Utrecht University

Institute for
Marine and Atmospheric
research Utrecht

Comparing models and observations of the surface accumulation zone of floating plastic in the North Atlantic subtropical gyre

A thesis submitted by

Sophie Schmiz

in partial fulfillment of the
requirements for the degree of

Master of Science

supervised by

Dr. Erik van Sebille
Darshika Manral, M.Sc.

in

Utrecht, August 2021

Contents

Abstract **iii**

Acknowledgements **iv**

1 Introduction **1**
1.1 A shallow dive into floating plastics 1
1.2 Transport of tracers by surface currents in the North Atlantic 3
1.3 Review of tracer accumulation models in the North Atlantic 4
1.4 Aim of this thesis 8

2 Observations **9**
2.1 Earlier data 9
2.2 Trawling for additional data 9
2.3 Combining the data sets 12

3 Simulations with Ocean Circulation Models **16**
3.1 Lagrangian simulations with Parcels 16
3.2 Ocean circulation models 18
3.3 Determining the simulation runtime 20
3.4 Final distributions 20

4 Combining Models and Observations **31**

5 Results **37**
5.1 Different release scenarios 37
5.2 Age of plastic particles 43
5.3 Influence of observations sampled along the coastline 50

6 Discussion **56**
6.1 Biases in the observational data 56
6.2 Notes on the simulations 57
6.3 Comparison of models and observations 58
6.4 Inclusion of subgrid-scale processes 59

7 Conclusion **60**

Appendix A Additional Figures **62**

Bibliography **69**

Abstract

Plastics have been invented as a durable, reusable alternative to other organic materials. Their mass production started only in the 20th century, but plastic pollution is already now one of humankind's most enduring legacies, penetrating virtually every ecosystem on this planet. Marine plastics are especially persistent and hard to clean up, as they often sink down to the seafloor or fragment into smaller pieces that are called microplastics if they are smaller than 5 mm in size, accounting for over 90% of all floating plastic particles. In order to protect the marine environment, it is therefore essential to understand the sources, transport and sinks of floating debris. Only this way, a comprehensive evaluation of the risk caused by plastic pollution to the marine ecosystem can be achieved. To do so, observations of plastic samples are interpolated with predictions from ocean general circulation models to create maps of the regions that are most vulnerable to plastic pollution, for instance the surface accumulation zones in the center of the subtropical gyres.

This thesis focusses on the surface accumulation zone of microplastics in the North Atlantic subtropical gyre, where previous modelling efforts have failed to agree with the observed accumulation pattern derived from microplastic samples. With the Lagrangian particle tracking tool Parcels, virtual particle distributions are generated in the North Atlantic, based on the hydrodynamic fields from four different ocean general circulation models: GlobCurrent, NEMO, SMOC (a CMEMS product) and HYCOM. The outcomes of these models are compared against each other, predicting the maximum of the surface accumulation in the North Atlantic subtropical gyre to lie between 25 and 55 °W. The models' prediction powers are tested against the most complete data set of microplastic samples in the North Atlantic yet, building up on the data set used in the global inventory of small floating plastic debris aggregated by van Sebille et al. (2015). Still, observations in the eastern part of the North Atlantic remain too sparse to show the true zonal extent of the surface accumulation zone. A linear regression between the observed concentrations and the kernel density estimate of the virtual particle distributions generated by a model measures the correlation between a model and the observations. Generally, the most realistic release scenario of virtual particles in the Lagrangian simulations, which includes a repeated coastal release and a sinking timescale, results in a surface accumulation zone that matches best to the observations in the North Atlantic subtropical gyre. Because observations are still lacking around the predicted maxima of the surface accumulation zone, it is too early to claim that one model corresponds best to the observations.

Acknowledgements

Robert F. Kennedy once said: "There is a Chinese curse which says, 'May he live in interesting times.' Like it or not, we live in interesting times."

Despite the extraordinary circumstances of the last year and a half, I was lucky enough to have had a very supportive team around me as I was working on this thesis. First, I would like to thank my thesis supervisors: Erik for introducing me to an inspiring community of oceanographers, for welcoming me into the Parcels group at Utrecht University and for pushing me whenever I needed it the most. Darshika, thank you for always being only a Teams call away and for solving all my technical problems so much faster than I would have alone. Everyone in the Parcels group contributed to an extremely friendly, approachable and supportive atmosphere, thank you all for that! More specific thanks go out to Reint for your patience and advice everytime I was struggling with Parcels and Gemini. Victor and Claudio, thanks for all the pieces of code you let me borrow, and also Mikael, Laura and Delphine, thank you all for allowing me to throw my ideas at you – whatever bounced back made its way into this thesis in a better shape. And finally, last but not least, Tycho and Ina, thank you so much for our caffeine-fuelled support sessions, without you, this experience would have only been half as fun.

This thesis has been written as part of the TOPIOS ("Tracking Of Plastic In Our Seas") and AtlantECO projects funded by the European Research Council under the European Union's Horizon 2020 programme for research and innovation.

Chapter 1

Introduction

1.1 A shallow dive into floating plastics

In the early 1960s, the Swedish engineer Sten Gustaf Thulin invented a method to produce a shopping bag from one simple, flat piece of plastic, shortly after the start of plastic mass production after World War II. Thulin intended for the plastic bag to become a reusable, durable alternative to the paper bag (Weston 2019), the production of which required chopping down forests and is to this day four times as energy intensive as the production of a plastic bag (Bell and Cave 2011). It was the advent of the "plastic age", when the possibilities of industrial plastic use seemed endless and the environmental repercussions negligible.

Fast forward 30 years further into the "plastic age" to 1997, when the American oceanographer Charles J. Moore set out on a cruise into the Pacific Ocean, where he discovered the consequences of the post-war plastic optimism: the Great Pacific garbage patch, an accumulation of slowly degrading, floating plastic debris between Hawaii and California. Today, it has an extent of 1.6 million km^2 (Lebreton et al. 2018) with more than 79 thousand tons of plastic floating inside it. Almost half of this mass stems from fishing nets, but microplastics – plastic pieces smaller than 5 mm in size – account for 94% of all floating plastic particles. On average, about 0.7 items float in one square meter of the Great Pacific garbage patch, which is therefore not really a plastic "island", as it is sometimes referred to, but more of a plastic soup. The Pacific, Atlantic and Indian Oceans all contain garbage patches, but none of those match the one in the North Pacific. The North Atlantic equivalent is considerably smaller because the ocean basin itself is smaller and the amount of plastic entering to ocean from the coastline does not compare to the immense amount released from the East Asian coastline. Early samples of microplastics in the Atlantic were mostly confined to the Western part of the North Atlantic (Law et al. 2010) where the highest concentrations reached about 0.2 particles per m^2. Summing up all ocean basins, van Sebille et al. (2015) estimated that the total amount of small, floating plastic debris lies between 15 and 51 trillion particles, weighing together about 93 to 236 thousand metric tons.

There are different types of marine plastic debris: On one hand, floating fishing debris, macroplastics and large microplastics down to sizes of 1 mm can be captured during efforts to clean up marine debris, according to The Ocean Cleanup, an organization that aims to remove plastic debris from the ocean surface. Smaller micro- and nanoplastics on the other hand, that are often fragments of the larger pieces, still evade these efforts. Microplastics are usually classified according to their generation, distinguishing between primary and secondary microplastics. Primary microplastics are produced as such, for example as small resin pellets intended for industrial use, microfibres from synthetic clothing, abrasive particles in cleaning agents or microbeads in cosmetics. After usage,

they often enter waste management plants where they should be filtered out, but a fraction still escapes (Talvitie et al. 2017). Secondary microplastics are fragments of larger pieces that break down in the marine environment because of mechanical abrasion from wind and waves or from the action of small organisms such as microbes, from the exposure to sunlight or heat or from hydrolysis in contact with water, as Andrady (2011) notes (see also Kaandorp, Dijkstra, and Sebille 2021). Larger pieces of plastic are often carried into the drainage systems and rivers by wind and rain, before they reach the oceans as mismanaged waste.

But microplastics are not only found in the ocean waters. They have been contaminating Antarctic glaciers (González-Pleiter et al. 2021), the sediments of remote deep sea trenches (Jamieson et al. 2019) and even pristine Tibetan glaciers (Zhang et al. 2021), since winds can transport land-borne particles to these remote places. Even now, humankind may never get rid of the microplastic pollution that has reached the planet's most remote corners in the last decades.

While some organisms adapt to the presence of microplastics to form the "plastisphere" (Zettler, Mincer, and Amaral-Zettler 2013), such as the sea skater who uses the small particles to deposit its eggs (Goldstein, Rosenberg, and Cheng 2012), micro- and nanoplastics in ocean waters remain mostly a threat to marine life, especially when they are ingested (Andrady 2011). They can be retained in the tissue of filter feeders like mussels and barnacles, altering their feeding behaviour and reducing their resilience. Often coated in toxins and absorbing other marine pollutants, microplastics also modify the brain structure of seabirds and the fetal development of other marine mammals upon consumption (Kühn, Bravo Rebolledo, and Franeker 2015).

Ultimately, micro- and nanoplastics can reach higher trophic levels, up until their ingestion by humans in sea food, where we are left with health hazards quite similar to what marine organisms are facing. The human hormonal balance is highly sensitive to the introduction of tiny particles in the bloodstream, which has led among other things to a decrease in human fertility rates (D'Angelo and Meccariello 2021). Microplastics have also been found in the human placenta where they impact embryonic development (Ragusa et al. 2021).

The urgency of these issues becomes even more evident considering that the global plastic production and the amount of plastic entering the oceans alike are still increasing exponentially (Wilcox, Hardesty, and Law 2019; Ostle et al. 2019), like many human activities as part of the so-called "Great Acceleration" of the Anthropocene (Brandon, Jones, and Ohman 2019). These challenges can only be tackled if the sources, the transport and the fate of microplastics are well understood, so that the marine plastic budget can be closed (Hardesty et al. 2017). Ocean general circulation models are valuable tools for studying and predicting the transport and accumulation of microplastics by modelling the physical processes that govern the ocean currents. Comparing transport routes and accumulation zones to environmental sensitivity maps or to distributions of marine wildlife (such as the microbial maps studied in AtlantECO) allows researchers to grasp the extent of the risks that microplastics pose to marine ecosystems. After all, healthy oceans play a key role in Earth's climate system due to their ability to absorb heat and dissolve atmospheric carbon dioxide.

But studying the transport of microplastics with ocean general circulation models is a two-way street: Observational records of floating particles can help validate the quality of a model or point out its shortcomings, thus aiding in better understanding the physical processes that govern the ocean circulation. The records of any kind of small flotsam act as a proxy for convergence zones of surface currents, where their buoyancy prevents them from downwelling. The most common tracers are microplastics and together with any other small, buoyant plastic particles that are heavy enough not to be affected by windage, they make up 99.9% of all debris counts (Lebreton et al. 2018).

Still, these small floating particle account for less than 1% of the estimated plastic reservoir in the oceans (Eriksen et al. 2014; Jambeck et al. 2015; Lebreton, Egger, and Slat 2019). Around 77%

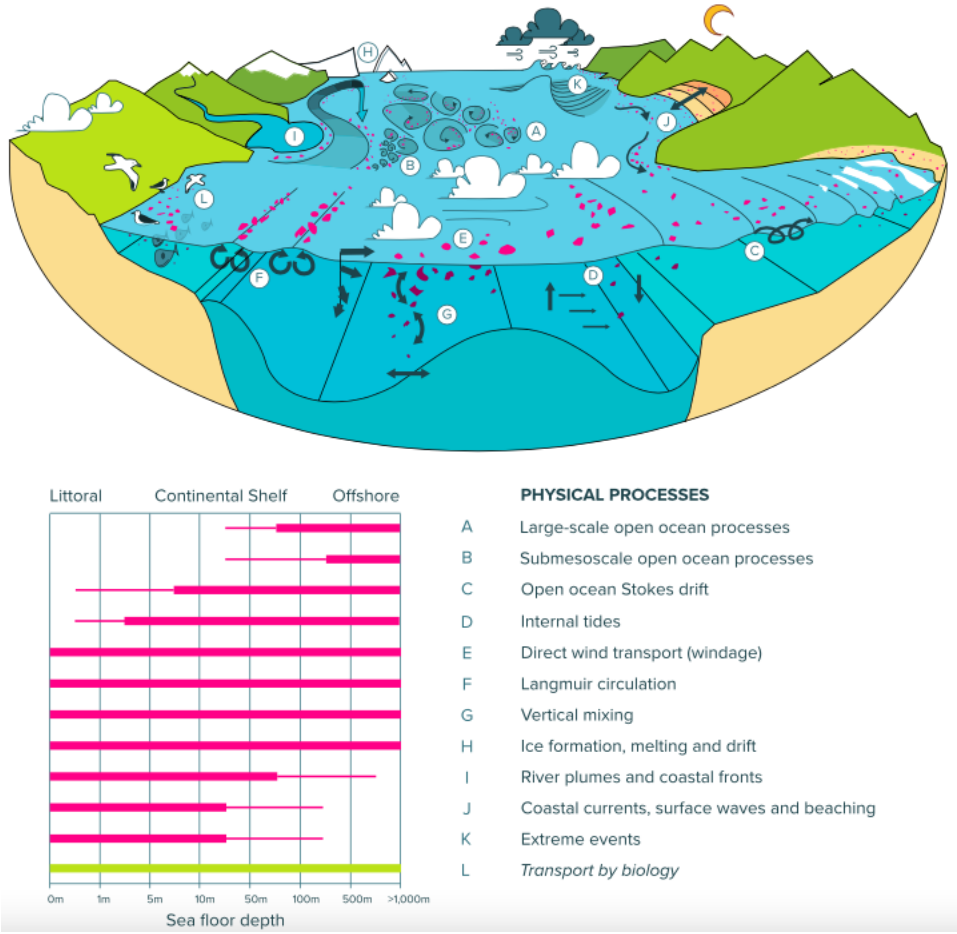


Figure 1.1: Illustration of the physical processes governing the transport of marine plastic debris (pink particles). The width of the pink line in the lower panel indicates the importance of a physical process in different regions, while the green line signifies that the transport by organisms is not a physical process. Figure taken from van Sebille et al. (2020).

of the global marine plastic is expected to beach at the continental shores or circulate in coastal waters (Onink et al. 2021; Olivelli, Hardesty, and Wilcox 2020). Another share is estimated to be suspended in the water column due to turbulent mixing (Kooi et al. 2016). Finally, many plastic particles lose their buoyancy when their hydrophobic surface is coated in a film of organic material or when particles aggregate with denser organic particles. They then sink to the seabed, where the particles are increasingly deposited in deep-sea sediments (Brandon, Jones, and Ohman 2019), especially in the Atlantic Ocean (Woodall et al. 2014).

1.2 Transport of tracers by surface currents in the North Atlantic

The transport of floating marine debris is governed by different processes at different scales (see Figure 1.1). In the open ocean, floating particles mainly accumulate inside the subtropical gyres. These gyres form in every major ocean basin where the Ekman transport associated with the easterly trade winds around the equator and with the westerlies in the mid-latitudes leads to converging water masses in the subtropics. The sea level as well as the hydrostatic pressure rise in the subtropics and a pressure gradient force develops, directed from high to low pressure areas.

Deflected by the Coriolis force, the resulting geostrophic current flows along the isobars, forming an anticyclonic gyre.

This is why the accumulation zone of floating particles is mainly determined by the wind-driven Ekman drift at the surface, converging at the same latitudes as the integrated Ekman transport around 30 °N and trapping the particles inside the gyre. Kubota, Takayama, and Namimoto (2005) have used satellite altimetry and wind data to study the large-scale circulation pattern in the North Atlantic and found that marine debris is first moved into the Ekman convergence zone by the Ekman surface currents, where it is then picked up by the geostrophic currents and transported eastwards. At the same time, Kubota, Takayama, and Namimoto (2005) wondered whether the Ekman drift associated with meridional winds along the basin boundary may be responsible for shifting the center of the accumulation zone towards the center of the basin.

Even though the geostrophic and the Ekman currents are the most well-studied contributions to the large-scale circulation, at the surface, the magnitude of the Stokes drift might be comparable to the Ekman drift (Maximenko, Hafner, and Niiler 2012). The Stokes drift is the net horizontal transport induced by surface waves as tracers move faster over wave crests than through the troughs. For light microplastics like polystyrene, windage is considered to be one of the main responsible factors for microplastic transport in addition to the Stokes drift (Chubarenko et al. 2016). Cardoso and Caldeira (2021) also found that wind plays a considerable role in the transport of surface particles over basin-scale distances.

The surface transport of buoyant debris also depends on the local eddy activity. Brach et al. (2018) note that in the North Atlantic, anticyclonic mesoscale eddies accumulate (and potentially transport) microplastics in their downwelling core. In the South Pacific, floating marine debris tends to accumulate in regions of lower eddy kinetic energy (Martinez, Maamaatuaiahutapu, and Taillandier 2009), but Onink et al. (2019) were not able to reproduce this finding in the North Atlantic, possibly because of the smaller basin size and consequently weaker eddy activity.

Finally, it is important to remember that the ocean basins are interconnected on a global scale. Therefore, an important sink for plastics released into the North Atlantic is the Arctic Ocean, the terminus of particles transported via the North Atlantic branch of the thermohaline circulation (Cózar et al. 2017) and by Stokes drift (Onink et al. 2019). Wichmann, Delandmeter, and Sebille (2019) found that the vertical shear of ocean currents also affects the horizontal distribution of microplastic transport, transporting submerged particles from the North Atlantic to the Arctic, as they do not accumulate in the surface convergence zones.

1.3 Review of tracer accumulation models in the North Atlantic

Ocean circulation models are often used to interpolate spatial distributions of microplastic observations in the oceans, in order to construct maps of the most vulnerable regions. The surface currents in the Pacific Ocean have been modelled in a number of studies (Lebreton, Greer, and Borrero 2012; Maximenko, Hafner, and Niiler 2012; van Sebille, England, and Froyland 2012; Kubota, Takayama, and Namimoto 2005; Martinez, Maamaatuaiahutapu, and Taillandier 2009; Onink et al. 2019) and generally agreed well with the patterns observed during extensive surface trawls for small, floating plastics (van Sebille et al. 2015). However, in the North (and South) Atlantic, the outputs of recent modelling experiments (Lebreton, Greer, and Borrero 2012; Maximenko, Hafner, and Niiler 2012; van Sebille, England, and Froyland 2012; Onink et al. 2019) did not seem to match the accumulation patterns inferred from micro- and mesoplastics surface trawls (van Sebille et al. 2015).

Law et al. (2010) were the first to present a systematic survey of plastic concentrations in the

North Atlantic, albeit limited to the Western North Atlantic and to the Caribbean Sea. They found a latitudinal band at 30 °N of high concentrations up to 0.2 particles per m^2, but did not make any estimates about the zonal extent of this accumulation zone, because they did not take any samples in the central or eastern part of the North Atlantic.

The pattern derived from these observations, stemming from more than 6000 surface trawls over a period of 22 years, was compared to trajectories from the Global Drifter Program, as well as to a numerical model based on these trajectories. After advecting an initially uniform distribution of virtual particles for a period of 10 years, Law et al. (2010) found that drifters passing through the Western North Atlantic around 30 °N, where observed plastic concentrations were highest, also tended to stay in this area, indicating an attractor of floating particles in this part of the subtropical gyre. Law et al. (2010) did not make any claims about where floating particles ended up that did not cross this region of highest concentrations.

Maximenko, Hafner, and Niiler (2012) also worked with drifter data from the Global Drifter Program to set up a probabilistic model of the ocean circulation using transition matrices. A transition matrix is often used to describe Markov chains, i.e. stochastic models in which the probability of an event only depends on the state of the system at the previous time step. Even though individual drifters may not always follow the exact currents, one can assume that a large number of drifters still gives an accurate representation of their final accumulation zone. This probabilistic approach extends the sparse observational data set inferred from the drifters that often do not travel long enough to reveal the large-scale ocean connectivity. Each element of the transition matrix states the probability that a drifter travels to another grid cell in a certain time frame, given the observed drifter trajectories. Maximenko, Hafner, and Niiler (2012) calculated the probability that in the next 5 days, a drifter would travel to a surrounding cell in a $0.5° \times 0.5°$ grid, assuming statistical stationarity. This, however, eliminated any seasonal variability in the resulting modelled circulation. Even though Maximenko, Hafner, and Niiler (2012) compared their model outcome to surface measurements of microplastics, their drifters were drogued at 15 m. Surprisingly, these drifters seemed to accumulate in the same region as their undrogued counterparts, suggesting that the currents at 15 m depth are somewhat representative of the surface currents. In terms of sinks, this model only included a very simplified version of beaching by removing particles that entered a coastal cell. No other removal processes (e.g. sinking due to biofouling) were considered.

Multiplying a global uniform distribution of virtual particles with the transition matrices for a circulation of 10 years, Maximenko, Hafner, and Niiler (2012) found accumulation zones in each of the subtropical gyres. The largest accumulation zone was generated in the South Pacific because the uniform release led to the largest number of particles circulating in this basin, although in reality, the largest input of plastics in the oceans is in the North Pacific (Jambeck et al. 2015). In the North Atlantic, a dynamic accumulation pattern built up in the subtropics around 30 °N after 2 to 3 years. It stretched across the entire zonal extent of the North Atlantic at this latitude, with its maximum around 55 °W. In the long run, though, diffusion allowed virtual particles to leave the accumulation zones, dissolving them over time.

Van Sebille, England, and Froyland (2012) also built a statistical model of the ocean surface circulation from the drifter data set, half the buoys of which remained drogued at 15 m, while the other half had lost their drogue. From their trajectories, transition matrices were determined on a $1° \times 1°$ grid for periods of 60 days to capture seasonal variabilities. Virtual particles were released along the coastlines in six pulses over the course of the first modelling year, scaled to the population density within 200 km of the coastline. The final model did not include any beaching nor sinking parametrization and thus conserved the number of particles over the course of the simulation. At the end, particles accumulated in the subtropical gyres and in the Barents Sea, while again the extent of the North Atlantic accumulation zone covered the entire width of the

basin. The dynamical aspect of the surface circulation was revealed after 1,110 years, when almost all of the dynamical features of the accumulation zones in the South Atlantic and Indian Ocean had disappeared, establishing the North Pacific subtropical gyre as the main attractor.

Lebreton, Greer, and Borrero (2012) tracked the trajectories of virtual particles in a Lagrangian simulation according to the flow field data of the HYCOM/NCODA model with a spatial resolution of 1/12° (see also Section 3.2). Virtual particles were seeded continuously at major river mouths, along coastlines scaled to the coastal population density, and along major shipping routes to account for a variety of different input scenarios, with the particle number increasing every year. The particles were then advected with the Lagrangian particle tracking model Pol3DD (more information about Lagrangian particle trackers can be found in Section 3.1) for 30 years, by looping data sets of 6 years. Again, neither beaching nor sinking were included in this model. Particles tended to accumulate mostly in the subtropical gyres of the Northern Hemisphere, implying that plastic input into the oceans scales with economic activity. In the North Atlantic, the highest concentration of virtual particles is observed around 30 °N and 55 °W.

The most recent modelling work in the North Atlantic has been done by Onink et al. (2019), who studied the role of the geostrophic current, the surface Ekman current and the Stokes drift in the accumulation of floating microplastics in subtropical gyres, following a suggestion by Maximenko, Hafner, and Niiler (2012) that the magnitude of the Stokes drift might be comparable to that of the surface Ekman drift. The geostrophic and Ekman current were modelled after satellite and in situ data from the GlobCurrent data set (see also Section 3.2), while the Stokes drift was estimated separately from the surface wave field of the WaveWatch III reanalysis. Adding the Stokes drift to the GlobCurrent data might have overestimated the effect of the Stokes drift, since it is at least partially included in the GlobCurrent observations already, but Onink et al. (2019) still considered the sum of the GlobCurrent and WaveWatch III data in a Lagrangian simulation. Initialized as a uniform $0.5° \times 0.5°$ grid, particles were advected over the course of 13 years.

Despite not including any subgrid-scale dispersion, nor beaching or sinking parametrization, Onink et al. (2019) still managed to recover the results by Kubota (1994), stating that in the North Atlantic, Ekman currents lead to surface accumulation around 30 °N, while geostrophic currents only contribute to the zonal dispersion of the accumulation zone. The angle between the Ekman currents from the GlobCurrent data set and the wind stress is 30.75°, which is smaller than the 45° angle predicted by Ekman theory. The increased zonal component of the Ekman current appeared to contribute to the zonal transport as well. The inclusion of Stokes drift generated transport to the Arctic and narrowed the peak of the accumulation zone to a region between 55 and 60 °W.

The three models published in 2012 (Maximenko, Hafner, and Niiler 2012, van Sebille, England, and Froyland 2012, Lebreton, Greer, and Borrero 2012) were reviewed and compared to the accumulation pattern revealed in the concentration pattern of small, buoyant plastics trawled at the sea surface (van Sebille et al. 2015). This global data set of small, floating plastics is largely based on the data collected by Law et al. (2010) in the North Pacific and North Atlantic and has been spatially interpolated and standardized to account for different sampling conditions (more details in Section 2.1). For better comparison with the other two models, the review of Lebreton, Greer, and Borrero (2012) was limited to the coastal scenario. In each of the three models, the North Atlantic subtropical gyre showed the lowest concentrations of virtual tracers compared to the other subtropical gyres. Across these three models as well as the model run by Onink et al. (2019), the highest concentration in the North Atlantic was consistently found around 30 °N and 55 °W, which did not coincide with the location of the peak in the observational data set around 40 °W.

For a more rigorous statistical comparison, the final distributions of the virtual particles were binned onto a $1° \times 1°$ grid, to allow for a point-by-point comparison at each sampling location, where the number of virtual particles aggregated from the virtual distribution was compared to the

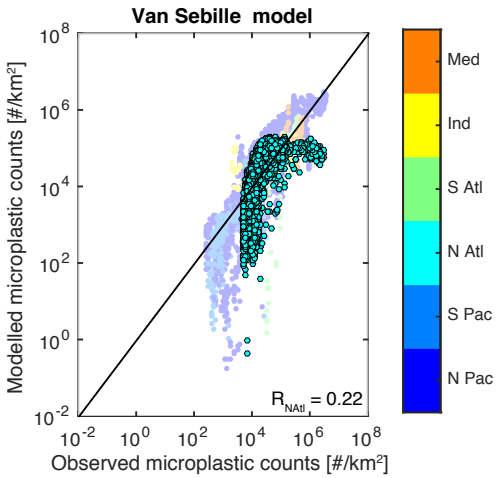


Figure 1.2: Scatter plot comparing the virtual particle distribution generated by the van Sebille, England, and Froyland (2012) model to the concentrations observed in surface trawls, as collected in van Sebille et al. (2015).

actual observed concentration. When these pairs are displayed in a scatter plot such as Figure 1.2, the data points show the correspondence of the model prediction to the observed values. A 1-to-1 correspondence would result in the bisector giving by the black line, but the North Atlantic data points follow a hook, or Γ-shape. The correlation coefficient r determined from a linear regression gives a quantitative measure for the goodness of correspondence between model and observations; the van Sebille, England, and Froyland (2012) model only reached $r = 0.22$ in the North Atlantic, which corresponds to a coefficient of determination of $R^2 = r^2 = 0.0484$ (Figure 1.2). The Lebreton, Greer, and Borrero (2012) and Maximenko, Hafner, and Niiler (2012) models gave regression plots that agreed qualitatively with Figure 1.2; see van Sebille et al. (2015) for more details.

Aside from the North Atlantic, a similarly large data set of observations has only been collected in the North Pacific. The models by Maximenko, Hafner, and Niiler (2012), van Sebille, England, and Froyland (2012), Lebreton, Greer, and Borrero (2012) and Onink et al. (2019) matched these observations rather well, as the location of the peak in the modelled distributions coincided with the peak in the observations around 30 °N and 50 °W. The van Sebille, England, and Froyland (2012) model was compared to the observational data set in Figure 1.2, in which the data points follow a straight line, at least for concentrations larger than 1,000 particles per km^2. It is thus reasonable to say that the models performed sufficiently in the North Pacific and that this approach to compare models to observations is valid.

