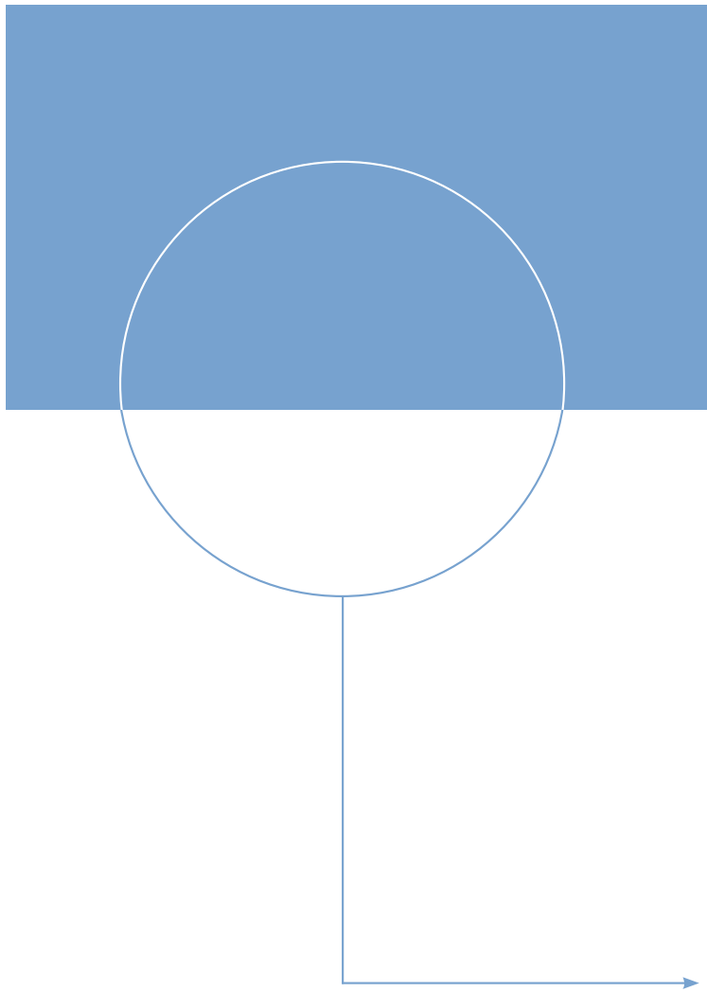


Master's thesis

2022

Syver Stenersen Kok

NTNU
Norwegian University of
Science and Technology
Faculty of Engineering
Department of Civil and Environmental Engineering

Syver Stenersen Kok

Assessment of GNSS Reflectometry's Capability to Predict the Global Distribution of Floating Ocean Microplastics

June 2022

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of
Science and Technology

NTNU

Norwegian University of
Science and Technology

Assessment of GNSS Reflectometry's Capability to Predict the Global Distribution of Floating Ocean Microplastics

Syver Stenersen Kok

Engineering and ICT
Submission date: June 2022
Supervisor: Hossein Nahavandchi
Co-supervisor: Mostafa Hoseini

Norwegian University of Science and Technology
Department of Civil and Environmental Engineering

Preface

This thesis finishes my master's degree in Geomatics in the study program Engineering and ICT at the Norwegian University of Science and Technology (NTNU).

I want to express my gratitude to the people that help me throughout this thesis. First and foremost, I am very grateful for all the assistance and guidance I received from my supervisor Professor Hossein Nahavandchi and my co-supervisor, Mostafa Hoseini. Additionally, I wish to thank Chris Ruf, the Principal Investigator of the NASA Cyclone Global Navigation Satellite System Earth Venture Mission, for taking the time to answer my questions. Lastly, I want to thank all my friends and family for their support throughout these five years in Trondheim.

S.S.K.

Abstract

As a severe environmental threat, the pollution of microplastics in the oceans affects marine life and often ends up in many parts of the food chain. The severity of the damage done to the ecosystem caused by microplastics in the world's oceans is hard to estimate accurately. In recent years, remote sensing techniques have been developed to predict floating microplastic distributions. This thesis aimed to investigate the potential of a novel remote sensing technique, i.e., the Global Navigation Satellite Systems (GNSS) Reflectometry (GNSS-R), for estimating the distribution of microplastics. Towards this goal, two available approaches are studied to assess their capability to monitor floating microplastics. NASA Cyclone GNSS (CYGNSS) observations over the oceans constitute the primary data set used in this thesis. Furthermore, several match-up data sets are used for modeling and performance assessment purposes. These data sets include wind measurements from the European Center for Medium-range Weather Forecast (ECMWF) Reanalysis-5 (ERA-5) data set, ocean surface current data from Ocean Surface Current Analyses Real-time (OSCAR), and available global microplastic models.

The conducted analysis reveals several issues within both methods and provides suggestions for improvement. Both approaches show promising results, highlighting GNSS-R's potential to detect possible accumulation zones of microplastics. This potential can draw particular attention since the GNSS-R technique can be efficiently implemented on small satellites. However, the method has limitations in the microplastic detection domain. Remote sensing of microplastics based on GNSS-R relies on a fundamental assumption that requires more verification research. The microplastics are assumed to share a similar transport mechanism with oceanic surfactants. Since surfactant concentrations leave detectable signatures in GNSS-R observations, detecting the accumulation zones of surfactants is considered to correlate with microplastic concentrations. Further research on the relationship between surfactants and floating microplastics is encouraged, which will be decisive for both methods' capability to indicate microplastic accumulation zones.

The code developed in this thesis can be found in this GitHub repository:
`https://github.com/Syverkok/Master_Thesis_Syver`

Sammendrag

Plastikkforurensing i havet påvirker maritimt liv og kan spores videre til mange deler av næringskjeden. Det er vanskelig å nøyaktig anslå hvor store skader økosystemet tar av mikroplastforurensingen av verdenshavene. Nylig har fjernmålingsteknikker blitt utviklet med mål om å predikere den globale distribusjonen av mikroplast. Denne masteroppgaven undersøkte to GNSS-R teknikker for å vurdere om de er kapable til å overvåke konsentrasjonen av mikroplast i havet. CYGNSS observasjoner over havet er den primære datakilden. Videre er flere datakilder brukt til modellering og vurdering av resulter. Disse datakildene inkluderer vindmålinger fra ERA-5 datasettet, havstrømsdata fra OSCAR datasettet, og eksisterende mikroplastmodeller.

Den utførte analysen belyser flere problemer og det foreslås forbedringsområder for begge metodene. Det presenteres også lovende resultater som understreker at GNSS-R kan være nyttige for å indikere mulige akkumuleringssoner av mikroplast. Teknikkene er spesielt interessante siden GNSS-R effektivt kan tas i bruk på små satellitter. GNSS-R har likevel begrensninger i mikroplastdeteksjonsdomenet. Fjernmåling av mikroplast i havet ved hjelp av GNSS-R avhenger fundamentalt av en usikker antagelse. Antagelsen er at mikroplast deler transportmekanismer med surfaktanter i havet. Siden surfaktanter kan spores i GNNS-R observasjoner blir akkumuleringssoner av surfaktanter assosiert med mikroplastansamlinger. Hvorvidt antagelsen om like transportmekanismer er troverdig, må videre undersøkes og blir avgjørende for metodenes evne til å indikere mikroplastakkumuleringssoner.

Koden som ble utviklet i forbindelse med masteroppgaven kan hentes fra dette GitHub-repositoriet:
`https://github.com/Syverkok/Master_Thesis_Syver`

Contents

List of Figures

List of Tables

Abbreviations

BRCS	Bi-static Radar Cross Section
CYGNSS	Cyclone Global Navigation Satellite System
DDMI	Delay Doppler Mapping Instrument
DDM	Delay Doppler Map
ECMWF	European Centre for Medium-Range Weather Forecasts
ERA5	ECMWF Reanalysis v5
GDAS	Global Data Assimilation System
GNSS-R	GNSS Reflectometry
GNSS	Global Navigation Satellite System
GMF	Geophysical Model Function
GPS	Global Positioning System
LHCP	Left-Handed Circular polarization
MSS	Mean Square Slope
NASA	National Aeronautics and Space Administration
NBRCS	Normalized Bi-static Radar Cross Section
OPENDAP	Open-source Project for a Network Data Access Protocol
OSCAR	Ocean Surface Current Analyses Real-time
PRN	Pseudorandom Noise
RHCP	Right-Handed Circular polarization
SP	Specular Point

Chapter 1

Introduction

The most common debris in lakes and oceans is plastic. Plastics are referred to as "microplastics" when smaller than five millimeters in length [US Department of Commerce and Administration, 2021]. Microplastics often originate from plastic debris being degraded into smaller particles by sun rays and other natural forces. An additional factor to why microplastics are so prevalent in the ocean is microbeads. Microbeads are tiny plastic products intentionally designed to be small and are often used in cosmetics and health products [US Department of Commerce and Administration, 2021]. They are illegal in a few countries, including the US. After consumers use these products, microbeads often end up in the sewer systems and later into the oceans. Researchers predict microplastics to double by 2030 in some ocean compartments [Hale et al., 2020].

It might be easy to only associate the pollution of marine life with leaking oil tankers or plastic waste dumped into the ocean. The small size of microplastics, being invisible to the naked eye, might make people overlook their existence and how it affects the ecosystem. In the 1970s, researchers began investigating the presence of tiny plastic particles in the ocean. However, not until a study was published in 2004 did research on the distribution and impact of microplastics effectively begin [Rochman, 2018].

The creator of the term "microplastics", marine biologist Richard Thompson, accompanied by professor of oceanography Kara Lavender Law, described the problem of microplastics in the following way: ≪Microplastics are of environmental concern because their size (millimeters or smaller) renders them accessible to a wide range of organisms at least as small as zooplankton, with potential for physical and toxicological harm≫ [Law and Thompson, 2014, p. 2]. Tiny organisms, fish, and other

animals ingesting microplastics make microplastics a part of the food chain. The combination of a contaminated food chain and the fact that most drinking water is contaminated with plastic particles make microplastics a part of most humans' everyday diet [Cox et al., 2019].

To assess how harmful microplastics are to the ecosystem, researchers need to obtain realistic estimates of the amounts and distribution of microplastics in the ocean. In 2015, Van Sebille et al. published a study that aimed to estimate the global distribution of microplastic mass and microplastic abundance on the ocean surface. The distribution models are well recognized by the scientific community. Despite their recognition, the models have uncertainties, as they are based on measurements made with surface trawling plankton synthesized with ocean circulation models.

Being able to globally monitor microplastic concentrations would be preferred. In 2021, Evans and Ruf published a study that aimed to detect microplastics and make global distributions from satellite measurements. A similar approach based on the same satellite data was studied by Strand in 2021.

Both approaches rely on reflected signals from the Global Navigation Satellite Systems (GNSS). The GNSS satellites are continuously emitting electromagnetic waves directed towards the earth's surface. The most prevalent GNSS is the American Global Positioning System (GPS) [Systems, 2021]. The American military developed GPS with the initial purpose of assisting the army. Despite this, the system has become a part of civilians' everyday life, being used in phones, cars, smartwatches, and many other products.

When the electromagnetic waves emitted by the satellites reach our devices, it helps us position ourselves and navigate. However, most signals never reach the products designed to receive them. The signals hitting trees, water bodies, or other vegetation reflect from the objects they hit. Already in the 1980s, it was proposed to sample reflected GNSS signals [Zavorotny et al., 2014]. The reflected signals carry information about the surface, demonstrated in many applications in recent years. Experts analyze the received signal strength, change in frequency, and change in travel time of the reflected signal to infer many geophysical parameters such as surface roughness and soil moisture. This concept makes up the foundation of GNSS Reflectometry (GNSS-R).

GNSS-R is relatively new in the microplastic detection domain. This thesis aims to further study if reflected GNSS signals carry information about floating microplastic concentrations when reflected on the ocean surface. This leads to the following research questions:

RQ1: What efforts have been made to estimate the global microplastic distribution in the oceans, and what are possible areas of improvement?

RQ2: Will suggested improvements to a GNSS-R approach to estimate the global microplastic distribution improve prior results?

RQ3: Is GNSS-R capable of measuring the global oceanic microplastic distribution?

Chapter 2

Theoretical Background

This chapter provides background information and theory considered prerequisite, aiming to mainly provide sufficient information on the origin of the observed data that fundaments the method.

2.1 GPS Signal Structure

GPS satellites use phase modulation to transmit information and are right-handed circular polarized (RHCP). Figure 2.1 shows a simplified version of what is modulated to highlight the relevant parts of the signal structure for this thesis. The illustration shows how the navigation message and the Pseudorandom noise (PRN) code are transmitted. GPS uses Modulo-2 addition to combine information from different sources. Modulo-2 addition results in 1 if the two sources of bits are the same and 0 otherwise. The combined signal is phase modulated by multiplying the carrier wave by -1 every time a bit contains 1, shifting the carrier by 180 degrees [Karaim, 2019].

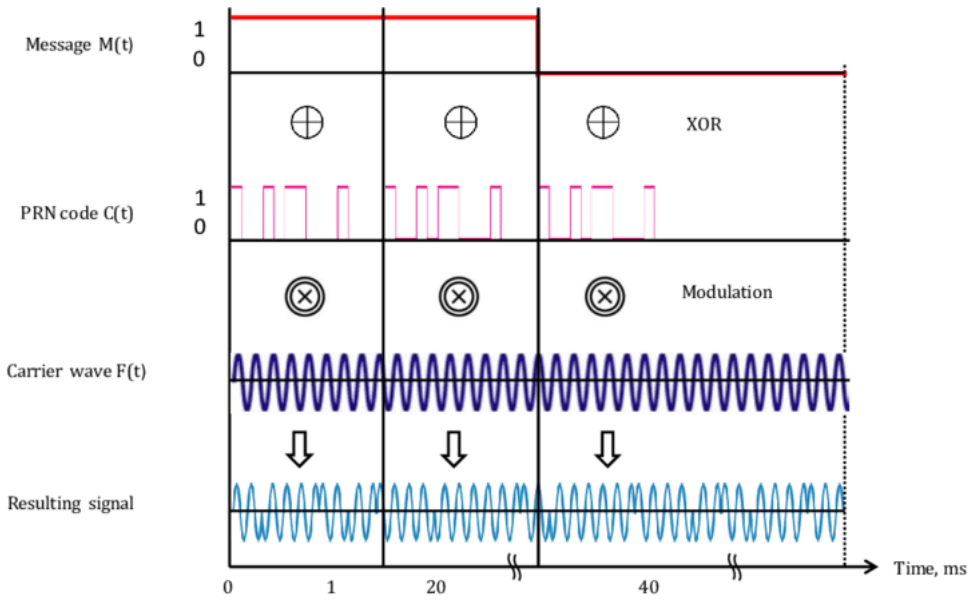


Figure 2.1: A simplified overview of the GPS signal structure. The transmitter combines the Navigation Message and the PRN code using Modulo-2 addition denoted with the XOR operation. The combined information is later modulated on the carrier wave using phase modulation.

Source: [Karaim, 2019]

A GPS receiver acquires GPS signals by first comparing the received signal with locally generated codes. The locally generated codes are PRN codes, identical to the PRN codes that the satellites in the GPS are transmitting. The PRN code functions as an identifier for a satellite. There does therefore exists a unique code for every satellite. The codes have a noise-like property for correlation purposes needed when differentiating the satellites from each other. The receiver can determine which satellite is broadcasting the received signal and retrieve the information by cross-correlating the received signal with the locally produced PRN codes. The highest cross-correlating value of all codes for all shifts and the received signal infer the broadcasting satellite and the broadcast time. The broadcast time is inferred as it is known by the receivers in advance the exact time each satellite broadcasts the code.

The modulated information can be reconstructed by despreading using Modulo-2 addition. The position of the transmitting satellites is known from the broadcast ephemerides, which are transmitted in the navigation message of the GPS signal. After acquiring the position of four satellites, both the distance equation and the GPS clock bias are solvable, giving the receiver sufficient information to infer its position [Wells et al., 1987].

2.2 GNSS Reflectometry

In GNSS-R, it is common to use a satellite arrangement where the transmitting satellite and the receiving satellite receiver are separated, as shown in Figure 2.2. A satellite arrangement like this is referred to as the Bi-static radar concept. One reason it has become the standard setup for GNSS-R is the low costs. The setup utilizes already existing satellites, which results in only the need to deploy receivers. The receivers are flying in low orbits and are less complex, often resulting in lower production and deployment costs [Willis, 2005]. Upon reflection, the polarization of the electromagnetic wave changes from RHCP to left-handed circular polarization (LHCP) [Ruf et al., 2016], making it easier for receivers to distinguish the direct signals from the reflected signals.

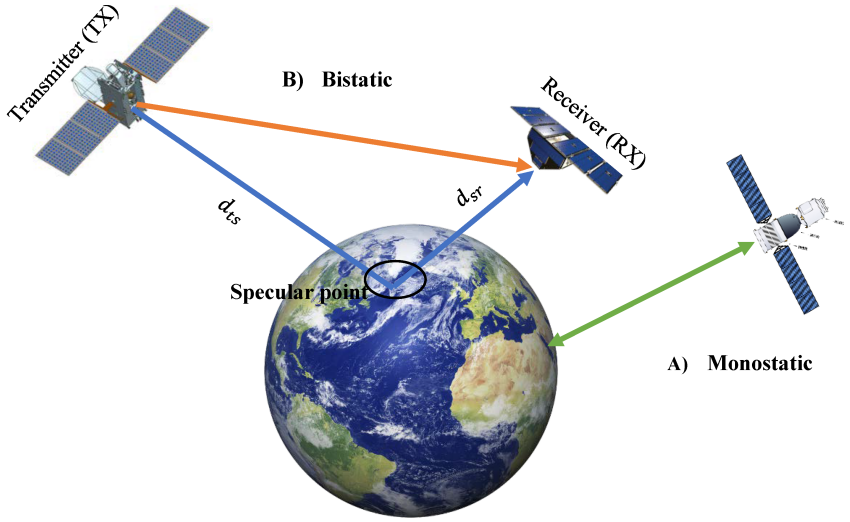


Figure 2.2: Illustration showing the (B) Bi-static and (A) mono-static radar concept.

Source: [Rajabi et al., 2020]

2.2.1 Cyclone Global Navigation Satellite System

Cyclone Global Navigation Satellite System (CYGNSS) is a satellite system owned and driven by National Aeronautics and Space Administration (NASA) that utilizes GPS. It is a constellation of eight low earth orbiting receivers launched to improve hurricane forecasting in tropical regions by using GNSS-R. The GPS and CYGNSS satellites form a Bi-static radar arrangement, where the CYGNSS satellites utilize both direct and reflected GPS signals. Figure 2.3 shows the spatial coverage of the CYGNSS satellites.

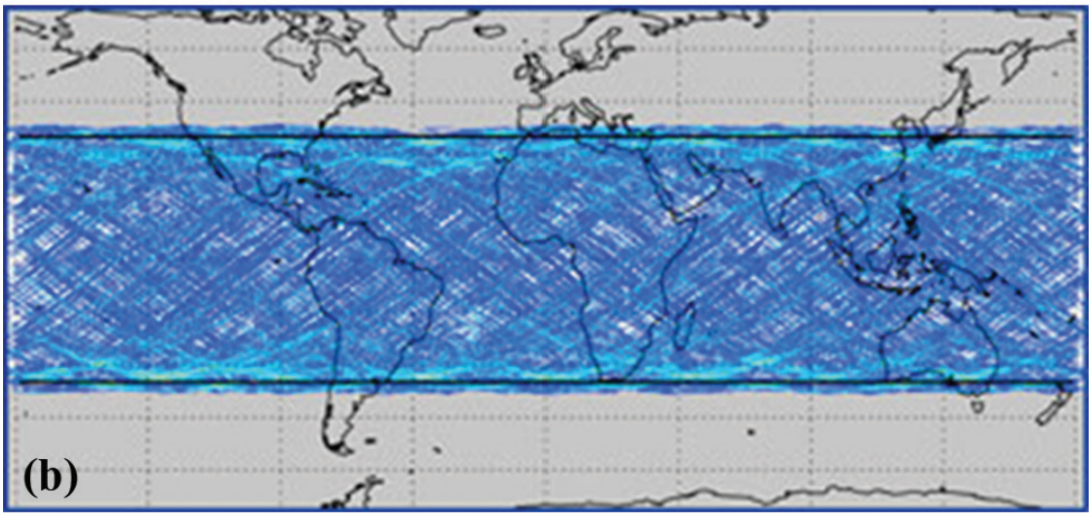


Figure 2.3: CYGNSS satellite tracks over a 24 hour period.

Source: [Ruf et al., 2016]

The data collected by CYGNSS is publicly available, making the scope of the application wider than measurements of wind speeds. NASA is publishing the data online on the NASA Physical Oceanography Data Active Archive Center (PO.DAAC). The data is published in four different levels, where the amount of processing is increasing per level. Level 1 data consists of delay Doppler maps (DDM) produced by the delay Doppler mapping instrument (DDMI) onboard each satellite. Level 2 data consists of Mean Square Slope (MMS) estimates, which is a measure of surface roughness. The two remaining levels are data over wind speed. Section 2.2.2 describes what DDMs are and briefly what is required to make them. Section 2.2.3 gives an overview of the procedure of how the CYGNSS satellites produce them.

2.2.2 Delay Doppler Maps

The fundamental measurement in GNSS-R is the DDM. It is produced by sampling the received reflected signal with different shifts in frequency and PRN code, where the shifts in PRN correspond to shifts in broadcast time [Gleason and Ruf, 2015]. Each unique shift in time and frequency combination results in various received signal power measurements, which constitutes the DDM. Two phenomena cause the spread of the signal. The relative motion between the satellite and receiver causes the spread in frequency, while the spread in time is caused by the different lengths that the signal travels. Figure 2.4 shows an example of a glistening zone which is the term used for the effective scattering region in GNSS-R [Yan et al., 2017]. The figure shows how the delay in time and shifts in Doppler are spread over the surface of

reflection. Reconstructing the glistening zone is feasible when both the location and velocity of the transmitter and receiver are known. The specular point (SP) is the location on the reflecting surface with the shortest distance between the transmitter via the reflecting surface to the receiver [Wu et al., 2018].

