length in order to estimate an effective FDS values. These are possible improvements that are under consideration for future development of the GMF in particular and the wind speed retrieval algorithm in general.

Another alternative approach might be to directly assimilate the L1 measurements into a coupled wind/wave model that is able to predict GNSS-R measurements given a known sea-state using an appropriate rough surface scattering model. However direct data assimilation will have its inherent difficulties in terms of physical assumptions and computational complexities.

7.2 GNSS-R for precipitation measurements ?

In Chapter 4 of the thesis, we proposed a 3-fold rain model that accounts for attenuation, surface effects of rain and rain induced local winds. This work utilized a large dataset of measurements made by the CYGNSS mission to individually evaluate these different effects of rain. The attenuation model suggested that a total of at least 96% transmissivity exists at L-Band up to a rain rate of 30 mm/h. This points to the fact that the attenuation component of rain have a very small/negligible impact on the signal strength. Next, a perturbation model was used to characterize the other 2 rain effects. It suggested that rain is accompanied by an overall reduction in the scattering cross-section of the ocean surface and, most importantly, this effect is observed only up to surface wind speeds of 15 m/s, beyond which the gravity capillary waves dominate the scattering in the quasi-specular direction. The results also suggested that, at very low wind speeds, the lower bound on wavenumber for the surface roughness spectrum that influences the measurements, deviates from the geometric optics approximation normally used. This work has taken a birds-eye view of several rain-related phenomena and has enhanced our overall understanding of rain effects on GNSS-R measurements.

The idea that all 3 of the effects of rain result in a net reduction in the radar cross-section can be used for purposes of the remote sensing of rain. Some other parallel research on this topic (*Asgarimehr et al.*, 2018), have discussed the possibility of using this signature to detect precipitation or other convective systems using GNSS-R measurements and have also reasoned that one way of detecting precipitation is to identify a shift in scattering mechanism from coherent to incoherent under low wind speed conditions. An interesting possible future study in this regard would be to try to use CYGNSS measurements to detect the presence of rain under low wind conditions.

7.3 Characterizing the nature of sea state inside hurricanes

Global Navigation Satellite System-Reflectometry measurements of the ocean surface are sensitive to roughness scales ranging from a few cms to several kms. Inside a hurricane the surface roughness changes drastically due to varying sea age and fetch length conditions and complex wave-wave interactions caused by its cyclonic rotation and translational motion. As a result, the relationship between the surface roughness at different scale sizes becomes azimuthally dependent, as does the relationship between scattering cross-section and wind speed as represented by the Geophysical Model Function (GMF) developed in Chapter 3. In Chapter 5 of this thesis, the impact of this azimuthal variation on the scattering cross-section was assessed. An empirical harmonic GMF was constructed using measurements by CYGNSS matched to HWRF reanalysis surface winds for 19 hurricanes in 2017 and 2018. The analysis revealed a 2-8% variation in scattering cross-section due to azimuthal location, and the magnitude of the azimuthal dependence was found to grow with wind speed.
The fact that the variation in the measured σ_0 vs. wind speed in the 4 different quadrants inside a hurricanes is consistent with the azimuthal variation of the local

wind wave directions in published directional wavenumber spectrum datasets suggest that the GNSS-R measurements can be used to study sea state development as well. The proposed azimuthally dependent GMF has several utilities in this regard. It can serve as an indicator of sea state development in the inner core of tropical cyclones, for use in process studies into air-sea and wind-wave interactions. It could also be used to improve wind speed retrieval algorithms in tropical cyclones that are based on GNSS-R observations. Retrieval algorithms essentially invert the GMF to estimate wind speed given the scattering crosssection, and a more physically representative GMF will allow for a more accurate inversion. Furthermore, this type of study can be extended to other non-TC weather systems with younger seas or scenarios with limited fetch conditions, for example, limited fetch on the lee side of major islands that may result in a different relationship between wind speed and MSS or σ_0 and therefore require a modified GMF for accurate wind speed retrieval.

7.4 Using ML filters for high winds

The CYGNSS mission aims at providing high quality global scale GNSS-R measurements that can be reliably used for scientific applications. To achieve this goal, strong quality control filters are needed which can detect and remove outlier measurements. In Chapter 6 of this thesis we developed a Neural Network based quality control filter for automated outlier detection of CYGNSS retrieved winds for fully developed seas. This filter uses various features that represent instrument, measurement geometry and surface related attributes to detect outliers. The proposed filter has a probability of outlier detection (PD) $> 75\%$ and False Alarm Rate (FAR) $< 20\%$ within a wind speed range of 5 m/s to 18 m/s. At least 75 % of the outliers with wind speed difference of at least 5 m/s are removed while $\sim 100\%$ of the outliers with wind speed difference of at least 10 m/s are removed. This quality filter was shown to signifi-

cantly improve the data quality. The standard deviation of wind speed retrieval error was reduced from 2.6 m/s without the filter to 1.7 m/s with it over a wind speed range of 0 to 25 m/s. The design space for this filter was also analyzed in this work to characterize trade-offs between PD and FAR.

The fact that such ML based filters have very good skill to detect and remove outliers motivates us to extend their application to more difficult datasets, such as the high wind measurements made inside hurricanes. The high wind measurements form a much smaller dataset and there is a possibility for the radar cross-section to go below the system noise level at very high winds, making them more prone to calibration errors. In addition, the low sensitivity of GNSS-R observations at high winds can lead to significant retrieval errors. All of these factors can result in a large number of outliers at very high wind speeds. Another important point to note is that some measurements might appear as outliers but can be an actual representation of the sea state as σ_0 is a more direct measure of sea state than wind speed, as seen in Chapter 5. Hence, for high wind applications, ML filters have to be devised carefully to take into account the above considerations and assess if they are beneficial under hurricane conditions as well.