The initialization of a simulation may also influence its final accumulation pattern. According to Wichmann et al. 2019, any information about the initial input location of a virtual particle should be lost when running a model for longer than the mixing time in a given basin, but Chenillat et al. (2021) found that the release scenario does in fact play a role in the accumulation of virtual particles in the subtropical gyre. They ran global Lagrangian simulations with GLORYS12V1 data with different source functions for 22 years, and noted that particles released at major river mouths ended up circulating in coastal waters retained by the rivers. Particles released more evenly along the coast, e.g. according to the population density within 200 km of the coastline (essentially as in van Sebille, England, and Froyland 2012), accumulated in the subtropical convergence zones, especially in the North Atlantic. This behaviour persisted even after running a global simulation for 22 years, which is longer than the mixing time in each ocean basin (Wichmann et al. 2019). Even though Chenillat et al. (2021) employed only a very simplified coastal parametrization that

risked falsifying the behaviour of virtual particles released in river mouths, their results indicate that the release scenario might substantially influence the accumulation of virtual particles in the subtropical gyres.

1.4 Aim of this thesis

In this thesis, I will try to unravel why previous modelling efforts so far have failed to reproduce the observed accumulation pattern of small, floating plastics in the North Atlantic. The North Atlantic has already been trawled extensively for micro- and mesoplastics, but mostly in the western part of the basin. Does this existing data set contain a bias that can be removed by extending it with novel observations? How does the modelled accumulation pattern change with different flow field data, and is there a model that aligns best with the observations? Does a more realistic input (and removal) scenario improve the correlation between modelled distributions and observations?

It should be noted that while observations are still very sparse in the South Atlantic, models show a comparably poor correlation to the observations like in the North Atlantic, even though the South Atlantic probably contains a larger amount of floating plastics (van Sebille et al. 2015). The analysis of this thesis needs a large set of observations, which is why it will only focus on the North Atlantic, but more efforts are needed to study the South Atlantic as well.

In Chapter 2, I will extend the existing data set van Sebille et al. (2015) with data collected since 2015 and see how this influences the observed accumulation pattern. In Chapter 3, I will run four different models, one based on satellite and in situ data and three numerical ones, with different input scenarios and examine the differences in their final distributions. In Chapter 4, I will outline the steps needed to compare a modelled distribution to observations, and I will show the results of this statistical analysis in Chapter 5. In Chapter 6, I will discuss any shortcomings of my analysis and propose outlooks for future studies. Chapter 7 summarizes the main findings of this thesis.

Chapter 2

Observations

2.1 Earlier data

Floating plastics have been sampled since the 1970s and are arguably one of the most well-studied kinds of marine debris (van Sebille et al. 2020). Van Sebille et al. set up a global inventory of small floating debris in their eponymous paper from 2015, containing records of small, floating plastics since the 1970s. Except for the Arctic, all the world's oceans and marginal seas had been combed in 17 different studies, resulting in over 11,000 measurements of the number density or mass density of small plastic particles, trawled with neuston and manta or other plankton nets with mesh sizes between 200 µm and 505 µm. More than 75% of this data was gathered on cruises in the Western North Atlantic and in the Caribbean Sea over the course of 22 years by undergraduate students and researchers from the Sea Education Association (SEA) (Law et al. 2010; see also Section 1.3).

This data set contains the raw data as presented in each of the studies, as well as the standardized data to account for varying sampling conditions such as the sampling year and location, the meteorological conditions and the tow distance. The standardization fitting best to the raw data was determined by a general additive model, using a spherical two-dimensional spline for spatial smoothing. The best fit was achieved when standardizing to no-wind sampling conditions in the year 2014, while taking into account the spatial discontinuity between the tropical Pacific Ocean and the Caribbean Sea, where the two basins are separated by a narrow piece of land only. Wind reanalysis data is available in the ERA5 data set of hourly data on single levels from 1979 to present from the Climate Data Store by the Copernicus Climate Change Service (Hersbach et al. 2018), but only for sampling years after 1979. This explains why the earliest data is from 1979 even though plastic had been sampled in the oceans since the early 1970s.

However, any further analysis of this data should use the raw data before standardization also given in van Sebille et al. (2015), especially if new data sets are added that would need to be standardized in the same way as described above. Moreover, statistical tools like general additive models and spatial smoothing functions merely consider the spatial distribution of the data, but fail to take into account any physical processes that shape this distribution.

2.2 Trawling for additional data

The data set presented by van Sebille et al. (2015) is heavily skewed towards the western part of the North Atlantic basin, especially towards the Sargasso and Caribbean Seas. In this data set, the highest concentrations lie around 40 °N, 30 °W, but the number of samples is much lower in this area than in the well-trawled western part of the basin. In order to determine whether the

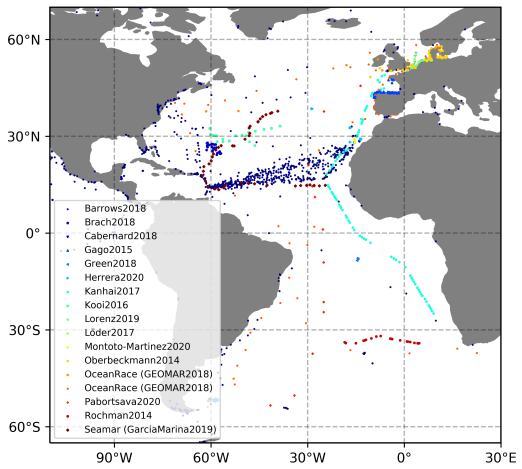


Figure 2.1: Atlantic locations of the novel plastic samples analyzed in this thesis.

location of the maximum concentration is biased by the sparse sampling outside the Western North Atlantic, the van Sebille et al. (2015) data set has to be extended with data collected in the central or Eastern North Atlantic.

I trawled through a myriad of online data bases and scanned a large number of publications about the presence of (micro-)plastics in the Atlantic Ocean to find observations that had not been included in the van Sebille et al. (2015) data set, either because they had been overlooked by the authors of the original data set or because they had been published after 2015. The search engines Google, Google Scholar and Google Dataset Search were a great help in this task. Another valuable database was LITTERBASE[1], an online portal compiling over 2,700 scientific publications with original marine litter samples, set up by the Alfred Wegener Institute for Polar and Marine Research. The locations of the new data points are mapped in Figure 2.1.

The data sets selected for this research needed to include the plastic concentrations, the coordinates of the sampling location and the sampling date. Since this thesis is concerned with the transport of small floating plastic particles, I limited my search not only to microplastics (commonly defined as particles smaller than 5 mm in diameter), but to particles slightly larger than this as well. In the end, a total of 648 new data points from 14 data sets were added to the data set by van Sebille et al. (2015) with particles sizes ranging from 32 µm to 25 mm. If plastic particles were sampled by nets, they were trawled mostly at the sea surface with bongo, neuston or manta nets with mesh sizes between 100 µm and 500 µm, although the typical mesh size for these trawls is 300 µm (Brach et al. 2018; Cabernard et al. 2018; Gago, Henry, and Galgani 2015, Herrera et al. 2020; **L“IeC −“”o″der2017enzymatic** Lorenz et al. 2019; Oberbeckmann et al. 2014; Rochman et al. 2014; García Marina et al. 2019). One data set (Kooi et al. 2016) was collected in a multi-level trawl in the water column. Other samples in the water column were obtained by pumping seawater through a mesh at the bottom of the vessel (Kanhai et al. 2017; Montoto-Martínez, Hernández-Brito, and Gelado-Caballero 2018; GEOMAR Helmholtz Centre for Ocean Research Kiel 2018; Pabortsava and Lampitt 2020). Because of contamination concerns, most data sets excluded fibres, except for Brach et al. (2018), Kanhai et al. (2017) and Montoto-Martínez, Hernández-Brito, and

[1] `https://litterbase.awi.de` (Last accessed: 14/05/2021)

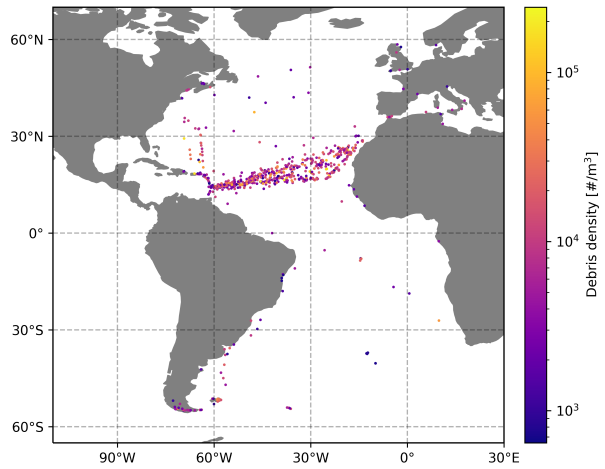


Figure 2.2: Surface concentrations as sampled in the bottle grabs.

Gelado-Caballero (2018). Although the mesh size of the trawling nets and in the pump filters generally sets a lower limit to the detectable particle size, it is not uncommon that smaller particles end up in the final sample as well, either as aggregates or by clogging the mesh (Lorenz et al. 2019).

Bottle grabs

Zooplankton sampling methods like neuston net sampling have inspired the ways that microplastics are sampled today (Green et al. 2018a). Barrows et al. (2017) note that net sampling can only detect plastic particles larger than 50 µm because the mesh size of the sampling net is limited. This is why Barrows et al. (2017) suggested a novel approach to sampling microplastics without being restricted by the mesh size: collecting sea water at the surface with 1 L bottles. Their bottle grabs were filtered over a 0.45 µm mesh, which is several orders of magnitude smaller than the sieves included in the underwater pumps with mesh size of about 25 µm (see e.g. Pabortsava and Lampitt 2020), or than the standard neuston net mesh size of 300 µm (see e.g. Brach et al. 2018).

Two microplastics data sets obtained by 1 L bottle grabs were considered in this research. Barrows, Cathey, and Petersen (2018) contributed 1393 new data points, while Green et al. (2018b) provided 16 measurements. These grab samples, however, were analyzed separately from the net and pump data. In fact, Barrows et al. (2017) note that the concentrations derived from these grab samples were three orders of magnitude larger than conventional net sampling at the same locations, largely because of the lower end of the size spectrum that can be sampled this way. Indeed, the bottle grabs from Barnes et al. (2018) and Green et al. (2018b) contained fibres and plastic particles ranging from 0.45 µm to 5 mm. As one aim of this thesis is the spatial distribution of small plastic particles, combining the concentrations from the net and pump samples with those from the bottle grabs would introduce a strong bias towards the bottle grabs. Nonetheless, the bottle grabs cover a substantial part of the previously undersampled region in the North Atlantic, as shown in Figure 2.2. Barrows et al. (2017) further indicate that the large variation within the bottle grabs may be due to quite a small sampling volume of 1 L.

2.3 Combining the data sets

Different data sets are often given in different units that need to be converted before the data sets can be combined. The data from van Sebille et al. (2015) is available in units of number of particles per km^2, converted to a total mass per km^2 using a range of values for a typical mass per particle. The plastic concentrations in the new data sets were recorded with different units, e.g. number of particles per m^3, number of particles per km^2, total mass per m^3 or total mass per km^2. The data sets recording the concentrations in total mass per m^3 or per km^2 (see Brach et al. 2018 and Gago, Henry, and Galgani 2015) also cited the concentrations in number of particles per m^3 or per km^2, hence the choice of number of particles as the first part of the unit to combine the different data sets into one. The next choice is between the concentrations given by volume or by surface area. The samples recorded in concentration by volume had often been obtained by counting the plastic particles in the water flux passing through an underwater pump, as in Kanhai et al. (2017), Montoto-Martínez, Hernández-Brito, and Gelado-Caballero (2018), GEOMAR Helmholtz Centre for Ocean Research Kiel (2018) and Pabortsava and Lampitt (2020). Since the dimensions of these pumps or the pumping times were not always recorded, it is impossible to infer the corresponding sampling area from these volumes. Therefore, the concentrations by surface area, resulting from net tows, are converted to concentrations by surface area by multiplying with an average submerged net height of 0.25 m, as suggested in Kukulka et al. (2012).

Some data sets do not give the surface concentrations, as underwater pumps or sub-surface trawls rather measure the concentration of buoyant plastic particles in the water column. While buoyant particles are usually expected to rise to the sea surface after submersion, their upwards-directed buoyancy is partially suppressed by wind-driven turbulent mixing in the so-called mixed layer (Kukulka et al. 2012). Surface heat fluxes and evaporation also homogenize the upper part of the water column, but wind and wave action remain the most important drivers (Chamecki et al. 2019). The magnitude of these effects determines the share of buoyant particles submerged in the mixed layer. Since the aim of this thesis is to assess the accumulation pattern of plastic particles at the surface, the plastic samples from the water column have to be extrapolated to the sea surface. In a first step, this is done following a parametrization by Kukulka et al. (2012) using the wind speed at the sampling point and time. None of the plastic data sets included meteorological parameters, but the wind speed is part of the ERA5 reanalysis of hourly data on single levels from 1979 to present from the Climate Data Store by the Copernicus Climate Change Service (Hersbach et al. 2018). The total wind speed U_{10} at 10 m above the sea surface is calculated from the zonal and meridional wind speeds at 10 m u_{10} and v_{10} as $U_{10} = \sqrt{u_{10}^2 + v_{10}^2}$.

The hydrodynamic model by Kukulka et al. (2012) describes the steady state of a buoyant particle in wind-driven turbulence. This leads to an exponential vertical distribution of buoyant plastic particles near the sea surface:

$$n_0 = n(z = 0) = n(z) \exp\left(-\frac{z w_b}{K_0}\right), \tag{2.1}$$

where n is the particle concentration and $z < 0$ is the sampling depth, so that $z = 0$ denotes the depth of the sea surface. The buoyant rise velocity of the plastic particles is given by $w_b = 1\,\mathrm{cm/s}$, a typical value inside the range of observed buoyant rise velocities for different sizes and shapes of particles (Kukulka et al. 2012). If breaking waves are the main factor for mixing near the surface and if the sea is fully developed with a wave age of $\beta = 35$, the vertical diffusion coefficient at the surface K_0 can be parametrized as given in Kukulka et al. (2012):

$$K_0 = 1.5 u_w \kappa H_s, \tag{2.2}$$

where $\kappa = 0.4$ is the von Kármán constant for flow velocities in turbulent fluids at the fluid boundary. $u_w = \sqrt{\tau/\rho_w}$ is the frictional velocity of the water with density ρ_w due to the wind stress τ. The wind stress can be determined from $\tau = \rho_{\text{air}} C_D U_{10}^2$ using the air density ρ_{air}, but the drag coefficient C_D has to determined empirically since it depends on the sea state. However, for average wind conditions around 12 knots or 6.1 m/s, $u_w = 0.65\,\text{m/s}$ is a reasonable assumption. Considering the wind speed at each sampling point in the water column (Hersbach et al. 2018), the average wind speed is $(6 \pm 3)\,\text{m/s}$ and therefore $u_w = 0.65\,\text{m/s}$ seems reasonable.

For a fully developed sea and in deep waters, the significant wave height is determined from the wind speed at 10 m according to Komen et al. (1996):

$$H_s = 0.243 \frac{U_{10}^2}{g}, \tag{2.3}$$

where $g = 9.81\,\text{m}^2/\text{s}^2$ is the gravitational acceleration. According to Kukulka et al. (2012), Equation (2.2) only holds for $z > -1.5H_s$, where H_s is the significant wave height. With a typical significant wave height of $(0.34 \pm 0.05)\,\text{m}$ averaged over each sampling point in the water column, the sampling depth cannot exceed 0.5 m. The sampling depths of the underwater pumps were much larger than this, though, ranging from 1.5 m (GEOMAR Helmholtz Centre for Ocean Research Kiel 2018) to 270 m (Pabortsava and Lampitt 2020). Kooi et al. (2016) also note that Kukulka's approach in Equation (2.2) often results in a potential overestimation of the inferred surface concentration, especially at low wind speeds.

This is why in his current work about an empirical Lagrangian parametrization for wind-driven mixing of buoyant particles at the ocean surface (reference updated when a DOI is available), Victor Onink investigates two different parametrizations of the vertical diffusion coefficient K_z in the mixed layer. He notes that neither of these approaches seems to have a clear advantage over the other. The first approach is based on the fact that for depths larger than $1.5H_s$, the eddy viscosity profile ν_z becomes depth-dependent. Assuming a linear relationship between ν_z and K_z, the so-called surface wave breaking (SWB) parametrization arises (Poulain-Zarcos, reference updated when a DOI is available):

$$K_z = \begin{cases} 1.5 u_w \kappa H_s + K_B & \text{if } z \geq -H_s \\ 1.5 u_w \kappa H_s^{5/2} |z|^{-3/2} + K_B & \text{if } z < -H_s. \end{cases} \tag{2.4}$$

The bulk dianeutral diffusion $K_B = 3 \times 10^{-5}\,\text{m}^2/\text{s}$ accounts for turbulent mixing below the mixed layer, outside the reach of surface waves.

The second approach considers the dependency of the vertical diffusion on the local mixed layer depth (MLD). In the so-called K-profile parametrization (KPP), Large, McWilliams, and Doney (1994) and Boufadel et al. (2020) give K_z as:

$$K_z = \frac{\kappa u_w \theta}{\phi} (|z| + z_0) \left(1 - \frac{z}{\text{MLD}} + K_B\right), \tag{2.5}$$

where $\phi = 0.9$ is used in the Monin-Obukov boundary layer theory, $\theta = 1$ parametrizes the Langmuir circulation, and z_0 is the roughness scale of turbulence, depending on the wind speed at 10 m U_{10} and on the wave age β (Zhao and Li 2019) (here, a fully developed sea is assumed and thus $\beta = 35$).

The wind speed at 10 meters used in these calculations is again obtained from the ERA5 data set (Hersbach et al. 2018). Holte et al. (2017) analyzed the records of over 2 million Argo drifters to compile a global monthly climatology of the mixed layer depths, averaged into $1° \times 1°$ bins. Since the mixed layer depth would have been an additional parameter but was not given for any

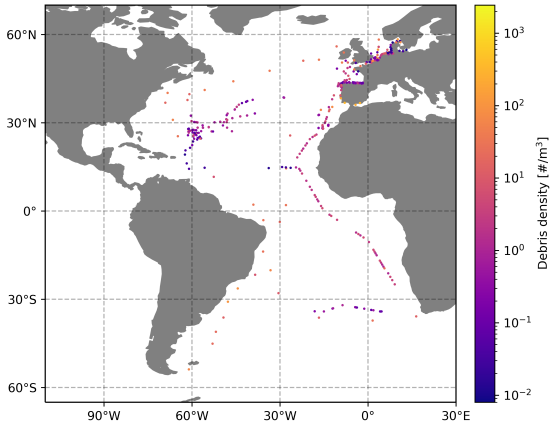


Figure 2.3: Adjusted surface concentrations as sampled in net trawls and by underwater pumps.

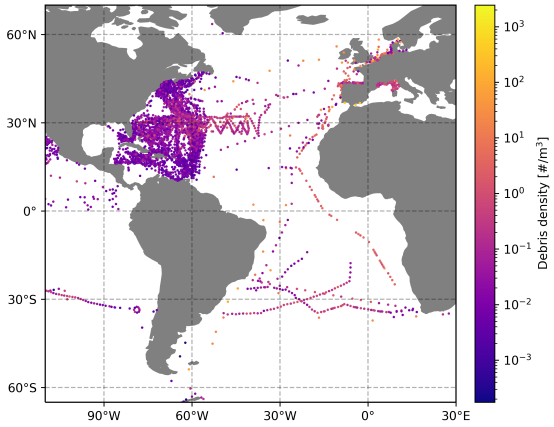


Figure 2.4: Combination of the surface concentrations compiled in van Sebille et al. (2015) and the adjusted surface concentrations as sampled in net trawls and by underwater pumps (Figure 2.3).

of the sub-surface data and would have to be obtained from a reanalysis product, I opted for the SWB parametrization. I decided to discard the Pabortsava and Lampitt (2020) data set because the data was collected well below the mixed layer at depths down to 240 m, where this mixed layer parametrization is no longer valid.

The vertical diffusion coefficient can be used to infer the share of particles collected at depth compared to the particles that would have been trawled at the surface, thus yielding a factor with which to correct the sub-surface observations taken with underwater pumps. In Lagrangian simulations that track virtual particles such as Parcels (more information in Section 3.1), subgrid-scale turbulence is often modelled by adding a stochastic term to a virtual particle's displacement (in a Markov model of order zero; "random walk"). In the case of wind-driven turbulence, a one-dimensional stochastic differential equation arises for the vertical coordinate z:

$$z(t + \mathrm{d}t) = z(t) + (w_r + \partial_z K_z)\mathrm{d}t + \sqrt{2K_z}\mathrm{d}W, \qquad z(0) = 0, \tag{2.6}$$

where $\mathrm{d}t$ is the time step, w_r is the buoyant rise velocity, K_z is the vertical diffusion coefficient and $\mathrm{d}W$ is the Wiener increment, a stochastic component distributed normally around a zero mean with a variance of $\mathrm{d}t$. Using the Parcels framework, the vertical diffusion of virtual particles initialized at the surface is simulated for a runtime of 12 hours and for a given wind speed at 10 m. Dividing the number of virtual particles at depth z by the total number of particles in the simulation gives the correction factor with which to adjust the measured concentration at depth z to infer the surface concentration.

Figure 2.3 shows the net and pump data adjusted for wind-driven mixing to give the surface concentrations. Still, compared to Figure 2.4 which shows the final net and pump data set including the van Sebille et al. (2015) data, the surface concentrations inferred from the sub-surface data sets (Kanhai et al. 2017; Montoto-Martínez, Hernández-Brito, and Gelado-Caballero 2018; GEOMAR Helmholtz Centre for Ocean Research Kiel 2018) seem rather high. Especially in the divergence zone around the equator, the surface concentrations are larger than expected. This might either be due to the finer mesh size used to filter particles collected in the pumps, compared to the nets used

in the surface trawls, or there might be an error in the modelling work of the surface concentration. I have assumed a typical buoyant rise velocity of $w_b = 1\,\mathrm{cm/s}$ for all these data sets, even though w_b depends strongly on the particle size and density (Kooi et al. 2016). Having said this, most of these data points do not seem to be part of the accumulation zone in the subtropical gyre as predicted in previous modelling work (Section 1.3), just the like the high concentrations found in the Northwestern European shelf. Disregarding these higher values, a local maximum of the observed accumulation zone lies at 40 °W, even though the lack of observations eastwards thereof leaves it unclear whether this is indeed the global maximum of the accumulation zone.

In the end, only 508 data points were added to the initial 11,103 data points to form the final net and pump data set. Many of these new data points were sampled around Europe and do not give much more information about the zonal extent of the garbage patch in the North Atlantic subtropical gyre. This is why in the following analysis that compares the observations to modelled predictions, I will investigate the influence of the samples from the coastal waters, especially because of the rather high concentrations around Europe. In the main part of the analysis, though, samples around Europe (north of 35 °N and east of 10 °W) or closer than 100 km to the continental coastlines are removed from the final data set.

Chapter 3

Simulations with Ocean Circulation Models

3.1 Lagrangian simulations with Parcels

Parcels[1] is a framework of Python classes and methods, designed to produce tracer simulations with the outputs from ocean general circulation models (OGCMs). The backronym Parcels stands for "Probably A Really Computationally Efficient Lagrangian Simulator", highlighting the Lagrangian approach used to follow active and passive tracers along their trajectories in hydrodynamic flow fields. First developed by Lange and van Sebille (2017) with additions by Delandmeter and van Sebille (2019), its current version (2.2) is a useful tool to simulate the movement of virtual, infinitesimal particles in a flow field, in analogy to small plastic particles in the oceans.

The trajectory of an infinitesimal water parcel or of a particle within the parcel follows the local advection prescribed by a Eulerian velocity field such that the particle position $\mathbf{X}(t)$ changes in time as:

$$\frac{\partial \mathbf{X}(t)}{\partial t} = \mathbf{v}(\mathbf{x}, t) \tag{3.1}$$

at the point $\mathbf{X}(t) = \mathbf{x}$. This connects the Lagrangian approach to the Eulerian form of the flow field, which determines the velocity at fixed points at space. Instead of updating the entire ocean grid as would be necessary in a Eulerian analysis, the Lagrangian approach only computes the particle's movement along its trajectory, which is especially useful when studying ocean connectivity. The trajectory of a particle is computed as:

$$\mathbf{X}(t + \Delta t) = \mathbf{X}(t) + \int_t^{t+\Delta t} \mathbf{v}(\mathbf{x}, \tau) \mathrm{d}\tau + \Delta \mathbf{X_b}(t), \tag{3.2}$$

where $\mathbf{X}(t)$ is the three-dimensional, Lagrangian positional vector of the particle at time t, Δt is the computational time step and $\mathbf{v}(\mathbf{x}, t)$ is the three-dimensional Eulerian velocity given by the OGCM at $\mathbf{x} = \mathbf{X}(\mathbf{t})$. $\Delta \mathbf{X_b}(t)$ is the change in position due to additional behaviour not included in the OGCM, which can be added to the basic Parcels code by the means of custom kernels.

This thesis only uses the horizontal components at the sea surface of the velocity fields generated by the OGCM, since the focus of this study lies on the surface accumulation pattern generated by the surface flow. The advection is computed with a Runge-Kutta scheme of fourth order and a

[1] `https://oceanparcels.org` (Last accessed: 26/05/2021)

time step of 5 min. This choice is motivated by the Courant-Friedrichs-Lewy (CFL) condition for the stability of a numerical time integration scheme:

$$\frac{v\Delta t}{\Delta x} < 1, \tag{3.3}$$

where v is the velocity magnitude, Δt is the integration time step and Δx is the spatial grid size. The largest flow velocities in the Atlantic Ocean are found in the Gulf Stream with a maximum of 9 km/h at the surface.[2] Δx ranges between 1/12° and 1/25° (see Section 3.2). The movement of the virtual particles is not restricted to the Atlantic basin and covers a large range of latitudes, along which the metric size of 1° varies with the cosine of the latitude. This is why the time step has to be chosen well below the order of one hour. The exact value of $\mathbf{v}(\mathbf{x}, \mathbf{t})$ at location x and time t is interpolated from the flow field data.

Most flow fields are given on grids that have a discontinuity around ±180 °E. To let virtual particles cross this discontinuity, the flow field is extended with a "halo" east and west of it, so that a custom kernel can move the particles from one side of the field to the other. Particles can not only get stuck at a discontinuity but also along the shore, either due to actual beaching or when they get trapped floating in the coastal zones (Onink et al. 2019). Beaching is a highly complex process that involves coastal dynamics at a much smaller scale than the basin-scale considered in this project. Modelling a realistic beaching process along the Atlantic coastline is therefore outside the scope of this work. Still, the number of particles getting stuck at the coast should be minimized in order to obtain meaningful statistics. Parcels is usually set up in a way that if particles enters a coastal cell, it stops moving, i.e. its zonal and meridional velocities are zero. A land mask can be created from all the grid cells at which the zonal and meridional velocities are zero, differentiating between coastal cells and inland cells by checking flow velocities in the grid cells around them. Similarly, the landward and seaward directions can be determined for the coastal cells. Now, each particle that enters a coastal cell is displaced seawards by a virtual "anti-beaching current" of 10 m/s, which for a time step of 5 min corresponds to a one-time seawards displacement of 3000 m back into the open ocean. This is enough to avoid that particles are trapped in most coastal grid cells, while at the same time it does not interfere much with the global circulation. Still, the anti-beaching current is often not enough to make sure that particles do not not accumulate in river mouths, as can be seen below in Figure 3.6, where particles end up in the mouth of the Amazon river. This is why at the end of each simulation, particles that are closer than 100 km to the continental coastlines are removed from the final accumulation pattern. Nevertheless, the assumption remains that particles that would in reality beach or stay in the coastal zone do not influence the shape of the basin-scale surface accumulation patterns.

In Lagrangian simulations, subgrid-scale processes are usually modelled stochastically, for example by adding a stochastic component to the particle displacement (Markov model of order zero; "random walk") or to its velocity (Markov model of order one; "random flight") (van Sebille et al. 2018). In the next section, I will describe the ocean circulation models used in my simulations; their spatial resolution lies between 1/4° and 1/12°, which is high enough to resolve mesoscale processes such as mesoscale eddies on length scales of 100 km. This is why I did not include any further subgrid-scale parametrization in my simulations.

