Studies of Land and Ocean Remote Sensing Using Spaceborne GNSS-R Systems

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Mohammad Mazen Al-Khaldi, B.S., M.S.

Graduate Program in Electrical and Computer Engineering

The Ohio State University

2020

Dissertation Committee:

Joel T. Johnson, Advisor

Fernando L. Teixeira

Ethan Kubatko

Abstract

Techniques for conducting spaceborne earth remote sensing are well established in the literature. Existing approaches include active and passive methods typically involving the launch of dedicated satellite platforms into orbit. More recently, there has been increasing interest in a relatively less mature mode of remote sensing, Global Navigation Satellite Signal Reflectometry (GNSS-R), which has opened new venues of investigation for the retrievals of geophysical parameters of interest at a global scale with unprecedented spatial and temporal coverage at a fraction of the cost compared to conventional satellite missions.

This dissertation aims to support the use of spaceborne GNSS-R observations for global land and ocean remote sensing through investigating the nature and dependencies on surface geophysical properties of these returns and by developing algorithms to retrieve those of interest. The utility of the proposed analyses and methodologies are investigated in the context of NASA's Earth Venture Mission, CYGNSS (Cyclone Global Navigation Satellite System).

For studies of land remote sensing, a time series retrieval method is introduced for near surface volumetric soil moisture content retrievals. This is supported by an analysis of the physical dependence of GNSS-R DDMs on land properties, showing that variations with soil moisture and composition, vegetation cover, and surface roughness are all to be expected. The proposed time series retrieval algorithm leverages

the slowly varying nature of many of these processes to retrieve soil moisture. Development of complementary approaches to reduce the corrupting effects of low SNR and coherent DDMs in addition to compensating for incidence angle variability are all presented. The utility of the proposed methodology is first explored with simulated GNSS-R measurements and is subsequently extended to CYGNSS measurements on the local, regional and global scales.

In support of the broader science community's land GNSS-R remote sensing investigations, a global coherence detection algorithm is developed. Coherence over land is expected to occur over the exceptionally flat surfaces typically manifested by inland water bodies. Due to the expectation that the correspondence of these returns to the geophysical properties of the surrounding land surface will be limited at best, it is crucial to be able to identify and separate these returns. The proposed methodology is shown to be highly effective at isolating coherent data, and tests over a large CYGNSS dataset suggests direct correlation of the prevalence of these returns with the presence of water bodies within the measurements' footprint. The ability to detect coherence on a global scale together with this correlation, provides a further opportunity to use GNSS-R measurements for the mapping of inland water bodies and analyzing their dynamics. This work therefore, also demonstrates the creation of dynamic inland water body masks through the use of the proposed coherence detector as part of a methodology that bypasses many of the uncertainties associated with existing techniques. The utility of the coherence detection algorithm over ocean surfaces is also explored.

For studies of ocean remote sensing, this work aims to improve storm feature characterization using spaceborne GNSS-R systems by demonstrating maximum hurricane wind speed retrievals through the use of forward models for GNSS-R measurements. The retrieval approach is based on the matching of observations to a synthetic dataset created for a similar track through synthetic storm models having maximum wind speeds and radius as the fundamental parameters. The efficacy of the proposed methodology and its dependencies on storm model and measurement delay extent are explored using CYGNSS special acquisitions, "Full DDMs" (delay-Doppler Maps) and "Raw I/F".

With the expectation that existing spaceborne GNSS-R systems will continue to operate under nominal conditions for the foreseeable future and will be followed by future missions offering further improvements to measurement resolution and that of the retrieved geophysical parameters' of interest as a consequence with a more comprehensive pole-to-pole coverage, this dissertation provides general frameworks and retrieval methodologies that leverage the sensitivity of GNSS-R measurements to various surface properties to support observing, understanding, modeling, and predicting the Earth's surface dynamics.

To the loving memory of my late grandparents and uncle

To my family

Acknowledgments

This dissertation appears in its current form due to the patience, guidance and support of a number of people to whom I would like to extend my appreciation.

To Professor Joel T. Johnson, I extend my sincere gratitude and indebtedness for the opportunity to be part of the Remote Sensing Group at the ElectroScience Laboratory, his guidance, patience and support have made me a better researcher and engineer. I could not have wished for a better mentor. I would also like to thank Professor Fernando Teixeira and Dr. Ethan Kubatko for their expert opinions which helped guide this dissertation into an improved final product. I am also grateful for the indispensable insights discussions with Dr. Steven J. Katzberg of NASA's Langley Research Center and Dr. Scott Gleason of COSMIC-UCAR, have provided. I am also appreciative of the constant support Ms. Patricia Toothman, ECE Academic Program Coordinator, has provided throughout my PhD journey.

My sincere gratitude and appreciation goes to my family for their unyielding support, continued sacrifice and never ending love. Without them, this work would not have been possible.

I would also like to thank all of my friends and colleagues at The Ohio State University, Texas A&M University and the American University of Sharjah for their support and encouragement.

Vita

June, 2015 ..... B.S. in Electrical Engineering.
American University of Sharjah,
Sharjah, United Arab Emirates

August, 2017 ..... M.S. in Electrical Engineering.
EM and Microwaves Laboratory,
Texas A&M University,
College Station, Texas

August, 2017 - Present ..... Graduate Research Associate.
ElectroScience Laboratory,
Department of ECE,
The Ohio State University,
Columbus, Ohio

August, 2019 ..... M.S. in Elec. and Comp. Engineering.
ElectroScience Laboratory,
The Ohio State University,
Columbus, Ohio

Publications

Journals

M. M. Al-Khaldi, J. T. Johnson, S. Gleason, C. Chew, C. Safdi, R. Shah and C. Zuffada, "Inland Water Body Mapping Using CYGNSS Coherence Detection," *IEEE Transactions on Geoscience and Remote Sensing (TGRS)*, Under Review, 2020.

M. M. Al-Khaldi, J. T. Johnson, S. Gleason, E. Loria, A. J. O'Brien and Y. Yi, "Development and Analysis of The Cyclone Global Navigation Satellite System Mission Level-1 Coherence Detection Algorithm," *IEEE Transactions on Geoscience and Remote Sensing (TGRS)*, Under Review, 2020.

S. Gleason, A. O'Brien, A. Russel, **M. M. Al-Khaldi** and J. T. Johnson, "Geolocation, Calibration and Surface Resolution of CYGNSS GNSS-R Land Observations," *Remote Sensing*, 2020.

M. M. Al-Khaldi, J. T. Johnson, Y. Kang, S. J. Katzberg, A. Bringer, E. Kubatko and D. Wood, "Track-based Cyclone Maximum Wind Retrievals Using The Cyclone Global Navigation Satellite System (CYGNSS) Mission Full DDMs," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS)*, 2019.

A. M. Balakhder, **M. M. Al-Khaldi** and J. T. Johnson, "On the Coherency of Ocean and Land Surface Specular Scattering for GNSS-R and Signals of Opportunity Systems," *IEEE Transactions on Geoscience and Remote Sensing (TGRS)*, pp. 1-11, 2019.

M. M. Al-Khaldi, J. T. Johnson, A. J. O'Brien, A. Balenzano and F. Mattia, "Time-Series Retrieval of Soil Moisture Using CYGNSS," *IEEE Transactions on Geoscience and Remote Sensing (TGRS)*, vol. 57, no. 7, pp. 4322-4331, 2019.

Conferences

M. Al-Khaldi, J. T. Johnson, S. J. Katzberg, Y. Kang and E. Kubatko, "Assessment of CYGNSS Characterization of Tropical Cyclones Using Matched Filter Retrievals,"

IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Jul 2020, Waikoloa, HI.

M. Al-Khaldi, J. T. Johnson, S. Gleason, E. Loria, A. J. O'Brien and Y. Yi, "Land and Ocean Coherence Detection Using The Cyclone Global Navigation Satellite System (CYGNSS) Mission," *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Jul 2020, Waikoloa, HI.

A. Bringer, **M. Al-Khaldi**, J. T. Johnson and J. Park, "Monitoring Rapid Change In The Atmosphere Using CYGNSS Wind Speed Measurements," *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Jul 2020, Waikoloa, HI.

M. Al-Khaldi, J. T. Johnson, S. J. Katzberg, Y. Kang and E. Kubatko, "Retrieving Hurricane Sustained Surface Winds Using The Cyclone Global Navigation Satellite System (CYGNSS) Mission's Special Modes of Operation," *American Meteorological Society (AMS) Annual Meeting*, Jan 2020, Boston, MA.

M. Al-Khaldi, J. T. Johnson, S. J. Katzberg, A. Bringer and E. Kubatko, "Maximum Sustained Surface Winds Retrieval Using The Cyclone Global Navigation Satellite System (CYGNSS) Constellation," *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Jul 2019, Yokohama, Japan.

M. Al-Khaldi and J. T. Johnson, "Retrieval Performance Analysis For Time Series Retrievals of Soil Moisture Under Dynamic Vegetation Canopies and Heterogeneous Land Cover Using The CYGNSS Constellation," *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Jul 2019, Yokohama, Japan.

Y. Kang, E. Kubatko, J. T. Johnson and **M. Al-Khaldi**, "Storm Surge Simulation with CYGNSS-derived Parametric Wind Fields," *U.S. National Congress on Computational Mechanics (USNCCM15)*, Jul 2019, Austin, TX.

M. Al-Khaldi, S. Wijesundara , J. T. Johnson, A. Balenzano and F. Mattia, "A Time Series Change Detection Algorithm for Spaceborne GNSS-R Soil Moisture Retrievals Over Complex Terrain," *Progress in Electromagnetic Research (PIERS)*, Jun 2019, Rome, Italy.

M. Al-Khaldi, J. T. Johnson, S. J. Katzberg, A. Bringer and E. Kubatko, "A Forward Modelling Approach to Improve Storm Feature Characterization Using Spaceborne GNSS-R Systems," *IEEE Specialist Meeting on Reflectometry using GNSS and other Signals of Opportunity (GNSS+R)*, Jul 2018, Benevento, Italy.

J. T. Johnson,**M. Al-Khaldi** and A. M. Balakhder , "On the Nature of GNSS-R Land Surface Specular Scattering," *IEEE Specialist Meeting on Reflectometry using GNSS and other Signals of Opportunity (GNSS+R)*, Jul 2018, Benevento, Italy.

M. Al-Khaldi and J. T. Johnson , "Time Series Soil Moisture Retrievals Using the CYGNSS Constellation," *USNC-URSI National Radio Science Meeting*, Jan 2019, Boulder, CO.

M. Al-Khaldi, A. Bringer, J. T. Johnson, S. J. Katzberg and E. Kubatko, "An Analysis of Maximum Hurricane Wind Retrievals Using Spaceborne GNSS-R Systems," *USNC-URSI National Radio Science Meeting*, Jan 2019, Boulder, CO.

A. M. Balakhder, **M. Al-Khaldi** and J. T. Johnson, "A Study of The Amplitude Distribution of Near-Specular Bistatic Sea Clutter," *International Conference on Electromagnetics in Advanced Applications (ICEAA)*, Sep 2018, Cartagena, Colombia.

M. Al-Khaldi, J. T. Johnson, E. Loria and A. O'Brien, "Multi-Temporal L-Band Estimation of Soil Moisture Using Bistatic GNSS Land Returns," *Progress in Electromagnetic Research (PIERS)*, Aug 2018, Toyama, Japan.

Balenzano, G. Satalino, F. Lovergine, F. Mattia, O. Cartus, M. Davidson, **M. Al-Khaldi** and J. T. Johnson, "Cross-Comparison of Three SAR Soil Moisture Retrieval Algorithms Using Synthetic and Experimental Data," *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Jul 2018, Valencia, Spain.

M. Al-Khaldi, J. T. Johnson, A. O'Brien, F. Mattia and A. Balenzano, "GNSS-R Time-Series Soil Moisture Retrievals from Vegetated Surfaces," *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Jul 2018, Valencia, Spain.

M. Al-Khaldi, J. T. Johnson and A. O'Brien, "Techniques in GNSS-Reflectometry Remote Sensing of Global Near-Surface Soil Moisture," *Consortium of Ohio Universities on Navigation and Timekeeping (COUNT)*, Apr 2018, Columbus, OH.

D. J. Posselt, C. S. Ruf, J. T. Johnson, T. Lang, J. Crespo, **M. Al-Khaldi**, M. Morris and D. Waliser, "CYGNSS Science Highlights from the First Year on Orbit," *American Meteorological Society (AMS) Annual Meeting*, Jan 2018, Austin, TX.

M. Al-Khaldi, J. T. Johnson, J. Park and A. O'Brien, "An Analysis of CYGNSS Returns Over Land," *USNC-URSI National Radio Science Meeting*, Jan 2018, Boulder, CO.

A. O'Brien, **M. Al-Khaldi** and J. T. Johnson, "Forward Modeling of CYGNSS GNSS-R Land Reflection Measurements," *USNC-URSI National Radio Science Meeting*, Jan 2018, Boulder, CO.

Workshops

M. Al-Khaldi, J. T. Johnson, S. Gleason, C. Chew, C. Safdi, R. Shah and C. Zuffada, "Inland Water Body Mapping Using CYGNSS Level-1 Coherence Detection," *NASA CYGNSS Science Team Meeting*, Jun 2020.

J. T. Johnson, Y. Yi, **M. Al-Khaldi**, A. Bringer, C. K. Shum and X. Wang, "Using CYGNSS to Observe Ocean Windspeed Dynamics," *NASA CYGNSS Science Team Meeting*, Jun 2020.

E. Kubatko, Y. Kang, J. T. Johnson, **M. Al-Khaldi** and S. J. Katzberg, "CYGNSS-based parametric wind fields for storm surge modeling," *NASA CYGNSS Science Team Meeting*, Jun 2020.

Y. Kang, **M. Al-Khaldi**, E. Kubatko, J. T. Johnson and S. J. Katzberg, "Effect of CYGNSS-Derived Parametric Wind Fields on Storm Surge Modelling," *NASA CYGNSS Science Team Meeting*, Jun 2019, Ann Arbor, MI.

A. Bringer, Y. Yi, **M. Al-Khaldi**, C. K. Shum, J. T. Johnson and X. Wang, "Investigating Atmospheric Variability over Multiple Time Scales with CYGNSS," *NASA CYGNSS Science Team Meeting*, Jun 2019, Ann Arbor, MI.

S. Gleason, A. O'Brien, J. T. Johnson, E. Loria, **M. Al-Khaldi**, T. Butler, A. Russel and C. Ruf, "CYGNSS Land Calibration and On-Orbit Reconfiguration Status," *NASA CYGNSS Science Team Meeting*, Jun 2019, Ann Arbor, MI.

M. Al-Khaldi, J. T. Johnson, A. Bringer and A. O'Brien, "Analysis of CYGNSS Land Returns and Soil Moisture Retrievals," *NASA CYGNSS Science Team Meeting*, Jan 2019, Pasadena, CA.

J. T. Johnson, A. Bringer, Y. Yi, **M. Al-Khaldi** and C. K. Shum, "Investigating Atmospheric Variability over Multiple Time Scales with CYGNSS," *NASA CYGNSS Science Team Meeting*, Jan 2019, Pasadena, CA.

E. Kubatko, J. T. Johnson, Y. Kang and **M. Al-Khaldi**, "Effects of CYGNSS-derived parametric wind fields on storm surge modelling," *NASA CYGNSS Science Team Meeting*, Jan 2019, Pasadena, CA.

C. K. Shum, Y. Jia, **M. Al-Khaldi**, Y. Yi, Y. Shi, W. Li, E. Cardellah, H. Ge, M. Ge, X. Wang, T. Wang, C. Kuo and H. Kao, "CYGNSS Reflectometry Altimetric Height Retrieval for Prospective Tropical Cyclone Studies," *NASA CYGNSS Science Team Meeting*, Jan 2019, Pasadena, CA.

M. Al-Khaldi, J. T. Johnson and A. O'Brien, "Development of CYGNSS Time Series Soil Moisture Retrieval Algorithm," *NASA CYGNSS Science Team Meeting*, Jun 2018, Ann Arbor, MI.

J. T. Johnson, J. Park, Y. Yi, A. O'Brien, **M. Al-Khaldi**, A. Bringer and C. K. Shum, "CYGNSS Cal/Val and Ocean Science Activities at OSU and UCLA," *NASA CYGNSS Science Team Meeting*, Jun 2018, Ann Arbor, MI.

M. Al-Khaldi and J. T. Johnson, "CYGNSS Soil Moisture Retrieval Studies," *NASA CYGNSS Science Team Meeting*, Dec 2017, Miami, FL.

Fields of Study

Major Field: Electrical and Computer Engineering

Studies in:

Applied Electromagnetics
Remote Sensing
Radar Systems
Signal Processing

Table of Contents

List of Tables

List of Figures

Figure **Page**

Chapter 1: Introduction

In the context of spaceborne Earth observing systems, remote sensing refers to the inference of information relating to the Earth's surface through the measurement of its naturally occurring emission/radiation or through observing its response to impinging radiation. While numerous ground based and airborne systems exist for the purposes of remote sensing, in this work spaceborne platforms are of most interest. This is primarily due to their ability to provide global coverage thereby furnishing access to remote and/or inhospitable environments over long periods of time, enabling the investigation of patterns, dynamics and seasonality of a wide range of lasting and ephemeral phenomena.

Choice of the mode of remote sensing relied upon, sensor used and frequency of operation are highly contingent on the intended application and how the constituents of the observed scene interact with the impinging radiation. At lower frequencies, the scene's collective wave reflection, scatter and emission is the result of electronic conduction and electron displacement dictated by the materials' dielectric properties giving rise to a composite received field. This occurs at the radio and microwave frequency bands and can be used to retrieve information about the physical properties of the observed surface, atmospheric constituents and precipitation using radars,

radiometers and scatterometers. At higher frequency bands electronic energy transitions between impinging radiation and interceding molecules at well defined spectral bands plays a more significant role in retrieving useful information about the observed scene. This occurs at and beyond the mid-infrared frequency band and is instrumental in the remote sensing of chemical composition of the observed surface, atmospheric gas distribution and vegetation cover using optical and hyperspectral sensors for example. In this work, it is the remote sensing of the Earth's surface within the microwave frequency band that is of most interest.

1.1 Motivation

Often the efficacy and utility of these remote sensing platforms is limited by practical considerations including cost, spatial coverage, spatial resolution, revisit time, mission life span, frequency of operation and several others. In contrast, Global Navigation Satellite System Reflectometry (GNSS-R) as an emerging spaceborne remote sensing technique offers to bypass many of these limitations. Its passive mode of remote sensing, taking advantage of already existing Global Navigation Satellite System (GNSS) transmitters and the construction of GNSS-R receivers from commercially available components, significantly reduces cost and project incubation time before launch. Further, the reduced cost enables the simultaneous launch of several receivers in orbit thereby allowing for improved spatial and temporal coverage.

Further, their L-band operation provides an ideal frequency window for the purposes of remote sensing of the Earth's surface. Due to the tendency of electromagnetic radiation to be perturbed by matter that is on the scale of a wavelength, the Earth's atmosphere is opaque at and beyond the sub-millimeter frequency bands. This is due

to the fact that, at those frequencies the size of a wavelength becomes comparable to that of atmospheric gases and therefore radiation is heavily perturbed by significant random scattering and absorption effects. For lower frequencies between 10 MHz and 10 GHz, the atmosphere is transparent providing a window for Earth-to-space and space-to-Earth exchange of radiation. At frequencies lower than 10 MHz, the ionosphere prevents the passage impinging waves. The L1 navigation signals, typically relied on for GNSS-R, are at a frequency of $\sim$1.575 GHz thereby allowing them to penetrate through the Earth's atmosphere, thick vegetation and heavy rain [1–3] making the L1 frequency ideal for land and ocean remote sensing studies.

In addition, the usefulness of traditional satellite remote sensing missions has often been limited to a single geophysical deliverable or a set of closely related retrieved surface parameters of interest. In contrast, current spaceborne GNSS-R measurements, made available through the Cyclone Global Navigation Satellite System (CYGNSS) mission, have shown sensitivity to a wide range of surface properties and processes of interest. Analyses of the extensive GNSS-R land and ocean datasets made available by the TDS-1 and CYGNSS missions [4–11] have shown their potential utility in a wide range of applications not limited to their core mission objectives, the estimation of wind speeds over ocean surface.

For the aforementioned reasons, this work has aimed to develop techniques to leverage the sensitivity of GNSS-R returns to surface properties for land and ocean applications. While the utility of the proposed methodologies can be generalized to any spaceborne GNSS-R system, they were applied in the context of CYGNSS given the availability of measurements from its eight satellite constellation.

1.2 Significance of Research Topics

Soil Moisture

Soil moisture plays an important role in a wide range of hydrological and geophysical processes on a global scale including vegetation and agriculture, land-atmosphere circulation and climatic conditions [12–15]. The ability to remotely sense soil moisture is therefore of vital importance. According to reports by the National Research Council (NRC) as part of their analysis of national imperatives for the future [16] the issue of accurate soil moisture determination and constant monitoring on a global basis was recognized as a high impact item due to the effect this is expected to have on human interests through influencing the ability to provide weather forecasting, assessment of agricultural productivity and the potential impact of droughts, and estimation of global water/energy fluxes.

Previous microwave remote sensing of soil moisture has exploited the sensitivity of the dielectric constant of soil at microwave frequencies to the presence of water, so that any measurements sensitive to the land surface permittivity can potentially be used for observing soil moisture. The current Soil Moisture Active Passive (SMAP) and Soil Moisture and Ocean Salinity (SMOS) missions, both operating within L-band, focus on passive microwave measurements, example other systems having demonstrated soil moisture sensing using radar backscatter measurements includes SMAP, Sentinel-1, ENVISAT/ASAR and TerraSAR-X1 [17–24]. The reduced sensitivity of passive microwave measurements to the confounding factors of surface roughness and vegetation have allowed more robust retrievals to be obtained from microwave radiometry; methods to achieve robust compensation for vegetation and surface roughness with radar backscatter remain of ongoing interest [20–24]. The recent availability of

spaceborne measurements of bistatic land surface specular scattering from systems based on GNSS-R are producing increasing interest in soil moisture remote sensing with GNSS-R [25–33].

In particular, studies analyzing land returns of the recently launched CYGNSS constellation [32,33] clearly demonstrate sensitivity to land surface properties including soil moisture. This has motivated further investigation into the means through which CYGNSS and future GNSS-R missions can be used for the purpose for soil moisture retrievals and land applications.

Early studies undertaken in [34] began investigating the utility of GNSS-R systems for soil moisture retrievals through airborne experimental campaigns. In these efforts a 12-channel GPS navigation receiver, modified to perform bistatic radar measurements, was fitted to an NCAR C-130 aircraft and used to receive reflected GPS signals off the Earth's surface. It was flown over sites of interest as part of the Soil Moisture Experiment 2002 (SMEX02) campaign. While no soil moisture retrieval algorithm was developed the authors illustrated the clear sensitivity of the received signals' SNR to the volumetric soil moisture content of the surface. Other more recent studies [27–33] have focused on investigating the sensitivity of spaceborne GNSS-R systems to near-surface soil moisture content and the potential of developing complementary retrieval algorithms.

The studies of [27] are concerned with examining correlations between TechDemoSat-1 (TDS-1) returns and soil moisture. This is conducted through fitting TDS-1 SNR and soil moisture provided by the Soil Moisture and Ocean Salinity (SMOS) mission. The authors specialize this process for different terrain types and quantify extent of correlation based on land cover. Subsequent studies focused on GNSS-R land

returns from CYGNSS where in [28] the authors develop a combined soil moisture product based on returns from CYGNSS and the SMAP mission. The retrieval is based on eliminating extremes within CYGNSS data by removing contributions of returns related to the most 'wet' and most 'dry' conditions for a given location over a predefined time period and subsequently weighting CYGNSS SNR with SMAP soil moisture to achieve an optimum retrieval. Other studies [29–33], some of which were amongst the first to report CYGNSS's sensitivity to near-surface soil moisture levels, have analyzed complementary patterns of change between CYGNSS SNR and surface soil moisture level and inundation. These investigations similarly focused on linear fits of CYGNSS SNR to SMAP soil moisture or related both quantities through the illustration of complementary patterns of change.

Robust retrieval algorithms are nonetheless complicated by the need to account for the complex combination of dependencies GNSS-R land returns have on soil moisture and composition, vegetation cover, and surface roughness as well as the presence of water bodies within the measurements' footprints. In the absence of accurate ancillary data provided with sufficiently low latency, an alternative retrieval algorithm based on change detection is proposed herein. The retrieval algorithm attempts to minimize dependence on ancillary data whilst limiting the corrupting effects of coherent returns on the retrieval.

Coherence Detection

In the context of bistatic radar systems, coherence is defined to be the result of a scattering process in which the fields scattered from the observed target arrive at the receiver having similar phase shifts and therefore add coherently. More specifically, in

the context of Earth observing bistatic radar systems, as is the case with GNSS-R, the presence of coherence refers to the unusual case where the area extensive target being observed has an exceptionally flat surface. Ocean surfaces are generally 'roughened' by surface winds of varying speeds such that incoherence is accepted as the dominant mode of scattering, with coherence arising only over harbored seas and areas with calm waters (i.e. low surface winds). Over land, the surface is predominantly roughened by large scale topographic features; due to the heterogeneity of the observed surfaces a mixture of coherent and incoherent returns exists. This poses a unique challenge. Due to the fact that coherent returns over land arise due to smooth surfaces, typically manifested by inland water bodies, the ability to distinguish between these returns is of vital importance for the ability to retrieve useful geophysical parameters of interest.

Beyond the interest in the use of spaceborne GNSS-R systems for the purpose of soil moisture retrievals [25–29], the use of GNSS-R observations over land surfaces has also received interest for several other purposes. This includes use to conduct vegetation and biomass retrievals [35] as well as flooding detection and inland water body mapping [36–39]. The success of studies relating to the retrieval of physical surface properties such as soil moisture and composition is contingent on the correlation of the observed surface scatter to that of the surface. Due to the expectation that dominantly coherent land surface returns arise primarily from inland water body contributions that are not directly representative of such surface properties, a need therefore arises for their identification and isolation. Similarly for vegetation and biomass studies, retrievals have often focused on the conformity of observed reflections to a set of scattering tendencies given by a forward model describing the expected behavior as a function of the scene's vegetation water content or vegetation

optical depth. Often with GNSS-R, this is expected to follow an exponential decaying behavior [40, 41]. The presence of coherent returns in this case may 'overwhelm' the incoherent components of the measured signals and cause significant deviations off the expected behavior. For this reason, elimination of measurements with a dominant coherent component is expected to be a crucial step towards retrieval of biomass and vegetation indices. On the other hand, for flood detection and inland water body mapping coherent returns are expected to be of most interest. Inclusion of incoherent returns as part of the retrieval process is expected to degrade results through the inclusion of returns expected to arise due to scatter from an area that is many times larger than the footprint of the targeted water body or flood zone. In this case a need arises to identify and isolate incoherent returns.

In support of the broader community's investigations concerned with land remote sensing using spaceborne GNSS-R systems, the work presented herein develops a global coherence detection algorithm which has been extended to CYGNSS measurements. The proposed algorithm leverages the correspondence of the extent of power spread across measured DDMs with the underlying scattering mechanisms to distinguish between returns dominated by coherent reflection or incoherent scatter.

Hurricane Characterization

According to reports by the National Oceanic and Atmospheric Administration (NOAA) [42–44], it is estimated that hurricanes Harvey, Irma, and Maria were associated with economic damage of $125 billion, $50 billion and $91.5 billion respectively in the 2017 Atlantic hurricane season. Similar analyses report a combined fiscal loss of $136 billion between the years 2018-2019 due to similar weather disaster [45]. The

ability to monitor storm properties multiple times per day and thereby improve storm track and intensity forecasts is therefore clearly of high importance to improve disaster readiness and to minimize the damage incurred, both to human life and to property. While extensive efforts have been devoted to monitoring cyclone properties, the high rains and heavy cloud cover near a hurricane eye continue to confound methods for remote sensing hurricane maximum winds. This challenge is compounded by the fact that storms can undergo substantial changes in their structure over a relatively short time span.

Systems such as the Hurricane Weather Research and Forecasting (HWRF) model can predict these rapid changes only with limited accuracy. Reconnaissance flights provide accurate storm wind information, but achieve only limited coverage in space and time during a storm's lifespan. In a similar fashion several existing spaceborne platforms can provide data on storm properties, such as the surface wind vector measurements of scatterometers such as ASCAT and the precipitation observations of the Global Precipitation Measurement (GPM) mission [46, 47]. However, their sparse spatial and temporal coverage as well as the impact of heavy rain on these systems (which use C-band or higher microwave frequencies) limits their ability to monitor storm maximum winds.

GNSS-R sensors are well suited to address these challenges. The potential for frequent revisits facilitated by these systems provides new opportunities for observing and forecasting storm properties. Further, their L-band frequency operation makes them resilient to adverse weather conditions including heavy rain, making observations of winds in the central core of storms possible. As a further opportunity, platforms like CYGNSS are capable of producing data products having extended delay

and Doppler measurement ranges during times of storm coverage under two special modes of acquisition. Due to these desirable properties and the mission's focus on cyclone wind sensing, this work has aimed to improve storm feature characterization using spaceborne GNSS-R systems by demonstrating maximum hurricane wind speed measurements based on matched filter retrievals.

1.3 Organization of Dissertation

Each chapter of this dissertation discusses models and techniques through which GNSS-R observations may be used for retrieving vital geophysical parameters of interest. While the usefulness of the proposed techniques may be generalized to any spaceborne GNSS-R system, their transition from concept and simulation to operation is explored in the context of the CYGNSS mission.

Chapter 2 presents an overview of the fundamental elements and theory of GNSS-R systems as Earth observing passive bistatic radars. Chapter 3 provides an overview of the CYGNSS mission and highlights the data products it provides under its nominal and special modes of operation. As an integral part of subsequent formulation, the CYGNSS End-to-End Simulator and its ability to predict CYGNSS DDMs with a high degree of accuracy are also described.

In Chapter 4, a time series global near surface volumetric soil moisture retrieval algorithm based on change detection is proposed. The use of the fundamental measurement GNSS-R provides, the delay-Doppler map (DDM), to derive the Normalized Radar Cross Section over land (NRCS) is discussed and the dependence this has on the composite contributions of soil moisture, vegetation attenuation and surface

roughness, distinguishing the physical mechanisms through which land returns exhibit dependence on soil moisture is elucidated. Issues of practical importance are also covered including contamination due to inland water bodies and the dependence of specular scattering on varied incidence angles, so that a correction for these can be developed. This is followed by an overview of the proposed time-series formulation for soil moisture estimation using bistatic GNSS-R systems. The extension of this to measured CYGNSS observations, related retrieval performance along with the results achieved are also described.

Chapter 5, discusses the question of the nature of GNSS-R land returns under varying surface roughness conditions. Approximate expressions are developed describing those conditions, and the maximum permissible level of surface roughness for coherent returns to be dominant is established. Analysis of this development over sites of interest is also discussed where it is shown that exceptionally flat surfaces are required for coherence to exist and incoherent scatter is otherwise expected to be dominant. The proposed global coherence detection algorithm is also discussed. Details and basic formulation of the proposed detection methodology are outlined and issues of practical relevance are also highlighted. In particular, the corrupting effects noise can have on the ability to distinguish and separate coherent/incoherent returns is discussed along with means through which its contributions may be minimized. Detection performance is compared against reference tracks and results obtained from applying the detection algorithm to a global CYGNSS dataset are reported. Dependencies on the prevalence of water bodies, varying levels of surface roughness, correspondence to Signal-to-Noise ratio and extended utility over ocean surfaces are all explored.

In Chapter 6, the wide range of applications for which the Level-1 coherence detection methodology is applicable, on a global scale, is highlighted including the mapping of inland water bodies, tracking and mapping large scale atmospheric phenomena, analysis of GNSS-R range resolutions under dominantly coherent reflection regimes as well as improvement of soil moisture retrievals.

In Chapter 7, an alternate methodology to using GNSS-R observations to retrieve storm maximum winds based on a matched filter approach is explored. This includes the tools used to forward model CYGNSS returns under different storm conditions, the parametric modelling of storms, and the retrieval process. The results obtained by extending the retrieval approach to both simulated and actual CYGNSS DDM observations over for several observed storms between 2017-2019 are summarized. This is complemented with a discussion relating to the dependencies of the retrieval process and its limitations.

Overall conclusions of the studies undertaken in this dissertation, recommendations for future work and potential for further investigation are all discussed in Chapter 8.

Chapter 2: GNSS-R As A Passive Bistatic Radar System

A radar is commonly defined as being an electromagnetic system that transmits a signal illuminating a scene or target and subsequently inferring information about it through examining the nature of the echo signal it detects [49]. A GNSS-R system in many respects can be regarded as a radar. It is a passive form of radar in that it does not transmit signals and instead relies on transmissions provided by navigation satellites. It is a bistatic form of radar in that the transmitter and receiver are not co-located. As opposed to observing a point target, the GNSS-R radar observes an area extensive target, a surface area that occupies an appreciable portion of the antenna footprint and its operation is affected by many of the same factors that impact traditional radars. This chapter provides an overview of some of the fundamental elements of GNSS-R systems and how returns ultimately translate into fundamental measurements and GNSS-R observables.

2.1 Global Navigation Satellite Systems

The term Global Navigation Satellite System (GNSS) refers to a satellite, or constellation of satellites, providing position and timing signals to receivers. Several GNSS constellations exist the first of which was Navstar, otherwise known as Global

Positioning System (GPS), and is operated by the United States Department of Defense. Others include GLObal Navigation Satellite System (GLONASS) operated by the Russian Federation, Galileo operated by the European Union, China's BeiDou (formerly Compass), the Indian Regional Navigation Satellite System (IRNSS) and Japan's Quasi-Zenith Satellite System (QZSS) [50–55]. Over the next decade in excess of 100 fully operational GNSS transmitters are expected to be in orbit [56] providing an abundance of transmitted signals for use in GNSS-R. Due to the fact that GPS remains the only fully operational GNSS system, the reflections of its transmissions off the Earth's surface are typically exclusively relied upon for GNSS-R and are therefore of most interest. For this reason, some of its properties are explored further in what follows.

Under nominal operating conditions, the GPS constellation includes 24 satellites placed in Medium Earth Orbit (MEO) at an orbital height of approximately 20,200 km. A number of additional GPS satellites are also kept in orbit as contingencies in the case of operational satellite failure. The satellites' altitudes are such that they are semisynchronous, allowing for two revisits per day. While the orbital configuration has undergone several revisions since its inception [57], the satellites are currently divided into six principle planes with each subset containing 4 satellites at an orbital inclination of 55 degrees an illustration of which is shown in Figure 2.1.

As a consequence, at any given time and location a minimum of 4 GPS satellites are expected to be in view. At any given instant, the information that a GPS satellite will attempt to convey constitutes four major components: self-identification of the transmitting satellite, current time, information about the transmitting satellite's orbit known as orbit ephemerides and its current status. This information in-whole

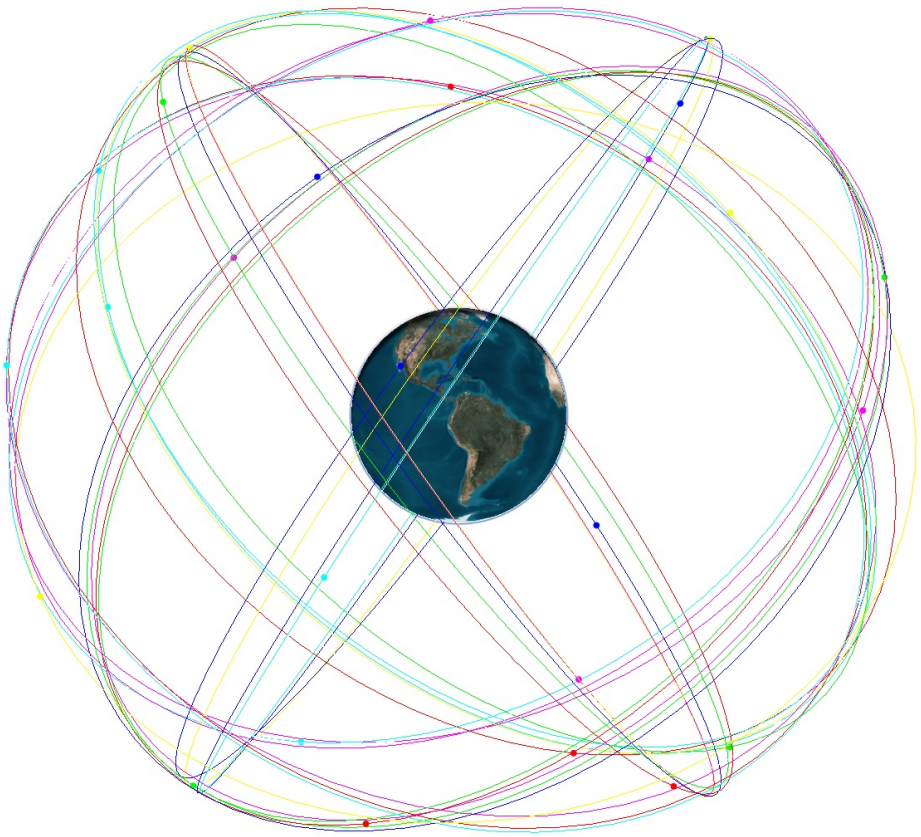


Figure 2.1: Illustration of orbital configuration of NAVSTAR constellation at six focal planes with an inclination of approximately 55 degrees. Configuration depiction is to scale.

is embedded within the transmitted GPS signal that is transmitted with a right hand circular polarization as to minimize its susceptibility to Faraday rotation.

The GPS signal includes three principle components L1, L2 and L5 at the center frequencies 1.575, 1.227 and 1.176 GHz generated synchronously. Notice, that all signal components current and future are within the L-band as to simplify transmission antenna design, minimize impact of ionospheric delay which can have an appreciable impact on lower frequency signals, minimize effects of weather and the atmosphere and to ensure adequate availability of bandwidth required by the coding schemes.

Typically, an unencrypted Coarse Acquisition (C/A) code is used to modulate the L1 carrier at a bit rate of 1.023 MHz and a military-use restricted access Precision code (P-code) is used to modulate its frequency at a bit rate of 10.23 MHz. In contrast L2 is modulated by the P-code only.

GNSS-R systems [4–11] have typically focused on the C/A code and therefore the L1 component of the GPS signal. Since each satellites' L1 signal is modulated by a unique C/A code, the term Pseudo-random Noise (PRN) is used interchangeably to refer to a certain C/A code or its transmitting satellite. The pseudo-randomness of the C/A codes is important for a variety of reasons including the fact that it makes the signal more resilient to interferences and jamming. More importantly, it allows multiple transmitted signals at the same frequency to be received by different GPS receivers. This is an integral part of the trilateration process used to conduct geolocation with GPS. As an example, the C/A code for PRN 15 is illustrated in Figure 2.2(a). The transition in states is known as chipping and each flip or chip is associated with a duration of 0.97μs with 1023 chips per code (i.e. chip duration is 1 ms). Its autocorrelation is high and cross-correlation among distinct PRN codes is low similar to white/Gaussian noise.

Therefore, with knowledge of the C/A codes a receiver discerns which GPS satellite transmitted a given signal. In the context of GNSS-R this can be done in two ways. The receiver can locally generate 'clean' or noise-free replicas of the transmitted signals serving as a reference for all PRNs which is the traditional approach taken for most GNSS-R systems. Alternatively, the receiver's zenith antenna can be used to capture the line of sight GPS signal which is later cross correlated with the reflected signal in what is known as interferometric GNSS-R or GNSS-IR.

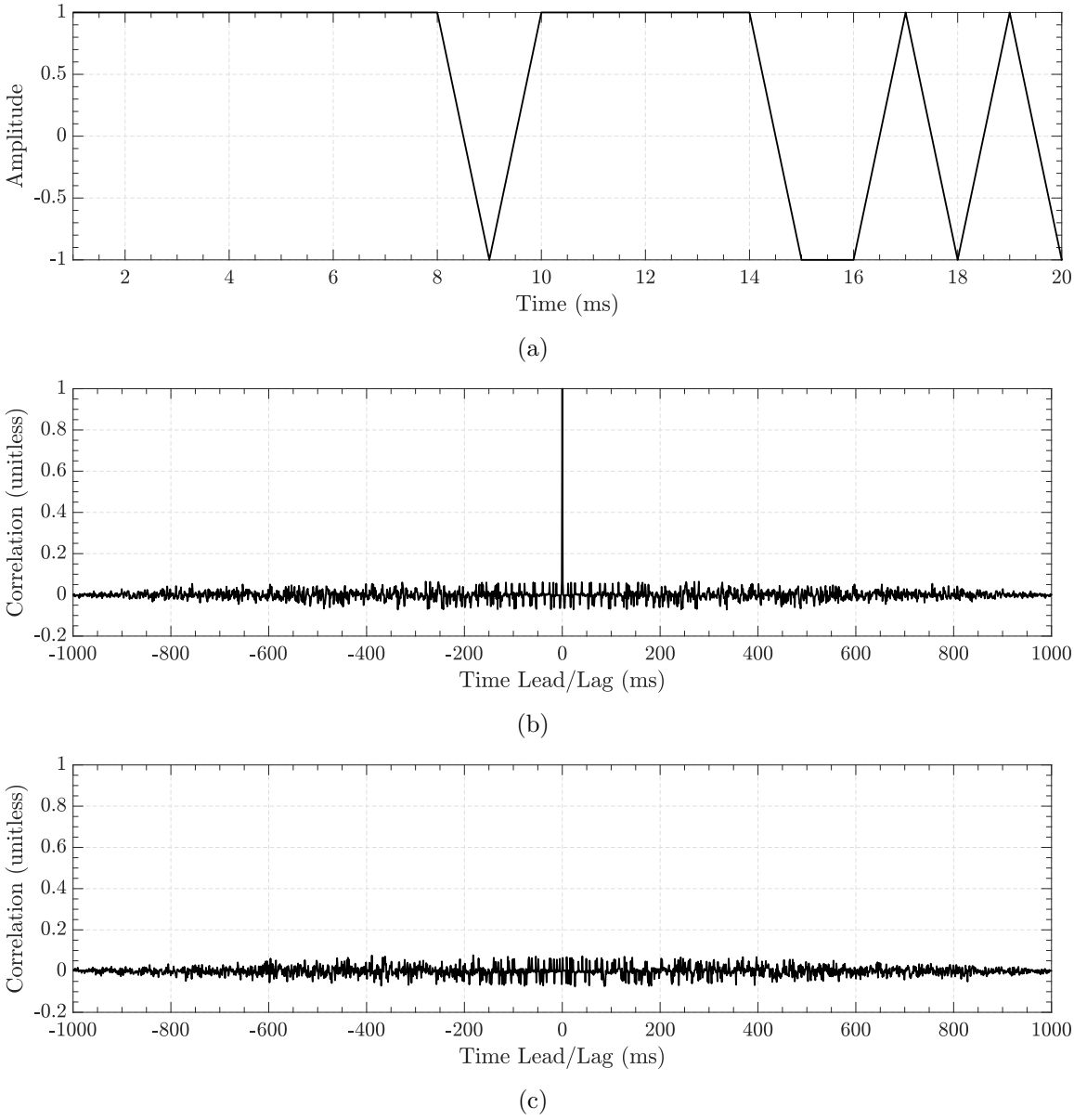


Figure 2.2: Illustration of C/A code pseudo-random noise-like properties (a) C/A code for PRN 15 (b) Auto-correlation of PRN 15 C/A code (c) Cross-correlation of PRN 15 and PRN 20 C/A codes

As part of an effort to modernize the GPS constellation, several other signals have been introduced. This includes the L5 signal currently supported by Block IIF

satellites in an effort to improve signal accuracy and resiliency to interference as well as provide redundancy. Other additions include the added modulation of L2 with C/A code giving rise to L2C for improved navigation accuracy through the minimization of ionospheric delay error. A fourth navigation signal is also expected to be added to L1 giving rise to L1C specifically targeting to improve the degraded GPS performance in urban and other similarly challenging environments [58].

2.2 Signal Scattering

2.2.1 Scattering Geometry

The typical GNSS-R geometry is depicted in Figure 2.3, comprising a GNSS-R receiver, a GNSS transmitter and a reflecting surface.

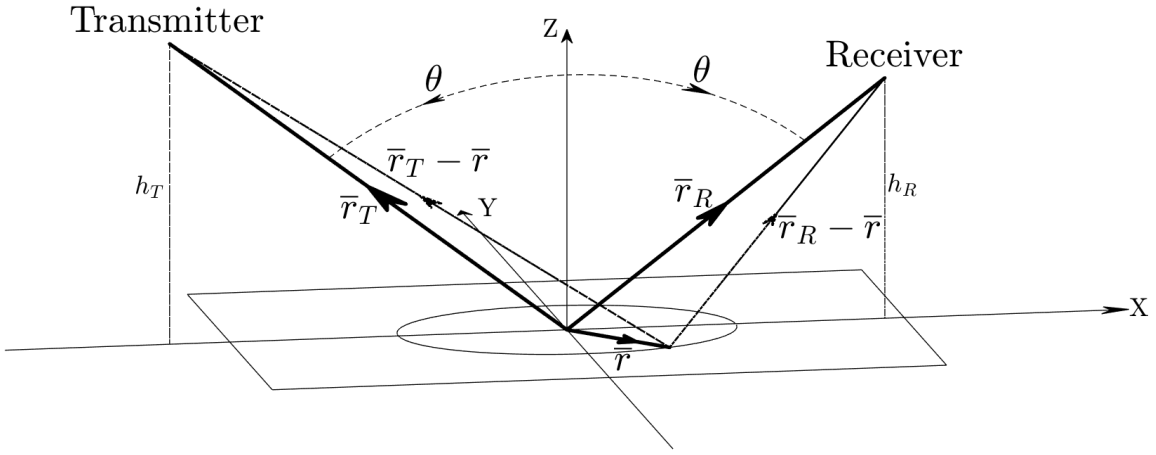


Figure 2.3: Illustration of specular bistatic geometry

where θ is the incidence and scattering angles, $\bar{r}_T$ is the radial vector between specular point and transmitter, $\bar{r}_R$ is the radial vector between specular point and receiver, $\bar{r}$ is the radial vector between specular point and an arbitrary surface point,

$\bar{h}_T$ is the height of transmitter above the plane tangential to the surface, and $\bar{h}_R$ is the height of receiver above the plane tangent to the Earth's surface at the specular point.

GNSS-R geometries are bistatic specular geometries in that the incidence θ_i and scattering θ_s angles are equal ($\theta_i = \theta_s = \theta$) with an azimuthal angular separation ϕ of 180°. The specular point is the reflecting surface point that gives rise to a minimum in total propagation distance of the signal as it is transmitted by the GNSS satellite, reflected off the Earth's surface and received by the GNSS-R receiver. The observed surface is typically 'roughened' by either topographic profiles within the measurement's footprint, as is the case with specular points over land, or by a forcing phenomena like surface wind, as is the case with ocean surfaces. The observed surface can be thought of as comprising randomly oriented scattering facets each with a height and tilt. For random media, as is the case with the Earth's surface, the exact configuration and properties of the scattering facets are not known and therefore statistical measures are typically relied upon to describe their distributions [59]. Depending on the orientation of each of the facets, some will scatter power toward the receiver and others will scatter power in random directions. The surface area contributing to the bulk of the signal scattering towards to the receiver is known as the glistening zone. Its size is dictated in large by the observed surface's levels of roughness. The rougher the surface is, the larger the glistening zone is expected to be; proportionally less power is reflected from the immediate vicinity of the specular point and more power is scattered towards the receiver from points farther away. Conversely, the smoother the surface is the smaller the glistening zone is expected to be where the bulk of the signal received is expected to have been scattered from

the specular point and the small area surrounding it. The properties of the received signal are therefore highly contingent on the properties of the scattering surface.

Due to the dependence of received signals on the nature of the random medium off which they reflect, it is important to have techniques that describe the scattering process. Several such techniques exist including the Small Perturbation Method (SPM), Second Order Small Slope Approximation (SSA), Method of Moments (MoM) and Phyiscal Optics (PO), or Kirchhoff, approximation [59–66]. It is important to note that the choice of which scattering theory to use depends in large on the relation between the size wavelength of the signal, and therefore its frequency, and the dominant scales of roughness that the observed surface manifests and off which the bulk of the received signal reflects. Due to the fact that with GNSS-R, the dominant scattering is expected to arise off roughness profiles with features much larger than a wavelength (with large slowly changing radius of curvature) the Kirchhoff approximation is of most interest and is explored further in what follows.

2.2.2 Received Signal Processing

In contrast to a traditional scenario in which a radar expects a single dominant reflection off the target which it is observing, with GNSS-R the target is a large footprint off which multiple 'echoes' or reflections are expected. Each reflection is expected to be associated with a delay τ and Doppler f_D. In the context of GNSS-R, delay diversity arises due to the varied paths scattered signals take from the different surface points within the measurement's footprint. Fundamentally, the delay on the surface is determined by the total path delay a transmitted signal undergoes as it propagates from a GNSS transmitter, reflects off a point on the surface and propagates

to the GNSS-R receiver. Computing delay as a function of spatial offset from the specular point for a smooth surface becomes a question of developing approximate expressions for the total distance travelled r_{tot} by the transmitted signal. Neglecting Earth's curvature this is given by:

$$\tau = \frac{r_{tot}}{c} = \frac{|\bar{r}_T - \bar{r}| + |\bar{r}_R - \bar{r}|}{c} \tag{2.1}$$

where c is the speed of light and:

$$\bar{r}_T = (-\hat{x}\sin\theta + \hat{z}\cos\theta) \cdot r_T \tag{2.2}$$

$$\bar{r}_R = (\hat{x}\sin\theta + \hat{z}\cos\theta) \cdot r_R \tag{2.3}$$

$$\bar{r} = (\hat{x}\cos\phi + \hat{y}\sin\phi) \cdot r \tag{2.4}$$

such that delay at any surface point as a function of offset from the specular point can be expressed as (2.5).

$$\tau \approx \frac{r_t + r_r}{c} + \frac{r^2}{2c}\left[\frac{(r_r + r_t)(1 - \sin^2\theta\cos^2\phi)}{r_r r_t}\right] \tag{2.5}$$

It is noted that delay is typically expressed relative to the specular point as opposed to absolute propagation delay. To do this the delay at the specular point, given by $(r_t + r_r)/c$ is subtracted from (2.5).

Illustrations of surface delay for a typical GNSS-R geometry at varying incidence angles are provided in in Figure 2.4.

In the general context of radar systems, Doppler shifts refer to the phenomenon relating to the change in frequency of the received signal(s) compared to that of the

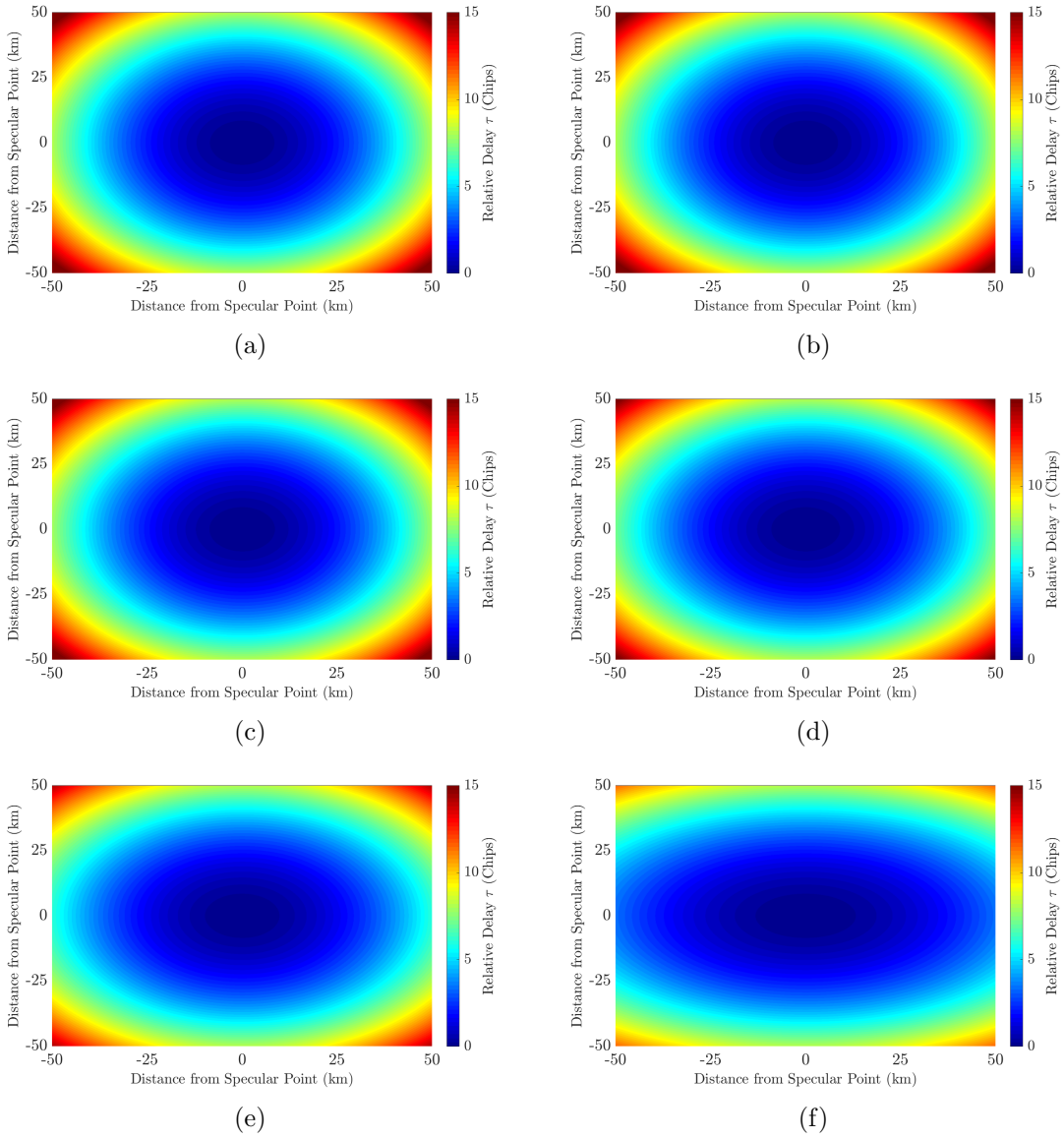


Figure 2.4: Illustration of delay on surface for CYGNSS-like bistatic geometry using the approximate expression developed (a) $\theta_i = 1°$ (b) $\theta_i = 5°$ (c) $\theta_i = 10°$ (d) $\theta_i = 20°$ (e) $\theta_i = 30°$ (f) $\theta_i = 50°$

transmitted signal(s) due to the relative motion between transmitter and target. A similar concept applies here where Doppler frequency shifts arise due to the relative motion between transmitter, receiver, and specular point on earth's surface. The

motion of the transmitter and receiver results in a temporal variation in radial separation(s) between a surface point r and the Tx/Rx pair thereby resulting in a variation in total phase change across the forward propagation path. The rate of change in radial separation (or range) is given by the transmitter and receiver velocity vectors. In practice, the velocity of the transmitter and receiver are not identical and therefore transmitter and receiver velocities are defined using (2.6, 2.7).

$$\bar{V}_T = \frac{d\bar{r}_t}{dt} \tag{2.6}$$

$$\bar{V}_R = \frac{d\bar{r}_r}{dt} \tag{2.7}$$

such that the Doppler shift f_D can be approximated by (2.8).

$$f_D = \frac{1}{\lambda}\left[\bar{V}_R \cdot \hat{k}_s - \bar{V}_T \cdot \hat{k}_i\right] \tag{2.8}$$

where λ is the electromagnetic wavelength ($\approx$ 19 cm for the L1 component of the GPS signal), k is the free space propagation constant, and $\hat{k}_i$ and $\hat{k}_s$ are the incident and scattering propagation unit vectors given by (2.9), (2.10).

$$\hat{k_i} = \frac{\bar{r} - \bar{r}_T}{|\bar{r}_T - \bar{r}|} \tag{2.9}$$

$$\hat{k_s} = \frac{\bar{r}_R - \bar{r}}{|\bar{r}_R - \bar{r}|} \tag{2.10}$$

Using (2.8)-(2.10), the Doppler shift can be approximated as (2.11).

$$f_D = \frac{1}{\lambda}\Bigg(\cos\theta\,(V_{rz} + V_{tz}) + \sin\theta\,(V_{rx} - V_{tx}) + r\Bigg(\sin\theta\cos\phi\Bigg(\frac{V_{rx}\sin\theta + V_{rz}\cos\theta}{r_r} + \frac{V_{tx}\sin\theta + V_{tz}\cos\theta}{r_t}\Bigg) - \frac{V_{rx}\cos\phi + V_{ry}\sin\phi}{r_r} - \frac{V_{tx}\cos\phi + V_{ty}\sin\phi}{r_t}\Bigg)\Bigg) \tag{2.11}$$

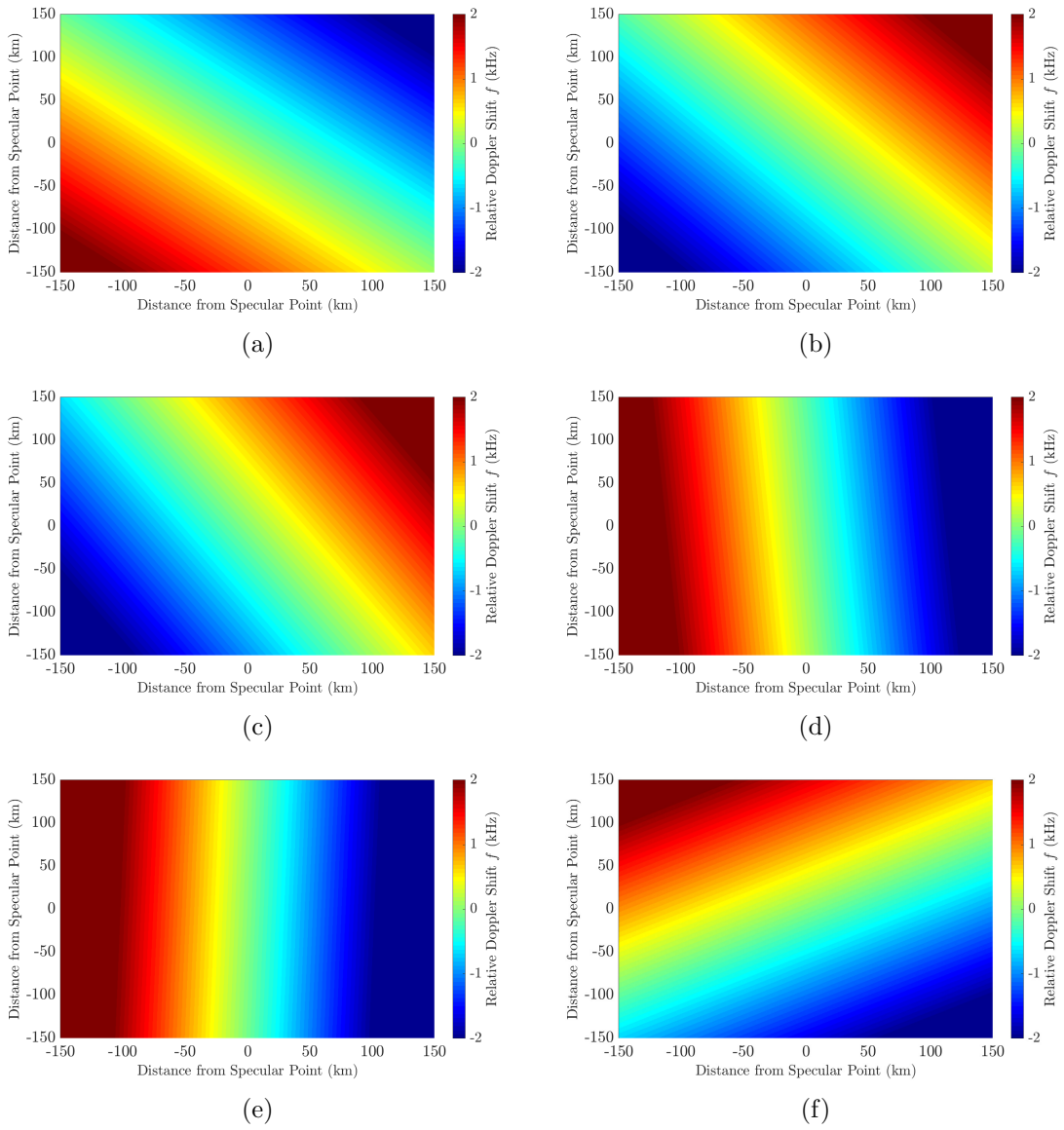


Figure 2.5: Illustration of magnitude of Doppler on surface for CYGNSS-like bistatic geometry using the approximate expression developed. They are based on actual CYGNSS geometries reported in Level-1 measured data (a) $\theta_i = 1^\circ$ (b) $\theta_i = 5^\circ$ (c) $\theta_i = 10^\circ$ (d) $\theta_i = 20^\circ$ (e) $\theta_i = 30^\circ$ (f) $\theta_i = 50^\circ$

Illustrations of surface Doppler shifts for typical GNSS-R geometries are provided in Figure 2.5.

The delay patterns on the surface represent ellipses of constant delay (iso-Delay lines) and the Doppler patterns represent constant Doppler hyperbolas (iso-Doppler lines). The scattering from each point on the surface is therefore associated with a delay-Doppler value the knowledge of which facilitates mapping power arising from a certain part of the observed footprint to delay-Doppler space. To do this, the received signal is cross-correlated with a reference signal (direct or locally generated) at varying time and frequency offsets as shown in (2.12), thereby giving rise to a 2-dimensional amplitude distribution $Y(\tau, f_D)$, mapping components of the total received signal into their respective delay τ and Doppler f_D bins as shown in (2.12):

$$Y(\tau, f_D) = \int_0^{T_i} a(t) \cdot r^*(t+\tau) e^{j2\pi f_D t} dt \tag{2.12}$$

where T_i is the coherent integration time, $r(t)$ is the reflected signal and $a(t)$ is the reference signal. Descriptions of the received power at a given delay and Doppler shift, vary in accordance with the properties of the observed surface. For sufficiently rough surfaces the bistatic radar equation [67,68] is typically relied upon for estimates of the incoherent component of the received power and is given by (2.13).

$$P_R^{inc}(\tau, f) = \frac{P_T \lambda^2}{(4\pi)^3} \int_A \frac{G_T G_R \sigma_0 \Lambda^2(\tau - \tau') S^2(f - f')}{|\overline{r}_T - \overline{r}|^2 \, |\overline{r}_R - \overline{r}|^2} d\overline{r} \tag{2.13}$$

where P_R^{inc} is the incoherent component of received power at delay and Doppler offsets τ and f from the specular point describing fields scattered with random phases, σ_0 is the surface normalized bistatic radar cross section (BRCS), and the $\frac{1}{R^2}$ terms account for the fall off of power with range on both the transmit and receive paths. Finally the product of the functions Λ and S are the delay and Doppler spreading functions. For ocean surfaces, σ_0 can be modeled using the geometric optics (GO) limit of the

Kirchhoff approximation previously described and is given by (2.14).

$$\sigma_0 = \pi \Gamma_{LR} \left(\frac{q_\perp}{q_z} \right)^2 P(\vec{s}) \tag{2.14}$$

where Γ_{LR} is the surface's Fresnel reflectivity in polarization LR (right hand circular incidence and left hand circular scattering for CYGNSS), $\vec{q}$ is the scattering vector and $P(\vec{s})$ slopes PDF assumed to follow an anisotropic bivariate Gaussian distribution. where $\vec{s} = \vec{q}_\perp / q_z$. The slope PDF is given by (2.15).

$$P(\vec{s}) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-b_{x,y}^2}} \times \exp\left(-\frac{1}{2\left(1-b_{x,y}^2\right)} \left(\frac{s_x^2}{\sigma_x} + \frac{s_y^2}{\sigma_y} - \frac{2b_{x,y}s_x s_y}{\sigma_x\sigma_y} \right) \right) \tag{2.15}$$

where $\sigma_{x,y}$ denotes the surface mean square slope (MSS) in an x-y coordinate frame and $b_{x,y}$ is the correlation between x-y slope components.

The ratio $\vec{q}_\perp / q_z$ is indicative of the extent of surface tilt necessary to direct scattered power towards the receiver and its magnitude can be approximated using (2.16).

$$q_\perp / q_z = \frac{\frac{r(r_t+r_r)}{r_t r_r}\sqrt{\cos^4\theta \cdot \cos^2\phi + \sin^2\phi}}{2\cos\theta + \frac{r(r_t-r_r)}{r_t r_r}\sin\theta \cdot \cos\theta \cdot \cos\phi} \tag{2.16}$$

The formulation above provides the fundamental steps necessary in the processing of the received signals giving rise to the fundamental GNSS-R measurement, the delay-Doppler Map.

2.3 Delay Doppler Maps and GNSS-R Observables

The fundamental GNSS-R measurement, the delay-Doppler Map, arises due to the mapping of received power from the spatial domain to delay-Doppler space through the processes described in the previous section. It is however not possible to produce a useful DDM using a single 'snap shot' measurement due to two closely related

considerations. Firstly, the power levels received are on the order of 1×10^{-18} W or smaller and are therefore comparable in magnitude to the measurement's noise floor. For this reason, coherent integration of the DDM is conducted as a means to allow sufficient signal build up. The coherent integration time dictates the duration over which this is done which in turn is determined by the time the surface is expected to be 'frozen'. That is, the period over which the observed footprint and its properties do not change appreciably such that scattering is highly correlated. Exact determination of this duration depends on the observed surface and the speed of the receiving platforms. Typically, tests of empirical performance are relied upon for setting this duration and for spaceborne platforms this is on the order of 1 ms [69]. Secondly, the scattered power is strongly influenced by speckle noise arising due to the random constructive and destructive interferences by the different scattering facets across the observed footprint. A single 'look' is therefore not expected to produce a discernible signal. An example of this is shown in Figure 2.6.

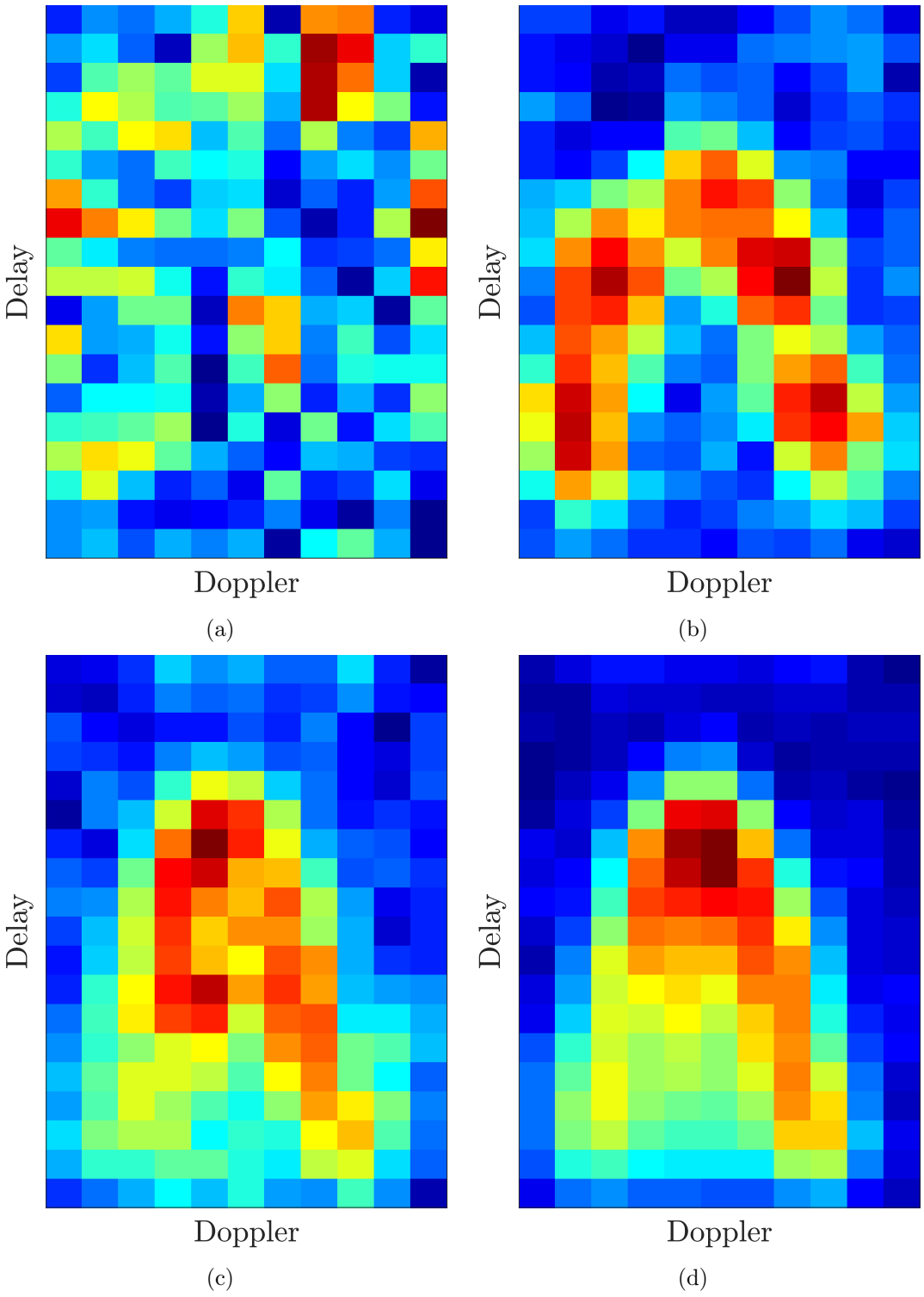


Figure 2.6: Delay Doppler maps formed as the result of a varying number of looks, explaining the need for incoherent summation (a) 1 look (b) 50 looks (c) 100 looks (d) 500 looks

Since beyond 1 ms, for spaceborne receivers, the scatter is assumed to have reached a decorrelation point multiple looks are incoherently summed to leverage the destructive interferences of uncorrelated noise. This effectively, 'beats down' noise while preserving the signal. Typically, 1000 looks are summed for spaceborne receivers thereby giving rise to a 1 s DDM measurement (1 ms $\times$ 1000) an example of which is shown in Figure 2.7.

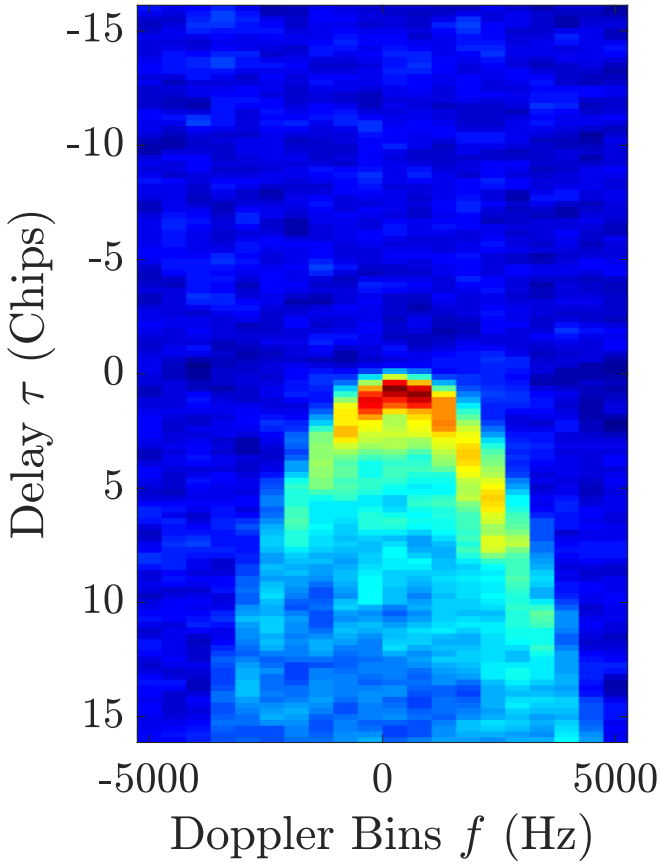


Figure 2.7: Measured delay-Doppler map that underwent 1 ms coherent integration and 1000 look incoherent summation over the ocean surface

It is noted that delay-Doppler bins associated with negative (relative to specular) delay typically include noise power and do not contain any elements of signal power. This is because these are associated with propagation paths that are shorter than that of the specular point. For ocean surfaces, off which the DDM in Figure 2.7 was measured, mean surface height/elevation is not expected to substantially depart from some mean sea level value, and therefore no physical propagation path allowing for the mapping of power into negative delay bins exists. It is noted, that in the use of GNSS-R measurements for land and ocean remote sensing numerous previous studies have focused on the use of 'delay waveforms' as opposed to the DDM [70–74]. This refers to the reduction of the 2-dimensional DDM measurement to a 1-dimensional waveform either by considering a single delay 'cut' at $f_D = 0$ Hz or by summing power values along the Doppler dimension. An example of this is shown in Figure 2.8.

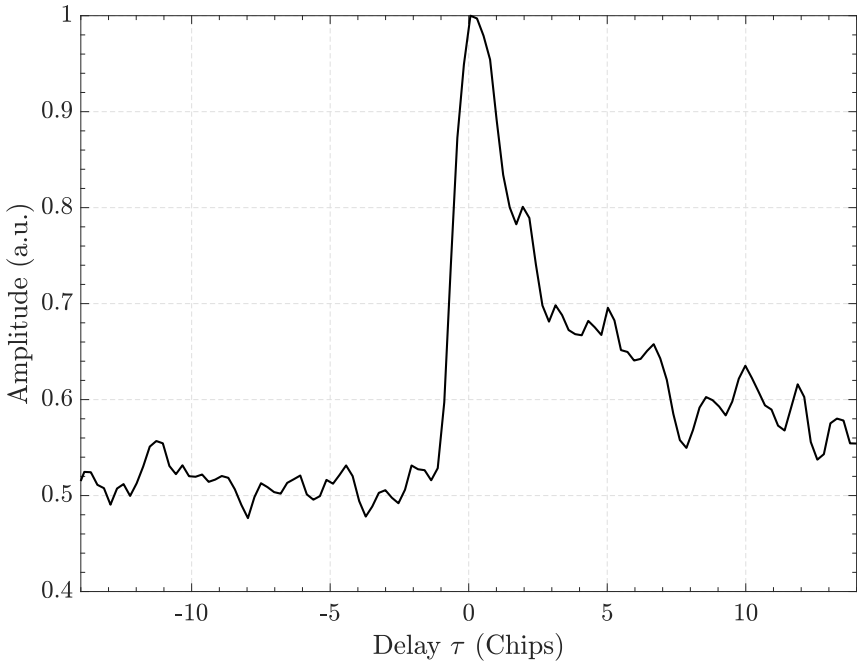


Figure 2.8: Power-delay waveform derived from delay-Doppler map

Studies attempting to then derive geophysical parameters of interest using GNSS-R data, are based on a series of observables derived from the whole of or part of the DDM. Those are summarized as follows:

- The maximum amplitude of the delay-Doppler Map, the careful calibration of which may be indicative of surface properties of interest.
- The DDM maximum normalized by the average noise floor value, also commonly referred to as the peak-to-mean ratio or Signal-to-Noise Ratio (SNR).
- The total or mean value of DDM power within a window of preset extent in delay and Doppler about the specular bin.
- The distance in delay for the DDM maximum to decay by a preset percentage.
- The offset center of gravity describing the distance in delay and Doppler the specular bin is away from the bin containing the DDM maximum.
- Leading edge slope, describing the absolute value of or a ratio of the slope of the line before the specular bin to that after the specular bin.
- Trailing edge slope, describing the absolute value of or a ratio of the slope of the line after the specular bin to that before the specular bin.

Variations in the magnitudes, shapes or patterns of any combination of these observables may be used as part of GNSS-R remote sensing schemes to estimate surface properties that may be of interest. For example, increases in the calibrated DDM peak value, either within a preset delay-Doppler window range or relative to the noise floor (SNR), over land may be indicative of increased surface volumetric

soil moisture content or the presence of standing water. On the other hand, a similar increase over the ocean is indicative of lower surface wind speed values, as more power gets reflected towards the GNSS-R receivers from the calmer ocean surface. Similarly, the slopes of power-vs-delay lines from/to the specular bin are telling of a wide range of surface properties that may be of interest. For example, steeper slopes are indicative of smoother surfaces as the bulk of received power, in delay-Doppler space, gets concentrated about the DDM peak and little, relative, power is received from facets at larger distances from the specular point. Conversely, smoother undulations away from the DDM peak are indicative of rougher surfaces with more power received from the larger footprint corresponding to the size of the glistening zone. Over the ocean, either, corresponds directly to the underlying roughening mechanism, surface wind speeds.

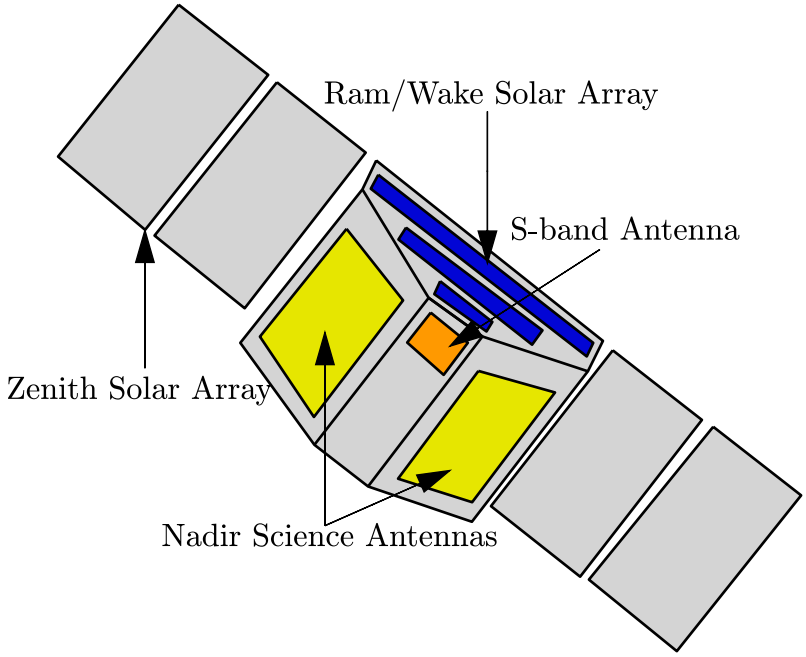
Nonetheless, the ability to rely exclusively on these general trends for the retrieval of geophysical quantities of interest, on a global scale, is often complicated by the confounding effects of a wide range of surface properties as well as calibration uncertainties. Robust retrieval algorithms must therefore be aware of the practical applicability of these general trends and remain cognisant of their limitations, the bulk of which are outlined in subsequent chapters. The development of retrieval algorithms that are able to leverage the correspondence of various manifestations related to surface properties within the measurements' delay-Doppler space, whilst minimizing the ambiguities introduced by the co-current confounding influences of different surface properties, need for additional ancillary information and resilience to calibration uncertainties will be of most interest in the work presented in this dissertation.

Chapter 3: Cyclone Global Navigation Satellite System Mission

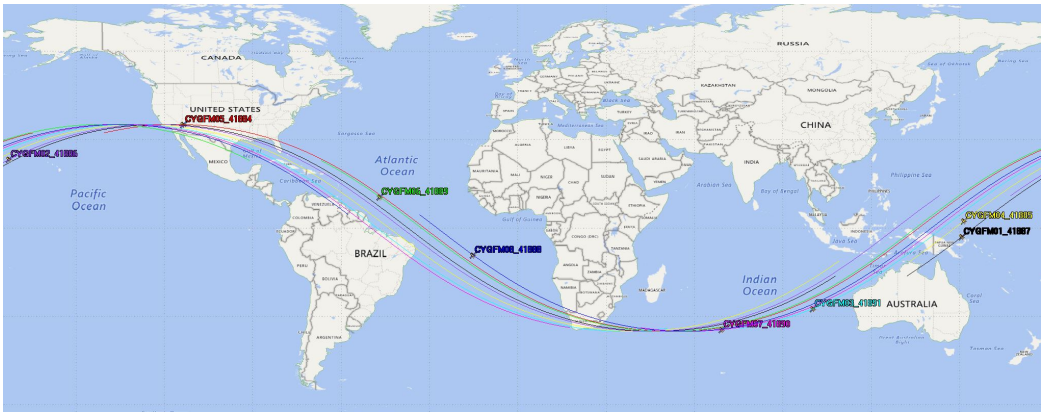
The proposed land and ocean remote sensing algorithms are transitioned from concept to operation in the context of the Cyclone Global Navigation Satellite System Mission. This chapter provides highlights of the most pertinent aspects of the mission for subsequent discussions.