7.5 Summary of research publications

7.5.1 Journal publications

Ruf, C. S., and Balasubramaniam, R. "Development of the CYGNSS geophysical model function for wind speed." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 12.1 (2018): 66-77.

Ruf, C. S., Chew, C., Lang, T., Morris, M. G., Nave, K., Ridley, A., and Bala-

subramaniam, R. (2018). A new paradigm in earth environmental monitoring with the CYGNSS small satellite constellation. Scientific reports, 8(1), 8782.

Ruf, C., Asharaf, S., Balasubramaniam, R., Gleason, S., Lang, T., McKague, D., and Waliser, D. (2019). In-Orbit Performance of the Constellation of CYGNSS Hurricane Satellites. Bulletin of the American Meteorological Society, (2019).

Balasubramaniam, Rajeswari, and Christopher Ruf. "Characterization of rain impact on L-Band GNSS-R ocean surface measurements." Remote Sensing of Environment 239 (2020): 111607.

Balasubramaniam, Rajeswari, and Christopher Ruf. "Azimuthal Dependence of GNSSR Scattering CrossSection in Hurricanes." Journal of Geophysical Research: Oceans 125.7 (2020): e2020JC016167.

Balasubramaniam, Rajeswari., and Christopher Ruf. Neural Network based quality control of CYGNSS wind retrieval Remote Sensing. (2020); 12(17):2859.

7.5.2 Conference presentations

Balasubramaniam, Rajeswari, et al. "Calibration and validation processing for the CYGNSS wind speed retrieval algorithm." 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, 2017.

Clarizia, M. P., Ruf, C. S., Gleason, S., Balasubramaniam, R., and McKague, D. (2017, July). Generation of cygnss level 2 wind speed data products. In 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (pp. 2647-2649). IEEE.

Balasubramaniam, R., Ruf, C. S., McKague, D. S., Clarizia, M. P., and Gleason, S. (2017, December). Early Calibration Results of CYGNSS Mission. In AGU Fall Meeting Abstracts.

McKague, D. S., Balasubramaniam, Rajeswari, and Christopher S. Ruf. "PMM Observations of Tropical Cyclones: PMM/CYGNSS Synergies." PMM Science Team Meeting, 2017.

Balasubramaniam, Rajeswari, and Christopher S. Ruf. "Early Assessment of CYGNSS Level 2 wind speed data product.GNSS+R specialist Meeting, 2017.

Balasubramaniam, Rajeswari, and Christopher S. Ruf. "Improved Calibration of CYGNSS Measurements for Downbursts in the Intertropical Convergence Zone." IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2018.

Balasubramaniam, Rajeswari, and Christopher S. Ruf. "High Wind Retrieval in Hurricanes Using CYGNSS Measurements." 69th International Astronautical Congress, Bremen, Germany, 2018.

Balasubramaniam, Rajeswari, and Christopher S. Ruf. "The impact of rain on L1 GNSS-R Radar Scattering Cross-section." IGARSS 2019 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2019.

Balasubramaniam, Rajeswari, and Christopher S. Ruf. "CYGNSS small satellite constellation mission for ocean science application." 26th IAA Symposium on Small

Satellite Missions, 70th International Astronautical Congress, Washington DC, US, 2019.

McKague, D. S., Christopher S. Ruf., and Balasubramaniam, Rajeswari. "Status of CYGNSS Level 2 winds." 100th American Meteorological Society Annual Meeting, Boston, US, 2020.

Balasubramaniam, Rajeswari, and Christopher S. Ruf. Azimuthal Dependence of Sea State development inside Tropical Cyclones as measured by CYGNSS." 100th American Meteorological Society Annual Meeting, Boston, US, 2020.

Balasubramaniam, Rajeswari, and Christopher S. Ruf. "Performance assessment of CYGNSS wind retrieval with improved GPS EIRP tracking." Virtual IEEE International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, 2020.

APPENDICES

APPENDIX A

Relationship Between Downdraft Winds and Final Velocity of Falling Rain Drops

Following is a derivation from (*Holleman*, 2001) of the relationship between downdraft winds and the terminal velocity of a falling raindrop, which is used in Chapter to understand the downdraft effect of rain on GNSS-R measurements.

The total vertical acceleration of an air parcel can be written using the vertical momentum equation as:

$$\frac{dw}{dt} \equiv \frac{\partial w}{\partial t} + w\frac{\partial w}{\partial z} \approx w\frac{\partial w}{\partial z} = \frac{1}{2}\frac{\partial w^2}{\partial z} \tag{A.1}$$

The time derivative can be neglected based on the assumption that the air parcel is stationary on the time scale of the downdraft. The effect of precipitation loading on the vertical motion of the air parcel is parameterized by the term gL, where g is the accelaration due to gravity and L is the precipitation mixing ratio (*Emanuel*, 1981; *Wolfson*, 1990; *Doswell*, 1994).