Finally, the virtual particles are advected for 4 years, a choice which is motivated in Section 3.3, starting in the Atlantic Ocean, but using the global flow field data. Even though the main focus lies on the North Atlantic, particles are seeded within the entire Atlantic basin.

[2] `https://oceanservice.noaa.gov/facts/gulfstreamspeed.html` (Last accessed: 23/06/2021)

3.2 Ocean circulation models

GlobCurrent

The GlobCurrent[3] data set is based on altimetry data from Archiving, Validation, and Interpretation of Satellite Oceanographic (AVISO) data and CNES-CLS11 Mean Sea Surface (MSS) altimetry data to construct a map of the geostrophic ocean currents derived from differences in sea surface height. The underlying geoid is obtained from the Gravity Field and Steady-State Ocean Circulation Explorer (GOCE), an ESA satellite operated between 2009 and 2013 that mapped the Earth's gravity field. Correcting these measurements by CNES-CLS11 altimeter MSS data, a so-called geodetic mean dynamic topography (MDT) is constructed.

The estimate of the geostrophic currents is further improved by interpolating the geostrophic measurements from in situ drifter velocities using the geostrophic and Ekman balance. In general, the work by Rio, Mulet, and Picot (2014) combines altimetry, gravimetry and in situ data to assess the geostrophic and Ekman currents, which are the two main contributors to the global ocean circulation. The Ekman currents are obtained from in situ data, which also allows the inclusion of small-scale structures, increasing the resolution of the final product. Among other in situ data sets are for example the Argo floats, the vertical profiles from Coriolis Ocean database ReAnalysis (CORA) and the drogued and undrogued buoys from the Surface Velocity Drifter Program (SVP), which is the precursor of the Global Drifter Program. The drogued buoys allow for velocity estimates below the surface at 15 m depth. The Ekman currents are also modelled from wind stress data from the ERA-Interim reanalysis: Subtracting the Ekman current from the drifter velocities and correcting for wind slippage, the geostrophic current is updated again, after removing ageostrophy from inertial oscillations, Stokes drift and tides with a 3 day low-pass filter. Since the inertial signal is more prominent around the equator, the low-pass filter might have been unable to remove it in a band of $\pm 10°$ around the equator.

This thesis uses the zonal and meridional components of the global, total surface current given as the sum of the geostrophic and Ekman currents in the third version of the GlobCurrent data set at level 4. It has a spatial resolution of $1/4°$ on a standard longitude/latitude grid and a temporal resolution of 24 h. The time range selected for this work is a 4 year period from 2002-01-01 00:00 until 2005-12-31 00:00.

NEMO12

I denote by "NEMO12" a product[4] of the ORCA0083 family of NEMO (Nucleus for European Modelling of the Ocean) models (Madec et al. 2017). The NEMO models integrate the momentum, temperature and salinity equations without data assimilation. The atmospheric forcing at the sea surface is given by the DRAKKAR Forcing Set 5.2 data, including winds, temperature, humidity, downward long- and shortwave radiative fluxes, as well as precipitation and river runoff. The NEMO data is given as a 5-day average on a curvilinear, tripolar ORCA grid where the poles have been placed on land to avoid singularities in the model domain, like a rectangular grid would have them at the North Pole. Along the equator, the grid cells have a size of $1/12°$ in longitude and latitude, but the grid is even finer at higher latitudes where the Rossby radius of deformation is smaller, which is crucial for a realistic representation of mesoscale eddies across a large range of latitudes. The NEMO data is available on 75 vertical levels, but I chose only the surface level in this work, running the model for 4 years between 2000-01-03 12:00 and 2003-12-29 12:00.

[3] `https://resources.marine.copernicus.eu/?option=com_csw&view=details&product_id=MULTIOBS_GLO_PHY_REP_015_004` (Last accessed: 14/06/2021)

[4] `https://gws-access.jasmin.ac.uk/public/nemo/runs/ORCA0083-N06/means/` (Last accessed: 14/06/2021)

SMOC

The SMOC[5] data set is a product from the global, high resolution Mercator Océan monitoring and forecasting sytem PSY4V3R1, provided by the Copernicus Marine Environment Monitoring Service (CMEMS). SMOC stands for Surface and Merged Ocean Currents, referring to the addition of surface waves and tides to a basic ocean currents model. This way, it aims to "reproduce the net velocity felt by a body at the sea surface" (Drillet et al. 2019).

SMOC is based on the NEMO ocean model (version 3.1) using a tripolar ORCA grid with a spatial resolution of 1/12° and 50 vertical levels (Gasparin et al. 2018). The model is based on the free surface formulation of momentum, salinity and temperature advection terms (Lellouche et al. 2018). It is forced diurnally by atmospheric fields from the European Centre for Medium-Range Weather Forecasts-Integrated Forecast System (ECMWF-IFS), using momentum and heat turbulent surface fluxes and precipitation fluxes. Similarly as for the GlobCurrent data set, a hybrid MDT is inferred from GOCE, MSS and the Mercator GLobal Ocean ReanalYsis and Simulation (GLORYS2V3).

After initializing the system on 11 October 2006 with zero velocity profiles and temperature and salinity profiles from the EN4 monthly climatology, the model spins up until the velocity field balances the density field. External gravity waves are filtered out, but isopycnal tracer diffusion and turbulent vertical mixing are included.

The model output is improved by assimilation with satellite and in situ data according to the Système d'Assimilation Mercator (SAM), e.g. by CMEMS altimetry and sea surface temperature (SST); sea level anomalies from the Data Unification and Altimeter Combination System (DUACS) altimeter and gridded operational sea surface temperature and sea ice analysis (OSTIA SST); vertical temperature and salinity profiles among others from Argo floats, CORA data and moorings.

SMOC also takes into account the Stokes drift, a type of surface current driven by waves that can reach the same velocities as the Ekman currents. It uses the 1/10° forecasting system by the Météo France WAve Model (MFWAM), which applies the spectral dissipation of wind-generated waves. Tides are added via the Finite Element Solution (FES2014) tide model, based on the spectral configuration of the shallow water equations. SMOC has difficulties in modelling unstable current systems; in the Atlantic Ocean, this is mainly the western boundary current. Like many numerical products, its forecast becomes more prone to errors around the equator, where the geostrophic approximation does not hold anymore.

The final zonal and meridional velocities are given in hourly intervals on a standard longitude/latitude grid with a cell size of 1/12°. Even though 50 depth levels are available, I ran the model at the surface only, for a period of 4 years between 2016-04-01 00:30 and 2020-03-31 23:30.

HYCOM

I use "HYCOM" as a shorthand for the Global Ocean Forecast System (GOFS) 3.1 reanalysis product[6] generated by the HYbrid Coordinate Ocean Model (HYCOM) coupled to the Community Ice CodE (CICE) model (Metzger et al. 2017). The model solves the momentum, temperature and salinity equations, using atmospheric forcing from the NAVy Global Environmental Model (NAVGEM) 1.4. It is updated daily by the Navy Coupled Ocean Data Assimilation (NCODA), which includes several in situ data sets such as Multi-Channel Sea Surface Temperature (MCSST) data and Argo floats. The hybridicity of the vertical coordinate arises from the use of isopycnic

[5] `https://resources.marine.copernicus.eu/?option=com_csw&product_id=GLOBAL_ANALYSIS_FORECAST_PHY_001_024&view=details` (Last accessed: 14/06/2021)

[6] `https://www.hycom.org/dataserver/gofs-3pt1/reanalysis` (Last accessed: 14/06/2021)

or ρ-level coordinates in the open, stratified ocean, while the mixed layer and unstratified regions are modelled on z-levels. Shallow coastal areas make use of terrain-following σ-levels. This unique vertical parametrization is generally used to model transitions from shallower to deeper waters. The HYCOM data is given at 41 levels on a standard longitude/latitude grid with a cell size of 1/12°, with a temporal resolution of 3 h. The model is run at the surface for 4 years between 2005-01-01 12:00 and 2008-12-30 21:00.

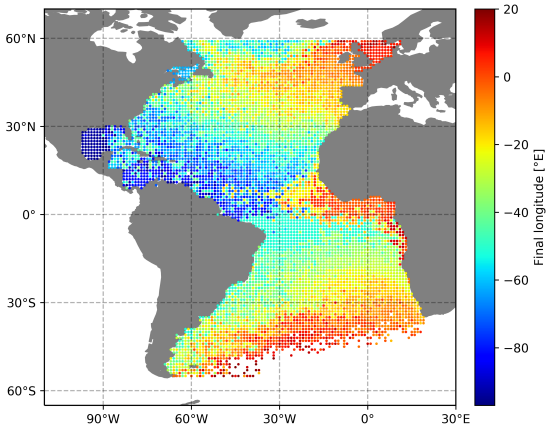
3.3 Determining the simulation runtime

In order to limit the computation time, simulations should be run for a minimum necessary time span only. For Lagrangian ocean analyses, this is often the so-called mixing time. Wichmann et al. (2019) note that due to the chaotic nature of the surface ocean circulation, any information about a virtual particle's initial position will be lost after the mixing time, which is why simulations longer than the mixing time should not depend on the input scenario anymore. Wichmann et al. (2019) calculated the mixing time in each basin using a transition matrix based on OGCM output to model the transport of virtual particles. The total variation distance of the distributions at a certain time compared to a stationary density corresponding the "garbage patch" will be small enough if the virtual particles have been advected for a period corresponding to the mixing time or longer. In the North and South Atlantic, the mixing time is then found to be 4 years, both when modelling the transport in the Atlantic Ocean only but also when global dynamics are considered.

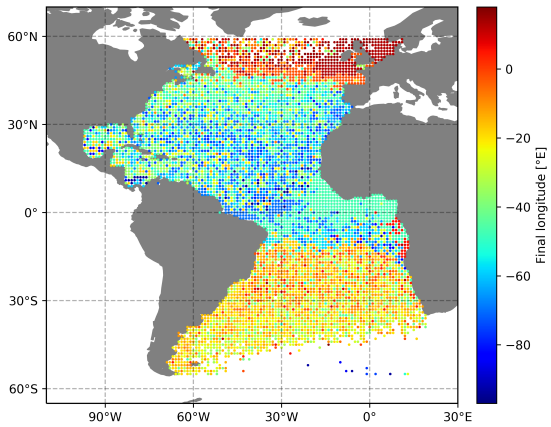
In order to verify this runtime of 4 years qualitatively, I ran a GlobCurrent simulation for runtimes between 1 and 10 years, releasing particles on a uniform $1° \times 1°$ grid within the boundaries of the Atlantic Ocean, i.e. between the longitudes of 100 °W and 20 °E and latitudes of 60 °N and 55 °S. I then produced maps of the initial position of the virtual particles coloured according to the final longitude that they had reached at the end of the simulation. Particles that left the domain of the Atlantic Ocean at the end of each simulation have been removed from the original release pattern. Supposing that longitudinal mixing is reached definitely after 10 years (see Figure 3.1d), a runtime of 4 years seems reasonable in order to ensure that the subtropical accumulation zone has reached a dynamic steady state, as shown in Figure 3.1b. Indeed, after 4 years the distribution of the final longitudes looks roughly similar to that after 10 years, even though around the equator and in the mid-latitudes, particles would still travel to more central respectively westward longitudes. All in all, differences in the final positions between runtimes of 4 years runtime and 10 years can probably be attributed to the dynamic nature of the steady state distribution. Figure 3.2 shows the same analysis to investigate the latitudinal mixing timescale, suggesting that a runtime of 4 years might be too short. Indeed, particles that circulate in equatorial waters (in turquoise in Figure 3.2b) after a runtime of 4 years would still move northward if the runtime were longer. Nevertheless, I limited the runtime to 4 years to restrict computation time and also because the SMOC data set was only available for 4 years.

3.4 Final distributions

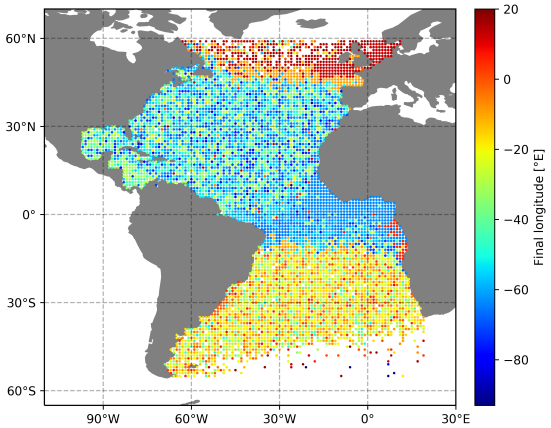
Figures 3.3 to 3.6 present the final virtual particle distribution generated by each of the four models after a runtime of 4 years. The left column in each figure shows the final snapshot using two different one-time release scenarios: the uniform seeding of virtual particles arranged in a $1° \times 1°$ grid (like in Maximenko, Hafner, and Niiler 2012), and the release of virtual particles along the coastline in every coastal grid cell. For the latter, the coastline is extracted by using the land mask of each



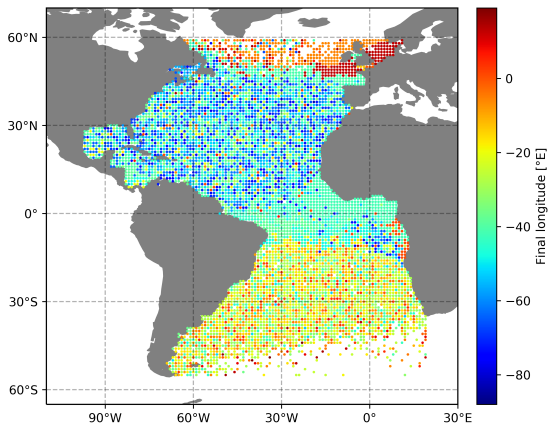
(a) After 1 year.



(b) After 4 years.

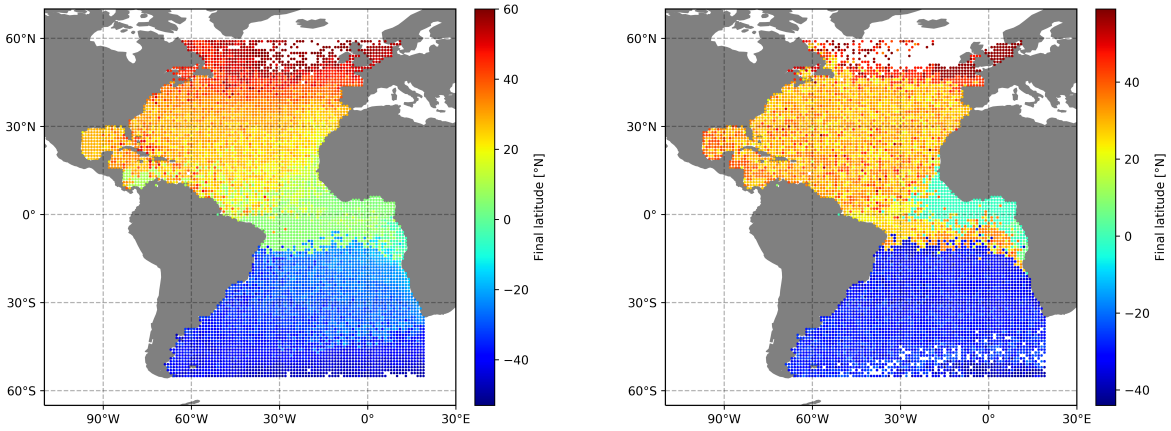


(c) After 7 years.



(d) After 10 years.

Figure 3.1: Maps of the initial positions of virtual particles released on a uniform $1^\circ \times 1^\circ$ grid, coloured according to their final longitude at the end of a simulation with the GlobCurrent flow field, for different runtimes.



(a) After 1 year.

(b) After 4 years.

(c) After 7 years.

(d) After 10 years.

Figure 3.2: Maps of the initial positions of virtual particles released on a uniform $1^\circ \times 1^\circ$ grid, coloured according to their final latitude at the end of a simulation with the GlobCurrent flow field, for different runtimes.

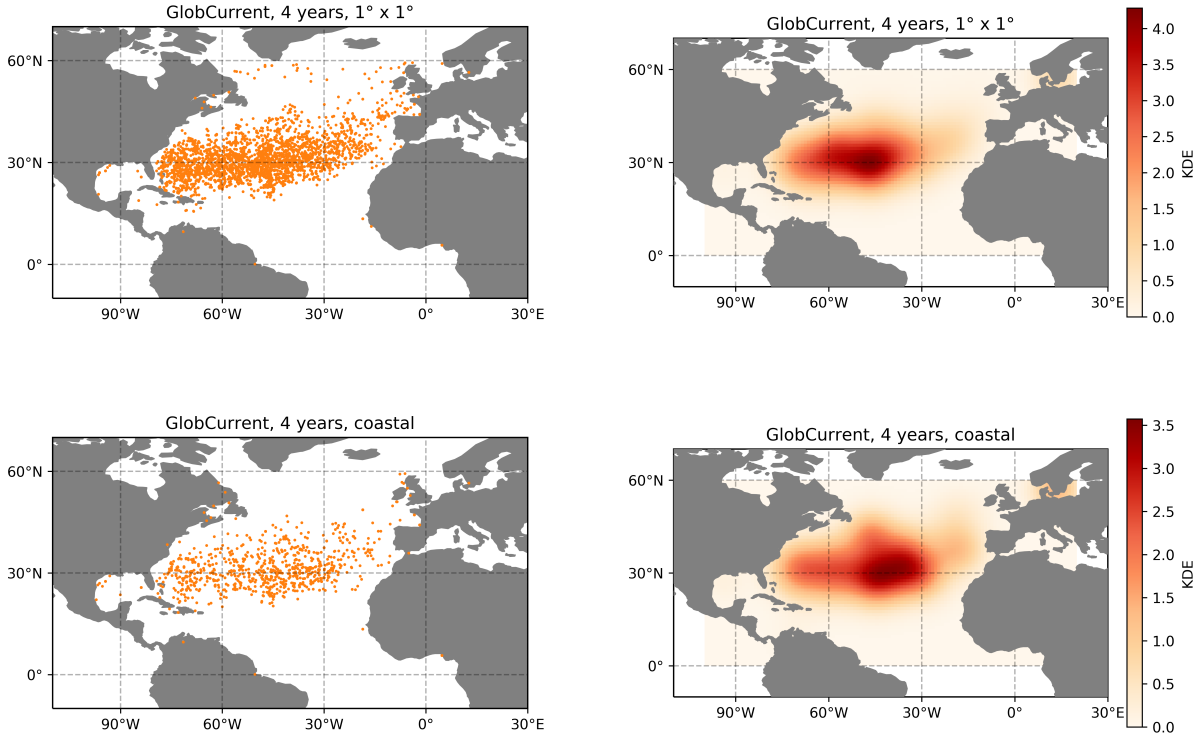


Figure 3.3: Maps of the virtual particle distribution *(left column)* and the associated KDE *(right column)* generated by the GlobCurrent model after a uniform release scenario *(top row)* and after a coastal release scenario *(bottom row)*. The KDE has been calculated after removing virtual particles that ended up closer than 100 km to the coastline at the end of the simulation.

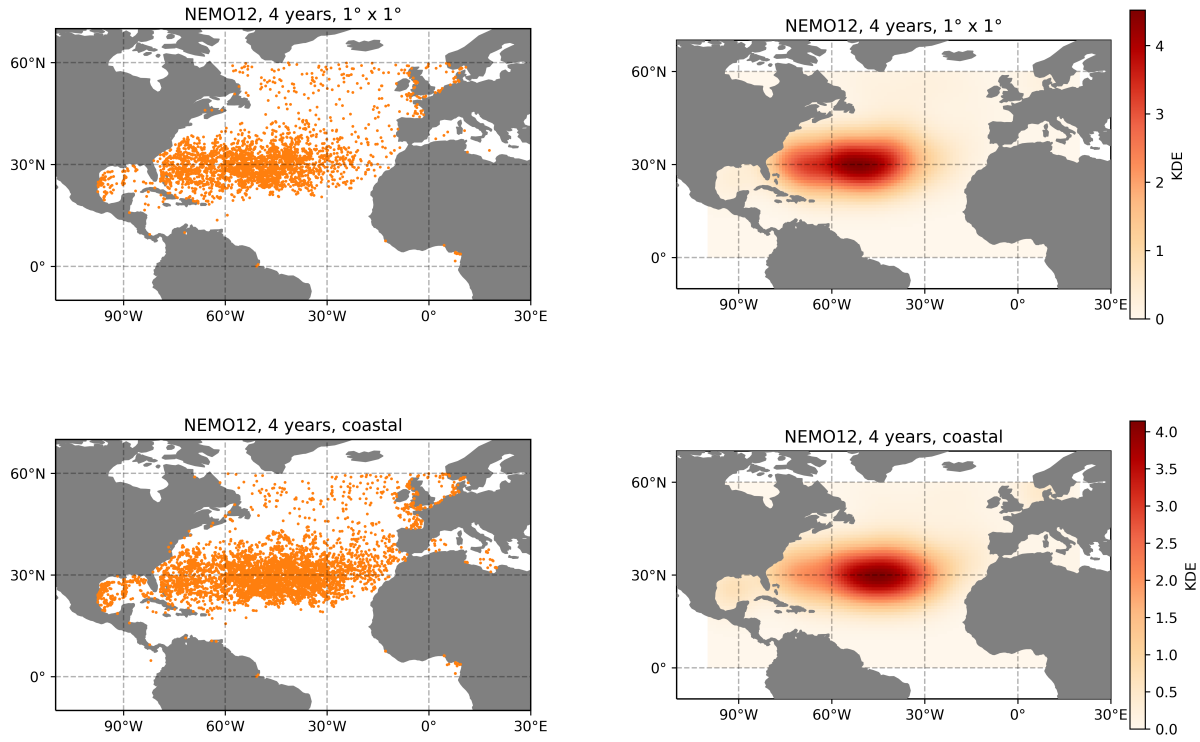


Figure 3.4: Maps of the virtual particle distribution *(left column)* and the associated KDE *(right column)* generated by the NEMO12 model after a uniform release scenario *(top row)* and after a coastal release scenario *(bottom row)*. The KDE has been calculated after removing virtual particles that ended up closer than 100 km to the coastline at the end of the simulation.

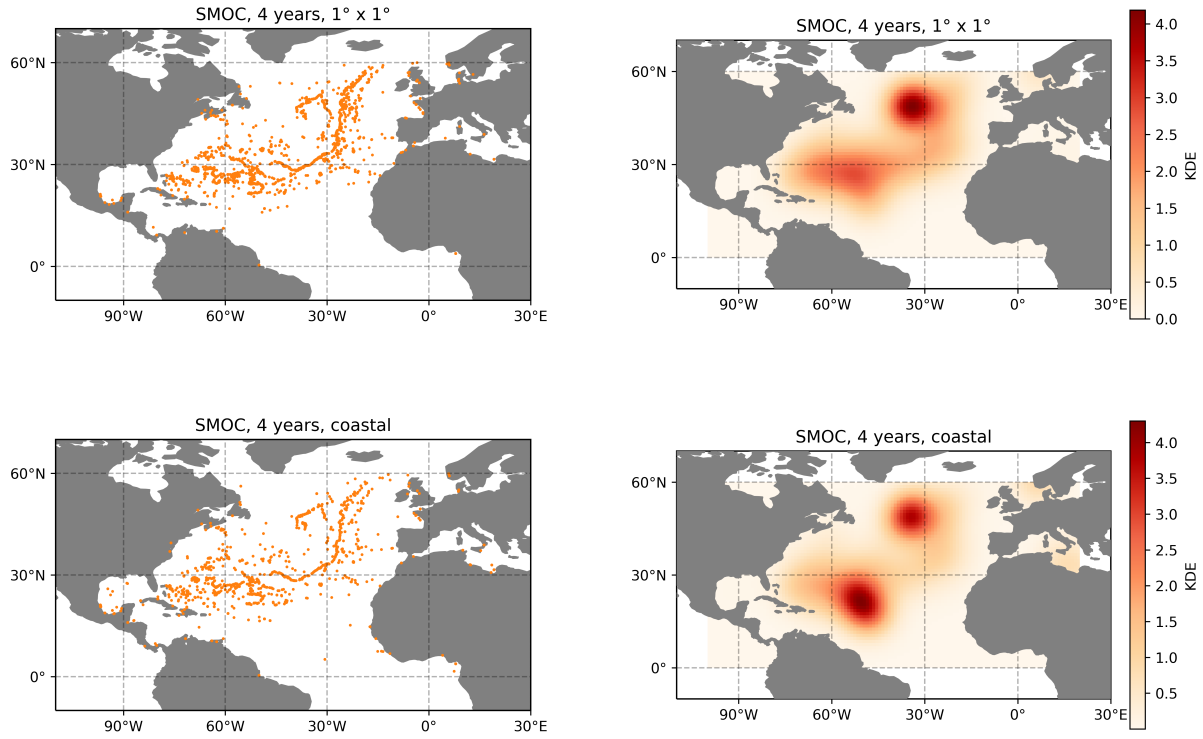


Figure 3.5: Maps of the virtual particle distribution *(left column)* and the associated KDE *(right column)* generated by the SMOC model after a uniform release scenario *(top row)* and after a coastal release scenario *(bottom row)*. The KDE has been calculated after removing virtual particles that ended up closer than 100 km to the coastline at the end of the simulation.

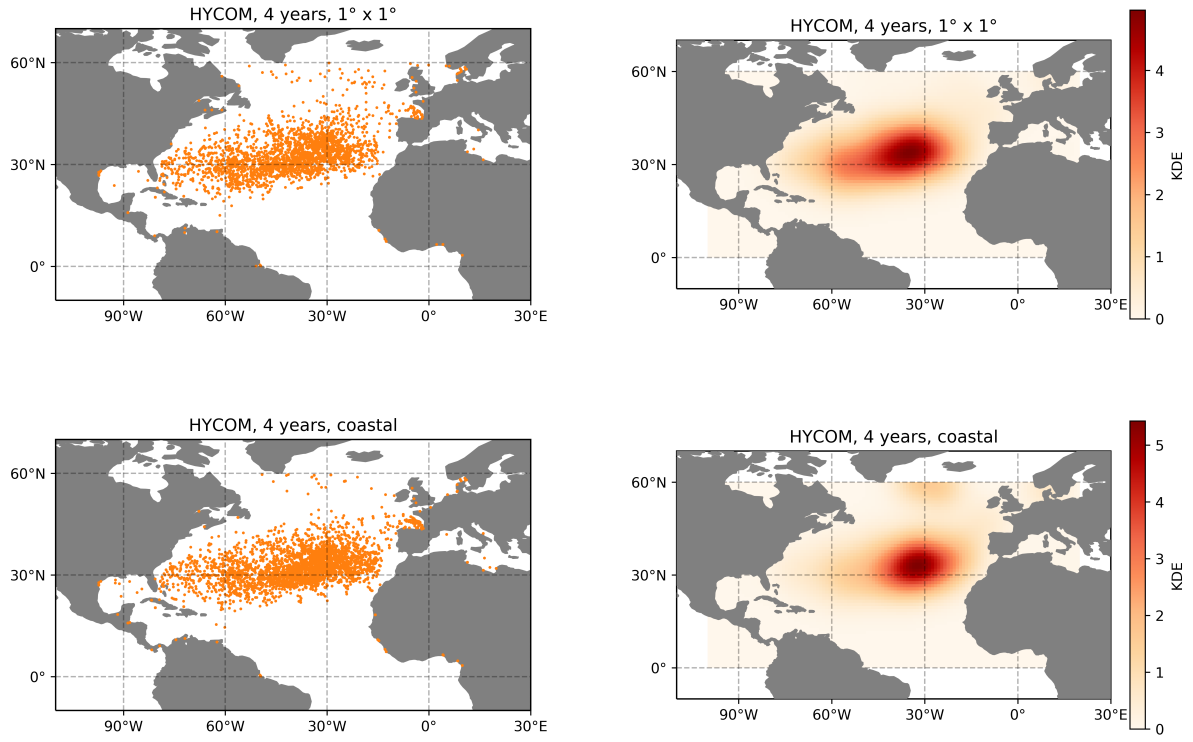


Figure 3.6: Maps of the virtual particle distribution *(left column)* and the associated KDE *(right column)* generated by the HYCOM model after a uniform release scenario *(top row)* and after a coastal release scenario *(bottom row)*. The KDE has been calculated after removing virtual particles that ended up closer than 100 km to the coastline at the end of the simulation.