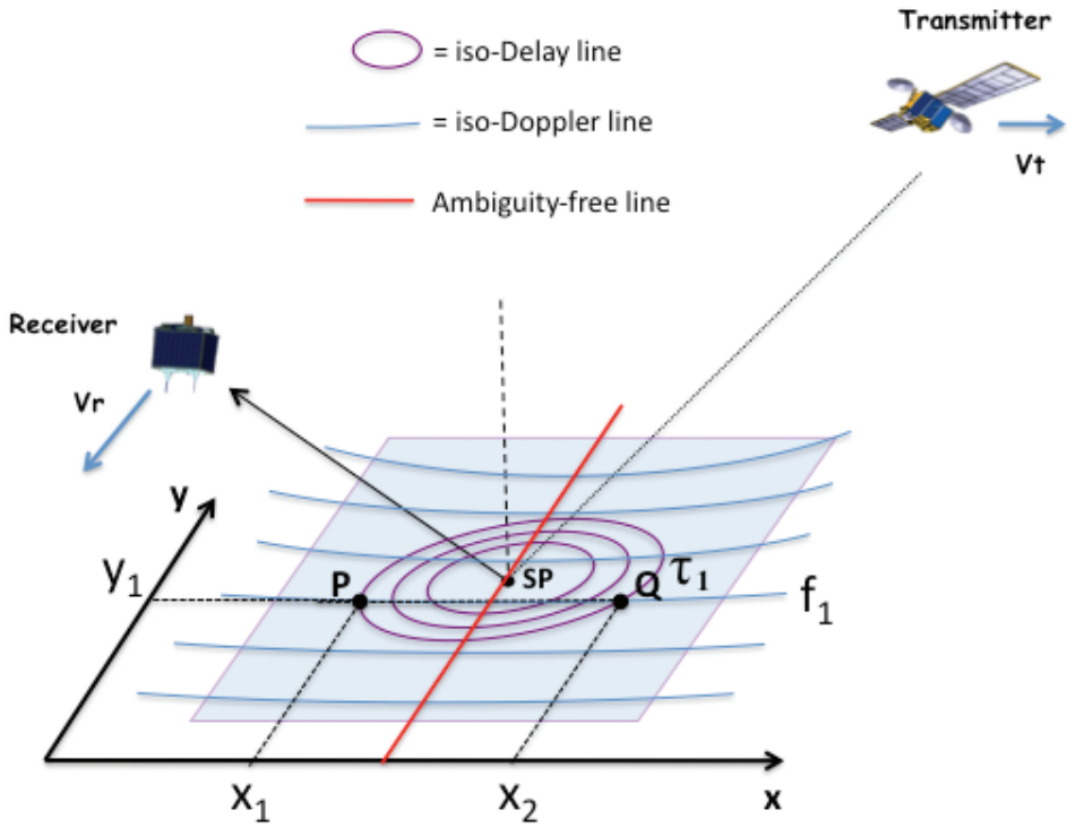


Figure 2.4: Illustration showing an example of spread in signal. The spread is caused by the different path lengths from transmitter to receiver via arbitrary reflection point and by the relative motion between the transmitter and receiver $\left\|\vec{V}_r - \vec{V}_t\right\|$. The two points, P and Q, reflect signals with the same shift in time and frequency to the receiver, as they are on the same iso-Delay τ_1 and iso-Doppler f_1 line, meaning that the path length and relative motion are equal for the two points.

Source: [Ruf et al., 2016]

Figure 2.5 shows two DDM produced by sampling reflected GNSS signals. The figure illustrates a typical pattern from reflections over different surfaces. A smoother surface like ice or calm water is resulting coherent reflections where the spread in signal in both time and frequency is small. On the other hand, a rougher surface like the ocean surface, which is affected by ocean winds, currents, and more, results in incoherent reflections where the spread is large. DDM's over the ocean often show the characteristic "horseshoe" shaped spread in the delay Doppler domain, like part

(b) of Figure 2.5 illustrates. The larger spread in the delay Doppler domain results in a larger glistening zone and weaker received signal power around the SP.

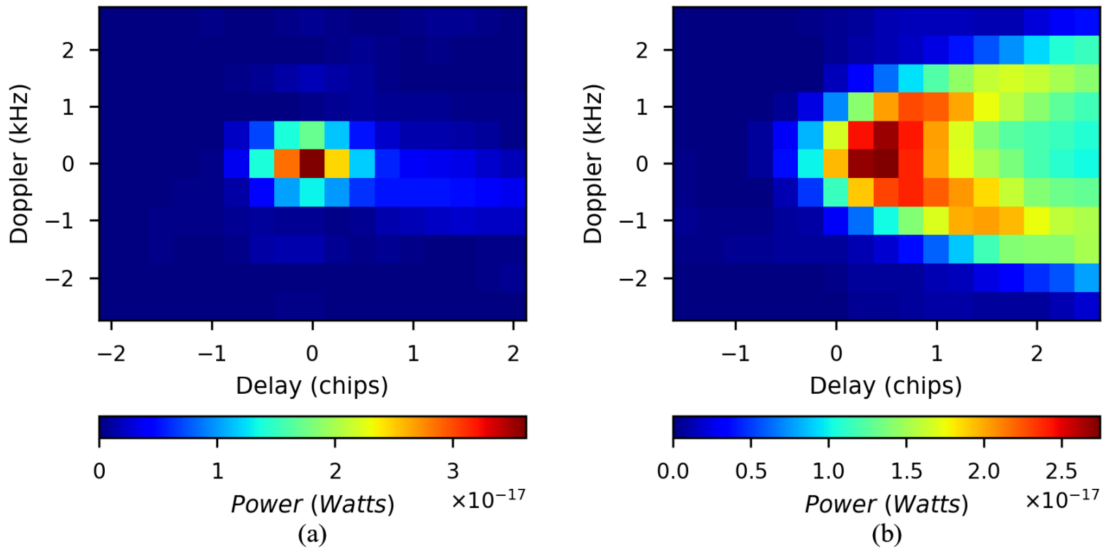


Figure 2.5: A calibrated delay Doppler Map over (a) ice/snow and (b) ocean, showing received signal power for different shifts in Delay and Doppler.

Source: [Dong and Jin, 2021]

DDMs are at acquisition showing how the unit "count" spreads in the time and Doppler domain. Unprocessed DDMs are referred to as raw DDMs. "Counts" are linearly proportional to signal power. The unit "count" lacks calibration for some errors and noise. The DDMs in Figure 2.5 are calibrated and are therefore showing the correctly obtained signal power in Watts. The calibration is done utilizing Equation 2.1

$$P_{g,\tau,f} = \frac{(C_{\tau,f} - C_N)(P_B + P_r)}{C_b}, \tag{2.1}$$

where $P_{g,\tau,f}$ is the sampled signal power in Watts for a specific time delay τ and Doppler f and C, τ, f is corresponding measured "count". The rest of the parameters are estimates and measurements for correction, where C_N is the noise measurement, P_B is the estimated blackbody load noise power measured, P_r is the calibrated receiver noise power, and C_B is the blackbody target measurement.

The level 1 data includes DDMs with different values. The first of the DDMs included shows the calibrated signal power $P_{g,\tau,f}$ and is referred to as L1A DDMs [Ruf et al., 2016]. The next sublevel, L1B data, consists of another DDM and a corresponding map of normalization areas. Similar to how noise and other error

estimates at every DDM time was prerequisite to the production of L1A data, there is information necessary for producing L1B data. The needed information consists of satellite GPS time, velocity, antenna gain, attitude information and position for the CYGNSS receiver, the position and velocity for the transmitting GPS satellite, and information about the transmitted power. Given all the above information, the signal power can further be utilized to produce another DDM with Bi-static Radar Cross Section (BRCS) values for each delay and Doppler. The calculation from signal power $P_{g,\tau,f}$ to BRCS $\bar{\sigma}_{\tau,f}$ are shown in Equation 2.2.

$$\bar{\sigma}_{\tau,f} = \frac{P_{g,t,f}(4\pi)^3 L_{a1} L_{a2} I_{\tau,f}}{P^T \lambda^2 G_{SP}^T G_{SP}^R R_{SP}^{Total}}, \tag{2.2}$$

where L_{a1} and L_{a2} are atmospherics loss corrections, $I_{\tau,f}$ are corrections for different losses caused by the DDMI, P^T is the transmitted power, G_{SP}^T is the antenna gain at the SP for the transmitter, G_{SP}^R is the antenna gain the SP for the receiver and R_{SP}^{Total} is the total range loss at the SP from the transmitter, via the surface, and to the receiver. The radar cross section does not have a physical meaning but is an essential measurement in electromagnetics, as it describes how detectable an object is by radar [Li et al., 2005].

The corresponding map of normalization areas is calculated by integrating two GPS ambiguities over the corresponding psychical areas related to each delay Doppler pair as shown in Equation 2.3.

$$\bar{A}_{\tau,f} = \iint_A \Lambda_{\tau;x,y}^2 S_{f;x,y}^2 \, dx \, dy, \tag{2.3}$$

where $\Lambda_{\tau;x,y}^2$ is the delay spreading function, and $S_{f;x,y}^2$ is the Doppler spreading function. The resulting effective area of surface scattering for each delay Doppler pair $\bar{A}_{\tau,f}$ and the BRCS $\bar{\sigma}_{\tau,f}$ can further be used to calculate the Normalized Bi-static Radar Cross Section (NBRCS) $\bar{\sigma}_0$ as shown in Equation 2.4.

$$\bar{\sigma}_0 = \frac{\bar{\sigma}_{total}}{\bar{A}_{total}} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{m} \bar{\sigma}_{\tau_i,f_j}}{\sum_{i=1}^{n} \sum_{j=1}^{m} \bar{A}_{\tau_i,f_j}} \tag{2.4}$$

By including all number of delay Doppler pairs from one observation used in Equation 2.4, the NBRCS is referred to as the scattering cross section. By only selecting the delay Doppler pairs surrounding the SP, the NBRCS is referred to as the maximum scattering cross section.

2.2.3 Delay Doppler Mapping Instrument

Each CYGNSS satellite is equipped with a DDMI, which produces the DDMs described in the previous section. The DDMI samples both the direct signals from higher orbiting GPS satellites and the reflected signals from the earth's surface simultaneously. An illustration describing the data flow from the unprocessed signals to DDM is shown in Figure 2.6. The signals are obtained in the DDMI by having two antennas, where one antenna points in the zenith direction sampling RHCP signals and the other one points in the nadir direction sampling the LHCP signals. The RHCP signals are direct signals from the higher orbiting GPS satellites. By acquiring and tracking the GPS signals, the position of both the transmitting satellites and the receiver is known. With the knowledge of these positions, the bi-static radar geometry can be calculated, which gives the location of the SP [Ruf et al., 2016].

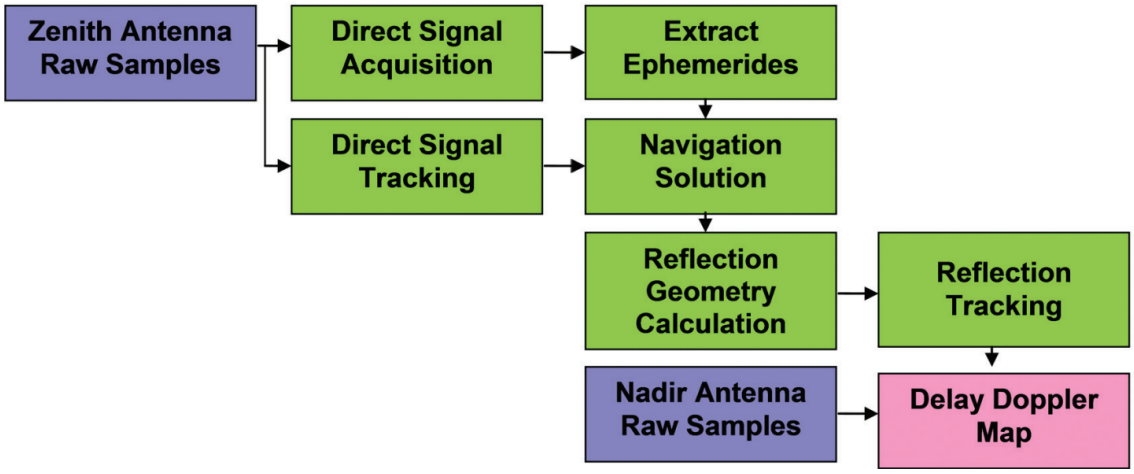


Figure 2.6: The Dataflow in the DDMI. The figure illustrates that the antenna pointing in the Zenith direction is responsible for calculating the bi-static radar geometry and geolocating the DDM, while the antenna in the Nadir direction produces the DDM.

Source: [Ruf et al., 2016]

The LHCP signals are signals reflected from the earth's surface. To process the reflected signals in real-time into DDM, the DDMI is equipped with a coprocessor and a regular navigation processor. The PRN code is generated locally on the receiver like a standard GPS receiver does, and it is used for both the acquisition of the direct signal and the processing of the reflected signal [Ruf et al., 2016].

Sampling the reflected signals and making raw DDMs consists of two main compo-

nents. Firstly, the carrier frequency is removed from the sampled signal by multiplying the data to a replica of the carrier. Secondly, the sampled data is cross-correlated with different shifts in time and Doppler to the locally generated PRN code, resulting in different count measurements for each shift, forming the raw DDM. Real-time processing is preferred as it reduces the amount of data to be stored. A computationally efficient way of producing DDMs is vital for real-time processing. In the DDMI, this is ensured by calculating the cross-correlation using matrix multiplication in the frequency domain. The use of Fast Fourier Transformations transforms the signal from the time domain to the frequency domain, which significantly accelerates calculation speed [Ruf et al., 2016].

2.2.4 DDMs over the Oceans Surface

The previous two sections, 2.2.2 and 2.2.3, gave a brief overview of what DDMs are and how they are produced. Figure 2.5 showed how the DDMs made from the CYGNSS satellites are affected by the surface on which the signal is reflected upon. Natural phenomena such as wind, swell, currents, surfactants, limited fetch, and bathymetry can alter the state of the ocean's surface [Ruf et al., 2016]. Wind is, however, the most impacting natural phenomenon and is contributing to roughening of the surface. The NBRCS derived from the DDM is sensitive to the roughness of the ocean's surface. Consequently, CYGNSS measurements can be used to estimate ocean wind speeds.

A geophysical model function (GMF) is a function that relates an observed variable to another geophysical variable of interest. The primary GMF used in the CYGNSS program relates the parameter correlated with received signal power, NBRCS, to ocean wind speed. It has been developed many wind speed retrieval algorithms, where one example shows promising results with a root mean square difference of only $1.51 m/s$ per second over wind speeds ranging up to $32 m/s$ [Reynolds et al., 2020]. The NBRCS are in these algorithms used to estimate ocean surface roughness, which in turn can estimate wind speed. The transition from ocean surface roughness to wind speed is done by assuming a nominal relationship between the estimated ocean roughness and wind speed [Evans and Ruf, 2021]. The measure of ocean roughness is mean square slope (MSS) and is defined in Equation 2.5.

$$MSS = \int_0^{k_c} k^2 S(k) dk, \tag{2.5}$$

where $S(k)$ is the omni-directional spectrum, k is the wavenumber, and k_c is the

cutoff wavenumber. The cutoff wavenumber depends on the carrier frequency from the reflected signal and the incidence angle. For the CYGNSS mission k_c is on average 7.5 rad m^{-1}. Equation 2.6 shows the relationship between the NBRCS $\bar{\sigma}_0$, which is measured in the CYGNSS mission, and MSS.

$$\bar{\sigma}_0 = \frac{\left|\mathfrak{R}(\theta)\right|^2}{MSS}, \tag{2.6}$$

where $\mathfrak{R}(\theta)$ is the Fresnel coefficient. The Fresnel coefficient is a function of incident angle θ and depends on signal polarization and a complex dielectric constant of the reflecting medium ϵ. The dielectric constant varies with the temperature and the salinity levels of the ocean.

Equation 2.5 is defining MSS as an integral over wavenumber. The top three images of Figure 2.7 show wavenumber versus theoretical wavenumber spectrum. The integral corresponds to the area under the curves. Other phenomena than wind can alter the state of the ocean's surface. The figures display that ocean currents, according to theory, impact MSS values. When considering the CYGNSS cutoff wavenumber, $k_c = 7.5$ or $10^{0.87}$ rad m^{-1}, the resulting changes in the area under the curve caused by surface currents are significant for all wind speeds.

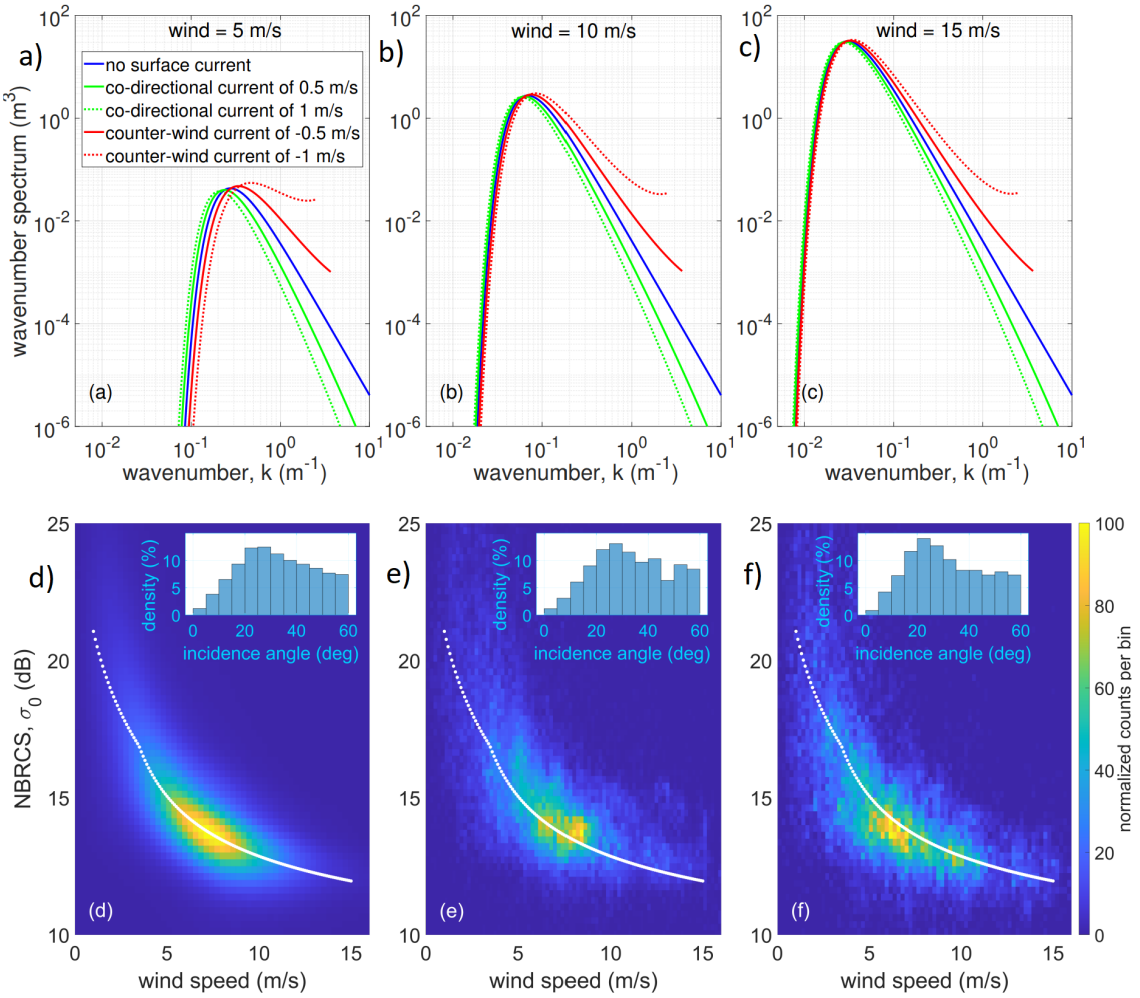


Figure 2.7: Theoretical wavenumber spectrum versus wavenumber for wind speeds of: (a) 5 m/s, (b) 10 m/s and (c) 15 m/s, under different current conditions based on the current-modified model from [Huang et al., 1972]. CYGNSS NBRCS measurements versus ERA5 wind speed with: (d) no surface current, (e) co-directional winds and currents, and (f) counter-wind currents. The histograms are the incidence angles associated with the measurements. The white line showing the [Katzberg et al., 2006] wind only MSS model.

Source: [Hoseini and Nahavandchi, 2022]

Equation 2.6 shows that NBRCS and MSS is inversely proportional. The figures in Figure 2.7 illustrates this. The top three figures show that counter directional wind and current produce higher MSS, as the area under the curve gets bigger. Part f) show a decrease in NBRCS for counter directional wind and current. Part e) shows that the same inversely proportional relationship exists for co-directional wind and current.

Chapter 3

Previous Works

This chapter briefly summarizes previous research which influenced the choice of working with remote detection of microplastics in this thesis. Three approaches that aim to estimate the global distribution of ocean microplastics are presented. After each method is presented, a brief assessment of drawbacks and possible improvement areas is highlighted. Firstly, well recognized and widely used estimations of microplastic distributions in the scientific community are presented. The chapter follows by presenting two different approaches taken in previous research to apply remote sensing to the microplastic detection domain. Lastly, the chapter presents a study that researches two possible causes of the observed reduction in ocean surface roughening, which is fundamental in the GNSS-R-based approaches that aim to detect floating ocean microplastics.

3.1 A Global Inventory of Small Floating Plastic Debris

In 2015, Van Sebille et al. published a study that aimed to estimate the global distribution of microplastic mass and microplastic abundance on the ocean surface. The distribution models presented in the study are well recognized by the scientific community. The study of new methods to detect microplastics often uses this study as a reference to compare results, which substantiates its recognition among researchers today [Evans and Ruf, 2021, Strand, 2021].

To obtain reasonable estimates, the study used as many of the microplastic measurements available at the time as possible. Data was identified by literature search

and were assembled directly from publications or by contacting the corresponding authors. The data retrieval resulted in 11854 observations from 27 studies.

The measurements were made with surface-trawling plankton nets of different grid sizes, ranging from $0.1mm$ to $0.5mm$. The net mouth size varied between studies, and the maximum particle size is not always reported. The measurements included debris larger than the net mesh and debris smaller than the net mouth. This resulted in the study changing what they referred to as microplastics to fit their data more conveniently. The definition was changed from plastic particles smaller than $0.5mm$ in size to all debris collected in the surface trawling plankton nets.

To further assess the reliability of the observations, Van Sebille et al. used a generalized additive model to evaluate the relationship between the variables associated with the observation and the measurement itself. The model showed that variables such as sampling year, wind speed, the distance of tow, and other factors directly impacted the concentration of microplastics or the representativeness of the observation. The resulting relationships were then used to remove variability associated with these factors, resulting in a standardized data set.

Figure 3.1 shows the resulting standardized data set of microplastic mass and abundance based on the surface trawling net measurements. The data set consists of observations from all major ocean basins except the Arctic Ocean. The figure shows the spatial imbalance in the data, where most observations lie in the Pacific Ocean.