Downdraft wind velocity is defined as the vertical velocity at the surface (w_0) which is deflected by the surface in the horizontal direction. Let the downdraft start at height

H above the surface, the velocity at the surface can be obtained from the integrated vertical momentum equation as given by :

$$\begin{aligned} \frac{1}{2}\frac{\partial w^2}{\partial z} &= -gL \\ w^2(z) = 2g\int_0^H Ldz &= \frac{2g}{\rho_0}\int_0^H H\rho_0 Ldz \\ w^2(z) = \frac{2gMH}{\rho_0} &= \frac{2gRH}{3600\rho_0 v_f} = 5.63\frac{RH}{v_f} \end{aligned} \tag{A.2}$$

where M is the precipitation loading per unit volume and ρ_0 is the vertically averaged air density. This averaged density is approximated as $0.968Kg/m^3$, which is the value for a standard atmosphere between 0 and $5km$. The maximum downdraft velocity can therefore be written as :

$$w_0 = \sqrt{5.63\frac{RH}{v_f}} \tag{A.3}$$

APPENDIX B

Controlled CYGNSS Dataset for Observing Precipitation Effects

In this section we describe the controlled dataset of 110 CYGNSS tracks used in Chapter 4. Each track can be uniquely identified by the DOY (Day Of Year) for the year 2017, FM (CYGNSS satellite ID), PRN (GPS satellite ID) and time range. We have also added additional information such as location of peak precipitation regions and locations of downdraft points for each track. Each row is a CYGNSS track and has 9 fields : DOY, FM, PRN, time_upper(UTC secs), time_lower(UTC secs), loc_pmax, loc_pmin, loc_DDpt1, locDDpt2. The details of these fields are given below:

DOY - Day Of Year in 2017

FM - CYGNSS satellite ID

PRN - GPS satellite ID

time_lower(UTC secs) - the start time of the CYGNSS track at UTC time.

time_upper(UTC secs)- end time of the CYGNSS track at UTC time.

loc_pmax - sample location along track with max precipitation (1 is the starting sample of the track).

loc_pmin - sample location along track with zero precipitation (1 is the starting

sample of the track).

loc_DDpt1 - closest sample location to peak precipitation with minimum precipitation for first downdraft measurement (1 is the starting sample of the track).

loc_DDpt2 - closest sample location to peak precipitation with minimum precipitation for second downdraft measurement (1 is the starting sample of the track).

DOY	FM	PRN	time_lower (UTC secs)	time_upper (UTC secs)	loc_pmax	loc_pmin	loc_DDpt1	loc_DDpt2
168	8	25	33238	34107	108	138	73	115
168	3	25	33073	33912	281	244	256	291
164	4	8	11081	11948	362	408	359	371
159	8	18	7865	8874	23	80	4	41
155	2	24	46275	46904	581	498	567	589
146	4	25	30739	31588	332	372	298	360
145	5	12	23407	24636	352	390	323	381
145	3	12	23069	24338	381	421	351	408
144	6	19	40471	41100	94	119	84	111
143	6	12	72061	72730	297	320	278	304
143	6	25	77710	78339	327	359	313	344
142	7	19	41439	41998	17	74	3	27
142	6	25	78641	79210	271	296	258	283
142	5	19	41157	41756	41	58	31	52
140	4	25	81449	82128	211	184	197	233
140	1	21	60367	61046	58	95	46	77
138	5	17	4486	5225	450	494	431	464
135	8	29	57933	58302	133	192	122	142
135	5	20	58457	59206	435	411	425	444

133	1	10	45997	46826	646	703	641	656
133	1	31	12296	13335	279	373	272	287
124	1	22	46387	47356	804	877	767	839
124	1	18	21146	21975	99	79	88	164
122	4	19	67059	67838	695	664	685	710
121	3	12	31913	32652	443	423	436	461
121	1	12	31766	32475	440	404	414	452
120	3	3	61582	62591	160	150	154	174
118	8	22	30480	31379	345	301	320	384
118	1	7	57171	57615	285	364	275	331
117	2	24	32760	34099	796	776	786	818
105	1	15	32477	32706	59	112	51	78
102	2	14	58270	59019	444	413	437	488
264	4	8	36608	36957	37	27	34	48
263	5	12	46732	47501	54	16	41	72
262	6	17	53991	54760	146	195	129	179
262	4	24	46397	47126	651	637	646	660
262	3	24	75907	76452	431	468	425	437
261	5	10	54032	54801	216	237	212	229
258	3	17	10024	10522	95	143	81	112
258	2	17	10403	10802	169	212	151	184
257	1	8	4645	5284	356	375	353	262
255	3	29	29772	30501	227	259	203	236
254	6	17	15351	16150	364	377	345	372
253	6	23	30071	30930	393	371	381	410
252	1	32	62795	63714	230	268	223	243

251	2	10	58814	59673	500	518	490	507
249	7	21	7746	8685	359	393	315	372
248	7	21	48505	49184	246	270	225	264
247	2	24	43953	44482	367	407	354	385
245	6	21	32601	33380	154	180	124	171
245	4	8	27788	28597	89	113	85	103
245	2	2	28144	29023	189	225	147	211
243	3	13	35153	36771	637	672	613	650
237	3	27	43597	44682	527	540	512	534
232	4	11	12458	13037	151	206	134	188
231	8	10	15030	15189	91	72	84	109
231	6	10	14821	15010	93	78	83	107
231	3	25	58543	59242	266	239	259	316
223	2	20	0	492	259	307	241	299
170	5	31	77832	78541	537	553	523	546
179	7	5	63187	63536	202	220	201	215
179	8	15	4062	4751	194	262	191	214
180	4	22	10199	10498	219	242	214	233
187	5	26	41267	42141	273	259	263	295
188	1	25	75265	76114	253	325	234	265
188	5	26	34131	34970	633	610	619	649
188	6	26	33773	34612	635	609	625	648
194	7	8	45471	46066	259	285	240	279
194	7	23	45067	45956	684	711	665	698
194	8	8	46074	46653	180	211	167	199
198	1	29	7350	8359	547	580	542	555