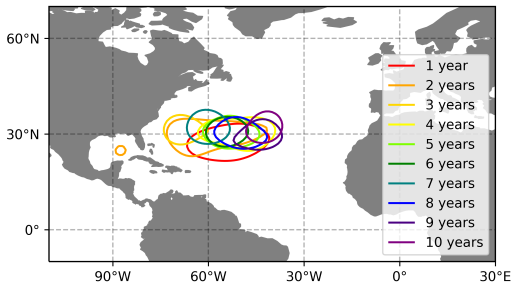


Figure 3.7: Dynamic change in the maximum of the surface accumulation zone over runtimes ranging from 1 to 10 years generated by the GlobCurrent model after a uniform release. The maximum of the accumulation zone is defined as the area where KDE values are at least 3/4 of the maximum KDE value.

model and checking which ocean grid cells lie immediately adjacent to a land cell. Table 3.1 gives the number of virtual particles seeded in each release scenario for each model.

Neither of these source functions is particularly realistic. A uniform, one-time release is easy to compute but does not happen in reality, even though some plastic is released into the oceans as fishing debris from vessels at sea. The coastal release starts at least with a more realistic spatial distribution, although the amount of plastic put into the ocean is hardly evenly distributed along the coast, but rather depends on a country's economic activity (Jambeck et al. 2015). Within one country, plastic is mostly entering the ocean in highly populated coastal areas and at beaches, but also via waterways (Lebreton, Greer, and Borrero 2012). The comparison of these two simplified release scenarios will, however, give a first indication as to whether the accumulation pattern after a runtime 4 years depends on the source function of a simulation, given that particles seeded at the coastline need to travel longer to reach the center of the basin.

The right column in Figures 3.3 to 3.6 shows the kernel density estimate (KDE) corresponding to the virtual particle distribution in the left column. The surface accumulation zone can be defined as the region of a higher virtual particle density, but rather than try to infer the shape and location of the accumulation zone from the discrete virtual particle distribution, it is useful to approximate the underlying continuous distribution, for instance by a kernel density estimation. This approach will be explained further in Chapter 4, but the KDE is shown here already, as its interpretation is rather straightforward: The general shape of the maximum of the accumulation zone is given by the highest values of the KDE, e.g. where the KDE values are at least 3/4 of the maximum KDE value, indicated in red in the heat map. The KDE has been determined after removing virtual particles that ended up closer than 100 km to the coastline at the end of the simulation.

The accumulation zones vary in shape and location for each of the models and release pattern. For the GlobCurrent model, the accumulation zone is centered around 30 °N and is confined to

Table 3.1: Number of virtual particles released on a uniform $1^\circ \times 1^\circ$ grid or along the coastline in every coastal grid cell, for simulations with the four different OGCMs.

	GlobCurrent	NEMO12	SMOC	HYCOM
Uniform	7,629	7,629	7,629	7,629
Coastal	2,416	10,628	9,031	11,052

the western and central part of the North Atlantic, stretching from the western boundary up to 30 °W (see Figure 3.3). The maximum of the accumulation zone lies around 45 °W after a uniform release and is shifted slightly eastwards after a coastal release. Since the exact location of the maximum depends on the densest part of the virtual particle distribution, it shifts in time due to the dynamic behaviour of the accumulation zone, as shown in Figure 3.7. After a runtime of at least 4 years, the maximum seems to remain more or less constant, which validates the runtime choice of 4 years. Still, after runtimes of 7 years (10 years), the center shifts slightly northwards and westwards (eastwards), which is consistent with the finding that the latitudinal (and longitudinal) mixing after 4 years did not correspond fully to that after 7 and 10 years (cf. Figures 3.2, 3.1).

The maximum of the surface accumulation zone lies around the same latitude and longitude as the highest concentrations observed around 30 °N and 40 °W, even though it is unclear whether this is the real maximum of the actual accumulation zone as there is still a lack of observations eastward of the highest concentrations.

For the NEMO12 model, the accumulation zone arises in the western and central part of the basin around 30 °N, with a maximum at 50 °W for the uniform release or at 45 °W for the coastal release. This is similar to the findings of Maximenko, Hafner, and Niiler (2012), van Sebille, England, and Froyland (2012) and Lebreton, Greer, and Borrero (2012) who locate the maximum between 55 and 60 °W. It is interesting to see that virtual particles tend to accumulate in the Gulfs of Guinea and Mexico. It could be an issue with the implementation of the "anti-beaching current" that would displace particles in a coastal cell seawards again, due to the concave shape of these Gulfs, so that displacing a coastal particle seawards would still trap it in another coastal cell. It should be noted, though, that the GlobCurrent model with a grid size three times that of NEMO12 does not show this behaviour, as the Gulfs of Guinea and Mexico are practically void of virtual particles. The same "anti-beaching current" of 10 m/s should be able to move particles outside coastal grid cells in the finer resolution NEMO12 model more so than in the GlobCurrent model. The trapping of particles in these bays seems to be a consequence of the hydrodynamic field, regardless of the coastal dynamics. This justifies the removal of particles closer than 100 km to the coastline at the end of each simulation to calculate the corresponding KDE.

The SMOC model shows a peculiar accumulation pattern (Figure 3.5). Around 60 °W, there seems to be a typical accumulation zone around 30 °N. Eastward of this, the virtual particles seem to be attracted to a long ridge that stretches northwards to Iceland, coinciding with the Mid-Atlantic Ridge. It is interesting to note that SMOC is based on the NEMO model, but the latter does not exhibit these convergence lanes (see Figure 3.4). In fact, SMOC is the only model studied in this thesis that includes the effect of tides, albeit only in the shallow water approximation which should not hold for the Mid-Atlantic Ridge at depths of 1500 to 3000 m. Still, internal tides reflected at the bottom topography are known to generate surface convergence zones parallel to the depth contours (van Sebille et al. 2020). In the appendix to this thesis, Figure A.2 shows the result of the same simulation but *without* tides or surface waves, so only the large-scale circulation from the PSY4V3R1 product. Indeed, the convergence at the Mid-Atlantic Ridge disappears, but similar braided structures or "convergence lanes" persist in the eastern part of the basin. As shown in the animation in Figure A.1, these structures are not seasonal and only appear right before the end of the total runtime of 4 years, after 197 weeks. The quality of the KDE suffers visibly from the presence of the convergence lanes, as these lanes are not captured in the overall shape of the KDE (see Figure 3.5).

Finally, Figure 3.6 shows the accumulation zone generated by the HYCOM model, with its maximum slightly north of 30 °N and around 30 °W, so more eastward than the for other models. After a coastal release, the accumulation zone is narrower than for the uniform release. In the latter case, the accumulation zone stretches westwards up to 60 °W. Lebreton, Greer, and Borrero (2012)

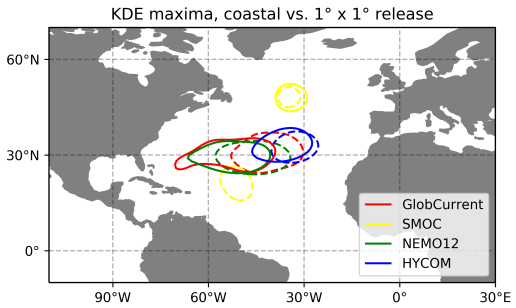


Figure 3.8: Change in the maximum of the surface accumulation zone for the GlobCurrent model after a coastal *(dashed line)* and after a uniform release *(solid line)*. The maximum of the accumulation zone is defined as the area where KDE values are at least 3/4 of the maximum KDE value.

also used HYCOM data, but found the center of the accumulation zone to lie more westwards.

Figure 3.8 summarizes the findings from above. It shows the center of the surface accumulation zone for each model and release scenario after a runtime of 4 years, suggesting that the coastal release scenario shifts the center of the accumulation zone to eastwards for GlobCurrent, NEMO12 and HYCOM. For the SMOC model, the peculiar locations for center of the accumulation zone arises mainly from the fact that the KDE fails to reproduce the pattern given by the virtual particle distribution.

Figure 3.9 shows the difference in KDE after a coastal and uniform release scenario for each model. As already seen in Figure 3.8, virtual particles accumulate preferably in the eastern part of the North Atlantic, but also in the coastal waters around Europe when a coastal release scenario is used, while a uniform release scenario increases the virtual particle density in the Western North Atlantic as well as in the Gulfs of Mexico and Guinea.

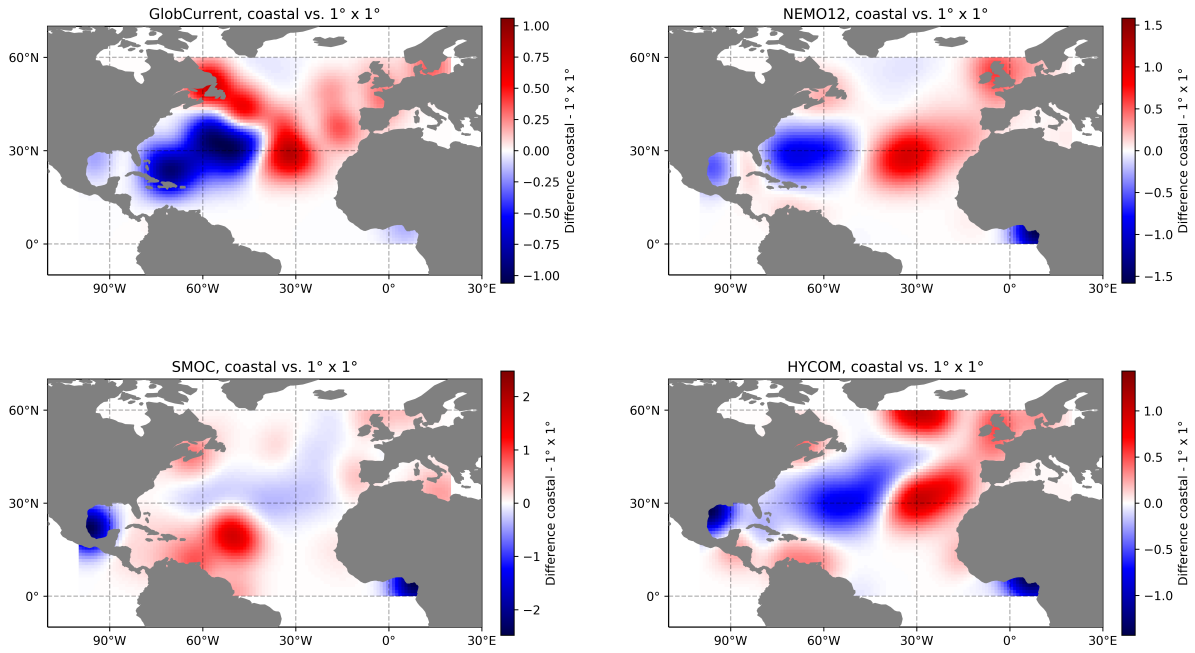


Figure 3.9: Difference between the KDE values after a coastal and uniform release scenario for each model.

Chapter 4

Combining Models and Observations

Both the virtual particle simulations and the concentrations from in situ plastic samples show the accumulation pattern of small, floating debris in the oceans. The observed and modelled patterns are not immediately comparable, though. It is already difficult enough to compare the observations amongst themselves, as I have shown in Chapter 2, for instance because the sampling methods differ from study to study. The sampling year also plays an important role in determining the plastic concentration at a given location, as more and more plastic is being released into the oceans every year (Wilcox, Hardesty, and Law 2019; Jambeck et al. 2015). With the sampling years ranging from 1979 to 2019, the particle detection methods, as well as the sample treatment procedures have improved over time as well (Jakob Strand, private communication, 14 December 2020). To account for the sampling year in terms of change in the large-scale circulation patterns, the simulations could have been run for a certain period of time after which the final positions of the virtual particles are compared to the observations from the same year. However, this would require large enough a data set sampled in the region where concentrations are expected to be highest over the course of the mixing time, which does not exist yet. In fact, the spatially most extensive data set in this region spans 22 years (Law et al. 2010). Finally, some meteorological phenomena that affect sampling conditions such as winter storms occur seasonally, influencing at which times sampling efforts are undertaken.

Virtual particle simulations have the advantage that the particles can be compared amongst themselves, since they are "sampled" at the same time and because they all have the same infinitesimal shape, size and density, contrary to their real counterparts. The virtual particles are distributed over the entire basin, independent of practical sampling issues such as rough sea states or remote sampling locations in the middle of the basin, and without any concerns for a spatial bias or blind spots in the virtual sampling. A virtual particle distribution can best be compared to the observed distribution if their spatial coordinates coincide. One way to achieve this is by binning either distribution into a two-dimensional histogram of a certain bin size. As can be seen in Figure 3.3, the order of magnitude of the number of virtual particles seems to small, as there are not enough virtual particles to cover the entire basin. This yields empty bins in the resulting histogram, as can be seen in the left panel of Figure 4.1, even though the probability of sampling a particle in these empty bins should not be zero. This issue arising from the discrete nature of the histogram can be circumvented by approximating the underlying continuous distribution, e.g. by the bivariate kernel density estimation (KDE).

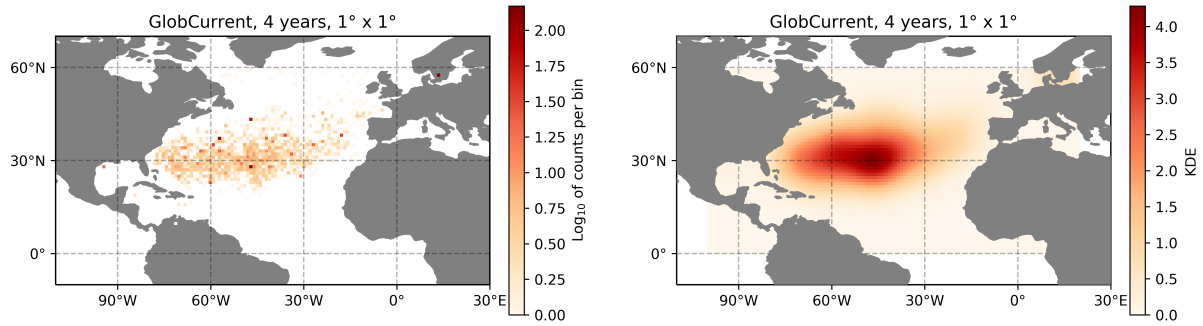


Figure 4.1: *Left:* Two-dimensional histogram of the virtual particle distribution generated by the GlobCurrent model when after a uniform release scenario (as shown in Figure 3.3) with a bin size of $1^\circ \times 1^\circ$. *Right:* KDE of the same virtual particle distribution, projected onto a $1^\circ \times 1^\circ$ grid. Both maps have been created after removing virtual particles that ended up closer than 100 km to the coastline at the end of the simulation.

The multivariate kernel density estimate is defined as (García-Portugués 2021):

$$\hat{f}_{\mathbf{H}}(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} K_{\mathbf{H}}(\mathbf{x} - \mathbf{x}_i), \tag{4.1}$$

where $\mathbf{x}$ is the location in the domain at which to evaluate the KDE. The multidimensional vectors $\mathbf{x}_i, i = 1, \ldots, n$ denote the final longitude and final latitude coordinates of the n virtual particles that remain inside the North Atlantic basin. The kernel function is given by K, a symmetric multivariate density. According to García-Portugués (2021) and Turlach (1993), a typical choice is the Gaussian kernel:

$$K(\mathbf{t}) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\mathbf{t}^2\right). \tag{4.2}$$

Here, $\mathbf{H}$ is the so-called bandwidth matrix. For two-dimensional data, it is a 2×2 symmetric positive-definite matrix that smoothes the final KDE. Finally, the scaled kernel is given by:

$$K_{\mathbf{H}}(\mathbf{x}) = |\mathbf{H}|^{-1/2} K(\mathbf{H}^{-1/2}\mathbf{x}). \tag{4.3}$$

The bandwidth matrix needs to be chosen such as to minimize the root mean square error (RMSE) between the KDE and the unknown underlying continuous distribution. It corresponds to a generalization of the definition of a scalar bandwidth h in the case of a univariate kernel density estimate (García-Portugués 2021):

$$\hat{f}_h(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right), \tag{4.4}$$

where the $x_i, i = 1, \ldots, n$ are the one-dimensional coordinates of a discrete distribution and the kernel function K is a symmetric univariate density. For multidimensional data, the anisotropy of the bandwidth is captured in a matrix. Trying out a range of bandwidths using the `sklearn.neighbors.KernelDensity` module from Python's `scikit-learn` package (Pedregosa et al. 2011), a quick visual evaluation of an exemplary virtual particle distribution (see the right panel in Figure 4.1)

shows that the same bandwidth can be chosen for both dimensions, as the KDE does not seem to be over- or underfitted in either dimension.

Being able to choose the same bandwidth for both dimensions reduces the complexity of the bandwidth determination. Algorithms such as `sklearn.model_selection.GridSearchCV` determine the bandwidth that corresponds best to the underlying data, but looping over a range of bandwidths between 0 and 1° worked well enough in this case. Still, it is important to note that the final shape of the KDE highly depends on the bandwidth choice, which is the main source of uncertainty for this statistical tool.

The KDE is evaluated on a standard grid between 100 °W and 20 °E, and between the equator and 60 °N. The grid cells are of size $1^\circ \times 1^\circ$, and therefore the observations need to be binned in the same way as well. The weighted mean concentration of all the observations in each grid cell is computed by dividing the arithmetic mean by the number of observations binned in the cell (i, j), n:

$$\bar{c}(i,j) = \frac{1}{n^2} \sum_{k=1}^{n} c_k(i,j). \tag{4.5}$$

Like in van Sebille et al. (2015), this is done to highlight regions with sparse sampling coverage. The weighted mean concentration in each $1^\circ \times 1^\circ$ bin containing observations can be compared to the corresponding value of the KDE in the same bin. A scatter plot is a helpful tool to visualize the relationship between the weighted mean observed concentrations and the corresponding KDE values. The coefficient of determination R^2 quantifies the correspondence between a model and the observations, as it gives the share of the variance in the dependent variable y that is predictable from the independent variable x. Here, the KDE values are chosen to be the dependent variable, because models are often assimilated to observations.

To calculate the coefficient of determination, the variance of the dependent variable is described by the total sum of squares:

$$\mathrm{SS}_{\mathrm{tot}} = \sum_{i=1}^{n} (y_i - \bar{y})^2, \tag{4.6}$$

where n is the number of data points and $\bar{y}$ is the arithmetic mean of the y values.

The sum of squares of residuals determines the difference of the dependent variable to the corresponding value of a fit function f:

$$\mathrm{SS}_{\mathrm{res}} = \sum_{i=1}^{n} \left(y_i - f(x_i)\right)^2. \tag{4.7}$$

Finally, the coefficient of determination is given as:

$$R^2 = 1 - \frac{\mathrm{SS}_{\mathrm{res}}}{\mathrm{SS}_{\mathrm{tot}}}. \tag{4.8}$$

R^2 will be 1 if the simulated values exactly coincide with the observed values. It can further be shown that if the fit function f is determined from a univariate linear regression, the coefficient of determination R^2 is equivalent to the square of Pearson's correlation coefficient r (Verbeek 2005), which is given as:

$$r^2 = \frac{\left(\mathrm{Cov}(x,y)\right)^2}{\sigma_x^2 \sigma_y^2} = \frac{\left(\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})\right)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}, \tag{4.9}$$

where $\mathrm{Cov}(x, y)$ is the covariance between x and y, and σ_x, σ_y are the standard deviations of x and y.

If surface accumulation zone generated by a model simulation agrees with the pattern determined from the observations, there should be a linear relationship between the KDE values and the observed concentrations. As mentioned in Chapter 2, the order of magnitude of the observed concentrations varies a lot, even within the results from the same sampling method. The relationship between KDE values and observed concentrations is therefore visualized best in a log-log plot. Here, a 1-to-1 linear correspondence, $f(x) = x$, is represented by a straight line of slope 1. A linear regression minimizing the sum of squared residuals for a linear relationship would put too large an emphasis on regions with higher modelled concentrations, which can be orders of magnitudes off from the observed concentration. Instead, a linear fit to the base 10 logarithm of the modelled and observed values is better suited to estimate the correspondence between models and observations. The best fit is determined by a least squares regression using Python's `scipy.stats.linregress`, which puts out Pearson's correlation coefficient r. I will interprete r as the coefficient of determination $R^2 = r^2$, because R^2 is easier to understand as the share of the variance in the dependent variable predicted from the independent variable.

To illustrate this comparison, Figure 4.2 shows the regression between the KDE values of the virtual particle distribution generated by the GlobCurrent model after a uniform release scenario and the concentrations sampled from bottle grabs, from the net/pump samples, and from the bottle grab, net and pump samples combined. The third panel shows the different orders of magnitude covered by the bottle grabs and net/pump data, resulting in a coefficient of determination of $R^2 = 0.1382$. Taking only the bottle grabs or the net/pump data to compare to the modelled concentrations, as shown in the top panels, the coefficients of determination are smaller than for the combined observations; they are $R^2 = 0.0135$ and $R^2 = 0.0125$ respectively. This is an example of a statistical fallacy called Simpson's paradox (Sprenger and Weinberger 2021): The statistical conclusion arising from combining two distinct data sets is fundamentally different than that from considering the data sets separately. In this case, the key to the paradox lies in the nature of the sampling methods, because the bottle grabs measured concentrations about three orders of magnitude larger than the sampled in the net trawls or with the underwater pumps. This is why the following analysis will consider the results for bottle grabs and net/pump data separately.

In the scatter plots in Figure 4.2, the modelled concentrations would correspond perfectly to the observed concentrations if the data points were to lie on the black line that simply states a 1-to-1 linear correspondence, $f(x) = x$. Since the modelled values are technically unitless as they arise from the KDE computation, they can be scaled to cover the same range as the observations. This is done by shifting the y-axis until the mean of the log-transformed modelled concentrations corresponds to the mean of the log-transformed observed concentrations. This allows to estimate the range of the predicted concentrations. Shifting data along a log-axis is equivalent to multiplying the data with a scalar, which does not affect the calculation of R^2 in Equation (4.8) or r^2 in Equation (4.9). It is essential to assign units to the modelled values, as these models are used to interpolate observations in order to assess maps of the surface plastic concentrations.

The best linear fit to the log-transformed values is indicated by a red line in Figure 4.2, as the coefficient of determination is calculated using this fit. Essentially, R^2 indicates the goodness of this linear fit, which is always different from the identity function. Any slope different from 1 indicates a power law relationship between modelled and observed values, so the closer the slope of the red line is to 1, the better the model fits to the observations.

The uncertainty on R^2 is determined by bootstrapping the data set. First, 1,001 bootstrap samples are built by sampling the original data set with replacement. Then, the coefficient of determination is calculated for each of these bootstrap samples. Assuming these R^2 are normally

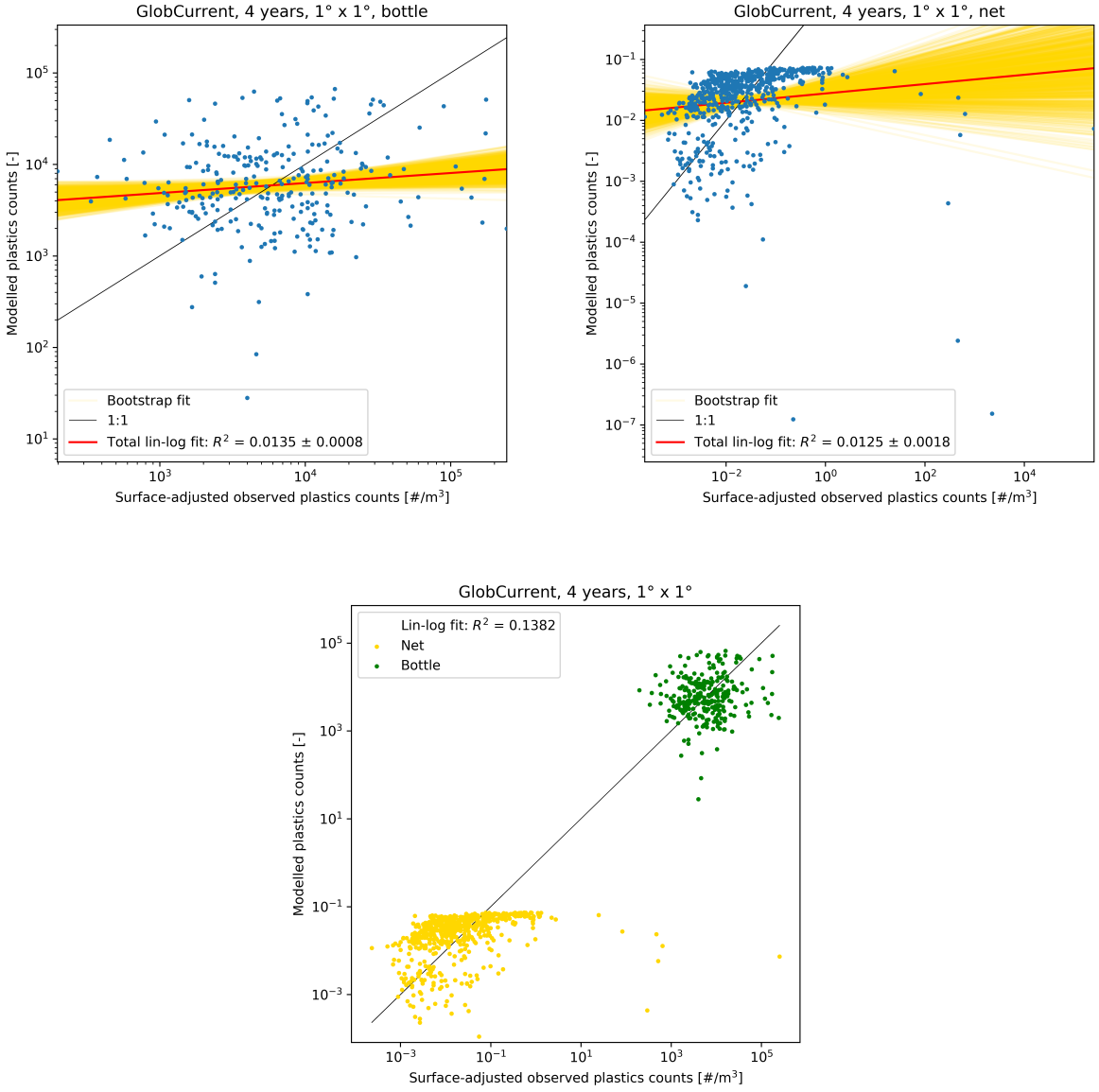


Figure 4.2: Illustration of Simpson's paradox. *Top left:* Regression between the KDE values of the virtual particle distribution generated by the GlobCurrent model after a uniform release scenario and the concentrations sampled from bottle grabs. *Top right:* Regression between the KDE values of the same simulation and the net/pump concentrations. *Bottom:* Regression between the KDE values of the same simulation and the concentrations from bottle grab, net and pump samples.

distributed, their standard deviation gives a measure for the uncertainty on R^2 of the original data set. The linear fit corresponding to each bootstrap sample is indicated by a golden line in Figure 4.2.

There are more elaborate statistical alternatives to compare two discrete, two-dimensional distributions than the method used in this thesis. The two-sample, two-dimensional Kolmogorov-Smirnov test measures the maximum distance between their empirical distribution functions over their domains in both dimensions (Peacock 1983). A similar statistic is the Cramér-von Mises criterion, defined as the integral of the difference between two (empirical) distribution functions. A slightly different formulation of this criterion emphasizes the tail of the distribution, and is known as the Anderson-Darling test. However, Babu and Feigelson (2006; 2019) note that all three tests may yield different results depending on how their statistics are calculated across all dimensions as they depend on the empirical distribution functions of the distributions that need to be compared. Indeed, an ill-defined statistic cannot help to determine the likeliness of two distributions. A statistical method that does not depend on the empirical distribution functions is the Wasserstein (Kantorovich) distance, also called earth-mover's distance, which describes the distance between two distributions as the minimal effort needed to convert one into the other, the same way one pile of earth would be shaped into another. The Kullback-Leibler divergence or relative entropy gives a relative comparison of two distributions. Both measures can be generalized to two and more dimensions. These methods are useful if there is no obvious bias in the two distribution that ought to be compared; this is not necessarily given with the observational data set, which is still biased towards the western part of the North Atlantic basin. These methods can only be used once the bias in the observations has been overcome.