3.1 Mission Overview

The Cyclone Global Navigation Satellite System (CYGNSS) is NASA's Earth Venture mission designed to support the investigation of emergent scientific priorities using orbital and sub-orbital instruments that are built, tested and launched over short incubation times and reduced cost. It was launched on December 15, 2016 and comprises a constellation of eight satellites receiving reflected GPS signals off the Earth's surface. Examples of a CYGNSS receiver and the constellation's orbital configuration are depicted in Figure 3.1.



(a)



(b)

Figure 3.1: Illustration of CYGNSS individual satellite and constellation (a) Bottom view of CYGNSS space vehicle (b) Hour long orbital configuration of CYGNSS constellation

Its operation is based on GNSS-R where CYGNSS together with GPS satellites form a bistatic radar geometry with GPS satellites acting as transmitters and

CYGNSS satellites acting as receivers. The constellation was placed at an orbital inclination of 35° offering global coverage for all longitudes and for latitudes within ±38°. The latter limit is due to the mission's primary focus on ocean winds and tropical cyclones which are most active between the aforementioned latitudes. Each CYGNSS space vehicle (SV) is equipped with an upward looking zenith antenna and two downward looking nadir antennas. The right hand circularly polarized zenith antenna is used to receive direct GPS signals used for positioning purposes and its utility has been more recently extended for improved absolute power calibration of CYGNSS measurements through the real-time monitoring of the transmitters' Equivalent Isotropically Radiated Power (EIRP) [75]. Each nadir 'science' antenna has a peak gain of $\sim$ 15 dB capable of tracking 4 simultaneous reflections such that at the original coherent integration time of 1 ms, the eight satellite constellation is capable of providing 32 DDM measurements per second. While the mission's primary focus is concerned with wind speed estimation over ocean surfaces, more recently the incoherent integration time was reduced to 0.5 s in support of land applications investigations such that the constellation now provides 64 DDM measurements at any given second.

CYGNSS's measurements occur as "tracks" on land and sea surfaces that can extend in excess of one thousand kilometers in length. The exact revisit times can be described statistically [76] in terms of their mean and median and vary in accordance with with latitude and the resolution of the spatial grid used. Under nominal operating conditions, CYGNSS provides median and mean revisit times of approximately 3 and 7 hours respectively.

3.2 Standard Products

As part of its official and standard data release the CYGNSS mission provides three levels of data products, Levels 1, 2 and 3 (L1, L2, L3). The L1 release aims to provide the fundamental quantities from which other geophysical and quantities of interest are derived beginning with the L0 DDMs in uncallibrated instrument 'raw counts'. These DDMs are cropped from a larger 'full' DDM formed on board the CYGNSS delay-Doppler Mapping Instrument (DDMI) and have a 17 bin extent in delay (with a ≈ 0.25 μs bin resolution) and an 11 bin extent in Doppler (with a $\approx$ 500 Hz bin resolution) [48]. The first 3-4 delay rows typically comprise thermal noise and therefore a DDM provided as part of the L1 data has a maximum delay extent of 3.5 chips, on average. The correspondence this has with the measurements' spatial footprint is explored in the next section. The L0 DDMs later undergo two levels of calibration [77, 78], where they are linearly regressed against received power. The calibration aims to convert L0 raw counts to power in watts and later to Bistatic Radar Cross Section (BRCS) through accounting for the effects of thermal noise, instrument gain and propagation losses. For a more comprehensive overview see [79]. In addition to the multiple versions of DDMs in uncalibrated counts, W and m^2 other engineering quantities of interest are included in L1 data which are not typically included in the higher level data products.

Level 2 data emphasizes the conversion of quantities derived from DDMs to ungridded wind speed and mean square slope estimates. To do this the Normalized Radar Cross Section (NRCS) derived from DDMs using the delay-Doppler Map Average (DDMA) [80–82] defined as an integration of the Bistatic Radar Cross Section (BRCS) within a predefined delay-Doppler windows about the specular DDM bin

normalized by the effective scattering area, is regressed against wind speeds provided by other instruments and/or models to derive a Geophysical Model Function (GMF). At a fundamental level, the GMF is therefore a function that expects an NRCS, or a related observable quantity as its input, and provides wind speed estimates as its output. A more comprehensive overview of the factors dictating the formulation and tuning of the CYGNSS GMFs is explored in [83]. Two types of L2 wind speed estimates are provided. The first is based on the Fully Developed Seas (FDS) which relates to a situation where the sea surface has reached a steady state response and wave age has sufficiently 'matured'. FDS estimates are generally capped at 30 m/s. The second is the Young Seas Limited Fetch (YSLF) and aims to describe a turbulent sea surface that has not fully developed through the matching of CYGNSS observables to ocean surface wind speed measurements provided by the stepped frequency microwave radiometer (SFMR) on board NOAA's hurricane hunter aircraft [84]. This GMF provides estimates that are not capped at 30 m/s and provides wind speed estimates for scenarios where higher wind speeds are expected to be observed (i.e. tracks overpassing hurricanes). Estimates of MSS are also provided by leveraging the inverse dependence between it and NRCS [85].

The data that Level 3 provides is limited to hourly wind speed maps within CYGNSS' coverage on a $0.2^\circ \times 0.2^\circ$ uniform grid, and pixel latitude and longitude. Due to CYGNSS' revisit times, each pixel can be associated with multiple wind speed estimates. The binning of those is done in such a way that produces a minimum variance estimate of mean wind speed within a given spatial and temporal interval [86]. An example of a daily CYGNSS L3 wind map is depicted in Figure 3.2

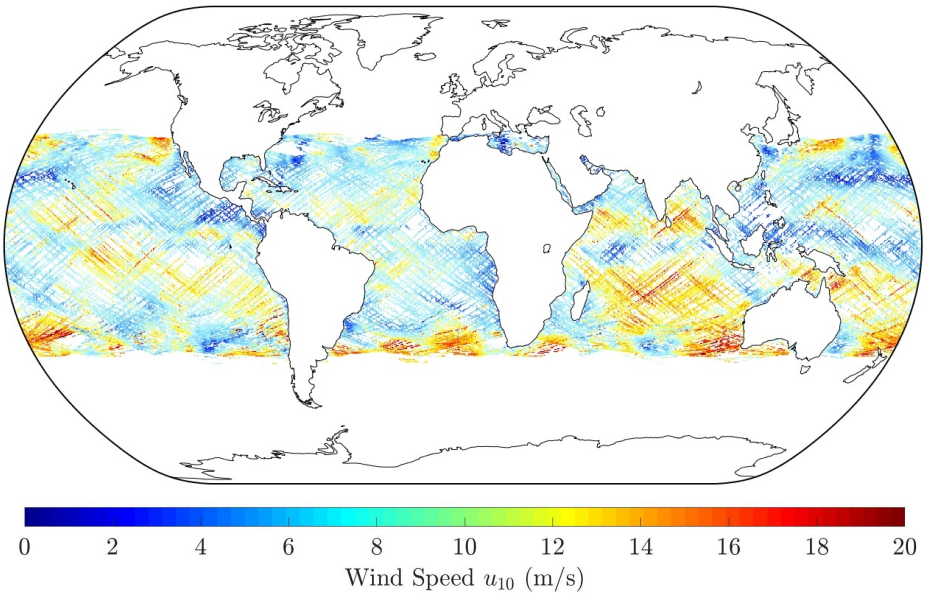


Figure 3.2: Illustration of mean Level-3 winds observed by the CYGNSS constellation on DOY-Year 159-2018 on $0.2° \times 0.2°$ uniform latitude/longitude grid

3.2.1 Theoretical Correspondence of Observables to Spatial Footprint

An important question that merits significant contemplation when attempting to use GNSS-R observables in the retrieval of geophysical parameters of interest relates to the spatial scale a given retrieved quantity is representative of. That is the range resolution defining the scale that relates the correspondence of GNSS-R observables to the spatial footprint. The spatial extent of a given quantity is largely dictated by the delay extent of the DDM, or DDM subset, from which it was derived. From Figure 2.4 it is clear that the trace of iso-Delay lines can be described by concentric ellipses defined by a semi-major axis a and semi-minor axis b. Using (2.5) the equations

describing a and b can then be shown to be (3.1), (3.2).

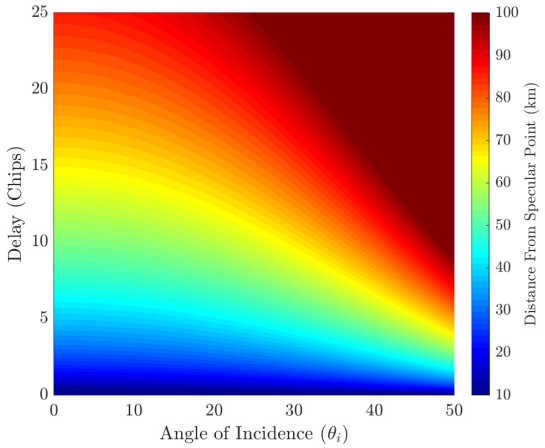
$$a = \sqrt{\frac{2\,c\,\tau_D\,R_r\,R_t\,sec^2\theta_i}{(R_r + R_t)}} \tag{3.1}$$

$$b = a\cos\theta_i \tag{3.2}$$

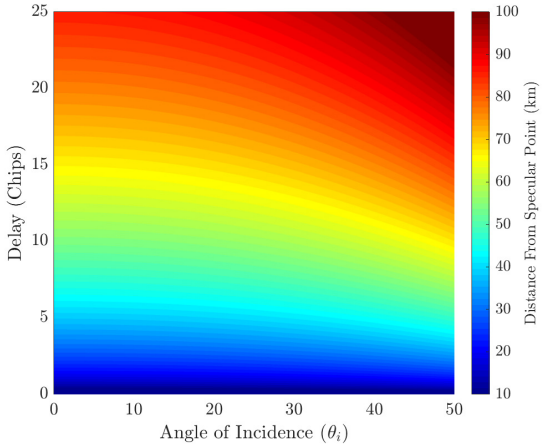
where τ_D is the excess delay relative to the specular path, θ_i is the incidence angle, R_r is the range from the receiver to the specular point and R_t is the range from the transmitter to the specular point. An effective diameter of the delay ellipse d can further be defined as d given by (3.3).

$$d \approx 2 \times \sqrt{a \cdot b} \tag{3.3}$$

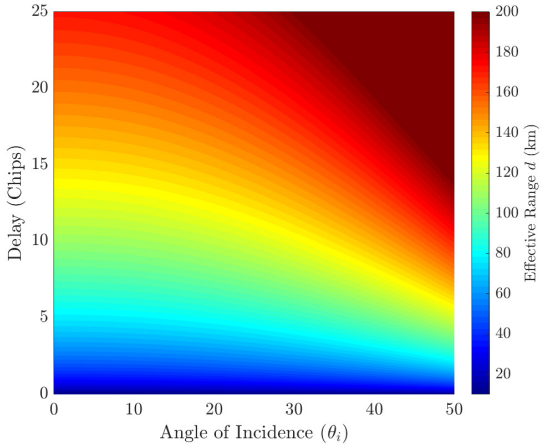
where the effective diameter describes an equivalent effective square range cell with vertices of length d. The response of a, b and d to a wide range of delays and incidence angles for a CYGNSS-like LEO orbit (elevation $\approx$ 520 km) is summarized in Figure 3.3.



(a)



(b)



(c)

Figure 3.3: (a) Ellipse major axis a (b) Ellipse minor axis b (c) Effective range d

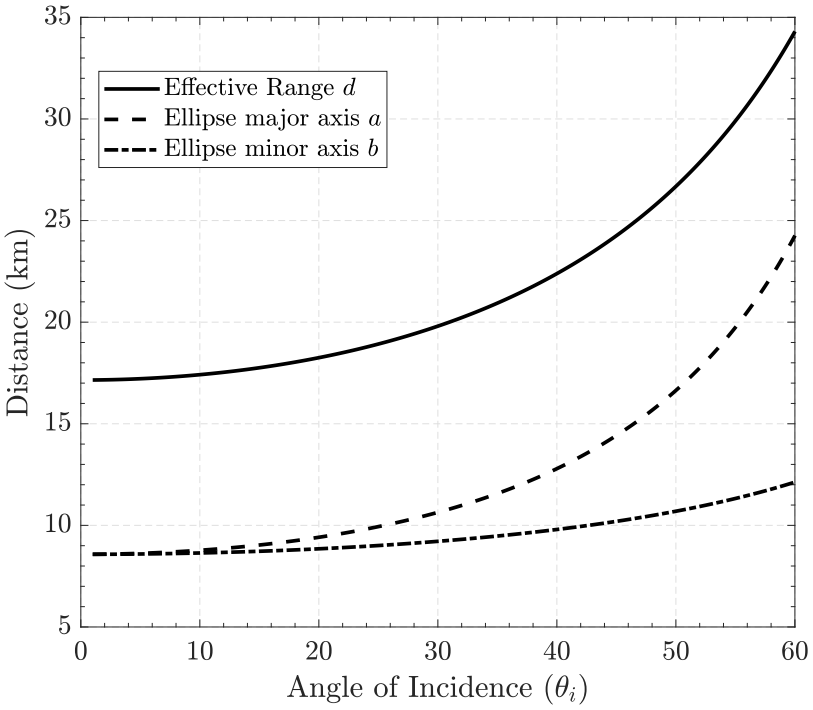


Figure 3.4: Illustration of angle dependence of constant delay ellipse at 0.25 chip delay

For larger delays the size of the iso-delay ellipses is expected to increase with the same applying to incidence angles. Following the same formulation, the size of a, b and d for the same delay increment used in CYGNSS DDMs ($\tau = 0.25$ μs) across a wide range of 'typical' CYGNSS incidence angles can be obtained and is shown in Figure 3.4

From this it follows that the reporting of CYGNSS DDMs at delay increment of 0.25 chips suggests that it is capable of resolving features as small as 17×17 km on the surface at nadir with a growth that can reach values as large as 34×34 km at an incidence angle of $\theta_i = 60°$ per delay pixel.

3.2.2 Practical Correspondence of Observables to Spatial Footprint

Practically, it is expected the there exists interbin correlation within DDMs suggesting an inability to resolve features unambiguously corresponding to a $\tau = 0.25$ μs scale and a more realistic estimate is expected to be coarser. Quantification of the extent of correlation using delay waveforms has received a nuanced and rigorous treatment in the works of [87–89]. A simpler Monte Carlo based approach is used here to describe the correlation.

A large number of DDM realizations are produced and a correlation matrix is constructed using (3.4):

$$\rho_{\tau_i f_j, \tau_m f_n} = \frac{\Sigma[P(\tau_i, f_j), P(\tau_m, f_n)]}{\sqrt{\Sigma[P(\tau_i, f_j), P(\tau_i, f_j)] \cdot \Sigma[P(\tau_m, f_n), P(\tau_m, f_n)]}} \tag{3.4}$$

where $\rho_{\tau_i f_j, \tau_m f_n}$ denotes correlation of a DDM pixel at delay bin τ_i and Doppler bin f_j with a pixel at delay bin τ_m and Doppler bin f_n such that $i, m = 1 \ldots. N_\tau$ and $j, n = 1 \ldots. N_f$ where N_τ and N_f are the number of delay and Doppler bins within the DDM respectively, $P(\tau_i, f_j)$ is a DDM pixel at a delay τ_i and Doppler f_j. Note that $\Sigma[...]$ denotes covariance and for $\Sigma[P(\tau_i, f_j), P(\tau_i, f_j)]$ this reduces to the pixel's variance the square root of which is the standard deviation σ_{τ_i, f_j}.

It follows that for a DDM of N_τ delay bins and N_f Doppler pixels a correlation matrix $\left[\rho_{\tau_i f_j, \tau_m f_n}\right]$ of dimension $N_\tau N_f \times N_\tau N_f$ can be constructed as shown in

expression (3.5).

$$\begin{bmatrix} 1 & \frac{\Sigma[P(\tau_1,f_1),P(\tau_1,f_2)]}{\sigma_{\tau_1,f_1}\times\sigma_{\tau_1,f_2}} & \cdots & \frac{\Sigma[P(\tau_1,f_1),P(\tau_2,f_1)]}{\sigma_{\tau_1,f_1}\times\sigma_{\tau_2,f_1}} & \cdots & \frac{\Sigma[P(\tau_1,f_1),P(\tau_{N_\tau},f_{N_f})]}{\sigma_{\tau_1,f_1}\times\sigma_{N_\tau},f_{N_f}} \\ \frac{\Sigma[P(\tau_1,f_2),P(\tau_1,f_1)]}{\sigma_{\tau_1,f_2}\times\sigma_{\tau_1,f_1}} & 1 & \cdots & \frac{\Sigma[P(\tau_1,f_2),P(\tau_2,f_1)]}{\sigma_{\tau_1,f_2}\times\sigma_{\tau_2,f_1}} & \cdots & \frac{\Sigma[P(\tau_1,f_2),P(\tau_{N_\tau},f_{N_f})]}{\sigma_{\tau_1,f_2}\times\sigma_{N_\tau},f_{N_f}} \\ \vdots & \ddots & 1 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & 1 & \vdots \\ \frac{\Sigma[P(\tau_{N_\tau},f_{N_f}),P(\tau_1,f_1)]}{\sigma_{N_\tau},f_{N_f}\times\sigma_{\tau_1,f_1}} & \cdots & \cdots & \cdots & \cdots & 1 \end{bmatrix} \tag{3.5}$$

Note that practically $\Sigma[P(\tau_i, f_j), P(\tau_m, f_n)]$ is computed as:

$$\Sigma[P(\tau_i, f_j), P(\tau_m, f_n)] = \mathbf{E}\left[P(\tau_i, f_j) \cdot P(\tau_m, f_n)\right] - \mathbf{E}\left[P(\tau_i, f_j)\right]\mathbf{E}\left[P(\tau_m, f_n)\right] \tag{3.6}$$

where for a given pair of pixels the expected value $\mathbf{E}(XY)$ is computed as the mean of the product of the realizations of the two pixels X and Y while the expected value $\mathbf{E}$[X], $\mathbf{E}$[Y] is simply the mean of the realizations for a given pixel. Together with (3.5), this captures the extent of correlation of a given pixel with any other pixel within the DDM. From the resulting correlation matrix $[\rho_{\tau_i f_j,\tau_m f_n}]$ 'correlation DDMs' are constructed for the purposes of illustrating the extent of correlation between a pixel and other delay-Doppler pairs across a given DDM. Examples of this together with delay/Doppler correlation cuts are depicted in Figure 3.5.

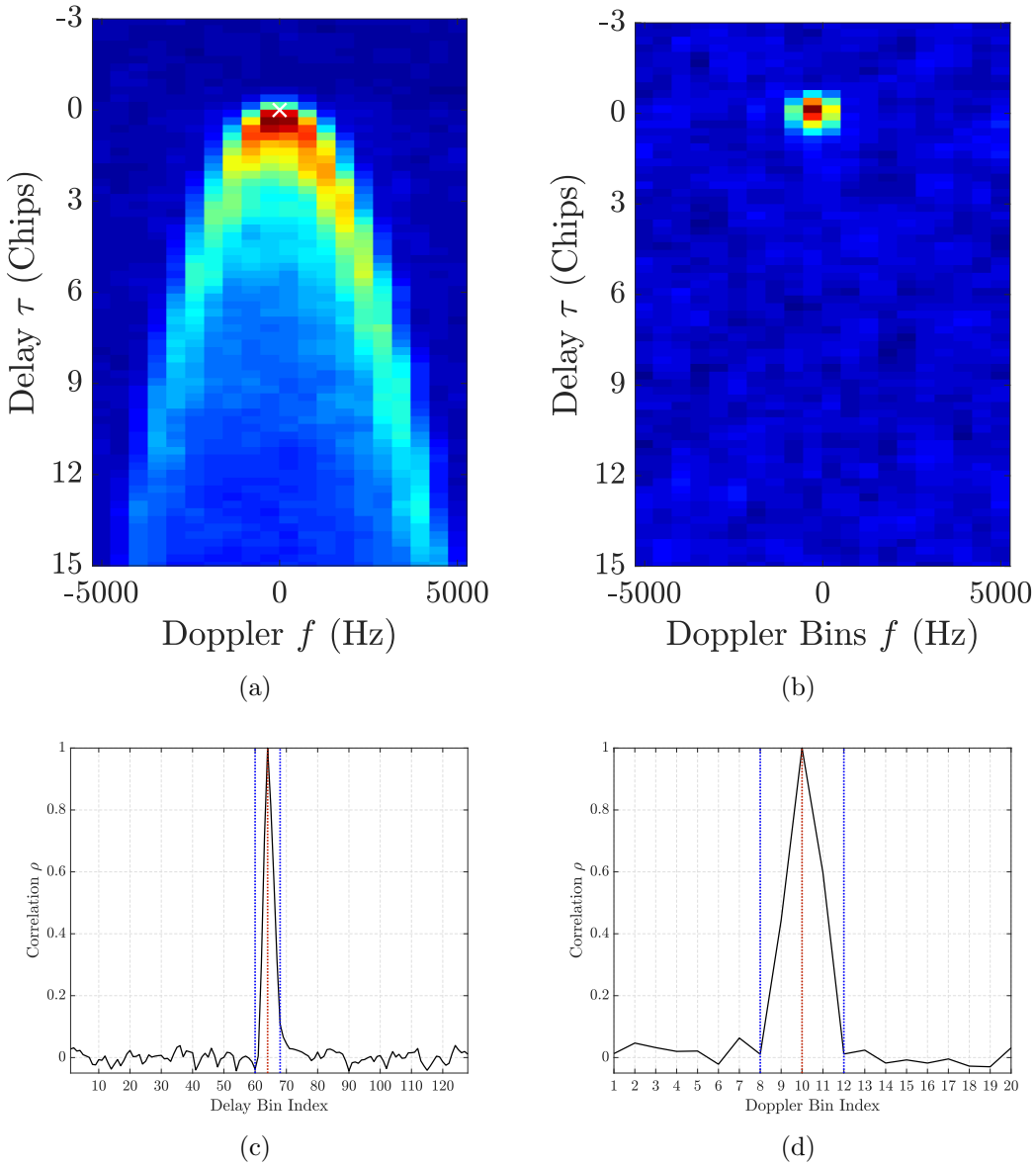


Figure 3.5: (a) DDM with pixel at $\tau = 0$ μs and Doppler $f = 0$ kHz (b) Correlation with other pixels within the DDM (c) Delay cut of correlation (d) Doppler cut of correlation

The results show that, consistently, a point of decorrelation is reached beyond $\pm(\tau_i - \tau_j) > 0.75$ μs or equivalently $\pm$ 3 delay bins. A point of decorrelation in

Doppler is reached beyond $\pm(f_i - f_j) > 1$ kHz or equivalently $\pm$ 2 Doppler bins. Note that the point of decorrelation is defined in delay as $\rho_{\tau_i,\tau_i}/e$ and in Doppler $\rho_{f_j,f_j}/e$ (i.e. the point(s) where correlation decays by a value of $1/e$ of the autocorrelation). Based on these results a model that describes inter-pixel correlation is given by (3.7).

$$\rho(\tau - \tau_0, f - f_0) \approx \Lambda^2(\tau - \tau_0) \cdot \Lambda^2(f - f_0) \tag{3.7}$$

where

$$\Lambda^2(\tau - \tau_0) = \begin{cases} \left(1 + \frac{\tau - \tau_0}{\tau_c}\right), & -\tau_c < \tau - \tau_0 \leq 0 \\ \left(1 - \frac{\tau - \tau_0}{\tau_c}\right), & 0 < \tau - \tau_0 \leq \tau_c \\ 0, & \text{otherwise} \end{cases} \tag{3.8}$$

and

$$\Lambda^2(f - f_0) = \begin{cases} \left(1 + \frac{f - f_0}{B}\right), & -B < f - f_0 \leq 0 \\ \left(1 - \frac{f - f_0}{B}\right), & 0 < f - f_0 \leq B \\ 0, & \text{otherwise} \end{cases} \tag{3.9}$$

where $B = 1/T_i$ and T_i is the coherent integration time of 1 ms. The correlation behaviour is similar to what one might expect to see when analyzing delay-Doppler behaviour of the CYGNSS Woodward Ambiguity Function (WAF) $\chi^2(\tau, f_D)$.

The inter-bin correlation therefore highlights the fact that the information within a set of adjacent DDM pixels does not furnish 'unique information' about the surface at $\tau = 0.25$ μs, $f = 0.5$ kHz bin size. Together with preceding results, this can be interpreted as an oversampling of the surface in delay by at least 3 folds and an oversampling in Doppler by at least 2 folds. A more realistic estimate of CYGNSS' spatial sampling capability is therefore at $\tau = 0.75$ μs where the angle dependent behaviour of a, b and d is now summarized in Figure 3.6.

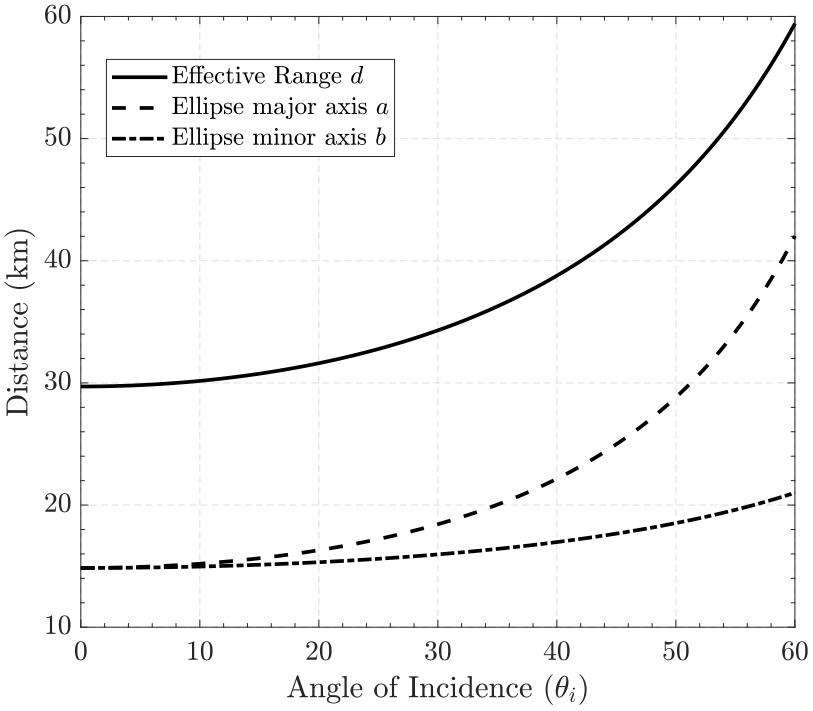


Figure 3.6: Illustration of angle dependence of constant delay ellipse at 0.75 chip delay

This suggests that a 'true' estimate of CYGNSS' spatial sampling capability ranges between an area of 30×30 km and 60×60 km.

Now to address the relevance of this to the geophysical quantities derived from CYGNSS L1 DDMs. The NRCS used to obtain L2 wind speeds derives from a DDMA with an area comprising 3 delay bins (or equivalently 0.75 μs in delay) and 5 Doppler bins (or equivalently 2 kHz in Doppler) centered about the specular bin. It follows that the spatial resolution for the DDMA ranges from 30×30 km and 60×60 km depending on incidence angle and the retrieved geophysical parameters are on the same spatial scale. It is also noted that practically this will be further coarsened due to the orbital

motion of the CYGNSS receivers and the subsequent smearing of the 0.75 μs delay ellipse. The spacecrafts' 6 km/s motion, translates to an 'effective' spatial resolution ranging between 35×35 km and 65×65 km depending on incidence angle. Note that the above formulation assumes diffuse scattering, in the case of coherent reflection the size of the range cell is expected to be on the order of a Fresnel zone smeared by the along track distance travelled by the CYGNSS receivers [90, 91].

Several practical considerations dictate the choice of the DDMA integration window. Some of which include the fact that for the DDMA to be a representative measure of the amount of scattering at and around the specular point, the integration region should be large enough such that the bulk of this power is captured. Increasing the size of the integration region in delay has the potential of improving signal-to-noise ratio (SNR) as more of the power scattered from the surface and mapped into the delay-Doppler space is accounted for but this comes at the cost of a degraded spatial resolution (due to increase in size of delay ellipse) as well as the inclusion of bins which can be mapped to points on the surface known to be less sensitive to wind speed (i.e. points at increasing distances from the specular point) containing no useful information. The choice of 0.75 μs as an integration window in delay strikes a balance between these considerations. On the other hand, two considerations dictate the choice of the Doppler integration window. If the Doppler integration range is too small, the iso-Dopp lines do not completely enclose the 0.75 μs ellipse and only part of the surface scatter arising from the area of interest gets accounted for in the DDMA. An example of this effect is shown in Figure 3.7.

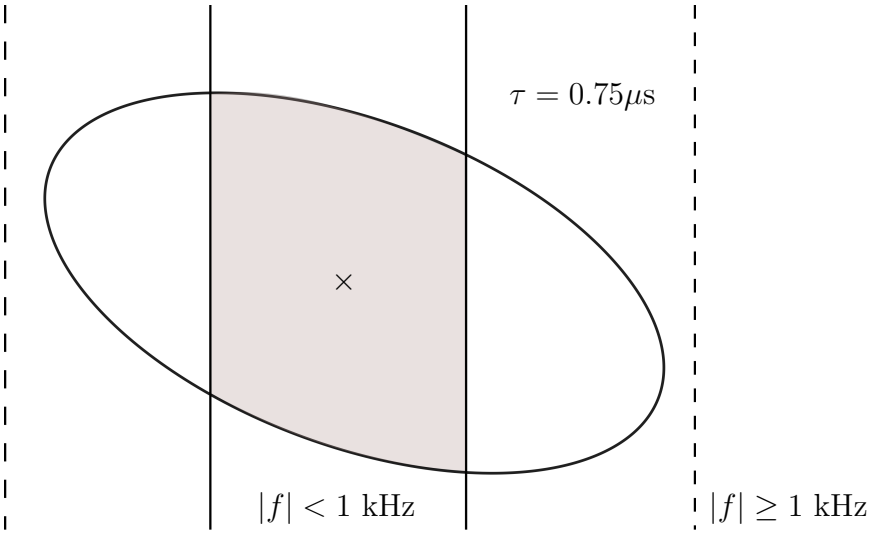


Figure 3.7: Explanation of effects associated with inadequate choice of Doppler integration window for DDMA formation. Shaded area is that which contributes to power mapped into DDM if Doppler window is too small at a given iso-delay ellipse

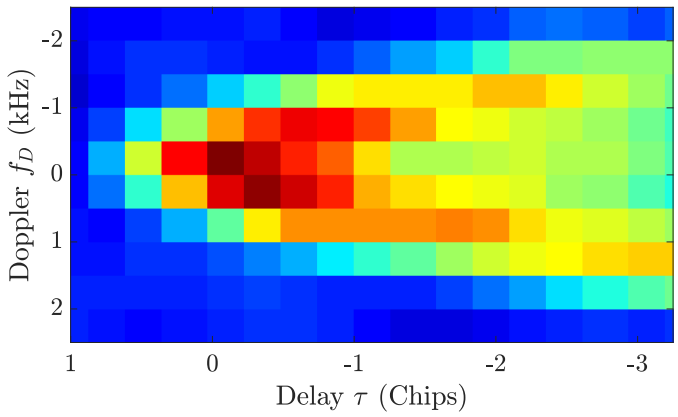
For cases where the choice of Doppler integration range does not completely enclose the 0.75 μs delay ellipse only part of the scattered power from the area of interest is accounted for. On the other hand, averaging over an area too large in Doppler allows for the inclusion of noise bins and risks reducing SNR. For these reasons, a Doppler integration range of 2 kHz strikes a balance between the aforementioned considerations whilst enclosing the 0.75 μs delay ellipse of interest in its entirety for all incidence angles [80] that are typically of interest when using CYGNSS data (namely $\theta_i \leq 60°$).

3.3 Special Modes of Operation

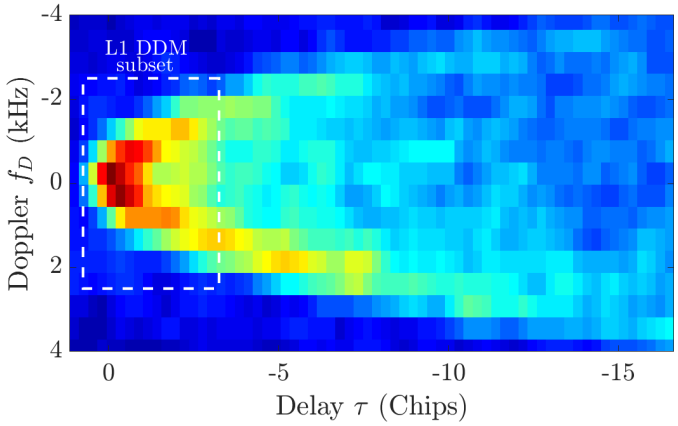
In addition to the standard science mode Levels 1-3 products provided continuously, the CYGNSS mission provides data under two special modes of operation. The

first is called the "Full DDM" and the second is the "Raw Intermediate Frequency" (Raw I/F) mode. Those are provided only over scenes of interest, for example when a CYGNSS receiver observes a scene of particular significance such as a tropical cyclone.

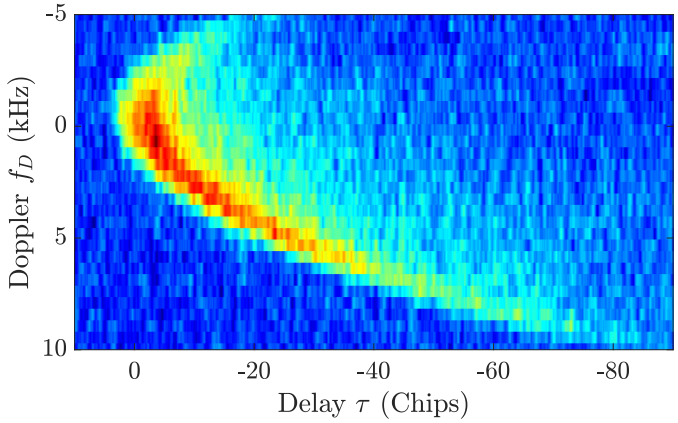
On board the receiver, observed surface scatter is mapped into a DDM comprising 128 delay bins (approximately $\pm$ 32 μs or $\pm$ 16 chips at the $\approx$0.25 μs CYGNSS sampling rate) and 20 Doppler bins (approximately $\pm$5 kHz at the 500 Hz CYGNSS sampling rate). This is the CYGNSS Full DDM, which is downlinked to ground stations when CYGNSS operates in the Full DDM mode. Over the ocean, one half of the delay bins contain receiver noise only with the first 45 delay values used to estimate the measurement's noise floor. A DDMI compression algorithm onboard the receiver then estimates the location of the specular bin as the location of the maximum value of the DDM about which a bounding box is formed. The box has a 17 bin extent in delay (approximately 3.5 chips with the first few delay rows containing thermal noise) and an 11 bin extent in Doppler (approximately $\pm$ 2.5 kHz). This is the L1 DDM, a subset of the Full DDM formed on board the DDMI. Finally, the Raw I/F mode records the CYGNSS intermediate frequency data stream in units of counts before any correlation with the GNSS code or integration. This enables the creation of DDMs of arbitrary extent and sampling in ground processing, at the cost of a greatly increased downlink datarate. It also allows for cross correlations against other GNSS codes as to investigate specular reflections not limited to GPS transmitters. Illustrations of the standard DDM, Full DDM and Raw I/F DDM data products are provided in Figure 3.8.



(a)



(b)



(c)

Figure 3.8: Illustration of different types of CYGNSS DDMs (a) Level-1 DDM observed by CYG01 on DOY-Year 246-2017 over the ocean at 24.41° N 62.56° W (b) Full DDM observed by CYG05 on DOY-Year 280-2017 over the ocean at 22.62° N 95.57° W (c) Raw I/F DDM observed by CYG03 on DOY-Year 249-2017 over the ocean at 13.15° N 65.99° W

The practical implications and potential benefits of using each of the products are explored further in Chapter 7.

3.4 Forward Modelling of CYGNSS Returns

The ability to predict the behaviour of GNSS-R returns in general, and more particularly CYGNSS returns, forms a crucial element of analyzing and understanding the nature of the observed surface and potentially retrieving parameters of interest [183,184]. To do this, predictions of the measurement under varying surface conditions of interest may be produced through the application of the bistatic radar equation described in [67, 68] and given by (2.13). The expected return under varying surface conditions may also be estimated through the use of probabilistic and statistical estimators that describe the expected mean pixel behaviour whilst accounting for co-variations between the different pixels. A description of both approaches is provided in what follows.

3.4.1 End-to-End Simulator

The CYGNSS End-to-End simulator (E2ES) does not use statistical estimators and instead computes simulated CYGNSS observations on a gridded numerical space through the modelling and dynamic orbit propagation of bistatic satellite geometries, GPS transmission and subsequent reflection off the sea surface and evaluation of surface scatter. The E2ES therefore aims to solve for power received due to incoherent scattering from the Earth's surface under the application of the GO limit of the Kirchhoff approximation given by (2.13) within each DDM delay-Doppler bin through the accurate computation and tracking of delay and Doppler at each surface patch used to discretize area surrounding a given specular point. Due to the fact that its

emphasis is on spaceborne observations as opposed to airborne observations [92, 93], it also accounts for the significant Doppler spreading which are expected to be caused by the CYGNSS receivers' ground speed. It also accounts for true orbital orientation at epoch. For a more comprehensive overview see [94–97]. An example of the highly accurate nature of the E2ES simulations of CYGNSS measurements is shown in Figure 3.9.

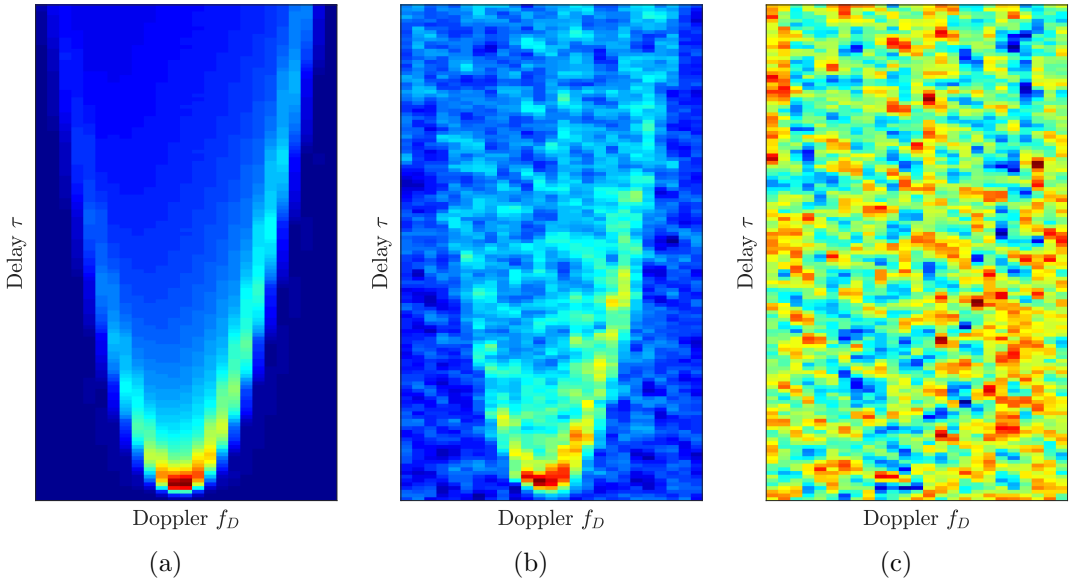


Figure 3.9: Comparisons of E2ES predictions to actual CYGNSS measurements (a) E2ES noise free DDM prediction (b) CYGNSS measurements (c) Noise only residual difference

The relevance of the E2ES to some of the studies conducted in this dissertation and extensions of its functionality are explored further in Chapter 6.

3.4.2 Probabilistic Descriptors

Many estimators proposed in the context of GNSS-R remote sensing, particularly relating to altimetry studies, rely on the precise statistical descriptions of simulated DDMs; those have received rigorous treatment in the studies of [87–89]. While somewhat complicated in terms of their presentations, many of these forward modelling approaches essentially aim to conduct comparisons between a set of measurements and simulated waveforms. When the difference between measurement and simulation is minimized, a retrieval is conducted. To outline the basics of this methodology, simpler descriptors are outlined in this section; ones that use the definition of the DDM pixels' mean vector and covariance matrix to completely define the DDM. This departs from the application of a Gaussian random vector model to the integrated "powers" of the CYGNSS DDM pixels. A Gaussian random vector can be regarded as a multivariate generalization of the Gaussian distribution such that its PDF is described by (3.10).

$$f(\bar{x}) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} e^{-\frac{1}{2}(\bar{x}-\bar{\mu}_x)^T \cdot \Sigma^{-1} \cdot (\bar{x}-\bar{\mu}_x)} \tag{3.10}$$

In the context of CYGNSS' one second measurements, each pixel defined by a delay τ and Doppler f can be described as a Gaussian random variable. For a DDM with N_τ delay bins and N_f Doppler bins, the vector $\bar{x}$ is of dimension $(N_\tau N_f \times 1)$ with each random variable $(x_1, x_2, x_3...)$ representing a DDM pixel with a Gaussian distribution. The vector $\bar{\mu}_x$ represents the expected value of each pixel $\mathbf{E}[P(\tau_1, f_1)...P(\tau_{N_\tau}, f_{N_f})]$ and is the mean value for a given pixel. The covariance matrix is given by Σ with $|\Sigma|$ being its determinant and Σ^{-1} representing its inverse and is of dimension $N_\tau N_f \times N_\tau N_f$.

The power that CYGNSS receives, in the absence of noise effects and calibration uncertainties, can be regarded as being purely a function of surface roughness state. As a consequence, the power received by CYGNSS has a deterministic outcome such that at a fixed level of surface roughness, that is at a fixed surface wind speed u_{10} assuming ocean observations, the power mapped into each delay-Doppler pixel will also have an outcome that does not change regardless of number of 'realizations' (in the absence of any form of random noise). An 'ideal' noise free DDM for a given wind speed can then be generated purely from the mean $\bar{\mu}_x$ vector whose elements are given by (2.13) at a given delay and Doppler. Examples of this are shown in Figure 3.10.

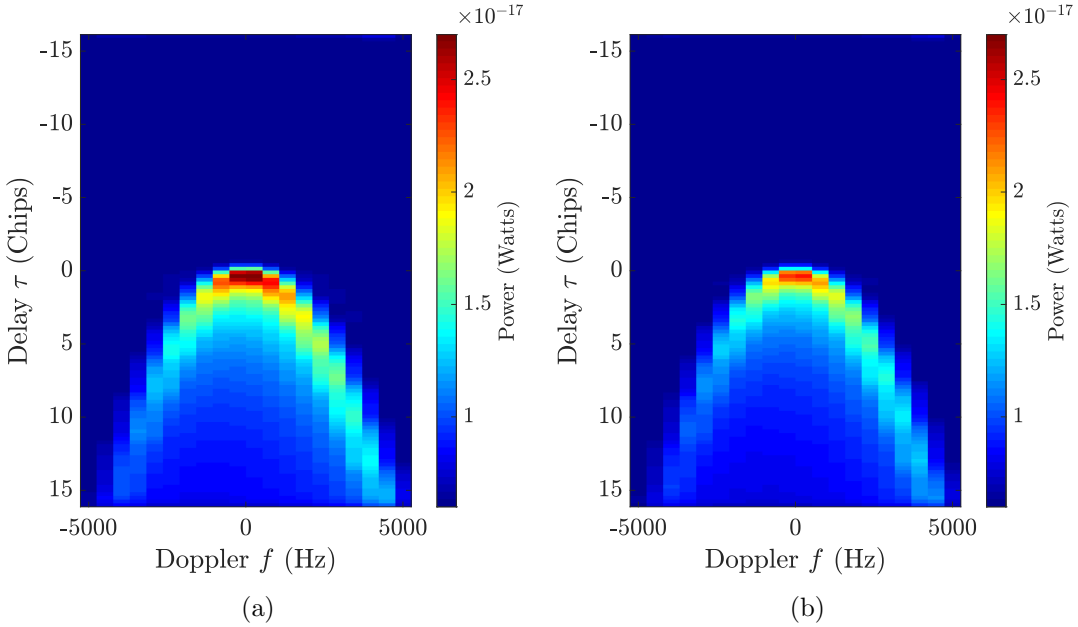


Figure 3.10: (a) Noise free DDM at $u_{10} = 6$ m/s (b) Noise free DDM at $u_{10} = 8$ m/s

Notice how the mean power for a given pixel, decreases as the wind speed increases. This is to be expected due to the increase in surface roughness.

Practically however, thermal noise will exist and is a random process in nature. Its mean value is its power P_N which is given by (3.11).

$$P_N = kTBF \tag{3.11}$$

where k is the Boltzmann constant (1.38×10^{-23} J/K), T is the temperature, B is the noise bandwidth and F is the receiver noise figure. The mean vector can now be obtained for random variables $x_1...x_{N_\tau \times N_f}$ and is the power at a given pixel $P(\tau, f; u_{10})$ plus thermal noise P_N and is denoted $\mu(\tau, f; u_{10})$.

To determine the covariance matrix, consider the simple example of covariance between two random variables $\Sigma(X, Y)$. There are a variety of ways through which this can be expressed but for the purposes of this analysis the form given by (3.12) is adopted.

$$\Sigma(X, Y) = \rho(XY) \times \sqrt{\Sigma(X, X) \times \Sigma(Y, Y)} \tag{3.12}$$

where $\rho(XY)$ is the correlation between the two random variables, $\Sigma(X, X)$ and $\Sigma(Y, Y)$ can be regarded as auto-covariance the square root of which reduces to the standard deviations σ_X and σ_Y. From this, the covariance matrix $[\Sigma(X, Y)]$ for the two random variables can be constructed as:

$$[\Sigma(X, Y)] = \begin{bmatrix} \sigma_X^2 & \rho(XY) \cdot \sigma_X \cdot \sigma_Y \\ \rho(YX) \cdot \sigma_Y \cdot \sigma_X & \sigma_Y^2 \end{bmatrix} \tag{3.13}$$

A similar principle can be extended to the case of the CYGNSS DDM where power at each pixel at delay τ and f is defined by the Gaussian distribution given by $N(\mu(\tau, f; u_{10}), \sigma^2(\tau, f; u_{10}))$, where $\sigma^2(\tau, f; u_{10})$ is the variance. Since the power follows an exponential distribution a pixel's standard deviation $\sigma(\tau, f; u_{10}) = \mu(\tau, f; u_{10})$.

The one second DDM measurement is a multilooked product with a number of looks $N_{looks} = 1000$ which in turn reduces standard deviation by a factor of $\sqrt{N_{looks}}$. The variance is then given by (3.14).

$$\sigma^2(\tau, f; u_{10}) = \mu^2(\tau, f; u_{10})/N_{looks} \tag{3.14}$$

Using the approximate model for describing inter-bin correlation (3.7), quantities required to define every element in the covariance matrix are now known. A matrix representation of the covariance matrix Σ is illustrated in (3.15).

$$\begin{bmatrix} \sigma^2_{\tau_1,f_1} & \sigma_{\tau_1,f_1}\sigma_{\tau_1,f_2}\rho(\tau_1,f_1;\tau_1,f_2) & \cdots & \sigma_{\tau_1,f_1}\sigma_{\tau_{N_\tau},f_{N_f}}\rho(\tau_1,f_1;\tau_{N_\tau},f_{N_f}) \\ \sigma_{\tau_1,f_2}\sigma_{\tau_1,f_1}\rho(\tau_1,f_2;\tau_1,f_1) & \sigma^2_{\tau_1,f_2} & \cdots & \sigma_{\tau_1,f_2}\sigma_{\tau_{N_\tau},f_{N_f}}\rho(\tau_1,f_2;\tau_{N_\tau},f_{N_f}) \\ \vdots & \ddots & \ddots & \vdots \\ \sigma_{\tau_{N_\tau},f_{N_f}}\sigma_{\tau_1,f_1}\rho(...) & \cdots & \cdots & \sigma^2_{\tau_{N_\tau},f_{N_f}} \end{bmatrix} \tag{3.15}$$

As such, noisy CYGNSS DDM measurements can be generated using the mean vector and covariance matrix developed using the preceding formulation with simple random Gaussian random number generation. To illustrate how this can be done, let's consider the simpler case of a set of Gaussian vectors $\bar{X}$. We would like to use this random number generation X to obtain a random vector Y such that:

$$Y = A\bar{X} + b \tag{3.16}$$

where

$$Y \sim N(\bar{\mu}_Y, [\Sigma_Y]) \tag{3.17}$$

The objective now is to solve for A and b that will achieve this. The first step is to find vector b. To do this consider the following:

$$\mathbf{E}[\bar{Y}] = \mathbf{E}[A\bar{X} + b] \tag{3.18}$$

$$= A\cancel{\mathbf{E}[\bar{X}]} + b \tag{3.19}$$

such that b = $\bar{\mu}_Y$. Next we note that the covariance matrix $[\Sigma_Y]$ can be written in terms of $[\Sigma_X]$.

$$[\Sigma_Y] = A \cdot [\Sigma_X] \cdot A^T \tag{3.20}$$

$$= AA^T \tag{3.21}$$

Next we note that any square matrix can be decomposed, through eigen-decomposition, into the product of eigen vectors V and eigen values Λ as shown in (3.22).

$$[\Sigma_Y] \cdot V = \Lambda \cdot V \tag{3.22}$$

where V is an $N_T \times N_T$ matrix and Λ is an $N_T \times N_T$ diagonal matrix. Now let $V = [\bar{v}_1 \bar{v}_2 ... \bar{v}_n]$. Matrix V is orthonormal in that it satisfies the condition described in (3.23):

$$\begin{aligned} V.V^T &= [\bar{v}_1 ... \bar{v}_n] \cdot \begin{bmatrix} \bar{v}_1 \\ \vdots \\ \bar{v_n} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \\ &= I_{N_T \times N_T} \end{aligned} \tag{3.23}$$

such that $V^T = V^{-1}$. From this it follows that expression (3.22) can be written in terms of $[\Sigma_Y]$ as described in (3.24).

$$\begin{aligned} [\Sigma_Y] &= V \Lambda V^T \\ &= [V \Lambda^{1/2}][\Lambda^{1/2} V^T] \end{aligned} \tag{3.24}$$

together with (3.21), A is described by (3.25).

$$A = V \Lambda^{1/2} \tag{3.25}$$

Applying the above formulation to CYGNSS' one second DDM measurements, N_T becomes the number of pixels within the DDM $N_T = N_\tau \times N_f$ and X is a set of multivariate normally distributed random numbers such that each row of X represents a DDM pixel and each column is a realization/observation. On the other hand, the vector $\bar{\mu}_Y$ comprises the means $\mu(\tau, f; u_{10})$ for each DDM pixel and is of dimension $(N_\tau N_f \times 1)$. The covariance matrix Σ_Y is the covariance matrix $[\Sigma_{\tau,f}]$ for the DDM pixels. Finally Y now represents a 'noisy' simulated DDM measurement which is here denoted by $M(\tau, f; u_{10})$. The noise free DDM is based on the mean vector $\bar{\mu}(\tau, f; u_{10})$ and can be regarded as an ideal 'reference' template. It is denoted by $S(\tau, f; u_{10})$. An example of how this formulation was used to generate a speckled DDM is shown in Figure 3.11.

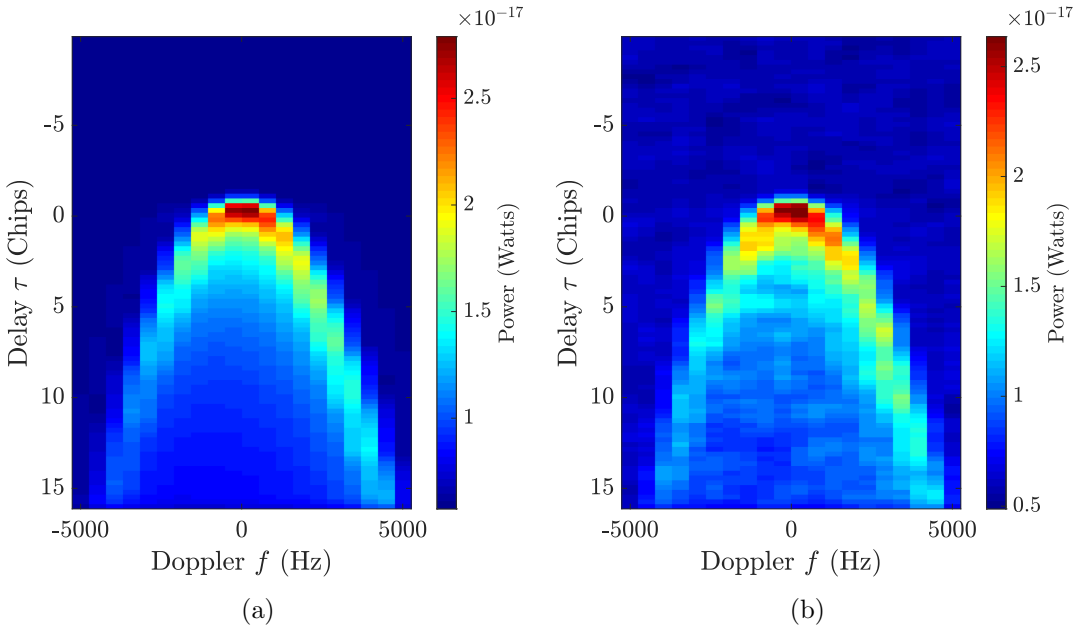


Figure 3.11: (a) Noise free DDM at $u_{10} = 6$ m/s (b) Noisy DDM at $u_{10} = 6$ m/s

3.5 Special Test Datasets

In support of studies exploring the utility of the CYGNSS constellation for land remote sensing, future releases of its data will include several improvements. In addition to improvements in CYGNSS antenna pattern estimates, GPS Effective Isotropic Radiated Power (EIRP) affecting all specular points (land and ocean alike) an improved land specular point location solver will be incorporated. This is of particular importance as the existing algorithm does not account for the significant local elevation variations that may occur over land using only the DTU10 digital elevation model currently being employed in the current v2.1 version of the L1 data. For specular points at high elevation this may result in appreciable errors in the geolocation process. To examine these improvements a special test set, called a 'sandbox' run, was created for the 12 day period from DOY-Year 240-2019 to 251-2019. The test run also provides added ancillary information not typically included in L1 data release, including estimates of static water body fraction within the measurement's footprint, Shuttle Radar Topography Mission (SRTM) digital elevation and slopes. This also includes local maps of interest, an example of some of which is shown in Figure 3.12.

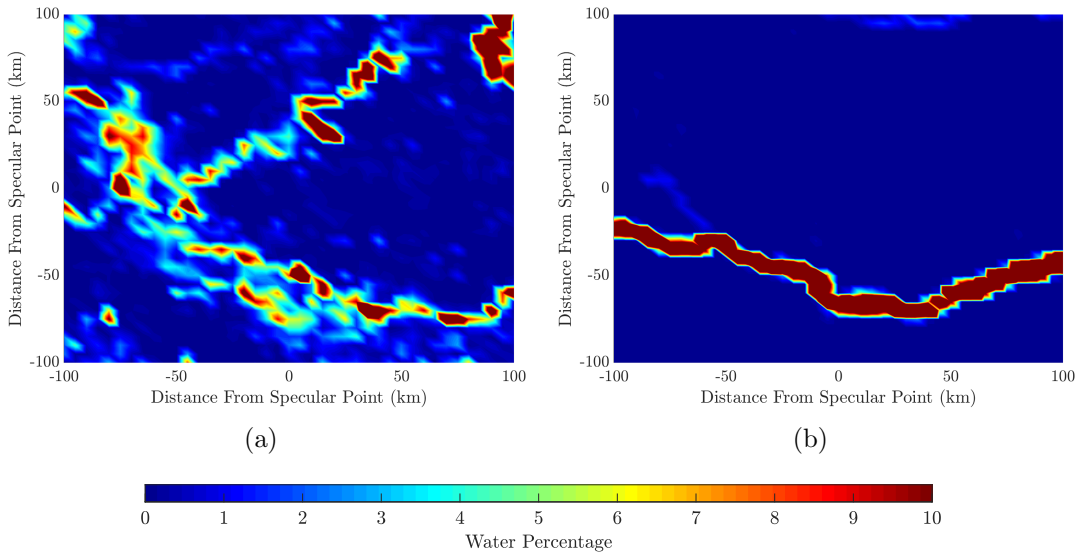


Figure 3.12: Added ancillary data for special land test dataset, sandbox019. Local area map of Pekel percentage of surface water. Each cell of the local area map is 5km × 5km (a) Map for a specular point at Latitude = 33.21° and Longitude = -87.86° observed by CYG01 on DOY, Year 240, 2019 (b) Map for a specular point at Latitude = -5.35° and Longitude = -53.3° observed by CYG01 on DOY, Year 240, 2019

The sandbox run includes DDMs for 16.5 million specular points over land, and 45 million specular points over the ocean across all eight CYGNSS receivers thereby providing a large dataset that will be of particular interest for the testing of the proposed coherence detection algorithm in Chapter 5.

3.6 Conclusions

This chapter provided a basic overview of the CYGNSS mission and several of its most pertinent operational aspects. An overview of the products it provides under its standard operating mode, Level 1-3, and those provided infrequently under its special modes of operation, Full DDM and raw I/F, was presented. Several issues of practical relevance were also highlighted including the correspondence of CYGNSS

observables to spatial footprint, the need to account for inter-bin correlation and how this ultimately coarsens the finest spatial features CYGNSS is capable of resolving.

The need to be able to predict GNSS-R returns in general, and more specifically CYGNSS returns, was highlighted and an overview of the two primary means through which this can be done was presented. The first relates to the use of simulation software capable of predicting CYGNSS returns through attempting to compute the bistatic radar equation at a given delay and Doppler whilst effectively accounting for the dependencies this has on observation geometry and surface properties. Alternatively, probabilistic descriptors of the DDM can be developed where it was shown that a Gaussian random vector model may be used to model DDM pixels where the procedure for generating noisy CYGNSS-like DDMs was introduced. It was shown, that knowledge of each of the DDMs' pixels mean and covariance were sufficient for this generation.

Chapter 4: Time Series Retrieval Algorithm for Global Near Surface Volumetric Soil Moisture

4.1 Motivation

Studies analyzing the land returns of the recently launched Cyclone Global Navigation Satellite System (CYGNSS) constellation clearly demonstrate sensitivity to numerous land surface properties including soil moisture. This has motivated further investigation into the means through which CYGNSS and future GNSS-R missions can be used for the purpose of soil moisture retrievals. Initial assessments highlight CYGNSS's SNR sensitivity to changes in soil moisture levels with depictions of this analysis summarized in Figure 4.1 in which changes in month-to-month average SNR values are compared to the corresponding month-to-month changes in soil moisture. A similar sensitivity is evidenced by a visual examination of Figure 4.2 in which the response of surface radar cross section (RCS) derived from measured CYGNSS DDMs, for a set of specular measurements over the Indian Peninsula, to varying levels of soil moisture derived from SMAP is compared. For this particular case study, the correlation between the two was found to be in excess of 80% lending further support to the expectation that the measurements of spaceborne GNSS-R systems may prove instrumental in providing supplementary provisions, to the existing SMAP, SMOS,

Sentinel and a number of other platforms for the monitoring global near-surface soil moisture.

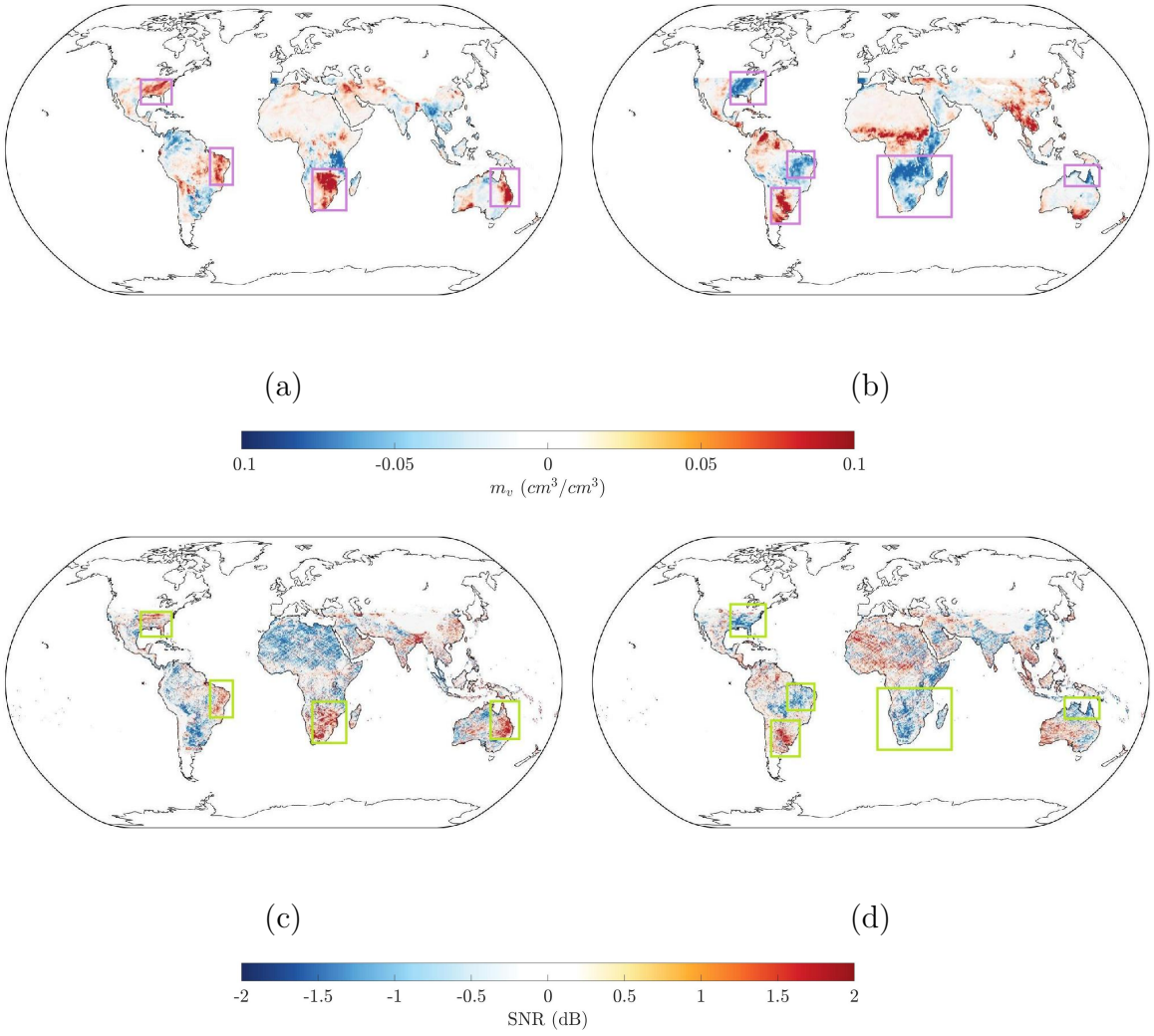
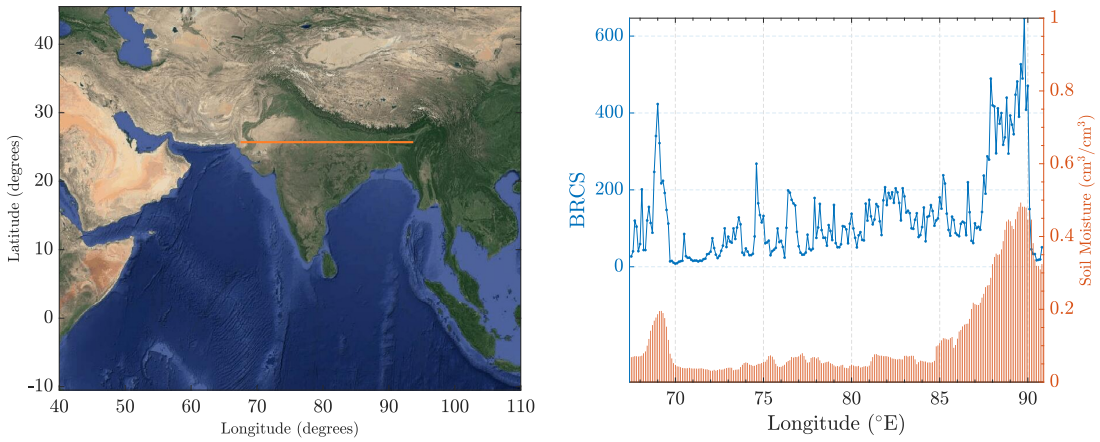


Figure 4.1: Comparisons of average month-to-month variations of radiometer derived soil moisture m_v ($\mathrm{cm}^3/\mathrm{cm}^3$) computed as ${m_v}^{j+1} - {m_v}^j$ and average month-to-month variations of CYGNSS Signal-to-Noise ratio (dB) over land computed as $\mathrm{SNR}^{j+1} - \mathrm{SNR}^j$ at $0.2^o \times 0.2^o$ resolution (a) Average change in soil moisture from January to February 2018 (b) Average change in soil moisture from April to May 2018 (c) Average change in SNR from January to February 2018 (d) Average change in SNR April to May 2018



(a) (b)

Figure 4.2: Sensitivity of measured CYGNSS observables to varying levels of near surface soil moisture (a) Illustration of horizontal transect over the Indian subcontinent for the purpose of CYGNSS BRCS and SMAP soil moisture comparison (b) Comparison between average CYGNSS BRCS and SMAP soil moisture for horizontal transect spanning (25.7°N and 67.5 - 93.7°E)

The presence of correlation between the CYGNSS measured SNR or RCS and soil moisture or patterns of complementary change thereof aids in establishing the presence of a soil moisture 'signal' embedded within CYGNSS's land measurements. While such results are indicative of general correlations between changes in CYGNSS signal-to-noise ratios and surface soil moisture, retrievals on a global scale require the separation of the dependence of the returns on surface roughness, vegetation, and soil moisture for a given scene and the reporting of soil moistures on time scales finer than one month. It is further important in retrieval attempts to compensate instrument related factors such as the incidence angle (which varies for each CYGNSS observation) and other parameters.

Due to the importance of global cost-effective remote sensing and monitoring of soil moisture relating to its playing a key role in global climate modulation, in land-atmosphere interactions, improved drought and flood prediction as well as other roles it plays in vegetation and agricultural applications and due to the need for retrieval approaches developed to address the confounding influences of surface roughness, land cover, and incidence angle variations an alternative time series soil moisture retrieval algorithm based on change detection is developed in this chapter. The method uses a time series of CYGNSS measurements under the assumption that vegetation and surface roughness effects are canceled when a ratio of two successive CYGNSS measurements is computed. The use of ratios makes the incorporation of ancillary information pertaining to vegetation or surface roughness unnecessary in the retrieval process.

This chapter aims to establish the formulation of the time series retrieval approach through a five step process which includes elimination of unwanted returns from daily CYGNSS measurements, computation of the Normalized Radar Cross Section (NRCS) for the remaining returns, the construction of a system of $(N-1) \times N$ equations involving ratios of successive CYGNSS Level-1 observables, solving this system for the land surface reflectivity, and the inversion of the reflectivity time series into soil moisture m_v through the use of a dielectric mixing model. The high correlations and low error levels associated with the soil moisture retrievals, relative to reference datasets, using the specular land measurements and the time series retrieval approach highlight the potential of using spaceborne bistatic GNSS-R for soil moisture estimation.

4.2 Normalized Radar Cross Section

The basic principles underlying the operation of spaceborne GNSS-R systems, outlined in Chapter 2, gives rise to a mapping of received signal power from the spatial domain to the fundamental GNSS-R measurement, the delay-Doppler map (DDM), as depicted in Figure 4.4(a) and conforms to the bistatic radar equation in its integral form given by (2.13). For measurements spanning large delay extents, and as a consequence are with a large spatial footprint, appreciable variations in the projection of the receiver(s)/transmitter(s) antenna patterns will be observed. The same is noted when analyzing variations in total propagation paths at surface points, increasingly distant from the specular point. However, due the limited delay extent of L1 DDMs in general and more specifically the spatial extent of the individual 0.25 chip bins it comprises, and as a consequence limited spatial extent, both the projection of antenna gains and ranges to transmitter(s) and receiver(s) may be assumed constant as shown in Figure 4.3. This then reduces (2.13) to (4.1) which enables a more straight forward inversion of the bistatic radar equation to obtain an estimate of the observed surface's normalized radar cross section (NRCS) through CYGNSS's calibration procedure [77–79].

$$P_R^{inc}(\tau, f) = \frac{P_T \lambda^2 G_T G_R \sigma_0 A_{eff}}{(4\pi)^3 r_T r_R} \tag{4.1}$$

The procedure results in an estimate of the power received at each "pixel" within the DDM, which can then be converted into an estimate of the bistatic radar cross

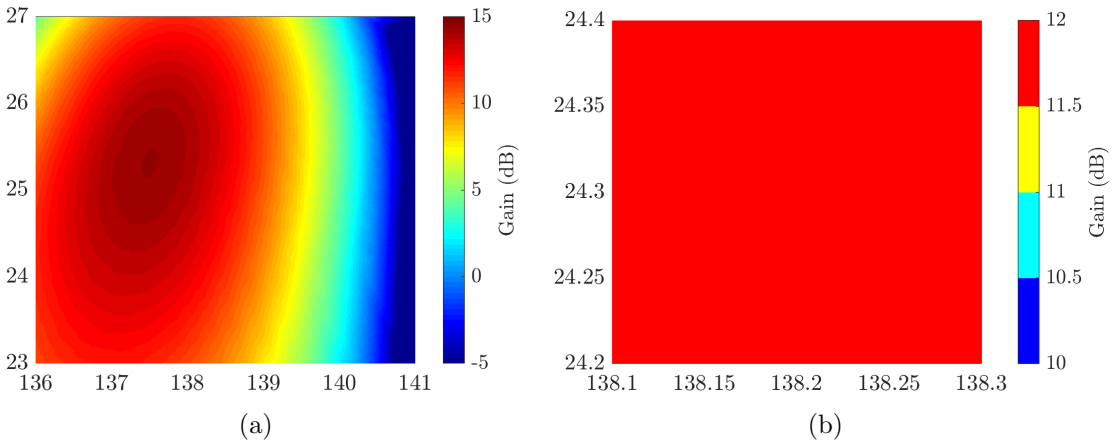
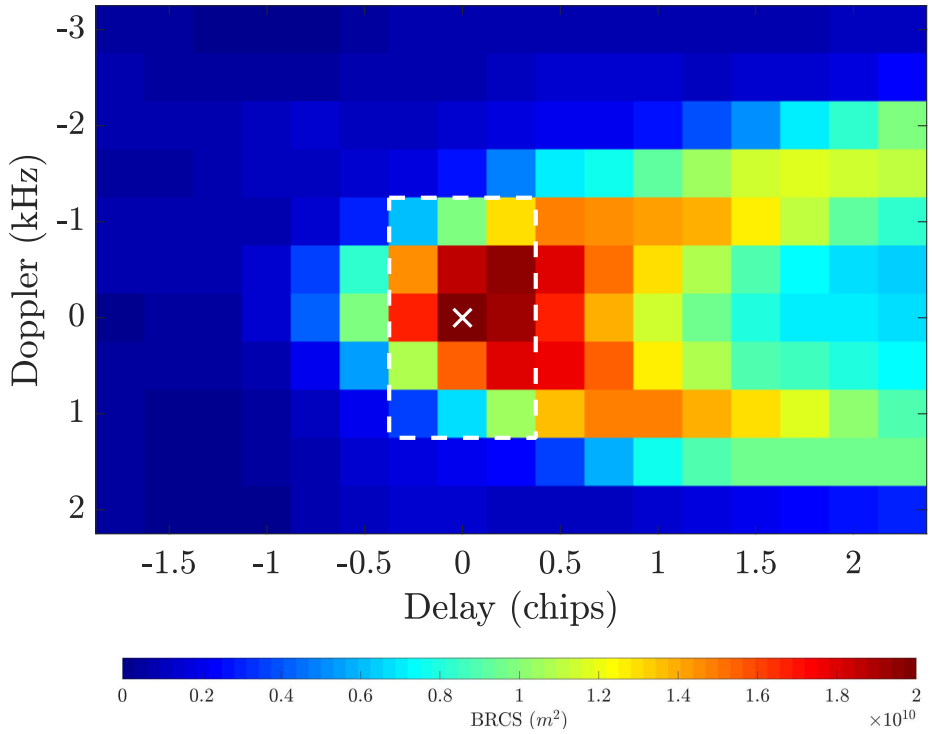


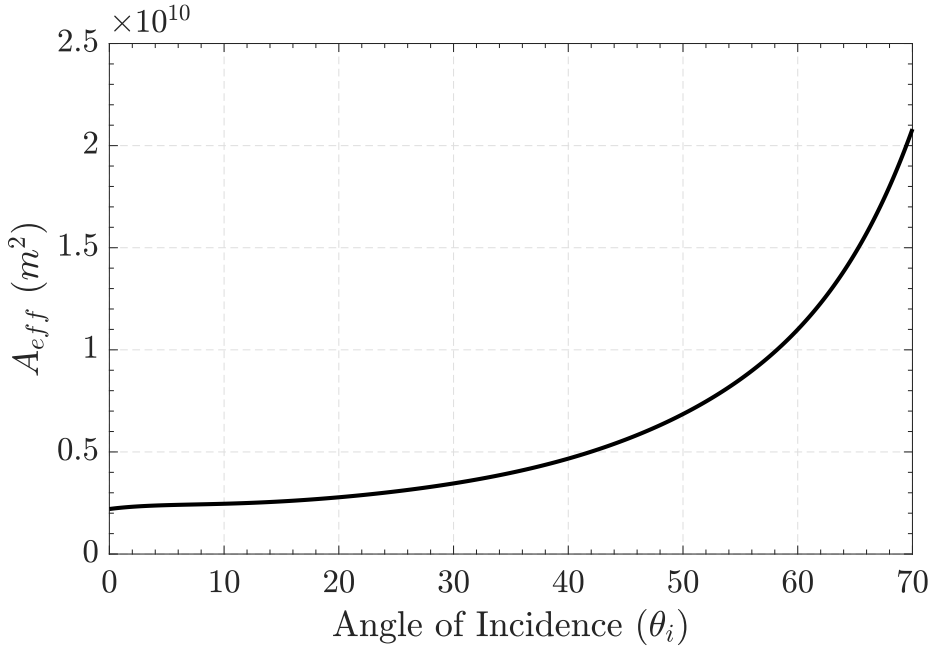
Figure 4.3: Explanation of rationale behind simplification of bistatic radar equation (a) Large scale projection of simulated CYGNSS antenna pattern centered about L1 specular point (b) Antenna pattern variability within spatial cell corresponding to 0.25 chip bin in Level-1 DDM

section (BRCS) by multiplying by known factors involving the CYGNSS receive antenna pattern, GPS transmit power and antenna pattern, as well as the ranges to the transmitter and receiver.

The importance of target (B)RCS derives from the fact that, after correcting for instrument (gains) and geometry (ranges) effects, it is a description of the fundamental scattering properties of a target which in turn are telling of a wide range of target properties including physical size, orientation and constituents. The targets observed by spaceborne passive, GNSS-R, systems are nonetheless area extensive. That is, the physical extent of the region illuminated by the transmit and receive antenna patterns is large enough to occupy a significant portion of the antennas' gain footprint with a large number of non-colocated points contributing reflections to total received surface scatter. Depending on incidence angles and observation geometry, the size and orientation of this footprint will undergo considerable variability. A need therefore



(a)



(b)

Figure 4.4: (a) Bistatic Radar Cross Section (BRCS) delay-Doppler Map (DDM) for diffuse/incoherent scattering (b) Illustration of effective scattering area (A_{eff}) best fit line as a function of CYGNSS incidence angle

arises to 'correct' (B)RCS assigned to observed targets, to account for the size of the area giving rise to surface scatter such that the measured quantity remains, solely, a function of surface properties (as opposed to it being 'biased' by different size footprints). The process of correcting BRCS to account for area effects gives rise to the Normalized Radar Cross Section (NRCS).

In order to estimate the surfaces' bistatic NRCS, a total BRCS representing the specular region must be divided by the effective scattering area on the surface. First the specular point (identified by a maximum) within the BRCS DDM (σ_{τ_i,f_j}) is tracked, and since interest is in the relatively small area surrounding the specular point (σ_{DDMA}), the sum over a 5 Doppler (approximately 2 kHz at the CYGNSS 500 Hz Doppler sampling rate) $\times$ 3 delay (approximately 0.75μsec at the CYGNSS 0.25μsec sampling rate) region is taken. The effective scattering area (A_{eff}) is obtained by a curve fit of CYGNSS reported A_{eff} values as a function of the incidence angle using a dataset of one million CYGNSS observations, with the resulting curve fit depicted in Figure 4.4(b). The CYGNSS NRCS (σ_0) is then given by (4.2).

$$\sigma_0 = \frac{\sigma_{DDMA}}{A_{eff}} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\sigma_{\tau_i,f_j}}{A_{eff}} \tag{4.2}$$

4.3 Physical Dependence of GNSS-R DDMs on Land Properties

Incoherent returns for land surfaces should exhibit dependence on the composite contributions of soil moisture, vegetation attenuation, and surface roughness. A simple model [99, 164] for the incoherent specular NRCS of land surfaces is:

$$\sigma_0 = \frac{\Gamma_{LR}(\epsilon_s, \theta_i)}{MSS}\ e^{-\tau_v sec\theta_i} \tag{4.3}$$

where σ_0 represents the specular NRCS observed by a GNSS-R measurement at incidence angle θ_i. Here $\Gamma_{LR}(\epsilon_s, \theta_i)$ is the surface's Fresnel reflectivity in polarization LR (right hand circular incidence and left hand circular scattering for CYGNSS)

$$\Gamma_{LR}(\epsilon_s, \theta_i) = \left| \frac{1}{2}(\Re_{VV} - \Re_{HH}) \right|^2 \tag{4.4}$$

with

$$\Re_{VV} = \frac{\epsilon_s \cos\theta_i - \sqrt{\epsilon_s - \sin^2\theta_i}}{\epsilon_s \cos\theta_i + \sqrt{\epsilon_s - \sin^2\theta_i}} \tag{4.5}$$

$$\Re_{HH} = \frac{\cos\theta_i - \sqrt{\epsilon_s - \sin^2\theta_i}}{\cos\theta_i + \sqrt{\epsilon_s - \sin^2\theta_i}} \tag{4.6}$$

and is a function of the surface relative complex permittivity ϵ_s.

It is well established that soil volumetric moisture content (m_v) impacts ϵ_s, with the real part of ϵ_s varying between that of dry sand ≈ 3 and water ≈ 80 depending on the volumetric water content and texture (soil bulk density, clay fraction, sand fraction) of the soil [100]. Increasing moisture content results in an increase in soil permittivity and therefore reflectivity, producing the dependence of GNSS land returns on soil moisture. Microwave soil moisture retrieval will therefore fundamentally retrieve ϵ_s which must then be inverted to soil moisture through the use of a permittivity model [100–105] and ancillary texture information. The models used in this study are briefly explored in the next section.

