

The background of the slide is a painting. The top half shows a sky with dark, swirling blue and grey clouds. Below the sky is a horizon line. The bottom half of the painting depicts a beach or shoreline. The ground is a mix of brown, yellow, and orange tones, suggesting sand and rocks. Scattered across the foreground are numerous pieces of plastic waste, including a prominent blue rectangular fragment in the center, and other smaller, colorful pieces of debris. The overall style is expressive and somewhat somber, reflecting the theme of marine plastic pollution.

Mikael Kaandorp

Closing the marine plastic mass budget

Using observational data to
inform numerical models

Closing the marine plastic mass budget

Using observational data to inform numerical models

Mikael Kaandorp

February 2023

Copyright © 2023 by Mikael Kaandorp

Institute for Marine and Atmospheric research Utrecht (IMAU)
Faculty of Science, Department of Physics, Utrecht University.
Princetonplein 5, 3584 CC Utrecht, The Netherlands.

Cover made using DALL-E: *an abstract oil painting of a beach with plastic litter.*

ISBN: 978-94-6419-741-9
Printed by Gildeprint



Utrecht University

Closing the marine plastic mass budget

Using observational data to inform numerical models

Een sluitende massabalans voor plastics in de oceaan

Het gebruik van observationele metingen om numerieke modellen te informeren

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof. dr. H.R.B.M. Kummeling, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op dinsdag 18 april 2023 des middags te 2.15 uur

door

Mikael Lauri Alexander Kaandorp

geboren op 24 april 1993

te Amsterdam

Promotoren:

Prof. dr. E. van Sebille
Prof. dr. ir. H.A. Dijkstra

Beoordelingscommissie:

Prof. dr. H.E. de Swart
Prof. dr. G.J. Reichart
Prof. dr. M. Dijkstra
Prof. dr. A.A. Koelmans
Prof. E. Chassignet

This work was conducted within the TOPIOS (Tracking Of Plastic In Our Seas) project and supported through funding from the European Research Council (ERC) under the European Union Horizon 2020 research and innovation programme (grant agreement No 715386).

Contents

| | Page |
|---|-----------|
| Samenvatting | v |
| Main Matter | |
| 1 Introduction | 1 |
| 1.1 Sources of marine plastics | 3 |
| 1.2 Sinks of marine plastics. | 4 |
| 1.2.1 Biofouling | |
| 1.2.2 Beaching | |
| 1.2.3 Degradation and fragmentation | |
| 1.3 Modelling transport of plastic particles. | 9 |
| 1.4 Combining models with observational data. | 12 |
| 1.5 Research questions and thesis outline. | 16 |
| 2 Estimating the mass budget of Dutch coastal litter | 19 |
| 2.1 Introduction | 20 |
| 2.2 Methods | 21 |
| 2.2.1 Data description and region of interest | |
| 2.2.2 Data preprocessing | |
| 2.2.3 Model | |
| 2.3 Results and discussion. | 32 |
| 2.3.1 Regression analysis | |
| 2.3.2 Spatial variability | |
| 2.3.3 Extrapolating litter quantities to the entire coastline | |
| 2.4 Conclusions and recommendations | 39 |
| 3 Modelling size distributions of fragmented marine plastics | 43 |
| 3.1 Introduction. | 44 |
| 3.2 Methods | 46 |
| 3.2.1 The cascading fragmentation model | |
| 3.2.2 Environmental box model | |

| | |
|--|----|
| 3.2.3 Applying the box model to the Mediterranean Sea | |
| 3.3 Results and discussion. | 53 |
| 3.3.1 Modelled environmental particle size distributions | |
| 3.3.2 Model limitations | |
| 3.3.3 Fragmentation models and size distribution data | |
| 3.4 Conclusions and recommendations | 61 |

4 Closing the Mediterranean mass budget of buoyant marine plastics 63

| | |
|---|----|
| 4.1 Introduction. | 64 |
| 4.2 Methods | 65 |
| 4.2.1 Lagrangian framework and forcing | |
| 4.2.2 Area of interest and field measurements used | |
| 4.2.3 Sources of plastics | |
| 4.2.4 Parameterisation of plastic particle properties | |
| 4.2.5 Sinks of plastics | |
| 4.2.6 Inverse modelling | |
| 4.3 Results and discussion. | 73 |
| 4.3.1 Parameter estimation | |
| 4.3.2 Mass balance | |
| 4.4 Conclusions and recommendations | 77 |

5 Closing the global mass budget of buoyant marine plastics 79

| | |
|---|----|
| 5.1 Introduction. | 80 |
| 5.2 Methods | 81 |
| 5.2.1 Data assimilation | |
| 5.2.2 Lagrangian model | |
| 5.2.3 Transition matrix model | |
| 5.3 Results and discussion. | 87 |
| 5.3.1 A 3D map of marine plastic litter | |
| 5.3.2 Particle size contributions to the marine plastic mass budget | |
| 5.4 Conclusions and outlook. | 91 |

6 Summary and Outlook 93

| | |
|--|----|
| 6.1 Answers to the research questions | 93 |
| 6.2 Discussion and outlook | 96 |
| 6.2.1 Combining numerical models with observational data | |
| 6.2.2 Observational data | |

Back Matter

| | |
|--------------------|------------|
| Appendices: | 102 |
|--------------------|------------|

A Supporting material Chapter 2 103

| | |
|--|-----|
| A.1 Observational and modelled data per year | 103 |
|--|-----|

| | | |
|----------|---|------------|
| A.2 | Extended results. | 103 |
| A.2.1 | Gini importance overview | |
| A.2.2 | Excluding Lagrangian model features | |
| A.2.3 | Effect of using only the top N features | |
| A.2.4 | Number of participants | |
| A.2.5 | Feature effect | |
| A.3 | Clustering dendrogram | 110 |
| A.4 | Model pipeline. | 110 |
| B | Supporting material Chapter 3 | 115 |
| B.1 | Additional results: parameter settings and Mass Size Distributions (MSDs). | 115 |
| B.2 | Box model: transition matrix formulation. | 119 |
| B.3 | Vertical mixing and size-dependent lateral transport | 121 |
| B.4 | Estimated municipal plastic waste properties | 125 |
| B.5 | Successive fragmentation model | 126 |
| C | Supporting material Chapter 4 | 129 |
| C.1 | Measurement corrections and variance | 129 |
| C.2 | Beaching time scale estimate. | 130 |
| C.3 | Inverse modelling implementation. | 132 |
| C.4 | Parameter estimation sensitivity study | 132 |
| C.5 | Model-measurement scatter plots and mass correction factor. . . | 135 |
| C.6 | Which sinks are neglected and why? | 139 |
| D | Supporting material Chapter 5 | 143 |
| D.1 | Observational data and model performance | 143 |
| D.2 | Supplementary results | 146 |
| D.3 | Supplementary methodological information. | 153 |
| | References | 159 |
| | Acknowledgments | 179 |
| | List of Publications | 181 |
| | Curriculum Vitae | 183 |

Samenvatting in het Nederlands

Plastics worden toegepast voor vele doeleinden, doordat het goedkoop te produceren is, lang meegaat en het bijvoorbeeld luchtdicht, waterdicht en steriel kan zijn. Doordat plastic zo lang meegaat ondervinden we echter veel vervuiling van deze materialen, onder meer in de oceaan. Veel plastic heeft een lagere dichtheid dan water en blijft dus drijven. Een inmiddels bekend fenomeen zijn de ‘drijvende vuilnisbelten’, in het Engels ‘floating garbage patches’ genoemd: dit zijn gebieden in het midden van de oceaan waar hoge concentraties aan plastic deeltjes te vinden zijn. Deze gebieden bevinden zich rond de subtropen: één van de bekendste drijvende vuilnisbelten is te vinden in het noorden van de Stille Oceaan, maar ook in het noorden van de Atlantische Oceaan is er bijvoorbeeld één te vinden.

Het is nog vrij onbekend waar het meeste plastic afval in de oceaan precies vandaan komt: komt het bijvoorbeeld vooral uit rivieren, vanaf de kustlijn, of van visserijen? Verder is het nog vrij onbekend welke processen ervoor zorgen dat plastic deeltjes ‘verdwijnen’ uit het oppervlak van de oceaan. Zo kan een deel van het plastic afval bijvoorbeeld aanspoelen op stranden, na een tijdje naar beneden zinken (bijvoorbeeld door aangroei van algen), of uit elkaar brokkelen in kleinere plastic stukjes. Hoeveel plastic afval er in de oceaan terechtkomt en hoeveel er na verloop van tijd verdwijnt is een groot vraagstuk: huidige schattingen van de hoeveelheid drijvend plastic afval in de oceaan zijn namelijk veel lager dan de geschatte hoeveelheid plastic afval dat elk jaar in de oceaan terecht komt. Dit proefstuk is geschreven in het kader van het TOPIOS project (Tracking Of Plastics In Our Seas), waarin we de massabalans voor plastic afval in de oceaan willen sluiten.

In dit proefschrift dragen we bij aan een sluitende massabalans voor plastic afval in de oceaan, door observationele metingen te combineren met numerieke (computer) modellen. Er zijn verschillende manieren om dit te doen. De afgelopen jaren zijn statistische methoden als machine learning steeds populairder geworden om voorspellingen mee te doen. In hoofdstuk 2 van dit proefschrift gebruiken we een regressiemethode (‘random forests’) om te voorspellen hoeveel afval er aanspoelt op de Nederlandse Noordzeekust. We maken hierbij gebruik van een dataset van Stichting De Noordzee, die in de jaren 2014–2019 schoonmaakacties hebben verricht op Nederlandse stranden en hebben bijgehouden hoeveel afval er bij deze acties werd opgeruimd. Wij koppelen de gemeten hoeveelheden afval aan de weers- en zeecondities rondom de schoongemaakte stranden: denk hierbij aan bijvoorbeeld de wind-

kracht en windrichting, het getij, de zeestromingen en de oriëntatie van de kustlijn. Met het regressiemodel voorspellen we vervolgens hoeveel afval er gemiddeld langs de hele kustlijn verwacht wordt: rond de 16 500–31 200 kilogram. We voorspellen gemiddeld genomen relatief hoge concentraties strandafval rond Egmond en Bergen en relatief lage concentraties in Zeeland.

Veel van de plastic deeltjes die in de oceaan worden gevonden zijn slechts enkele millimeters groot en zijn fragmenten van grotere plastic objecten. Er zijn echter niet veel modellen in de literatuur te vinden die het fragmentatieproces van plastic deeltjes beschrijven. In hoofdstuk 3 ontwikkelen we een nieuw model dat beschrijft hoe plastic deeltjes in de loop van tijd fragmenteren. Hiermee modelleren we de lengte van plastic fragmenten in de oceaan en op stranden. We laten zien dat het meten van de lengte van plastic deeltjes ons mogelijk informatie kan geven over bijvoorbeeld de tijdschaal waarop fragmentatie van plastic deeltjes plaatsvindt.

Het zijn niet alleen de ‘drijvende vuilnisbelten’ in het midden van de oceaan waar de concentratie van plastic deeltjes hoog is. Ook in gebieden als de Middellandse Zee, omgeven door kustlijn, is dit het geval. In hoofdstuk 4 stellen we een massabalans op voor drijvend plastic afval in de Middellandse Zee. We definiëren mogelijke bronnen voor plastic afval, mogelijke processen die plastic deeltjes verwijderen uit het wateroppervlak (het aanspoelen en zinken van plastic deeltjes) en kalibreren dit model vervolgens met metingen uit de literatuur van concentraties plastic aan het wateroppervlak (door middel van planktonnetmetingen). We gebruiken dit gekalibreerde model om een schatting te geven van de belangrijkste bronnen van plastic vervuiling in de Middellandse Zee, en waar het meeste plastic afval terecht komt.

Ten slotte doen we een soortgelijke analyse in hoofdstuk 5, maar dan op een globale schaal, met een complexer model en meer metingen. Net als in hoofdstuk 4 kalibreren we parameters in een numeriek model, gerelateerd aan bronnen, transport en processen die plastic deeltjes verwijderen uit het wateroppervlak. We doen deze kalibratie door middel van data-assimilatie, waar een groep modellen met verschillende parameterwaarden wordt vergeleken met observationele metingen. Hiervoor gebruiken we netmetingen van plastic concentraties aan het wateroppervlak en dieper in de waterkolom, en metingen van concentraties plastic op stranden. Een belangrijke conclusie uit dit hoofdstuk is dat de hoeveelheid plastic massa in de oceaan (ongeveer 3 000 000 ton) waarschijnlijk veel hoger is dan voorheen geschat (250 000 ton): mogelijk zagen voorgaande studies grotere plastic objecten over het hoofd, die bijdragen aan het grootste deel van de totale massa. Verder schatten we de bronnen lager in dan voorgaande studies, en verwachten we dat plasticvervuiling een langdurig probleem zal zijn, zelfs als er plotseling geen nieuw plastic afval in de oceaan terecht zou komen.

Introduction

In the 20th century, fossil fuels revolutionized all industry sectors by serving as the source for synthetic plastics, making it possible to create the durable, cheap, and sterile objects that are used in our daily lives. Harmful side-effects of this ‘Plastic Revolution’ are directly observable, as the highly durable plastic objects started to accumulate in the marine environment. Some of the earliest scientific accounts of marine plastic pollution come from the seventies, from measurement campaigns finding high concentrations of plastic particles in the North Atlantic Ocean (Carpenter & Smith Jr., 1972; Colton et al., 1974; Morris, 1980) and the North Pacific Ocean (Wong et al., 1974; Shaw & Mapes, 1979). More studies followed in the eighties, some noting that marine life was getting entangled in plastic items such as fishing gear, and suggesting that the presence of plastics in the stomach of animals might be a cause of mortality (Fowler, 1987; Ryan, 1987). In the late nineties marine plastic pollution started to gather more attention, as more observations in the North Pacific showed large plastic concentrations in the subtropical gyre (Moore et al., 2001), an area that would later be called the ‘Great Pacific Garbage Patch’ (Young et al., 2009; Lebreton et al., 2018). The direct visibility of plastics as one type of anthropogenic pollution led to worldwide media attention, awareness campaigns, and cleanup initiatives over the years.

The North Pacific subtropical gyre is one of the 5 subtropical gyres in which floating matter accumulates at the ocean surface. Large-scale regions of downwelling at the ocean surface are driven by the Ekman currents, caused by the surface winds and the Coriolis effect (Kubota, 1994). Due to the fact that most plastic particles are buoyant and remain at the ocean surface, they accumulate within these regions. The dispersion of plastic particles is furthermore affected by the geostrophic ocean

currents, caused by pressure gradients (variations in sea surface elevation) and the Coriolis effect. Although the geostrophic currents are nondivergent and do not lead to accumulation on their own, they have a strong impact on the overall accumulation patterns (Onink et al., 2019). Plastic concentrations at the ocean surface as estimated in van Sebille et al. (2015) are presented in Figure 1.1. Clearly visible are the five subtropical gyres: two gyres in the Northern hemisphere (the Pacific and Atlantic), and three in the Southern hemisphere (the Pacific, Atlantic, and Indian Ocean). Besides high plastic concentrations in the subtropical gyres, a large amount of plastic particles can be found in enclosed basins such as the Mediterranean Sea, and relatively high concentrations are predicted around coastlines in South-East Asia and around the Barents Sea. The plastic concentrations as predicted in van Sebille et al. (2015) reach values of more than one million particles per square kilometer. These high numbers, combined with photos of marine plastic pollution in the media, led to the persistent myth of literal plastic islands or even plastic continents in the middle of the oceans. With a mean mass of roughly 0.01 grams per particle (van Sebille et al., 2015) this is far from the reality, however: most of the plastic items that are found in the ocean are millimeter-sized plastic fragments (Cózar et al., 2014).

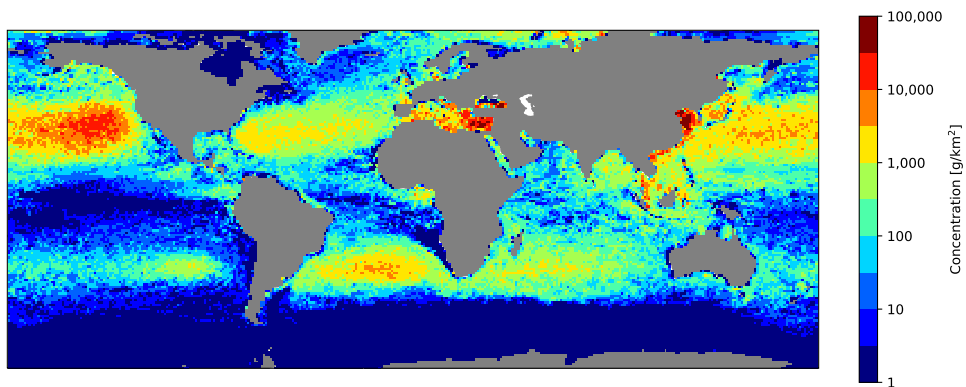


Figure 1.1: Modelled plastic mass concentrations at the ocean surface from van Sebille et al. (2015). High concentrations can be observed in the five subtropical gyres, as well as in the Mediterranean Sea, around the Barents Sea, and near the coastlines of South-East Asia. Figure adapted from van Sebille et al. (2015).

Surprisingly, even though marine plastic pollution is easily observable, little is known about the sources of the pollution: how much of the plastic waste that we produce is littered, how much of this enters the ocean, and where? Additionally, little is known about the sinks: are there processes that remove the plastic pollution from the marine environment, and on what time scales is this happening? These issues are directly related to a long-standing question in the marine plastic research community: what happens with the majority of the plastics that enter the ocean (Thompson et al., 2004)? This question gained traction due to the fact that the amount of plastics estimated to float at the ocean surface (in the order of hundreds of thousands metric tonnes; Eriksen et al., 2014; van Sebille et al., 2015) is roughly two orders of

magnitude smaller than the estimated amount of plastics entering the ocean every year (in the order of tens of millions metric tonnes; Jambeck et al., 2015; Borrelle et al., 2020).

The work presented in this thesis is part of the TOPIOS project: Tracking Of Plastic In Our Seas. The aim of this project is to make advances in closing the plastic budget, by creating a novel modelling framework that tracks plastic transport in the ocean. In this thesis we contribute to this project, by using observational data of plastics in the marine environment to inform numerical models. By doing this, we make an effort in obtaining a better understanding of marine plastic sources, pathways, and sinks. This is important to make future prevention strategies and cleanup efforts more effective. By knowing where and when plastic litter accumulates in the marine environment, we can make more targeted removal efforts, and get a clearer overview of potential impacts on marine ecosystems. By knowing where and when plastic litter enters the marine environment, policymakers can target the sources that are responsible for a large part of the plastic problem, and mitigation strategies can be developed to remove plastic pollution before it actually enters the ocean and can cause harm to marine ecosystems.

1.1 Sources of marine plastics

The production of plastic items has grown rapidly the past couple of decades, from roughly 50 million metric tonnes per year in the 1970's, to an estimated 400 million metric tonnes in 2015. In a similar fashion, the plastic waste generation has been increasing by roughly 5% per year to 300 million metric tonnes in 2015. Only a small part (9%) of this waste is estimated to be recycled, with a large part (79%) likely ending up on landfills or in the natural environment (Geyer et al., 2017). The term 'Mismanaged Plastic Waste' (MPW) is used to define the plastics that are available to enter the natural environment, for example via littering, or by ending up on uncontained open dumps (Jambeck et al., 2015; Geyer et al., 2017).

How much of this mismanaged plastic waste actually enters the ocean is an open question, and the quantities and pathways vary widely between different studies. A part of the mismanaged plastic waste could enter the oceans along coastal regions. Jambeck et al. (2015) estimated that 4.8 million to 12.7 million metric tonnes of plastic enter the ocean every year this way. Plastic waste can be transported from inland regions towards the oceans via rivers. These estimates range widely, from 0.8 million to 2.7 million metric tonnes (Meijer et al., 2021), to 6 kilotonnes per year (Weiss et al., 2021). We will come back later to the large discrepancies between these studies in Chapter 5. Borrelle et al. (2020) estimated 19 million to 23 million metric tonnes of plastic to enter aquatic ecosystems every year (i.e. oceans, lakes, and major rivers combined). Fishing activity is thought to be a major source of plastic pollution as well. On Dutch beaches, about 40% of the plastic items are thought to originate from the fishing industry, which we will come back later on in Chapter 2. In global scenarios, the fishing industry is thought to contribute to about 18% of the total plastic litter (Lebreton et al., 2018).

Input via coastal regions, rivers, and fishing activity are estimated to be the three major sources of marine plastic pollution (Lebreton et al., 2018). It is very difficult to put exact percentages on their contributions. First, they might vary region to region. Secondly, whereas fishing related items are sometimes possible to identify (e.g. fishing nets or buoys), this is much harder for coastal plastics and riverine plastics. One of the objectives in this thesis is to attempt to put numbers to the source percentages. In Chapter 5, Figure 5.1 we show the three main plastic source categories on a global map. The three sources are distributed differently on the map, with for example relatively high expected coastal plastic input in South-East Asia (Jambeck et al., 2015), and a relatively high amount of fishing hours in the North Sea (Kroodsma et al., 2018). We will use numerical models to predict how plastic items from these different sources will disperse through the ocean, see Section 1.3. By varying the magnitude of the different sources in our numerical models, we expect to change the global distribution of marine plastic pollution. In Chapter 4 and Chapter 5, we calibrate the marine plastic sources in such a way that we achieve the best possible model fit to observational data.

1.2 Sinks of marine plastics

There are different mechanisms that are hypothesized to remove plastics from the surface water of the ocean or the marine environment altogether, which could explain the mismatch between the estimated plastic input and the amount of plastics floating in the ocean. We will briefly discuss some of these: sinking and sedimentation due to biofouling, plastics ending up on coastlines (beaching), degradation, and fragmentation. We will refer to these processes as ‘sinks’ in this thesis, understanding that this term is dependent on the time scale and particle size that is considered. For example, plastic particles might get buried on beaches or end up in marine sediments. It is still unknown how long plastic particles actually remain in these reservoirs, which can make them temporary sinks when considering long time scales. Fragmentation can act as a sink for marine plastics, by breaking down particles below typically measured sizes.

1.2.1 Biofouling

The majority of plastics that are produced are expected to float in sea water, since they are created from polymers with a density lower than sea water, or due to their foamed structure. These buoyant polymers include polyethylene (PE), polypropylene (PP), and polystyrene (PS), and are expected to make up roughly 60% of all plastics entering the marine environment (Andrady, 2011; Lebreton et al., 2018). Even though many plastics are initially buoyant, a biofilm can develop on the surface of the particles that is denser than seawater. This can make plastic items negatively buoyant on time scales of just weeks to months (Lobelle & Cunliffe, 2011; Fazey & Ryan, 2016). An example of biofouled plastic particles from the experi-

mental study in Lobelle & Cunliffe (2011) are shown in Figure 1.2. The larger the surface area to volume ratio, the more biofilm can attach per volume of plastic. In Fazey & Ryan (2016) larger surface to volume ratios were therefore observed to lead to shorter sinking time scales of polyethylene samples. Field data supports the theory that biofouling can indeed make initially buoyant particles to sink down, as for example Egger et al. (2020b) found high percentages of polyethylene and polypropylene (i.e. initially buoyant) particles in the deep North Pacific waters. The microbial community that lives on plastic items and forms the biofilm has been called the ‘plastisphere’ in the literature (Amaral-Zettler et al., 2020). These communities form highly complex ecosystems that can vary particle to particle, and change over time. Diatoms and Rhodobacteraceae are commonly found in biofilms, and are thought to be early and one of the dominant colonizers (Carson et al., 2013; Amaral-Zettler et al., 2020). It is possible that microbes by themselves are not enough to make all types of plastic items sink: Amaral-Zettler et al. (2021) found that particles with a low surface to volume ratio might require the attachment of multicellular organisms before getting negatively buoyant, such as bivalves, barnacles, and worms.

In Kooi et al. (2017) biofouling was hypothesized to lead to interesting dynamical behaviour of plastic particles in the water column. As plastic particles start sinking due to biofouling, the light intensity and water temperature start to decrease, slowing down the growth of algae. Due to algae mortality and respiration, the particles might start to become positively buoyant again, leading to a vertical oscillatory movement of plastic particles. The idealized model by Kooi et al. (2017) was later extended to more realistic scenarios in the ocean in Lobelle et al. (2021) and Fischer et al. (2022). We implement the model by Fischer et al. (2022) in Chapter 5 to capture the effect of biofouling on vertical mixing of plastic particles in the water column.

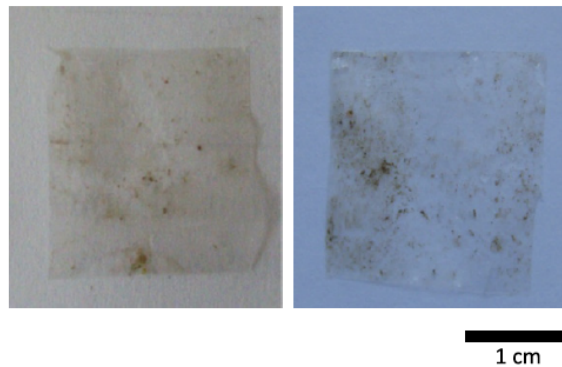


Figure 1.2: Sheets of polyethylene plastic that have been submerged in seawater for 2–3 weeks (Plymouth, United Kingdom) (Lobelle & Cunliffe, 2011). Photo courtesy of Delphine Lobelle.

1.2.2 Beaching

Coastlines are thought to be an important reservoir for marine plastic pollution (Isobe & Iwasaki, 2022). Onink et al. (2021) estimated that the majority (67-77%) of plastics reside on beaches or in coastal waters up to 10 km offshore. Anecdotal evidence that plastic concentrations are high on beaches can be obtained when going to any beach in the Netherlands. When walking along the beach at the location where floating matter has accumulated during the latest flood, one is very likely to find items such as fishing nets, bottle caps, and packaging. It is likely one can find more plastic mass on a couple of square meters of beach compared to the open ocean, where it is typical to find a lot of small plastic fragments. Some example photographs are shown in Figure 1.3. Given the possibility that beaches play an important role as a reservoir for marine plastics, we will take a closer look at beached plastics in the Netherlands in Chapter 2.

Estimating plastic concentrations on beaches remains a difficult task. One challenge is the fact that the hydrodynamics in coastal waters are highly different from those in the open ocean, with the (increased) influence of for example (breaking) waves, tides, and coastline morphology (van Sebille et al., 2020). Some studies have tried to link observed beach plastic concentrations to conditions related to the environment and land usage: for example looking at distance to nearby population centers (Ryan et al., 2018; Olivelli et al., 2020), looking at the wind speed and ocean current components towards the land (Eriksson et al., 2013; Thepwilai et al., 2021), and looking at local tide and wave conditions (Pawlowicz, 2020; Williams & Tudor, 2001). Other experimental studies have estimated how long it takes for plastic particles to move from coastal waters onto the beach (Pawlowicz, 2020), and how long particles typically remain on beaches before washing back into the ocean (resuspension), parameterising these processes using time scales (Hinata et al., 2020). These kind of parameterisations can be informed by mark-recapture experiments, where a large amount of easily recognizable items are released on the beach or in the coastal water, after which the amount of these items on the beach is kept track of over time (Hinata et al., 2017). It is very difficult to estimate how all these different studies generalize to other regions, since their underlying methodologies, models, considered variables, and amount of data differ.

Another challenge in estimating beach plastic concentrations is the lack of standardisation of measurements: studies differ in considered particle sizes, perform measurements per unit length or per unit area, use different units (number of particles versus mass), vary in the length/width of beach that is sampled, vary in the use of replicates, and use different platforms/formats to provide the data. Two examples of large-scale measurement datasets are the OSPAR dataset from Europe (OSPAR, 2010), and the MDMAP dataset from the United States (Burgess et al., 2021). Even though these datasets are similar in many ways, one can not directly compare the resulting plastic concentrations between both datasets, since different particle sizes are considered (all sizes and >50 cm in OSPAR data, versus >2.5 cm in MDMAP data). Both datasets only consider the number of plastic particles as opposed to weight. This can introduce measurement error into the datasets, as the number of



(a) Microplastics filtered from $\sim 900 \text{ m}^2$ (b) Plastics located on the beach of Texel, the surface water of the North Atlantic Ocean. Netherlands

Figure 1.3: Plastic concentrations can be relatively high on coastlines compared to the open ocean: a couple of square meters of Dutch beach yield more plastic mass compared to roughly 900 m^2 in the North Atlantic subtropical gyre in this anecdotal example. *Own material*

plastic particles will be heavily dependent on the smallest items that cleanup participants will look for, which we will further discuss in Chapter 5.

1.2.3 Degradation and fragmentation

Looking at Figure 1.3, it is clear that degradation and fragmentation of plastics in the marine environment is important to take into account: the majority of plastics particles found in the open ocean are small millimeter-sized fragments (Cózar et al., 2014). There are different mechanisms for plastic degradation: photodegradation by UV-light, thermal degradation by visible light, and thermal oxidation by infrared light (Amaral-Zettler et al., 2020). Furthermore, biological degradation is possible under the influence of enzymes: PET and PUR are susceptible for this type of degradation, PE and PP are likely less susceptible (Amaral-Zettler et al., 2020). It is commonly thought that degradation of plastics is dominant on beaches, due to higher temperatures, and more UV-radiation and oxidation (Efimova et al., 2018; Andrady, 2011). Degradation makes plastic particles more brittle, which enhances the fragmentation by mechanical abrasion (Song et al., 2017). Studies have shown that some polymer types such as polystyrene could be sensitive to photochemical oxidation into carbon dioxide and dissolved organic carbon. Ward et al. (2019) estimated half-life times of polystyrene in the marine environment due to this photochemical oxidation in the order of decades. Polypropylene, and to a lesser extend polyethylene, have additionally been reported to oxidize into dissolved organic carbon in lab studies (Zhu et al., 2020).

Despite the clear effect of fragmentation on the number of observed plastic particles,

relatively little is known about the modelling of this process. Observational data show an increase in the number of particles for a decreasing particle size. Some of the earliest works looking closer at abundance-size distributions are C  zar et al. (2014, 2015), for which the measured data are shown in Figure 1.4. In C  zar et al. (2014) a fragmentation model was developed to explain the power law that is visible in the data for particles larger than approximately one millimeter, which is visible as a straight line on a log-log scale. This power law was modelled by hypothesizing scale-invariance of the fragmentation process: each time a plastic object fragments (e.g. the length of the fragments are half the original object length), the amount of fragments increases by a similar amount (e.g. an increase in abundance by 8 when considering a three-dimensional cubic object). The slope of the power law on a log-log scale would then directly be related to the spatial dimension of the fragmented object: e.g. two-dimensional sheet-like objects, or three-dimensional cube-like objects. A drop in abundance can be seen in Figure 1.4 for particles smaller than roughly one millimeter, which was hypothesized to be caused by the removal of plastic particles from the surface water by processes such as biofouling in C  zar et al. (2014). The fragmentation model from C  zar et al. (2014) has some limitations and inconsistencies, however. We will come back to this in Chapter 3, where we will present a new fragmentation model for marine plastics.

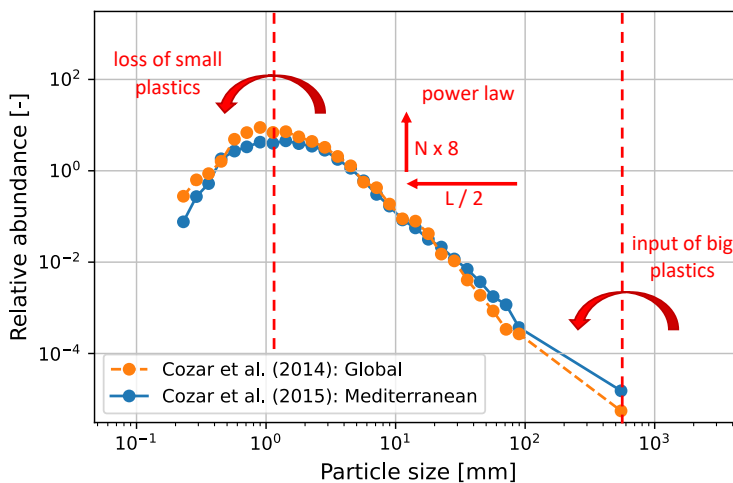


Figure 1.4: An increasing amount of particles can be found in the ocean surface for a decreasing particle size, down to about one millimeter. Two examples are shown with data from the global ocean and the Mediterranean Sea (C  zar et al., 2014, 2015). The amount of particles on the y-axis is normalized to allow for the comparison between both datasets. The hypothesized cause for the power law behaviour is furthermore illustrated for a three dimensional example (a cube-like object), where each time that the particle size decreases by a factor 2, the amount of fragments increases by a factor of 8.

1.3 Modelling transport of plastic particles

The dispersion of floating matter such as plastics can be studied in two different reference frames. The Eulerian reference frame is fixed in space, and particles are described in terms of concentrations. Dispersion of particles can be modelled by a tracer advection-diffusion equation as for example done in Mountford & Morales Maqueda (2019). One can also have a much more simplified representation of plastic transport in the marine environment by using box models. In Koelmans et al. (2017) a box model was used to represent the fragmentation of plastics in the global ocean, in Lebreton et al. (2019) this was extended to include transport globally between beaches, coastal waters, and the open ocean. In Chapter 3 we use a similar box model to represent fragmentation and transport of plastics in the Mediterranean Sea.

In a Lagrangian reference frame we follow individual particle trajectories. We can study the transport of particles by making use of Lagrangian drifter data, or by numerically integrating the trajectories of individual virtual particles.

More than 600 satellite-tracked Lagrangian drifters have been present in the global ocean in any given month since 1993 (Pazan & Niler, 2004), and about 1000 new drifters are deployed every year by the NOAA global drifter program (Lumpkin et al., 2016). One way to capture drifter transport statistics is by constructing stochastic matrices, which describe how drifters are transported between different grid cells of the ocean in terms of probabilities. These stochastic matrices have been used to simulate how floating matter such as plastic particles are transported in the global ocean, as was done in for example Maximenko et al. (2012) and van Sebille et al. (2012). This methodology was furthermore used in van Sebille et al. (2015) to estimate the total global mass of marine plastic pollution which was presented in Figure 1.1.

To construct the stochastic matrices, first the ocean is divided into a set of bins, e.g. by constructing an equidistant grid in the zonal and meridional direction. Each bin is given an unique index. The stochastic matrices P are then calculated by counting the amount of objects n moving to bin i from bin j , for a given time window Δt . The columns are furthermore normalized by dividing by the total amount of particles starting in bin j :

$$P_{ij} = \frac{n_{\text{to } i, \text{ from } j}}{n_{\text{from } j}} \quad (1.1)$$

As the matrix columns are normalized to 1, these stochastic matrices describe a Markov process with transition probabilities from bin j to bin i . The matrix P can then be used to model the evolution of a tracer concentration ρ over the time window Δt :

$$\rho(t + \Delta t) = P \rho(t) \quad (1.2)$$

For a more elaborate overview of stochastic matrices in oceanography, and possible applications such as using them to find stationary densities and dispersion time scales see e.g. Wichmann (2021), Wichmann et al. (2019b), and van Sebille (2014). We will use stochastic matrices in this thesis in Chapter 5 to capture the transport of

plastic tracers in the global ocean.

In Lagrangian particle tracking models we numerically integrate the trajectories of virtual particles in the ocean. In this case we need an ocean velocity field, which is normally provided by an ocean general circulation model (OGCM). OGCMs generally resolve the primitive equations, containing an approximation of the Navier-Stokes equations for the rotating earth (using the Boussinesq approximation and hydrostatic approximation), the continuity equation, heat and salt conservation equations, and an empirically determined equation of state relating sea water density to temperature, salinity, and pressure (Randall & Zehnder, 2001). The ocean velocity fields used in this thesis are derived from the NEMO primitive equation ocean model (Madec & the NEMO team, 2017). We use ocean reanalysis products in this thesis, which means that the OGCM has been assimilated to optimally match historical observational data. This can for example be historical data of sea surface height and sea surface temperature, from satellite or in-situ observations.

The evolution of a tracer with concentration ρ in the ocean can be described using the advection-diffusion equation:

$$\frac{\partial \rho(x, t)}{\partial t} = -\mathbf{v}(x, t) \cdot \nabla \rho(x, t) + \nabla \cdot (\mathbf{K} \cdot \nabla \rho(x, t)), \quad (1.3)$$

under the influence of a resolved velocity component \mathbf{v} , and a diffusivity tensor \mathbf{K} which models the effect of unresolved (sub-grid) scale velocity components. This tracer equation can be rewritten in terms of a Fokker-Planck (or Kolmogorov forwards) equation, describing the time evolution of the tracer probability density function (van Sebille et al., 2018). These equations are in an Eulerian reference frame that is fixed in space.

In the Lagrangian reference frame, a cloud of Lagrangian particles with stochastic trajectories can be used to approximate the evolution of the (Eulerian) Fokker-Planck equation. The Lagrangian particles are integrated according to a stochastic ordinary differential equation, where they experience a deterministic drift due to currents acting on scales larger than the grid size, as well as a stochastic forcing capturing uncertainty introduced by sub-grid processes. The deterministic drift is relatively trivial, where we can use for example a Runge-Kutta ordinary differential equation solver to solve for the particle advection. Different methods exist to model the stochastic forcing. In zeroth-order Markov models, stochastic noise is added to the particle positions. In first-order Markov models, stochastic forcing is added to the particle velocities. When one is interested in the dispersion of particles within relatively short time scales (e.g. days or even weeks) Markov-1 models can be more suitable (Reijnders et al., 2022). In this thesis, we are interested in the evolution of plastic tracers in the order of years or longer, in which case Markov-0 models are suitable and the easiest to implement. The stochastic differential equation that is used in this thesis to integrate Lagrangian particle trajectories that remain at the surface ocean is given by:

$$\mathbf{x}(t + \Delta t) = \mathbf{x}(t) + \int_{t+\Delta t}^t \mathbf{v}(\mathbf{x}, \tau) d\tau + R\sqrt{2K\Delta t}, \quad (1.4)$$

where \mathbf{x} are the Lagrangian particle locations, R is a random number from the normal distribution with zero mean and unit variance, and K is the tracer diffusivity. In this thesis and many other Lagrangian particle studies (e.g. Delandmeter & van Sebille, 2019) the diffusivity is assumed to be uniform in space and time, where the value depends on the grid resolution (Neumann et al., 2014). When K is not uniform, the random term in Eq. (1.4) needs to be modified. Non-uniform diffusivity fields can for example be estimated by considering drifter buoy trajectories and their separation over time (see e.g. Rühs et al., 2017), which is out of scope for this thesis. The grid resolution of the ocean circulation models from which the ocean velocity components are obtained in this thesis are $1/12^\circ$ and finer. Most turbulent dispersion should therefore already be accounted for, as mesoscale eddies, containing most of the oceanic kinetic energy, are resolved (Ferrari & Wunsch, 2009; Reijnders et al., 2022).

The ocean velocity data obtained from ocean circulation models generally contain the effects of the Ekman currents and the geostrophic currents. Additional transport can be induced on particles at the ocean surface by surface waves (Stokes drift) and wind drag (Kubota, 1994; Cunningham et al., 2022). These components can be added to the deterministic drift term in Eq. (1.4). Wind drag can be important for very light plastics such as foamed polystyrene (Chubarenko et al., 2016). Wind drag is often neglected, however, assuming that most plastic objects float just below the water surface (Lebreton et al., 2012; Macias et al., 2019). This is supported by modelling studies not finding an increased correlation between model output and measurements of plastic concentrations when adding wind drag (Lebreton et al., 2018). Many studies superimpose a field containing the Stokes drift onto the Ekman and geostrophic current fields. In Onink et al. (2019), no closer agreement between observations and model output was observed when superimposing the Stokes drift onto the Ekman and geostrophic velocity components, explained by overestimation of the Stokes drift caused by non-independence of the Ekman and Stokes drift components. In Cunningham et al. (2022), it was shown that the interaction of Stokes drift and the Coriolis force in the upper-ocean boundary layer results in a unsteady flow that can counteract the effect of Stokes drift. This finding was still unpublished when implementing Chapter 5. Since simply adding Stokes drift to the velocity components does not increase the model-observation match, we do not add it as a separate component there.

Vertical displacement of plastic particles in the water column can have an effect on the overall accumulation patterns, due to differences in the ocean currents in different water layers (Wichmann et al., 2019a), or due to processes only acting near the ocean surface such as Stokes drift (Breivik et al., 2014). Vertical displacement in the water column can be caused by particles rising or sinking due to a difference in density with respect to the surrounding sea water, or by vertical turbulent mixing. We can add these vertical transport terms to the stochastic differential equation in Eq. (1.4), where the vertical turbulent mixing scheme from Onink et al. (2022a) is

used:

$$\begin{aligned} \mathbf{x}(t + \Delta t) = & \mathbf{x}(t) + \int_{t+\Delta t}^t \mathbf{v}(\mathbf{x}, \tau) d\tau + R\sqrt{2K\Delta t} \\ & + \mathbf{w}_b(t)\Delta t + (\partial \mathbf{K}_z(\mathbf{x}, t)/\partial z) \Delta t + R\sqrt{2\mathbf{K}_z(\mathbf{x}, t)\Delta t} \end{aligned} \quad (1.5)$$

where \mathbf{w}_b is the particle sinking velocity induced by (negative) buoyancy, and \mathbf{K}_z is the vertical diffusivity field which can be obtained from the ocean circulation model. The particle rise velocity is usually estimated by assuming a certain shape and size (e.g. sphere or ellipsoid; Poulain et al., 2019), and by calculating the difference in density with respect to the sea water. Different formulations for estimating the rise velocity exist, where we use the one by Dietrich (1982) in Chapter 5 after calibrating it to experimental data from Poulain et al. (2019). We furthermore allow the rise velocity to vary over time, due to the biofouling process changing the effective particle density and radius (Fischer et al., 2022).

1.4 Combining models with observational data

In this thesis, we will explore a range of different techniques that can be used to combine numerical models with observational data. The most appropriate method to combine numerical models with observational data depends on the quantity of observational data that are available, and the availability of accurate and reliable numerical models.

We show a scale of some available techniques in Figure 1.5. On the left end of the spectrum we have the case where a numerical model is available that is able to accurately capture the underlying process of the problem. For example, a numerical solver is available for a problem where three criteria are satisfied: 1) the underlying physics are well known, 2) all relevant scales of the problem can be resolved, and 3) the boundary and initial conditions are well known. Think for example a simplified flow case where the Navier-Stokes equations can be resolved up to the Kolmogorov microscales. If some observational data are available, these can be used to validate the model.

In practical circumstances it is hard to satisfy all these three criteria. In the middle of the spectrum we have a scenario where the (physical) models cannot meet all three criteria, but where more observational data are available. In this case, data assimilation methods are a good option. One well known example are numerical weather prediction models. These models are often non-linear, making them chaotic and sensitive to initial conditions. Data assimilation algorithms can be used to calibrate the initial conditions, such that the model output optimally match the available observational data. Afterwards, the model with the calibrated initial conditions can be used to provide a more accurate forecast.

Most applications of data assimilation relate to the state estimation problem: for the problem presented in this thesis the state vector could contain the plastic concentrations ρ . A numerical model \mathbf{m} is used to calculate how this state evolves over time, for example making use of the stochastic matrix presented in Eq. (1.2), or by solving the (Lagrangian approximation of the) advection diffusion equation in Eq. (1.3).

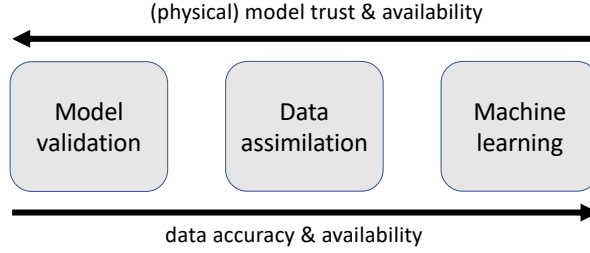


Figure 1.5: A trade-off between model and data availability is driving the selection of the most appropriate data-driven modelling method. Figure adapted, courtesy of Alberto Carrasi.

We know that the state ρ predicted by our model is not perfect. On the other hand we might have some observations \mathbf{d} : these can be relatively accurate (with a given amount of observational error), but sparse in space and/or time. We can use these observations to create an updated estimate of our state. Most data assimilation algorithms can be derived in some way from Bayes theorem, by which we can write the state estimation problem as:

$$f(\rho|\mathbf{d}) = \frac{f(\mathbf{d}|\rho) f(\rho)}{f(\mathbf{d})}, \quad (1.6)$$

i.e. we estimate the posterior probability density function f of our state ρ given the observations \mathbf{d} . In the nominator we multiply the likelihood, $f(\mathbf{d}|\rho)$, with the prior estimate of the state, $f(\rho)$. The term in the denominator acts as a normalization constant.

Data assimilation can also be used to estimate parameter values θ . This is relevant for this thesis: in our plastic dispersion models we have unknown parameters defining, for example, the time scales on which plastic particles start sinking down due to the biofouling process. Combined state and parameter estimation is also a possibility (Evensen, 2009). In that case, the set of parameters can simply be appended to the state vector, after which the data assimilation algorithm is used to estimate both simultaneously. We will denote the vector that contains the state ρ and/or the parameters θ as \mathbf{z} , consistent with Evensen et al. (2022) in which a clear and comprehensive overview of data assimilation and state estimation methods is given.

Some approaches, like Markov-Chain Monte-Carlo (MCMC) methods, try to sample the full posterior probability density function f in Eq. (1.6). To reliably sample the

full posterior probability density function, one needs a lot of model evaluations, however, which can be prohibitively expensive. Many data assimilation approaches do not try to estimate the full posterior probability density function in Eq. (1.6). When one assumes that the prior probability density function is a Gaussian, the maximum a posteriori (MAP) estimate, or the most probable solution for the state-parameter vector \mathbf{z} can be calculated (Evensen et al., 2022):

$$\mathbf{z}_{MAP} = \underset{\mathbf{z}}{\operatorname{argmax}} (f(\mathbf{z}|\mathbf{d})). \quad (1.7)$$

This problem can be solved by minimizing a cost function J that is given by:

$$J(\mathbf{z}) = \frac{1}{2} (\mathbf{z} - \mathbf{z}_{\text{prior}})^T \mathbf{C}_{zz}^{-1} (\mathbf{z} - \mathbf{z}_{\text{prior}}) + \frac{1}{2} (\mathbf{g}(\mathbf{z}) - \mathbf{d})^T \mathbf{C}_{dd}^{-1} (\mathbf{g}(\mathbf{z}) - \mathbf{d}), \quad (1.8)$$

where the first term compares the difference between the \mathbf{z} and its prior estimate $\mathbf{z}_{\text{prior}}$, weighted by the prior error covariance matrix \mathbf{C}_{zz} ; and where the second term compares the observations \mathbf{d} with the predicted output from the model $\mathbf{g}(\mathbf{z})$, weighted by the measurement error covariance matrix \mathbf{C}_{dd} . The term \mathbf{g} is called the observation operator, which maps the state vector with size N_z to the observation vector with size N_d : $\mathcal{R}^{N_z} \rightarrow \mathcal{R}^{N_d}$. Taking the example in this thesis: the observation operator would extract the plastic concentration from the state vector at the specific times and locations for which the corresponding measurements are available.

In variational data assimilation methods, gradients of Eq. (1.8) with respect to \mathbf{z} are calculated to solve the minimization problem, by using for example Gauss-Newton iterations (Evensen et al., 2022). This problem can be hard when \mathbf{z} is very large, and when no tools are available to automatically differentiate the model. Ensemble Kalman Filter methods are very popular in data assimilation, which approximately solve the minimization of Eq. (1.8) by using an ensemble of model runs. These methods are easy to implement since the model to be assimilated can be used as given, and work well for high-dimensional problems.

A part of this thesis consists of exploring what data assimilation and parameter estimation techniques are feasible for constructing plastic mass budget analyses. In Chapter 4 we use a MAP approach to estimate a set of parameters defining sources and sinks of plastic pollution, where Eq. (1.8) is minimized using a gradient based method. We furthermore explore whether MCMC methods are feasible and compare the results. In Chapter 5 we use an ensemble based approach called the Ensemble Smoother with Multiple Data Assimilation (ESMDA; Emerick & Reynolds, 2013). This method is similar to the Ensemble Kalman Filter, but works better for non-linear models.

Finally, on the right side of the spectrum in Figure 1.5 we have a scenario where it is hard to have an underlying numerical model for the problem at hand, and where we try to predict underlying processes purely based on observational data. These kind of problems can be resolved using statistical learning or machine learning techniques. In this thesis we will consider supervised learning, where an algorithm

is provided an input vector, a corresponding output vector, with the goal of creating a map between the two. More specifically, we consider regression algorithms, as opposed to predicting a finite amount of classes (classification). A broad range of machine learning algorithms are available, where a good overview is given in Hastie et al. (2008). A very basic example of a regression algorithm is linear regression, where a linear map is made from a explanatory variable, called a feature in machine learning (e.g. wind speed) to an observed quantity (e.g. plastic concentration). Decision trees and random forests are popular methods due to their robustness, and will be used in this thesis in Chapter 2. Artificial neural networks are probably the most commonly known machine learning technique, due to their high versatility and being able to capture highly complex relations, which is for example necessary for image processing applications.

In machine learning, the input vector typically consists of multiple (N_f) features. Taking the example that will be presented later in Chapter 2: these can be the wind speed and the height of the tide. For every observation (e.g. measured plastic concentration at a given location and time), the corresponding set of features are calculated. We then let the regression algorithm create the map from the input vector containing the features, to the output vector containing the observations: $\mathcal{R}^{N_f \times N_d} \rightarrow \mathcal{R}^{N_d}$. This can be done by minimizing a cost function very similar to Eq. (1.8). Where g in Eq. (1.8) was called the observation operator in the case of data assimilation, this is now a regression algorithm such as an artificial neural network, that takes the N_f features as input, and by optimizing its internal parameters (e.g. weights connecting the nodes of the artificial neural network) tries to reproduce the observations as well as possible. An illustration showing the similarities between machine learning and data assimilation is shown in Figure 1.6.

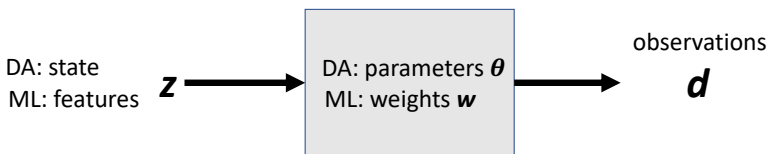


Figure 1.6: Data assimilation and machine learning can be quite similar in their set-up and available methods, but differ in terminology. In data assimilation, a state evolves over time as a function of model parameters θ , where the observation operator g maps the state to the observations d . In machine learning, a vector containing the features is transformed by e.g. a neural network with parameters (weights) w into the observations d . Both methods minimize a similar underlying cost function. Figure based on Geer (2021).

In machine learning, there is (typically) no underlying physical model, which means that the parameters (e.g. weights of an artificial neural network) have no direct physical meaning or interpretation. The parameters are therefore often not assigned a prior probability density function, which means that only the second term of the

cost function in Eq. (1.8) remains, and the weighing with C_{dd} can be ignored. There are exceptions, however, such as Bayesian neural networks where the neural network nodes and/or weights are given a prior probability density function (Jospin et al., 2022). Many machine learning algorithms furthermore employ regularization techniques (Hastie et al., 2008), which has a similar effect as assigning a prior probability density function to parameter values.

1.5 Research questions and thesis outline

In this thesis, we will focus on one methodological research question:

How can we use data-driven modelling to quantify the marine plastic mass budget, where we bring together numerical models with observational data?

Here, we define data-driven modelling as using data to derive the functional form or the parameters of a model.

This methodological question is used to answer an underlying scientific research question:

What processes cause the large mismatch between the expected marine plastic input and the estimated marine plastic standing stock? Is there a missing sink of marine plastics?

In **Chapter 2**, we start close to home, by looking at litter on beaches in the Netherlands. As touched upon in Section 1.2.2, concentrations of plastics are relatively high on coastlines. We will try, for the first time in the literature, to estimate the total amount of litter on Dutch beaches. We make use of a dataset provided by *Stichting De Noordzee*, who organize the *Boskalis Beach Cleanup Tour* every year, keeping track of the amount of litter that was gathered. Data of litter concentrations from the years 2014–2019 are used to train a random forest regression algorithm, which takes the environmental conditions at the location and time of the cleanup stage as input features. Afterwards, the regression algorithm is used to extrapolate the litter concentrations along the entire Dutch North Sea coastline. We will also investigate whether some particular environmental features play an important role in estimating the litter concentrations. As mentioned in Section 1.2.2 little is known about which physical and environmental parameters are the main drivers behind the beaching of litter. No physical model exists at the moment which capture all relevant parameters for the beaching process at an affordable computational cost. The dataset from *Stichting De Noordzee* is quite unique in the field of plastic research, because of the amount of datapoints, consistency, and underlying effort. We therefore use machine learning in this chapter, located on the right side of the data-driven spectrum in Figure 1.5.

Many of the plastics that are found in the marine environment are small millimeter-sized fragments. While developing the plastic source-sink-transport models, it turned out that existing fragmentation models for marine plastics were relatively limited,

and contained some inconsistencies. A new fragmentation model is presented in **Chapter 3**, capturing how plastic particles fragment over time. This model is based on existing fragmentation models capturing how rocks break down over time (Charalambous, 2015), which has been used to model and understand grain size distributions. We will use the fragmentation model to explain the shape of the plastic particle size distribution that was shown in Figure 1.4. We will furthermore analyze how different parameters in the fragmentation model, for example related to the fragmentation time scale and dimensionality of the plastic objects, affect this particle size distribution. This can give us information whether parameters such as the fragmentation time scale can be estimated based on observational data.

In **Chapter 4**, we develop the methodology used to estimate the marine plastic mass budget, by combining tracer transport models with observational data of plastic concentrations. We move towards the middle of the data-driven spectrum in Figure 1.5, where we use observational data to calibrate a set of parameters defining sources and sinks of marine plastics. To slightly simplify the problem, we limit the spatial domain and focus on the surface water of the Mediterranean Sea. After calibrating the set of parameters, we give an updated estimate on the total amount of plastic entering the Mediterranean Sea, and give an overview where most of the plastic pollution is expected to end up.

This analysis is extended to the global ocean in **Chapter 5**, where we consider longer time scales and furthermore increase the model complexity by adding biofouling dynamics, adding the fragmentation process from Chapter 3, increasing the amount of particle size classes, and increasing the amount of data from different marine reservoirs (surface ocean, deep ocean, and beaches). This calibrated model is used to answer what is most likely causing the large mismatch between the expected marine plastic input and the estimated amount of plastics at the ocean surface.

Finally, we end the thesis with a summary in **Chapter 6** by answering the research questions stated above, and provide an outlook on future research coming forth from our findings.

Estimating the mass budget of Dutch coastal litter using beach cleanup data and machine learning

Coastlines potentially harbor a large part of litter entering the oceans such as plastic waste. The relative importance of the physical processes that influence the beaching of litter is still relatively unknown. Here, we investigate the beaching of litter by analyzing a data set of litter gathered along the Dutch North Sea coast during extensive beach cleanup efforts between the years 2014–2019. This data set is unique in the sense that data is gathered consistently over various years by many volunteers (a total of 14,000), on beaches which are quite similar in substrate (sandy). This makes the data set valuable to identify which environmental variables play an important role in the beaching process, and to explore the variability of beach litter concentrations. We investigate this by fitting a random forest machine learning regression model to the observed litter concentrations. We find that especially tides play an important role, where an increasing tidal variability and tidal height lead to less litter found on beaches. Relatively straight and exposed coastlines appear to accumulate more litter. The regression model indicates that transport of litter

This chapter has been published as:

Kaandorp, M.L.A., Ypma, S.L., Boonstra, M., Dijkstra, H.A., van Sebille, E. (2022): Using machine learning and beach cleanup data to explain litter quantities along the Dutch North Sea coast. *Ocean Science*, 18(1):269-293. <https://doi.org/10.5194/os-18-269-2022>

through the marine environment is also important in explaining beach litter variability. By understanding which processes cause the accumulation of litter on the coast, recommendations can be given for more effective removal of litter from the marine environment, such as organizing beach cleanups during low tides at exposed coastlines. We estimate that 16,500–31,200 kilograms (95% confidence interval) of litter are located on the 365 kilometres of Dutch North Sea coastline.

2.1 Introduction

The accelerated release of mismanaged plastic waste into the global ocean gives rise to the need for effective cleanup strategies (Ogunola et al., 2018). In order to minimize the negative impact of plastic pollution on the environment, cleanup strategies need to be optimized to target the most impacted areas while limiting the economic costs (Haarr et al., 2019; Newman et al., 2015). Recent studies indicate that plastics remain trapped in coastal zones (Koelmans et al., 2017; Lebreton et al., 2019; Kaandorp et al., 2021; Morales-Caselles et al., 2021), with at least 77% of buoyant marine plastic debris beaching or floating in coastal waters (Onink et al., 2021). Therefore, beach cleanups have the potential to be a highly effective mitigation measure.

In addition, the plastic concentrations found on beaches are generally higher compared to other environmental compartments such as the surface water or the seafloor (Morales-Caselles et al., 2021), making beaches favorable locations for cleanup activities. Furthermore, by limiting the resuspension of plastic items by removal, the overall plastic concentration on the beach decreases over time and the formation of microplastic is reduced (Andrady, 2011; Haarr et al., 2020; Lebreton et al., 2019). At the same time, as cleanup activities generally involve a large number of volunteers, awareness of the plastic pollution problem increases, leading to a reduction of plastic waste in the local environment (Kordella et al., 2013).

Although the benefits of beach cleanups are well known, the location and timing of these activities are often not optimized. Haarr et al. (2019) identified accumulation zones of beached plastic using the shoreline curvature and gradient in the Lofoten, Norway, and showed that high-accumulation areas are often missed by cleanup actions. Other coastal properties like substrate and backshore type have been found to influence debris quantities as well (Hardesty et al., 2017; Brennan et al., 2018), with more litter accumulating in areas with increased backshore vegetation. Additionally, physical processes play an important role in the beaching of plastics and should be considered when selecting effective sites for beach cleanups.

However, the relative importance of the various physical processes involved and how these can be parameterised remains so far unknown (van Sebille et al., 2020; Pawlowicz, 2020). Studies have addressed the importance of the landward wind direction for debris accumulation rates (Eriksson et al., 2013; Critchell et al., 2015; Hengstmann et al., 2017; Moy et al., 2018), the landward ocean circulation direction (Thepwilai et al., 2021), the role of tides (Eriksson et al., 2013; Pawlowicz, 2020) and waves (Williams & Tudor, 2001). The spatial and temporal variability of the sources, e.g. rivers, population density and the fishing industry, also play an important role for the accumulation of plastic on beaches (Rech et al., 2014;

Critchell & Lambrechts, 2016).