Chapter 5

Results

5.1 Different release scenarios

As mentioned in Section 1.3, Chenillat et al. (2021) note that the release scenario of virtual particles may influence the final accumulation pattern generated by a simulation. Figure 3.8 also shows that the maximum of the surface accumulation zone shifts eastwards after a coastal release as opposed to seeding virtual particles on a uniform grid. In this chapter, I will examine the influence of different release scenarios on the correlation between modelled and observed concentrations.

Figures 5.1 to 5.4 show the correlation between the KDE values of the virtual particle distributions generated by the each model and the observed concentration pattern, after releasing the virtual particles uniformly across the basin or along the coast, compared to the net/pump concentrations and to the concentrations sampled from bottle grabs. The scatter plots display the observed concentrations on the x-axis, binned into 1° cells, while the y-axis shows the corresponding modelled value in the same grid cell. The log-log plot reveals the wide range of magnitudes of the observed concentrations.

As I explained in Chapter 4, the y-axis has been shifted to display the modelled concentrations in the same units as the observed concentrations. The black line indicates a 1-to-1 correspondence, while the red line shows the best linear fit to the log-transformed data. The golden lines arise from bootstrapping the data set in order to estimate the uncertainty on the coefficient of determination R^2.

The scatter plots of the GlobCurrent model are shown in Figure 5.1. For the bottle data, there does not appear to be any visual correlation between the modelled and the observed concentrations. The coefficient of determination is rather small as well for both the uniform and the coastal release, with $R^2 = 0.0135 \pm 0.0008$ and $R^2 = 0.0326 \pm 0.0011$. Considering the map of the concentrations sampled with the bottle grabs in Figure 2.2, there also does not seem to be any pattern or clear maximum in the spatial distribution. Indeed, Barrows et al. (2017) note that the large variability among their samples might arise because the sampling volume of 1 L is quite small, but it also good to note that the sampling area was below the accumulation zone of 30 °N.

For the net and pump data, the scatter plot displays a similar hook that looks like the Γ-shape in Figure 1.2, when van Sebille et al. (2015) performed a similar regression analysis on their original observations with the van Sebille, England, and Froyland (2012) model. Since these observations also form the largest part of the net and pump used in this thesis, it is no surprise that the same shape arises. The hook shows a cut-off of the modelled concentrations at around 0.1 particles per m^3, indicating that the modelled values are too low in locations where observed values are highest. The modelled maximum seems to lie at a location that has not been thoroughly sampled

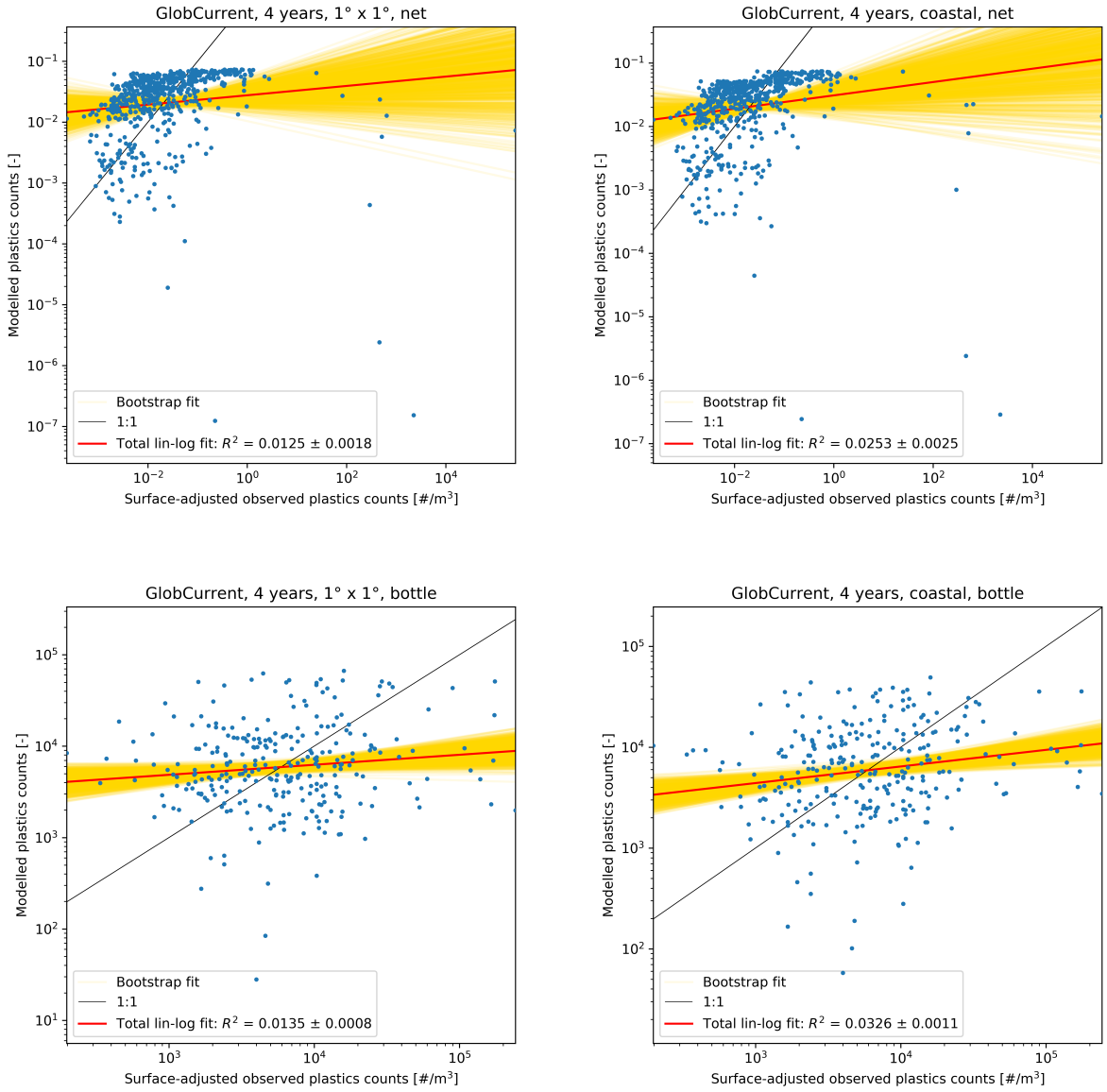


Figure 5.1: Regression between the KDE values of the virtual particle distribution generated by the GlobCurrent model after a uniform release scenario *(left column)* and after a coastal release scenario *(right column)*, and the net/pump concentrations *(top row)* and the concentrations sampled from bottle grabs *(bottom row)*.

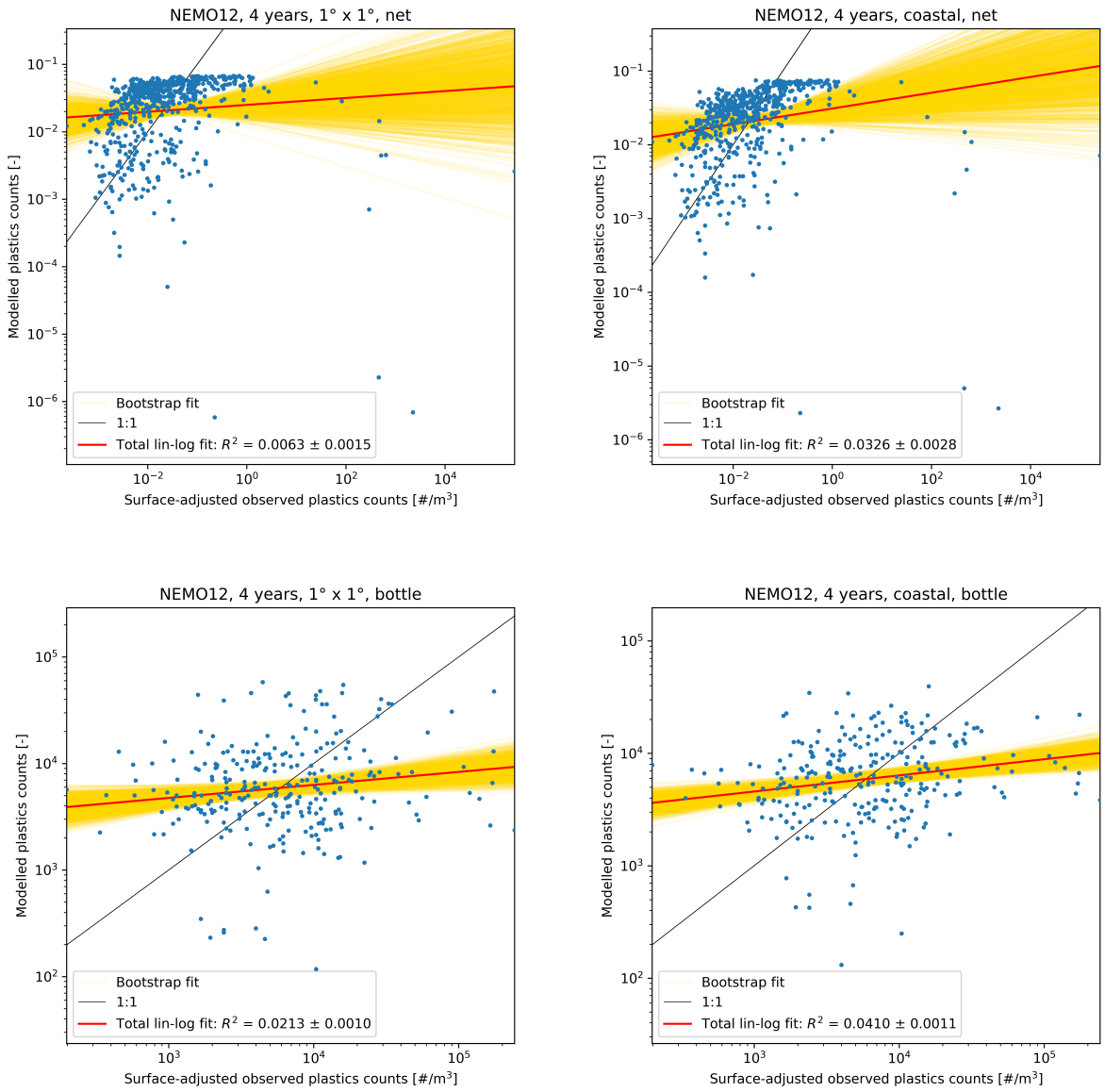


Figure 5.2: Regression between the KDE values of the virtual particle distribution generated by the NEMO12 model after a uniform release scenario *(left column)* and after a coastal release scenario *(right column)*, and the net/pump concentrations *(top row)* and the concentrations sampled from bottle grabs *(bottom row)*.

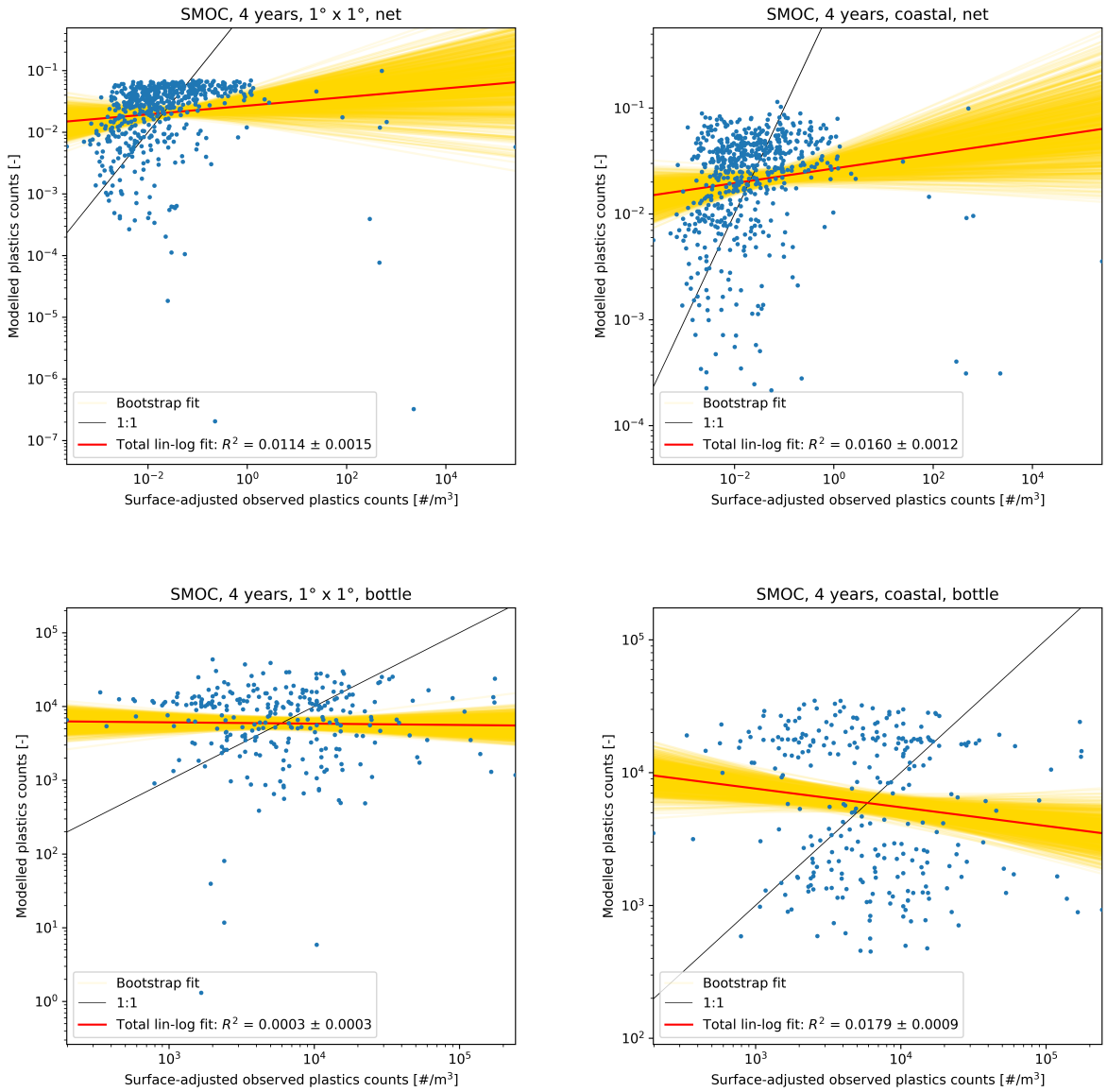


Figure 5.3: Regression between the KDE values of the virtual particle distribution generated by the SMOC model after a uniform release scenario *(left column)* and after a coastal release scenario *(right column)*, and the net/pump concentrations *(top row)* and the concentrations sampled from bottle grabs *(bottom row)*.

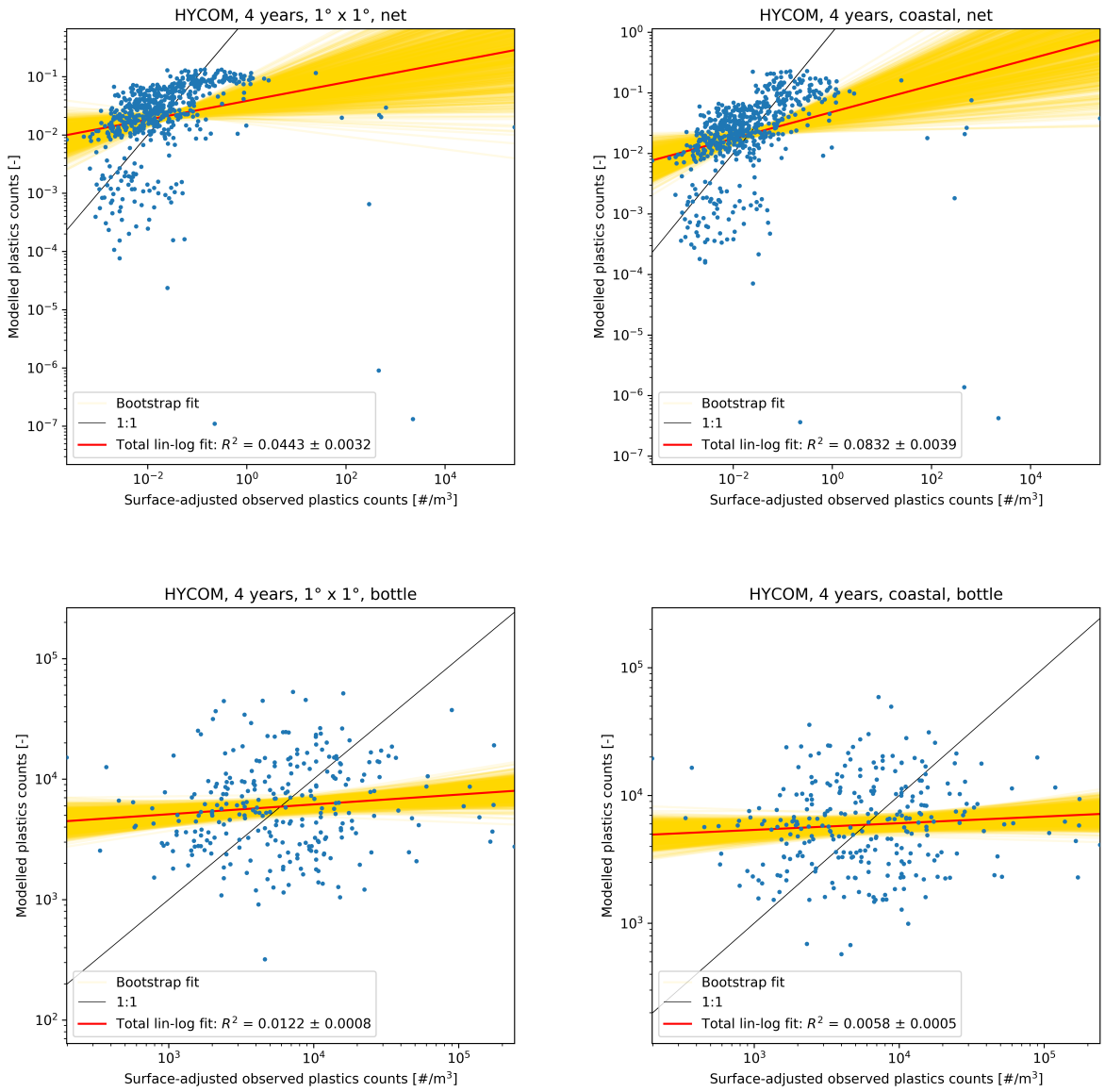


Figure 5.4: Regression between the KDE values of the virtual particle distribution generated by the HYCOM model after a uniform release scenario *(left column)* and after a coastal release scenario *(right column)*, and the net/pump concentrations *(top row)* and the concentrations sampled from bottle grabs *(bottom row)*.

yet. The high observations at lower model values probably stem from the surface-adjusted pump data (cf. Section 2.3). Again, there seems to be little difference in the coefficient of determination between the uniform and the coastal release, with $R^2 = 0.0125 \pm 0.00018$ and $R^2 = 0.0253 \pm 0.0025$ remaining too small to be meaningful.

The scatter plots for NEMO12 are shown in Figure 5.2. Comparison with the bottle data again results in a broad distribution with $R^2 = 0.0213 \pm 0.0010$ for the uniform release and $R^2 = 0.0410 \pm 0.0011$ for the coastal release. The scatter plots comparing the different release scenarios to the net data again show the Γ-shape. For the uniform release, $R^2 = 0.0063 \pm 0.0015$, and for the coastal release, $R^2 = 0.0326 \pm 0.0028$.

Figure 5.3 shows the scatter plots for the SMOC model. Again, the comparison of each release scenario with the bottle grabs results in a dispersed point cloud with $R^2 = 0.0003 \pm 0.0003$ for the uniform release and $R^2 = 0.0179 \pm 0.0009$ for the coastal release, for which the slope of the best linear fit is negative although it should be positive. Comparing the uniform release to the net and pump data, the data points are again arranged in a Γ-shape. This is not the case for the coastal release, because the higher modelled concentrations sampled at the locations of the observations are not cut off at the same value due to the local KDE maximum around 50 °W, 10 °S (Figure 3.5). For the uniform release, $R^2 = 0.0037 \pm 0.0003$, and for the coastal release, $R^2 = 0.0275 \pm 0.0007$.

For the HYCOM model, the scatter plots in Figure 5.4 again show the dispersed point clouds for the bottle data, with $R^2 = 0.0122 \pm 0.0008$ for the uniform release and $R^2 = 0.0058 \pm 0.0005$ for the coastal release. The comparison with the net data shows the Γ-shape after the uniform release, for which $R^2 = 0.0443 \pm 0.0032$, but gives a slightly different shape after the coastal release, as low observations in the Western North Atlantic are shifted to lower modelled values, because the accumulation zone is modelled to be confined to the eastern part of the basin. For the coastal release, $R^2 = 0.0832 \pm 0.0039$.

For each of the four models, the comparison with the bottle data does not give any information about the surface accumulation zone, as there is no clear spatial pattern to be detected in the concentrations sampled with bottle grabs. This may be due to the sampling method and locations, as sampling was performed mostly well below 30 °N outside the convergence zone in the mid-latitudes.

In the comparison with the net data, the coefficients of determination remain below 0.1 and are too small to make any assumption about which model matches the observations best. NEMO12, a numerical model, seems to predict a similar accumulation zone in the western and central part of the North Atlantic as GlobCurrent, which is mainly based on altimetry and in situ measurements and has a lower spatial resolution. The underlying physics of these models reproduce the same results, giving a similar hook shape in the scatter plots of Figures 5.1 and 5.2, independent of the spatial or temporal resolution. This is surprising in a sense that the varying grid size of NEMO models allows for an unparalleled resolution of eddies on a basin scale. These eddies might have disappeared because NEMO data is given as a 5-day-average.

HYCOM is another numerical model with the same spatial resolution as NEMO12, but with a much higher temporal resolution of 3 hours, which may be responsible for modelling the accumulation in the east. Still, Onink et al. (2019) modelled the North Atlantic accumulation zone with 3 hourly GlobCurrent data, which remained in the western and central part of the basin. Both GlobCurrent and HYCOM predict the accumulation zone in the North Pacific quite well (Onink et al. 2019; Lebreton, Greer, and Borrero 2012), which validates their modelling power. However, the hook shape present in the scatter plots of Figures 5.1 and 5.2 is much less pronounced for HYCOM, implying that the maximum of the accumulation zone is modelled better.

SMOC is the only model that includes the effect of surface waves and tides. Given that SMOC predicts a braided accumulation pattern with convergence lanes, even when waves and tides are not

included in the model, it is difficult to gauge how these effects influence the accumulation pattern.

Similarly, even though for every model R^2 is slightly larger if virtual particles are released along the coastlines, it remains inconclusive whether an even more realistic release scenario will improve the correlation between the models and the observations. Table 5.1 summarizes the change in R^2 for each model and release scenario compared to the net/pump data and to the bottle grabs. As it seems redundant to repeat any further statistical analysis with both scenarios, I will only consider the coastal release in the next sections, since this release pattern is the most realistic after all.

Table 5.1: Change in R^2 computed from the KDE values of the virtual particle distribution generated by each model after a one-time uniform and coastal release, compared to the net/pump concentrations and to the concentrations sampled from bottle grabs.

		GlobCurrent	NEMO12	SMOC	HYCOM
Uniform	Net/pump	0.0125 ± 0.0018	0.0063 ± 0.0015	0.0114 ± 0.0015	0.0443 ± 0.0032
	Bottle	0.0135 ± 0.0008	0.0213 ± 0.0010	0.0003 ± 0.0003	0.0122 ± 0.0008
Coastal	Net/pump	0.0253 ± 0.0025	0.0326 ± 0.0028	0.0160 ± 0.0012	0.0832 ± 0.0039
	Bottle	0.0326 ± 0.0011	0.0410 ± 0.0011	0.0179 ± 0.0009	0.0058 ± 0.0005

Maps of the ratio between models and observations

Figure 5.5 shows the ratio between the modelled and observed concentrations for each model, after deleting observations around Europe and closer than 100 km to the coastline. The closer models and observations match in a certain location, the whiter the colour of the bin (and the closer a data point lies to the 1-to-1 line in the corresponding scatter plot). All models seem to overpredict the concentrations west of 55 °W, while they underpredict them between 45 and 50 °W. HYCOM also overpredicts the concentrations around 40 °W, as it locates the maximum of the accumulation zone more to the east than the other models. In the Caribbean Sea, SMOC over- and HYCOM underpredicts the concentrations. There are some outliers indicated in dark blue, for which all models strongly underpredict the concentration. The samples at these locations have mostly been taken by an underwater pump.

5.2 Age of plastic particles

Impact of the sampling year

Both of the previous release scenarios are one-time releases that do not take into account the increase in plastic production and input into the ocean over the last decades (Wilcox, Hardesty, and Law 2019). There are multiple ways to include temporal variability into the analysis. One way is to assume that the exponential plastic production, doubling globally every 11 years, reflects the plastic input into the oceans as well. In the North Atlantic specifically, the doubling time lies around 8 years (Wilcox, Hardesty, and Law 2019).

Since the observational records span a range of sampling years from 1979 to 2020, it is reasonable to assume that if particles do not sink, more recent samples would show higher concentrations than earlier samples at the same location, simply because the abundance of plastic has increased over time. When comparing a sample at a specific location to the modelled distribution, the sampling year will therefore influence how well the modelled and observed concentrations match. In order to address this, I adjusted the value of the modelled concentration with a factor corresponding

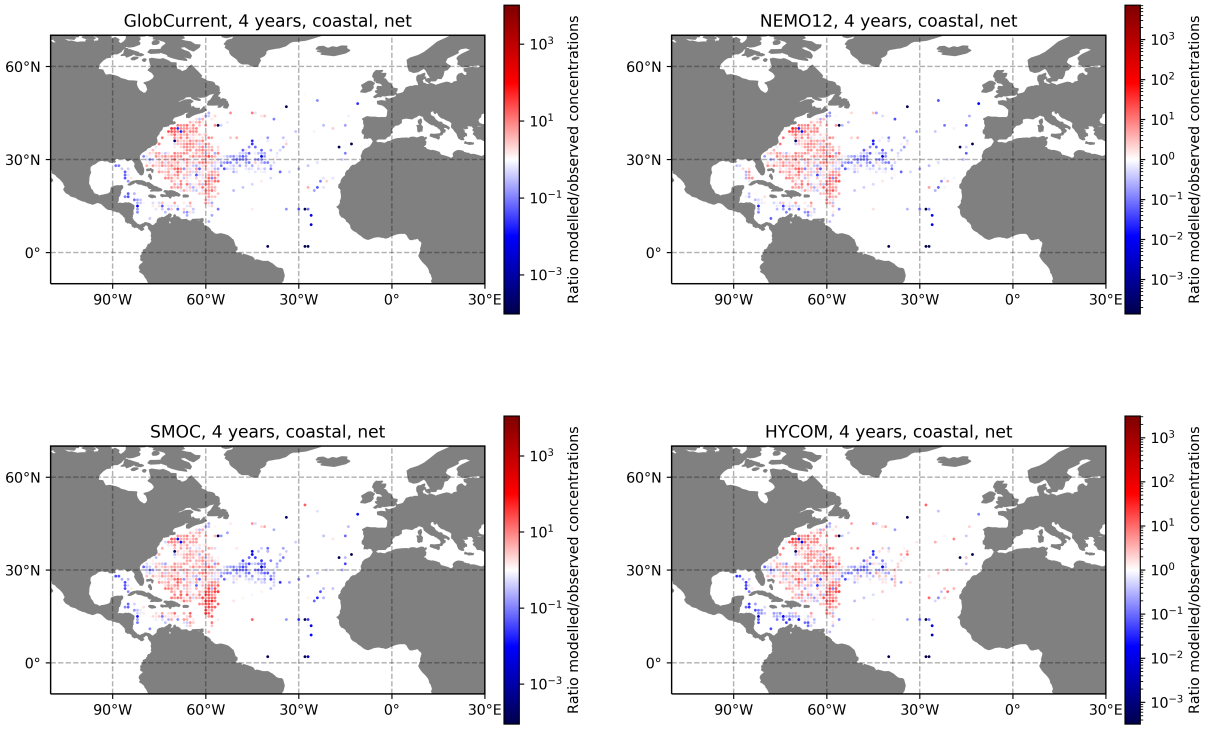


Figure 5.5: Ratio between the modelled concentrations after a coastal release scenario for each model and the observed net/pump concentrations. Observations around Europe and closer than 100 km to the continental coastlines have been deleted.