199	2	5	43554	44343	470	497	460	479
199	3	5	44332	44961	507	537	497	519
206	2	1	13224	14033	158	187	130	176
208	5	11	21802	22391	131	189	118	153
207	6	3	14375	15163	590	638	582	637
208	7	5	75456	76265	101	54	81	114
105	1	15	32477	32706	59	112	51	78
113	1	3	61757	62266	417	464	388	431
114	2	28	83691	84420	381	437	373	403
113	3	28	67783	68562	694	646	680	712
116	3	29	68753	70032	222	192	221	270
117	3	20	75653	76332	140	184	124	178
118	1	22	51607	52526	424	451	404	440
118	3	13	68833	69052	147	211	109	194
118	7	13	69369	70218	34	119	19	102
118	8	10	59339	59608	89	128	74	108
120	4	12	35378	36137	15	88	4	51
121	8	10	16780	17429	64	100	49	91
127	1	31	83627	84336	161	206	151	175
132	7	20	22559	23548	114	62	108	153
140	1	25	79667	80306	75	114	56	93
140	3	9	42618	43277	325	361	315	332
140	5	20	53587	54446	463	528	456	481
141	3	24	58359	58978	454	513	433	464
140	5	9	42736	43425	326	362	325	338
141	1	7	42437	42926	361	391	355	370

141	4	16	26399	27038	431	451	405	438
142	7	19	41439	41998	17	74	8	27
142	4	3	11449	12198	94	52	79	157
142	5	19	41157	41756	41	58	31	52
142	6	30	48101	48480	308	295	299	332
142	6	25	78641	79210	271	296	258	283
142	8	29	17544	18183	68	94	64	74
143	3	17	59748	60727	66	126	52	108
143	6	25	77710	78339	327	359	313	344
143	6	12	72061	72730	297	320	278	304
143	6	20	71801	72540	574	597	557	582
143	7	17	60168	61207	102	173	81	122
143	6	18	18201	19140	322	382	304	357

Table B.1: Controlled CYGNSS dataset

BIBLIOGRAPHY

BIBLIOGRAPHY

Aggarwal, C. C., and S. Sathe (2017), *Outlier ensembles: An introduction*, Springer.

Andersson, E., A. Persson, and I. Tsonevsky (2015), User guide to ecmwf forecast products, *ECMWF, v2*, *1*, 121.

Annane, B., B. McNoldy, S. M. Leidner, R. Hoffman, R. Atlas, and S. J. Majumdar (2018), 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*, *146*(7), 2221–2236.

Apel, J. R. (1994), An improved model of the ocean surface wave vector spectrum and its effects on radar backscatter, *Journal of Geophysical Research: Oceans*, *99*(C8), 16,269–16,291.

Asgarimehr, M., V. Zavorotny, J. Wickert, and S. Reich (2018), Can GNSS Reflectometry detect precipitation over oceans?, *Geophysical Research Letters*, *45*(22), 12–585.

Asgarimehr, M., J. Wickert, and S. Reich (2019), Evaluating impact of rain attenuation on space-borne gnss reflectometry wind speeds, *Remote Sensing*, *11*(9), 1048.

Atlas, D., and C. W. Ulbrich (1977), Path-and area-integrated rainfall measurement by microwave attenuation in the 1–3 cm band, *Journal of Applied Meteorology*, *16*(12), 1322–1331.

Balasubramaniam, R., and C. S. Ruf (2018), Improved calibration of CYGNSS measurements for downbursts in the Intertropical Convergence Zone, in *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, pp. 3987–3990, IEEE.

Bass, F., I. Fuks, A. Kalmykov, I. Ostrovsky, and A. Rosenberg (1968), Very high frequency radiowave scattering by a disturbed sea surface part II: Scattering from an actual sea surface, *IEEE Transactions on Antennas and Propagation*, *16*(5), 560–568.

Bliven, L., P. Sobieski, and C. Craeye (1997), Rain generated ring-waves: Measurements and modelling for remote sensing, *International Journal of Remote Sensing*, *18*(1), 221–228.

Bosilovich, M., R. Lucchesi, and M. Suarez (), Merra-2: File specification. 2015.

Brown, G. S. (1990), Quasi-specular scattering from the air-sea interface, in *Surface waves and fluxes*, pp. 1–39, Springer.

Cardellach, E., G. Ruffini, D. Pino, A. Rius, A. Komjathy, and J. L. Garrison (2003), Mediterranean balloon experiment: Ocean wind speed sensing from the stratosphere, using gps reflections, *Remote Sensing of Environment*, *88*(3), 351–362.

Chang, P. S., and L. Li (1998), Ocean surface wind speed and direction retrievals from the ssm/i, *IEEE transactions on geoscience and remote sensing*, *36*(6), 1866–1871.

Chelton, D. B., and P. J. McCabe (1985), A review of satellite altimeter measurement of sea surface wind speed: With a proposed new algorithm, *Journal of Geophysical Research: Oceans*, *90*(C3), 4707–4720.

Clarizia, M., V. Zavarotny, and C. Ruf (), Level 2 wind speed retrieval algorithm theoretical basis document, *CYGNSS Project Document*, pp. 148–0138.