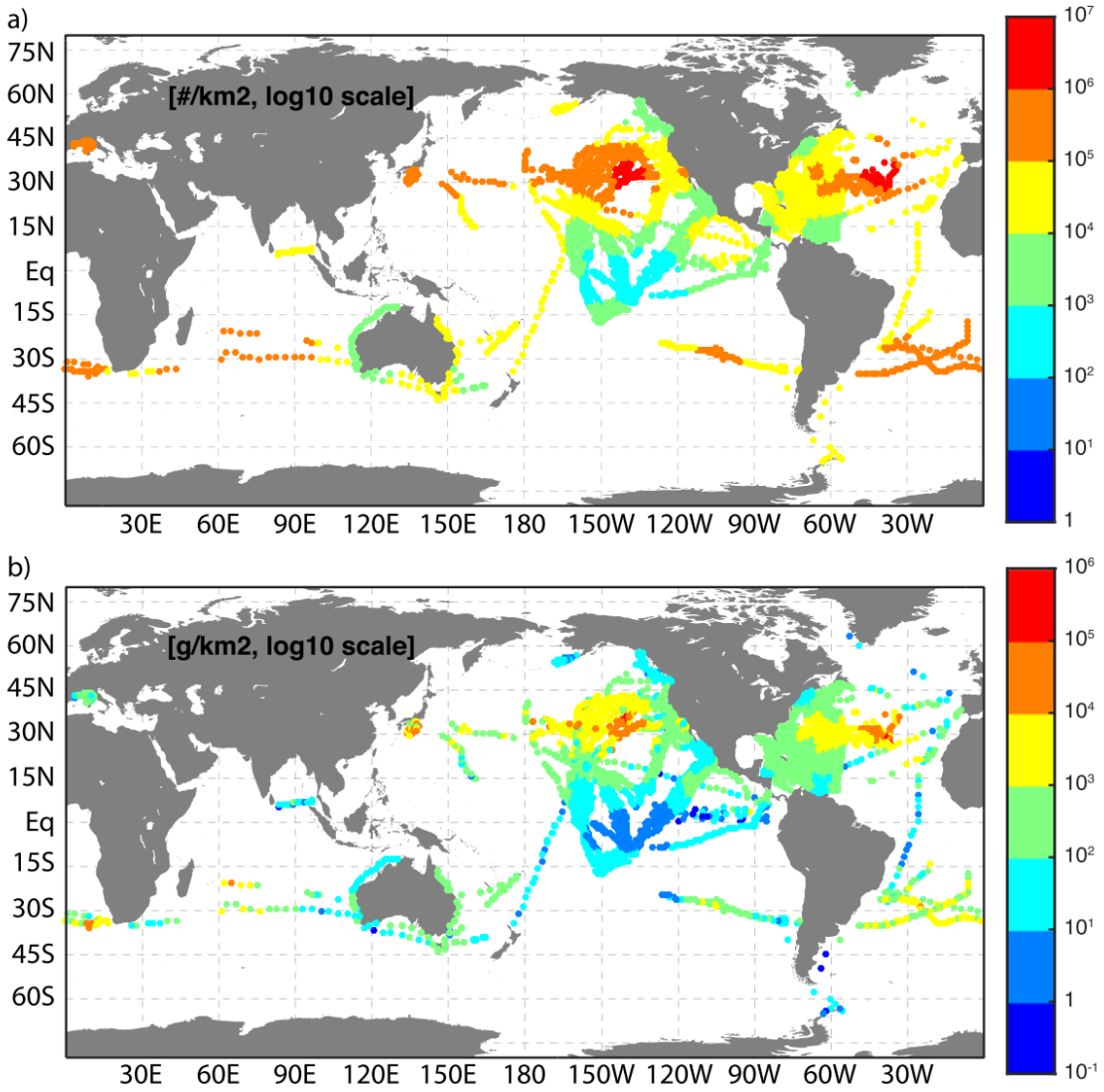


Figure 3.1: The standardized microplastic (a) count and (b) mass made from measurements with surface trawl nets.

Source: [Van Sebille et al., 2015]

The study aimed to produce global distributions of microplastics. Spatial interpolation is needed to go from the non-uniformly distribution in Figure 3.1 to global distribution. Van Sebille et al. hypothesized that more realistic results are obtained by synthesizing the observations with the predictions from ocean circulation models, compared to using standard interpolation methods such as Linear or Kriging interpolation. In practice, this means that they ran simulations with virtual microplastics based on the surface trawling observations put into the ocean current models. They developed three microplastic distribution models based on three different ocean circulation models. The circulation models they selected were highly independent of each other. The selection of models independent of each other was

useful. It enabled the authors to monitor how impacting the choice of the ocean circulation model was on the final distributions. The ocean circulation models they used in the study were made in [Maximenko et al., 2012], [Lebreton et al., 2012] and [Van Sebille et al., 2012].

Humanity has not stopped polluting the oceans with plastics. The authors tried to compensate for the continuous addition of plastics to the oceans. The different microplastic models they developed make different assumptions about how more microplastics enter the oceans. Lastly, the models differed from each other by making different assumptions about whether plastic particles could exit the system or not.

The Lebreton model assumes that microplastics originate from coastlines scaled by the size population living there, major river mounts scaled by urban development, and major shipping routes scaled by the amount of traffic among the route. The simulation is based on ocean velocity fields advecting the virtual microplastics and was run for 30 years. Virtual plastic particles were continuously added to the oceans throughout the simulation. The amounts of plastics added to the system were based on a data set from 2009 containing data on global plastic production. Plastics were not able to exit the system during the Lebreton simulation.

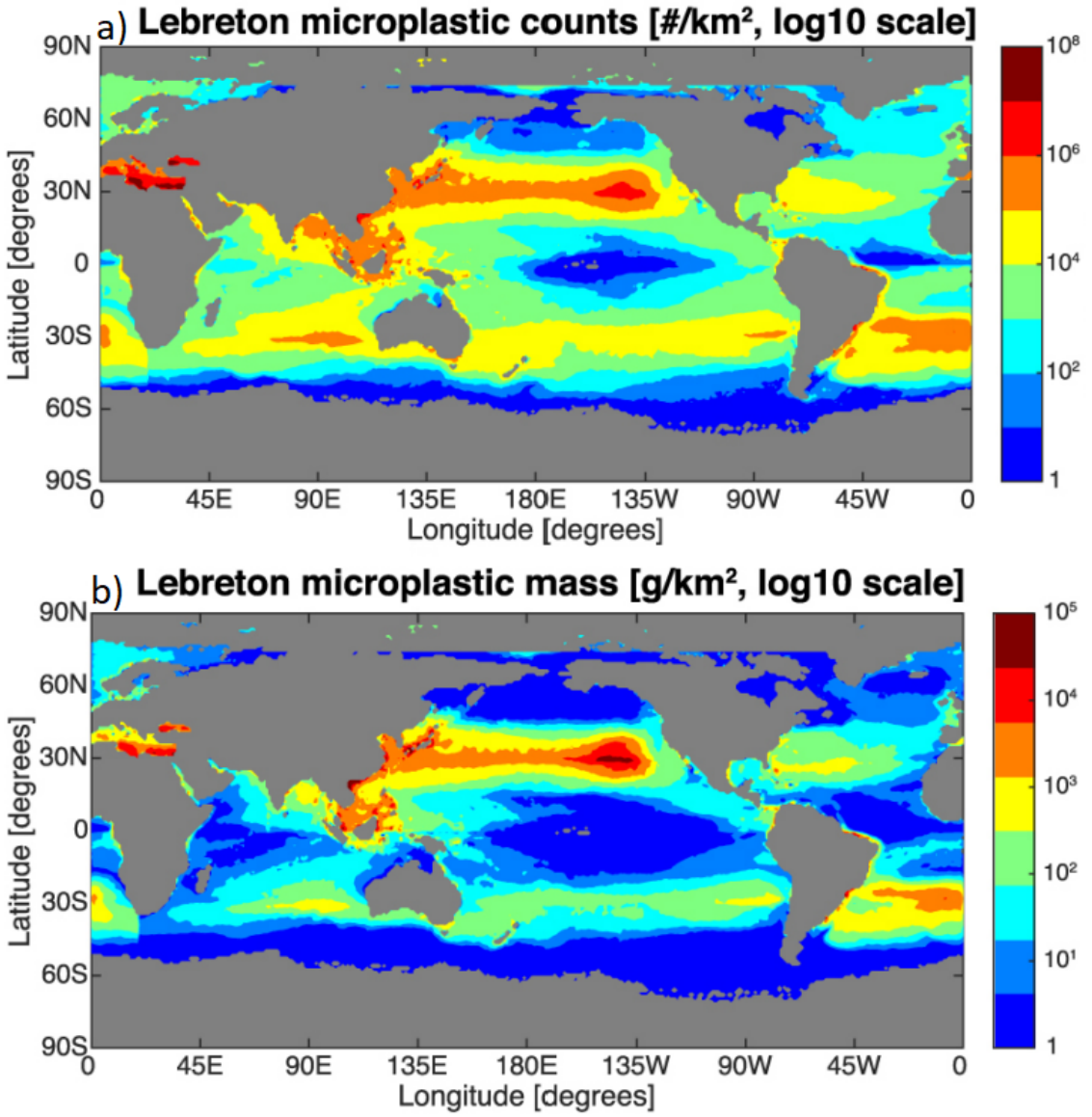


Figure 3.2: The distribution of microplastic (a) count and (b) mass predicted by the Lebreton model.

Source: [Van Sebille et al., 2015]

The van Sebille model assumes that microplastics originate from coastlines where the amount is proportional to the population size within 200 kilometers from the coast. In addition, the amount is scaled by how good countries are at waste management. The virtual microplastics were advected by applying a transition matrix based on a historical database derived from satellite tracking of drifting buoys. The simulation was run for 50 years, where plastics were continuously added to the oceans. The amounts of plastics added to the system were based on a data set from 2013 containing data on global plastic production. Plastics were not able to exit the system during the van Sebille simulation.

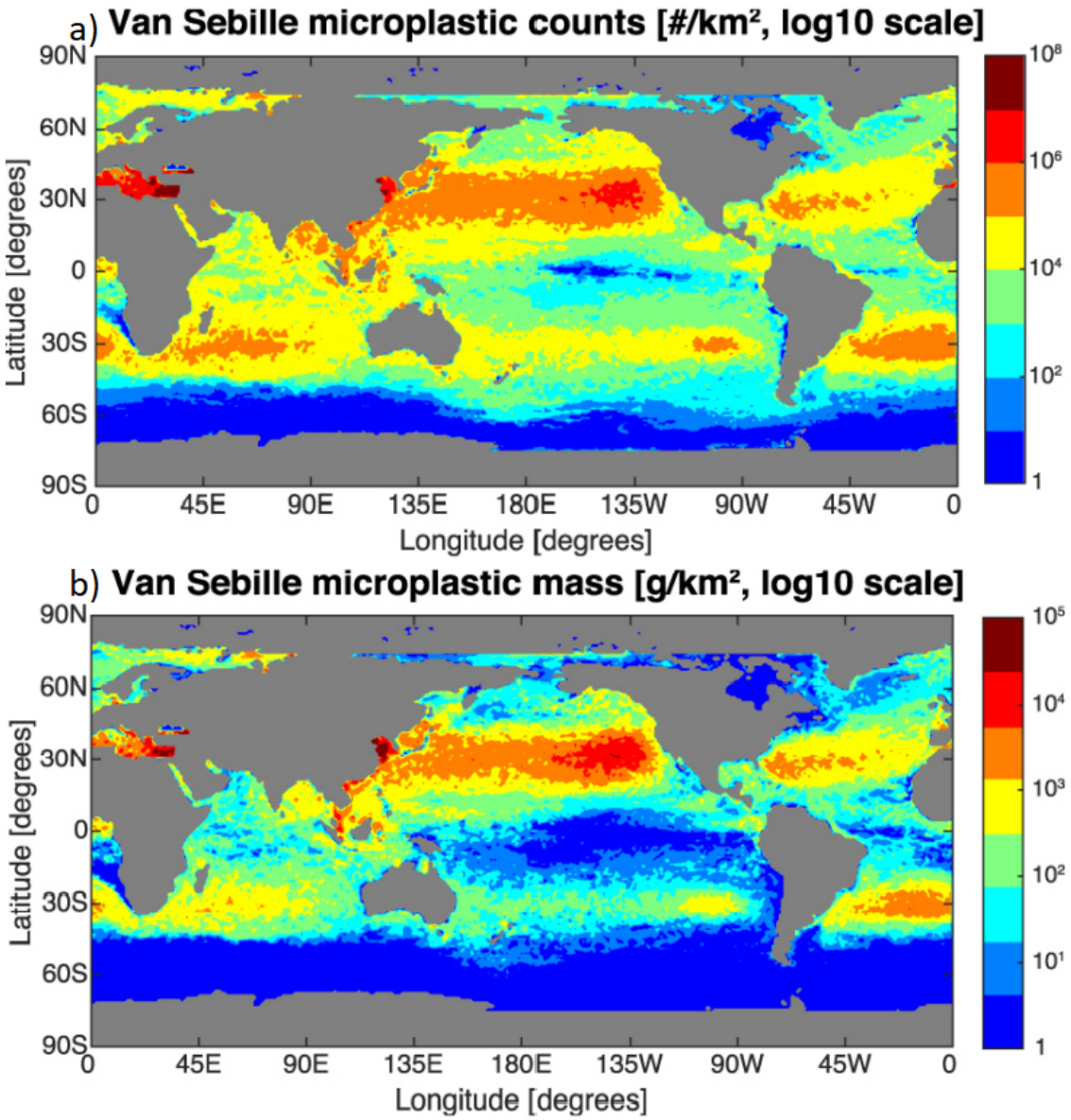


Figure 3.3: The distribution of microplastic (a) count and (b) mass predicted by the Van Sebille model.

Source: [Van Sebille et al., 2015]

The Maximenko model differs from the other models because it assumes that the source of microplastics is uniformly distributed over the ocean. Like the van Sebille model, the same transition matrix approach was used based on the same historical data. The Maximenko model was the only model that allowed microplastics to exit the simulation in the form of being "washed ashore".

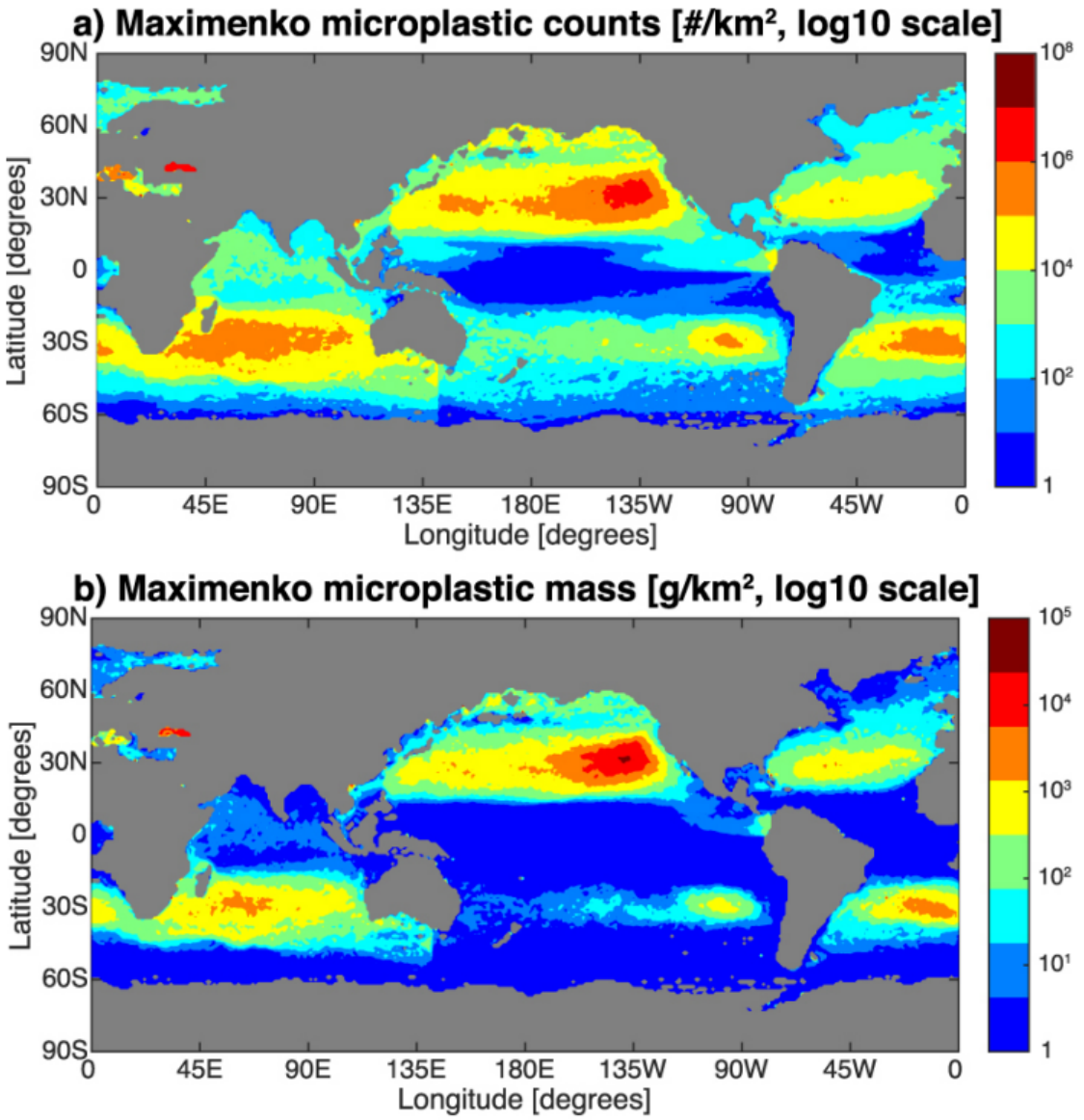


Figure 3.4: The distribution of microplastic (a) count and (b) mass predicted by the Maximenko model.

Source: [Van Sebille et al., 2015]

3.1.1 Drawbacks and Possible Improvement Areas

The quality of data and chosen research methodology affect the uncertainty of the estimated microplastic distributions. The bullet points list factors affecting the uncertainties in the models and is further elaborated below.

Data Quality

- Number of measurements
- Imbalanced spatial coverage
- Various debris captured
- Various data sources

Firstly, the estimates in the research are based on data from surface trawling nets which has its drawbacks. Fewer measurements generally result in worse estimates. An issue with the data is the lack of global coverage. Figure 3.1 shows the imbalance in the data, making estimates likely worse in regions with low coverage. The surface trawling nets also collect other buoyant debris than microplastics. The various sizes of collected debris, constrained by the net mouth and mesh size, also introduce errors. The study relied on data from multiple sources with no common standard of how measurements were made and at different points in time.

Methodology

- Data standardization
- Ocean circulation model
- Assumptions of origin and exiting

The standardization was done to improve the quality of the data. How good it was done affects the estimates. The choice of ocean circulation model and the assumptions of origin and exiting of microplastics affect the results, which can be seen by comparing the differences between the three resulting microplastic models.

3.2 Toward the Detection and Imaging of Ocean Microplastics with a Spaceborne Radar

In 2021, Evans and Ruf published a study that aimed to detect microplastics and make global distributions from satellite measurements. The study was based on one year of CYGNSS measurements supplemented with Global Data Assimilation System (GDAS) wind data, where the data ranged from June 1, 2017, until May 31, 2018.

The authors hypothesized that surfactants share similar transport mechanisms as microplastics, making surfactants a good tracer for microplastics. The hypothesis further states that surfactants can suppress the effects of wind-driven roughening of the ocean's surface, which MSS measures. Followingly, their presented approach relies on MSS observations from the CYGNSS satellites. Evans and Ruf based their hypothesis on presented evidence in the paper, where [Alpers and Hühnerfuss, 1989, Spivak et al., 2002, Kiefhaber et al., 2015] was listed as evidence for surfactants contributing to the suppression of wind-driven roughening and [Van Sebille et al., 2020] as evidence that surfactants share similar ocean transport mechanisms as microplastics.

Wind as the primary driver behind the roughening of the ocean's surface can be used to model MSS and vice versa. However, there are anomalies between the two, which the authors related to the presence of surfactants or microplastics.

The following paragraphs present how the authors tried to distinguish the contribution of wind and surfactants to MSS observations from each other in an attempt to model microplastics based on MSS. In trying to isolate the effects of wind on the MSS observations, the authors colocated CYGNSS MSS observations with GDAS wind data in two regions with believed low microplastic concentration using the full year of data. Figure 3.5 shows the selected regions with coordinates [10°–25°S, 105°–120°E] and [10°–20°N,128°–143°E]. The two control regions were chosen based on the global microplastic distribution models produced by Van Sebille et al.. The Van Sebille microplastic model shows areas with lower concentrations than the two control regions chosen. Despite this, the authors made their selection by arguing that these regions were the best candidates because they were outside the intertropical convergence zone, where MSS observations are more unreliable [Balasubramaniam and Ruf, 2018].

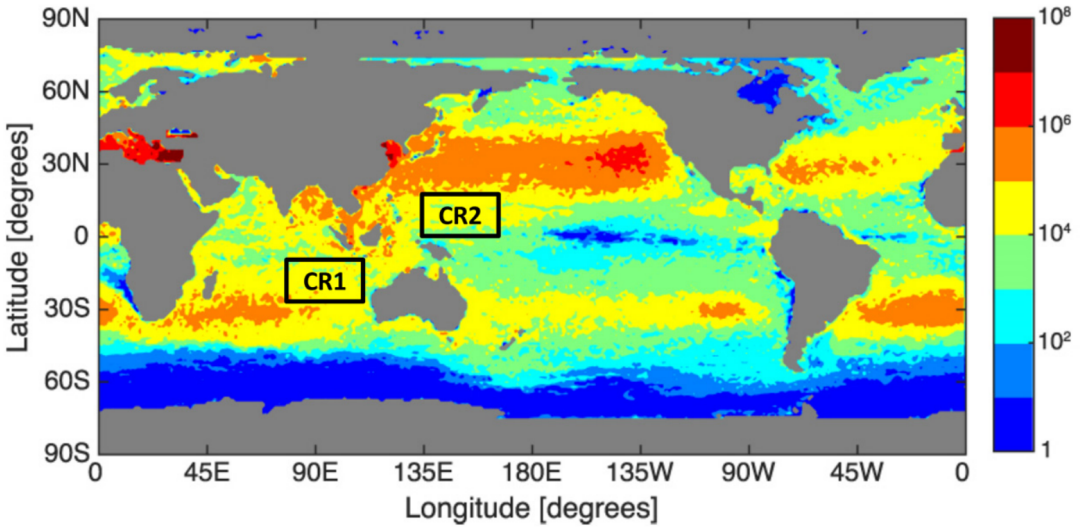


Figure 3.5: The two control regions shown on the distribution predicted by the Van Sebille microplastic model from Figure 3.3.

Source: [Evans and Ruf, 2021]

Evans and Ruf used the colocated GDAS wind data to build an empirical model with the intention of predicting the MSS that CYGNSS satellites observe based on wind data. Figure 3.6 shows a density scatter plot of GDAS wind speed versus CYGNSS MSS observations from the two control regions. The black line shows the developed empirical model.