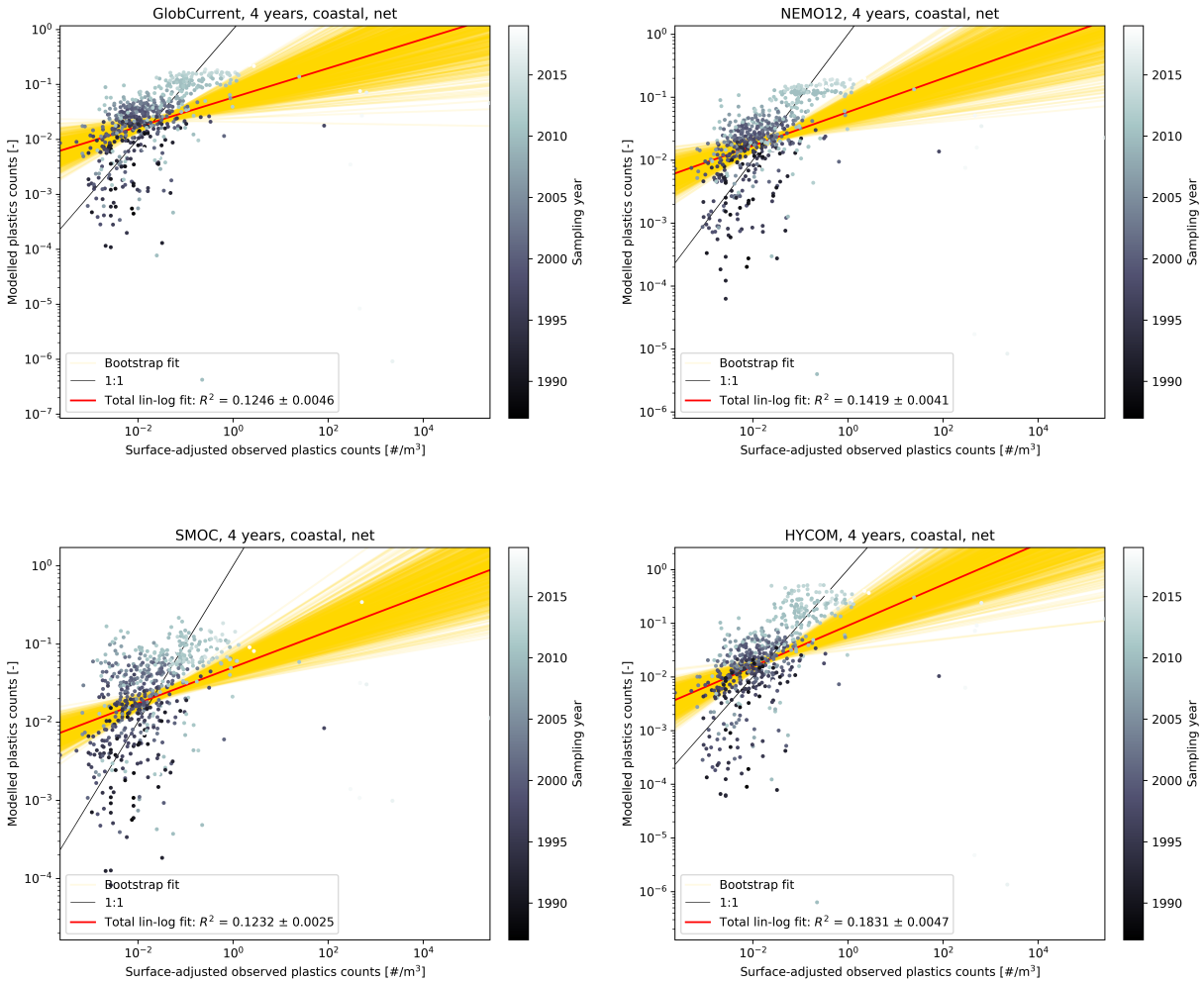


Figure 5.6: Regression between the KDE values of the virtual particle distribution generated by the each model after a coastal release scenario and the net/pump concentrations, after adjusting the modelled concentrations with the sampling year of the observations at the same location.

to $2^{(t-t_0)/T_2}$, where t is the sampling year at the corresponding location, $t_0 = 1979$ and T_2 is the doubling time of 8 years.

Figure 5.6 shows the scatter plots comparing each of the four models using the coastal release to the net/pump data, after adjusting the modelled concentrations with the sampling year of the observation at the same location. The most striking feature in each of these plots is the increase in observed concentrations for recent sampling years. Wilcox, Hardesty, and Law (2019) already noticed the increase in the abundance of floating particles the North Atlantic, using the data from Law et al. (2010) that is also analyzed in this thesis. Nevertheless, the same trend is also found across the extended data set.

The unusually high concentrations from recent pump data are now paired with higher modelled values as well, aligning the data points closer around the black line indicating 1-to-1 correspondence. This is probably the main reason why the coefficients of determination increase for each model when the modelled values are adjusted according to the sampling year of the observations at the same location, from $R^2 = 0.0253 \pm 0.0025$ to $R^2 = 0.1246 \pm 0.0046$ for GlobCurrent, from $R^2 = 0.0326 \pm 0.0028$ to $R^2 = 0.1419 \pm 0.0041$ for NEMO12, from $R^2 = 0.0160 \pm 0.0012$ to $R^2 = 0.1232 \pm 0.0025$ for SMOC, and from $R^2 = 0.0832 \pm 0.0039$ to $R^2 = 0.1831 \pm 0.0047$ for HYCOM.

Repeated release and sinking

A second option to incorporate temporal variability is to include a sink for virtual particles in the analysis. Particles do not beach in my simulations because the coastal dynamics are too complicated to resolve them in a meaningful parametrization in the scope of this thesis. But buoyant plastic particles may be lost from the sea surface by sinking through the water column and even down to the seafloor as "marine snow" as they stick to organic debris or after being excreted in fecal pellets (van Sebille et al. 2020). Another important sinking mechanism is biofouling, when a film of algae forms around the usually hydrophobic surface of a plastic particle, raising the density of the compound and sinking it (Kaiser, Kowalski, and Waniek 2017). Kooi et al. (2017) hypothesized that this sinking might actually result in an oscillatory movement, when below the euphotic layer, algae growth stops as photosynthesis becomes impossible. The biofilm vanishes, so that the plastic particle rises towards the sea surface, where a new biofilm starts to develop.

Lobelle et al. (2021) did not consider these oscillations when they set up a three-dimensional model to study the sinking characteristics of biofouled microplastics, using data from NEMO-MEDUSA at 1/12° that includes not only physical but also biological effects. For different particle sizes and densities, the sinking timescale, defined as the number of days before the vertical velocity of a particle is negative, was mapped globally. Here, I will only consider the sinking timescales for particles larger than 10 µm, as this is the order of magnitude of the smallest particles included in the net and pump observations. At this size, the overall pattern of the sinking timescales in the Atlantic is consistent over a range of different densities. Due to downwelling around 30 °N, the center of a subtropical gyre is oligotrophic and algae concentration are low, which is why in this region the sinking timescale is too large to be determined during the model runtime of 90 days. Biofouling becomes more important at the equator because of the strong upwelling in this area, but in the Atlantic Ocean, the sinking timescale seems to increase in a narrow band just below the equator. Most of the sinking happens at latitudes north of 30 °N and south of 30 °S, with an average timescale that lies around 30 days.

In the following analysis, I assume an exponential sinking behaviour so that for a particle, the

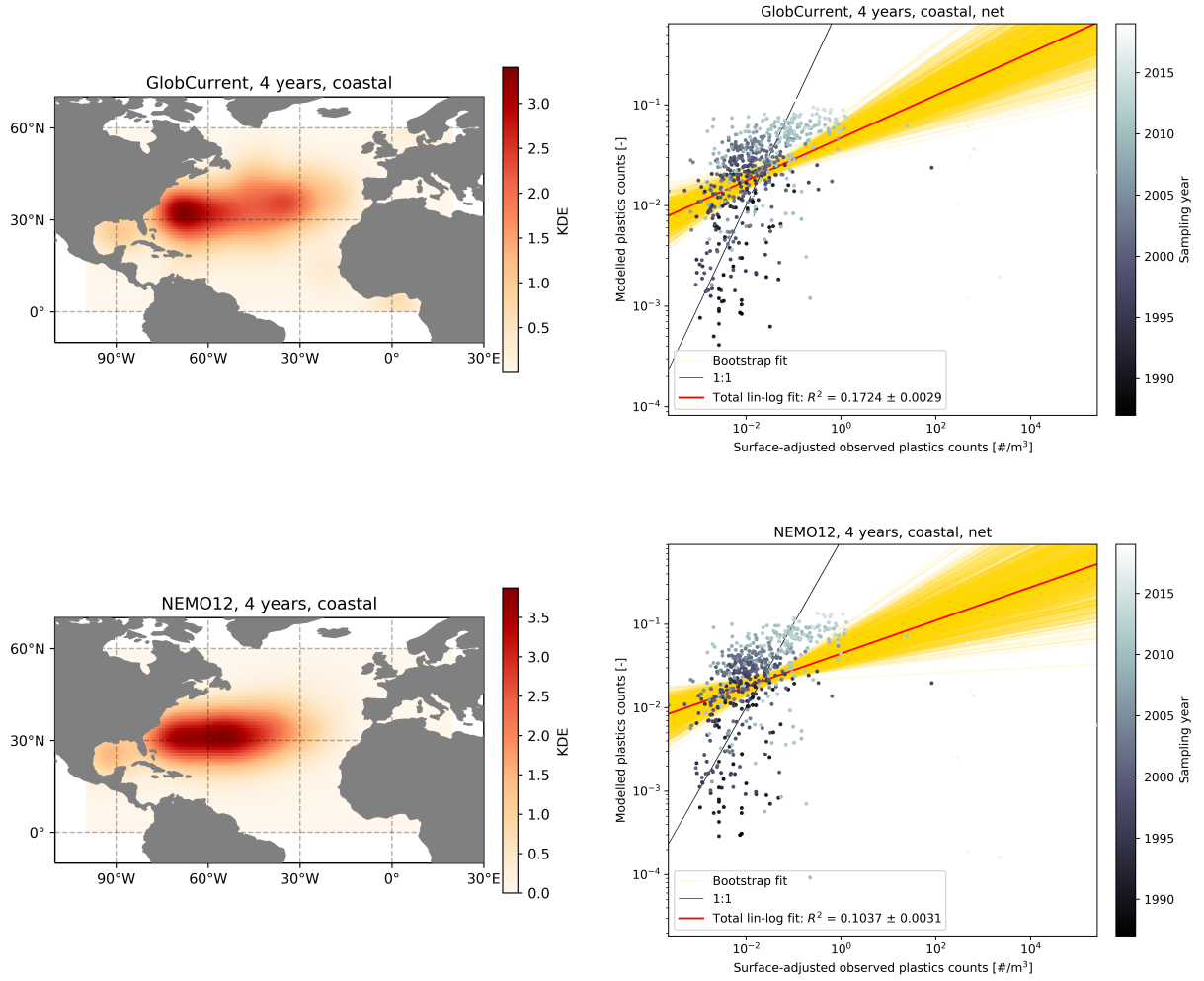


Figure 5.7: *Left column:* Maps of the KDE associated to the virtual particle distribution generated by the GlobCurrent model *(top row)* and the NEMO12 model *(bottom row)* after a repeated coastal release and sinking due to biofouling. The KDE has been calculated after removing virtual particles that ended up closer than 100 km to the coastline at the end of the simulation. *Right column:* Associated regression between the KDE values and the net/pump concentrations, after adjusting the modelled concentrations with the sampling year of the observations at the same location.

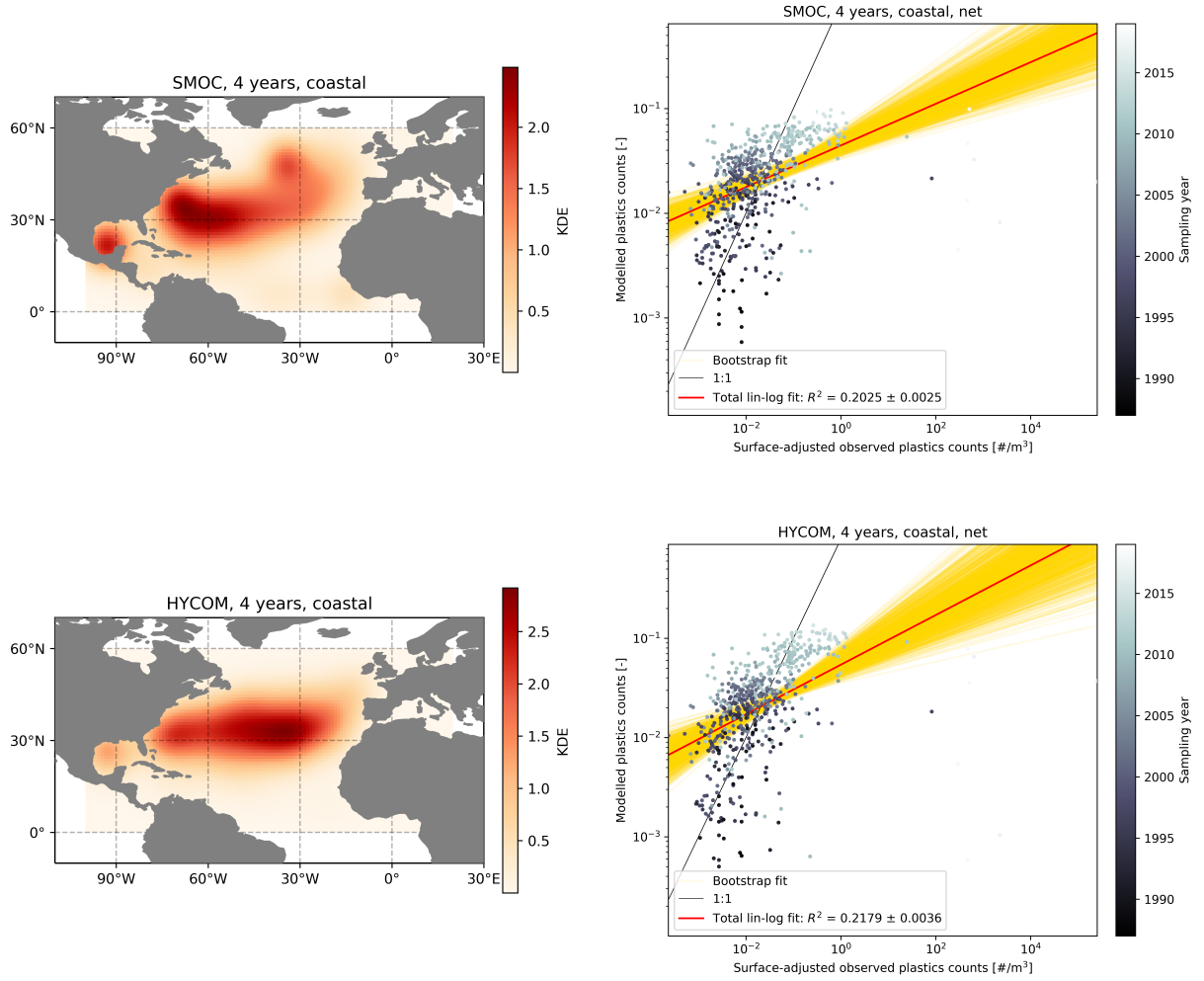


Figure 5.8: *Left column:* Maps of the KDE associated to the virtual particle distribution generated by the SMOC model *(top row)* and the HYCOM model *(bottom row)* after a repeated coastal release and sinking due to biofouling. The KDE has been calculated after removing virtual particles that ended up closer than 100 km to the coastline at the end of the simulation. *Right column:* Associated regression between the KDE values and the net/pump concentrations, after adjusting the modelled concentrations with the sampling year of the observations at the same location.

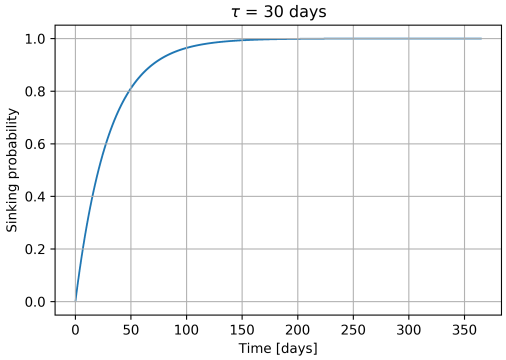


Figure 5.9: Probability that a virtual particle with an initial latitude north of 30 °N or south of 30 °S sinks over time, given a sinking timescale of 30 days.

probability to sink p_s increases over time as:

$$p_s(t) = 1 - \mathrm{e}^{-t/\tau}, \tag{5.1}$$

where the e-folding timescale τ is assumed to be the timescale as defined in Lobelle et al. (2021). In the limit $t \to \infty$, the probability to sink will reach 1 (see Figure 5.9). The analysis of the sinking behaviour remains rather qualitative, as I assume $\tau = 30$ days at latitudes north of 30 °N and south of 30 °S, and that particles in between these latitudes will never sink. This is in line with the observation by Wilcox, Hardesty, and Law (2019) that the some plastics seem to remain in the ocean for decades.

This parametrization of the sinking probability includes some spatial variability already. In order to incorporate a temporal variability, simulations of different runtimes can be concatenated to model the effect of releasing additional virtual particles as time passes. Because of the large runtime of my simulations, I had already split them up into yearly runs, by taking the final positions of one run as the starting positions for the following year (more information in Section 6.2). This way, I was able to concatenate the different slices to total runtimes of 1 year, 2 years, 3 years and 4 years, while sinking particles north of 30 °N and south of 30 °S according to the probability at the end of these runtimes in the post-processing of the simulation outputs. Finally, I added up the final positions of each of these concatenations to give the final distribution that included particles of ages between 1 to 4 years. This linear superposition of distributions is possible because the virtual particles move independently and do not interact. For an even more realistic source function, the number of particles released each year should have increased exponentially, which I did not consider in this analysis.

Figures 5.7 and 5.8 explore the effect of releasing virtual particles along the coast every year over the course of 4 years, while also sinking them due to biofouling as in Lobelle et al. (2021). For each model, the KDE associated to the final distribution is shown next to the scatter plot comparing the modelled concentrations and net/pump concentrations. Wilcox, Hardesty, and Law (2019) note that the plastic input must be larger than its removal from the surface, which is why the modelled concentrations are again adjusted to take into account the sampling year (Section 5.2), assuming a doubling time of 8 years.

For the GlobCurrent model, Figure 5.7 shows a rather broad accumulation zone with its maximum close to the western boundary at around 70 °W, contrary to the one-time release, for which the maximum of the accumulation zone is located more towards the center of the basin. In the case of the NEMO12 model, as shown in Figure 5.7, the repeated release shifts the accumulation

zone around slightly westwards, compared to the one-time release. The discerning feature in the map of the SMOC model in Figure 5.8 is the local accumulations in the Gulf of Mexico. The map of the HYCOM model in Figure 5.8 shows a more spread out accumulation zone, compared to the one-time coastal release.

For GlobCurrent, SMOC and HYCOM, the coefficient of determination increases even more compared to the one-time release if the sampling year is taken into account: $R^2 = 0.1724 \pm 0.0029$ for GlobCurrent, $R^2 = 0.2025 \pm 0.0025$ for SMOC and $R^2 = 0.21790.0036$ for HYCOM, suggesting that for these models, a more realistic simulation increases the correlation between the modelled and the observed concentrations. For the NEMO12 model, however, $R^2 = 0.1037 \pm 0.0031$ is smaller than for the previous simulations. Table 5.2 summarizes the change in R^2 for each model compared to the net/pump data, for the one-time coastal release with and without taking into account the sampling year, and for a repeated coastal release including sinking, while also taking into account the sampling year.

Table 5.2: Change in R^2 computed from the KDE values of the virtual particle distribution generated by each model after a one-time coastal release with and without taking into account the sampling year, and after a repeated coastal release including sinking, while also taking into account the sampling year, compared to the net/pump concentrations.

	GlobCurrent	NEMO12	SMOC	HYCOM
Coastal	0.0253 ± 0.0025	0.0326 ± 0.0028	0.0160 ± 0.0012	0.0832 ± 0.0039
Sampling year	0.1246 ± 0.0046	0.1419 ± 0.0041	0.1232 ± 0.0025	0.1831 ± 0.0047
Repeated	0.1724 ± 0.0029	0.1037 ± 0.0031	0.2025 ± 0.0025	0.2179 ± 0.0036

5.3 Influence of observations sampled along the coastline

In this section, I will analyze the influence of the samples taken along the coastline and around Europe (north of 35 °N and east of 10 W) on the correlation between modelled and observed concentrations. I had previously removed these samples from the analysis because the continental coastal waters and the Northwest European shelf seas are not part of the subtropical gyre. However, especially the samples around Europe show higher concentrations (Figure 2.4), which may change the overall correlation between modelled and observed concentrations. The following figures repeat the statistical analysis as seen above, but now including samples taken along the continental coastlines and around Europe.

Figures 5.10 and 5.11 show the correlation between the KDE values of the virtual particle distribution generated by each model after releasing virtual particles in one pulse uniformly across the basin (left column), and along the coast (right column), and the observed net/pump concentrations. For each model, the scatter plot is extended towards higher observations. The coefficient of determination remains below 0.1 for both the one-time uniform or coastal release (Table 5.3). For NEMO12, the best linear fit of the log-transformed data has a negative slope, implying that the model unexpectedly predicts higher concentrations where samples are lower and vice versa. Again, SMOC does not show a hook shape for the coastal release. Compared to the previous analysis, it seems like the samples taken along the coastline and around Europe mask the modest improvement in the coefficient of determination after a coastal instead of a uniform release scenario.

Figure 5.12 shows the correlation between the KDE values of the virtual particle distribution generated by each model after a one-time coastal release and the observed net/pump concentrations, after adjusting the modelled concentrations with the sampling year of the observation at the same

location. For each model, the coefficient of determination increases compared to the simulations that do not take into account the sampling year but still, R^2 mostly remains below 0.1, except for HYCOM.

Finally, Figure 5.13 displays the correlation between the KDE values of the virtual particle distribution generated by each model after releasing virtual particles along the coast every year over the course of 4 years, while sinking them due to biofouling as given in Lobelle et al. (2021), and the observed net/pump concentrations. For NEMO12, the best linear fit of the log-transformed data has a negative slope again. Contrary to the previous findings when samples taken along the coastline and around Europe were excluded, the coefficient of determination is now smaller than for the one-time coastal release when taking into account the sampling year, suggesting that the simple sinking parametrization implemented in this work does not represent the sinking behaviour of plastic particles closer than 100 km to the coastline or around Europe in a realistic way. This finding is not surprising since the sinking parametrization was implemented with a focus on the open ocean, as Lobelle et al. (2021) focussed on a large-scale, global assessment of the sinking timescales of biofouled microplastics.

Table 5.4 summarizes the changes in R^2 for each model compared to the net/pump data, after a one-time coastal release with and without taking into account the sampling year, and after a repeated coastal release including sinking, while also taking into account the sampling year.

Table 5.3: Change in R^2 computed from the KDE values of the virtual particle distribution generated by each model after a one-time uniform and coastal release, compared to the net/pump concentrations. Samples around Europe and closer than 100 km have not been deleted in this analysis.

	GlobCurrent	NEMO12	SMOC	HYCOM
Uniform	0.0016 ± 0.0004	0.0017 ± 0.0003	0.0008 ± 0.0003	0.0283 ± 0.0012
Coastal	0.0111 ± 0.0008	0.0090 ± 0.0007	0.0007 ± 0.0002	0.0956 ± 0.0021

Table 5.4: Change in R^2 computed from the KDE values of the virtual particle distribution generated by each model after a one-time coastal release with and without taking into account the sampling year, and after a repeated coastal release including sinking, while also taking into account the sampling year, compared to the net/pump concentrations. Samples around Europe and closer than 100 km have not been deleted in this analysis.

	GlobCurrent	NEMO12	SMOC	HYCOM
Coastal	0.0111 ± 0.0008	0.0090 ± 0.0007	0.0007 ± 0.0002	0.0956 ± 0.0021
Sampling year	0.0950 ± 0.0020	0.0996 ± 0.0018	0.0730 ± 0.0013	0.2022 ± 0.0026
Repeated	0.0289 ± 0.0008	0.0084 ± 0.0005	0.0155 ± 0.0007	0.0145 ± 0.0008

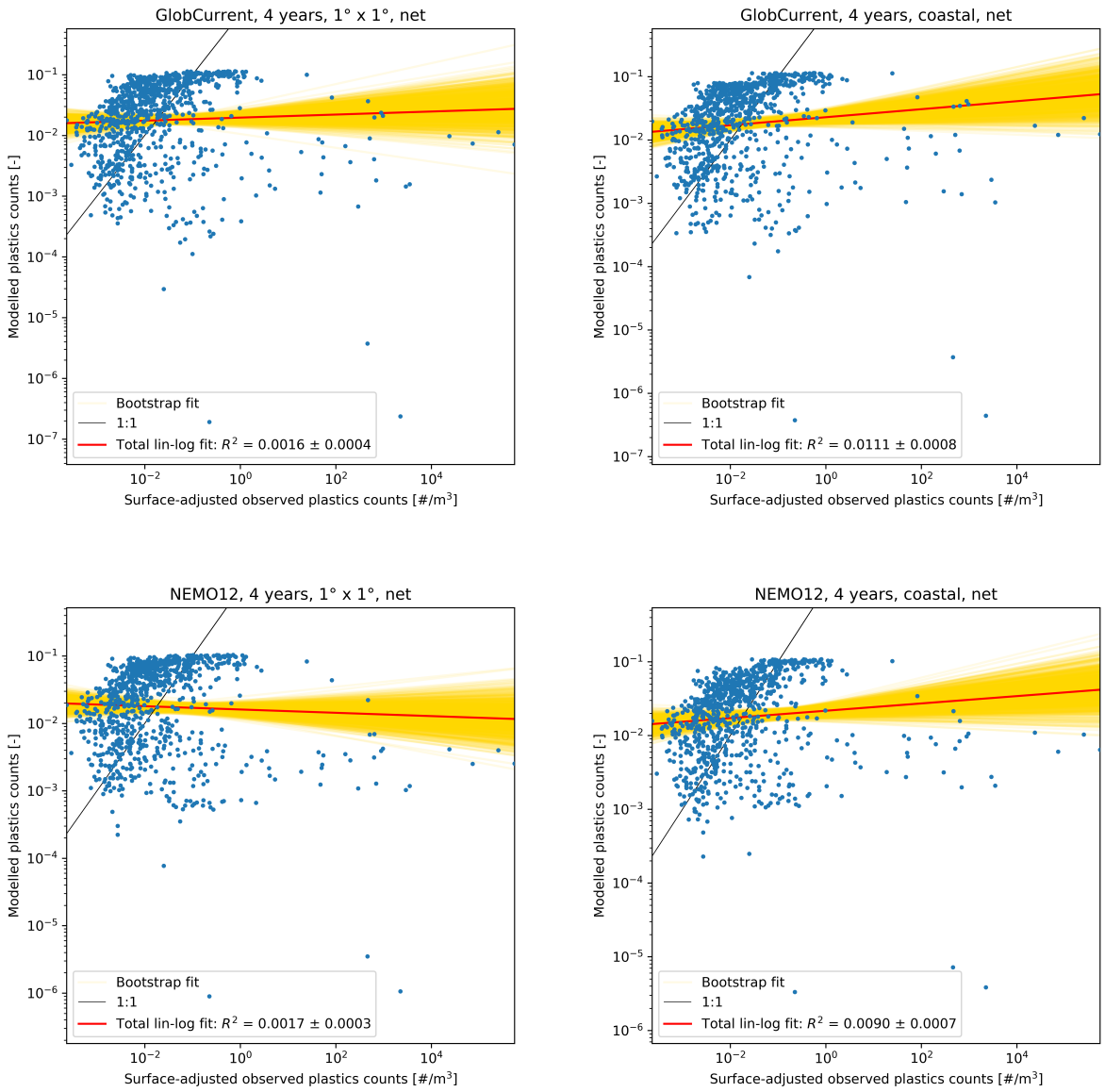


Figure 5.10: Regression between the KDE values of the virtual particle distribution generated by the GlobCurrent model *(top row)* and the NEMO12 model *(bottom row)* after a uniform release scenario *(left column)* and after a coastal release scenario *(right column)*, and the net/pump concentrations. Samples around Europe and closer than 100 km have not been deleted in this analysis.