In Equation (4.3), MSS represents the "mean square slope" of surface roughness, which under the Hagfors' model of land surface scattering [99] can be a function of the angle of incidence but here is approximated as independent of the observation angle. It is here assumed that no ancillary information on MSS is readily available.

Finally, the exponential term in (4.3) describes attenuation due to vegetation cover in terms of the incidence angle and the nadiral vegetation optical depth (VOD) τ_v. The VOD includes terrain-specific scattering and attenuation contributions, but is frequently represented in terms of the vegetation water content (VWC) W_c as

$$\tau_v = b \cdot W_c \tag{4.7}$$

The relationship between τ_v and W_c should vary in principle from one crop type to another depending on a number of vegetation-specific physical and dielectric properties [109].

4.4 Dielectric Descriptions of Soil

As a non-magnetic medium, it is generally the permittivity ϵ_s of soil that is of most interest when describing the dielectric properties of a given land surface. All models, to varying degrees of accuracy, will ultimately aim to capture the behaviour depicted in Figure 4.5 namely relating to the monotonic increase in soil's permittivity as a function of volumetric soil moisture approaching that of water. It follows that, drier sandy soils that tend to be more porous and have a lower permittivity on average. As a consequence, drier soils are less reflective. In contrast, soils with a higher clay content often constitute a higher volumetric moisture content and, when bare, exhibit significantly higher surface reflectivity that is fundamentally driven by its increased ϵ_s. Other complexities in describing the dielectric properties of soil arise however as a result of the variability of soil density, salinity, differing proportions of sand/clay fractions, temperature and frequency of impinging radiation. Different dielectric models attempt to account for this variability through semi-empirical descriptions of these coupled dependencies. The importance of choosing an accurate descriptor of soil

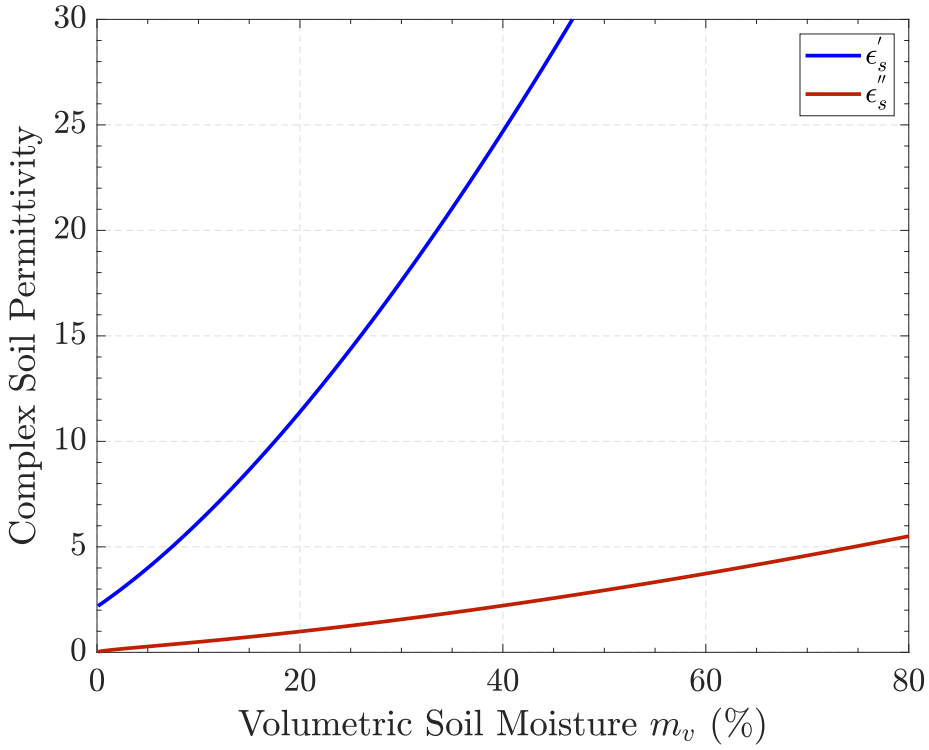


Figure 4.5: Dependence of soil complex permittivity on varying levels of soil moisture

permittivity derives from the fact that it facilitates location dependent translations, based on (4.3)-(4.4), between surface $\Gamma_{LR} - \epsilon_s$ which ultimately dictate m_v retrievals. That is, given an observed or retrieved Γ_{LR} and incidence angle, an estimate of a 'retrieved' ϵ_s is obtained and using (4.4)-(4.6) and through an iterative approach involving the set of dielectric models' function(s), soil moisture may be retrieved. In this work two models have been considered, the Peplinski/Dobson/Ulaby and the Mironov models.

4.4.1 Peplinski/Dobson/Ulaby Model

An impinging wave is perturbed by the constituents of soil which can be described as a mixture of water, soil and air pockets, with the resulting permittivity being representative of the compound properties of the three. The Peplinski/Dobson/Ulaby model is a 'mixing' model, in that it recognizes this inherent heterogeneity of soil composition and aims to describe the permittivity of soils through weighting the relevant contributions of the three main constituents of a soil mixture in a manner that is representative of their respective dielectric contributions. To do this, a comprehensive empirical study is conducted in [101,102] to determine the variability in the dielectric properties of soil as a function of wave frequency, temperature, volumetric water/air contents of the soil as well as soil density. This is accomplished by studying the perturbations introduced by soil samples with different properties to guided waves and inference of soil properties is accomplished by comparing those to anticipated properties of modes (the design of the waveguide is typically such that the principle TE_{10} mode is the only one capable of propagation) travelling in an air-filled wave guide. This practice is known as waveguide transmission technique. Clear patterns emerge as a function of the aforementioned parameters and the result is the semi-empirical Peplinski/Dobson/Ulaby model given by (4.8)-(4.10) dielectric mixing model.

$$\epsilon_s = \epsilon_s^{'} - j\epsilon_s^{''} \tag{4.8}$$

where

$$\epsilon_s^{'} = \left[1 - \frac{\rho_b}{\rho_s}(\epsilon_m^{\alpha}) + m_v^{\beta^{'}}\epsilon_{fw}^{'\alpha} - m_v\right]^{1/\alpha} \tag{4.9}$$

$$\epsilon_s^{''} = \left[m_v^{\beta^{''}} e_{fw}^{''\alpha} \right]^{1/\alpha} \tag{4.10}$$

where ρ_b is the bulk density of soil, ρ_s is the specific density of soil particles, $\alpha = 0.65$ is an empirically determined constant, $\beta^{'}$ and $\beta^{''}$ also empirically determined to weight (4.9)-(4.10) in a manner that accounts for different soil types by factoring in the soil's sand S and clay C fractions given by (4.11)-(4.12).

$$\beta^{'} = 1.2748 - 0.519S - 0.152C \tag{4.11}$$

$$\beta^{''} = 1.33797 - 0.603S - 0.166C \tag{4.12}$$

ϵ_m is another empirically derived relationship, given by (4.13) and is aimed at accounting for the effects of varying levels of soil specific density, and as a result the specific gravity of soil (the ratio of the mass of equal volumes of soil to distilled water).

$$\epsilon_m = (1.01 + 0.44\rho_s)^2 - 0.062 \tag{4.13}$$

The dielectric response of water to varying temperatures, frequencies and salinity levels is captured by $\epsilon_{fw}^{'}$ and $\epsilon_{fw}^{''}$ given by (4.14)-(4.15).

$$\epsilon_{fw}^{'} = \epsilon_{w\infty} + \frac{\epsilon_{w0} - \epsilon_{w\infty}}{1 + (2\pi f \tau_w)^2} \tag{4.14}$$

$$\epsilon_{fw}^{''} = \frac{2\pi f \tau + w(\epsilon_{w0} - \epsilon_{w\infty})}{1 + (2\pi f \tau_w)^2} + \frac{\sigma_{eff}}{2\pi \epsilon_0 f} \frac{\rho_s - \rho_b}{\rho_s m_v} \tag{4.15}$$

where f is the frequency of operation which for the L1 course-acquisition codes relied on for the the CYGNSS mission is 1.575 GHz, ϵ_{w0} is the low frequency limit of water and $\epsilon_{w\infty}$ is water's permittivity as the frequency of operation approaches

infinity, τ_w is water's relaxation time and σ_{eff} is an empirically derived estimate of water's effective conductivity as it is suspended by soils with differing textural properties and is given by (4.16).

$$\sigma_{eff} = -1.645 + 1.939\rho_b - 2.25622S + 1.594C \tag{4.16}$$

The model is semi-empirical in that its description of soil permittivity is based on a series of empirically derived expressions, weighted by parameters that have been designed to fit the general experimental tendencies exhibited by different soil samples whilst coupling them with more theoretical models, in this case to capture the dielectric properties of water within the soil mixture, like the Debye dispersion relations (4.14)-(4.15). It also follows that through this semi-decoupled integration of the water's permittivity, it is indirectly implied that soil water's complex dielectric constant is independent of soil mineralogy and type, accounting for this variability only partially through the weighting of the water's effective conductivity using (4.16). As a result, the use of this model requires ancillary location-specific inputs that relate to soil salinity, temperature, bulk density as well as clay and sand fractions. This together, with expressions (4.8)-(4.16) completely define the permittivity of soil under all conditions that are likely to be of interest. Due to the relative simplicity of this model, robustness across a wide range of soils and frequencies of operation, as well as the availability of the bulk of the limited amount of ancillary information its implementation requires, it is this model that is used in all the results that follow.

4.4.2 Mironov Model

The second, and equally well established, model considered in this work is the Mironov model [103–105]. In contrast to the Peplinski/Dobson/Ulaby semiemperical

dielectric mixing model, that developed by Mironov et al. is a generalized refractive mixing model in which the determination of the soil's permittivity is facilitated through a set of empirically derived coefficients based on spectroscopic analysis of a soil sample's refractive index and normalized attenuation coefficient which is used to derive a second-order polynomial weighted by soil temperature, clay fraction and volumetric moisture content using (4.17)-(4.19) to describe the soil's complex permittivity.

$$\begin{aligned} \epsilon_s &= \epsilon_s^{'} - j\epsilon_s^{''} \\ &= (n_s^2 - \kappa_s^2) - j2n_s\kappa_s \end{aligned} \tag{4.17}$$

with

$$n_s = \begin{Bmatrix} n_d + (n_b - 1)m_v, & m_v \leq m_v^t \\ n_d + (n_b - 1)m_v^t + (n_u - 1)(m_v - m_v^t), & m_v > m_v^t \end{Bmatrix} \tag{4.18}$$

$$\kappa_s = \begin{Bmatrix} \kappa_d + \kappa_b m_v, & m_v \leq m_v^t \\ \kappa_d + \kappa_b m_v^t + \kappa_u(m_v - m_v^t), & m_v > m_v^t \end{Bmatrix} \tag{4.19}$$

where n and κ represent the samples' refractive indices and attenuation coefficients respectively, with the subscripts s, d, b, u indicating relation to moist soil, dry soil, bound soil water and free soil soil water respectively. The maximum fraction of bound water within a given sample is given by m_v^t. The Mironov model places particular emphasis on the need to account for proportions of water that are bound to soil molecules, the thin layer of water surrounding soil minerals, and free water relating to freely flowing water within the sample that has not formed any molecular or electrically attractive bonds with the soil. Since electrical permittivity is ultimately a measure of the opaqueness of a given medium with respect to the size of a wavelength,

the increased 'coating' of a medium's molecules with water described by an increase in bound water, can have the highest proportional impact on increases of its permittivity. The distinction is of particular significance when a higher degree of accuracy in the representation of soil permittivity is needed to the differing effects the two induce to impinging electrical radiation and as a consequence, the surface's permittivity. Making these distinctions however, requires an empirical determination of a given samples' energies of activation, entropies of activation, temperature coefficients of ion conductance and a number of other parameters for the range of soil types of interest. No theoretical or empirical relationship exists that enable extrapolating the results of a given spectroscopy study at one frequency to a range of other frequencies such that its use requires the resource intensive undertaking of a near complete spectroscopic characterization [106] of soils at the operation frequency of interest. Within the L-band however, the coefficients necessary have been previously determined and are given by (4.20)-(4.26), having been fitted to a range of soil temperatures (T) up to 40°C and clay fractions up to 70%.

$$m_v^t = 0.0286 + 0.00307C \tag{4.20}$$

$$n_d = 1.634 - 0.00539C + 2.75 \times 10^{-5} C^2 \tag{4.21}$$

$$\begin{aligned} n_b &= (8.86 + 0.00321T) + (-0.0644 + 7.96 \times 10^{-4} T)C \\ &\quad + (2.97 \times 10^{-4} - 9.6 \times 10^{-6} T)C^2 \end{aligned} \tag{4.22}$$

$$\begin{aligned} n_u &= (10.3 - 0.0173T) + (6.5 \times 10^{-4} + 8.82 \times 10^{-5} T)C \\ &\quad + (-6.34 \times 10^{-6} - 6.32 \times 10^{-7} T)C^2 \end{aligned} \tag{4.23}$$

$$\kappa_d = 0.0395 - 4.038 \times 10^{-4} C \tag{4.24}$$

$$\begin{aligned} \kappa_b &= (0.738 - 0.00903T + 8.57 \times 10^{-5} T^2) + (-0.00215 + 1.47 \times 10^{-4} T)C \\ &\quad + (7.36 \times 10^{-5} - 1.03 \times 10^{-6} T + 1.05 \times 10^{-8} T^2) C^2 \end{aligned} \tag{4.25}$$

$$\begin{aligned} \kappa_u &= (0.7 - 0.017T + 1.78 \times 10^{-4} T^2) + (0.0161 + 7.25 \times 10^{-4} T)C \\ &\quad + (-1.46 \times 10^{-4} - 6.03 \times 10^{-6} T - 7.87 \times 10^{-9} T^2) C^2 \end{aligned} \tag{4.26}$$

In what follows, retrievals on the global scale were attempted using both the Mironov and Peplinski/Dobson/Ulaby models. In spite of the ability of the Mironov model to reduce the number of ancillary parameters required to facilitate a soil moisture - soil permittivity - surface reflectivity interchange, by rendering information relating to soil bulk density and sand fraction as unnecessary, the relative complexity of its formulation, the inflexibility of its usage across other frequencies that may be of interest together with the fact that the bulk of the simulation studies, retrievals conducted and comparisons of results which incorporated the two models against reference data did not suggest any particular advantages of using one model over the other and did not on average exceed 5% difference in retrievals, an observation which is consistent with the studies of [100] all motivated the use of the Peplinski/Dobson/Ulaby semi-empirical dielectric mixing model in CYGNSS retrievals.

4.5 Compensation for Incidence Angle Dependence

Previous studies investigating soil moisture retrievals have often dealt with systems making observations at one or a relatively small set of incidence angles. An example of this is the SMAP mission where observations are always made with a fixed $\theta_i = 40°$.

GNSS-R observations in contrast are made at a wide range of incidence angles. A robust soil moisture retrieval method must therefore address the dependence of the NRCS on incidence angle. An initial approach for doing so is presented in this section.

To provide context for CYGNSS, Figure 4.6 is a histogram of the angles of incidence at which the CYGNSS constellation made observations over a nine month period from April 2017 to December 2017. From this it can be observed that the angles at which observations are made vary significantly with an standard deviation of $\cong 16.70°$. For CYGNSS, a degradation in data quality occurs at larger incidence angles [48], and CYGNSS measurements obtained beyond $\theta_i = 60°$ are commonly disregarded for poor quality.

The dependence of the Fresnel reflectively $\Gamma_{LR}(\epsilon_s, \theta_i)$ on m_v and θ_i is depicted in Figures 4.7(a) and 4.7(b) for typical loam soil parameters.

Figure 4.7 clearly demonstrates the increase in reflectivity with soil moisture that is coupled with a dependence on θ_i. A simple normalization to the value at nadir as in Figure 4.7(b) shows that the curves become almost identical, they 'collapse'. This result motivates a remapping of the reflectivity at θ_i to its corresponding value at $\theta_i = 0$ by representing the reflectivity as a product of two functions, one that is entirely m_v dependent given by (4.4-4.6) at $\theta_i = 0$ and another that is exclusively angle dependent and given by a simple curve fit of the function $f(\theta_i)$:

$$\Gamma_{LR}(\epsilon_s, \theta_i) = \Gamma_{LR}(\epsilon_s) \cdot f(\theta_i) \tag{4.27}$$

Expression (4.27) provides an angle correction by which NRCS returns made at different θ_i can be used to create a corrected NRCS value σ_c referenced to nadir incidence

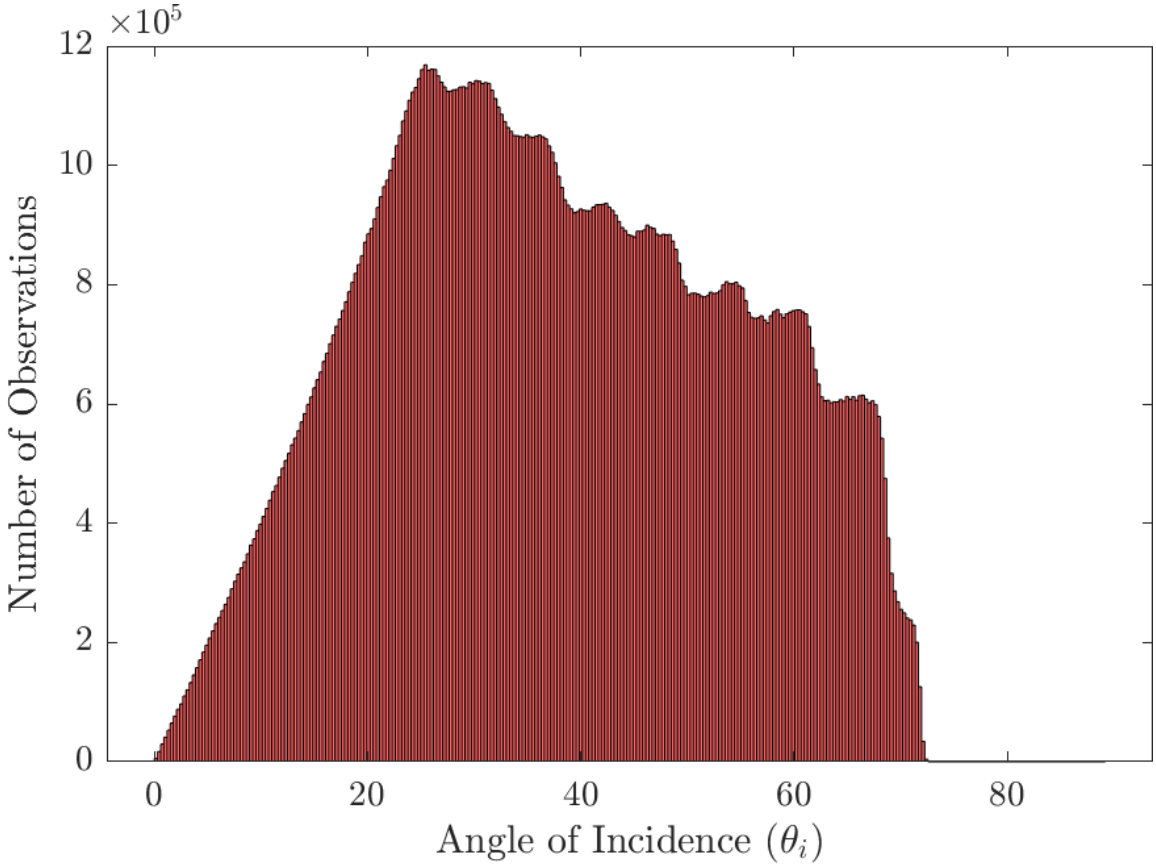


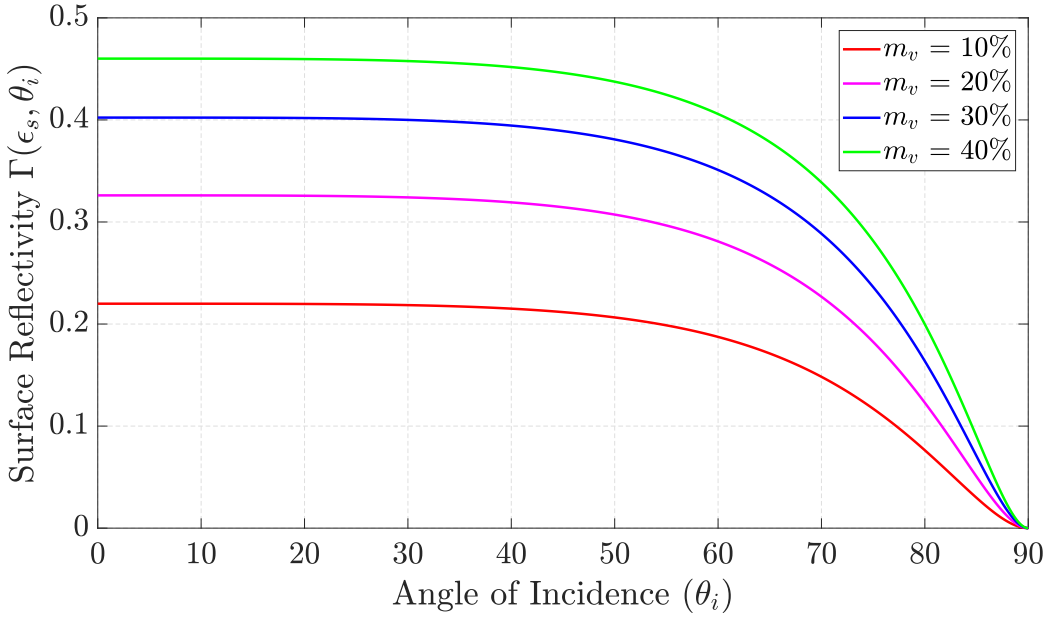
Figure 4.6: Histogram of the angle of incidence of CYGNSS land measurements over a nine month period from April 2017 to December 2017 for all longitudes and latitudes within $\pm 40^\circ$

for the purposes of direct comparison:

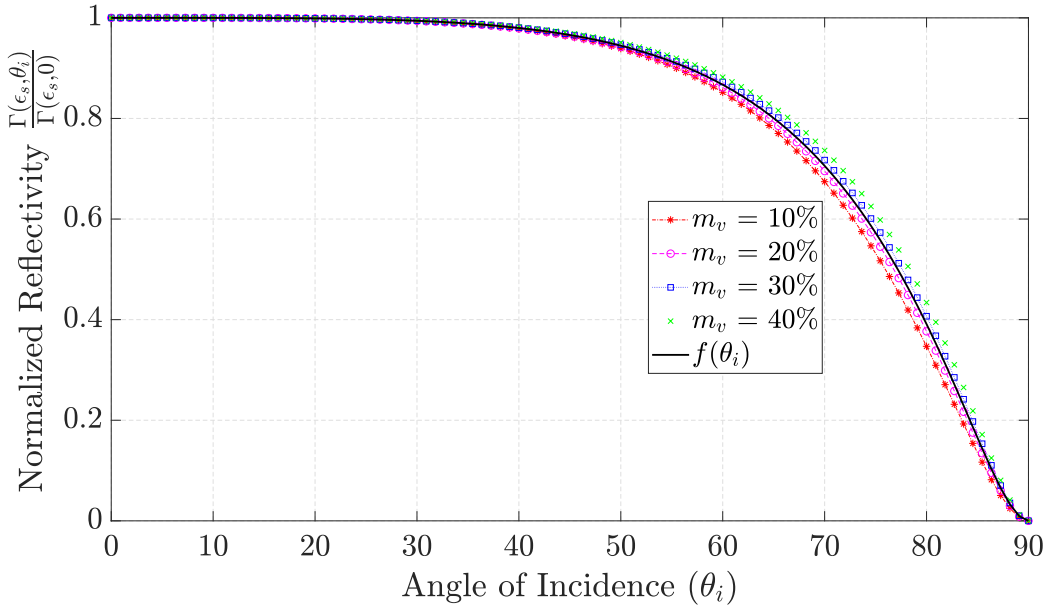
$$\sigma_c = \frac{\sigma_0}{f(\theta_i)} \tag{4.28}$$

This approach should largely mitigate the effects of incidence angle dependencies related to the Fresnel reflectivity.

Beyond the Fresnel reflectivity, Equation (4.3) also shows that the vegetation attenuation contributions can also vary with the incidence angle. Furthermore, ratios of consecutive measurements would not cancel vegetation attenuation due to its



(a)



(b)

Figure 4.7: Illustration of coupled θ_i-m_v dependence and angle correction (a) incidence angle and moisture dependence of reflectivity curves for typical loam soil (b) incidence angle and moisture dependence of reflectivity normalized by reflectivity at nadir incidence

dependence on incidence angle. While vegetation attenuation could potentially be compensated by incorporating ancillary information on the τ_v parameter, for simplicity the algorithm to be described will assume that vegetation attenuation remains constant on successive measurements, even those at differing incidence angles. It should be expected that this assumption will have the most significant impact for scenes having larger vegetation attenuation.

4.6 Algorithm Formulation

The model developed thus far describes the land surface incoherent specular NRCS as a product of terms accounting for the dependence on soil moisture, surface roughness, and vegetation. The retrieval method then adds an underlying assumption that surface roughness and vegetation are slowly changing processes compared to changes in soil moisture. These same assumptions for the case of backscatter radar have led to a time series soil moisture retrieval algorithm [20–22] that can be extended to GNSS-R land returns. Past studies have invoked the method with measurement latencies as large as two weeks; existing GNSS-R platforms like CYGNSS offer latencies of three days or less. Previous radar studies obtain a dependence on soil moisture in terms of alpha scattering coefficients specialized for backscatter [107, 108]. In contrast for a bistatic specular geometry, the ratio of successive angle corrected measurements at t_j and t_{j+1} is:

$$\frac{\sigma_c^{t_{j+1}}}{\sigma_c^{t_j}} \approx \frac{\Gamma_{LR}^{t_{j+1}}(\epsilon_s^{t_{j+1}})}{\Gamma_{LR}^{t_j}(\epsilon_s^{t_j})} \tag{4.29}$$

where the superscripts now indicate the acquisition time and the assumption that vegetation attenuation cancels has been invoked.

For a time series of N acquisitions, a matrix system of equations of size $(N-1)$ by N can now be constructed having the form $Cx = d$:

$$[C] \cdot \begin{bmatrix} \Gamma_{LR}^{t_1} \\ \Gamma_{LR}^{t_2} \\ \Gamma_{LR}^{t_3} \\ \vdots \\ \Gamma_{LR}^{t_{N-1}} \\ \Gamma_{LR}^{t_N} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \tag{4.30}$$

and

$$[C] = \begin{bmatrix} 1 & -\frac{\sigma_c^{t_1}}{\sigma_c^{t_2}} & 0 & \cdots & 0 & 0 \\ 0 & 1 & -\frac{\sigma_c^{t_2}}{\sigma_c^{t_3}} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -\frac{\sigma_c^{t_{N-1}}}{\sigma_c^{t_N}} \end{bmatrix} \tag{4.31}$$

The matrix in (4.31) is an underdetermined system of equations and cannot be solved without the incorporation of additional information. This additional information is provided in the form of maximum and minimum bounds on the retrieved reflectivities. Previous studies as in [22] have considered trial and error approaches for bound selection, but for global studies on large time scales this is not practically feasible. Instead, maximum and minimum limits for the reflectivites in the retrieval were obtained as a function of space by examining the maximum and minimum reflectivities computed using SMAP soil moisture time histories at a specified location. Given these bounds, a constrained linear least-squares solution of the matrix equation is applied to obtained times series of reflectivities, which can then be mapped into the corresponding soil moisture time series using ancillary soil texture information.

In the results to be shown CYGNSS soil moistures are retrieved daily for a one-month times series (i.e. the matrix solved at each location is 29×30) on a $0.2^o \times 0.2^o$ lat/lon grid. The retrieval process begins with the removal of anomalous DDM measurements, identified through the process outlined in Section 4.9 and is followed with

the introduction of the incidence angle correction terms outlined in Section 4.5. From the remaining pool of DDM measurements, the NRCS is derived and for a given location it is the median value that is used in the time series retrieval. The same process is repeated for every day of the time series retrieval at every, surface, grid cell.

4.7 Ancillary Datasets

In what follows the utility of the time series algorithm outlined in Section 4.6 for the purposes of GNSS-R retrieval of soil moisture on a global scale is illustrated, initially through simulation studies, and subsequently through its extension to measured CYGNSS Level-1 DDMs. The dependence of retrieval performance on various surface properties is also investigated. Doing this, requires a limited range of ancillary data sets that are outlined in this section.

Relating CYGNSS observables to various levels of soil moisture, is ultimately facilitated in the time series algorithm through solving for estimates of surface reflectivity at varying instants in time over the duration of the time series at a given location. The interchange between this, and soil moisture is accomplished through inverting reflectivity estimates to m_v using the dielectric mixing model, outlined in Section 4.4.1, used to interchange between the real and complex dielectric constants and soil moisture. Doing so, requires descriptions of soil attributes including sand and clay fractions as well as bulk density provided on a global scale, which exhibit spatial variation but are temporally invariant. Examples of the three attributes are shown in Figures 4.8-4.10.

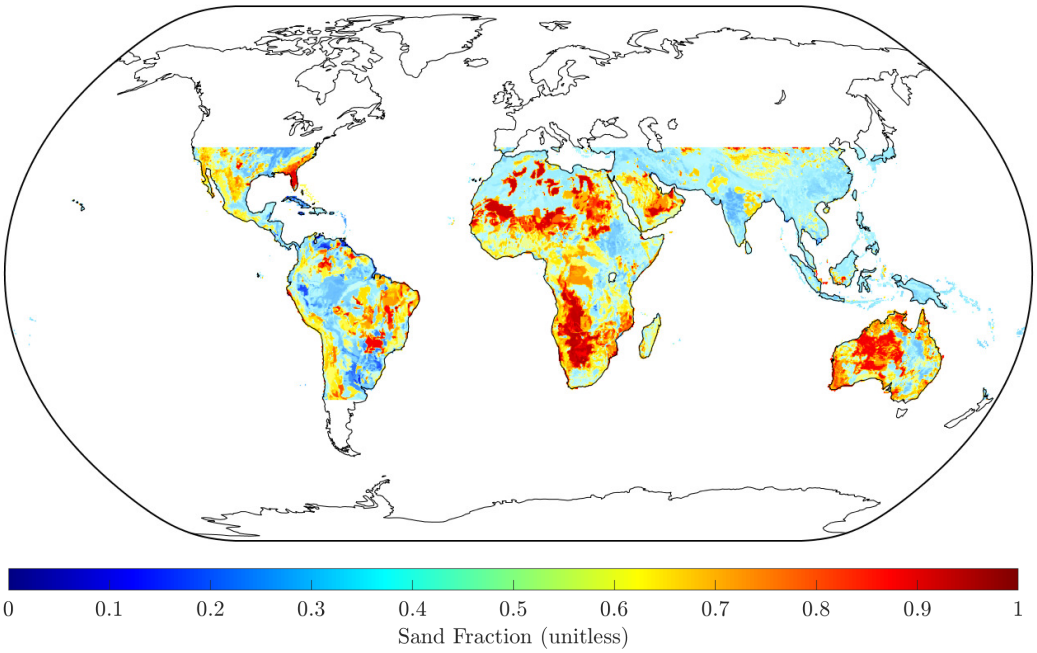


Figure 4.8: Soil attributes necessary for implementation of Peplinski/Dobson/Ulaby soil complex permittivity model: sand fraction

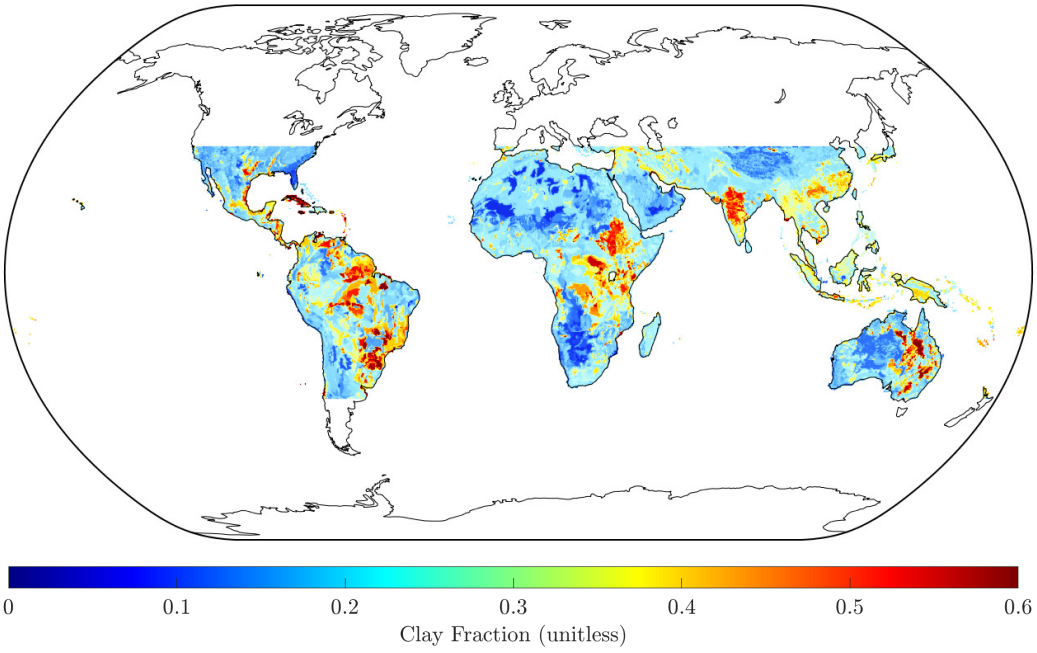


Figure 4.9: Soil attributes necessary for implementation of Peplinski/Dobson/Ulaby soil complex permittivity model: clay fraction

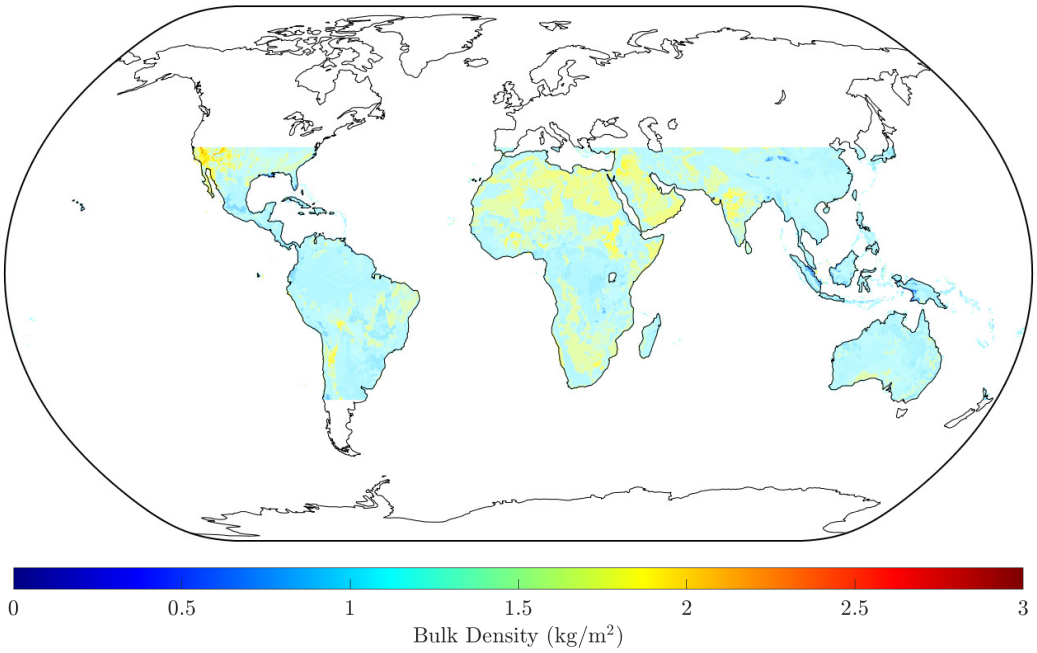


Figure 4.10: Soil attributes necessary for implementation of Peplinski/Dobson/Ulaby soil complex permittivity model: bulk density

The maps shown are based on the same static maps used as part of the SMAP retrieval algorithm [110] and are derived from a combination of inputs from established soil databases including FAO Soil Map of the World (SMW) dataset, Harmonized World Soil Databse (HWSD), State Soil Geographic (STATSGO), National Soil Database Canada (NSDC) and Australian Soil Resource Information System (ASTRIS). In this work the maps have been re-gridded and projected on grids with resolutions ranging between 1-36 km (equal area) and $0.2\times0.2^\circ$ degrees inline with the CYGNSS Level-3 product native resolution.

The above soil attributes are also coupled with ancillary information pertaining to the levels of silt and granular mineral sedimentation mixed with soil and unbound water, for points within CYGNSS's coverage. Silt estimates used in this work are

obtained from the Global Land Data Assimilation System (GLDAS) product, which in turn bases its estimates in large on the FAO SMW dataset. This data is not used in the forward modelling or in enhancing the retrieval but is instead coupled with estimates of soil and clay fractions to determine the percentages of differing soil textures available within CYGNSS coverage, and to later asses the dependence of retrieval error, if any, on soil texture.

A total of 15 soil texture categories exist [111] each of which is defined in accordance with the varying levels of sand, clay and silt a given location comprises. All 15 categories and their relative prevalence for land points within CYGNSS's coverage are identified in Table 4.1.

Table 4.1: Soil texture categories and their relative prevalence over land points within CYGNSS's coverage. Category 14, indicative of standing water, is excluded

Category Code	Soil Texture	Occurrence (%)
1	Sand	3.86
2	Loamy Sand	3.49
3	Sandy Loam	18.15
4	Silt Loam	1.12
5	Silt	0.00
6	Loam	33.05
7	Sandy Clay Loam	12.66
8	Silty Clay Loam	0.10
9	Clay Loam	15.86
10	Sandy Clay	0.29
11	Silty Clay	0.05
12	Clay	10.63
13	Organic Materials	0.17
14	Water	-
15	Bedrock	0.34

In addition to the information defining soil attributes, the use of a dielectric mixing model requires estimates of surface temperature. The dependence of soil's complex

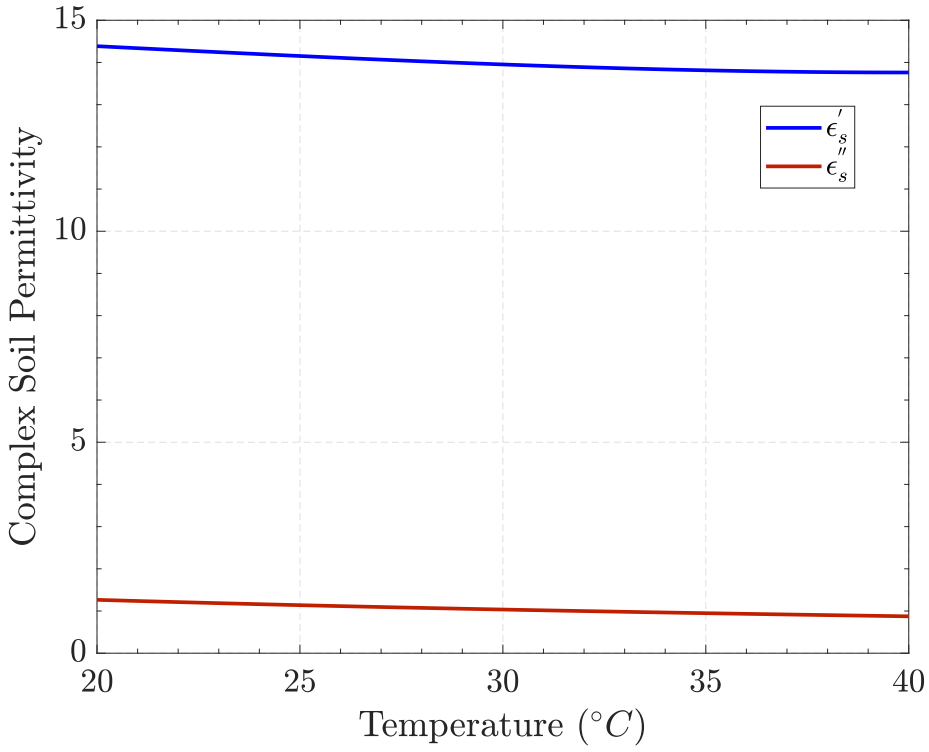


Figure 4.11: Dependence of soil complex permittivity on varying levels of surface temperature

permittivity on surface temperature is weak where, as shown in Figure 4.11, a 100% increase in surface temperature is associated with a less than 5% variation in ϵ_s. Therefore, instead of using daily surface temperature estimates, monthly mean maps derived from NASA's Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) are used which further limits the amount of ancillary information required to conduct the time series retrieval. An example of these monthly temperature maps is depicted in Figure 4.12.

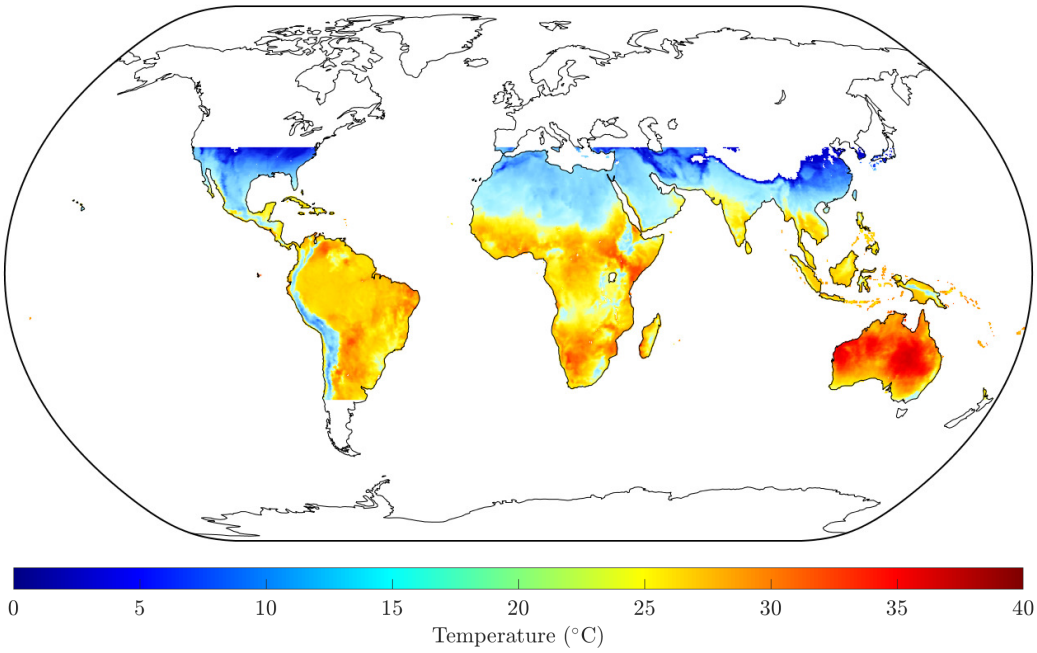


Figure 4.12: Monthly mean temperature estimates for land points within CYGNSS's coverage based on MERRA-2 reanalysis estimates on 0.2×0.2 degree grid

With the fractional sand/clay estimates, bulk density and surface temperature information, the ancillary information required to describe the soil's complex permittivity and surface reflectivity at varying levels of soil moisture is complete.

Other important ancillary information pertains to the levels of soil moisture with which the solutions of the systems of equations described in Section 4.6 are bounded. This is derived from daily estimates of SMAP m_v information. The m_v bounds are used indirectly through solving for maximum and minimum Γ_{LR} with which the underdetermined time series systems of equations is bounded. Examples of the bounds obtained are depicted in Figures 4.13(a) and (b), and are temporally invariant over the duration of an entire month over which the retrievals are typically attempted.

The bounds are nonetheless updated month-to-month to account for their seasonal variability

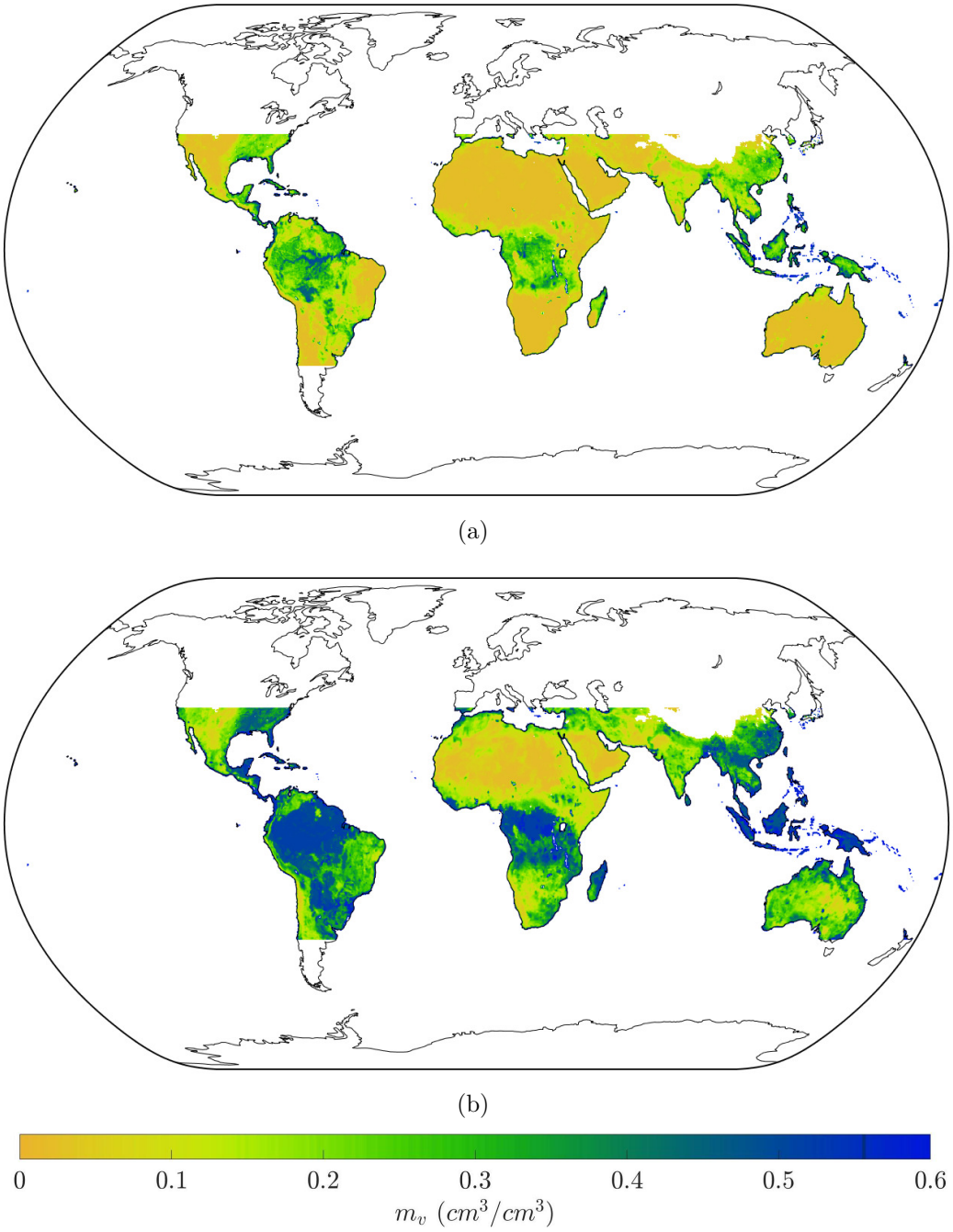


Figure 4.13: Location specific (a) monthly minimum and (b) maximum soil moisture estimates derived from the SMAP observatory over January 2018 used to bound underdetermined time series system of equations

The incorporation of monthly SMAP maximum and minimum soil moistures in the retrieval process makes the retrieval dependent on SMAP ancillary soil moisture information. Therefore the retrievals obtained are not completely independent of the SMAP soil moistures used in their assessment. However, the incorporated SMAP information provides only monthly bounds on the soil moisture retrieval, and does not directly induce correlations between CYGNSS and SMAP retrieved soil moistures within a one month retrieval period. Significant correlations between CYGNSS and SMAP soil moisture time series therefore remain indicative of the soil moisture information provided by CYGNSS. Future work will attempt to reduce the dependence of the retrieval approach on soil moisture information from SMAP, either through the incorporation of soil moisture climatology or other methods for setting bounds on monthly soil moisture variations.

No additional ancillary information is incorporated as part of the retrieval process with actual CYGNSS measurements. However, the forward modelling of CYGNSS measurements as part of the simulation studies explored in Section 4.8 two additional datasets are incorporated to enable accounting for the various effects of surface vegetation following the simple model (4.3). The first relates to a static map communicating the global variation of the b parameter in (4.7) which is based on an empirical relationship [112] relating the amount of water content present within a canopy and the extent of attenuation, impinging microwave radiation experiences. The b parameter is the slope of the straight line relating vegetation optical depth and vegetation water content thereby varying in accordance with land cover class. The b parameter map is depicted in Figure 4.14.

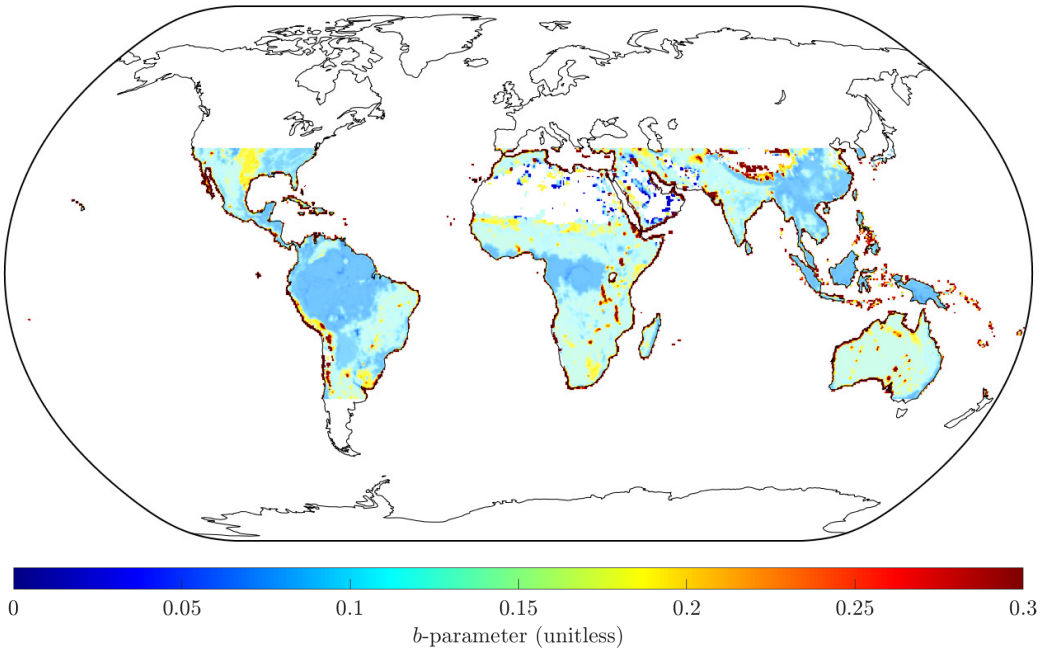


Figure 4.14: Static map of the b-parameter relating the slope of vegetation water content and optical depth of impinging microwave radiation

The second variable required to capture the effects of vegetation is vegetation water content, together with b, enabling the computation of τ_v and the expected attenuation due to vegetation at varying incidence angles based on CYGNSS observations. Estimates of vegetation water content are obtained indirectly from a combined Terra/Aqua MODIS product [113]. The W_c estimates are based on a climatology study that aims to relate Normalized Difference Vegetation Index (NDVI) obtained from the Terra/Aqua MODIS product to VWC through a polynomial fit. Strictly defined, NDVI is an estimate of the prevalence of the color green within an observation footprint and is expressed on a scale ranging between -1 and 1. Regions covered with thick or 'healthy' vegetation are expected to have a higher abundance of chlorophyll which in turn reflects green light in the near infrared (NIR) band whilst absorbing

incoming radiation within the infrared (IR) band. Values closer to 1 are indicative of denser vegetation, values of zero indicate the absence of vegetation and values of -1 are indicative of standing water. Higher NDVI values are therefore indicative of a higher VWC content since healthier vegetation is likely to constitute more moisture within its barks, stems and branches. The second order polynomial [113] used to relate the two that is described by (4.32)

$$\text{VWC} = 1.9134 \times \text{NDVI}^2 - 0.3215 \times \text{NDVI} + \text{SF} \times \frac{\text{NDVI}_{max} - \text{NDVI}_{min}}{1 - \text{NDVI}_{min}} \tag{4.32}$$

and is based on the 10 year historical maximum and minimums of NDVI_{max} and NDVI_{min} with SF describing a factor accounting for the peak amount of water within the stems of available vegetation, varying based on land cover class.

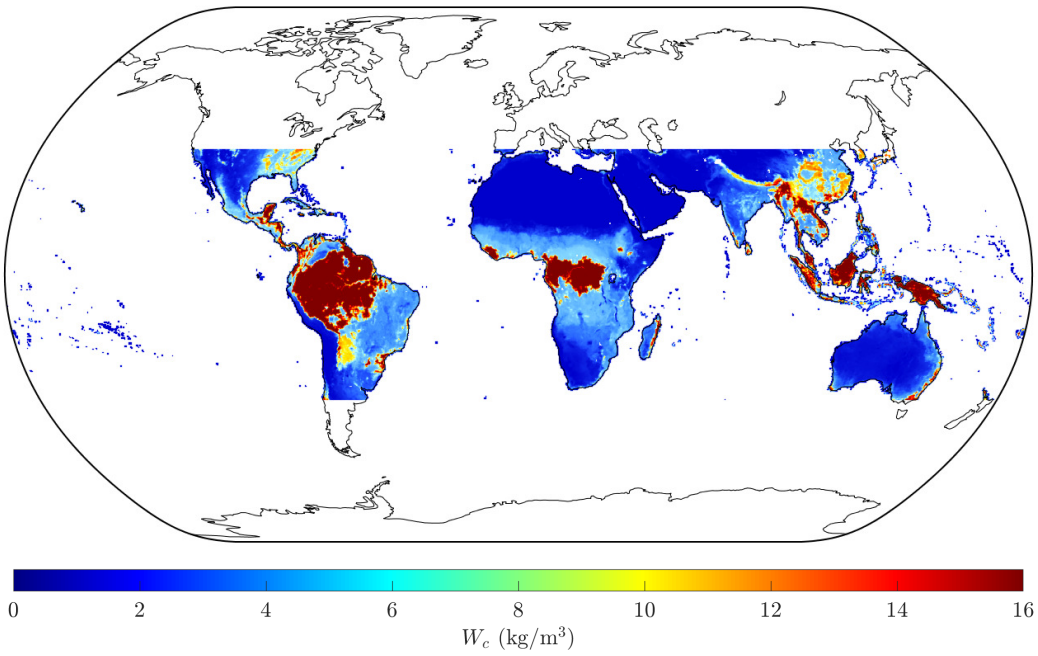


Figure 4.15: Vegetation water content map used in the forward modelling of CYGNSS's measurements to account for attenuation due to vegetation

An example of these VWC maps is depicted in Figure 4.15 where the correlation between high vegetation water content and areas with the thickest vegetation cover, like the Amazon and Congo forests, within CYGNSS's coverage is readily identifiable. Identification of these region is of a particular importance in the assessment of the quality of retrievals. Areas with thick vegetation (VWC $\geq$ 5 kg/m^2) typically comprise forests with a thick canopy known to compromise the quality of retrievals across a variety of platforms. This is primarily due to the heavy levels of attenuation the canopy induces in the forward scattering to the CYGNSS receivers. A further complexity arises due to the multiple scattering effects these thick forests introduce where a reflected GNSS signal may be the result of any number of propagation paths (1) from the transmitter, reflected off the ground and received by CYGNSS (2) from the transmitter, reflecting off any portion of the tree and onwards towards CYGNSS (3) from the transmitter, reflecting off any portion of the tree, to the ground and from the ground towards CYGNSS along with several other potential paths all of which compromise the representativeness of the measured signal(s) of true surface properties. As a result, it is desirable to exclude these regions from representations of error statistics which is accomplished by SMAP through the enforcement of VWC thresholds (and through the exclusion of additional points susceptible to RFI). An example of a quality flag map that was formed based on points not "recommended for retrieval" is shown in Figure 4.16(a) and for a given pixel to be incorporated into the land mask, it needed to have been flagged for poor quality for (at least) an entire month over all months a given retrieval is considered.

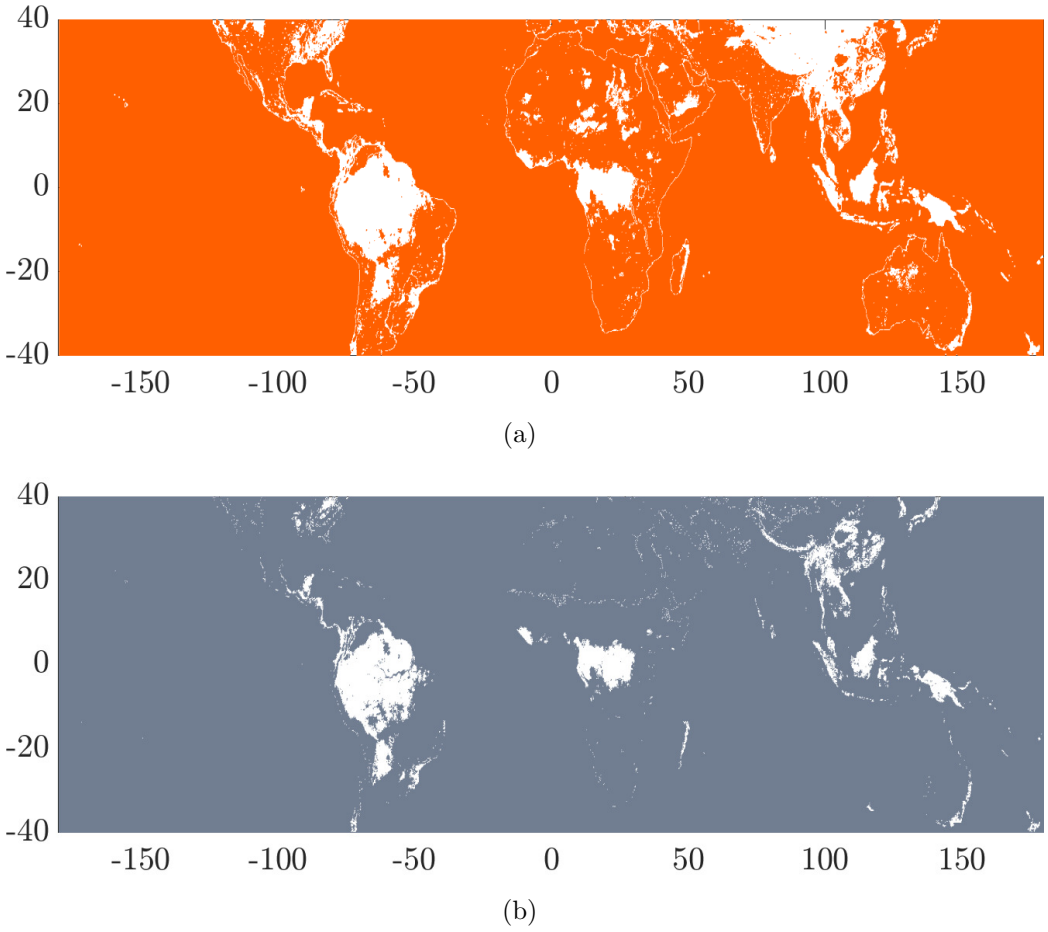


Figure 4.16: Land masks; points eliminated are in white (a) Based on SMAP retrieval quality (b) Based on IGBP land classes

Due to the desire to reduce dependence on dynamic SMAP data, the development of an independent quality flag mask is sought in this work that is based on the International Geosphere-Biosphere Program (IGBP) [114] land cover classification system, in which points indicative of a land cover comprising thick forests are excluded as shown in Figure 4.16(b). With the exception of the Tibetan Plateau, where retrievals are not attempted due to the high elevation of all points within the Plateau, the points identified for exclusion based on IGBP land cover class are in agreement over more

than 90% of locations expected to result in degraded retrievals with the SMAP based quality flag mask. The 17 IGBP land classes and their relative occurrence for all land points within CYGNSS's coverage are summarized in Table 4.2. Regions with classes indicative of thick forests are those that are expected to be associated with higher error levels and a per-IGBP land class error analysis is presented in the results section of this chapter.

Table 4.2: IGBP land class categories and their relative prevalence over land points within CYGNSS's coverage. Category 17, indicative of standing water, is excluded

Category Code	Land Cover	Occurrence (%)
1	Evergreen needleleaf forests	0.12
2	Evergreen broadleaf forests	15.37
3	Deciduous needleleaf forests	0.02
4	Deciduous broadleaf forests	0.98
5	Mixed forests	3.10
6	Closed shrublands	0.32
7	Open shrublands	14.30
8	Woody savannas	9.48
9	Savannas	11.08
10	Grasslands	9.80
11	Permanent wetlands	0.91
12	Croplands	7.99
13	Urban and built-up lands	0.63
14	Cropland/natural vegetation mosaics	5.18
15	Snow and ice	0.18
16	Barren	20.54
17	Water bodies	-

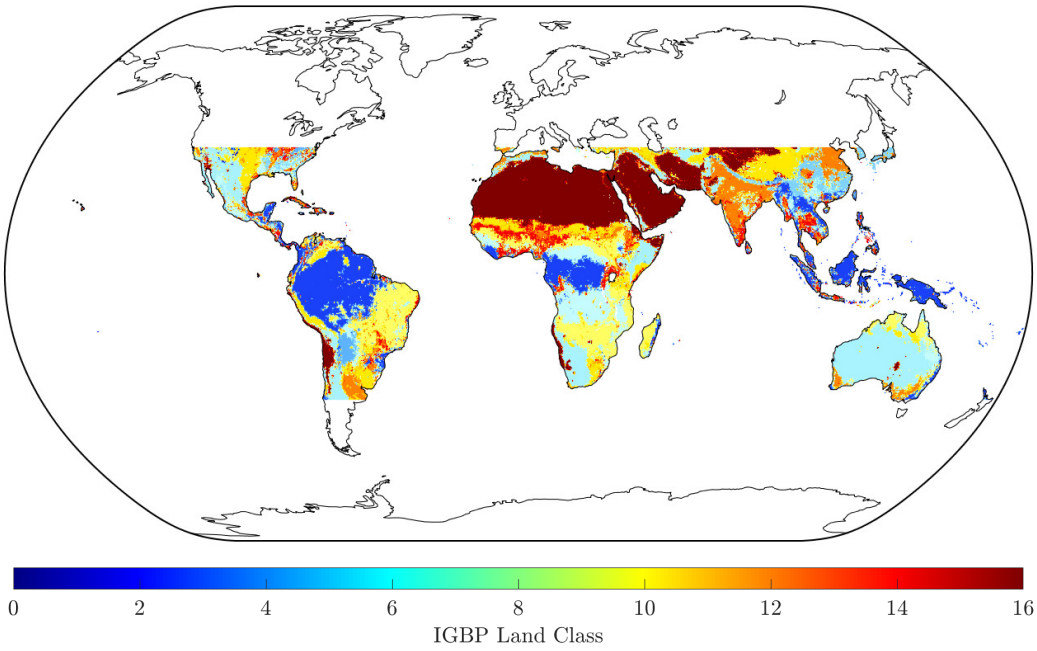


Figure 4.17: IGBP land class map

4.8 Simulation Studies of Time Series Retrieval

A nine month simulation study (April to December 2017) was performed to assess the usefulness of the proposed retrieval methodologies prior to testing with measured CYGNSS data, on both the local and global scales. The simulation study modeled land surface NRCS returns using Equation (4.3) with a fixed MSS value of 0.01 for all measurements, soil moisture values determined from SMAP, and incidence angle determined from the CYGNSS sampling pattern with speckle noise added by describing the measurement as a Gaussian random variable with a variance equal to the mean value divided by the number of incoherent averages (1000 for CYGNSS). Vegetation attenuation was included using τ_v values computed from W_c obtained from a combined Terra/Aqua MODIS product that is temporally varying [113], and a

static global map based on the terrain class was utilized for the b-parameter. Retrieval performance was compared for measurements at time varying CYGNSS angles (for which vegetation attenuation would vary on successive measurements) and a fixed CYGNSS incidence angle (for which vegetation attenuation would remain static for short revisit times). Simulated CYGNSS NRCS measurements were averaged over a 3 day time interval to produce a single time series entry. A time series length of N = 10 (i.e. approximately 1 month) was then used in the linear least-squares solution of Equation (12) to achieve the 3 day retrieved CYGNSS soil moisture. This same system was then solved for each month of the study. Reflectivities given by the solution of the system were finally inverted to ϵ_s and later to m_v through the Peplinski/Dobson/Ulaby dielectric mixing model.

Scatter plots depicting the results of the simulation study for sites of interest are depicted in Figure 4.18 with performance metrics summarized in Table 4.3. To test the effects of coupled VOD-θ_i variability, and not correcting for those to limit the amount of ancillary information used as part of the retrieval, the study is conducted over the 9 month period with and without added ancillary information pertaining to surface vegetation and land cover to enable the incorporation of additional corrections to the retrieval process.

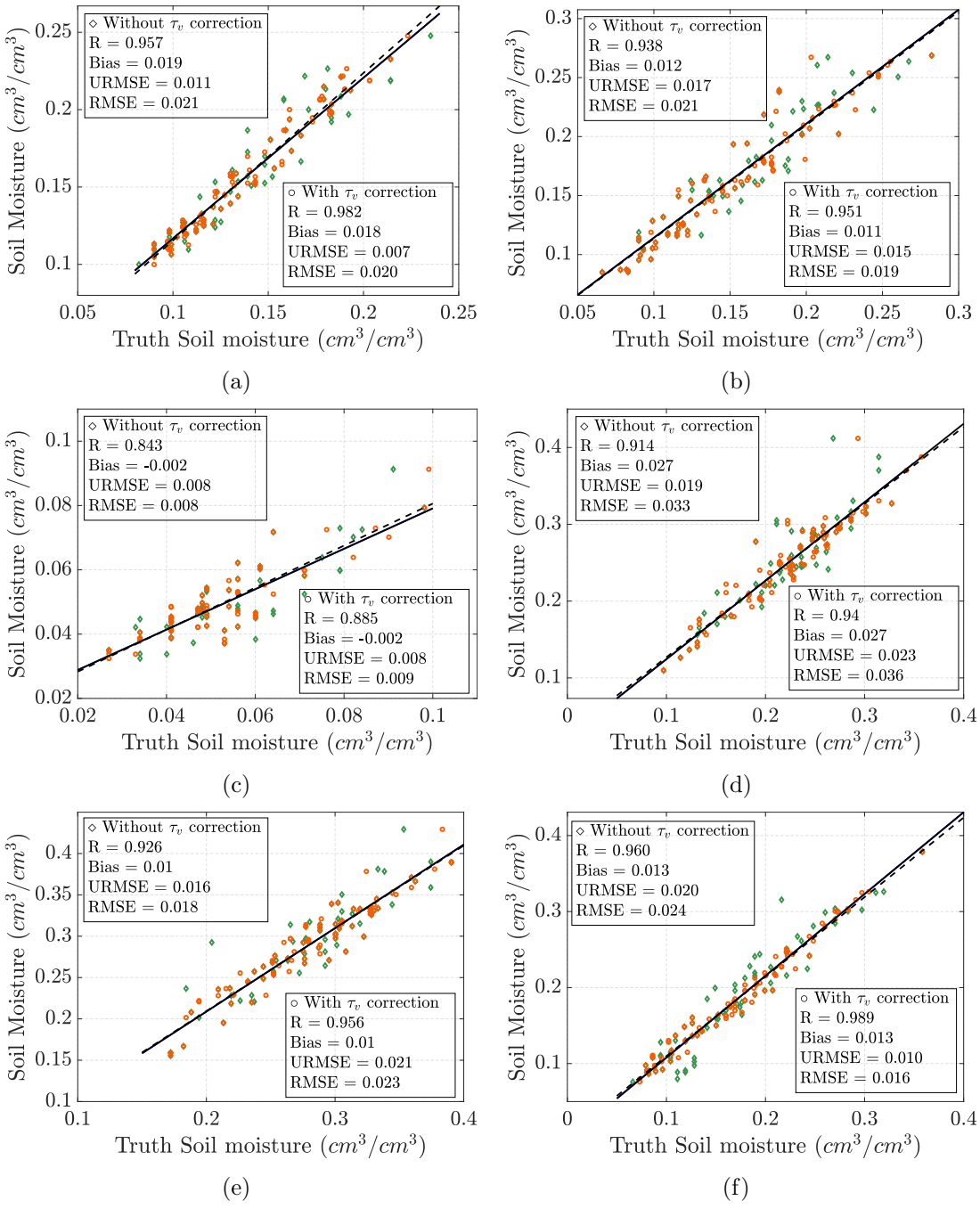


Figure 4.18: Results of simulation study to test time-series retrieval framework with modelled CYGNSS data for sites of interest. Sites are based on co-ordinates and surface properties of SMAP soil moisture calibration and validation sites (a) TxSON, Texas (b) Little Washita, Oklahoma (c) Kuwait, Kuwait (d) Bell Ville, Argentina (e) Monte Buey, Argentina (f) Yanco, Australia

The performance metrics of most interest include correlation, root mean square error (RMSE), unbiased root mean square error (URMSE) and bias. The correlation R is based on the pearson correlation coefficient given by (4.33).

$$R = \frac{\left|\left\langle m_v^S - \overline{m_v^S}, m_v^C - \overline{m_v^C} \right\rangle\right|^2}{\left\langle m_v^S - \overline{m_v^S}, m_v^S - \overline{m_v^S} \right\rangle \left\langle m_v^C - \overline{m_v^C}, m_v^C - \overline{m_v^C} \right\rangle} \tag{4.33}$$

where m_v^S is the reference SMAP soil moisture vector and $\overline{m_v^S}$ is its mean over the length of the time series T_s. The set of retrieved CYGNSS moisture estimates are given by m_v^C and their mean is $\overline{m_v^C}$.

Estimates of RMSE and URMSE are given by (4.34) and (4.35), respectively.

$$\text{RMSE} = \sqrt{\frac{1}{T_s} \sum_{i=1}^{T_s} \left(m_v^S[i] - m_v^C[i]\right)^2} \tag{4.34}$$

in which the bracket notation refers to a point-wise multiplication of the 'truth' SMAP soil moisture vector and retrieved CYGNSS soil moisture vector followed by a summation over the length of the time series.

$$\text{URMSE} = \sqrt{\frac{1}{T_s} \sum_{i=1}^{T_s} \left(m_v^S[i] - \overline{m_v^S} - m_v^C[i] + \overline{m_v^C}\right)^2} \tag{4.35}$$

The bias is also given by (4.36).

$$\text{Bias} = \frac{1}{T_s} \sum_{i=1}^{T_s} \left(m_v^S[i] - m_v^C[i]\right) \tag{4.36}$$

Table 4.3: Performance metrics on local scales of time series soil moisture retrievals using simulated CYGNSS observations.

Site	(a)	(b)	(c)	(d)	(e)	(f)	Overall
Without τ_v Correction							
R	0.96	0.94	0.84	0.91	0.93	0.96	0.923
Bias	0.02	0.01	0.00	0.03	0.01	0.01	0.012
URMSE	0.01	0.02	0.01	0.02	0.02	0.02	0.012
RMSE	0.02	0.02	0.01	0.03	0.02	0.02	0.020
With τ_v Correction							
R	0.98	0.95	0.89	0.94	0.96	0.99	0.952
Bias	0.02	0.01	0.00	0.03	0.01	0.01	0.013
URMSE	0.01	0.02	0.01	0.02	0.02	0.01	0.011
RMSE	0.02	0.02	0.01	0.04	0.02	0.02	0.020

It can be observed that the proposed technique is highly effective at retrieving truth soil moisture levels. However, slightly increased errors are found to be associated with multilook effects in which the same or slightly changed landcover offers a different contribution at different θ_i. However, across metrics of interest, the results of the simulation study showed that soil moisture retrieval performance statistics such as RMSE and correlation were impacted only at the 1-1.5% level by the correction for VOD-θ_i effects, justifying the neglecting of vegetation attenuation variations in what follows. Nonetheless, future studies will attempt to improve the compensation of vegetation attenuation through other approaches or the incorporation of ancillary information. Retrieval error level on the global scale is depicted in Figure 4.19 where on average it is observed that all retrievals are within acceptable error metrics, namely 6% RMSE. Points found to exceed this were exclusively limited to regions known to be prohibitive to reliable soil moisture retrievals due to thick vegetation canopies which

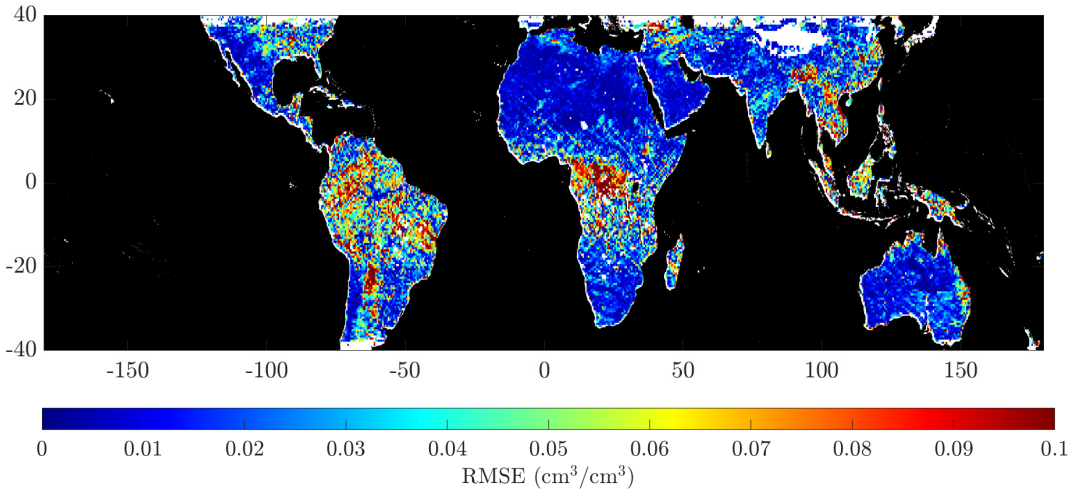


Figure 4.19: RMSE for all land points within CYGNSS coverage based on retrievals throughout 9 months simulation study period

typically induce significant attenuation to the observables of spaceborne platforms, can be associated with significant daily m_v retrieval variability and are associated with soil moisture monthly bounds m_v^b that can approach the extremes of m_v values (ranging between $0 \leq m_v^b \leq 0.8$ cm^3/cm^3).

4.9 Data Refinement

While the results of the analysis presented as part of the simulation study are highly encouraging, their compatibility with results based on the extension of the retrieval methodology to measured CYGNSS data is inherently limited to the applicability of (4.3) for describing the specular scattering of land surfaces. The preceding formulation assumes that the powers from differing surface portions add incoherently, i.e. that the scattering from the land surface is incoherent. This assumption should

be expected to be reasonable for most land surfaces having reasonable elevation variations due to the rapid attenuation of the coherent reflection coefficient with surface roughness, a question that is explored further in subsequent chapters. The fact that numerous examples of coherent reflections have been reported from GNSS-R spaceborne observations motivates careful examination of this assumption as additional GNSS-R land observations are obtained. Initial analyses, explored in further detail in Chapter 5, suggests that coherent reflections over land surfaces are largely associated with the presence of inland water bodies in some portion of the GNSS-R measurement footprint. Coherent returns are characterized by a sharp peak about the specular point with a peak amplitude that is typically many times larger than standard incoherent returns.

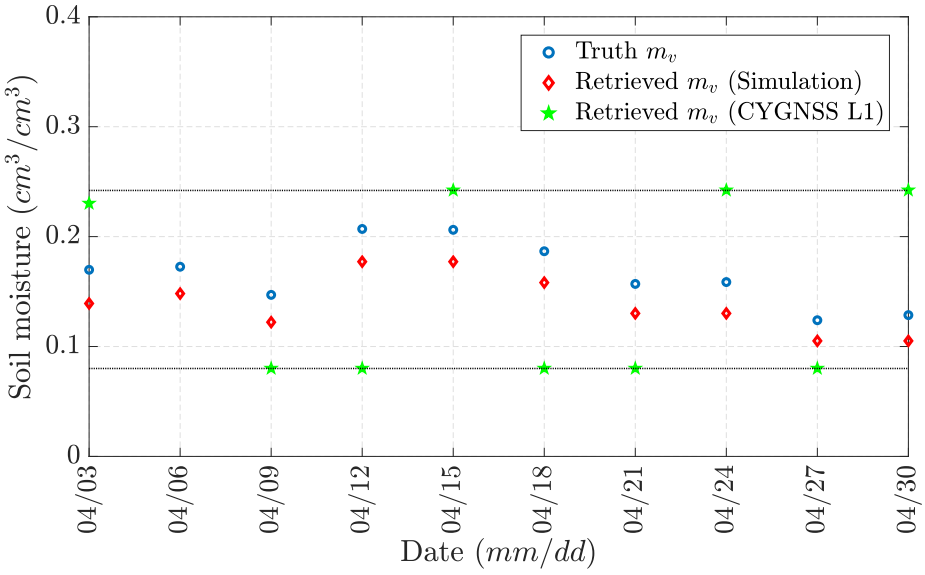


Figure 4.20: Impact of not filtering coherent returns prior to extending time series retrieval methodology to measured CYGNSS data based on a month long retrieval

Including coherent measurements in the time-series retrieval algorithm would be disruptive since such contributions are strong and are not directly representative of soil moisture variations. An example of the impact this has on retrieval performance is shown in Figure 4.20. Due to the disruptive contributions of coherent returns, with magnitudes many times larger than those dominated by incoherent scattering, often retrievals with measured CYGNSS data reverted to the soil moisture bound closest to the expected 'truth' value with abrupt up-down variations in retrieval values. In the absence of external stimulus, such as heavy rain events or flooding, smooth transitions in retrieved m_v such as those obtained by the results of the simulation study with a high degree of correlation to the truth m_v values are expected and sought after. Because of these considerations, a data refinement scheme that is cognisant of these issues is devised and applied to NRCS derived from Level-1 CYGNSS DDMs prior to its inputting into the system of equations (4.31).

Since coherent returns arise due to mirror-like reflection off the Earth's surface, their correlation to the CYGNSS Woodward Ambiguity Function (WAF) will be high relative to incoherent returns. An example of this is depicted in Figure 4.21. The WAF $\chi^2(\tau, f_D)$ has the form:

$$\chi^2(\tau, f_D) \approx \Lambda^2(\tau) \cdot |S(f_D)|^2 \tag{4.37}$$

$$|S(f_D)|^2 = \left| \frac{\sin(\pi f_D T_i)}{\pi f_D T_i} \right|^2 \tag{4.38}$$

$$\Lambda^2(\tau) = \begin{matrix} \left(1 - \frac{|\tau|}{\tau_c}\right)^2 & |\tau| \leq \tau_c \\ 0 & |\tau| > \tau_c \end{matrix} \tag{4.39}$$

for CYGNSS, where T_i is coherent integration time (1 msec) and τ_c represents the C/A code chip duration of 0.97 μsec. To eliminate coherent returns, the cross correlation

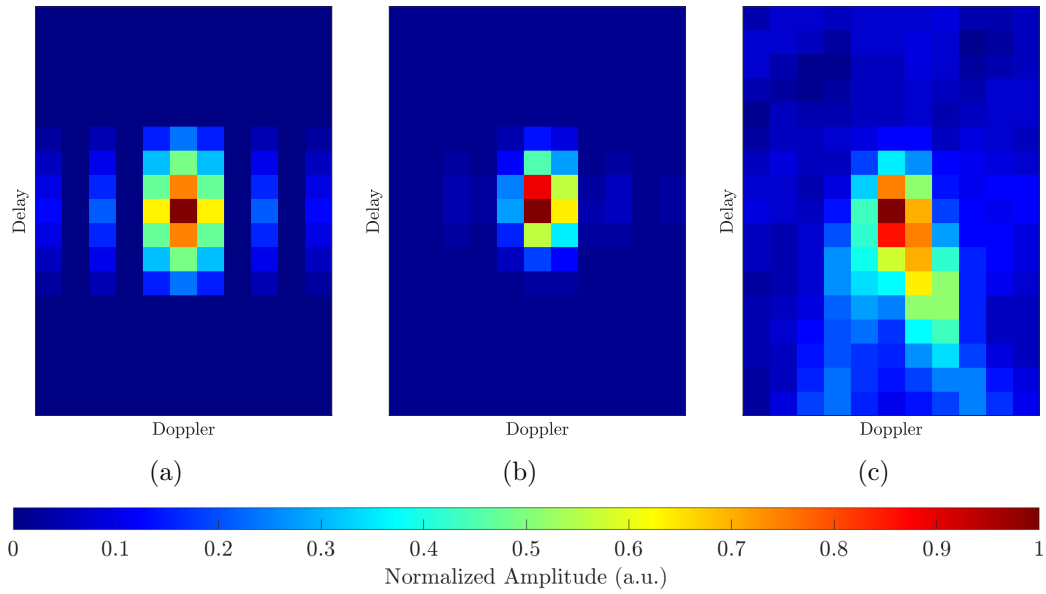


Figure 4.21: Matched filter coherence detection concept for initial elimination of coherent returns as part of change detection soil moisture retrievals (a) Woodward Ambiguity Function in delay-Doppler space (b) Measured coherent delay-Doppler map over land (c) Measured incoherent delay-Doppler map over land

between the WAF and the normalized BRCS DDM is computed, and DDMs with this cross correlation exceeding a maximum correlation threshold (COR_{max}) are discarded.