In addition to the study by Haarr et al. (2019), there are several other studies that assess the prediction or monitoring of beached plastic items using machine learning methods. These algorithms can be useful in discovering complex relations between environmental variables and litter concentrations. In Granado et al. (2019), a marine litter forecasting model was made using Bayesian networks, involving various variables like wave height and period, wind velocity and directions, precipitation, and river flows. Neural networks have been used to quantify litter categories in Balas et al. (2004) and Schulz & Matthies (2014), and deep learning methods have been used to automatically identify debris on beaches (Song et al., 2021).

In order to make data-driven methods work, relatively large and consistent data sets are necessary, whereas most observational data is sparse. Beach cleanups and citizen science initiatives can potentially provide valuable information for scientific studies on marine pollution (Zettler et al., 2017), as these data are based on a considerable amount of person hours. Examples of citizen science data used in marine pollution research are e.g. Hidalgo-Ruz & Thiel (2013), where schoolchildren in Chile documented the distribution and abundance of plastic debris on beaches, and Ribic et al. (2010, 2012), where amounts of marine debris were measured by volunteer teams on beaches in the Pacific and Atlantic.

Here, we will build upon past data-driven studies by using an unprecedented data set obtained from beach cleanup efforts organized along the Dutch North Sea coast between 2014–2019. The number of participants (about 14,000), person-hours (about 84,000 hours), the length of beach sampled (about 1,400 kilometres) and the fact that all beaches sampled were similar in substrate (sandy), make this data set unique and very appropriate to apply data-driven methods. Furthermore, a large set of explanatory variables will be created, based on environmental conditions and modelled transport of marine litter. We will fit a random forest regression model to the observed litter concentrations as a function of these explanatory variables, and investigate which ones are important to explain the variability in beach litter. This allows us to investigate which variables are important predictors for the amount of litter present on beaches, to get a better understanding of marine pollution, and to increase the efficacy of beach cleanups by creating a predictive model that could aid future cleanup efforts.

2.2 Methods

2.2.1 Data description and region of interest

Since 2013 the North Sea Foundation, a Dutch environmental non-governmental organisation (NGO) advocating the protection and sustainable use of the North Sea marine ecosystem, has organised the national Boskalis Beach Cleanup Tour. During this tour, every year in August, the entire Dutch North Sea Coast is cleaned up by volunteers. It is the largest cleanup campaign in The Netherlands. The tour is divided into stages along the North Sea coast. The length of each stage is between 8-10 kilometres. The midway points of all stages are plotted in Figure 2.1 using the black crosses.

During the first three editions (2013-2015), the tour was organised over a period of a month, with one stage per day. From 2016 on, the tour took 15 days, with simultaneous cleaning of two stages per day. One cleanup team started on the Wadden Island Schiermonnikoog (most eastern cross in Figure 2.1), the other team started in the southwestern province Zeeland in Cadzand (most western cross in Figure 2.1). On day 15, both teams met halfway in Zandvoort ($\approx 4.5^\circ\text{E}$). The cleanups started around 10.00am and ended around 4.00pm, with total cleanup times between 4-6 hours for each stage. The volunteers were guided by cleanup teams of the North Sea Foundation, which consist of professional employees of the North Sea Foundation and trained volunteers.

At each stage, all litter present on the beach was collected in plastic bags and weighed. The weighing of the collected litter was done using analogue and/or digital scales (during the stage or at the end of the stage) and carried out by one of the members of the cleanup team. Most of the litter found was plastic (estimated percentage between 80-90% in terms of numbers). The years over which weights of collected litter are available for each stage are plotted in Figure 2.1 using the coloured squares. For most stages, weights are available for all years, in some cases stages were added in later years. Figures with the observed amount of litter per location per year are presented in the supporting material, Figure A.1a and Figure A.1b. To get an impression of the mean environmental conditions along the Dutch North Sea coast, the mean surface currents are plotted in Figure 2.1 using the arrows (Global Monitoring and Forecasting Center, 2021), and the mean wind speed and direction are plotted using the wind rose (Hersbach et al., 2020), all averaged over August between 2014–2019. The wind is predominantly coming from the southwest. Generally, the currents move from southwest to northeast along the North Sea coast. The effect of fresh water influxes from rivers is visible around the southern province of Zeeland ($< 52^\circ\text{N}$). The effect of this fresh water influx can be observed over considerable distances along the Dutch coast, for example in the form of fresh water lenses travelling downstream (de Ruijter et al., 1997; Rijnsburger et al., 2021). Ricker & Stanev (2020) found that locations with high salinity gradients due to a fresh water influx can act as a barrier for neutrally buoyant particles, possibly causing accumulation of litter along these fronts. Finally, not plotted in the figure, tidal currents move along the coast to the northeast during flood tide and southwest during ebb tide.

2.2.2 Data preprocessing

Different sources of marine litter exist, such as mismanagement of waste near the coast, input from rivers, or fishing gear which is lost at sea. The litter is then transported through the environment, and can eventually end up on beaches, influenced by various factors such as ocean currents and winds. How all these variables combined influence the beaching of litter is unknown, however. A regression model is used here to relate various environmental variables to the observed litter concentrations. We will assess whether it is possible to use the regression model to make predictions on the amount of beached litter, and if so, which environmental variables are important predictors to take into account.

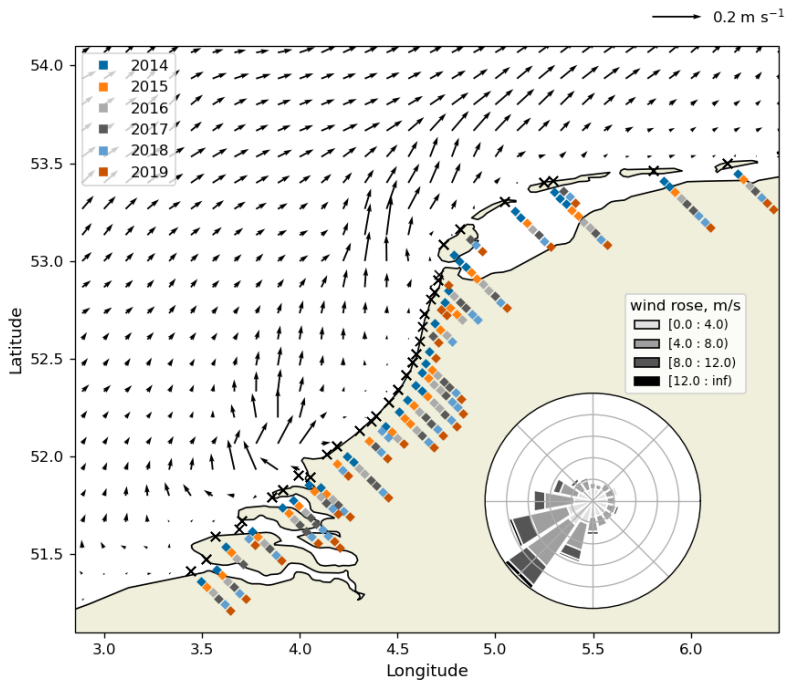


Figure 2.1: Locations of the midway points for each cleanup tour stage (black crosses); and in which year data are available (the coloured squares). For stages with multiple data points per year, different stretches of beach were cleaned (e.g. once the northern side, once the southern side). Also plotted are the mean surface currents (arrows) (Global Monitoring and Forecasting Center, 2021), and the wind rose (Hersbach et al., 2020), calculated over August 2014–2019.

Table 2.1: An overview of the numerical hydrodynamic and wind data used to derive the variables for the regression analysis. The data set name, temporal/spatial resolution, data used to assimilated the numerical models, and corresponding references are presented. For variables with an asterisk (*) data are used from July up to September 2014–2019. For data with a double asterisk (**) data are used for all months from January 2011 up to September 2019, as these are used for the Lagrangian model simulations as well.

| Variables | Data set name | Spatial res. | Temporal res. | Assimilated data | Reference |
|--------------------------|-------------------------------|--------------------------------|------------------------|---|--|
| U_{curr}^{**}, S^* | North West Shelf reanalysis | $1/9^\circ \times 1/15^\circ$ | daily-mean | temperature, salinity observations | (Global Monitoring and Forecasting Center, 2021) |
| U_{Stokes}^{**}, H_s^* | Global Ocean Waves reanalysis | $1/5^\circ \times 1/5^\circ$ | 3-hourly-instantaneous | H_s and directional wave spectra observations | (Global Monitoring and Forecasting Center, 2020) |
| U_{tide}^*, h_{tide}^* | FES2014 | $1/16^\circ \times 1/16^\circ$ | spectral | altimetry data, tidal gauges | (Lyard et al., 2021) |
| U_{wind}^* | ERA5 global reanalysis | $1/4^\circ \times 1/4^\circ$ | daily-mean | various observations | (Hersbach et al., 2020) |

For the environmental variables, three classes of data are used. First of all, hydrodynamic data (ocean currents, ocean surface waves, tides) and wind data are used (Section 2.2.2). Furthermore, we use Lagrangian simulation data, capturing transport of virtual particles representing floating litter. These simulations are used to estimate fluxes of litter onto beaches (Section 2.2.2). Finally, we use data of the coastal geometry and orientation (Section 2.2.2). Environmental variables are calculated for various lead times and distances from the measurement locations (expressed as radii around the stage midway points). These variables are then fed into a random forest algorithm to make the regression model.

Hydrodynamic and wind data

Numerical model data are used to specify the state of the sea and wind around the beach cleanup locations, as these factors have been found to likely play a role in the accumulation of beach litter (Eriksson et al., 2013; Thepwilai et al., 2021; Williams & Tudor, 2001). Reanalysis data are used, where historical observational data have been assimilated in numerical models.

Information on the ocean surface currents ($U_{curr.}$), salinity (S), Stokes drift (U_{Stokes}), and significant wave height (H_s) are derived from E.U. Copernicus Marine Environmental Monitoring Service Information data. High frequency tidal forcing has been used to produce the ocean current data, but output is only provided daily. To capture the effects of tides on a high temporal resolution, FES2014 data are used. Tidal currents (U_{tides}) and heights (h_{tide}) are calculated, taking the M_2 , S_2 , K_1 , and O_1 constituents into account (Sterl et al., 2020), as well as the M_4 and M_6 components which have been shown to play an important role in transport of suspended particles in the North Sea (Gräwe et al., 2014). The wind velocity field at 10m (U_{wind}) is taken from ERA5 reanalysis data. ERA5 data are used for the atmospheric forcing in the North West Shelf reanalysis product from which the surface current data are obtained, making these data sets consistent. Further details on the temporal/spatial resolution and assimilated data are given in Table 2.1.

Lagrangian model setup

While data on the sea state and wind might explain the litter accumulating on beaches to some extent, it misses information on possible sources of litter, and how this litter is transported through the marine environment. We therefore include estimates of beached litter fluxes in our analysis based on Lagrangian particle simulations.

Using the OceanParcels Lagrangian ocean analysis framework (Delandmeter & van Sebille, 2019), we model the trajectories of virtual buoyant particles at the sea surface using a Runge-Kutta 4 integration scheme. These virtual particles represent floating litter such as plastics. For the trajectories we consider a domain between 20°W–13°E, and 40°N–65°N, see Figure 2.2. We simulate a total of about 380,000 trajectories over the years 2011–2019. When particles move out of the specified domain they are removed, which mainly happens after particles move northward along the Norwegian coast. The ocean surface currents and Stokes drift from the hydrodynamic data are used to move the virtual particles around. We do not add additional tidal forcing to the Lagrangian model (Sterl et al., 2020) since the net effect of tides is already included in the ocean surface current data set (Global Monitoring and Forecasting Center, 2021). It is assumed that particles move just below the surface water, and do not experience a direct wind drag (Lebreton et al., 2018; Macias et al., 2019; Kaandorp et al., 2020). Effects of subgrid-scale phenomena are parameterised using a zeroth-order Markov model (van Sebille et al., 2018). The tracer diffusivity is set to a constant value of 10 m²/s, appropriate for the given mesh size (Neumann et al., 2014).

Virtual particles are released daily at river mouths, proportional to the estimated monthly riverine outflow of plastic waste based on the model by Lebreton et al. (2017). These sources are plotted using green circles in Figure 2.2. Particles are released daily in the sea, proportional to the amount of fishing hours based on Kroodsmas et al. (2018), shown in blue in Figure 2.2. These data are dependent on fishing vessel transponders, which are not equally present over the years. We therefore release a constant input of virtual particles from this source each day. Finally, there is a constant daily release of particles along coastlines, proportional to the amount of estimated land-based mismanaged plastic waste within a radius of 50km from the coastline (Jambeck et al., 2015; SEDAC et al., 2005). These sources are plotted in red-brown in Figure 2.2.

A beaching time scale τ_{beach} parameterises how quickly litter moves from the sea onto the beach when residing near the coast (for more details see Chapter 4). Here, the probability of beaching P_{beach} is given by:

$$P_{beach} = 1 - e^{-t_{coast}/\tau_{beach}}, \quad (2.1)$$

where t_{coast} is the time that particles spend in the model ocean cell adjacent to the coast. Various values for τ_{beach} are tested here, from $\tau_{beach} = 25$ days estimated for plastic particles and $\tau_{beach} = 75$ days estimated for drifter buoys later in Chapter 4, to a more conservative value of $\tau_{beach} = 150$ days. While in reality τ_{beach} might vary significantly both in space and time, it is unknown how this can be best parameterised (Onink et al., 2021). We use the Lagrangian model simulations to capture the large-scale transport of litter, and allow the regression model to pick the most

appropriate value for τ_{beach} later on. Only direct pathways of litter through the surface water are considered here and resuspension of litter from beaches (Onink et al., 2021) is ignored. Particles are tracked until they have lost more than 99% of their initial mass in the most conservative scenario of $\tau_{beach} = 150$ days. This means that particles are deleted when they have spent more than 691 days near the coast.

Each virtual particle starts with a unit mass. Each time step that a virtual particle spends near the coast, a fraction of its mass is lost due to the beaching process. This means that as t_{coast} increases for a virtual particle, a fraction of its mass is lost, which is calculated using Eq. (2.1). For each virtual particle, we calculate where and when it loses mass due to the beaching process. These masses lost to beaching are binned in a $1/9^\circ \times 1/15^\circ$ beaching flux histogram for each day. These beaching fluxes are denoted by F_{beach} , and are calculated for each particle source: $F_{beach,fis.}$, $F_{beach,riv.}$ and $F_{beach,pop.}$ for fishing activity, river inputs, and mismanaged plastic waste from coastal population, respectively.

Coastal orientation and geometry

Coastal orientation, geometry, and substrate are likely to influence the amount of litter that actually beaches on coastlines (Brennan et al., 2018; Andrades et al., 2018; Hardesty et al., 2017). Although the substrate of beaches in the Netherlands is relatively similar (sandy), there are local variations in the coastline orientation with respect to the large-scale coastline. We take this into account by including information on how the hydrodynamic and wind data are oriented with respect to the local coastline.

The Natural Earth data set is used here at a 1:10 million resolution (Kelso & Patterson, 2010), which is fine enough to estimate the general orientation of the beaches on which the cleanup stages have taken place. Two locations are not present in the coastal geometry of this data set (two man-made beaches along dams: Brouwersdam and Neeltje Jans); the coastal orientations of these locations were determined manually.

Normal vectors to the coastline (denoted by \mathbf{n}) are estimated by fitting a tangent plane through the points defining the coastline segments. Using a singular value decomposition we minimize the orthogonal distance between these points and the plane. All points within a box of 10×10 km centered around the stage midway point are selected (roughly the length scale of the beach cleanup tours). One example is plotted in Figure 2.3a, where the dotted box is the selection around the stage midway point, and the coastline segments within this box are indicated in orange. The resulting normal vector to this coastline segment is plotted using the orange arrow.

Dot products are calculated for vector fields (e.g. current velocity) with respect to the coastline normal vectors, to quantify how much a vector points on-shore (positive dot product), or off-shore (negative dot product). An example is presented in Figure 2.3b. At a given stage midway point, the numerical data within a certain radius are selected. For each of the cells we can then calculate the dot product of the vector data with respect to the coastline normal vector. In the example of Figure 2.3b, the normal vector points towards the northeast. Cells where the velocity vector points in roughly the same direction (on-shore) are coloured red, the opposite

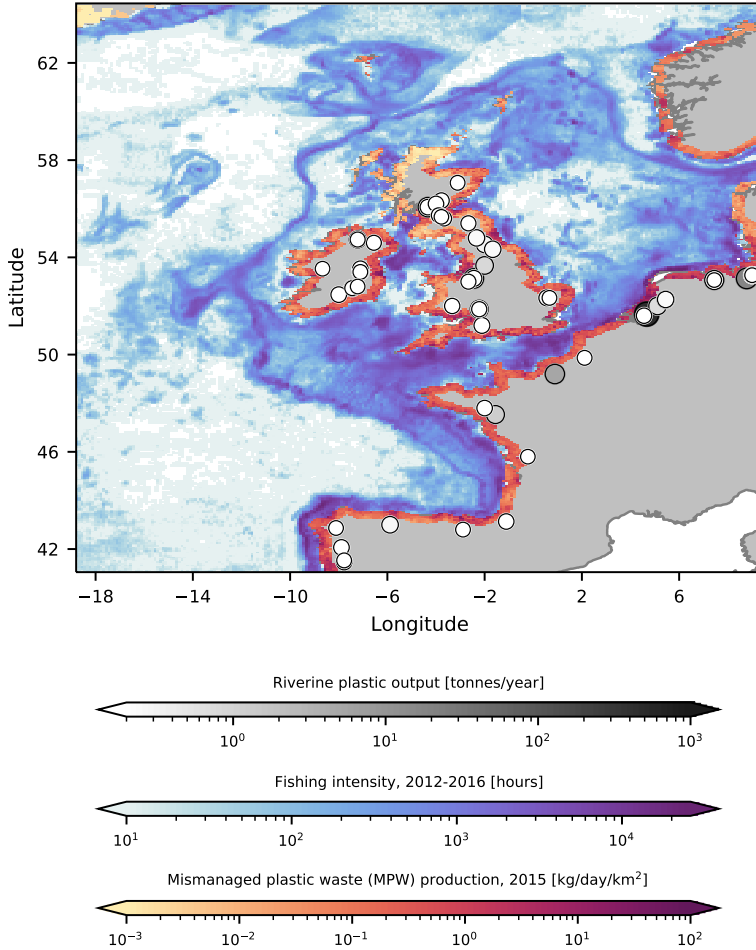


Figure 2.2: Input scenarios used to seed virtual litter particles in the Lagrangian simulations. Riverine input is indicated by the green circles, the amount of fishing hours in blue, and the coastal mismanaged plastic waste density in red. Note the log scale used for all input scenarios. While all rivers from Lebreton et al. (2017) are included in our analysis, only rivers predicted to transport more than 0.2 tonnes of plastic litter into the ocean are plotted here.

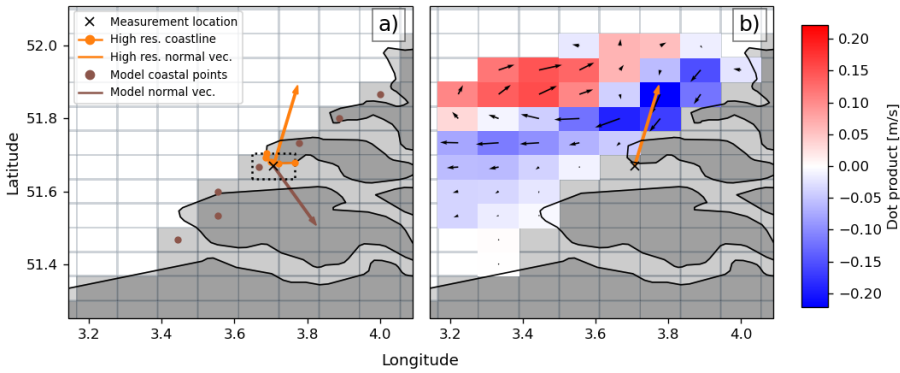


Figure 2.3: Illustration of the methodology used to calculate the directional variables. In the left panel (a), we show the high resolution coastline points and the derived normal vector (\mathbf{n}) in orange, located around the stage midway point (the black cross). Also shown are the numerical model coastline points and the derived normal vector (\mathbf{n}_{grid}) in brown. In the right panel (b), it is shown how the dot product variables are calculated. In a radius around the stage midway point, the dot product of the vector field is calculated with respect to the high resolution coastline normal vector (\mathbf{n}), where off-shore components are indicated in blue, and on-shore components in red.

directions (off-shore) are coloured blue. In Figure 2.3b the example is presented for only one time snapshot: the quantities can be calculated for various lead times. We then save derived quantities such as the mean, maximum, or minimum dot product over the lead time in a given radius, which will be further explained in Section 2.2.3. The coastal normal vectors are also used to estimate the misalignment between the numerical model coastline and the high resolution coastline. In Figure 2.3a, the numerical model grid cell centers at the coast are plotted using the brown dots. A singular value decomposition is used again to estimate the coastline normal vector of the numerical grid (\mathbf{n}_{grid} , indicated by the brown arrow). At each stage midway point, the dot product is taken of \mathbf{n}_{grid} with respect to the high resolution coastline normal vector \mathbf{n} , to obtain a measure for the misalignment. In the example plotted in Figure 2.3a there would be a large amount of misalignment between \mathbf{n}_{grid} and \mathbf{n} , resulting in a negative dot product between the two quantities.

Finally, the coastline length per grid cell is estimated. For each cell of the numerical model, we take the coastline segments within the given cell, and calculate their total length. Since coastlines show fractal behavior (Kappraff, 1986) their Euclidian length is not well defined. This means that the lengths calculated here are estimates, and that their value would increase when taking a higher model resolution.

Spatial variability

Information on spatial variability of beached litter can be useful for cleanup campaigns to target areas which are likely to be the most polluted. One might expect

that cleanup locations close to each other show more similar litter concentrations, compared to locations that are further apart. Furthermore, it is important for modelling studies to know the subgrid-scale variability which is not captured by the (discrete) numerical data. Finally, observing how spatial variability changes for different length scales could give us clues which physical processes are important for the dispersion of litter.

We will quantify the spatial variability of litter found on the coast as a function of the separation distance between the different cleanup locations using an empirical variogram. To compute the empirical variogram, all pairs of measurements within a certain distance of each other are compared, defined by $h \pm \delta$, where h is called the separation distance, and δ is half the bin width used to discretize the separation distance. The empirical variance $\hat{\gamma}(h)$ of the measurements separated by $h \pm \delta$ is calculated using (Bachmaier & Backes, 2011):

$$\hat{\gamma}(h \pm \delta) = \frac{1}{2N(h \pm \delta)} \sum_{(i,j) \in N(h \pm \delta)} (z_i - z_j)^2, \quad (2.2)$$

where $N(h \pm \delta)$ denotes the number of samples in the given separation distance bin, and z is the quantity of interest.

We calculate the empirical variogram on the \log_{10} values of the measured plastic concentrations in kg km^{-1} . Confidence intervals of the calculated variogram are estimated using a jackknife parameter estimation (Shafer & Varljen, 1990).

Measured litter concentrations are subject to both spatial and temporal variability. To remove temporal variability as much as possible from the empirical variance estimates, we only use data pairs within a certain time separation. Decreasing the time separation window reduces the effect of the temporal variability, but also reduces the number of available data pairs. We use a time separation of 3 days here, for which it was found there are still enough available data pairs to compute the empirical variogram.

2.2.3 Model

Machine learning features

The variables described in sections 2.2.2–2.2.2 are used to create a set of explanatory variables, which are related to the observed beach litter quantities. It is, however, not obvious what kind of lead time should be considered for the variables, and over which spatial scale the variables will have an influence on beach littering. We therefore calculate a large set of combinations for the explanatory variables by varying the radius of influence and/or the lead time. For the radii, we will consider the variable data closest to the stage midway point (which we will denote by a radius of 0 km), and variable data within radii of 50 and 100 kilometres. For lead times, we will consider 1, 3, 9, and 30 days. As shown in Eriksson et al. (2013) and Ryan et al. (2014), the turnover of litter on beaches generally happens within time scales of days, meaning that with this range of lead times we should be able to capture most of the litter accumulation. Furthermore, a lead time of 30 days also captures all tidal variability up to and including the spring-neap cycle. The combinations of vari-

ables, lead times, and radii will be called features, which are fed into the regression algorithm.

An overview of the features is given in Table 2.2. Three categories are defined: scalar features; directional features, which contain information on the direction of various vector fields with respect to the coastline; and features derived from the Lagrangian model simulations.

For the scalar features, we look at H_s , and the magnitude of U_{Stokes} , U_{wind} , $U_{curr.}$, and U_{tides} . We calculate the mean and the maximum of these quantities using all data points within the given radii and lead times.

We calculate a number of features derived from the tidal height h_{tide} . First of all, the maximum tidal height and the standard deviation of the tidal height over the given lead times are calculated, taking the closest data point from the stage midway point. Furthermore, a quantity is defined giving information in which period of the spring-neap tidal cycle the stage was monitored ($h_{tide,deriv.}$). The maximum tidal height at the stage day, and the maximum tidal height at the given lead time are calculated. We calculate the temporal derivative by subtracting both values and dividing by the lead time. A positive value means we are approaching the spring tide, a negative value means we are approaching the neap tide. Since spring tides occur roughly every two weeks, only lead times of 1 and 3 days are used for this feature. Finally, the minimum and maximum tidal height encountered during each stage are calculated, since these might contribute to how much beach was sampled during that day.

The total coastline length within a given radius is calculated (l_{coast}), using the Natural Earth data set as explained in Section 2.2.2. To include possible local sources of litter, the population within a given radius ($n_{pop.}$) is included as a feature (SEDAC et al., 2005), as well as the total fishing activity (Kroodsmas et al., 2018) within a given radius ($n_{fis.}$). Additionally, we want to include information on whether river mouths are present upstream of the cleanup stage. We use salinity (S) as a proxy for this, as a low salinity will indicate a nearby river mouth. The mean and minimum salinity are calculated over the various radii and lead times.

The number of participants for each stage is used as a feature ($n_{part.}$), to assess whether a lower percentage of litter is captured at stages with less participants. These data are available for 2017–2019. For 2014–2016 only the total number of participants per year is available. To estimate the number of participants per stage for these years, we first calculate the participant fractions per location over 2017–2019. These fractions are then scaled with the total number of participants over 2014–2016.

For the directional features, we calculate the dot product of the Stokes drift, wind, ocean currents, and tides with respect to the coastline normal vector (\mathbf{n}). Again, the mean and maximum are calculated, as well as the minimum, since this gives us additional information whether there have been strong off-shore components. These features are calculated for all radii and lead times. Furthermore, the misalignment of the numerical model coastline normal vector (\mathbf{n}_{grid}) with respect to the coastline normal vector is specified as a feature.

Finally, the total fluxes of beached litter from the Lagrangian particle simulations are given as features, from fisheries ($F_{beach,fis.}$), riverine input ($F_{beach,river.}$), and

Table 2.2: An overview of the machine learning features used. For each set of variables in each column, derived quantities are calculated such as the maximum, sum, or mean, over the given radius and lead time. Directional features are dot products of a given vector field with respect to the coastline normal vector \mathbf{n} . For parameters with an asterisk*, further explanation is given in the main text. For the last category (Lagrangian model features), the radius, lead time and the beaching time scale (τ_{beach}) are all varied.

| Categ. | Scalar features | | | | | | Directional features | | Lagrangian model features |
|-------------------|---|-------------------|--------------|---|-----------|-------------|--|--------------------------------------|---|
| Var. | H_s , $ \mathbf{U}_{Stokes} $, $ \mathbf{U}_{wind} $, $ \mathbf{U}_{curr.} $, $ \mathbf{U}_{tides} $ | h_{tide} | h_{tide} | l_{coast} , $n_{pop.}$, $n_{fis.}$ | S | $n_{part.}$ | $\mathbf{U}_{Stokes} \cdot \mathbf{n}$, $\mathbf{U}_{wind} \cdot \mathbf{n}$, $\mathbf{U}_{curr.} \cdot \mathbf{n}$, $\mathbf{U}_{tides} \cdot \mathbf{n}$ | $\mathbf{n}_{grid} \cdot \mathbf{n}$ | $F_{beach, fis.}$, $F_{beach, riv.}$, $F_{beach, pop.}$, $\tau_{beach} = 25, 75, 150d.$ |
| Qtt. | mean, max | max, std, deriv.* | max, min | sum | mean, min | - | mean, max, min | - | sum |
| Radii [km] | 0,50,100 | 0 | 0 | 0,50,100 | 0,50,100 | - | 0,50,100 | 0 | 0,50,100 |
| Lead times [days] | 1,3,9,30 | 1,3,9,30 | during tour* | - | 1,3,9,30 | - | 1,3,9,30 | - | 1,3,9,30 |

mismanaged waste from the coastal population ($F_{beach, pop.}$). These features are calculated for different beaching time scales τ_{beach} , all radii, and all lead times. The features are divided by the appropriate l_{coast} corresponding to the radius, to get the estimated beached litter fluxes per unit length of coast. One benefit of adding beached litter fluxes from the Lagrangian particle simulations, is that potential sources of litter far away from the beaching location can be included. While the radius of influence for all features goes up to 100 kilometres, the Lagrangian model features can still include information from further away, since the virtual particles are tracked indefinitely as explained in Section 2.2.2.

Regression model. The features and corresponding response (the measured amount of litter in kg km^{-1}) are used to fit a random forest regression algorithm (Pedregosa et al., 2011). This model allows us to capture non-linear relations between the features and response. It is a non-parametric model, and does not require prior knowledge on the model structure. These are both important reasons to choose the specific algorithm: coastal processes affecting dispersion of marine litter are highly complex (van Sebille et al., 2020), so we do not know a priori how the different environmental variables might interact, and how non-linear these interactions might be. The random forest regression model can aid in scientific knowledge discovery (Bortnik & Camporeale, 2021): it gives us Gini importances for all features (Nembrini et al., 2018). This is another reason for choosing this specific algorithm, as it provides us information which processes are important for predicting beached litter concentrations.

In total we have 342 features from all variable, radius, and lead time combinations. There are a total of 175 measured litter concentrations. The large number of features in comparison to the measurements makes it difficult to interpret the feature importance and could lead to overfitting. Therefore, k-fold cross validation is used to validate and test the model on a reduced amount of features, which are selected from a set of clusters.

Some features correlate as these are, for example, derived from the same variable,

but for a different radius or lead time. However, we do not know a priori which of these radii and lead times are the most appropriate predictors for the beached litter quantities. For example, litter concentrations might be influenced by long-term processes, slowly increasing the standing stock of litter on the beach, or the concentrations could be better predicted by conditions on the day leading up to the cleanup stage. Since we do not know this, we let the algorithm select the most appropriate variables. Features which are highly correlated will be assigned to clusters. We use hierarchical Ward-linkage clustering for this, based on Spearman rank-order correlations (McCann et al., 2019; Cope et al., 2017). This way, the total set of features is reduced to 66 feature clusters. For further details and interpretation of the clusters see the supporting material A.3.

Nested 5-fold cross validation is used for optimal feature selection from the clusters, and to assess the model performance on a test data set. In the outer loop, we use 80% of the data to train the model, and use the remaining 20% to test the model performance. This is repeated for each fold, i.e. 5 times. In the inner loop, 80% of the training data (i.e. 64% of the total data) is used to train the model, and 20% of the training data (i.e. 16% of the total data) is used to calculate the importance of the features, also repeated 5 times. Since in the inner loop none of the test data are used to train the model, we do not overpredict the model performance (Hastie et al., 2008). As all features in our regression model are continuous (i.e. there is no bias from categorical features (Nembrini et al., 2018)) we use the random forest Gini importance. After the inner loop is complete, we then select the feature with the highest Gini importance from each cluster. The random forest is trained using the selected features, and its performance is evaluated using the test data. We keep track of which features from the clusters are estimated to be the most important. The entire process is repeated 10 times, to obtain consistent feature importance estimates. A schematic of the model pipeline is presented in the supporting material A.4.

2.3 Results and discussion

2.3.1 Regression analysis

The regression model shows reasonable correspondence with the measured litter concentrations, where the Pearson correlation coefficient (R) based on the repeated cross validation is 0.72 ± 0.08 . A scatter plot with the measured litter concentrations on the x-axis and the predicted litter concentrations on the y-axis is shown in Figure 2.4. The points are coloured according to their test folds. As the 5-fold cross validation is repeated 10 times, only one realization is shown here, where every data point is plotted once.

In the same figure, the variability is shown that can be expected for length- and time scales smaller than the numerical data resolution. Using the empirical variogram, we calculate that $\hat{\gamma} = 0.08$ for lag distances of $h = 5 \pm 5$ km. This lag distance is at the lower side of the grid resolution for the numerical data (approximately 7 km for the ocean current data), so the model is not able to capture variations below this length scale. Therefore a 1:1 line is plotted ± 2 standard deviations based on this variance, as an indication of the optimal performance that can be

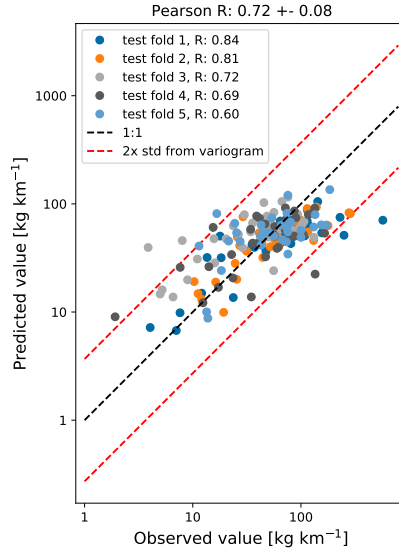


Figure 2.4: Scatter plot of the observed litter quantities (x-axis), and the modelled litter quantities (y-axis), both log-transformed. The points are coloured according to the 5 test folds used in the analysis. The 1:1 line is plotted using the black dashed line, and the estimated uncertainty based on the small-scale variance ($\pm 2\sigma$) is plotted using the red dashed lines.

expected. In this case, 94% of the predicted values lie inside the $\pm 2\sigma$ interval, indicating that the model is close to the optimal performance that can be expected for the given spatial and temporal resolution. It can be seen that there are two kinds of outliers in Figure 2.4: low observed litter concentrations not captured by the model (points in the upper left corner of the scatter plot), and high observed litter concentrations not captured by the model (points in the lower right corner of the scatter plot). This can be explained by the fact that the model is not able to capture all variability contained in the observations. As the hydrodynamic and wind data in the model have a limited resolution, subgrid-scale effects are missing (see Section 2.3.2). Furthermore, local point sources of litter (both spatially and temporally, e.g. shipping container accidents (van der Molen et al., 2021)) are not captured by the model.

In Figure 2.5 we show box-plots for the 10 most important features based on the Gini importance, picked out of the total 66 feature clusters. Importance scores for all 66 feature clusters are plotted in the supporting material A.2. The model indicates that tides play an important role for predicting the amount of beached litter. The most important feature is related to the long term variability of the tidal height, with a lead time of 30 days. Short term behavior is also seen as important, as the second most important feature is the maximum tidal height encountered within a lead time of 3 days. Furthermore, the maximum tidal height encountered during the

tour is the 6th most important feature, and the dot product of the tidal currents with respect to the coastline is the 8th most important feature. In general, a higher tidal maximum and variability lead to less litter measured on the coastline (see the supporting material A.2.5 for further details). A higher tide during or preceding the cleanup could re-suspend some of the litter from the beach. Furthermore, a higher tide encountered during the cleanup stage reduces the beach width that can be sampled. Perhaps a stronger variability in the tidal height leads to less persistent high strandlines where the highest litter concentrations are normally found (Heo et al., 2013). It has been shown in numerical studies that residual tidal currents can lead to a net transport of both suspended and floating matter (Gräwe et al., 2014; Børve et al., 2021; Schulz & Umlauf, 2016). While the regression model indicates that tides play an important role, it is difficult to separate the causal relations between all these different effects and the litter quantities found on beaches. To quantify this in more detail, further experimental and numerical studies are required.

The coastline length in the neighborhood of the cleanup stage (l_{coast}) is ranked as the 4th most important feature. This feature can describe multiple effects on litter concentrations. More coastline per unit area means that litter concentrations are possibly spread out over longer stretches of beach, reducing the amount of litter per kilometer of beach. Furthermore, an increasing l_{coast} indicates an increasing irregularity of the nearby coastline shape. This is for example the case around the province of Zeeland in the southwest ($< 52^\circ\text{N}$ in Figure 2.1): in these regions with irregular coastlines, more sheltered beaches can be found compared to regions with a long straight coastline, influencing the litter concentrations. Coastal orientation, $\mathbf{n}_{grid} \cdot \mathbf{n}$, plays an important role given its 5th highest Gini importance. When the coastline section tends to be more directly located towards the open sea, the large scale coastal geometry (\mathbf{n}_{grid}) aligns with the small scale coastal geometry (\mathbf{n}) at the locations used here. In e.g. Haarr et al. (2019) and Hardesty et al. (2017), it was reported that large scale headlands tend to enhance catchment of litter compared to large scale sheltered areas. This is in line with our findings, with an increasing $\mathbf{n}_{grid} \cdot \mathbf{n}$ leading to more predicted litter (see the supporting material A.2.5).

Results suggest that transport of marine litter is important to take into account, as the 3rd and 7th most important features are beaching fluxes from the Lagrangian model simulations from fishing activity and coastal mismanaged waste. These features implicitly contain information on various hydrodynamic variables and sources of litter, explaining why these are ranked above most other scalar and directional features related to wind, currents, and waves. Also interesting to notice, is that they are all ranked above the nearby fishing activity ($n_{fis.}$) and population density ($n_{pop.}$), which are the 10th and 14th most important features respectively (see Figure A.2). This could indicate that transport of litter through the marine environment is important to take into account, as opposed to only considering local terrestrial sources. From the three possible sources of litter used in the model, transport from fisheries is the most important. This is consistent with the litter composition found on Dutch beaches, which consists for a major part of fishing related items (40%; van Duinen et al., 2022).

Finally, the dot product of $U_{curr.}$ with respect to the coastline is seen as important, at place 9. This feature is related to small-scale/long-term behavior, which might

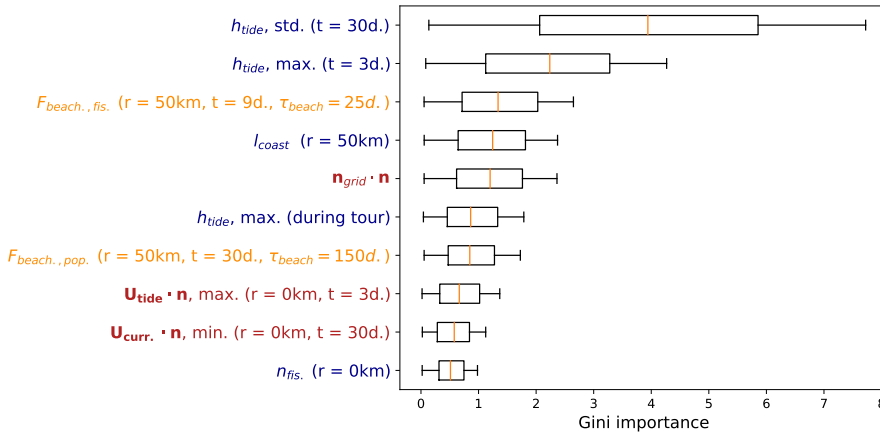


Figure 2.5: Box plots for the feature Gini importances from the random forest regression algorithm. Only the top 10 features are plotted here, an overview of all features can be found in the supporting material A.2. The label colours correspond to the variable categories in Table 2.2, where scalar features are indicated in blue, directional features in red, and Lagrangian model features in orange. The radius and lead time are indicated in the brackets when applicable.

give an indication whether there are currents present moving the litter on-shore to the cleanup stage location.

Changes in predictive capability are relatively small when leaving out the Lagrangian model simulation features, see Figure A.3. The Pearson correlation coefficient R in this case is 0.72 ± 0.10 , which is not significantly less than the full model. This suggests that to some extent information on transport of litter is also contained in other variables such as the currents, waves, and wind magnitude and direction. Directional information seems to play an important role, as when leaving out the Lagrangian model simulation features, 4 out of the 10 most important features are related to the dot products of currents, tides, and Stokes drift with respect to the coastline (see Figure A.4).

It is estimated that the number of participants taking part in the tour does not have a large influence on the amount of litter that is found, see the supporting material A.2 for further details. This suggests that with an average of 77 participants per campaign, adding more participants would not necessarily lead to more litter being cleaned up. No clear patterns emerge regarding lead times and radii for the most important features. This could indicate that litter found on beaches is an ensemble of objects with different moments of beaching and residence times. Features regarding wind and significant wave height are seen as less important, being ranked 18th and lower, see Figure A.2. It is possible that this information is already contained in the Stokes drift, or that they play a lesser role in the transport of litter. One explanation is that most of the litter found during the cleanup tour has a relatively low wind drag coefficient in the water, which was also observed in Lebreton et al. (2018) for litter in the Great Pacific Garbage Patch.

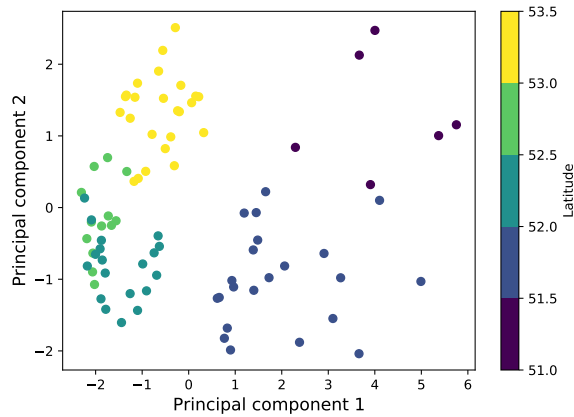


Figure 2.6: The two principal components based on the five most important features (see Figure 2.5). The points are coloured according to their latitude, from which the separation of measurements into three different clusters (51–52°N, 52–53°N, and 53–53.5°N) becomes evident.

Having the full set of 66 feature clusters is not necessary for predictive capability. In Figure A.5 we show that the model performs well when only picking the top 8 features (Pearson correlation coefficient $R: 0.79 \pm 0.04$). Increasing the amount of features does not increase the model performance. For an operational model it would therefore be recommended to stick to a lower amount of features, as this keeps the model simple and easier to interpret. We investigate if the most important variables are related to certain locations by performing a principal component analysis, taking these 8 most important features in the full model (Figure 2.5). A scatter plot of the first two principal components is presented in Figure 2.6, where the dots are coloured according to their latitude. The two principal components explain 50% and 17% of the total variance respectively. What can be seen, is that the points separate into roughly three different regions: measurements taken at lower latitudes around the province of Zeeland (51–52°N), measurements taken between 52–53°N, and measurements obtained near the Wadden Islands (53–53.5°N). The first principal component shows the highest absolute correlation (Pearson $R: 0.45$) with long-term tidal variability (with a lead time of 30 days). The second principal component shows the highest absolute correlation (Pearson $R: -0.58$) with the nearby coastal length (within a radius of 50km). As the measurements taken between 52–53°N are clustered quite closely together, this indicates that conditions regarding tides and coastline geometry are relatively similar for these locations. Variations of the tidal height are relatively large for 51–52°N. The coastal geometry is also more irregular here compared to the rest of the Netherlands. These factors combined likely lead to less litter on beaches here: for $< 52^\circ\text{N}$ we find on average 52 kg km^{-1} , for $> 52^\circ\text{N}$ we find on average 73 kg km^{-1} , calculated over 2014–2019.

2.3.2 Spatial variability

To assess which length scales are important for the spatial variability of beached litter, we calculate the empirical variogram for different lag distances. Spatial variability remains relatively constant for lag distances up to about 100 km, with a mean of $\hat{\gamma} = 0.07$, see Figure 2.7. For the smallest lag distance ($h = 5 \pm 5$ km), we find $\hat{\gamma} = 0.08$. This variance estimate was also used to create the error bars in Figure 2.4. Around $h = 125$ km there seems to be an increase in the variance, to about $\hat{\gamma} = 0.2$ – 0.3 . At this lag distance there is also a large uncertainty in the estimates however, and fewer unique data pairs to calculate the empirical variance.

Interestingly, some periodic behavior seems to be present, with a length scale of about 25 kilometers. One possible explanation could be the typical spacing of the Dutch islands and peninsulas. As shown in the previous section, coastline orientation likely plays an important role in the amount of observed litter. This effect can also present itself in the variogram with, for example, measurements in sheltered areas (e.g. coves) being more correlated with each other, compared to nearby exposed locations (e.g. headlands).

The grid sizes used for our numerical data ranges from about 7 km (the surface current data), to about 20 km (the wind data). This means that the variance at and below these length scales is not captured by the numerical data. The variance calculated for lag distances up to 20 km is quite substantial ($\hat{\gamma} = 0.05$ – 0.12). As can be seen in Figure 2.4 the values corresponding to the lower and upper 95% confidence interval vary by about an order of magnitude. This is essential to consider when using observational data to inform numerical models: due to the amount of variability at the subgrid-scale level, relatively large sets of observational data are required to extract information. A large number of physical processes could induce variability below length scales of 20 km, such as Langmuir circulations, or processes in the coastal zone such as wave breaking, rip currents, and longshore currents (van Sebille et al., 2020). Finally, it is important to consider that spatial variability is inherent to data obtained from cleanup campaigns such as analysed here, due to e.g. different participants having slightly different strategies for finding litter on beaches.

2.3.3 Extrapolating litter quantities to the entire coastline

The random forest regression model can be used to extrapolate how much litter is likely beached along the entire Dutch coastline. First, a regression model is trained using the top 8 features listed in Figure 2.5. We then divide the Dutch North Sea coastline into $1/9^\circ \times 1/15^\circ$ sections (roughly 7 by 7 kilometers). For each of the sections the top 8 features are computed, as well as the total coastline length contained in each section. In total we have 65 separate sections, and a total coastline length of 365 kilometers, which matches the total length of the Dutch North Sea coastline from literature (Roomen et al., 2008). We choose to use a model trained using the top 8 features for the extrapolations, as increasing the amount of features does not increase the predictive performance (see Figure A.5). Furthermore, reducing the amount of features simplifies the computations, as we do not need to compute all 391 variables again for all coastline sections.

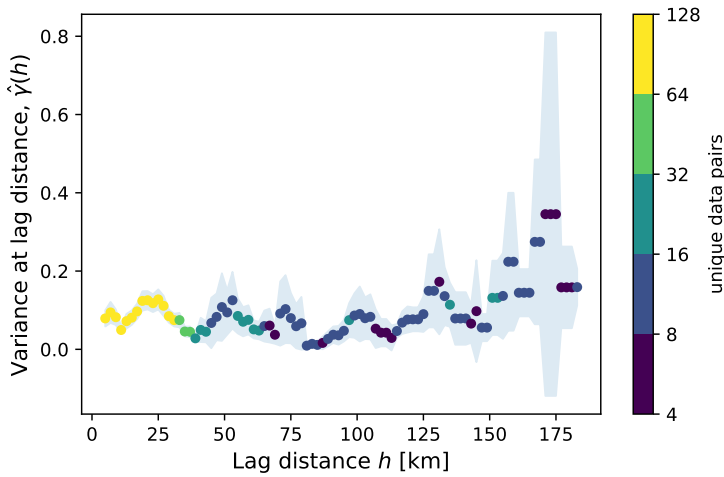


Figure 2.7: Variogram calculated for the \log_{10} of the measured litter quantities in kg km^{-1} , with the lag distance h on the x-axis, and the empirical variance $\hat{\gamma}(h)$ on the y-axis, only taking data pairs into account with a maximum of 3 days temporal separation. For the separation distance half bin width $\delta = 5$ km is used. The points are coloured by the number of unique data pairs used to calculate the variance, the jackknife uncertainty estimate ($\pm\sigma$) is shaded in blue.

For each section, the litter concentrations in kg km^{-1} are predicted per day over August 2014–2019. Predictions are only made for August since all cleanup campaigns were organized during this period, and making predictions for other months might induce seasonal biases. The mean concentrations per coastline section are plotted in Figure 2.8. For each day, the total litter quantities are computed by multiplying the litter concentrations by the coastline length per section. Monte Carlo estimates of the confidence bounds are calculated by randomly adding noise proportional to the estimated variance ($\hat{\gamma} = 0.08$), which is repeated 1000 times per day per section.

We find a total of 16,500–31,200 kg litter along the Dutch North Sea coastline based on the 95% confidence interval. It must be noted that this only accounts for the visible litter on the beach surface. The cleanup efforts are likely to miss a substantial amount of beached litter which is buried in beach sediment or located at the back of the beach (e.g. in vegetation). This was for example noted in Lavers & Bond (2017) for a remote island in the South Pacific, where in terms of mass about 68% of the litter was located on the beach surface, 27% at the back of the beach in and around vegetation, and 5% buried in beach sediment. Further research is necessary to quantify how these numbers translate to Dutch beaches.

The total amount of litter gathered during the cleanup campaigns, and the total amount of kilometers sampled per year is presented in Table A.1. The total amount of litter gathered varies from 9,872 to 20,078 kilograms. This is in line with the expected total amount of litter predicted by the model, since the majority of the coastline (222–262 kilometers out of 365 kilometers) was covered during the cleanup

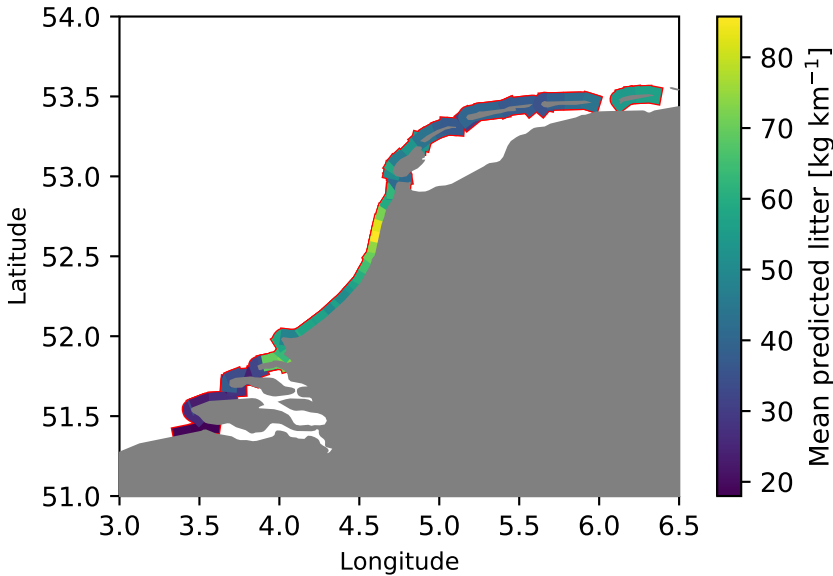


Figure 2.8: Mean litter concentrations over August 2014–2019 extrapolated to the entire Dutch coastline.

campaigns.

2.4 Conclusions and recommendations

Using data from beach cleanup efforts in the Netherlands between the years 2014–2019, we analysed which variables are important for predicting litter on beaches, and what spatial variability this litter has. In order to do this, we fitted a regression model to the observed litter quantities, as a function of variables related to wind, waves, currents, tides, coastal geometry, and simulated oceanic transport. We find that tides play an important role, where an increasing tidal variability and increasing tidal maximum lead to less observed litter on beaches. Other important variables are whether the local orientation of a beach corresponds to the large-scale coastline orientation, and the total nearby coastal length, which can both be seen as measures of how exposed a beach is. These factors are likely explanations why the observed litter quantities are relatively low in the southwestern part of the Netherlands compared to the other parts. Additionally, transport of litter through the marine environment is seen as important to take into account by the regression model. Rivers, fishing activity, and mismanaged plastic waste along coastlines were taken into account as possible sources of litter in the transport model, where the regression analysis attributed relatively much importance to litter originating from fishing activity. This is in line with findings in van Duinen et al. (2022), as approximately 40% of the litter found on the Dutch North Sea coastline is estimated to originate from the fishing

industry.

We compute that spatial variability of the observed litter concentrations is substantial on length scales less than 10 kilometers, causing model $\pm 2\sigma$ confidence bounds to vary by about an order of magnitude. Due to this significant variability, large observational data sets are necessary if they are to be used to inform numerical models. Finally, based on extrapolation of the regression model, we estimate that the Dutch North Sea coastlines contain a total of 16,500–31,200 kilograms (95% confidence interval) of litter on the beach surface.

Estimating the spatial variability of beached litter can give us information for efficient monitoring of pollution. It can be used to constrain estimates of litter concentrations based on observations elsewhere. We found that the variance for lag distances smaller than 125 km is relatively constant around $\hat{\gamma} = 0.08$. As an example, if one measures a relatively high amount of 200 kg km^{-1} at the northern tip of the mainland near Den Helder ($\approx 53^\circ\text{N}$ in Figure 2.1), one can expect at least 54 kg km^{-1} of litter elsewhere in the Northern part of the Netherlands, taking the 95% confidence interval. After 125 km, the estimated variance seems to increase, meaning that this observation becomes less informative for locations further away.

For future studies on quantifying beach litter variability, it would be interesting to segment the beach cleanup tours into smaller stretches. One idea would be to organize some stages where the litter quantities are weighed per 1 kilometer, 100 meter, or even shorter stretches. This way it would be possible to estimate the variance on sub-kilometer scale. Ryan et al. (2020) reported significant correlations between measurements taken roughly 50 meters apart (Spearman rank correlation of about 0.9). It would be interesting to see how this changes up to the kilometer scale. This can give us valuable insights into which processes might be causing the high amount of variability between litter observations, and what length scales should be taken into account to capture this variability with models. We see relatively few data points in Figure 2.7 for larger lag distances. Performing the cleanup stages in a randomized order would provide a more even coverage of data points over the given lag distances.

Future studies could further investigate the causal relations between the variables seen as important predictors by the regression model and the litter concentrations found on beaches. This is especially the case for tides, which constitute the two most important features in the regression model (see Figure 2.5). Experimental studies could further determine whether lower litter concentrations at locations with higher tidal variability are mainly caused by litter re-suspending back into the sea, or for example due to the fact that less area of the beach is sampled during high tide. It should additionally be investigated how these effects compare to the role of (residual) tidal currents, as it has been shown that this can play an important role in transporting suspended matter towards the shore (Schulz & Umlauf, 2016). Experimental investigations can be done in combination with numerical studies of the nearshore marine environment, to capture the interactions between processes such as tides, waves, and particle sizes (Alsina et al., 2020).

It should be investigated how the results found here generalize to other geographic regions, and how the importance of explanatory variables vary globally. The model

itself can not directly be used for other geographic regions, since the features used to train the algorithm are specific to the region of interest. The model is likely to perform poorly when making extrapolations for conditions not present in the training data. As an example, the substrate of beaches is likely to have a large impact on litter concentrations (Hardesty et al., 2017), which are relatively uniform in this analysis (all sandy beaches). According to our regression model, wind is not a very important variable to take into account. Perhaps some of the high-windage litter has been beached before reaching the Dutch waters. It should be noted, however, that wind indirectly affects other variables such as the ocean currents and therefore also the Lagrangian particle simulations. It would be interesting to re-do this analysis with data obtained nearby the English channel and check if wind plays a more important role there, as in the Lagrangian model simulations many virtual particles pass this region.

It is necessary to further investigate the effect of regular cleaning of beaches by municipalities and other volunteer groups or individuals. This effect was left out in this analysis due to unavailability of these data. It is likely that mainly the beaches near densely populated areas are regularly cleaned. Since data on population density has been included in the features, it is possible that this effect is taken into account by the regression model, but further analysis is necessary. Furthermore, effects of tourism can be taken into account in the future when these data are available, as this affects the local population density seasonally.

Regarding effective cleanup of beaches, it is recommended to perform beach cleanups during low tide, preferably in a week around the neap tide, when the tidal variability is lower. If limited resources are available, one can focus on exposed shorelines which generally accumulate more litter. Additionally, more litter can be expected on relatively straight shorelines, compared to more irregular geometries where litter is distributed over longer stretches of beach. We saw no effect from the number of participants per beach cleanup tour on the amount of gathered litter, with an average of 77 participants per tour. One possible improvement to clean up more litter could therefore be to spread out participants over different stages, avoiding that parts of the beach are inspected multiple times.

Modelling size distributions of marine plastics under the influence of continuous cascading fragmentation

Field studies in the global ocean have shown that plastic fragments make up the majority of plastic pollution in terms of abundance. It is not well understood how quickly plastics in the marine environment fragment, however. Here, we study the fragmentation process in the oceanic environment by considering a model which captures continuous fragmentation of particles over time in a cascading fashion. With this cascading fragmentation model we simulate particle size distributions (PSDs), specifying the abundance or mass of particles for different size classes. The fragmentation model is coupled to an environmental box model, simulating the distributions of plastic particles in the ocean, coastal waters, and on the beach. We demonstrate the capabilities of the model by calibrating it to estimated plastic transport in the Mediterranean Sea, and compare the modelled PSDs to available observations in this region. Results are used to illustrate the effect of size-selective processes such as vertical mixing in the water column and resuspension of particles from the beach into coastal waters. The model quantifies the role of fragmentation

This chapter has been published as:

Kaandorp, M.L.A., Dijkstra, H.A., van Sebille, E. (2021): Modelling size distributions of marine plastics under the influence of continuous cascading fragmentation. *Environmental Research Letters*, 16:054075. <https://doi.org/10.1088/1748-9326/abe9ea>

on the marine plastic mass budget: while fragmentation is a major source of secondary plastic particles in terms of abundance, it seems to have a minor effect on the total mass of particles larger than 0.1 mm. Future comparison to observed PSD data allow us to understand size-selective plastic transport in the environment, and potentially inform us on plastic longevity.

3.1 Introduction

Studies have shown that fragments make up the majority of marine plastic litter in terms of abundance in the global ocean (Cózar et al., 2014; Suaria et al., 2016). The large amount of fragments is evident from particle size distribution (PSD) data, specifying the abundance or mass of particles for different size classes. An overview of PSD data from various studies is given in Kooi & Koelmans (2019); some examples are presented in Figure 3.1. What is commonly observed in PSD data is a power law for larger fragments (>1 mm in Figure 3.1), see e.g. Cózar et al. (2014, 2015), Erni-Cassola et al. (2017), and Enders et al. (2015). This shows as a straight line on a log-log scale as can be seen in Figure 3.1. Oftentimes, a maximum in the PSD is observed at smaller particle sizes (~ 1 mm in Figure 3.1), ending the power law regime. This maximum has been observed to vary, and has been attributed to, for example, the distance to the nearest coast (Pedrotti et al., 2016; Isobe et al., 2014).

It is necessary to further investigate the fragmentation process if we want to explain the particular shapes of measured PSD data. Fragmentation of plastics is likely dominant on beaches or inland water bodies such as rivers, where plastics are subjected to UV-radiation, oxidation, and higher temperatures, embrittling the particles, which enhances the breaking down of particles by mechanical abrasion (Andrady, 2011; Kalogerakis et al., 2017; Song et al., 2017; Efimova et al., 2018). Fragmentation models have been proposed in e.g. Cózar et al. (2014), hypothesising that the PSD slope depends on whether particles break down in a three-dimensional fashion (i.e. like a cube), or more in a two-dimensional fashion (like a thin sheet). It has been shown that the polymer type influences how plastic particles fragment (e.g. due to differences in the surface cracks, see Andrady, 2011), and how quickly plastic particles fragment (Song et al., 2017), hence directly influencing how PSDs evolve.