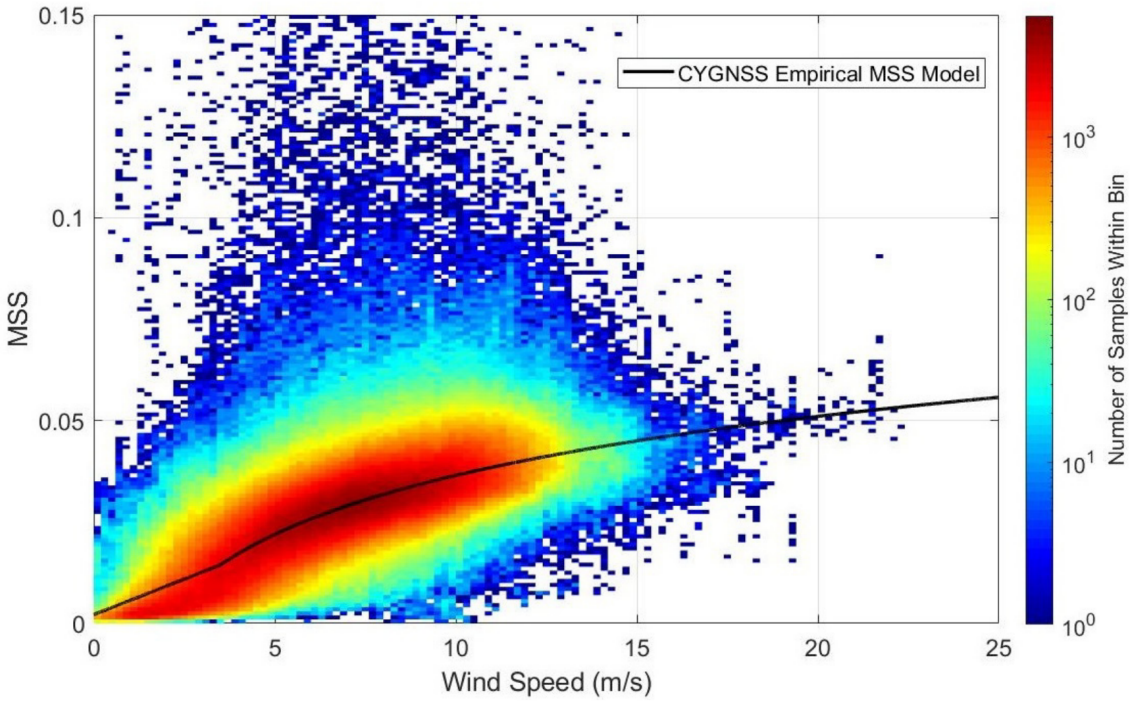


Figure 3.6: CYGNSS MSS observations versus GDAS wind speed.

Source: [Evans and Ruf, 2021]

The empirical relationship was developed by least-squares refitting the empirical L-Band MSS model proposed by Katzberg et al. in 2006. The refitting resulted in the model for MSS in Equation 3.1, which gives modeled MSS based on the current wind conditions.

$$MSS_{mod} = \begin{cases} 0.0035(\left|\vec{U}_{10}\right| + 0.62), & \text{if } \left|\vec{U}_{10}\right| \leq 3.49 m/s \\ 0.0035(6ln(\left|\vec{U}_{10}\right|) - 3.39), & \text{if } \left|\vec{U}_{10}\right| > 3.49 m/s \end{cases} \tag{3.1}$$

It was expected that the model in Equation 3.1 would predict a rougher surface when applied to other areas where microplastic concentrations are higher compared to the observed MSS values from CYGNSS. The hypothesis explains the expectance of higher predictions, as it states that microplastics are accompanied by surfactants, which suppresses wind-driven roughening. The developed microplastic predictor is fundamentally based on principle, as it quantifies microplastic concentration based on how much lower the observed MSS values are compared to the modeled values. The predictor is referred to as the MSS anomaly and is defined as

$$MSS_{anomaly} = \frac{MSS_{obs} - MSS_{mod}}{MSS_{mod}}, \tag{3.2}$$

where MSS_{mod} is defined in Equation 3.1 and MSS_{obs} is the observed MSS from the CYGNSS measurement. The normalized anomaly was used, enabling the authors to directly compare observations of MSS anomaly with different degree of roughness with each other. A negative MSS anomaly indicates that the surface is smoother than expected and was therefore being related to the presence of microplastics, where an increasingly negative anomaly yields increasingly more microplastics, according to the hypothesis.

After developing the empirical relationship between wind and MSS in believed low microplastic concentration regions, normalized MSS anomaly for each CYGNSS observation for the entire year of data was calculated. Later they spatially averaged the anomaly observations into a 1° x 1° grid and temporally averaged the observations over the entire year of data. A comparison of annually gridded normalized MSS anomaly observations averaged into bins of width 0.005, and microplastic number density is shown in Figure 3.7.

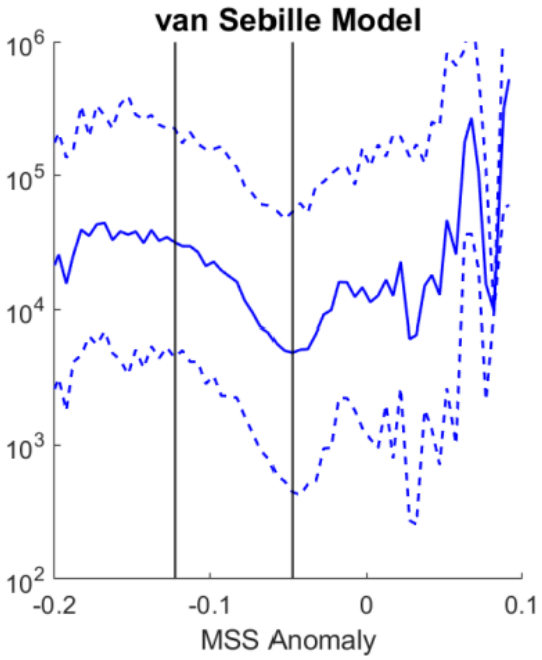


Figure 3.7: Annually gridded normalized MSS anomaly observations separated into bins of width 0.005 versus microplastic count. The dashed lines show the minimum and maximum value within the bins. The black lines denote the region of anomaly ranging from -0.1227 and -0.0478.

Source: [Evans and Ruf, 2021]

Figure 3.7 shows that there is little evidence for any correlation between the two variables over the entire range. The comparison does, however, indicate a strong

correlation between the variables in the sub-range denoted by the black lines in the figure, where ~ 61 % of the samples lie.

An empirical relationship was developed between the anomaly observations within the sub-range denoted by the black lines in Figure 3.7 and the estimated microplastic model from Van Sebille et al. to rescale the normalized MSS anomaly observations to fit believed concentrations better. The empirical relationship was based on the microplastic model shown in Figure 3.3, as it showed the highest correlation. According to the authors, selecting the other models produced similar results. Equation 3.3 shows the empirical relationship.

$$p = 2035exp(-23.18MSS_{anom}), \tag{3.3}$$

The relationship is exponential even though it first appears linear in Figure 3.7, as the microplastic count is logarithmically scaled on the axis in the figure. Figure 3.8 shows the resulting microplastic concentration maps produced by the study compared to the Van Sebille model from Figure 3.3. The authors drew the boxes on the distributions to argue for good model agreement.

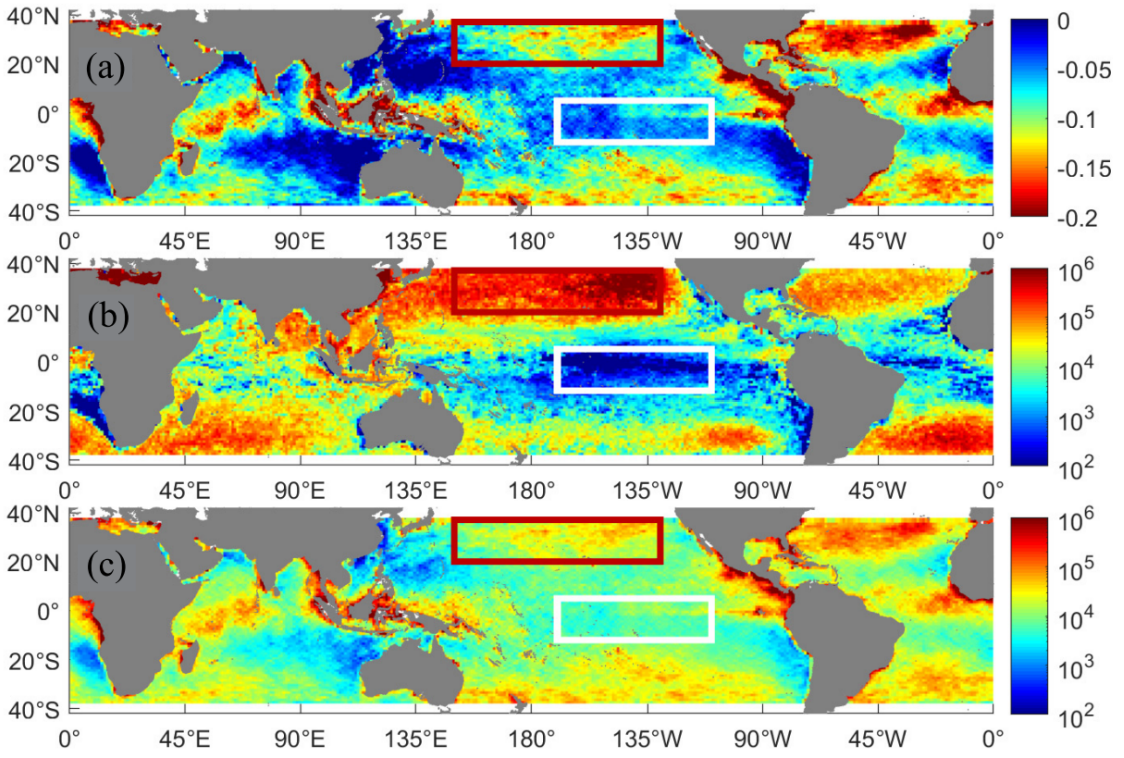


Figure 3.8: Global distribution of the (a) MSS anomaly observations, (b) predicted microplastic concentration made by the Van Sebille microplastic model from Figure 3.3 and the (c) predicted microplastic concentration based on MSS anomaly observations inputted into Equation 3.3.

Source: [Evans and Ruf, 2021]

The proposed method is unlike previous attempts to create global microplastic distributions trying to estimate microplastic distributions on a shorter temporal scale. This enabled the research of seasonal dependencies in the data. Figure 3.9 shows a times series of how monthly averaged predicted microplastic concentration varies globally and in the North- and South Pacific Oceans. The authors reported that the concentrations are highest in both hemispheres in their respective summer months and lowest in their winter months.

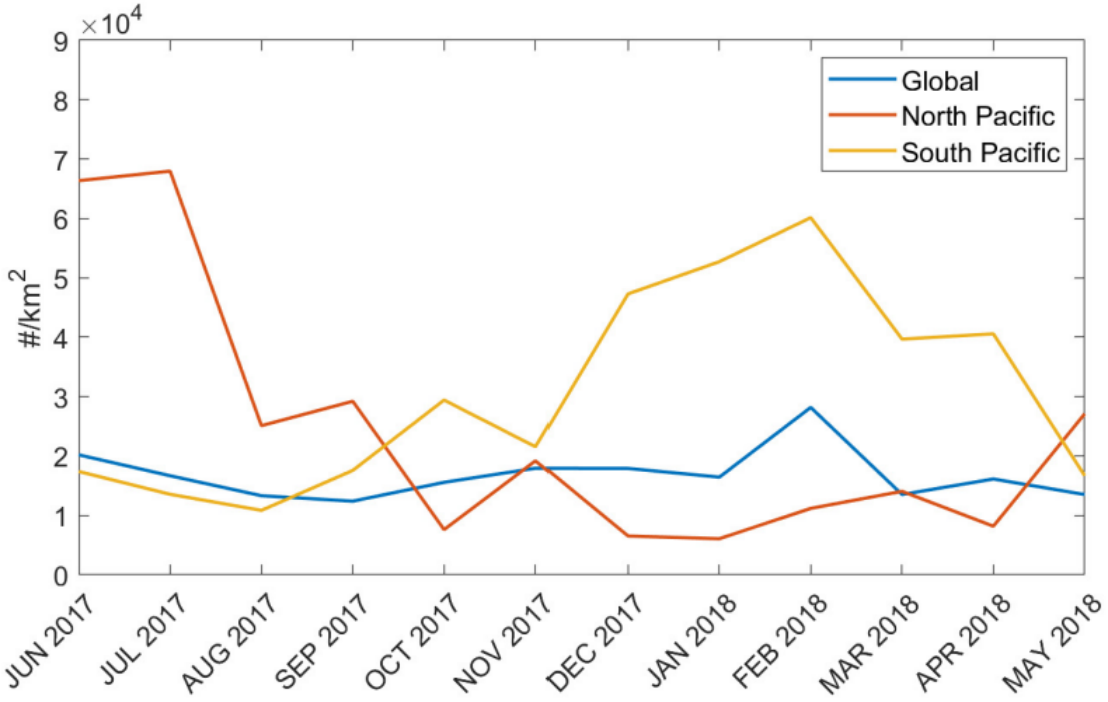
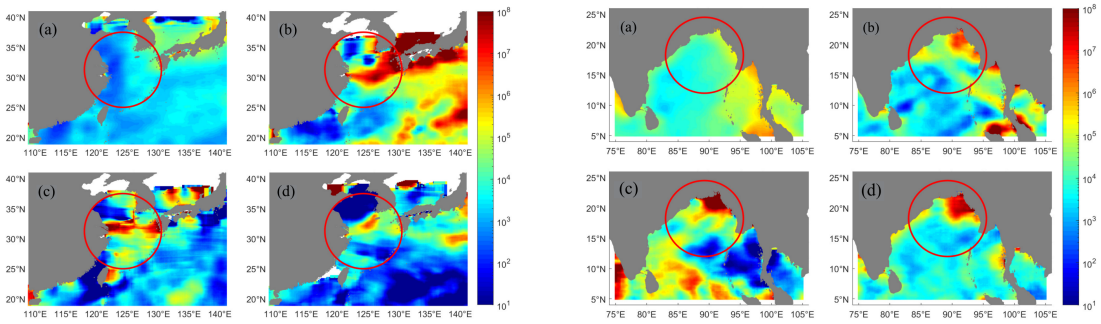


Figure 3.9: Monthly averaged predicted microplastic concentration for the North- and South Pacific Ocean. The global average is also shown.

Source: [Evans and Ruf, 2021]

The shorter temporal scale also enabled better investigations of local phenomena, like activities around river mounts like Figure 3.10 is showing.



(a) East China Sea (b) Bay of Bengal

Figure 3.10: Predicted microplastic concentrations. In both figures: a) is showing the annual average, while b), c) and d) is showing weekly averages. The drawn red circle shows major river mouths.

Source: [Evans and Ruf, 2021]

3.2.1 Drawbacks and Possible Improvement Areas

Figure 3.7 is arguably the main finding of the research. Most samples lie within the boundaries indicating a strong correlation between Van Sebille et al. estimated microplastics and the MSS anomaly. The 39 percent lying outside the relationship is described as corresponding areas highly sensitive/insensitive to wind forcing. No further investigations are done to understand why close to half of the observations violate the assumed relationship. Plotting the MSS anomaly observations outside the -0.1227 to -0.00478 range where good correlation is seen may partly explain the undesired behavior. What if most observations are along coastlines, where bathymetry is more impacting on the ocean surface, or in areas with other known natural phenomena altering the ocean surface, like strong ocean currents? Perhaps the authors investigated it and saw that the undesired observations were randomly distributed, which speaks against the method's validity.

On the other hand, the microplastic models of Van Sebille et al. are only estimated distributions, not actual distributions. Hence, they are interesting for comparison, but it is incorrect to consider them ground truth. This may partly explain the 39 percent violating the assumed relationship. Nevertheless, it may seem like the authors are effect by slight confirmation bias when not further investigating the undesired behavior.

Wind is the primary driver of ocean surface roughness, but other natural phenomena like currents, eddies, and bathymetry contribute to MSS variations. Even when assuming that the hypothesis is entirely accurate, these phenomena need to be modeled to compensate for their contribution to MSS observations. The task of modeling all the natural phenomena affecting surface roughness is not feasible, implying that perfect results in impossible. This drawback exists for both GNSS-r based approaches.

Solving many of the above-mentioned issues is complicated or simply not feasible. There do, however, exist possible improvements to the method, which is easier to consider that should improve estimations. Possible improvements are listed below.

- Little of the available data is used. More observations should give better estimates.
- The Katzberg model is old and likely outdated. Today's state-of-the-art regression models could improve the wind-based MSS modeling.
- The CYGNSS data used is version 2.1, which is not compensated for certain errors. Version 3.0 data should be preferred.

3.3 The Application of Spaceborne GNSS-R for the Study of Ocean Microplastics

In 2021, Strand studied remote detection of ocean microplastics in his thesis. His work was similar to Evans and Ruf's work, also based on CYGNSS and wind data. Strand did, however, collocate his data with ERA5 wind speeds instead of GDAS wind speeds. He based his study on CYGNSS and ERA5 data ranging from March 18, 2017, to May 31, 2021. Strand hypothesize that the presence of microplastics increases the NBRCS value, which is measured by the CYGNSS satellites. His explanation for the phenomena is based on the assumption that microplastics can suppress ocean surface roughness.

In the approach of trying to isolate the impact of microplastics, NBRCS values along a CYGNSS satellite track are monitored. A satellite track is a set of repeated measurements that follows the orbit of one of the satellites in the constellation. Strand split a track if the distance between two measurements is more than 100 kilometers. In addition, a track was not considered valid if it contained less than 50 measurements. Microplastics are assumed to cause the peaks in NBRCS along a track.

The measurements are first smoothed to continuous values to more easily detect peaks in NBRCS. The measurements are smoothed by averaging neighboring points. The number of neighboring points that are considered is not reported. Later peaks are extracted from the smoothed continuous data. The number of peaks in the data is dependent on the threshold set for what defines a peak. Two parameters commonly define a peak in signal theory. The first is width, which describes the

width of the peak, and the second is prominence, which describes how much the peak stands out from the baseline of the signal. Strand used the function: *find_peaks(data, prominence, width)*, from the *Scipy.signal* python package to detect peaks, where width of 5 and prominence of 1 used. Figure 3.11 shows an example of the resulting NBRCS peaks based on the peak detection method.

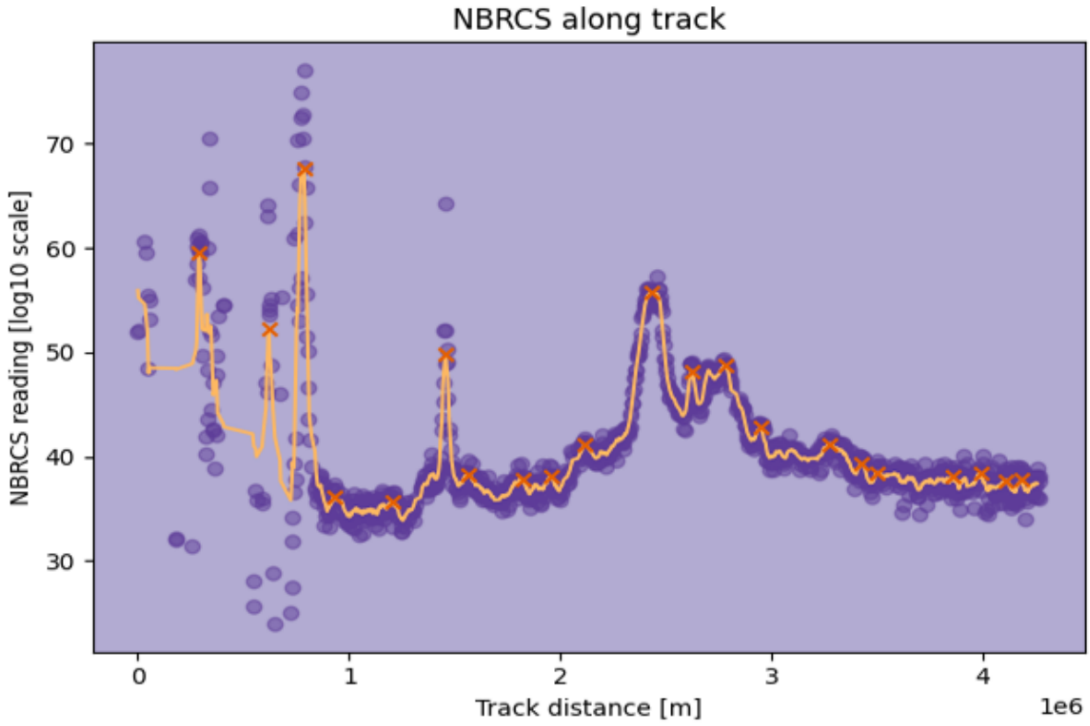


Figure 3.11: NBRCS peak detection along a CYGNSS satellite track. NBRCS measurements are shown as purple dots, the smoothed NBRCS values as a yellow curve and the peaks marks with red crosses.

Source: [Strand, 2021]

After developing the peak detection method, Strand made observations of peaks using all the data. The observations were then spatially averaged, and the number of peaks was counted in 1° x 1° bins. Global peak distributions are shown in Figure 3.12. More peaks were believed to indicate more microplastics.

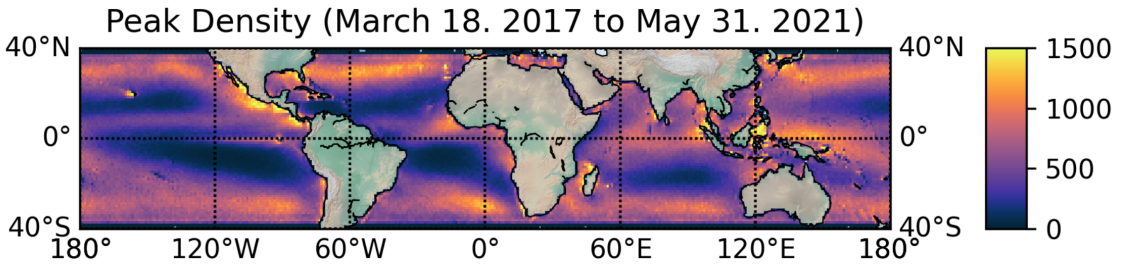


Figure 3.12: Global average peak density.

Source: [Strand, 2021]

Strand stated that ocean surface currents could contaminate the results by affecting the number of peaks in areas with strong currents. He later argued that the effects of ocean surface currents are less noticeable during lower wind speeds. He then presented global density peak distributions where data are filtered out based on wind speeds. All observations outside the range of 3 to 5 meters per second wind speeds were removed. The resulting filtered peak distribution is shown in Figure 3.13. Strand claimed to see changes in areas with strong surface currents and little change in other regions.