Measurements with low SNR are also common over land, due to either the location of the specular point within the CYGNSS antenna pattern or due to a low level of surface scattering. It is important also to eliminate low SNR quantities from the time-series retrieval process due to their lower measurement quality. DDMs having SNR less than a minimum SNR threshold (SNR_{min}) are therefore also discarded.

Retrieval performance was found to be optimized if the threshold values and SNR_{min} were allowed to vary spatially, likely due to the varying presence of inland water bodies and low surface scatter returns at differing locations. All thresholds are, nonetheless, temporally invariant. Ultimately, the threshold combination is chosen

to minimize the retrieval errors; maps of the resulting COR_{max} (ranging from 0.7 to 0.85) and SNR_{min} (ranging from 1 to 3 dB) values are depicted in Figure 4.22.

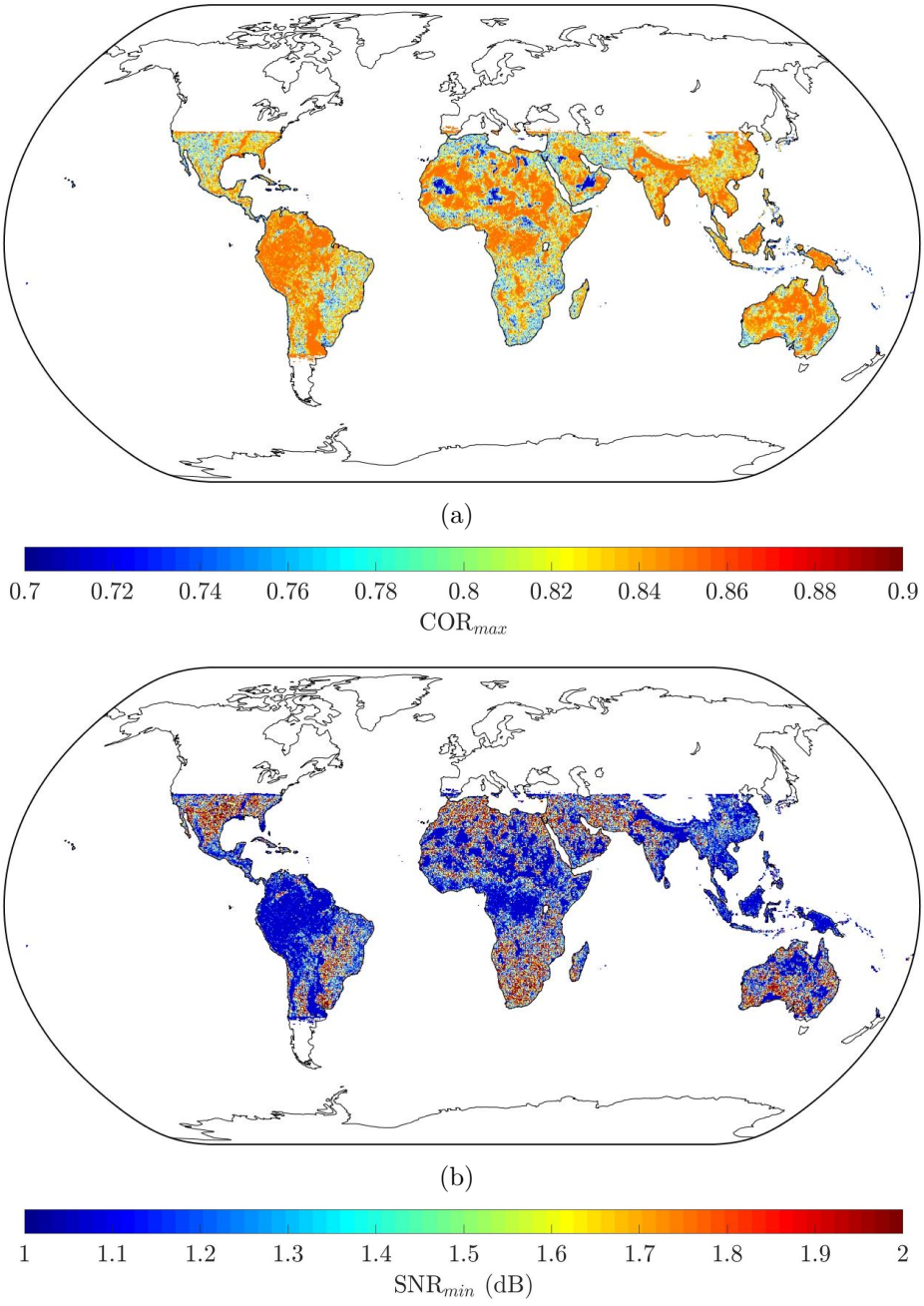


Figure 4.22: Pixel-based thresholds used to eliminate returns from time-series (a) COR_{max} (b) SNR_{min}

4.10 Extension to CYGNSS Land Observations

Using the outlined time series retrieval framework preceded by the refinement of CYGNSS data, the methodology was extended to CYGNSS measurements on the local, regional and global scales.

4.10.1 Site Specific Studies

Soil moistures were retrieved for CYGNSS measurements from December 2017 to Nov 2018. Tests of retrieval performance were used to select COR_{max} and SNR_{min} thresholds at each location that varied from 0.7 to 0.85 and 1 to 3 dB, respectively, to eliminate coherent and noisy returns (Figure 4.22). The site-specific retrievals were for locations of interest for the same time period with thresholds set as to minimize RMSE for each site. For both the global and site-specific twelve month retrievals, COR_{max} and SNR_{min} thresholds are fixed with no temporal variation over the twelve month analysis period.

Daily CYGNSS retrieved soil moistures were averaged over 3 days for comparisons with SMAP Level-2 radiometer derived soil moistures (Data Set ID: SPL2MP_E) obtained from the National Snow and Ice Center (NSIC-DAAC) for all points within CYGNSS's spatial coverage i.e. for latitudes between $\pm\ 40°$.

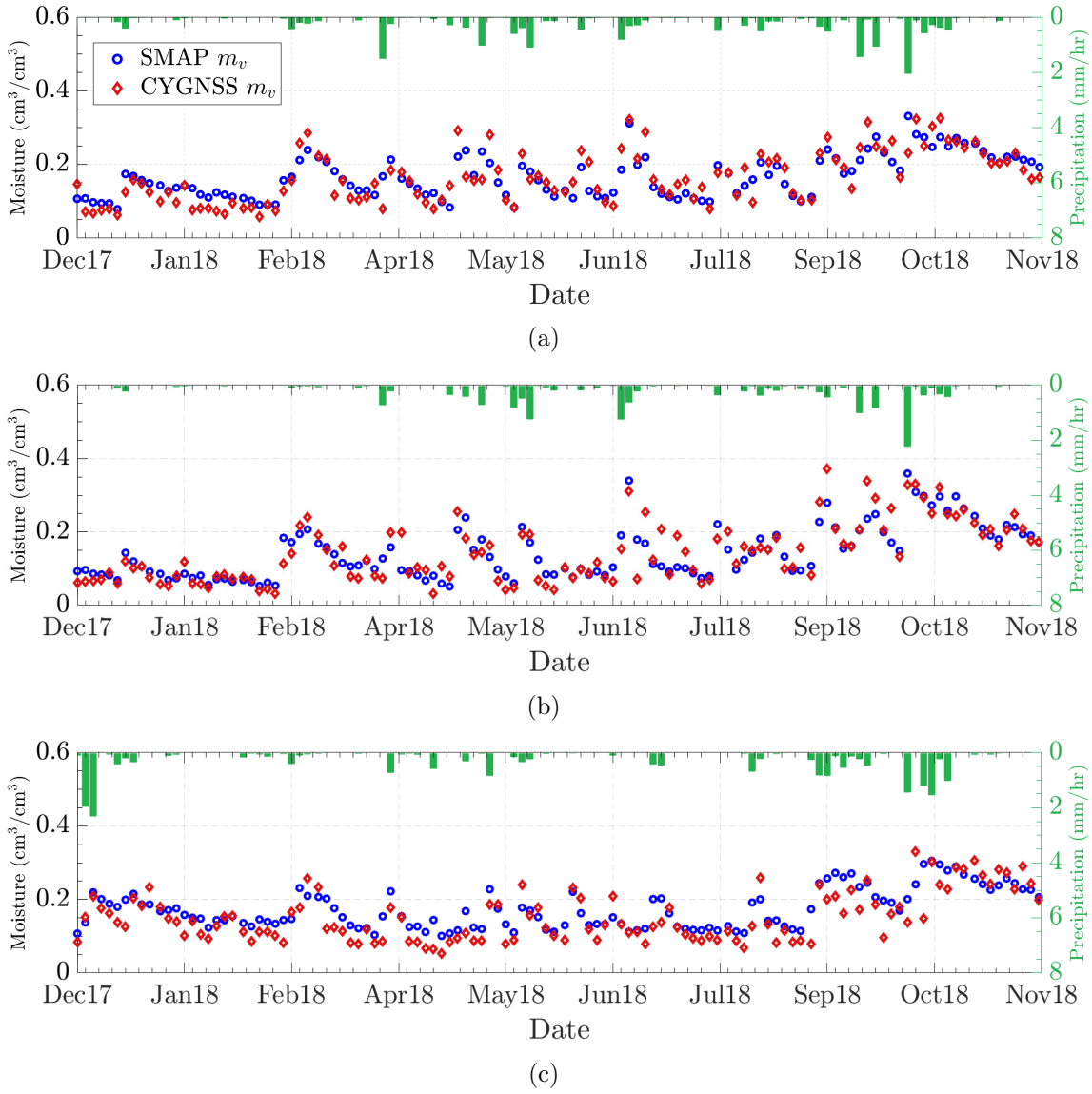


Figure 4.23: Comparisons between SMAP and CYGNSS soil moisture time series for the twelve month period from December 2017 to November 2018 with precipitation data from Global Precipitation Mission (GPM) products (a) Little Washita, Oklahoma (b) Fort Cobb, Oklahoma (c) TxSON, Texas (d) Walnut Gulch, Arizona (e) Yanco, Australia (f) Little River, Georgia

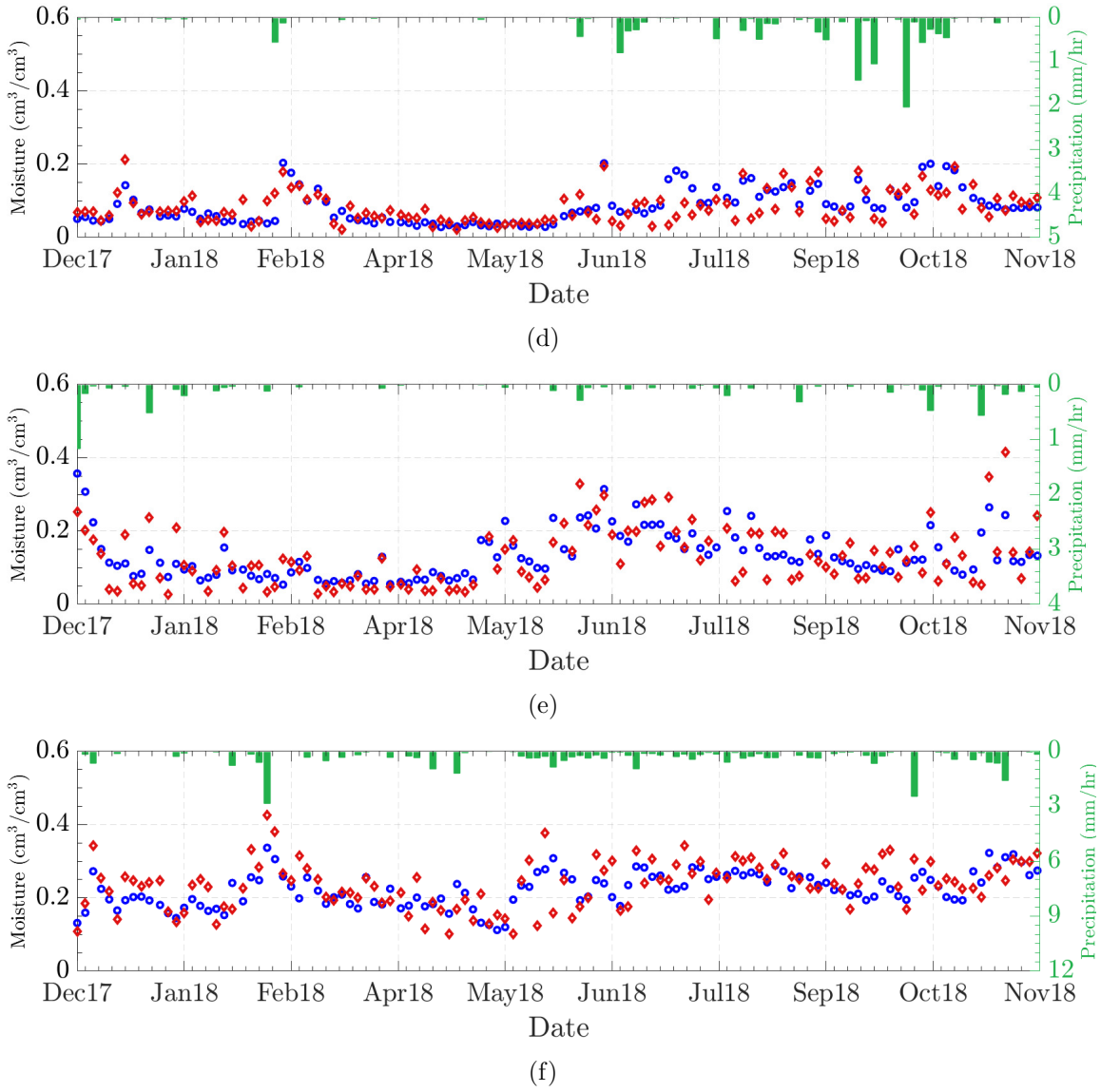


Figure 4.23: Comparisons between SMAP and CYGNSS soil moisture time series for the twelve month period from December 2017 to November 2018 with precipitation data from Global Precipitation Mission (GPM) products (a) Little Washita, Oklahoma (b) Fort Cobb, Oklahoma (c) TxSON, Texas (d) Walnut Gulch, Arizona (e) Yanco, Australia (f) Little River, Georgia

Table 4.4: Sites considered for time series soil moisture retrieval on local scales

Site	Site Name	Location
(a)	Little Washita	Oklahoma, USA
(b)	Fort Cobb	Oklahoma, USA
(c)	TxSON	Texas, USA
(d)	Walnut Gulch	Arizona, USA
(e)	Yanco	New South Wales, Australia
(f)	Little River	Georgia, USA

Table 4.5: Time series retrieval results description and performance metrics. Site-specific retrieval results are over SMAP calibration and validation sites within CYGNSS's coverage.

Site Code	(a)	(b)	(c)	(d)	(e)	(f)
R	0.857	0.889	0.816	0.686	0.739	0.636
Bias	0.001	-0.001	0.020	0.002	0.007	-0.014
URMSE	0.036	0.036	0.038	0.034	0.053	0.051
RMSE	0.036	0.036	0.043	0.035	0.054	0.052
NDVI_{min}	-0.174	-0.195	-0.286	-0.172	-0.724	-0.234
NDVI_{max}	0.636	0.650	0.743	0.500	0.456	0.736
Elevation	0.365	0.482	0.392	1.497	0.128	0.108
Land Cover	Grass	Grass	Grass	Shrub	Grass	Crop
Climate Regime	Temperate	Temperate	Temperate	Arid	Arid	Temperate
Soil Texture	Loam	Silt Loam	Sandy Loam	Loam	Clay Loam	Loamy Sand

The results depicted in Figure 4.23 compare soil moisture time-series for the twelve month analysis period at the six cal-val sites, with Table 4.5 summarizing the corresponding performance metrics. The sites in (1) Oklahoma, Little Washita (2) in Oklahoma, Fort Cobb (3) in Texas, TxSON are characterized by a temperate climate regime comprised predominantly of crop-lands and experiences a wide dynamic range

of soil moisture values with several rain events (and related dry-downs) during the observation period. The site Walnut Gulch on the other hand has an arid climate and consists of open shrublands with foliage that can either be deciduous or evergreen. The Yanco site in Australia consists of mixed croplands and grasslands with a semi-arid regime with mixtures of flat barren surfaces and systems of water bodies in and around the site. Finally, Little River in Georgia is with a temperate climate regime with a surface covered with croplands and natural mosaics [115].

Analysis of Advanced Very High Resolution Radiometer (AVHRR) data also shows a wide variation in vegetation cover, with Normalized Difference Vegetation Index (NDVI) ranging from -0.724 to 0.736 on the scale of -1 to 1, across the sites over the observation period. In spite of heterogeneity of climate regime, land cover and soil texture a significant overall correlation ($R \approx 0.85$ value) and low RMSE (≈ 0.043 cm^3/cm^3) are observed.

The observed correlations between CYGNSS and SMAP soil moisture time series confirm that CYGNSS is providing information on soil moisture. Comparisons against rain rates obtained from the Global Precipitation Measurement (GPM) mission show a clear response to rain events characterized by abrupt spikes in soil moisture content, followed by dry downs after these events. Overall retrieval performance is summarized in Figure 4.24. The success of these retrievals lends support for the efficacy, and validity of the underlying assumptions, associated with the proposed retrieval methodology; in particular those pertaining to the slow evolving nature of surface roughness and vegetation and the ability to eliminate those contributions (4.40) through ratios of successive measurements for a given location.

$$\frac{\sigma_c^{t_{j+1}}}{\sigma_c^{t_j}} \approx \frac{\Gamma_{LR}^{t_{j+1}}(\epsilon_s^{t_{j+1}})}{\Gamma_{LR}^{t_j}(\epsilon_s^{t_j})} \frac{\cancel{\mathrm{MSS}^{t_j}}}{\cancel{\mathrm{MSS}^{t_{j+1}}}} \frac{\cancel{e^{-\tau_o sec\theta_i^{t_{j+1}}}}}{\cancel{e^{-\tau_o sec\theta_i^{t_j}}}} \tag{4.40}$$

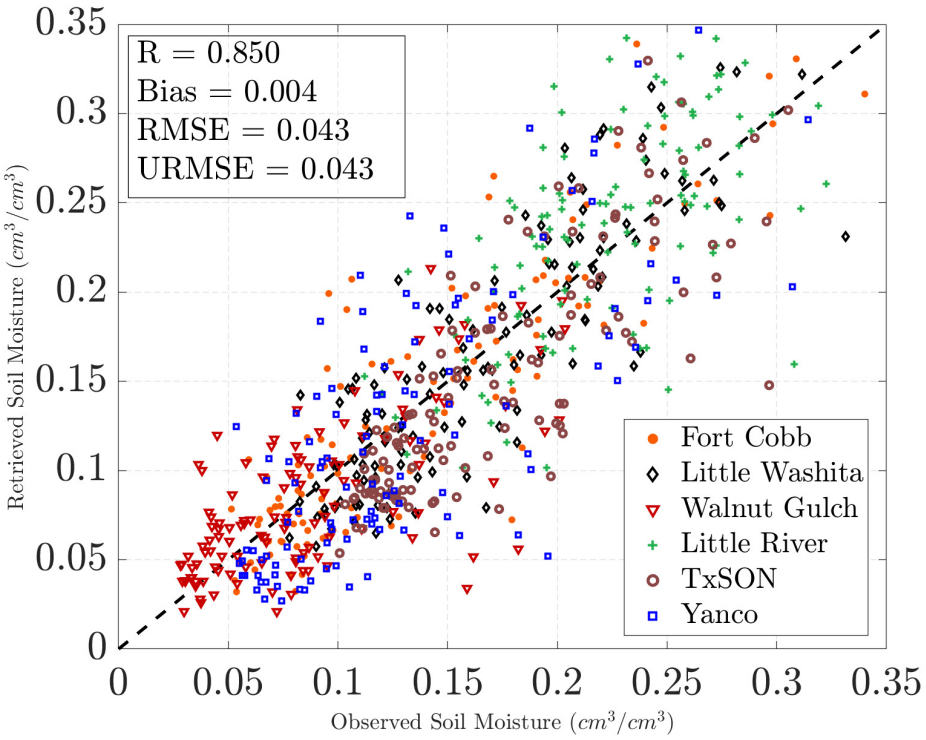


Figure 4.24: Scatter plot of retrieved CYGNSS and SMAP soil moisture retrievals at six SMAP calibration and validation sites, December 2017 to May 2018.

4.10.2 Regional Retrievals

Extreme weather events such as hurricanes are often associated with elevated levels of rain fall, causing increased levels of soil moisture over short periods of time. To test the responsiveness of CYGNSS measurements and time series retrieval approach to known external stimulus, regional retrievals were conducted over the Carolinas

during September of 2018. Over the analysis period Hurricane Florence developed in the Atlantic basin reaching a category of 5 in terms of level of development.

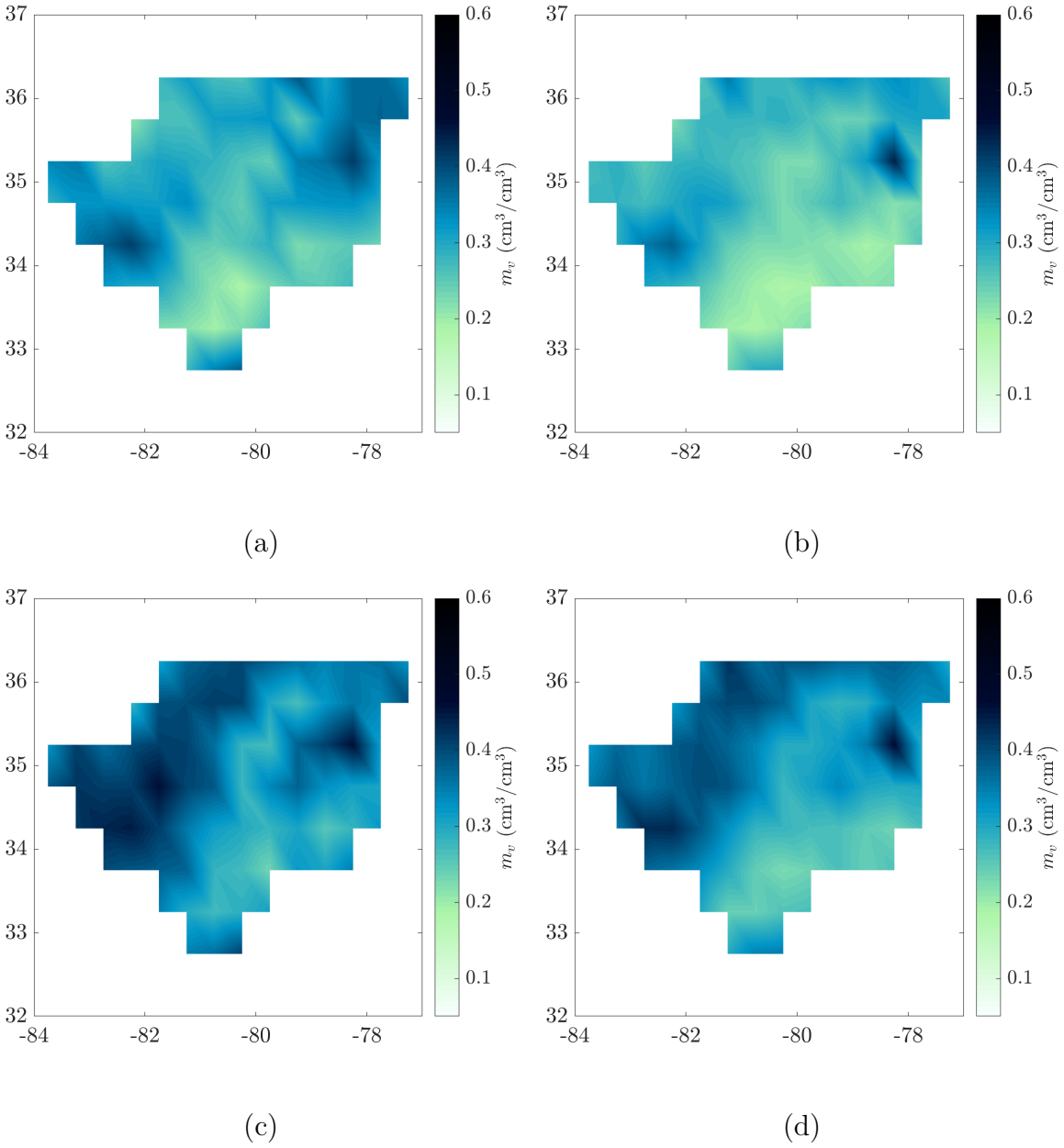


Figure 4.25: Retrieval response to elevated rain rates due to Hurricane Florence across North/South Carolina at 0.5°×0.5° lat/lon resolution (a) CYGNSS m_v (09/22/18 - 09/24/18) (b) SMAP m_v (09/22/18 - 09/24/18) (c) CYGNSS m_v (09/28/18 - 09/30/18) (d) SMAP m_v (09/28/18 - 09/30/18)

As the storm decayed to a level 1, its path moved it to a near-stationary position off the coasts of the Carolinas bringing about significant levels of rainfall prior to moving further inland and bringing about significant levels of storm surge.

To-date, Hurricane Florence resulted in the most moisture being transferred inland within North and South Carolina compared to any other storm on record. As a result, over the course of the month of September the mean m_v surface distribution shifted upwards by more than 30%. Analysis of GPM rain data suggests that this increase was associated with rain rates as high as 30 mm/hr. The response of the retrieval algorithm to this series of events is depicted in Figure 4.25 depicting the results across both states at a $0.5^\circ \times 0.5^\circ$ lat/lon resolution over the same time period. The impact of increased rainfall and the responsiveness of the algorithm to this is clear when comparing Figures 4.25 (a) and (c) which illustrate m_v before and after the storm event. The success of the retrieval methodology is further evidenced by a visual examination of plots (a)-(c) depicting spatial patterns of retrieved m_v values relative to (b)-(d) depicting reference SMAP m_v; where the highly complimentary nature of the retrievals both during the dry and wet phases are readily identifiable both in terms of magnitude and spatial distribution. Further, the highly complementary spatial patterns between retrieved m_v using CYGNSS and SMAP were associated with an RMSE of ≈ 0.05 cm^3/cm^3 over the analysis period.

4.10.3 Global Retrievals

Global retrievals were also attempted over a six month analysis period from December 2017 - May 2019. Figs 4.26-4.28 provide maps of CYGNSS and SMAP retrievals on three different time scales of one day, three days and 1 month at a spatial resolution of $0.2^o \times 0.2^o$.

Areas with thick vegetation are identified and excluded using the flagging methodologies depicted in Figure 4.16 in quantitative comparisons; those correspond to land points of classes 2, 4 and 5 indicative of evergreen broadleaf forests, deciduous broadleaf forests and mixed forests.

Globally, a high degree of correlation is noted with a global average RMS difference over the eighteen month period of ≈ 0.040 cm^3/cm^3, global average pixel based correlation over the same time period ≈ 0.71 and correlation of all retrievals in excess of 20 million points of ≈ 0.96. The extended coverage in a single day provided by CYGNSS as compared to SMAP also suggests the potential of CYGNSS at providing more frequent revisit for soil moisture measurements. Combined, Figures (4.23)-(4.28) illustrate the potential performance of GNSS-R based soil moisture retrievals.

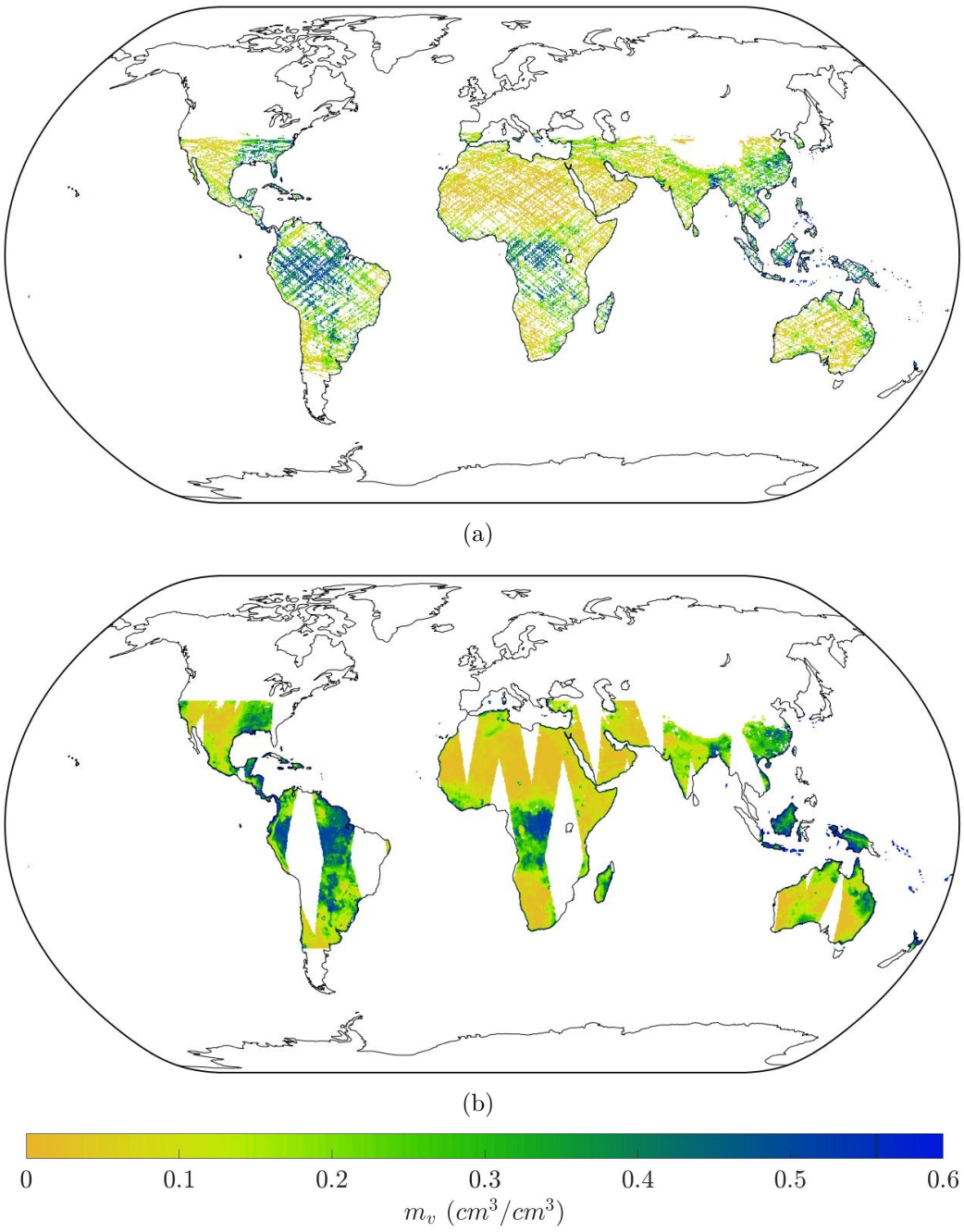


Figure 4.26: Comparisons of reference SMAP and retrieved CYGNSS global soil moistures m_v on time scale of 1 day, 02/03/2018 on a $0.2^o \times 0.2^o$ lat/lon grid (a) CYGNSS (b) SMAP

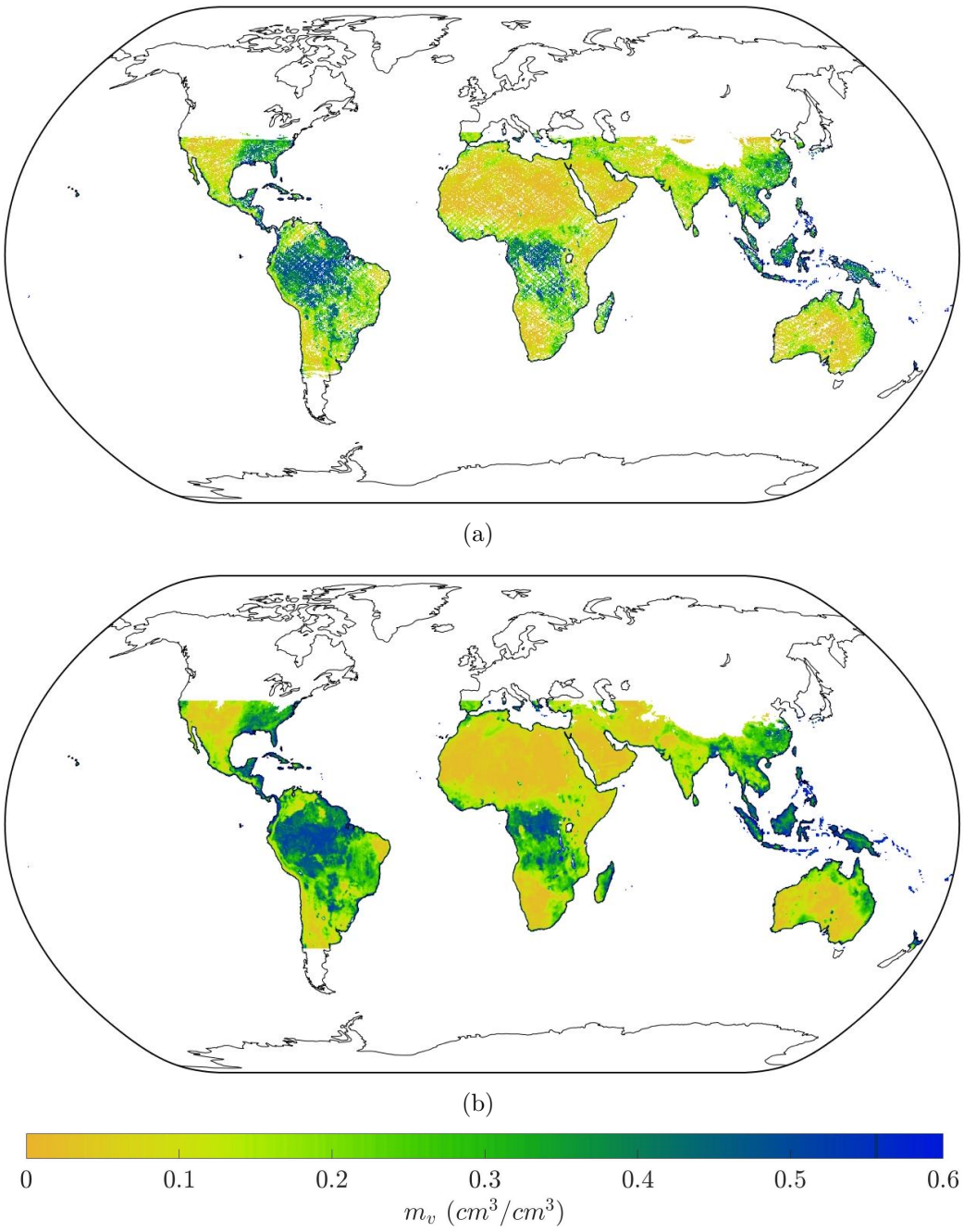


Figure 4.27: Comparisons of reference SMAP and retrieved CYGNSS global soil moistures m_v on time scale of 3 days, from 02/03/2018 to 02/05/2018 on a $0.2^o \times 0.2^o$ lat/lon grid (a) CYGNSS (b) SMAP

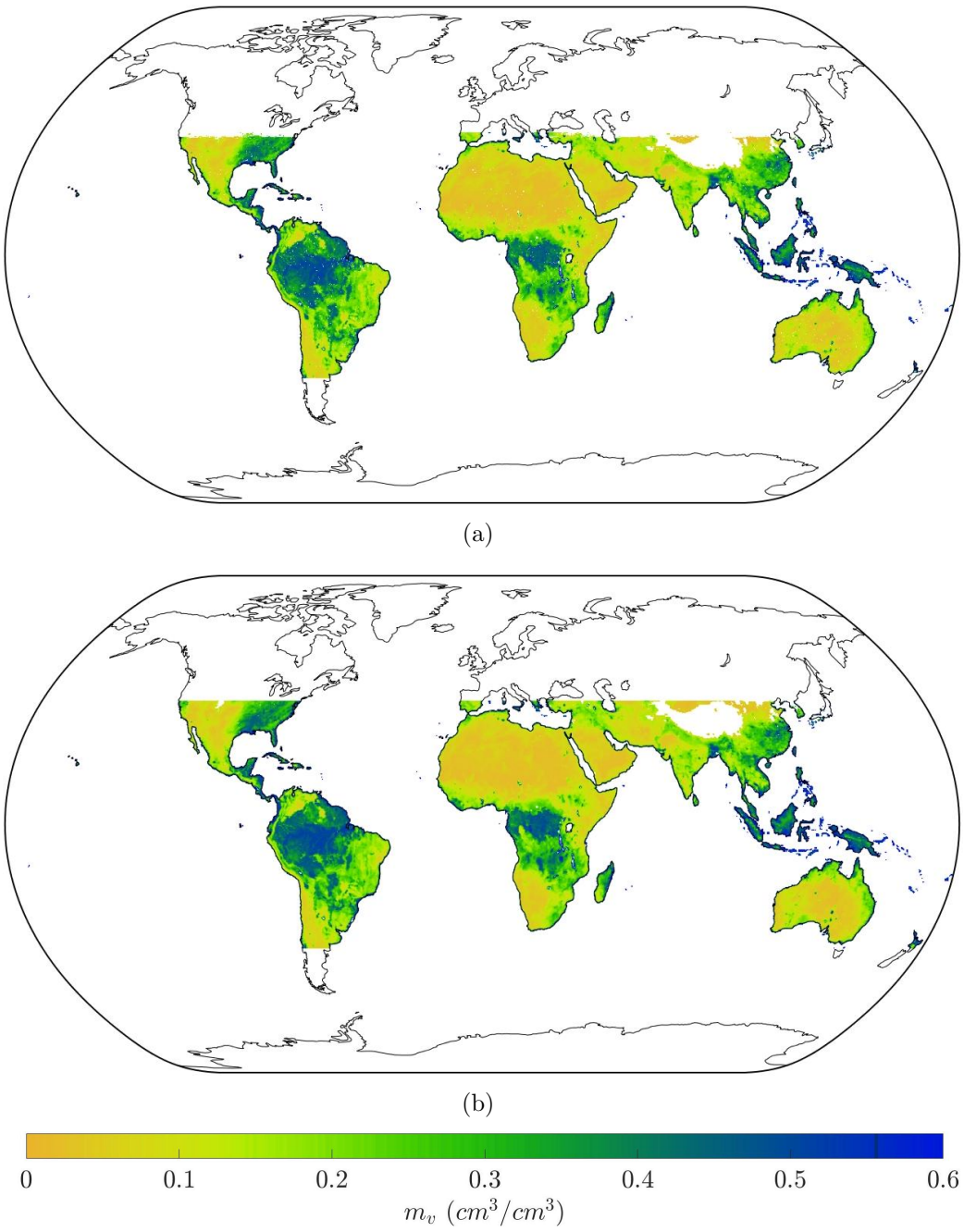


Figure 4.28: Comparisons of reference SMAP and retrieved CYGNSS global soil moistures m_v on time scale of 1 month, from 02/01/2018 to 02/28/2018 on a $0.2^o \times 0.2^o$ lat/lon grid (a) CYGNSS (b) SMAP

Initial studies aimed at investigating the impact of the use of fixed monthly soil moisture maximum and minimum bounds were also performed for site-specific, regional and global examples. In these studies, a single location specific maximum and minimum soil moisture bound was derived using the twelve month SMAP record. The results showed that the degradation in the obtained RMSE varies depending on location. For example, sites like Fort Cobb experience limited degradation due to the limited variability in monthly maximum and minimum soil moistures, whereas sites like Yanco undergo more appreciable degradation due to the larger variability in the monthly soil moisture range. Global retrievals using fixed max/min m_v bounds (i.e. with bounds that do not vary from one month to another) showed the RMSE to increase to ≈ 0.06 cm^3/cm^3. These results highlight the sensitivity of the algorithm to the choice of bounds.

To investigate the correspondence of retrieval error to the physical properties of various surfaces, RMSE may be 'binned' by both IGBP land classes and soil texture classes as shown in Figures 4.29 and 4.30.

Errors, based on retrievals spanning a significant number of points (in excess of 20 million), suggest that the largest error levels with RMSE 0.05 cm^3/cm^3 are over points with deciduous broadleaf forests (land class 4). This is attributed to a number of reasons which include the heavy levels of attenuation this particular land cover class induces to power levels within Level-1 DDMs from which surface NRCS is derived using (4.2). This results in proportionally less power, relative to noise levels, representative of surface scatter within the bins relevant for the computation of σ_0 and as a result a lower correlation to surface properties. Furthermore, from simulation based analysis it was evident that the larger the separation of moisture, or

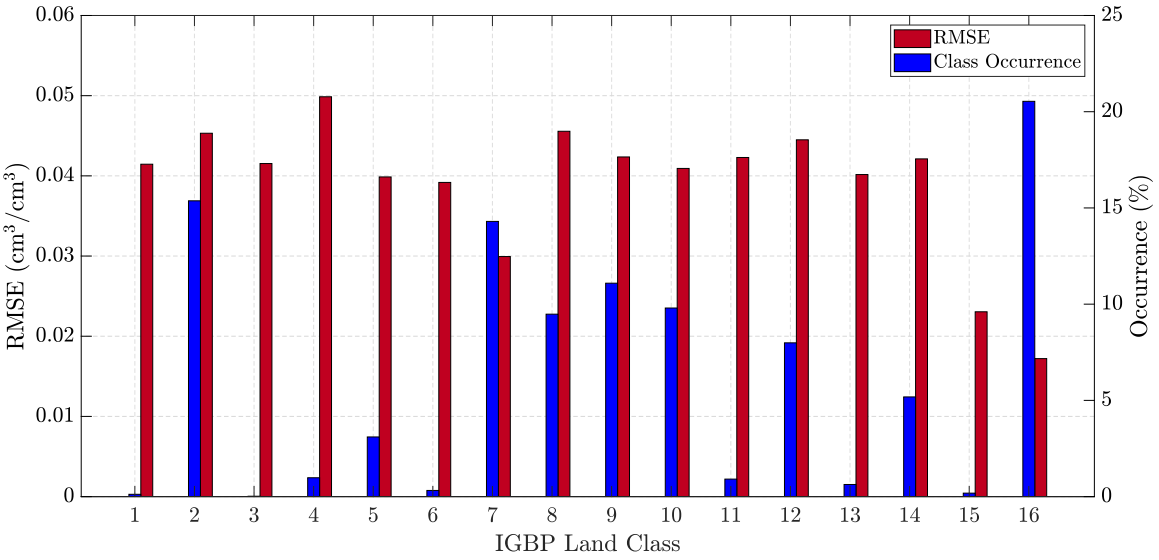


Figure 4.29: Root mean square error for global time series retrievals over 18 months analysis period binned by IGBP land class

surface reflectivity, bounds the more degraded performance gets as a consequence of the underdetermined nature of the system being solved and inverted into a retrieval solution. Over regions with thick forestry, the bounds derived from radiometer reported moistures are global maximums and often near the extreme limits of bounding values (0% or perfectly dry and 100% or comprising volumetrically entirely of water) over the course of just a month. As a result, retrievals over regions with thick forests are found to be associated with above average error levels. Other contributing factors relate to the fact that SMAP 'truth' moistures in these areas are known to be of lower quality and are flagged as being 'not recommended for retrieval', owing to many of the same factors previously mentioned, and may undergo anomalous variations. As a consequence, comparisons to the lower-quality SMAP product in these regions is not as informative owing to the inherent limitations 'truth' and retrievals experience

in these areas. The second highest error level was associated with locations covered with woody savannas (IGBP class 8). In these areas the nature of vegetation may be degrading retrievals not through attenuation but through the introduction of volume scattering effects not accounted for by the simple NRCS model (4.3) that forms the basis for the proposed change detection time series retrieval methodology. The best performing land class, in terms of RMSE, are barren land surfaces (with IGBP class 16). Here, a number of factors contribute to the 60% below average error levels including the fact that, in these regions the absence of vegetation renders considerations relating to multiple scattering effects and attenuation as irrelevant and the received forward scatter is expected to be more directly proportional to land surface properties. Furthermore, areas with land class 16 undergo a more limited dynamic variation of soil moisture and therefore the enforcement of the bounds are more constrictive in terms of potential solutions. Nonetheless, resulting correlations between CYGNSS and SMAP soil moisture time series and low error levels across a wide range of land covers remain indicative of the soil moisture information provided by CYGNSS.

Similarly, error levels are considered as a function of soil texture where it is observed that some of the highest error levels were for points with soil texture classes indicative of various clay formations including sandy clay loam (class 7), silty clay loam (class 8), sandy clay (class 10), silty clay (class 11) and clay (class 12). Because clay inherently entails the saturation of the surface mixture by water, the effects of variations in near-surface soil moisture on surface reflectivity are expected to be, relatively modest as water saturation year round plays a more central role in dictating surface reflectivity. Retrievals associated with these points are nonetheless expected to proportionally contribute less to the overall retrieval error due to their

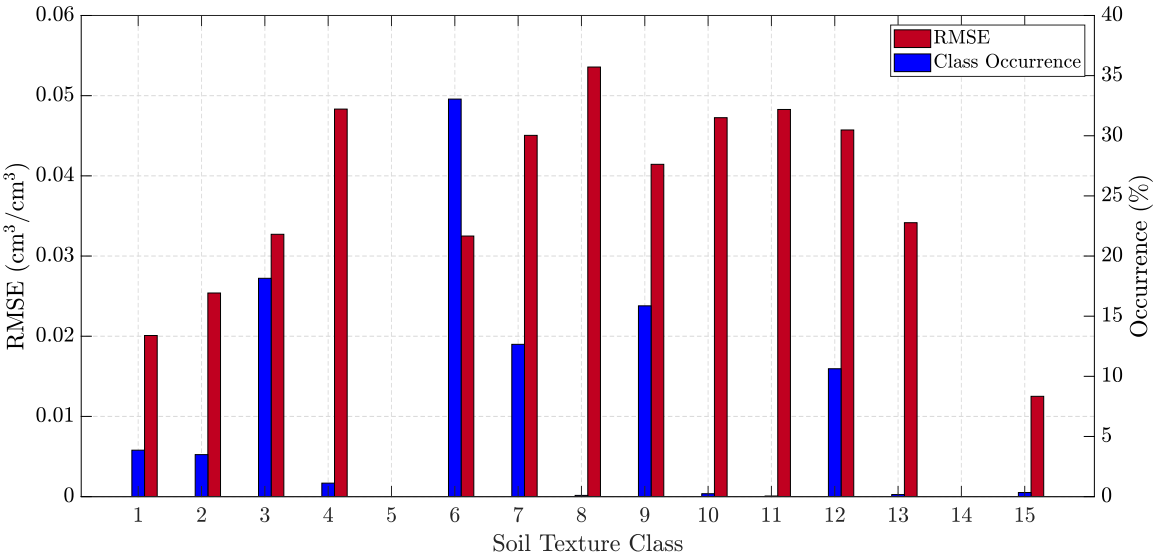


Figure 4.30: Root mean square error for global time series retrievals over 18 months analysis period binned by soil texture class

lower occurrence for points within CYGNSS's coverage. Below average error levels were associated, in large, to soils comprising larger proportions of sand by volume including sand (class 1), loamy sand (class 2) and sandy loam (class 3). The drier, and more porous, nature of these surfaces results in a minimal effect of their composition on overall surface reflectivity and the bulk of surface reflectivity variability becomes the result of near-surface soil moisture variability. Small variations in near-surface soil moisture manifest themselves in significant variations in surface reflectivity, in relative terms, and are therefore more easily related to underlying volumetric soil moisture as part of the retrieval.

4.11 Conclusions

The studies presented in this chapter provide a framework through which soil moisture from barren and vegetated terrains can be retrieved through a time-series approach that leverages the slow changing nature of vegetation and surface roughness compared to soil moisture. This is used to construct a system of equations solved using constrained linear least-squares optimization to obtain surface reflectivities at instants corresponding to the length of the time-series. Obtained reflectivities are inverted to soil moisture through the use of a dielectric mixing model. Further, particular emphasis is placed on practical considerations associated with GNSS-R systems namely the distinguishing characteristic of highly variable incidence angles for returns. Means of mitigation are introduced through which reflectivity is divided into multiplicative products of reflectivity at nadir incidence and a correction function at a given incidence angle with applicability generalized to all soil moisture levels. Results of simulation studies indicate high levels of correlation between estimated and truth soil moisture levels with modest error levels. The effects and means through which multilook effects may be addressed are also discussed, with the same studies suggesting that their impact is, within error tolerance levels, negligible.

The results of extending the retrieval methodology to CYGNSS measurements illustrate the potential utility of CYGNSS observations for retrieving land surface soil moisture through the use of a time-series retrieval algorithm. The one year retrieval results shown provide further information on retrieval performance and the impact of varying land surface conditions, as well as the ability of the CYGNSS constellation to provide observations of soil moisture as it is impacted by extreme weather events. Current GNSS-R systems, namely the CYGNSS constellation, are primarily tasked

with wind speed estimation over ocean surfaces but upcoming land calibration efforts coupled with the results shown herein promise potential for use for the purpose of near-surface soil moisture retrievals.

Chapter 5: Development of A Global Coherence Detection Algorithm

5.1 Motivation

While CYGNSS's primary mission objective is the measurement of ocean wind speeds [117–120], CYGNSS measurements are now being applied for land applications including soil moisture [25, 26, 28–31] and biomass retrievals [35]. The high sensitivity of CYGNSS inland measurements to the presence of water bodies has also motivated studies of flood detection, the measurement of wetland extent, and inland water body mapping [36–39, 121]. Over land, the bistatic specular forward scatter CYGNSS observes may be dominated by coherent reflection or incoherent scattering, with the influence of these contributions determined by a complex combination of terrain properties [122] such as topography, small scale surface roughness, and the presence of inland water bodies. Because coherent reflection and incoherent scattering are differing physical effects that have distinct relationships to scene and sensor properties, it is important to distinguish these effects in measured data so that retrieval methods can be developed based on the applicable physical process. An example of the variability this mixture induces is shown in Figure 5.1 for a 1 day set of CYGNSS NRCS measurements within 50 km of the TxSON measurement site in Texas.

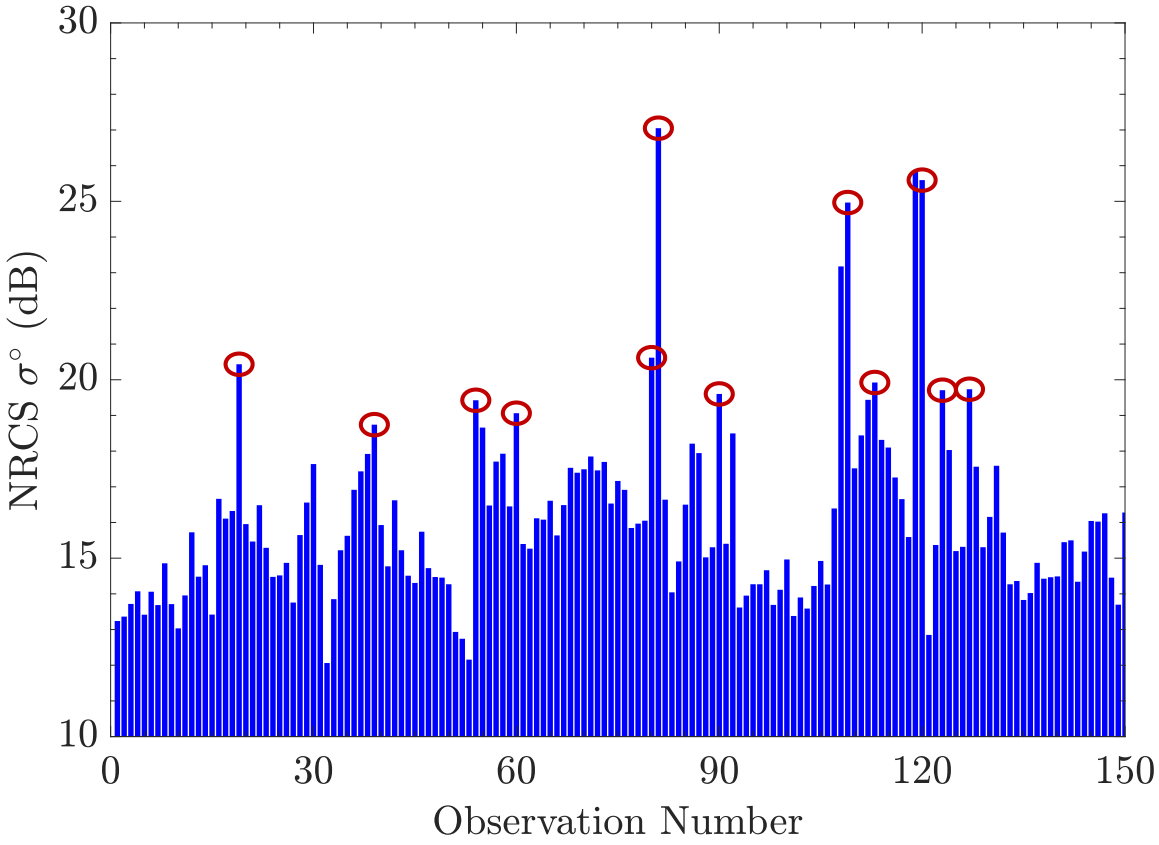


Figure 5.1: Example of bistatic land returns (NRCS) derived from CYGNSS constellation CYG01-CYG08 operating at L-band over site of interest (TxSON, Texas) within a 50 km radius of site location on January 10th, 2019. The variability within a relatively small footprint highlights the impact of the presence of very flat surfaces giving rise to coherent reflections

The results show an NRCS variability exceeding 10 dB, with the circled points suggesting the presence of coherent reflections from the land surface that are less present at other points in the time series. The differing footprints corresponding to a measurement dominated by coherent reflection compared to one dominated by incoherent scatter results in a significant dependence of the returns on the exact location of the track within the region included in Figure 5.1. The development

of methods to identify CYGNSS measurements that are dominated by coherent or incoherent effects is therefore important for supporting science investigations over Earth's land surface.

This Chapter describes an algorithm for detecting coherence in CYGNSS measurements that can be applied for both land and ocean measurements. Because coherent and incoherent returns show differing "spreads" in CYGNSS delay doppler map (DDM) measurements, the algorithm detects coherence using the extent of power spread across a DDM. Detection thresholds are developed through comparisons with CYGNSS "Raw I/F" mode measurements, which are capable of measuring signal phase and therefore provide a more unambiguous detector for coherence. The performance of the detector and selected thresholds is then analyzed over a set of test sites and also over a larger set of CYGNSS measurements.

5.2 Coherent Component of Received Power

The meaning of the the term 'coherence' may be ambiguous. In some contexts, the use of the term coherence may be used to focus on the 'coherence time' of an observed signal, an important quantity for determining the performance of GNSS-R measurements in many applications. The coherence time is impacted by the velocities of the transmitters and receivers, which can vary significantly between airborne and spaceborne observations. In this work however, coherence refers to a scattering situation in which returns from large portions of the Earth's surface arrive at the receiver having identical phase shifts, and therefore add coherently. When time is considered frozen, this phenomenon does not depend on the velocities of the transmitter or receiver. While signals received under a dominantly incoherent scattering regime

may be described using the bistatic radar equation (2.13), the coherent component of received power is described by the Friis transmission (5.1).

$$P_R^{coh} = \frac{P_T \lambda^2 G_R G_T}{(4\pi)^2 (r_R + r_T)^2} \cdot |\Re|^2 \cdot F_{ratio} \cdot e^{-4k_0^2 h^2 \cos^2\theta} \tag{5.1}$$

where P_R^{coh} is the coherent component of the received power and P_T is the transmitted power. Also G_T and G_R are the transmit and receive antenna gains in the direction of the specular point, and $\Re$ is the surface Fresnel reflectivity for incidence angle θ_i and the appropriate polarization combination. The term F_{ratio} is the size of the target contributing coherent reflections, relative to the size of the measurements' footprint. The dominance of one mode of scattering relative to another is contingent upon a number of factors, of which the surface's roughness levels, RMS height h, relative to a wavelength is of most interest. The roughness scales necessary for one component of received power to dominate another may be investigated through the ratio of coherent (5.1) to incoherent power (2.13). This gives rise to the Rician K-factor (K) given by (5.2), defined as the ratio of power corresponding to the dominant/direct (coherent) signal power divided by the mean (incoherent) signal power.

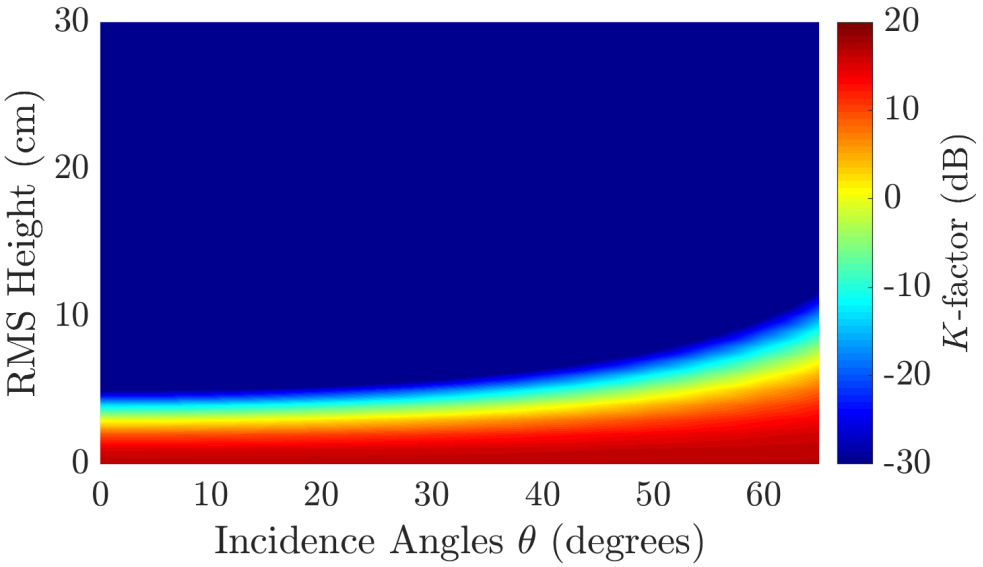
$$K = \frac{P_{coh}}{P_{inc}} = 4\pi s^2 F_{ratio} r_R^2 \cdot e^{-4k_0^2 h^2 \cos^2\theta} \tag{5.2}$$

Expression (5.2) assumes that $r_T >> r_R$ (valid since GNSS-R and SOoP receivers are usually in LEO and GNSS transmitters ar in MEO) and that surface NRCS may be approximated using (5.3)

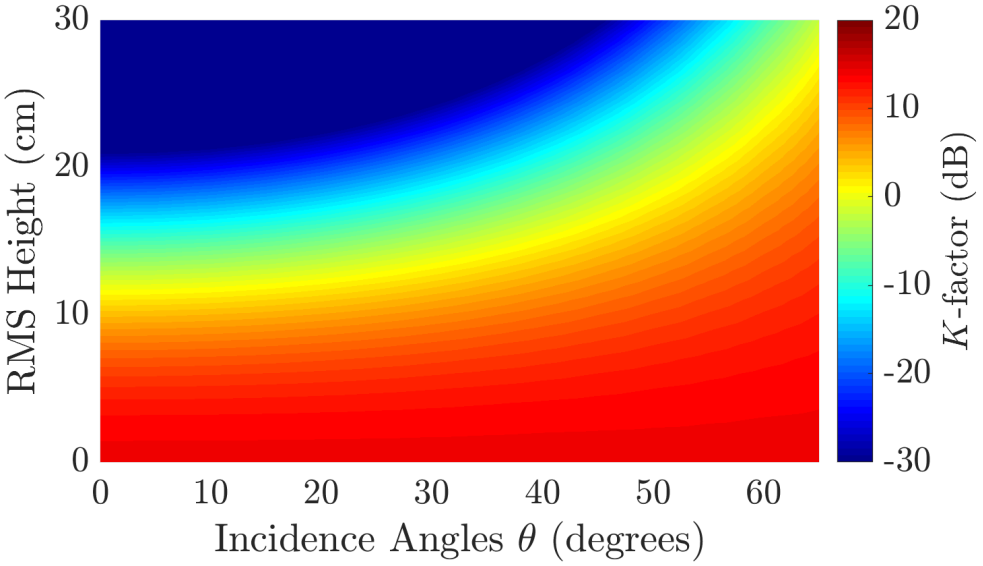
$$\sigma_0 \approx \frac{|\Re|^2}{s^2} \tag{5.3}$$

with s^2 describing surface mean square slope.

Equation (5.2) can be examined as a function of the surface rms height to obtain some insight into expectations for land surface returns.



(a)



(b)

Figure 5.2: K-factor as a function of RMS height and incidence angle (a) 1.575 GHz (b) 360 MHz

Figure 5.2 plots the K factor as a function of rms height and incidence angle for a CYGNSS-like measurement (i.e. 1.575 GHz, 1 MHz bandwidth, and $r_R \approx$ 500 km $<< r_T$) and for a P-band system (360 MHz, 0.5 MHz bandwidth, similar orbit properties) to illustrate expected variability as a function of lower frequency of operation for SOoP systems, assuming that $s^2 = 0.01$. For L-band, the results show a rapid reduction in the importance of coherence for surface rms heights greater than 4-8 cm (as a function of incidence angle) within the first Fresnel zone region. Similar results are reported in [136] where a surface RMS height of ≈ 3 cm was identified as the boundary beyond which coherent returns were not observed. The range of rms heights for which coherence is expected at P-band is expanded by approximately the ratio of the frequencies, but coherence still decreases even in this case for rms heights greater than 15-30 cm.

Figure 5.2 makes clear that terrain rms heights are required to be within a few cm over regions approximately the size of the first Fresnel zone, i.e. hundreds of meters, for the example of CYGNSS. For coherent reflections to dominate, a requirement for terrain heights to be very flat over these scales exists. The prevalence of coherent returns for P-band spaceborne measurements should be expected to be greater than at L-band, but the assumption that terrain surface rms heights remain within $\sim$ 15-30 cm over $\sim$ 1-2 km length scales remains to be validated due to the impact of Earth's natural topography over such scales. It is also noted that while at L-band the penetration depth is limited to near-surface ($\approx$ top 5 cm layer) depths, at P-band the reflected wave will also be impacted by any sub-surface layers or volume scattering. Such contributions may further reduce the coherence of P-band returns.

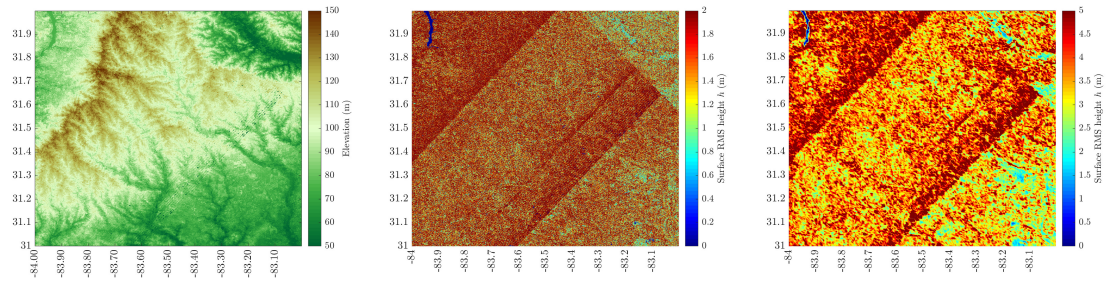
To examine typical roughness scales over sites of interest, reported digital elevation maps are used to estimate RMS heights h given by (5.4) within 100 and 500 meter grid cells.

$$h = \sqrt{\mathbf{E}[\mathrm{DEM}^2] - \mathbf{E}[\mathrm{DEM}]^2} \tag{5.4}$$

An example of this for Little River, Georgia using Shuttle Radar Topography Mission (SRTM) digital elevation maps (DEMs) with a native resolution of 30 meters is depicted in Figure 5.3. RMS heights are found to range between 1-5 meters with the only occurrence of h less than 50 cm being associated with known water bodies, in this case Lake Blackshear centered about 31.9°N 83.95°W. It is nonetheless noted that height estimation errors associated with SRTM estimates are known to be on the order of several meters motivating conducting a similar analysis using higher resolution, less erroneous DEMs. This is explored in Figures 5.4-5.6 using 1 meter resolution LiDAR DEMs, produced as part of the USGS 3DEP initiative, for the sites Fort Cobb, TxSON and Little Washita on similar spatial scales. The absolute limits of 3DEP DEM accuracies range between 5-11 cm which is still higher than the 1-3 cm scale variations this analysis investigates, but nonetheless remain indicative of the mean RMS heights that prevail over much of the world's surface. All RMS height estimates using the 1 meter DEMs, were also found to range between 1-10 meters with a limited number of occurrences of estimated RMS roughness values less than 10 cm. The lower h estimates were, again, exclusively associated with inland water bodies.

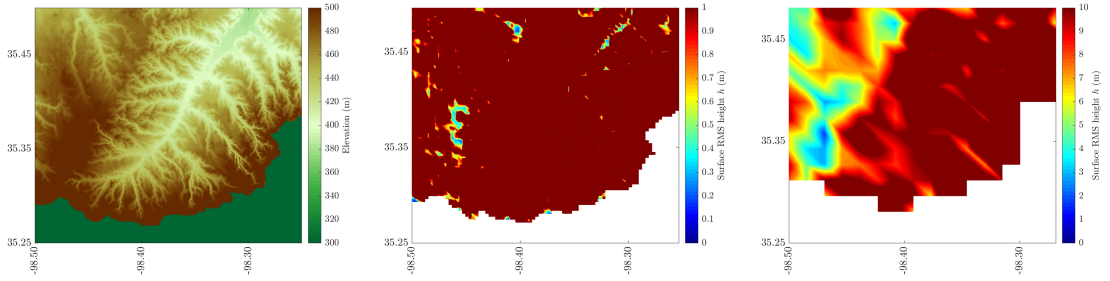
An examination of numerous digital elevation maps on the global scale over these footprint sizes typically shows RMS heights on the order of meters, although it is noted that the errors in existing digital elevation maps are large compared to the cm

length scales of interest here. Nevertheless, the requirement for terrain heights to be very flat over these scales suggests that land surface coherent returns at L-band are most likely to arise from inland water bodies since it is only for such surfaces that rms heights can be assumed to be at the required levels.



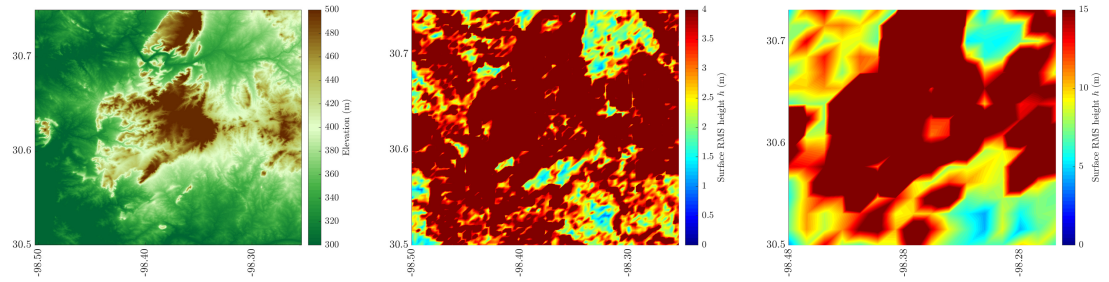
(a) (b) (c)

Figure 5.3: Land RMS height considerations relating to varied levels of coherence prevalence. Maps are for SMAP calibration and validation site Little River, Georgia and are based on 30 meter SRTM DEMs (a) Digital Elevation Map (b) RMS height based on 100 meter height standard deviations (c) RMS height based on 500 meter height standard deviations



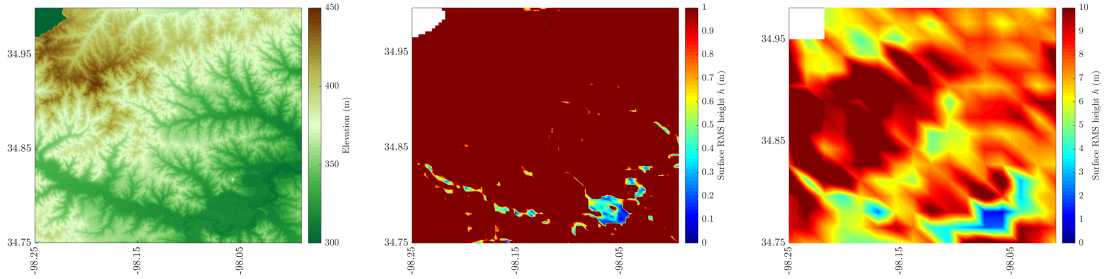
(a) (b) (c)

Figure 5.4: Land RMS height considerations relating to varied levels of coherence prevalence. Maps are for SMAP calibration and validation site Fort Cobb, Oklahoma and are based on 1 meter USGS 3DEP DEMs (a) Digital Elevation Map (b) RMS height based on 100 meter height standard deviations (c) RMS height based on 500 meter height standard deviations



(a) (b) (c)

Figure 5.5: Land RMS height considerations relating to varied levels of coherence prevalence. Maps are for SMAP calibration and validation site TxSON, Texas and are based on 1 meter USGS 3DEP DEMs (a) Digital Elevation Map (b) RMS height based on 100 meter height standard deviations (c) RMS height based on 500 meter height standard deviations



(a) (b) (c)

Figure 5.6: Land RMS height considerations relating to varied levels of coherence prevalence. Maps are for SMAP calibration and validation site Little Washita, Oklahoma and are based on 1 meter USGS 3DEP DEMs (a) Digital Elevation Map (b) RMS height based on 100 meter height standard deviations (c) RMS height based on 500 meter height standard deviations

In light of the distinctly dissimilar behaviours exhibited by coherent reflection and incoherent scattering, the next sections explore the means with which the two may be identified and separated.

5.3 Available Coherence Detection Methodologies

Several previous studies have attempted to develop techniques through which coherence may be detected as a means of distinguishing sea-ice from sea surface measurements [123–127] using standard downlinked GNSS-R DDMs from both the TDS-1 and CYGNSS missions. These detectors fall under one or a combination of the following categories:

1. Signal-To-Noise Ratio: SNR, that is the DDM's peak to mean noise floor ratio, has previously been used as an indicator of the underlying scattering regime contributing the bulk of received power within a given measurement. Because of the expectation that exceptionally flat surfaces give rise to coherent reflections, high SNR values are used as an indication of the presence of coherent reflection. Other variations of this include a similar peak-to-noise ratio, whilst introducing terms correcting for the effects of transmitter/receiver antenna gain patterns and ranges from/to the specular point.

2. Delay Doppler Map Average: This indicator is formed by computing the average of total power about the specular bin, or DDM peak, within delay-Doppler windows of varied extents. Similar to SNR, higher DDMA values are used as indicators of dominant coherent reflection.

3. DDM Maximum Amplitude: This is simply the peak value of a given DDM and corresponds, like SNR, to the state of a given measurement.

4. Average Noise Floor Value: This is obtained by analyzing the mean noise floor values at points corresponding to negative delay (top rows of a Full DDM).

Consider CYGNSS Full DDMs as an example, over the ocean the location of the specular bin will be at delay bin 64 (± 1-2 delay bins). Because this bin is associated with the path of shortest propagation distance, no power maps to the negative delay points and delay rows beyond row 64, which comprises thermal noise power. In contrast, the mean elevation over land is above the nominal sea surface heights used to estimate the location of the specular bin, and as a consequence significant amounts of power can map to the upper (typically noise only) delay rows thereby increasing the average noise floor value. The use of this as an indicator of coherence inherently implies that all GNSS-R returns over land are by default, coherent.

5. Distance to decay: DDMs dominated by coherent reflection are typically associated with sharp peaks about the specular bin and a rapid decay of power levels beyond ± 1-2 delay bins. The distance to decay indicator typically reduces the two-dimensional DDM measurements to a 1-dimensional (power-vs-delay) waveform. If power levels decay by a preset percentage within a given number of delay bins, this is taken to indicate coherence. In contrast, longer decay lengths are used as indicators of incoherence.

6. Leading/Trailing Edge Slope: Similar to the distance to decay metric, this approach recognizes the peakedness of coherent DDMs and uses the steep power slopes before/after the specular bin as indicators of coherence. This can include the power slope leading to the specular bin, the power slope trailing the specular bin or ratios of the two (which are expected to be close to ≈1 for coherent returns).

7. Combined Approach: This relates to the use of a weighted combination of some or all of the aforementioned metrics, typically as part of an Artificial Neural Network (ANN) framework. Because the bulk of previous studies have sought to develop coherence detectors as a means of identifying sea ice, an independent truth dataset exists, and therefore the ANN is used to weight the estimates of each of the aforementioned metrics in a manner that achieves the best fit to the reference dataset.

Challenges nonetheless arise with the use of many of these metrics on the global scale, using CYGNSS Level-1 DDMs. All approaches that use the magnitude of GNSS-R observables as indicators of coherence including SNR, DDMA and DDM peak neglect the facts that coherence can exist at any peak power level. Furthermore, their use makes the success of the detection methodology contingent on the accuracy of power calibration which remains limited. Furthermore, while 'ideal' coherent and incoherent DDMs are associated with some levels of characteristic slope behaviour, realistic measurements undergo a wide range of distortions that renders the ability of selecting a fixed detection threshold that is valid on the global scale difficult. For example, high SNR incoherent DDMs exhibit a large noise floor - peak separation resulting in high slope values that would typically be used as indications of coherence. In contrast, low SNR coherent DDMs exhibit a low noise floor - peak separation and the transitions to/from the peak bin would be, in relative terms, smooth. As a result, this would be characterized as being incoherent. Finally, while ANN based combined approaches have been associated with near-ideal detection capability in numerous reported scholarly investigations, their success is inherently contingent on the presence of a 'truth' dataset used to tune the weighting of each of the coherent

metrics used as part of the training process. In the absence of a global truth dataset, the ability to extend a similar approach to the global scale remains limited. Taking into account some of the limitations of established approaches, subsequent sections highlight a new detector capable of detecting coherence on the global scale using CYGNSS's Level-1 measurements.

5.4 DDM Power Spread Detector

This work proposes a method related to the extent of power spread across the DDM, and compares with a subset of previously developed methods.

Coherent returns typically arise due to a mirror like reflection of the transmitted signal off the Earth's surface. This is due to the smoothness of the surface within the observed footprint, with properties that are explored in [90]. Due to the low level of surface roughness, the coherent received field is reflected from the specular point and the small surface area surrounding it on the order of the first Fresnel zone in size [91] with dimensions varying in accordance with the observing geometry. Nonetheless, it is noted that other studies [128] have highlighted the complexities associated with exact quantification of the spatial extent of this small surface area under a scattering regime dominated by coherent reflections, with similar ambiguities arising due to varying observation geometries [129]. Coherence therefore refers to a scattering situation in which returns from large portions of the measurement's footprint arrive at the receiver having identical phase shifts, and therefore add coherently. The DDM is then characterized by a sharp peak at the specular bin and a concentration of power within $\pm$ 1 delay bin and $\pm$ 2 Doppler bins about the DDM peak. In contrast, rougher surfaces give rise to incoherent returns in which the received power is scattered from

points at significantly farther distances from the specular point. In delay-Doppler space this translates into a larger spread in power at delay and Doppler values beyond the specular point.

CYGNSS land measurements show a variety of behaviors, including cases with power spread throughout the DDM as well as mixed DDMs indicative of partial coherence in which the power from incoherent scattering and coherent reflection are both observable. These mixed cases also have appreciable power spread compared to coherent DDMs. In what follows, incoherent, mixed and low SNR DDMs will be classified as non-coherent, so that the detected DDMs can be regarded as "purely" or "dominantly" coherent. The proposed DDM Power Spread Detector (DPSD) compares the power concentrated in the limited delay-Doppler space about the specular bin to the total power outside this region, which is expected to be indicative of the nature of the return due to the correspondence of the physical mechanism of scattering and the extent of power spread. An estimate of power spread is developed through a power ratio (PR):

$$PR = \frac{C_{in}}{C_{out}} \tag{5.5}$$

with

$$C_{in} = \sum_{i=-1}^{1} \sum_{j=-2}^{2} \mathrm{DDM}(\tau_M + i, f_M + j) \tag{5.6}$$

$$C_{out} = \sum_{i=1}^{N_\tau} \sum_{j=1}^{N_f} \mathrm{DDM}(i, j) - C_{in} \tag{5.7}$$

where C_{in} is the integrated power in raw counts within the (3×5) delay-Doppler region about the specular bin, C_{out} is the integrated power across the DDM excluding

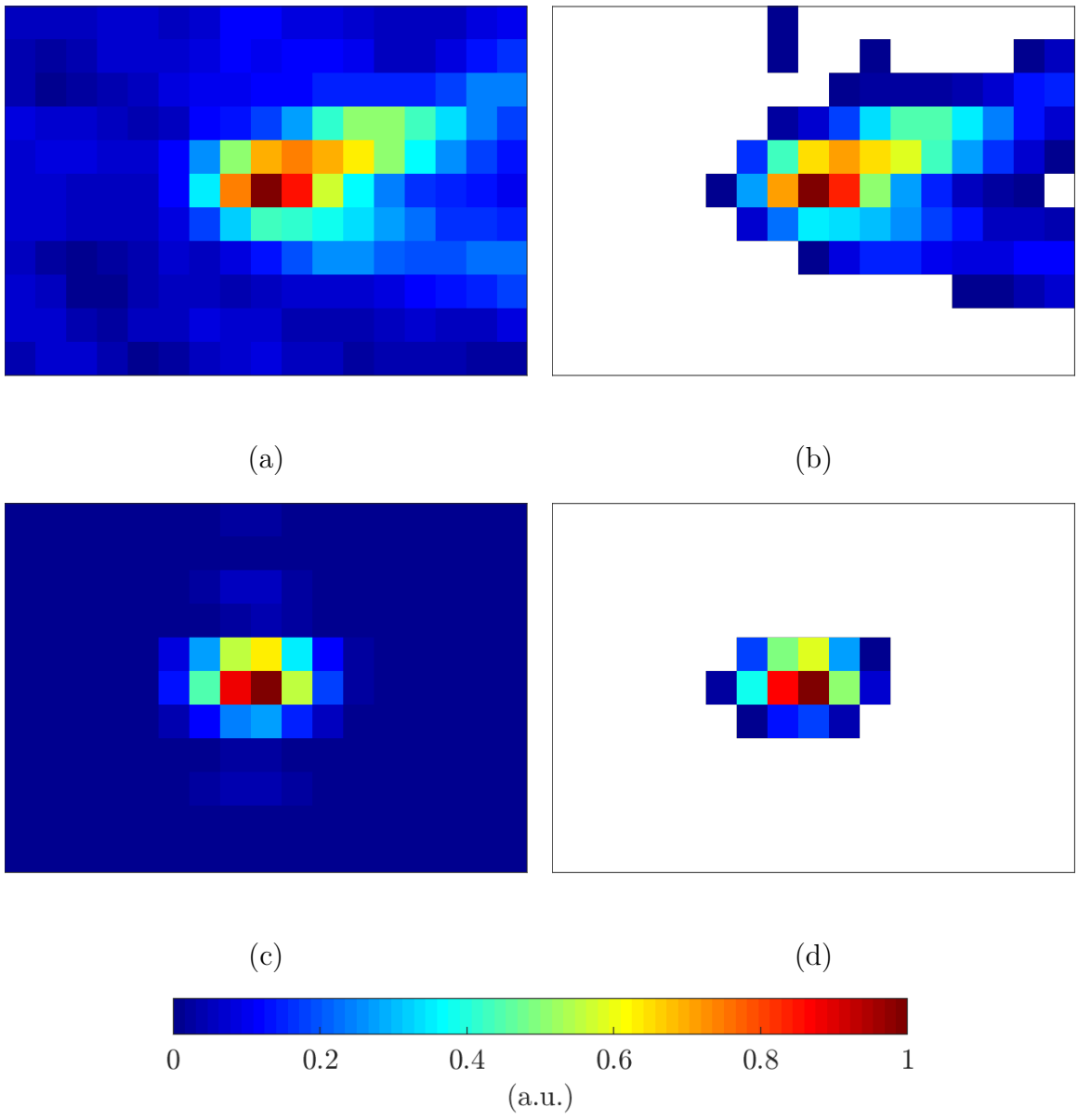


Figure 5.7: Illustration of noise pixel exclusion effects. Regions in white are excluded from the computation of C_{out} (a) Incoherent CYGNSS L1 DDM - no noise exclusion (b) Incoherent CYGNSS L1 DDM - with 30% noise exclusion threshold (c) Coherent CYGNSS L1 DDM - no noise exclusion (d) Coherent CYGNSS L1 DDM - 10% exclusion threshold

the space spanned by C_{in}, N_τ is the number of delay bins, N_f is the number of Doppler bins, and τ_M and f_M are the delay Doppler indices of the specular bin respectively. Here τ_M and f_M correspond to the location of the maximum value within a DDM. For coherent returns, the received power will be concentrated within C_{in} with little power in C_{out} so that PR is large. For non-coherent DDMs, the increased spreading of DDM power should result in a lower value for PR. DDMs are flagged as dominatly coherent when $PR \geq \rho_0$ with ρ_0 the detection threshold. The idea of power spread across the DDM had been previously explored in [130] but in a manner that focused on absolute power levels within the entire DDM exceeding a preset threshold as opposed to relating power extents within the C_{in} and C_{out} regions.