Clarizia, M., C. Gommenginger, S. Gleason, M. Srokosz, C. Galdi, and M. Di Bisceglie (2009), Analysis of GNSS-R Delay-Doppler maps from the UK-DMC satellite over the ocean, *Geophysical Research Letters*, *36*(2).

Clarizia, M. P., and C. S. Ruf (2016a), On the spatial resolution of gnss reflectometry, *IEEE Geoscience and Remote Sensing Letters*, *13*(8), 1064–1068.

Clarizia, M. P., and C. S. Ruf (2016b), Wind speed retrieval algorithm for the cyclone global navigation satellite system (cygnss) mission, *IEEE Transactions on Geoscience and Remote Sensing*, *54*(8), 4419–4432.

Clarizia, M. P., C. S. Ruf, P. Jales, and C. Gommenginger (2014), Spaceborne gnss-r minimum variance wind speed estimator, *IEEE transactions on geoscience and remote sensing*, *52*(11), 6829–6843.

Contreras, R. F., and W. J. Plant (2006), Surface effect of rain on microwave backscatter from the ocean: Measurements and modeling, *Journal of Geophysical Research: Oceans*, *111*(C8).

Craeye, C., P. Sobieski, and L. Bliven (1997), Scattering by artificial wind and rain roughened water surfaces at oblique incidences, *International Journal of Remote Sensing*, *18*(10), 2241–2246.

Darbellay, G. A., and I. Vajda (1999), Estimation of the information by an adaptive partitioning of the observation space, *IEEE Transactions on Information Theory*, *45*(4), 1315–1321.

Doswell, C. (1994), Extreme convective windstorms: current understanding and research, in *Report of the Proceedings of the US–Spain Workshop on Natural Hazards (Barcelona, Spain, 8–11 June 1993)*, pp. 44–55.

Doviak, R. J., et al. (2006), *Doppler radar and weather observations*, Courier Corporation.

Draper, D. W., and D. G. Long (2004), Simultaneous wind and rain retrieval using SeaWinds data, *IEEE Transactions on Geoscience and Remote Sensing*, *42*(7), 1411–1423.

Droppleman, J. (1970), Apparent microwave emissivity of sea foam, *Journal of Geophysical Research*, *75*(3), 696–698.

Ebuchi, N., H. C. Graber, and M. J. Caruso (2002), Evaluation of wind vectors observed by quikscat/seawinds using ocean buoy data, *Journal of Atmospheric and Oceanic Technology*, *19*(12), 2049–2062.

Elfouhaily, T., B. Chapron, K. Katsaros, and D. Vandemark (1997), A unified directional spectrum for long and short wind-driven waves, *Journal of Geophysical Research: Oceans*, *102*(C7), 15,781–15,796.

Emanuel, K. A. (1981), A similarity theory for unsaturated downdrafts within clouds, *Journal of the Atmospheric Sciences*, *38*(8), 1541–1557.

Fan, Y., I. Ginis, T. Hara, C. W. Wright, and E. J. Walsh (2009), Numerical simulations and observations of surface wave fields under an extreme tropical cyclone, *Journal of Physical Oceanography*, *39*(9), 2097–2116.

Figa-Saldaña, J., J. J. Wilson, E. Attema, R. Gelsthorpe, M. R. Drinkwater, and A. Stoffelen (2002), The advanced scatterometer (ascat) on the meteorological operational (metop) platform: A follow on for european wind scatterometers, *Canadian Journal of Remote Sensing*, *28*(3), 404–412.

Foti, G., C. Gommenginger, P. Jales, M. Unwin, A. Shaw, C. Robertson, and J. Rosello (2015), Spaceborne gnss reflectometry for ocean winds: First results from the uk techdemosat-1 mission, *Geophysical Research Letters*, *42*(13), 5435–5441.

Foti, G., C. Gommenginger, and M. Srokosz (2017a), First spaceborne gnss-reflectometry observations of hurricanes from the uk techdemosat-1 mission, *Geophysical Research Letters*, *44*(24), 12–358.

Foti, G., C. Gommenginger, M. Unwin, P. Jales, J. Tye, and J. Roselló (2017b), An assessment of non-geophysical effects in spaceborne gnss reflectometry data from the uk techdemosat-1 mission, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(7), 3418–3429.

Fung, A., and K. Lee (1982), A semi-empirical sea-spectrum model for scattering coefficient estimation, *IEEE Journal of Oceanic Engineering*, *7*(4), 166–176.

Gaiser, P. W., et al. (2004), The windsat spaceborne polarimetric microwave radiometer: Sensor description and early orbit performance, *IEEE Transactions on Geoscience and Remote Sensing*, *42*(11), 2347–2361.

Garrison, J. L., S. J. Katzberg, and M. I. Hill (1998), Effect of sea roughness on bistatically scattered range coded signals from the global positioning system, *Geophysical research letters*, *25*(13), 2257–2260.

Garrison, J. L., A. Komjathy, V. U. Zavorotny, and S. J. Katzberg (2002), Wind speed measurement using forward scattered gps signals, *IEEE Transactions on Geoscience and Remote Sensing*, *40*(1), 50–65.

Gelaro, R., et al. (2017), The modern-era retrospective analysis for research and applications, version 2 (merra-2), *Journal of Climate*, *30*(14), 5419–5454.