While the main driver behind the PSD might be fragmentation, physical processes can have a size-selective influence on plastic particles (van Seville et al., 2020). Vertical turbulent mixing, induced by for example the wind, has been shown to mix smaller particles with lower rise velocities to larger depths (Reisser et al., 2015; Poulain et al., 2019; Kukulka et al., 2012; Chor et al., 2018). This can reduce smaller size fractions in PSDs measured by nets at the ocean surface (typically submerged ± 10 –50 centimeters depending on net type, see e.g. Pedrotti et al., 2016; Cózar et al., 2015; Suaria et al., 2016). Furthermore, bigger (more buoyant) particles likely experience more influence from Stokes drift, given its limited depth of influence (van den Bremer & Breivik, 2017; Breivik et al., 2016). Model studies have indicated that Stokes drift tends to push plastic particles towards coastal areas (Iwasaki et al., 2017; Onink et al., 2019; Delandmeter & van Seville, 2019).

In Isobe et al. (2014) it was observed that for coastal seas near Japan, overabundances of larger plastic particles were found close to the coast versus more offshore, see Figure 3.1b. Coastal processes, such as beaching and resuspension, can be size-selective. In Hinata et al. (2017), residence times of particles on beaches in Japan were estimated using tagged litter. Higher particle rise velocities in the water were related to longer residence times, as these particles are more likely to be pushed to the backshore by wave swash. This could mean that larger objects remain longer on beaches, and hence experience more weathering (Hinata et al., 2020). Finally, PSDs could be influenced by size selective sinking, induced by for example biofouling (Ryan, 2015). Biofouling models predict that smaller particles, which have a larger surface to volume ratio, tend to sink more quickly (Kooi et al., 2017). This has been observed in experimental studies as well (Fazey & Ryan, 2016).

Previous studies, such as the ones by Koelmans et al. (2017) and Lebreton et al. (2019), have tried to quantify marine plastic mass budget using conceptual models. In both of these works, fragmentation is purely defined as a rate, breaking down a mass percentage of a macroplastics category into a microplastics category over time. How, and how quickly plastics fragment is still a very uncertain factor however.

In this work, we consider a fragmentation model based on fractal theory (Turcotte, 1986; Charalambous, 2015), modelling a large range of different size classes. A benefit of modelling a range of size classes is that we can calibrate the model to experimental fragmentation studies such as the one by Song et al. (2017), where different polymers were subjected to laboratory conditions simulating weathering in the marine environment. By modelling a range of size classes, we can furthermore compare model output to measured PSDs in the environment, such as the ones presented in Figure 3.1.

We couple our fragmentation model to an idealized box model where the marine environment is split into three different compartments, similar to Lebreton et al. (2019): the beach, coastal water, and open ocean. By considering a range of size classes, we can study size-dependent processes in the marine environment mentioned earlier and their influence on the fragmentation process and resulting PSDs. Finally, our model allows quantification of PSDs both in terms of the amount of particles in each size class, and the particle mass in each size class. We will make a distinction between the two, and call them the number (i.e. abundance) size distribution (NSD), and the mass size distribution (MSD). We will use the term PSD when talking about size distributions in general (i.e. either NSD or MSD). We will show that with these MSD data, our model can contribute to obtain a better understanding of the plastic mass budget. Similar to the models in Koelmans et al. (2017) and Lebreton et al. (2019), the idealized model presented here allows us to efficiently test hypotheses regarding fragmentation, sources, sinks, and transport of marine plastics. We will demonstrate this by applying the model to different marine plastic scenarios in the Mediterranean Sea. The goal is to have an analysis which is consistent with current experimental data of the fragmentation process, observational data in terms of plastic concentrations and plastic PSD data, and current knowledge on marine plastic sources, sinks, and transport.

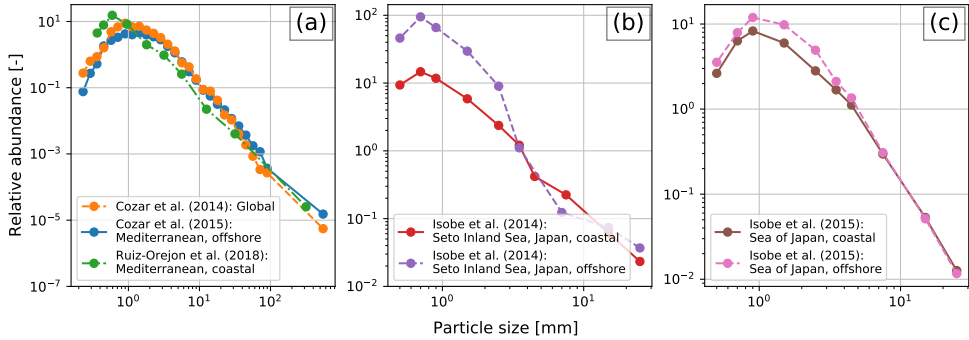


Figure 3.1: Observed number size distributions. a): Samples from C  zar et al. (2014) obtained around the world, and samples from C  zar et al. (2015) and Ruiz-Orej  n et al. (2018) obtained in the Mediterranean Sea, see the supporting material B.3 for locations. b) and c): Samples from Isobe et al. (2014, 2015) obtained from the Seto Inland Sea (Japan) and the Sea of Japan. Coastal samples are defined as less than 15 kilometers from the shoreline. All particle size distributions have been normalized relative to the total abundance of large items (> 10 mm) to show differences for small particle sizes clearly. Most datasets seem to follow a power law for the larger particle size classes (see Figure 3.7 below). Coastal samples tend to have relatively few small particles (or: relatively many large particles), and all distributions show a peak in abundance around 1 mm, instead of a monotonic increase with smaller sizes.

3.2 Methods

3.2.1 The cascading fragmentation model

The fragmentation model discussed here is based on simple fractal geometries. We define the spatial dimension as D_N . When $D_N = 3$, we start with a cube with a size of $L \times L \times L$ which we call the parent object, see Figure 3.2a. This cube can be split in eight equally-sized cubes, which can each be recursively split again. The size class of the parent object is defined as $k = 0$, the size class of the cubes with length $L/2$ is defined as $k = 1$, and so on. When $D_N = 2$, the starting object is a sheet instead of a cube, which can be split in four smaller sheets each time the size class increases. C  zar et al. (2014) presented a fragmentation model where objects are broken down into a set of smaller (equally sized) fragments in a series of successive fragmentation events. This fragmentation model was used to explain why measured PSDs often resemble power laws, i.e. functions of the form

$$n(l) = Cl^{-\alpha}, \quad (3.1)$$

where n is the abundance, l is the particle size, α is the power law slope, and C is a constant. However, this fragmentation model requires a constant input of new parent objects to achieve a power law, while laboratory experiments have shown that power laws in the PSD also appear after fragmenting a single input of parent objects (Song et al., 2017).

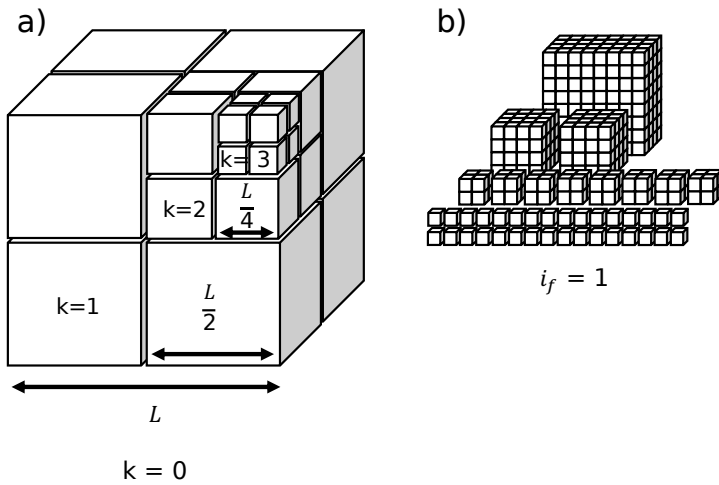


Figure 3.2: a) Illustration of a parent cube (size class $k = 0$), consisting of successively smaller cubes, based on Turcotte (1986). Only 3 iterations until size class $k = 3$ are shown here, the size class can increase indefinitely. b) Illustration of the cascading fragmentation model with $p = 0.5$ and after one fragmentation event i_f , based on Turcotte (1986) and Charalambous (2015)

The fragmentation model used here builds upon the work of Turcotte (1986), where it was noted that scale-invariance of the fragmentation process, whether it be caused by weathering, explosions, or impacts, leads to such a power law. The idea behind the model of Turcotte (1986) can be illustrated using Figure 3.2b. Following one fragmentation event i_f , a certain fraction p of the original cube (size class $k = 0$) splits off. For example, if $p = 0.5$, this results in 4 fragments of $k = 1$ splitting off, leaving 0.5 object in size class $k = 0$. This process is assumed to be the same on all length scales: a fraction p will split off from the fragments in size class $k = 1$ as well: 16 fragments of size class $k = 2$ are created, and $4 \times 0.5 = 2$ fragments are left in size class $k = 1$. This process is repeated indefinitely.

Bird et al. (2009) and Gregory et al. (2012) extended this model by including a temporal component, with each fragment breaking down further as i_f progresses. Charalambous (2015) showed that repeatedly breaking down fragments over discrete steps of i_f is a sequence of independent and identical Bernoulli trials with a chance of success p , yielding a negative binomial distribution. This is rewritten in terms of a continuous fragmentation index f (instead of the discrete i_f), yielding a probability density function giving the mass m in size class k at fragmentation index f as:

$$m(k; f, p) = \frac{\Gamma(k + f)}{\Gamma(k + 1)\Gamma(f)} p^k (1 - p)^f, \quad (3.2)$$

where Γ is the gamma function. We will call this model, introduced in Charalambous (2015), the cascading fragmentation model. We assume that f is directly proportional to time in the environment, and will review this assumption in the discussion. The amount of fragments in a given size class is estimated by multiplying the mass with $2^{D_N k}$, a factor determining how many fragments of size class k fit inside the parent object:

$$n(k, f, p) = 2^{D_N k} m(k; f, p) \quad (3.3)$$

We use $D_N = 3$ as the baseline. However, this factor is $D_N = 2$ for purely flat objects like plastic sheets and $D_N = 1$ for fibers or lines. As real-world samples contain a combination of these objects, the value for D_N in the environment can be a non-integer between 1 and 3. The value of D_N is only influenced by the shape of the objects. The material properties (e.g. polymer type) only affect the value of p and how quickly f progresses in time.

Figure 3.3a shows the NSD resulting from the cascading fragmentation model at various fragmentation indices f . We start with one cube with a length L of 1 mm at $f = 0$. The continuous description in Eq. (3.3) allows us to model the amount of fragments at a very small fragmentation index of $f = 0.001$. There are few larger fragments ($> 10^{-2}$ mm) per parent object at this stage. At $f = 1$ we have exactly a power law in the NSD, equivalent to the model by Turcotte (1986) which only considers a single discrete fragmentation event i_f . A fractal dimension D_f of the object formed by all fractions can be defined, relating to D_N and p by:

$$D_f = \log_2 (2^{D_N p}). \quad (3.4)$$

The NSD power law slope at $f = 1$ is given by this fractal dimension.

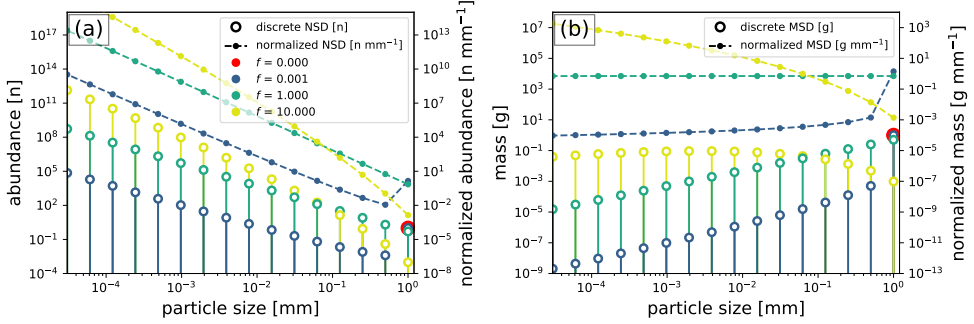


Figure 3.3: Particle size distributions from the cascading fragmentation model, using $p = 0.5$. Distributions for different fragmentation indices f are shown. Figure a: the discrete NSD (left y-axis) and the NSD normalized with respect to the particle size (right y-axis). Figure b: the discrete MSD (left y-axis) and the normalized MSD (right y-axis)

Fragments can be broken down further, eventually resulting in the NSD shown for $f = 10$. This is not a power law anymore, and the slope of this curve has increased significantly, with relatively many particles in the small size classes. The NSD (units: n) can be normalized, by dividing the amount of fragments by the size class bin width (units: $n \text{ mm}^{-1}$). These normalized NSDs are presented by the dashed lines. Because of the log-scale on the x-axis, the distance between the given particle sizes increases by a constant factor. This increases the magnitude of the normalized NSD slopes by 1 compared to the discrete NSD. The slope of these normalized NSDs is not dependent on the size class bins used for the measurements, allowing for comparison between different studies.

Figure 3.3b shows the same analysis in terms of mass, i.e. the MSD, starting with one cube of 1 g and 1 mm^3 . As fragmentation progresses, mass shifts from the large fragments towards smaller fragments. At $f = 1$ we have a power law: the difference in the slope between the NSD and MSD is 3, resulting from the $2^{D_N k}$ term in Eq. (3.3), with $D_N = 3$.

3.2.2 Environmental box model

With the cascading fragmentation model we can now simulate PSDs over time. Different particle sizes will undergo different types of forcing and transport in the environment. The combination of fragmentation and size-selective transport is investigated using a box model, presented in Figure 3.4. The boxes in this model represent three different environmental regions: the beach, coastal waters, and open ocean. Particles can move between the different environments, defined by a set of transition probabilities (P): particles can move to a different environmental box (the arrows on the right in Figure 3.4), remain in the current box (recurring arrows on the left), or vanish from the system (i.e. a sink, arrows on the left). Subscripts in Figure 3.4 denote ocean, coast, beach, or sink (O , C , B , and S respectively).

Besides different environmental regions, we have different particle size classes. For a given size class, certain mass fractions will move to smaller size classes under the

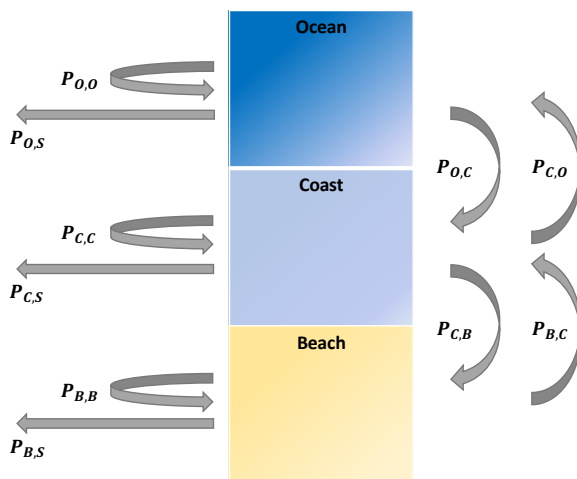


Figure 3.4: Environmental box model used to simulate Particle Size Distributions in the marine environment. Particles move between and within the ocean, coast, and beach boxes, and for each box there is a probability of being removed from the system.

influence of fragmentation. These fractions are estimated by evaluating Eq. (3.2) for the given time step of the box model. Similarly, Eq. (3.3) is evaluated to determine the abundance of fragments moving to smaller size classes. Fragmentation is assumed to only happen on the beach, where degradation is expected to be much more effective than in the sea (Andrady, 2011; Andrady et al., 1993).

Each environmental box contains a range of different particle size classes. The combination of environmental transition probabilities and fragmentation is modelled using a transition matrix. For example, taking 15 different size classes and 3 environmental compartments leads to a transition matrix of size 45×45 . Further details are given in the supporting material B.2.

3.2.3 Applying the box model to the Mediterranean Sea

We will demonstrate the capabilities of the environmental box model using a set-up based on the Mediterranean Sea. Environmental transition probabilities are derived from the literature on plastic transport in the Mediterranean Sea as much as possible. The different parameters, the studies on which they are based, and the areas of these studies are shown in table 3.1.

Transport in the marine environment. A Lagrangian simulation of floating plastic in the Mediterranean Sea (see Chapter 4) is used to determine transition probabilities within and between the ocean and coast ($P_{O,O}$, $P_{O,C}$, $P_{C,C}$, and $P_{C,O}$). The coast is defined as the ocean within 15 kilometers of the coastline.

Previous model studies have indicated that Stokes drift is able to push floating particles towards the coast, e.g. in the North Atlantic (Onink et al., 2019), the North Sea (Delandmeter & van Sebille, 2019), and in the Sea of Japan (Iwasaki et al., 2017). It has been hypothesised that this leads to near-shore trapping of larger

Table 3.1: Environmental box model parameters and fragmentation parameters, references used to estimate the parameter values, the respective study areas, and the estimated baseline parameter values.

| Parameter | Reference study or data | Reference study area | Baseline parameter value |
|-----------|--|----------------------|---|
| $P_{O,O}$ | Kaandorp et al. (2020) | Mediterranean | $7.2 \cdot 10^{-1} \text{ week}^{-1}$ |
| $P_{O,C}$ | Kaandorp et al. (2020) | Mediterranean | $2.7 \cdot 10^{-1} \text{ week}^{-1}$ |
| $P_{C,O}$ | Kaandorp et al. (2020) | Mediterranean | $3.4 \cdot 10^{-2} \text{ week}^{-1}$ |
| $P_{C,C}$ | Kaandorp et al. (2020) | Mediterranean | $8.3 \cdot 10^{-1} \text{ week}^{-1}$ |
| $P_{C,B}$ | Menna et al. (2017), Kaandorp et al. (2020) | Mediterranean | $1.3 \cdot 10^{-1} \text{ week}^{-1}$ |
| $P_{B,C}$ | Hinata et al. (2017) | Japan | $3.2 \cdot 10^{-2} \text{ week}^{-1}$ |
| $P_{B,B}$ | Hinata et al. (2017) | Japan | $9.6 \cdot 10^{-1} \text{ week}^{-1}$ |
| P_S | Cózar et al. (2015), Kaandorp et al. (2020) | Mediterranean | $5.1 \cdot 10^{-3} \text{ week}^{-1}$ |
| input | Kaandorp et al. (2020) | Mediterranean | $2,500 \text{ t year}^{-1}$ |
| p | Song et al. (2017) | South Korea | 0.4 |
| λ | Song et al. (2017) | South Korea | $1.8 \cdot 10^{-2} f \text{ week}^{-1}$ |

plastic particles, as more bouyant particles would tend to reside closer to the water surface, where they experience more influence from the Stokes drift (Isobe et al., 2014; Iwasaki et al., 2017).

We investigate the effect of this near-shore trapping of larger plastic particles and its influence on fragmentation. First, we calculate the transition probabilities assuming all particle sizes reside at the ocean surface, where they experience the maximum Stokes drift (corresponding to the transition probabilities between the ocean and coast in table 3.1). We then compute how differently sized particles are vertically distributed in the water column, and how this influences the lateral transport induced by Stokes drift. The approach from Poulain et al. (2019) is used to estimate particle rise velocities w_b for different particle sizes, see the supporting material B.3 for further details. From these rise velocities we calculate the median particle depth, using the particle density profiles from Kukulka et al. (2012). The Stokes drift is estimated at this depth, assuming a Stokes profile based on the Phillips wave spectrum (Breivik et al., 2016). For this Stokes drift, the transition probabilities are calculated using Lagrangian model simulations with different Stokes drift factors. In the end, this gives us different transition probabilities ($P_{O,O}$, $P_{O,C}$, $P_{C,C}$, and $P_{C,O}$) for each particle size. More information and resulting transition probability values are given in the supporting material B.1 and B.3.

Using the same approach as presented in the next Chapter (see supporting material C.2), we estimate $P_{C,B}$ by analysing drifter buoy data: from a set of 1682 drifters in the Mediterranean (Menna et al., 2017), we calculate how much time these drifters spend near the coast before beaching. For drifter buoys within 15 kilometers of the coastline, the beaching rate is estimated to be about $6.7 \cdot 10^{-3} \text{ day}^{-1}$ (corresponding to an e-folding time-scale τ_{CB} of 149 days). In Chapter 4 we estimate that τ_{CB} for plastic particles is about 3 times lower than that for drifter buoys. We will therefore use $\tau_{CB} = 50$ days as the baseline estimate here.

We use data from Hinata et al. (2017) to estimate residence times τ_{BC} of plastic particles on beaches, to obtain $P_{B,C}$ and $P_{B,B}$. This study was conducted on a beach

in Japan, no information could be found for the Mediterranean Sea specifically. We therefore assume that the Japanese setting is representative of the Mediterranean too: a sensitivity study for τ_{BC} is presented in the supporting material B.1. As a baseline estimate we use $\tau_{BC} = 211$ days, reported for small plastic floats (corresponding to the baseline $P_{B,C}$ in table 3.1). We will then investigate the effect of size-selective resuspension, for which the empirical relation from Hinata et al. (2017) is used, i.e.,

$$\tau_{BC}(w_b) = 2.6 \cdot 10^2 w_b + 7.1, \quad (3.5)$$

where τ_{BC} is given in days, and w_b in m s^{-1} .

The box model also requires transition probabilities for removal of particles: $P_{O,S}$, $P_{C,S}$, $P_{B,S}$. We assume these are the same in all compartments, denoted by P_S . A given value for P_S yields a certain amount of steady-state mass in the system. We take the estimated input of waste into the Mediterranean from Chapter 4 (2,500 metric tonnes for the year 2015), and the estimated total floating mass from C  zar et al. (2015) (2,000 metric tonnes). The value for P_S is iterated until this mass balance is satisfied, see the supporting material B.2 for more information.

Finally, we need to specify the plastic input into the marine environment in terms of location and shape. We assume that new plastic objects are introduced on the beach. This assumption does not affect results significantly, see the supporting material B.1. We use an initial length of 200 mm based on typical dimensions of municipal plastic waste in the Netherlands (Jansen et al. (2015), see the supporting material B.4, assuming that plastic product dimensions are similar to those used around the Mediterranean Sea. We use $D_N = 3$ as the baseline, i.e. cubical-shaped objects. The model time step is set to one week.

Fragmentation parameters. We use data from Song et al. (2017) to estimate the fragmentation parameter p , and to estimate the fragmentation rate λ specifying how much f increases per unit time.

In Song et al. (2017), plastic pellets were subjected to different levels of UV exposure and to 2 months of mechanical abrasion with sand, simulating a beach environment. The data for polyethylene (PE) and polypropylene (PP) pellets (26 mm^3 and 19 mm^3 respectively) are used, as these are the most abundant polymers in the Mediterranean surface waters (PE: 52–76%, PP: 7–16%; Pedrotti et al., 2016; Suaria et al., 2016).

We assume a single p value per material, and $D_N = 3$. The fragmentation index f is allowed to vary between the different levels of UV exposure when fitting the data. By fixing p and varying f , we get a robust estimate for the unknown parameter p for which we need a plausible value in the box model. We can expect that f is larger for particles subjected to longer periods of UV exposure, since embrittlement will make it easier for the mechanical abrasion to wear down the particles.

Resulting NSD fits using weighted least squares are presented in Figure 3.5, top row, fitted values for f are presented in the legend. For PE particles, the best fit results in $p = 0.39$, for PP particles $p = 0.45$. The experimental data are still at the early stage of fragmentation ($f < 1$), with few fragments per parent pellet, except for small fragment sizes.

There is a good fit for the PE data, with almost all simulations within the data errorbars (one relative standard deviation). For PP there is a good fit for 0, 2 and 6 months of UV exposure. At 12 months of UV exposure there is more mismatch for the smallest size class (0.05–0.10 mm). This is also the only case where the estimated f is lower than for the previous level of UV exposure.

The bottom row of Figure 3.5 compares the estimated volume fractions of the parent pellets and the fragments. Generally, the modelled volume fraction of the parent pellet is estimated reasonably well, although there is some overprediction for PE with 12 months of UV exposure. The modelled fragment volumes are higher than the ones estimated in Song et al. (2017). A possible explanation is that some of the larger fragments could have been missed in the experimental setting since there are very few of these per parent object (e.g. tenths or hundredths). One can see in Figure 3.3 that at an early stage of fragmentation ($f < 1$), the larger fragments contribute little to the total abundance of fragments (Figure 3.3a), but a lot to the total volume or mass of fragments (Figure 3.3b). In the experimental setting ten parent objects per sample were used, and fragments were counted under magnification on 0.7%–5% of the filter paper area. The larger fragments could therefore have been missed, or even have a low probability of actually existing.

Following this analysis, we set the baseline value of p in the box model to 0.4, within the range of the fitted p for PE and PP. A fragmentation rate λ needs to be chosen, specifying how much f increases per unit time. We assume that λ is a constant, meaning that the amount of fragmentation f is directly proportional to the time that particles spend on the beach.

In Song et al. (2017) 12 months of laboratory UV exposure were roughly related to 4.2 years of environmental exposure, representing beach conditions in South Korea. Regarding UV exposure, these conditions are quite similar to the Mediterranean (Herman et al., 1999). Taking our estimated fragmentation indices for PE and PP results in fragmentation rates λ of $1.8 \cdot 10^{-2} f \text{ year}^{-1}$ to $6.9 \cdot 10^{-2} f \text{ year}^{-1}$. The value for PE is used as the baseline here, given it is the most common polymer in the Mediterranean surface water (Pedrotti et al., 2016).

In Song et al. (2017) low-density polyethylene pellets are used: future studies are necessary to analyse how the results change for high-density polyethylene. We acknowledge that the fragmentation rate λ is still very uncertain, and more experimental research is necessary to verify whether the assumption that f varies linearly in time is a good approximation.

3.3 Results and discussion

3.3.1 Modelled environmental particle size distributions

Now that we have estimates for transition probabilities in the box model and estimates for the fragmentation parameters, we will simulate PSDs using a scenario based on the Mediterranean Sea. We will quantify the power law slope α of the results by numerically maximizing the log-likelihood ℓ of the data (Virkar & Clauset,

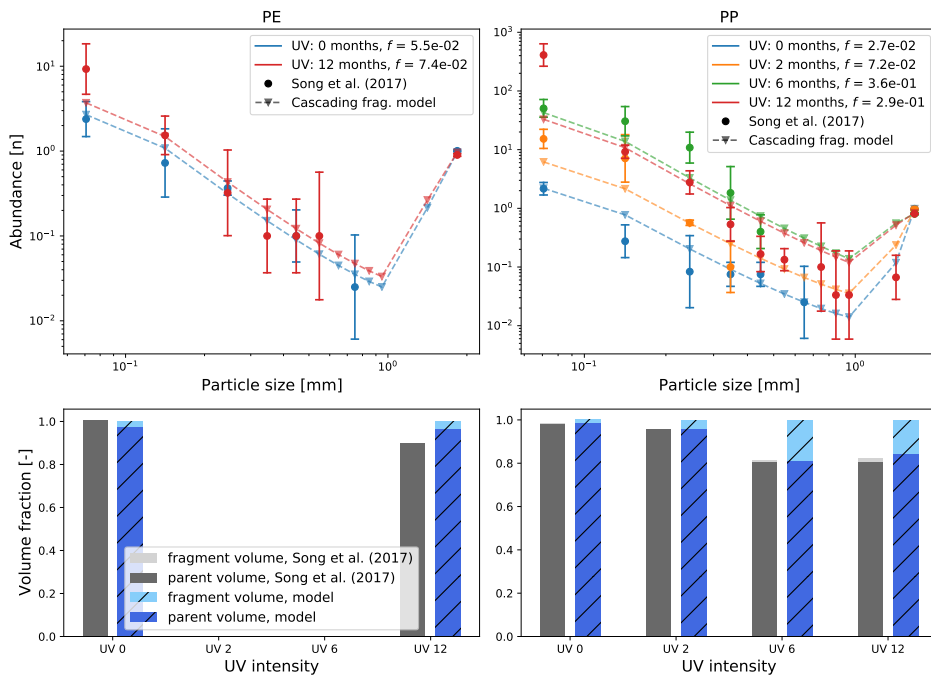


Figure 3.5: Calibration of the cascading fragmentation model with experimental data from Song et al. (2017). The experimental data is shown using the circles, including errorbars representing the relative error (one time the sample standard deviation). The fragmentation model results are shown using the triangles and dashed lines.

2014):

$$\ell = n(\alpha - 1) \ln b_{\min} + \sum_{i=\min}^k n_i \ln \left(b_i^{(1-\alpha)} - b_{i+1}^{(1-\alpha)} \right), \quad (3.6)$$

where b are the bin boundaries used to discretize the data, containing n_i samples in the bin with index i , and $n = \sum n_i$. In some cases, not the entire particle size range adheres to a power law. The lower bound of the power law domain is estimated by minimizing the Kolmogorov-Smirnov statistic between the modelled NSD and the theoretical power law NSD (Virkar & Clauset, 2014).

NSDs resulting from the box model are shown in Figure 3.6, corresponding MSDs and a table with parameter settings can be found in the supporting material B.1. Fragmentation is expected to increase the fraction of small particles, increasing α over time (see Figure 3.3). However, environmental sinks limit the magnitude of α : assuming a constant removal rate of plastic particles, smaller fragments, which tend to be older, have a higher probability of being removed from the environment. This combination of fragmentation and environmental sinks eventually leads to an equilibrium, or statistical steady state. This is illustrated in Figure 3.6a using the box model with the baseline parameters described in Section 3.2.2. As time progresses, the relative proportion of fragments to parent objects increases. In this scenario, it takes on the order of years for the NSD to resemble the steady state (red dashed line). The magnitude of the environmental sinks is high enough to avoid long persistence of fragmented particles: there are still relatively many parent objects, and $\alpha = 2.57$ is still below the value derived from the fractal dimension of $\alpha = 2.67$ from Eq. (3.4).

Steady state NSDs for different scenarios are presented in Figure 3.6b. Results for the baseline parameters (blue lines) almost overlap with the results where size-selective lateral transport is added to the box model, induced by vertical mixing and Stokes drift (orange lines). In the baseline scenario $\alpha = 2.57$ for all three NSDs. When adding size-selective ocean transport, larger particles tend to move more frequently from the ocean to the coast. This results in slightly more small particles in the ocean box, increasing the power law slope here to $\alpha = 2.73$. Adding size-selective resuspension of particles (Hinata et al., 2017) has a strong effect (green lines). Bigger objects have longer residence times on the beach, and therefore undergo more fragmentation. This produces a large number of smaller fragments with shorter residence times, which therefore move more rapidly to the coastal and ocean cells. This near-shore trapping of larger plastic objects was already hypothesized in e.g. Isobe et al. (2014). The empirical resuspension relation Eq. (3.5) causes the model to deviate from a power law, the domain over which α is calculated is shaded in green in Figure 3.6. The model yields $\alpha = 2.69$ on the beach, which is lower than in the coastal and ocean cells (both $\alpha = 3.37$). A scenario where the fragmentation rate is based on PP instead of PE is presented (red lines). Fragmentation breaks down the particles more quickly: a monotonic relation between particle size and abundance is observed, with $\alpha = 3.03$. Finally, a scenario is presented (purple lines) where the input of plastic waste into the Mediterranean is 100,000 tonnes per

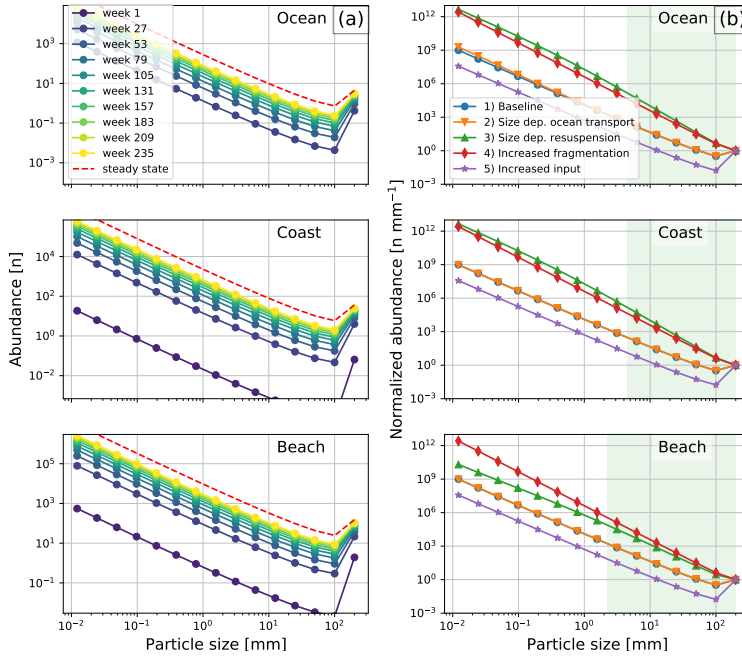


Figure 3.6: Modelling NSDs using the environmental box model. Column a: transient response to a constant input of particles into the model (baseline scenario). Column b: steady state normalized NSDs for different environmental scenarios, normalized to the amount of parent particles (200 mm). The approach from Virkar & Clauset (2014) is used to determine the power law size range: for most scenarios, α is calculated over the entire size range except for scenario 3, which is shown using the green shading.

year (Liubartseva et al., 2018; Jambeck et al., 2015), instead of the aforementioned 2,500 tonnes per year (from Chapter 4). The magnitude of the sinks needs to be much larger now to attain a mass balance based on 2,000 tonnes of floating plastics (Cózar et al., 2015). Fragmentation has little time to break down the particles, resulting in relatively few fragments per parent object.

In Figure 3.7, we compare PSDs resulting from the box model with observed ones in the Mediterranean Sea. In the model results we include both size dependent ocean transport and resuspension. Fragmentation parameters are set to $\lambda = 2 \cdot 10^{-4} \text{ f week}^{-1}$, and $D_N = 2.5$, resulting in good agreement with the observed PSDs. The effect of vertical turbulent mixing of fragments using the model from Poulain et al. (2019) is shown as well (calm, $U_{10} \approx 4 \text{ m s}^{-1}$, and above average, $U_{10} \approx 7 \text{ m s}^{-1}$ conditions based on the 30% and 70% quantile of Mediterranean sea weather conditions (Hersbach et al., 2020), see supporting material B.3, assuming a submerged net depth of 25 centimeters (similar to e.g. Cózar et al., 2015)).

Figure 3.7a presents the resulting NSDs. We compare model results in the ocean cell

with measurements by C  zar et al. (2015), and results in the coastal cell with measurements by Ruiz-Orej  n et al. (2018), as these were mainly obtained further away from the coast or close to the coast respectively (see the supporting material B.3, Figure B.7). Under calm wind and wave conditions there is good agreement between the modelled and observed NSDs. Vertical mixing causes the modelled NSDs to deviate from the power law around <3 millimeters, similar to the measured NSDs. Many of the smaller fragments are expected to be mixed below the net depth, resulting in measuring only a fraction of small fragments. This, combined with a size detection limit effect where elongated particles escape from meshes smaller than their maximum length (Abeynayaka et al., 2020; Enders et al., 2015; Tokai et al., 2021), could explain a part of the underabundance of sub-millimeter fragments in observations. Measurement campaigns with much smaller size-detection limits than the standard neuston nets (see e.g. Enders et al., 2015 or Kooi & Koelmans, 2019) show increasing abundances for sub-millimeter fragments. It is therefore unlikely that the underabundance of sub-millimeter fragments is explained by an increased loss of these particles, suggested in some studies (C  zar et al., 2014; Pedrotti et al., 2016).

Including vertical mixing has a strong effect on the estimated power law slope α . An overview of the estimated α and the power law size range for Figure 3.7a is given in the supporting material B.1. In the ocean cell, $\alpha = 2.73$ without vertical mixing, and decreases to $\alpha = 2.63$ for calm conditions ($U_{10} \approx 4$ m/s), and to 2.37 for above average conditions ($U_{10} \approx 7$ m/s). A similar decrease is observed for the coastal cell: $\alpha = 2.69$ without mixing, decreasing to $\alpha = 2.60$ ($U_{10} \approx 4$ m/s) and $\alpha = 2.34$ ($U_{10} \approx 7$ m/s). Similar to the model, a slightly lower α is calculated for the measurements near the coast (2.49 ± 0.06) compared to measurements further away from the coast (2.53 ± 0.04), although this difference is not significant.

Few PSD measurements are available for beaches. Two examples are shown in Figure 3.7a: one from the Mediterranean (Constant et al., 2019), and one for which both the NSD and MSD were available (Fok et al., 2017). The measurements on beaches have much lower power law slopes ($\alpha < 1.60$) compared to measurements in the water. This is also captured by the model, meaning size-selective resuspension indeed seems to play an important role. In the beach cell the modelled power law slope of $\alpha = 2.02$ is higher than the measured ones, which might indicate that size-selective beaching should be taken in account as well.

Figure 3.7b presents the modelled MSDs. Vertical mixing has a large influence on the measured mass for small particle sizes: even under calm conditions, the measured mass for particles of 0.1 mm is almost three orders of magnitude lower than without mixing. Unfortunately, there is very limited observational data reporting MSDs, so the comparison to data is more limited than for the NSDs in Figure 3.7a. On beaches, the model matches the set of measurement well, but more data are necessary to further verify this. Large fragments are expected to dominate in terms of mass on beaches. In the water, α seems to be approximately zero on average. This would mean that the mass contribution would scale roughly quadratically for an increasing size class k , i.e. large fragments also dominate in terms of mass here.

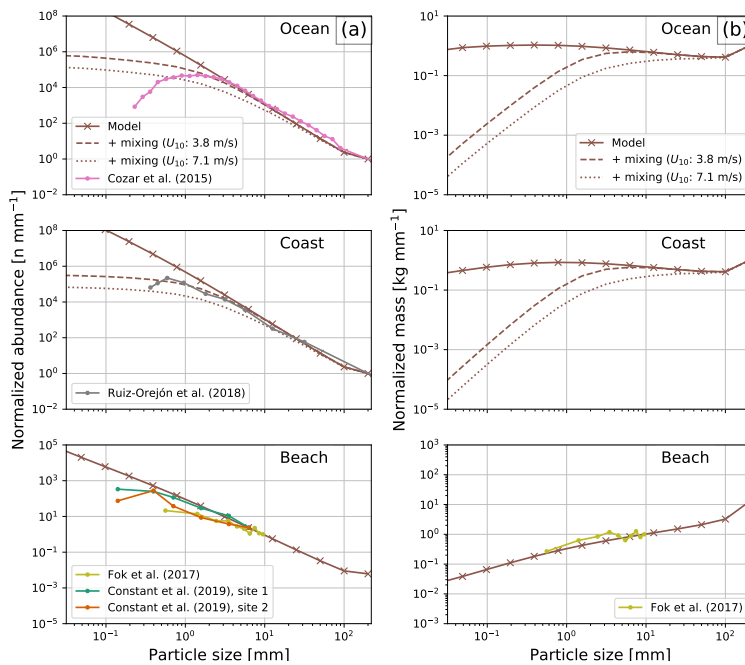


Figure 3.7: Comparison of modelled and measured PSDs, along with effects of vertical mixing induced by turbulence on measured particle sizes. PSDs are normalized to the maximum observed size class, if no measurements are available to the maximum modelled size class. Column a: comparison of NSDs. Column b: comparison of MSDs.

The environmental box model used to model the PSDs is a useful tool for future mass balance studies. The steady-state with the model settings used for Figure 3.7, gives that about 98% of the mass in the system is on the beach, about 2% in the coastal surface water, and about 0.2% in the surface open ocean. This large fraction of plastics stranding is in good agreement with previous mass balance estimates (Lebreton et al., 2019). It should be noted that other environmental regions, like the ocean floor, are not included in these numbers as these are part of the sinks in the box model (P_S), which continuously take up more mass over time. Secondary microplastics generation can be estimated: for the same model settings, about $6.5 \cdot 10^{-5}\%$ of the macroplastic (> 5 mm) mass breaks down into microplastics per week, about $2.0 \cdot 10^{-6}\%$ of microplastics become smaller than 0.1 mm. This is orders of magnitude smaller than the estimated sinks, taking up about $5.0 \cdot 10^{-2}\%$ of the plastic mass per week. Longevity of plastics can be estimated: taking a sudden stop of new plastics entering the marine environment, it would take about 176 years for 99% of the plastic mass to disappear from the surface water and beaches. This is a much longer time scale than given by the conceptual model in Koelmans et al. (2017), where for a similar stop of new plastics, almost all plastic mass was removed from the ocean surface layer within three years. Plastic residence times are highly dependent on the input scenarios: in Koelmans et al. (2017), 3% of the world plastic pollution was estimated to enter the ocean. Here, the input scenario from Chapter 4 is used for the Mediterranean Sea, where less than 0.1% of plastic waste from coastal population was estimated to enter the marine environment. These differences show the importance of further mass balance studies to constrain this number.

3.3.2 Model limitations

We will give a brief overview of the fragmentation model and environmental box model limitations in this section, which can be addressed in future studies.

The fragmentation model presented here only has few parameters (p , f , and D_N), and assumes that the fragmentation process is scale-invariant. One example where the assumption of scale-invariance might not hold is when only the plastic surface layer gets brittle with microcracks (Andrady, 2011). This possibly increases the fragmentation rate below a certain length scale, dependent on how far UV radiation penetrates the polymer and the polymer type. It is assumed that f is directly proportional to the time that particles spend on the beach, leading to a constant fragmentation rate λ . In e.g. Charalambous (2015), it was shown that grinding can become less efficient as particles become smaller, which might lead to e.g. a logarithmic relation instead.

One source of uncertainty in the fragmentation model is the fact that parameters are calibrated using experimental data for low-density polyethylene and polypropylene pellets only (Song et al., 2017). More research is necessary to quantify how fragmentation differs between low-density and high-density polyethylene, and for other polymers found in the environment not taken in account here (e.g. polyamides and polystyrene, which form a substantial fraction of polymers found in the Mediterranean Sea surface water (Pedrotti et al., 2016)). PSDs are quite sensitive to the choice of λ (see Figure 3.6) and p (see the supporting material B.1), for which the values are still quite uncertain. More experimental data are required to further con-

strain these parameters and to estimate how they vary globally. Finally, it is still uncertain which processes can extend the degradation process to finer scales than mechanical abrasion, and the magnitude of their influence (e.g. photochemical oxidation (Ward et al., 2019), or biodegradation (Gerritse et al., 2020; Molitor et al., 2019)).

The environmental box model is an idealized section of the marine environment, only considering overall transport between cells representing the open ocean, coastal water, and beach. Regional influences are not taken in account and should be investigated in the future: think, for example, about different types of coasts (Weideman et al., 2020) with different particle residence times and beaching timescales (Samaras et al., 2014). A number of assumptions were made to arrive at the environmental box model. We assume that the majority of plastic particles are fragments (Cózar et al., 2014; Suaria et al., 2016), which means that the influence of primary plastics such as resin pellets (Turner & Holmes, 2011) or plastic beads from consumer products (Fendall & Sewell, 2009) is neglected. It is assumed that fragmentation is dominant on beaches (Andrady, 2011). Fragmentation in the water column is neglected, which might be for example induced by hydrolysis and biodegradation (Gerritse et al., 2020), or ingestion and scraping by marine organisms (Reisser et al., 2014; Mateos-Cárdenas et al., 2020). It is furthermore assumed that new plastic particles are introduced on the beach. In reality there will be a combination of inputs into the different environmental compartments depending on the sources. In the supporting material B.1 it is shown that this assumption has no significant effect on the results. Finally, it is assumed that the rate of plastic removal from the marine environment is constant. Although we are looking at time scales in the order of years here, seasonality might have an effect on the removal rate by influencing, for example, biological activity.

We demonstrated the model capabilities using a set-up based on the Mediterranean Sea. One source of uncertainty is that the resuspension time scale is obtained from Hinata et al. (2017), based on experiments at a Japanese beach. A sensitivity study for this parameter is given in the supporting material B.1. Future studies should look at how this parameter varies for different beaches globally. The size of new plastics introduced into the marine environment is still uncertain: it is fixed to 200 mm here, while in reality this will be a spectrum of different sizes, see the supporting material B.1 and B.4 for more information.

3.3.3 Fragmentation models and size distribution data

The cascading fragmentation model by Charalambous (2015) used in this work, shows quite good correspondence with experimental data from Song et al. (2017) (see Figure 3.5). A benefit of the fragmentation model presented here, is the ability to model the mass size distribution (MSD) as well, which can help us obtain a better understanding of the marine plastic mass budget. MSD data to validate the model is currently lacking however, for example to verify whether the larger size classes indeed make up most of the environmental plastic mass. More PSD data from beaches would allow for better constraining residence times of plastic particles on beaches and in coastal waters, and more data from marine sediment might give insight in the role of size-selective sinking, induced by e.g. biofouling (Kooi et al., 2017).

Fragmentation models for plastics have been introduced in previous works, such as in C  zar et al. (2014). They focused mainly on spatial dimensionality: $\alpha = 3$ in the NSD was related to three-dimensional fragmentation, i.e. a cube splitting into 8 smaller cubes. Care should be taken in future studies that when working with logarithmic binning, the normalized NSD (units: n mm^{-1}) slope decreases by one compared to the discrete NSD (units: n), see Figure 3.3. This was overlooked in C  zar et al. (2014): $\alpha = 3$ would correspond to two-dimensional fragmentation with their model, see the supporting material B.5 for further explanation. Normalization is also important to take into account when describing plastic particle size in terms of a probability density function (Kooi & Koelmans, 2019), specifying the probability per unit length (units: mm^{-1}). Finally, estimating α is not trivial: fitting straight lines on log-log transformed data induces large biases, maximum likelihood approaches are more suitable, see e.g. Newman (2005) and Virkar & Clauset (2014).

3.4 Conclusions and recommendations

In this work, we modelled particle size distributions (PSDs) of plastics in the marine environment, by considering a cascading fragmentation model, and a box model taking in account size-selective transport between the open ocean, coastal water, and beach. We showed that the cascading fragmentation model is able to explain the power-law observed in PSDs from the environment and experimental fragmentation studies, that size-selective transport plays an important role near the coast, and that vertical mixing in the water column has a strong impact on measured PSDs.

Understanding the nature of PSDs and how they differ in environmental regions can help us get a better understanding of the marine plastic mass budget. Previous conceptual mass balance studies, such as the ones by Koelmans et al. (2017) and Lebreton et al. (2019), did consider fragmentation, but only for 2 or 3 categories (macro-, micro- and nanoplastics). Here, we model a range of size classes. This way, we can not only predict which environmental compartment contains most plastic mass, but also which size ranges.

We applied the combined fragmentation and environmental box model to a scenario based on the Mediterranean Sea. For the steady-state, we estimate that of buoyant plastics about 98% of plastics reside on beaches, about 2% in coastal surface waters, and about 0.2% in the surface open ocean. On one hand, the model predicts fragmentation to play an important role in terms of generating a large number of plastic fragments. On the other hand, fragmentation seems to play a minor role in the mass budget compared to other environmental sinks, by moving mass from large to small particle size classes.

Overall, the idealized model presented here is a valuable tool to efficiently test hypotheses regarding the marine plastic mass budget. It can be checked whether a certain hypothesis leads to results which are consistent with current knowledge of plastic sources, sinks, transport, fragmentation, and observational data of plastic concentrations and PSDs. At a later stage, the fragmentation model could also be applied to more complex physical models, taking into account spatial and temporal variability of plastic transport in the marine environment.

Closing the Mediterranean mass budget of buoyant marine plastics: inverse modelling of sources and sinks

Estimates of plastic inputs into the ocean are orders of magnitude larger than what is found in the surface waters. This can be due to discrepancies in the sources of plastic released into the ocean, but can also be explained due to the fact that it is not well known what the most dominant sinks of marine plastics are, and on what time scales these operate. To get a better understanding on possible sources and sinks, an inverse modelling methodology is presented here for a Lagrangian ocean model, estimating floating plastic quantities in the Mediterranean Sea. Field measurements of plastic concentrations in the Mediterranean are used to inform parameterisations defining various sources of marine plastics, and removal of plastic particles due to beaching and sinking. The parameters of the model are found using inverse modelling, by comparison of model results and measurements of floating plastic concentrations. Time scales for the sinks are found, and likely sources of plastics can be ranked in importance. A new mass balance is made for floating plastics in the Mediterranean: for 2015 there is an estimated input of 2,100–3,400 tonnes, and

This chapter has been published as:

Kaandorp, M.L.A., Dijkstra, H.A., van Sebille, E. (2020): Closing the Mediterranean Marine Floating Plastic Mass Budget: Inverse Modeling of Sources and Sinks. *Environmental Science and Technology*, 54(19):11980–11989. <https://doi.org/10.1021/acs.est.0c01984>

of plastics released since 2006, about 170–420 tonnes remain afloat in the surface waters, 49–63% ended up on coastlines, and 37–51% have sunk down.

4.1 Introduction

It is currently not well known what happens with plastics once they end up in the marine environment. Studies have shown that only a fraction of plastics which are expected to enter the oceans remains afloat in the surface water. The total mass of floating plastics in the global ocean has been estimated to be from 93–236 thousand tonnes (van Sebille et al., 2015), to at least 269 thousand tonnes (Eriksen et al., 2014). This is significantly different from the total input estimates into the marine environment, which range from 4.8–12.7 million tonnes per year from coastal population (Jambeck et al., 2015), to 1.15–2.41 million tonnes from rivers only (Lebreton et al., 2017). This is not taking into account other possible sources resulting from activities such as fishing, aquaculture, and shipping (Lebreton et al., 2018).

One can investigate different environmental compartments where the remainder of plastics might reside, such as shorelines. Another possibility is the deep ocean and marine sediments: biofilm formed by micro- and macro-organisms (Fazey & Ryan, 2016; Lobelle & Cunliffe, 2011) and fecal pellets (Cole et al., 2013) can cover plastic particles, increasing the average density and therefore induce sinking. Plastic particles might be present in biota: e.g. zooplankton, fish, or birds (Cole et al., 2013; Güven et al., 2017; Franeker et al., 2011). Oxidation caused by UV-exposure can make polymers more brittle, enhancing the fragmentation of plastics in the environment (Song et al., 2017): particles might become too small to measure using conventional techniques.

Estimates have been made in which environmental compartments the marine plastics are likely to reside. Marine sediments are likely to contain a major percentage of plastics, e.g. more than 90% of microplastics in terms of numbers for a global scenario (Booth et al., 2017), with abundances of about four orders of magnitude higher per unit volume of sediment than that found in surface waters in the oceanic gyres (Woodall et al., 2014). Other studies cite the possibility that shorelines store the majority of plastics (Lebreton et al., 2019), and that coastal fluxes possibly dominate mass fluxes to the sea bottom (Liubartseva et al., 2018).

In this chapter, a framework is presented to close the plastic mass budget, by combining the strength of numerical models and in-situ measurements (Maximenko et al., 2019). Models allow us to estimate plastic concentrations continuously over time on a large spatial domain, but there are still a number of unknowns regarding processes affecting the dispersal of marine debris (van Sebille et al., 2020). Measurements of plastic concentrations as obtained by e.g. neuston net trawls give us more reliable information at a given instance at a specific location. However, these are expensive to carry out, and can be prone to high variation due to a relatively small area covered, high heterogeneity of plastic concentrations (Pedrotti et al., 2016), and presence of waves (DiBenedetto, 2020). By using an inverse modelling approach, the best of modelling and observations is combined.

Here, parameters in a numerical model governing sources and sinks of marine plastics are estimated using observed plastic concentrations in surface waters. A Bayesian framework is used, where prior information can be specified for the parameters based on previous (experimental) findings. After the posterior step is done, the estimated parameters are used to quantify where, and in which environmental compartments, most of the marine plastics are expected to reside. Here, we choose to focus on the Mediterranean, which is an interesting test case because of two reasons. First of all, numerical studies and field measurements suggest that there are no stable plastic retention areas in the basin due to variability of the surface currents (Cózar et al., 2015), making it important to take time-varying processes into account. Secondly, a large number of field studies measuring plastic concentrations are available, providing valuable information that can be used to train numerical models. We choose to focus on two major sinks of plastics: sinking down of plastics, and plastics ending up on coastlines (beaching). Other sinks, such as fragmentation, degradation, and ingestion of plastics by animals are neglected, based on the assumption that the removal rates of these sinks are likely at least an order of magnitude smaller.

4.2 Methods

4.2.1 Lagrangian framework and forcing

In the Lagrangian framework presented here, virtual particles represent floating plastics. The *OceanParcels* Lagrangian ocean analysis framework (Delandmeter & van Sebille, 2019) is used to calculate the movement of floating plastic within a given velocity field. Trajectories are integrated using a Runge-Kutta 4 scheme. The velocity field is derived from E.U. Copernicus Marine Service Information reanalysis data for the Mediterranean currents at a $1/16^\circ$ resolution (Simoncelli et al., 2019), and hindcast data for the Stokes drift at a $1/24^\circ$ resolution (Korres et al., 2019), both spanning the years 2006–2016. Like other Lagrangian modelling studies (Macias et al., 2019; Lebreton et al., 2018), it is assumed that the plastic particles move just below the water surface, and hence do not experience a direct wind drag.

The effects of subgrid-scale phenomena such as submesoscale eddies are parameterised using a zeroth-order Markov model (van Sebille et al., 2018), with a constant tracer diffusivity K . While some experimental estimates have been done estimating this diffusivity parameter (Okubo, 1971), it is difficult to determine an appropriate value, also because it will vary spatially (Rühs et al., 2017). Here, three different (constant) values for K are used, namely $K = 1 \text{ m}^2 \text{ s}^{-1}$, $K = 10 \text{ m}^2 \text{ s}^{-1}$, and $K = 100 \text{ m}^2 \text{ s}^{-1}$, to determine the sensitivity to this parameter.

The number of virtual particles should be large enough for the results to be statistically significant. First, a baseline simulation is done with about 1.2 million particles. A certain percentage of the plastic particles will disappear from the surface water over time due to sinking and particles ending up on coastlines. A time threshold at which 99.9% of the plastic particles in the baseline simulation are removed was determined to be approximately 50 days. Subsequent simulations were done with about 7.2 million particles, where particles were removed well above this threshold

Table 4.1: Data used of plastic concentration measurements in the Mediterranean.

| Reference | n km ⁻² | g km ⁻² | Sampling year | Size classes measured |
|--|--------------------|--------------------|---------------|-----------------------|
| Collignon et al. (2012) | ✓ | | 2010 | < 5 mm |
| Collignon et al. (2014) | ✓ | | 2011-2012 | < 5 mm, > 5 mm |
| Cózar et al. (2015) | ✓ | ✓ | 2013 | < 5 mm, > 5 mm |
| Fossi et al. (2012) | ✓ | | 2011 | < 5 mm |
| Gajšt et al. (2016) | ✓ | ✓ | 2012-2014 | < 5 mm, > 5 mm |
| Galgani (2011) (unpublished, Gerigny et al., 2018) | ✓ | ✓ | 2011 | < 5 mm, > 5 mm |
| Galgani (2012) (unpublished, Gerigny et al., 2018) | ✓ | ✓ | 2012 | < 5 mm, > 5 mm |
| Gündoğdu & Çevik (2017) | ✓ | | 2016 | < 5 mm |
| Gündoğdu et al. (2018) | ✓ | | 2016-2017 | < 5 mm |
| Güven et al. (2017) | ✓ | | 2015 | < 5 mm |
| de Haan et al. (2019) | ✓ | ✓ | 2015 | < 5 mm, > 5 mm |
| van der Hal et al. (2017) | ✓ | | 2013-2015 | < 5 mm, > 5 mm |
| Pedrotti et al. (2016) | ✓ | | 2013 | < 5 mm, > 5 mm |
| Ruiz-Orejón et al. (2016) | ✓ | ✓ | 2011-2013 | < 5 mm, > 5 mm |
| Ruiz-Orejón et al. (2018) | ✓ | ✓ | 2014 | < 5 mm, > 5 mm |
| Suaria et al. (2016) | ✓ | ✓ | 2013 | < 5 mm, > 5 mm |
| Zeri et al. (2018) | ✓ | | 2014-2015 | < 5 mm |

(after 180 days), see the supporting material C.4.

The beaching of particles is parameterised using a model presented later in this chapter. Particles should therefore not move from mesh cells belonging to the ocean onto land cells due to other processes, such as interpolation errors or Stokes drift. This is ensured by pushing particles back towards the closest ocean cell when they have ended up on the land, identical to what is done in Delandmeter & van Sebille (2019).

4.2.2 Area of interest and field measurements used

The area studied here is the Mediterranean. The high spatio-temporal variability of the currents in this basin causes that there are no known plastic retention areas (Cózar et al., 2015). In order to get a better picture of the flow field, the time-mean surface currents over 2006-2016 have been plotted as vectors in Figure 4.1. In the same figure, locations of the measurements used here are plotted for which references are shown in the legend. Two types of measurements are used here (Table 4.1): manta trawl or neuston net samples reported in terms of abundance (counts per square kilometer, n km⁻²) and in terms of mass (grams per square kilometer, g km⁻²). A majority of the measurements were taken in the western basin of the Mediterranean. There are much fewer measurements in the eastern basin, which are mainly found in front of the coast of Turkey and Israel.

Two types of correction factors are used for the measurements: one for wind-induced vertical mixing, and one accounting for different measured particle sizes. For wind-induced vertical mixing the correction factor from Kukulka et al. (2012) is used, see the supporting material C.1.

We want to account for all plastic particle sizes which are larger than the mesh-size of the neuston nets. If the data are available, measurements of microplastics (< 5 mm) and macroplastics (> 5 mm) are combined. If the data are given for < 5 mm only, a correction factor is used. This correction factor is calculated from the available measurements reporting both size classes. In terms of abundance a correction factor of 1.14 (standard deviation: 0.14) was calculated. In terms of

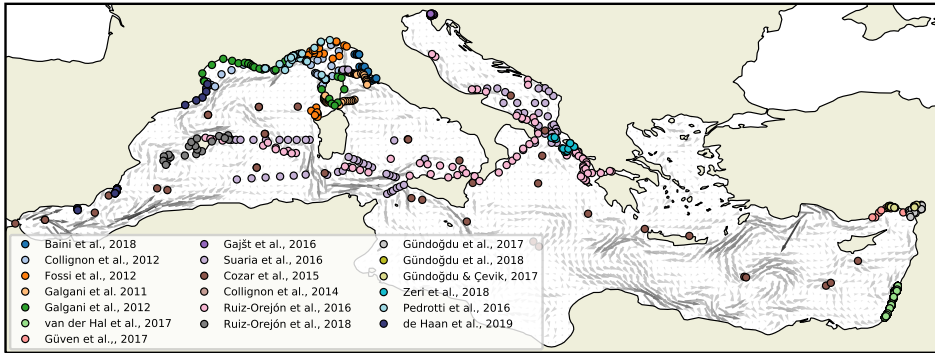


Figure 4.1: Available plastic measurements used here (coloured dots), and the time-mean surface currents over 2006-2016 (grey arrows)

mass only measurements are used where both microplastics and macroplastics were reported, so no correction factor is necessary in this case. Table 4.1 presents the size-classes reported for each study.