5.5 Reducing The Impacts of Thermal Noise

The sensitivity of the DPSD method can be improved by refining the computation of C_{out} to reduce the impact of DDM pixels suspected of containing only thermal noise contributions [131]. The mapping of received power from the spatial domain to each pixel within the DDM is dictated by the intersection on Earth's surface of constant delay ellipses with constant Doppler frequency hyperbolas. This process gives rise to the characteristic "horse-shoe" shape of an incoherent DDM. Some portions of the DDM fail to map onto Earth's surface; these DDM pixels contain primarily thermal noise contributions. Because such pixels convey no information on coherence, it is desirable to exclude them from the C_{out} summation. A simple method was therefore developed in which pixels having powers less than a predefined percentage of the DDM maximum are excluded from the C_{out} summation.

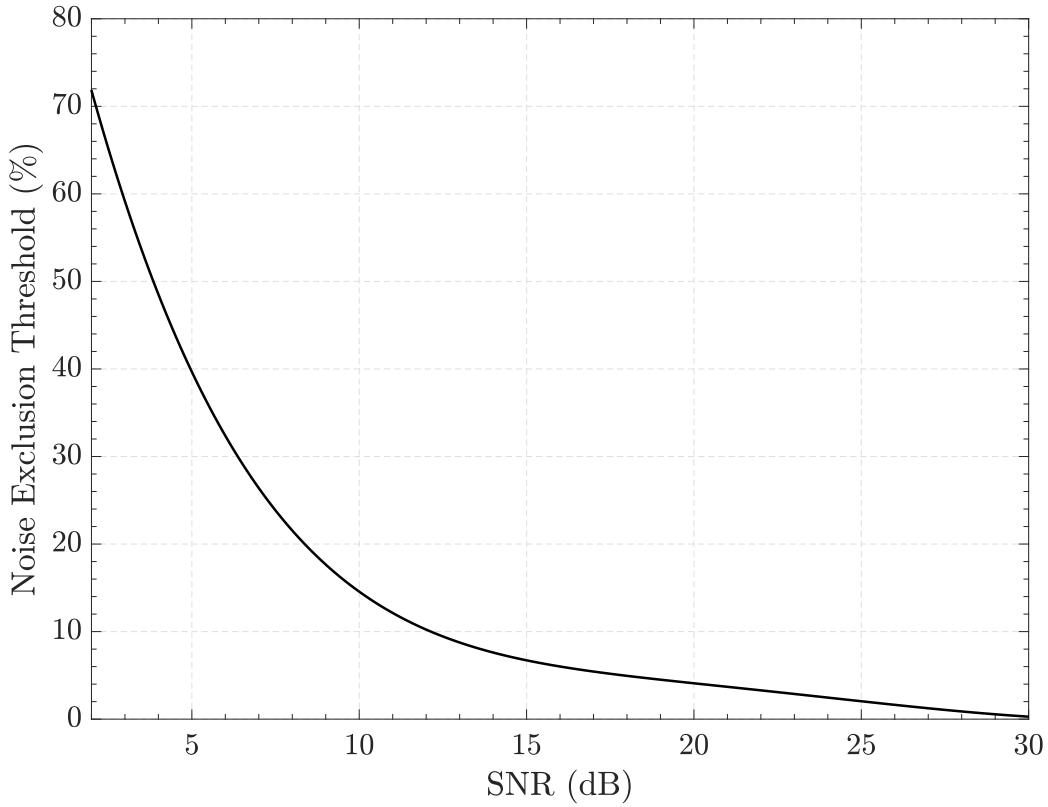


Figure 5.8: Noise exclusion threshold as a function of CYGNSS measurement Signal-to-Noise ratio

Example incoherent and coherent DDMs are illustrated in Figure 5.7 to clarify this process. Plots (a) and (b) demonstrate an example incoherent DDM in which all DDM pixels outside the specular region or only those at least 0.3 times the DDM maximum are used to compute C_{out}, respectively. The corresponding PR changes from 0.119 to 0.116 in these cases. Plots (c) and (d) similarly show an example coherent DDM before and after pixel exclusion, in this case retaining pixels having amplitudes at least 0.1 times the maximum. PR in this case is amplified from 1.437 to 3.543 following the exclusion process. The limited effect on PR of the exclusion for non-coherent DDMs and the significant effect for coherent DDMs motivates the

use of this method. Note that an exclusion threshold that is proportional to the DDM maximum value implies that thermal noise may still be retained in low SNR situations. For this reason, the noise exclusion threshold (the percentage of the DDM maximum that pixel powers must exceed to be included in C_{out}) was specified as a function of the DDM SNR. Tests varying this threshold were used to produce the empirically derived function shown in Figure 5.8, which is used in all results to follow.

5.6 Extension to CYGNSS Observations

Results from the DPSD method are compared in what follows to those based on DDM correlation with the CYGNSS Woodward Ambiguity Function (WAF) [26, 123]. The WAF is defined as the DDM response induced by a point scatterer. The correlation metric is defined here as

$$R = \frac{\left|\langle \mathrm{DDM}(\tau, f_D), \chi(\tau, f_D)\rangle\right|^2}{\langle \mathrm{DDM}(\tau, f_D), \mathrm{DDM}(\tau, f_D)\rangle \langle \chi(\tau, f_D), \chi(\tau, f_D)\rangle} \tag{5.8}$$

in which the bracket notation refers to a point-wise multiplication of DDM "pixels" followed by a summation of all pixel products.

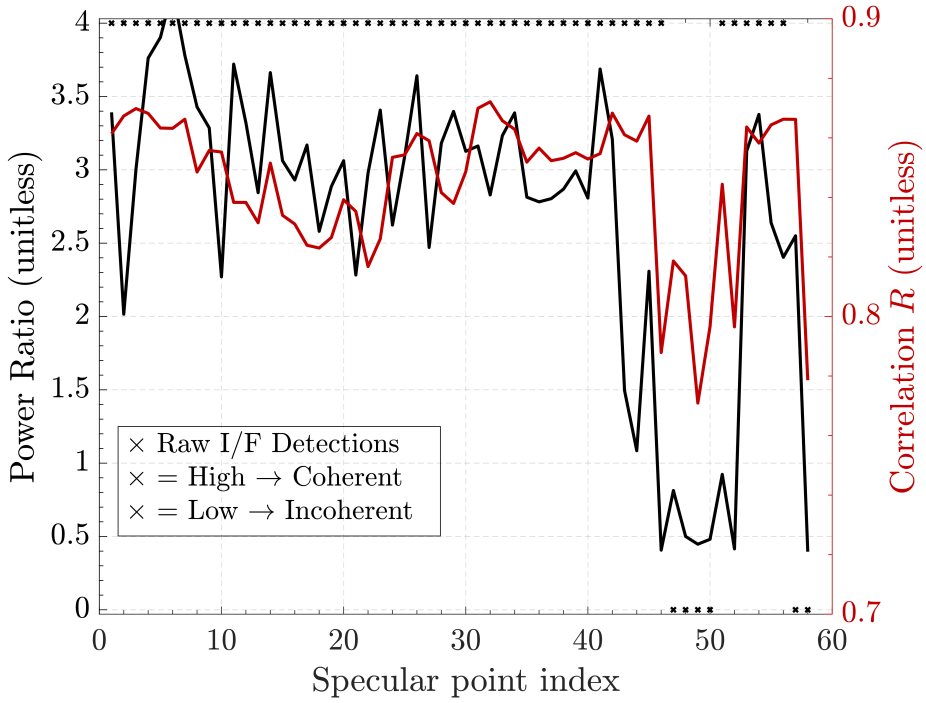
The correlation metric R also neglects amplitude information and instead focuses on the overall similarity of the DDM to the WAF. Coherent DDMs are expected to show a high correlation with the WAF that should reduce for incoherent DDMs, and a threshold can be placed on R above which coherency is declared.

5.6.1 Raw Intermediate Frequency Tracks

CYGNSS satellites are capable of reporting the observed intermediate frequency raw data stream (in uncalibrated counts) over short periods of time. This "Raw I/F"

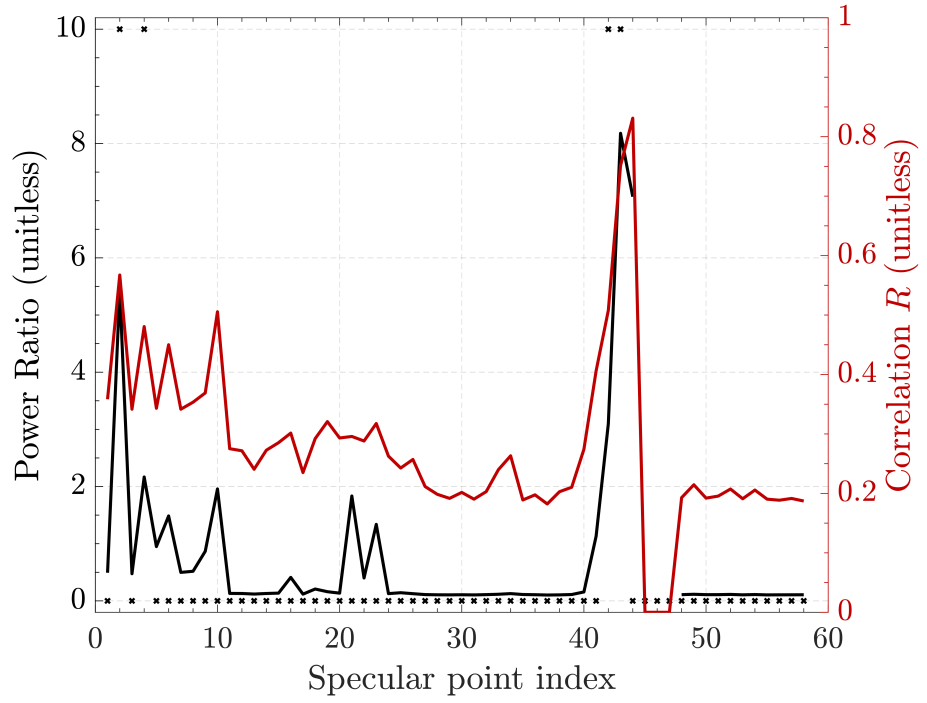
mode data can be post processed into DDMs having arbitrary integration times and delay extents, and can provide DDM phase information that is very sensitive to the presence of coherence.

Reference [132] reports an algorithm for detecting coherence in Raw I/F mode DDMs based on the growth in coherent DDM power as the integration time is extended. Detections obtained from this "gold standard" method were compared with those obtained from the DPSD approach for three reference tracks for which Raw I/F mode data were available. Figure 5.9 illustrates the comparisons.

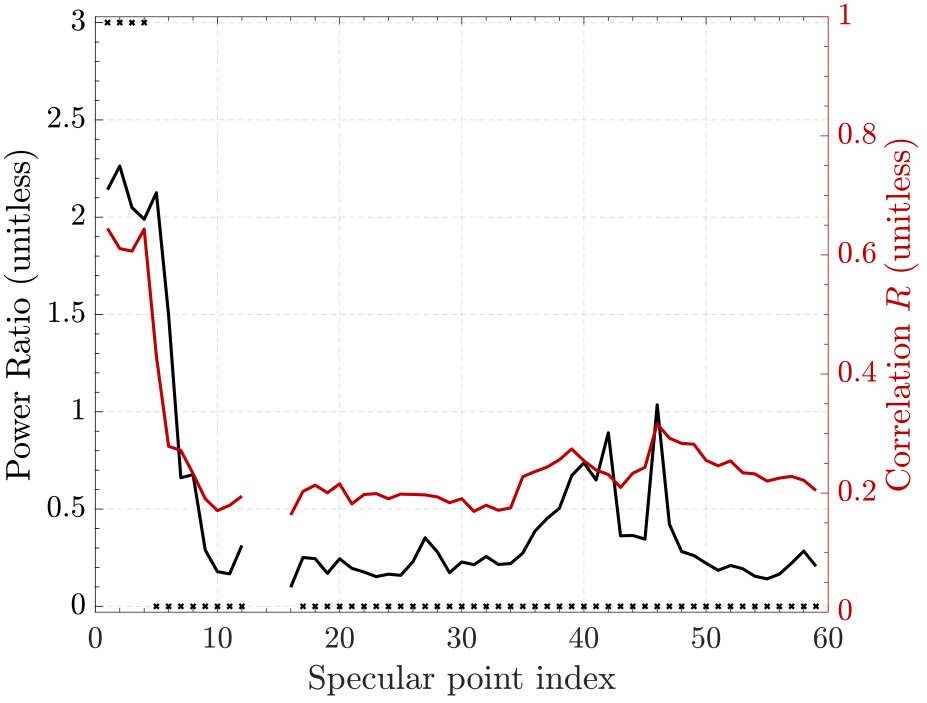


(a)

Figure 5.9: Comparison of Level-1 DPSD power ratio and MFD correlation against Raw I/F coherence/non-coherent detection (a) Track 1, dominantly coherent (b) Track 2, dominantly non-coherent (c) Track 3, dominantly non-coherent



(b)



(c)

Figure 5.9: Comparison of Level-1 DPSD power ratio and MFD correlation against Raw I/F coherence/non-coherent detection (a) Track 1, dominantly coherent (b) Track 2, dominantly non-coherent (c) Track 3, dominantly non-coherent

For track 1 (plot a) the gold standard method detected coherence in 53 out of 59 1 second periods (detections are indicated by the $\times$ symbols), with brief intervals of non-coherency occurring in later portions of the 1 minute dataset. For track 2 (plot b) the gold standard method detects coherence in only 4 out of 59 intervals. For the third track a brief interval of coherence is detected comprising 4 consecutive specular points at the beginning of the track followed by 55 non-coherent DDMs. The PR quantity in plot (a) is found to fall below a value of 2 only for those points flagged as incoherent (with one exception), while the PR quantity is small for the incoherent cases in plot (b). Both the DPSD and MFD approaches appear to be able to capture the overall coherence/non-coherence tendencies compared to the Raw I/F detections. However, for track 1 the average PR and R values over coherent intervals are 3.12 and 0.85, respectively, while the non-coherence interval averages are 0.77 (PR) and 0.82 (R), respectively. A similar response is noted for track 3, in which coherent detections were associated with a 2.11 average power ratio, falling below a value of 2 when transitioning to a dominantly non-coherent mode of scattering where the average PR for the non-coherence interval was 0.38. The higher sensitivity and increased separation in PR between dominantly coherent and dominantly incoherent DDMs motivate its selection for use in coherence detection. The results of these tests suggest that ρ_0 should be set to approximately 2 to provide a good match to the three Raw I/F mode tracks investigated.

The ability to generate DDMs, and corresponding coherence estimates, at arbitrary integration times using Raw I/F streams have motivated further investigation into some of the underlying causes for the discrepancies observed. An example of

this is shown in Figure 5.10 for a Level-1 DDM that the power spread detector declared coherent, but was reported by the 'gold standard' as being incoherent. Here it is noted that the Level-1 DDMs, before July 2019, were a 1 second product arising from a 1 ms coherent integration time, and 1000 incoherent averages. The means with which the Raw I/F based detector produces coherence estimates on the same temporal scales is by comparing the mean power ratios of 10 ms coherently integrated DDMs by 10, 1 ms DDMs, averaged 10 times for a total of 10 intervals (i.e. $10\times10\times10$ ms = 1 s). The estimates corresponding to each of the 10 intervals are those that are shown in Figure 5.10. In it, a very dominant coherent component exists at interval 5 but for the remaining 9 coherence estimates are noticeably lower suggesting that incoherence is dominant. Because the Level-1 DDM does not allow for accounting for these nuances (due to its fixed integration time) the contributions of the coherent reflector in this case 'overwhelms' all other incoherent scattering effects and the final DDM is declared coherent by the Level-1 coherence detection methodology.

Similarly, the behaviour of the proposed detector may be further analyzed by varying detection thresholds and assessing detection probabilities, and the probability of false alarm giving rise to the receiver operating characteristic (ROC) curve depicted in Figure 5.11. This was undertaken using a much large set of Raw I/F downlinks comprising ≈8000, 50 ms DDMs, with the results suggesting that with the appropriate selection of detection threshold, probabilities of detection as high as $P_d = 90\%$ and simultaneously a false alarm probabilities as low as $P_{fa} = 5\%$ are achievable.

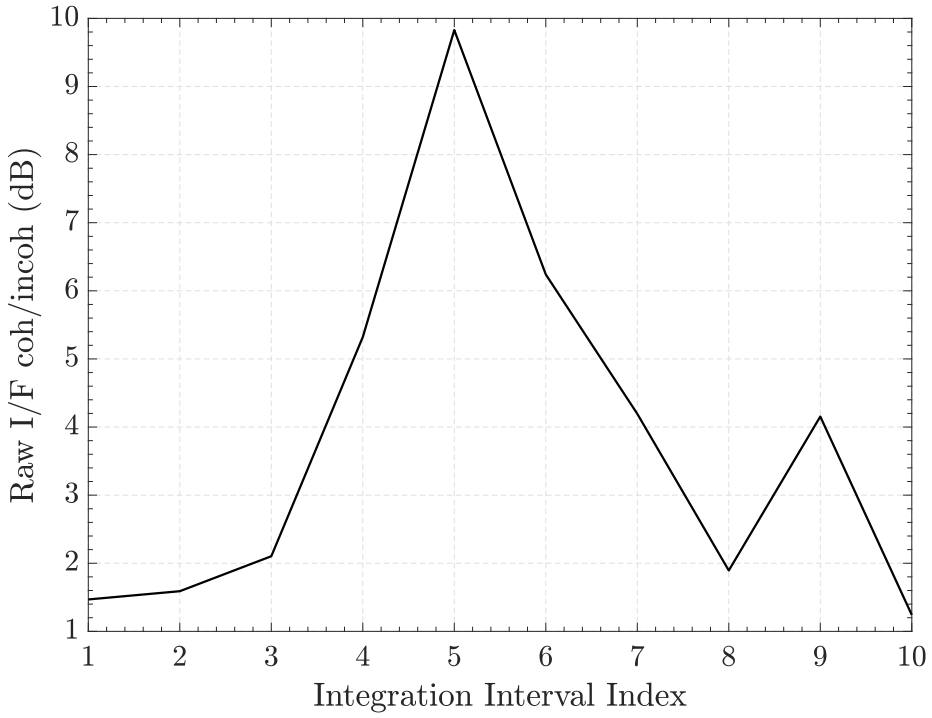


Figure 5.10: Coherence estimates at temporal resolutions finer than standard Level-1 DDM integration times

5.6.2 Land Sandbox Tests

The Raw I/F mode comparisons motivates use of the PR metric to detect coherence. However, the limited Raw I/F mode dataset available does not allow complete determination of the detection threshold ρ_0. DPSD algorithm detections were therefore examined as ρ_0 was varied for the 12 day CYGNSS "sandbox" dataset. The resulting percentages of inland DDMs detected as coherent are shown in Figure 5.12 as a function of ρ_0. The resulting percentages range from $\approx$ 6.6%-11.4% for ρ_0 ranging from 2.4 to a much more lenient threshold of 1.8. Results from the Raw I/F comparison and manual examinations of numerous DDMs suggest setting the detection threshold ρ_0 as 2.0, with the result that $\approx$ 8.9% of all inland points, a total of 1.5

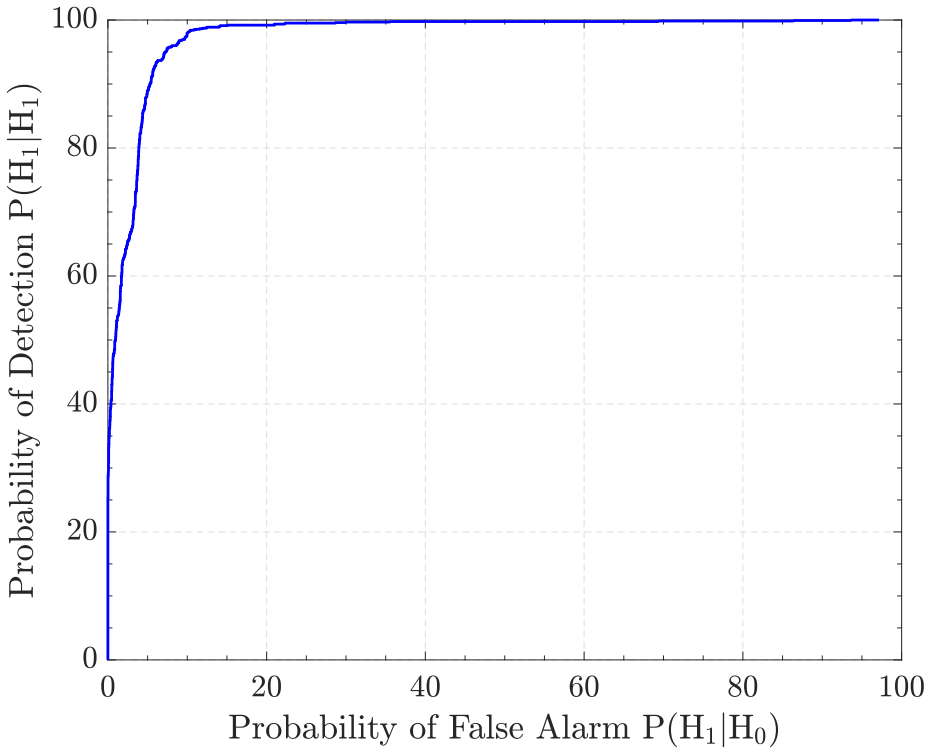


Figure 5.11: Power spread detector receiver operating characteristic curve relative to Raw I/F coherence detection approach estimates

million, are declared coherent. Extension of the detection methodology to the larger Level-1 dataset, comprising >800 million land specular points, resulted in a comparable prevalence of coherence. It is recognized that there remains some uncertainty in the threshold value used, and that a slightly larger or smaller percentage of inland returns would be flagged as coherent if ρ_0 were decreased or increased, respectively. The optimal value for the detection threshold is nonetheless expected to be limited to the range $1.95 \leq \rho_0 \leq 2.05$. Extensive testing showed that the selected 2.0 value appears to provide a good tradeoff between detection sensitivity and false detections.

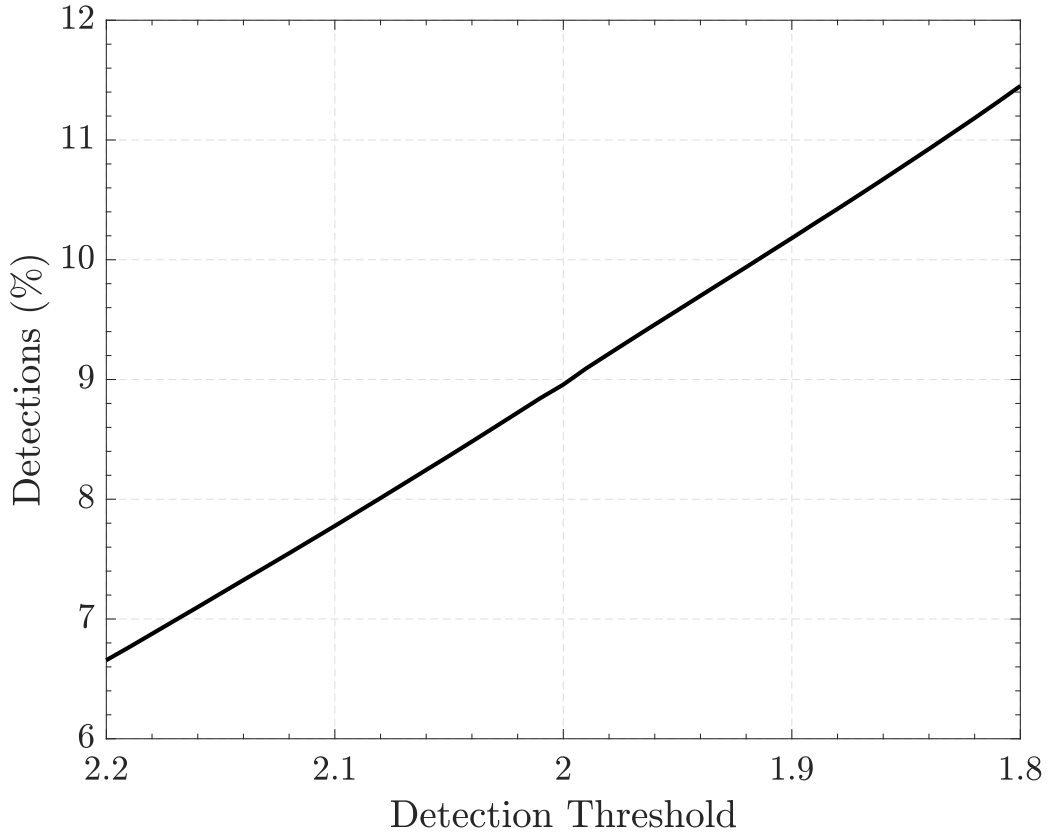
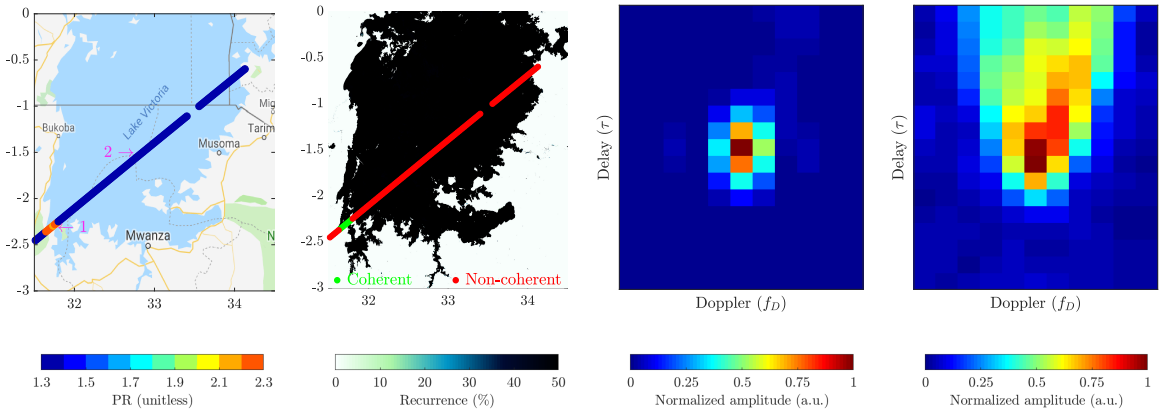


Figure 5.12: Percentage of CYGNSS land-surface DDMs detected as coherent for varying ρ_0

5.6.3 Case Studies

Example detection results are shown for tracks through Lake Victoria and the Everglades region of Florida in Figures 5.13 and 5.14, respectively. Most specular points within the track over Lake Victoria are classified as non-coherent. The small portion flagged as coherent apparently corresponds to calmer waters close to shore; an example DDM from position 1 is provided in plot (c). As the track moves across the lake, typical horseshoe shaped DDMs indicative of incoherent scatter are observed; this should be expected as the large size of Lake Victoria allows the development

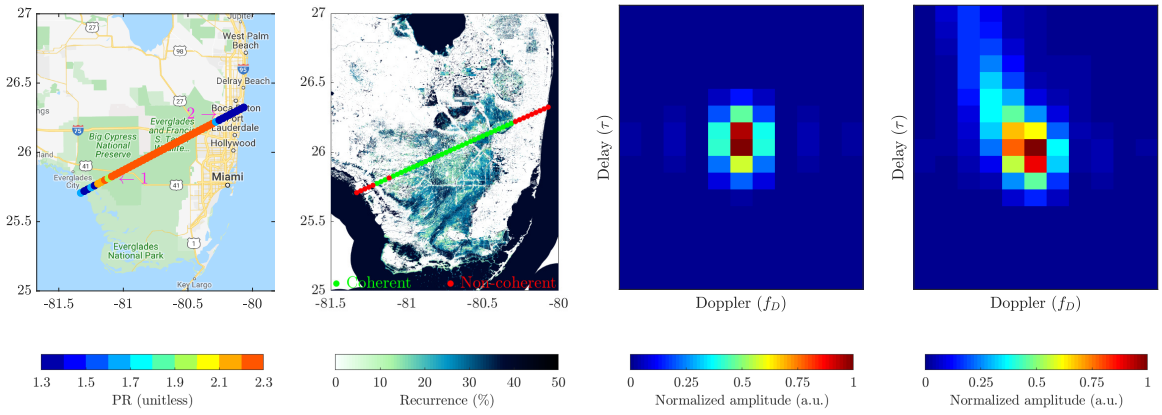


(a) (b) (c) (d)

Figure 5.13: Case study for CYGNSS L1 coherence detection at Lake Victoria (a) Power ratios across CYGNSS specular point tracks (b) Yearly water recurrence percentage (c) DDM detected as coherent at position 1. DDM SNR is 19.07 dB and its PR is 2.83 (d) DDM detected as non-coherent at position 2. DDM SNR is 5.48 dB and its PR is 0.28

of significant wave heights that result in incoherent scattering. An example of the resulting effects are shown in the DDM from position two in plot(d)

Conversely, the majority of DDMs for the Everglades track were declared coherent, suggesting that scattering surfaces sufficiently flat to allow coherent reflections are widespread in this region. The annual water recurrence map [133] depicted in Figure 5.14(b) shows high values in this region, indicating the presence of water bodies that are static or present for most of the year. The coherent returns detected therefore are associated with the presence of water bodies. The average PR value for all coherent detections across this track was ≈ 3.48. A limited number of DDMs were classified as non-coherent, all of which correspond to areas with little to no water recurrence. The DDM at position 2 (Figure 5.14(d)) has $PR = 0.99$ and therefore

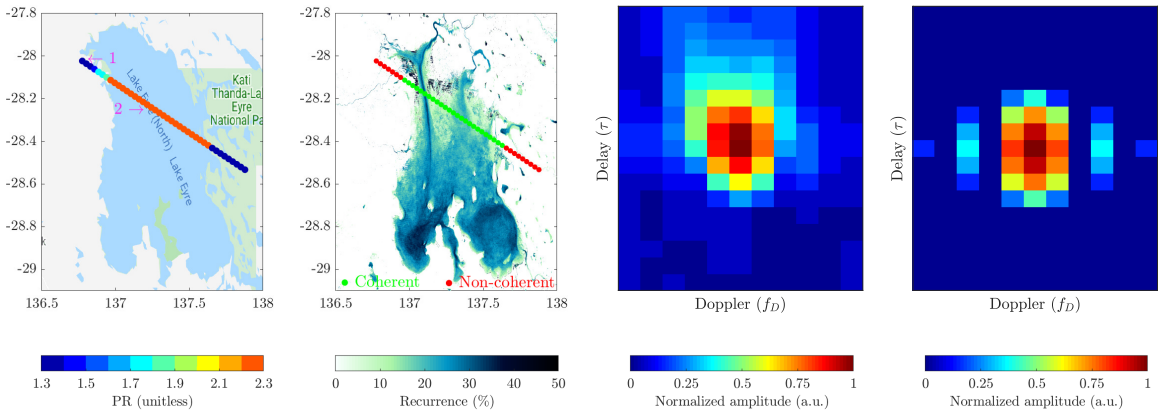


(a) (b) (c) (d)

Figure 5.14: Case study for CYGNSS L1 coherence detection at the Florida Everglades (a) Power ratios across CYGNSS specular point tracks (b) Yearly water recurrence percentage (c) DDM detected as coherent at position 1. DDM SNR is 23.18 dB and its PR is 2.65 (d) DDM detected as non-coherent at position 2. DDM SNR is 17.73 dB and its PR is 0.99

is declared non-coherent, but is likely representative of a mixed return with both coherent contributions from water bodies and incoherent scattering from the land surface. Such returns have been found to be associated with high SNR (>15 dB) and low PR (<2.0). The algorithm developed classifies such returns as non-coherent given the "purely coherent" definition used for the DPSD algorithm.

Other example detection results for the Australian Lakes Eyre and Frome are shown in Figures. 5.15-5.16.

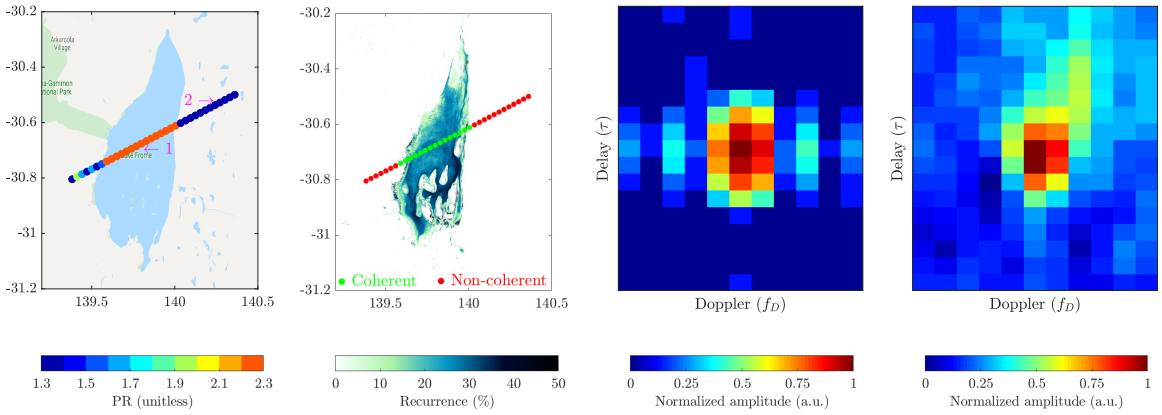


(a) (b) (c) (d)

Figure 5.15: Case study for CYGNSS L1 coherence detection over Lake Eyre (a) Power ratios across CYGNSS specular point tracks (b) Yearly water recurrence percentage (c) DDM detected as non-coherent at position 1. DDM SNR is 13.30 dB and its PR is 1.35 (d) DDM detected as coherent at position 2. DDM SNR is 26.25 dB and its PR is 2.76

From panel (b) it is readily identifiable that CYGNSS's measurement switch to dominantly coherent when the respective tracks traverse the lakes' surfaces, switching back to a dominantly non-coherent mode of scattering over the shores. Two factors are expected to contribute to coherence in the case of the Australian lakes. The first is the prevalence of water (with low recurrence) within the lake. While both lakes rarely completely fill with water, they are never completely empty and the presence of inland water contained within the lakes is expected to present the roughness levels required for dominant coherent reflection to be present.

The second contributing factor for coherence over Lakes Eyre and Frome is the presence of large shallow salt pans which are also expected to be flat enough to be conducive to persistent coherence. Globally, a limited set of locations were identified



(a) (b) (c) (d)

Figure 5.16: Case study for CYGNSS L1 coherence detection over Lake Frome (a) Power ratios across CYGNSS specular point tracks (b) Yearly water recurrence percentage (c) DDM detected as coherent at position 1. DDM SNR is 19.83 dB and its PR is 3.05 (d) DDM detected as non-coherent at position 2. DDM SNR is 4.04 dB and its PR is 0.48

to similarly pose persistent (highly recurring) coherence in spite of the absence of inland water bodies. The clear sensitivity to the presence of these surfaces, evidenced by the precise switches in coherence states at the shoreline, is of particular interest in exploring means with which the Level-1 coherence detector may be used for the purposes of inland water body mapping. This is explored further in Chapter 6.

5.7 Persistent Land Coherence

Specular land surface scattering depends on surface small scale roughness, large scale topography, land cover, surface permittivity, and any presence of water bodies. CYGNSS DDMs from some locations will exhibit a higher likelihood of being classified as coherent due to these factors, particularly the presence of water bodies.

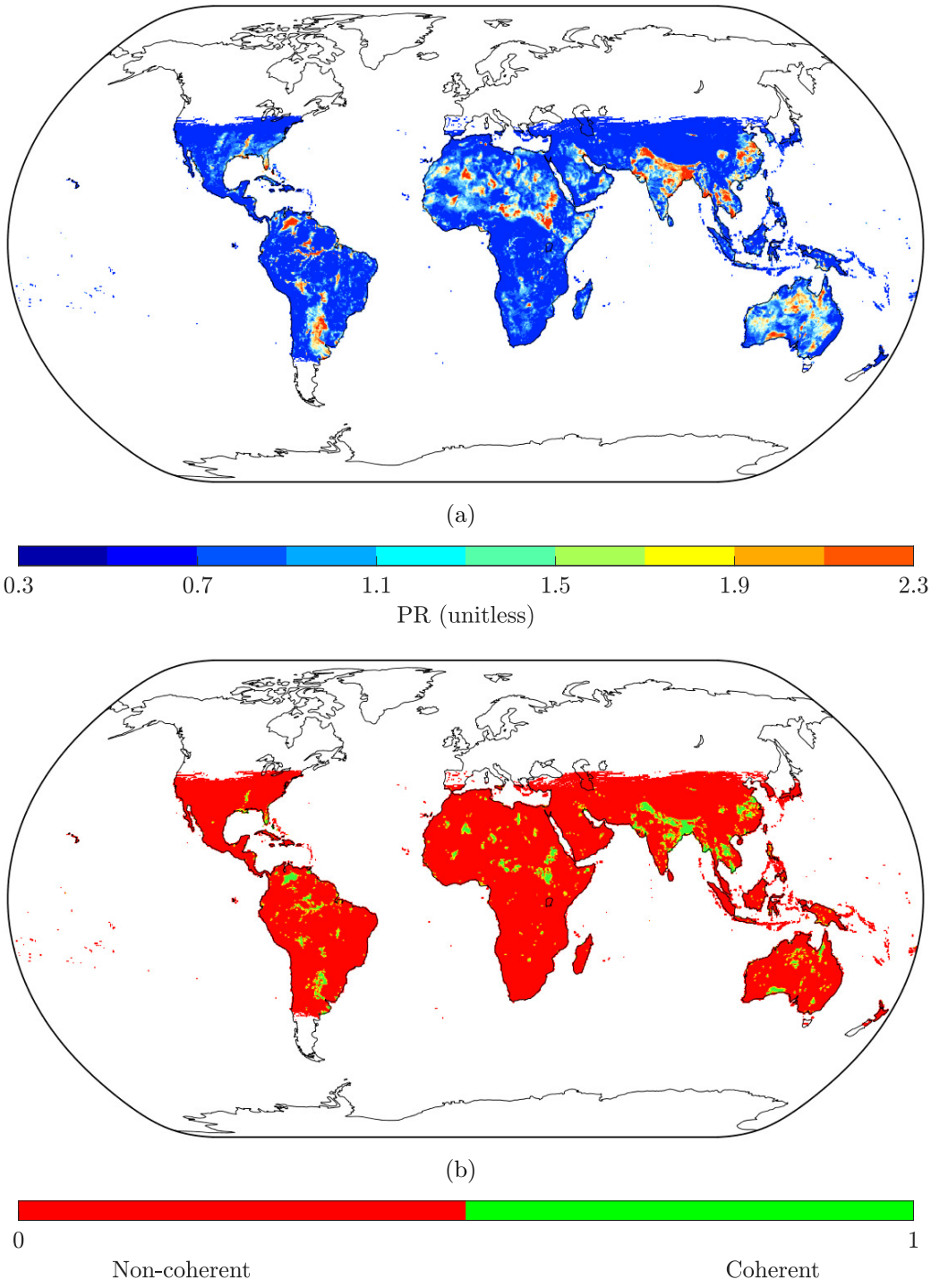


Figure 5.17: Illustration highlighting persistent coherence over land points (a) Median power ratio over sandbox run 12 day period (b) Median detection using $\rho_0 = 2.0$ over sandbox run 12 day period

To investigate these factors, Figure 5.17 presents the median PR value over the 12 day sandbox dataset following the projection of measured DDMs onto a 36 km Equal-Area Scalable Earth (EASE) grid [134, 135]. The resulting detector output using the median PR value is also shown. The results identify locations that exhibit a higher likelihood of coherence, including the Florida Everglades, Louisiana Wetlands, and Mississippi River wetlands in the USA, the Orinoco Wetlands in Venezuela, the Iberá Wetlands in Argentina, the vicinity of the Congo River in the DRC, the Rann of Kutch lake and salt marsh in India/Pakistan. The region bordering the Himalayas extending from Bangladesh through Nepal similarly contains extensive networks of water bodies including the Padma River, Brahmaputra River and irrigation networks. All of these regions have a significant prevalence of water bodies that provide reflecting surfaces sufficiently flat to produce coherent reflections. A further analysis showed that, globally, approximately 81% of DDMs flagged as coherent included a non-zero water fraction within the first Fresnel zone.

Other frequently coherent locations include parts of the Sahelian Zone in Mauritania, the Saharan Zone in Mauritania, the Algerian Desert in Algeria, the Libyan Desert in Libya and the semi-arid savanna in Africa. Water bodies are not widespread in these regions, indicating that the terrain surface must be sufficiently flat to cause coherent reflections. Surface RMS heights (including topography within the first Fresnel zone) need to be on the order of approximately 5 cm or less for coherence to be dominant at L-band [90,136] as shown in Section 5.2. An analysis of the SRTM digital elevation model showed that these locations are associated with global RMS height minima within a first Fresnel zone sized region. However an exact quantification of surface roughness would require measurements made with cm (or less) accuracy

that, currently, are not available for the locations of interest. The results obtained nevertheless suggest these areas are likely to be sufficiently flat for coherence to be dominant. It is noted that other similar desert regions are not identified as producing highly coherent returns; analyses showed terrain rms heights for such locations to be larger than those at locations regularly producing coherency.

5.7.1 Correspondence of DDM Power Ratio to Signal-to-Noise Ratio

The DDM SNR or the SNR corrected for antenna gain and range effects has been previously considered in land studies [28, 30, 31], with values that may be indicative of the nature of the observed return. Figure 5.18 provides a scatter plot of PR versus the DDM (uncorrected) SNR for the sandbox dataset. For DDMs having $PR > 2.0$, the SNR is observed to range between 2.0 and approximately 40 dB, so that the DPSD provides information beyond that contained in the SNR alone. Cases having SNR less than 1.5 dB were largely too noisy to allow clear classification; the high exclusion thresholds used at low SNR (Figure 5.8) prevent detections from occurring in this case.

Approximately 2% of the measurements have $PR < 2.0$ but SNR values greater than 15 dB. Detailed examinations of such cases showed the DDMs typically to contain both coherent and incoherent contributions, and therefore not to fall into the "purely coherent" category. Future efforts could potentially consider a combination of the SNR and PR metrics to identify these "mixed" cases.

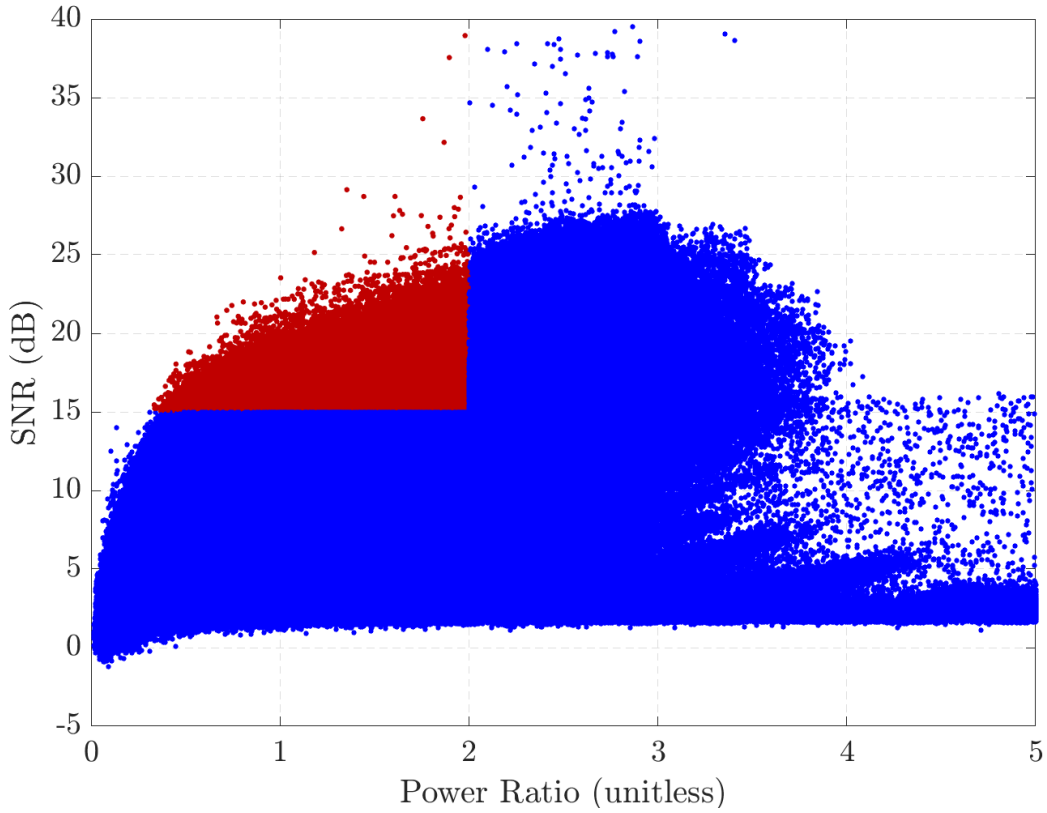


Figure 5.18: Correspondence of SNR to DPSD Power Ratio for 16.5 million land specular points. Scatters in red are high SNR low PR subset

5.7.2 Assessment of Probability of False Alarm and Ocean Coherence

Evaluating the DPSD detector's probability of false alarm P_{fa} is difficult given the absence of truth information on the presence of coherence in the sandbox dataset. However, the expectation that ocean surface DDMs should largely be incoherent provides one opportunity for estimating P_{fa}. To test this, the DPSD was applied across a subset of ocean returns over the 12 day sandbox run period. The resulting P_{fa} was found to range between $1.0 \times 10^{-5} \leq P_{fa} \leq 1.5 \times 10^{-5}$ for a threshold range

of $1.8 \leq \rho_0 \leq 2.2$. The variability of detections versus detection threshold over the ocean is depicted in Figure 5.19.

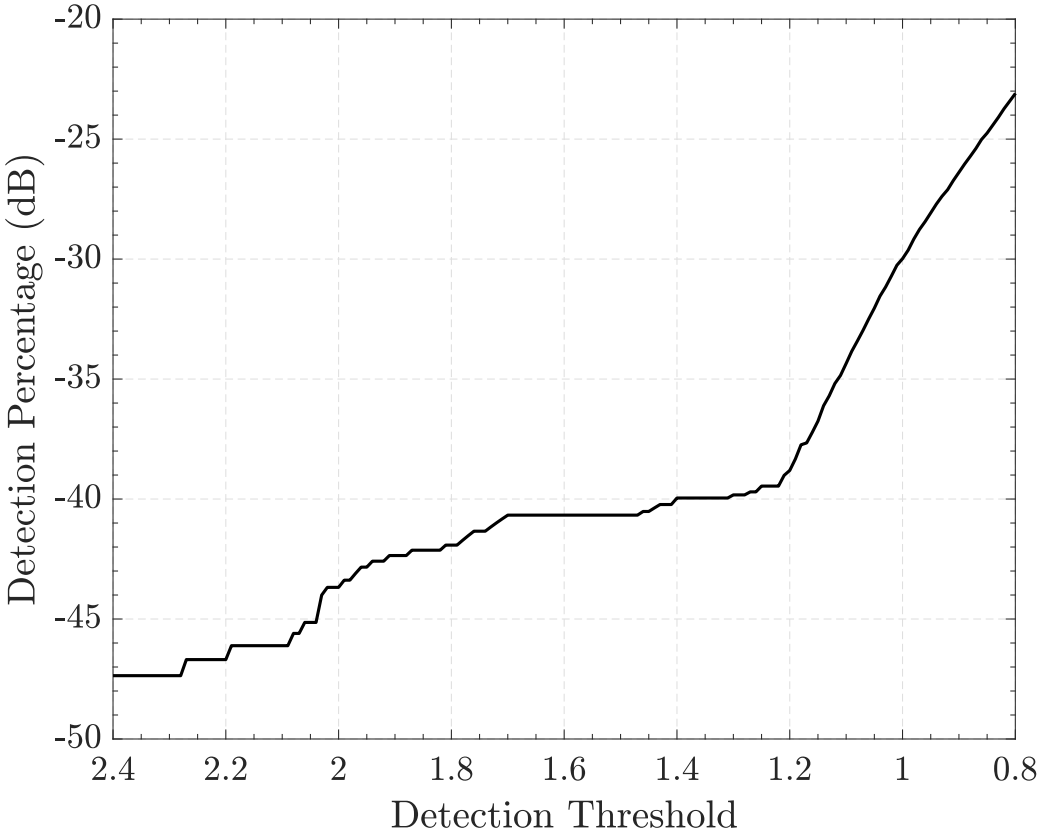


Figure 5.19: Percentage of all ocean points detected as being coherent at varying thresholds

The subset used excluded areas with harbored waters. For the selected $\rho_0 = 2$ threshold, P_{fa} for the ocean subset is approximately 1.2×10^{-5}, and the typical range of PR is from 0.15 to 0.45. It is noted that this value for P_{fa} is small, so that the DPSD detector is unlikely to cause false alarms for "purely incoherent" scenes as typically occur for sea observations.

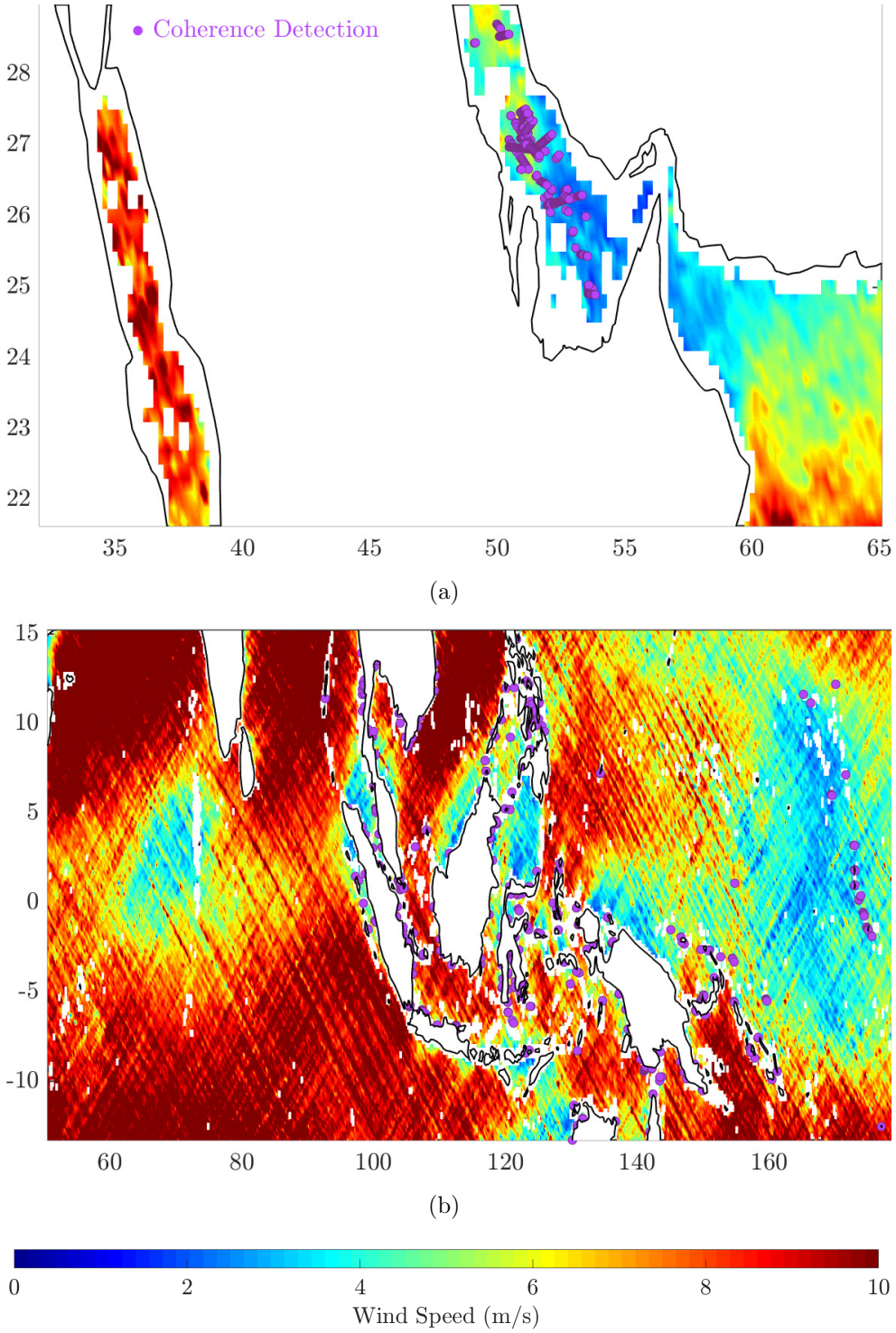


Figure 5.20: Coherence detection over the ocean during 12 day sandbox run overlaid on mean CYGNSS Level-3 winds (a) Coherence detection over The Gulf region (b) Coherence detection over Maritime Continent

Although ocean measurements are expected to be incoherent, coherency can occur for sea returns when wind speeds are less than 5 m/s [90] in regions that are sheltered from non-local swell waves. Coherence detections in ocean observations primarily occurred in a small set of sheltered locations; these locations were excluded from the ocean data subset used in assessing P_{fa}. The utility of the proposed detector can therefore also be extended to ocean surfaces. Examples of coherent detections in The Gulf region and the Maritime Continent are shown in Figure 5.20 along with the corresponding mean wind speed obtained from the CYGNSS L3 wind retrievals during the 12 day dataset. The results confirm that low wind speeds in sheltered areas can increase the likelihood of coherence for ocean observations. The Gulf and Maritime Continent regions account for 60% of the ocean coherence observed over the 12 day test dataset. Detections typically occurred for sea areas enclosed between two or more land masses. Other locations where coherence was observed include the Alboran Sea, Gulf of Mexico, Caribbean Sea and the Gulf of California. The locations over which coherence was observed over the ocean, were therefore highly complementary to the regions with higher probabilities of calm conditions, depicted in [137].

The ability to detect coherence over the ocean is expected to be of particular interest for future missions with the objective of conducting altimetry, the mapping of ocean surface topography by providing estimates of sea surface heights. This includes future receivers like the European Space Agency's Passive Reflectometry and Dosimetry (Pretty) Mission. Similar to traditional radar altimeters, GNSS-R measurements can be used to estimate sea surface heights through the computation of precise time delays across a given forward propagation and comparing this to

some reference (geoid) level. The extraction of carrier phase information necessary to do this however, requires, the calm surface conditions typically associated with the dominance of coherence. As a result, the Level-1 detection methodology may prove instrumental in the identification of ocean coherence, necessary as a preliminary step for phase altimetry.

5.8 Conclusions

An algorithm for detecting coherence in Cyclone Global Navigation Satellite System mission Delay Doppler Maps (DDMs) was presented. Because CYGNSS Level-1 DDMs report only the observed power without phase information, the algorithm uses estimates of power "spread" within the DDM to flag coherency. Since the estimate used is a ratio of the powers in differing portions of the DDM, it is less sensitive to absolute power calibration and to the GPS C/A code type observed, and is applied to CYGNSS Level-1 uncalibrated DDMs. The basic detector formulation was described along with modifications to improve performance in lower signal-to-noise ratio (SNR) situations. The required detection thresholds are determined using matchups with CYGNSS "Raw I/F" mode measurements for which the DDM phase can be computed and used to identify coherence more precisely. Application of the final detector over a large CYGNSS dataset suggests that approximately 8.9% of all inland returns are coherent. Inland regions persistently identified as coherent were found largely to be associated with the presence of water bodies. A smaller set of desert locations apparently having very low surface roughness were also found to be associated with persistent coherence. The detector was also applied to a set of ocean measurements,

with the results showing that persistent coherence is limited to areas with sheltered waters.

Chapter 6: On The Applications of Coherence Detection Over Land

6.1 Motivation

The properties of measurements dominated by coherent reflection are fundamentally dissimilar to those dominated by incoherent scatter, both in terms of the underlying physical scattering mechanism as well as the properties exhibited. Interest in separating and using coherent reflections for a wide range of applications has been expressed in a number of scholarly investigations, but the ability to make progress has typically been limited by the inability to reliably detect coherence on the global scale, as previous undertakings concerned with the analysis of coherent reflections have often relied upon the less frequently available Raw I/F streams. The preceding Chapter has outlined the unprecedented capability of coherence detection on a global scale, and therefore this Chapter aims to outline the range of needs for which the Level-1 power spread coherence detector may be used to address.

6.2 Inland Water Body Mapping

Mapping inland water bodies and their dynamics is vital for a number of applications including water resource management [138], assessment of climate change

along with the development of mitigation techniques [139,140], agricultural productivity [141] and the modelling of land-atmosphere exchange [142]. Accurate knowledge of inland water and its variation is also key to the operational aspects of spaceborne remote sensing missions such as the Soil Moisture Active Passive (SMAP) and Soil Moisture and Ocean Salinity (SMOS) missions [143, 144], the Advanced Very High Resolution Radiometer (AVHRR) [145], and several others. This is due to the variability standing water causes in satellite observables and the significant errors water bodies can induce to their respective geophysical retrievals. Due to this importance a number of previous studies have attempted to explore the means through which spaceborne optical, hyperspectral, infrared and microwave systems may be used for the purposes of inland water body mapping. While several of these systems, such as Landsat, Sentinel and Terra ASTER [146–148], are capable of providing inland water body maps with resolutions on the order of 50 meters or less, their temporal resolution is limited to revisits on the order of two weeks or coarser. Other instruments including the Moderate Resolution Imaging Spectroradiometer (MODIS), Advanced Along Track Scanning Radiometer (AATSR) and Visible Infrared Imaging Radiometer Suite (VIIRS) [149–151] are capable of providing inland water body estimates at spatial resolutions 0.25 to 1 km at significantly shorter time intervals. However, their ability to do so is limited by cloud cover and thick vegetation which can obscure the Earth's surface.

The sensitivity of spaceborne Global Navigation Satellite System Reflectometry (GNSS-R) to various land surface properties offers the potential for dynamic inland water body mapping at moderate spatial resolutions. The extensive land data set made available by the CYGNSS mission enables investigation of this application.

The "coherent" nature of GNSS-R returns from inland water bodies suggests that resolutions on the order of 200 meters are achievable [90, 153, 154]. The frequent revisits provided by CYGNSS within its latitude coverage region [76] indicates the potential ability to produce water body maps updated on bi-weekly to annual time scales.

This Section describes a method for detecting inland water bodies through analysis of the coherence of CYGNSS's Level-1 delay-Doppler map (DDM) products. The approach is based on the expectation that inland water surfaces are sufficiently flat at CYGNSS's L-band wavelength to cause coherent reflection to dominate the observed specular scattering, and uses a coherence detector previously reported. The method is shown to be highly effective in its mapping of inland water bodies, whilst avoiding a number of the uncertainties associated with the use of the CYGNSS signal-to-noise ratio (SNR) as in [36, 155, 177] because the DDM coherence detector used is less sensitive to calibration uncertainties and to the impact of vegetation cover and low levels of surface roughness.

6.2.1 Properties of CYGNSS Returns Over Inland Water

To date, CYGNSS has provided approximately 1 billion land surface measurements; numerous previous studies have explored the sensitivity of CYGNSS measurements to various land surface properties. CYGNSS land surface DDMs show more complex behaviors than the "incoherent horseshoe" shape typical over the ocean, due to the heterogeneity of land surfaces and a range of factors including small scale surface roughness, local topography, vegetation cover, soil moisture, and, for inland water bodies, water depth and surface winds. Land surface DDMs show behaviors

indicating scattering that ranges from purely coherent reflection to purely incoherent scattering to mixtures of both, as well as highly variable signal-to-noise ratios related to these factors.

A key factor is the change in the CYGNSS spatial resolution associated with coherent reflection, in which returns can be modeled to zeroth order as arising from the first Fresnel zone, with dimensions given by (6.1)-(6.5), surrounding the specular point [90,91,153], with some variability observed in accordance with the precise nature of the observed scene relative to the CYGNSS/GPS observation geometry [128]. The relevant length scales are:

$$F_1 = \sqrt{\frac{\lambda\,(r_R r_T)}{(r_R + r_T)}} \tag{6.1}$$

$$D_x = \sqrt{1 + 2\frac{F_1}{a_{eff}}\frac{F_1}{\lambda\cos\theta_i}} \tag{6.2}$$

$$D_y = \sqrt{1 + 2\frac{F_1}{a_{eff}}\frac{F_1\cos\theta_i}{\lambda}} \tag{6.3}$$

$$F_{1x} = \frac{F_1}{D_x\cos\theta_i} \tag{6.4}$$

$$F_{1y} = \frac{F_1}{D_y} \tag{6.5}$$

where F_{1x} and F_{1y} are the first Fresnel zone semi-major and semi-minor axes respectively, a_{eff} is Earth's radius of curvature at the specular point, with D_x and D_y representing "divergence factors" caused by Earth's curvature. The electromagnetic wavelength is given by λ, r_R is the range from the receiver to the specular point and r_T is the range from the transmitter to the specular point. In contrast, under a dominantly incoherent scattering regime the expected measurement's range resolution is described by contours of constant range defining concentric ellipses with the

CYGNSS/GPS pair being at the two foci for this bistatic geometry, with dimensions described by (3.1)-(3.2).

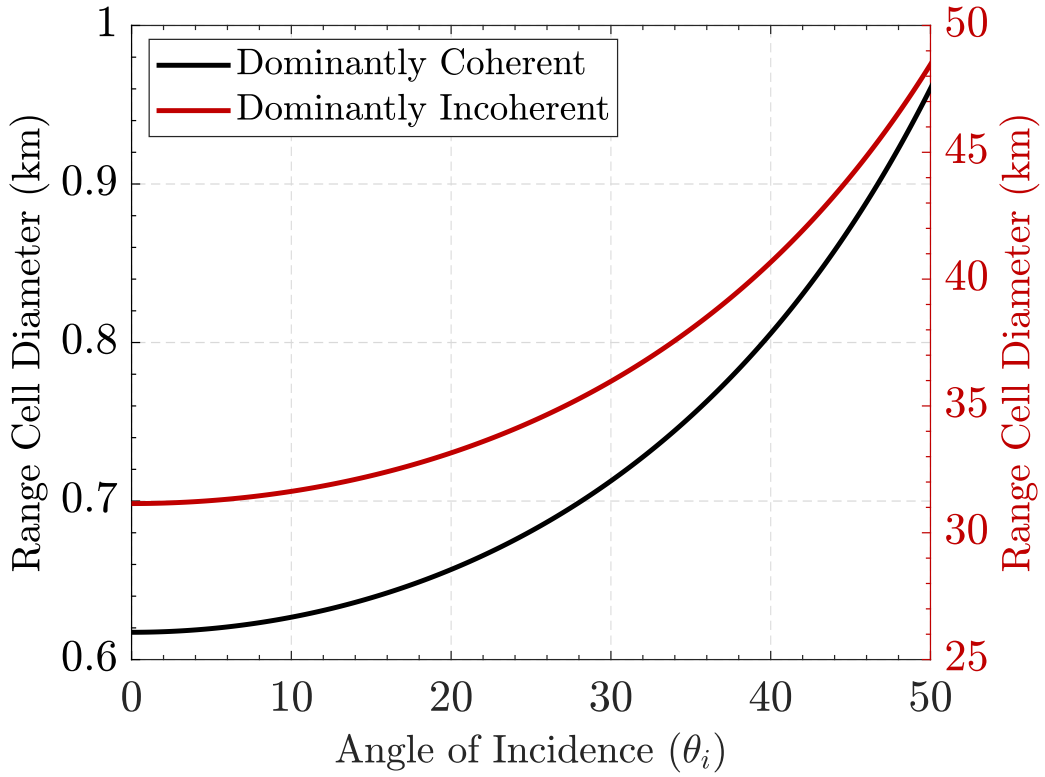


Figure 6.1: Comparisons between measurement range resolutions for dominantly coherent and dominantly incoherent CYGNSS measurements

Figure 6.1 compares the Fresnel zone diameter F_{1m} (defined as being twice the geometric mean $F_{1m} = 2\sqrt{F_{1x} \times F_{1y}}$ of the ellipses' semi-major F_{1x} and semi-minor F_{1y} axes) as a function of incidence angle with the spatial footprint diameter associated with incoherent returns [90], and shows that spatial resolutions from 600 m to 1 km may be expected when coherent reflection dominates the scattering process. This is of particular interest for inland water body mapping, since the surface roughness

of inland water bodies is typically small compared to the C/A code wavelength ($\lambda \approx$ 19 cm). Coherence is therefore expected to be the dominant mode of reflection for inland water body surfaces.

When coherent reflection dominates, the received power P_R^{coh} can be described using the Friis transmission (5.1). Surface properties affecting the received power include Γ_{LR}, the LR (right hand circular incidence and left hand circular scattering) polarized Fresnel reflectivity, h, the surface root-mean-square (RMS) height, and τ_v, the vegetation optical depth (VOD) associated with any vegetation cover (frequently related to the vegetation water content (VWC) [109].)

Past studies of inland water body dynamics with CYGNSS [36, 39, 155, 177] have focused on the use of a corrected CYGNSS signal-to-noise ratio defined as:

$$SNR_{cor} = \frac{P_R^{coh}}{P_{noise}} \frac{(4\pi)^2 (r_R + r_T)^2}{P_T \lambda^2 G_R G_T} \tag{6.6}$$

$$\approx \frac{\Gamma_{LR}}{P_{noise}} \cdot e^{-2\tau_v \sec\theta_i} \cdot e^{-4k_0^2 h^2 \cos^2\theta_i} \tag{6.7}$$

or variants thereof that may assume for example that P_T is constant for all measurements, or that may apply differing scale factors in range if incoherent scattering is assumed. Here P_{noise} is an estimate of the noise power in the DDM, obtained from an average over portions of the DDM expected to contain no Earth surface scattering. Note equation (6.7) is applicable only if it is assumed that coherent reflection dominates the scattering process.

The generally higher values of SNR_{cor} that result in the presence of inland water bodies have enabled their mapping using threshold values of SNR_{cor}. However equation (6.7) also makes evident that SNR_{cor} remains subject to variability caused

by vegetation cover, surface roughness, and is also subject to uncertainties related to calibration corrections for P_T, G_T, G_R , and other factors.

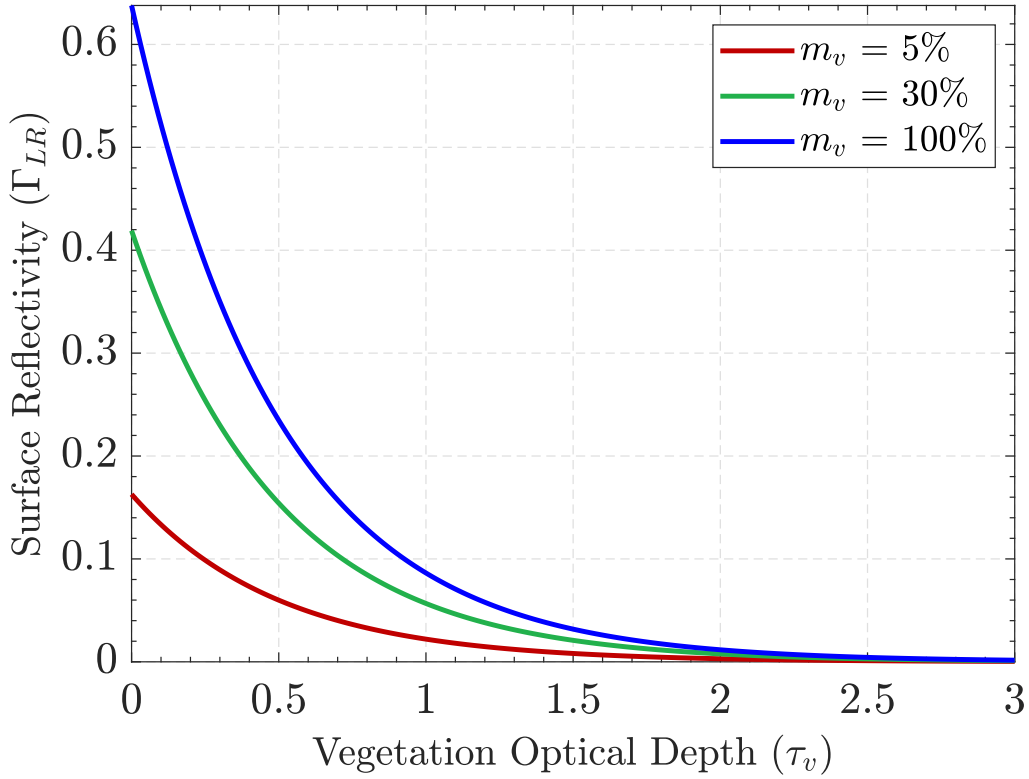


Figure 6.2: Impact of vegetation attenuation on surface reflectivity at various volumetric water contents

As an example, $\Gamma_{LR} \cdot e^{-2\tau_v \sec\theta_i}$ is plotted in Figure 6.2 versus vegetation optical depth for soil surfaces having varying volumetric soil moisture levels m_v as well as surface (standing) water with $m_v = 100\%$. The significant impact of vegetation on the observed SNR_{cor} is apparent, so that reliance on SNR_{cor} is expected to be subject to errors associated with vegetation coverage in detecting inland water bodies. Due

to these challenges, an alternative approach based on the detection of coherence in CYGNSS Level-1 DDMs is explored in the next section.

6.2.2 Inland Water Body Mapping Methodology

The proposed inland water body mapping method is based on the prevalence of CYGNSS DDM coherence on a gridded map of the CYGNSS coverage region. An example of the prevalence of coherence detections relative the total number of measurements made, on a 3 km grid and using all CYGNSS Level-1 data over the two year period from 2018-2019, is depicted in Figure 6.3. The results aim to make clear the underlying methodology with which inland water bodies are mapped in this work, relating to the expectation that inland water bodies support the recurrence (repeating) of coherence; a property that is almost entirely unique to them. The coherence detector provides a detection outcome for every CYGNSS measurement over Earth's land surface. By gridding these measurements onto a grid of a specified spatial resolution, the percentage of CYGNSS measurements within a grid cell over a specified time duration can be computed.

The results to be shown in the next Section flag grid cells as inland water bodies if more than 20% of the observations are detected as coherent. This threshold for "recurrent coherence" is used independent of location or the spatial or temporal resolution of a particular map, and was determined based on an analysis of the range of recurrences associated with the world's major water bodies. Observed recurrence percentages vary in accordance with river width, water depth and surrounding land cover. For example, over much of the Amazon river the recurrence of coherence ranged between 90-100% as shown in Figure 6.3. For other rivers, undergoing varied levels of

seasonal change and those more susceptible to surface wind roughening effects however, the recurrence percentage was found to undergo a wider range of variability between 40-100%. This nonetheless remains many folds larger than recurrence percentages to dry land which is, over much of the world's surface, <<1%. The 20% recurrence threshold is therefore set to accommodate a wide range of water bodies.

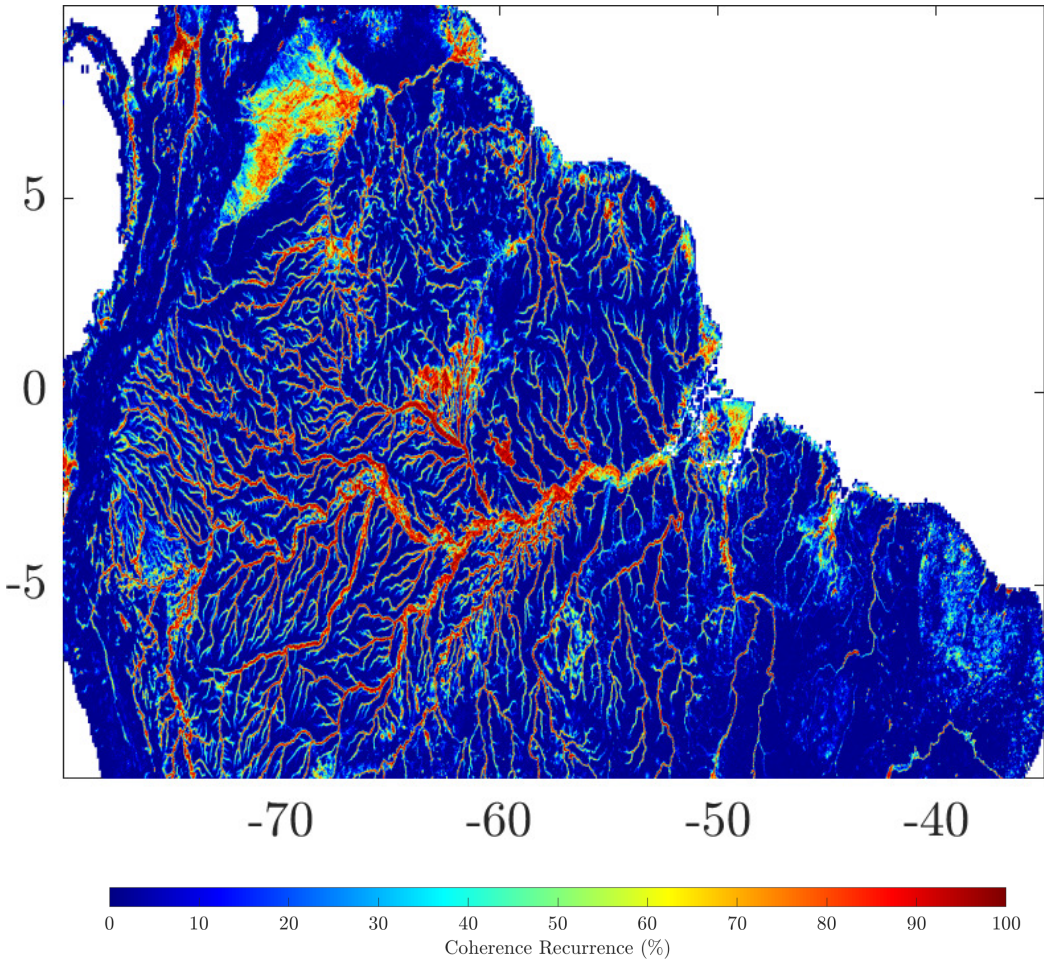


Figure 6.3: Percentage of points declared coherent relative to total measurements made over the Amazon delta on 3 km grid using all Level-1 measurements over the two year period spanning 2018/2019

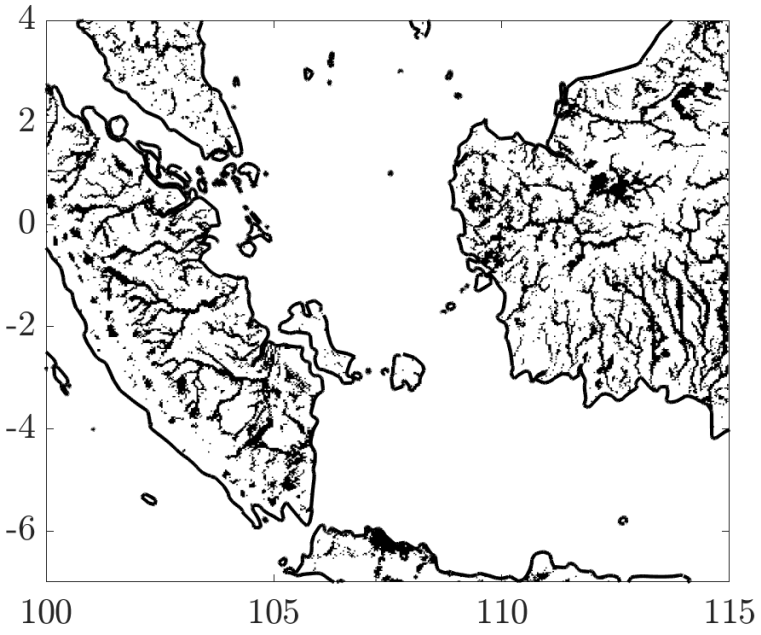
6.2.3 Inland Water Body Mapping Results

The efficacy of the proposed mapping methodology is explored on the regional and global scales using CYGNSS measurements in what follows. To asses mapping performance both quantitatively and qualitatively, comparisons are made with the Pekel occurrence water masks [133]. The Pekel masks are among the most comprehensive descriptors of inland surface water and the long-term dynamic changes it undergoes on the global scale, reporting on the presence of permanent water and its seasonality at spatial resolutions of 30 meters using data that derives from the ≈35 year data record of available Landsat imagery. Given the differences in the time history (35 years for the Pekel masks and 2 years for the CYGNSS datasets) as well as the spatial sampling of these masks, a process was developed to facilitate the intercomparison. In particular, for a specified EASE grid resolution (typically 1-3 km) on which the CYGNSS water body detection was performed, a grid cell was marked as flagged by the Pekel product if any 30 m Pekel grid cell within the EASE grid cell showed a 90% or greater (35 year) water body recurrence in the Pekel dataset. This approach was used to ensure that the comparison dataset included only "highly likely" Pekel water body locations. It is noted that the process implemented is only one of many possibilities, and that the intercomparison of these two distinct products should be interpreted carefully given the different datasets from which they arise. It is also noted that Pekel mask occurrence percentages may be indicative of true seasonality of water bodies, but can also be associated with the occlusion of their detection due to cloud cover or thick vegetation cover.