Ghavidel, A., and A. Camps (2016), Impact of rain, swell, and surface currents on the electromagnetic bias in GNSS-Reflectometry, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(10), 4643–4649.

Glazman, R. E., and A. Greysukh (1993), Satellite altimeter measurements of surface wind, *Journal of Geophysical Research: Oceans*, *98*(C2), 2475–2483.

Gleason, S. (2013), Space-based gnss scatterometry: Ocean wind sensing using an empirically calibrated model, *IEEE transactions on geoscience and remote sensing*, *51*(9), 4853–4863.

Gleason, S. (2014), Level 1b ddm calibration algorithm theoretical basis document, *CYGNSS Project Document*, pp. 148–0137.

Gleason, S., S. Hodgart, Y. Sun, C. Gommenginger, S. Mackin, M. Adjrad, and M. Unwin (2005), Detection and processing of bistatically reflected gps signals from low earth orbit for the purpose of ocean remote sensing, *IEEE Transactions on Geoscience and Remote Sensing*, *43*(6), 1229–1241.

Gleason, S., D. Gebre-Egziabher, and D. G. Egziabher (2009), Gnss applications and methods.

Gleason, S., C. S. Ruf, M. P. Clarizia, and A. J. O'Brien (2016), Calibration and unwrapping of the normalized scattering cross section for the cyclone global navigation satellite system, *IEEE Transactions on Geoscience and Remote Sensing*, *54*(5), 2495–2509.

Gleason, S., C. S. Ruf, A. J. OBrien, and D. S. McKague (2018a), 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*, *12*(1), 37–49.

Gleason, S., V. U. Zavorotny, D. M. Akos, S. Hrbek, I. PopStefanija, E. J. Walsh, D. Masters, and M. S. Grant (2018b), Study of surface wind and mean square slope correlation in hurricane ike with multiple sensors, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(6), 1975–1988.

Graf, J. E., W.-y. Tsi, and L. Jones (1998), Overview of quikscat mission-a quick deployment of a high resolution, wide swath scanning scatterometer for ocean wind measurement, in *Proceedings IEEE Southeastcon'98'Engineering for a New Era'*, pp. 314–317, IEEE.

Hall, C., and R. Cordey (1988), Multistatic scatterometry, in *International Geoscience and Remote Sensing Symposium,'Remote Sensing: Moving Toward the 21st Century'.*, vol. 1, pp. 561–562, IEEE.

Harris Jr, G. N., K. P. Bowman, and D.-B. Shin (2000), Comparison of freezing-level altitudes from the NCEP reanalysis with TRMM precipitation radar brightband data, *Journal of climate*, *13*(23), 4137–4148.

Hasselmann, K., et al. (1973), Measurements of wind-wave growth and swell decay during the joint north sea wave project (jonswap), *Ergänzungsheft 8-12*.

Hawkins, D. M. (1980), *Identification of outliers*, vol. 11, Springer.

Hersbach, H., A. Stoffelen, and S. de Haan (2007), An improved c-band scatterometer ocean geophysical model function: Cmod5, *Journal of Geophysical Research: Oceans*, *112*(C3).

Holleman, I. (2001), Estimation of the maximum velocity of convective wind gusts, *Internal KNMI report*.

Hollinger, J. P., and R. C. Lo (1983), Ssm/i (special sensor microwave/imager) project summary report, *Tech. rep.*, NAVAL RESEARCH LAB WASHINGTON DC.

Huffman, G. J., D. T. Bolvin, D. Braithwaite, K. Hsu, R. Joyce, P. Xie, and S.-H. Yoo (2015), NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG), *Algorithm theoretical basis document, version*, *4*, 30.

Hwang, P. A., Y. Fan, F. J. Ocampo-Torres, and H. García-Nava (2017), Ocean surface wave spectra inside tropical cyclones, *Journal of Physical Oceanography*, *47*(10), 2393–2417.

Hwang, P. A., Y. Fan, and E. J. Walsh (2018), Hurricane hunter observations of wind and wave spectral properties: Implications on tropical cyclone remote sensing, in *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, pp. 149–152, IEEE.

Jales, P., and M. Unwin (2015), Mission description-gnss reflectometry on tds-1 with the sgr-resi, *Surrey Satellite Technol. Ltd., Guildford, UK, Tech. Rep. SSTL Rep*, *248367*.

Jones, W. L., P. G. Black, V. E. Delnore, and C. T. Swift (1981), Airborne microwave remote-sensing measurements of hurricane allen, *Science*, *214*(4518), 274–280.

Katzberg, S. J., R. A. Walker, J. H. Roles, T. Lynch, and P. G. Black (2001), First gps signals reflected from the interior of a tropical storm: Preliminary results from hurricane michael, *Geophysical Research Letters*, *28*(10), 1981–1984.

Katzberg, S. J., O. Torres, and G. Ganoe (2006), Calibration of reflected gps for tropical storm wind speed retrievals, *Geophysical Research Letters*, *33*(18).

Kawanishi, T., et al. (2003), The advanced microwave scanning radiometer for the earth observing system (amsr-e), nasda's contribution to the eos for global energy and water cycle studies, *IEEE Transactions on Geoscience and Remote Sensing*, *41*(2), 184–194.

Komjathy, A., M. Armatys, D. Masters, P. Axelrad, V. Zavorotny, and S. Katzberg (2004), Retrieval of ocean surface wind speed and wind direction using reflected gps signals, *Journal of Atmospheric and Oceanic Technology*, *21*(3), 515–526.