The model output and measurements are transformed to a \log_{10} scale for comparison. Measured values of plastic concentrations span multiple orders of magnitude. Not transforming the data would lead to high outliers dominating the inverse modelling process, while discrepancies at lower concentrations are just as relevant as those at higher concentrations.

As shown in e.g. de Haan et al. (2019), replicate samples taken of plastic concentrations reveal a lot of variability. This variability was calculated on a specific length and time scale using an empirical variogram, see the supporting material C.1. The model used here has a spatial resolution of $1/16^\circ$, and a temporal resolution of one day. The variance of the measurements at this length and time scale, denoted by γ , is $\gamma_n = 0.1376$ (units: $[\log_{10}(n \text{ km}^{-2})]^2$) for the abundance measurements, and $\gamma_m = 0.2201$ (units: $[\log_{10}(g \text{ km}^{-2})]^2$) for the mass measurements. When comparing model output to the observations, this variance is used to specify the measurement uncertainty, since fluctuations on length and time scales smaller than these are not resolved by the model.

4.2.3 Sources of plastics

Different release scenarios for plastics entering the marine environment are considered here. In modelling studies, the sources of marine plastics are often divided into different classes. In one recent example for a global scenario (Lebreton et al., 2018), 59.8% was estimated to come from the coastal population ($<50\text{km}$ from the coastline), 12.1% from inland population by riverine transport, 17.9% from fisheries, 1.3% from aquaculture, and 8.9% from shipping. Since the proportions for

the Mediterranean might be significantly different, we make no such assumption here. Instead, the model selects the appropriate fractions of input waste such that there is a good fit of the model with the observed plastic concentrations, consistent with their error estimates. The model can select from three major possible plastic sources as shown in Figure 4.2, which were estimated to be the biggest sources of pollution in Lebreton et al. (2018).

Firstly, input from rivers is considered, using the results from Lebreton et al. (2017). In Figure 4.2 the yearly waste is plotted using green circles, where only rivers estimated to release more than 0.2 tonnes of plastic per year are shown. In the model, monthly estimates of plastic emissions are used for all rivers available. In Lebreton et al. (2017), lower, mid, and upper estimates for riverine plastic input were given. This is represented in our model by including a parameter varying from -1 to 0 to 1, corresponding to the lower, mid, and upper estimates respectively. The parameter is allowed to vary continuously in phase space, linear interpolation is used to determine the riverine output for intermediate values. The input from rivers is given instantaneously at the river mouth, possible delayed response due to e.g. transport in the river itself (van Emmerik & Schwarz, 2020; Castro-Jiménez et al., 2019) are not taken in account.

Another possible source of marine plastics is fishing activity (shown in blue in Figure 4.2). Data for the fishing intensity were obtained from the global fishing watch (Kroodsmas et al., 2018). These data are based the automatic identification system (AIS) installed on vessels. This system has not been equally present on fishing vessels over the years, and no data was available from before 2012. It was therefore decided to assume a constant fishing intensity over the years, based on the years 2012-2016.

Finally, land-based mismanaged plastic waste (MPW) from coastal population is considered (shown in red in Figure 4.2 for 2010). The mismanaged plastic waste density was estimated by overlaying population density data (SEDAC et al., 2005) with the estimated mismanaged plastic waste per capita per country (Jambeck et al., 2015). The estimated population density data are available from 2000 to 2020 in increments of 5 years. It is linearly interpolated to estimate the population density at a given moment in time. The number of particles released along the coast is proportional to the MPW production within 50 km, similar to Jambeck et al. (2015).

Input of plastics from the Atlantic is neglected here. In Cózar et al. (2015) a plastic concentration of 159 g km^{-2} was reported inside the Strait of Gibraltar. Taking the width of the strait, and the mean surface current which was calculated to be 0.55 m/s, would theoretically lead to about 40 tonnes of plastic per year. Since this is small compared to the previously mentioned sources, and as it would require a separate model to estimate how this source varies over time, it is not taken in account here.

4.2.4 Parameterisation of plastic particle properties

Each virtual particle in the Lagrangian framework represents a certain abundance (n) and mass (g) of plastic particles, similar to the super-individual approach used for microbial modelling (Hellweger & Bucci, 2008). The concentration of plastics in

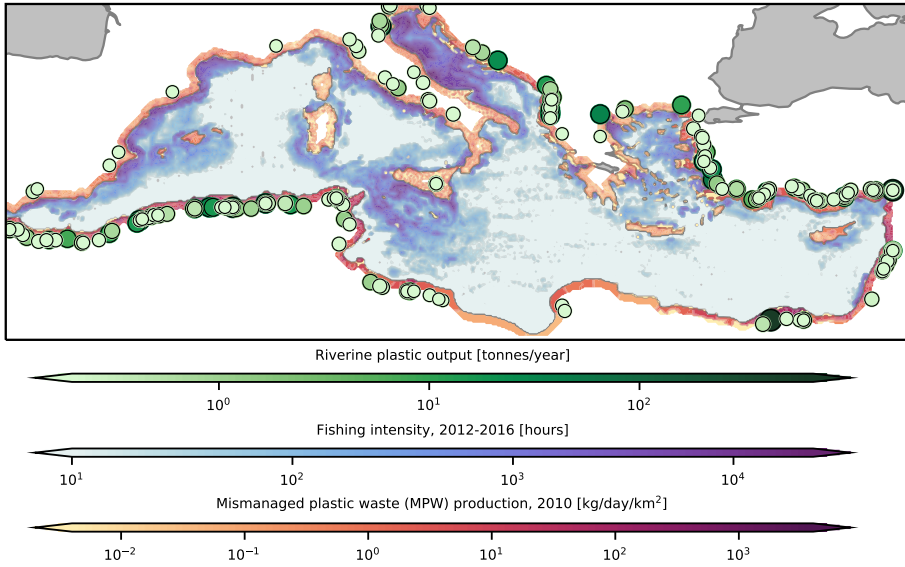


Figure 4.2: Sources of marine plastics used in the model. The amount of virtual particles released from each source is proportional to the magnitude of the source as plotted here. Virtual particles released from rivers can directly be expressed in terms of mass, the other sources are defined in relative terms to the riverine input.

n km^{-2} and g km^{-2} is calculated by taking a weighted kernel density estimate (KDE) (Guillamón et al., 1998) of all virtual particles, weighted by the total abundance or total mass of plastic particles inside the virtual particle respectively.

Initially, the abundance and mass of plastic particles inside the virtual particle depends on the particle's source, since one source might contribute more to the total plastic pollution compared to the other. Over time, the abundance and mass of the virtual particle are modified by sinks acting upon it. It is e.g. assumed that the collection of plastics inside the virtual particle has a constant probability of beaching over time when nearby the coast. This leads to an exponential reduction of the abundance and mass of the virtual particle on a certain time scale τ_{beach} .

This Lagrangian approach, which assigns an abundance and weight to the virtual particles, allows for a relatively quick evaluation of different parameter sets compared to e.g. a continuum approach with a plastic tracer concentration. Another benefit is that it is easy to use reanalysis data sets for the forcing fields which have already been assimilated with observational data.

Two sinks of plastic particles are considered: beaching and sinking. Each sink has its own fraction defining what percentage of the plastics is still floating and not taken away by the sink, denoted by f_{beach} and f_{sink} respectively. The weight of the virtual particle is the product of the weight at its source (w_{source}), with these different

factors:

$$w_{ptcl} = w_{source} \cdot f_{beach}(x, t) \cdot f_{sink}(t). \quad (4.1)$$

Here, w_{source} can be expressed in mass ($w_{source,m}$) or in abundance ($w_{source,n}$); the same holds for the particle weight ($w_{ptcl,m}$ or $w_{ptcl,n}$). The value of w_{source} depends on which of the three sources the particle comes from; this source is kept track of for each particle during the simulation. For riverine sources there is an estimate available of their individual pollution per month in tonnes (Lebreton et al., 2017): $w_{source,m}$ can directly be calculated. The rest of the sources are expressed in relative terms to the riverine sources, to convert these to tonnes as well. This leads to two parameters in the model defining the source ratios: $S_{pop:riv}$ and $S_{fis:riv}$, where the subscripts *pop*, *riv*, and *fis* denote the sources from coastal population, rivers, and fisheries respectively. A prior probability density function needs to be defined for these parameters in the Bayesian framework used here. Bounds for the prior, defined in terms of the 99.7th percentile of a Gaussian distribution, are set to enable a very wide range of possibilities ($\frac{1}{20}$ –20), such that each source can contribute at most to 95% of the total mass. This easily captures the possible release scenarios mentioned in the previous section (Lebreton et al., 2018).

For the total abundance of particles emitted by different sources no estimates could be found. In order to express $w_{ptcl,n}$ in terms of abundance (n), a linear fit through the origin is made of the modelled (unitless) concentrations versus the measured concentrations, see also van Sebille et al. (2015). The slope of this fit is used to assign abundances (n) to $w_{ptcl,n}$ of the virtual particles, allowing us to calculate the density field in terms of $n \text{ km}^{-2}$.

One possible sink of floating plastics not taken into account here is fragmentation and degradation of plastics. Fragmentation eventually leads to particles being smaller than the detection limit (here: neuston net mesh size). This likely acts on a significantly longer time scale (order of years) than beaching and sinking of particles. In Song et al. (2017), polyethylene pellets, the material which forms the majority of plastics found in the Mediterranean (Suaria et al., 2016; Pedrotti et al., 2016), were subjected to 12 months of UV exposure and 2 months of mechanical abrasion. It was estimated that this might translate to more than 4 years in the natural environment. This weathering resulted in a volume loss of about 10%, and produced about 20 fragments per polyethylene pellet. Photochemical oxidation might also play a direct role in plastic degradation, converting plastic polymers into carbon dioxide and dissolved organic carbon. In Ward et al. (2019) it was reported that this process might play a role on decadal time scales. Both processes are unlikely to have a significant effect on the results presented here, since their removal rates are expected to be at least an order of magnitude smaller than what is necessary for a mass balance: see the supporting material C.6 for a detailed discussion. Nevertheless, taking fragmentation and degradation in account might be a next step for future modelling studies.

Another sink not taken in account is the presence of plastics in biota. To our knowledge the total amount of plastics in biota has not been quantified thoroughly yet. In Booth et al. (2017) the total amount of plastics in fish was estimated to be about six orders of magnitude lower than the amount of plastics in the surface water, hence we neglect this possible sink.

The goal of this work is to have a surface mass balance: particles are removed once they start sinking down, and only surface measurements are used to infer the model parameters. The water column and marine sediments therefore do not need to be taken in account as separate sinks.

4.2.5 Sinks of plastics

The parameterisation of sinks is kept simple in order to avoid the problem from becoming too under-determined (i.e. multiple sets of parameters fitting the data equally well). Time scales define how quickly particles are removed from the surface water due to the different sinks, with the goal of having a first order estimation on their influence.

Beaching

The process of beaching takes place in the mesh cell adjacent to land, which will be referred to as the coastal cell. It is assumed that plastic particles have a constant probability of beaching when inside this coastal cell. The cumulative probability of beaching for a set of plastic particles will follow an exponential distribution as a function of the time that the particles spend in the coastal cell, denoted by t_{coast} :

$$f_{beach} = 1 - P_{beach} = e^{-\frac{t_{coast}}{\tau_{beach}}}, \quad (4.2)$$

where τ_{beach} is the time scale on which beaching occurs. The value for τ_{beach} is one of the parameters which is estimated in the inverse modelling process. The larger the time scale τ_{beach} , the longer the particles will remain in the water. This beaching time scale should be interpreted as a time at which particles remain permanently on the coastline (e.g. due to burial) and are not washed back to sea any more.

For drifter buoys, the beaching time scale is calculated to be about 76 days, see supporting material C.2. However, floating plastic particles do not necessarily behave like drifter buoys close to the shore. Therefore the prior probability density function for τ_{beach} is defined on the \log_{10} of the values to cover a wide range of possibilities (10^1 – 10^3 days), with the beaching time scale for the drifter buoys approximately at the mode of the prior probability density function (10^2 days).

Sinking

For the sinking of particles, a similar approach is used as for beaching of particles, where a time scale τ_{sink} determines how quickly plastic particles are removed from the surface water. A majority of plastics is buoyant: the fraction of initially non-buoyant plastics is defined as $P_{sink,0}$. Due to the formation of a biofilm, initially buoyant particles can start sinking down. Similar to Fazey & Ryan (2016), the probability that particles sink due to biofouling is modelled using a logistic function, where over time the growing biofilm will increase the sinking probability:

$$f_{sink} = (1 - P_{sink,0}) \cdot (1 - P_{sink}), \quad 1 - P_{sink} = \frac{1}{1 + e^{\frac{r_{sink}}{\tau_{sink}}(t_{age} - \tau_{sink})}}, \quad (4.3)$$

where t_{age} is the age of the particle, τ_{sink} is a time scale when 50% of the initially buoyant particles will have sunk, and r_{sink} is the inverse rate at which this happens

(i.e. the slope of the logistic function at the inflection point) in terms of days. As a first order approximation, sinking is assumed to be permanent: the effects of potential oscillations in the water column due to fouling/defouling (Kooi et al., 2017) are assumed to be small.

Data from Fazey & Ryan (2016) are used to estimate parameter bounds for the priors governing the biofouling process. The prior should cover a wide range of values, since differences in the fouling process can be induced by factors such as the particle size used in the experiment, the material, tethered vs. free-floating samples, and differences in fouling communities for different geographical regions. The prior probability density function of τ_{sink} is defined on the \log_{10} of the value to cover a wide range of possibilities. The lower bound is set to the lowest fouling time found in Fazey & Ryan (2016) of 2 weeks. The upper bound is set to a value of 1 year, which is much longer than the experimentally found fouling times, to allow for possible differences in fouling behavior as described above. For r_{sink} bounds on the prior are set to the smallest and largest values calculated using the reported experimental data (3–15 days). The initial fraction of positively buoyant plastic particles is estimated by computing the fraction of polymers produced with a density lower than water (Bond et al., 2018). Since it is not known for all materials whether it will float or sink (e.g. the ‘other materials’ category, or polystyrene, which often appears in its foamed version), this information is used to estimate a lower and upper bound on the initial sinking fraction (0.17–0.44).

4.2.6 Inverse modelling

Parameters governing the sources and sinks are estimated using an inverse modelling approach: parameters are chosen such that the model fit is consistent with the observed plastic concentrations, whilst trying to adhere to the prior parameter bounds specified in the previous sections.

There is relatively little information available on what kind of distribution is the most suitable for the prior information. In most cases there are only point estimates available for possible parameter values as obtained from previous modelling studies or lab experiments. These estimates might differ for our modelling scenario due to different geographical and environmental conditions. We do want to use these estimates as prior knowledge however, since they tell us at least what orders of magnitude we should look at. We choose Gaussian prior distributions here, and assume Gaussian statistics for the model and measurement errors. This allows us to formulate the problem as a least-squares problem, which is computationally much less costly than e.g. using Monte Carlo methods, see supporting material C.3. The cost function of the least-squares problem to be minimized as a function of the model parameters \mathbf{m} is defined as (Tarantola, 2005):

$$S(\mathbf{m}) = \frac{1}{2}(\mathbf{g}(\mathbf{m}) - \mathbf{d}_{obs})^T \mathbf{C}_D^{-1} (\mathbf{g}(\mathbf{m}) - \mathbf{d}_{obs}) + \frac{1}{2}(\mathbf{m} - \mathbf{m}_{prior})^T \mathbf{C}_M^{-1} (\mathbf{m} - \mathbf{m}_{prior}). \quad (4.4)$$

The first term on the right hand side is the mismatch between the modelled plastic concentrations $\mathbf{g}(\mathbf{m})$ and the observations \mathbf{d}_{obs} , weighted by the measurement covariance matrix \mathbf{C}_D . The last term is the deviation of \mathbf{m} from the prior \mathbf{m}_{prior} , weighted by the covariance matrix defining uncertainty of the prior model parameters.

ters \mathbf{C}_M . This term is derived from assuming Gaussian prior distributions. It has the benefit of acting as a regularization term, which can help for solving ill-posed problems (Tikhonov, 1963). Both \mathbf{C}_D and \mathbf{C}_M are diagonal matrices: it is assumed there is no correlation between the measurements. The diagonal entries of \mathbf{C}_D contain the small-scale measurement variance presented before ($\gamma_n=0.1376$, $\gamma_m=0.2201$). Bounds on the model parameters as mentioned in the text are used for the entries in \mathbf{C}_M .

The cost function is minimized by linearising the forward model around an estimate for the parameters \mathbf{m} , and iteratively updating the parameters using a quasi-Newton method, see supporting material C.3.

4.3 Results and discussion

4.3.1 Parameter estimation

Results are presented here for the simulation with a tracer diffusivity of $K = 10 \text{ m}^2 \text{ s}^{-1}$, which was calculated to be the most appropriate value for the grid resolution used here (Neumann et al., 2014; Liubartseva et al., 2018). See supporting material C.4 for further details, along with a discussion on the sensitivity of the results to the value of K , and entries of \mathbf{C}_D and \mathbf{C}_M .

Figure 4.3 shows the probability density function of the prior and the updated (posterior) estimates for each parameter. The most likely value of the posterior for τ_{beach} is 24 days. This is lower than the τ_{beach} estimated for drifter buoys (76 days). The reason may be that floating plastic particles are more severely influenced by e.g. wave action compared to the (drogued) drifters. The most likely estimate for τ_{sink} is approximately 81 days. This is a bit higher than estimates found in Fazey & Ryan (2016) ranging from 17 to 66 days for polyethylene samples. One explanation could be that the Mediterranean is relatively oligotrophic (Stambler, 2014), causing slow growth of the biofilm. For r_{sink} there is not much difference between the prior and posterior. The available data does not seem to contain much information about this parameter (see supporting material C.4 for further discussion). For $P_{sink,0}$ the most likely estimate is 0.36. This corresponds well to the estimated value in Lebreton et al. (2019), where 65.5% of all polymers are expected to be positively buoyant (i.e. $P_{sink,0}=0.345$).

The inverse model suggest that most plastics are likely to originate from coastal population: the most likely value specifies about 1.9 times the total riverine input. This is slightly lower than the value range (3.2–17.6) calculated from Lebreton et al. (2018) for a global scenario. Fisheries are expected to emit less plastics: the most likely value specifies about 0.2 times the total riverine input. This is at the lower end of the global scenario range (0.2–4.9) (Lebreton et al., 2018). In terms of percentages, 61% of marine plastics in the Mediterranean originate from coastal population, 32% from rivers, and 6% from fisheries according to the most likely posterior estimates.

The inverse model finds the low-end estimate of riverine input given in Lebreton et al. (2017) to be the most likely, see supporting material C.4. Scatter plots of the modelled versus measured plastic concentrations can also be found here (S5), where

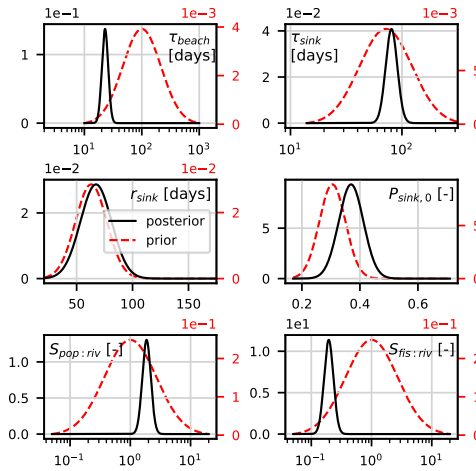


Figure 4.3: Prior (red dashed lines, right y-axes) and posterior (black solid lines, left y-axes) probability density functions for the estimated parameters defining sources and sinks of floating plastic particles. For probability density functions plotted using a logarithmic x-axis, parameters were defined in terms of the \log_{10} of the values.

it can be seen that correlation between the model and measurements is somewhat low. This is difficult to overcome with the highly variable water surface measurements used here. Recommendations to address this in the future are given in the outlook.

4.3.2 Mass balance

The posterior parameter estimates as obtained using the observational data can be plugged into the model. The now calibrated model is used to create a map where plastics are removed from the surface water. The resulting fluxes due to beaching and sinking are shown for the most likely estimates in Figure 4.4. Please note that the beaching fluxes are given in terms of the amount entering the coastal cells of the model, i.e. per unit area. This way no assumptions have to be made about the coastal length inside the cells.

Some beaches which appear to be heavily polluted are located along the North African coast, areas with high estimated amounts of mismanaged plastic waste (Jambeck et al., 2015). Another area is the eastern coast of the Mediterranean. A significant amount of plastics are predicted to be emitted at the coast of Egypt, with predominantly eastward currents following the coastlines. Other major sources of plastics are thought to be the Seyhan and Ceyhan rivers in Turkey, where coastlines in the vicinity are predicted to be heavily polluted as well. Adding to the various sources of plastics, many surface currents end in the eastern basin due to downwelling (Figure 4.1), enhancing the problem at these locations. Patterns of beaching are different on islands depending on which side one looks at: e.g. more beaching is estimated on the western face of Sardinia, and the northern face of Crete which

was also reported in observations (Karkanorachaki et al., 2018).

The highest fluxes of sinking of more than $1 \text{ kg km}^{-2} \text{ day}^{-1}$ occur just next to the coast, where the non-buoyant plastics immediately sink down. Further away from the coast the fluxes are significantly less. In the centre of the Adriatic Sea, relatively high sinking fluxes are predicted of more than $1 \text{ g km}^{-2} \text{ day}^{-1}$. In the western basin there is a large area around the Balearic Islands spanning the Algerian to the Spanish coast with relatively high sinking fluxes, of $0.1\text{--}1 \text{ g km}^{-2} \text{ day}^{-1}$ in the open water. Some qualitative similarities can be observed when comparing with the previous modelling study from Liubartseva et al. (2018), which also found high sinking fluxes around the Balearic, the western coast of the Adriatic, south of the Ionian sea, and the southern coast of Turkey. However, we find higher sinking fluxes in the Gulf of Lion compared to its surroundings, and high sinking fluxes along the Eastern Adriatic coast, and between Tunisia and Sicily.

There is an estimated total plastic input of about 25,600 tonnes over 2006–2016 (2,500 tonnes for the last complete model year 2015). The floating mass stays relatively constant during the simulation, while the sinks keep taking up mass introduced to the basin. Approximately 54% of all plastics eventually ends up on coastlines, 45% starts sinking down. The most likely estimate for the total floating mass in 2015 ranges from 110–190 tonnes. This has a small caveat: the model misses some variance compared to the measurements, and since model output is produced on a \log_{10} scale this results in an underestimation of the total mass, see supporting material C.5 for further discussion. Correcting for this missing variance leads to an estimate of 190–340 tonnes of floating plastics. This is somewhat lower than the estimate from Cózar et al. (2015), where it was estimated to be 756–2969 tonnes.

The numbers presented above are for the most likely posterior estimates. We can also estimate the posterior covariance matrix, see supporting material C.3, allowing us to estimate likely mass balance ranges using Monte Carlo sampling. For 2015 this results in a total plastic input in the range of 2,100–3,400 tonnes; a floating mass of 170–420 tonnes; 1,200–1,900 tonnes of plastics beaching (49–63%); and 900–1,500 tonnes of plastics sinking (37–51%); all reported in terms of the 95% confidence interval (80 samples).

Given the results presented here, it seems likely that at least for the Mediterranean previous estimates of plastics entering the marine environment ($>100,000$ tonnes; Liubartseva et al., 2018; Jambeck et al., 2015) are too high. Observed floating plastic concentrations could in these cases only be explained by having time scales for the sink terms which are much lower than estimated here. The estimated beaching time scale for floating plastics is already lower than the one calculated for drifter buoys. While the sinking time scale could in theory be lower than the estimated 12 weeks, it is very unlikely it will fall much below the minimum 2 weeks reported in experimental studies (Fazey & Ryan, 2016). We do not expect that sinks neglected here such as fragmentation and degradation of plastic could explain a large part of the discrepancy, since time scales of these processes are expected to be relatively high. See the supporting material C.6 for a detailed discussion. Using the approach from Jambeck et al. (2015) and the same conversion rates of mismanaged plastic waste to marine debris (15–40%, 50km radius), we get a plastic input into the

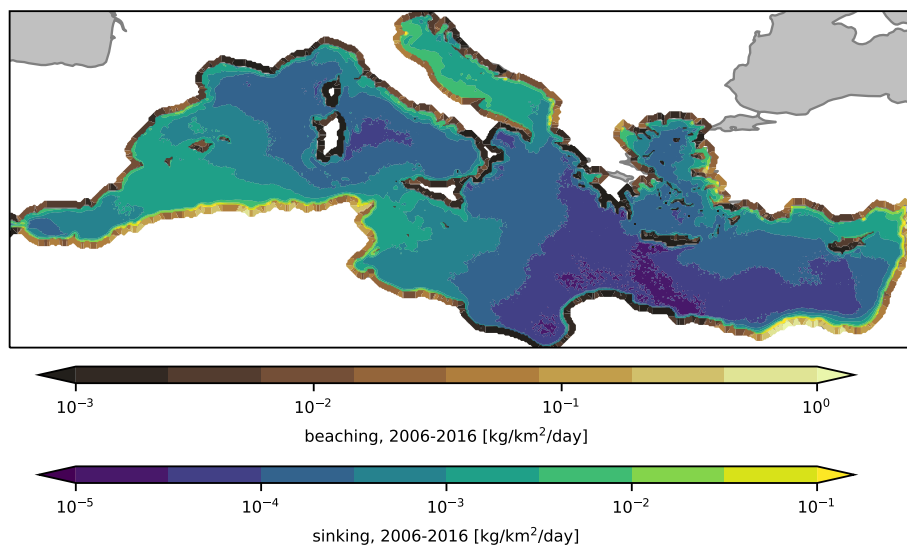


Figure 4.4: Locations in the Mediterranean where beaching and sinking of plastic particles is expected to occur, calculated over 2006-2016. Beaching fluxes are given for the coastal grid boxes ($1/16^\circ$ by $1/16^\circ$), hence no assumptions are made about coastal segment lengths or widths (i.e. coastal lengths contained in the grid boxes will vary, the map does not represent fluxes per stretch of beach in $\text{kg km}^{-1} \text{day}^{-1}$ directly).

Mediterranean water of about 340,000–910,000 tonnes for 2015. Our estimated plastic input from coastal population (1,100–2,300 tonnes for 2015) would correspond to a conversion rate of 0.05–0.10%, which is about two orders of magnitude lower. Neuston net measurements missing the larger plastic pieces could explain some of this discrepancy, which should be quantified in the future. We do not expect that this will explain all of the discrepancy however: in Lebreton et al. (2018) it was found that in the North Pacific accumulation zone, the megaplastics category ($> 50 \text{ cm}$) consists for 92% of fishing nets, ropes, and lines, which are more likely attributed to fishing related activities than land based mismanaged plastic waste.

4.4 Conclusions and recommendations

In this work, inverse modelling was used to calibrate parameters governing sources and sinks of floating plastics in the Mediterranean, by making use of neuston net observations of plastic concentrations. The mass balance of floating plastics resulting from this calibrated model was presented, which gives us insight where we expect most plastics to enter and leave the surface water.

A major step which needs to be taken in future work is ensuring that there is enough reliable data to inform the model, and making sure there is good correlation between the model and measurements. Here, the correlation is somewhat low due to the high measurement variability, which is further discussed in the supporting material C.5. Observed plastic concentrations are highly variable as discussed earlier in this text, see for example de Haan et al. (2019) and the supporting material C.1. Measurement variability is further increased by the fact that different sampling campaigns might have slightly different methodologies. The Mediterranean features highly dynamic currents, making it relatively difficult to model plastic concentrations accurately compared to a domain with a more steady-state structure, like the accumulation zones in the subtropical gyres (van Sebille et al., 2015).

We can look at including more types of measurements, such as observations from beaches, marine sediments, particle size distributions and possibly data of plastic ingestion by animals. Some of these measurements are of a more cumulative nature, such as plastics gathered in sediment traps over time. Perhaps this could alleviate some of the high temporal variability, allowing for more accurate comparison of model output against observational data, helping to constrain the model parameters more accurately. Furthermore, this can result in a better understanding of the sinks neglected here such as fragmentation, degradation, and ingestion. We expect that these processes have a minor influence on the total mass balance (see supporting material C.6 for a detailed discussion). However, how particle sizes evolve over time due to degradation and fragmentation might be important to consider when extending the model to consider size-dependent processes.

Decreasing the mismatch of the model with respect to the measurements will also involve making the model more complex. Only three parameters define the magnitude and ratio of the different plastic sources. In future work the number of sources could be extended, and local uncertainty in the input could be taken into account spatially, e.g. correcting for errors in the estimated mismanaged plastic waste per country. The output from individual rivers could be estimated more accurately on smaller temporal scales, possibly taking into account variations on outflow and precipitation. Extending the amount of parameters defining the sources makes the problem more under-determined however. This means it will be necessary to have more accurate *a priori* knowledge, or more measurement data and/or reduced measurement errors.

For some parameters, spatial and temporal variability is likely important to consider in the future. Biological productivity in the Mediterranean has temporal variability (e.g. seasonal blooms) and spatial variability (e.g. productivity related to upwelling, Macias et al., 2015). This likely influences the sinking time scale of plastics, and could be taken in account in the future by using data from biochemistry models.

Similarly, coastlines along the basin vary in type, which might influence the beaching time scale. Spatial variability could e.g. be taken in account by estimating whether a beach is more ‘rocky’ or ‘sandy’. On the other hand, a parameter like $P_{sink,0}$ might remain relatively constant both spatially and temporally, assuming the types of plastics discarded in different countries are relatively similar.

Some larger plastic objects, like fishing nets, might not be captured by the neuston net measurements used to calibrate the model. The input in terms of mass might therefore in reality be larger than estimated here. It might be a good idea to combine data used here with visual observations of litter as was for example done in Eriksen et al. (2014) to account for the larger plastic items, if the mass of these objects could be estimated.

In future work, effects caused by e.g. washing away of particles from beaches, defouling of particles, different forcings for different particle sizes and shapes can be taken in account by using a more elaborate data-assimilation scheme. This would also allow for better separating the effects of primary and secondary sources of plastics (Browne, 2015). As a final point this work can be extended to other geographical regions where measurements are available.

Closing the global mass budget of buoyant marine plastics: large long-lived debris dominates

One of the longstanding puzzles concerning marine plastic is where plastic ends up after it enters the ocean. Latest estimates of the oceanic input of plastic are one to two orders of magnitude larger than the amount floating at the surface. This discrepancy could be due to overestimation of input estimates, processes removing plastic from the surface ocean (sinking, sedimentation, and beaching), or fragmentation and degradation. Here we present a global buoyant marine plastic mass budget to address this discrepancy. We assimilate unprecedented amounts of observational data from different marine reservoirs (coastlines, the ocean surface, and the deep ocean) considering an extensive particle size range (0.1 mm–1.6 m). We show that larger plastics (>2.5 cm) contribute to most marine plastic mass (3,100 out of 3,200 kilotonnes). Our model estimates an annual ocean plastic input of 500 kilotonnes, which supports the theory that previous estimates (5,600–25,000 kilotonnes) were likely overestimated. Our findings of a reduced oceanic plastic input and increased standing stock lead to much higher residence times of plastics in the marine environment compared to previous model studies, in line with observational findings. With longer-lived marine plastics come drastic consequences without countermeasures and prevention strategies.

This chapter has been submitted as:

Kaandorp, M.L.A., Lobelle, D., Kehl, C., Dijkstra, H.A., van Sebille, E. (2023): The global mass of buoyant marine plastics is dominated by large long-lived debris.

5.1 Introduction

An estimated 250 metric kilotonnes (250 million kilograms) of plastic pollution floats on the surface of the global ocean (Eriksen et al., 2014; van Sebille et al., 2015). A much larger amount of plastic pollution is estimated to enter the ocean every year, in the order of 800–2,400 kilotonnes from rivers (Meijer et al., 2021), and 4,800–23,000 kilotonnes from coastal regions (Jambeck et al., 2015; Borrelle et al., 2020). We assess what causes the misalignment between the estimated plastic input and the total floating plastic mass, by assimilating unprecedented amounts of observational data into a state-of-the-art 3D global transport model for marine plastics, considering timescales in the order of decades (1980–2020). Our dataset includes concentrations in terms of both number (in n m^{-3} in the ocean and n m^{-1} on beaches) and mass (in g m^{-3} in the ocean and g m^{-1} on beaches). In total we use 14,977 measurements from the surface water, 7,114 from beaches, and 120 from the deep ocean (for an overview see the supporting material D.1). From 2,303 beach measurements we additionally use the fraction of fishing related items such as fishing nets. We expand on the mass budget studies in Chapter 4 by increasing the model complexity, incorporating numerous recently developed models for different processes affecting marine plastic transport: sinking via biofouling, beaching, turbulent vertical mixing, and fragmentation. By using a Bayesian framework, our model results match well with both observed plastic concentrations across different marine reservoirs and different size classes (see Figure D.2) and the latest understanding of processes removing plastic from the surface ocean.

Sinking of plastic particles and sedimentation (particles settling on the ocean floor) likely play an important role in removing plastic mass from the surface water (Kvale et al., 2020; Martin et al., 2022). Initially buoyant items can start sinking due to the growth of biofilm on their surface, on timescales of weeks to months (Ye & Andrady, 1991; Fazey & Ryan, 2016; Lobelle et al., 2021). We consider various biofouling scenarios, including fouling-defouling cycles (Kooi et al., 2017; Fischer et al., 2022). Model studies (Lebreton et al., 2019; Onink et al., 2021) have suggested that the majority (67–77%) of plastics reside on beaches or in coastal waters up to 10 km offshore. We therefore include models for beaching and resuspension of plastics back to the ocean (Onink et al., 2019). Surface measurements are currently the only large global observational datasets available in the ocean. Mixing of plastic particles in the water column is hypothesized to be an explanation for the relatively low estimates of plastic mass found in surface net trawls (Poulain et al., 2019; Fischer et al., 2022). We account for this by resolving plastic transport three-dimensionally, including the modelling of vertical turbulent mixing in the water column (Onink et al., 2022a).

Fragmentation plays an important role in explaining the increasing number of plastic particles for smaller particle sizes (Cózar et al., 2014; Song et al., 2017) and can furthermore affect mass budget analyses by breaking down plastic items into particles smaller than typically measured sizes. We therefore include the fragmentation model presented in Chapter 3, including a size spectrum of plastic particles (0.1 mm–1.6 m). This allows us to assimilate different types of observations such as net trawls that mainly capture microplastics (<5 mm, with a typical mesh size of

0.2 mm (Cózar et al., 2015)), as well as measurements of larger plastics (>25 mm) from shipboard observations and beach cleanup campaigns. With this size spectrum we can also more accurately link concentrations in terms of number of plastic particles to concentrations in terms of plastic mass, as a biased conversion between the two has been shown to have a big impact on mass budget estimates (Weiss et al., 2021).

We focus on plastics that are initially buoyant when entering the marine environment, such as polyethylene (PE), polypropylene (PP), and polystyrene (PS). These polymers have been shown to make up the majority of items in the ocean's surface (Bond et al., 2018), deeper layers (Egger et al., 2020b), and beaches (Vianello et al., 2013; Frias et al., 2014; Carson et al., 2011). This means we do not consider polymers denser than seawater such as PVC and PET, which are estimated to make up about 35–40% of the plastic mass entering the marine environment (Andrady, 2011; Lebreton et al., 2018)

5.2 Methods

We use a hybrid Lagrangian-Eulerian model to efficiently advect a virtual plastic tracer through the global ocean, under various environmental forcings. First, we use Lagrangian simulations to advect a globally dense set of virtual plastic particles over a timespan of one month. We run simulations for different particle sizes, different biofouling scenarios, and different months. We then use each of these simulations to construct transition matrices: linear systems that define the probability that plastic particles move from one grid cell of the ocean to another (van Sebille, 2014; Wichmann et al., 2019b). Parameterised sources and sinks for marine plastic pollution are then directly added into the transition matrix model. This allows us to efficiently evaluate different source and sink scenarios, which is necessary during the data assimilation step where we calibrate the set of unknown parameters to optimally match the observational data. For a more detailed explanation, see the supporting material D.3.1. With the hybrid Lagrangian-Eulerian approach we have a parsimonious model which can explain the sparse observed data with as few parameters as possible (16 in total; see Figure D.7).

5.2.1 Data assimilation

Parameters defining sources, transport, and sinks of plastic pollution are given a plausible range (i.e. the Bayesian 'prior') in accordance with current understanding of these processes as discussed in the next sections. Gaussian probability density functions are used to define the ranges, where the 95% confidence intervals define the lower and upper parameter estimates. Measurements contain an error, due to both instrument errors (e.g., differences between campaigns in sampling) and representation errors (e.g., due to unresolved scales and processes (Evensen et al., 2022)). Measurement error is estimated by calculating variograms of the observational data (the same approach as Section 2.2.2 and Section C.1). We use an ensemble smoother with multiple data assimilation (ES-MDA) to update the model parameter values with the observational data (Emerick & Reynolds, 2013). An en-

semble of 55 members (iterated 8 times) is used to estimate the most likely posterior parameter values and confidence intervals. The ensemble members are furthermore used to quantify uncertainty ranges for the estimated plastic concentrations and fluxes. The modelled plastic concentrations represent a mean state, where subgrid-scale variability is not captured. We estimate the subgrid-scale variability from the model-observation mismatch after the data assimilation procedure. The subgrid-scale variability is accounted for in the uncertainty ranges by performing a Monte Carlo analysis (as was also done for the Mediterranean Sea mass budget, see Section C.5), where the plastic concentrations in each ensemble member are perturbed 100 times.

5.2.2 Lagrangian model

To generate the transition matrices, we advect virtual plastic particles in three dimensions in the global ocean using OceanParcels (Delandmeter & van Sebille, 2019), with the Mercator Ocean PSY4 analysis product at $1/12^\circ$ resolution as forcing (Gasparin et al., 2018). This forcing product has been assimilated with various data sources (including altimetry, sea surface temperature, salinity and temperature vertical profile data) and includes freshwater fluxes (Lellouche et al., 2018). Particles are released horizontally on a hexagonal grid with an average hexagon edge length of 22 km, and vertically at 12 logarithmically spaced depth layers between 0.5 m and 5000 m. This release is repeated every month, for five years (2015–2019). Transport is resolved for six different particle sizes (diameter), using an increment of a factor 4 (0.1 mm, 0.4 mm, 1.6 mm, 6.4 mm, 2.6 cm, 10.2 cm). These particles experience a varying amount of influence from vertical turbulent mixing, which can affect their horizontal dispersion (Wichmann et al., 2019a). Analysis showed that of these sizes, the largest particles (10.2 cm) experience negligible effect from vertical mixing in the water column due to their high buoyancy. This is therefore the largest particle size for which we calculate the advection in OceanParcels. We assume similar transport for larger particles (up to 1.6 meter) when constructing the transition matrices (see Section 5.2.3) since these all remain at the ocean surface. Recent studies have shown that simply adding the Stokes drift velocity (Cunningham et al., 2022) or a windage term (Lebreton et al., 2018) to Lagrangian particle simulations representing plastic transport does not increase the match with observational data, which was verified in a preliminary analysis. These effects are therefore not included in our model. Generally, Lagrangian particle simulations include a stochastic (diffusive) term (van Sebille et al., 2018) to account for missing subgrid-scale effects (e.g. submesoscale eddies). This stochastic term is not included as the transition matrices calculated from the Lagrangian transport already introduce diffusion in the dynamics (Wichmann et al., 2019b). Transport is resolved for four different vertical transport scenarios under influence of turbulence and biofouling as described in the next section. The total number of particle trajectories across all simulations for the different months, sizes, and transport scenarios, is 1.7 billion.

Vertical motions: mixing and biofouling

We consider four different scenarios for the vertical behaviour of plastic particles in the ocean. In all four scenarios, vertical diffusion due to turbulence is included,

using a Markov-0 random walk model (Onink et al., 2022a) forced by the PSY4 vertical diffusivity fields.

First, we consider plastic particles that remain positively buoyant, with a rise velocity that is dependent on the particle size (Poulain et al., 2019). Spherical particles (Dietrich, 1982) with a density of 1010 kg m^{-3} are used as a baseline, giving the best match with observational data (see Figure D.13). In reality, environmental plastics have a range of densities and shapes (Kooi & Koelmans, 2019). For a given particle size in the model, we take a linear combination of the six differently sized baseline particles, to model an assemblage of particles with different rise velocities. This linear combination is calibrated during the data assimilation procedure to give an optimal match with the observational data, see the supporting material D.3.4 for further details. Using spherical particles as a baseline keeps the rise velocity model consistent with our biofouling implementation (Fischer et al., 2022), and keeps the calculation procedure computationally cheap (as opposed to some recent iterative procedures for calculating rise velocities of non-spherical particles (Poulain et al., 2019; Waldschläger & Schüttrumpf, 2019), which would add extra computational costs).

Second, we include two scenarios for biofouling, using a recently developed Lagrangian model (Fischer et al., 2022). Biochemistry fields from the Mercator Ocean BIOMER4 analysis product are used at $1/4^\circ$ resolution. Biofilm on plastic particles is gained via collisions and growth with algae in this model, and is lost via respiration. Biofilm loss via grazing and viral lysis is neglected to keep the amount of free model parameters limited, since this effect is suggested to be minor (Fischer et al., 2022). Due to the growth and loss of biofilm, plastic particles can oscillate vertically in the water column. This oscillatory behavior is only theoretical and has not been experimentally observed. We therefore also include a scenario where the fouling of particles is permanent, neglecting the respiration loss term. In our fourth scenario, particles become neutrally buoyant, e.g. due to a balance in the fouling and defouling processes, or slightly negatively buoyant particles reaching an equal density isopycnal surface.

5.2.3 Transition matrix model

The statistics of the Lagrangian particle transport (probabilities that particles move from one grid cell of the ocean to another) are stored in transition matrices, with time-windows of 30 days. The Uber H3 grid is used to construct the transition matrix bins horizontally, where each cell has an edge length of approximately 60 km. Each horizontal bin is furthermore divided vertically into four depth bins, with boundaries at 0, 5, 50, 500 meters deep and at the ocean floor. Additional cells are introduced into the transition matrix system representing the coastline segments inside the coastal cells (see Figure D.10), to model transport between the ocean and beaches. The resulting transition matrices have a size of $121,000 \times 121,000$.

In the next two sections we discuss how sources, transport, and sinks of marine plastics are parameterised in the transition matrix model. We touch upon plausible ranges for each parameter value, as these are used to define the prior probability density functions in the Bayesian analysis, when assimilating observational data into the model (see the supporting material D.2.3).

Parameterisation of sources: rivers, coastlines, and fisheries

We consider three major types of marine plastic pollution sources in our model: rivers, coastlines, and fishing activity (Lebreton et al., 2018), which are shown on a map in Figure 5.1.

Current estimates of riverine plastic inputs vary widely. Global estimates based on modelling studies calibrated to observational data range from 1,150–2,410 kilotonnes per year (Lebreton et al., 2017) to 800–2,700 kilotonnes per year (Meijer et al., 2021). However, it was recently argued that these values might be overestimates (Weiss et al., 2021), giving a much lower estimate of about 6.1 kilotonnes per year. Reasons for potential overestimations are biases in conversion from number concentrations to mass concentrations due to too high particle mass estimates, and mixing different sampling techniques without accounting for varying lower size detection limits. We use the most recent global estimate of riverine inputs (Meijer et al., 2021), given that they find a reasonable correlation to observational riverine data ($r^2 = 0.74$), and their data are publicly available. We take possible biases into account by scaling the total input with a factor $S_{riv.}$, where the bounds of this prior are chosen to capture both the high-end (2,700 kilotonnes per year (Meijer et al., 2021)) and low-end (6.1 kilotonnes per year, Weiss et al., 2021) estimates.

For coastal mismanaged plastic waste (MPW) we make use of a global MPW dataset per country (Jambeck et al., 2015) in terms of kilotonnes per year per capita. Combined with the estimated population density within 50 kilometers from the coast (SEDAC et al., 2005), this gives us the coastal MPW per unit area per year. A parameter $S_{pop.}$ defines the relative input of coastal MPW with respect to the total riverine input.

We estimate plastic loss per fishing hour by scaling a globally estimated fishing hours dataset (Kroodsmas et al., 2018) with a parameter $S_{fis.}$ that defines the relative input of fishing-related plastic with respect to the total riverine input. The prior bounds for $S_{pop.}$ and $S_{fis.}$ are defined using previously estimated input ranges for different waste categories (Lebreton et al., 2018). This gives 2.7–7.3 times the riverine input for coastal MPW, and 0.2–2.0 times the riverine input for fishing related plastic.

Larger items make up the majority of plastic mass found in the marine environment (Lebreton et al., 2018; Ryan et al., 2020), while small fragments dominate in terms of the number of particles (Cózar et al., 2015). It is not yet well known which particle size dominates new plastic items introduced into the marine environment. We parameterise the plastic input size using a log-normal distribution, capturing the dominant particle sizes of plastic packaging in municipal solid waste sorting facilities (about 0.2 meters, Jansen et al. (2015)) and the dominant sizes of plastic items found in rivers (about 0.2–0.3 meters, van Emmerik et al. (2019) and Vriend et al. (2020)), see the supporting material D.3.2 for more details.

Plastic waste generation has been increasing exponentially the last decades (Geyer et al., 2017). We use an exponential function to parameterise the possibility that this has led to an increasing amount of waste entering the ocean. The midpoint estimate for the exponential growth rate (GR_{in}) prior is calibrated to plastic waste production estimates (Geyer et al., 2017), and the lower bound is set to zero to allow for the possibility of no increased input into the ocean (e.g. due to more efficient collection and processing of waste).

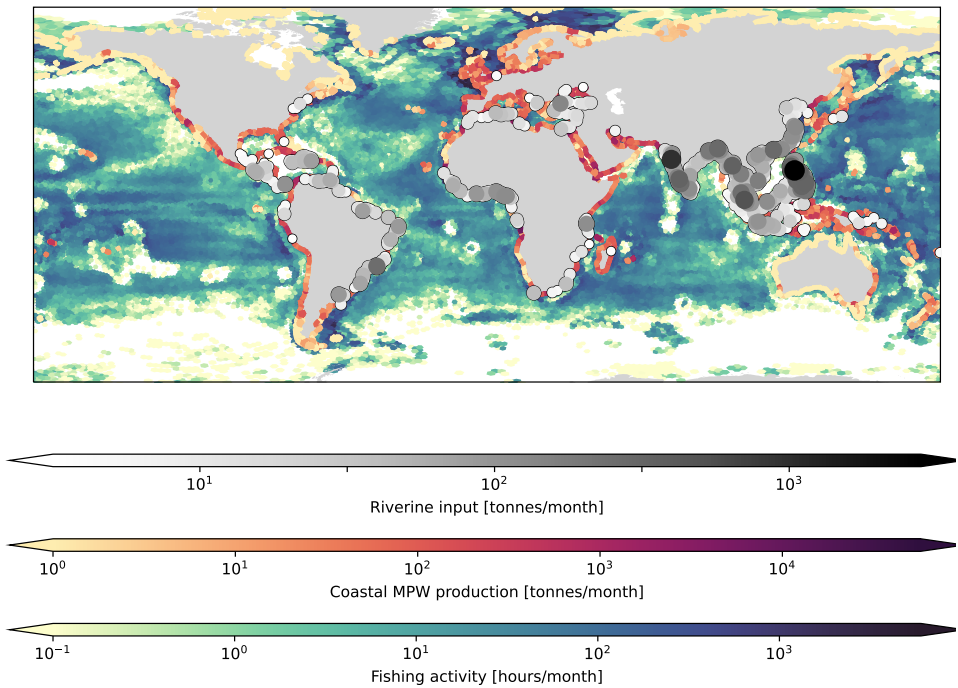


Figure 5.1: The three major sources of marine plastic pollution: rivers (Meijer et al., 2021), coastal mismanaged plastic waste (Jambeck et al., 2015), and fishing activity (Kroodsma et al., 2018).

Parameterisation of transport and sinks: biofouling, beaching, and fragmentation

The four different vertical transport scenarios (Section 5.2.2) yield four different transition matrices. We assume particles in the ocean are an assembly of these four scenarios. How much each scenario contributes is parameterised using three fractions: f_{of} (oscillatory fouling/defouling), f_{pf} (permanent fouling), f_{nb} (neutrally buoyant), with the remaining fraction being the positively buoyant particles. In the case of permanent fouling, we keep track of particles hitting the ocean floor (which is one of our sinks), in which case they are classified as ‘sedimented’ and removed from the system. Prior bounds for f_{pf} are set to 1.7%–97% by comparing previous estimates of plastic seafloor export (Kvale et al., 2020) to the estimated plastic mass at the ocean surface (Eriksen et al., 2014). The remaining fractions are given equal prior probabilities, with the maximum fraction values set to 95%.

The transition matrix contains separate cells for the coastline segments. The probability that plastic particles beach, i.e. move from a coastal ocean cell onto the dry land, is parameterised using a beaching time scale τ_{beach} (Onink et al., 2021; Kaandorp et al., 2020). The prior probability density function for τ_{beach} is defined on the \log_{10} of the value to cover a wide range of (positively valued) possibilities. Parameter bounds are based on the findings from Chapter 4, with a mid-point estimate of 100 days, and a lower bound set to 25 days. Resuspension time scales determining how quickly differently sized plastic move from the beach to the ocean are based on experimental findings (Hinata et al., 2017). A probability $p_{removal}$ is defined for plastics being removed from beaches (e.g. due to burial (Ryan et al., 2020), cleanup efforts, or direct degradation of plastic material such as oxidation (Ward et al., 2019)). We use a removal rate of 0.2% based on the results from Chapter 3, and allow it to vary an order of magnitude, capturing the removal rates from other global mass budget studies (0.8%–4% per month, Isobe & Iwasaki, 2022). This is the second ‘sink’ in which particles are permanently removed from our simulations. The coastline length inside each grid cell, necessary to calculate litter concentrations per unit length of beach, is computed using the natural earth dataset (Kelso & Patterson, 2010). Coastlines have a fractal structure, which can lead to different along-shore lengths in beach surveys compared to the discrete map data. We use a correction factor (Isobe & Iwasaki, 2022) to account for the 1.27 fractal dimension of coastlines (Husain et al., 2021). The typical beach survey resolution is set to 100 meters, and the coastline segment resolutions are calculated directly from the natural earth map data. To account for the fact that less litter might beach in grid cells with only a small amount of coastline, a parameter $l_{beach,min}$ is introduced. Below this value, the beaching probability decreases linearly down to zero.

We assume fragmentation of plastic items is dominant on beaches due to higher temperatures, oxidation, UV-radiation, and mechanical abrasion (Andrady, 2011; Kalogerakis et al., 2017; Song et al., 2017; Efimova et al., 2018). Previous studies (Onink et al., 2022b) show that neglecting ocean fragmentation is justified, as long as plastics fragment at the same rate or slower in the ocean compared to on beaches. The fragmentation model from Chapter 3 is used here to simulate how plastic items break down into smaller particles over time. Parameters to be estimated are the fragmentation rate λ_f , and the shape factor d_N which is used to represent the di-

dimensionality of plastic items (2 for flat objects, 3 for cubes, non-integer values for mixtures of differently shaped objects). For plastic items the fragmentation rate is still not well known. Parameter bounds for λ_f are based on experimental data (up to $1.9 \cdot 10^{-4} \text{ d}^{-1}$, Song et al., 2017) and results from Chapter 3 (down to $2.9 \cdot 10^{-5} \text{ d}^{-1}$), and are defined on the \log_{10} of the value to cover a wide range of possibilities. Bounds for d_N are based on observational data of plastic particle sizes and masses (see the supporting material D.3.3). Fragmentation is the third ‘sink’ for plastic particles in our simulations, where they are removed when reaching a size smaller than 0.1 mm.

In our model we consider a full size spectrum from 0.1 mm to 1.6 m, using increments of a factor 2. Lagrangian particle transport is resolved for six sizes (0.1 mm to 10.2 cm, see Section 5.2.2). For intermediate sizes in the spectrum, the available transition matrices are interpolated linearly. For larger plastics (>0.1 m) similar transport is assumed, as these particles remain predominantly at the ocean surface. We model both the particle size distribution in terms of the number of particles and in terms of mass, see Chapter 3 for more details. This way we can quantify which particle size contributes to most of the marine plastic pollution. For smaller plastic particles data are available on typical masses (Lenz, 2020; Egger et al., 2020a). For bigger items the particle mass m_p is extrapolated from the particle length l_p using $m_p \propto l_p^{d_N}$, consistent with the fragmentation model. See the supporting material D.3.3 for more details.

5.3 Results and discussion

5.3.1 A 3D map of marine plastic litter

We estimate a total amount of initially buoyant plastics in the 3D global ocean of 3,200 kilotonnes (95% confidence interval: 3,000–3,400 kilotonnes) for the year 2020 based on our assimilated model. A three-dimensional global map of the estimated marine plastic pollution for the complete modelled particle size spectrum (0.1 mm to 1.6 m) is shown in Figure 5.2. The largest fraction of plastic mass is located at the ocean surface: 59%–62%. More than a third of the mass resides deeper in the ocean (36–39%), and the remainder is located on beaches (1.5%–1.9%).

These results, as well as the the estimated fluxes into and out of the marine environment are summarized in the schematic overview in Figure 5.3. We calculate a total marine plastic input of 500 kilotonnes per year (95% confidence interval: 470–540 kilotonnes, for the reference year 2020), originating from coastlines (39–42%), from fishing activity (45–48%), and from rivers (12–13%). The total input we predict increases by about 4% per year, which is consistent with the estimated increase in global plastic waste generation of about 5% per year (Geyer et al., 2017), and with the observed increase of plastic concentrations in the Pacific Ocean (Lebreton et al., 2018) and the North Atlantic Ocean (Wilcox et al., 2020). Our calculated global riverine input of 57–69 kilotonnes per year is lower than previous estimates (800–2,700 kilotonnes per year (Meijer et al., 2021).) The 190–220 kilotonnes of input from coastlines is furthermore at least an order of magnitude smaller than previous estimates (4,800–12,700 kilotonnes per year, Jambeck et al. (2015)). These much

lower input estimates for rivers and coastlines are consistent with recent modelling and observational studies (Ryan et al., 2020; Schöneich-Argent et al., 2020; Weiss et al., 2021). The estimated input of 220–260 kilotonnes from fishing activity is somewhat lower than previous estimates of 640 kilotonnes per year (Li et al., 2016). Our modelled concentrations of fishing related plastics are consistent with the observed amount of items on beaches, see the supporting material D.2.1, and match qualitatively with review studies showing that the majority of plastic litter in the open ocean originates from the ocean (Morales-Caselles et al., 2021) (e.g., items such as nets, ropes, and buoys).

Biological processes (such as biofouling) play an important role in the dynamics and export of plastic waste from the ocean surface. We estimate that 220 kilotonnes of plastics are exported to marine sediments per year, of which 6 kilotonnes are microplastics (<5 mm), which is close to the 7–420 kilotonnes per year from previous modelling studies (Kvale et al., 2020). We estimate that 6,200 kilotonnes of initially buoyant plastics have ended up in marine sediments since 1950, which is less than a recent estimate of 25,000–900,000 kilotonnes (Martin et al., 2022) for all plastic (buoyant and non-buoyant). Our model shows that about half of the plastic particles in the marine environment experience so much biofouling that they start sinking or become neutrally buoyant. The high amount of plastic particles in the deep ocean (Egger et al., 2020b), for a large part consisting of low density polymers such as PE and PP, cannot be explained without this fouling.

We estimate a plastic sink of 3 kilotonnes per year at coastlines, due to processes such as burial, cleanup efforts, and direct UV degradation. Additionally, a substantial amount of plastics are fragmented into particles smaller than 0.1 mm; about 73 kilotonnes per year. We estimate that 2.2% of plastics larger than 5 mm fragment into particles smaller than 5 mm per year, which is very close to previous model estimates of about 3% per year (Lebreton et al., 2019).

5.3.2 Particle size contributions to the marine plastic mass budget

One of our key results is that the majority of plastic mass is contained in the large plastic items (>2.5 cm): 90%–98% (2,800–3,300 kilotonnes). Microplastics (<5 mm) and plastics between 5 mm–2.5 cm form the small remainder at 49–53 kilotonnes and 150–170 kilotonnes respectively, which is in the same order of magnitude as previous estimates for small (0.3 mm – 20 cm) floating plastics (93–236 kilotonnes, van Sebille et al., 2015). These findings are also consistent with an analysis of the Great Pacific Garbage Patch (Lebreton et al., 2018), where it was found that microplastics (<5 mm) only make up 8% of the total plastic mass at the ocean surface.

Particle size distributions (Figure 5.4a) reveal an increase in the total plastic mass with particle size according to a power law, up to an estimated dominant particle size of about 0.4 meters. Few particles above this length scale are expected to enter the marine environment. The model results indicate that most of the plastic mass for particles smaller than 0.8 mm is below the ocean surface. The total plastic mass on coastlines is about an order of magnitude less compared to the surface and

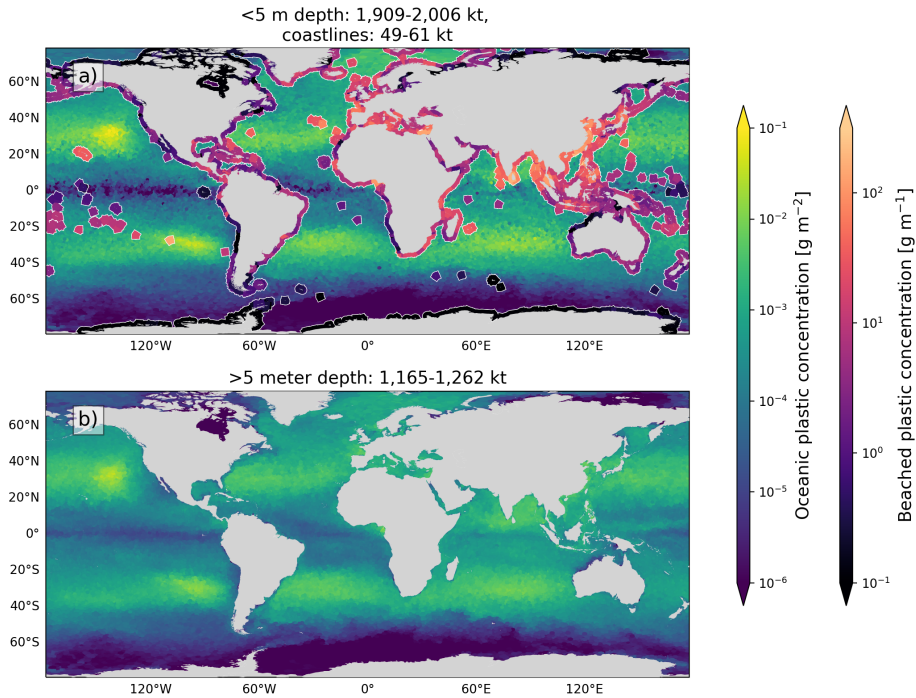


Figure 5.2: A map of the predicted plastic concentrations in the marine environment. The predicted concentrations of plastic items (0.1 mm – 1.6m) are shown for the most likely parameter estimates in the ocean surface (0–5 m depth, panel a) and below the ocean surface (panel b) in terms of g m^{-2} . Predicted plastic concentrations on beaches (in purple to red in panel a, white delineation) are shown in terms of g m^{-1} . The estimated concentrations are shown for the year 2020.