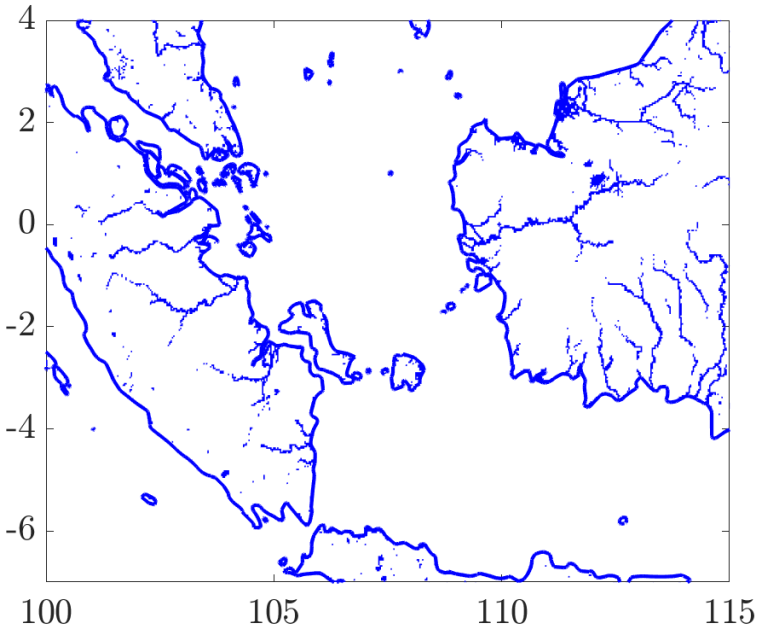
6.2.4 Regional Mapping

Example mapping results for a portion of the Maritime Continent and the south-eastern United States are provided in Figures 6.4 and 6.5 respectively. The success of the mapping approach over the Maritime Continent in Figure 6.4 for inland water bodies within Indonesia, Malaysia and Singapore is evidenced by a visual examination of plots (a) and (b) comparing the CYGNSS based water map for a 3 month analysis period on a 3 km EASE grid to a downsampled Pekel occurrence mask on the same resolution.

The probability of detection, that is the probability of the CYGNSS based mapping approach of correctly detecting water bodies identified by the Pekel occurrence mask, over the Maritime Continent is found to be 78.67%. Within this test region, the CYGNSS map suggests that 16.78% of all land points equivalent to an area of 14,351 km^2 are inland water bodies while the Pekel map detects only 5.59% or equivalently an area 5,046 km^2. While some percentage of false alarms, that is the percentage of land points declared as being water bodies by CYGNSS and not identified by Pekel, is expected to be present, many of the CYGNSS detections were found to be associated with significantly lower occurrence percentages within the Pekel map ranging between $30 - 80\%$. An example of this for the Maritime Continent is depicted in Figure 6.6(a). In particular, further analysis showed that the CYGNSS-detected fine tributaries have widths on the order of 200 meters or less for the rivers Miru, Musi, Lematang, Pawan, Kapuas, Sungai Seruyan and others.

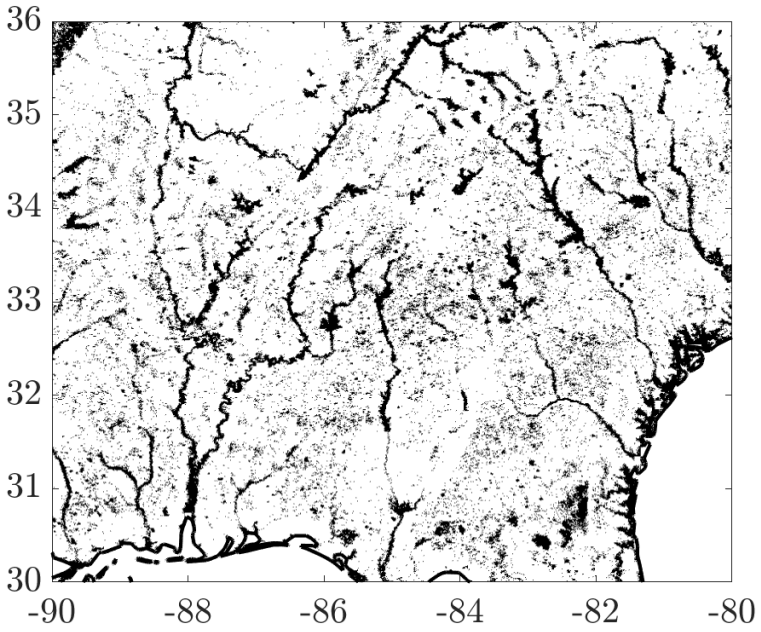


(a)

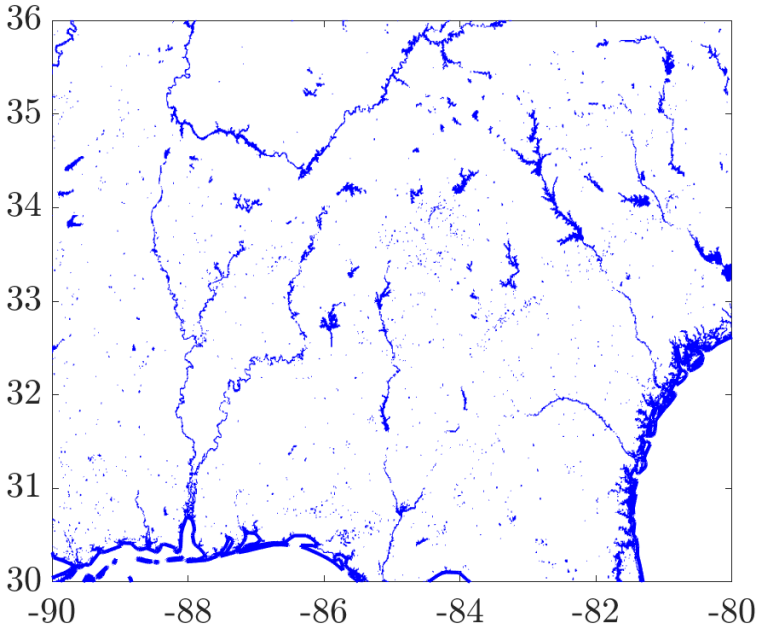


(b)

Figure 6.4: Case study for a segment of the Maritime Continent (a) CYGNSS based watermask on 3 km EASE grid generated using 3 months of L1 data from October 2019 to December 2019 (b) Pekel occurrence water mask down sampled to 3 km resolution

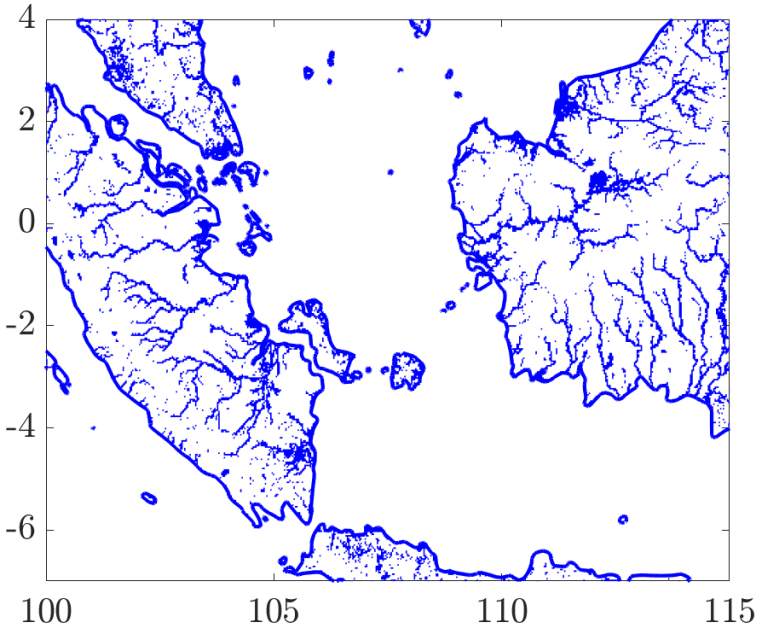


(a)

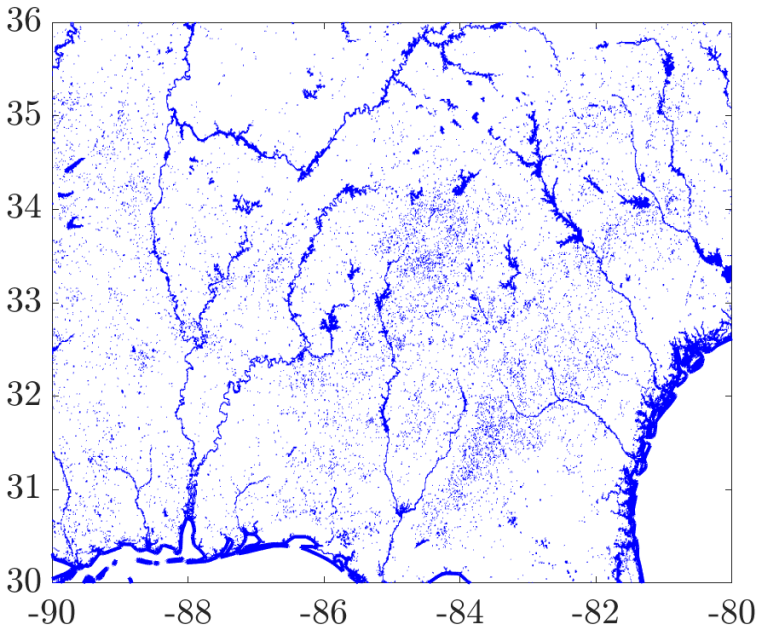


(b)

Figure 6.5: Case study for a portion of the United States (a) CYGNSS based water-mask on 1 km EASE grid generated using 1 year of L1 data from January 2019 to December 2019 (b) Pekel occurrence water mask down sampled to 1 km resolution



(a)



(b)

Figure 6.6: Pekel occurrence maps with lower threshold percentages for regional inland water body mapping study (a) Pekel occurrence water mask down sampled to 3 km resolution with 50% occurrence threshold over Maritime Continent (b) Pekel occurrence water mask down sampled to 1 km resolution with 80% occurrence threshold over USA

Their lower occurrence within the Pekel maps is attributed to the fact that the environment within this region is largely opaque for the optical Landsat imagery from which the Pekel mask derives, predominantly comprising International Geosphere-Biosphere Program (IGBP) evergreen broadleaf forest land class (woody vegetation with canopy heights exceeding 2 meters [114] and having an average VWC exceeding 15 kg/m^2). The distribution of the CYGNSS SNR for points detected as being coherent over lakes, rivers and their tributaries over the 3 month analysis period is found to range between 1.5 to 26.4 dB. Observations dominated by coherent reflection within the lower SNR range of $1.5 \leq \text{SNR} \leq 3.5$ dB were found to comprise 25% of all measurements, corresponding to water bodies obscured by vegetation, highlighting some of the compromises that may arise from the reliance on a single SNR threshold.

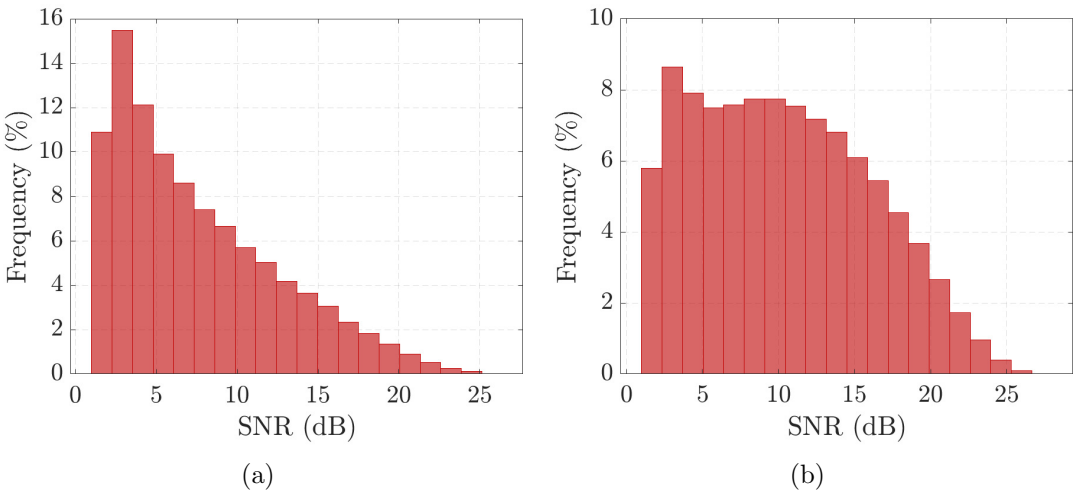


Figure 6.7: SNR distribution for measurements exclusively over the lakes, rivers and their tributaries detected as being dominantly coherent over the analysis periods illustrating ambiguities with reliance on observables' magnitude for inland water body mapping (a) Over Maritime Continent case study (b) Over the analysis period over southeastern USA case study

While water masks produced with a resolution finer than 3 km are desirable, there exists an inherent limitation between the resolution of the surface grid on which CYGNSS's data is projected and the duration over which the analysis is undertaken. Coarser grids capture more CYGNSS measurements within a grid cell, allowing for more frequent updates and examination of the seasonality of water bodies, their growth and extinction. Finer grids in contrast require longer durations to acquire a sufficient number of points per grid cell. As an example, Figure 6.5 depicts a 1 km inland water body map generated using a 12 month record of CYGNSS's L1 data over the southeastern United States. The computed probability of detection as compared to the Pekel map is found to be 90.29%. The CYGNSS map indicates 17.41% of the test region is inland water, while the Pekel map obtains 6.035%. Similar interpretations of the discrepancies, relating to the occlusion of detection by optical imagery of these narrow water bodies by their environment, arise examples of which are depicted in Figure 6.6(b). This includes CYGNSS detections for the Leaf River forking at approximately 30.98°N -88.73°W in Mississippi, the Pearl River forking at approximately 30.31°N -89.64°W between Louisiana and Mississippi, parts of the Tallapoosa River branching out at approximately 32.50°N -86.26°W in Alabama , and the Oconee River branching out at approximately 31.96°N -82.54°W in Georgia. All of these discrepancies coincide with rivers having widths that were on average less than 150 meters and were surrounded or covered by vegetation. Many of the other smaller discrepancies were found to be associated with smaller water bodies, but the absence of accurate *in-situ* information limited a detailed investigation of these cases. The same set of water bodies were also found to be identified within the Pekel dataset, with occurrences ranging between 60-80%. The SNR over all detected

CYGNSS water bodies within this region again shows a wide distribution of SNRs ranging from 1.5 to 15 dB.

Although the mapping process uses a threshold of only 20% for "recurrent" coherence, the maps created for both the Maritime Continent and the United States on average found that approximately 70% of DDMs in a grid cell were declared as being dominantly coherent for the detected water body locations. The non-detection of the remaining 30% for these grid cells can be attributed to a variety of factors. One important factor is the motion of the specular point through and across grid cells that occurs within the 1 second (prior to July 2019) or 500 msec (after July 2019) period of a CYGNSS measurement. Given the CYGNSS satellite orbital velocity of approximately 6000 m/s, the resulting "spatial smearing" of the CYGNSS measurement over heterogeneous scenes can obscure water body returns, particularly for those of small size. Fluctuations in coherent returns associated with the specific water body shape and the orientation of the specular track relative to this shape can also influence the dominance of coherent contributions [90, 128].

A visual examination of plots (a) makes clear that the mapping methodology successfully captures all major occurrences of the water bodies reported within the reference Pekel occurrence maps in plots (b). However, due to the mapping of the observed surface by the CYGNSS constellation one specular point at a time, as opposed to large swaths of the surface, an insufficient accumulation of points associated with dominant coherence for individual pixels may occur. As a result, individual pixels within the major inland water formations may be flagged as non-water by the mapping procedure and result in the ≈80-90% probabilities of detection reported for both case studies. A simple nearest neighbor interpolation is an effective tool to address

this, that is declaring individual non-detected pixels surround by detected pixels by all sides as being inland water, and is found to lead to detection probabilities of 97.97% and 92.57% for the Maritime Continent and the United States respectively; those are expected to be the 'true' representative statistics of the probability of detection as evidenced by a visual examination of Figures 6.4 and 6.5. This method is applied in the results that follow.

6.2.5 Mapping on Larger Spatial Scales

A map for the entire CYGNSS coverage region (Figure 6.8) was then created on a 1 km EASE grid using two years of CYGNSS data. However, initial analysis of the resulting map showed two areas that required improvement. The first was related to false alarms over desert regions having exceptionally flat surfaces.

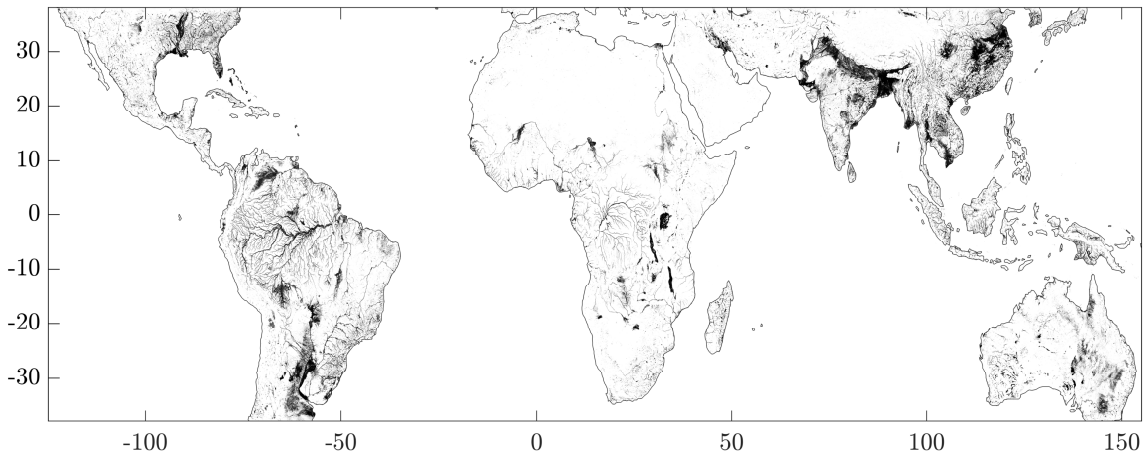


Figure 6.8: Global inland water body mask generated using 2 years of CYGNSS Level-1 data from 01/18 to 12/19 on 1 km EASE grid

The largest of these regions is the Nullarbor Plain in southern Australia spanning an area in excess of 200,000 km^2. Others, shown in Figure 6.9, include small sections of the Kuwaiti Desert, the Sahelian Zone in Mauritania, the Saharan Zone in Mauritania, the Algerian Desert in Algeria, the Libyan Desert in Libya and the semi-arid savanna in Africa.

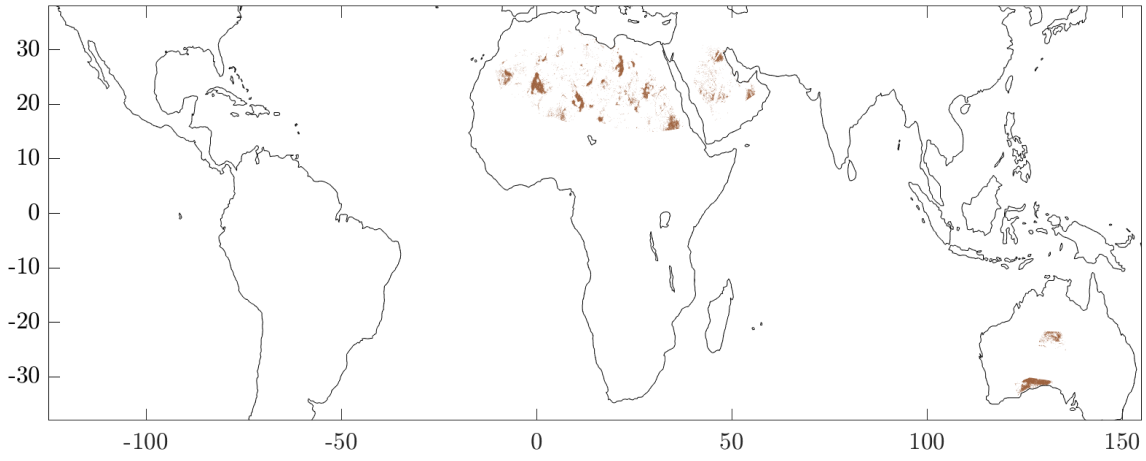


Figure 6.9: Regions conducive to recurrent coherence that are known not to be hosts of any inland water bodies and are therefore with an inland water body occurrence of 0%

Analysis of available digital elevation map datasets showed these regions to have values of RMS roughness near the minimum resolvable for the maps used. Because the coherence for these regions is associated with the land surface as opposed to water surface topography (with water surface roughness being much more subject to changes in wind speed), the resulting recurrence of coherence showed little to no temporal variability. The known absence of inland water within these regions enabled their removal from Figure 6.8 by removing locations having a 0% Pekel occurrence of water. Similar false alarms over large portions of the Levant, Arabian Peninsula

and the Sahara have been reported in other surface water datasets derived from both active and passive microwave remote sensed data [156].

The absence in the created map of the largest lakes within CYGNSS's coverage and the third, sixth and ninth largest lakes in the world (Lakes Victoria, Tanganyika and Malawi respectively) was the second issue requiring further investigation. This is attributed to the larger size of these water bodies, such that wind roughening can cause coherence no longer to be the dominant mode of scattering. An examination of DDMs for these cases showed the frequent presence of weak incoherent scatter with an equally pronounced coherent component occasionally present. This "partially coherent" state can be identified using $0.2 \leq PR < 2$ and SNR$\geq$15 dB [154]. Further analysis highlighted that the local invocation of this threshold enabled the successful mapping of 80.11%, 87.14% and 96.17% of the surfaces of lakes Victoria, Tanganyika and Malawi respectively.

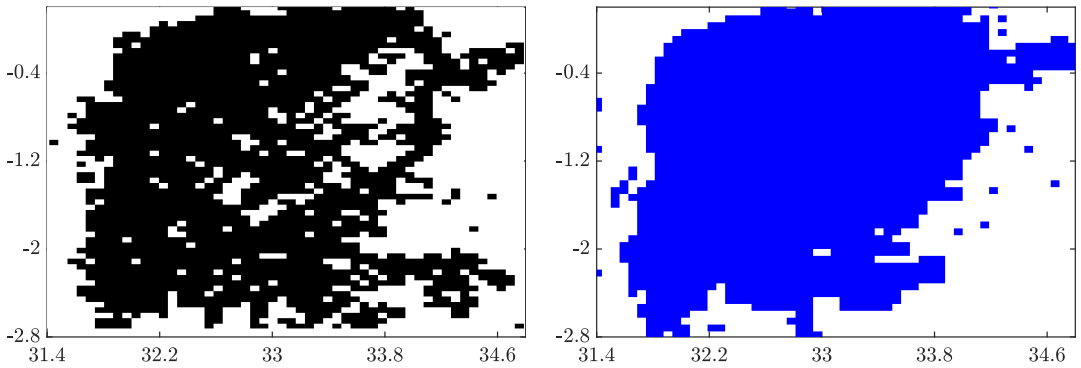


Figure 6.10: Mapping of Lake Victoria by invoking mixed coherence detection (a) Lake Victoria mapped using 6 months of CYGNSS data from 09/19 to 02/20 on 6 km EASE grid (b) Reference Pekel occurrence water masks for Lake Victoria down sampled to a 6 km spatial resolution

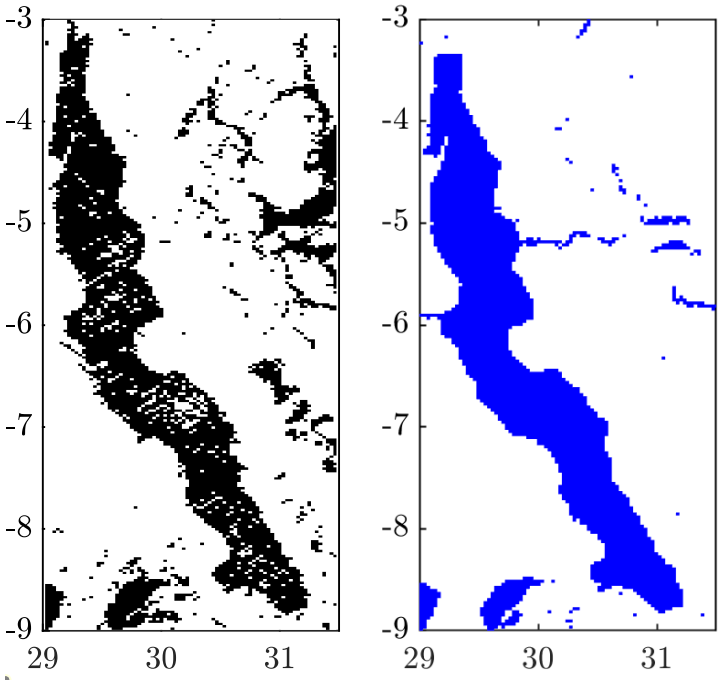


Figure 6.11: Mapping of Lake Tanganyika by invoking mixed coherence detection (a) Lake Tanganyika mapped using 8 months of CYGNSS data from 07/19 to 02/20 on 3 km EASE grid (b) Reference Pekel occurrence water masks for Lake Tanganyika down sampled to 3 km spatial resolution

The mapping approach's probability of detection relative to Pekel occurrence masks was found to also be on the order of 80%, with a true detection probability (after interpolation) of ≈96%, over most of the major rivers within CYGNSS's coverage including the Amazon River and its extensive network of tributaries, the Congo River, the river Nile and the Niger River. The utility of the CYGNSS inland mapping approach in producing a global water mask is illustrated in Figure 6.8 on a 1 km EASE grid using a 2 year record of CYGNSS's L1 measurement. While an exact determination of the finest spatio-temporal global inland water body masks attainable

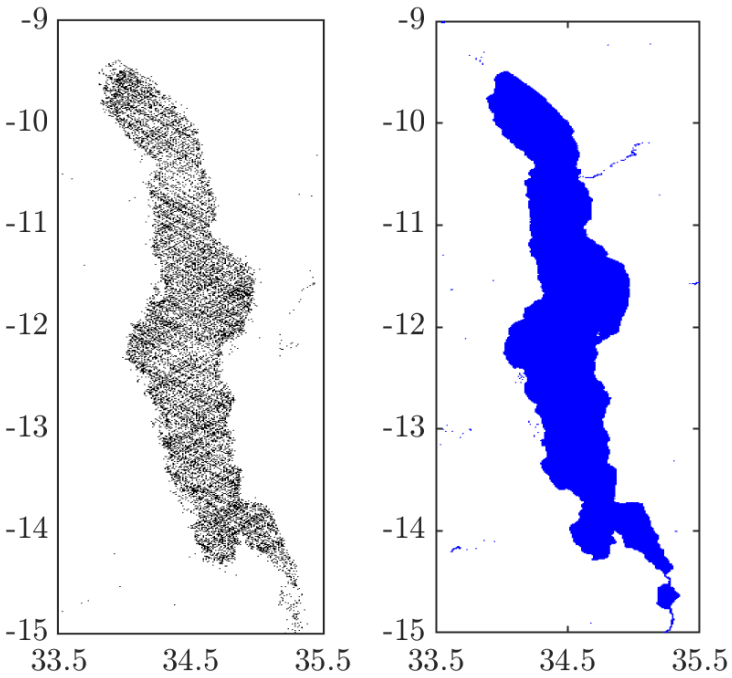


Figure 6.12: Mapping of Lake Malawi by invoking mixed coherence detection (a) Lake Malawi mapped using 26 month record of CYGNSS data from 01/18 to 02/20 on 1 km EASE grid (b) Reference Pekel occurrence water masks for Lake Malawi down sampled to 1 km spatial resolution

using the CYGNSS inland mapping approach requires a more extensive analysis with a longer record of CYGNSS data as it becomes available, initial estimates suggest that a coverage similar to that depicted in Figure 6.8 is possible at spatial/temporal resolutions of 1 km updated every 1 year, 3 km updated every 3 months and 6 km updated every 2 weeks.

6.3 Large Scale Atmospheric Dynamics Mapping

Global climate and its circulation is modulated by a number of large scale processes of which the South Asian monsoon plays a key role [158], affecting through its dynamics convection life cycles, water/energy fluxes and weather forecasting [159–161]. The implications this process has in terms of extended periods of drought and several months of torrential rainfall has on the assessment of agricultural productivity, food security, sudden flooding and disaster preparedness are of high importance [162,163]. Owing to the vastness of the regions affected by the dynamics of the South Asian monsoon, spanning an area greater than 4 million km^2, a need arises for monitoring its different stages of development and effects using spaceborne sensing platforms, due to the availability of their measurements on a global scale. A similar need arises to incorporate these observations into relevant prediction and reanalysis models [164–166].

The properties of the Cyclone Global Navigation Satellite System (CYGNSS) mission make it well poised to address this need. This section attempts to develop a methodology that is used to map the dynamic changes the monsoon season brings to the Indian subcontinent in near real time by specializing the function of the recently developed Level-1 coherence detection methodology [154] described in Chapter 5 to mark the start, end and evolution of the monsoon's characteristic phases. This is pursued by identifying transitions of the dominant scattering mode within different parts of the Indian subcontinent and relating this to the development and dissipation of different phases of the monsoon. Through the insensitivity of the proposed methodology to calibration uncertainties, different levels of surface roughness as well as land cover, this work aims to show the potential for using spaceborne GNSS-R systems in an operational capacity in the mapping of these large scale weather phenomena.

6.3.1 Description of the Monsoon Season

The South Asian (Indian) monsoon is a consequence of the interactions of near surface ocean air, varied levels of air moisture and impinging solar radiation. As sunlight heats the ocean's surface near the equator, evaporated water mixes with warmed air creating a flow of wind propagating northwards reaching an upper zonal limit towards the pole. The transfer of the large mass of warm moist air, creates a region of low pressure with dry air known as the Intertropical Convergence Zone (ITCZ) [167]. This stage of the solar-surface interaction brings about the dry stage of the monsoon season characterized by the absence of rainfall and the bringing about of arid conditions throughout the Indian subcontinent predominately affecting Pakistan, India, Nepal, Bhutan, Bangladesh, Sri-Lanka and Myanmar. The drought phase typically prevails for a total of 8 months starting from October and ending with May of every year. As the warmed, humid air reaches its zenith, it cools and begins its descent towards the subtropics at $\approx 30^{\circ}$ where a process of condensation ensues marking the beginning of the wet phase, bringing about heavy rainfall culminating in transient flooding and inundation over different parts of the Indian subcontinent. The wet phase typically lasts for a period of 4-5 months ranging between June to early October. This cyclical atmospheric circulation, the Hadley Circulation [168], is the fundamental driver for all observed manifestations of the monsoon season. Both phases of this process create unique challenges.

During the dry phase a need arises to monitor its onset and end for water resource management, water conservation and distribution practices as well as the assessment of agricultural productivity which in turn can have profound implications on food security. Similarly, tracking the dynamics of the wet phase is of particular importance

for disaster preparedness due to the heavy and sustained rainfall often resulting in flooding leading to the loss of life and destruction of property [162,163]. The ability to predict and monitor monsoon dynamics is nonetheless complicated by its sudden onset and the linkage of its underlying circulation mechanisms, the Hadley Circulation, with other larger scale phenomena such as the El Nino and Southern Oscillation (ENSO) phenomenon [167]. As a result the monsoon onset and evolution prediction ability of current models remains limited at best [169, 170], motivating exploring the use of systems like the CYGNSS constellation to address this need due to its global coverage and ability to provide low latency measurements.

6.3.2 Sensitivity to Transient Flooding and Inundation

Numerous previous studies have explored the use of radiometers [171, 172], synthetic aperture radars [173, 174] and optical [175] systems for the purposes of examining the various manifestations of monsoon effects on the observables of spaceborne receivers and means by which the two may be related. In the case of passive radiometers, the flooding and inundation effects associated with the monsoon wet phase may be 'detected' through the anomalous dampening this induces to observed brightness temperatures which may be used together with emissivity indices through the enforcement of various detection thresholds to provide indications of monsoon dynamics [171]. In contrast, studies analyzing SAR backscattered normalized radar cross section for vegetated land surfaces have modelled received signals as [20,176] the sum of three mechanisms describing the aggregated contributions of surface, volume and surface-to-volume scatter. Through varied interactions between the three fundamental contributions, flooding and inundation reduces backscatter levels thereby

revealing the presence of inundation through 'darker' areas within the SAR image. In spite of the presence of inundation effects within a given scene, varied levels of 'darkness' may be perceived depending on levels of surface roughness and variety of land cover types. On the other hand, the sensitivity of CYGNSS's measurements to inundation is through the direct proportionality of peak received DDM power to surface reflectivity [68, 90] given by (4.4).

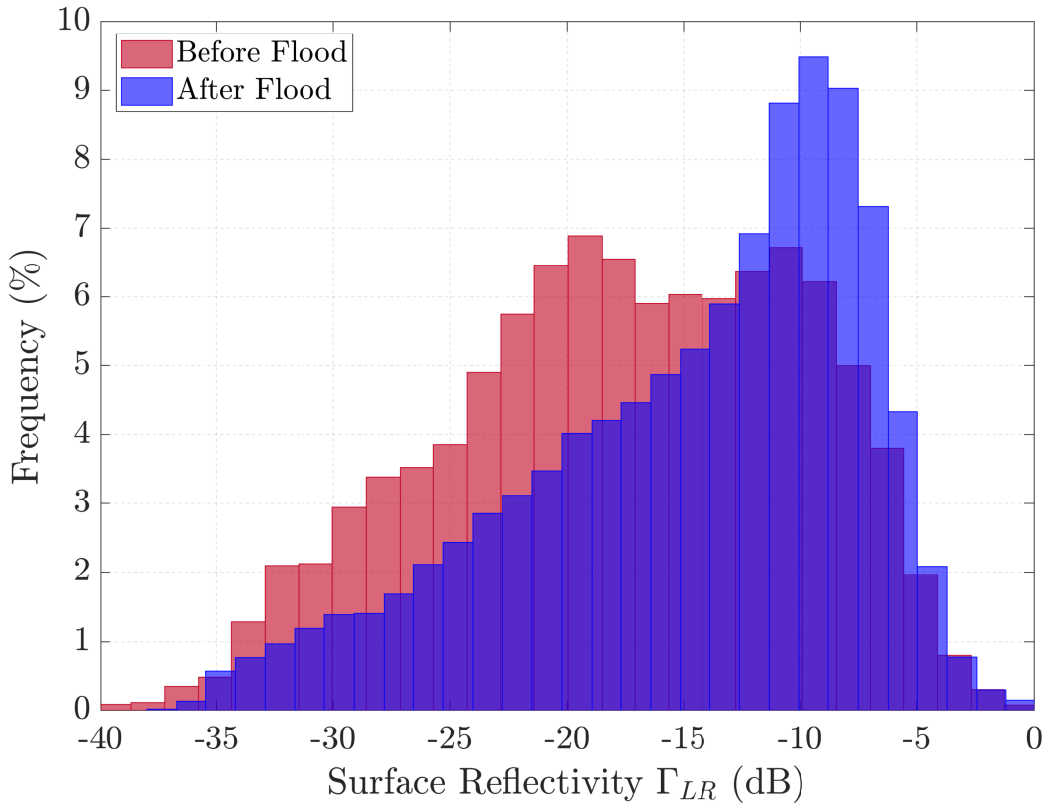


Figure 6.13: Changes in CYGNSS surface reflectivity in response to flooding in the vicinity of the Indus River coinciding with monsoon wet phase

Surface reflectivity, in turn, is dictated by the constituents of soil through its dependence on ϵ_s, but undergoes the most significant variability based on its volumetric

soil moisture content [26,100]. As the monsoon's wet phase begins, the increased levels of rainfall results in a significant increase in moisture content and a complementary increase in the soil's complex permittivity approaching that of water ≈80. Surface reflectivity undergoes complementary seasonal variations, experiencing increases during flooding and inundation events and a steady decline as volumetric moisture content diminishes during the monsoon's dry phase, accounting for the success of numerous previous studies in using CYGNSS measurements to study seasonal inundation on smaller scales, over various wetlands and as a result of storm surge [36, 39, 155, 177].

Table 6.1: Summary of relevant statistics for case study illustrating the sensitivity of CYGNSS measurement surface reflectivity to flooding event detected by MODIS NRT in the basin on the Indus River, Pakistan.

	Before Flood	After/During Flood
Start Date	July 1st, 2019	August 1st, 2019
End Date	July 31st, 2019	August 31st, 2019
Number of Measurements	33,069	29,550
Mean Γ_{LR} (dB)	-17.59	-14.59
Std Γ_{LR} (dB)	7.63	7.27
Peak Flood Level (%)	0.00	> 50
Mean NDVI	0.06	0.11
Std NDVI	0.05	0.09

The sensitivity of CYGNSS measurements to these effects is illustrated in Figure 6.13 where surface reflectivity derived from calibrated CYGNSS Level-1 DDMs, corrected for instrument and geometry related parameters, is depicted for two, 1 month long periods before (over July 2019) and after (during August 2019) a monsoon related flooding event in the vicinity of the Indus River took place. From Table 6.1 it follows that other surface properties including roughness and vegetation remained

relatively unchanged as evidenced by the modest 2.5% increase in Normalized Difference Vegetation Index (NDVI), expressed on a scale from -1 to +1, derived from the Advanced Very High Resolution Radiometer (AVHRR) within the targeted region throughout the two month analysis period. Within the same area, the Moderate Resolution Imaging Spectroradiometer (MODIS) Near Real Time (NRT) suggests that over the month of August, flood occurrence exceeded 50% within the test region. The occurrence percentage is determined based on the portion of MODIS measurements associated with the detection of water relative to total measurements made, with flood state declared based the number of water observations exceeding 'normal' water levels for a given region based on yearly MODIS water masks. The reported flooding event is expected to be the primary driver for the observed 17.06% change in mean CYGNSS surface reflectivity during/after the flooding event relative to the period before it occurred.

While this example highlights the sensitivity of CYGNSS measurements to the effects of the monsoon season, via its sensitivity to the flooding events which it brings about, it also highlights a number of challenges, many of which are not unique to CYGNSS relating to the reliance on the magnitude of observables to aid in the mapping of monsoon effects and dynamics. In particular, the significant overlap of the observed reflectivities pre/post the flooding event, captured by the 49.83% Γ_{LR} standard deviation relative to the mean, makes clear the ambiguities that may arise from attempting to enforce detection thresholds on the magnitude of observables. Furthermore, while NDVI on the local scale remains relatively unchanged over the analysis period, over all regions within the Indian subcontinent affected by the monsoon season levels of vegetation range between -0.100 and 0.989 (or equivalently a 54.45%

dynamic range). This together with limited knowledge of surface roughness, creates additional variability in observables' magnitude [178]. This is further compounded by the need for precise power calibration of measurements and while CYGNSS's are continuing to advance with time [75, 79], the difficulties of achieving precise absolute calibration for eight observatories acquiring the reflections of 32 distinct GPS transmitters has created a need to explore robust detection and mapping methodologies that are insensitive to these effects. Those are explored in the next section.

6.3.3 Proposed Mapping Methodology

Instead of using variations in the magnitude of CYGNSS observables or their rate of change with time, the proposed methodology aims to map the dynamics of the South Asian monsoon by marking the onset of its wet phase, its persistence, the subsiding of its effects and ultimately the transition to its dry phase through leveraging the ability to detect the dominant underlying GNSS-R scattering mechanism contributing the bulk of received power within a given set of DDM measurements, and relating the interchange between one mode to another to monsoon dynamics.

For a given location a transition can nonetheless occur from dominant incoherent scattering to dominant coherent reflection. This work therefore aims to specialize the function of the Level-1 detector described in Chapter 5 in order to investigate these transitions and use them as indicators of monsoon dynamics. This is then coupled with establishing the dominant underlying scattering mechanism, and any changes that it might undergo within a given grid cell. This is explored in an example of this process depicted in Figure 6.14 for a 36 km grid cell within Bihar, in the northern districts of India.

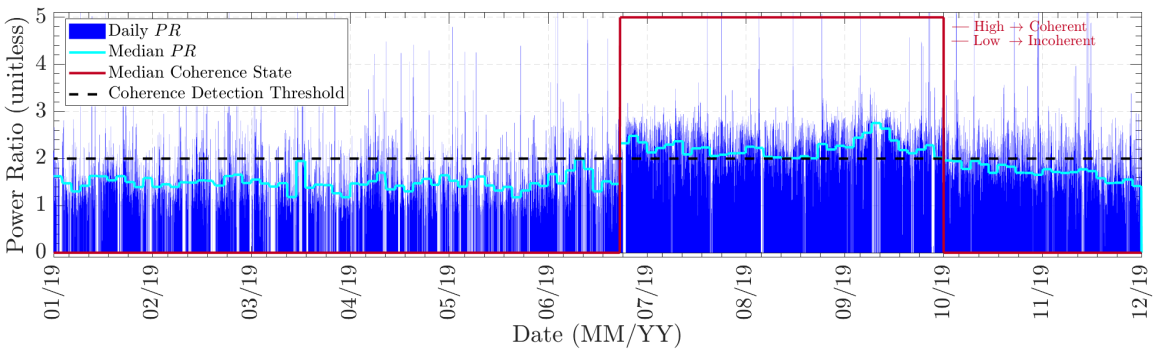


Figure 6.14: Level-1 coherence metric for all CYGNSS observations made throughout 2019 for a 36 km footprint centered about 25.76°N 84.09°E. Illustration depicts transition from dominant incoherent scattering state during the dry phase of the monsoon to a dominant coherent scattering state during the wet phase of the monsoon.

The analysis presented depicts coherence estimates for all Level-1 measurements made within this pixel throughout 2019. The pixel spans parts of the Ganges River, Ghaghara River, Suraha Lake as well as drier, rough, terrain within which no known water bodies exist. A total of ≈20,000 measurements were made for this location, or equivalently an average of 55 per day, and depending on the precise locations of the specular points within this grid cell any number of them may be dominated by coherent reflection or incoherent scatter as evidenced by the numerous transitions above/below the 2.0 detection threshold in Figure 6.14. Therefore, the reliance on the mere presence of coherence as an indication of monsoon flooding effect is expected to give rise to considerable uncertainty. Instead, the dominant mode of scattering within a given grid cell is established based on the daily median PR value. The latter is indicative of the most recurring, 'dominant' mode of reflection within a series of measurements for a given location. For days with a median power ratio below the detection threshold, incoherence is declared to be dominant and is taken

to indicate the onset or persistence of the dry phase. The transition to a median value that exceeds the detection threshold is indicative of the onset or persistence of the monsoon's wet phase. In the case of the example in Figure 6.14 the transition to a dominant mode of coherence, ergo the monsoon's wet phase, begins July 9th and ends October 25th, 2019. The persistence of a dominant incoherent state throughout other parts of the year marks the start, end and continuation of different cycles of the monsoon's dry phase.

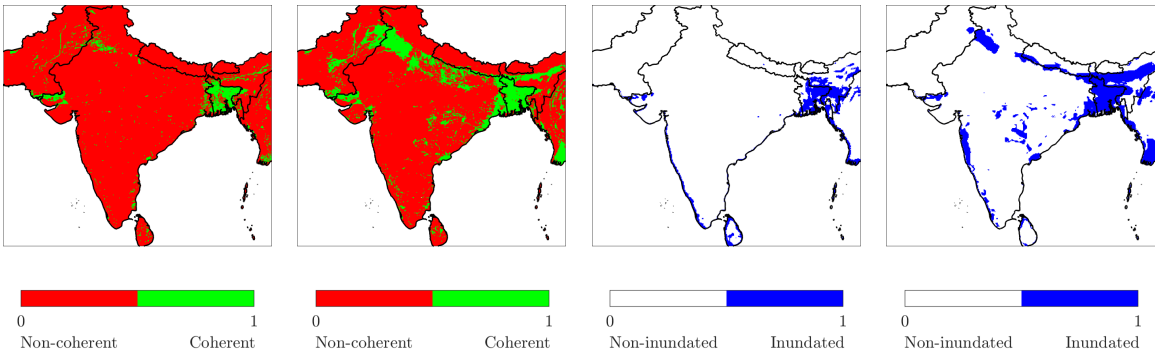
It is further noted that the increase in the time series' density in terms of available measurements beyond July 1st, 2019-present is the result of halving the Level-1 DDMs' incoherent integration time thereby doubling data rates from 1 Hz to 2 Hz. The transition was directly in support of land investigations as it reduced along track spatial smearing, particularly for DDMs dominated by coherent reflection, from a distance of 6 km to 3 km.

6.3.4 Dynamics Mapping Results and Discussion

The ability of the proposed methodology to map the dynamics of monsoon induced flooding and inundation are explored using the multi-year Level-1 data record made available by the CYGNSS mission extending from Day of Year (DOY) 77, 2018 to DOY 130, 2020.

Spatial Trends

Consider the results summarized in Figure 6.15, plots (a) and (b), depicting the areas where a dominant mode of coherence is detected before (May 2018) and after (July 2018) the onset of the monsoon's wet phase on a 9 km spatial grid using 30 day medians.



(a) (b) (c) (d)

Figure 6.15: Illustration of spatial trends associated with dry and wet phases of the monsoon (a) Distribution of coherent and incoherent points on 9 km grid during dry phase based on 30 day medians throughout May 2018 (b) Distribution of coherent and incoherent points on 9 km grid during wet phase based on 30 day medians throughout August 2018 (c) Points associated with elevated SMAP soil moistures indicative of inundation on 9 km grid using 30 day mean throughout May 2018 (d) Points associated with elevated SMAP soil moistures indicative of inundation on 9 km grid using 30 day mean throughout August 2018

During the dry phase a total of 5.61% of all points within the Indian subcontinent were associated with coherent reflection. Transient flooding and inundation brought about by the wet phase resulted in a growth of this percentage to 16.26%. In the absence of ground truth, provided with comparable spatial and temporal scales, the observed effects are compared to reported Soil Moisture Active Passive (SMAP) mission soil moistures on a 9 km grid (based on enhanced L2 radiometer half-orbit 9 km EASE-grid soil moisture SPL2MPE product). Points with reported moistures that exceed 45% are declared inundated. During the dry phase, the SMAP maps in plot (c) suggest that 4.57% of all points are inundated while in plot (d), over the wet phase this is found to increase to 12.51%. The utility of the proposed methodology in

detecting and mapping the effects of the monsoon are evidenced by a visual examination of plots (a-b) with (c-d) where the spatial distributions of locations reported as being dominantly coherent using CYGNSS data and those declared inundated using SMAP data are found to be highly comparable. The additional detections made by CYGNSS however, particularly during the dry phase, are the result of its superior spatial resolution under a dominantly coherent reflection regime ($\geq$ 200m) [153] where in the border region between Pakistan and India a number of additional detections are observed. Those are associated with known inland water bodies which include the Ramn of Kutch Lake, Shakoor Lake, Indus River and Chehab River. The difference of additional coherence detections made in plot (b) relative to the signature of the dry phase in plot (a) enables the estimation of the additional area inundated solely due to the monsoon's wet phase. In the case of the 2018/2019 cycle depicted in Figure 6.15, the wet phase resulted in transient flooding and inundation that spanned 10.65% of the Indian subcontinent or equivalently 43,110 km^2 while the reference SMAP dataset suggests that this is 7.94% or equivalently 32,193 km^2.

Temporal Variability

While the preceding results highlight the large scale spatial trends associated with the monsoon's characteristic phases, of most interest is the ability to track its temporal dynamics. An example of this is shown in Figure 6.16 based on analysis that targets India, projecting the relevant CYGNSS data on a 12 km grid and updating coherence/inundation estimates every 10 days.

A clear monsoon wet phase signal is observed characterized by a significant, relative to the dry phase months, increase in the total percentage area detected as being inundated reaching peaks of $\approx$25% during the monsoon's wet phase compared to a

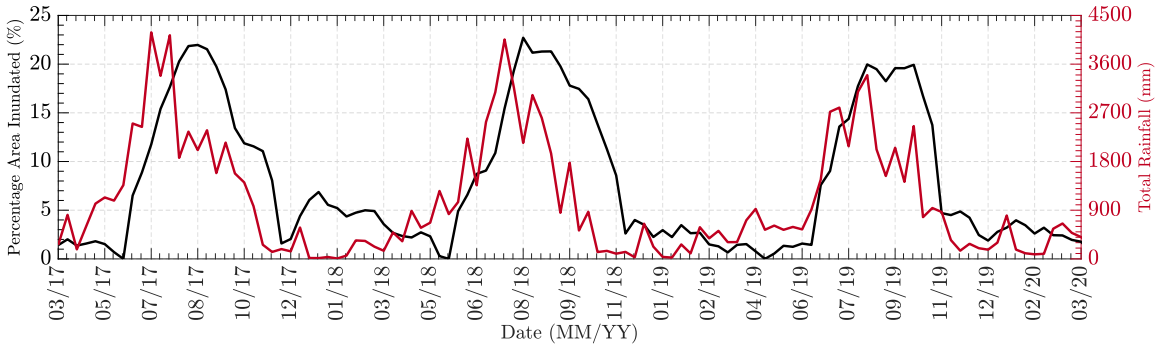


Figure 6.16: Case study over India. Detected total inundation percentage relative to total area, on 12 km grid using 10 day medians, and total GPM reported rainfall over four dry phases and three wet phases of the monsoon season.

prevalence of coherence that is generally limited to <5% during the dry phase. A number of other recurrent trends are observed. During the dry phase, the percentage of points detected as being coherent undergoes a monotonic decrease beyond the end of any given wet phase. This is expected to be the result of prolonged periods of drought resulting in a sustained drying of previously inundated, and water logged, areas. A minimum, of <1% of all grid cells being conducive to dominant coherence, is typically reached marking the end of any given dry phase and the beginning of a monotonic increase in the total number of points detected as being inundated marking the onset of the wet phase. Examples of this are the minimums reached on June 5th and May 1st prior to the 2017 and 2018 wet phases, respectively. A second feature relates to the lag between the onset of elevated rain rates and the detection of inundation, and as a consequence CYGNSS declarations of the beginning of the wet phase. The lag is to be expected, as the proposed mapping methodology detects monsoon dynamics indirectly through its impact on surface roughness as opposed to directly detecting rain events. As a result, the lag arises due to the time it takes

for sufficient rain fall to occur. During the 2017 wet phase, peak rain occurred on July 20th whilst peak inundation was detected on August 14th suggesting a lag of 25 days. During 2018 the lag was 21 days long. During 2019, however, peak rain and inundation detection coincide on the same day, August 14th. In this case, peak rain is not expected to be the primary driver for the peak in inundation, as the latter was preceded by prolonged periods of above average rainfall, compared to previous years, throughout India.

While the results depicted in Figures 6.15-6.16 are indicative of the success of the proposed methodology in capturing the large scale spatial trends associated with the monsoon's two major phases, dry and wet, they are associated with non-trivial 1 month and 10 day latencies. This is the result of the need to allow a sufficient accumulation of measurements within a given grid cell in order to investigate the dominance of one mode of scattering over another. Finer spatial grids are therefore inherently associated with coarser temporal resolutions. Nonetheless, near real time mapping of monsoon dynamics is possible using coarser spatial grids.

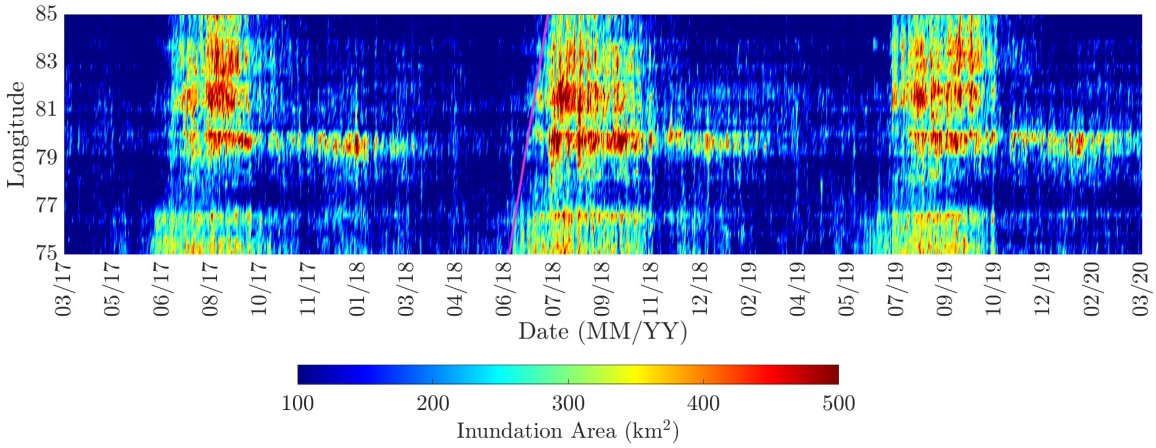


Figure 6.17: Hovmoller diagram illustrating total area associated with inundation brought about by the monsoon's wet phase. Trends indicative of the monsoon's dry and wet phases with the slope of the magenta line used to compute eastward propagation velocity.

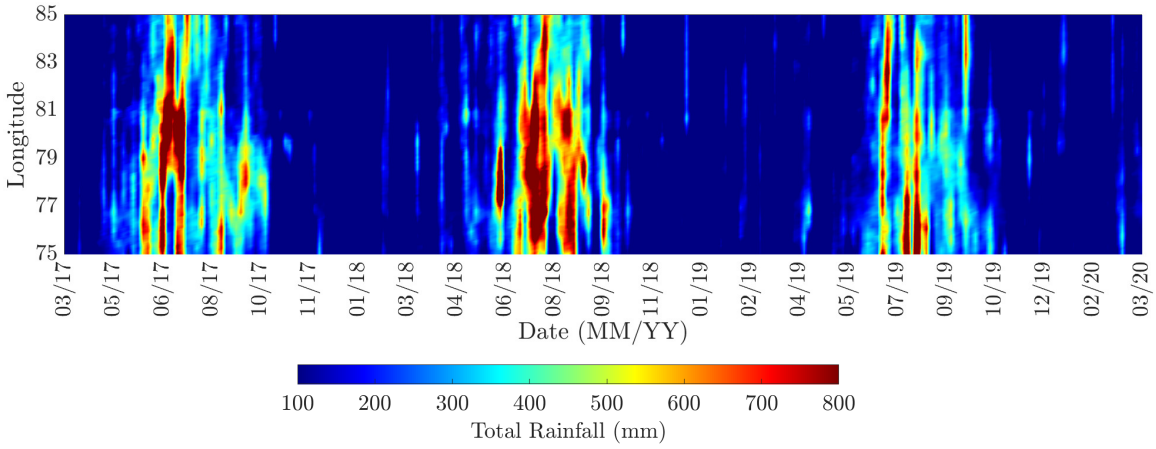


Figure 6.18: Hovmoller diagram of GPM derived rain rates illustrating high rain rates characteristic of monsoon's wet phase and low rain rates characteristic of monsoon's dry phase.

To illustrate this, a 25 km grid is used enabling daily updates of coherence estimates and inundation extent using CYGNSS measurements. The temporal dynamics of the detection process and its correspondence to the different stages of the monsoon, are summarized by the Hovmöller diagram in Figure 6.17 providing a comprehensive view of the start, end, growth and dissipation of the monsoon's dry and wet phases over the three year period from 2017-2020.

The results in Figure 6.17 are associated with the mapping of detection and mapping of three wet phases over the observation period and four dry phases in between. In order to limit inundation area estimates to those corresponding only to the monsoon's wet phase, the mean area associated with the preceding dry phase is subtracted from a given wet phase. The start, end and peak of each of the phases are compared with rain rates reported by the Global Precipitation Mission (GPM) in Figure 6.18 over the same time record. The highly complementary temporal variability of CYGNSS inundation estimates, and reported rain rates are indicative of the success of the ability to map the monsoon's dynamics. The results in Figure 6.17 suggest that the end of the peak of the arid conditions brought about by the monsoon's dry phase was reached on June 2nd, 2017. This was followed by a transitional phase that lasted 3 weeks where the Indian subcontinent experienced a combination of transient flooding and inundation. The transitional dry-wet phase is brought about by a beginning of rain events starting June 6th, 2017. The wet phase during the same year is found to range from June 18th, 2017 to September 23rd, 2017, reaching its peak on August 22nd, 2017. A roughly 30 day lag between peak rain, occurring on July 18th, and peak inundation, occurring on August 22nd, is observed. The variability in the onset of these phases nonetheless is indicative of the need for the near real time

observation. For example, the onset of the following wet phase is found to be delayed by 20 days relative to the previous year starting on July 8th, 2018. The durations of the three wet phases are also found to undergo variability lasting 98 days in 2017, 85 days in 2018 and 91 days in 2019.

From Figure 6.17 it is also noted that detected inundation undergoes gradual propagation over the subcontinent as is evidenced by the lagging onset of comparable levels of inundation along longitude. A further capability the proposed methodology provides relates to the near real time monitoring of the velocity with which the effects of the monsoon's wet phase propagates, where the eastward propagation of these effects may be 'tracked' via the slope of the solid line in magenta. In 2017 the wet phase propagated eastward with a speed 31.35 km/day, with 2018 being associated with a growth in propagation velocity to 54.86 km/day whilst over 2019 this was 42.40 km/day.

Limitations

In spite of the ability of CYGNSS measurements and the proposed methodology to effectively respond to the dynamics of the South Asian monsoon, two major limitations have been identified. The first relates to the limited ability to accurately monitor inundation, that is specifically induced by the monsoon's wet phase, for locations within the subcontinent that are conducive to dominant coherence year-long. More specifically, this relates to 70-90% of Bangladesh's surface which experiences a combination of monsoon floods, river bank floods, rain fed floods and flash floods throughout the year. This is compounded by the prevalence of water logged soil and as a result during both the dry and wet phases of the monsoon, much of Bangladesh is

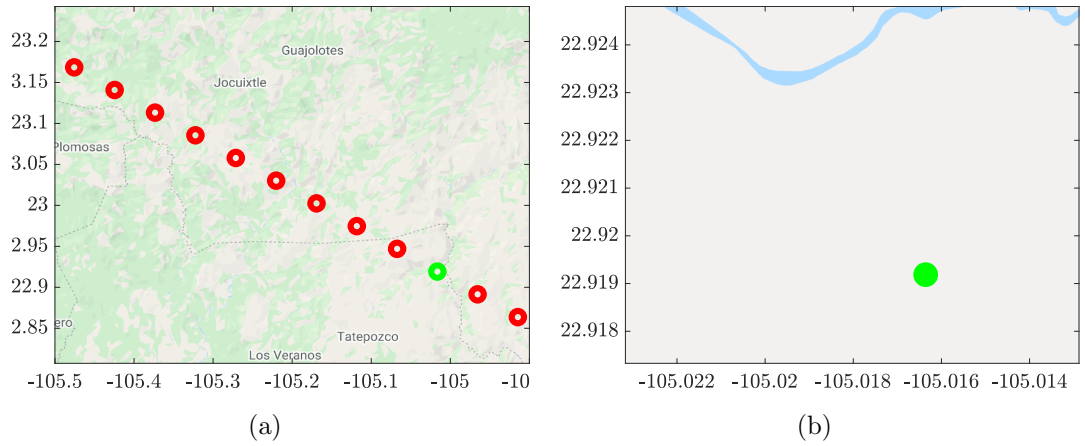
declared coherent. This is further evidenced by the examination of inundation maps based on reported SMAP soil moisture in Figure 6.15.

The second limitation relates to the elevation of some of the areas (Bhutan, for example) affected by the monsoon. During December 2017, an update was made to all eight CYGNSS receivers to expand the specular search window to include the whole of the 128 bin delay extent of the DDMI's Full DDM in support of land investigations. As a result, the maximum surface elevation for which useful measurements are downlinked has been expanded significantly affecting measurements within the whole of Nepal and some parts of India and Pakistan. Nonetheless, the mean 4 km elevation over much of Bhutan renders ≈75% of all specular measurements made beyond the minimum permissible delay and as a consequence much of the downlinked measurements are thermal noise only DDMs. As a result, in spite of the fact that significant impacts of the monsoon's wet phase, in terms of rain rates providing >70% of its total annual rain, mapping its evolution over Bhutan is not possible.

6.4 Estimation of Range Resolution Under Dominant Coherent Reflection

The resolution corresponding to measurements dominated by coherent reflections is typically assumed to be on the order of a first Fresnel zone. However little effort has been invested in attempting to develop techniques with which this may be estimated using CYGNSS measurements. This section outlines how the Level-1 power spread detector may be used to assess CYGNSS's spatial resolution under a dominantly coherent reflection regime. The analysis begins by defining the spatial resolution as an inequality bounded by two limits. The upper relates to the distance travelled over the total integration time of a single Level-1 DDM. For 1 Hz DDMs the upper limit

is expected to be 6 km, taking into account the 6000 m/s velocity of the CYGNSS receivers, and for 2 Hz CYGNSS data the same upper limit is expected to be 3 km. While the upper limits are also validated as part of this investigation, it is the determination of the lower limit of this inequality that is of most interest and is defined as the finest feature a coherent CYGNSS measurement is capable of resolving under a dominantly coherent reflection regime. To assess this, the width corresponding to the finest river the coherence detector identifies (declarations of coherence associated with river crossings) is used as an indirect measure of the measurements' range resolutions due to the expectation that rivers are sufficiently flat to support coherence. An example of the expected response is provided in Figures 6.19-6.20.



(a) (b)

Figure 6.19: River crossing used to estimate surface resolution under coherent reflection regime for 27 meter wide river. Points in red depict incoherent DDMs and points in green depict coherent DDMs (a) Complete CYGNSS track (b) Specular point detected as being coherent

The depictions are for a river in Nayarit, Mexico with an estimated width of ≈27 meters.

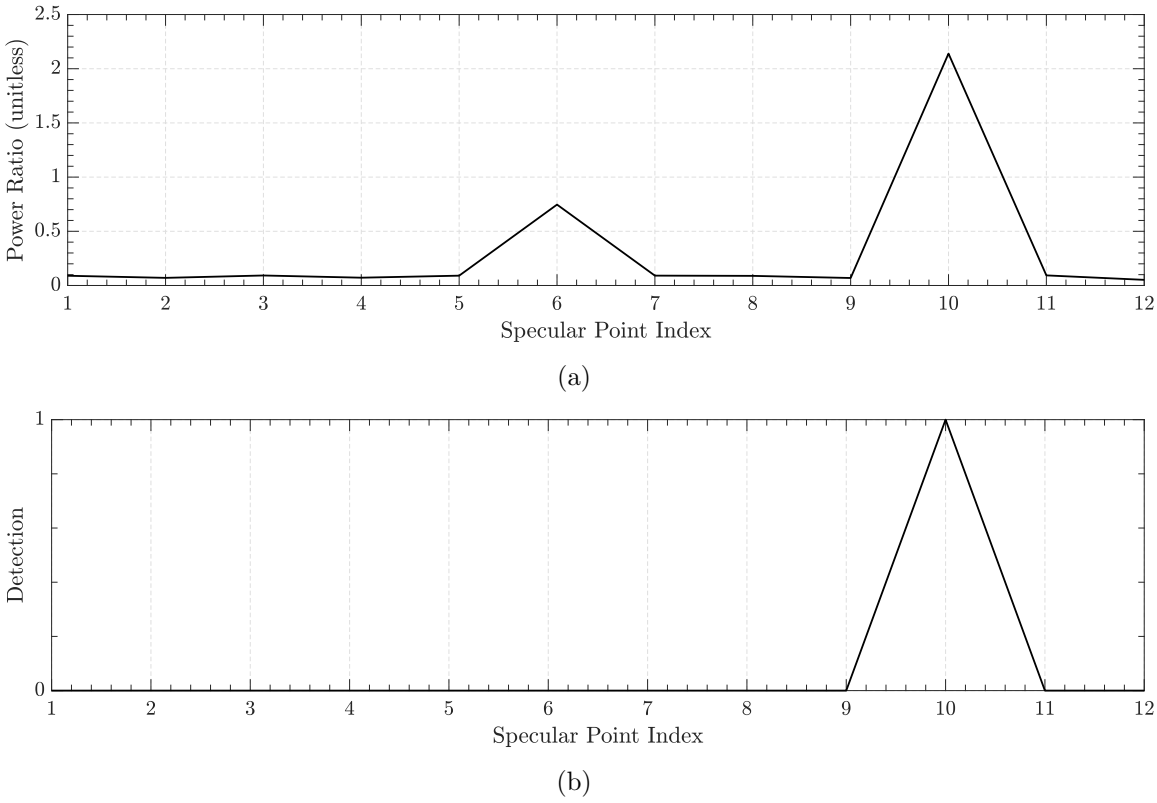


Figure 6.20: Response of Level-1 coherence metrics to river crossing using to estimate surface resolution under coherent reflection regime for 27 meter wide river (a) Power ratio (b) Binary coherence state

Estimates of the Level-1 coherence metric in Figure 6.20(a) suggest that all points within this track have the characteristic behaviors of incoherent DDMs and are therefore with power ratios less than the 2.0 detection threshold, with the exception of the specular point at index 10 corresponding to the point after the river crossing. This particular detection was during the time period when the constellation operated in a 1 Hz mode and therefore the fact that 27 meter wide river was detected suggests that the lower bound of the range resolution under a dominantly coherent reflection

regime is 0.45% of the upper limit, or equivalently the width of the river $\approx$ 27 meters. Nonetheless, detection results derived from a single river crossing cannot be generalized to all CYGNSS measurements as the footprint corresponding to the measurement's resolution is expected to undergo a wide range of variability in accordance with CYGNSS/GPS geometry relative to river orientation, width and surrounding topography. Therefore, to estimate the mean surface resolution a statistical study is undertaken using the North American River Width Data Set (NARWidth) [179]. The dataset provides estimates of tens of thousands of river widths throughout North America thereby providing a reference for the locations expected to produce coherent reflection.

The statistical study is subject to the limitations relating to the fact that river widths are known to change dynamically, seasonally and during flooding events. The NARWidth data set provides an estimate of the river width at mean discharge, thus it does not account for natural seasonal variations. The analysis undertaken here focuses on using CYGNSS data over two eight month periods, the first starting with July 2018 and ending with February 2019 and the second starting with July 2019 and ending with February 2020. This included a total of 50 million measured CYGNSS specular points, within the NARWidth dataset's coverage, providing ample data to generate statistically significant results. While variations in the size of the coherent reflection surface due to natural fluctuations in river widths will contribute errors into the analysis, the mean correspondence of measures of interest to varying river widths is expected to remain indicative of CYGNSS's spatial resolution under conditions of coherent reflection. Furthermore, previous analysis of CYGNSS data has shown that river surfaces are generally smooth enough to result in coherent forward reflections.

However, there will be cases when a river surface is rough relative to the reflecting L-band wavelength (19cm) and will not generate coherent forward reflections. In these circumstances, the area of surface scattering will be significantly larger than an integrated Fresnel zone and not result in (coherent) detection of the river crossing.

Using the available CYGNSS measurements together with reported river widths, the mean correspondence of the Level-1 coherence metric, the power ratio PR, to varying river widths is examined and summarized in Figure 6.21 for specular points within 500 meters of reported river locations.

Approximately 2000 unique river widths existed for points points that are within CYGNSS coverage, and for every day of the analysis period the coherence metric for up to 100 rivers per unique river width was used to examine the ability to detect the rivers with varying widths. The choice of rivers is randomized for every unique river width and for every day of the analysis period. Increasing river widths were found to be associated with a steady increase in the Level-1 coherence metric and the probability of detecting the rivers within the NARWidth data set as being coherent, with the detection probability approaching 80% for 1 Hz data and 100% for 2 Hz, over rivers with widths exceeding $\approx$1.2 km. This is to be expected as rivers with larger widths present a target more conducive to readily identifiable dominant coherence characterized by a significant concentration of power about the DDM peak. They are also less susceptible to ambiguities introduced by the L1 integration time(s) as well as perturbations associated with more heterogeneous scenes. From the analysis presented in Figure 6.21 it is observed that the power ratios exceed the operational 2.0 coherence detection threshold, on average, for rivers with a minimum width exceeding 250 meters and 200 meters for 1 Hz and 2 Hz CYGNSS data respectively. This places

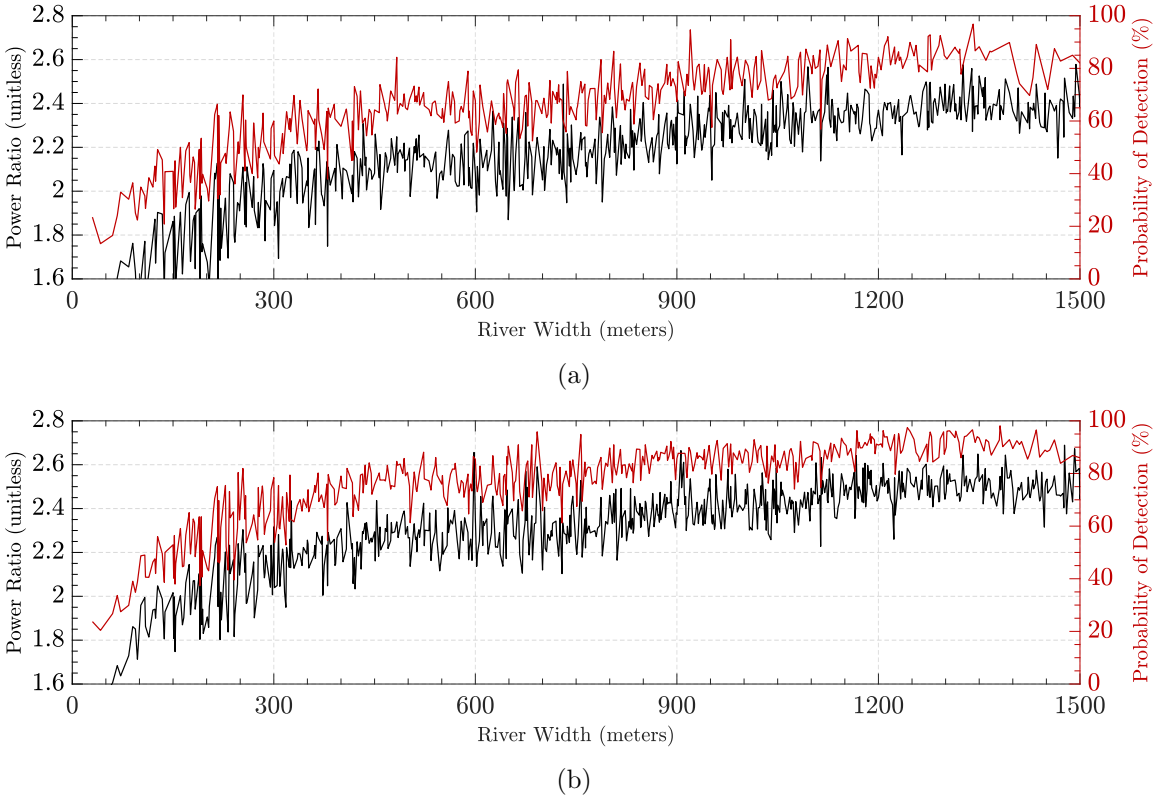


Figure 6.21: Analysis deriving from the NARWidth data set illustrating effects of varied river widths on Level-1 coherence metric and probability of detection (a) Across a sixth month period from 09/01/2018 to 02/28/2019, exclusively comprising 1 Hz Level-1 data (b) Across a sixth month period from 09/01/2019 to 02/28/2020, exclusively comprising 2 Hz Level-1 data

a minimum bound on the finest spatial features CYGNSS measurements are able to resolve, on average.

The upper bound of the constellation's ability to resolve features on the surface, under a coherent reflection regime, is expected to be governed by the incoherent integration time the reduction of which minimizes along track smearing of L1 DDMs thereby improving the overall spatial resolution. To illustrate this, CYGNSS measurements overlooking a 500 meter wide section, on average, of the Courantyne River,

an outline of which is depicted in Figure 6.22, were examined over 8 month periods when the constellation operated exclusively in 1 Hz and 2 Hz modes.

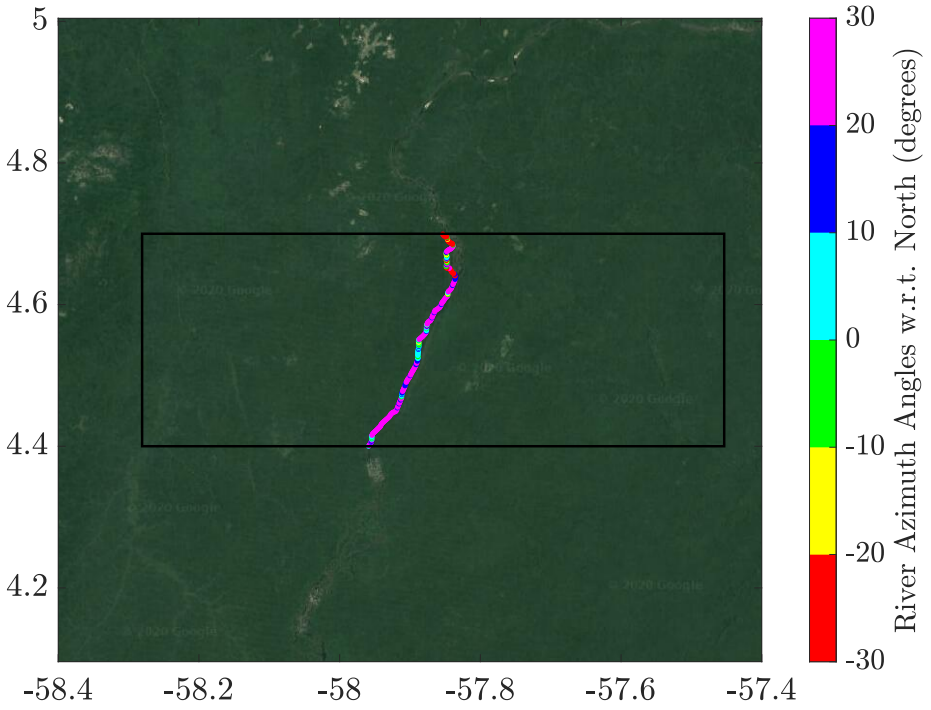
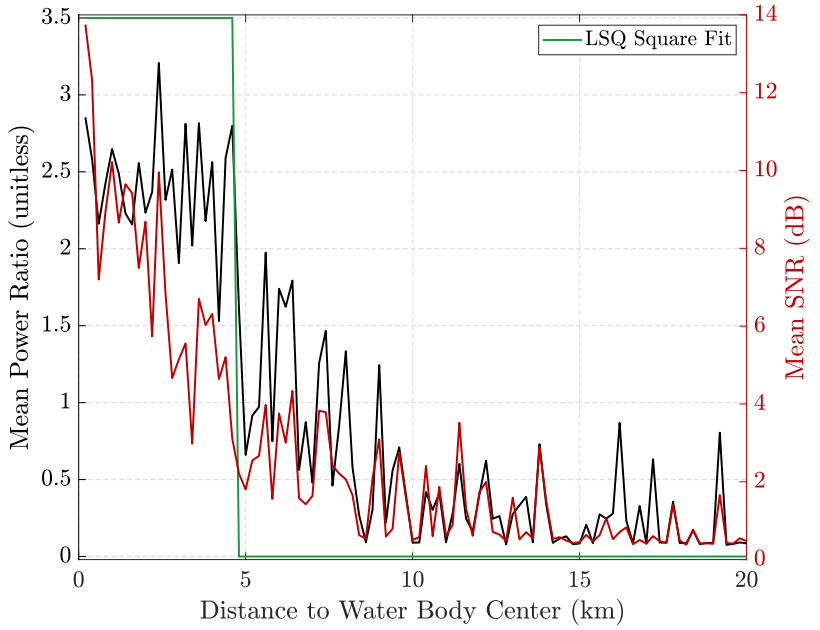


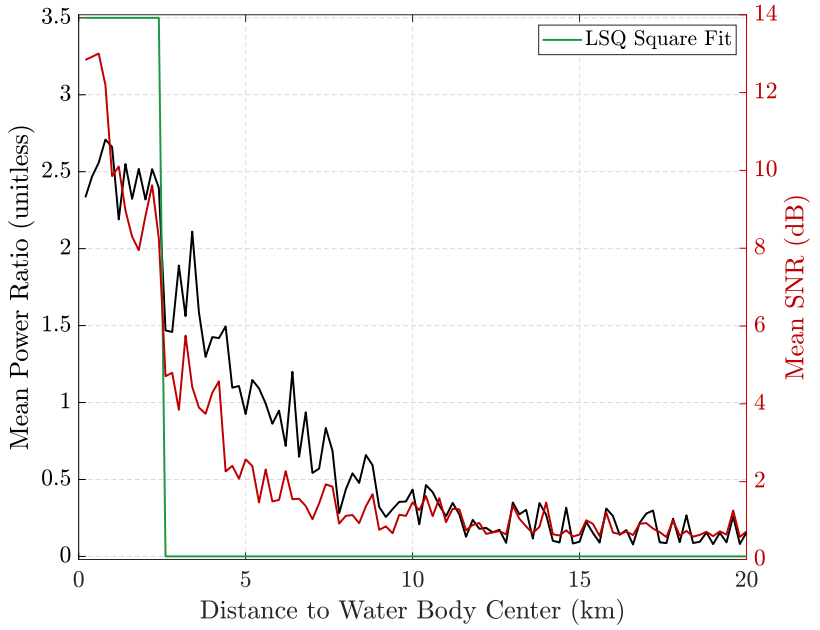
Figure 6.22: Test region centered about a 35 km section of the Courantyne River between Guyana and Suriname

It is observed that the 1 Hz CYGNSS data mean power ratios remain above the detection threshold for distances up to 5 km, based on the Least Squares Fit (LSQ) slightly less than the ≈6km distances travelled during the 1 second integration time. Within this 5 km radius, the PR is maintained above the detection threshold as the result of dominant coherent reflections arising from the 500 meter wide river. The contributions of the the river get 'smeared' along track thereby degrading the resolution of the spatial scales to which the relevant L1 measurements correspond.

This is contrasted with the behaviour of measurements made after the switch to 2 Hz sampling over the same region where it is observed that due to the halving of the incoherent integration time, the along track smearing is reduced to approximately the distance travelled along track between two consecutive specular points of ≈3km. As a result, the power ratios continue to be declared 'dominantly coherent' to 2.4 km away from the river center resulting in an average refinement in spatial resolution of ≈48% brought about by the reduction in integration time.



(a)



(b)

Figure 6.23: Effects of varying the CYGNSS mission's Level-1 product integration time for observations in the vicinity of Courantyne River on Level-1 coherence metric and Signal-to-Noise Ratio at varying distances averaged within 200 meter bins (a) Exclusively comprising 1 Hz Level-1 data (b) Exclusively comprising 2 Hz Level-1 data

6.5 Supporting Near Surface Soil Moisture Estimation Techniques

The sensitivity of the time series retrieval algorithm to the variability of land returns, comprising mixtures of incoherent and coherent DDMs, was outlined in Chapter 4. To adress the limitations this creates, a matched filter approach was introduced aimed at identifying coherent reflections based on the likeness of Level-1 DDMs to WAF templates. This resulted in the introduction of the COR_{max} and SNR_{min} thresholds which were allowed to vary spatially, to accommodate local surface topographic variability, whilst being fixed temporally. The choice of the thresholds was based on an iterative approach that aimed at enabling (or training) the thresholds to achieve optimal performance relative to truth SMAP soil moisture. In this section, the impact of substituting both thresholding techniques with the Level-1 power spread detector is highlighted. The analysis retrieves soil moisture using the proposed time series techniques over previously identified locations over a 27 month period with results that are depicted in Figure 6.24 (the two month gap is the result of a SMAP service interruption). The viability of the substitution is evidenced by the complementary transitions of CYGNSS m_v relative to SMAP m_v throughout the multi-year analysis period. The corrections introduced by the detector are of particular significance as they substitute location varying thresholds, see Figure 4.22, with the globally fixed $\rho_0 = 2.0$ coherence detection threshold over all locations, time instants and grid sizes. Furthermore, it reduces the computational time by 99% as it makes redundant the need to train data refinement thresholds relative to SMAP, thereby bringing about a further benefit of reducing the reliance of the time series approach on independent soil moisture data. It nonetheless, still uses soil moisture bounds derived from SMAP.