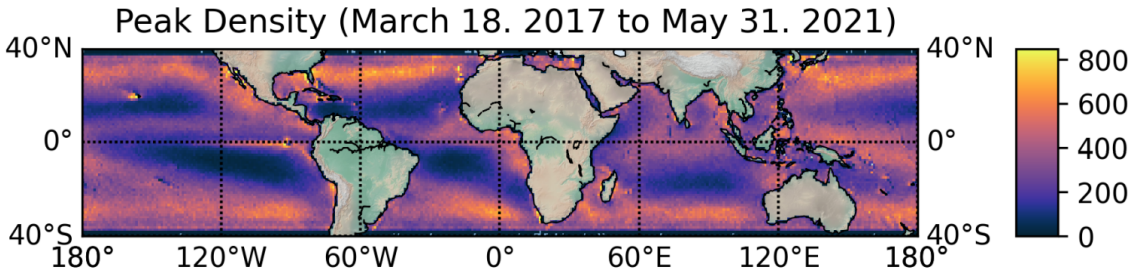


Figure 3.13: Global average peak density for observations with wind speeds ranging between 3 to 5 meters per second.

Source: [Strand, 2021]

Similar to the other GNSS-R method, this approach enabled the investigation of the microplastics indicator on shorter temporal scales. Figure 3.14 shows a times series of how the peak density varies globally and in the North- and South Pacific Oceans.

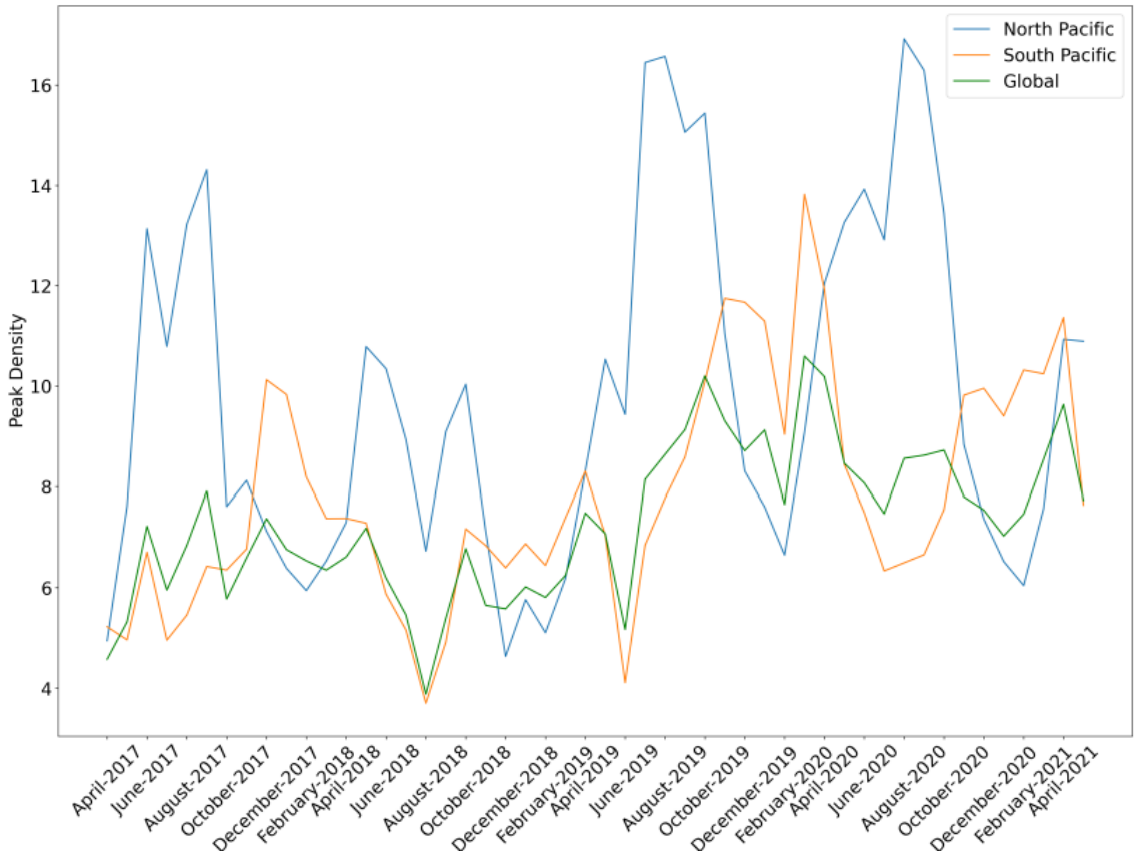


Figure 3.14: Monthly averaged peak density for the North- and South Pacific Ocean and globally.

Source: [Strand, 2021]

3.3.1 Drawbacks and Possible Improvement Areas

Strand is inspired by Evans and Ruf and states that microplastics can suppress ocean surface roughness. The statement is bold and overly simplifies their proposed hypothesis. Nevertheless, the simplification does not affect the logic of the proposed approach. Strand spends little time explaining why his proposed method is believed to be able to detect microplastics. A suggested explanation could be the following:

"I (Strand) hypothesize that other phenomena impacting the surface roughness (primarily wind, maybe currents, and others) are relatively constant within small temporal and spatial changes. I, therefore, examine sequential NBRCS measurements made by a satellite within a track. Given the hypothesis, sudden spikes are not expected in the NBRCS profile. Therefore, these sharp peaks are believed to be caused by the presence of surfactants/microplastics, as surfactants reduce roughness and may share similar transport mechanisms as microplastics. I expect regions

with high surfactant concentrations will cause a high frequency of detected peaks. Therefore, I relate more peaks to more microplastics".

As more peaks are being related to more microplastics, how a peak is determined impacts the final predictions. Strand reports that peaks are detected with width of 5 and prominence of 1 in the $find\_peaks(data, prominence, width)$ function from the *Scipy.signal* python package. The values are seemingly arbitrary set based on what is reported.

Strand also reports that the effect of ocean currents on surface roughening is less impacting for lower wind speeds. He, therefore, omits observations with colocated ERA5 wind speed of under 3 meters per second. The specific threshold set still seems arbitrary, as other values may produce better results. As well as omitting observations with under 3 meters per second, observations with over 5 meters per second were removed. No reason for this selection is reported.

When defining a peak and selecting observations based on wind speeds, the choices made both appear arbitrary. Width, prominence, max wind, and min wind should be considered hyperparameters, and different combinations should be evaluated against the microplastic models of Van Sebille et al., as they are considered ground truth by Strand in the analysis.

The location of the SP is determined by the position of both the transmitting GPS and receiving CYGNSS satellite. The revisit time describes how often a CYGNSS observation is being made at a specific location. As the GPS and CYGNSS satellites are asynchronous, the revisit time varies between different locations. The median revisit time is 2.8 hours, while the mean revisit time is 7.2 hours. Figure 2.3 shows the imbalanced revisit times. The proposed method is counting the number of detected peaks. Therefore, an undesirable bias is introduced in the approach, as revisit time impacts the number of peaks. The introduced bias is probably even more significant when filtering out observations based on wind speeds. The mean value of wind speed differs in different regions. Consequently, more observations will be made over areas where winds often are within the constraining range than in areas with other winds. More observations will result in more detected peaks, and fewer observations will result in fewer peaks. The bias can be compensated by scaling the number of detected peaks by the number of observations over the same region.

Strand evaluates his results qualitatively by presenting maps showing the density of peaks and Van Sebille et al. microplastic distributions next to each other. He later argues for good model agreement between the two. As a reader, it is hard to compare

the two as the colors are not aligned. No scales indicating value ranges are presented either. Unlike Evans and Ruf, Strand presents no quantitative comparisons between the microplastic predictor and prior estimations of microplastic concentrations, like Figure 3.7 shows for Evans and Ruf's results. This would be preferred to assess the correlation. In addition, quantitatively comparison of the variables enables the development of empirical models, like Evans and Ruf did in Equation 3.3. Strand describes his developed approach as a method to detect microplastics. Without building an empirical relationship between the two variables, peaks should only be referred to as an indicator for microplastics.

One factor favoring Strand's approach over Evans and Ruf is that his methodology does not necessarily rely on matchup wind data. Firstly, this simplifies the data retrieval and preprocessing. Secondly, the wind data are just estimates. Estimates will always bring uncertainties into the results.

3.4 Effects of Microplastics and Surfactants on Surface Roughness of Water Waves

In 2021, Sun et al. submitted a study, currently under scientific review. The study investigated the underlying reasons for the observed reduction in ocean surface roughness reported by Evans and Ruf in their research on microplastic detection using GNSS-R.

Two possible reasons behind the observed reduction in surface roughness are researched. The study assessed the possibility that microplastics cause the reduction either as floating particles or by surfactants that share similar transport mechanisms as microplastics in the ocean [Van Sebille et al., 2020]. Figure 3.15 shows a semantic sketch of the experimental setup. The wind-wave tank is located at the Marine Hydro-dynamics Laboratory at the University of Michigan.

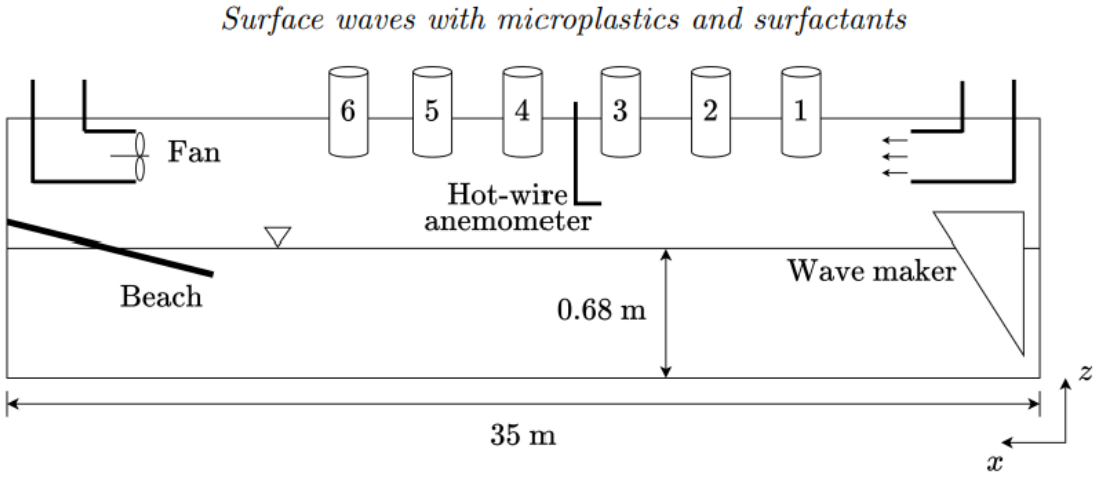


Figure 3.15: Semantic sketch showing the experimental setup. The different elements are denoted by text in the figure, except the six ultrasonic sensors only marked with numbers. The sketch is also showing the water depth.

Source: [Sun et al., 2021]

An approximation of a fully developed sea state was simulated by producing artificial waves based on the empirically developed Bretschneider spectrum [Bulian et al., 2004]. The waves are produced either by a wind from a fan in an open-loop tunnel or by a mechanical wavemaker. MSS was monitored and compared to a clear water situation with the same wave conditions for all experiments with either surfactants or plastic particles to infer a possible reduction in MSS.

3.4.1 Experiment with Plastic Particles

The experiment with plastic particles was only conducted with waves generated by the mechanical wavemaker. The authors reported that it was difficult to control the amounts of plastics on the surface due to drift by the wind when producing wind waves. The microplastics used in the study were made of polypropylene, which is one of the two major materials found in oceanic microplastics [Cózar et al., 2014]. Two different particle types were used, where one type had an average diameter of $5mm$ with irregular shape, and the other type was shaped regularly with $8mm$ average diameter. The experiment with plastic particles is conducted with varying amounts of particles in the tank. The amount of microplastics was reported according to the

average area function shown in Equation 3.4

$$C = \frac{N_p S_p}{WL}, \tag{3.4}$$

where N_p is the number of particles, S_p is the particle area size, W is the width of the tank, and L is the average length of the spreading of particles in the tank. C was throughout the experiment varying between 0.1% and 20%. 0.1% is reported to be close to realistic microplastic concentration in the ocean.

The experiment showed that particles only affect MSS when the average area of particles is above $C = 5\% - 10\%$, and that particle size is inconsequential. The effect was however reported to be small, and negligible when considering the cutoff wavenumber (k_c) used in the CYGNSS measurements ($k_c = 7.5$ rad m^{-1}) even for the highest concentrations tested ($C = 23.35\%$). This can be seen in Figure 3.16, where MSS is plotted versus cutoff wavenumber with average area of particles ($C = 18.69\%$ and 23.35%) way exceeding what is estimated in the world's oceans ($C = 0.1\%$).

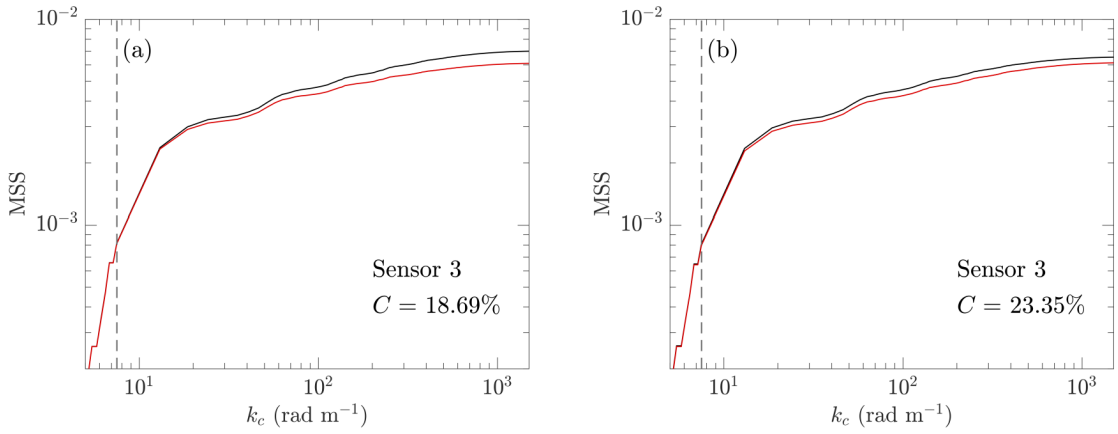


Figure 3.16: MSS versus cutoff wavenumber from the experiment with plastic particles and mechanical waves. Average area of plastic particles and particle diameter are respectively given in (a): $C = 18.69\%$, $D_p = 5mm$ and in (b): $C = 23.35\%$, $D_p = 8mm$. Red line shows the results for water with plastic particles, and black line for the reference clear water. Dashed line shows the cutoff wavenumber, $k_c = 7.5$ rad m^{-1}, used in the CYGNSS satellites.

Source: [Sun et al., 2021]

As a result, the authors concluded that the observed reduction seen in the CYGNSS measurements is caused by another phenomenon than the presence of plastic particles.

3.4.2 Experiment with Surfactants

The experiment with surfactants was conducted with both mechanical waves and wind-generated waves. The surfactant used in the study, Triton X-100, is reported to be used in many previous studies on surfactants. The reference clear water was denoted with Γ_0, and the other surfactant concentrations were denoted from Γ_1 to Γ_8 with increasing surfactant concentration for larger numbers, as shown in Table 3.1.

Table 3.1: Reference symbol, concentration, and corresponding surface tension for the researched surfactant concentrations.

Symbol	Γ_0	Γ_1	Γ_2	Γ_3	Γ_4	Γ_5	Γ_6	Γ_7	Γ_8
$\Gamma \times 10^5$ (mol 1^{-1})	0	0.05	0.20	0.31	0.61	0.81	1.53	3.11	5.09
σ (mN m^{-1})	72.0	69.1	60.1	57.5	54.4	53.5	49.9	45.5	43.5

The experiment with mechanical waves was conducted with the following surfactant concentrations: Γ_0, Γ_1, Γ_2, Γ_6, Γ_7, Γ_8. Figure 3.17 shows the main findings from the experiment. Part (a) of Figure 3.17 shows MSS versus cutoff wavenumber for the clear water reference Γ_0 and the highest surfactant concentration Γ_8. Part (b) shows MSS versus the different surfactant concentrations tested at cutoff wavenumber, k_c = 7.5 rad m^{-1}, used in the CYGNSS mission.

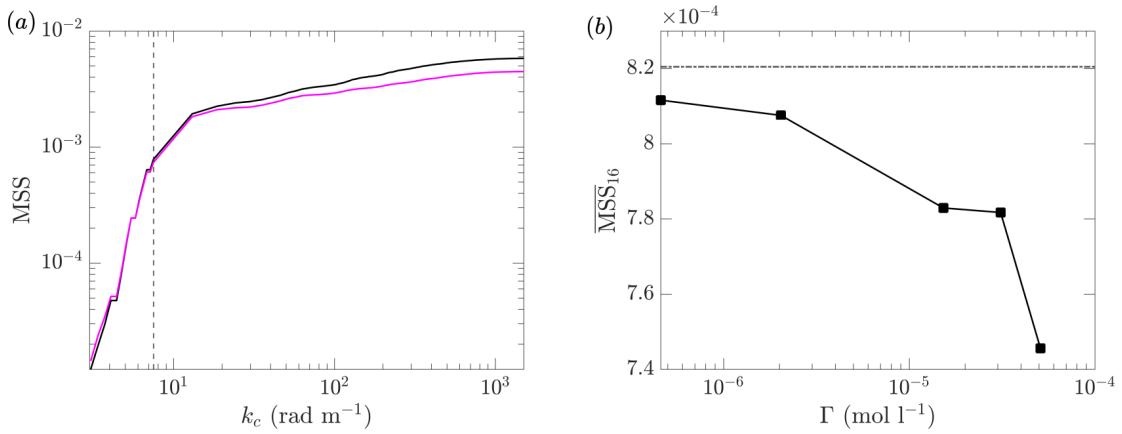


Figure 3.17: (a) MSS versus cutoff wavenumber from the experiment with surfactants and mechanical waves. Pink line shows results for water with surfactant concentration Γ_8, black line shows the reference clear water Γ_0. Dashed line shows the cutoff wavenumber, $k_c = 7.5$ rad m^{-1}, used in the CYGNSS satellites. (b) MSS versus different surfactant concentrations for mechanically generated waves at cutoff wavenumber, $k_c = 7.5$ rad m^{-1}.

Source: [Sun et al., 2021]

The experiment with wind-generated waves was conducted with the following surfactant concentrations: Γ_0 - Γ_7. Figure 3.18 shows the main findings from the experiment. MSS reduction was significantly more present for all wind speeds (part a-c) researched here than in the other experiments with plastic particles or with surfactants and mechanically generated waves. A higher surfactant concentration led to a more significant reduction in MSS, which can be seen in Figure 3.18. Nevertheless, the experiment with a moderate surfactant concentration, Γ_3, showed a reduction of about 17% in MSS for wind speeds of about 9 ms^{-1}.

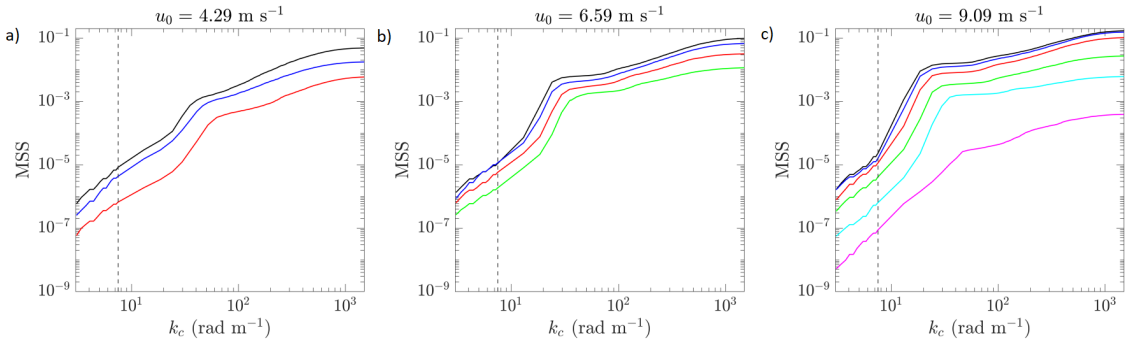


Figure 3.18: MSS versus cutoff wavenumber from experiment with surfactants and wind generated waves, with wind speeds: (a) ; $u_0 = 4.29\ ms^{-1}$, (b) ; $u_0 = 6.59\ ms^{-1}$, (c) ; $u_0 = 9.09\ ms^{-1}$. The colors representing the different surfactant concentrations are: black ; Γ_0, blue ; Γ_1, red ; Γ_2, green ; Γ_3, turquoise ; Γ_4, pink ; Γ_5.

Source: [Sun et al., 2021]

Figure 3.18 and Figure 3.17 show that surfactants have a more significant damping effect on wind-driven roughening than on mechanically-driven roughening. The authors, therefore, concluded that surfactants in a windy sea are the most impacting contributor to the reduction in MSS observed by the CYGNSS satellites, which is reported to be about 20%.

Chapter 4

Data and Methodology

The following chapter begins with an overview of the technical part of the thesis. Later the data and methodology are presented.

Figure 4.1 aims to give the reader an overview of the technical aspect of this thesis. The figure reads from left to right and highlights every major data processing step, which is primarily done in Python. The data sets are put into context to give the reader an understanding of the path from unprocessed data to the microplastic predictor.

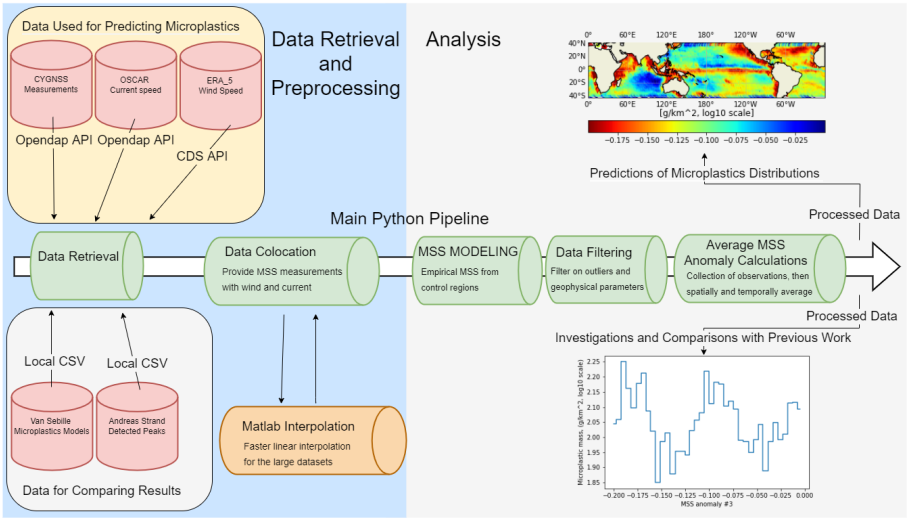


Figure 4.1: The blue part of the illustration shows the retrieval and preprocessing of data, which is presented in section 4.1. The grey part shows an overview of the analysis that was applied to the prepossessed data, which is further described in section 4.2.