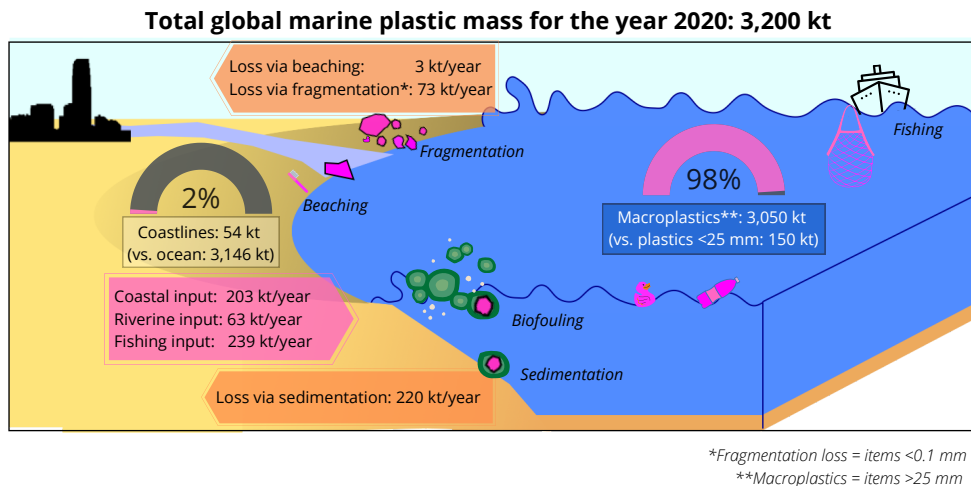


Figure 5.3: Schematic overview of the estimated marine plastic fluxes and standing stocks.

deep ocean for all particle sizes. The number of particles increases for decreasing particle size (Figure 5.4b) according to a power law, as has been shown in previous observational studies (Cózar et al., 2014; Kooi & Koelmans, 2019).

Our estimate of the total plastic mass for particles smaller than 5 mm (53 kilotonnes) is similar to previous studies (35.5 kilotonnes, Eriksen et al., 2014). However, our mass estimate for particles between 5 mm and 200 mm, 700 kilotonnes, is much higher than previously estimated (30.6 kilotonnes, Eriksen et al., 2014). This is even more so for particles larger than 200 mm, where our estimate of 2,500 kilotonnes vastly exceeds the previously estimated 202.8 kilotonnes (Eriksen et al., 2014). This difference can largely be explained by the very broad size intervals used previously (Eriksen et al., 2014). To elaborate, due to the scaling of particle mass with size ($m \propto l^2$ for flat objects, $m \propto l^3$ for cubic objects, see the methodology for further details), using the same mean particle mass for a large size interval will vastly underestimate the total mass. Biases in mass budget estimates due to incorrect usage of mean particle masses were also observed for riverine plastic studies (Weiss et al., 2021). As an example it can be seen in Figure 5.4 (left panel) that the larger particles around 200 mm by far dominate the total mass in the size range 4.76–200 mm, whereas these particles are very sparse in terms of total number (right panel). These large differences in total mass estimates underline the importance of treating different particle sizes very carefully in mass budget studies: preferably a (semi-) continuous size spectrum should be used.

We recommend that future plastic measurement campaigns and mass budget studies treat number and mass measurements more carefully. The amount of plastic particles increases exponentially with decreasing particle sizes. Reporting the number of particles in observational studies can be unreliable when no strict lower limit of the

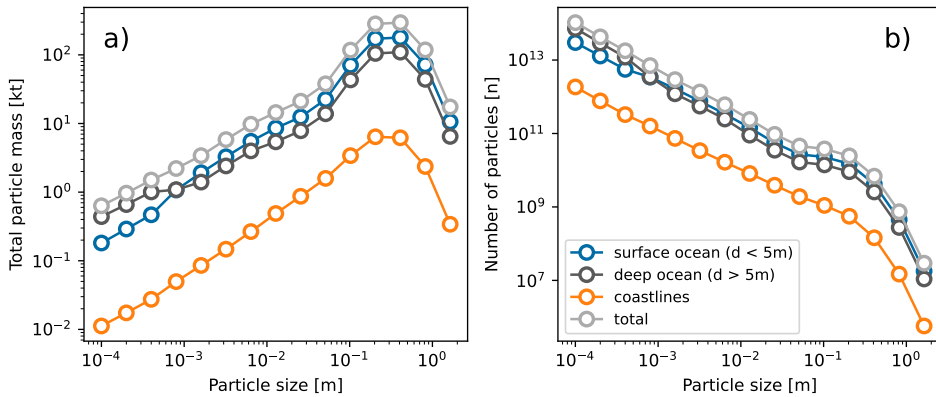


Figure 5.4: Particle size distributions of marine plastics. The total mass in kilotonnes (a) and number of particles (b) are shown for different size classes (from 0.1 mm to 1.6 m in terms of diameter). Note the logarithmic axes.

particle size is used. This is especially true for visual observations (the main source of data for beaches, OSPAR, 2010; Burgess et al., 2021), as the lower detection limit likely varies per person. Measuring the total mass of items is more reliable in those cases, as most of the plastic mass in the marine environment is contained in the larger particle sizes which are more easily observed.

5.4 Conclusions and outlook

In summary, we find that the total amount of marine plastic litter, 3,000–3,400 kilotonnes, is much higher than previous estimates (Eriksen et al., 2014), which for a large part can be explained by better representing large plastic object masses. We also find a plastic input into the marine environment of 470–540 kilotonnes per year, at least an order of magnitude less than previous estimates (Jambeck et al., 2015; Meijer et al., 2021; Borrelle et al., 2020). The decreased input and increased standing stock suggest that there is no ‘missing sink’ for marine plastic pollution, which has been the focus of many recent papers (Weiss et al., 2021; Isobe & Iwasaki, 2022; Kvale et al., 2020; Ryan et al., 2020). Our mass budget estimate is consistent with observed plastic concentrations in different marine reservoirs, and with our latest understanding of processes removing plastic particles from the surface ocean, such as biofouling and sedimentation, beaching, fragmentation, and mixing.

Our finding of a lower plastic input into the marine environment and a higher standing stock means that the residence time of plastics in the marine environment is much higher than previously estimated. For example, some studies (Koelmans et al., 2017) predicted that given an instantaneous stop of plastic emissions, more than 95% of the plastic mass would be removed from the ocean surface within 1–2 years due to fragmentation and sinking. We show a similar analysis for a sudden stop of

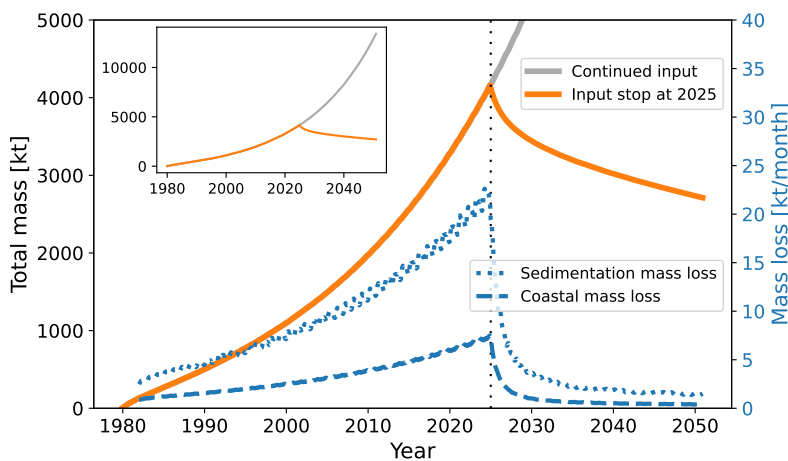


Figure 5.5: Scenarios for the future marine plastic mass budget. A scenario for a sudden stop of new plastics introduced into the marine environment in 2025 is shown in orange, a business-as-usual scenario with an exponentially increasing input in grey. The inset presents a zoom out of both scenarios. Also shown are the expected marine plastic mass losses along coastlines (dashed blue line) and due to sedimentation (dotted blue line), showing that the mass loss after a sudden input stop is expected to decrease rapidly.

new plastics introduced into the marine environment in 2025 in Figure 5.5 using our data-assimilated model. We expect that in this scenario only 10% of the plastic mass would be removed from the marine environment within 2 years (orange line). The removal rate is expected to decrease rapidly over time, as plastics move away from coastal regions towards areas such as the subtropical gyres. As there is no beaching and little sinking of marine plastics in these low algal areas compared to coastal waters (Lobelle et al., 2021; Fischer et al., 2022), the plastic particles become highly persistent (see the supporting material D.2.2). We estimate that the plastic input into the marine environment is likely still growing around 4% per year. Previous studies (Galgani et al., 2021) found no conclusive temporal trends regarding the amount of plastic litter in the marine environment. Establishing temporal trends is difficult due to the high variance in measured plastic concentrations (de Haan et al., 2019). Our estimated growth rate of 4% per year gives the best match with observational data over the past decades (see the supporting material D.2.4), but might change in the future under the influence of mitigation strategies and cleanup efforts. Without further mitigation strategies and cleanup efforts, our estimated growth rate of 4% per year has the potential to double the plastic standing stock within decades, as shown in the inset in Figure 5.5. The combination of a projected exponentially increasing input and long persistence of marine plastics means a likely increasing negative impact of marine plastic pollution on ecosystems in the future.

Summary and Outlook

In this chapter, I first look back on the research questions from Chapter 1. Afterwards, I provide a more general discussion on the field of plastic research and the use of data-driven modelling, and what steps could be taken in future research.

6.1 Answers to the research questions

| |
|---|
| How can we use data-driven modelling to quantify the marine plastic mass budget, where we bring together numerical models with observational data? |
|---|

In this thesis, we used different data-driven modelling techniques such as machine learning and data assimilation, where machine learning can be used to create regression models purely based on observational data, and where data assimilation can be used to optimally combine numerical models with observational data. We showed a schematic representation of the differences between these data-driven modelling techniques in Figure 1.5. Purely data-driven modelling tools such as machine learning can be useful for marine litter studies, when no numerical model is available that is able to capture the underlying physical processes of the problem at hand. Data assimilation methods are suitable when the underlying physics can be modelled, but where parts of the model are uncertain due to e.g. unknown parameter values. With an underlying physical model the results are generally easier to interpret.

In Chapter 2 we explored how the amount of litter beaching on the Dutch North Sea coastline can be predicted. It is not feasible to use traditional physics-based models in this case, due to the complex physical processes and their unknown inter-

actions near the coastline. We therefore used machine learning, where we trained a statistical algorithm on a large set of explanatory variables and the corresponding litter concentrations. We showed that the machine learning model is able to make reasonable predictions of the plastic concentrations. We analysed the underlying regression model to assess which variables contributed to the model performance, indicating that tides, coastline orientation and length, and oceanic transport are important factors to take into account. In Chapter 4 and Chapter 5 we explored the assimilation of marine plastic transport models with observational data. The benefit of these methods is that we are able to get a reasonable match to the observational data with relatively little parameters, which all have a clear interpretation as well.

The principle of parsimony, also referred to as Occam's razor, means that the simplest model that explains the data is preferred (Hoffmann et al., 1997). A Bayesian perspective on this principle is given in Ghahramani (2013): when one has a set of models m and observations d , the posterior probability of a model m is given by:

$$P(m|d) \propto P(d|m)P(m). \quad (6.1)$$

Highly complex models cover a much wider range of possible datasets compared to simple models. Their likelihood $P(d|m)$ is therefore much more 'spread out'. Since $P(m|d)$ is a probability density function and hence normalized, the posterior probability of a complex model that is able to capture observations d is lower than a simple model that is able to do the same. Non-parsimonious models are likely to overfit data used for training, which means that they will not generalize well to new data.

Based on this argument, constructing a model with a set of interpretable parameters and assimilating it with observational data is a good first choice for data-driven modelling of the marine plastic mass budget. While machine learning was a good option for Chapter 2, the model has a lot of parameters (the amount of parameters in a random forest algorithm generally scales with the amount of data). It is unlikely that the specific model that was trained in Chapter 2 generalizes well to other regions of the world. This does not mean that the underlying methodology is not useful: the machine learning algorithm is easily retrained, which could be done in future studies using additional data from other regions. These kind of methods can be useful in an operational setting, where one is less interested in the underlying physics, and more interested in making predictions as accurate as possible. The machine learning method in Chapter 2 was useful to identify which environmental variables likely play an important role in predicting beached litter quantities, as this was not known a priori. This information could be used in future Lagrangian particle tracking studies to improve the parameterisation of the beaching process.

The benefit of the data assimilation approach used in Chapter 5 lies in the fact that different data from all over the marine environment can be used to inform parameters. As an example, we calibrated fragmentation parameters in Chapter 3 based on observational data of particle size distributions alone. In the data assimilation framework from Chapter 5, there are additional data that can inform the fragmentation process, such as the amount of large plastic items relative to the amount of

small plastic items found in the surface water or on beaches.

What processes cause the large mismatch between the expected marine plastic input and the estimated marine plastic standing stock? Is there a missing sink of marine plastics?

We estimate that the mismatch between the expected marine plastic input and the floating plastic standing stock is caused by previous studies missing the contribution of large plastic items, and by previous input estimates being too high.

We could already see in Chapter 3 that large plastic items play an important role in the mass budget, when studying particle size distributions. We estimated that the normalized mass size distribution of marine plastics in the surface water is approximately flat, see Figure 3.7b. A flat normalized particle size distribution indicates that the amount of mass in a size class scales with the amount of length units within this size class: for example, particles between 1–10 mm would contribute to 9 times as much mass compared to particles between 0–1 mm.

Looking at Figure 5.4 from our global mass budget estimate, it comes to no surprise that previous studies overlooked the contribution of larger items. When measuring plastics at the ocean surface, millimeter-sized objects are calculated to be 1,000–10,000 times more abundant than objects between 0.1–1 meter. However, we estimate that these large items contribute orders of magnitude more to the total marine plastic mass. In many regions of the global ocean, fishing related items contribute to a big fraction of these large plastic items, see Figure D.3.

In Chapter 4 we estimated the floating plastic mass budget in the Mediterranean Sea, by calibrating a plastic transport model to neuston net measurements. Chronologically, this was the first work to be completed in this thesis. In the discussion in Section 4.4 we argued that our Mediterranean plastic mass estimates could be underestimated, due to neuston net measurements not capturing the larger items. Indeed, we saw later in Chapter 5 an order of magnitude difference between previous global mass budget estimates calibrated to neuston net measurements (van Sebille et al., 2015) and our own mass budget estimate calibrated to data including large plastic items. It is therefore likely that the input mass estimates for the Mediterranean Sea given in Chapter 4 are underestimates: they are probably more representative of small plastic items (roughly <10 mm in size) typically captured in neuston net trawls (Morét-Ferguson et al., 2010).

Previous estimates of plastics entering the ocean have likely been too high due to a lack of data, or due to biases in model calibration. For land-based plastic input usually the estimate by Jambeck et al. (2015) is used (4.8–12.7 million tonnes). In one step of this study, the fraction of mismanaged plastic waste entering the ocean was estimated, based on calibration data from the San Francisco Bay in the United States. Due to a lack of calibration data elsewhere in the world, this makes the global extrapolation quite uncertain. As addressed in Chapter 5, previous riverine input estimates (Lebreton et al., 2017; Meijer et al., 2021) could be overestimated due to biases in the model calibration data: too high particle mass estimates when

converting from number concentrations to mass concentrations, and mixing different sampling techniques without accounting for varying lower size detection limits (Weiss et al., 2021). We will come back to this in the next section.

Plastic mass is removed from the marine environment by sinks, such as biofouling, beaching, and degradation and fragmentation. Coastal waters and beaches likely play an important role in removing plastics from the surface water, as the biological activity is relatively high there, and as the majority of degradation and fragmentation of plastics is expected to occur there (Song et al., 2017; Onink et al., 2022a). In Chapter 5, we estimated that biofouling is the predominant sink globally, removing about 220 kilotonnes of plastic from the surface water per year. We estimated that about 73 kilotonnes of plastic fragment into particles smaller than 0.1 mm per year, and estimated about 3 kilotonnes of plastic to be lost along coastlines due to effects such as burial, degradation, and cleanup efforts. We calculated that beaches contain a relatively small part of the marine plastic standing stock globally (2%). In Chapter 2 we estimated that the Dutch North Sea coastlines contain about 17–31 tonnes of litter, of which about 80–90% are plastic items. To put this into context, based on the analysis in Chapter 5, we calculate roughly 350 tonnes of plastics to enter the North Sea per year from Dutch rivers and coastlines. This number is a flux and not a standing stock, but still indicates that it is likely a relatively small part of the Dutch marine litter that remains on beaches.

6.2 Discussion and outlook

Several challenges remain in the field of marine plastic research, some of which were already addressed in the previous chapters. I will discuss challenges related to the modelling and observational data of marine plastic pollution, and will touch upon steps that can be taken in future research to tackle these challenges.

6.2.1 Combining numerical models with observational data

A better understanding of (subgrid-scale) variability is essential. There is a high amount of variability between observed plastic concentrations on temporal and spatial scales that are not captured by the numerical model. We explored variability in Chapter 2, where we constructed a variogram using the observational data of litter on Dutch beaches. Variability is important to quantify: numerical models will always have a limited temporal and spatial resolution, which means that a certain amount of subgrid-scale variability will be missed by the model. We used the calculated subgrid-scale variability to estimate the amount of uncertainty that the model can be expected to have. We did a similar variability analysis in Chapter 4, where we constructed variograms of number and mass concentrations of plastic particles in the surface water. Ideally, we would correct for different factors that introduce variability between measurements, to make the observational data more informative for the numerical model.

Different physical processes can influence small-scale accumulation patterns of plastics in the ocean. Since this introduces increased variability between measurements in observational datasets, it would be beneficial to log the relevant physical vari-

ables. One process that could have an influence on small-scale accumulation patterns are windrows, where bubbles and floating debris accumulate in lines aligned with the wind, see e.g. Cózar et al. (2021). In this case it is likely that the variability between measurements (e.g. net trawls) is influenced by the wind speed and direction, and the direction of the measurement path. These are therefore factors that would be valuable to note down in observational datasets. The wind speed, wave height, and mixed layer depth furthermore influence how deep plastic particles mix in the water column (Kukulka et al., 2012; Onink et al., 2022a) and would therefore be valuable to log during measurement campaigns if possible. Studies have suggested that plastic concentrations might differ between anticyclonic and cyclonic mesoscale eddies (Brach et al., 2018). When using coarser ocean circulation model data, or a model that has not been assimilated with sea surface height, it might be good to correct for this effect by using for example satellite altimetry data or derived products.

On a similar note, further standardization of observational data would allow for better intercomparison of data from different laboratories (Maximenko et al., 2019). We already touched upon the issue of standardization of beach measurements in the introduction (Section 1.2.2): to make observational efforts of beach plastic more comparable (see e.g. OSPAR, 2010 and Burgess et al., 2021), efforts are necessary to standardize for example the considered particle sizes, the length/width of beach that is sampled, and whether this is done per unit length or per unit area.

With standardized observational data, and by correcting for subgrid-scale effects that introduce variability between measurements, observational data can be made more representative of the modelled plastic concentrations. An illustration of a possible framework is shown in Figure 6.1. Numerical models produce predictions d at a given spatial resolution Δx and a temporal resolution Δt . The raw observational data d_{raw} can be made representative of these spatial and temporal scales, where we denote these corrected observations by \bar{d} . In Chapter 5, we already correct observed plastic concentrations for vertical turbulent mixing using the correction factor from Kukulka et al. (2012). By doing this, we make the observed plastic concentrations at the ocean surface representative of the modelled plastic concentrations for the given vertical grid resolution Δz . In the framework in Figure 6.1 we could try to make measurements representative in all spatial directions by correcting for other factors as well, such as the aforementioned windrows, and differences in concentrations between cyclonic and anticyclonic eddies. The measurement error σ_d could furthermore be made dependent on the sampling technique (e.g. the neuston net type and the filtered area/volume) and sampling conditions. This way, measurements taken under highly variable conditions can be made less influential in the data assimilation or machine learning step.

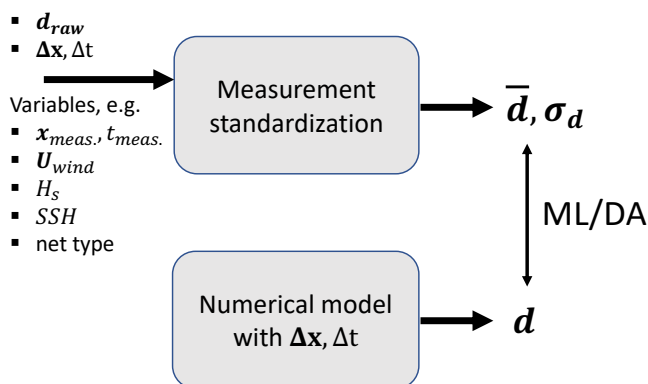


Figure 6.1: A proposed framework to standardize observational measurements to make them more representative of the numerical model output. The raw measurements (d_{raw}) can be standardized for various environmental and sampling conditions to make them more representative of the scales resolved by the numerical model, resulting in the standardized measurements \bar{d} and their estimated measurement error σ_d .

More complex models will require more complex data. In this thesis, we represented marine plastic transport using tracers. These tracers describe a large ensemble of different types of plastics. Plastics in the ocean are very complex: they differ for example in shape, size, density, and polymer type, which in turn affects processes such as biofouling, beaching, mixing, the fragmentation process, and how particles are transported horizontally and vertically. When we want to further disentangle the effects that these different plastic particle properties have on transport and fate in the marine environment, we need more complex data for model validation and calibration.

Transport of plastics between the ocean and coastlines plays a role in all chapters of this thesis. Beaches are one marine reservoir where more extensive and detailed data are necessary. We parameterised transport of plastic particles between coastal water and beaches using a beaching and resuspension time scale. Based on experimental studies (Hinata et al., 2017) it is thought that larger plastic items remain longer on beaches, due to their higher buoyancy, which makes them more likely to wash ashore by waves. When we want to create more complex parameterisations for the beaching and resuspension processes, extensive data on plastic particles in coastal waters and nearby beaches, with measured sizes, masses, and rise velocities would help greatly.

In Chapter 4 and Chapter 5 we estimated that the biofouling process has a substantial influence on the vertical transport and sinking of marine plastics. More observational data on biofilms would help to further constrain biofouling model parameters. Morét-Ferguson et al. (2010) already noted that plastic particles on beaches tend to

have a lower mass per unit volume than those in the open ocean. Besides biofouling, it was hypothesised that density changes could be caused by degradation of the polymers as well. Measuring particle densities and rise velocities from different regions of the global ocean with biofilm and after digesting the biofilm can help us to quantify to what extent vertical transport is influenced by the biofouling process.

The fragmentation model presented in Chapter 3 is simplified to keep the amount of free parameters limited. We tried to find fragmentation parameters that are approximately valid for all marine plastics. As a next step, one could try to model plastic transport and fragmentation for different polymer types separately. This would require more particle size distribution data separated by polymer type, which to our knowledge has not been published in the literature so far.

The influence of particle shape on transport is an active field of research, as shape can influence the orientation and tumbling behaviour of particles in the water, which in turn affects particle settling velocities and how particles are transported by e.g. waves (Byron et al., 2015; Clark et al., 2020; DiBenedetto et al., 2018). Data on particle shapes in the global ocean could quantify whether these kind of effects lead to preferential accumulation of differently shaped particles in different regions of the ocean, and whether more efforts should be taken in parameterising the effect of shape on large scale plastic transport.

Data-driven modelling of plastic pollution allows us to learn more about the marine environment. Using plastic pollution as an oceanic tracer has already allowed us to study surface ocean dynamics. For example, studies have compared numerical model output to observed plastic concentrations, to assess the influence of different types of forcings such as the Ekman and geostrophic ocean currents, and wave-induced currents (Onink et al., 2019; Cunningham et al., 2022). Hopefully, when more reliable observational data and more elaborate numerical models become available in the future, we can use data-driven models to learn more about the deep ocean and coastal processes as well.

Plankton play an important role in the global climate, by pumping down carbon to the deep sea, and thus influencing the amount of carbon dioxide in the atmosphere (Falkowski, 2012). Current biochemical models which simulate plankton ecosystems and carbon and nutrient cycles have a large amount of parameters. Data to validate these models come from satellite images used to estimate chlorophyll concentrations at the ocean surface, and very sparse in-situ observations (Aumont et al., 2015). Perhaps these kind of biochemical models can be coupled to plastic tracer models in the future, parameterising and calibrating the growth of plankton using observational data of biofilms on plastic particles from the ocean surface, the deep ocean, and ocean sediments.

Similarly, plastic pollution might allow us to learn more about coastal dynamics. Satellite-tracked surface drifters have been used in the past to study submesoscale processes and the effects of wind on coastal and estuarine systems (Davis, 1985; Meyerjürgens et al., 2019). By using machine learning methods as in Chapter 2, or by using data assimilation in coastal transport models, plastic particles could perhaps replace these kind of drifter experiments by serving as a readily available and cheap tracer.

6.2.2 Observational data

Particle sizes matter. Future observational reporting should be very clear what part of the particle size spectrum (see e.g. Figure 5.4) is actually measured, both regarding the lower and the upper size limit: the total measured mass is sensitive to the upper detection limit, while the total measured number is very sensitive to the lower detection limit. Coming back to the standardization issue of beach measurements, a strict lower size limit that is easily observed by volunteers needs to be set when reporting the number of litter items on beaches. Measuring the mass would provide a more reliable number, as the larger plastic items that contribute to most of the plastic mass are more easily observable.

Particle size distributions on a log-log scale can be misinterpreted easily. We saw in Chapter 3 that previous studies attributed the slope of particle size distributions to the dimensionality of plastic objects: e.g. a slope of 3 in the case of fragmentation of three-dimensional objects (Cózar et al., 2014). The fact that the slope of the particle size distribution changes on a log-log scale when showing the abundance on the y-axis (n , i.e. like a histogram), versus showing the normalized abundance on the y-axis ($n \text{ mm}^{-1}$, i.e. like a probability density function) was overlooked. Incorrectly linking particle size distributions with a slope of ~ 3 to the fragmentation of three-dimensional objects is still observed in some recent studies, e.g. Sorasan et al. (2022) and Egger et al. (2020b). We hope that future studies can avoid this oversight, and possibly make use of the fragmentation model that was presented in this thesis.

More automation of plastic item detection will help greatly. Automating the detection of plastic items from (hyperspectral) images would greatly help all aforementioned issues in this chapter.

The first motivation for more automatic detection lies in the fact that more data are necessary of large plastic items in the ocean, which are calculated to be very sparse, but which contribute to most of the plastic mass in the marine environment. Installing camera systems on vessels and automatically detecting floating plastic objects has been shown to be possible in e.g. de Vries et al. (2021). These methods can be used to quantify the amount of large floating objects, and allow for automatically estimating the object dimensions. With enough labelled data, these systems can also be trained to give insights in the type of objects (e.g. nets and buoys), which is useful to further quantify the contribution of fisheries to the marine plastic pollution problem. Remote sensing of marine plastic debris using satellites is still under development, but could additionally provide a continuous stream of global data in the future (Martínez-Vicente et al., 2019).

Secondly, automating plastic detection can help to create the complex datasets that will be necessary for future model calibration. Hyperspectral camera systems are being developed, which can classify the polymer type of plastic particles (Henriksen et al., 2022). With particle detection algorithms applied to the hyperspectral images, properties such as length and shape can automatically be quantified. This information can be used to gain a better understanding of the fragmentation process per polymer type as mentioned earlier, and size-dependent or shape-dependent transport.

Finally, automation will help greatly in standardization, when the same algorithms are used to quantify plastic particle properties across studies instead of measuring these by hand. The plastic research community would greatly benefit from a centralized database where observational data can be stored, such that different detection and analysis methods can be tested, cross-checked, and standardized.

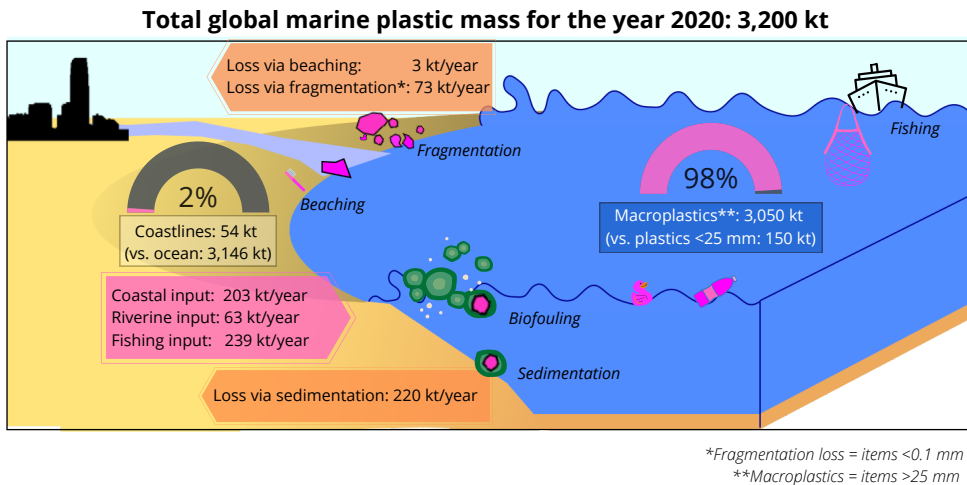


Figure 6.2: A closed mass budget for marine plastics. Much of the work presented in this thesis built up to Chapter 5, where we calculated the global marine plastic mass budget. In this schematic, we present an overview of the most important sources and sinks of plastics in the global marine environment.

Supporting material Chapter 2

A.1 Observational and modelled data per year

Figure A.1a and Figure A.1b present the modelled litter quantities (left columns) and the raw observational data (right columns) per year per cleanup stage. The litter concentrations are plotted using circles, where the color and size correspond to the litter quantities (note the logarithmic scale here). Table A.1 presents the total gathered litter per year.

A.2 Extended results

A.2.1 Gini importance overview

A complete overview of the Gini importance for all features is presented in Figure A.2. The numbers in the feature labels give information on the radius (in kilometers) and lead time (in days) if applicable, and in this order. See Table 2.2 for the radius and lead time combinations used for the variables. The Lagrangian model features (orange labels) are indicated by 'beaching_p', 'beaching_r', 'beaching_f', for litter sources originating from mismanaged coastal plastic waste (p), rivers (r), and fishing activity (f) respectively.

Table A.1: Overview of the total amount of litter gathered per year during the beach cleanup tours.

| Year | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 |
|----------------------------|--------|-------|--------|--------|--------|--------|
| Total litter gathered [kg] | 20,078 | 9,872 | 19,203 | 14,863 | 11,163 | 10,991 |

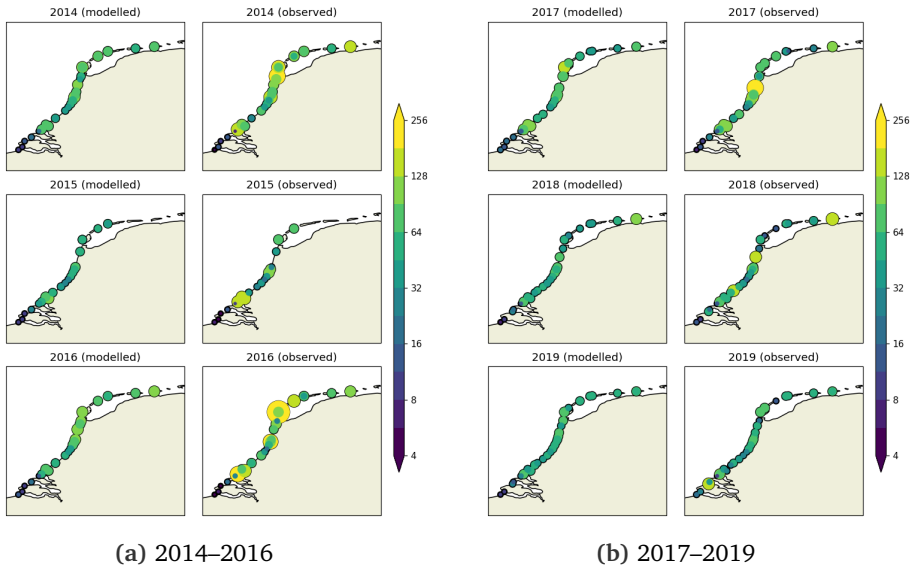


Figure A.1: Modelled (left column) and observed (right column) litter concentrations in kg km^{-1} per individual location and year. Circles are scaled and colored according to the litter concentrations.

A.2.2 Excluding Lagrangian model features

A scatter plot of the measured litter concentrations versus the predicted values is presented in Figure A.3, where Lagrangian model features have been excluded from the feature set. As described in the main text, no significant decrease in the correlation is observed compared to the case where Lagrangian model features have been included (0.70 ± 0.10 versus 0.71 ± 0.11).

The complete overview of the feature Gini importances corresponding to the case without Lagrangian model features is presented in Figure A.4. As mentioned in the main text, more features related to the currents and Stokes drift orientation with respect to the coastline are seen as important now, compared to Figure A.2. This could be explained due to these features taking over the role of the Lagrangian model features in capturing the effect of marine litter transport.

A.2.3 Effect of using only the top N features

It is not necessary to include all 66 feature clusters for predictive capability of the model. In Figure A.5 we present the Pearson correlation coefficient R as a function of the number of features included in the random forest algorithm, both with and without using the Lagrangian model features. Each time only the top features (corresponding to Figure A.2 and Figure A.4) are used to train and test the model, using 10 times repeated 5-fold cross validation. Generally the model performs well with about 7–8 features used. Performance is quite stable in the case when Lagrangian model features are used, some outliers with lower Pearson correlation coefficients can be observed when not taking into account these features. The Pearson corre-

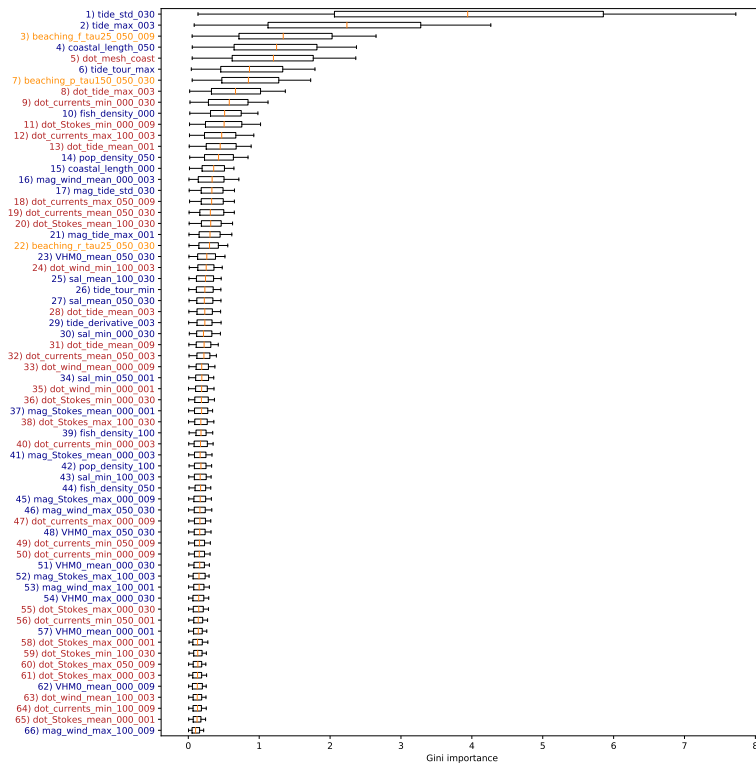


Figure A.2: Gini importance overview of all features; again labels are colored according to the feature categories in Table 2.2.

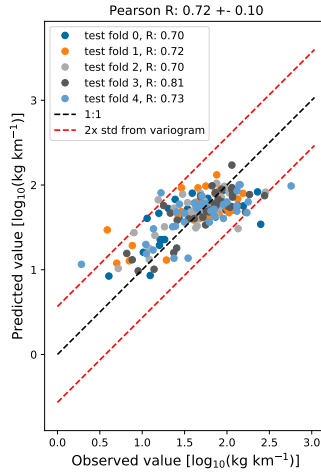


Figure A.3: Scatter plot of the observed litter quantities (x-axis), and the modelled litter quantities (y-axis), when not taking Lagrangian model features into account. Litter quantities are log-transformed, and points are colored according to the 5 test folds used in the analysis.

lation coefficient when using all 66 features, corresponding to Figure 2.4, is shown using the red error bar. In this case the Pearson correlation coefficient is slightly smaller than when using, for example, the top 8 features, which could indicate a small amount of overfitting, although this difference is not significant.

In Figure A.6, we analyse the effect of leaving out certain feature categories on the model performance. The random forest can create a highly non-linear map between the features and corresponding response. It is therefore possible that when using a large set of features and leaving out one important explanatory variable, it will use a combination of the remaining features to still obtain a good fit. We therefore only use the top 10 features in this analysis, and exclude the Lagrangian model variables, as these implicitly contain information on the other features. As can be seen, leaving out a certain category of features reduces the model performance. This can especially be observed when leaving out all features regarding tides, and the two features regarding coastal properties (l_{coast} and $n_{grid} \cdot n$). The mean Pearson correlation coefficient decreases and the variance of the model performance increases.

A.2.4 Number of participants

As mentioned in the main text, the number of participants is not seen as a important in terms of the Gini importance. The number of participants is correlated with the population density in the neighborhood of the stage, and is therefore assigned to the same feature cluster as the population density, for more details see Appendix A.3. The number of participants was not picked out of this cluster as one of the most important features during the k-fold cross validation. In order to separate the effect of the number of participants per cleanup stage, a model run was done without the

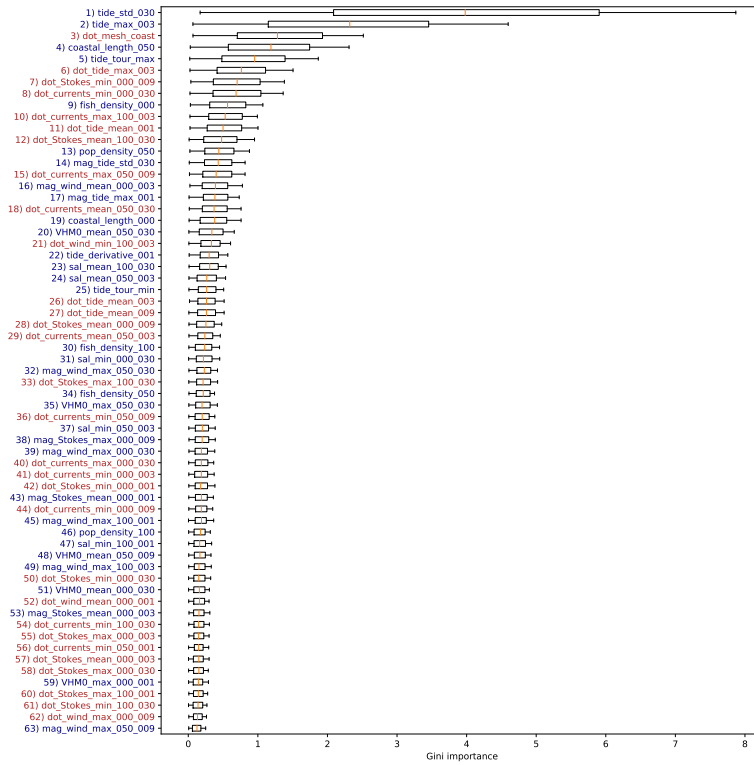


Figure A.4: Gini importance overview when not taking into account the Lagrangian model features, where labels are colored according to the feature categories in Table 2.2.

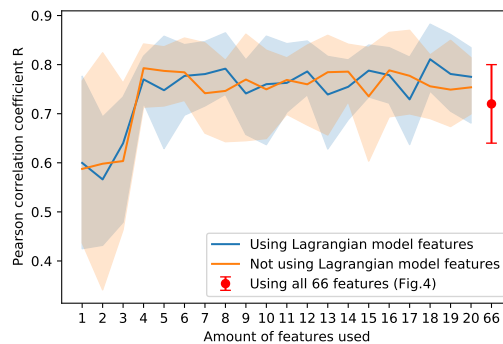


Figure A.5: The effect of the number of included features on the Pearson correlation coefficient R , where the mean is plotted using the solid line, and the filled area represents the 10% to 90% quantile. Both the case with and without using Lagrangian model features is presented (blue and orange lines respectively). The case using all 66 features (corresponding to Figure 2.4) is shown using the red error bar

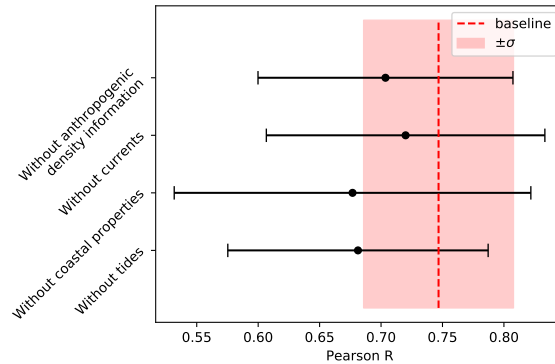


Figure A.6: Analysis where some of the feature categories have been left out. The top 10 features have been used without the Lagrangian model features (see Figure A.4, as these implicitly contain information on all feature categories). As can be observed, leaving out a set of features generally decreases the predictive performance of the model, and increases the variability of the prediction quality.

nearby population densities as features. A summary of the resulting Gini importances is shown in Figure A.7, where only the top 10 features and the number of participants are plotted.

A.2.5 Feature effect

The general effect of some features was described in the main text, such as the fact that an increasing tidal variability, and misalignment of the high resolution coastline with respect to the numerical model coastline ($\mathbf{n}_{\text{grid}} \cdot \mathbf{n}$) lead to less observed litter. Figure A.8 illustrates this, by varying one feature on the x-axis, and plotting the resulting predictions on the y-axis. In the decision trees of the random forest, decision boundaries are made at optimal splitting locations, making the resulting model highly non-linear. This makes it difficult to interpret the regression model. In Figure A.8, we 'fix' all features except the one listed on the x-axis. This feature is then varied from its minimum until its maximum encountered value. Since the random forest result can depend highly on the exact value of the other features, noise is introduced. Each other feature is varied uniformly between its 0.4–0.6 quantile, to illustrate whether the found relation for the given feature on the x-axis is robust. Features which show relatively robust relations are related to tidal height, where an increasing variability, and a higher maximum decrease the predicted litter concentrations. The effect for $\mathbf{n}_{\text{grid}} \cdot \mathbf{n}$ also seems to be robust, with increasing values leading to more predicted litter. For the coastal length in the neighborhood (l_{coast}) an increasing value seems to lead to less litter, although there is a sudden drop observed here. This might be caused by the fact that there are relatively little data points available where this feature has a high value (most of the stages were conducted on relatively straight coastline sections), so the model has trouble learning a relation here. For the Lagrangian model features, increasing values lead to more predicted litter as expected. For the mismanaged coastal plastic waste (indicated by 'beaching_p_tau25_050_009'), the results are quite dependent on the values of

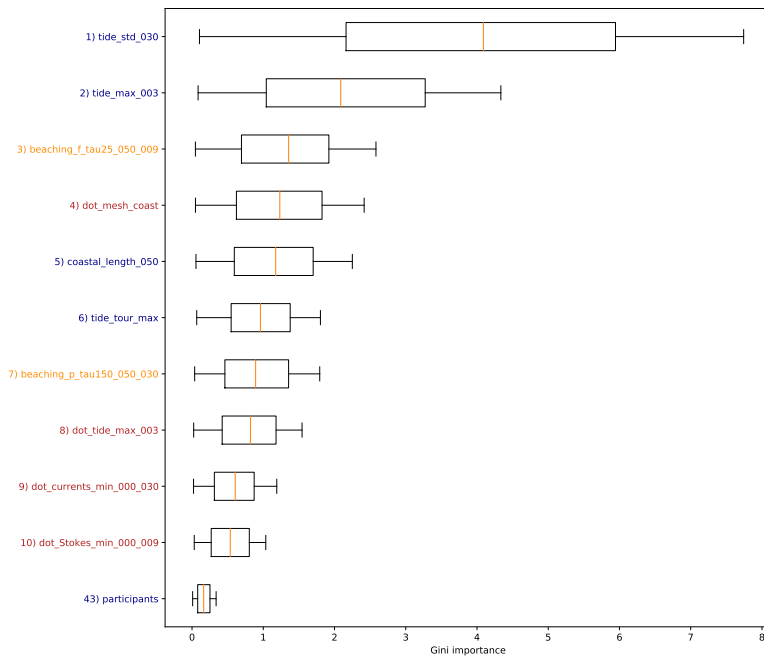


Figure A.7: Gini importance overview when not using nearby population densities as features, to separate the effect of the number of participants per cleanup stage. In that case, it is the 28th most important feature.

other features, as a lot of noise can be seen here. Generally, the model indicates there are increasing litter concentrations for increasing currents and on-shore Stokes drift.

A.3 Clustering dendrogram

Correlated features are put into clusters using hierarchical Ward-linkage clustering (McCann et al., 2019; Cope et al., 2017). An overview of the resulting dendrogram is shown in Figure A.9. A threshold is chosen to make a cut in the dendrogram. This was selected by hand to be a value of 2.3, at which the clusters remain relatively interpretable (e.g. separate clusters for coastal properties and tidal properties). The cut is shown in the figure by the red dashed line. Some general patterns regarding the clusters are indicated in the dendrogram.

A.4 Model pipeline

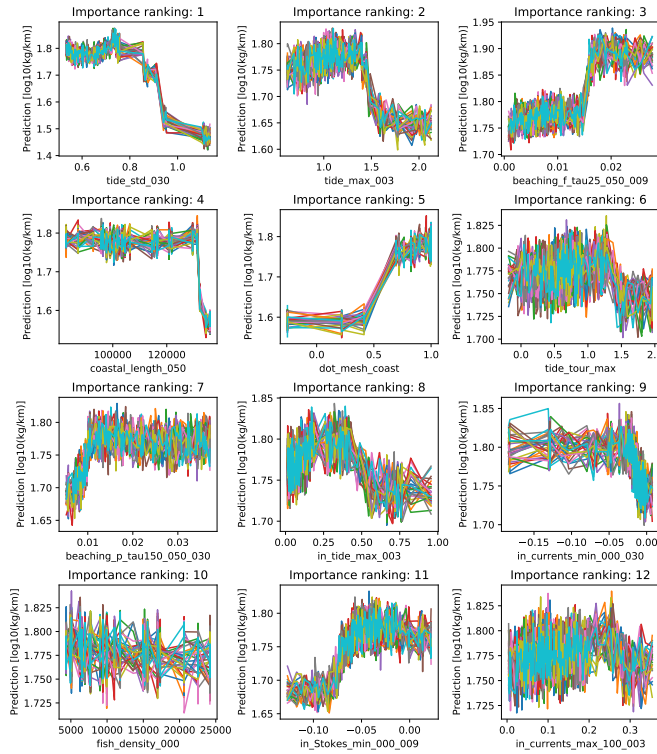


Figure A.8: Illustrated effect of the 12 most important features (x-axes) on the litter concentrations (y-axes) according to the random forest regression model. For the 12 important features, we vary their value from the minimum to maximum encountered value. All other features are fixed, and some noise is added to illustrate robustness of the relations.

A

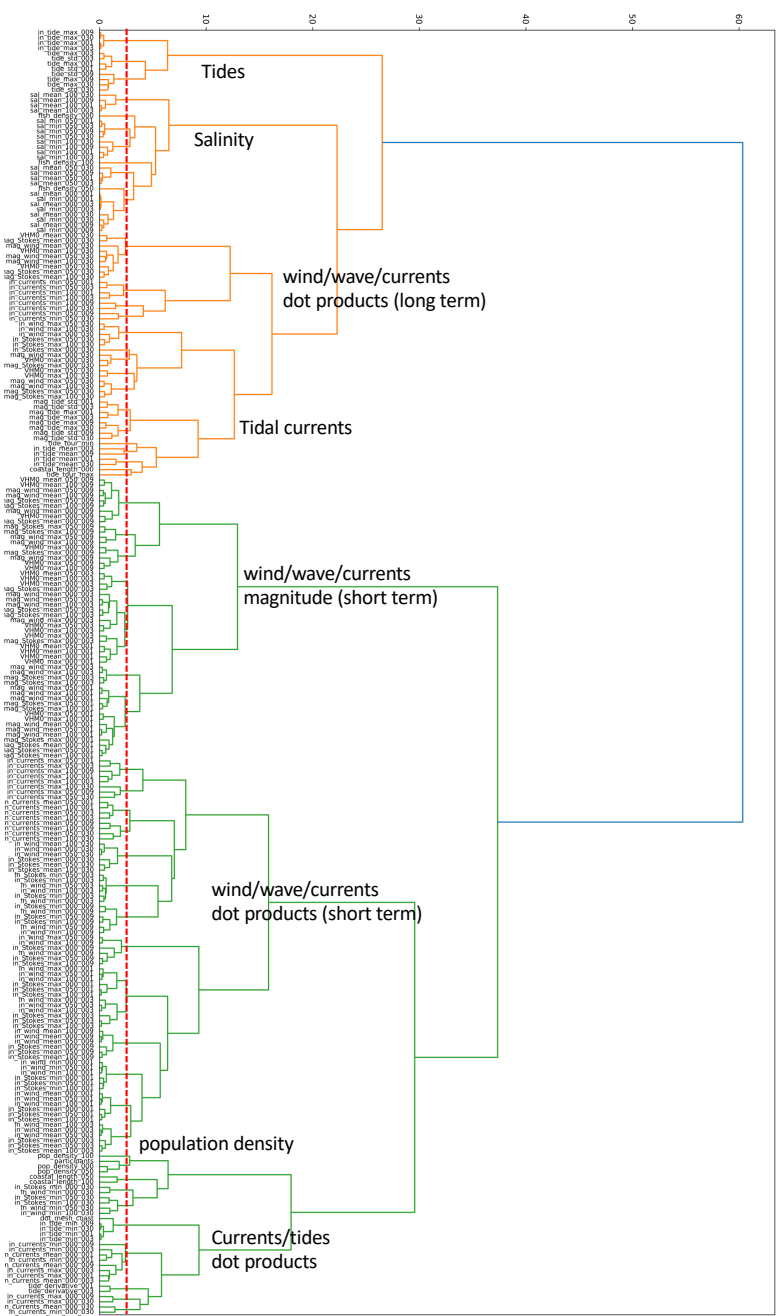


Figure A.9: Dendrogram used to construct the feature clusters

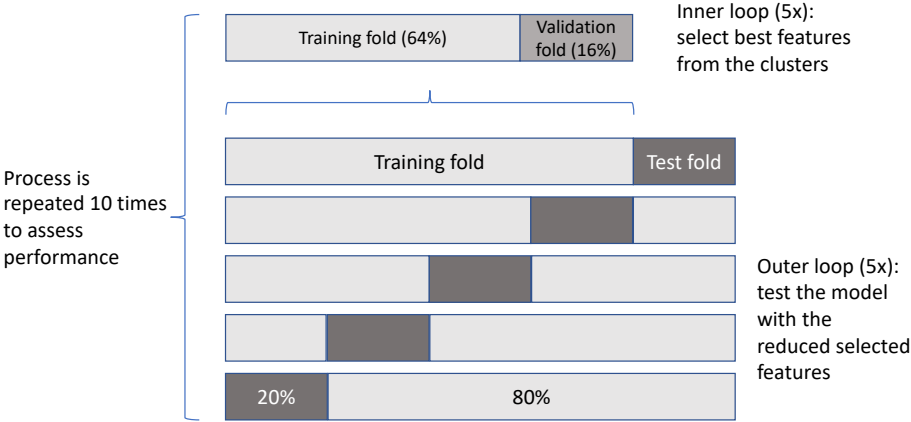


Figure A.10: Pipeline to train and test the random forest regression model. Nested k-fold cross validation is used to select the best feature from each cluster (inner loop), and to evaluate the model trained with the best features on the test data set (outer loop). The process is repeated to assess the average performance.

Supporting material Chapter 3

B.1 Additional results: parameter settings and Mass Size Distributions (MSDs)

The main text presents the transient NSD results for the baseline scenario, and various steady state NSD results for scenarios where model parameters are varied. An overview of the box model parameters is given in table B.1. The corresponding MSDs are presented in Figure B.1. Results show clearly that the scenarios have a strong effect on the estimated mass for given particle sizes. Increased fragmentation, either resulting from increased fragmentation rates or increased transport towards beaches, results in the smaller particles making up more of the total mass. The results of these simulations should be verified in the future with more in-situ measurements, as this will give us more insight in the marine plastic mass budget.

Table B.1: Environmental box model parameters used to simulate the PSDs in Figure 3.6 and Figure 3.7 of the main text, and Figure B.1. Transition probabilities are given per week.

| | Baseline | Size-dependent ocean transport | Size-dependent resuspension | Increased frag. | Increased input | fig.7 main text |
|-----------------------------------|---------------------|---|---|---------------------|---------------------|---|
| $P_{O,O}$ | $7.2 \cdot 10^{-1}$ | $7.3 \cdot 10^{-1}$ – $8.7 \cdot 10^{-1}$ | $7.2 \cdot 10^{-1}$ | $7.2 \cdot 10^{-1}$ | $6.5 \cdot 10^{-1}$ | $7.4 \cdot 10^{-1}$ – $8.8 \cdot 10^{-1}$ |
| $P_{O,C}$ | $2.7 \cdot 10^{-1}$ | $1.2 \cdot 10^{-1}$ – $2.6 \cdot 10^{-1}$ | $2.7 \cdot 10^{-1}$ | $2.7 \cdot 10^{-1}$ | $2.4 \cdot 10^{-1}$ | $1.2 \cdot 10^{-1}$ – $2.6 \cdot 10^{-1}$ |
| $P_{C,O}$ | $3.4 \cdot 10^{-2}$ | $3.2 \cdot 10^{-2}$ – $3.4 \cdot 10^{-2}$ | $3.4 \cdot 10^{-2}$ | $3.4 \cdot 10^{-2}$ | $3.0 \cdot 10^{-2}$ | $3.2 \cdot 10^{-2}$ – $3.4 \cdot 10^{-2}$ |
| $P_{C,C}$ | $8.3 \cdot 10^{-1}$ | $8.3 \cdot 10^{-1}$ – $8.3 \cdot 10^{-1}$ | $8.3 \cdot 10^{-1}$ | $8.3 \cdot 10^{-1}$ | $7.4 \cdot 10^{-1}$ | $8.4 \cdot 10^{-1}$ – $8.4 \cdot 10^{-1}$ |
| $P_{C,B}$ | $1.3 \cdot 10^{-1}$ | $1.3 \cdot 10^{-1}$ | $1.3 \cdot 10^{-1}$ | $1.3 \cdot 10^{-1}$ | $1.2 \cdot 10^{-1}$ | $1.3 \cdot 10^{-1}$ |
| $P_{B,C}$ | $3.2 \cdot 10^{-2}$ | $3.2 \cdot 10^{-2}$ | $2.0 \cdot 10^{-3}$ – $4.5 \cdot 10^{-1}$ | $3.2 \cdot 10^{-2}$ | $2.9 \cdot 10^{-2}$ | $2.0 \cdot 10^{-3}$ – $4.5 \cdot 10^{-1}$ |
| $P_{B,B}$ | $9.6 \cdot 10^{-1}$ | $9.6 \cdot 10^{-1}$ | $5.5 \cdot 10^{-1}$ – $9.9 \cdot 10^{-1}$ | $9.6 \cdot 10^{-1}$ | $8.6 \cdot 10^{-1}$ | $5.5 \cdot 10^{-1}$ – $9.9 \cdot 10^{-1}$ |
| λ [f year ⁻¹] | $1.8 \cdot 10^{-2}$ | $1.8 \cdot 10^{-2}$ | $1.8 \cdot 10^{-2}$ | $3.3 \cdot 10^{-1}$ | $1.8 \cdot 10^{-2}$ | $1.0 \cdot 10^{-2}$ |
| input [t yr ⁻¹] | 2,500 | 2,500 | 2,500 | 2,500 | 100,000 | 2,500 |
| P_S | $5.1 \cdot 10^{-3}$ | $5.1 \cdot 10^{-3}$ | $5.6 \cdot 10^{-4}$ | $5.1 \cdot 10^{-3}$ | $1.1 \cdot 10^{-1}$ | $5.0 \cdot 10^{-4}$ |

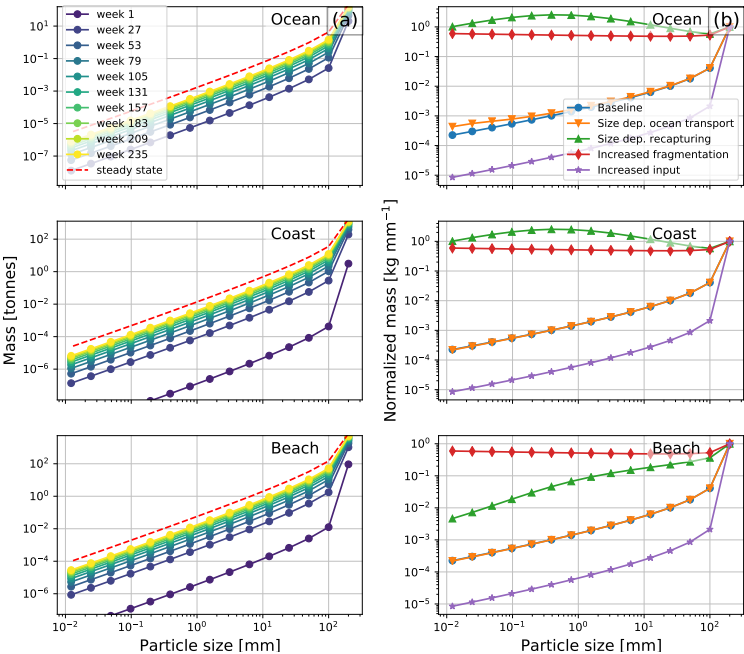


Figure B.1: Modelling MSDs using the environmental box model. Column a: transient response to a constant input of particles into the model (baseline scenario). Column b: steady state normalized MSDs for different environmental scenarios, normalized to the amount of parent particles (200 mm)

Table B.2: Power law slopes α obtained by maximizing equation (6) in the main text to the modelled and measured PSDs. The size range for which the power law holds is estimated by minimizing the Kolmogorov-Smirnov statistic of the observed data with respect to the fitted power law, see Virkar & Clauset (2014). For beach samples no lower bound is estimated given the low amount of reported bins. Uncertainties in α (± 1 times the standard deviation) are not calculated for the model, since this depends on the amount of samples in the bins, which is undefined in this case.

| Domain | α | Minimum size |
|--|-------------------------------------|----------------|
| Ocean | 2.73 | ≥ 3.13 mm |
| Ocean + mixing ($U_{10} \approx 4$ m/s) | 2.63 | ≥ 6.25 mm |
| Ocean + mixing ($U_{10} \approx 7$ m/s) | 2.37 | ≥ 12.5 mm |
| Cózar et al. (2015) | 2.53 ± 0.04 | ≥ 3.16 mm |
| Coast | 2.69 | ≥ 3.13 mm |
| Coast + mixing ($U_{10} \approx 4$ m/s) | 2.60 | ≥ 6.25 mm |
| Coast + mixing ($U_{10} \approx 7$ m/s) | 2.34 | ≥ 12.5 mm |
| Ruiz-Orejón et al. (2018) | 2.49 ± 0.06 | ≥ 2.50 mm |
| Beach | 2.02 | ≥ 1.56 mm |
| Fok et al. (2017) | 1.60 ± 0.01 | ≥ 0.32 mm |
| Constant et al. (2019) | $1.47 \pm 0.01,$ 1.45 ± 0.03 | ≥ 0.06 mm |

An overview of resulting power law slopes α for Figure 3.7 in the main text are presented in table B.2, as well as the estimated power law size range. The model predicts a slightly larger α in the ocean compared to the coast. A similar difference is seen between the chosen sets of measurements near the coast (Ruiz-Orejón et al., 2018) and further away from the coast (Cózar et al., 2015), although this difference is not significant. Including vertical mixing has a strong effect on the estimated α . The modelled values for α are somewhat higher than the measured ones with no or calm wind conditions. Some differences can be expected since the environmental box model is highly idealized, so comparison between the various model simulations are of more interest. The model is quite sensitive to the fragmentation parameters p and λ , which will be discussed later in Figure B.2. More experimental data is necessary to further constrain these parameters, hopefully reducing mismatch between the model and measurements.