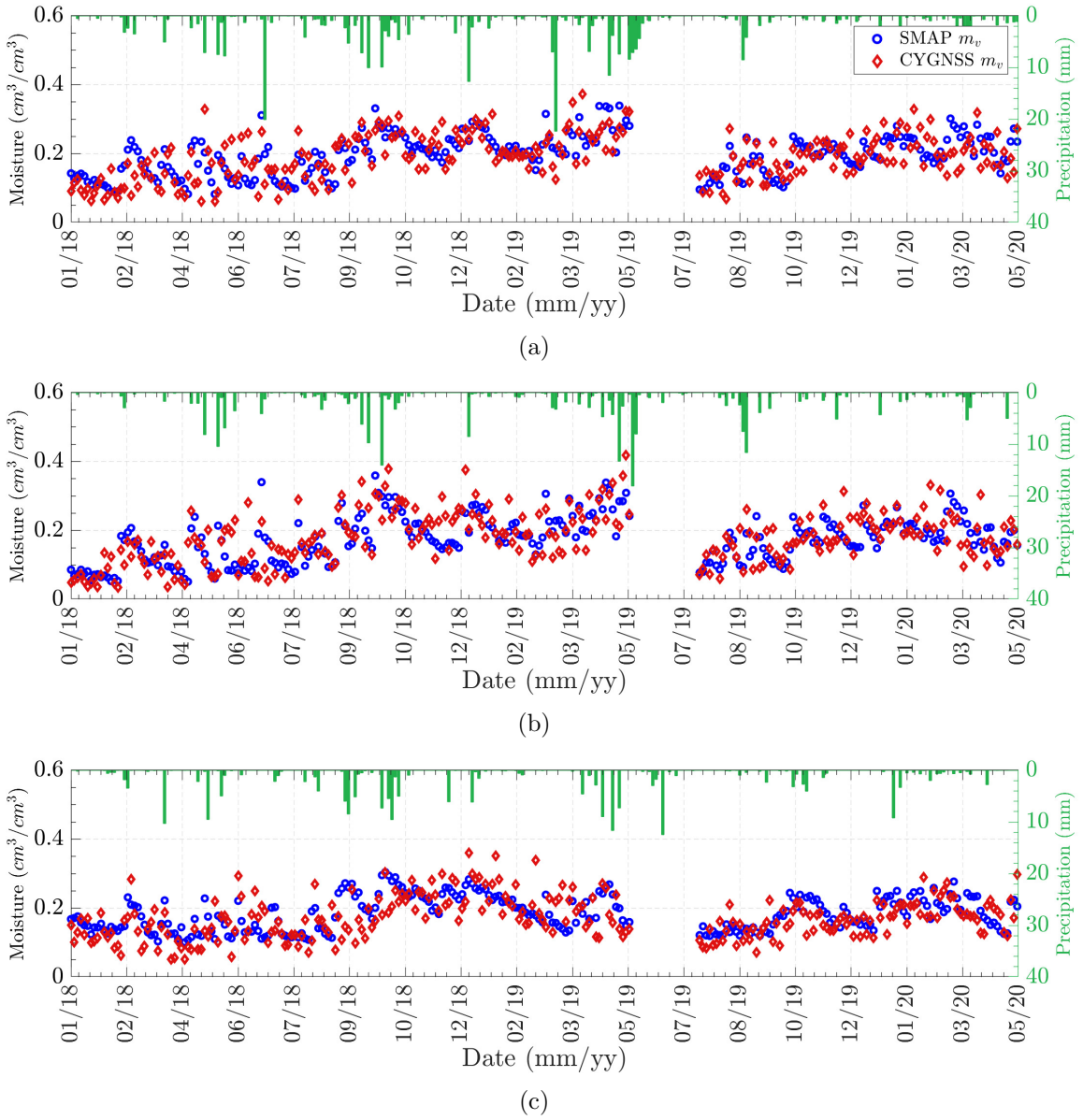


Figure 6.24: Comparisons between SMAP and CYGNSS soil moisture time series for the twelve month period from January 2018 to May 2020 with precipitation data from Global Precipitation Mission (GPM) products after application of Level-1 power spread coherence detector (a) Little Washita, Oklahoma (b) Fort Cobb, Oklahoma (c) TxSON, Texas (d) Walnut Gulch, Arizona (e) Yanco, Australia (f) Little River, Georgia

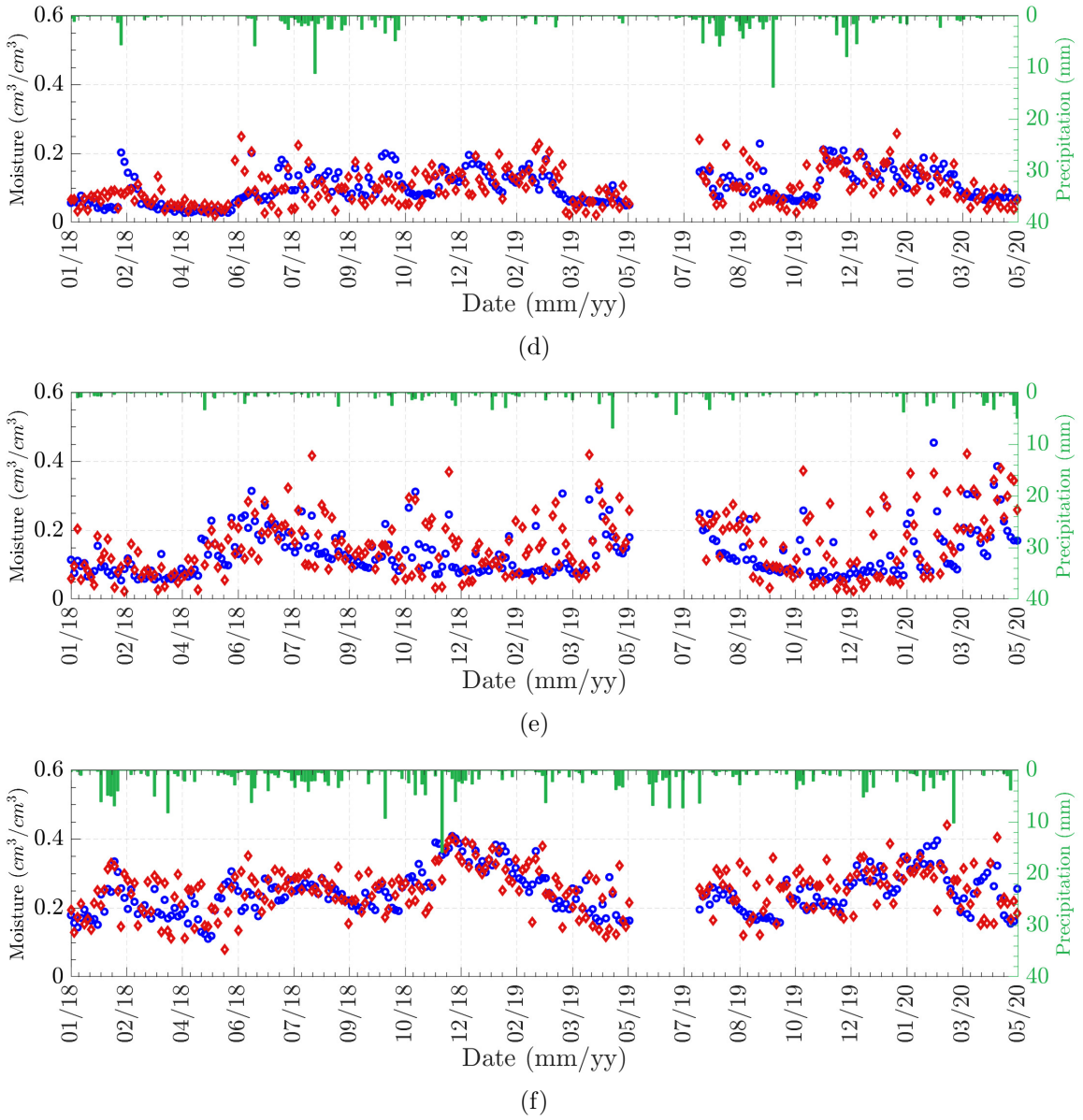


Figure 6.24: Comparisons between SMAP and CYGNSS soil moisture time series for the twelve month period from January 2018 to May 2020 with precipitation data from Global Precipitation Mission (GPM) products after application of Level-1 power spread coherence detector (a) Little Washita, Oklahoma (b) Fort Cobb, Oklahoma (c) TxSON, Texas (d) Walnut Gulch, Arizona (e) Yanco, Australia (f) Little River, Georgia

From Table 6.2 a decline in performance metrics is evident across all sites of interest, relative to those previously reported in Chapter 4, with varied extents depending on site specific surface properties and prevalence of water bodies. Nonetheless, retrieval errors remain within acceptable thresholds with overall performance summarized in Figure 6.25 indicating an overall CYGNSS-SMAP correlation $R = 73.1\%$ and a retrieval RMSE = 5.8%. The inability to reliably retrieve soil moisture using CYGNSS measurements without compensating for the effects of coherent DDMs is not unique to the time series approach and was reported by subsequent investigations [40] attempting to retrieve soil moisture using a geophysical model function. The Level-1 coherence detector is also expected to be instrumental in supporting the operation of additional soil moisture retrieval algorithms.

Table 6.2: Time series retrieval results description and performance metrics after application of Level-1 power spread coherence detector.

Site	(a)	(b)	(c)	(d)	(e)	(f)
R	0.623	0.693	0.563	0.523	0.664	0.539
Bias	0.005	-0.005	0.011	0.004	-0.006	-0.006
URMSE	0.054	0.056	0.052	0.048	0.067	0.054
RMSE	0.054	0.056	0.054	0.048	0.071	0.054

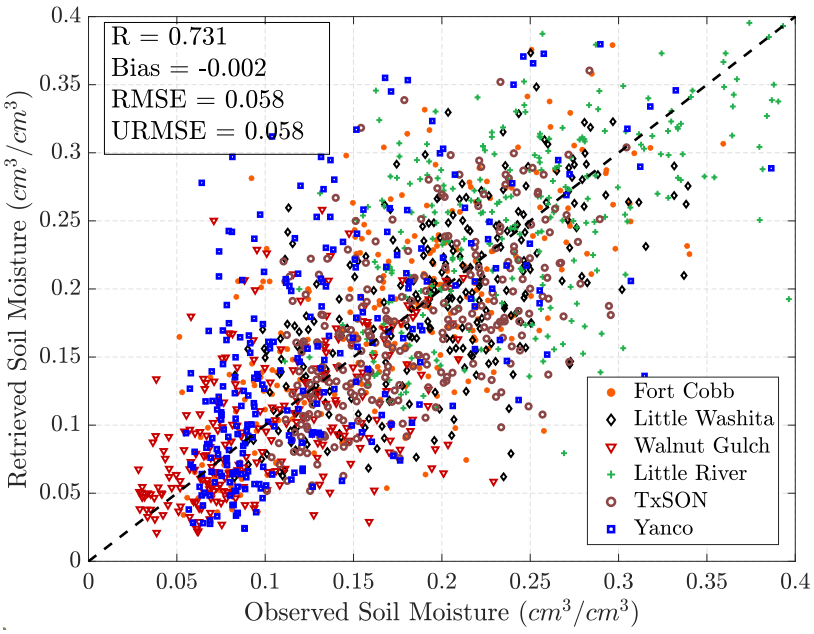


Figure 6.25: Overall retrieval performance after application of L1 coherence detector

6.6 Conclusions

This chapter outlined applications of interest using the Level-1 coherence detector, proposed in Chapter 5, as part of various remote sensing.

This work demonstrates the creation of dynamic inland water body masks at spatial resolutions ranging from 1 to 3 km through the use of a recently developed coherence detector for the delay-Doppler maps produced by the CYGNSS constellation. The use of the coherence of the observed measurements reduces many of the uncertainties associated with previous signal-to-noise ratio based water body detection approaches for CYGNSS. Using data from January 2018 to February 2020 and producing maps representing time intervals ranging from 3 months to 2 years, the water body masks created are found to be associated with a probability of detection that exceeds 80% as compared to the Pekel water mask developed from Landsat

observations. The analysis presented in this work highlights the potential of using spaceborne GNSS-R systems for dynamic inland water body mapping.

The sensitivity of CYGNSS's land measurement to transient flooding and inundation was also used as an indirect indicator of monsoon dynamics. To bypass ambiguities and uncertainties associated with relying on the magnitude of its observables, the function of a Level-1 coherence detection methodology is specialized to mark the start, end and evolution of the monsoon's dry and wet phases. The coherence detection methodology separates DDM measurements in uncalibrated raw counts into the coherent and incoherent categories. Because incoherence is expected to be the dominant mode of scattering over most of the world's locations, its persistence over much of the Indian subcontinent is used as an indicator of the monsoon's dry phase. The transition from a mode of incoherence to coherence is used as an indicator of the beginning of the monsoon's wet phase, as observed surfaces begin to experience transient flooding and inundation brought about by the heavy rain fall typically associated with this phase. Its end, is detected by a second transition to a dominant mode of incoherence. Comparisons with SMAP soil moistures during the dry and wet phases of the monsoon indicate highly complementary spatial distributions, whilst the ability to map the dynamics of the monsoon in near real time with latencies as short as 1 day have been illustrated over CYGNSS's multi-year dataset from 2017-2020 and compared to reference GPM rain rates. The highly complementary nature of the two further illustrates the ability to use spaceborne GNSS-R systems to support, observation and these large scale cyclical events.

An investigation was also performed to quantify the detection resolution of coherently reflected land signals using the NARWidth Data Set of North American river

widths. This analysis resulted in an empirical estimate of 200 m as the minimum robustly detectable river crossing width of the CYGNSS GNSS-R constellation, with occasional detections of river crossings down to 20 m. The utility of the Level-1 coherence detector in support of soil moisture retrievals was also highlighted.

Chapter 7: Track Based Cyclone Sustained Surface Wind Retrievals

7.1 Motivation

Characterizing storm features in general and more specifically the retrieval of their maximum hurricane wind speeds (V_{max}) is directly inline with the core mission objectives of spaceborne GNSS-R systems [48] and was therefore the subject of some investigation [180–184], during the early stages of the CYGNSS mission. These studies have focused on the use of the standard retrieved Level-2 wind speeds [185] from the CYGNSS mission in order to produce a general storm characterization comprising a retrieved V_{max}.

While CYGNSS's wind retrievals are continuing to advance with time, the usefulness of the standard Level-2 retrieved winds for estimating storm features is limited by several issues of practical relevance. This includes the limited spatial extent of the standard Level-1 DDM from which wind speeds are derived, therefore a CYGNSS track will need to be within close proximity to the storm center for this methodology to be effective. This is further complicated by the tendency of the received signal power to decrease monotonically at a slow rate with increasing surface wind speed. An illustration of the peak received power by the CYGNSS space vehicles, based on

simulated orbits, surface conditions and varied surface wind speeds is shown in Figure 7.1.

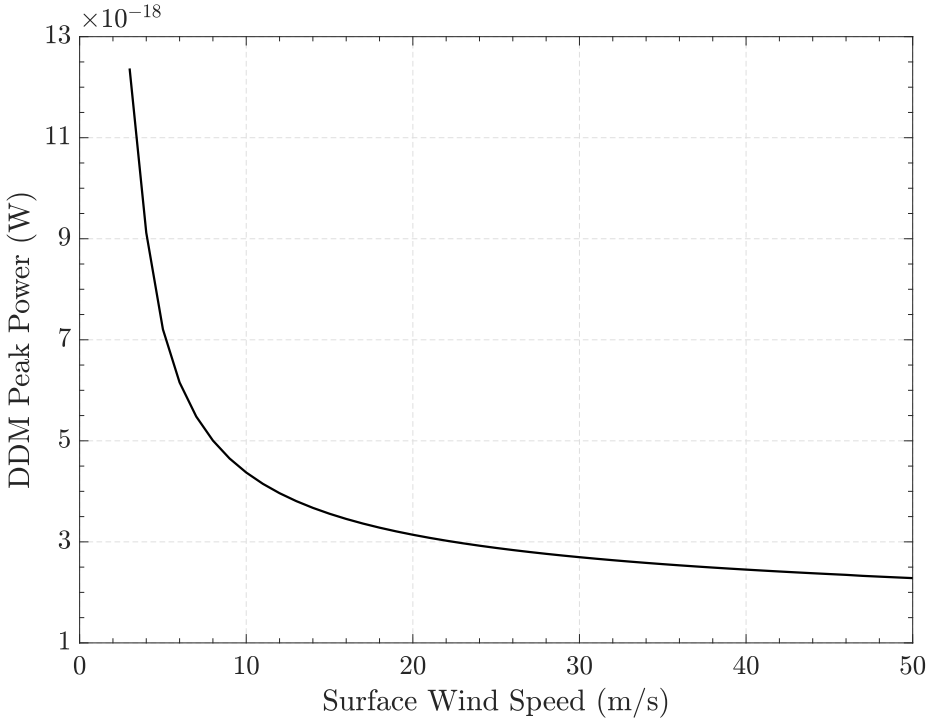


Figure 7.1: Illustration of monotonic decrease of simulated DDM peak power as surface wind speed is increased

At their core, all electrical sensors attempt to relate a measured signal, be it in the form of voltage, current or field, to an equivalent physical quantity of interest by leveraging the correspondence of the two through a linear or polynomial fit. Considering the limiting case where the slope of the line relating the two approaches zero, the measurement is said to saturate. This is because the problem becomes ill conditioned where the measured signal at varying amplitudes corresponds to multiple wind speeds

at which point the standard retrieval process fails to provide unique surface wind estimates. Standard retrieval efforts therefore often fail in obtaining accurate estimates of high winds, which are often associated with the formation and development of tropical cyclones. An example of this is depicted in Figure 7.2 summarizing retrieved L2 wind profiles for a storm during 2017, for which CYGNSS tracks were estimated to have come within 20 km of the cyclone's center. This is overlaid over storm models generated based on 'truth' wind conditions. Ideally, the retrieved profiles across the L2 CYGNSS tracks would 'blend' with the truth background. However, the differences between retrieved and reported profiles are readily identifiable with L2 retrieval errors ranging between 10 and 50 m/s depending on location along the track. These effects are further compounded by the complications of the process of using the Level-1 DDMs to obtain wind speeds, involving a wide range calibration uncertainties, to compute the Normalized Radar Cross Section (NRCS) and the use of a Geophysical Model Function (GMF) to relate NRCS to wind speeds.

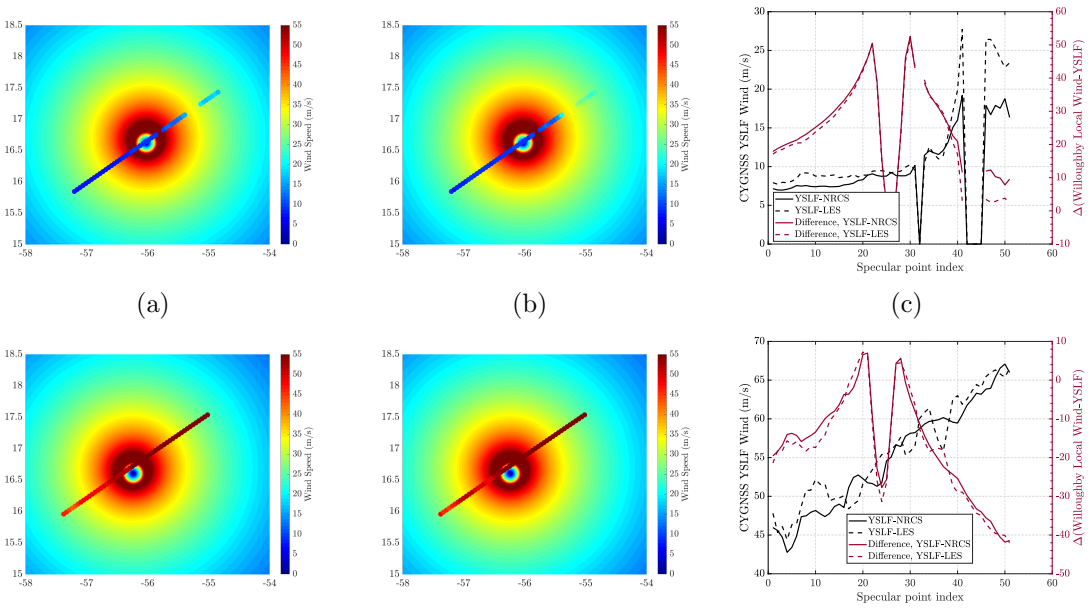


Figure 7.2: (a) CYGNSS YSLF NRCS-based Wind estimate, background is Willoughby model wind with V_{max} = Best Track Estimate on DOY 247 - CYG05 (b) CYGNSS YSLF LES-based Wind estimate, background is Willoughby model wind with V_{max} = Best Track Estimate on DOY 247 - CYG05 (c) Differences for DOY 247 - CYG05 (d) CYGNSS YSLF NRCS-based Wind estimate, background is Willoughby model wind with V_{max} = Best Track Estimate on DOY 247 - CYG02 (e) CYGNSS YSLF LES-based Wind estimate, background is Willoughby model wind with V_{max} = Best Track Estimate on DOY 247 - CYG02 (f) Differences for DOY 247 - CYG02

The work presented herein attempts to bypass limitations associated with absolute power calibration through the adoption of a matched filter approach extending the simulation studies of [183, 184] to CYGNSS. In this work the authors presented results from early simulation studies using modelled CYGNSS returns and reference libraries. The authors then used the one dimensional measurement power-vs-delay (as opposed to the full delay-Doppler map) to conduct retrievals and used no amplitude information, thereby relying exclusively on waveform shape information. Since

forward modelling (simulation of CYGNSS returns) plays a crucial role in this approach, particular emphasis is placed on the End-to-End Simulator (E2ES) described in Chapter 3.

7.2 Matched Filter Retrieval Concept

Several recent works have investigated the utility of CYGNSS measurements for the purposes of retrieving one or more storm parameters. Many such efforts have relied on on the use of the standard retrieved Level-2 wind speeds [181] from the CYGNSS mission in order to produce a storm characterization having parameters such as the storm maximum wind speed V_{max}, measures of storm radii (R_{64}, R_{50} and R_{34}), and the radius of maximum surface wind speed R_{max} through the fitting of retrieved CYGNSS wind tracks to modelled storms [186]. Due to the reliance of these methods on CYGNSS L2 winds, they make use of only the small portion of the DDM used to form the delay-Doppler Map Average (DDMA) [80, 81, 85]. This portion typically extends over 5 Doppler bins (approximately 2 kHz at the CYGNSS 500 Hz Doppler sampling rate) by 3 delay bins (approximately ≈ 0.75 μsec at the CYGNSS 0.25 μsec sampling rate) surrounding the specular return. The surface area represented by the DDMA varies with angle, but is generally $\leq 50\times50$ km. While these efforts continue to show promise, highlighting CYGNSS's potential in improving storm feature characterization, more recent investigations using a similar approach have highlighted challenges caused by the dependence of DDMA retrieved wind speeds on CYGNSS absolute power calibration [75,77,78,116] and its associated uncertainties [186].

The method used in this Chapter retrieves storm V_{max} by matching CYGNSS Full DDM measurements to a library of simulated delay-Doppler maps created for a storm parametric model as the parameters of the storm are varied. The retrieved storm parameter values then correspond to those of the library storms whose DDMs best match a particular Full DDM measurement or track of Full DDM measurements. Both the measured and reference DDMs are normalized by their maximum values before the matching is performed, so that the retrieval approach focuses on the use of DDM "shape" rather than "amplitude". The impact of any uncertainties associated with CYGNSS absolute power calibration on retrieval performance is therefore eliminated.

The retrieval is based on use of the E2ES forward model for CYGNSS returns, which can produce predicted returns for synthetic storm models having varying storm features, in particular the maximum sustained wind speed (V_{max}). A "matched filter" approach is then adopted by comparing predicted returns with those observed throughout the time when a CYGNSS track was within, approximately, $\pm$150 km of the storm center. The retrieval concept departs from the Cauchy-Schwarz inequality, where it follows that maximum correlation between a measured $M(x)$ and simulated $S(x)$ pair of waveforms is maximized when the two only differ by a constant value.

The matching is performed between the predicted $S(\tau, f_D; V_{max})$ and measured $M(\tau, f_D; V_{max})$ DDMs normalized by their rms amplitudes. Multiple metrics for assessing agreement were examined. The first measure finds the storm V_{max} parameter that maximizes the correlation between waveforms over a track of P specular points:

$$V_{max}^{*} = \underset{\mathrm{V}_{max}}{\operatorname{argmax}} \sum_{p=1}^{P} R\left(V_{max}\right) \tag{7.1}$$

where

$$R\left(V_{max}\right) = \frac{\left|\left\langle M_p^{\tau,f_D}(V_{max}), S_p^{\tau,f_D}(V_{max})\right\rangle\right|^2}{\left\langle M_p^{\tau,f_D}(V_{max}), M_p^{\tau,f_D}(V_{max})\right\rangle \left\langle S_p^{\tau,f_D}(V_{max}), S_p^{\tau,f_D}(V_{max})\right\rangle} \tag{7.2}$$

in which the bracket notation refers to a point-wise multiplication of DDM "pixels" followed by a summation of all pixel products.

A secondary measure for assessing agreement based on the root mean square error (RMSE) between $S(\tau, f_D; V_{max})$ and $M(\tau, f_D; V_{max})$ was also examined:

$$V_{max}^* = \underset{\mathrm{V}_{max}}{\operatorname{argmin}} \sum_{p=1}^{P} \sqrt{\frac{1}{N\tau}\frac{1}{N_f}\sum_{i=1}^{N_\tau}\sum_{j=1}^{N_f}\left|M_p(\tau^i, f_D^j; V_{max}) - S_p(\tau^i, f_D^j; V_{max})\right|^2} \tag{7.3}$$

where N_τ and N_f and the total number of delay and Doppler bins in a single DDM respectively. Due to the potential of slight misalignment of the exact locations of the specular bins within the predicted and measured DDMs, the predicted DDM is shifted, relative to the measured DDM, over a limited range in delay and Doppler until the point of best alignment between the two for a given specular point is reached.

The retrieval method can also be extended to retrieve additional parameters beyond maximum wind speed due to the extensive amount of data used in the retrieval process. The retrieval of storm location was also examined, in terms of offsets in the storm center location from the ancillary information in latitude and longitude. Ultimately, a principle factor dictating retrieval value is the ability to describe the sea surface under different conditions as part of the forward modelling process. Different surface wind conditions, produce varied perceived levels of roughness affecting surface mean square slope which directly affects the mapping of received power into different delay-Doppler pairs within reference DDMs. While the absolute power levels within each bin do not affect final retrieved V_{max} values, the normalized relative power distribution across the different delay-Doppler bins affects the shape of the DDM and

therefore can have a significant effect on the retrieval process. Descriptors of the sea surface are discussed in what follows.

7.3 Synthetic Storm Models

The effects of varying levels of surface winds on the sea surface as part of the forward modelling process are not direct, and are captured through variations in surface mean square slope representing surface roughness. The interchange between surface wind and sea surface roughness are made possible through two main techniques. The first entails describing the sea surface using semi-empirical spectra $\Psi(\bar{\kappa})$ describing the sea's roughness scales or surface elevations as a function of wave number κ and surface wind. The variance of slopes, or MSS, can then be derived using (7.4) entailing the integration of spectrum's spectral density up to some cutoff wavenumber κ_*.

$$MSS_{x,y} = \iint_{\kappa<\kappa_*} \kappa_{x,y}^2 \Psi(\bar{\kappa}) d^2\kappa \tag{7.4}$$

Numerous examples of ocean surface spectra exist, of which the Elfouhaily spectrum [187] has been the most widely adopted.

An alternative approach to describe surface MSS is through employing purely empirical relationships in which measured surface MSS is fitted to surface winds. Due to its simplicity, conformity to other empirical models/spectra and computational effectiveness the latter approach is used in this work. The relation between wind and MSS is based on a slightly modified empirically derived relationship in [92, 93] and is given by (7.5-7.7).

$$\sigma_{\parallel}^2 = 0.45 \times \left(\left(3.16 \times 10^3\right) f\right) \tag{7.5}$$

$$\sigma_{\perp}^{2} = 0.45 \times \left(0.003 + \left(1.92 \times 10^{-3}\right) f\right) \tag{7.6}$$

with

$$f = \begin{cases} u_{10}, & u_{10} < 3.49 \\ 6 \times \log\left(u_{10}\right) - 4, & 3.49 \leq u_{10} < 46 \\ 0.13 \times u_{10} + 13, & u_{10} \geq 46 \end{cases} \tag{7.7}$$

where $\sigma_{\parallel}$ is the upwind MSS component, $\sigma_{\perp}$ is the cross wind MSS component, and u_{10} is the wind speed measured at a surface height of 10 m. Total MSS σ_t is given by (7.8).

$$\sigma_t = 2 \cdot \sigma_{\perp} \cdot \sigma_{\parallel} \tag{7.8}$$

The resulting MSS-wind relationship is depicted in Figure 7.3 where the interchange from surface wind to MSS is made possible by describing upwind and cross wind MSS components under three regimes of low winds ($u_{10} < 3.49$), intermediate winds ($3.49 \leq u_{10} < 46$) and high winds ($u_{10} \geq 46$).

The uniqueness of the storm SSW retrieval problem nonetheless derives from the non-uniformity of surface winds. As a CYGNSS track traverses different parts of a storm, each DDM along the track encounters a different region of the storm so that the measurement footprint extends over portions of the sea surface having different winds, thereby furnishing information about the storm structure. The non-uniformity of storm wind fields in space is therefore implicit in CYGNSS measurements, in contrast to the baseline Level-2 retrieval that assumes uniform winds within a footprint that are retrieved independently at each point on a CYGNSS track. For example, as CYGNSS traverses through a storm, the storm center will be present at different delays from the specular point and therefore will be mapped to different pixels in delay-Doppler space

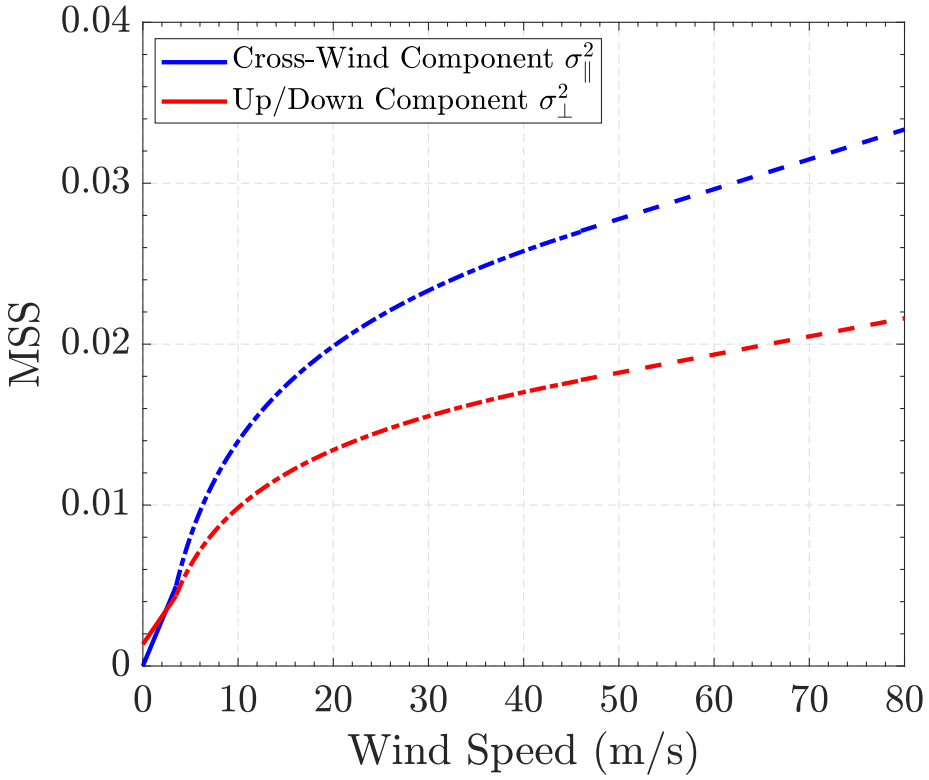


Figure 7.3: Katzberg emperical MSS, surface relationship used to describe surface mean square slope within E2ES as a function of local grid cell surface winds

for different DDMs. The same analogy extended to the mapping of different parts of the storm within the DDM footprint shows that the storm structure will manifest in varying 'shapes' of the DDM. This entails implementing representative wind models on the E2ES surface grids, describing cyclone storm structure and surface winds under different storm conditions. A need therefore arises for synthetic or parametric storm models. Parametric wind fields or synthetic storms refer to a set of model functions that describe wind field distributions within a particular domain. They are 'synthetic' (simulated) representations of actual storms. The E2ES functionality has been extended to incorporate internally generated model storms such that DDM

formation is over non-uniform wind fields with properties dictated by a set of model functions, described in subsequent subsections, and user-definables.

7.3.1 Willoughby Model

A wide range of synthetic storm models is available in the relevant literature and the choice of which model(s) to explore was dictated by the need for computational efficiency of the model implementation whilst ensuring that it describes realistic storm structure. The reported improvements that the Willoughby storm [188,189] has achieved over preceding models in [184] has motivated its use in initial investigations as it facilitates a reasonable compromise between these goals.

The model is a function of storm latitude φ and SSW V_{max}, and has been shown to outperform preceding models such as those in [190] and [191] as shown in [188,189].

The Willoughby model divides descriptions of tropical cyclone surface winds into three domains. The first V_{in} describes winds within the eye wall radius given by:

$$V_{in} = V_{max} \left(\frac{r}{R_{max}} \right)^n \tag{7.9}$$

where r is the radial separation from a surface point to storm center, n is an empirical fit 'power' parameter and R_{max} is the radius of maximum winds. The second domain V_{out} describes surface winds beyond the transitional region, the region between the eye wall and the 'steady state' background wind field, and is given by:.

$$V_{out} = V_{max} \left[(1 - A) e^{-\frac{r - Rmax}{X_1}} + A e^{-\frac{r - Rmax}{X_2}} \right] \tag{7.10}$$

where $A \geq 0$ is a mixing weight factor, X_1 is a decay length constant and X_2 is the outer vortex length of decay. The winds within the third, transitional region, V_{tr} is

in essence a weighted mixture of the winds in the inner and outer domains.

$$V_{tr} = V_{in}(1 - \omega) + \omega V_{out} \tag{7.11}$$

where ω is the weighting mixing factor. The parameters R_{max}, n, X_1, X_2, A and ω are all described by empirical fits expressed as a function of storm V_{max} and storm latitude φ. The Willoughby model is azimuthally symmetric, but extensions may be incorporated to its model functions to introduce some level of asymmetry about its eyewall, using (7.12) which entails incorporating storm translational speed (ground speed in direction of propagation) and translational direction (heading, measured in degrees clockwise from north).

$$V_f = \sqrt{(-V_m \cos\theta_t)^2 + \left[\frac{V_t \sin\theta_t}{R_{max}^2 + r^2}\right]^2} \tag{7.12}$$

where V_f is the final model description of surface wind, V_m is the output wind using the three model functions, θ_t is the storm's translational direction and V_t is its translational speed. Willoughby renditions of hurricanes observed by the CYGNSS constellation during the operation of their receivers under standard and special modes of operations are depicted in Figure 7.4.

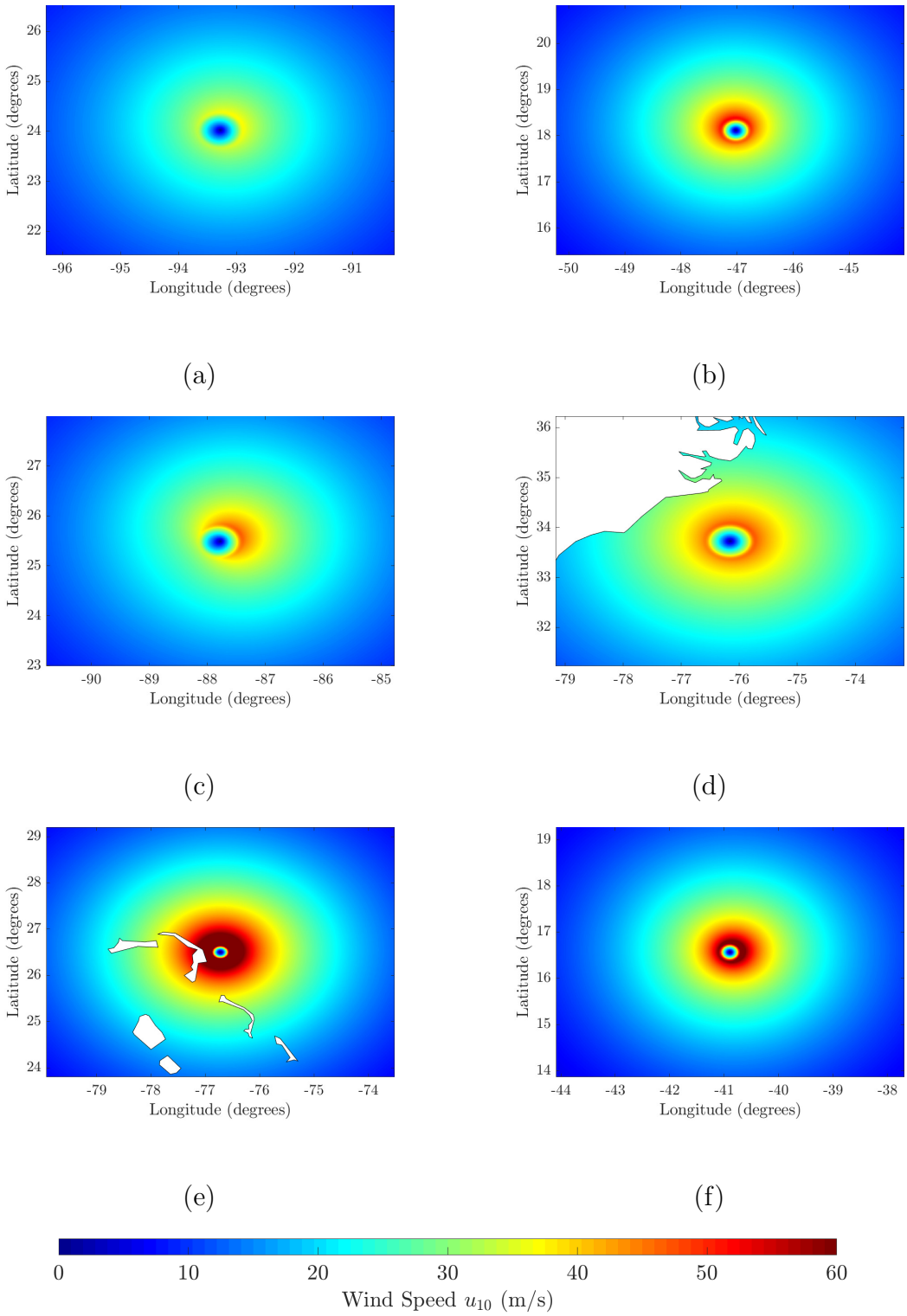


Figure 7.4: Willoughby renditions of storms for which standard and special CYGNSS downlinks have been made available. (a) Harvey on DOY 236, 2017 (b) Irma on DOY 246, 2017 (c) Nate on DOY 280, 2017 (d) Florence on DOY 256, 2018 (e) Dorian on DOY 244, 2019 (f) Lorenzo on DOY 269, 2019

7.3.2 Generalized Asymmetric Holland Model

The E2ES has also been extended to include a second internally generated synthetic storm model, the Generalized Asymmetric Holland Model (GAHM). GAHM is a quadrant specific generalization of the 1980 Holland model in which storm profiles are described by parametric relationships given by rectangular hyperbolas scaled by a shape parameter and location parameter. This allows for the introduction of true asymmetries within the storm profile, in contrast to the Willoughby model. This work will therefore also aim to assess the sensitivity of the retrieval approach to the employed storm model.

The GAHM model divides the storm into 4 quadrants each with 3 isotachs, contours of constants wind values at varying radial separations from the storm center. In total, the model will therefore attempt to solve for surface wind values under 12 different regimes. To do this, at every polar angle θ the parameters Ψ_g and B_g are described by:

$$\Psi_g(\theta) = 1 + \frac{V_{max} R_{max} f}{B\left(V_{max}^2 + V_{max} R_{max} f\right)} \tag{7.13}$$

$$B_g(\theta) = \frac{\left(V_{max}^2 + V_{max} R_{max} f\right) \rho e^{\Psi}}{\Psi\left(P_n - P_c\right)} \tag{7.14}$$

and the surface wind V_g and pressure P_g fields are then given by:

$$V_g(r,\theta) = \sqrt{\left(V_{max}^2 + V_{max} R_{max} f\right) e^{\Psi\left(1 - R_{max}/r\right)^B} + \left(\frac{rf}{2}\right)^2} - \left(\frac{rf}{2}\right) \tag{7.15}$$

$$P_g(r,\theta) = P_c + \left(P_n - P_c\right) e^{-\Psi\left(R_{max}/r\right)^B} \tag{7.16}$$

where B_g is the Holland-B parameter describing overall shape of surface wind contours within a given domain, f is the Coriolis force, P_c is the storm's central

pressure and P_n is its ambient pressure. The coupled nature of the GAHM model functions require a significantly more complicated, and computationally intensive process involving brute-force marching of the parameters Ψ_g and B_g, forcing them to converge to radial conditions dictated by the radius of maximum winds and isotach radii within each of the four quadrants. This also requires, as inputs, additional ancillary information pertaining to storm structure obtained from National Hurricane Center (NHC) Best Track forecasts. The more sophisticated nature of the GAHM storm model adds some level of complexity to the forward modelling process, but provides the advantage of producing significantly more realistic renditions of synthetic storms. Examples of this are shown in Figure 7.5 for a total of six storms for which the CYGNSS constellation provided downlinks under its standard and special modes of operation. The contrasts between this and the Willoughby renditions in Figure 7.4 are readily identifiable where significant distortions and asymmetries, varying from one storm quadrant to another mimicking realistic storm behaviour are observed. The necessity of the GAHM model, and the impact its use has on the retrieval process are explored in more detail in Section 7.6.

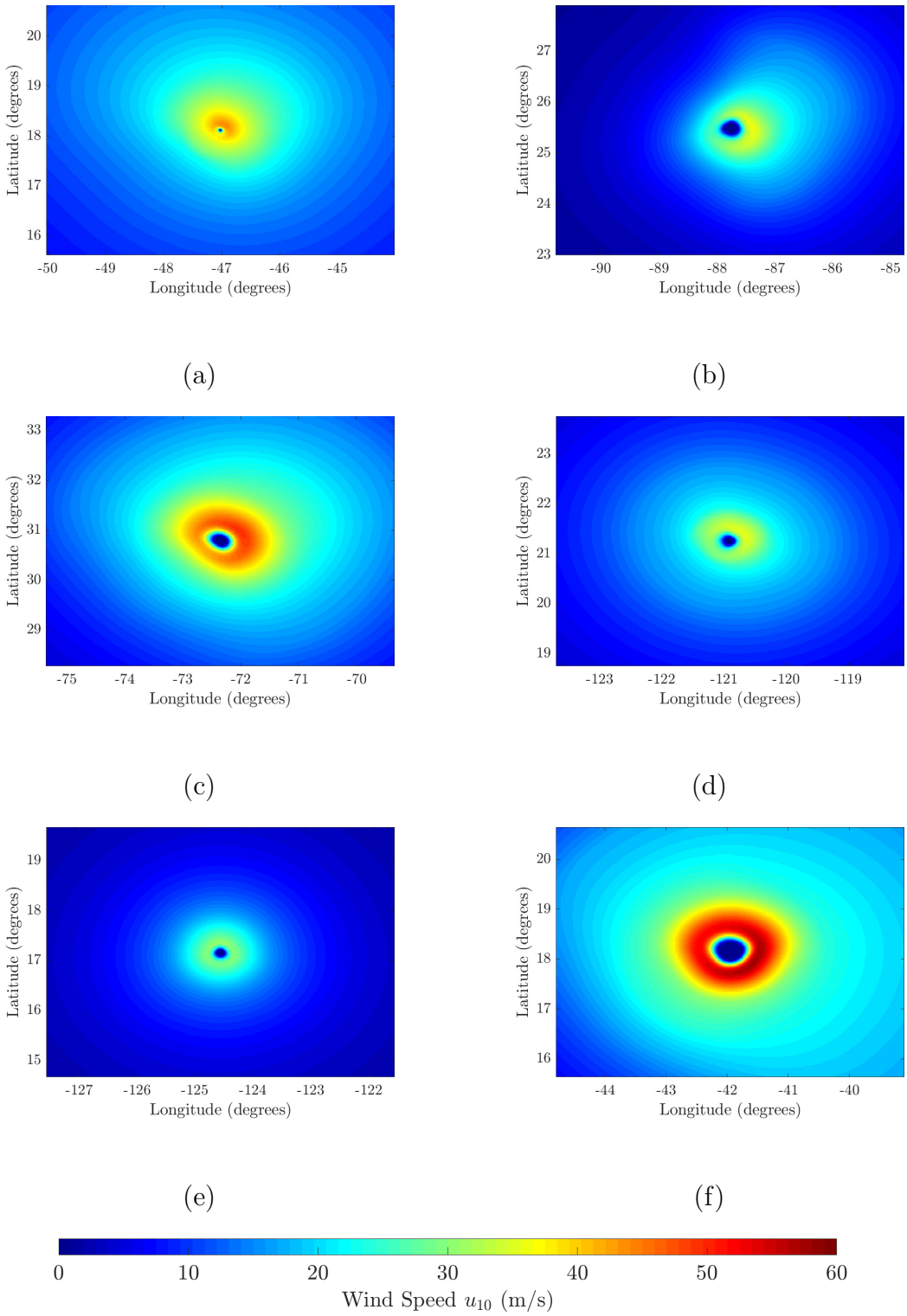


Figure 7.5: GAHM renditions of storms for which standard and special CYGNSS downlinks have been made available. (a) Irma on DOY 246, 2017 (b) Nate on DOY 280, 2017 (c) Florence on DOY 255, 2018 (d) Dorian on DOY 248, 2019 (e) Kiko on DOY 260, 2019 (f) Lorenzo on DOY 270, 2019

7.4 Effect of Surface Winds on Reference DDMs

To assess the impact of including a parametric storm model in the DDM computation, Figure 7.6 illustrates the difference between the DDM predicted using a uniform wind field equal to the wind speed at the specular point and the non-uniform synthetic storm case. The particular geometry considered is shown in the upper portion of the figure. The differences obtained in this case (e.g. strictly positive and strictly negative differences below and above zero Doppler, respectively) are affected by the position of the specular point relative to the storm center. The differences at the individual specular point level are modest, typically on the $\pm 10\%$ levels. However the effect of these differences can be amplified by performing the retrieval over all points on a CYGNSS track within a specified distance of the storm center instead of for a single one second CYGNSS measurement. This approach is adopted in what follows. The small differences obtained between the uniform and non-uniform winds cases makes clear that a highly accurate forward model such as the E2ES is necessary to conduct successful storm parameter retrievals using this method. Initial results to be shown will demonstrate that use of the Willoughby model can produce reasonable retrievals. It is also possible to apply the retrieval methodology with other parametric wind field descriptions such as GAHM. Such efforts have been motivated by some of the limitations of the Willoughby model that will be discussed further in Section 7.6.1.

7.5 Simulation Studies

The non-uniform wind fields produced by the Willoughby model for a range of SSW values (and a-priori known translation speed and direction) are input to the

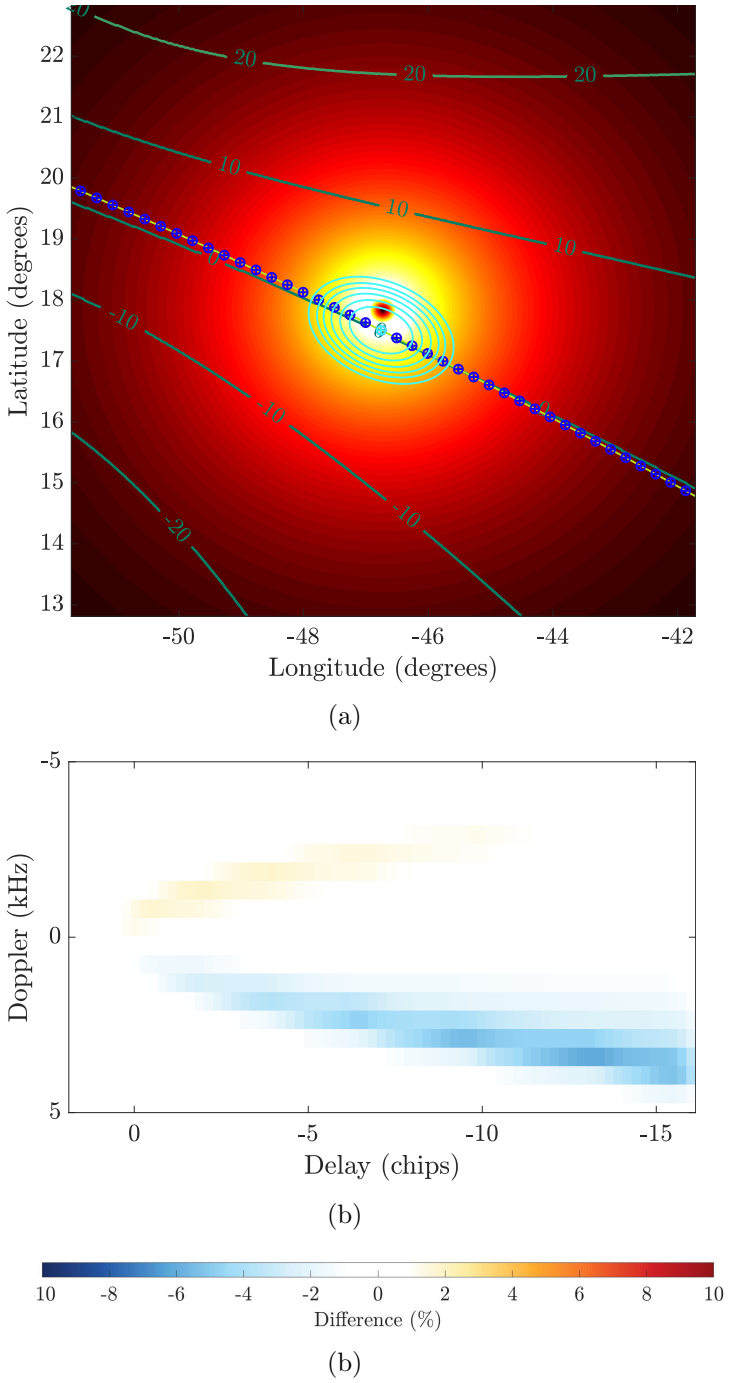


Figure 7.6: (a) CYGNSS track with projection of iso-Doppler (green) lines in kHz and iso-Delay (cyan) ellipsis corresponding to maximum Full DDM spatial extent (b) Illustration of DDM percentage shape difference under the assumptions of Willoughby model based surface wind fields and uniform wind fields

CYGNSS End-to-End (E2ES) simulator, which serves as the forward model. Out of the pools of standard and special acquisition measurements made available by CYGNSS, potentially promising tracks are identified by interpolating storm location(s) in space/time and collocation with CYGNSS tracks using a-priori information based on Best Track and/or ECMWF-EPS data. Using known storm locations reported in these forecasts, storm translational speed and direction are computed. Subsequently, the E2ES accounts for all important aspects of the CYGNSS instrument, including bistatic ocean scattering calculations driven by the varying wind speeds at every pixel within the surface grid, as well as the CYGNSS antenna pattern and other effects.

Prior to applying the method with CYGNSS measured DDMs, a simulation study was conducted to examine the performance obtained. Results from this simulation are depicted in Figure 7.7 a for simulated CYGNSS track at ≈ 20 km from the hurricane Irma storm center. The synthetic storm was produced using a Willoughby model representation of hurricane Irma based on the Best Track data for this storm, which includes the "truth" SSW value for the simulated storm. A total of 51 specular points were then used to simulate measurements of the truth storm that are corrupted by thermal and speckle noise representative of CYGNSS measurements. These corrupted measurements were then compared to a library of reference DDMs (without noise corruption) constructed for each specular point for SSW ranging from $30 \leq V_{max} \leq 80$ (m/s) in 1 m/s steps. The SSW was then retrieved for each of the 51 specular points individually by matching the simulated noise corrupted measurement with the library at the same location, and a final retrieved SSW value was obtained based on the correlation with the library over the entire track of DDMs.

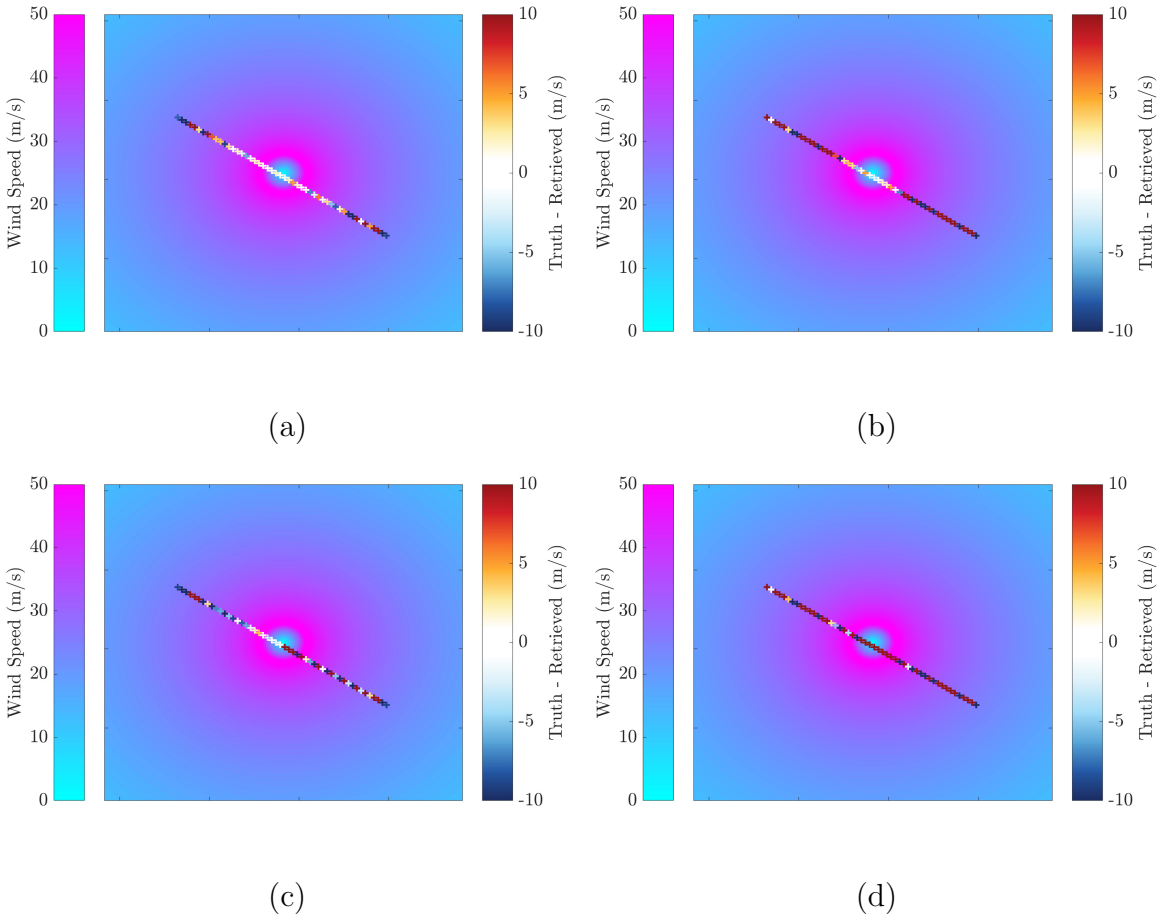


Figure 7.7: Simulation study results for hurricane Irma on September 3rd, 2017 (truth V_{max} is 51 m/s) (a) SSW retrievals for FD CYGNSS track (b) SSW retrievals for SD CYGNSS track (c) SSW retrievals for FD CYGNSS track with 30 km storm center location error (d) SSW retrievals for SD CYGNSS track with 30 km storm center location error

The same process was then repeated using either simulated standard (SD) or Full DDM (FD) products. An SSW estimate is also computed based on overall track correlation with a reference track. Here, for every SSW step the overall track correlation is computed for all specular points along the track. A retrieved SSW is then reported based on the maximum along-track correlation. Figures 7.7(a, b)

illustrate the synthetic storms for the FD and SD cases, respectively, as well as the error in the retrieved V_{max} parameter for the retrieval at each location. Ideally, for every specular point the retrieved V_{max} would be equal to the truth $V_{max} = 51$ m/s. The results also show that the retrieved V_{max} from the individual DDMs is consistently closer to the truth value in the FD case as compared to the SD case. In particular, the FD case obtained error in V_{max} no greater than $\pm$ 3 m/s for specular points as far as $\approx \pm$ 100 km from storm center, whereas in the SD case this performance is limited to a distance of $\approx \pm$ 35 km from storm center. The reduced spatial extent captured by SD likely contributes to this performance difference. A final V_{max} is also retrieved based on overall along-track correlation. In both cases the retrieved V_{max} = 50 m/s from the overall track correlation for "truth" value 51 m/s.

The impact of uncertainties in the location of the storm relative to the CYGNSS track was also examined. Errors in knowledge of the storm location were introduced to the simulation study and the retrieval approach repeated. Figures 7.7(c, d) show the impact of this error when the storm center is shifted 30 km with the corresponding wind field shifted accordingly relative to the CYGNSS track. The track correlation result for the FD product remains $V_{max} = 50$ m/s whereas the SD product obtains an erroneous value of 32 m/s. This initial simulation study highlights the superiority of the Full DDM measurement compared to that corresponding to the standard Level-1 product and the resilience of its final SSW estimates to potential errors in ancillary information. This motivates the reliance on the Full DDM product to retrieve storm V_{max} in what follows.

Other simulation studies conducted as part of this work highlight the potential of CYGNSS based retrievals to correct erroneous storm forecasts. An example of this

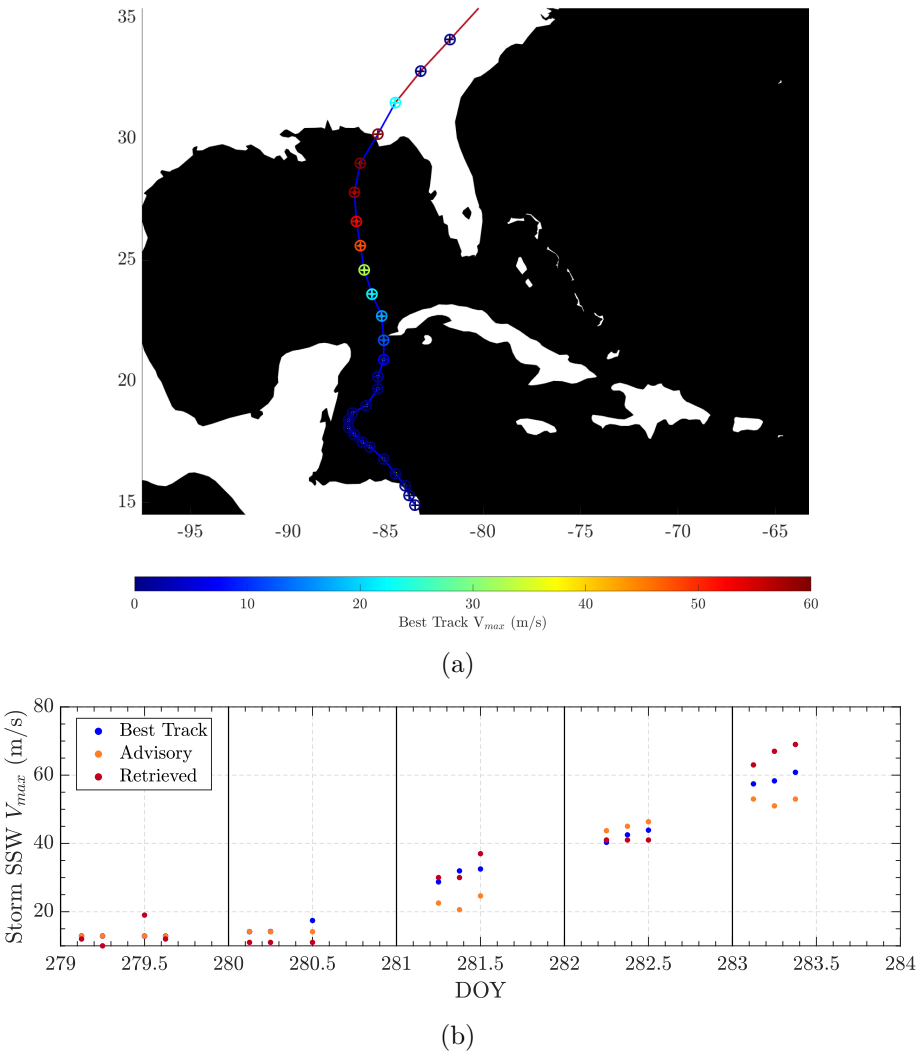


Figure 7.8: Simulation study of V_{max} estimates throughout the life cycle of hurricane Michael from DOY 279-284, 2018 (a) Hurricane Michael track and Best Track V_{max} estimates (b) Time series of V_{max} estimates compared to interpolated Best Track and advisory V_{max} reports

is shown in Figure 7.8 in which both the GAHM and Willoughby models are used to simulate CYGNSS measurements and retrievals throughout hurricane Michael's life cycle. Best Track forecasts routinely undergo updates as prediction models correct their storm structure estimates to account for true variations in storm location, speed,

heading and structure based on satellite observations and reconnaissance flights. Advisory reports issued early in the storm's life cycle often offer erroneous estimates of storm speed and structure compared to final Best Track forecasts. To illustrate the potential for the CYGNSS mission to provide 'measured' CYGNSS observations are modelled based on the GAHM model using erroneous advisories as 'truth'. The simulation of CYGNSS observations is based on actual receiver/transmitter geometries and the occurrence of specular tracks in the storm's vicinity throughout DOYs 279-284, 2018. The retrieval procedure is then repeated and a V_{max} estimate based on maximum 'likeness' between the two is obtained. The results in Figure 7.8 showed that the CYGNSS based estimates were consistently closer to final Best Track estimates, reducing RMSE by roughly 20%.

7.6 Retrievals Using CYGNSS Full DDM Measurements

Retrievals were conducted for a range of storms of interest under both the Willoughby and GAHM model using CYGNSS Full DDM observations. An illustration of these tracks over a single acquisition is illustrated in Figure 7.9 overlooking hurricane Irma during the 2017 Atlantic Hurricane season. The 15 individual tracks each contain more than 500 specular points spanning distances of more than 3000 km; only a portion of the measurements are within sufficient distance of the storm to be useful for retrieval.

To identify the relevant measurements, *a priori* storm location information from National Hurricane Center (NHC) Best Track forecasts [43] is used. This information is updated on 6 hour intervals based on numerical modelling and reconnaissance, and

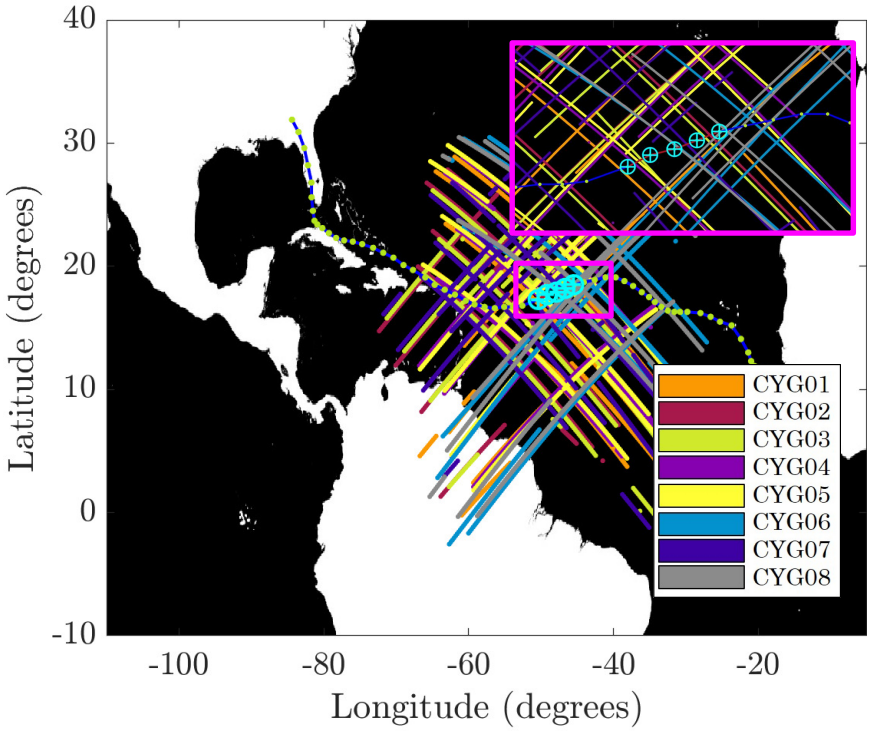


Figure 7.9: Full DDM Coverage of Hurricane Irma on September 3rd, 2017. Storm locations denoted by cyan circles at 00:00:00, 06:00:00, 12:00:00, 18:00:00 and 00:00:00 right to left

the maximum sustained winds from the forecasts are later used in retrieval performance assessment. Using known Best Track storm locations, the translational speed and direction of the storm is computed at the same temporal resolution, based on a WGS84 curved earth, and used as a-priori information for the use as part of the implementation of synthetic storm model on the E2ES surface grid. The storm is also propagated in space assuming a constant translational velocity over 6-hour intervals to obtain approximate knowledge of the storm's location in space and time at 1 second intervals. For each selected Full DDM track, 51 specular points centered on the closest point to the storm's center are used in the retrieval, representing a

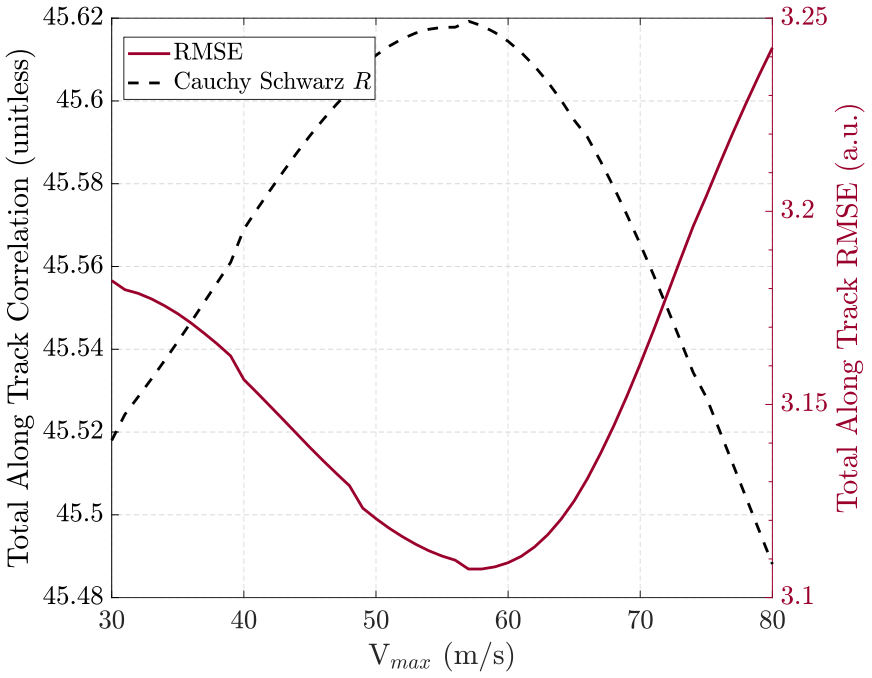


Figure 7.10: Retrieval metric profile across range of storm maximum winds across which reference DDMs are generated

$\approx$ 300 km transect of the storm. The reference library of waveforms is generated for $30 \leq V_{max} \leq 80$ m/s at 1 m/s increments; note this is the expected range over which the storm will have developed from a a tropical depression and/or storm to a Category 1-5 hurricane based on the Saffir-Simpson Hurricane Scale [192]. To enable storm location retrievals, the reference library also includes sweeps over a $\approx 200 \times 200$ km grid sampled every 20 km range of potential storm centers centered about the estimated storm center location at the time of the CYGNSS track. A simultaneous retrieval of storm location and maximum winds is possible by searching for the V_{max} and offset parameters that produce the best fit. However, better performance was found by first retrieving V_{max} and then subsequently performing a second retrieval process to find location offsets using the V_{max} value already determined.

7.6.1 Retrieval Based on The Willoughby Model

Retrievals have been attempted using the Willoughby model for numerous storms for which Full DDM downlinks have been made available. As examples, this section explores those which pertain to Hurricanes Irma, Hagibis, Lorenzo and Harvey. Starting with Hurricane Irma, Figure 7.10 plots the obtained RMSE and correlation values obtained as a function V_{max} for a CYGNSS track through the storm on DOY 246, 2017. The two metrics obtain similar peaks in correlation or minimum in RMSE for $V_{max} \approx 58$ m/s. Other tests also showed the two statistics to provide similar wind speed retrieval results, so an additional approach was adopted of averaging the V_{max} values from the two methods (7.2, 7.3) to obtain a final averaged value.

Table 7.1 summarizes results for the 10 tracks considered with each of the observing CYGNSS space vehicles' (SV) mean along track signal-to-noise ratio (SNR) and incidence angle θ_i listed.

Table 7.1: Retrieval results based on CYGNSS Full DDM observations for Hurricane Irma and use of the Willoughby model as part of the forward modelling procedure. All V_{max} estimates are in m/s

Track #	DOY	CYG SV#	PRN	SNR (dB)	θ_i (°)	V_{max}	BT V_{max}	Δ
1	246	CYG06	27	3.34	48.95	57	49.33	7.67
2	246	CYG08	26	3.08	50.21	59.5	49.40	10.10
3	246	CYG07	27	3.87	44.62	57.5	49.88	7.62
4	247	CYG04	27	2.62	25.93	60.5	52.03	8.47
5	247	CYG05	24	4.07	18.81	60	52.07	7.93
6	247	CYG01	24	3.58	20.76	63	52.15	10.85
7	247	CYG04	27	4.72	34.63	60.5	61.16	-0.66
8	247	CYG05	15	6.59	35.61	62.5	61.20	1.30
9	247	CYG03	27	4.73	32.19	61	61.40	-0.40
10	247	CYG02	15	4.35	36.34	60.5	61.43	-0.93

The results show a tendency of the retrieval method to overestimate the Best Track values with differences generally contained to less than 10 m/s differences. The root mean square error (RMSE) was 6.89 m/s, unbiased-RMSE was 5.20 m/s and mean difference was 4.52 m/s, for an average truth wind speed of approximately 60 m/s. This performance is within the CYGNSS performance goal of 10 percent error for wind speeds greater than 20 m/s, although it is noted that the "track-based" retrieval process used is significantly different from the "snapshot" 1 second standard CYGNSS products for which this requirement applies. A direct comparison with the standard CYGNSS Level 2 (L2) products for 8 of the 10 tracks listed in Table 7.1 was not possible due to the fact that the space vehicles were observing reflections from block IIF PRNs for which the current v2.1 retrieval algorithm does not provide wind speed estimates due to limited knowledge of their variable EIRP levels [75]. For tracks 8 and 10 observing reflections from block IIRM PRNs, the maximum along track CYGNSS L2 Young Seas Limited Fetch (YSLF) Normalized Radar Cross Section (NRCS) based estimates were 19.2 and 67.0 m/s, respectively, resulting in a difference of -42.0 and 5.57 m/s difference relative to NHC estimates.

Beyond retrieval errors, two other considerations are of particular interest. The first relates to the need for consistency in retrievals over periods for which a given storm is not expected to have evolved, developed or dissipated in any substantial manner. An opportunity to examine the consistency of the retrieval methodology is afforded by observations of the same storm within relatively short time spans. While the occurrence of such scenarios is uncommon, Full DDM downlinks made for tracks observing Typhoon Hagibis made this possible.

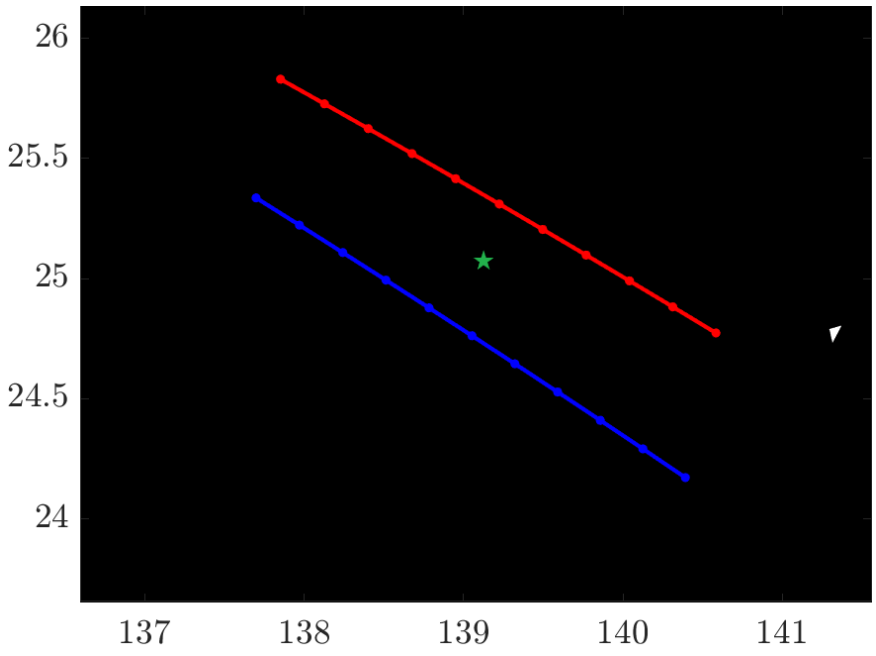


Figure 7.11: Near co-located tracks observing Typhoon Hagibis on DOY 283, 2019. Green pentagram is the interpolated storm center estimate at the time of CYGNSS tracks.

Typhoon Hagabis was amongst the most devastating pacific storms on record, starting as a tropical depression in the pacific ocean and developing into a category 5 super typhoon inflicting significant damage to Guam, the Mariana Islands and Japan. Numerous downlinks have been made for tracks observing Typhoon Hagibis but those made on DOY 283-2019 depicted in Figure 7.11 were of most interest. The two were made in tandem, approximately 20 minutes apart, both roughly 20 km on opposite sides of the storm center estimate. A test of retrieval consistency in this case requires that the two retrievals have very small or zero differences. The retrieved storm profiles using the Willoughby model for both tracks is depicted in Figure 7.12.

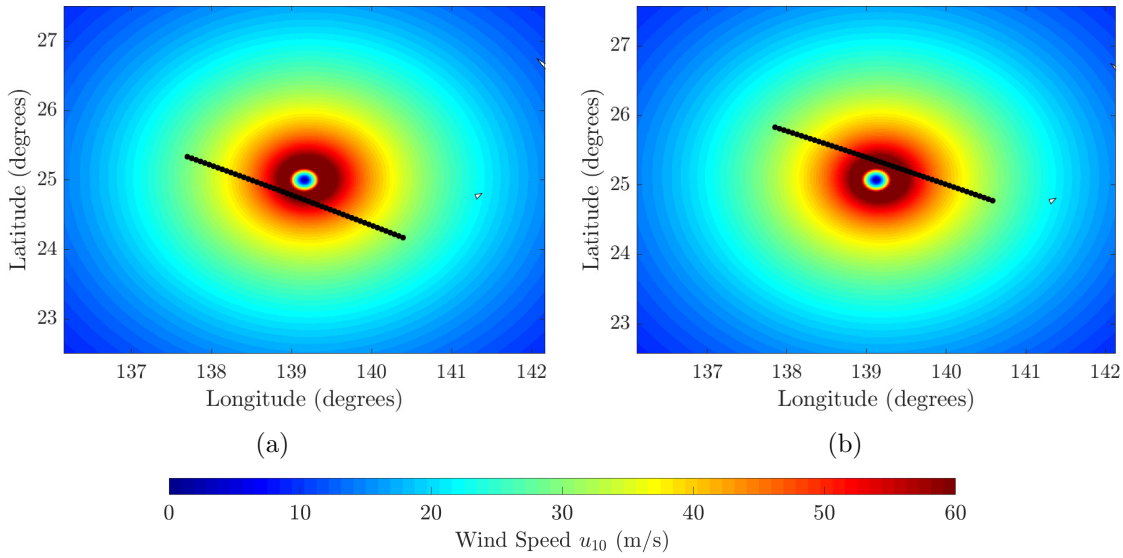


Figure 7.12: Retrieval results for two co-located tracks overlooking Typhoon Hagabis on DOY 283, 2019 (a) Willoughby rendition of storm based on track 1 with V_{max} = 70 m/s retrieval value (b) Willoughby rendition of storm based on track 2 with V_{max} = 70 m/s retrieval value

Both tracks were associated with an identical retrieval value of V_{max} = 70 m/s; relative to the reported 71.5 m/s truth SSW value both results correspond to an error of 1.5 m/s or equivalently 2.14% relative to storm SSW. As a result, the retrieval methodology for this case study passes both the consistency and low error (Δ <10%) requirements.

The second consideration relates to the responsiveness of both the CYGNSS measurements and the retrieval methodology to various levels of developments in storm structure. Due to the infrequent availability of Full DDM measurements, it is often the case that retrievals capture a single snapshot of the storm's behaviour during a single instant of the storm's life cycle offering little insight to its growth and decay. Full DDM downlinks for Hurricane Lorenzo however were made available over four

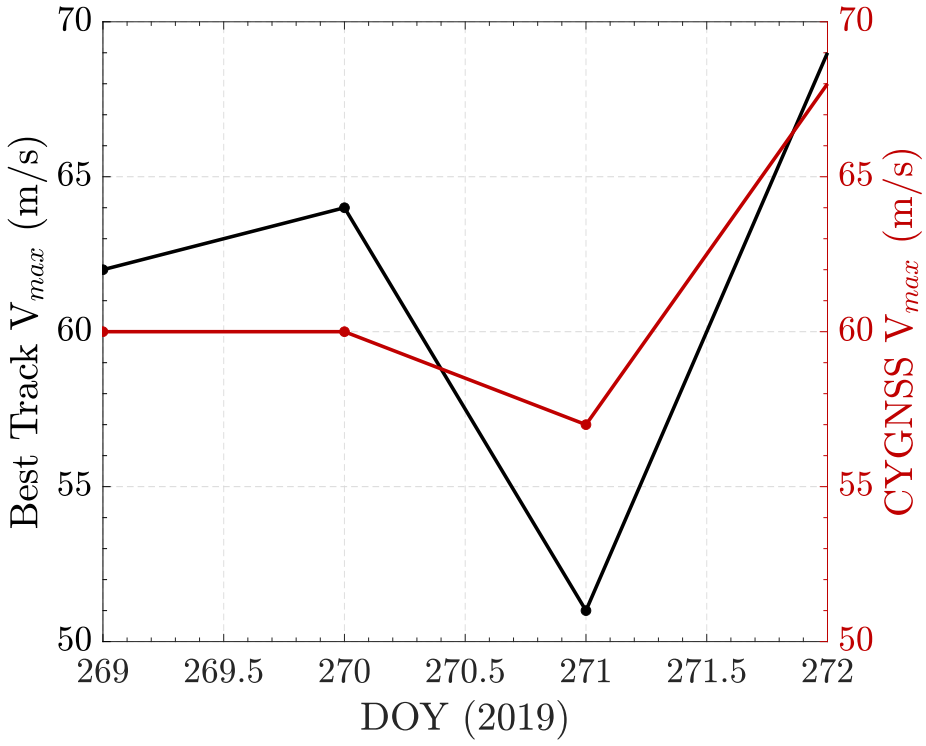


Figure 7.13: Retrieval results across four days of Hurricane Lorenzo's life cycle

consecutive days during 2019. The results of retrieval attempts, compared to truth Best Track data are depicted in Figure 7.13 where it can be observed that the storm exhibited the relatively uncommon behaviour, of growth, decay followed by another growth phase. A response to these tendencies is clearly implicitly captured within the Full DDM measurements and successfully translated through the matched filter retrieval methodology to successful SSW retrievals where over the four day period mean retrieval error was $\Delta_{mean} = 3.25$ m/s.