Kummerow, C., W. Barnes, T. Kozu, J. Shiue, and J. Simpson (1998), The tropical rainfall measuring mission (trmm) sensor package, *Journal of atmospheric and oceanic technology*, *15*(3), 809–817.

Leidner, S. M., B. Annane, B. McNoldy, R. Hoffman, and R. Atlas (2018), 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*, *35*(8), 1571–1584.

Li, X., J. R. Mecikalski, and T. J. Lang (2020), A study on assimilation of cygnss wind speed data for tropical convection during 2018 january mjo, *Remote Sensing*, *12*(8), 1243.

Lippmann, R. (1987), An introduction to computing with neural nets, *IEEE Assp magazine*, *4*(2), 4–22.

Lowe, S. T., J. L. LaBrecque, C. Zuffada, L. J. Romans, L. E. Young, and G. A. Hajj (2002), First spaceborne observation of an earth-reflected gps signal, *Radio Science*, *37*(1), 7–1.

Luntama, J.-P., G. Kirchengast, M. Borsche, U. Foelsche, A. Steiner, S. Healy, A. von Engeln, E. O'Clerigh, and C. Marquardt (2008), Prospects of the eps gras mission for operational atmospheric applications, *Bulletin of the American Meteorological Society*, *89*(12), 1863–1876.

Marshall, and . Mc. K. Palmer (1948), The distribution of raindrops with size, *Journal of Meteorology*, *5*, 165–166.

Masters, D. (2004), Surface remote sensing applications of gnss bistatic radar: Soil moisture and aircraft altimetry, Ph.D. thesis, Citeseer.

Maul, G. A. (2012), *Introduction to satellite oceanography*, vol. 3, Springer Science & Business Media.

Mayers, D., and C. Ruf (2019), Tropical cyclone center fix using cygnss winds, *Journal of Applied Meteorology and Climatology*, *58*(9), 1993–2003.

McNoldy, B., B. Annane, S. Majumdar, J. Delgado, L. Bucci, and R. Atlas (2017), Impact of assimilating cygnss data on tropical cyclone analyses and forecasts in a regional osse framework, *Marine Technology Society Journal*, *51*(1), 7–15.

Mehrotra, K. G., C. K. Mohan, and H. Huang (2017), *Anomaly detection principles and algorithms*, Springer.

Meissner, T., and F. Wentz (2002), An updated analysis of the ocean surface wind direction signal in passive microwave brightness temperatures, *IEEE Transactions on Geoscience and Remote Sensing*, *40*(6), 1230–1240.

Moon, I.-J., I. Ginis, T. Hara, H. L. Tolman, C. Wright, and E. J. Walsh (2003), Numerical simulation of sea surface directional wave spectra under hurricane wind forcing, *Journal of physical oceanography*, *33*(8), 1680–1706.

Nielsen, S. N., and D. G. Long (2008), A wind and rain backscatter model derived from AMSR and SeaWinds data, *IEEE Transactions on Geoscience and Remote Sensing*, *47*(6), 1595–1606.

OBrien, A., S. Gleason, J. Johnson, and C. Ruf (2014), The End-to-End simulator for the Cyclone GNSS (CYGNSS) mission, *IEEE J. Sel. Topics Appl. Earth Obs.*

Pierson Jr, W. J., and L. Moskowitz (1964), A proposed spectral form for fully developed wind seas based on the similarity theory of sa kitaigorodskii, *Journal of geophysical research*, *69*(24), 5181–5190.

PO.DAAC, C. . (2018), Cygnss level 1 science data record version 2.1. ver. 2.1., *PO.DAAC, CA, USA*.

Portabella, M., A. Stoffelen, W. Lin, A. Turiel, A. Verhoef, J. Verspeek, and J. Ballabrera-Poy (2012), Rain effects on ascat-retrieved winds: Toward an improved quality control, *IEEE Transactions on Geoscience and Remote Sensing*, *50*(7), 2495–2506.

Ricciardulli, L., and F. J. Wentz (2015), A scatterometer geophysical model function for climate-quality winds: Quikscat ku-2011, *Journal of Atmospheric and Oceanic Technology*, *32*(10), 1829–1846.

Robinson, I. S. (2010), *Discovering the Ocean from Space: The unique applications of satellite oceanography*, Springer Science & Business Media.

Ruf, C., M. Unwin, J. Dickinson, R. Rose, D. Rose, M. Vincent, and A. Lyons (2013), CYGNSS: Enabling the future of hurricane prediction [remote sensing satellites], *IEEE Geoscience and Remote Sensing Magazine*, *1*(2), 52–67.

Ruf, C., et al. (2016), Cygnss handbook, *Ann Arbor, MI, Michigan Pub*, *154*.

Ruf, C. S., and R. Balasubramaniam (2018), Development of the CYGNSS Geophysical Model Function for wind speed, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, (99), 1–12.

Ruf, C. S., C. Chew, T. Lang, M. G. Morris, K. Nave, A. Ridley, and R. Balasubramaniam (2018a), A new paradigm in earth environmental monitoring with the cygnss small satellite constellation, *Scientific reports*, *8*(1), 1–13.

Ruf, C. S., S. Gleason, and D. S. McKague (2018b), Assessment of cygnss wind speed retrieval uncertainty, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *12*(1), 87–97.

Soisuvarn, S., Z. Jelenak, F. Said, P. S. Chang, and A. Egido (2016), The GNSS-Reflectometry response to the ocean surface winds and waves, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(10), 4678–4699.