4.1 Data

This section begins by describing the data that constitutes the method and the data used to compare and assess results. Later the procedure of how the remote data was collected is presented and the processing steps applied prior to the analysis. The section finishes by describing the methodology.

4.1.1 Data Set Descriptions

The purpose of the analysis is to investigate the relationship between satellite measurements of MSS, wind data, and ocean current data to the concentration of microplastics. Data containing MSS measurements, wind estimates, current estimates, and microplastic distribution models are needed to assess the relationship. Previous results from Strand are also retrieved for comparison. The leftmost part of Figure 4.1 shows an overview of the data used. A description of all the data used is provided below, where the bullet points correspond to attributes retrieved from each data set.

CYGNSS

Data from 18 March 2017 to 30 January 2022 are retrieved. Section 2.2.1 further describes the origin of the data. The Level 2, version 3.0 MSS data is used. The data set provides MSS observations daily. Some days between the two dates are missing because of internal server errors.

- `mean square slope` Average MSS value of a 25 x 25 kilometers area centered at lon, lat
- `sample_time` Average of seconds since midnight of the MSS observations
- `lon` Average longitude of the SP of the MSS observations ranging from 0 to 360
- `lat` Average latitude of the SP of the MSS observations ranging from -60 to 60

ERA5 Wind

Data from 1 March 2017 to 31 March 2022 are retrieved. Wind from the ECMWF Reanalysis 5th Generation (ERA5) data set is used. The data is produced by the

European Centre for Medium-Range Weather Forecasts (ECMWF) and includes wind estimations from 1979 to the present. The data set consists of global estimates of wind speeds at every hour given on a grid of 0.25° by 0.25°, which sufficiently provides reliable wind estimations for every CYGNSS measurement. The data consists of wind speed estimations 10 meters above the surface, which are assumed to be similar to the wind speeds at surface levels.

- `u10` Eastward wind velocity 10 meters over the surface given in meters per second
- `v10` Northward wind velocity 10 meters over the surface given in meters per second
- `longitude` Longitude of wind estimate ranging from 0 to 360 degrees
- `latitude` Latitude of wind estimate ranging from -90 to 90 degrees
- `time` Time of wind estimate given on format yyyy-mm-dd hh-mm-ss

Oscar Current

Data from 18 March 2017 to 4 February 2022 are retrieved. Estimates of current speeds from the Ocean Surface Current Analysis (Oscar) Third Degree Resolution data set are used. The data consists of current speed estimations at 15 meters depth. However, the current speeds are assumed to be similar to the currents at surface levels for the analysis. The data set has approximately five days between each current estimate.

- `u` Zonal current velocity at 15 meters depth given in meters per second
- `v` Meridional current velocity at 15 meters depth given in meters per second
- `longitude` Longitude of current estimate ranging from 20 to 420 degrees
- `latitude` Latitude of current estimate ranging from -80 to 80 degrees
- `time` Time of current estimate given in days since 05.10.1992

Van Sebille Microplastic Models

The microplastic models described in section 3.1 produced by Van Sebille et al. are retrieved for comparison. The data was provided by Ph.D. candidate Hoseini, Mostafa, as a CSV file. The file consists of three different models, where all models

provide an estimate of both the number and mass of microplastics. The data is stationary distributions.

- `Abundance` Number of microplastics per square km
- `Mass` Weight of microplastics per square km
- `longitude` Longitude of microplastics estimate ranging from 0 to 360
- `latitude` Latitude of microplastics estimate from -70 to 79

Strand Detected Peaks

Data from 18 March 2017 to 31 May 2021 are retrieved. The detected peaks described in section 3.3 produced by Strand are retrieved to further investigate the approach and for comparison. The data was provided by Ph.D. candidate Hoseini, Mostafa as multiple CSV files, where each day had its own file. Each row in the CSV file corresponded to a detected peak.

- `timestamp_utc` Seconds since midnight of the NBRCS observation
- `wind_speed` Nearest interpolated ERA5 wind speed
- `longitude` Longitude of detected peak ranging from -180 to 180
- `latitude` Latitude of detected peak ranging from -38 to 38

4.1.2 Remote Data Retrieval

This section presents code that retrieves data remotely from the two API services used in this thesis and stores it on a local disk. The developed code is written in the Python programming language and requires the installment of specific packages, included at the beginning of both snippets.

CDS API

The Climate Data Store (CDS) API is used for the retrieval of ERA5 wind data. The following code snippet sequentially sends requests to the API service via the API client installed in the Python package "cdsapi" and asks for monthly wind data. Before running the code, it is required to register an account at CDS to get an API key. The key later needs to be referred to in the Python environment.

```
import pandas as pd
import cdsapi
import xarray as xr
from urllib.request import urlopen

c = cdsapi.Client()
# Dates to fetch
years = ['2017','2018','2019','2020','2021', '2022']
months = ['01', '02', '03', '04', '05', '06',
          '07', '08', '09','10','11', '12']
days = ['01', '02', '03','04', '05', '06', '07', '08','09', '10', '11',
        '12','13', '14', '15','16', '17', '18', '19', '20', '21','22',
        '23', '24','25', '26','27', '28', '29', '30','31']
# Iterate over years and months to fetch monthly ERA5
for year in years:
    for month in months:
        # Setting API parameters prior to request
        params = {'product_type': 'reanalysis',
                'variable': ['10m_u_component_of_wind',
                             '10m_v_component_of_wind',],
                'year': year,
                'month': month,
                'day': days,
                'time': [
                    '00:00', '01:00', '02:00','03:00', '04:00', '05:00',
                    '06:00', '07:00', '08:00','09:00', '10:00', '11:00',
                    '12:00', '13:00', '14:00','15:00', '16:00', '17:00',
                    '18:00', '19:00', '20:00', '21:00', '22:00', '23:00',
                ],
                'area': [38, -180, -38, 180],
                'format': 'netcdf',
```

```
        }
    # Retrieve data from API and store to local disk as CSV
    fl = c.retrieve('reanalysis-era5-single-levels', params)
    with urlopen(fl.location) as f:
        ds = xr.open_dataset(f.read())
        era_5_df = ds.to_dataframe()
        year_month_string = year + "_" + month
        era_5_df.to_csv("era_5/" + year_month_string + ".csv"
                        ,index=False)
```

OpenDap API

The Open-source Project for a Network Data Access Protocol (OpenDap) API is used for retrieving both the CYGNSS and Oscar data. The following code snippet sequentially sends requests to the API service via the API client installed in the Python package "pydap" and asks for daily CYGNSS MSS data.

```
from pydap.client import open_url
import numpy as np
import pandas as pd
from datetime import datetime
from datetime import timedelta, date

# Funtion for retrieveing CYGNSS data for one day
def get_cygnss_data_one_day(year, month, day):
    cygnss_df = pd.DataFrame()
    test_data_url, test_clickable_url = generate_url(year, month, day)

    df = pd.DataFrame()
    mss = np.array(dataset.mean_square_slope[:])
    lat = np.array(dataset.lat[:])
    lon = np.array(dataset.lon[:])
    sample_time = np.array(dataset.sample_time[:])

    df['mss'] = mss.tolist()
    df['lat'] = lat.tolist()
    df['lon'] = lon.tolist()
    df['time'] = sample_time.tolist()
    return df

# Sub function used when getting CYGNSS data for one day.
# Generates API URL for level 2 v 2.1 CYGNSS data
def generate_url(year, month, day):

    day_of_year = datetime(year, month, day).timetuple().tm_yday
    date_string = str(year) + str(month).zfill(2) + str(day).zfill(2)

    base_url = 'https://podaac-opendap.jpl.nasa.gov/opendap/hyrax/'
```

```
                + 'allData/cygnss/L2/v2.1/'

    specific_url = str(year) + '/' + str(day_of_year).zfill(3) +
                   '/cyg.ddmi.s' + date_string + '-000000-e' +
                   date_string + '-235959.l2.wind-mss.a21.d21.nc'


    data_url = base_url + specific_url
    clickable_url = base_url + specific_url + '.html'

    return data_url + '?lat,lon,mean_square_slope,sample_time', clickable_url

sdate = date(2017, 3, 18)   # start date
edate = date(2022, 1, 3)   # end date
delta = edate - sdate
# Iterate over each daty between the two dates to fetch CYGNSS data
# and store to local disk
for i in tqdm(range(delta.days + 1)):
    day = sdate + timedelta(days=i)
    df = get_cygnss_data_one_day(day.year, day.month, day.day)
    df.to_csv("level_2_mss_v2.1/" + str(day.year) + "_" +
              str(day.month).zfill(2) + "_" +
              str(day.day).zfill(2) + ".csv" ,index=False)
```

4.1.3 Preprocessing

The following section presents the processing steps that were applied prior to the analysis. The amounts of data that needed to be processed were large (more than 1 Terabyte). The computers used did not have sufficient RAM to process all the data simultaneously. As a result, the preprocessing was done in batches.

The main goal of the preprocessing is to supply CYGNSS MSS measurements with current and wind estimates. The wind and current are provided on a grid with different resolutions. In addition, the variation of time needs to be accounted for. The measurements made from the CYGNSS satellites do not follow a grid-like structure and are being made every half a second. The exact wind and current estimations for each longitude, latitude, and observation time are not available, as it would be an infinite-sized data set. The lack of available data is solved by interpolation, as it can provide reasonable wind and current estimations for every measurement that varies both temporally and spatially.

Data Modification

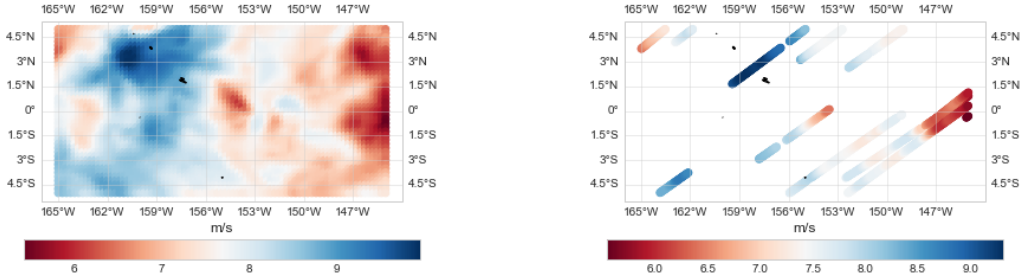
The data from the different data sets had missing values at retrieval. One could replace measurements with missing values with average values or try to infer the missing values by other methods. However, the missing values were only dropped in this project, as the number of missing values was insignificant. In addition, the interpolation later fills in the removed values for both current and wind making the problem less consequential.

Prior to interpolation, the variables geolocating the data need a standard reference. This means that MSS, wind, and current need an equal representation of longitude, latitude, and time. The representation of latitude coincides in all three data sets. Time is, however, represented in seconds since midnight in the CYGNSS data set, as a string containing date and time in the ERA5 data set, and as the number of days since 05.10.1992 in the Oscar data set. The different representations of time are changed by converting to hours since a fixed reference point for all data sources. Longitude coincides in the CYGNSS and ERA5 data sets but is represented differently in the Oscar data set. The representation is explained in the following quote: «However the data represents longitude as 20° to 420°, i.e. the entire Earth is represented in the data as 20° to 380°, and the data repeats itself from 380.33° to 420°. Thus 390° is actually 30°» [Hausman et al., 2009, p. 4]. This choice of representation was made to have continuous major ocean basins, which is irrelevant for this thesis. The representation is modified to match the other data sources by dropping all data points with longitudes over 380° and replacing the longitude with the remainder of the original longitude after dividing it by 360.

Data Interpolation

After the three key variables were represented in the same way, current and wind data was interpolated. This was done using linear interpolation, which assumes a linear relationship both in time and space between the different wind and current estimates. The assumption seems reasonable, as natural phenomena seldom have sudden changes with minor spatial and temporal variations.

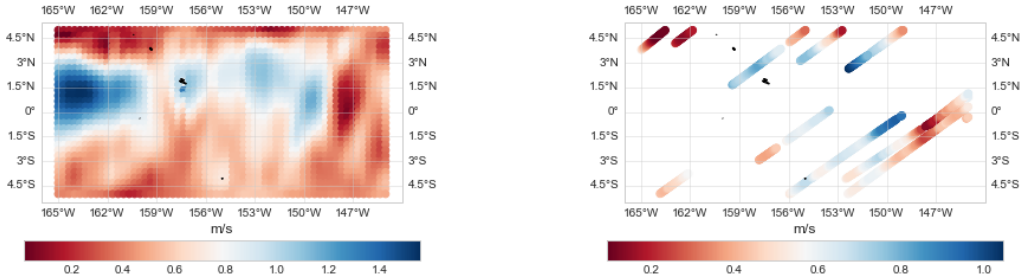
Figure 4.2 and Figure 4.3 are showing the results of the interpolation, resulting in CYGNSS observations with provided wind and current estimates at the location and time of measure. The location and time selected for what was displayed here were randomly chosen for verification purposes. The resemblance between the (a) and (b) parts of the figures indicate that the interpolation was successful.



(a) ERA5 Wind

(b) Interpolated ERA5 Wind

Figure 4.2: ERA5 wind data over a region around the Line Islands on 01.10.2021.



(a) Oscar Current

(b) Interpolated Oscar Current

Figure 4.3: Oscar ocean current data over a region around the Line Islands on 01.10.2021.

The interpolation of wind and current data was unlike the rest of the project done in Matlab, as Figure 4.1 is denoting, using the built-in function "scatteredInterpolant". Various Python interpolation techniques were too slow when processing significant amounts of data, making Matlab the preferred option.

4.2 Method

Both Strand and Evans and Ruf presented approaches where efforts were made to infer microplastic distributions based on CYGNSS measurements. Section 3.3.1 and section 3.3.1 highlighted possible improvement areas in both approaches. The preliminary analysis that was conducted, described in section 5.1, showed promising results for the approach presented by Evans and Ruf. As a consequence, this study aimed to improve their approach, which is described in section 3.2.

To investigate the possible improvement areas, observations of MSS anomaly were collected using all the data described in section 4.1, extending the data compared to what Evans and Ruf used. The same definition of MSS anomaly, which is the microplastic predictor, was used and shown in Equation 4.1.

$$MSS_{anomaly} = \frac{MSS_{obs} - MSS_{mod}}{MSS_{mod}}, \tag{4.1}$$

where MSS_{obs} is the CYGNSS measurements of MSS and MSS_{mod} is modeled MSS. This approach researched possible improvements to both variables. Unlike Evans and Ruf, MSS_{obs} were based on version 3.0 CYGNSS data compared to version 2.1, which should reduce some GPS related errors. Multiple definitions of MSS_{mod} were tested to study the impact of modeling technique, choice of control regions, and modeling argument. The Katzberg model refitted by Evans and Ruf will be used for reference as MSS_{mod} and is shown in Equation 4.2.

$$MSS_{mod} = \begin{cases} 0.0035(\left|\vec{U}_{10}\right| + 0.62), & \text{if } \left|\vec{U}_{10}\right| \leq 3.49 m/s \\ 0.0035(6ln(\left|\vec{U}_{10}\right|) - 3.39), & \text{if } \left|\vec{U}_{10}\right| > 3.49 m/s \end{cases}, \tag{4.2}$$

where U_{10} is ERA5 wind speed, instead of GDAS wind speed which they used. The ERA5 data set was used, as GDAS is believed to provide equally good wind estimations combined with the fact that it is easier to retrieve. Remodeling of MSS_{mod} were done based on the control regions shown in Figure 5.1, like Evans and Ruf did, as well as in the two regions shown in Figure 4.4.

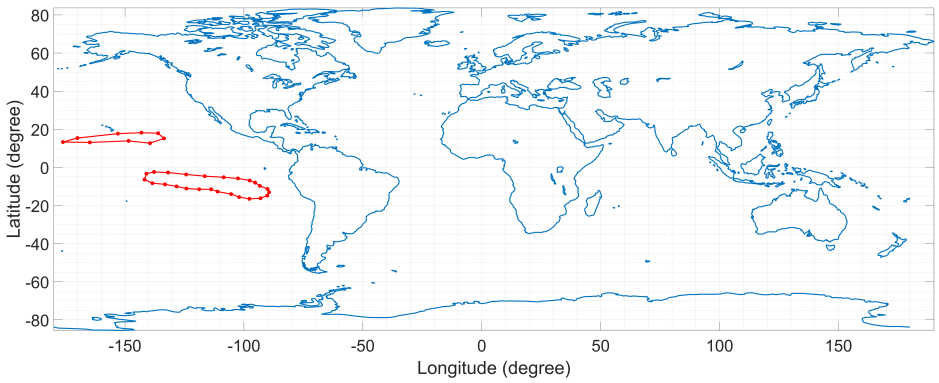


Figure 4.4: New suggestions for control regions based on indications of low microplastics regions from Strand's detected peaks.

Different regressions models were tested when remodeling MSS_{mod}. This included the Katzberg model refitted with the python Library *lmfit* and multiple regression models from the Python Library *sklearn*. Root-mean-square error (RMSE) is used to compare modeling performance. The modeling techniques that were tested are listed below.

- Katzberg
- LinearRegression
- Lasso
- ElasticNet
- KNeighborsRegressor
- DecisionTreeRegressor
- GradientBoostingRegressor

MSS was also modeled using the relative wind and current speed δ, defined in Equation 4.3, as the model argument instead of just wind speed.

$$\delta = \left\| \vec{U}_{10} - \vec{U}_c \right\|, \tag{4.3}$$

where U_c is Oscar current speed. The possible effect of ocean currents was included in the analysis, which aimed to compensate for surface current's contribution to MSS variations shown in Figure 2.7.

$MSS_{anomaly}$ observations with different combinations of modeling technique, choice of control region, and modeling argument for MSS_{mod} were calculated. Later, the observations were spatially and temporally averaged to 1 x 1 degree over the whole period of data. The temporally and spatially averaged MSS anomalies were compared to the microplastic distribution models from Van Sebille et al. as well as Strand's detected peaks.

Strand's detected peaks were not recalculated, as only the processed data was retrieved. Consequently, only some of the improvement areas highlighted in section 3.3.1 are researched. This was done to better understand the approach's capability better to measure ocean microplastics. Quantitative comparisons between detected peaks and Van Sebille microplastic mass were made as well as improved visualization of the resulting peak distributions.

Chapter 5

Numerical Analysis

This chapter first presents the preliminary analysis, which was conducted in the fall of 2021. Later the results from MSS modeling based on wind and relative wind-current speed are presented. The resulting MSS models are further used in Equation 4.1 to produce MSS anomaly observations. The observations are later compared to the reference microplastic models in section 5.3. Improved investigations of Strand's detected peaks complete the chapter.

5.1 Preliminary Analysis

A preliminary analysis with a subset of the retrieved data was conducted that aimed to replicate Evans and Ruf's approach described in section 3.2. The analysis deviated from their work as it used ERA5 wind data instead of GDAS wind, as well as level 1 CYGNSS data, and Equation 2.6 to get MSS observations instead of using the level 2 data. The deviations are, however, insignificant and should produce similar results.

The effect of ocean current was also investigated. This was done by using the relative wind-current speed δ defined in Equation 4.3 to predict MSS, which was done by inputting δ instead of wind speed in Equation 3.1. The wind speed was also used to predict MSS for comparison.

Figure 5.1 shows the selected region with coordinates [20°–40°S, 100°–140°W]. Data ranging from 01.10.2021 to 19.11.2021 in this region was investigated.

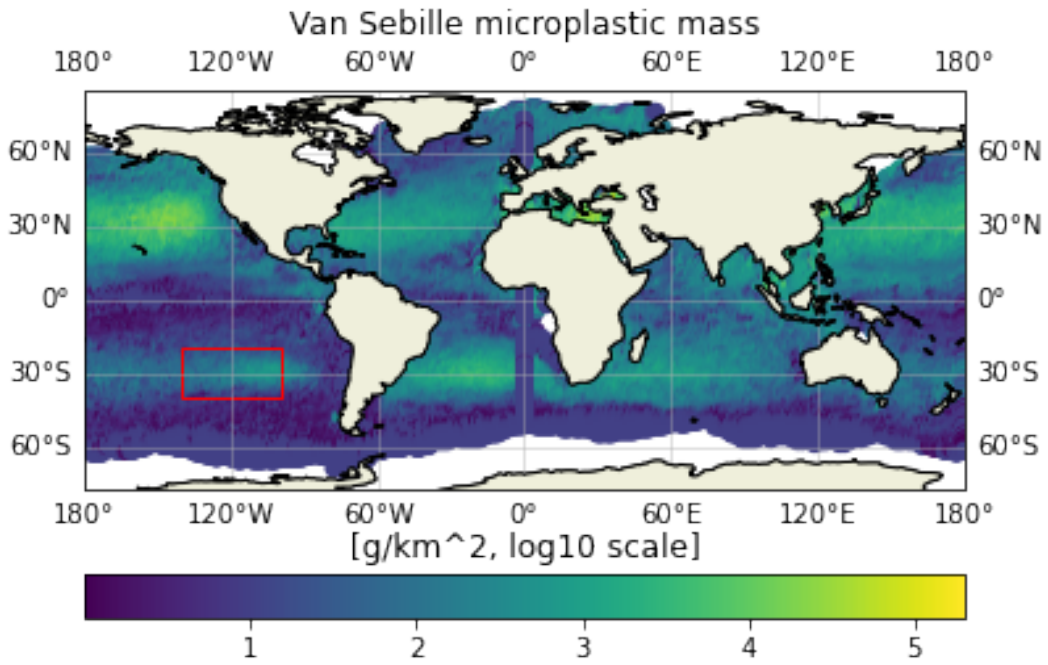
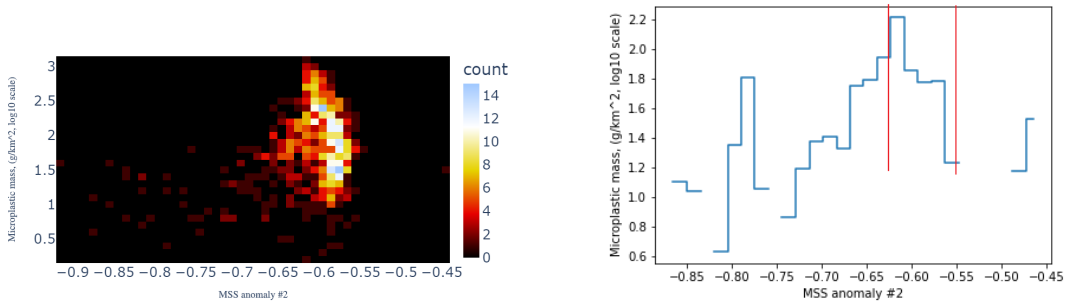


Figure 5.1: Map showing the distribution of microplastic mass based on the Van Sebille model. The selected region that was investigated are highlighted with a red box.