Additional results were generated to quantify the sensitivity of parameter value estimates and assumptions made in the main text. Results are presented in Figure B.2 and Figure B.3. We only show results for the ocean cell in the box model, since the sensitivity studies include no size-dependent transport and hence show no qualitative differences between the different cells.

We first present the baseline scenario introduced in the main text, for which the parameters were shown in table B.1. For the baseline scenario, we assume all new plastic waste is introduced on the beach. In reality this is of course more complicated, as rivers might directly transport waste into coastal waters, and there might be sources in the open ocean due to, for example, fishing activity and shipping. Regarding plastic fragments found in the surface waters, land-based waste and rivers are likely the main sources (see Chapter 4). We test what happens when we introduce new plastic particles on both the beach and coastal waters assuming a 50%-50% split,

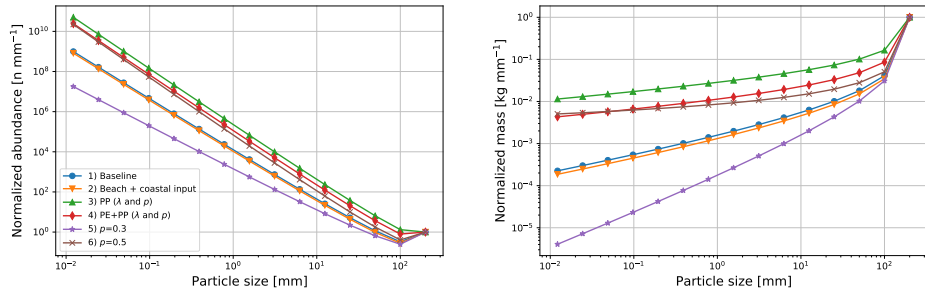


Figure B.2: Sensitivity study regarding the input scenario (50%-50% split of new plastics entering the beach and coastal waters), and sensitivity studies regarding polymer choice and fragmentation parameters.

indicated by scenario 2 in Figure B.2. There is little effect from this assumption. This can be explained due to relatively large exchange from the coastal cell to the beach cell, compared to that from the coastal cell to the ocean cell (see the transition probabilities in table B.1), leading to only a small delay in fragmentation.

Polymer choice has a significant effect on the modelled PSDs. We compare the baseline (based on PE), to a case where λ and p are based on PP (scenario 3 in Figure B.2). The slope of the NSD significantly increases, due to the increased fragmentation rate λ and a higher p . We can test what happens when a combination of PE and PP is used (scenario 4 in Figure B.2). We use a mixture of 60% PE and 40%, based on European plastic demand (Bond et al., 2018). In that case, we get a combination of the baseline (PE) results, and the results for PP. We furthermore show sensitivity to the value of p (scenario 5 and 6 in Figure B.2). The value of p has a significant effect on the slope of the PSDs. Future work should further look at constraining the fragmentation parameters (p and λ), as these clearly have a significant effect on the results.

In Figure B.3 we first analyze what happens when taking the length of the plastic input objects as 400mm instead of 200mm from the baseline scenario. Qualitatively there is not much difference, except that the PSD shifts to the right. The slope remains the same, so this assumption has no big influence on the results presented in the main text. In scenario 8 and scenario 9 we analyze what happens when adjusting the resuspension time, as this was based on an analysis in Japan (Hinata et al., 2017), and might be different for our test case in the Mediterranean Sea. We see that decreasing $\tau_{B,C}$ from 211 days (baseline) to 100 days leads to a decrease in the time that particles spend on the beach, leading to less fragmentation, and fewer small fragments. The opposite holds for increasing $\tau_{B,C}$ to 400 days, leading to more small fragments. There is almost no effect on the slope of the NSD, only on the amount of fragments per parent object, so the assumption of the baseline value of $\tau_{B,C}$ has little effect on the results in the main text.

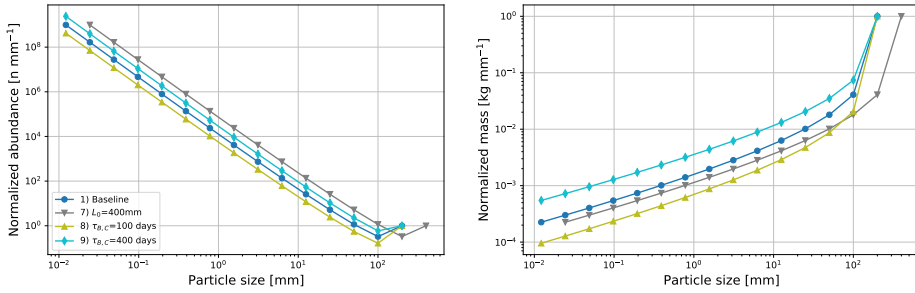


Figure B.3: Sensitivity study for the initial size of new plastics, and residence times of plastic particles on the beach.

B.2 Box model: transition matrix formulation

In this section we describe how the transition matrix underlying the environmental box model is set up, and how the steady-state is calculated. As explained in the main text, the transition matrix captures the effect of transport in the environment, as well as the effect of fragmentation. The transition matrix describes these effects for a given time step, set to one week. As a simple example, we present the case where the fragmentation rate λ is $1 f \text{ week}^{-1}$, and we take $p = 0.5$, $D_N = 3$. For size class $k = 0, 1, 2, \dots$ the cascading fragmentation model at $f = 1$ yields in terms of abundance $n = 0.5, 2, 8, \dots$ fragments per parent particle, and in terms of mass fractions $m = \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots$.

The transition matrix in terms of abundance (\mathbf{T}_n) is presented in Figure B.4, the one in terms of mass (\mathbf{T}_m) in Figure B.5. The example is shown for three different size classes only. For notation, we use O, C, and B to denote ocean, coast, and beach respectively, subscripts are used to denote the size class k , superscripts are used for time indices i . On the left hand side of Figure B.4 and Figure B.5, we have the abundance or mass at time $i + 1$, on the right hand side we have the transition matrix multiplying the abundance or mass at time i . Cells in the transition matrix containing zeros are left empty. The two coloured blocks (red and yellow) indicate the entries where fragmentation is active. As can be seen the matrix is quite sparse. This is because fragmentation is only active on the beach, so particles do not move from one size class to another in the coastal water or ocean. On the beach, particles can move to a higher or equal size class k under influence of fragmentation. The upper right entries of the red and yellow blocks are zero, since particles can not move to a lower size class (i.e. they cannot increase in size). If one has an infinite amount of size classes, the operation in Figure B.5 would be mass conserving. Due to a finite amount of size classes, there is some mass loss due to fragmentation however. Finally, the entire transition matrices in Figure B.4 and Figure B.5 are multiplied with the scalar P_S , as particles are removed from the system by sinks (e.g. plastic particles sinking down, or being removed from beaches). As explained in the main, text P_S is assumed to be the same in all environmental boxes, however this value could be varied per environmental box if indications for this become available in future studies.

B

| | | | | | | | | | | | |
|--------------------|---|-----------|-----------|-----------|-----------|-----------|-----------|-------------------------|-------------------------|-------------------------|----------------|
| $O, n_{k=0}^{i+1}$ | = | $P_{O,O}$ | | | $P_{C,O}$ | | | | | | $O, n_{k=0}^i$ |
| $O, n_{k=1}^{i+1}$ | | | $P_{O,O}$ | | | $P_{C,O}$ | | | | | $O, n_{k=1}^i$ |
| $O, n_{k=2}^{i+1}$ | | | | $P_{O,O}$ | | | $P_{C,O}$ | | | | $O, n_{k=2}^i$ |
| $C, n_{k=0}^{i+1}$ | | $P_{O,C}$ | | | $P_{C,C}$ | | | $\frac{1}{2} * P_{B,C}$ | | | $C, n_{k=0}^i$ |
| $C, n_{k=1}^{i+1}$ | | | $P_{O,C}$ | | $P_{C,C}$ | | | $\frac{2}{* P_{B,C}}$ | $\frac{1}{2} * P_{B,C}$ | | $C, n_{k=1}^i$ |
| $C, n_{k=2}^{i+1}$ | | | | $P_{O,C}$ | | $P_{C,C}$ | | $\frac{4}{* P_{B,C}}$ | $\frac{2}{* P_{B,C}}$ | $\frac{1}{2} * P_{B,C}$ | $C, n_{k=2}^i$ |
| $B, n_{k=0}^{i+1}$ | | | | | $P_{C,B}$ | | | $\frac{1}{2} * P_{B,B}$ | | | $B, n_{k=0}^i$ |
| $B, n_{k=1}^{i+1}$ | | | | | | $P_{C,B}$ | | $\frac{2}{* P_{B,B}}$ | $\frac{1}{2} * P_{B,B}$ | | $B, n_{k=1}^i$ |
| $B, n_{k=2}^{i+1}$ | | | | | | | $P_{C,B}$ | $\frac{4}{* P_{B,B}}$ | $\frac{2}{* P_{B,B}}$ | $\frac{1}{2} * P_{B,B}$ | $B, n_{k=2}^i$ |

$* P_S$

Figure B.4: Box model transition matrix in terms of abundance, for three different size classes $k = 0, 1, 2$, and after $f = 1$

Size-dependent transport is easily added to the transition matrix, by varying the transition probabilities per size class. All columns are properly normalized (i.e. column sums of 1), such that there would be conservation of mass when there would be no fragmentation or sinks in the system.

Since there is a constant input of new parent objects into the system, the mass at time $i + 1$ is given by the mass at time i and the addition of new mass \mathbf{m}_{in} :

$$\mathbf{m}^{i+1} = \mathbf{T}_m \mathbf{m}^i + \mathbf{T}_m \mathbf{m}_{in} \quad (\text{B.1})$$

We want to estimate the steady state mass of the box model (\mathbf{m}^{ss}), which is done by setting

$$\mathbf{m}^{i+1} = \mathbf{m}^i = \mathbf{m}^{ss}, \quad (\text{B.2})$$

which yields:

$$\mathbf{m}^{ss} - \mathbf{T}_m \mathbf{m}^{ss} = \mathbf{T}_m \mathbf{m}_{in} \quad (\text{B.3})$$

$$\mathbf{m}^{ss} = (\mathbf{I} - \mathbf{T}_m)^{-1} \mathbf{T}_m \mathbf{m}_{in}. \quad (\text{B.4})$$

The transition matrix \mathbf{T}_m has one unknown parameter, P_S . This value is optimized such that the steady state floating mass is equal to the midpoint estimate for the Mediterranean sea of 2,000 metric tonnes by C  zar et al. (2015).

| | | | | | | | | | | |
|--------------------|---|-----------|-----------|-----------|-----------|-----------|-------------------------|-------------------------|-------------------------|----------------|
| $O, m_{k=0}^{i+1}$ | = | $P_{O,O}$ | | | $P_{C,O}$ | | | | | $O, m_{k=0}^i$ |
| $O, m_{k=1}^{i+1}$ | | | $P_{O,O}$ | | | $P_{C,O}$ | | | | $O, m_{k=1}^i$ |
| $O, m_{k=2}^{i+1}$ | | | | $P_{O,O}$ | | | $P_{C,O}$ | | | $O, m_{k=2}^i$ |
| $C, m_{k=0}^{i+1}$ | | $P_{O,C}$ | | | $P_{C,C}$ | | $\frac{1}{2} * P_{B,C}$ | | | $C, m_{k=0}^i$ |
| $C, m_{k=1}^{i+1}$ | | | $P_{O,C}$ | | | $P_{C,C}$ | $\frac{1}{4} * P_{B,C}$ | $\frac{1}{2} * P_{B,C}$ | | $C, m_{k=1}^i$ |
| $C, m_{k=2}^{i+1}$ | | | | $P_{O,C}$ | | $P_{C,C}$ | $\frac{1}{8} * P_{B,C}$ | $\frac{1}{4} * P_{B,C}$ | $\frac{1}{2} * P_{B,C}$ | $C, m_{k=2}^i$ |
| $B, m_{k=0}^{i+1}$ | | | | | $P_{C,B}$ | | $\frac{1}{2} * P_{B,B}$ | | | $B, m_{k=0}^i$ |
| $B, m_{k=1}^{i+1}$ | | | | | | $P_{C,B}$ | $\frac{1}{4} * P_{B,B}$ | $\frac{1}{2} * P_{B,B}$ | | $B, m_{k=1}^i$ |
| $B, m_{k=2}^{i+1}$ | | | | | | $P_{C,B}$ | $\frac{1}{8} * P_{B,B}$ | $\frac{1}{4} * P_{B,B}$ | $\frac{1}{2} * P_{B,B}$ | $B, m_{k=2}^i$ |
| $* P_S$ | | | | | | | | | | |

Figure B.5: Box model transition matrix in terms of mass, for three different size classes $k = 0, 1, 2$, and after $f = 1$

B.3 Vertical mixing and size-dependent lateral transport

B.3.1 Modelling mixing and Stokes drift

Vertical mixing of plastic particles is estimated using the approach from Poulain et al. (2019). The rise velocity w_b is estimated by solving (Poulain et al., 2019; Clift et al., 1978):

$$\frac{240}{\pi Re} (1 + 0.138 Re^{0.792}) w_b^2 = \frac{2}{15} l \left(1 - \frac{\rho_p}{\rho_w} \right) g, \quad (\text{B.5})$$

where Re is the particle Reynolds number ($Re = l w_b / \nu$, with $\nu = 1 \cdot 10^{-6}$), l is the particle size, ρ_w is the density of sea water (set to 1.029 kg m^{-3}), and g is the gravitational constant (9.81 m s^{-2}).

In our model, the rise velocities are estimated assuming an ellipsoidal shape, with a density ρ_p of 950 kg m^{-3} . This combination gives a good match with reported densities (Morét-Ferguson et al., 2010) of plastic particles, and observed rise velocities (Poulain et al., 2019), see Figure B.6. In Poulain et al. (2019) a particle equivalent length was used, for ellipsoids defined as $L_{eq} = \sqrt{L_1 L_2 / 4}$, where L_1 is the length and L_2 is the width. We assume that these are approximately equal ($L_1 \approx L_2$), to simplify the calculations in terms of a particle size l .

Taking a balance between the upward particle flux caused by buoyancy, and a Reynolds-averaged turbulent flux, results in an exponential profile for the amount of particles as a function of depth ($n(z)$):

$$n(z) \sim \exp(z w_b A_0^{-1}), \quad (\text{B.6})$$

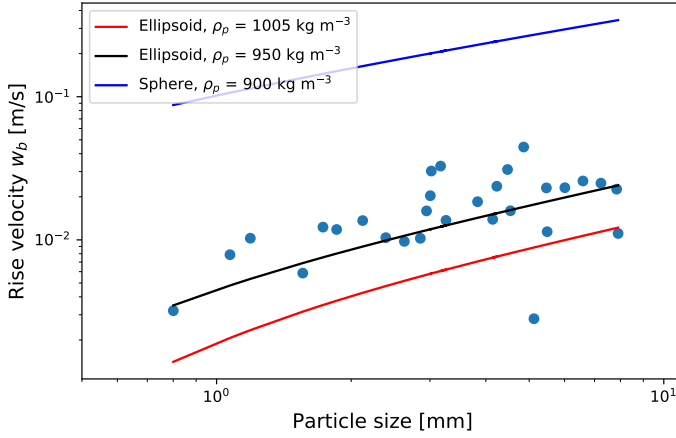


Figure B.6: Estimated rise velocities using the model from Poulain et al. (2019). In Poulain et al. (2019) it was shown that most rise velocities of plastic particles fall in between those calculated for spheres (900 kg m^{-3}) and ellipsoids (1005 kg m^{-3}). Here, we use ellipsoids with $\rho_p = 950 \text{ kg m}^{-3}$, which fit the measured rise velocities from Poulain et al. (2019) well, see the solid black line.

for $z \leq 0$, where A_0 is an eddy-viscosity parameter estimated using $A_0 = 1.5u_{*w}\kappa H_s$ (Thorpe et al., 2003), with u_{*w} being the water friction velocity, $\kappa = 0.4$ the von Karman constant, and H_s the significant wave height (Kukulka et al., 2012). Particles at different depths will be affected differently by Stokes drift. The influence of Stokes drift with depth, $v_{Stokes}(z)$, is estimated using the approach presented in Breivik et al. (2016), where the wave spectrum is approximated using the Phillips spectrum:

$$v_{Stokes}(z) = \frac{2\alpha g}{\omega_p} \left(\exp(2k_p z) - \sqrt{-2\pi k_p z} \operatorname{erfc}(\sqrt{-2k_p z}) \right), \quad (\text{B.7})$$

where α is Phillips' parameter ($8.3 \cdot 10^{-3}$), ω_p is the wave peak frequency, and $k_p = \omega_p^2/g$ is the peak wavenumber.

ERA5 reanalysis data (Hersbach et al., 2020) are used to calculate typical weather conditions for the Mediterranean Sea. In order to evaluate the vertical mixing and the Stokes drift profiles, we need data for the peak wave period ω_p , the significant wave height H_s , and the water friction velocity u_{*w} . The water friction velocity is estimated from the wind velocity at 10 meters U_{10} . First the air friction velocity u_{*a} is estimated from the drag coefficient C_D using (Thorpe et al., 2003; Geernaert, 1990):

$$u_{*a}^2 = C_D U_{10}^2, \quad C_D = 10^{-3} (0.75 + 0.067 U_{10} [m s^{-1}]), \quad (\text{B.8})$$

afterwards continuous stress at the water surface is assumed, yielding

$$\rho_w u_{*w}^2 = \rho_a u_{*a}^2. \quad (\text{B.9})$$

Statistics for ω_p , H_s , and U_{10} are calculated from ERA5 reanalysis data (Hersbach et al., 2020). Various quantiles are calculated from the data, in order to be able to

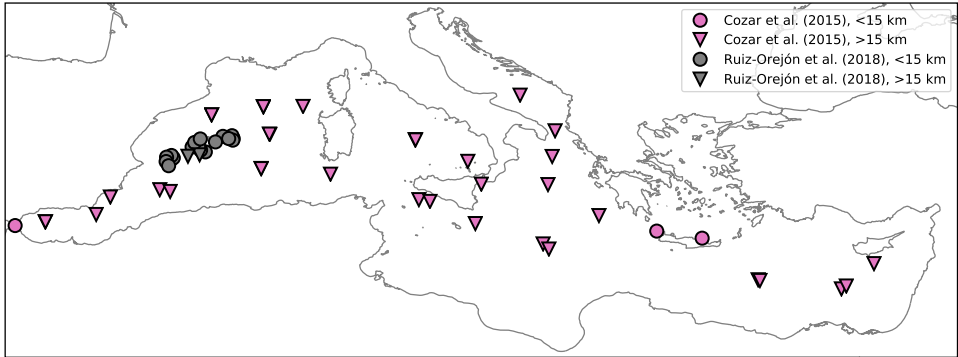


Figure B.7: Dividing the ocean into a coastal region (≤ 15 km from the coastline), and an open ocean region (> 15 km) divides the measurements from C  zar et al. (2015) and Ruiz-Orej  n et al. (2018) quite well. The majority of samples from C  zar et al. (2015) are classified as open ocean this way (triangles), the majority of samples from Ruiz-Orej  n et al. (2018) as coastal samples (circles). These measurements are used to compare the box model results with in the main text.

evaluate the environmental box model for different weather conditions as will be explained in the next Section.

B.3.2 From mixing and Stokes drift to transition probabilities

In order to determine transition probabilities between the ocean and coastal water, we use the Lagrangian model run for the Mediterranean sea from Chapter 4. The domain of the Mediterranean Sea is split into offshore and coastal regions based on the distance to the closest coastline, with a threshold of 15 km. The 15 kilometer threshold was chosen as this divides the two sets of measurements with mainly coastal samples (Ruiz-Orej  n et al., 2018) and mainly offshore samples (C  zar et al., 2015) well. This is shown in Figure B.7: coastal samples are plotted using dots, and offshore samples using triangles.

The Lagrangian particle simulation is run with the results found in Chapter 4 for five years starting from January the 1st 2010. From the total of approximately 770,000 particles, 66% are released based on the estimated mismanaged plastic waste at the coast (50 km radius)(Jambeck et al., 2015; SEDAC et al., 2005), 29% are released at river mouths based on Lebreton et al. (2017), and 6% are released at locations with high fishing activity (Kroodasma et al., 2018). Transition probabilities of particles moving between offshore and coastal cells of the mesh are calculated from this simulation.

The Stokes drift influence is varied in the simulation to estimate how this influences the transition probabilities between the ocean and coastal water. Six situations are simulated, ranging from no Stokes influence, to various fractions of Stokes drift, to full Stokes drift influence. These different situations can be translated to a certain depth, by evaluating the Stokes drift over depth using Eq. (B.7) and comparing this value to the surface Stokes drift (i.e. $\frac{v_{Stokes}(z)}{v_{Stokes}(0)}$). Resulting transition probabilities from the model runs are presented in Figure B.8. As can be seen, increas-

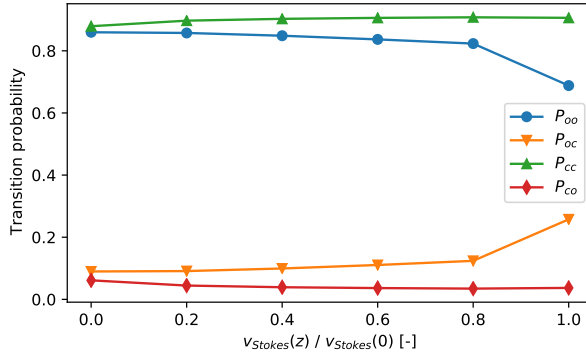


Figure B.8: Transition probabilities between coastal and ocean cells for various values of $v_{Stokes}(z) / v_{Stokes}(0)$ as calculated by the Lagrangian model run, for $dt = 7$ days

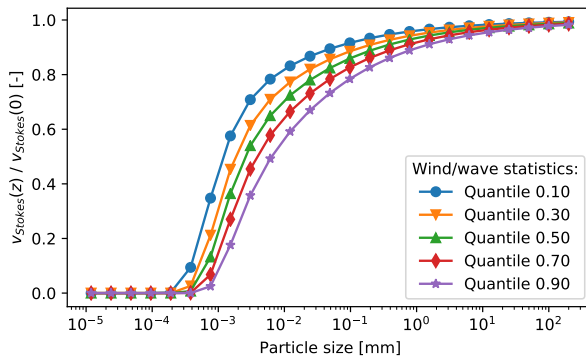


Figure B.9: Stokes influence fraction (Stokes drift at a given depth divided by the surface Stokes drift, i.e. $v_{Stokes}(z) / v_{Stokes}(0)$) for different particle sizes, under the influence of different wind/wave conditions.

ing the Stokes drift increases the probability that particles move from the ocean to coastal waters ($P_{O,C}$), decreases the probability that particles remain in the ocean ($P_{O,O}$), slightly increases $P_{C,C}$, and slightly decreases $P_{C,O}$ as expected. Even without Stokes drift, transport from the open ocean towards coastal waters can be observed. This accumulation of particles in the coastal zone means that the coastal transition probabilities are relatively insensitive to adding more Stokes drift.

Different particle sizes will be mixed to different depths using Eq. (B.5) and Eq. (B.6). The median depth of the particles is calculated, and the Stokes drift influence (i.e. the Stokes drift at this depth divided by the surface Stokes drift) is calculated using Eq. (B.7). This yields the results presented in Figure B.9. As wind speed and wave height increase, particles tend to be more mixed, and hence experience less Stokes drift.

Data in Figure B.8 are linearly interpolated for the Stokes drift influence at a given particle size as given in Figure B.9. This yields transition probabilities per particle

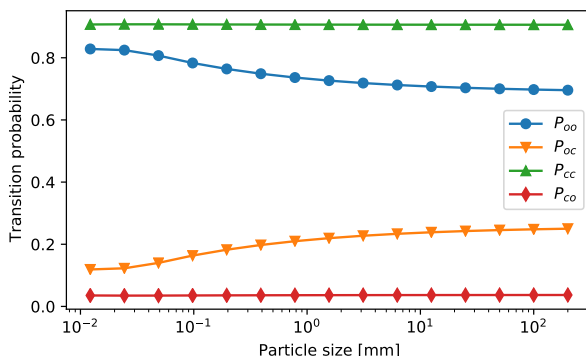


Figure B.10: Transition probabilities for the environmental box model as a function of the particle size, calculated for the median wind/wave conditions

size, results are presented in Figure B.10. These transition probabilities are used in the transition matrix of the environmental box model.

B.4 Estimated municipal plastic waste properties

Typical properties of municipal plastic waste were estimated using data from Jansen et al. (2015), in which waste mass fractions for different size classes were reported from sorting facilities in the Netherlands. Fine screen apertures are used (50-65 mm), midsize apertures (140 mm), and coarse apertures (220-250 mm) to sort the plastic waste. We take the midpoints here as the estimated aperture sizes (57.5 mm, 140 mm, and 235 mm). We assume the lower size limit to be 0 mm, and the upper size limit to be 500 mm, which is an arbitrary choice, but does not have a large impact on the results. We use the reported mass data for polyethylene and polypropylene in Jansen et al. (2015), and use the environmental polymer fractions from Suaria et al. (2016) to estimate their contribution (76% and 24% respectively). Data are parsed using the WebPlotDigitizer tool Rohatgi (2020), and are shown in Figure B.11a.

We normalize the mass percentages by the bin widths, to obtain an estimate for the normalized MSD. The point estimates are given in Figure B.11b using the dots with errorbars. From this we argue that particles around 200 mm likely contribute to most of the new plastic mass entering the environment. Of course this is still a first order estimate, and more research is needed to quantify the mass, abundance, and dimensionality of 'new' plastics entering the environment. As an example we fit a lognormal distribution through the points in Figure B.11b, which fits the larger particle size classes well, but underestimates the smallest size class. It is important to further quantify the shape of this probability density function, since it will influence the particle size distributions measured in the environment.

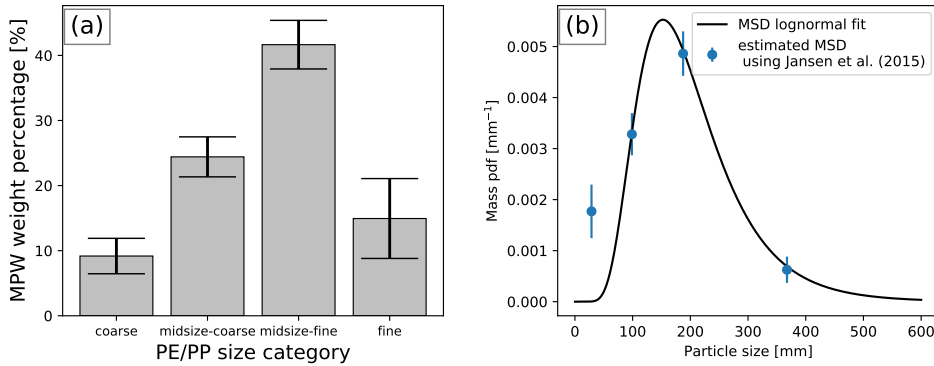


Figure B.11: Quantification of municipal plastic waste dimensions, based on data from Jansen et al. (2015). Figure a: histogram of MPW masses reported for 4 different size categories. Figure b: estimated MSD using these data.

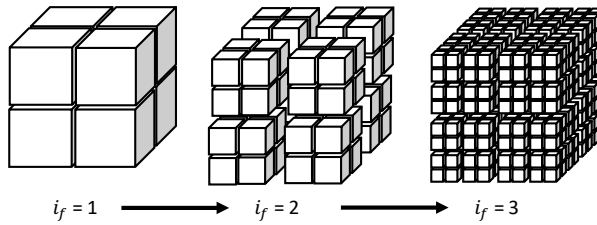


Figure B.12: Illustration of the fragmentation model by Cózar et al. (2014)

B.5 Successive fragmentation model

In the main text we mention the fragmentation model by Cózar et al. (2014), which we will call the successive fragmentation model here. We summarise the basic details here, and the difference between the resulting NSD and normalized NSD.

Objects in the successive fragmentation model are successively broken down into a set of smaller (equally sized) fragments. Some iterations are shown in Figure B.12: a cube with a length of L can be broken down into eight equally-sized smaller cubes with length $L/2$, which can again be broken down into 8 smaller cubes with length $L/4$. A resulting histogram counting the amount of particles at fragmentation index f is presented in Figure B.13.

The grey dashed line in Figure B.13 is obtained with a constant input rate of large objects with length $L = 1$ mm. The slope of this line on a log-log scale is -3. This slope is purely dependent of the spatial dimension considered for the fragmentation process. In case of the fragmentation of a 2-dimensional object (i.e. a square sheet with negligible thickness), this line would have a slope of -2 on a log-log scale.

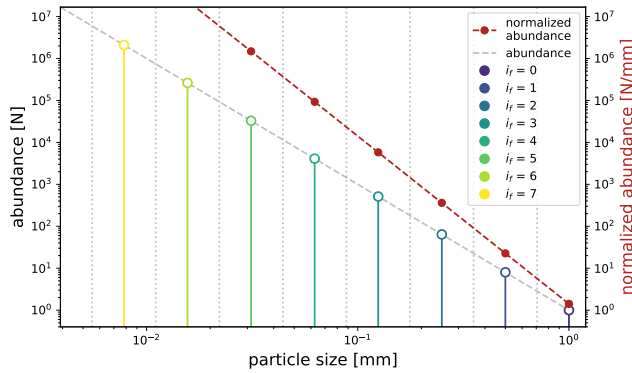


Figure B.13: Illustration of the particle size distribution resulting from the successive breakage model

Results do not change when including a shape factor (α , relating the volume (V) of the particles to their characteristic size (l) by $V = \alpha l^3$), and does not change when adjusting the amount of fragmentation events per iteration (which is 1 in this example).

This grey dashed line is not the resulting NSD however. It only connects the amount of fragments at a given discrete particle size. We can bin the amount of fragments per size class, e.g. using the bin boundaries as depicted by the vertical grey dotted lines. By dividing the amount of fragments in each bin by the width of the bin, we obtain the normalized abundance (red dashed line in Figure B.13), which is the particle size distribution normally reported in literature. Important to note is that the slope of this NSD is -4 instead of -3, due to the fact that the bin width decreases by a factor of two when going to the left on a log scale. This has been overlooked in e.g. Cózar et al. (2014), where NSD's with a slope of -3 were thought to be resulting from three-dimensional fragmentation, while in fact they correspond to two-dimensional fragmentation in the successive fragmentation model.

Supporting material Chapter 4

C.1 Measurement corrections and variance

Before using the plastic concentration measurements, they are corrected for vertical mixing using the correction factor introduced by Kukulka et al. (2012):

$$\frac{N}{N_{tow}} = \frac{1}{1 - e^{-\frac{dW_b}{A_0}}}, \quad (\text{C.1})$$

where N_{tow} is the amount of plastics measured by the net tow, N is the corrected amount of plastic particles, d is the submerged net depth, W_b is the rise velocity of the plastic particles, and A_0 is a parameter defining the near-surface turbulence. The value for d was not provided for all references used here. Where possible, it was calculated by comparing measurements in terms of volume with measurements in terms of area. Otherwise the most commonly occurring values in the data-set were used, which is $d = 0.1$ m for nets with a height of 0.2 m, and $d = 0.15$ m for nets with a height larger than 0.2 m. The value for W_b was set to 0.0053 m/s, the median rise velocity of plastic particles found in Reisser et al. (2015). Finally, A_0 is given by

$$A_0 = 1.5u_{*w}\kappa H_s, \quad (\text{C.2})$$

where u_{*w} is estimated to be 0.0012 times the wind velocity at 10 meters (U_{10}) as given by Pugh (1987), with U_{10} as obtained from wave watch III data (Tolman et al., 2014); κ is the von Karman constant with a value of 0.4; and H_s is the significant wave height. The significant wave height was estimated from U_{10} by using the relation from Rossby & Montgomery (1935).

Here we use a tool from spatial statistics called the variogram, in order to assess the variance γ of plastic measurements separated by a given spatial lag distance h .

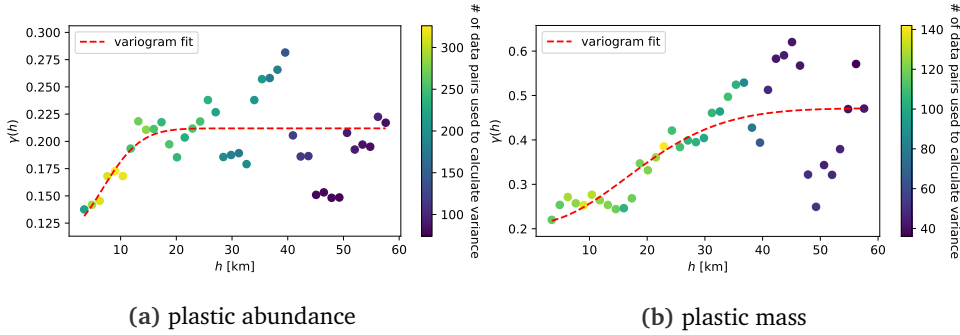


Figure C.1: Variograms constructed from the available plastic measurements

This can be used to estimate the amount of subgrid scale variance which can not be captured by the model. By assessing the variance of the plastic measurements below these scales, an indication can be obtained of the maximum achievable accuracy. The variogram is calculated only for samples which were taken within a day of each other, since this is the temporal resolution of the Lagrangian model output. Since sample information is not available at every location, an empirical variogram is constructed, considering the variance of measurements z separated by a certain distance $h \pm \delta$. The variance for the samples within this separation distance is then calculated using:

$$\gamma(h \pm \delta) = \frac{1}{2N(h \pm \delta)} \sum_{(i,j) \in N(h \pm \delta)} (z_i - z_j)^2, \quad (\text{C.3})$$

where $N(h \pm \delta)$ denotes the amount of samples in a given separation distance bin. The variogram can be calculated for both the plastic abundance and mass. Results are shown in Figure C.1. The samples for $\gamma(h \pm \delta)$ are colored by the amount of data point pairs used for the calculations in this figure. A Gaussian variogram was fitted through the points to show the trends more clearly. Several important observations can be made. For both types of measurements there seems to be a clear increase in the beginning, until about 20 to 40 kilometers. Afterwards this increasing trend seems to level off, although this is difficult to be certain about since there is a lack of data points for larger lag distances h , which also shows in the larger spread for $\gamma(h \pm \delta)$. What is also important to notice, is that the variance at a given lag distance is larger for the mass measurements compared to the abundance measurements. This means that more variance is expected when comparing the model to the measurements in terms of mass, and that it is more difficult to obtain information from the individual mass measurements compared to the abundance measurements for the inverse modelling process.

C.2 Beaching time scale estimate

In order to get a feeling for possible parameter values of τ_{beach} , a beaching analysis was done for a set of drifters in the Mediterranean (Menna et al., 2017). A total of

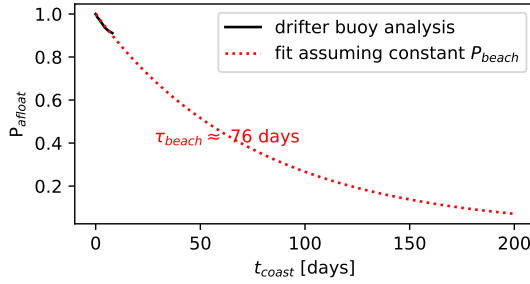


Figure C.2: Beaching time scale τ_{beach} estimated from drifter buoy data

1682 drifters were analysed, to see how long these drifters in general spend time near the coast before ending up on the beach. The mesh used for the currents has a spacing of $1/16^\circ$, corresponding to roughly 6.9 km for the Mediterranean. This distance was used to define whether a buoy is close to the coast. For each buoy the time it spend near the coast was tracked: the coastal time was binned, and for each interval the fraction was calculated of the buoys not having beached. A drifter was assumed to have beached, when at its end-of-life one of the four cells surrounding the drifter had a positive elevation. A 30 arc-second bathymetry dataset was used for this (Becker et al., 2009). Out of all buoys, 195 buoys were estimated to have beached at the end-of-life.

This analysis of drifters doesn't have to reflect the behaviour of plastic particles near the coast accurately. First of all, the particles have very different sizes, shapes, and properties. Some of the drifters are drogued, and will therefore be less sensitive to e.g. wave action and windage compared to floating plastic particles. Secondly, the beaching time scale τ_{beach} will implicitly have to take in account the recapturing of plastic particles by the water, i.e. it should be interpreted more as a time scale at which plastic particles remain permanently buried in the sand. The buoy trajectory dataset was designed to cut off when the buoy was estimated to be beached; this means that the time scale calculated from this data set should only account for one beaching event.

From the drifter buoy analysis, it followed that for each day spend near the coast, the drifter buoys have a probability of beaching of 1.2%. This translates into a time scale τ_{beach} of roughly 76 days, when assuming this probability remains constant when spending time near the coast. Since the drifter data set is sparse especially for longer coastal times, this is only a very first order estimate. For the inverse modelling step of the floating plastics, the prior for τ_{beach} was centred at 10^2 days which is close to the buoy estimate. The inverse modelling was done on \log_{10} of the values in order to span multiple orders of magnitude, with the bounds on the prior (99.7th percentile) ranging from 1 (10^1 days) to 3 (10^3 days).

C.3 Inverse modelling implementation

The most likely values for the parameters are estimated by solving a minimization problem. The cost function can be minimized by linearising the forward model around an estimate for the parameters \mathbf{m} , and iteratively updating the parameters using e.g. a gradient descent algorithm. The update steps were done by using a quasi-Newton method (see Tarantola (2005)):

$$\mathbf{m}_{n+1} = \mathbf{m}_n - \mu_n \left(\mathbf{G}_n^T \mathbf{C}_D^{-1} \mathbf{G}_n + \mathbf{C}_M^{-1} \right)^{-1} \left(\mathbf{G}_n^T \mathbf{C}_D^{-1} (\mathbf{d}_n - \mathbf{d}_{obs}) + \mathbf{C}_M^{-1} (\mathbf{m}_n - \mathbf{m}_{prior}) \right), \quad (\text{C.4})$$

where $\mu_n \leq 1$ defines the step-size of the update, and \mathbf{G} is a matrix containing the partial derivatives of the forward model with respect to the model parameters (i.e. the local linearisation of the problem): $(\mathbf{G}_n)_\alpha^i = (\partial g^i / \partial m^\alpha)_{m_n}$. This matrix was calculated using forward finite differences.

After finding the most likely posterior estimate using the quasi-Newton method presented above, it is also possible to estimate the posterior covariance matrix $\tilde{\mathbf{C}}_M$, using:

$$\tilde{\mathbf{C}}_M \simeq \mathbf{C}_M - \mathbf{C}_M \mathbf{G}^T (\mathbf{G} \mathbf{C}_M \mathbf{G}^T + \mathbf{C}_D)^{-1} \mathbf{G} \mathbf{C}_M. \quad (\text{C.5})$$

A second method was used for the inverse modelling step to verify the results from the least-squares solution. A Markov Chain Monte Carlo (MCMC) method was used to explore the parameter space which best matches the measurements and the given prior model parameters. The posterior of the model parameters is explored by using a random walk, where steps in the direction of likely parameter values have a higher probability of being accepted by making use of Metropolis-Hastings algorithm. This method does not rely on linearisation of the problem and is able to handle non-Gaussian assumptions on model parameters and measurement uncertainties. The MCMC method was implemented using the PyMC3 package in python (Salvatier et al., 2016). The same prior is used for the MCMC method as for the least-squares approach. Results comparing the MCMC method to the least-squares approach are presented in the next section, see Figure C.3.

In Table C.1 the different parameters tuned in the inverse modelling process are presented. The lower and upper values for the prior were defined as 3 times the standard deviation. An extra parameter not discussed in the text is the kernel density estimation bandwidth (KDE_{bw}). Lower and upper estimates for this parameters were set by hand, by seeing whether the predicted field for the plastic concentrations was not overly noisy or overly smoothed out. The parameter *river_{low-high}* corresponds to the parameter representing the riverine input from Lebreton et al. (2017) in terms of the lower, mid, and upper estimate as explained in the main text.

C.4 Parameter estimation sensitivity study

Figure C.3 presents the sensitivity study done for various tracer diffusivity values K , and different time thresholds at which particles are deleted. As explained in the

Table C.1: Parameters optimized in this paper, possible values, and lower/upper bounds used in the inverse modelling process. For parameters denoted by a power of 10, the inverse modelling process was done on the exponent, in order to be able to cover a wide range of scales. Values displayed using \pm indicate one times the standard deviation.

| Parameter | Prior, lower estimate | Prior, upper estimate | notes |
|------------------------|-----------------------|-----------------------|---|
| τ_{beach} [days] | 10^1 | 10^3 | See drifter buoy analysis |
| τ_{sink} [weeks] | $10^{0.30}$ | $10^{1.72}$ | Corresponds to 2–52 weeks |
| r_{sink} [days] | 3 | 15 | |
| $P_{sink,0}$ [-] | 0.17 | 0.44 | |
| $river_{low-high}$ [-] | -1 | 1 | Lower to higher riverine input estimate from Lebreton et al. (2017) |
| $S_{pop:riv}$ [-] | $10^{-1.3}$ | $10^{1.3}$ | Maximum contribution each source: 95% |
| $S_{fis:riv}$ [-] | $10^{-1.3}$ | $10^{1.3}$ | See above |
| KDE_{bw} [-] | 0.05 | 0.2 | |

main text, three different values for K were studied: $K = 1 \text{ m}^2/\text{s}$, $K = 10 \text{ m}^2/\text{s}$, and $K = 100 \text{ m}^2/\text{s}$, of which $K = 10 \text{ m}^2/\text{s}$ is likely the most appropriate value, given e.g. the relation used in Neumann et al. (2014):

$$K = K_0 \cdot (l/l_0)^{\frac{4}{3}}, \quad (\text{C.6})$$

where K_0 is set to $1 \text{ m}^2/\text{s}$, l is the local grid resolution ($1/16^\circ$, or approximately 7 km in our case), and l_0 is set to 1 km.

Most differences are observed when selecting a high tracer diffusivity of $K = 100 \text{ m}^2/\text{s}$. A higher diffusivity leads to less particles remaining close to the coastal zone in the model. The beaching time scale remains relatively similar, which leads to a reduction in the sinking time scale in order to keep the right mass balance in the basin. For all simulations all source importance ratios S are relatively equal, except from $S_{fis:riv}$ for $K = 100 \text{ m}^2/\text{s}$. There is an increase in the Kernel Density Estimate bandwidth (KDE_{bw}) when the tracer diffusivity increases. Possibly this could mean that a lower tracer diffusivity leads to a more realistic simulation, as results with a higher diffusivity are more smoothed out to reach a better agreement with the measured plastic concentrations.

Little differences are observed between the simulation where no particles are deleted, the simulation where particles are removed after 300 days, and where particles are removed after 180 days, all at $K = 10 \text{ m}^2/\text{s}$. In order for the mass balance to hold, it was calculated that after about 50 days, 99.9% of the mass input is gone from the surface water. Keeping track of all particle trajectories indefinitely is therefore not necessary, and induces a lot of computational costs. Instead, by deleting the particles after a certain amount of time, a lot more particles can be released for less computational effort. This is important in an inverse modelling study, as multiple model iterations are necessary to converge to the right parameter estimate.

Results from the Markov Chain Monte Carlo (MCMC) analysis are consistent with the results found using the least-squares approach. Since the parameter $river_{low-high}$ converged consistently to the lower bound for the least-squares analysis, it was fixed to -1. For the rest of the parameters, minimal differences are found for the most likely MCMC values compared to the least-squares analysis for $K = 10 \text{ m}^2/\text{s}$ and

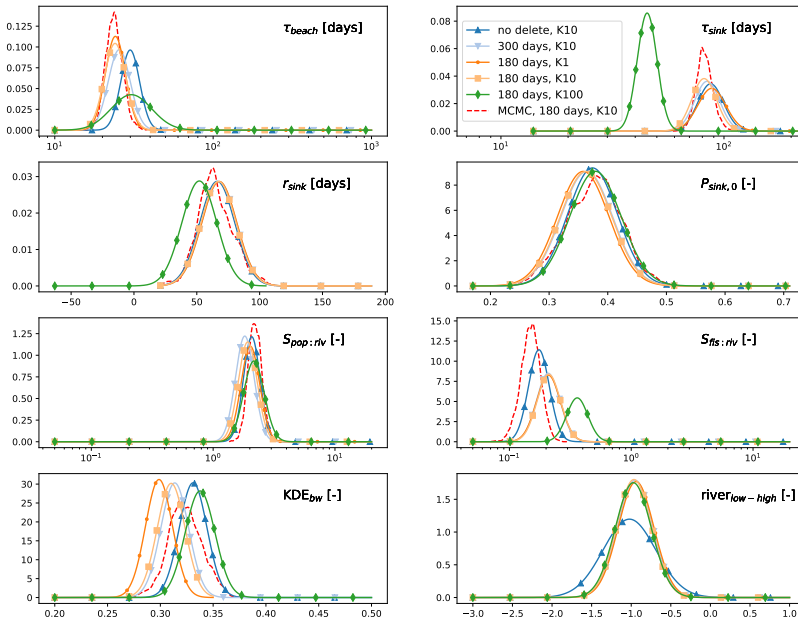


Figure C.3: Sensitivity study: days until deletion and K , and the MCMC analysis

removal at 180 days. Values in order of appearance in Figure C.3 (left to right, top to bottom) for the MCMC and least-squares analysis respectively are: 24 days (24 days), 80 days (81 days), 63 days (67 days), 0.38 (0.36), 2.2 (1.9), 0.2 (0.2), 0.32 (0.31).

A summary of the total input of plastics and where plastics end up, for different tracer diffusivity values is presented in Table C.2. As explained before, the higher tracer diffusivity leads to less particles remaining close to the coast, and hence less particles ending up on coastlines. Differences are quite minimal between $K = 1$ m²/s and $K = 10$ m²/s. The total input of plastics is relatively similar for all three tracer diffusivity values, as there is not much difference in the source importance ratios. Finally, the cost function value resulting from the inverse modelling process is presented. Although the mismatch is slightly lower for $K = 1$ m²/s, the difference is too little to prefer this tracer diffusivity value over the others from this result alone.

Figure C.4 presents the sensitivity study when adjusting the specified confidence interval of the parameter bounds, and adjusting the specified variation of the measurements (all for $K = 10$ m²/s, and deletion of particles after 180 days). The baseline simulation is presented, where the parameter bounds specify the 3σ confidence interval, and the variation of the measurements on a log₁₀ scale is specified

Table C.2: Sensitivity study tracer diffusivity, input and sinks presented for 2006–2016 in terms of percentages and metric tonnes

| Tracer diffusivity | Beaching | Sinking | Total input | Cost function value |
|--------------------------------|--------------------|--------------------|---------------|---------------------|
| $K = 1 \text{ m}^2/\text{s}$ | 56%, 14,300 tonnes | 43%, 11,000 tonnes | 25,500 tonnes | 1253 |
| $K = 10 \text{ m}^2/\text{s}$ | 54%, 13,800 tonnes | 45%, 11,500 tonnes | 25,600 tonnes | 1255 |
| $K = 100 \text{ m}^2/\text{s}$ | 33%, 9,400 tonnes | 67%, 19,100 tonnes | 28,400 tonnes | 1255 |

as $\gamma_n = 0.1376$ (measurements reported in terms of abundance), and $\gamma_m = 0.2201$ (measurements reported in terms of mass).

When changing the parameter bounds to the 2σ confidence interval, this means that less weight is given to deviations from the prior; when changing to the 4σ confidence interval more weight is given to deviations from the prior. Two sets of measurement variation levels were tested: one where $\gamma_n = \gamma_m = 0.32$, which comes from the reported coefficient of variation for abundance measurements from de Haan et al. (2019), and $\gamma_n = 0.32$ and $\gamma_m = 0.75$ as calculated from data reported in de Haan et al. (2019) for abundance and mass measurements separately. These measurement variation levels are larger than the one calculated here using the variogram method, which means that less weight is given to deviations of the model from measurements, hence it will have a similar effect as putting more weight on deviations from the prior (the C.I. = 4σ case). The main observed effect is that putting more weight to the prior will cause the Kernel Density Estimate bandwidth to be smaller. The initial sinking fraction $P_{\text{sink},0}$ gets higher for the C.I. = 2σ case as well. In turn, to keep the right mass balance, the beaching time scale seems to increase. Almost no changes are observed in the sinking time scale, and minimal variations are present for the source importance ratios and the expected riverine input.

In the main text it was mentioned that there is likely not enough information in the data to say something about the parameter r_{sink} : its posterior is almost the same as the prior. To verify this, we did the same analysis, but changed the prior of r_{sink} to a much higher value (16 weeks instead of 9 weeks). Results are presented in Figure C.5. As can be seen the posterior distribution of r_{sink} is still almost exactly the same as the prior distribution, further underlining our hypothesis there is no information in the data about this parameter.

C.5 Model-measurement scatter plots and mass correction factor

Scatter plots showing agreement between the model and measurements (the baseline case: $K = 10 \text{ m}^2/\text{s}$, deletion of particles after 180 days, $\gamma_n = 0.1376$, $\gamma_m = 0.2201$) are presented in Figure C.6, for abundance measurements (left) and mass measurements (right). As can be seen, there is quite a lot of variability, and the correlation is not very high.

As discussed in the main text, variability is high due to several reasons. First of all, there is natural subgrid-scale variability which can not be captured by the model. Secondly, variance in the measurements is likely further increased due to the fact that different measurement campaigns might have slightly different measurement

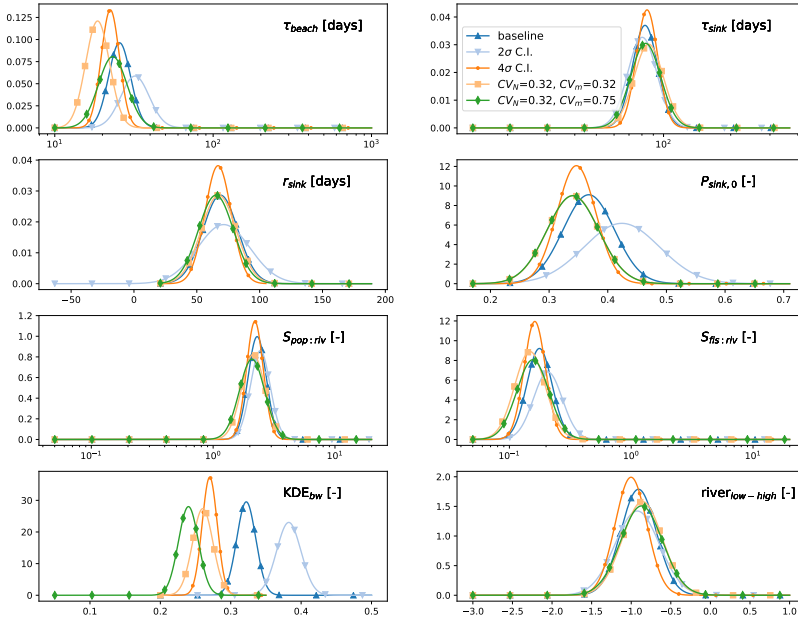


Figure C.4: Sensitivity study: specified confidence interval for the prior, and coefficient of variation specified for the measurements

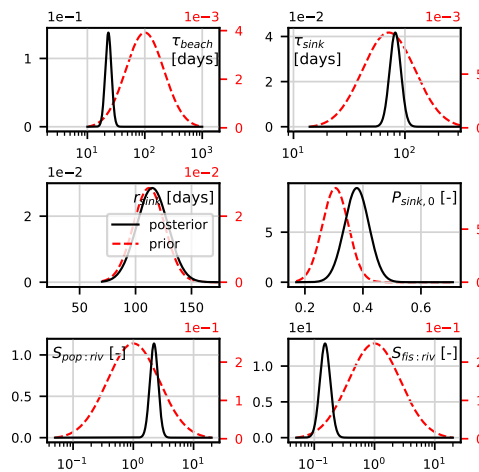


Figure C.5: After changing the prior of r_{sink} , the posterior of r_{sink} is still almost exactly the same as its prior after running the inverse modelling process

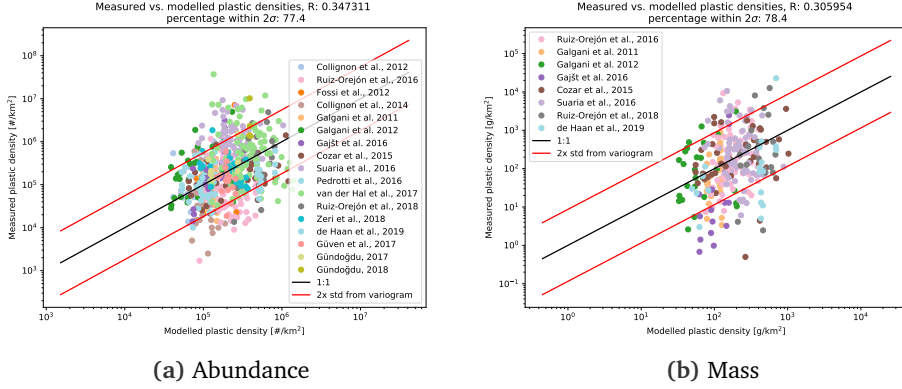


Figure C.6: Model versus measurements: $K = 10 \text{ m}^2/\text{s}$, deletion at 180 days, $\gamma_n = 0.1376$, $\gamma_m = 0.2201$

methodologies. The expected subgrid-scale variability was estimated in section S1 using the variogram method.

The maximum achievable correlation between the model and measurements is limited due to the high subgrid-scale variability. This is illustrated using the red solid lines in Figure C.6, which represent the model ± 2 times the standard deviation as estimated using the variogram. In case of a perfect model and accurate variability estimates, 95% of the data should lie between the red solid lines. In this case, this is 77.4% for the abundance measurements and 78.4% mass measurements. There is still room for improvement of the model, but even when 95% of the points would fall between the red lines, the maximum achievable correlation remains limited.

Even with high variability and low correlation, there is still information to be obtained from the observational data. In order to illustrate this, we varied the value of one parameter, τ_{sink} around its most likely posterior value (81 days) to observe what happens with the correlation and the cost function. Results are shown in Figure C.7. As can be seen, the cost function value clearly attains its minimum at the most likely posterior value. The maximum correlation values are not attained at exactly the same value for τ_{sink} , but the optimization algorithm converges to a value where both are relatively high. We see however that when moving away from the most likely posterior point, there is reduction in either the correlation for the abundance concentrations or the mass concentrations.

What is also important to note, is that we are not only interested in the correlation, but also the bias of the model with respect to the measurements. By keeping the bias at a minimum, we ensure that the mass balance in the basin is consistent with the measured mass concentrations. In Figure C.8 we illustrate this by choosing a too small value for τ_{sink} . Not only do we see a decreased correlation which was also illustrated in Figure C.7, we also see that the model generally underpredicts the mass (most of the points now lie above the 1:1 line), leading to an incorrect mass balance.

In future work, it is recommended to look for sets of measurement campaigns with

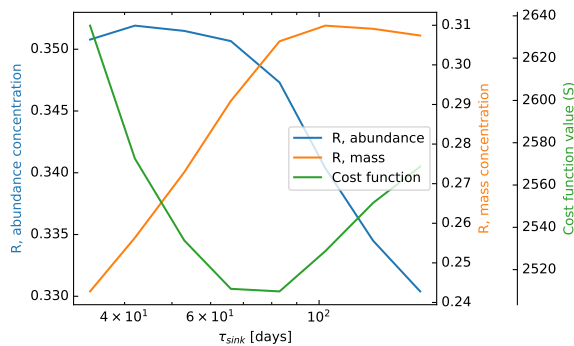


Figure C.7: Effect of varying τ_{sink} around the most likely posterior value on the correlation and cost function value

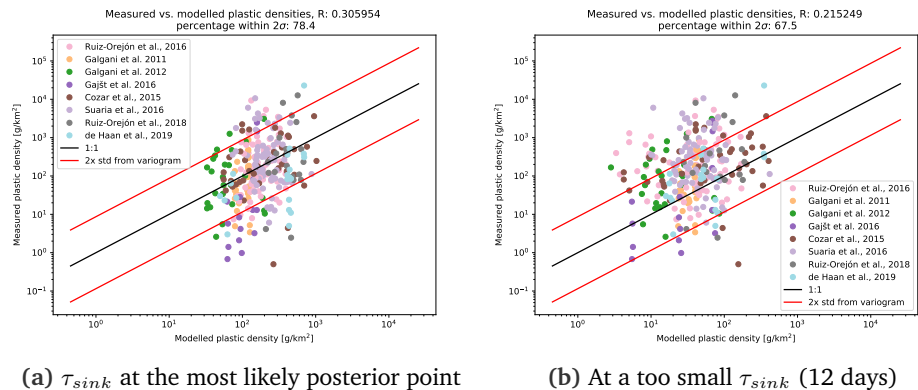


Figure C.8: Moving away from the most likely posterior point not only impacts the correlation between the model and measurements negatively, it also creates a bias in the mass balance

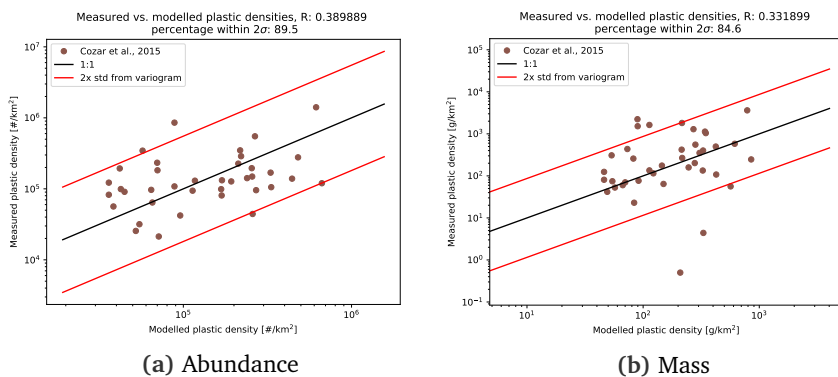


Figure C.9: Model versus measurements: $K = 10 \text{ m}^2/\text{s}$, deletion at 180 days, $\gamma_n = 0.1376$, $\gamma_m = 0.2201$ for measurements by Cózar et al. (2015) only

a larger amount of data points, which are taken in a consistent manner. To illustrate this, we plotted the same scatter plots for the measurement campaign by Cózar et al. (2015) only, which is the only dataset which spans both the western and the eastern Mediterranean basin. Although this is only one sample study, we do see in this case an increased correlation, and almost 90% of the data points falling within the expected 2σ bounds.