Stiles, B. W., and S. H. Yueh (2002), Impact of rain on spaceborne Ku-band wind scatterometer data, *IEEE Transactions on Geoscience and Remote Sensing*, *40*(9), 1973–1983.

Tallapragada, V., et al. (2014), Hurricane weather research and forecasting (hwrf) model: 2013 scientific documentation, *HWRF Development Testbed Center Tech. Rep*, *99*.

Tang, W., S. Yueh, A. Fore, G. Neumann, A. Hayashi, and G. Lagerloef (2013), The rain effect on Aquarius' L-band sea surface brightness temperature and radar backscatter, *Remote sensing of environment*, *137*, 147–157.

Tang, W., S. H. Yueh, A. Hayashi, A. G. Fore, W. L. Jones, A. Santos-Garcia, and M. M. Jacob (2015), Rain-induced near surface salinity stratification and rain roughness correction for Aquarius SSS retrieval, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *8*(12), 5474–5484.

Tournadre, J., and Y. Quilfen (2003), Impact of rain cell on scatterometer data: 1. theory and modeling, *Journal of Geophysical Research: Oceans*, *108*(C7).

Uhlhorn, E. W., P. G. Black, J. L. Franklin, M. Goodberlet, J. Carswell, and A. S. Goldstein (2007), Hurricane surface wind measurements from an operational stepped frequency microwave radiometer, *Monthly Weather Review*, *135*(9), 3070–3085.

Ulaby, F., D. Long, W. Blackwell, C. Elachi, A. Fung, C. Ruf, K. Sarabandi, H. Zebker, and J. Van Zyl (2014), Microwave radar and radiometric remote sensing. the university of michigan.

Voronovich, A. G., and V. U. Zavorotny (2017), Bistatic radar equation for signals of opportunity revisited, *IEEE Transactions on Geoscience and Remote Sensing*, *56*(4), 1959–1968.

Walsh, E., D. Hancock III, D. Hines, R. Swift, and J. Scott (1985), Directional wave spectra measured with the surface contour radar, *Journal of physical oceanography*, *15*(5), 566–592.

Wan, W., B. Liu, Z. Zeng, X. Chen, G. Wu, L. Xu, X. Chen, and Y. Hong (2019), Using cygnss data to monitor chinas flood inundation during typhoon and extreme precipitation events in 2017, *Remote Sensing*, *11*(7), 854.

Weissman, D., B. Stiles, S. Hristova-Veleva, D. Long, D. Smith, K. Hilburn, and W. Jones (2012), Challenges to satellite sensors of ocean winds: Addressing precipitation effects, *Journal of Atmospheric and Oceanic Technology*, *29*(3), 356–374.

Wentz, F. J. (1975), A two-scale scattering model for foam-free sea microwave brightness temperatures, *Journal of Geophysical Research*, *80*(24), 3441–3446.

Wentz, F. J., and D. K. Smith (1999), A model function for the ocean-normalized radar cross section at 14 ghz derived from nscat observations, *Journal of Geophysical Research: Oceans*, *104*(C5), 11,499–11,514.

Wetzel, L. (1990), On the theory of electromagnetic scattering from a raindrop splash, *Radio science*, *25*(06), 1183–1197.

Wickert, J., T. Schmidt, G. Beyerle, R. König, C. Reigber, and N. Jakowski (2004), The radio occultation experiment aboard champ: Operational data analysis and validation of vertical atmospheric profiles, *Journal of the Meteorological Society of Japan. Ser. II*, *82*(1B), 381–395.

Wilheit, T. T. (1979), A model for the microwave emissivity of the ocean's surface as a function of wind speed, *IEEE Transactions on Geoscience Electronics*, *17*(4), 244–249.

Wolfson, M. M. (1990), Understanding and predicting microbursts, Ph.D. thesis, Massachusetts Institute of Technology.

Wright, C. W., E. Walsh, D. Vandemark, W. Krabill, A. Garcia, S. Houston, M. Powell, P. Black, and F. Marks (2001), Hurricane directional wave spectrum spatial variation in the open ocean, *Journal of Physical Oceanography*, *31*(8), 2472–2488.

Xu, F., X. Li, P. Wang, J. Yang, W. G. Pichel, and Y.-Q. Jin (2014), A backscattering model of rainfall over rough sea surface for Synthetic Aperture Radar, *IEEE Transactions on Geoscience and Remote Sensing*, *53*(6), 3042–3054.

Young, I. R. (1999), *Wind generated ocean waves*, Elsevier.

Young, I. R. (2006), Directional spectra of hurricane wind waves, *Journal of Geophysical Research: Oceans*, *111*(C8).

Young, I. R. (2017), A review of parametric descriptions of tropical cyclone wind-wave generation, *Atmosphere*, *8*(10), 194.

Zavorotny, V., and A. Voronovich (2018), GNSS-R modeling results obtained with improved bistatic radar equation, in *2018 International Conference on Electromagnetics in Advanced Applications (ICEAA)*, pp. 35–38, IEEE.

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

Zavorotny, V. U., and A. G. Voronovich (2014), Recent progress on forward scattering modeling for gnss reflectometry, in *2014 IEEE Geoscience and Remote Sensing Symposium*, pp. 3814–3817, IEEE.

Zhang, G., X. Li, W. Perrie, B. Zhang, and L. Wang (2016), Rain effects on the hurricane observations over the ocean by C-band Synthetic Aperture Radar, *Journal of Geophysical Research: Oceans*, *121*(1), 14–26.