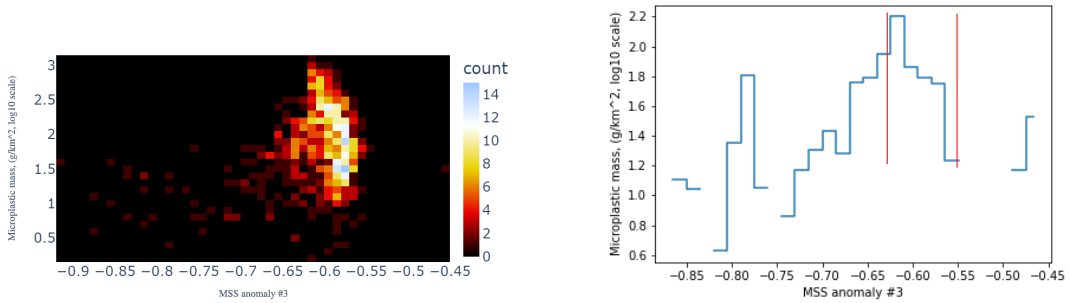
Figure 5.2 and Figure 5.3 show that a similar relationship was found between the spatial and temporal averaged MSS anomaly and the Van Sebille estimated microplastic mass over a sub-range, like Evans and Ruf found in Figure 3.7.



(a) Density scatter plot

(b) Separation into 50 bins

Figure 5.2: Density scatter plot and average bin plot both showing spatial and temporal averaged MSS anomaly with expected MSS defined in Equation 3.1 versus microplastic mass. The red line is denoting the range that shows good correlation.



(a) Density scatter plot

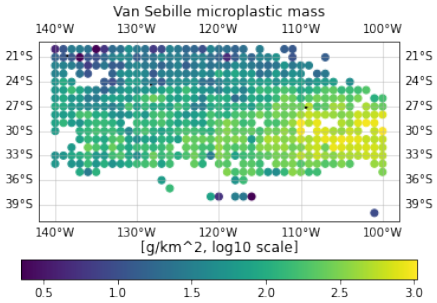
(b) Separation into 50 bins

Figure 5.3: Density scatter plot and average bin plot both showing spatial and temporal averaged MSS anomaly with expected MSS defined in Equation 3.1 with δ as input versus microplastic mass. The red line is denoting the range that shows good correlation.

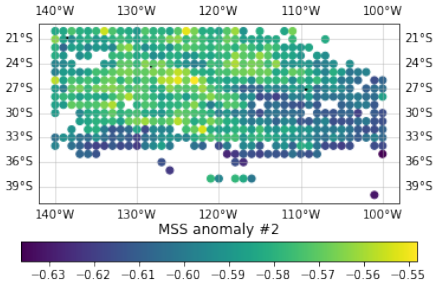
Table 5.1: Correlation between the different spatial and temporal averaged MSS anomaly observations over the sub-ranges and the microplastic concentration and mass.

	Wind MSS Anomaly	δ based MSS anomaly
Wind MSS Anomaly	1.000000	0.996464
δ MSS anomaly	0.996464	1.000000
micro_mass	-0.456341	-0.469553
abundance	-0.456237	-0.469495

Figure 5.4, and Figure 5.5 show microplastic mass and spatial and temporal averaged MSS anomaly filtered on the sub-ranges. The relationship is negatively correlated, resulting in the dark regions in the *(a)* part of Figure 5.4 and Figure 5.5 often being light in the *(b)* part and vice versa.

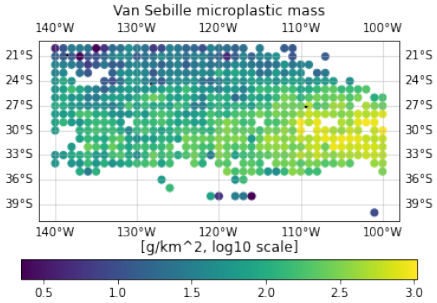


(a) Van Sebille microplastic mass

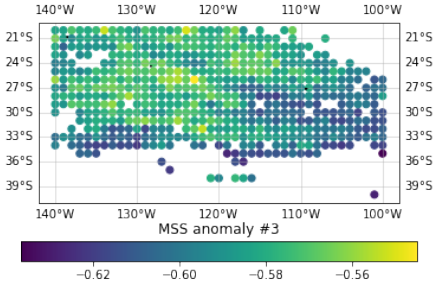


(b) MSS anomaly with expected MSS defined in Equation 3.1

Figure 5.4: Comparison of MSS anomaly with expected MSS defined in Equation 3.1 and microplastic mass.



(a) Van Sebille microplastic mass



(b) MSS anomaly with expected MSS defined in Equation 3.1 with δ as input

Figure 5.5: Comparison of MSS anomaly with expected MSS defined in Equation 3.1 with δ as input and microplastic mass.

The MSS anomaly model with the highest correlating sub-range used the relative wind current speed δ to predict MSS, indicating that the effect of current may be a slight improvement to the approach. Figure 5.6 shows the density scatter plot of microplastic mass versus the *delta* based MSS anomaly over the sub-range. A negative log-linear relationship is seen, which agrees with what was expected from the prior research. The relationship is, however, found in a more negative region of values, compared to Figure 3.7, which can be a result of using a different matchup wind data set.

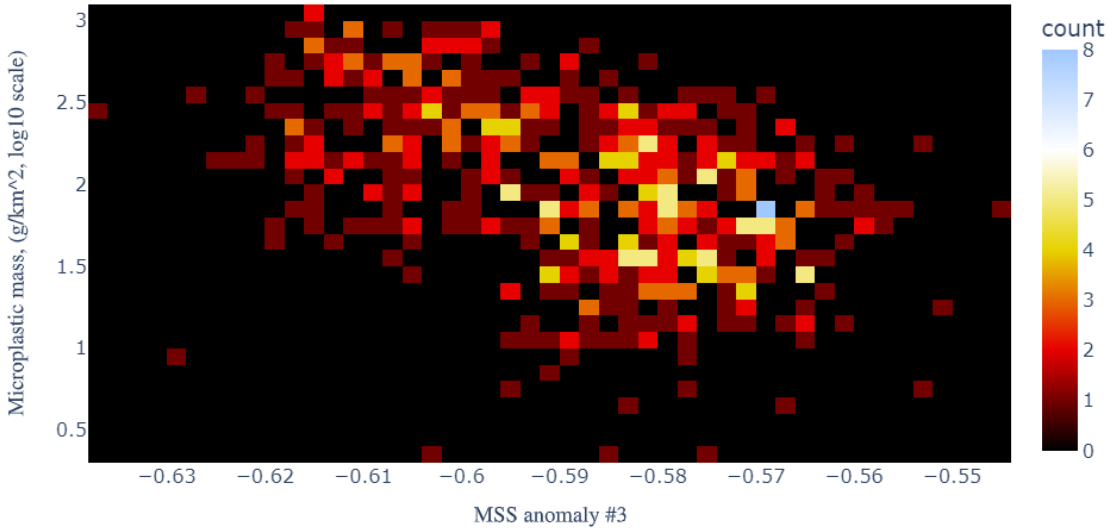


Figure 5.6: MSS anomaly with expected MSS defined in Equation 3.1 with δ as input over the sub-ranges shown in part *(b)* of Figure 5.5 versus microplastic mass.

5.2 MSS Modeling

Table 5.2 shows that changing the modeling argument to relative wind-current speed δ brings no improvement, as the correlation between MSS and δ is insignificantly lower for both control regions compared to the correlation with wind only. Consequently, the analysis focuses on MSS modeling solely on ERA5 wind speed.

Table 5.2: Correlation between MSS, ERA5 wind and relative wind-current speed δ. Left table shows values for the regions shown in Figure 4.4, right for Figure 3.5.

	era_wind	delta
MSS	0.537275	0.537235
era_wind	1.000000	0.997622
delta	0.997622	1.000000

	era_wind	delta
MSS	0.746921	0.745927
era_wind	1.000000	0.998008
delta	0.998008	1.000000

Figure 5.7 and Figure 5.8 shows MSS versus wind speed for both regions with believed low microplastic concentration. Wind is seemingly more dominant on surface roughness in the regions used in previous research, as it shows higher correlation. This may not be the case, as the two other regions lie in the intertropical convergence zone, where MSS observations are known to be more unreliable [Evans and Ruf, 2021]. Other factors may also impact the relationship.

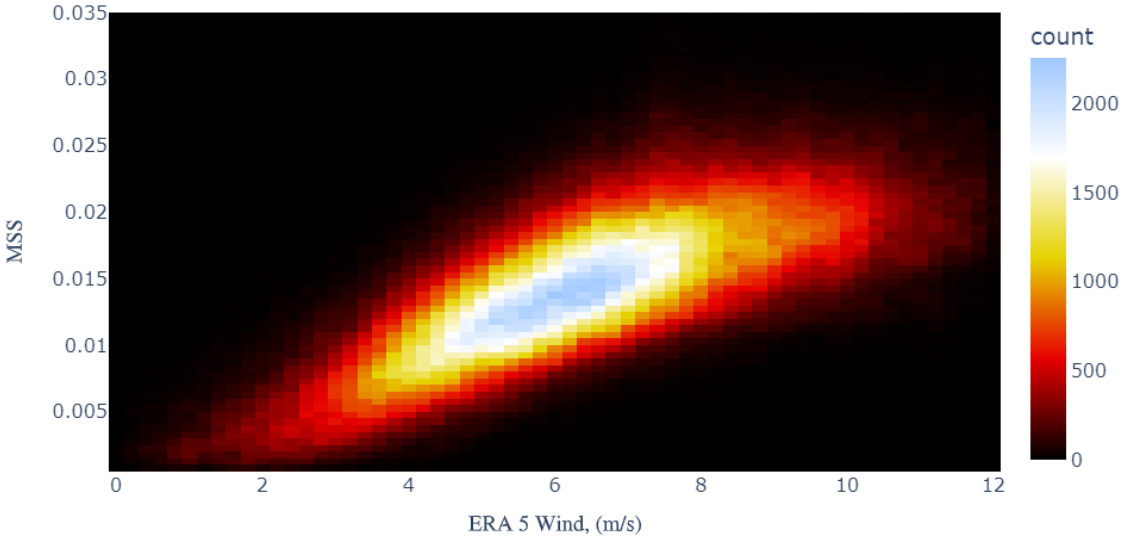


Figure 5.7: MSS versus ERA5 Wind in regions shown in Figure 3.5.

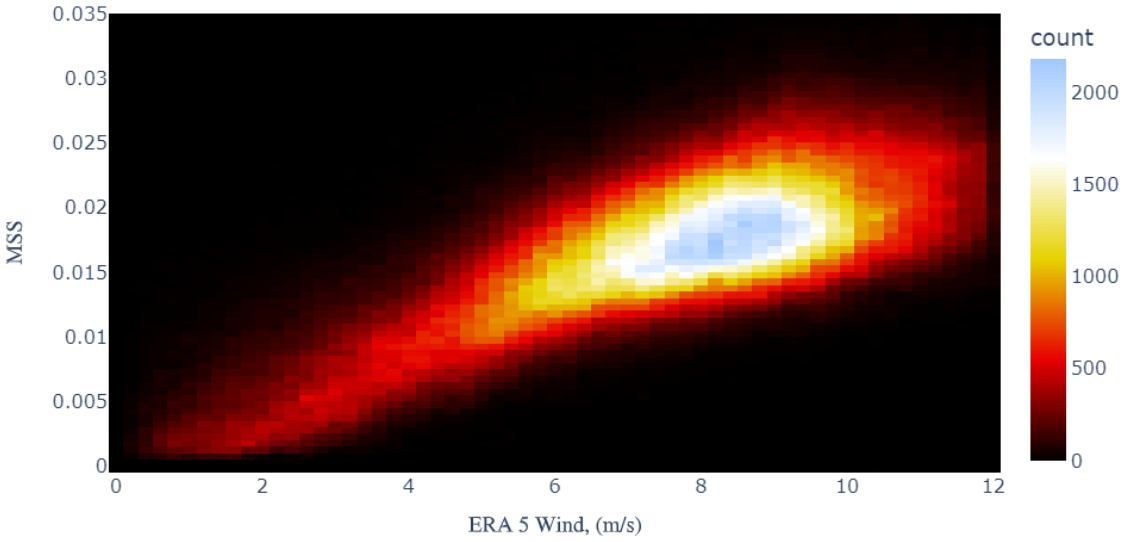


Figure 5.8: MSS versus ERA5 Wind in regions shown in Figure 4.4.

Table 5.3 shows the results after fitting different models with wind speeds in the two regions. MSS takes an average value of approximately 0.015 in the regions. The resulting RMSE varies based on the model selection, with relative RMSE varying between 1/1000 and 2/1000. However, most models produce similar errors, like Katzberg, Linear, KNeighborsRegressor, and GradientBoostingRegressor, all showing promising results. This argues that the modeling technique is less decisive, which is reasonable considering that there only is one feature.

Table 5.3: RMSE for different models fitted in the regions shown in Figure 3.5 and Figure 4.4.

Model	Figure 4.4 RMSE	Figure 3.5 RMSE
Katzberg	0.000017	0.000016
LinearRegression	0.000016	0.000016
Lasso	0.000030	0.000027
ElasticNet	0.000030	0.000027
KNeighborsRegressor	0.000018	0.000018
DecisionTreeRegressor	0.000028	0.000027
GradientBoostingRegressor	0.000015	0.000015

The resulting MSS models developed were wind-based, as relative wind-current speed shows no improvement. The GradientBoostingRegressor showed (insignificantly) best results and is therefore used. Two resulting wind-based GradientBoostingRegressor models fitted over the different regions shown in Figure 4.4 and Figure 3.5 is therefore further utilized for MSS anomaly calculations. The reference Katzberg model fitted by Evans and Ruf will be used for comparison.

5.3 MSS Anomaly

Figure 5.9 shows averaged MSS anomaly over all the data for the three different models. The figures to the left show that outliers exist and removing these produce the figures to the right. Evans and Ruf did not report removing outliers.

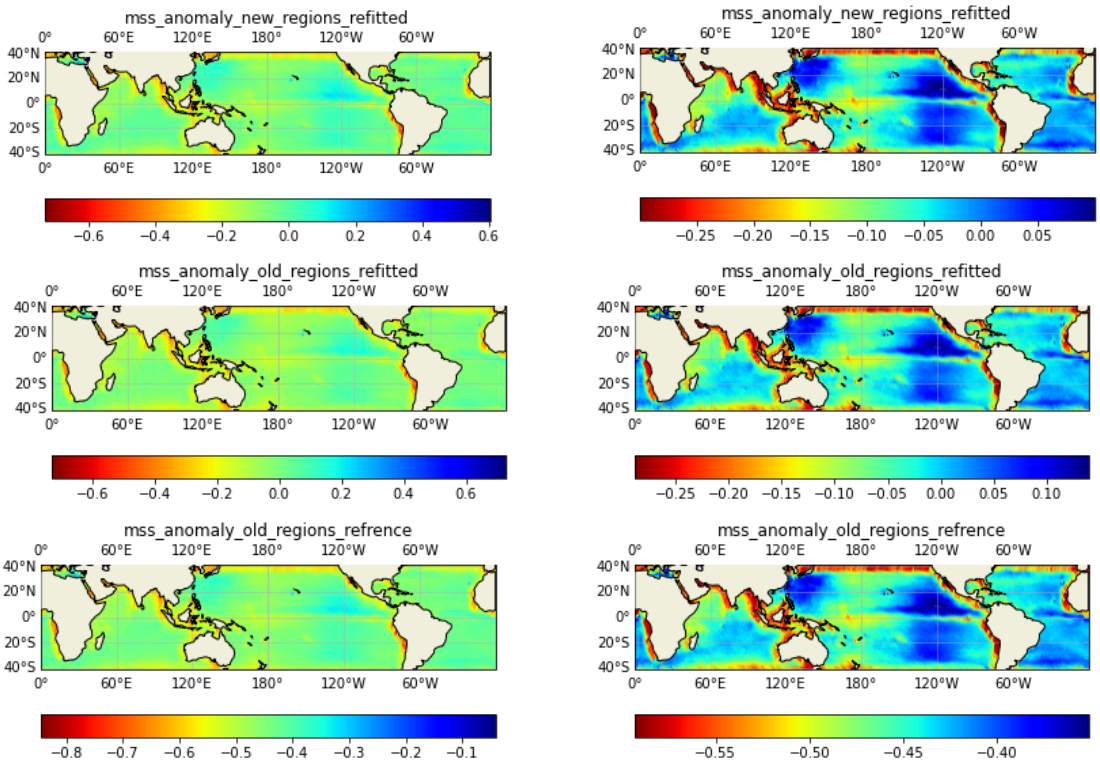
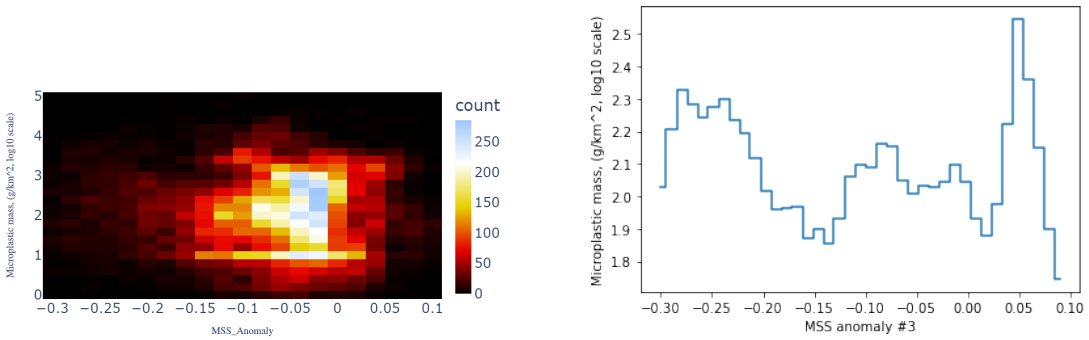


Figure 5.9: Averaged MSS anomaly over all the data. Left showing original values, right filtered for outliers. Top figures use modeled MSS based on regions shown in Figure 4.4, middle figures use modeled MSS based on regions shown in Figure 3.5, and bottom figures is using the reference Katzberg model.

As the models show good agreement with each other, the modeled MSS seems not to be decisive. Consequently, the analysis only compares one of them to the reference microplastic models. It is noteworthy that the reference model shows the same pattern but has a shift in the value range of MSS anomaly seen in the bars in Figure 5.9, which is probably caused by the different matchup wind data used.

Figure 5.10 shows the MSS anomaly versus the Van Sebille microplastic model. Figure 3.7 shows that a good correlation between the variables is expected in a range where the majority of samples lie in the reference study. However, this is not the case here, as most samples, which lie between 0 to -0.1 in MSS anomaly, have reference microplastics values relatively constant around 2 $g/km^2(log10scaled)$.



(a) Density scatter plot

(b) Separation into 40 bins

Figure 5.10: Density scatter plot and average bin plot both showing spatial and temporal averaged MSS anomaly versus the Van Sebille microplastic model.

Figure 5.11 shows the MSS anomaly compared to the three microplastic models developed by Van Sebille et al.. As Figure 5.10 indicated no model agreement with the Van Sebille model, it was not expected to see agreement with the others either, as they show relatively good agreement with each other. All the models from Van Sebille et al. agree that there exist accumulation zones, often referred to as garbage patches. The garbage patches are not visible in the MSS anomaly domain, where almost all high negative values, which are expected to be seen there, are observed along coastlines.

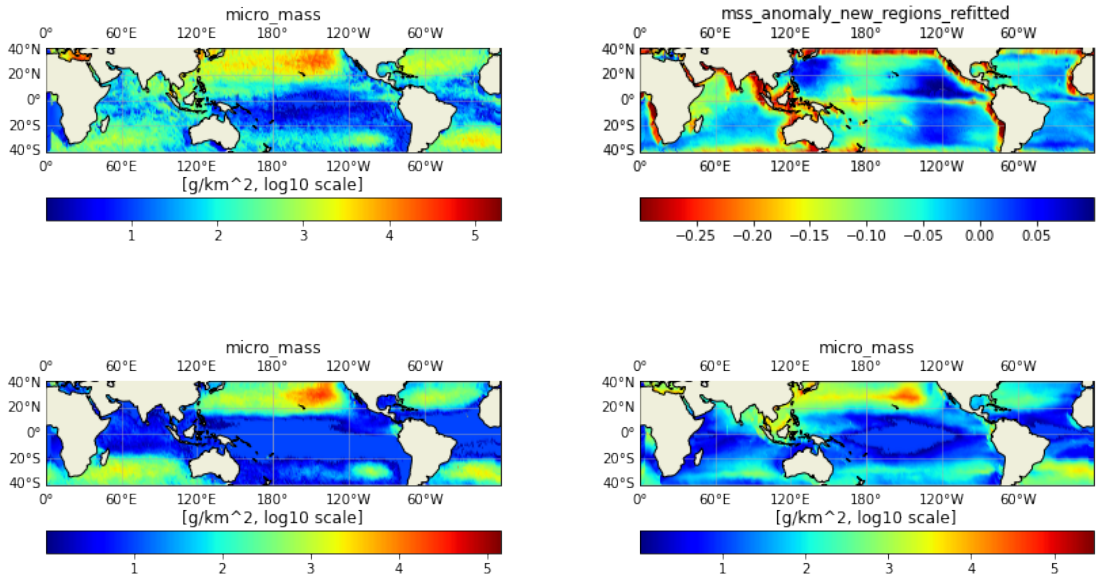


Figure 5.11: In the top-left the Van Sebille microplastic model, top-right MSS anomaly, bottom-left the Maximenko microplastic model, bottom-right the Lebreton microplastic model.

Figure 5.11 concluded that this effort of predicting microplastics showed poor agreement with believed estimations. Figure 3.8 shows that different results were obtained in the reference study. Despite the deviating results, it is interesting to note that similar seasonal variations like Evans and Ruf reported in Figure 3.9 are observed. The figure shows that their predicted microplastic concentrations were highest in both hemispheres in their respective summer months. Figure 5.12 shows the same similar pattern, where average values are oscillating. More negative anomaly means more predicted microplastic concentration. The South Pacific has the highest negative average anomaly every January, and the North Pacific has the highest negative average anomaly every July. Strand reported the same oscillating pattern in Figure 3.14 with the number of peaks being highest in their respective summer months. This similarity indicates that the different approaches are capturing a similar phenomenon.