Figure C.6 also illustrates more clearly why there is an underestimation of the total floating mass in the model compared to previous estimates based on measurements only by Cózar et al. (2015). In the mass scatter plot, the model data points range from roughly 10^1 to 10^3 . The measurement data points have a higher range due to their high variability, ranging from roughly 10^0 to 10^4 . Since we are working on a log-log scale, the missing of the upper tails in the model distribution has a relatively high influence on the total estimated mass. A mass correction was applied, by introducing noise to the modelled plastic concentrations, where the noise is equal to the subgrid-scale variance calculated using the variogram for the mass measurements.

C.6 Which sinks are neglected and why?

We will illustrate why certain sinks are neglected in this work by considering a simple mass balance model, where the floating mass inside the basin is modelled over time using

$$\frac{dm(t)}{dt} = \text{Source} - \text{Sink}. \quad (\text{C.7})$$

We will consider two cases here: the *a priori* case, taking numbers from previous studies for the unknowns, and an *a posteriori* case, where we plug in our own estimates for the sources and sinks to see whether our assumptions are still valid, or whether they should be adjusted in future studies.

C.6.1 a priori

As *a priori* knowledge, we can use the numbers from Cózar et al. (2015) where the total amount of floating plastic in the Mediterranean was estimated to be approximately 1000-3000 tonnes. Furthermore, we can use data from previous modelling studies to get an estimate on how much plastic is thought to enter the basin *a priori*: >100,000 tonnes per year (Liubartseva et al., 2018; Jambeck et al., 2015). Since the estimated input is much larger than the amount which is floating, we assume we are at least in a quasi-equilibrium, and that the estimated floating mass is approximately the steady-state value. We assume that the magnitude of the sink is dependent on the amount of floating mass: each unit of plastic mass has a certain constant probability of being removed from the surface water per unit time, defined by the sinking rate p_{sink} , from which we write:

$$\frac{dm(t)}{dt} = \text{Source} - p_{sink} \cdot m(t) = 0. \quad (\text{C.8})$$

The value of p_{sink} can be related to a time scale τ by assuming a constant probability over time:

$$p_{sink} = 1 - e^{-1/\tau}. \quad (\text{C.9})$$

Solving Eq. (C.8) for the *a priori* estimated input and estimated floating mass, yields a sinking rate p_{sink} of about 0.1 to 0.3 day⁻¹, or a time scale τ of about 3 to 10 days. One question is whether we can neglect fragmentation in our model. In the main text it was already addressed that in Song et al. (2017) weathering experiments were done, from which we can get some feeling for possible time scales. We take the volume loss of 10% for polyethylene samples subjected to 12 months of UV radiation and 2 months of mechanical abrasion, which was estimated to correspond roughly to 4 years of weathering in the environment. We can again assume this volume loss is constant over time, defined by the fragmentation rate $p_{frag.}$ (or time scale $\tau_{frag.}$). This results in a $p_{frag.}$ of about $7 \cdot 10^{-5}$ day⁻¹. We can separate the fragmentation rate from the other sinks, which will be denoted by $p_{sink,other}$ and $\tau_{sink,other}$:

$$\frac{dm(t)}{dt} = \text{Source} - p_{sink,other} \cdot m(t) - p_{frag.} \cdot m(t) = 0. \quad (\text{C.10})$$

From this it is clear that $p_{sink,other}$ and $p_{frag.}$ should add up to p_{sink} in Eq. (C.8) in order to have a mass balance, meaning that fragmentation has no significant effect in terms of mass.

We can also analyse whether fragmentation could act as a significant source in terms of abundance, due to the production of secondary plastics. In Song et al. (2017) about 20 fragments per parent particle were observed for the same scenario as described above (roughly 4 years of weathering in the environment). We write a similar balance in terms of abundance:

$$\frac{dN(t)}{dt} = \text{Source} - p_{sink,other} \cdot N(t) + p_{frag.} \cdot N(t) = 0. \quad (\text{C.11})$$

We assume that the sinks have a similar influence in terms of mass as in terms of abundance: every single unit of mass, and every single plastic particle, has the same

rate of being removed from the system (i.e. p_{sink} is still about 0.1 to 0.3 day⁻¹). In this case, 20 fragments over 4 years would translate to a $p_{frag.}$ of about $2 \cdot 10^{-3}$ day⁻¹. This is significantly more than the influence in terms of mass, but still two orders of magnitude lower than the magnitude of $p_{sink,other}$ that we expect.

Finally, UV radiation might convert plastic polymers into carbon dioxide and dissolved organic carbon by itself: in Ward et al. (2019) the average lifetime until partial oxidation of polystyrene was estimated to range between 10 to 50 years. Assuming this process is constant over time, would lead to a UV degradation rate p_{UV} of about $3.8 \cdot 10^{-5}$ to $1.9 \cdot 10^{-4}$ day⁻¹. This is about three orders of magnitude lower than the p_{sink} required for the mass balance, and therefore neglected.

C.6.2 a posteriori

We can do the same analysis with the values from the mass balance obtained from the inverse modelling study. We estimated a total input of about 2,100–3,400 tonnes, and a total floating mass of 170–420 tonnes. Using these numbers, we get a required sink magnitude ranging from $p_{sink} = 1.3 \cdot 10^{-2}$ to $p_{sink} = 5.5 \cdot 10^{-2}$ day⁻¹. Using the same reasoning as above, this would mean that in terms of mass, the fragmentation rate is still expected to be significantly lower ($7 \cdot 10^{-5}$ day⁻¹, so at least two orders of magnitude) than the sinks required for the mass balance. In terms of time scales, we have τ_{sink} ranging from approximately 17 to 72 days. Removing the influence of fragmentation from the total sink would result in τ_{sink} increasing by at most 0.5%. On a similar note, considering partial oxidation of polymers (which was calculated to be $1.9 \cdot 10^{-4}$ day⁻¹ for the decadal time scale reported in Ward et al. (2019)), could lead to an increase in τ_{sink} of about 1.4%. Although these effects are much more important now than for the *a priori* case, it will not change the inverse modelling results dramatically.

In terms of abundance, fragmentation starts to play a bigger role, with a rate which is approximately five times lower than the lowest estimated sink magnitude ($p_{frag.} = 2 \cdot 10^{-3}$ versus $p_{sink} = 1.3 \cdot 10^{-2}$).

For a first order estimate we think that leaving out fragmentation is reasonable. Its influence was estimated to be mainly in terms of an increase in abundance. However, fragmentation will act as a net source (in terms of abundance) of secondary plastics only at the beginning of the fragmentation process. After a while a lot of the fragments will become too small to measure (i.e. below the neuston net mesh size), leading to a net loss of particles. What is also seen in previous studies (Onink & Laufkötter, 2020; Liubartseva et al., 2018), is that beaching patterns in generally correspond quite well with the pattern of the source itself (i.e. a majority of plastics will not travel far from its original source). This means it is likely the pattern of secondary plastic sources is quite similar to the primary sources. This likely diminishes the influence of fragmentation on the patterns of plastic concentrations in the water as modelled here. The combined effect of fragmentation as a sink for mass and a source in terms of abundance is very interesting, and is recommended to be taken in account in future studies, also because it is clear from the *a posteriori* calculations it is more important than one would think from the numbers given for the *a priori* case.

We can also turn the line of reasoning around: perhaps the discrepancy between

our estimated input and the one from literature is due to neglecting fragmentation and degradation. We can take the estimated value for $p_{frag.}$ ($7 \cdot 10^{-5} \text{ day}^{-1}$) and the most conservative (largest) value for p_{UV} ($1.9 \cdot 10^{-4} \text{ day}^{-1}$), keep $p_{sink, other}$ at our *a posteriori* range ($1.3 \cdot 10^{-2}$ to $5.5 \cdot 10^{-2} \text{ day}^{-1}$), and estimate the source strength which would result in a steady-state floating mass of 170–420 tonnes. In the extreme case of taking the smallest $p_{sink, other}$ ($1.3 \cdot 10^{-2}$) and the maximum floating mass (420 tonnes) we would get a required input in the system of 8,500 tonnes. This is much higher than our estimated input (2,100 - 3,400), but one has to keep in mind that here we combined our very highest floating mass (420 tonnes), with the $p_{sink, other}$ corresponding to the very lowest input mass (2,100). These extremes correspond to the upper and lower tail of the 95% confidence interval: the probability of both occurring would be less than 0.05%. Still this increased number does not explain the difference with respect to previous studies (Liubartseva et al., 2018), where the input was estimated to be $>100,000$ tonnes per year.

Supporting material Chapter 5

D.1 Observational data and model performance

D.1.1 Data overview

An overview of the data used to calibrate the numerical model is presented in Figure D.1, sources and details of the data are given in Table D.1.

A large amount of net trawl data come from previous studies that aggregated observational data for numerical model calibration, both globally (van Sebille et al., 2015) and for the Mediterranean Sea (see Chapter 4). To account for the fact that neuston net measurements can be biased due to wind mixing, the Kukulka correction factor (Kukulka et al., 2012) is applied using ERA5 reanalysis wind speeds (Hersbach et al., 2020). Data for large plastic items at the ocean surface (>0.2 m) come from previous studies using visual observations (Eriksen et al., 2014; de Vries et al., 2021).

For beach measurements, MDMAP (Burgess et al., 2021) and OSPAR (OSPAR, 2010) observations are included in terms of n m^{-1} , focusing on larger plastic items (>2.5 cm and >50 cm respectively). These datasets contain information on the type of litter items, which is used to define a minimum percentage of fishing related plastics per measurement location. These values can then be compared to the predicted model fishing item percentages. Plastic litter types containing words such as ‘fish’, ‘rope’, or ‘net’ are categorized as fishing related items. We gathered additional beach measurements from the literature in terms of g m^{-1} , using the criterion that the entire beach width was sampled for comparability between samples.

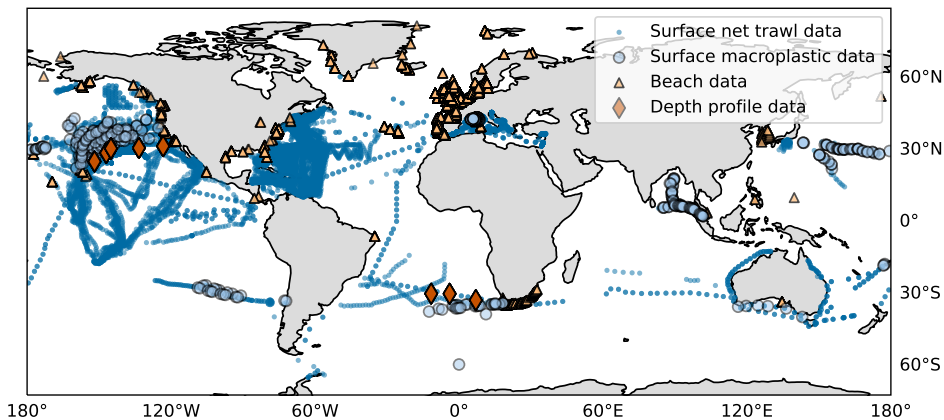


Figure D.1: Observational data across different marine reservoirs used to assimilate into the numerical model. The markers show the locations of the available observational data, from neuston net measurements at the ocean surface (blue dots), visual observations of large plastic items (>0.2m, blue circles), the deep ocean (orange diamonds), and beaches (orange triangles).

Table D.1: Overview of the observational data used to calibrate the numerical model

| Reference | Measurement type | Notes |
|---|------------------------------------|--|
| van Sebille et al., 2015 | Net trawls, surface ocean | Compilation of 37 measurement campaigns for small plastic particles (>0.335 mm) in terms of mass and number, 11,854 data in total |
| Kaandorp et al., 2020 | Net trawls, surface ocean | Compilation of 18 measurement campaigns for small plastic particles in the Mediterranean Sea (>0.335 mm) in terms of mass and number, 1170 data in total |
| Egger et al., 2020b,a | Net trawls, surface ocean | Data on 4 size classes available in terms of mass and number, 1390 data in total |
| Egger et al., 2020b; Zhao et al., 2022 | Net trawls, deep ocean | Data on 4 size classes available in terms of mass and number, 120 data in total |
| Eriksen et al., 2014 | Visual observations, surface ocean | Visual detection for large plastics (>0.2 m) in terms of number, 553 data in total |
| de Vries et al., 2021 | Visual observations, surface ocean | Automatic floating object detection for large plastics (>0.5 m) in terms of number, 10 data in total |
| OSPAR, 2010 | Beach measurements | OSPAR data (>0.5 m) in terms of number, 1692 data in total |
| Burgess et al., 2021 | Beach measurements | MDMAP data (>25 mm) in terms of number, 5022 data in total |
| Hong et al., 2014; Kaandorp et al., 2022; Edyvane et al., 2004; Lee & Sanders, 2015 | Beach measurements | Additional beach measurements in terms of mass, 238 data in total |
| Ryan et al., 2018 | Beach measurements | Data on small plastics (2–25 mm) in terms of mass and number, 162 data in total |

D.1.2 Model performance

We compare measured plastic concentrations to the assimilated model results (Figure D.2), both in terms of number of particles (upper row), and mass (lower row). The red dashed lines indicate the expected measurement error (95% confidence intervals). The Pearson R correlation coefficients between the model and the various datasets are presented in Table D.2.

The left column compares the model output to measurements from the Pacific Ocean (Egger et al., 2020a,b), where different particle size classes were measured. When considering all particle sizes together (0.5–50 mm) there is a good fit ($R = 0.85$) for the number concentrations. This indicates that our fragmentation model is able to capture the progression from large to small particle sizes accurately. The somewhat poorer performance ($R = 0.59$) for the mass concentrations is likely due to varying shapes and densities of plastic particles introducing increased variability into the dataset, as was already noted in earlier studies (de Haan et al., 2019; Kaandorp et al., 2020). When looking at the size classes separately we see a lower performance for larger particle sizes ($R = 0.50$ for 15–50 mm). One explanation is the fact that larger particles tend to be sparse, introducing variability between tows capturing only a few or no particles (neuston net measurements typically show a cut-off around $1 \cdot 10^{-4} \text{ n m}^{-3}$ caused by the lower detection limit of the method, see Figure D.2b). Measurements in the 15–50 mm category span about three orders of magnitude, as opposed to about four orders of magnitude for measurements in the 0.5–1.5 mm category, which makes it more difficult to achieve the same level of correlation with the amount of variability present in the datasets. Another explanation would be a potential different influence by open ocean Stokes drift (Egger et al., 2020a) which is not captured by our model.

The good correlation ($R = 0.80$; second column) between the model results and the plastic concentrations in the global ocean from neuston net measurements shows that the model is able to capture the magnitude of plastic concentrations globally. Again a lower correlation is observed for larger plastic items ($R=0.41$ for the $>0.2 \text{ m}$ category), which can be explained due to their relative sparsity (the measurements span fewer orders of magnitude). Furthermore additional windage effects might play a role for the largest plastic items that are not completely submerged, which is not captured by our model.

The concentrations in the deep global ocean, where only few neuston net measurements are available, are lower in our model than in the observations. Measurement detection limits are a likely cause for the difference. A large fraction of the measurements that we use here (Egger et al., 2020b; Zhao et al., 2022) did not capture any plastic fragments during the trawl ($>60\%$, depending on the measured size class). The study providing most deep ocean measurements (Egger et al., 2020b) already noted that the absence of detectable plastic fragments induces large uncertainties in measured plastic concentrations.

Finally, the relative number and mass for different particle sizes on beaches is captured well by the model ($R=0.75$ and $R=0.87$ for all particle sizes combined; fourth column), where we use measurements for three size classes: $>0.5 \text{ m}$ (OSPAR, 2010), $>25 \text{ mm}$ (Burgess et al., 2021), and $2\text{--}25 \text{ mm}$ (Ryan et al., 2018). When looking at

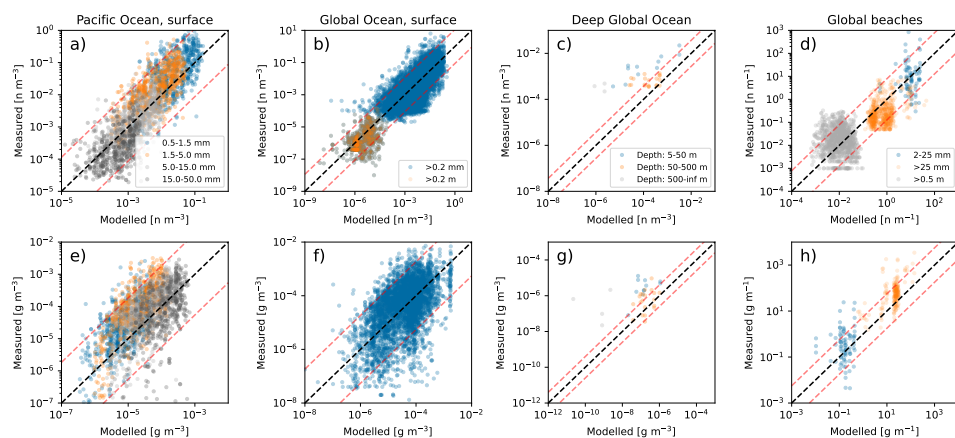


Figure D.2: Scatter plots for modelled versus measured plastic concentrations. First column: measurements for separate size classes at the ocean surface (Egger et al., 2020a,b). Second column: measurements for small plastics (>0.2 mm) and large plastics (>0.2 m) at the ocean surface. Third column: measurements below 5 m depth. Fourth column: measurements from beaches. The upper row presents number concentrations (n m^{-3} in the ocean and n m^{-1} on beaches), the lower row presents mass concentrations (g m^{-3} in the ocean and g m^{-1} on beaches). The 1:1 line is shown using the dashed black line, the expected amount of measurement error ($\pm 2\sigma$) is shown using the dashed red lines. The Pearson R correlation coefficients between the modelled and measured concentrations is shown in Table D.2.

the separate size classes we find no correlation for the 2–25 mm category. This can be explained due to the fact that these measurements were taken only around one location (South Africa), whereas our model is only able to differentiate over larger spatial (global) scales. The same holds for the >0.5 m category, which were obtained over a relatively limited area. Additionally, we discussed in the main text that number concentrations are highly dependent on the lower size limit that is used. This likely adds to the fact that there is no correlation in the >0.5 m category data, which are largely obtained using volunteers (OSPAR, 2010). Nonetheless, these data are valuable to constrain the order of magnitude of large plastic items on beaches. We do see positive correlations for the >25 mm category: $R = 0.35$ and $R = 0.56$ for number and mass concentrations, respectively. This further underlines our recommendation in the main text of measuring the mass of plastic items on beaches when performing visual observational studies.

D.2 Supplementary results

D.2.1 Estimated amount of fishing related plastics

In Figure D.3 we present the percentage of large fishing related plastic items (>2.5 cm) in terms of numbers as predicted by the calibrated model. The calculated percentage of fishing related items is relatively high in sparsely populated areas near the polar regions, where little plastic is expected to enter the ocean from coastlines or rivers. Furthermore we expect relatively high quantities of fishing related items in

Table D.2: Pearson R correlation coefficients between the modelled and observed plastic concentrations

| Domain | Details | Pearson R |
|------------------------|------------------------------------|-----------|
| Pacific Ocean, surface | n m^{-3} , all sizes | 0.85 |
| | n m^{-3} , 0.5–1.5 mm | 0.64 |
| | n m^{-3} , 1.5–5.0 mm | 0.70 |
| | n m^{-3} , 5.0–15.0 mm | 0.62 |
| | n m^{-3} , 0.5–50.0 mm | 0.47 |
| | g m^{-3} , all sizes | 0.59 |
| | g m^{-3} , 0.5–1.5 mm | 0.64 |
| | g m^{-3} , 1.5–5.0 mm | 0.71 |
| | g m^{-3} , 5.0–15.0 mm | 0.62 |
| | g m^{-3} , 15.0–50.0 mm | 0.50 |
| Global Ocean, surface | n m^{-3} , all sizes | 0.80 |
| | n m^{-3} , >0.2 mm | 0.80 |
| | n m^{-3} , >0.2 m | 0.41 |
| | g m^{-3} , >0.2 mm | 0.62 |
| Deep Global Ocean | n m^{-3} , all depths | 0.47 |
| | n m^{-3} , 5–50 m depth | 0.88 |
| | n m^{-3} , 50–500 m depth | 0.19 |
| | g m^{-3} , all depths | 0.33 |
| | g m^{-3} , 5–50 m depth | 0.53 |
| | g m^{-3} , 50–500 m depth | 0.37 |
| Global beaches | n m^{-1} , all sizes | 0.75 |
| | n m^{-1} , 2–25 mm | 0.09 |
| | n m^{-1} , >25 mm | 0.35 |
| | n m^{-1} , >0.5 m | -0.06 |
| | g m^{-1} , all sizes | 0.87 |
| | g m^{-1} , 2–25 mm | 0.07 |
| | g m^{-1} , >25 mm | 0.56 |

the South Pacific. The high amount of large fishing related items in the open ocean is in agreement with overview studies indicating that the majority of plastic litter in the open ocean originates from the ocean (Morales-Caselles et al., 2021).

In Figure D.4 we present the same percentage for small plastics (<2.5 cm) only. We see an increase in the non-fishing related plastic percentage for smaller particle sizes. This can be explained due to the input locations of the different sources. Plastic input from rivers and coastlines tends to spend more time around coastlines and is therefore expected to undergo more fragmentation compared to plastics introduced from fishing activity further away from the coastlines.

From the observational data on beaches (OSPAR, 2010; Burgess et al., 2021) we calculate the fraction of fishing related items in terms of numbers. We compare these fractions to the model output in Figure D.5. Not all items can be accurately categorized as fishing related items (e.g. fragments). The observed fractions of fishing related items are therefore treated as minimum values: points in the lower right half of the plot are plausible given the observational data. Most model results fall within the plausible value ranges based on the observational data. Some interesting differences can be observed in the fishing related item percentages from European data (OSPAR, 2010) versus data mainly coming from the United States (Burgess et al., 2021). Higher fishing related item percentages are modelled and observed in Europe, which matches the expectation of high fishing related item input in this region, see Figure 5.1 in the main text.

D.2.2 Environmental removal

In the main text we showed high persistence of plastic waste in the marine environment: given a sudden stop of new plastic items introduced into the marine environment, we expect that only 9% of the plastic mass is lost within two years. The

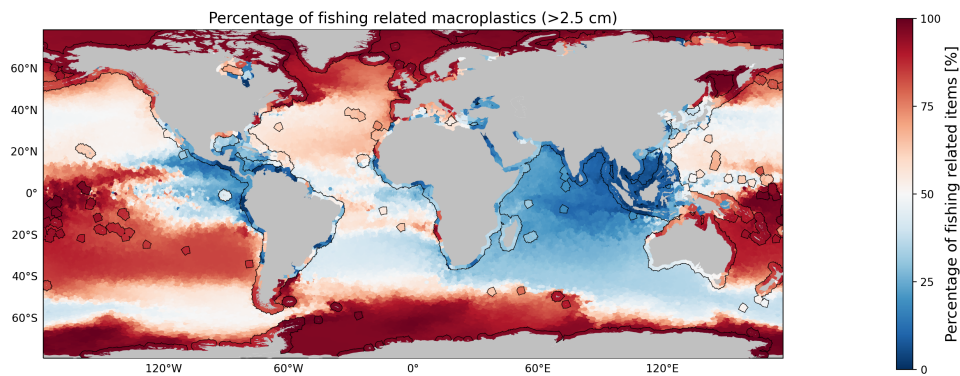


Figure D.3: Estimated global percentage of large fishing related plastics (> 2.5 cm, in terms of number of items)

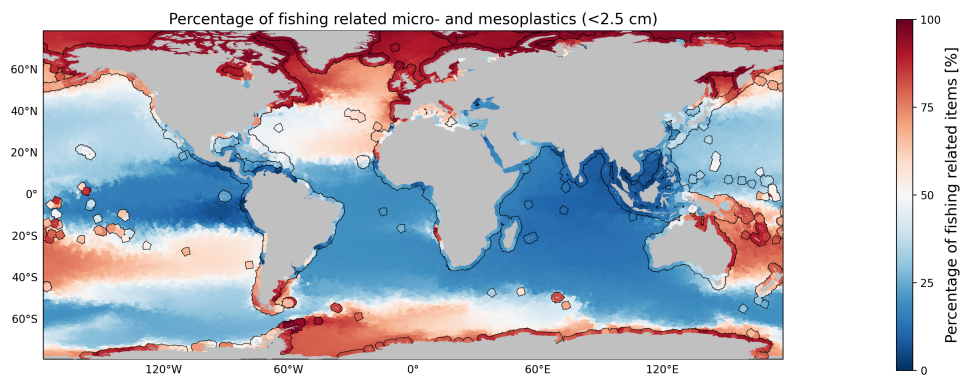


Figure D.4: Estimated global percentage of small fishing related plastics (<2.5 cm, in terms of number of items)

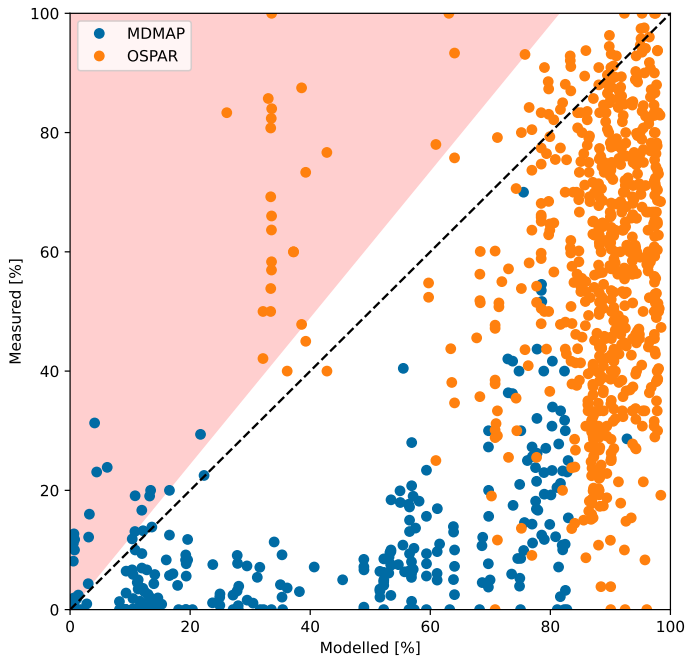


Figure D.5: Modelled versus measured percentage of fishing related items on beaches. Measurement data come from OSPAR (OSPAR, 2010) and MDMAP (Burgess et al., 2021) beach cleanup data. Since not all items can be accurately categorized as fishing related items (e.g. fragments), the observed percentage of fishing related items is used as a lower bound. Shaded in red are the outliers, where more fishing related items are measured than predicted by the model. Since the variability of number concentrations is defined on a log scale, this area deviates from the 1:1 line for higher concentrations.

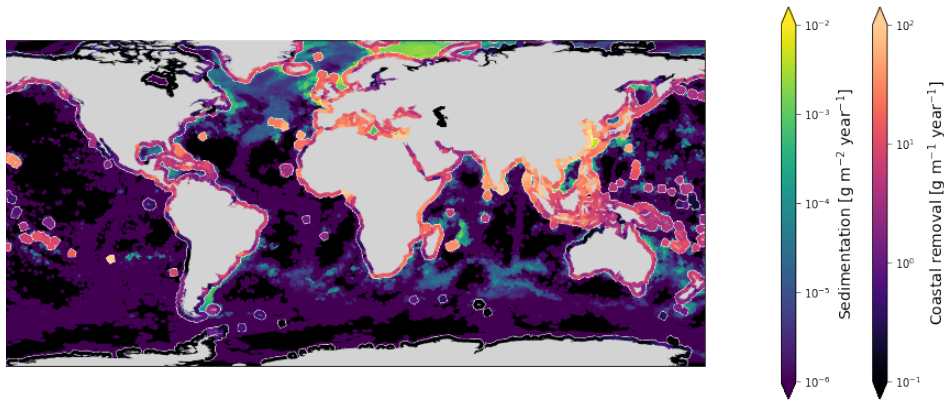


Figure D.6: Estimated sedimentation fluxes and coastal removal of marine plastics

reason for this slow removal can be seen in Figure D.6. Most of the sedimentation and coastal removal of plastic mass is expected near coastlines. We know, however, that a lot of plastic mass eventually ends up in the subtropical gyres (Wichmann et al., 2019b). Removal in these areas is relatively low. We therefore see a substantial decrease in the total marine mass during the first couple of years after the input stop in Figure 5.5 of the main text. As the plastic pollution reaches waters further offshore the sedimentation and coastal removal decreases substantially.

D.2.3 Prior and posterior parameter estimates

In Chapter 5 we bring together different models and data sources that have been developed in the past. In total 16 parameters are used to describe the sources, transport, and sinks of marine plastics. We summarize the different scenarios that are modelled and their corresponding parameters in Table D.3, and provide references for further detailed explanations.

The prior probability distribution functions for the model parameters, and the (updated) posterior estimates are presented in Figure D.7. Similar to the model study for the Mediterranean Sea (see Chapter 4), we calculate a lower input from rivers compared to previous studies (Lebreton et al., 2017; Meijer et al., 2021). We estimate roughly 3 times the input from coastlines compared to the input from rivers, which is close to the estimated 2 times for the Mediterranean Sea (see Chapter 4). However, much higher fishing related quantities are estimated (about 3 times as much as the riverine input, versus 0.2 times for the Mediterranean Sea). These differences can be due to the specific region of interest, and can be caused by better representation of bigger item sizes in this study. Figure D.3 already showed that a lot of the bigger plastic items are expected to originate from fishing activity, but also that the fraction of fishing related items can vary substantially depending on the region. The beaching time scale τ_{beach} is estimated to be 74 days. This matches well with the Mediterranean Sea study, where it was estimated to be around 25 days for plastic particles, and 75 days for drifter buoys. A higher value for plastic particles (74

Table D.3: Summary table of the model parameters

| Scenario | Parameters | Section description | Limitations |
|-----------|--|---------------------|---|
| Sources | S_{riv} , S_{pop} , S_{fis} , $\mu_{release,log10}$, GR_{in} | 5.2.3, D.3.2 | Seasonal variability is only available for the fishing activity. A single exponential growth rate of plastic input is estimated, which might vary over time under influence of mitigation strategies |
| Beaching | τ_{beach} , $l_{beach,min}$, $p_{removal}$ | 5.2.3, D.3.1 | Beaching is modelled using a global time scale, whereas this might vary spatially depending on various factors, e.g. coastline orientation/geometry, tides (Kaandorp et al., 2022) |
| Sinking | f_{of} , f_{nb} , f_{pf} | 5.2.3 | Biofouling is assumed to be the main driver of positively buoyant particles becoming negatively buoyant, biofilm growth is modelled assuming spherical particles using algal concentration and primary productivity fields (Fischer et al., 2022) |
| Mixing | a_{Leff} , a_{Leff} | 5.2.2, D.3.4 | Vertical mixing is modelled using a random walk model (Onink et al., 2022a) with baseline rise velocities estimated using spherical particles |
| Fragment. | λ_f , p_f , d_N | 5.2.3, D.3.3 | Fragmentation parameters are calibrated for marine plastics in general, while they might vary for different polymers (Kaandorp et al., 2021) |

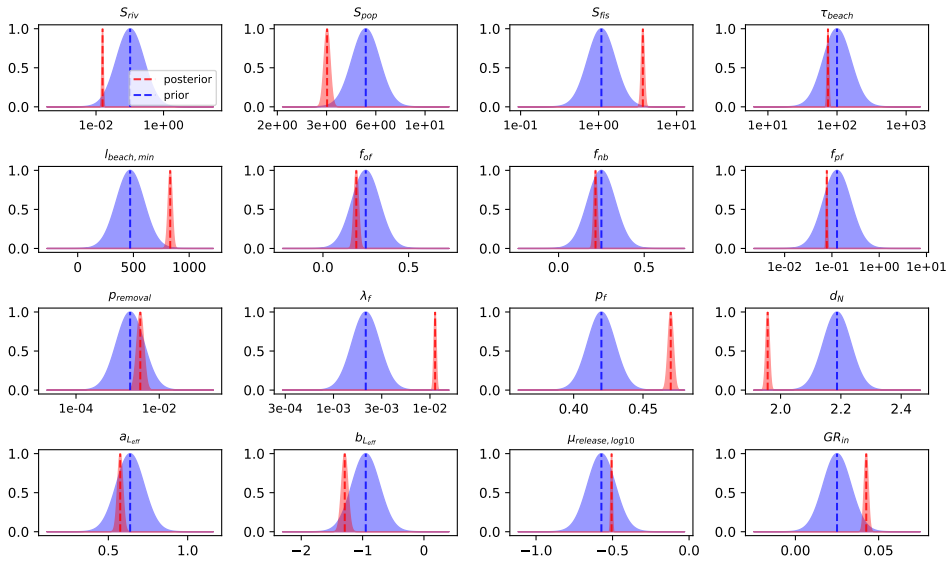


Figure D.7: Given prior probability density functions (blue), and the updated posterior probability density functions (red) for the model parameters.

versus 25 days) can be expected here, since the grid size used in the global model is much coarser (approximately 60 kilometers versus 7 kilometers). As touched upon in the main text, biofouling is expected to have a substantial influence on the particle transport when looking at the calculated fouling fractions. Fragmentation is calculated to happen more quickly than previously expected (Kaandorp et al., 2021). This can be due to the fact that environmental plastics are residing in the environment much longer than what experimental weathering experiments allow for (Song et al., 2017), on which these previous fragmentation rates were based. Longer residence times in the environment can lead to more oxidation and embrittlement, possibly enhancing the fragmentation rate. The estimated object dimension d_N of around 2 suggests that most environmental plastics are roughly 2-dimensional: items with a negligible thickness compared to their length and width. Mixing parameters ($a_{L_{eff}}$, $b_{L_{eff}}$, see Section D.3.4) are close to their prior values, indicating that the baseline estimates for the amount of mixing that positively buoyant particles experience match the observational data. Finally, the estimated exponential growth rate of 6% per year is close to the estimated global plastic waste production of 4% per year in (Geyer et al., 2017), indicating that the amount of plastic waste entering the global ocean is likely proportional to the total global plastic waste production.

D.2.4 Plastic input growth rate

We estimate that the amount of plastic waste entering the ocean has increased by about 4% per year. Establishing temporal trends in the amount of plastic pollution in the marine environment is difficult, due to the high variance in the observations. Previous studies (Galgani et al., 2021) have found no conclusive temporal trends

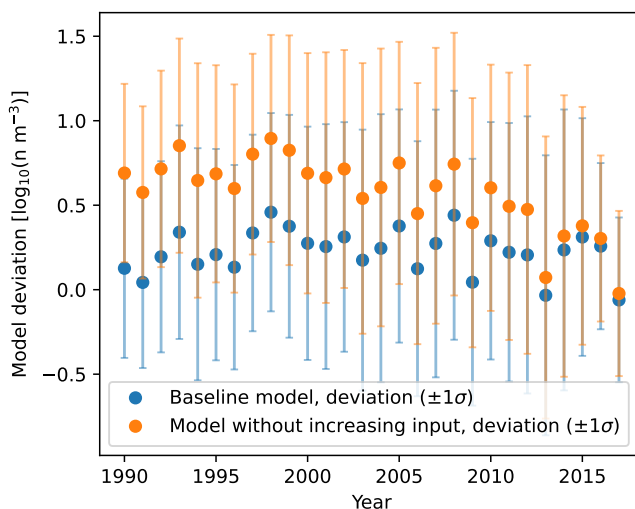


Figure D.8: Model deviation per year for the baseline model where the plastic input increases by 4% per year (in blue), and a model run where the amount of plastic entering the ocean does not increase over time (in orange)

regarding the amount of litter in the marine environment.

In Figure D.8 we plot the deviation between the modelled and observed plastic concentrations per year in blue. We take neuston net measurements in terms of number of particles, for which the the largest and longest record is available. The mean model deviation is relatively steady over time. There is a high variance in the model deviation (which could also be seen in Figure D.2). This is expected, as studies taking replicate plastic measurements showed coefficients of variation up to 80% (de Haan et al., 2019). In orange we plot the model deviation without the 4% increase in plastic input per year. As can be seen, the results are relatively similar for the most recent years. Without the increase in plastic input over time, the mean model deviation increases for earlier years, however. This suggests that the amount of plastic entering the marine environment has increased over the past decades as predicted by the model.

D.3 Supplementary methodological information

D.3.1 Transition Matrix model schematic

A schematic representation of the transition matrix model is shown in Figure D.9, using for simplicity an ocean model with 8 grid cells in total. Of these grid cells, 2 cells represent the coastlines, and 6 cells represent different depth layers in the ocean.

First, we use OceanParcels (Delandmeter & van Sebille, 2019) to resolve transport of virtual plastic particles in the ocean as explained in the main text. This transport is

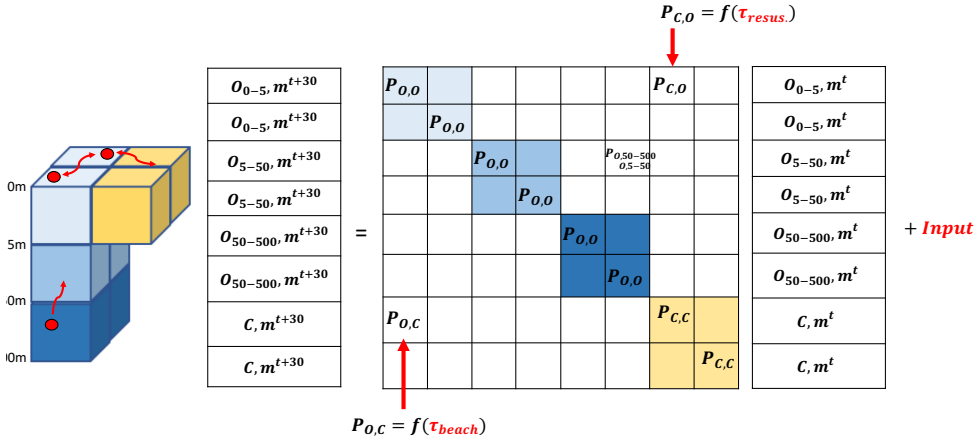


Figure D.9: Schematic illustration of the transition matrix model setup, with O and C denoting ocean and coastline (beach) grid cells respectively; $P_{C,O}$ and $P_{O,C}$ representing the transition probabilities from the coastline to the ocean and vice versa; $P_{O,O}$ and $P_{C,C}$ representing the transition probabilities that particles remain within the ocean or coastline.

then captured using a transition matrix (with size 6×6 in the case of 6 ocean cells). The transition matrix defines the probability that particles move from one grid cell of the ocean to another grid cell. The columns of the transition matrix sum up to 1 since they define probabilities. In the example, the two elements of the coastline are represented by 2 cells as well, resulting in a 8×8 transport matrix. Elements in the transition matrix around the diagonal (indicated by the different shadings) represent transport within the same depth layer. Transport from the surface ocean towards the coast is captured by the lower left elements in the matrix (beaching), transport from the coast to the ocean in the upper right (resuspension). These elements are directly calculated from the given beaching and resuspension timescales, which are parameters in the model. The columns are renormalized to 1 afterwards to have a total transport probability of 1. Each time step, new plastics are introduced to the system, determined by the three source parameters ($S_{pop.}$, $S_{fis.}$, $S_{riv.}$). This transition matrix is calculated for six different particle sizes, and different months as explained in the main text.

To further clarify the beaching and resuspension model, Figure D.10 is shown. Plastic particles in the coastal grid cell have a probability of beaching, determined by the beaching time scale $\tau_{beach.}$ The fraction of tracer beaching ($p_{beach.}$) is calculated using $p_{beach.} = 1 - e^{-\frac{30}{\tau_{beach.}}}$ given a transition matrix time step of 30 days. The same is done for the resuspension of tracer back into the ocean, given the resuspension time scale $\tau_{resus.}$. The tracer concentration on the beach is calculated by dividing the amount of tracer by the length of the coastline within the grid cell, shaded in red. This length is furthermore corrected for the fractal structure of the coastline

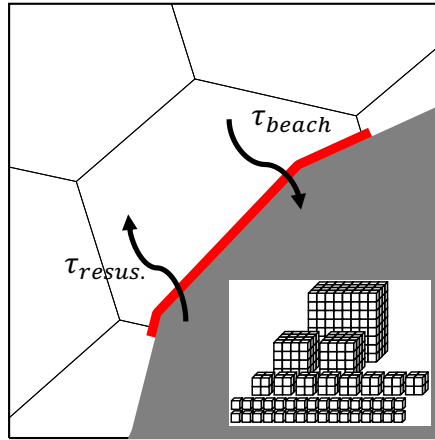


Figure D.10: Schematic illustration of the beaching parameterisation.

(Isobe & Iwasaki, 2022) as explained in the main text. Finally, plastic particles move from one size class to smaller size classes under the influence of fragmentation (Kaandorp et al., 2021). Each size class has a different transition matrix system as explained above.

D.3.2 Input Mass Size Distribution

The input size distribution is parameterised using a Gaussian distribution on the \log_{10} of the particle size. Data from municipal solid waste sorting facilities (Jansen et al., 2015) are used to estimate typical particle sizes of plastic waste. Furthermore data from river studies (van Emmerik et al., 2019; Vriend et al., 2020) are used to estimate which particle sizes contribute to most of the plastic mass entering the ocean from rivers. These studies found dominant object sizes of roughly 0.3 meters (Vriend et al., 2020) and 5–50 centimeters (van Emmerik et al., 2019). The available data are plotted in Figure D.11. Data from municipal solid waste sorting facilities might be biased: plastic waste is deformed as it is transported in larger (compressed) bales. We set the upper bound for the mean of the input size distribution to a conservative 0.5 meters, capturing both the size distribution from the municipal solid waste sorting facilities, and the upper bound for the dominant particle size from rivers (van Emmerik et al., 2019).

D.3.3 Particle size mass ratio

In order to couple measurements in terms of number of particles and in terms of mass, we need a relation between particle size and mass. Data on individual plastic particle sizes and masses were obtained from measurements in the North Atlantic (Lenz, 2020), see the blue dots in Figure D.12. In order to extrapolate these data, a powerlaw relation is assumed for the particle mass, given that it depends on the effective spatial dimension d_N of the plastic objects: $m \sim l^2$ for flat objects (the dash-

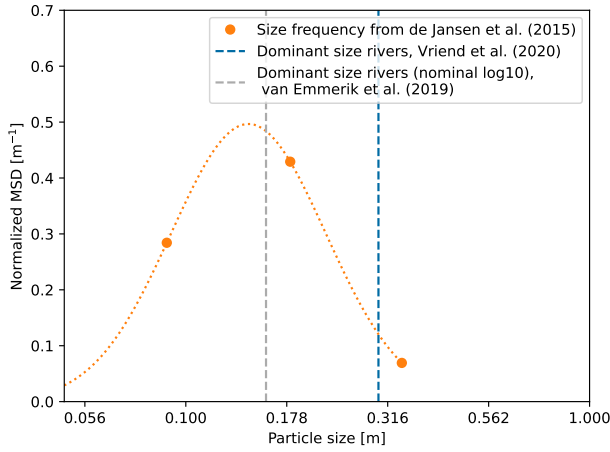


Figure D.11: Estimated input mass size distributions given data from waste treatment plants (Jansen et al., 2015) and visual observations of large plastic items (de Vries et al., 2021).

dotted orange line), $m \sim l^3$ for cubic objects (the dashed orange line). In reality marine plastics will be a combination of differently shaped objects, which means that this spatial dimension is likely (Kaandorp et al., 2021) to have an intermediate value around 2–3. The value of d_N is introduced as a parameter in our model, where we calibrate it such that the predicted mass and number concentrations are consistent with the available measurement data.

D.3.4 Particle rise velocity estimates

Particle rise velocity estimates were obtained from experimental data (Poulain et al., 2019), see Figure D.13 (left panel). We use the rise velocity formula for spheres from Dietrich (1982) to interpolate this data, where the particle density is calibrated to optimally fit the measurements ($\rho = 1010 \text{ kg m}^{-3}$). The estimated rise velocities were used to resolve transport at different particle sizes. We denote the particle size corresponding to the best fit in the left panel as the effective particle size L^* .

In reality, particles have different shapes and densities, and thus different rise velocities. We assume that for a given ‘true’ particle size L , rise velocities vary, due to the different shapes and densities. We estimate this relation by looking at the difference between the measured particle size (black markers in Figure D.13a) with the particle size corresponding to the best fit (blue line in Figure D.13a). The result is shown in Figure D.13b. We approximate this relation by a straight line in log-log space, where the slope ($a_{L_{eff}}$) and intercept ($b_{L_{eff}}$) are uncertain, and introduced as parameters to the model to control the vertical mixing behavior. Furthermore, the variance of the fit is estimated, such that each particle size L in the model corresponds to an assembly of particles with different effective particle sizes L^* as resolved by the Lagrangian simulations.

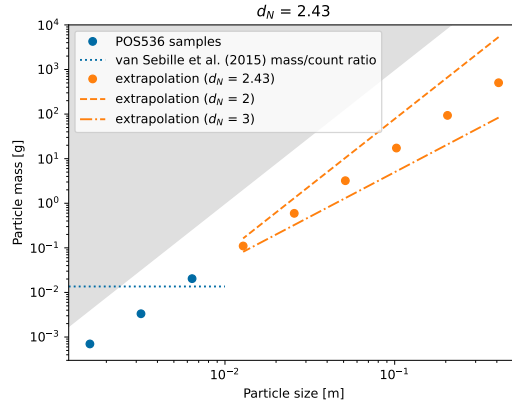
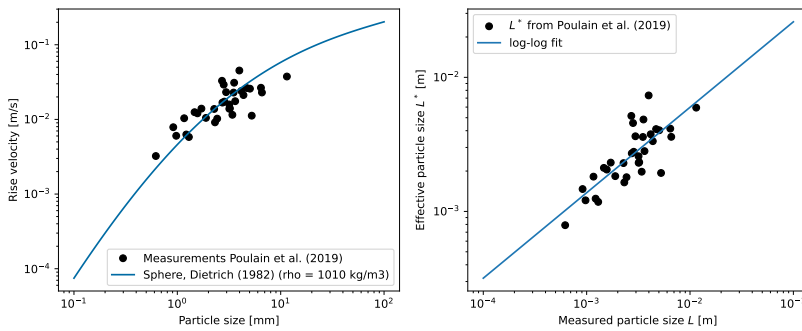


Figure D.12: Calculated and extrapolated particle size versus particle mass

Figure D.13: Particle rise velocity calculations: particle size versus measured rise velocities from Poulain et al., 2019 (left panel), and the map made from the measured particle size to the effective particle size of a spherical particle with density 1010 kg m^{-3} (right panel).

References

- Abeynayaka, A. et al. (2020), “Rapid sampling of suspended and floating microplastics in challenging riverine and coastal water environments in Japan”, *Water*, vol. 12, no. 1903, DOI: [10.3390/w12071903](https://doi.org/10.3390/w12071903).
- Alsina, J. M., C. E. Jongedijk & E. van Sebille (2020), “Laboratory Measurements of the Wave-Induced Motion of Plastic Particles: Influence of Wave Period, Plastic Size and Plastic Density”, *Journal of Geophysical Research: Oceans*, vol. 125, no. 12, DOI: [10.1029/2020JC016294](https://doi.org/10.1029/2020JC016294).
- Amaral-Zettler, L. A., E. R. Zettler & T. J. Mincer (2020), “Ecology of the plastisphere”, *Nature Reviews Microbiology*, vol. 18, no. 3, pp. 139–151, DOI: [10.1038/s41579-019-0308-0](https://doi.org/10.1038/s41579-019-0308-0).
- Amaral-Zettler, L. A., E. R. Zettler, T. J. Mincer, M. A. Klaassen & S. M. Gallagher (2021), “Biofouling impacts on polyethylene density and sinking in coastal waters: A macro/micro tipping point?”, *Water Research*, vol. 201, p. 117289, DOI: [10.1016/j.watres.2021.117289](https://doi.org/10.1016/j.watres.2021.117289).
- Andrades, R., R. G. Santos, J. C. Joyeux, D. Chelazzi, A. Cincinelli & T. Giarrizzo (2018), “Marine debris in Trindade Island, a remote island of the South Atlantic”, *Marine Pollution Bulletin*, vol. 137, pp. 180–184, DOI: [10.1016/j.marpolbul.2018.10.003](https://doi.org/10.1016/j.marpolbul.2018.10.003).
- Andrady, A. L., J. E. Pegram & Y. Song (1993), “Studies on enhanced degradable plastics. II. Weathering of enhanced photodegradable polyethylenes under marine and freshwater floating exposure”, *Journal of Environmental Polymer Degradation*, vol. 1, no. 2, pp. 117–126, DOI: [10.1007/BF01418205](https://doi.org/10.1007/BF01418205).
- Andrady, A. L. (2011), “Microplastics in the marine environment”, *Marine Pollution Bulletin*, vol. 62, no. 8, pp. 1596–1605, DOI: [10.1016/j.marpolbul.2011.05.030](https://doi.org/10.1016/j.marpolbul.2011.05.030).
- Aumont, O., C. Ethé, A. Tagliabue, L. Bopp & M. Gehlen (2015), “PISCES-v2: An ocean biogeochemical model for carbon and ecosystem studies”, *Geoscientific Model Development*, vol. 8, no. 8, pp. 2465–2513, DOI: [10.5194/gmd-8-2465-2015](https://doi.org/10.5194/gmd-8-2465-2015).
- Bachmaier, M. & M. Backes (2011), “Variogram or Semivariogram? Variance or Semivariance? Allan Variance or Introducing a New Term?”, *Mathematical Geosciences*, vol. 43, no. 6, pp. 735–740, DOI: [10.1007/s11004-011-9348-3](https://doi.org/10.1007/s11004-011-9348-3).

- Balas, C. E., A. Ergin, A. T. Williams & L. Koc (2004), “Marine litter prediction by artificial intelligence”, *Marine Pollution Bulletin*, vol. 48, no. 5-6, pp. 449–457, DOI: [10.1016/j.marpolbul.2003.08.020](https://doi.org/10.1016/j.marpolbul.2003.08.020).
- Becker, J. J. et al. (2009), “Global Bathymetry and Elevation Data at 30 Arc Seconds Resolution : SRTM30 _ PLUS Global Bathymetry and Elevation Data at 30 Arc Seconds Resolution : SRTM30 PLUS”, *Marine Geodesy*, vol. 32, pp. 355–371, DOI: [10.1080/01490410903297766](https://doi.org/10.1080/01490410903297766).
- Bird, N. R. A., A. M. Tarquis & A. P. Whitmore (2009), “Modeling Dynamic Fragmentation of Soil”, *Vadose Zone Journal*, vol. 8, no. 1, pp. 197–201, DOI: [10.2136/vzj2008.0046](https://doi.org/10.2136/vzj2008.0046).
- Bond, T., V. Ferrandiz-mas, M. Felipe-sotelo & E. van Sebille (2018), “The occurrence and degradation of aquatic plastic litter based on polymer physicochemical properties : A review”, *Critical Reviews in Environmental Science and Technology*, vol. 48, no. 7-9, pp. 685–722, DOI: [10.1080/10643389.2018.1483155](https://doi.org/10.1080/10643389.2018.1483155).
- Booth, A. M. et al. (2017), *Microplastic in global and Norwegian marine environments: Distributions, degradation mechanisms and transport*, tech. rep.
- Borrelle, S. B. et al. (2020), “Predicted growth in plastic waste exceeds efforts to mitigate plastic pollution”, *Science*, vol. 369, no. 6510, pp. 1515–1518, DOI: [10.1126/science.aba3656](https://doi.org/10.1126/science.aba3656).
- Bortnik, J. & E. Camporeale (2021), “Ten Ways to Apply Machine Learning in Earth and Space Sciences”, *Eos*, vol. 102, DOI: [10.1029/2021E0160257](https://doi.org/10.1029/2021E0160257).
- Børve, E., P. E. Isachsen & O. A. Nøst (2021), “Rectified tidal transport in Lofoten–Vesterålen, northern Norway”, *Ocean Science*, vol. 17, no. 6, pp. 1753–1773, DOI: [10.5194/os-17-1753-2021](https://doi.org/10.5194/os-17-1753-2021).
- Brach, L. et al. (2018), “Anticyclonic eddies increase accumulation of microplastic in the North Atlantic subtropical gyre”, *Marine Pollution Bulletin*, vol. 126, no. December 2017, pp. 191–196, DOI: [10.1016/j.marpolbul.2017.10.077](https://doi.org/10.1016/j.marpolbul.2017.10.077).
- Breivik, Ø., P. A. Janssen & J. R. Bidlot (2014), “Approximate stokes drift profiles in deep water”, *Journal of Physical Oceanography*, vol. 44, no. 9, pp. 2433–2445, DOI: [10.1175/JPO-D-14-0020.1](https://doi.org/10.1175/JPO-D-14-0020.1).
- Breivik, Ø., J. R. Bidlot & P. A. Janssen (2016), “A Stokes drift approximation based on the Phillips spectrum”, *Ocean Modelling*, vol. 100, pp. 49–56, DOI: [10.1016/j.ocemod.2016.01.005](https://doi.org/10.1016/j.ocemod.2016.01.005).
- van den Bremer, T. S. & Ø Breivik (2017), “Stokes Drift”, *Philosophical Transactions A*, vol. 376, no. 20170104.
- Brennan, E., C. Wilcox & B. Denise (2018), “Science of the Total Environment Connecting flux , deposition and resuspension in coastal debris surveys”, *Science of the Total Environment*, vol. 644, pp. 1019–1026, DOI: [10.1016/j.scitotenv.2018.06.352](https://doi.org/10.1016/j.scitotenv.2018.06.352).
- Browne, M. A. (2015), “Sources and Pathways of Microplastics to Habitats”, in: *Marine Anthropogenic Litter*, ed. by Bergmann, M., Gutow, L. & Klages, M., pp. 229–244.
- Burgess, H. K., C. E. Herring, S Lippiatt, S Lowe & A. V. Uhrin (2021), *NOAA Marine Debris Monitoring and Assessment Project Shoreline Survey Guide*, tech. rep., DOI: [10.25923/g720-2n18](https://doi.org/10.25923/g720-2n18).

- Byron, M., J. Einarsson, K. Gustavsson, G. Voth, B. Mehlig & E. Variano (2015), “Shape-dependence of particle rotation in isotropic turbulence”, *Physics of Fluids*, vol. 27, no. 3, DOI: [10.1063/1.4913501](https://doi.org/10.1063/1.4913501), arXiv: [1412.3166](https://arxiv.org/abs/1412.3166).
- Carpenter, E. J. & K. L. Smith Jr. (1972), “Plastics on the Sargasso sea surface”, *Science*, vol. 175, no. 4027, pp. 1240–1241.
- Carson, H. S., S. L. Colbert, M. J. Kaylor & K. J. McDermid (2011), “Small plastic debris changes water movement and heat transfer through beach sediments”, *Marine Pollution Bulletin*, vol. 62, no. 8, pp. 1708–1713, DOI: [10.1016/j.marpolbul.2011.05.032](https://doi.org/10.1016/j.marpolbul.2011.05.032).
- Carson, H. S., M. S. Nerheim, K. A. Carroll & M. Eriksen (2013), “The plastic-associated microorganisms of the North Pacific Gyre.”, *Marine pollution bulletin*, vol. 75, no. 1-2, pp. 126–132, DOI: [10.1016/j.marpolbul.2013.07.054](https://doi.org/10.1016/j.marpolbul.2013.07.054).
- Castro-Jiménez, J., D. González-fernández, M. Fornier, N. Schmidt & R. Sempéré (2019), “Macro-litter in surface waters from the Rhone River : Plastic pollution and loading to the NW Mediterranean Sea”, *Marine Pollution Bulletin*, vol. 146, no. September, pp. 60–66, DOI: [10.1016/j.marpolbul.2019.05.067](https://doi.org/10.1016/j.marpolbul.2019.05.067).
- Charalambous, C. (2015), “On the Evolution of Particle Fragmentation With Applications to Planetary Surfaces”, PhD thesis, Imperial College London, DOI: doi.org/10.25560/32780.
- Chor, T., D. Yang, C. Meneveau & M. Chamecki (2018), “A Turbulence Velocity Scale for Predicting the Fate of Buoyant Materials in the Oceanic Mixed Layer”, *Geophysical Research Letters*, vol. 45, DOI: [10.1029/2018GL080296](https://doi.org/10.1029/2018GL080296).
- Chubarenko, I., A. Bagaev, M. Zobkov & E. Esiukova (2016), “On some physical and dynamical properties of microplastic particles in marine environment”, *Marine Pollution Bulletin*, vol. 108, no. 1-2, pp. 105–112, DOI: [10.1016/j.marpolbul.2016.04.048](https://doi.org/10.1016/j.marpolbul.2016.04.048).
- Clark, L. K., M. H. Dibenedetto, N. T. Ouellette & J. R. Koseff (2020), “Settling of inertial nonspherical particles in wavy flow”, *Physical Review Fluids*, vol. 5, no. 12, p. 124301, DOI: [10.1103/PhysRevFluids.5.124301](https://doi.org/10.1103/PhysRevFluids.5.124301).
- Clift, R., J. R. Grace & M. E. Weber (1978), *Bubbles, Drops, and Particles*, Academic Press, ISBN: 012176950X.
- Cole, M. et al. (2013), “Microplastic Ingestion by Zooplankton”, *Environmental Science & Technology*, vol. 47, pp. 6646–6655, DOI: [10.1021/es400663f](https://doi.org/10.1021/es400663f).
- Collignon, A., J. H. Hecq, F. Glagani, P. Voisin, F. Collard & A. Goffart (2012), “Neustonic microplastic and zooplankton in the North Western Mediterranean Sea”, *Marine Pollution Bulletin*, vol. 64, no. 4, pp. 861–864, DOI: [10.1016/j.marpolbul.2012.01.011](https://doi.org/10.1016/j.marpolbul.2012.01.011).
- Collignon, A., J. H. Hecq, F. Galgani, F. Collard & A. Goffart (2014), “Annual variation in neustonic micro- and meso-plastic particles and zooplankton in the Bay of Calvi (Mediterranean-Corsica)”, *Marine Pollution Bulletin*, vol. 79, no. 1-2, pp. 293–298, DOI: [10.1016/j.marpolbul.2013.11.023](https://doi.org/10.1016/j.marpolbul.2013.11.023).
- Colton, J. B., F. D. Knapp & B. R. Bums (1974), “Plastic Particles in Surface Waters of the Northwestern Atlantic”, *Science*, vol. 185, no. 4150, pp. 491–497.