While the results of the retrieval process using the Willoughby model indicates satisfactory performance, tests of retrieval errors associated with storms with lower levels of development showed significantly higher error levels. An example of this is

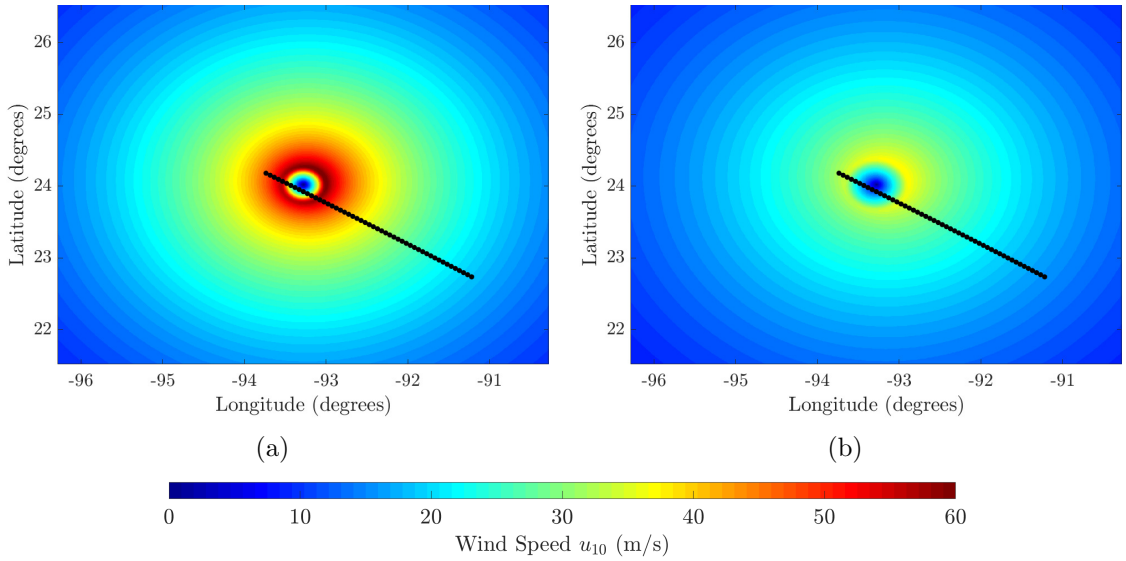


Figure 7.14: Example of erronious storm V_{max} retrieval using Willoughby model. Black scatter provides an outline of CYGNSS specular point track. (a) Retrieved Willoughby storm profile based on $V_{max} = 59$ m/s (b) Willoughby storm profile based on truth $V_{max} = 35$ m/s

shown in Figure 7.14 for an attempted retrieval for a Full DDM track overlooking hurricane Harvey on DOY 236, 2017. The retrieved storm profile in plot (a) was associated with a significant overestimation of storm V_{max} approximately 20 m/s higher than the reported truth value for which the Willoughby storm profile is depicted in plot (b).

The significant error levels are expected to be in large attributable to the inability of the Willoughby model to describe storm structure with lower levels of development with sufficient accuracy. The results depicted in Figure 7.14 clearly highlight the importance of the parametric storm model providing an accurate representation of the storm structure.

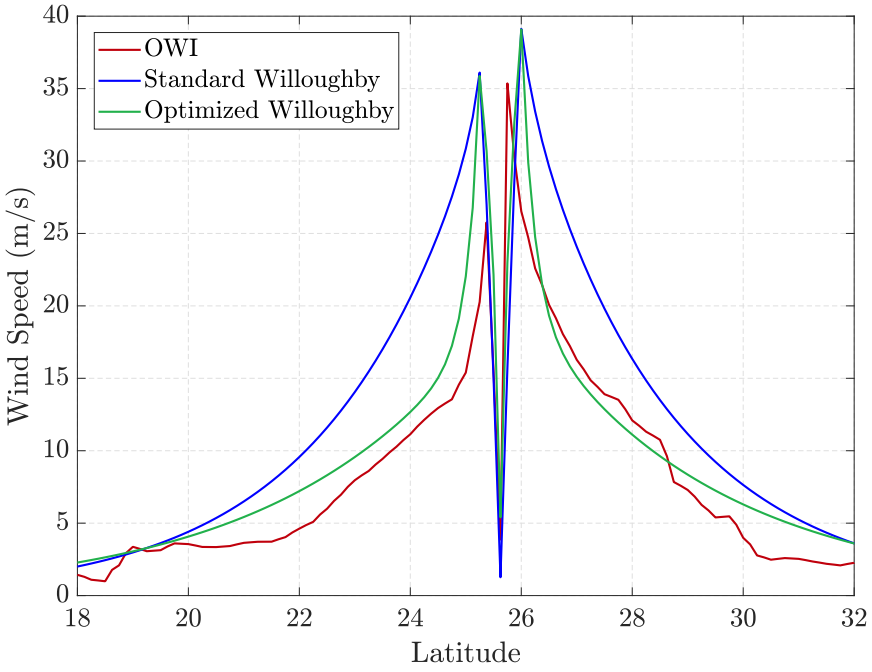


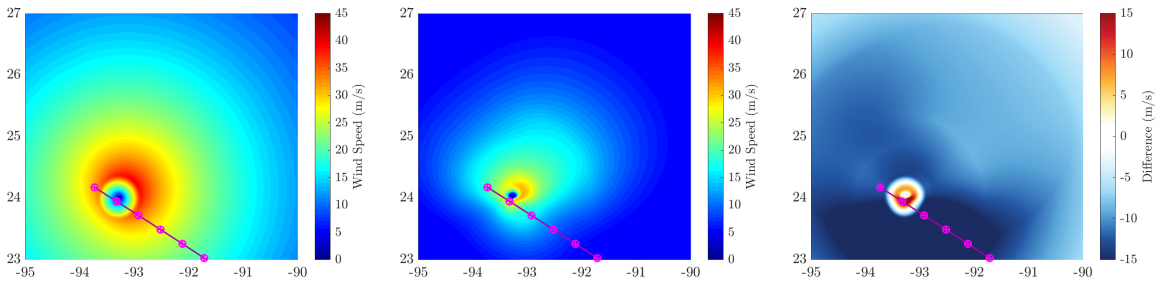
Figure 7.15: Comparison of radial profile of Willoughby storm against reanalysis product

While the Willoughby model is known to outperform some descriptions, it is also known to have a tendency to overestimate surface wind speeds. As an example, Figure 7.15 compares the Willoughby model radial profile of wind speed with wind fields from Ocean Weather Inc. (OWI) [193]. A consistent overestimation, at times exceeding 15 m/s, is evident.

Further, while azimuthal asymmetries are included in the model used through the incorporation of translational speed/direction, the changes do not introduce large quadrant specific variations across the various storm radii (R_{64}, R_{50} and R_{34}). This motivates exploring the combined use of the Willoughby model with more sophisticated models like GAHM to account for these effects. This is explored in the next section.

7.6.2 Retrieval Based on The GAHM Model

Consider the case study for a track observing hurricane Harvey during 2017 shown in Figure 7.16.



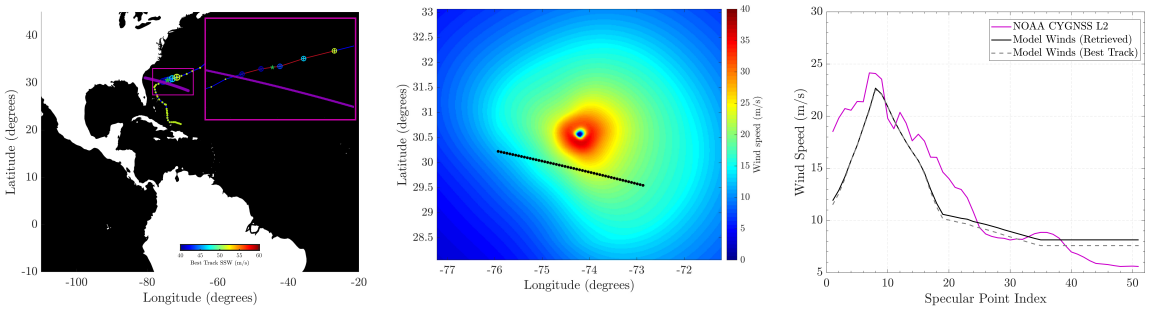
(a) (b) (c)

Figure 7.16: Illustration of different synthetic storm models for hurricane Harvey on DOY-Year 236-2017, CYGNSS track is in magenta (a) Representation of hurricane Harvey using Willoughby model (b) Representation of hurricane Harvey using GAHM model (c) Difference in surface wind distributions

Variations in storm structure are readily identifiable, where it can be observed that the GAHM model is significantly more 'contained' whereas Willoughby suggests noticeably higher wind speeds at distances farther than 50 km from the storm center. Further, through ingesting Best Track forecast information pertaining to storm radii as part of its implementation, the GAHM model is significantly more asymmetric compared to the Willoughby model and as a consequence results in significant changes in wind field distributions. Comparisons against Ocean Weather Inc [193] wind fields suggests that the GAHM model is more accurate in this case. Depending on location within the grid, local wind variations between the models exceed 20 m/s. As a result,

variations exceeding 15% based on implementations of the two storm models were noted within the 'shape' of the reference DDMs. Retrieval errors were subsequently found to range between 5 m/s using GAHM and 25 m/s using Willoughby.

The success of the GAHM model in describing storm structure with low levels of development, and the implications this has in terms of retrievals with low errors, is not unique to hurricane Harvey, and has been observed repeatedly throughout numerous attempted retrievals. For example, in Figure 7.17 the retrieval approach is conducted for a track of CYGNSS FDs during the 2019 Atlantic hurricane season overlooking hurricane Humberto.



(a) (b) (c)

Figure 7.17: (a) Storm CYGNSS track on DOY 260, 2019 (b) Synthetic storm based on GAHM Model and retrieved storm V_{max} using CYGNSS Full DDMs (c) Local wind profile

During this time, Humberto was a category 2 hurricane with an interpolated Best Track V_{max} of ≈43 m/s. The retrieved V_{max} based on CYGNSS measurements was 40.5 m/s (an average of 40 m/s based on maximum R and 41 m/s based on minimum RMSE) suggesting an error of ≈2.5 m/s. Using the retrieved V_{max}, coupled with

a-priori Best Track data a complete storm profile was then constructed as shown in plot (b) using the Generalized Asymmetric Holland Model (GAHM) [194], which was used in this case as the synthetic storm model in the forward modelling of CYGNSS' observations. As an additional step, the wind field grid in plot (b) can be sampled at the CYGNSS specular point locations to produce a local, along track profile of wind speeds that are compared against a version of CYGNSS L2 winds, produced by the National Oceanic and Atmospheric Administration (NOAA), that underwent track-wise de-biasing. The de-biasing process includes incorporating correction factors across a contiguous track to minimize the impact of calibration uncertainty on retrieved L2 surface wind estimates, based on incorporation of significant wave height (SWH) estimates and comparisons against reference model surface winds. The modest error levels with respect to the retrieved V_{max} together with the large extent of similarity between the two profiles are evidence of the success of the retrieval process where across the 'retrieved' and the L2 winds profiles RMSE was 2.73 m/s, URMSE was 2.45 m/s, bias was 1.16 m/s. A high degree of correlation estimated at 92.7% is also noted. It is also shown that due to the minimum track to storm center separation of 75 km in this case, the 2.5 m/s difference in retrieved vs Best Track V_{max} difference did not result in appreciable variations in the local wind profiles across the two as shown in plot (c). Results from other retrieval attempts based on the use of the GAHM model for storms with low levels of development are summarized in Table 7.2.

Table 7.2: Retrieval results based on CYGNSS Full DDM observations for storms with low levels of development and use of the GAHM model as part of the forward modelling procedure.

Track	DOY	Year	Storm	Storm Category	$\lvert\Delta(\text{CYGNSS} - \text{Best Track})\rvert$
1	237	2017	Harvey	2	2.84
2	280	2017	Nate	3	3.08
3	280	2017	Nate	3	3.08
3	255	2018	Florence	3	3.74
4	242	2019	Dorian	3	3.52
5	269	2019	Lorenzo	1	2.25

The retrieval successes using the GAHM model for storms with low levels of development and Willoughby for storms with higher levels of development motivates their use in as part of a combined V_{max} estimation strategy. Tests across a wide range of Full DDM downlinks suggests that optimal performance is achieved when the GAHM model is used for storms with levels of development D_s ranging between $1 \leq D_s \leq 3$ and using the Willoughby model within the range $4 \leq D_s \leq 5$.

The results of this strategy are depicted in Figure 7.18 for a total of 43 Full DDM retrievals spanning 11 storms over the 17-19 hurricane seasons where mean retrieval difference $\Delta_{mean} = 5.04$ m/s, RMSE = 6.50 m/s, URMSE = 6.42 m/s, bias is 1.06 and overall correlation $R = 89\%$. Relative to storm V_{max}, retrieval error was found to be on average 9.79% which is within the 10% mission requirement, both highlighting the success of the retrieval methodology and emphasizing the potential role of GNSS-R systems to provide accurate estimates of storm structure and its evolution.

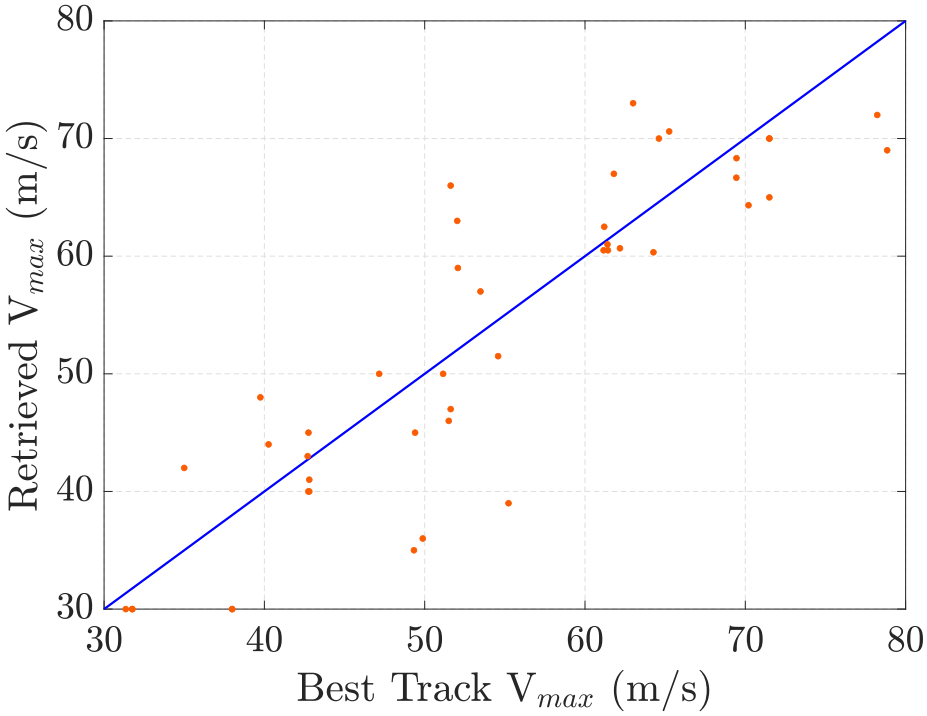


Figure 7.18: Scatter plot of retrieved CYGNSS based storm V_{max} results for 43 Full DDM tracks across 11 storms over the 17-19 hurricane seasons based on the combined use of the GAHM and Willoughby models

It is nonetheless emphasized, that the superiority of the GAHM model in providing more realistic descriptions of storm structure is facilitated only by the ingestion of a significantly increased number of ancillary parameters relating to storm structure. The number of ancillary parameters that GAHM requires is contrasted against those required by Willoughby in Table 7.3.

While the success of the retrieval strategy relative to reported Best Track V_{max} estimates are indicative of the sensitivity of CYGNSS measurements to various storm properties, it is desirable to reduce the number of required ancillary parameters as

Table 7.3: Ancillary parameters required to implement Generalized Asymmetric Holland Model and Willoughby Model as part of E2ES forward modelling procedure. Asterisk indicates that need for parameter is contingent on level of storm development.

GAHM	Willoughby
Storm center latitude	Storm center latitude
Storm center longitude	Storm center longitude
Minimum sea level pressure	-
Pressure of the last closed isobar	-
Radius of maximum winds	-
Radius in quadrant 1 at 34kt isotach	-
Radius in quadrant 2 at 34kt isotach	-
Radius in quadrant 3 at 34kt isotach	-
Radius in quadrant 4 at 34kt isotach	-
*Radius in quadrant 1 at 50kt isotach	-
*Radius in quadrant 2 at 50kt isotach	-
*Radius in quadrant 3 at 50kt isotach	-
*Radius in quadrant 4 at 50kt isotach	-
*Radius in quadrant 1 at 64kt isotach	-
*Radius in quadrant 2 at 64kt isotach	-
*Radius in quadrant 3 at 64kt isotach	-
*Radius in quadrant 4 at 64kt isotach	-

to advance matched filter storm SSW retrievals towards a more near-real time operational capacity (which is currently limited by 1-2 day latencies, as required ancillary information becomes available). A number of such venues were investigated which may provide guidance for future retrieval efforts. Examples of this include modifying the Willoughby model functions to compensate for the overestimation of winds. An example of this is shown in Figure 7.19 where correction terms were introduced to the storm's decay lengths as well as weighted exponential mixing factors.

A consistent error reduction, at times exceeding 80%, from doing so is observed. Peak errors in storm profile within a 250 km of the storm center of $\approx$ 12 m/s are

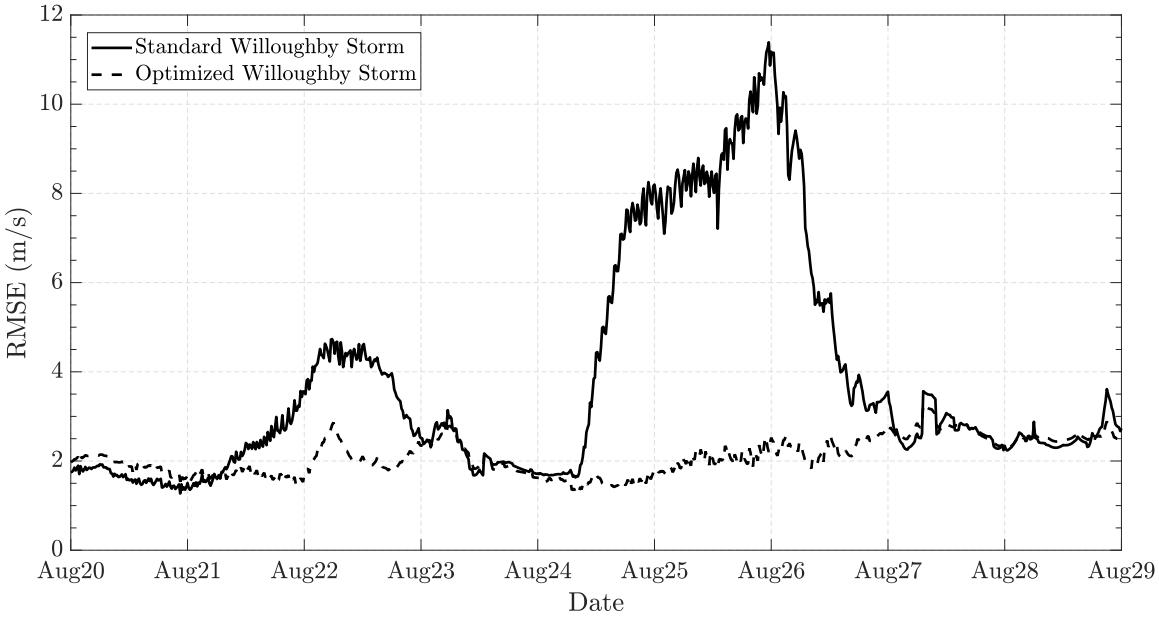


Figure 7.19: Time series of Willoughby model root mean square error for Hurricane Harvey relative to reference dataset before and after the modification of its model functions

observed using the original model functions whilst those using the modified terms peak at 2.5 m/s. Using a sufficiently large reference dataset, retrained Willoughby parametric relationships may be derived. A further example of the advantages this methodology provides is given in Figure 7.20 where a significantly more 'contained' storm rendition is observed after tuning.

The improvement in overall storm profile is beneficial for V_{max} retrievals as it enables the reference data set to better mimic the surface conditions which affected the formation of measured CYGNSS DDMs but is of particular importance for closely related studies such as storm surge estimation using CYGNSS V_{max} estimates. The original Willoughby model, through its significant overestimation of winds beyond

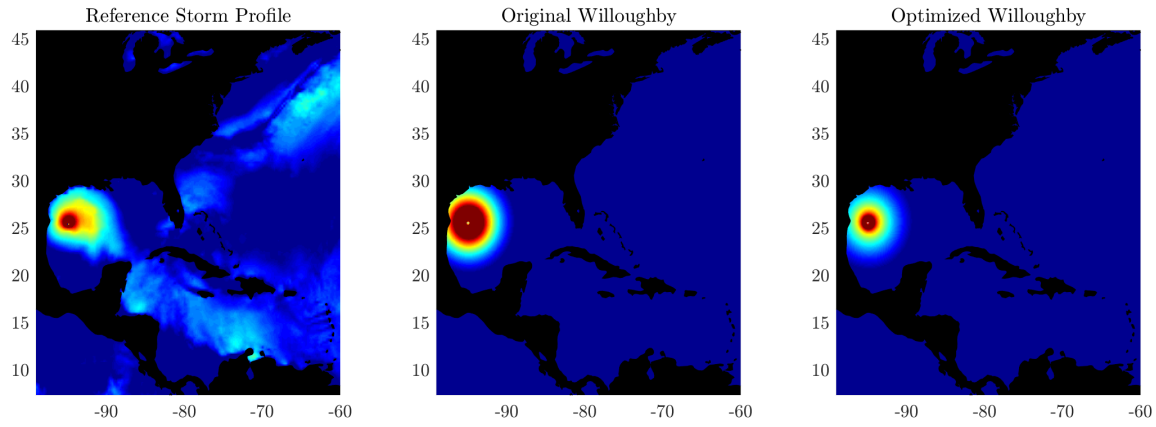


Figure 7.20: Effects of tuning Willoughby model function for Hurricane Harvey on DOY 237, 2017

70 km from the storm center, in this case, can lead to overestimates in excess of 5 meters.

Other approaches with which the need for an increased number of ancillary storm structure parameters may be minimized is through the merging of simpler models like the 1980 Holland model with the Willoughby model. The use of the original Holland model is of particular interest, as it is known to be significantly more 'contained' compared to other descriptions.

7.7 Retrieval Using Raw I/F Data

Other potential venues for improvement are explored in the context of measurement delay extent. Analysis relating to the extension of the retrieval approach to standard Level-1 CYGNSS DDMs, motivated by their more frequent availability, has shown that varying storm V_{max} introduces limited changes to the 11×17 delay-Doppler region thereby resulting in a very 'flat' response in correlation/RMSE across

the varying wind speeds. As a consequence, this did not allow successful retrievals to be conducted using the SDs. On the other hand, based on initial simulation results a very clear decline in error is observed with increasing measurement delay extent owing to the increased information this facilitates about the observed storm. The impact of conducting retrievals at increased delays is therefore explored further in the context of CYGNSS Raw I/F measurements.

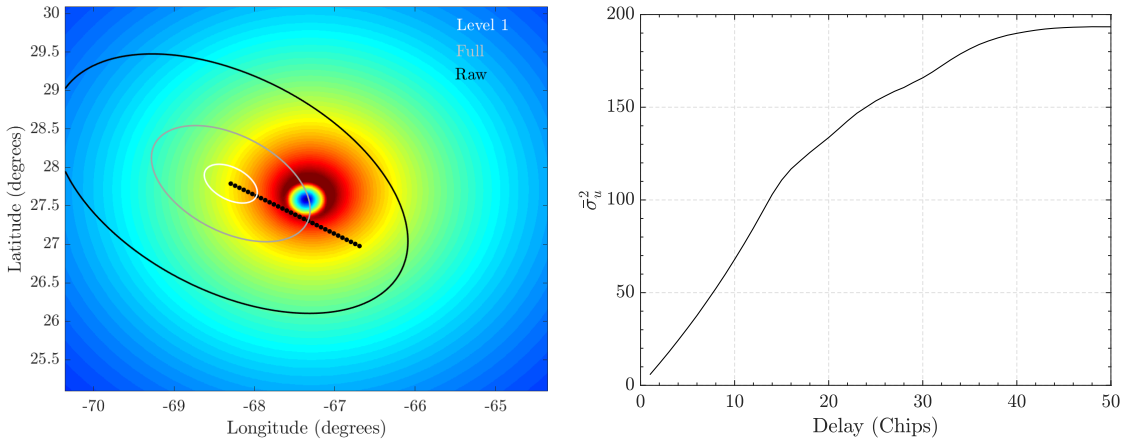


Figure 7.21: Projections of spatial footprint of the three different CYGNSS DDM products (a) Projection of spatial equivalent of maximum measurement delay extent for standard and special CYGNSS DDM products for a geometry resembling a Raw I/F mode downlink over hurricane Florence (b) Variance of wind speed within measurement footprint at increasing delay extent

The primary distinction between the different DDM products is their varying extents in delay. L1 DDMs are limited to a 3.5 chip maximum delay extent, Full DDMs are limited to 16 chips, and Raw I/F streams enable the creation of DDMs with an arbitrary maximum delay extent up to even 50 chips or more. An example of the surface footprint this translates to for a CYGNSS Raw I/F track observing

hurricane Florence is depicted in Figure 7.21(a). The increased delay extent of the Full DDM and Raw I/F modes is particularly advantageous from a retrieval perspective for a number of reasons. In observation scenarios where CYGNSS tracks do not align exactly with the storm center (i.e. "near miss" geometries) and instead pass through the storm at varying radial separations from its center, the effectiveness of the L1 DDM may be limited as it only furnishes information about a small portion of the storm. The increased spatial span of the special modes of acquisition enable spatial extents that instead "sample" a larger portion of the storm profile so that information from the storm center, transition region and the portion of the storm that is beyond the transition region is included. The larger "wind variance" mapped within each footprint is shown in Figure 7.21(b). As a result, measurements with increased delay extent hold the potential to improve retrievals and further CYGNSS's ability to improve storm feature characterization.

To investigate this, all available storm observing Raw I/F tracks have been processed, and those relevant for the retrievals through the enforcement of a similar +/- 30 minute CYGNSS/Storm separation and a +/- 200 km distance separation thresholds have been identified. Retrievals were then attempted at varying delay extents ranging between 1 chip to 50 chips to investigate the dependence of retrieval error relative to delay extent. Sample results are shown in Figure 7.22; a clear downward trend in error is observed as the measurement delay extent increases thereby showing support for the value of measurements over larger delay extents in improving the sensing of cyclone maximum winds. Beyond roughly 46 chips, an upward trend in error is observed. This is not a general result and is specific to the track geometry analyzed as part of the example in Figure 7.22 and arises due to the fact that delay

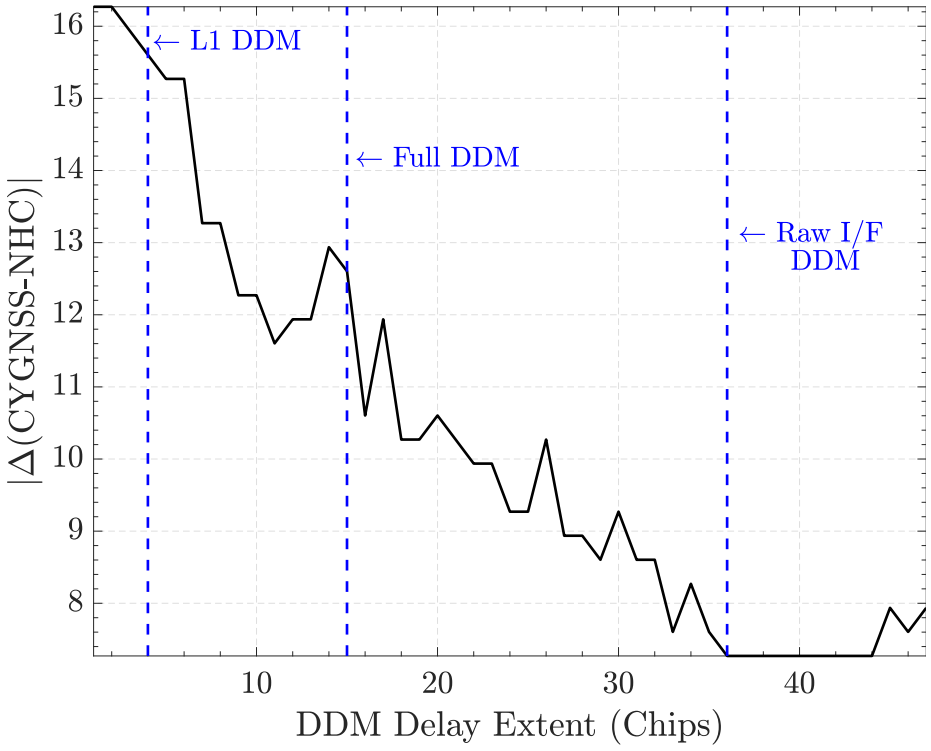


Figure 7.22: Maximum hurricane wind retrieval error expressed as a function of maximum measurement delay extent using CYGNSS Raw I/F DDMs for a track observing hurricane Florence on DOY 254, 2018

bins corresponding to levels exceeding 46 chips fall into the lower gain portion of the receiver's antenna footprint such that the DDMs' shape is no longer sufficiently defined and becomes diluted.

7.8 Conclusions

Cyclone maximum wind retrievals were demonstrated using tracks of full delay-Doppler map measurements of the Cyclone Global Navigation Satellite System constellation. The method is based on the production of a library of simulated delay-Doppler maps for a synthetic storm model as the parameters of the synthetic storm

are varied. A matched filter approach between normalized simulated and measured DDMs is then applied to retrieve storm parameters. Because the method uses normalized DDMs in the retrieval, it is insensitive to the absolute power calibration of the DDM and instead relies on the DDM shape. The impacts of using different storm models as part of the forward modelling procedure were also analyzed, where variations in local winds exceeding 25 m/s were observed depending on which parametric storm model was used. The two models incorporated as part of the E2ES forward modelling process include Willoughby and GAHM. The Willoughby model describes an exponentially decaying wind speed function within the eye wall radius R_1 given by V_{in}, a transition wind speed function between R_1 and the radial extent of the transition region R_2 given by V_{tr}, and a second exponentially decaying wind speed function beyond R_2 given by V_{out} where V_{tr} is a combination of V_{in} and V_{out}. The second model is GAHM and follows an iterative procedure to generate a storm model with true quadrant by quadrant asymmetries. Optimal performance was shown to be achieved when the GAHM model is used for storms with levels of development D_s ranging between $1 \leq D_s \leq 3$ and using the Willoughby model within the range $4 \leq D_s \leq 5$. Results from the method were shown for CYGNSS Full DDM cyclone overpasses for a total of 43 Full DDM retrievals spanning 11 storms over the 2017-2019 hurricane seasons with mean retrieval difference $\Delta_{mean} = 5.04$ m/s, RMSE = 6.50 m/s, URMSE = 6.42 m/s, bias is 1.06 and overall correlation $R = 89\%$. Relative to storm V_{max}, error was found to be on average 9.79% (and is within the CYGNSS 10% mission requirement). The success of the storm SSW retrieval methodology using Full DDMs motivates moving towards making their availability more frequent over

storms, as part of CYGNSS's routine mission operations. The effects of measurement delay extent on retrieval error were also analyzed in the context of CYGNSS Raw I/F downlinks due to the facility of creating DDMs with arbitrary delay extents their use provides. Significant and near-monotonic decrease was observed in retrieval error associated with increased measurement extent. This was attributed to the increased information DDMs with longer 'tails' provide about the storm's surface through sampling a larger portion of its structure within a single measurement.

Chapter 8: Conclusions and Future Work

The work presented in this dissertation aimed to develop new theory and techniques to support and facilitate the retrieval of geophysical quantities of interest over land and ocean surfaces using spaceborne GNSS-R systems. While the utility of the bulk of the proposed methodologies was investigated in the context of the Cyclone Global Navigation Satellite System mission, the proposed techniques are applicable to any spaceborne GNSS-R system.

Chapters 2 and 3 provided an overview of the fundamental signal processing and physical underpinnings of Global Navigation Satellite System Reflectometry highlighting all pertinent aspects relevant to their operation as passive bistatic radar systems. This also included an overview of the basic operation, properties, objectives and data products of the CYGNSS constellation.

Chapter 4 develops a time series soil moisture retrieval technique, based on change detection. The study begins by recognizing that the influence of roughness and vegetation significantly complicates attempts to retrieve soil moisture information and subsequently develops a framework which leverages the slow changing nature of processes like vegetation and roughness relative to soil moisture such that with sufficiently low latency of acquisition, ratios of NRCS will be directly proportional to surface reflectivity and therefore to soil moisture ratios. The retrieval algorithm begins with

the elimination of corrupting returns for a given day of the time series, followed by the computation of the NRCS and subsequently the construction of an underdetermined (N-1) by N system of equations which constitutes ratios of successive NRCS observables. Due to the fact that the system is not of full rank, a bounded least-squares optimization scheme is adopted to solve for reflectivities at various instants within the time series. Finally, the reflectivity solutions are inverted into soil moistures through a dielectric mixing model. The results of testing the retrieval algorithm over a wide range of canopies and land covers made clear its efficacy. High degrees of correlation between soil moisture values retrieved using CYGNSS land returns relative to SMAP were observed and are indicative of the potential for using GNSS-R systems for the purposes of retrieving soil moisture.

Numerous avenues for future investigation, still exist. In particular, future studies should aim to reduce the dependence of the method on SMAP ancillary soil moisture information. Further, efforts aimed at conducting spaceborne remote sensing of soil moisture are motivated by the desire to incorporate retrieved estimates into weather prediction models. In spite of accounting for less than 1% of global volumetric water content, soil moisture modulates atmospheric conditions through the role it plays in partitioning net incoming radiation into latent heat flux (associated with evapotranspiration) and sensible heat fluxes. This and its strong memory extending over inter-annual time scales in some cases enables it to play the outsized role of impacting weather evolution and unpredicted variability if not accounted for accurately in weather forecast models. The expectation is that accurate knowledge of soil moisture

can have profound implications on the ability to predict weather and large scale atmospheric variability. Future efforts should also aim to incorporate CYGNSS based soil moisture into coupled land-atmosphere models.

Chapter 5 develops a methodology that enables the detection of coherent returns on the global scale, using CYGNSS Level-1 DDM measurements. The proposed approach aimed to leverage the correspondence of the increase in power spread across the varying types of DDMs through a coherence metric, the power ratio PR, to classify CYGNSS Level-1 DDMs as dominantly coherent or dominantly non-coherent. To do this, the DDMs are decomposed into two segments. The first comprises all the delay-Doppler pairs within the ± 1 delay bins and $\pm$ 2 Doppler bin region about the specular bin and a second segment comprising all bins excluding this region. The power in the 'inside' region is then summed and divided by the summed power in the 'outside' region resulting in PR. Because of the limited power spread in a dominantly coherent scenario the power ratio is expected to be high while for the incoherent case due to the increased power spread it is expected to be low. As a further step to reduce the corrupting impact thermal noise can have on this procedure a 'noise exclusion' threshold was introduced to eliminate pixels as a preliminary step prior to computing PR. Extension of the detection approach to approximately 1 billion CYGNSS land specular point suggests coherence is, relative to incoherence, uncommon comprising 8.9% of all measurements. Further investigation suggested that on the global scale, the prevalence of coherence was in large limited to inland water bodies, inundated areas and flat desert regions.

In spite of the transition of the proposed technique to the CYGNSS mission's science operations center there remains a number of questions meriting further research.

Initial tests made clear that the setting of the appropriate detection threshold is affected by DDMs integration times. This is primarily due to the balance of the Level-1 coherence metric's numerator and denominator and the means with which varied integration time primarily impacts the denominator through altering overall noise levels. While the Level-1 integration time is fixed, interest in introducing modes with reduced integration times are of interest. Future studies should aim to establish a framework with which the appropriate detection threshold may be set as a function of DDM integration time and resolutions in delay and Doppler. In the absence of a global truth dataset, future efforts should also aim to analyze the detector's receiver operating characteristics through a comprehensive simulation framework that incorporates all factors that contribute to the formation of a measured DDM including accounting for the effects of the observation geometry, antenna footprint, surface roughness, land cover, soil moisture and canopy.

Chapter 6 illustrated the functionality of the Level-1 coherence detection methodology, developed in the previous Chapter, beyond being an indicator of dominant coherent reflection. This included mapping inland water bodies on the global scale, monsoon dynamics mapping, assessment of GNSS-R surface resolution under a dominantly coherent reflection regime and potential for improving soil moisture estimation.

Chapter 7 aimed to bypass the limitations of standard GNSS-R wind retrievals, through a matched filter approach to estimate maximum hurricane winds. The method relies on comparing the 'likeness' of CYGNSS measurements to a forward modeled delay-Doppler map library of its observations under varying storm parameters. The matched filter approach between simulated and measured DDMs is then applied to retrieve storm parameters. The extension of this retrieval methodology

to numerous CYGNSS observations of storms of interest highlights the promise of this approach in improving storm feature characterization using spaceborne GNSS-R systems. Particular emphasis in this study was placed on the assessment of retrieval sensitivity to storm models used in the forward modelling of CYGNSS returns as well as measurement delay extent highlighting the needs for highly accurate storm models used as part of the forward modelling procedure and the benefits increased measurement delay can have on retrieval performance.

Here too a number of avenues for future improvements exists. First, while retrieval performance was found to be optimized by incorporating both the GAHM and Willoughby models at different stages of storm development, the use of the GAHM model entails the ingestion of a significant amount of ancillary information pertaining to storm structure based on NHC Best Track forecasts. This is then inputted to a highly complex, and computationally demanding, iterative procedure aimed at implementing the GAHM model on the E2ES surface grid. Future efforts should aim to simplify this process by either substituting GAHM with an equally effective but less complex storm model, or aim to devise a single unified storm model that substitutes both GAHM and Willoughby whilst retaining the strengths of both. Secondly, future iterations of the E2ES equipped with internally generated synthetic storm models should also aim to include the effects of land, both on CYGNSS measurements and sea state near coastal regions; more specifically on storm structure as it approaches land. Thirdly, towards meeting the objectives outlined as part of this Chapter the E2ES underwent several rounds of revision, each with a number of iterative improvements aimed at enabling better computational efficiency. Relative to the first modified version of the E2ES including internally generated Willoughby models, the current

version's processing time, for conducting a complete retrieval, is less than 0.1%. This nonetheless still equates to a duration of 3-6 hours depending on the configuration of a given simulation and future studies should explore venues with which near real time maximum hurricane wind matched filter estimation may be achieved.

Bibliography

[1] J. T. Johnson, J. Park, Y. Yi and A. O'Brien, "Investigating Atmospheric Variability over Multiple Time Scales with CYGNSS," *NASA CYGNSS Science Team Meeting AOML/NOAA*, Dec 2017, Miami, FL.

[2] Z. Jelenak, S. Soisuvarn, F. Said and P. S. Chang, "L2 Wind Validation," *NASA CYGNSS Science Team Meeting AOML/NOAA*, Dec 2017, Miami, FL.

[3] J. T. Johnson, J. Park, Y. Yi, A. O'Brien, M. Al-Khaldi, A. Bringer and C. K. Shum, "CYGNSS Cal/Val and Ocean Science Activities at OSU and UCLA," *NASA CYGNSS Science Team Meeting*, Jun 2018, Ann Arbor, MI.

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

[5] R. de Vos van Steenwijk, M. Unwin and P. Jales, "Introducing the SGR-ReSI: A next generation spaceborne GNSS receiver for navigation and remote-sensing," *5th ESA Workshop on Satellite Navigation Technologies and European Workshop on GNSS Signals and Signal Processing (NAVITEC)*, 2010.

[6] M. Unwin, P. Jales, P. Blunt and S. Duncan, "Preparation for the first flight of SSTL's next generation space GNSS receivers," *6th ESA Workshop on Satellite Navigation Technologies (Navitec 2012) & European Workshop on GNSS Signals and Signal Processing*, 2012.

[7] C. Ruf et al., "The CYGNSS nanosatellite constellation hurricane mission," *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2012.

[8] J. Dickinson, J. Alvarez, L. McDaniel, J. Pruitt, B. Walls and K. Zajicek, "CYGNSS command and data subsystem and electrical power subsystem phase A and B developments," *IEEE Aerospace Conference*, 2014.

[9] R. Rose, C. Ruf, D. Rose, M. Brummitt and A. Ridley, "The CYGNSS flight segment; A major NASA science mission enabled by micro-satellite technology," *IEEE Aerospace Conference*, 2013.

[10] S. Miller, R. Killough, S. Baldor, M. Araujo, Z. Dischner and D. Rose, "Juggling Spacecraft: Similarities and differences of eight microsatellites for the CYGNSS mission," *IEEE Aerospace Conference*, 2016.

[11] S. Gleason and C. Ruf, "Overview of the Delay Doppler Mapping Instrument (DDMI) for the cyclone global navigation satellite systems mission (CYGNSS)," *IEEE MTT-S International Microwave Symposium*, 2015.

[12] C. Barthlott and N. Kalthoff, "A Numerical Sensitivity Study on the Impact of Soil Moisture on Convection-Related Parameters and Convective Precipitation over Complex Terrain," *Journal of the Atmospheric Sciences*, vol. 68, no. 12, pp. 2971-2987, 2011.

[13] C. Hohenegger, P. Brockhaus, C. Bretherton and C. Schär, "The Soil Moisture–Precipitation Feedback in Simulations with Explicit and Parameterized Convection," *Journal of Climate*, vol. 22, no. 19, pp. 5003-5020, 2009.

[14] D. Bae and K. Georgakakos, "Climatic variability of soil water in the American Midwest: Part 1. Hydrologic modeling," *Journal of Hydrology*, vol. 162, no. 3-4, pp. 355-377, 1994.

[15] J. Chang and P. Wetzel, "Effects of Spatial Variations of Soil Moisture and Vegetation on the Evolution of a Prestorm Environment: A Numerical Case Study," *Monthly Weather Review*, vol. 119, no. 6, pp. 1368-1390, 1991.

[16] National Research Council, Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond, pp. 400, 2007.

[17] S. Paloscia, S. Pettinato, E. Santi, C. Notarnicola, L. Pasolli and A. Reppucci, "Soil moisture mapping using Sentinel-1 images: Algorithm and preliminary validation," *Remote Sensing of Environment*, vol. 134, pp. 234-248, 2013.

[18] C. Pathe, W. Wagner, D. Sabel, M. Doubkova and J. Basara, "Using ENVISAT ASAR Global Mode Data for Surface Soil Moisture Retrieval Over Oklahoma, USA," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 2, pp. 468-480, 2009.

[19] M. Aubert, N. Baghdadi, M. Zribi, A. Douaoui, C. Loumagne, F. Baup, M. El Hajj and S. Garrigues, "Analysis of TerraSAR-X data sensitivity to bare soil moisture, roughness, composition and soil crust," *Remote Sensing of Environment*, vol. 115, no. 8, pp. 1801-1810, 2011.

[20] J. Ouellette, J. T. Johnson, A. Balenzano, F. Mattia, G. Satalino, S. Kim, R. Dunbar, A. Colliander, M. Cosh, T. Caldwell, J. Walker and A. Berg, "A

Time-Series Approach to Estimating Soil Moisture From Vegetated Surfaces Using L-Band Radar Backscatter," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 6, pp. 3186-3193, 2017.

[21] J. Ouellette, J. T. Johnson, S. Kim, J. van Zyl, M. Moghaddam, M. Spencer, L. Tsang and D. Entekhabi, "A Simulation Study of Compact Polarimetry for Radar Retrieval of Soil Moisture," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 9, pp. 5966-5973, 2014.

[22] A. Balenzano, F. Mattia, G. Satalino and M. Davidson, "Dense Temporal Series of C- and L-band SAR Data for Soil Moisture Retrieval Over Agricultural Crops," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 4, no. 2, pp. 439-450, 2011.

[23] S. Kim, L. Tsang, J. T. Johnson, S. Huang, J. van Zyl and E. Njoku, "Soil Moisture Retrieval Using Time-Series Radar Observations Over Bare Surfaces," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 5, pp. 1853-1863, 2012.

[24] S. Kim, M. Arii and T. Jackson, "Modeling L-Band Synthetic Aperture Radar Data Through Dielectric Changes in Soil Moisture and Vegetation Over Shrublands," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 11, pp. 4753-4762, 2017.

[25] M. Clarizia, N. Pierdicca, F. Costantini and N. Floury, "Analysis of CYGNSS Data for Soil Moisture Retrieval," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 7, pp. 2227-2235, 2019.

[26] M. Al-Khaldi, J. T. Johnson, A. O'Brien, A. Balenzano and F. Mattia, "Time-Series Retrieval of Soil Moisture Using CYGNSS," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1-10, 2019.

[27] A. Camps, H. Park, M. Pablos, G. Foti, C. Gommenginger, P. Liu and J. Judge, "Sensitivity of GNSS-R Spaceborne Observations to Soil Moisture and Vegetation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 10, pp. 4730-4742, 2016.

[28] H. Kim and V. Lakshmi, "Use of Cyclone Global Navigation Satellite System (CyGNSS) Observations for Estimation of Soil Moisture," *Geophysical Research Letters*, 2018.

[29] C. Chew, A. Colliander, R. Shah, C. Zuffada and M. Burgin, "The Sensitivity of Ground-Reflected GNSS Signals to Near-Surface Soil Moisture, as Recorded by Spaceborne Receivers," *International Geoscience and Remote Sensing Symposium (IGARSS)*, Jul 2017, Dallas, TX.

[30] C. Ruf, C. Chew, T. Lang, M. Morris, K. Nave, A. Ridley and R. Balasubramaniam, "A New Paradigm in Earth Environmental Monitoring with the CYGNSS Small Satellite Constellation," *Nature*, vol. 8, no. 1, 2018.

[31] C. Chew and E. Small, "Soil Moisture Sensing Using Spaceborne GNSS Reflections: Comparison of CYGNSS Reflectivity to SMAP Soil Moisture," *Geophysical Research Letters*, vol. 45, no. 9, pp. 4049-4057, 2018.

[32] C. Chew, E. Small, K. Larson and V. Zavorotny, "Effects of Near-Surface Soil Moisture on GPS SNR Data: Development of a Retrieval Algorithm for Soil Moisture," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 1, pp. 537-543, 2014.

[33] C. Chew, R. Shah, C. Zuffada, G. Hajj, D. Masters and A. Mannucci, "Demonstrating soil moisture remote sensing with observations from the UK TechDemoSat-1 satellite mission," *Geophysical Research Letters*, vol. 43, no. 7, pp. 3317-3324, 2016.

[34] D. Masters, "Initial results of land-reflected GPS bistatic radar measurements in SMEX02," *Remote Sensing of Environment*, vol. 92, no. 4, pp. 507-520, 2004.

[35] C. Safdi and C. Ruf, "CYGNSS sensitivity to leaf surface wetness," CYGNSS Science Team Meeting, Jan 2019, Pasadena, CA.

[36] M. Morris, C. Chew, J. T. Reager, R. Shah and C. Zuffada, "A Novel Approach To Monitoring Wetland Dynamics Using CYGNSS: Everglades Case Study," *Remote Sensing of Environment*, 2019.

[37] E. Loria, V. Zavorotny, M. Lavalle, C. Chew and R. Shah, "Analysis of Wetland Extent Retrieval Accuracy Using CYGNSS," *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Jul 2019, Yokohama, Japan.

[38] C. Safdi and C. Ruf, "Improved global seasonal water masks with CYGNSS," CYGNSS Science Team Meeting, Jun 2019, Ann Arbor, MI.

[39] K. Jensen, K. McDonald, E. Podest, N. Rodriguez-Alvarez, V. Horna and N. Steiner, "Assessing L-Band GNSS-Reflectometry and Imaging Radar for Detecting Sub-Canopy Inundation Dynamics in a Tropical Wetlands Complex," *Remote Sensing*, vol. 10, no. 9, p. 1431, 2018.

[40] S. Yueh, R. Shah, A. Hayashi, A. Colliander, X. Xu and J. Chaubell, "Dependence of CYGNSS Reflectivity on Vegetation Water Content and Surface Roughness," *NASA CYGNSS Science Team Meeting*, Jun 2019, Ann Arbor, MI.

[41] M. Al-Khaldi, J. T. Johnson, J. Park and A. O'Brien, "An Analysis of CYGNSS Returns Over Land," *USNC-URSI National Radio Science Meeting*, Jan 2018, Boulder, CO.

[42] E. S. Blake and D. A. Zelinsky, "National Hurricane Center Tropical Cyclone Report Hurricane Harvey (AL092017)," National Oceanic and Atmospheric Administration (NOAA), 2018.

[43] J. P. Cangialosi, A. S. Latto and R. Berg, "National Hurricane Center Tropical Cyclone Report Hurricane Irma (AL112017)," National Oceanic and Atmospheric Administration (NOAA), 2018.

[44] R. J. Pasch, A. B. Penny and R. Berg, "National Hurricane Center Tropical Cyclone Report Hurricane Maria (AL152017)," National Oceanic and Atmospheric Administration (NOAA), 2018.

[45] U.S. Billion-Dollar Weather & Climate Disasters 1980-2019, NOAA National Centers for Environmental Information, 2019.

[46] Eumetsat, "ASCAT Product Guide", 2016.

[47] W. S. Olson and G. Masunaga, "GPM Combined Radar-Radiometer Precipitation Algorithm Theoretical Basis Document", 2016.

[48] CYGNSS Handbook, Cyclone Global Navigation Satellite System: Deriving Surface Wind Speeds in Tropical Cyclones, 2016.

[49] M. Skolnik, Introduction to radar systems. McGraw-Hill Education, 1981.

[50] United States Department of Defense, "Global Positioning System Standard Positioning Service Performance Standard 4th Edition," Washington, DC, 2008.

[51] T. Hobiger, R. Haas and J. Löfgren, "GLONASS-R: GNSS reflectometry with a Frequency Division Multiple Access-based satellite navigation system," *Radio Science*, vol. 49, no. 4, pp. 271-282, 2014.

[52] S. Yalvac and M. Berber, "Galileo satellite data contribution to GNSS solutions for short and long baselines," *Measurement*, vol. 124, pp. 173-178, 2018.

[53] O. Montenbruck and P. Steigenberger, "The BeiDou Navigation Message," *Journal of Global Positioning Systems*, vol. 12, no. 1, pp. 1-12, 2013.

[54] A. Sarma, Q. Sultana and V. Srinivas, "Augmentation of Indian Regional Navigation Satellite System to Improve Dilution of Precision," *Journal of Navigation*, vol. 63, no. 2, pp. 313-321, 2010.

[55] H. Namie, O. Okamoto, N. Kubo and A. Yasuda, "Initial performance evaluation of centimeter-class augmentation system using Quasi-Zenith Satellite System," *Electronics and Communications in Japan*, vol. 101, no. 8, pp. 3-10, 2018.

[56] S. T. Lowe, "Earth Remote Sensing using Surface-Reflected GNSS Signals (GNSS-Reflectometry)," *Space Based Positioning Navigation & Timing National Advisory Board*, Nov 2017, Redondo Beach, CA.

[57] B. Hofmann-Wellenhof, H. Lichtenegger and J. Collins, Global positioning system. New York: Springer, 2001.

[58] C. Jeffrey, An introduction to GNSS. Calgary: NovAtel, 2010.

[59] A. Ishimaru, *Wave Propagation and Scattering in Random Media*. Saint Louis: Elsevier Science, 2014.

[60] J. T. Johnson, "An efficient two-scale model for the computation of thermal emission and atmospheric reflection from the sea surface," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 3, pp. 560-568, 2006.

[61] J. T. Johnson, "Comparison of the physical optics and small slope theories for polarimetric thermal emission from the sea surface," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 2, pp. 500-504, 2002.

[62] M. Demir and J. T. Johnson, "Fourth- and higher-order small-perturbation solution for scattering from dielectric rough surfaces," *Journal of the Optical Society of America*, vol. 20, no. 12, p. 2330, 2003.

[63] M. Gilbert and J. T. Johnson, "A study of the higher-order small-slope approximation for scattering from a Gaussian rough surface," *Waves in Random Media*, vol. 13, no. 2, pp. 137-149, 2003.

[64] J. T. Johnson, R. Shin, Jin Au Kong, Leung Tsang and Kyung Pak, "A numerical study of the composite surface model for ocean backscattering," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 1, pp. 72-83, 1998.

[65] J. T. Johnson, R. Shin, J. Eidson, Leung Tsang and Jin Au Kong, "A method of moments model for VHF propagation," *IEEE Transactions on Antennas and Propagation*, vol. 45, no. 1, pp. 115-125, 1997.

[66] T. Elfouhaily and J. T. Johnson, "Extension of the local curvature approximation to third order and improved tilt invariance," *Waves in Random and Complex Media*, vol. 16, no. 2, pp. 97-119, 2006.

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

[68] A. G. Voronovich and V. U. Zavorotny, "Bistatic radar equation for signals of opportunity revisited," *IEEE Transactions on Geosciennce and Remote Sensing*, vol. 56, No.3, pp. 1959-1968, April, 2018.

[69] S. Gleason, "Remote Sensing of Ocean, Ice and Land Surfaces Using Bistatically Scattered GNSS Signals From Low Earth Orbit," Ph.D dissertation, University of Surrey, 2006.

[70] F. Said, S. Soisuvarn, S. Katzberg, Z. Jelenak and P. Chang, "Estimation of maximum hurricane wind speed using simulated CYGNSS measurements," *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2015.

[71] S. Katzberg, O. Torres, M. Grant and D. Masters, "Utilizing calibrated GPS reflected signals to estimate soil reflectivity and dielectric constant: Results from SMEX02," *Remote Sensing of Environment*, vol. 100, no. 1, pp. 17-28, 2006.

[72] V. Zavorotny and A. Voronovich, "Bistatic GPS signal reflections at various polarizations from rough land surface with moisture content," *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2000.

[73] M. Islam, D. Yang, S. Farooq, S. Masum and S. Siddika, "Sea Wind Measurement from Power Waveforms of GNSS-R Delay Signals of Beidou GEO Satellite," *IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, 2018.

[74] W. Li, A. Rius, F. Fabra, E. Cardellach, S. Ribo and M. Martin-Neira, "Revisiting the GNSS-R Waveform Statistics and Its Impact on Altimetric Retrievals," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 5, pp. 2854-2871, 2018.

[75] T. Wang, C. Ruf, B. Block, D. McKague and S. Gleason, "Design and Performance of a GPS Constellation Power Monitor System for Improved CYGNSS L1B Calibration," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 1, pp. 26-36, 2019.

[76] C. Ruf, R. Atlas, P. Chang, M. Clarizia, J. Garrison, S. Gleason, S. Katzberg, Z. Jelenak, J. Johnson, S. Majumdar, A. O'brien, D. Posselt, A. Ridley, R. Rose and V. Zavorotny, "New Ocean Winds Satellite Mission to Probe Hurricanes and Tropical Convection," *Bulletin of the American Meteorological Society*, vol. 97, no. 3, pp. 385-395, 2016.

[77] S. Gleason, "Algorithm Theoretical Basis Document (ATBD): Level 1A DDM Calibration Technical Memo," 2018.

[78] S. Gleason, "Algorithm Theoretical Basis Document (ATBD): Level 1B DDM Calibration Technical Memo," 2018.

[79] S. Gleason, C. Ruf, A. O'Brien and D. McKague, "The CYGNSS Level 1 Calibration Algorithm and Error Analysis Based on On-Orbit Measurements," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, pp. 1-13, 2018.

[80] M. Clarizia and C. Ruf, "On the Spatial Resolution of GNSS Reflectometry," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 8, pp. 1064-1068, 2016.

[81] M. Clarizia, C. Ruf, P. Jales and C. Gommenginger, "Spaceborne GNSS-R Minimum Variance Wind Speed Estimator," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 11, pp. 6829-6843, 2014.

[82] M. P. Clarizia, "Investigating the Effect of Ocean Waves on GNSS-R Microwave Remote Sensing Measurements," Ph.D dissertation, University of South Hampton, 2012.

[83] C. Ruf and R. Balasubramaniam, "Development of the CYGNSS Geophysical Model Function for Wind Speed," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 1, pp. 66-77, 2019.

[84] E. Uhlhorn, P. Black, J. Franklin, M. Goodberlet, J. Carswell and A. Goldstein, "Hurricane Surface Wind Measurements from an Operational Stepped Frequency Microwave Radiometer," *Monthly Weather Review*, vol. 135, no. 9, pp. 3070-3085, 2007.

[85] M. P. Clarizia, V. Zavorotny and C. Ruf, "Algorithm Theoretical Basis Document (ATBD): Level 2 Wind Speed Retrieval Technical Memo," 2018.

[86] C. Ruf, "Algorithm Theoretical Basis Document (ATBD): Level 3 Gridded Wind Speed," 2018.

[87] J. Garrison, "Modeling and simulation of bin-bin correlations in GNSS-R waveforms," *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2012.

[88] J. Garrison, "A Statistical Model and Simulator for Ocean-Reflected GNSS Signals," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 10, pp. 6007-6019, 2016.

[89] F. Martin, S. D'Addio, A. Camps and M. Martin-Neira, "Modeling and Analysis of GNSS-R Waveforms Sample-to-Sample Correlation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 5, pp. 1545-1559, 2014.

[90] A. Balakhder, M. Al-Khaldi and J. Johnson, "On the Coherency of Ocean and Land Surface Specular Scattering for GNSS-R and Signals of Opportunity Systems," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1-11, 2019.

[91] A. Camps, "Spatial Resolution in GNSS-R Under Coherent Scattering," *IEEE Geoscience and Remote Sensing Letters*, pp. 1-5, 2019.

[92] S. Katzberg, J. Dunion and G. Ganoe, "The use of reflected GPS signals to retrieve ocean surface wind speeds in tropical cyclones," *Radio Science*, vol. 48, no. 4, pp. 371-387, 2013.

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

[94] A. O'Brien, "Algorithm Theoretical Basis Document (ATBD): End-to-End Simulator Technical Memo," 2014.

[95] A. O'Brien, A. Ridley and D. J. Posselt, "Cyclone Global Navigation Satellite System End to End Simulator User Guide," 2015.

[96] A. O'Brien, S. G. Gleason and J. T. Johnson , "The CYGNSS end-to-end simulator," GNSS+R, May 2017, Postdam, Germany.

[97] A. O'Brien and J. T. Johnson, "Comparing the CYGNSS simulator forward scattering model with TDS-1 and CYGNSS on-orbit observations," *IEEE International Geoscience and Remote Sensing Symposium* (IGARSS), Jul 2017, Fort Worth, TX.

[98] R. De Roo and F. Ulaby, "Bistatic specular scattering from rough dielectric surfaces," *IEEE Transactions on Antennas and Propagation*, vol. 42, no. 2, pp. 220-231, 1994.

[99] J. Johnson, K. Warnick and Peng Xu, "On the Geometrical Optics (Hagfors' Law) and Physical Optics Approximations for Scattering From Exponentially Correlated Surfaces," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 8, pp. 2619-2629, 2007.

[100] A. Mialon, P. Richaume, D. Leroux, S. Bircher, A. Bitar, T. Pellarin, J. Wigneron and Y. Kerr, "Comparison of Dobson and Mironov Dielectric Models in the SMOS Soil Moisture Retrieval Algorithm," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 6, pp. 3084-3094, 2015.

[101] N. Peplinski, F. Ulaby and M. Dobson, "Dielectric properties of soils in the 0.3-1.3-GHz range," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, no. 3, pp. 803-807, 1995.

[102] M. Hallikainen, F. Ulaby, M. Dobson, M. El-rayes and L. Wu, "Microwave Dielectric Behavior of Wet Soil-Part 1: Empirical Models and Experimental Observations," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 23, no. 1, pp. 25-34, 1985.

[103] V. Mironov, L. Kosolapova and S. Fomin, "Physically and Mineralogically Based Spectroscopic Dielectric Model for Moist Soils," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 7, pp. 2059-2070, 2009.

[104] V. Mironov and S. Fomin, "Temperature and Mineralogy Dependable Model for Microwave Dielectric Spectra of Moist Soils," *PIERS Online*, vol. 5, no. 5, pp. 411-415, 2009.

[105] V. Mironov, P. Bobrov and S. Fomin, "Multirelaxation Generalized Refractive Mixing Dielectric Model of Moist Soils," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 3, pp. 603-606, 2013.

[106] V. Mironov, Y. Kerr, J. Wigneron, L. Kosolapova and F. Demontoux, "Temperature- and Texture-Dependent Dielectric Model for Moist Soils at 1.4 GHz," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 3, pp. 419-423, 2013.

[107] S. O. Rice, "Reflection of electromagnetic waves from slightly rough surfaces," *Communications on Pure and Applied Mathematics*, vol. 4, no. 2-3, pp. 351-378, 1951.

[108] D. E. Barrick and W. H. Peake, "Scattering from Surfaces with different Roughness Scales: Analysis and Interpretation," *Battelle Memorial Institute publication*, vol. 10, no. 3, pp. 197A, 1967.

[109] T. Jackson and T. Schmugge, "Vegetation effects on the microwave emission of soils," *Remote Sensing of Environment*, vol. 36, no. 3, pp. 203-212, 1991.

[110] N. Das, "Ancillary Data Report: Soil Attributes," California Institute of Technology, 2013.

[111] M. Rodell et al., "The Global Land Data Assimilation System," *Bulletin of the American Meteorological Society*, vol. 85, no. 3, pp. 381-394, 2004.

[112] SMAP Ancillary Data Set, Jet Propulsion Laboratory, Pasadena, CA.

[113] S. Chan, R. Bindlish, and J. Kimball,"Vegetation Water Content," *SMAP Science Document*, Jet Propulsion Laboratory, Pasadena, CA, 2011.

[114] M. Friedl et al., "MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets," *Remote Sensing of Environment*, vol. 114, no.

[115] A. Colliander et al., "Validation of SMAP surface soil moisture products with core validation sites," *Remote Sensing of Environment*, vol. 191, pp. 215-231, 2017

[116] S. Gleason, C. Ruf, A. O'Brien and D. McKague, "The CYGNSS Level 1 Calibration Algorithm and Error Analysis Based on On-Orbit Measurements," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, pp. 1-13, 2018.

[117] M. Clarizia and C. Ruf, "Wind Speed Retrieval Algorithm for the Cyclone Global Navigation Satellite System (CYGNSS) Mission," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 8, pp. 4419-4432, 2016

[118] C. Ruf and R. Balasubramaniam, "Development of the CYGNSS Geophysical Model Function for Wind Speed,"

[119] C. Ruf, S. Gleason and D. McKague, "Assessment of CYGNSS Wind Speed Retrieval Uncertainty," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 1, pp. 87-97, 2019.

[120] C. Ruf et al., "In-Orbit Performance of the Constellation of CYGNSS Hurricane Satellites," *Bulletin of the American Meteorological Society*, vol. 100, no. 10, pp. 2009-2023, 2019.

[121] S. Nghiem et al., "Wetland monitoring with Global Navigation Satellite System reflectometry," *Earth and Space Science*, vol. 4, no. 1, pp. 16-39, 2017.

[122] H. Carreno-Luengo, G. Luzi and M. Crosetto, "First Evaluation of Topography on GNSS-R: An Empirical Study Based on a Digital Elevation Model," *Remote Sensing*, vol. 11, no. 21, p. 2556, 2019

[123] A. Alonso-Arroyo, V. Zavorotny and A. Camps, "Sea Ice Detection Using U.K. TDS-1 GNSS-R Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 9, pp. 4989-5001, 2017

[124] J. Cartwright, C. Banks and M. Srokosz, "Sea Ice Detection Using GNSS-R Data From TechDemoSat-1," *Journal of Geophysical Research: Oceans*, 2019.

[125] J. Strandberg, T. Hobiger and R. Haas, "Coastal Sea Ice Detection Using Ground-Based GNSS-R," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 9, pp. 1552-1556, 2017.

[126] C. Chew, C. Zuffada, R. Shah, and A. Mannucci, "Mapping sea ice using refected GNSS signals in a bistatic radar system," *EGU General Assembly Conference Abstracts*, vol. 18, p. 10574, Apr. 2016.

[127] N. Rodriguez-Alvarez, B. Holt, S. Jaruwatanadilok, E. Podest and K. Cavanaugh, "An Arctic sea ice multi-step classification based on GNSS-R data from the TDS-1 mission," *Remote Sensing of Environment*, vol. 230, p. 111202, 2019.

[128] E. Loria and A. O'Brien , "Simulation Study of CYGNSS Retrieval Algorithms for Wetland Extent," *USNC-URSI National Radio Science Meeting*, Jan 2019, Boulder, CO.

[129] H. Carreno-Luengo, G. Luzi and M. Crosetto, "Impact of the Elevation Angle on CYGNSS GNSS-R Bistatic Reflectivity as a Function of Effective Surface Roughness over Land Surfaces," Remote Sensing, vol. 10, no. 11, p. 1749, 2018.

[130] Q. Yan and W. Huang, "Spaceborne GNSS-R Sea Ice Detection Using Delay-Doppler Maps: First Results From the U.K. TechDemoSat-1 Mission," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 10, pp. 4795-4801, 2016.

[131] S. Gleason, J. Johnson, C. Ruf and C. Bussy-Virat, "Characterizing Background Signals and Noise in Spaceborne GNSS Reflection Ocean Observations," *IEEE Geoscience and Remote Sensing Letters*, pp. 1-6, 2019.

[132] E. Loria, A. O'Brien and I. Gupta, "Detection & Separation of Coherent Reflections in GNSS-R Measurements Using CYGNSS Data," *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2018.

[133] J. Pekel, A. Cottam, N. Gorelick and A. Belward, "High-resolution mapping of global surface water and its long-term changes," *Nature*, vol. 540, no. 7633, pp. 418-422, 2016.

[134] M. Brodzik, B. Billingsley, T. Haran, B. Raup and M. Savoie, "EASEGrid 2.0: Incremental but Significant Improvements for Earth-Gridded Data Sets," *ISPRS International Journal of GeoInformation*, vol. 1, no. 1, pp. 32-45, 2012.

[135] M. Brodzik, B. Billingsley, T. Haran, B. Raup and M. Savoie, "Correction: Brodzik, M.J., et al. EASEGrid 2.0: Incremental but Significant Improvements for EarthGridded Data Sets. ISPRS International Journal of Geo-Information

2012, 1, 32–45," *ISPRS International Journal of Geo-Information*, vol. 3, no. 3, pp. 1154-1156, 2014.

[136] A. Egido, S. Paloscia, E. Motte, L. Guerriero, N. Pierdicca, M. Caparrini, E. Santi, G. Fontanelli and N. Floury, "Airborne GNSS-R Polarimetric Measurements for Soil Moisture and Above-Ground Biomass Estimation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 5, pp. 1522-1532, 2014.

[137] E. Cardellach et al., "First Precise Spaceborne Sea Surface Altimetry With GNSS Reflected Signals," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 102-112, 2020.

[138] S. Postel, "Entering An Era of Water Scarcity: The Challenges Ahead," *Ecological Applications*, vol. 10, no. 4, pp. 941-948, 2000.

[139] D. Bastviken, L. Tranvik, J. Downing, P. Crill and A. Enrich-Prast, "Freshwater Methane Emissions Offset the Continental Carbon Sink," *Science*, vol. 331, no. 6013, pp. 50-50, 2011.

[140] L. Strelich, "Great Lakes Hold Sway over Water and Carbon Cycling," *Eos*, vol. 96, 2015.

[141] M. Blango, R. Cooke and J. Moiwo, "Effect of soil and water management practices on crop productivity in tropical inland valley swamps," *Agricultural Water Management*, vol. 222, pp. 82-91, 2019.

[142] L. Tranvik et al., "Lakes and reservoirs as regulators of carbon cycling and climate," *Limnology and Oceanography*, vol. 54, no. 62, pp. 2298-2314, 2009.

[143] S. Chan, E. Njoku and A. Colliander, "Soil Moisture Active Passive: Algorithm Theoretical Basis Document (ATBD) for Level 1C Radiometer Data Product," Jet Propulsion Laboratory, Pasadena, CA, 2014.

[144] A. Mahmoodi and I. Buchan, "Algorithm Theoretical Basis Document (ATBD) for the SMOS Level 2 Soil Moisture Processor Development Continuation Project," Toronto, Canada, 2011.

[145] EUMSAT, "AVHRR Level 1b Product Guide," Darmstadt, Germany, 2011.

[146] V. Ihlen, "Landsat 8 (L8) Data Users Handbook ," South Dakota, USA, 2019.

[147] European Space Agency, "Sentinel-2User Handbook," 2015.

[148] M. Abrams, S. Hook and B. Ramachandran, "Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) User Handbook," Pasadena, CA, USA.

[149] M. Carroll et al., "Development of an operational land water mask for MODIS Collection 6, and influence on downstream data products," *International Journal of Digital Earth*, vol. 10, no. 2, pp. 207-218, 2016.

[150] E. Borg, K. Guenther and B. Fichtelmann, "Dynamic self-learning water-masking algorithm for AATSR, MERIS, and SPOT VEGETATION," *International Journal of Business Intelligence and Data Mining*, vol. 1, no. 1, p. 1, 2017.

[151] C. Huang, Y. Chen, S. Zhang, L. Li, K. Shi and R. Liu, "Spatial Downscaling of Suomi NPP–VIIRS Image for Lake Mapping," Water, vol. 9, no. 11, p. 834, 2017.

[152] S. Kim, J. Ouellette, J. van Zyl and J. Johnson, "Detection of Inland Open Water Surfaces Using Dual Polarization L-Band Radar for the Soil Moisture Active Passive Mission," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 6, pp. 3388-3399, 2016.

[153] S. Gleason, A. O'Brien, A. Russel, M. M. Al-Khaldi and J. T. Johnson, "Geolocation, Calibration and Surface Resolution of Spaceborne GNSS-R Land Observations," *Remote Sensing*, 2020.

[154] M. M. Al-Khaldi, J. T. Johnson, S. Gleason, E. Loria, A. J. O'Brien and Y. Yi, "An Algorithm for Detecting Coherence in Cyclone Global Navigation Satellite System Mission Level 1 Delay Doppler Maps," Submitted to IEEE Transactions on Geoscience and Remote Sensing, Dec 2019.

[155] C. Gerlein-Safdi and C. Ruf,"A CYGNSS-Based Algorithm for the Detection of Inland Waterbodies," *Geophysical Research Letters*, vol. 46, no. 21, pp. 12065-12072, 2019.

[156] R. Schroeder et al., "Development and Evaluation of a Multi-Year Fractional Surface Water Data Set Derived from Active/Passive Microwave Remote Sensing Data," *Remote Sensing*, vol. 7, no. 12, pp. 16688-16732, 2015.

[157] E. Loria, A. O'Brien, V. Zavorotny, B. Downs and C. Zuffada, "Analysis of Scattering Characteristics from Inland Bodies of Water Observed by CYGNSS," *Remote Sensing of Environment*, 2020.

[158] P. J. Webster and S. Yang, "Monsoon and ENSO: Selectively interactive systems," *Royal Meorological Society*, vol. 118, pp. 877–926, 1992.

[159] B. N. Goswami, South Asian Monsoon, Intraseasonal Variability of the Atmosphere-Ocean Climate System, W. K. M. Lau and D. E. Waliser, Eds. Berlin, Germany: Springer-Verlag, 2005, ch. 2, pp. 19–61.

[160] P. J.Webster et al., "Monsoon: Processes, predictability, the prospects for prediction," *Journal of Geophysical Research*, vol. 103, pp. 14451–14510, 1998

[161] C. Zhang, "Madden–Julian Oscillation: Bridging Weather and Climate," *Bulletin of the American Meteorological Society*, vol. 94, no. 12, pp. 1849-1870, 2013.

[162] E. Douglas et al., "Changes in moisture and energy fluxes due to agricultural land use and irrigation in the Indian Monsoon Belt," *Geophysical Research Letters*, vol. 33, no. 14, 2006.

[163] World Meteorological Organization. *The Global Climate 2001-2010: A Decade of Climate Extremes–Summary Report*, 2013.

[164] V. Rakesh, R. Singh and P. Joshi, "Impact of four dimensional assimilation of satellite data on long-range simulations over the Indian region during monsoon 2006," *Advances in Space Research*, vol. 46, no. 7, pp. 895-908, 2010.

[165] M. Ramamurthy and F. Carr, "Four-Dimensional Data Assimilation in the Monsoon Region. Part II: Role of Temperature and Moisture Data," *Monthly Weather Review*, vol. 116, no. 10, pp. 1896-1913, 1988.

[166] R. Singh, C. Kishtawal, S. Ojha and P. Pal, "Impact of assimilation of Atmospheric InfraRed Sounder (AIRS) radiances and retrievals in the WRF 3D-Var assimilation system," *Journal of Geophysical Research*, vol. 117, no. 11, p. 1-17, 2012.

[167] B. Goswami and S. Chakravorty, "Dynamics of the Indian Summer Monsoon Climate," *Oxford Research Encyclopedia of Climate Science*, 2017.

[168] Q. Wu and Q. Hu, "Atmospheric circulation processes contributing to a multidecadal variation in reconstructed and modeled Indian monsoon precipitation," *Journal of Geophysical Research*, vol. 120, no. 2, pp. 532-551, 2015.

[169] B. Wang et al., "Rethinking Indian monsoon rainfall prediction in the context of recent global warming," *Nature Communications*, vol. 6, no. 1, 2015.

[170] A. Saha, S. Ghosh, A. Sahana and E. Rao, "Failure of CMIP5 climate models in simulating post-1950 decreasing trend of Indian monsoon," *Geophysical Research Letters*, vol. 41, no. 20, pp. 7323-7330, 2014.

[171] T. Antony, C. Suresh Raju, N. Mathew and K. Krishna Moorthy, "Flood Extent Analysis Over the Major River Basins in the Indian Subcontinent Using Satellite Microwave Radiometric Data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 9, pp. 4373-4378, 2015.

[172] F. Papa, C. Prigent, F. Durand and W. Rossow, "Wetland dynamics using a suite of satellite observations: A case study of application and evaluation for the Indian Subcontinent," *Geophysical Research Letters*, vol. 33, no. 8, 2006.

[173] N. Pierdicca, L. Pulvirenti, M. Chini, L. Guerriero and L. Candela, "Observing floods from space: Experience gained from COSMO-SkyMed observations," *Acta Astronautica*, vol. 84, pp. 122-133, 2013.

[174] R. Hoque, D. Nakayama, H. Matsuyama and J. Matsumoto, "Flood monitoring, mapping and assessing capabilities using RADARSAT remote sensing, GIS and ground data for Bangladesh," *Natural Hazards*, vol. 57, no. 2, pp. 525-548, 2010.

[175] J. Sanyal and X. Lu, "Application of Remote Sensing in Flood Management with Special Reference to Monsoon Asia: A Review," *Natural Hazards*, vol. 33, no. 2, pp. 283-301, 2004.

[176] L. Pulvirenti, M. Chini, N. Pierdicca and G. Boni, "Use of SAR Data for Detecting Floodwater in Urban and Agricultural Areas: The Role of the Interferometric Coherence," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 3, pp. 1532-1544, 2016.

[177] C. Chew, J. Reager and E. Small, "CYGNSS data map flood inundation during the 2017 Atlantic hurricane season," *Scientific Reports*, vol. 8, no. 1, 2018.

[178] M. M. Al-Khaldi, J. T. Johnson, S. Gleason, C. Chew, C. Gerlein-Safdi, R. Shah and C. Zuffada, "Inland Water Body Mapping Using CYGNSS Coherence Detection," *IEEE Transactions on Geoscience and Remote Sensing*, Under Review, 2020.

[179] G. Allen and T. Pavelsky, "Patterns of river width and surface area revealed by the satellite-derived North American River Width data set," *Geophysical Research Letters*, vol. 42, no. 2, pp. 395-402, 2015.

[180] M. Morris and C. Ruf, "Estimating Tropical Cyclone Integrated Kinetic Energy with the CYGNSS Satellite Constellation," *Journal of Applied Meteorology and Climatology*, vol. 56, no. 1, pp. 235-245, 2017.

[181] M. Morris and C. Ruf, "Determining Tropical Cyclone Surface Wind Speed Structure and Intensity with the CYGNSS Satellite Constellation," *Journal of Applied Meteorology and Climatology*, vol. 56, no. 7, pp. 1847-1865, 2017.

[182] A. Warnock, C. Ruf and M. Morris , "Storm Surge Modeling with CYGNSS winds," GNSS+R, May 2017, Ann Arbor, MI.

[183] F. Said, S. Soisuvarn, Z. Jelenak and P. Chang, "Performance Assessment of Simulated CYGNSS Measurements in the Tropical Cyclone Environment," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS)*, vol. 9, no. 10, pp. 4709-4719, 2016.

[184] F. Said, S. Katzberg and S. Soisuvarn, "Retrieving Hurricane Maximum Winds Using Simulated CYGNSS Power-Versus-Delay Waveforms," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 8, pp. 3799-3809, 2017.

[185] M. Clarizia and C. Ruf, "Wind Speed Retrieval Algorithm for the Cyclone Global Navigation Satellite System (CYGNSS) Mission," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 8, pp. 4419-4432, 2016.

[186] M. Morris and C. Sampson, "CYGNSS-based TC Characterization," CYGNSS Science Team Meeting, Jun 2019, Ann Arbor, MI.

[187] T. Elfouhaily, B. Chapron, K. Katsaros and D. Vandemark, "A unified directional spectrum for long and short wind-driven waves," *Journal of Geophysical Research: Oceans*, vol. 102, no. 7, pp. 15781-15796, 1997.

[188] H. Willoughby and M. Rahn, "Parametric Representation of the Primary Hurricane Vortex. Part I: Observations and Evaluation of the Holland (1980) Model," *Monthly Weather Review*, vol. 132, no. 12, pp. 3033-3048, 2004.

[189] H. Willoughby, R. Darling and M. Rahn, "Parametric Representation of the Primary Hurricane Vortex. Part II: A New Family of Sectionally Continuous Profiles," *Monthly Weather Review*, vol. 134, no. 4, pp. 1102-1120, 2006.

[190] R. W. Schloemer, "Analysis and synthesis of Hurricane wind patterns over Lake Okeechobee," *U.S. Department of Commerce Weather Bureau*, Washington, DC, USA, Technical Report 31, 1954.

[191] G. Holland, "An Analytic Model of the Wind and Pressure Profiles in Hurricanes," *Monthly Weather Review*, vol. 108, no. 8, pp. 1212-1218, 1980.

[192] "The Saffir/Simpson Hurricane Scale", NOAA's National Climatic Data Center, Asheville, NC.

[193] A. T. Cox and V. J. Cardone, "20 Years of Operational Forecasting at Oceanweather," *7th International Workshop on Wave Hindcasting and Forecasting*, 2002, Alberta, Canada.

[194] J. Gao, R. Leuttich and J. Fleming, "Development and Initial Evaluation of A Generalized Asymmetric Tropical Cyclone Vortex Model in ADCIRC," *ADCIRC Users Group Meeting*, 2013, Vicksburg, MS.