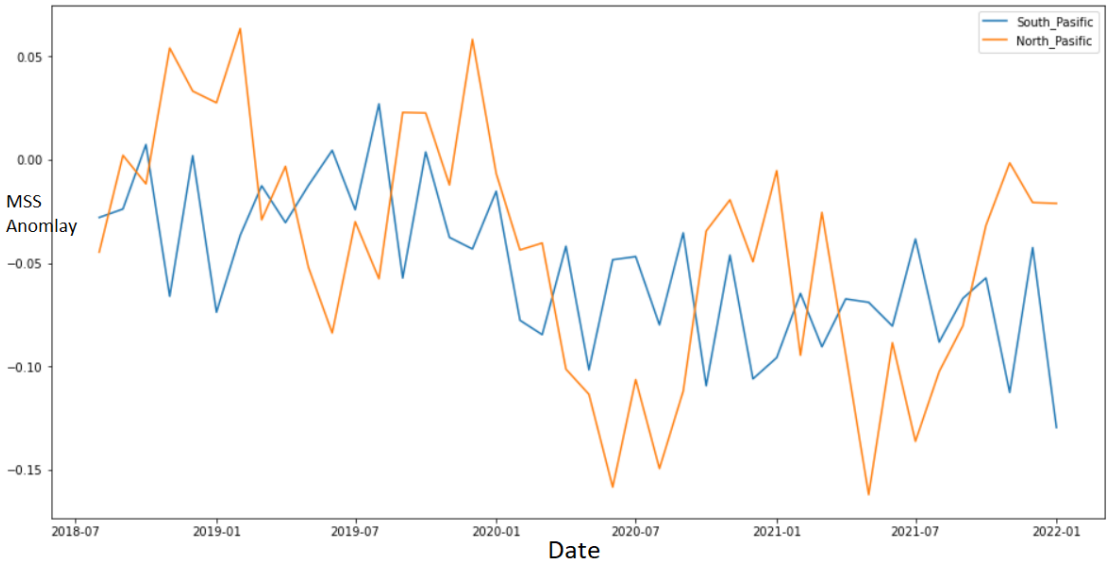
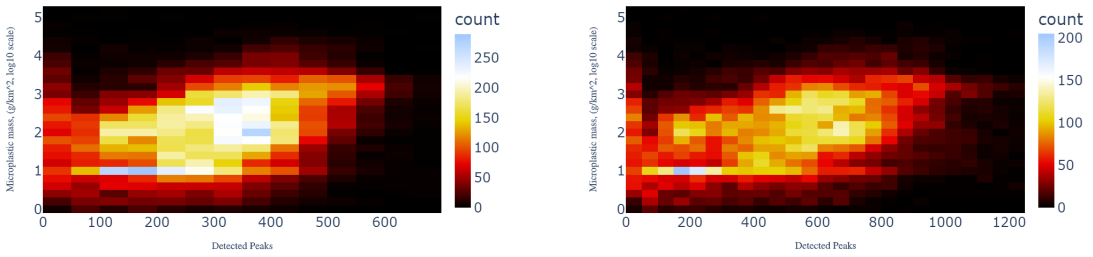


Figure 5.12: Monthly Averaged MSS anomaly over South and North Pacific.

5.4 Detected Peaks

Section 3.3.1 highlighted possible improvement areas to Strand's work. Among the things mentioned is the lack of quantitative comparisons of peaks to prior estimations. Figure 5.13 shows detected peaks versus the Van Sebille microplastic mass. Strand states that he sees good model agreement between peaks and microplastic mass and that filtering on wind speeds improve results, as it reduces the peak density in regions with strong ocean surface currents. By investigating his data, his claim seems to be incorrect. The correlation between peaks and microplastic mass is 0.314560 when low wind speed observations are filtered out, compared to 0.322997 on all observations.



(a) Observations filtered on wind speed

(b) All observations

Figure 5.13: Density scatter plot showing detected peaks versus microplastic mass.

Figure 5.14 shows detected peaks for all wind speeds and the different Van Sebille et al. microplastic models. Compared to Strand's resulting distribution shown in Figure 3.12, the colors are aligned, and the figure is centered in the Pacific. This improves qualitative comparisons to reference microplastic models.

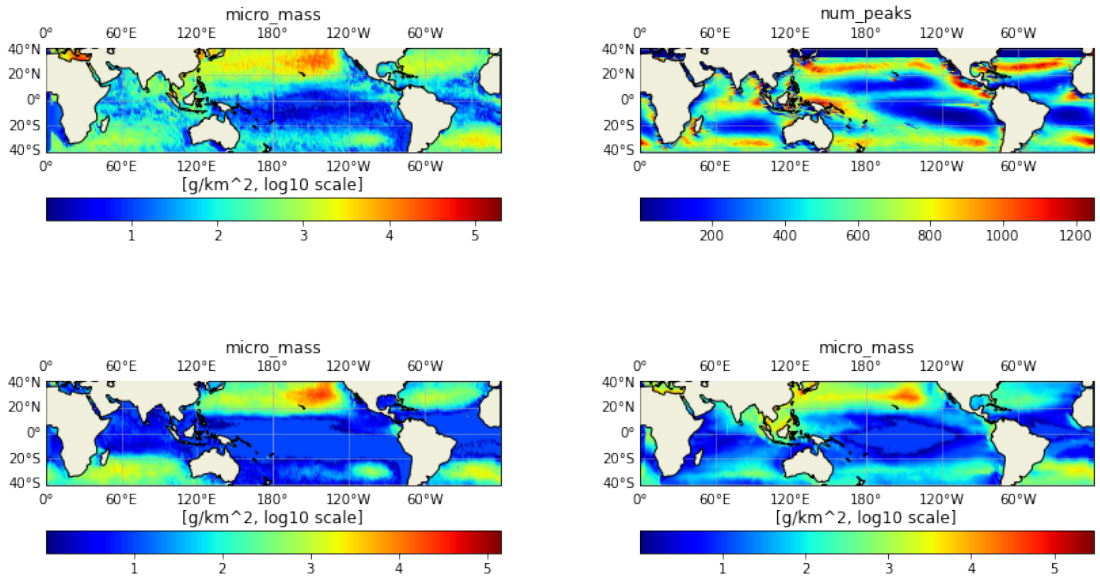


Figure 5.14: In the top-left the Van Sebille microplastic model, top-right number of peaks, bottom-left the Maximenko microplastic model, bottom-right the Lebreton microplastic model.

The number of peaks shows better agreement with the reference models than results from the conducted MSS anomaly approach. High densities of peaks are observed in relative proximity to the garbage patches. However, the number of peaks does disagree in many regions, like in many coastal areas and most equatorial regions.

Chapter 6

Conclusions and Further Work

This chapter summarizes the outcomes of this thesis aiming to answer the research questions based on the presented material and the conducted research. The thesis completes by suggesting future research topics.

6.1 Practical Contribution

Addressing the issue of ocean pollution and making informed decisions demands timely, high-quality measurements. This research investigated the potential of spaceborne GNSS-R for the remote sensing of floating oceanic microplastics. The GNSS-R technique, based on the bistatic radar concept, is particularly interesting since it can be implemented on small satellites and be inexpensively scaled for high spatiotemporal global coverage.

The analysis presented in this thesis utilized a data set from a dedicated reflectometry mission called CYGNSS, which NASA launched in 2016. The mission has produced spaceborne measurements since March 2017. In addition, various ancillary datasets were included for justification and validation purposes which provided measurements of ocean surface wind, surface currents, and estimates of ocean microplastic concentration. The results showed the prospects of spaceborne GNSS-R when intending to retrieve the global concentration map of microplastics while highlighting several issues to be addressed for more reliable data products. The conclusions of this thesis can be articulated by answering the following research questions raised in the Introduction chapter.

RQ1: *What efforts have been made to estimate the global microplastic distribution*

in the oceans, and what are possible areas of improvement?

This thesis presented three approaches to measuring global oceanic microplastic concentrations. Van Sebille et al.'s approach is based on surface trawling measurements and ocean circulation models and is used as a benchmark for the two GNSS-R approaches presented. The two GNSS-R approaches rely on an assumed reduction in ocean surface roughening in areas with microplastics. Possible improvement areas to the GNSS-R based approaches were highlighted in section 3.3.1 and in section 3.2.1.

RQ2: *Will suggested improvements to a GNSS-R approach to estimate the global microplastic distribution improve prior results?*

An attempt to improve Evans and Ruf's approach described in section 3.2 was made. Possible improvement areas that were researched were improved MSS modeling, extending the data used, and using a better calibrated CYGNSS product. MSS was modeled with different modeling techniques, choice of control regions, and modeling argument. Table 5.2 shows that the MSS observations were more correlated with wind than relative wind-current speed, making wind the preferred modeling argument. Table 5.3 shows that the selection of modeling technique is not decisive, as multiple techniques show equally good performance. Lastly, the choice of control region also seems to not be decisive as Figure 5.9 show similar results for different MSS anomaly models based on different regions. The figure also shows that changing the matchup wind data source produces similar results, except that the range of values is shifted to a more negative region.

The microplastic models of Van Sebille et al. was like in previous attempts considered ground truth. Figure 5.10 and Figure 5.11 shows that the produced results show no model agreement, unlike what Evans and Ruf reported in Figure 3.7.

RQ3: *Is GNSS-R capable of measuring the global oceanic microplastic distribution?*

The question is impossible to fully answer, as no one knows the true global oceanic microplastic distribution. Van Sebille et al. created models that are well recognized, but they are only estimates, with uncertainties as section 3.1.1 highlighted. The GNSS-R methods use his models for comparison, and the models do agree in some areas.

A problem with both GNSS-R approaches, which argue against their capability to measure microplastics, is that they rely on the assumption that a reduction in surface roughening is seen in areas where microplastics are present. This is because peaks in NBRCS and negative MSS anomaly both, to some extent, capture the

same phenomenon, as NBRCS is inversely proportional to MSS. Peaks in NBRCS correspond to a sudden smoother surface. A negative MSS anomaly indicates a smoother surface than the given wind conditions, where wind is the primary driver behind ocean surface roughening. Both more peaks and more negative anomaly are being related to the presence of microplastics. Both sources report the same seasonal patterns, arguing that they measure a similar phenomenon. To answer the question, how justifiable the assumption is must be assessed. The results from section 3.4 concluded that plastic particles themselves do not have a damping effect on waves unless the surface coverage is unrealistically high (above 20 %). Even when considering unrealistic amounts, the effects were reported to be negligible.

The results from the experiment with surfactants did, however, conclude that the surfactants have a considerable damping effect on wind-driven roughening, as also earlier research suggested [Alpers and Hühnerfuss, 1989, Spivak et al., 2002, Kiefhaber et al., 2015]. As a consequence of the particles themselves not having a damping effect on waves, both approaches heavily rely on the assumption that surfactants and plastic particles share similar transport mechanisms. Evans and Ruf present [Van Sebille et al., 2020] as evidence suggesting they share similar transport mechanisms.Van Sebille et al. cites [D'Asaro et al., 2018] and states that "Surfactants, man-made and natural, accumulate in surface current convergence zones, where debris will also accumulate". Figure 3.10 highlights a problem with this as it shows seemingly high concentrations of microplastics being washed into the sea by rivers in highly populated areas. The figure perfectly fits the narrative, as highly populated areas in developing countries with poor waste management are causing substantial plastic pollution. However, the observed reduction of surface roughening, leading to the high predictions of microplastics flowing into the sea, is likely caused by something else. One reason for this is that the evidence suggesting surfactants share similar transport mechanisms is based on the accumulation of floating material caused by ocean surface current convergence zones. This does not mean that polluting a river with surfactants is followed by the pollution of microplastics. The assumption of shared transport mechanisms relies on the not instantaneous mechanisms of the oceans, causing accumulation over time. Dialogues with the Evans and Ruf revealed that they agreed on the issue and believed surfactants from agricultural spills likely caused the observed phenomenon. Consequently, at best, GNSS-R is only a tool that may help discover the accumulation zones.

Even when considering a hypothetical world where surfactants and microplastics always accompany each other, there exist problems with both methods. One example is the intertropical convergence zone, where MSS/NBRCS observations are known

to be more unreliable. Another example is that the world's oceans are complex systems, and that the ocean's surface is affected by many forces and phenomena. Currents, eddies, bathymetry, and more contribute to MSS/NBRCS variations, affecting the results. These effects should be modeled to compensate for their contribution, which is probably not feasible. How impacting these effects are, is hard to know. In Strand's approach, it may be possible that these effects do not vary significantly enough in proximity areas to cause sharp peaks very often? It may also be that Evans and Ruf's approach reduces these effects when averaging the MSS anomaly temporally.

Both Strand's and Evans and Ruf's methods show promising results despite the above-mentioned problems. Section 3.3.1 and section 3.2.1 highlighted possible areas of improvement. As both approaches attempt to accomplish the same goal, their relationship should be further examined to answer which method is superior to the other. Quantitative comparisons of Strand's results produced in this thesis reveal that his peaks show a correlation of 0.32 with existing models, which obviously is not perfect. Despite this, it does, however, demonstrates that GNSS-R, to some extent, is capturing a phenomenon related to microplastics. This is emphasized by the fact that the other GNSS-R approach produced Figure 3.7, where most samples of MSS anomaly show good model agreement.

GNSS-R may help indicate where microplastics accumulate, as there is evidence that surfactants and debris accumulate in surface current convergence zones. Despite the drawbacks of both approaches, they are, in my opinion, highly interesting for further research. Manually measuring the global microplastic concentration is a severe and probably unfeasible task. If the satellites, to some extent, could aid in monitoring possible accumulation zones, GNSS-R could become a valuable tool to help streamline the cleaning of our precious oceans. The CYGNSS mission, which provides data for both GNSS-R methods, was initially only designed with a two-year lifespan but is still functioning after approximately five years [Ruf et al., 2016]. It is, however, likely that the life of the satellites is reaching its end sooner than later. To better understand and emphasize the threat of microplastics to our world, techniques like remote sensing should be researched, as they can help uncover the severity of the problem. Hopefully, raising awareness will result in actions taken to preserve today's diverse and vulnerable ecosystem.

6.2 Further Work

To improve the results from both GNSS-R approaches, it is likely reasonable to smoothen the microplastic predictor spatially as the value of neighboring grid points often varies more in the MSS anomaly/detected peaks domain than in the microplastic domain. Additionally, as mentioned above, the MSS/NBRCS is affected by several different parameters, like currents, eddies, bathymetry, and more. These parameters should be further researched and examined synchronously to compensate for their contribution to MSS/NBRCS variations. To get a more comprehensive understanding of what phenomenon the GNSS-R techniques are capturing, I encourage further research on the relationship between surfactants and plastic particles. To what extent the assumption that they share similar transport mechanisms is correct dictates how capable the methods are of detecting microplastics.

As for the possible improvement areas highlighted in section 3.3.1 of Strand's approach, I believe that they are both feasible and reasonable to implement. Recalculating peaks scaled by the number of observations is very logical. Considering width, prominence, and wind speeds as hyperparameters and evaluating results against Van Sebille's models is also reasonable. Additionally, developing a microplastic predictor based on detected peaks is lacking in Strand's approach. By regressing exponentially to microplastic mass, RMSE can be calculated, which can be used to compare the two methods.

Considering Evans and Ruf's results, I encourage further investigations of the samples violating the desired relationship, seen in Figure 3.7. Plotting these samples may reveal spatial patterns in MSS due to, for instance, known local phenomena.

Bibliography

Werner Alpers and Heinrich Hühnerfuss. The damping of ocean waves by surface films: A new look at an old problem. *Journal of Geophysical Research: Oceans*, 94(C5):6251–6265, 1989.

Rajeswari Balasubramaniam and Christopher S Ruf. Improved calibration of cygnss measurements for downbursts in the intertropical convergence zone. In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, pages 3987–3990. IEEE, 2018.

Gabriele Bulian, Alberto Francescutto, and Claudio Lugni. On the nonlinear modeling of parametric rolling in regular and irregular waves. *International shipbuilding progress*, 51(2-3):173–203, 2004.

Kieran D Cox, Garth A Covernton, Hailey L Davies, John F Dower, Francis Juanes, and Sarah E Dudas. Human consumption of microplastics. *Environmental science & technology*, 53(12):7068–7074, 2019.

Andrés Cózar, Fidel Echevarría, J Ignacio González-Gordillo, Xabier Irigoien, Bárbara Úbeda, Santiago Hernández-León, Álvaro T Palma, Sandra Navarro, Juan García-de Lomas, Andrea Ruiz, et al. Plastic debris in the open ocean. *Proceedings of the National Academy of Sciences*, 111(28):10239–10244, 2014.

Zhounan Dong and Shuanggen Jin. Evaluation of the land gnss-reflected ddm coherence on soil moisture estimation from cygnss data. *Remote Sensing*, 13(4):570, 2021.

Eric A D'Asaro, Andrey Y Shcherbina, Jody M Klymak, Jeroen Molemaker, Guillaume Novelli, Cédric M Guigand, Angelique C Haza, Brian K Haus, Edward H Ryan, Gregg A Jacobs, et al. Ocean convergence and the dispersion of flotsam. *Proceedings of the National Academy of Sciences*, 115(6):1162–1167, 2018.

Madeline C Evans and Christopher S Ruf. Toward the detection and imaging of

ocean microplastics with a spaceborne radar. *IEEE Transactions on Geoscience and Remote Sensing*, 2021.

Scott Gleason and Christopher Ruf. Overview of the delay doppler mapping instrument (ddmi) for the cyclone global navigation satellite systems mission (cygnss). In *2015 IEEE MTT-S International Microwave Symposium*, pages 1–4. IEEE, 2015.

Robert C Hale, Meredith E Seeley, Mark J La Guardia, Lei Mai, and Eddy Y Zeng. A global perspective on microplastics. *Journal of Geophysical Research: Oceans*, 125(1):e2018JC014719, 2020.

J Hausman, F Bonjean, and K Dohan. Ocean surface current analysis (oscar) third degree resolution user's handbook. *California Institute of Technology contract with the National Aeronautics and Space Administration and the National Oceanic and Atmospheric Administration. California*, 2009.

Mostafa Hoseini and Hossein Nahavandchi. The potential of GNSS Reflectometry for detecting ocean surface currents. *submitted to the journal of Remote Sensing of Environment*, 2022.

Norden E Huang, Davidson T Chen, Chi-Chao Tung, and James R Smith. Interactions between steady won-uniform currents and gravity waves with applications for current measurements. *Journal of Physical Oceanography*, 2(4):420–431, 1972.

Malek Karaim. *Ultra-tight GPS/INS Integrated System for Land Vehicle Navigation in Challenging Environments*. PhD thesis, Queen's University (Canada), 2019.

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

D Kiefhaber, Christopher J Zappa, and B Jähne. Influence of natural surfactants on short wind waves in the coastal peruvian waters. *Ocean Science Discussions*, 12(4):1291–1325, 2015.

Kara Lavender Law and Richard C Thompson. Microplastics in the seas. *Science*, 345(6193):144–145, 2014.

LC-M Lebreton, SD Greer, and Jose Carlos Borrero. Numerical modelling of floating debris in the world's oceans. *Marine pollution bulletin*, 64(3):653–661, 2012.

Hsueh-Jyh Li, Yean-Woei Kiang, and WK Chen. Radar and inverse scattering. *The Electrical Engineering Handbook*, pages 671–690, 2005.

Nikolai Maximenko, Jan Hafner, and Peter Niiler. Pathways of marine debris derived from trajectories of lagrangian drifters. *Marine pollution bulletin*, 65(1-3):51–62, 2012.

Mahmoud Rajabi, Hossein Nahavandchi, and Mostafa Hoseini. Evaluation of cygnss observations for flood detection and mapping during sistan and baluchestan torrential rain in 2020. *Water*, 12(7):2047, 2020.

Jennifer Reynolds, Maria Paola Clarizia, and Emanuele Santi. Wind speed estimation from cygnss using artificial neural networks. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13:708–716, 2020.

Chelsea M Rochman. Microplastics research—from sink to source. *Science*, 360 (6384):28–29, 2018.

Christopher Ruf, Derek Posselt, Sharanya Majumdar, Scott Gleason, M.P. Clarizia, Derek Starkenburg, Damen Provost, Valery Zavorotny, John Murray, Stephen Musko, Zorana Jelenak, Paul Chang, and Mary Morris. *CYGNSS Handbook*. 04 2016. ISBN 978-1-60785-380-0.

B Spivak, J-M Vanden-Broeck, and T Miloh. Free-surface wave damping due to viscosity and surfactants. *European Journal of Mechanics-B/Fluids*, 21(2):207–224, 2002.

Andreas Gaustadnes Strand. The application of spaceborne gnss-reflectometry for the study of ocean microplastics, 2021. URL `https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2826369`.

Yukun Sun, Christopher Ruf, Thomas Bakker, and Yulin Pan. Effects of microplastics and surfactants on surface roughness of water waves. *arXiv preprint arXiv:2111.07021*, 2021.

Other Global Navigation Satellite Systems. Other global navigation satellite systems, 2021. URL `https://www.gps.gov/systems/gnss/`.

National Oceanic US Department of Commerce and Atmospheric Administration. What are microplastics?, 2021. URL `https://oceanservice.noaa.gov/facts/microplastics.html`.

Erik Van Sebille, Matthew H England, and Gary Froyland. Origin, dynamics and evolution of ocean garbage patches from observed surface drifters. *Environmental Research Letters*, 7(4):044040, 2012.

Erik Van Sebille, Chris Wilcox, Laurent Lebreton, Nikolai Maximenko, Britta Denise Hardesty, Jan A Van Franeker, Marcus Eriksen, David Siegel, Francois Galgani, and Kara Lavender Law. A global inventory of small floating plastic debris. *Environmental Research Letters*, 10(12):124006, 2015.

Erik Van Sebille, Stefano Aliani, Kara Lavender Law, Nikolai Maximenko, José M Alsina, Andrei Bagaev, Melanie Bergmann, Bertrand Chapron, Irina Chubarenko, Andrés Cózar, et al. The physical oceanography of the transport of floating marine debris. *Environmental Research Letters*, 15(2):023003, 2020.

David Wells, Norman Beck, Alfred Kleusberg, Edward J Krakiwsky, Gerard Lachapelle, Richard B Langley, Klaus-peter Schwarz, James M Tranquilla, Petr Vanicek, and Demitris Delikaraoglou. Guide to gps positioning. In *Canadian GPS Assoc.* Citeseer, 1987.

Nicholas J Willis. *Bistatic radar*, volume 2. SciTech Publishing, 2005.

Fan Wu, Wei Zheng, Zhaowei Li, and Zongqiang Liu. Improving the gnss-r specular reflection point positioning accuracy using the gravity field normal projection reflection reference surface combination correction method. *Remote Sensing*, 11 (1):33, 2018.

Qingyun Yan, Weimin Huang, and Giuseppe Foti. Quantification of the relationship between sea surface roughness and the size of the glistening zone for gnss-r. *IEEE Geoscience and Remote Sensing Letters*, 15(2):237–241, 2017.

Valery U Zavorotny, Scott Gleason, Estel Cardellach, and Adriano Camps. Tutorial on remote sensing using gnss bistatic radar of opportunity. *IEEE Geoscience and Remote Sensing Magazine*, 2(4):8–45, 2014.