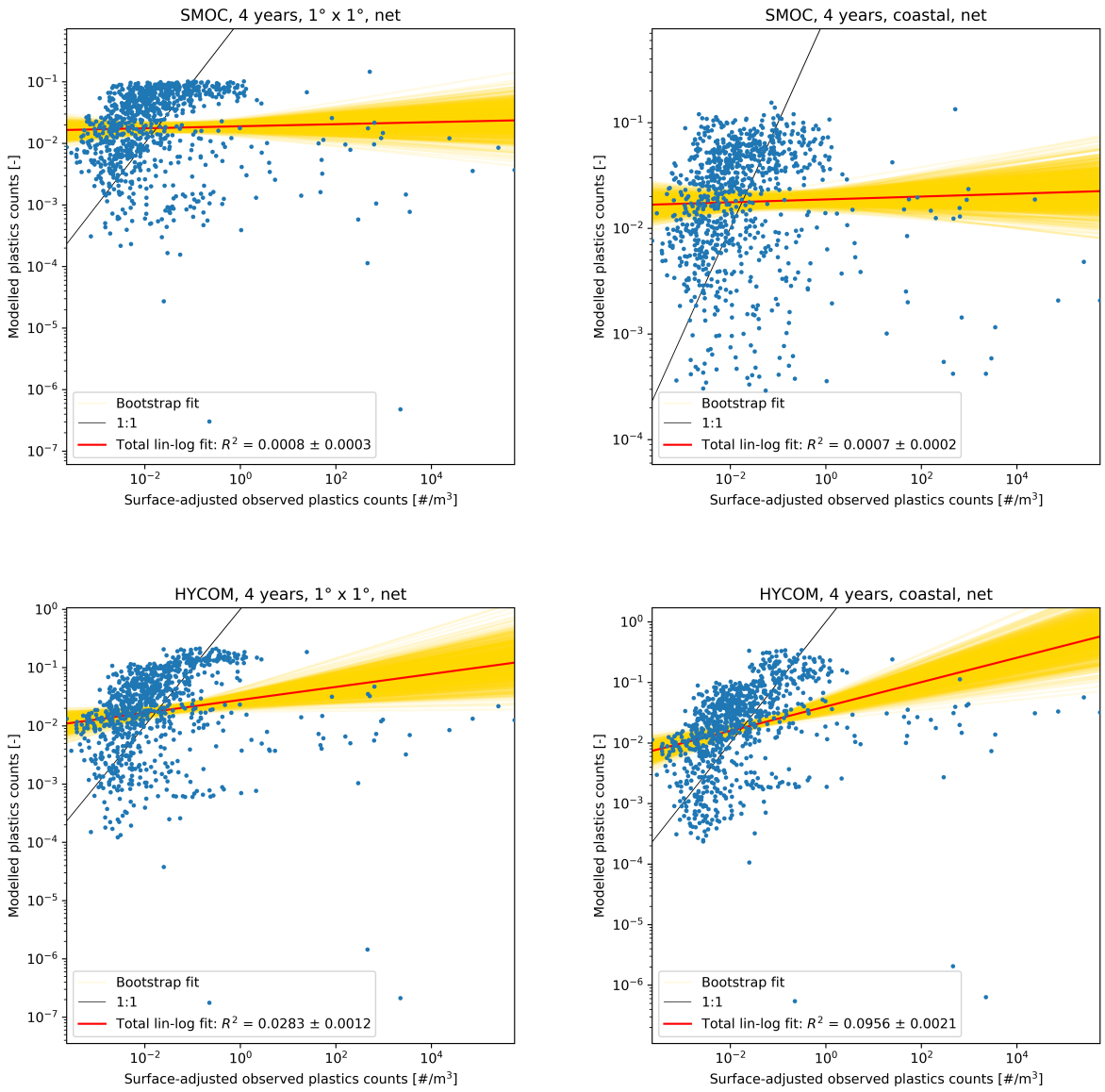


Figure 5.11: Regression between the KDE values of the virtual particle distribution generated by the SMOC model *(top row)* and the HYCOM model *(bottom row)* after a uniform release scenario *(left column)* and after a coastal release scenario *(right column)*, and the net/pump concentrations. Samples around Europe and closer than 100 km have not been deleted in this analysis.

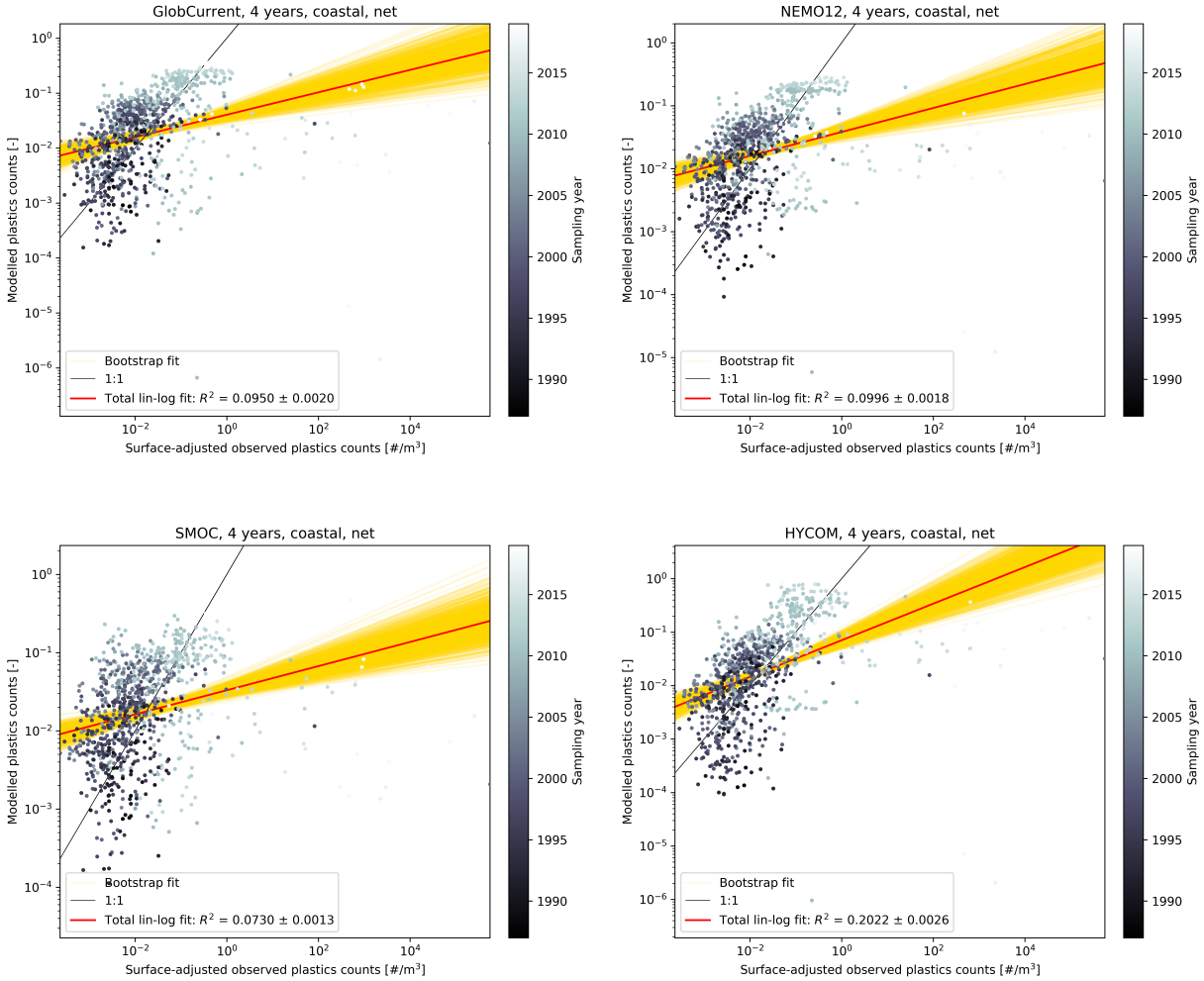


Figure 5.12: Regression between the KDE values of the virtual particle distribution generated by the each model after a coastal release scenario and the net/pump concentrations, after adjusting the modelled concentrations with the sampling year of the observations at the same location. Samples around Europe and closer than 100 km have not been deleted in this analysis.

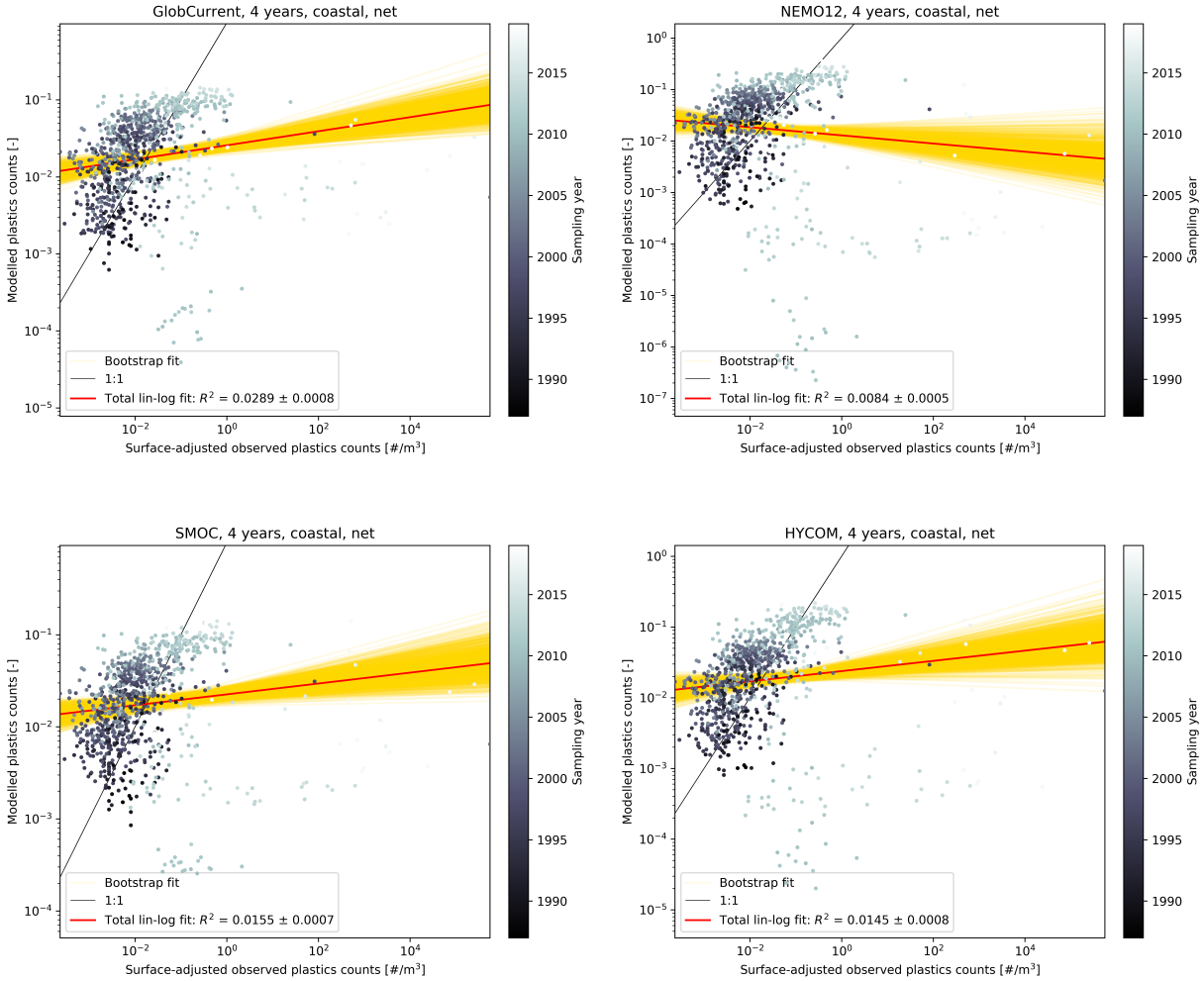


Figure 5.13: Regression between the KDE values of the virtual particle distribution generated by the each model after a repeated coastal release and sinking due to biofouling and the net/pump concentrations, after adjusting the modelled concentrations with the sampling year of the observations at the same location. Samples around Europe and closer than 100 km have not been deleted in this analysis.

Chapter 6

Discussion

6.1 Biases in the observational data

The standardization of the observational data sets has proven to be quite a challenge, as these data sets have been collected over a time span of 40 years during which sampling and detection methods have improved drastically. More and more plastic has entered the oceans during that period, too, so that earlier samples show lower concentrations than more recent samples at the same location. This has to be accounted for during the analysis; either by adjusting the observational data set itself like van Sebille et al. (2015) did by using a general additive model to standardize for the sampling year, or by using an estimate for the input rate like in this thesis.

A couple of new data sets containing samples of plastic measurements were added to the existing data set by van Sebille et al. (2015), but much more data is needed in the central and eastern part of the North Atlantic, to finally get a better picture of the true zonal extent of the surface accumulation zone. A large part of the new data consisted of observations around the Northwestern European shelf where concentrations were high, the waters of which are still part of the Atlantic Ocean, but certainly not of the accumulation zone in the subtropical gyre. The new data also included samples collected below the surface in the mixed layer by underwater pumps that needed to be adjusted for their sampling depth. A wind-dependent parametrization is currently being developed (Onink, reference updated when a DOI is available), but the implementation in this thesis probably still yielded too large values (see Figure 5.5). This may also be due to the finer mesh size used to filter particles collected in the pumps or simply to the fact that I assumed the same buoyant rise velocity for all particles, even though it varies throughout the data sets. These high concentrations in locations where models predict lower concentrations distort the correspondence between models and observations significantly, as found in Section 5.3.

More surface trawls should be undertaken in the central and eastern part of the North Atlantic, employing rigorous standardized reporting conventions as proposed by Galgani et al. (2021) so that the results from the individual trawls can easily be combined, as Law (2017) mentions. Even though bottle grabs yield in higher concentrations than the net or pump data as they catch a larger number of very small particles, their limited sampling volume is a source for high variability among samples (Barrows et al. 2017). Bottle grabs are a promising tool to help determine the absolute amount of plastics inside the oceans, but they still need to be scaled up to larger sampling volumes.

6.2 Notes on the simulations

The Parcels framework was used to run simulations advecting virtual particles with the flow fields given by four different ocean circulation models. Other than displacing particles seawards that came too close to shore, I did not include any coastal parametrization, even though 77% of particles in the oceans wash up on beaches or circulate in coastal waters (Onink et al. 2021; Olivelli, Hardesty, and Wilcox 2020). Moreover, the "beaching" parametrization of the simulations performed in this thesis seems to be at least partially flawed as particles still tend to accumulate in rivers mouths (cf. the Amazon river in Figure 3.6), even though this behaviour is less visible in a coarser spatial grid (cf. the first panel in Figure **??**), for which an error in this approach should be more apparent.

Contrary to the suggestion by Wichmann et al. (2019), a simulation runtime of 4 years was probably not long enough (see Figure 3.2). As the SMOC data set was only available for 4 years, though, each simulation was run for 4 years to enable intercomparison between the different models.

At the end of a simulation, I took the final snapshot to be representative of the dynamic steady state of the spatial virtual particle distribution. It would have been wiser to run the simulation for a year longer than the minimum mixing time, and then average over a larger number of snapshots of the virtual particle density distribution in that year.

I only implemented two different source functions to release virtual particles into the oceans. The coastal seeding seems slightly more realistic than the uniform seeding in the sense that it corresponded slightly better to the observations, while still being a simplified release scenario. A next step towards more realistic simulations includes a coastal release scenario that is proportional to the coastal population and its amount of mismanaged waste (Jambeck et al. 2015), as well as to the amount of plastics entering the oceans via rivers (Lebreton et al. 2017) or as fishing debris along shipping routes. It would also be interesting to study what kind of varying temporal release is the most realistic.

But the main bottleneck of this thesis were the model simulations runs with Parcels (version 2.2.2) on the secure shell server Gemini that is maintained by Utrecht University. While the GlobCurrent simulation proved relatively easy to run for long time periods, I had some difficulties running the higher resolution flow field data without my simulations being aborted either because they took up too much memory themselves (> 150 GB) or because they were running for days. The high spatial and temporal resolution but also the basin-scale nature of the simulations were the main challenges, even when loading only the flow field data that would need to be accessed by the virtual particles in a certain time step. This dynamic loading is called "field chunking" in Parcels but it did not always resolve the memory issues, for example in the case of the SMOC flow field. This is a Parcels issue that still needs to be investigated further.

Field chunking might reduce the memory taken up during a simulation, but it increases the runtime. I solved this issue by running the simulation for a year, then taking the final positions as the starting positions for the next year and so on, until I reached the total runtime of 4 years.

In light of these memory problems, I took care not to release too many virtual particles, so only in the order of 10,000 in the entire Atlantic (and only 2,000 virtual particles for the GlobCurrent model), about half of which were then included in analysis of the North Atlantic. These numbers are still too low, though, since the histogram of the virtual particle distributions looks rather patchy. Moreover, GlobCurrent shows a different accumulation pattern after a repeated release scenario which results in fourfold amount of virtual particle (compare the lower right panel in Figure 3.3 to the top left panel in Figure **??**), even though this difference also might be due to the inclusion of sinking into the analysis.

Parcels computes the main advection C which should be relatively fast, compared to the loading of input and saving of output in Python. I presume that the last bottleneck happened as the output

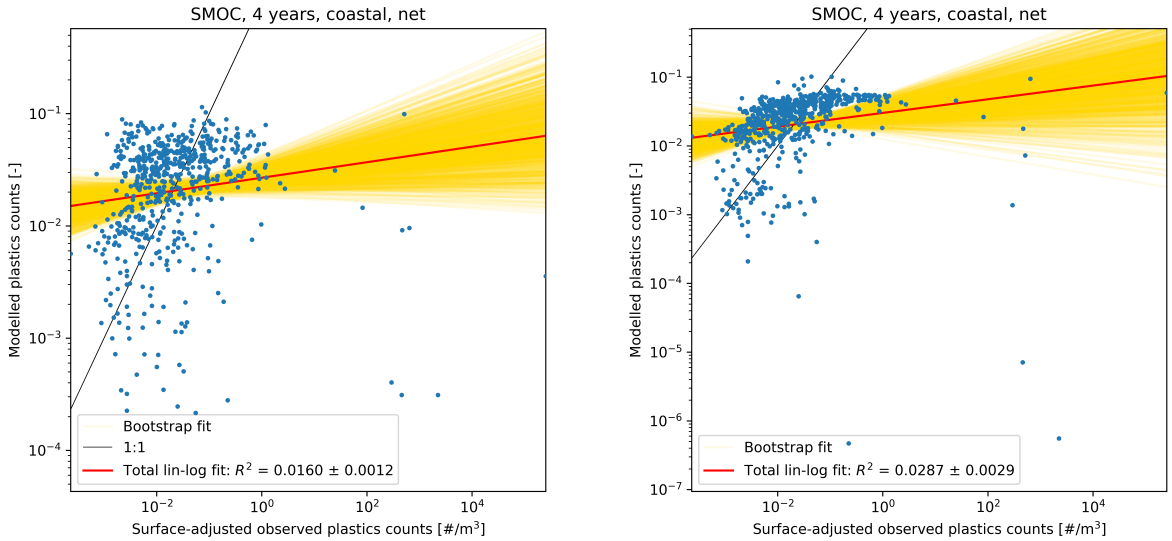


Figure 6.1: *Left:* Regression between the KDE values of the virtual particle distribution generated by the SMOC model after a uniform release scenario and the net/pump concentrations. *Right:* Same as left panel, but only using the large-scale circulation component, excluding the effect of waves or tides.

was saved hourly in a netCDF file, when a weekly time step should have sufficed for my analysis. I did not try saving any of the longer simulations in weekly slices, so this remains a hypothesis only.

6.3 Comparison of models and observations

The ocean circulation models all seem to perform comparably well, except for the SMOC model, in which the virtual particles converge in lanes right before the end of the simulation. This model also points out an underlying flaw in the analysis, as shown in Figure 6.1. The left panel displays the scatter plot comparing the SMOC model including the wave and tide action to the net data, while the scatter plot in the right panel uses only the large-scale component of the surface circulation. The developers of the SMOC model show that the product compares better to drifter data if it includes waves and tides (Drillet et al. 2019). However, comparing the coefficients of determination which are both lower than 0.1, Figure 6.1 seems to indicate that SMOC corresponds best to the observations if waves and tides are not included ($R^2 = 0.0160 \pm 0.0012$ vs. $R^2 = 0.0287 \pm 0.0029$). This further supports the claim that at low R^2 values, one cannot say whether one model performs better than the other.

Yet also the KDE seems to have its own flaws in this analysis. I decided to use the KDE as an alternative to a two-dimensional histogram, since I did not achieve full coverage of the subtropical gyre with the number of virtual particles used in my simulations (cf. e.g. Figure 3.3). Indeed, a histogram using around 4,000 particles has a lot of empty bins and therefore looks rather patchy (Figure 4.1), which does not allow for any statements about the location of the maximum of the accumulation zone. However, the computation of the KDE also presents some challenges. Its main disadvantage is the dependency on the bandwidth that determines the resolution of individual features in the particle distribution. Even though the chosen bandwidth seems to yield an acceptable KDE for some models (cf. Figures 3.3, 3.4, 3.6), this is not the case for SMOC (cf. Figure 3.5).

6.4 Inclusion of subgrid-scale processes

Subgrid-scale processes evolve on length scales that are too small to be resolved by a model's spatial resolution. In numerical models, they are often parametrized as diffusion. For instance, SMOC, NEMO12 and HYCOM take into account horizontal and vertical diffusion. In this thesis, I used models with spatial resolutions down to 1/12°, which is good enough to resolve mesoscale structures of sizes between 10 and 200 km. Smaller-scale features such as submesoscale eddies of sizes between 1 and 10 km are not resolved, even though they might play an important role in the transport of floating debris. In fact, D'Asaro et al. (2018) found that cyclonic submesoscale eddies accumulate floating particles.

Mesoscale eddy kinetic energy (EKE) scales with the basin size (Spall 2000) and should therefore be lower in the Atlantic than in the Pacific Ocean, but is only slightly lower with a maximum value of $4000\,\mathrm{cm}^2/\mathrm{s}^2$ compared to $4500\,\mathrm{cm}^2/\mathrm{s}^2$ (Ducet and Le Traon 2001). Furthermore, Martinez, Maamaatuaiahutapu, and Taillandier 2009 found that in the South Pacific, floating marine debris tends to accumulate in regions of low EKE, probably because eddies transport floating particles as they advance through the oceans. Onink et al. 2019 were not able to confirm that virtual particles converge in regions of low EKE in the North Atlantic, attributing this to the absence of a local minimum in EKE in this basin.

Since mesoscale eddies are already resolved in the models, it may be interesting to study submesoscale eddies. Indeed, Gula, Molemaker, and McWilliams 2016 note that frontal submesoscale eddies generated by the Gulf Stream tend to have a stronger upwelling core when resolved properly. NEMO at 1/60°, HYCOM at 1/48° and eNATL60 at 1/60° are interesting models in this regard, even though the submesoscale parametrization in HYCOM at 1/12° was supported by HYCOM at 1/48° results (Haza et al. 2012). Because the computational effort increases with $\mathrm{d}x^3$, it would be worth to see how the current implementation of subgrid-scale diffusion differs from the actual submesoscale activity on the scale of the gyre center only, which is about 10° in longitude and latitude.

Chapter 7

Conclusion

The aim of this thesis was to compare the observed accumulation pattern of small, floating plastics in the North Atlantic to its modelled counterparts, but it is important to note that the data set of observed surface concentrations is still very sparse in the eastern part of the predicted accumulation zone. In order to know the true zonal extent of the surface accumulation, more observations are needed in the central and eastern part of the Western North Atlantic. Galgani et al. (2021) propose a standardized sampling protocol that will facilitate the comparison of different data sets. As more and more plastic is entering the oceans, the sampling year strongly influences the amount of plastics sampled in a specific location. This must be taken into account when analyzing the data set of surface observations. Concentrations measured inside the mixed layer can be taken as proxies for the surface concentration, but need to be carefully extrapolated by taking into account the size and density of a sample. Samples collected by bottle grabs present a high spatial heterogeneity.

The observation-based GlobCurrent model and the numerical NEMO12 model predict a similar surface accumulation in the western and central part of the North Atlantic, despite being quite different in their computation and resolution. The numerical HYCOM model locates the maximum of the surface accumulation slightly northwards and more to the east, while the numerical SMOC model shows floating particles aggregating into lanes at the end of the simulation, starting in early 2020. For GlobCurrent and NEMO12, a coastal release scenario shifts the center of the surface accumulation zone from the west more towards the center of the basin, compared to a simpler uniform release scenario. For HYCOM, it is shifted towards the center as well, but from the eastern location given after a uniform release. Comparing the improved SMOC model that includes wave and tide action to only its large-scale circulation component shows that the regression analysis fails to notice any difference. This points towards a flaw either in the statistical analysis or in the computation of the KDE, which, in its present form, does not capture any small-scale structures like the "convergence lanes" observed in the final snapshots of the SMOC simulation.

A measure for the correspondence between models and observations is the coefficient of determination R^2 that is calculated from a linear fit between the log-transformed observed concentrations and the modelled concentrations sampled at the location of the observations. Because of the large variability of the bottle grabs, the models were mainly compared to the net/pump data only. For each model, R^2 increases with increasingly realistic modelling scenarios. The correlation improves only modestly after a one-time coastal seeding of virtual particles compared to a one-time uniform release, but R^2 reaches values above 0.1 once the modelled concentrations are adjusted with the sampling year of the observation at the same location. Again, this supports the claim that the sampling year is one of the most important parameters to consider when analysing a data set of observations that spans decades. Finally, a repeated release over 4 years combined with a simple

Table 7.1: Change in R^2 computed from the KDE values of the virtual particle distribution generated by each model after a one-time uniform release without taking into account the sampling year, and after a repeated coastal release including sinking, compared to the net/pump concentrations.

	GlobCurrent	NEMO12	SMOC	HYCOM
One-time uniform	0.0253 ± 0.0025	0.0326 ± 0.0028	0.0160 ± 0.0012	0.0832 ± 0.0039
Repeated coastal	0.1724 ± 0.0029	0.1037 ± 0.0031	0.2025 ± 0.0025	0.2179 ± 0.0036

sinking parametrization gives the best correlation for GlobCurrent, SMOC and HYCOM, while R^2 decreases again for NEMO12. Table 7.1 shows the range of improvement for R^2, from the simplest to the most complex release scenario.

Since the coastal release seems to improve the correlation (albeit only modestly), the center of the accumulation zone probably lies between 25 and 55 °W in the center of the basin. More observations in this regions are direly needed to support this claim. Even though HYCOM gives the best overall coefficient of determination with the observation, the center of its accumulation zone around 35 °W is not modelled well enough. In Figure 5.5, it seems like HYCOM overpredicts the concentrations in this area.

Observations that were collected closer than 100 km to the continental coastlines or around Europe have been removed in the analysis described above, as they were not sampled in the subtropical gyre. Since the aggregation of new data sets was an important part of this thesis, it was important to study their influence on the overall correlation between models and observations. The results are rather inconclusive because for almost each release scenario, R^2 remains below 0.1, except for the HYCOM model after a coastal release, when accounting for the sampling year. This is probably due to the high concentrations sampled around Europe, which extend the range of observed concentrations. The correlation did not improve further with a repeated release over 4 years combined with a simple sinking parametrization, suggesting that the sinking parametrization implemented in this thesis does not represent the behaviour of particles closer than 100 km to the coastline or around Europe in a realistic way. Indeed, this sinking parametrization originates from a global, large-scale assessment of the sinking timescales for biofouled microplastics and does not consider coastal dynamics.

Appendix A

Additional Figures

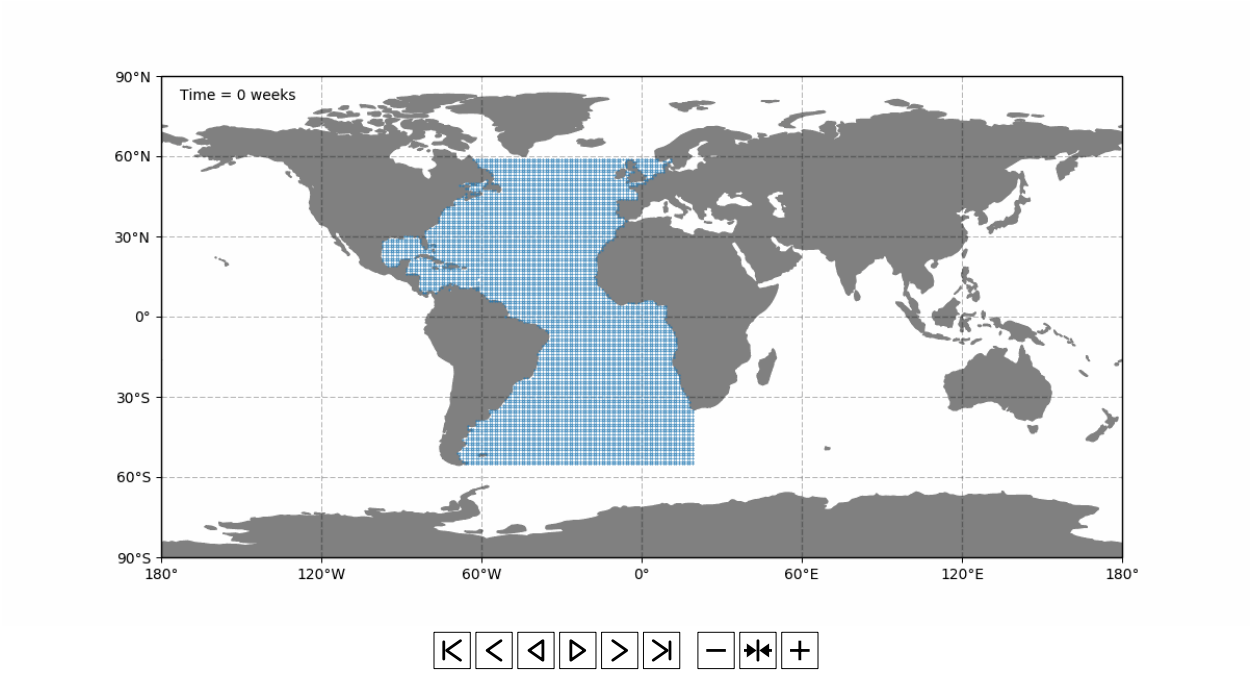


Figure A.1: Animation of the advection of virtual particles over 4 years by the large-scale component of the SMOC model, without the effect of waves or tides, after a uniform release scenario. To display the animation, please open this PDF document with Adobe Acrobat Reader.

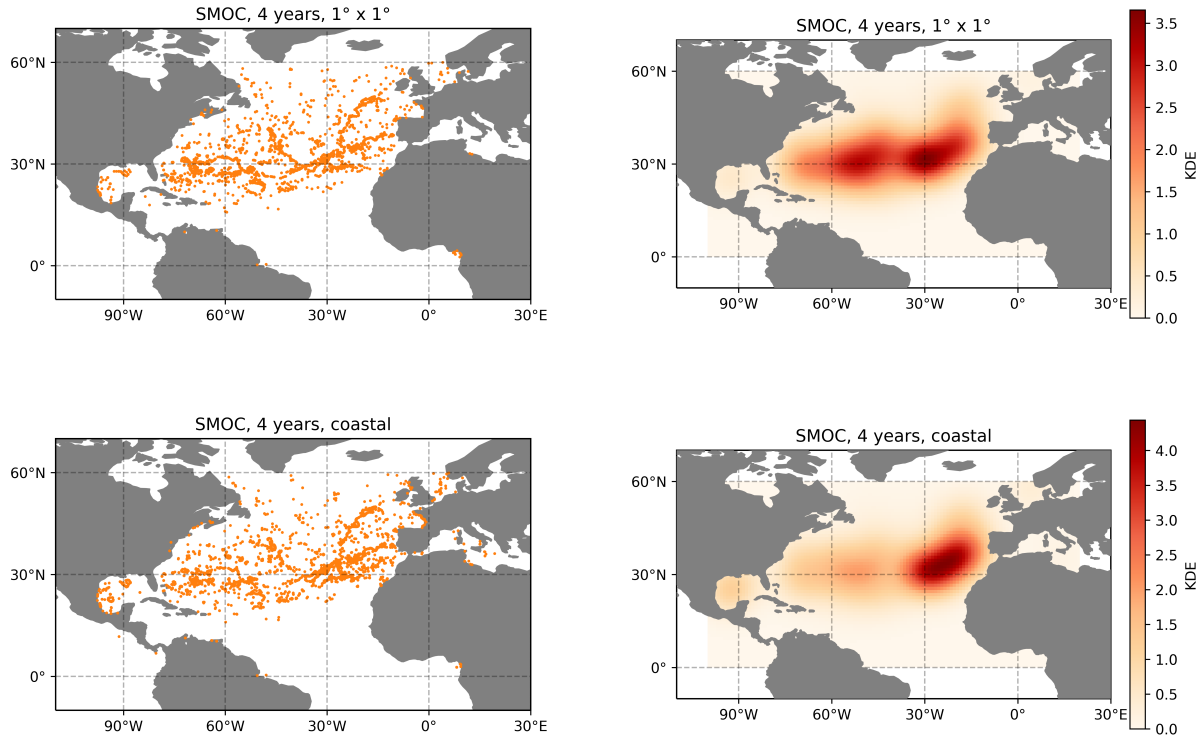


Figure A.2: Maps of the virtual particle distribution *(left column)* and the associated KDE *(right column)* generated by the large-scale circulation component of the SMOC model without the effect of waves or tides, after a uniform release scenario *(top row)* and after a coastal release scenario *(bottom row)*. The KDE has been calculated after removing virtual particles that ended up closer than 100 km to the coastline at the end of the simulation.

Bibliography

Andrady, A. L. (2011). "Microplastics in the marine environment". In: *Marine Pollution Bulletin* 62.8, pp. 1596–1605. DOI: 10.1016/j.marpolbul.2011.05.030.

Babu, G. J. and E. D. Feigelson (2006). "Astrostatistics: Goodness-of-fit and all that!" In: *Astronomical Data Analysis Software and Systems XV*. Vol. 351, p. 127.

— (2019). "Beware the Kolmogorov-Smirnov test!" In: *Astrostatistics and Astroinformatics Portal*. URL: https://asaip.psu.edu/articles/beware-the-kolmogorov-smirnov-test/.

Barnes, D. K. et al. (2018). "Marine plastics threaten giant Atlantic marine protected areas". In: *Current Biology* 28.19, R1137–R1138. DOI: 10.1016/j.cub.2018.08.064.

Barrows, A. P., S. E. Cathey, and C. W. Petersen (2018). "Marine environment microfiber contamination: Global patterns and the diversity of microparticle origins". In: *Environmental Pollution* 237, pp. 275–284. DOI: 10.1016/j.envpol.2018.02.062.

Barrows, A. P. et al. (2017). "Grab vs. neuston tow net: A microplastic sampling performance comparison and possible advances in the field". In: *Analytical Methods* 9.9, pp. 1446–1453. DOI: 10.1039/C6AY02387H.

Bell, K. and S. Cave (2011). "Comparison of environmental impact of plastic, paper and cloth bags". In: *Northern Ireland Assembly: Research and Library Service Briefing Note*.

Boufadel, M. et al. (2020). "Transport of oil droplets in the upper ocean: Impact of the eddy diffusivity". In: *Journal of Geophysical Research: Oceans* 125.2, e2019JC015727. DOI: 10.1029/2019JC015727.

Brach, L. et al. (2018). "Anticyclonic eddies increase accumulation of microplastic in the North Atlantic subtropical gyre". In: *Marine Pollution Bulletin* 126, pp. 191–196. DOI: 10.1016/j.marpolbul.2017.10.077.

Brandon, J. A., W. Jones, and M. D. Ohman (2019). "Multidecadal increase in plastic particles in coastal ocean sediments". In: *Science Advances* 5.9, eaax0587. DOI: 10.1126/sciadv.aax0587.

Cabernard, L. et al. (2018). "Comparison of Raman and Fourier transform infrared spectroscopy for the quantification of microplastics in the aquatic environment". In: *Environmental Science & Technology* 52.22, pp. 13279–13288. DOI: 10.1021/acs.est.8b03438.

Cardoso, C. and R. M. A. Caldeira (2021). "Modeling the exposure of the Macaronesia islands (NE Atlantic) to marine plastic pollution". In: *Frontiers in Marine Science* 8, p. 332. DOI: 10.3389/fmars.2021.653502.

Chamecki, M. et al. (2019). "Material transport in the ocean mixed layer: Recent developments enabled by large eddy simulations". In: *Reviews of Geophysics* 57.4, pp. 1338–1371. DOI: 10.1029/2019RG000655.

Chenillat, F. et al. (2021). "Fate of floating plastic debris released along the coasts in a global ocean model". In: *Marine Pollution Bulletin* 165, p. 112116. DOI: 10.1016/j.marpolbul.2021.112116.

Chubarenko, I. et al. (2016). "On some physical and dynamical properties of microplastic particles in marine environment". In: *Marine Pollution Bulletin* 108.1-2, pp. 105–112. DOI: 10.1016/j.marpolbul.2016.04.048.

Cózar, A. et al. (2017). "The Arctic Ocean as a dead end for floating plastics in the North Atlantic branch of the Thermohaline Circulation". In: *Science Advances* 3.4. DOI: 10.1126/sciadv.1600582.

D'Angelo, S. and R. Meccariello (2021). "Microplastics: A threat for male fertility". In: *International Journal of Environmental Research and Public Health* 18.5, p. 2392. DOI: 10.3390/ijerph18052392.

Delandmeter, P. and E. van Sebille (2019). "The Parcels v2. 0 Lagrangian framework: New field interpolation schemes". In: *Geoscientific Model Development* 12.8, pp. 3571–3584. DOI: 10.5194/gmd-12-3571-2019.

Drillet, Y. et al. (2019). *SMOC: A new global surface current product containing the effects of general circulation, waves and tides*. Poster presented at EGU 2019, session 0S4.11/X4.94.

Ducet, N. and P.-Y. Le Traon (2001). "A comparison of surface eddy kinetic energy and Reynolds stresses in the Gulf Stream and the Kuroshio Current systems from merged TOPEX/Poseidon and ERS-1/2 altimetric data". In: *Journal of Geophysical Research: Oceans* 106.C8, pp. 16603–16622. DOI: 10.1029/2000JC000205.

D'Asaro, E. A. et al. (2018). "Ocean convergence and the dispersion of flotsam". In: *Proceedings of the National Academy of Sciences* 115.6, pp. 1162–1167. DOI: 10.1073/pnas.1718453115.

Eriksen, M. et al. (2014). "Plastic pollution in the world's oceans: More than 5 trillion plastic pieces weighing over 250,000 tons afloat at sea". In: *PLoS ONE* 9.12, e111913. DOI: 10.1371/journal.pone.0111913.

Gago, J., M. Henry, and F. Galgani (2015). "First observation on neustonic plastics in waters off NW Spain (spring 2013 and 2014)". In: *Marine Environmental Research* 111, pp. 27–33. DOI: j.marenvres.2015.07.009.

Galgani, F. et al. (2021). "Guidelines and forms for gathering marine litter data: Beach and seafloor trawlings". In: *EMODnet Chemistry*. Updated version 11/01/2021, p. 58. DOI: 10.6092/15c0d34c-a01a-4091-91ac-7c4f561ab508.

García Marina, J. et al. (2019). *Sampling microplastics across the North Atlantic Ocean*. Last accessed: 26/01/2021. URL: https://www.seamarproject.org/en/sailing-for-the-knowledge-of-microplastics.html.

García-Portugués, E. (2021). *Notes for Nonparametric Statistics*. Version 6.4.4. ISBN 978-84-09-29537-1. Last accessed: 04/06/2021. URL: https://bookdown.org/egarpor/NP-UC3M/.

Gasparin, F. et al. (2018). "A large-scale view of oceanic variability from 2007 to 2015 in the global high resolution monitoring and forecasting system at Mercator Océan". In: *Journal of Marine Systems* 187, pp. 260–276. DOI: 10.1016/j.jmarsys.2018.06.015.

GEOMAR Helmholtz Centre for Ocean Research Kiel (2018). *Microplastic measurement journey around the world from the Volvo Ocean Race 2017/2018, Turn the Tide on Plastic and AkzoNobel*. Last accessed: 09/01/2021. URL: https://www.emodnet-ingestion.eu/submissions/submissions_details.php?menu=39&tpd=232&step=0103.

Goldstein, M. C., M. Rosenberg, and L. Cheng (2012). "Increased oceanic microplastic debris enhances oviposition in an endemic pelagic insect". In: *Biology letters* 8.5, pp. 817–820. DOI: 10.1098/rsbl.2012.0298.

González-Pleiter, M. et al. (2021). "A pilot study about microplastics and mesoplastics in an Antarctic glacier". In: *The Cryosphere* 15.6, pp. 2531–2539. DOI: 10.5194/tc-15-2531-2021.

Green, D. S. et al. (2018a). "A comparison of sampling methods for seawater microplastics and a first report of the microplastic litter in coastal waters of Ascension and Falkland Islands". In: *Marine Pollution Bulletin* 137, pp. 695–701. DOI: 10.1016/j.marpolbul.2018.11.004.

— (2018b). "Data for: A comparison of sampling methods for seawater microplastics and a first report of the microplastic litter in coastal waters of Ascension and Falkland Islands". In: *Mendeley Data*, p. V1. DOI: 10.17632/g3pj8wkc5y.1.

Gula, J., M. J. Molemaker, and J. C. McWilliams (2016). "Submesoscale dynamics of a Gulf Stream frontal eddy in the South Atlantic Bight". In: *Journal of Physical Oceanography* 46.1, pp. 305–325. DOI: 10.1175/JPO-D-14-0258.1.

Hardesty, B. D. et al. (2017). "Using numerical model simulations to improve the understanding of micro-plastic distribution and pathways in the marine environment". In: *Frontiers in Marine Science* 4, p. 30. DOI: 10.3389/fmars.2017.00030.

Haza, A. C. et al. (2012). "Parameterization of particle transport at submesoscales in the Gulf Stream region using Lagrangian subgridscale models". In: *Ocean Modelling* 42, pp. 31–49. DOI: 10.1016/j.ocemod.2011.11.005.

Herrera, A. et al. (2020). "First evaluation of neustonic microplastics in the Macaronesian region, NE Atlantic". In: *Marine Pollution Bulletin* 153, p. 110999. DOI: 10.1016/j.marpolbul.2020.110999.

Hersbach, H. et al. (2018). "ERA5 hourly data on single levels from 1979 to present". In: *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. Last accessed: 12/03/2021. DOI: 10.24381/cds.adbb2d47.

Holte, J. et al. (2017). "An Argo mixed layer climatology and database". In: *Geophysical Research Letters* 44.11, pp. 5618–5626. DOI: 10.1002/2017GL073426.

Jambeck, J. R. et al. (2015). "Plastic waste inputs from land into the ocean". In: *Science* 347.6223, pp. 768–771. DOI: 10.1126/science.1260352.

Jamieson, A. J. et al. (2019). "Microplastics and synthetic particles ingested by deep-sea amphipods in six of the deepest marine ecosystems on Earth". In: *Royal Society Open Science* 6.2, p. 180667. DOI: 10.1098/rsos.180667.

Kaandorp, M. L., H. A. Dijkstra, and E. van Sebille (2021). "Modelling size distributions of marine plastics under the influence of continuous cascading fragmentation". In: *Environmental Research Letters*. Accepted manuscript. DOI: 10.1088/1748-9326/abe9ea.

Kaiser, D., N. Kowalski, and J. J. Waniek (2017). "Effects of biofouling on the sinking behavior of microplastics". In: *Environmental Research Letters* 12.12, p. 124003. DOI: 10.1088/1748-9326/aa8e8b.

Kanhai, L. D. K. et al. (2017). "Microplastic abundance, distribution and composition along a latitudinal gradient in the Atlantic Ocean". In: *Marine Pollution Bulletin* 115.1-2, pp. 307–314. DOI: 10.1016/j.marpolbul.2016.12.025.

Komen, G. J. et al. (1996). *Dynamics and Modelling of Ocean Waves*. Cambridge University Press, p. 532.

Kooi, M. et al. (2016). "The effect of particle properties on the depth profile of buoyant plastics in the ocean". In: *Scientific Reports* 6.1, pp. 1–10. DOI: 10.1038/srep33882.

Kooi, M. et al. (2017). "Ups and downs in the ocean: Effects of biofouling on vertical transport of microplastics". In: *Environmental Science & Technology* 51.14, pp. 7963–7971. DOI: 10.1021/acs.est.6b04702.

Kubota, M. (1994). "A mechanism for the accumulation of floating marine debris north of Hawaii". In: *Journal of Physical Oceanography* 24.5, pp. 1059–1064. DOI: 10.1175/1520-0485(1994)024<1059:AMFTAO>2.0.CO;2.

Kubota, M., K. Takayama, and D. Namimoto (2005). "Pleading for the use of biodegradable polymers in favor of marine environments and to avoid an asbestos-like problem for the future". In: *Applied Microbiology and Biotechnology* 67.4, pp. 469–476. DOI: 10.1007/s00253-004-1857-2.

Kühn, S., E. L. Bravo Rebolledo, and J. A. van Franeker (2015). "Deleterious effects of litter on marine life". In: *Marine Anthropogenic Litter*. Ed. by M. Bergmann, L. Gutow, and M. Klages. Cham: Springer International Publishing, pp. 75–116. ISBN: 978-3-319-16510-3. DOI: 10.1007/978-3-319-16510-3_4.

Kukulka, T. et al. (2012). "The effect of wind mixing on the vertical distribution of buoyant plastic debris". In: *Geophysical Research Letters* 39.7. DOI: 10.1029/2012GL051116.

Lange, M. and E. van Sebille (2017). "Parcels v0.9: Prototyping a Lagrangian ocean analysis framework for the petascale age". In: *Geoscientific Model Development* 10.11, pp. 4175–4186. DOI: 10.5194/gmd-10-4175-2017.

Large, W. G., J. C. McWilliams, and S. C. Doney (1994). "Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization". In: *Reviews of Geophysics* 32.4, pp. 363–403. DOI: 10.1029/94RG01872.

Law, K. L. (2017). "Plastics in the marine environment". In: *Annual review of marine science* 9, pp. 205–229. DOI: 10.1146/annurev-marine-010816-060409.

Law, K. L. et al. (2010). "Plastic accumulation in the North Atlantic subtropical gyre". In: *Science* 329.5996, pp. 1185–1188. DOI: 10.1126/science.1192321.

Lebreton, L., M. Egger, and B. Slat (2019). "A global mass budget for positively buoyant macroplastic debris in the ocean". In: *Scientific reports* 9.1, pp. 1–10. DOI: 10.1038/s41598-019-49413-5.

Lebreton, L. C.-M., S. D. Greer, and J. C. Borrero (2012). "Numerical modelling of floating debris in the world's oceans". In: *Marine Pollution Bulletin* 64.3, pp. 653–661. DOI: 10.1016/j.marpolbul.2011.10.027.

Lebreton, L. C.-M. et al. (2017). "River plastic emissions to the world's oceans". In: *Nature Communications* 8.1, pp. 1–10. DOI: 10.1038/ncomms15611.

Lebreton, L. C.-M. et al. (2018). "Evidence that the Great Pacific Garbage Patch is rapidly accumulating plastic". In: *Scientific Reports* 8.1, pp. 1–15. DOI: 10.1038/s41598-018-22939-w.

Lellouche, J.-M. et al. (2018). "Recent updates to the Copernicus Marine Service global ocean monitoring and forecasting real-time 1/12° high-resolution system". In: *Ocean Science* 14.5, pp. 1093–1126. DOI: 10.5194/os-2018-15.

Lobelle, D. et al. (2021). "Global modeled sinking characteristics of biofouled microplastic". In: *Journal of Geophysical Research: Oceans* 126.4. DOI: 10.1029/2020JC017098.

Lorenz, C. et al. (2019). "Spatial distribution of microplastics in sediments and surface waters of the Southern North Sea". In: *Environmental Pollution* 252, pp. 1719–1729. DOI: 10.1016/j.envpol.2019.06.093.

Madec, G. et al. (2017). "NEMO ocean engine". In: *Notes du Pôle de modélisation de l'Institut Pierre-Simon Laplace* 27. DOI: 10.5281/zenodo.3248739.

Martinez, E., K. Maamaatuaiahutapu, and V. Taillandier (2009). "Floating marine debris surface drift: Convergence and accumulation toward the South Pacific subtropical gyre". In: *Marine Pollution Bulletin* 58.9, pp. 1347–1355. DOI: 10.1016/j.marpolbul.2009.04.022.

Maximenko, N., J. Hafner, and P. Niiler (2012). "Pathways of marine debris derived from trajectories of Lagrangian drifters". In: *Marine Pollution Bulletin* 65.1-3, pp. 51–62. DOI: 10.1016/j.marpolbul.2011.04.016.

Metzger, E. J. et al. (2017). *Global Ocean Forecast System 3.1 Validation Testing*. Naval Research Laboratory, NRL/MR/7320–17-9722.

Montoto-Martínez, T., J. J. Hernández-Brito, and M. D. Gelado-Caballero (2018). *Pump-underway ship intake: Microplastic concentrations on oceanic and coastal seawaters off the Canary Islands (subtropical NE Atlantic)*. DOI: 10.17882/70158.

Oberbeckmann, S. et al. (2014). "Spatial and seasonal variation in diversity and structure of microbial biofilms on marine plastics in Northern European waters". In: *FEMS Microbiology Ecology* 90.2, pp. 478–492. DOI: 10.1111/1574-6941.12409.

Olivelli, A., B. D. Hardesty, and C. Wilcox (2020). "Coastal margins and backshores represent a major sink for marine debris: Insights from a continental-scale analysis". In: *Environmental Research Letters* 15.7, p. 074037. DOI: 10.1088/1748-9326/ab7836.

Onink, V. et al. (2019). "The role of Ekman currents, geostrophy, and stokes drift in the accumulation of floating microplastic". In: *Journal of Geophysical Research: Oceans* 124.3, pp. 1474–1490. DOI: 10.1029/2018JC014547.

Onink, V. et al. (2021). "Global simulations of marine plastic transport show plastic trapping in coastal zones". In: *Environmental Research Letters*. DOI: 10.1088/1748-9326/abecbd.

Ostle, C. et al. (2019). "The rise in ocean plastics evidenced from a 60-year time series". In: *Nature Communications* 10.1, pp. 1–6. DOI: 10.1038/s41467-019-09506-1.

Pabortsava, K. and R. S. Lampitt (2020). *Abundance and size of microplastic particles in filtered seawater samples collected on a north-south transect in the Atlantic Ocean during cruise AMT26 (JR16001), September–November 2016*. DOI: 10/d4q8.

Peacock, J. A. (1983). "Two-dimensional goodness-of-fit testing in astronomy". In: *Monthly Notices of the Royal Astronomical Society* 202.3, pp. 615–627. DOI: 10.1093/mnras/202.3.615.

Pedregosa, F. et al. (2011). "Scikit-learn: Machine learning in Python". In: *Journal of Machine Learning Research* 12, pp. 2825–2830.

Ragusa, A. et al. (2021). "Plasticenta: First evidence of microplastics in human placenta". In: *Environment International* 146, p. 106274. DOI: 10.1016/j.envint.2020.106274.

Rio, M.-H., S. Mulet, and N. Picot (2014). "Beyond GOCE for the ocean circulation estimate: Synergetic use of altimetry, gravimetry, and in situ data provides new insight into geostrophic and Ekman currents". In: *Geophysical Research Letters* 41.24, pp. 8918–8925. DOI: 10.1002/2014GL061773.

Rochman, C. M. et al. (2014). "Polybrominated diphenyl ethers (PBDEs) in fish tissue may be an indicator of plastic contamination in marine habitats". In: *Science of The Total Environment* 476, pp. 622–633. DOI: 10.1016/j.scitotenv.2014.01.058.

Spall, M. A. (2000). "Generation of strong mesoscale eddies by weak ocean gyres". In: *Journal of Marine Research* 58.1, pp. 97–116. DOI: 10.1357/002224000321511214.

Sprenger, J. and N. Weinberger (2021). "Simpson's Paradox". In: *The Stanford Encyclopedia of Philosophy*. Ed. by E. N. Zalta. Metaphysics Research Lab, Stanford University.

Talvitie, J. et al. (2017). "Solutions to microplastic pollution – Removal of microplastics from wastewater effluent with advanced wastewater treatment technologies". In: *Water Research* 123, pp. 401–407. DOI: 10.1016/j.watres.2017.07.005.

Turlach, B. A. (1993). "Bandwidth selection in kernel density estimation: A review". In: *CORE and Institut de Statistique*. Citeseer.

van Sebille, E., M. H. England, and G. Froyland (2012). "Origin, dynamics and evolution of ocean garbage patches from observed surface drifters". In: *Environmental Research Letters* 7.4, p. 044040. DOI: 10.1088/1748-9326/7/4/044040.

van Sebille, E. et al. (2015). "A global inventory of small floating plastic debris". In: *Environmental Research Letters* 10.12, p. 124006. DOI: 10.1088/1748-9326/10/12/124006.

van Sebille, E. et al. (2018). "Lagrangian ocean analysis: Fundamentals and practices". In: *Ocean Modelling* 121, pp. 49–75. DOI: 10.1016/j.ocemod.2017.11.008.

van Sebille, E. et al. (2020). "The physical oceanography of the transport of floating marine debris". In: *Environmental Research Letters* 15.2, p. 023003. DOI: 10.1088/1748-9326/ab6d7d.

Verbeek, M. (2005). *A Modern Guide to Econometrics*. Wiley.

Weston, P. (2019). *Plastic bags were created to save the planet, inventor's son says*. Last accessed: 16/01/2021. URL: https://www.independent.co.uk/climate-change/news/plastic-bags-pollution-paper-cotton-tote-bags-environment-a9159731.html.

Wichmann, D., P. Delandmeter, and E. van Sebille (2019). "Influence of near-surface currents on the global dispersal of marine microplastic". In: *Journal of Geophysical Research: Oceans* 124.8, pp. 6086–6096. DOI: 10.1029/2019JC015328.

Wichmann, D. et al. (2019). "Mixing of passive tracers at the ocean surface and its implications for plastic transport modelling". In: *Environmental Research Communications* 1.11, p. 115001. DOI: 10.1088/2515-7620/ab4e77.

Wilcox, C., B. D. Hardesty, and K. L. Law (2019). "Abundance of floating plastic particles is increasing in the Western North Atlantic Ocean". In: *Environmental Science & Technology* 54.2, pp. 790–796. DOI: 10.1021/acs.est.9b04812.

Woodall, L. C. et al. (2014). "The deep sea is a major sink for microplastic debris". In: *Royal Society Open Science* 1.4, p. 140317. DOI: 10.1098/rsos.140317.

Zettler, E. R., T. J. Mincer, and L. A. Amaral-Zettler (2013). "Life in the "plastisphere": Microbial communities on plastic marine debris". In: *Environmental Science & Technology* 47.13, pp. 7137–7146. DOI: 10.1021/es401288x.

Zhang, Y. et al. (2021). "Microplastics in glaciers of the Tibetan Plateau: Evidence for the long-range transport of microplastics". In: *Science of The Total Environment* 758, p. 143634. DOI: 10.1016/j.scitotenv.2020.143634.

Zhao, D. and M. Li (2019). "Dependence of wind stress across an air–sea interface on wave states". In: *Journal of Oceanography* 75.3, pp. 207–223. DOI: 10.1007/s10872-018-0494-9.