- Constant, M. et al. (2019), “Beached microplastics in the Northwestern Mediterranean Sea”, *Marine Pollution Bulletin*, vol. 142, no. March, pp. 263–273, DOI: [10.1016/j.marpolbul.2019.03.032](https://doi.org/10.1016/j.marpolbul.2019.03.032).
- Cope, T. E. et al. (2017), “Artificial grammar learning in vascular and progressive non-fluent aphasia”, *Neuropsychologia*, vol. 104, pp. 201–213, DOI: [10.1016/j.neuropsychologia.2017.08.022](https://doi.org/10.1016/j.neuropsychologia.2017.08.022).
- Cózar, A. et al. (2014), “Plastic debris in the open ocean.”, *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, no. 28, pp. 10239–44, DOI: [10.1073/pnas.1314705111](https://doi.org/10.1073/pnas.1314705111).
- Cózar, A. et al. (2015), “Plastic accumulation in the mediterranean sea”, *PLoS ONE*, vol. 10, no. 4, pp. 1–12, DOI: [10.1371/journal.pone.0121762](https://doi.org/10.1371/journal.pone.0121762).
- Cózar, A. et al. (2021), “Marine Litter Windrows: A Strategic Target to Understand and Manage the Ocean Plastic Pollution”, *Frontiers in Marine Science*, vol. 8, no. February, pp. 1–9, DOI: [10.3389/fmars.2021.571796](https://doi.org/10.3389/fmars.2021.571796).
- Critchell, K et al. (2015), “Modelling the fate of marine debris along a complex shoreline: Lessons from the Great Barrier Reef”, *Estuarine, Coastal and Shelf Science*, vol. 167, pp. 414–426, DOI: [10.1016/j.ecss.2015.10.018](https://doi.org/10.1016/j.ecss.2015.10.018).
- Critchell, K. & J. Lambrechts (2016), “Modelling accumulation of marine plastics in the coastal zone; what are the dominant physical processes?”, *Estuarine, Coastal and Shelf Science*, vol. 171, pp. 111–122, DOI: [10.1016/j.ecss.2016.01.036](https://doi.org/10.1016/j.ecss.2016.01.036).
- Cunningham, H. J., C. Higgins & T. S. van den Bremer (2022), “The Role of the Unsteady Surface Wave-Driven Ekman–Stokes Flow in the Accumulation of Floating Marine Litter”, *Journal of Geophysical Research: Oceans*, vol. 127, no. 6, pp. 1–19, DOI: [10.1029/2021jc0018106](https://doi.org/10.1029/2021jc0018106).
- Davis, R. E. (1985), “Subsurface Float”, *Journal of Geophysical Research*, vol. 90, pp. 4741–4755.
- Delandmeter, P. & E. van Sebille (2019), “The Parcels v2 . 0 Lagrangian framework : new field interpolation schemes”, *Geoscientific Model Development*, vol. 12, pp. 3571–3584, DOI: [doi.org/10.5194/gmd-12-3571-2019](https://doi.org/doi.org/10.5194/gmd-12-3571-2019).
- DiBenedetto, M. H. (2020), “Non-breaking Wave Effects on Buoyant Particle Distributions”, vol. 7, p. 148, DOI: [10.3389/fmars.2020.00148](https://doi.org/10.3389/fmars.2020.00148).
- DiBenedetto, M. H., N. T. Ouellette & J. R. Koseff (2018), “Transport of anisotropic particles under waves”, *Journal of Fluid Mechanics*, vol. 837, pp. 320–340, DOI: [10.1017/jfm.2017.853](https://doi.org/10.1017/jfm.2017.853).
- Dietrich, W. E. (1982), “Settling velocity of natural particles”, *Water Resources Research*, vol. 18, no. 6, pp. 1615–1626, DOI: [10.1029/WR018i006p01615](https://doi.org/10.1029/WR018i006p01615).
- van Duinen, B., M. L. Kaandorp & E. van Sebille (2022), “Identifying Marine Sources of Beached Plastics Through a Bayesian Framework: Application to Southwest Netherlands”, *Geophysical Research Letters*, vol. 49, no. 4, pp. 1–9, DOI: [10.1029/2021GL097214](https://doi.org/10.1029/2021GL097214).
- Edyvane, K. S., A. Dalgetty, P. W. Hone, J. S. Higham & N. M. Wace (2004), “Long-term marine litter monitoring in the remote Great Australian Bight, South Australia”, *Marine Pollution Bulletin*, vol. 48, no. 11-12, pp. 1060–1075, DOI: [10.1016/j.marpolbul.2003.12.012](https://doi.org/10.1016/j.marpolbul.2003.12.012).

- Efimova, I., M. Bagaeva, A. Bagaev, A. Kileso & I. P. Chubarenko (2018), “Secondary Microplastics Generation in the Sea Swash Zone With Coarse Bottom Sediments : Laboratory Experiments”, *Frontiers in Marine Science*, vol. 5, no. 313, DOI: [10.3389/fmars.2018.00313](#).
- Egger, M. et al. (2020a), “A spatially variable scarcity of floating microplastics in the eastern North Pacific Ocean”, *Environmental Research Letters*, vol. 15, no. 11, DOI: [10.1088/1748-9326/abb4f](#).
- Egger, M., F. Sulu-Gambari & L. Lebreton (2020b), “First evidence of plastic fallout from the North Pacific Garbage Patch”, *Scientific Reports*, vol. 10, no. 1, pp. 1–10, DOI: [10.1038/s41598-020-64465-8](#).
- Emerick, A. A. & A. C. Reynolds (2013), “Ensemble smoother with multiple data assimilation”, *Computers and Geosciences*, vol. 55, pp. 3–15.
- van Emmerik, T. & A. Schwarz (2020), “Plastic debris in rivers”, *WIREs Water*, vol. 7, no. e1398, DOI: [10.1002/wat2.1398](#).
- van Emmerik, T., E. Strady, T. C. Kieu-Le, L. Nguyen & N. Gratiot (2019), “Seasonality of riverine macroplastic transport”, *Scientific Reports*, vol. 9, no. 1, pp. 1–9, DOI: [10.1038/s41598-019-50096-1](#).
- Enders, K., R. Lenz, C. A. Stedmon & T. G. Nielsen (2015), “Abundance , size and polymer composition of marine microplastics > 10 μm in the Atlantic Ocean and their modelled vertical distribution”, *Marine Pollution Bulletin*, vol. 100, no. 1, pp. 70–81, DOI: [10.1016/j.marpolbul.2015.09.027](#).
- 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”, *PLoS ONE*, vol. 9, no. 12, pp. 1–15, DOI: [10.1371/journal.pone.0111913](#).
- Eriksson, C., H. Burton, S. Fitch, M. Schulz & J. van den Hoff (2013), “Daily accumulation rates of marine debris on sub-Antarctic island beaches”, *Marine Pollution Bulletin*, vol. 66, no. 1-2, pp. 199–208, DOI: [10.1016/j.marpolbul.2012.08.026](#).
- Erni-Cassola, G., M. I. Gibson, R. C. Thompson & J. A. Christie-Oleza (2017), “Lost, but Found with Nile Red: A Novel Method for Detecting and Quantifying Small Microplastics (1 mm to 20 μm) in Environmental Samples”, *Environmental Science and Technology*, vol. 51, no. 23, pp. 13641–13648, DOI: [10.1021/acs.est.7b04512](#).
- Evensen, G. (2009), “The ensemble Kalman filter for combined state and parameter estimation”, *IEEE Control Systems*, vol. 29, no. 3, pp. 83–104, DOI: [10.1109/MCS.2009.932223](#).
- Evensen, G., F. C. Vossepoel & P. J. van Leeuwen (2022), *Data Assimilation Fundamentals*, Springer, ISBN: 9783030967086.
- Falkowski, P. (2012), “Ocean Science: The power of plankton”, *Nature*, vol. 483, no. 7387, S17–S20, DOI: [10.1038/483S17a](#).
- Fazey, F. M. & P. G. Ryan (2016), “Biofouling on buoyant marine plastics: An experimental study into the effect of size on surface longevity”, *Environmental Pollution*, vol. 210, pp. 354–360, DOI: [10.1016/j.envpol.2016.01.026](#).
- Fendall, L. S. & M. A. Sewell (2009), “Contributing to marine pollution by washing your face: Microplastics in facial cleansers”, *Marine Pollution Bulletin*, vol. 58, no. 8, pp. 1225–1228, DOI: [10.1016/j.marpolbul.2009.04.025](#).

- Ferrari, R. & C. Wunsch (2009), "Ocean circulation kinetic energy: Reservoirs, sources, and sinks", *Annual Review of Fluid Mechanics*, vol. 41, pp. 253–282, DOI: [10.1146/annurev.fluid.40.111406.102139](#).
- Fischer, R. et al. (2022), "Modelling submerged biofouled microplastics and their vertical trajectories", *Biogeosciences*, vol. 19, no. 8, pp. 2211–2234, DOI: [10.5194/bg-19-2211-2022](#).
- Fok, L., P. K. Cheung, G. Tang & W. C. Li (2017), "Size distribution of stranded small plastic debris on the coast of Guangdong, South China", *Environmental Pollution*, vol. 220, pp. 407–412, DOI: [10.1016/j.envpol.2016.09.079](#).
- Fossi, M. C. et al. (2012), "Are baleen whales exposed to the threat of microplastics? A case study of the Mediterranean fin whale (*Balaenoptera physalus*)", *Marine Pollution Bulletin*, vol. 64, no. 11, pp. 2374–2379, DOI: [10.1016/j.marpolbul.2012.08.013](#).
- Fowler, C. (1987), "Marine Debris and Northern Fur Seals : a Case Study", *Marine pollution bulletin*, vol. 18, no. 68, pp. 326–335.
- Franeker, J. A. V. et al. (2011), "Monitoring plastic ingestion by the northern fulmar *Fulmarus glacialis* in the North Sea", *Environmental Pollution*, vol. 159, no. 10, pp. 2609–2615, DOI: [10.1016/j.envpol.2011.06.008](#).
- Frias, J. P., V. Otero & P. Sobral (2014), "Evidence of microplastics in samples of zooplankton from Portuguese coastal waters", *Marine Environmental Research*, vol. 95, pp. 89–95, DOI: [10.1016/j.marenvres.2014.01.001](#).
- Gajšt, T., T. Bizjak, A. Palatinus, S. Liubartseva & A. Kržan (2016), "Sea surface microplastics in Slovenian part of the Northern Adriatic", *Marine Pollution Bulletin*, vol. 113, no. 1-2, pp. 392–399, DOI: [10.1016/j.marpolbul.2016.10.031](#).
- Galgani, F. et al. (2021), "Are litter, plastic and microplastic quantities increasing in the ocean?", *Microplastics and Nanoplastics*, vol. 1, no. 1, pp. 8–11, DOI: [10.1186/s43591-020-00002-8](#).
- 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", *Journal of Marine Systems*, vol. 187, no. June, pp. 260–276, DOI: [10.1016/j.jmarsys.2018.06.015](#).
- Geer, A. J. (2021), "Learning earth system models from observations: Machine learning or data assimilation?", *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 379, no. 2194, DOI: [10.1098/rsta.2020.0089](#).
- Geernaert, G. (1990), "Bulk Parameterizations for the Wind Stress and Heat Fluxes", in: *Surface Waves and Fluxes*, ed. by Geernaert, G. & Plant, W., Kluwer Academic, pp. 91–172.
- Gerigny, O., M. Brun, C. Tomasino, C. Lacroix, L. Kerambrun & F. Galgani (2018), *Évaluation du descripteur 10 "Déchets marins" en France métropolitaine. Rapport scientifique pour l'évaluation 2018 au titre de la DCSMM (ministère de l'environnement). Rapport interne IFREMER/CEDRE*, tech. rep., p. 349.
- Gerritse, J., H. A. Leslie, C. A. de Tender, L. I. Devriese & A. D. Vethaak (2020), "Fragmentation of plastic objects in a laboratory seawater microcosm", *Scientific Reports*, vol. 10, no. 1, pp. 1–16, DOI: [10.1038/s41598-020-67927-1](#).

- Geyer, R., J. R. Jambeck & K. L. Law (2017), “Production, use, and fate of all plastics ever made”, *Science Advances*, vol. 3, no. 7, pp. 19–24, DOI: [10.1126/sciadv.1700782](https://doi.org/10.1126/sciadv.1700782).
- Ghahramani, Z. (2013), “Bayesian non-parametrics and the probabilistic approach to modelling”, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 371, no. 1984, DOI: [10.1098/rsta.2011.0553](https://doi.org/10.1098/rsta.2011.0553).
- Global Monitoring and Forecasting Center (2020), *Global Ocean Waves Reanalysis WAVERYs product, E.U. Copernicus Marine Service Information [Data set]*.
- Global Monitoring and Forecasting Center (2021), *Atlantic-European North West Shelf Ocean Physics Reanalysis product, E.U. Copernicus Marine Service Information [Data set]*.
- Granado, I. et al. (2019), “Beach litter forecasting on the south-eastern coast of the Bay of Biscay: A bayesian networks approach”, *Continental Shelf Research*, vol. 180, pp. 14–23, DOI: [10.1016/j.csr.2019.04.016](https://doi.org/10.1016/j.csr.2019.04.016).
- Gräwe, U., H. Burchard, M. Müller & H. M. Schuttelaars (2014), “Seasonal variability in M2 and M4 tidal constituents and its implications for the coastal residual sediment transport”, *Geophysical Research Letters*, vol. 41, no. 15, pp. 5563–5570, DOI: [10.1002/2014GL060517](https://doi.org/10.1002/2014GL060517).
- Gregory, A. S., N. R. A. Bird, C. W. Watts & A. P. Whitmore (2012), “An assessment of a new model of dynamic fragmentation of soil with test data”, *Soil and Tillage Research*, vol. 120, no. April, pp. 61–68, DOI: [10.2136/vzj2008.0046](https://doi.org/10.2136/vzj2008.0046).
- Guillamón, A., J. Navarro & J. M. Ruiz (1998), “Kernel density estimation using weighted data”, *Communication in Statistics - Theory and Methods*, vol. 27, no. 9, pp. 2123–2135, DOI: [10.1080/03610929808832217](https://doi.org/10.1080/03610929808832217).
- Gündoğdu, S. & C. Çevik (2017), “Micro- and mesoplastics in Northeast Levantine coast of Turkey: The preliminary results from surface samples”, *Marine Pollution Bulletin*, vol. 118, no. 1-2, pp. 341–347, DOI: [10.1016/j.marpolbul.2017.03.002](https://doi.org/10.1016/j.marpolbul.2017.03.002).
- Gündoğdu, S., C. Çevik, B. Ayat, B. Aydoğan & S. Karaca (2018), “How microplastics quantities increase with flood events? An example from Mersin Bay NE Levantine coast of Turkey”, *Environmental Pollution*, vol. 239, pp. 342–350, DOI: [10.1016/j.envpol.2018.04.042](https://doi.org/10.1016/j.envpol.2018.04.042).
- Güven, O., K. Gökdağ, B. Jovanović & A. E. Kideys (2017), “Microplastic litter composition of the Turkish territorial waters of the Mediterranean Sea, and its occurrence in the gastrointestinal tract of fish”, *Environmental Pollution*, vol. 223, pp. 286–294, DOI: [10.1016/j.envpol.2017.01.025](https://doi.org/10.1016/j.envpol.2017.01.025).
- de Haan, W. P., A. Sanchez-Vidal & M. Canals (2019), “Floating microplastics and aggregate formation in the Western Mediterranean Sea”, *Marine Pollution Bulletin*, vol. 140, no. December 2018, pp. 523–535, DOI: [10.1016/j.marpolbul.2019.01.053](https://doi.org/10.1016/j.marpolbul.2019.01.053).
- Haarr, M. L., L. Westerveld, J. Fabres, K. R. Iversen & K. E. T. Busch (2019), “A novel GIS-based tool for predicting coastal litter accumulation and optimising coastal cleanup actions”, *Marine Pollution Bulletin*, vol. 139, pp. 117–126, DOI: [10.1016/j.marpolbul.2018.12.025](https://doi.org/10.1016/j.marpolbul.2018.12.025).

- Haarr, M. L., M. Pantalos, M. K. Hartviksen & M. Gressetvold (2020), "Citizen science data indicate a reduction in beach litter in the Lofoten archipelago in the Norwegian Sea", *Marine Pollution Bulletin*, vol. 153, p. 111000, DOI: [10.1016/j.marpolbul.2020.111000](https://doi.org/10.1016/j.marpolbul.2020.111000).
- van der Hal, N., A. Ariel & D. L. Angel (2017), "Exceptionally high abundances of microplastics in the oligotrophic Israeli Mediterranean coastal waters", *Marine Pollution Bulletin*, vol. 116, no. 1-2, pp. 151–155, DOI: [10.1016/j.marpolbul.2016.12.052](https://doi.org/10.1016/j.marpolbul.2016.12.052).
- Hardesty, B. D., T. J. Lawson, T. van der Velde, M. Lansdell & C. Wilcox (2017), "Estimating quantities and sources of marine debris at a continental scale", *Frontiers in Ecology and the Environment*, vol. 15, no. 1, pp. 18–25, DOI: [10.1002/fee.1447](https://doi.org/10.1002/fee.1447).
- Hastie, T., R. Tibshirani & J. Friedman (2008), *The Elements of Statistical Learning*, 2nd, Springer.
- Hellweger, F. L. & V. Bucci (2008), "A bunch of tiny individuals — Individual-based modeling for microbes", *Ecological Modelling*, vol. 220, pp. 8–22, DOI: [10.1016/j.ecolmodel.2008.09.004](https://doi.org/10.1016/j.ecolmodel.2008.09.004).
- Hengstmann, E., D. Gräwe, M. Tamminga & E. K. Fischer (2017), "Marine litter abundance and distribution on beaches on the Isle of Rügen considering the influence of exposition, morphology and recreational activities", *Marine Pollution Bulletin*, vol. 115, no. 1-2, pp. 297–306, DOI: [10.1016/j.marpolbul.2016.12.026](https://doi.org/10.1016/j.marpolbul.2016.12.026).
- Henriksen, M. L., C. B. Karlsen, P. Klarskov & M. Hinge (2022), "Plastic classification via in-line hyperspectral camera analysis and unsupervised machine learning", *Vibrational Spectroscopy*, vol. 118, no. November 2021, p. 103329, DOI: [10.1016/j.vibspec.2021.103329](https://doi.org/10.1016/j.vibspec.2021.103329).
- Heo, N. W. et al. (2013), "Distribution of small plastic debris in cross-section and high strandline on Heungnam beach, South Korea", *Ocean Science Journal*, vol. 48, no. 2, pp. 225–233, DOI: [10.1007/s12601-013-0019-9](https://doi.org/10.1007/s12601-013-0019-9).
- Herman, J. R., N. Krotkov, E. Celarier, D. Larko & G. Labow (1999), "Distribution of UV radiation at the Earth's surface from TOMS-measured UV-backscattered radiances", *Journal of Geophysical Research*, vol. 104, no. D10, pp. 12,059–12,076.
- Hersbach, H. et al. (2020), "The ERA5 global reanalysis", *Quarterly Journal of the Royal Meteorological Society*, vol. 146, pp. 1999–2049, DOI: [10.1002/qj.3803](https://doi.org/10.1002/qj.3803).
- Hidalgo-Ruz, V. & M. Thiel (2013), "Distribution and abundance of small plastic debris on beaches in the SE Pacific (Chile): A study supported by a citizen science project", *Marine Environmental Research*, vol. 87-88, pp. 12–18, DOI: [10.1016/j.marenvres.2013.02.015](https://doi.org/10.1016/j.marenvres.2013.02.015).
- Hinata, H., K. Mori, K. Ohno, Y. Miyao & T. Kataoka (2017), "An estimation of the average residence times and onshore-offshore diffusivities of beached microplastics based on the population decay of tagged meso- and macrolitter", *Marine Pollution Bulletin*, vol. 122, no. 1-2, pp. 17–26, DOI: [10.1016/j.marpolbul.2017.05.012](https://doi.org/10.1016/j.marpolbul.2017.05.012).
- Hinata, H., N. Sagawa, T. Kataoka & H. Takeoka (2020), "Numerical modeling of the beach process of marine plastics : A probabilistic and diagnostic approach with a particle tracking method", *Marine Pollution Bulletin*, vol. 152, no. October 2019, p. 110910, DOI: [10.1016/j.marpolbul.2020.110910](https://doi.org/10.1016/j.marpolbul.2020.110910).

- Hoffmann, R., V. I. Minkin & B. K. Carpenter (1997), "Ockham's Razor and chemistry", *Hyle-International Journal for the Philosophy of Chemistry*, vol. 3, pp. 3–28, DOI: [10.1093/oso/9780199755905.003.0010](https://doi.org/10.1093/oso/9780199755905.003.0010).
- Hong, S., J. Lee, D. Kang, H. W. Choi & S. H. Ko (2014), "Quantities, composition, and sources of beach debris in Korea from the results of nationwide monitoring", *Marine Pollution Bulletin*, vol. 84, no. 1-2, pp. 27–34, DOI: [10.1016/j.marpolbul.2014.05.051](https://doi.org/10.1016/j.marpolbul.2014.05.051).
- Husain, A., J. Reddy, D. Bisht & M. Sajid (2021), "Fractal dimension of coastline of Australia", *Scientific Reports*, vol. 11, no. 1, pp. 1–10, DOI: [10.1038/s41598-021-85405-0](https://doi.org/10.1038/s41598-021-85405-0).
- Isobe, A. & S. Iwasaki (2022), "The fate of missing ocean plastics: Are they just a marine environmental problem?", *Science of the Total Environment*, vol. 825, p. 153935, DOI: [10.1016/j.scitotenv.2022.153935](https://doi.org/10.1016/j.scitotenv.2022.153935).
- Isobe, A., K. Kubo, Y. Tamura, E. Nakashima & N. Fujii (2014), "Selective transport of microplastics and mesoplastics by drifting in coastal waters", *Marine Pollution Bulletin*, vol. 89, no. 1-2, pp. 324–330, DOI: [10.1016/j.marpolbul.2014.09.041](https://doi.org/10.1016/j.marpolbul.2014.09.041).
- Isobe, A., K. Uchida, T. Tokai & S. Iwasaki (2015), "East Asian seas : A hot spot of pelagic microplastics", *Marine Pollution Bulletin*, vol. 101, no. 2, pp. 618–623, DOI: [10.1016/j.marpolbul.2015.10.042](https://doi.org/10.1016/j.marpolbul.2015.10.042).
- Iwasaki, S., A. Isobe, S. Kako, K. Uchida & T. Tokai (2017), "Fate of microplastics and mesoplastics carried by surface currents and wind waves: A numerical model approach in the Sea of Japan", *Marine Pollution Bulletin*, vol. 121, no. 1-2, pp. 85–96, DOI: [10.1016/j.marpolbul.2017.05.057](https://doi.org/10.1016/j.marpolbul.2017.05.057).
- Jambeck, J. R. et al. (2015), "Plastic waste inputs from land into the ocean", *Science*, vol. 347, no. 6223, pp. 768–771, DOI: [10.1126/science.1260352](https://doi.org/10.1126/science.1260352).
- Jansen, M., U. van Velzen & T. Pretz (2015), *Handbook for sorting of plastic packaging waste concentrates*, Wageningen UR - Food & Biobased Research, ISBN: 9789462575295.
- Jospin, L. V., H. Laga, F. Boussaid, W. Buntine & M. Bennamoun (2022), "Hands-On Bayesian Neural Networks - A Tutorial for Deep Learning Users", *IEEE Computational Intelligence Magazine*, vol. 17, no. 2, pp. 29–48, DOI: [10.1109/MCI.2022.3155327](https://doi.org/10.1109/MCI.2022.3155327), arXiv: [2007.06823](https://arxiv.org/abs/2007.06823).
- Kaandorp, M. L. A., H. A. Dijkstra & E. van Sebille (2020), "Closing the Mediterranean Marine Floating Plastic Mass Budget: Inverse Modeling of Sources and Sinks", *Environmental Science & Technology*, vol. 54, no. 19, pp. 11980–11989, DOI: [10.1021/acs.est.0c01984](https://doi.org/10.1021/acs.est.0c01984).
- Kaandorp, M. L. A., H. A. Dijkstra & E. van Sebille (2021), "Modelling size distributions of marine plastics under the influence of continuous cascading fragmentation", *Environmental Research Letters*, vol. 16, no. 5, p. 54075, DOI: [10.1088/1748-9326/abe9ea](https://doi.org/10.1088/1748-9326/abe9ea), arXiv: [2011.01775](https://arxiv.org/abs/2011.01775).
- Kaandorp, M. L. & R. P. Dwight (2020), "Data-driven modelling of the Reynolds stress tensor using random forests with invariance", *Computers & Fluids*, vol. 202, p. 104497, DOI: [10.1016/j.compfluid.2020.104497](https://doi.org/10.1016/j.compfluid.2020.104497).

- Kaandorp, M. L., S. Menzel & S. Schmitt (2017), “An Aerodynamic Perspective on Shape Deformation Methods”, *18th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, DOI: [10.2514/6.2017-3145](https://doi.org/10.2514/6.2017-3145).
- Kaandorp, M. L., S. L. Ypma, M. Boonstra, H. A. Dijkstra & E. Van Sebille (2022), “Using machine learning and beach cleanup data to explain litter quantities along the Dutch North Sea coast”, *Ocean Science*, vol. 18, no. 1, pp. 269–293, DOI: [10.5194/os-18-269-2022](https://doi.org/10.5194/os-18-269-2022).
- Kalogerakis, N. et al. (2017), “Microplastics Generation : Onset of Fragmentation of Polyethylene Films in Marine Environment Mesocosms”, *Frontiers in Marine Science*, vol. 4, no. 84, DOI: [10.3389/fmars.2017.00084](https://doi.org/10.3389/fmars.2017.00084).
- Kappraff, J. (1986), “The Fractal Geometry of coastlines: a study in fractals”, *Comp. & Maths. with Appls.* Vol. 12B, pp. 655–671.
- Karkanorachaki, K., S. Kiparissis, G. C. Kalogerakis, E. Yiantzi, E. Psillakis & N. Kalogerakis (2018), “Plastic pellets , meso- and microplastics on the coastline of Northern Crete : Distribution and organic pollution Plastic pellets , meso- and microplastics on the coastline of Northern Crete : Distribution and organic pollution”, *Marine Pollution Bulletin*, vol. 133, pp. 578–589, DOI: [10.1016/j.marpolbul.2018.06.011](https://doi.org/10.1016/j.marpolbul.2018.06.011).
- Kelso, N. V. & T. Patterson (2010), “Introducing Natural Earth Data - Naturalearth-data.com”, *Geographia Technica*, pp. 82–89.
- Koelmans, A. A., M. Kooi, K. L. Law & E. van Sebille (2017), “All is not lost: Deriving a top-down mass budget of plastic at sea”, *Environmental Research Letters*, vol. 12, no. 11, DOI: [10.1088/1748-9326/aa9500](https://doi.org/10.1088/1748-9326/aa9500).
- Kooi, M. & A. A. Koelmans (2019), “Simplifying Microplastic via Continuous Probability Distributions for Size, Shape, and Density”, *Environmental Science and Technology Letters*, vol. 6, pp. 551–557, DOI: [10.1021/acs.estlett.9b00379](https://doi.org/10.1021/acs.estlett.9b00379).
- Kooi, M., E. H. van Nes, M. Scheffer & A. A. Koelmans (2017), “Ups and Downs in the Ocean: Effects of Biofouling on Vertical Transport of Microplastics”, *Environmental Science and Technology*, vol. 51, no. 14, pp. 7963–7971, DOI: [10.1021/acs.est.6b04702](https://doi.org/10.1021/acs.est.6b04702).
- Kordella, S, M Geraga, G Papatheodorou, E Fakiris & I. M. Mitropoulou (2013), “Litter composition and source contribution for 80 beaches in Greece, Eastern Mediterranean: A nationwide voluntary clean-up campaign”, *Aquatic Ecosystem Health & Management*, vol. 16, no. 1, pp. 111–118, DOI: [10.1080/14634988.2012.759503](https://doi.org/10.1080/14634988.2012.759503).
- Korres, G., M. Ravdas & A. Zacharioudaki (2019), *Mediterranean Sea Waves Hind-cast (CMEMS MED-Waves) [Data set]*, DOI: doi.org/10.25423/CMCC/MEDSEA_HINDCAST_WAV_006_012.
- Kroodsma, D. A. et al. (2018), “Tracking the global footprint of fisheries”, *Science*, vol. 359, no. 6378, pp. 904–908, DOI: [10.1126/science.aao5646](https://doi.org/10.1126/science.aao5646).
- Kubota, M. (1994), “A Mechanism for the Accumulation of Floating Marine Debris North of Hawaii”, *Journal of Physical Oceanography*, vol. 24, no. 5, pp. 1059–1064, DOI: [10.1175/1520-0485\(1994\)024<1059:AMFTA0>2.0.CO;2](https://doi.org/10.1175/1520-0485(1994)024<1059:AMFTA0>2.0.CO;2).
- Kukulka, T., G. Proskurowski, S. Morét-Ferguson, D. W. Meyer & K. L. Law (2012), “The effect of wind mixing on the vertical distribution of buoyant plastic debris”, *Geophysical Research Letters*, vol. 39, no. 7, pp. 1–6, DOI: [10.1029/2012GL051116](https://doi.org/10.1029/2012GL051116).

- Kvale, K. F., A. E. Friederike Prowe & A. Oschlies (2020), "A Critical Examination of the Role of Marine Snow and Zooplankton Fecal Pellets in Removing Ocean Surface Microplastic", *Frontiers in Marine Science*, vol. 6, no. January, pp. 1–8, DOI: [10.3389/fmars.2019.00808](https://doi.org/10.3389/fmars.2019.00808).
- Lavers, J. L. & A. L. Bond (2017), "Exceptional and rapid accumulation of anthropogenic debris on one of the world's most remote and pristine islands", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 114, no. 23, pp. 6052–6055, DOI: [10.1073/pnas.1619818114](https://doi.org/10.1073/pnas.1619818114).
- Lebreton, L. et al. (2018), "Evidence that the Great Pacific Garbage Patch is rapidly accumulating plastic", *Scientific Reports*, vol. 8, no. 1, pp. 1–15, DOI: [10.1038/s41598-018-22939-w](https://doi.org/10.1038/s41598-018-22939-w).
- Lebreton, L. C., S. D. Greer & J. C. Borrero (2012), "Numerical modelling of floating debris in the world's oceans", *Marine Pollution Bulletin*, vol. 64, no. 3, pp. 653–661, DOI: [10.1016/j.marpolbul.2011.10.027](https://doi.org/10.1016/j.marpolbul.2011.10.027).
- Lebreton, L., M. Egger & B. Slat (2019), "A global mass budget for positively buoyant macroplastic debris in the ocean", *Scientific Reports*, vol. 9, p. 12922, DOI: [10.1038/s41598-019-49413-5](https://doi.org/10.1038/s41598-019-49413-5).
- Lebreton, L. C., J. Van Der Zwet, J. W. Damsteeg, B. Slat, A. Andrady & J. Reisser (2017), "River plastic emissions to the world's oceans", *Nature Communications*, vol. 8, pp. 1–10, DOI: [10.1038/ncomms15611](https://doi.org/10.1038/ncomms15611).
- Lee, R. F. & D. P. Sanders (2015), "The amount and accumulation rate of plastic debris on marshes and beaches on the Georgia coast", *Marine Pollution Bulletin*, vol. 91, no. 1, pp. 113–119, DOI: [10.1016/j.marpolbul.2014.12.019](https://doi.org/10.1016/j.marpolbul.2014.12.019).
- 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", *Ocean Science*, vol. 14, no. 5, pp. 1093–1126, DOI: [10.5194/os-14-1093-2018](https://doi.org/10.5194/os-14-1093-2018).
- Lenz, M. (2020), *Cruise Report POS536*, tech. rep., DOI: [10.3289/GEOMAR_REP_NS_56_2020](https://doi.org/10.3289/GEOMAR_REP_NS_56_2020).
- Li, W. C., H. F. Tse & L. Fok (2016), "Plastic waste in the marine environment: A review of sources, occurrence and effects", *Science of the Total Environment*, vol. 566–567, pp. 333–349, DOI: [10.1016/j.scitotenv.2016.05.084](https://doi.org/10.1016/j.scitotenv.2016.05.084).
- Liubartseva, S., G. Coppini, R. Lecci & E. Clementi (2018), "Tracking plastics in the Mediterranean : 2D Lagrangian model", *Marine Pollution Bulletin*, vol. 129, no. February, pp. 151–162, DOI: [10.1016/j.marpolbul.2018.02.019](https://doi.org/10.1016/j.marpolbul.2018.02.019).
- Lobelle, D. & M. Cunliffe (2011), "Early microbial biofilm formation on marine plastic debris", *Marine Pollution Bulletin*, vol. 62, no. 1, pp. 197–200, DOI: [10.1016/j.marpolbul.2010.10.013](https://doi.org/10.1016/j.marpolbul.2010.10.013).
- Lobelle, D. et al. (2021), "Global Modeled Sinking Characteristics of Biofouled Microplastic", *Journal of Geophysical Research: Oceans*, vol. 126, no. 4, pp. 1–15, DOI: [10.1029/2020JC017098](https://doi.org/10.1029/2020JC017098).
- Lumpkin, R., L. Centurioni & R. C. Perez (2016), "Full access fulfilling observing system implementation requirements with the global drifter array", *Journal of Atmospheric and Oceanic Technology*, vol. 33, no. 4, pp. 685–695, DOI: [10.1175/JTECH-D-15-0255.1](https://doi.org/10.1175/JTECH-D-15-0255.1).

- Lyard, F. H., D. J. Allain, M. Cancet, L. Carrère & N. Picot (2021), “FES2014 global ocean tide atlas: Design and performance”, *Ocean Science*, vol. 17, no. 3, pp. 615–649, DOI: [10.5194/os-17-615-2021](https://doi.org/10.5194/os-17-615-2021).
- Macias, D, A Cózar, E Garcia-gorriz, D González-fernández & A Stips (2019), “Surface water circulation develops seasonally changing patterns of floating litter accumulation in the Mediterranean Sea . A modelling approach”, *Marine Pollution Bulletin*, vol. 149, p. 110619, DOI: [10.1016/j.marpolbul.2019.110619](https://doi.org/10.1016/j.marpolbul.2019.110619).
- Macias, D. M., E. Garcia-gorriz & A. Stips (2015), “Productivity changes in the Mediterranean Sea for the twenty-first century in response to changes in the regional atmospheric forcing”, *Frontiers in Marine Science*, vol. 2, p. 79, DOI: [10.3389/fmars.2015.00079](https://doi.org/10.3389/fmars.2015.00079).
- Madec, G. & the NEMO team (2017), *NEMO ocean engine*, DOI: [10.5281/ZENODO.3248739](https://doi.org/10.5281/ZENODO.3248739).
- Martin, C., C. A. Young, L. Valluzzi & C. M. Duarte (2022), “Ocean sediments as the global sink for marine micro- and mesoplastics”, *Limnology and Oceanography Letters*, vol. 7, no. 3, pp. 235–243, DOI: [10.1002/lol2.10257](https://doi.org/10.1002/lol2.10257).
- Martínez-Vicente, V. et al. (2019), “Measuring marine plastic debris from space: Initial assessment of observation requirements”, *Remote Sensing*, vol. 11, no. 20, pp. 8–14, DOI: [10.3390/rs11202443](https://doi.org/10.3390/rs11202443).
- Mateos-Cárdenas, A., J. O'Halloran, F. N. van Pelt & M. A. Jansen (2020), “Rapid fragmentation of microplastics by the freshwater amphipod *Gammarus duebeni* (Lillj.)”, *Scientific Reports*, vol. 10, no. 1, pp. 1–12, DOI: [10.1038/s41598-020-69635-2](https://doi.org/10.1038/s41598-020-69635-2).
- Maximenko, N., J. Hafner & P. Niiler (2012), “Pathways of marine debris derived from trajectories of Lagrangian drifters”, *Marine Pollution Bulletin*, vol. 65, no. 1–3, pp. 51–62, DOI: [10.1016/j.marpolbul.2011.04.016](https://doi.org/10.1016/j.marpolbul.2011.04.016).
- Maximenko, N. et al. (2019), “Toward the Integrated Marine Debris Observing System”, *Frontiers in Marine Science*, vol. 6, p. 447.
- McCann, A. et al. (2019), “Exploratory analysis of covariation of microbiota-derived Vitamin K and cognition in older adults”, *American Journal of Clinical Nutrition*, vol. 110, no. 6, pp. 1404–1415, DOI: [10.1093/ajcn/nqz220](https://doi.org/10.1093/ajcn/nqz220).
- Meijer, L. J., T. van Emmerik, R. van der Ent, C. Schmidt & L. Lebreton (2021), “More than 1000 rivers account for 80% of global riverine plastic emissions into the ocean”, *Science Advances*, vol. 7, no. 18, pp. 1–14, DOI: [10.1126/sciadv.aaz5803](https://doi.org/10.1126/sciadv.aaz5803).
- Menna, M., R. Gerin, A. Bussani & P.-m. Poulain (2017), *The OGS Mediterranean Drifter Dataset: 1986-2016*, tech. rep., DOI: [10.6092/7a8499bc-c5ee-472c-b8b5-03523d1e73e9](https://doi.org/10.6092/7a8499bc-c5ee-472c-b8b5-03523d1e73e9).
- Meyerjürgens, J., T. H. Badewien, S. P. Garaba, J. O. Wolff & O. Zielinski (2019), “A state-of-the-art compact surface drifter reveals pathways of floating marine litter in the German Bight”, *Frontiers in Marine Science*, vol. 6, no. FEB, pp. 1–15, DOI: [10.3389/fmars.2019.00058](https://doi.org/10.3389/fmars.2019.00058).
- van der Molen, J., S. M. van Leeuwen, L. L. Govers, T. van der Heide & H. Olff (2021), “Potential Micro-Plastics Dispersal and Accumulation in the North Sea,

- With Application to the MSC Zoe Incident”, *Frontiers in Marine Science*, vol. 8, no. June, DOI: [10.3389/fmars.2021.607203](https://doi.org/10.3389/fmars.2021.607203).
- Molitor, R., A. Bollinger, S. Kubicki, A. Loeschcke, K. E. Jaeger & S. Thies (2019), “Agar plate-based screening methods for the identification of polyester hydrolysis by *Pseudomonas* species”, *Microbial Biotechnology*, vol. 13, no. 1, pp. 274–284, DOI: [10.1111/1751-7915.13418](https://doi.org/10.1111/1751-7915.13418).
- Moore, C. J., S. L. Moore, M. K. Leecaster & S. B. Weisberg (2001), “A comparison of plastic and plankton in the North Pacific Central Gyre”, *Marine Pollution Bulletin*, vol. 42, no. 12, pp. 1297–1300, DOI: [10.1016/S0025-326X\(01\)00114-X](https://doi.org/10.1016/S0025-326X(01)00114-X).
- Morales-Caselles, C. et al. (2021), “An inshore–offshore sorting system revealed from global classification of ocean litter”, *Nature Sustainability*, vol. 4, no. 6, pp. 484–493, DOI: [10.1038/s41893-021-00720-8](https://doi.org/10.1038/s41893-021-00720-8).
- Morét-Ferguson, S., K. L. Law, G. Proskurowski, E. K. Murphy, E. E. Peacock & C. M. Reddy (2010), “The size, mass, and composition of plastic debris in the western North Atlantic Ocean”, *Marine Pollution Bulletin*, vol. 60, no. 10, pp. 1873–1878, DOI: [10.1016/j.marpolbul.2010.07.020](https://doi.org/10.1016/j.marpolbul.2010.07.020).
- Morris, R. J. (1980), “Plastic debris in the surface waters of the South Atlantic”, *Marine Pollution Bulletin*, vol. 11, no. 6, pp. 164–166, DOI: [10.1016/0025-326X\(80\)90144-7](https://doi.org/10.1016/0025-326X(80)90144-7).
- Mountford, A. S. & M. A. Morales Maqueda (2019), “Eulerian Modeling of the Three-Dimensional Distribution of Seven Popular Microplastic Types in the Global Ocean”, *Journal of Geophysical Research: Oceans*, vol. 124, no. 12, pp. 8558–8573, DOI: [10.1029/2019JC015050](https://doi.org/10.1029/2019JC015050).
- Moy, K. et al. (2018), “Mapping coastal marine debris using aerial imagery and spatial analysis”, *Marine Pollution Bulletin*, {{SI}}: {{Japanese Tsunami Debris}}, vol. 132, pp. 52–59, DOI: [10.1016/j.marpolbul.2017.11.045](https://doi.org/10.1016/j.marpolbul.2017.11.045).
- Nembrini, S., I. R. König & M. N. Wright (2018), “The revival of the Gini importance?”, *Bioinformatics*, vol. 34, no. 21, pp. 3711–3718, DOI: [10.1093/bioinformatics/bty373](https://doi.org/10.1093/bioinformatics/bty373).
- Neumann, D., U. Callies & M. Matthies (2014), “Marine litter ensemble transport simulations in the southern North Sea”, *Marine Pollution Bulletin*, vol. 86, no. 1-2, pp. 219–228, DOI: [10.1016/j.marpolbul.2014.07.016](https://doi.org/10.1016/j.marpolbul.2014.07.016).
- Newman, M. E. (2005), “Power laws, Pareto distributions and Zipf’s law”, *Contemporary Physics*, vol. 46, no. 5, pp. 323–351, DOI: [10.1080/00107510500052444](https://doi.org/10.1080/00107510500052444), arXiv: [0412004](https://arxiv.org/abs/0412004) [cond-mat].
- Newman, S., E. Watkins, A. Farmer, P. ten Brink & J.-P. Schweitzer (2015), “The Economics of Marine Litter”, in: *Marine Anthropogenic Litter*, ed. by Bergmann, M., Gutow, L. & Klages, M., Cham: Springer International Publishing, pp. 367–394, ISBN: 978-3-319-16509-7 978-3-319-16510-3, DOI: [10.1007/978-3-319-16510-3_14](https://doi.org/10.1007/978-3-319-16510-3_14).
- Ogunola, O. S., O. A. Onada & A. E. Falaye (2018), “Mitigation measures to avert the impacts of plastics and microplastics in the marine environment (a review)”, *Environmental Science and Pollution Research*, vol. 25, no. 10, pp. 9293–9310, DOI: [10.1007/s11356-018-1499-z](https://doi.org/10.1007/s11356-018-1499-z).

- Okubo, A. (1971), “Oceanic diffusion diagrams”, *Deep-Sea Research and Oceanographic Abstracts*, vol. 18, no. 8, pp. 789–802, DOI: [10.1016/0011-7471\(71\)90046-5](https://doi.org/10.1016/0011-7471(71)90046-5).
- Olivelli, A., B. D. Hardesty & C. Wilcox (2020), “Coastal margins and backshores represent a major sink for marine debris: insights from a continental-scale analysis”, *Environmental Research Letters*, vol. 15, no. 7, p. 074037, DOI: [10.1088/1748-9326/ab7836](https://doi.org/10.1088/1748-9326/ab7836).
- Onink, V. & C. Laufkötter (2020), “Modelling the Global Distribution of Beaching of Marine Plastic”, in: *Proceedings of the 2nd International Conference on Microplastic Pollution in the Mediterranean Sea*, ed. by Cocca, M. et al., pp. 299–305, DOI: doi.org/10.1007/978-3-030-45909-3_48.
- Onink, V., D. Wichmann, P. Delandmeter & E. van Sebille (2019), “The Role of Ekman Currents, Geostrophy, and Stokes Drift in the Accumulation of Floating Microplastic”, *Journal of Geophysical Research: Oceans*, vol. 124, no. 3, pp. 1474–1490, DOI: [10.1029/2018JC014547](https://doi.org/10.1029/2018JC014547).
- Onink, V., C. E. Jongedijk, M. J. Hoffman, E. van Sebille & C. Laufkötter (2021), “Global simulations of marine plastic transport show plastic trapping in coastal zones”, *Environmental Research Letters*, vol. 16, no. 6, DOI: [10.1088/1748-9326/abecbd](https://doi.org/10.1088/1748-9326/abecbd).
- Onink, V., E. van Sebille & C. Laufkötter (2022a), “Empirical Lagrangian parametrization for wind-driven mixing of buoyant particles at the ocean surface”, *Geoscientific Model Development*, vol. 15, no. 5, pp. 1995–2012, DOI: [10.5194/gmd-15-1995-2022](https://doi.org/10.5194/gmd-15-1995-2022).
- Onink, V., M. L. A. Kaandorp, E. van Sebille & C. Laufkötter (2022b), “Influence of Particle Size and Fragmentation on Large-Scale Microplastic Transport in the Mediterranean Sea”, *Environmental Science & Technology*, vol. 56, no. 22, pp. 15528–15540, DOI: [10.1021/acs.est.2c03363](https://doi.org/10.1021/acs.est.2c03363).
- OSPAR (2010), *Guideline for monitoring marine litter on the beaches in the OSPAR Maritime Area*, tech. rep., p. 84.
- Pawlowicz, R. (2020), “The Grounding of Floating Objects in a Marginal Sea”, *Journal of Physical Oceanography*, pp. 537–551, DOI: [10.1175/jpo-d-20-0183.1](https://doi.org/10.1175/jpo-d-20-0183.1).
- Pazan, S. E. & P. Niiler (2004), “New global drifter data set available”, *Eos, Transactions American Geophysical Union*, vol. 85, no. 2, pp. 17–17, DOI: [10.1029/2004E0020007](https://doi.org/10.1029/2004E0020007).
- Pedregosa, F et al. (2011), “Scikit-learn: Machine Learning in Python”, *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, DOI: [10.5555/1953048.2078195](https://doi.org/10.5555/1953048.2078195).
- Pedrotti, M. L. et al. (2016), “Changes in the floating plastic pollution of the mediterranean sea in relation to the distance to land”, *PLoS ONE*, vol. 11, pp. 1–14, DOI: [10.1371/journal.pone.0161581](https://doi.org/10.1371/journal.pone.0161581).
- Poulain, M. et al. (2019), “Small Microplastics As a Main Contributor to Plastic Mass Balance in the North Atlantic Subtropical Gyre”, *Environmental Science & Technology*, vol. 53, pp. 1157–1164, DOI: [10.1021/acs.est.8b05458](https://doi.org/10.1021/acs.est.8b05458).
- Pugh, D. (1987), “Tides, Surges and mean sea-level”, *Marine and Petroleum Geology*, vol. 5, no. 3, p. 301, DOI: [10.1016/0264-8172\(88\)90013-X](https://doi.org/10.1016/0264-8172(88)90013-X).

- Randall, D. & J. Zehnder (2001), “General Circulation Model Development: Past, Present, and Future. International Geophysics Series, Vol 70”, *Applied Mechanics Reviews*, vol. 54, no. 5, B94–B94, DOI: [10.1115/1.1399682](https://doi.org/10.1115/1.1399682).
- Rech, S, V Macaya-Caquilpán, J. Pantoja, M. Rivadeneira, D Jofre Madariaga & M Thiel (2014), “Rivers as a source of marine litter – A study from the SE Pacific”, *Marine Pollution Bulletin*, vol. 82, pp. 66–75, DOI: [10.1016/j.marpolbul.2014.03.019](https://doi.org/10.1016/j.marpolbul.2014.03.019).
- Reijnders, D., E. Deleersnijder & E. van Sebille (2022), “Simulating Lagrangian Subgrid-Scale Dispersion on Neutral Surfaces in the Ocean”, *Journal of Advances in Modeling Earth Systems*, vol. 14, no. 2, pp. 1–25, DOI: [10.1029/2021MS002850](https://doi.org/10.1029/2021MS002850).
- Reisser, J. et al. (2015), “The vertical distribution of buoyant plastics at sea: An observational study in the North Atlantic Gyre”, *Biogeosciences*, vol. 12, no. 4, pp. 1249–1256, DOI: [10.5194/bg-12-1249-2015](https://doi.org/10.5194/bg-12-1249-2015).
- Reisser, J. et al. (2014), “Millimeter-sized marine plastics: A new pelagic habitat for microorganisms and invertebrates”, *PLoS ONE*, vol. 9, no. 6, pp. 1–11, DOI: [10.1371/journal.pone.0100289](https://doi.org/10.1371/journal.pone.0100289).
- Ribic, C. A., S. B. Sheavly, D. J. Rugg & E. S. Erdmann (2010), “Trends and drivers of marine debris on the Atlantic coast of the United States 1997-2007”, *Marine Pollution Bulletin*, vol. 60, no. 8, pp. 1231–1242, DOI: [10.1016/j.marpolbul.2010.03.021](https://doi.org/10.1016/j.marpolbul.2010.03.021).
- Ribic, C. A., S. B. Sheavly, D. J. Rugg & E. S. Erdmann (2012), “Trends in marine debris along the U.S. Pacific Coast and Hawai’i 1998-2007”, *Marine Pollution Bulletin*, vol. 64, no. 5, pp. 994–1004, DOI: [10.1016/j.marpolbul.2012.02.008](https://doi.org/10.1016/j.marpolbul.2012.02.008).
- Ricker, M. & E. V. Stanev (2020), “Circulation of the European northwest shelf: A Lagrangian perspective”, *Ocean Science*, vol. 16, no. 3, pp. 637–655, DOI: [10.5194/os-16-637-2020](https://doi.org/10.5194/os-16-637-2020).
- Rijnsburger, S., R. P. Flores, J. D. Pietrzak, A. R. Horner-Devine, A. J. Souza & F. Zijl (2021), “The Evolution of Plume Fronts in the Rhine Region of Freshwater Influence”, *Journal of Geophysical Research: Oceans*, vol. 126, no. 7, pp. 1–28, DOI: [10.1029/2019jc015927](https://doi.org/10.1029/2019jc015927).
- Rohatgi, A. (2020), *WebPlotDigitizer version 4.3*.
- Roomen, M. V., G. Keijl, B. Koks & K. Mostert (2008), “Numbers of wintering waders on the North Sea coast of the Netherlands in January 1998”, in: *International Wader Studies. International Wader Study Group, Thetford, UK*, 18, pp. 55–58.
- Rossby, C. G. & R. B. Montgomery (1935), “The Layer of Frictional Influence in Wind and Ocean Currents”, *Papers in Physical Oceanography and Meteorology*, vol. 3, no. 3.
- Rühs, S., V. Zhurbas, I. M. Koszalka, J. V. Durgadoo & A. Biastoch (2017), “Eddy Diffusivity Estimates from Lagrangian Trajectories Simulated with Ocean Models and Surface Drifter Data—A Case Study for the Greater Agulhas System”, *Journal of Physical Oceanography*, vol. 48, no. 1, pp. 175–196, DOI: [10.1175/jpo-d-17-0048.1](https://doi.org/10.1175/jpo-d-17-0048.1).
- de Ruijter, W. P., A. W. Visser & W. G. Bos (1997), “The Rhine outflow: A prototypical pulsed discharge plume in a high energy shallow sea”, *Journal of Marine Systems*, vol. 12, no. 1-4, pp. 263–276, DOI: [10.1016/S0924-7963\(96\)00102-9](https://doi.org/10.1016/S0924-7963(96)00102-9).

- Ruiz-Orejón, L. F., R. Sardá & J. Ramis-Pujol (2016), "Floating plastic debris in the Central and Western Mediterranean Sea", *Marine Environmental Research*, vol. 120, pp. 136–144, DOI: [10.1016/j.marenvres.2016.08.001](https://doi.org/10.1016/j.marenvres.2016.08.001).
- Ruiz-Orejón, L. F., R. Sardá & J. Ramis-Pujol (2018), "Now, you see me: High concentrations of floating plastic debris in the coastal waters of the Balearic Islands (Spain)", *Marine Pollution Bulletin*, vol. 133, no. June, pp. 636–646, DOI: [10.1016/j.marpolbul.2018.06.010](https://doi.org/10.1016/j.marpolbul.2018.06.010).
- Ryan, P. G. (1987), "The effects of ingested plastic on seabirds: Correlations between plastic load and body condition", *Environmental Pollution*, vol. 46, no. 2, pp. 119–125, DOI: [10.1016/0269-7491\(87\)90197-7](https://doi.org/10.1016/0269-7491(87)90197-7).
- Ryan, P. G. (2015), "Does size and buoyancy affect the long-distance transport of floating debris?", *Environmental Research Letters*, vol. 10, no. 8, p. 84019, DOI: [10.1088/1748-9326/10/8/084019](https://doi.org/10.1088/1748-9326/10/8/084019).
- Ryan, P. G., A. Lamprecht, D. Swanepoel & C. L. Moloney (2014), "The effect of fine-scale sampling frequency on estimates of beach litter accumulation", *Marine Pollution Bulletin*, vol. 88, pp. 249–254, DOI: [10.1016/j.marpolbul.2014.08.036](https://doi.org/10.1016/j.marpolbul.2014.08.036).
- Ryan, P. G., V. Perold, A. Osborne & C. L. Moloney (2018), "Consistent patterns of debris on South African beaches indicate that industrial pellets and other mesoplastic items mostly derive from local sources", *Environmental Pollution*, vol. 238, pp. 1008–1016, DOI: [10.1016/j.envpol.2018.02.017](https://doi.org/10.1016/j.envpol.2018.02.017).
- Ryan, P. G., E. A. Weideman, V. Perold & C. L. Moloney (2020), "Toward Balancing the Budget: Surface Macro-Plastics Dominate the Mass of Particulate Pollution Stranded on Beaches", *Frontiers in Marine Science*, vol. 7, p. 575395, DOI: [10.3389/fmars.2020.575395](https://doi.org/10.3389/fmars.2020.575395).
- Salvatier, J., T. V. Wiecki & C. Fonnesbeck (2016), "Probabilistic programming in Python using PyMC3", *PeerJ Computer Science*, vol. 2, e55, DOI: [10.7717/peerj-cs.55](https://doi.org/10.7717/peerj-cs.55).
- Samaras, A. G., M. De Dominicis, R. Archetti, A. Lamberti & N. Pinardi (2014), "Towards improving the representation of beaching in oil spill models: A case study", *Marine Pollution Bulletin*, vol. 88, no. 1-2, pp. 91–101, DOI: [10.1016/j.marpolbul.2014.09.019](https://doi.org/10.1016/j.marpolbul.2014.09.019).
- Schöneich-Argent, R. I., K. Dau & H. Freund (2020), "Wasting the North Sea? – A field-based assessment of anthropogenic macrolitter loads and emission rates of three German tributaries", *Environmental Pollution*, vol. 263, DOI: [10.1016/j.envpol.2020.114367](https://doi.org/10.1016/j.envpol.2020.114367).
- Schulz, K. & L. Umlauf (2016), "Residual transport of suspended material by tidal straining near sloping topography", *Journal of Physical Oceanography*, vol. 46, no. 7, pp. 2083–2102, DOI: [10.1175/JPO-D-15-0218.1](https://doi.org/10.1175/JPO-D-15-0218.1).
- Schulz, M. & M. Matthies (2014), "Artificial neural networks for modeling time series of beach litter in the southern north sea", *Marine Environmental Research*, vol. 98, pp. 14–20, DOI: [10.1016/j.marenvres.2014.03.014](https://doi.org/10.1016/j.marenvres.2014.03.014).
- van Sebille, E. (2014), "Adrift.org.au - A free, quick and easy tool to quantitatively study planktonic surface drift in the global ocean", *Journal of Experimental Marine Biology and Ecology*, vol. 461, pp. 317–322, DOI: [10.1016/j.jembe.2014.09.002](https://doi.org/10.1016/j.jembe.2014.09.002).

- van Sebille, E., M. H. England & G. Froyland (2012), “Origin, dynamics and evolution of ocean garbage patches from observed surface drifters”, *Environmental Research Letters*, vol. 7, no. 4, DOI: [10.1088/1748-9326/7/4/044040](https://doi.org/10.1088/1748-9326/7/4/044040).
- van Sebille, E. et al. (2015), “A global inventory of small floating plastic debris”, *Environmental Research Letters*, vol. 10, no. 12, p. 124006, DOI: [10.1088/1748-9326/10/12/124006](https://doi.org/10.1088/1748-9326/10/12/124006).
- van Sebille, E. et al. (2018), “Lagrangian ocean analysis: Fundamentals and practices”, *Ocean Modelling*, vol. 121, pp. 49–75, DOI: [10.1016/j.ocemod.2017.11.008](https://doi.org/10.1016/j.ocemod.2017.11.008).
- van Sebille, E. et al. (2020), “The physical oceanography of the transport of floating marine debris”, *Environmental Research Letters*, vol. 15, no. 2, DOI: [10.1088/1748-9326/ab6d7d](https://doi.org/10.1088/1748-9326/ab6d7d).
- SEDAC, CIESIN - Center for International Earth Science Information Network - Columbia University, FAO - United Nations Food and Agriculture Programme & CIAT-Centro Internacional de Agricultura Tropical (2005), *Gridded Population of the World, Version 3 (GPWv3): Population Count Grid [Data set]*, DOI: <http://dx.doi.org/10.7927/H4639MPP>.
- Shafer, J. M. & M. D. Varljen (1990), “Approximation of confidence limits on sample semivariograms from single realizations of spatially correlated random fields”, *Water Resources Research*, vol. 26, no. 8, pp. 1787–1802, DOI: [10.1029/WR026i008p01787](https://doi.org/10.1029/WR026i008p01787).
- Shaw, D. G. & G. A. Mapes (1979), “Surface circulation and the distribution of pelagic tar and plastic”, *Marine Pollution Bulletin*, vol. 10, no. 6, pp. 160–162, DOI: [10.1016/0025-326X\(79\)90421-1](https://doi.org/10.1016/0025-326X(79)90421-1).
- Simoncelli, S. et al. (2019), *Mediterranean Sea Physical Reanalysis (CMEMS MED-Physics) [Data set]*, DOI: https://doi.org/10.25423/MEDSEA_REANALYSIS_PHYS_006_004.
- Song, K., J. Y. Jung, S. H. Lee & S. Park (2021), “A comparative study of deep learning-based network model and conventional method to assess beach debris standing-stock”, *Marine Pollution Bulletin*, vol. 168, p. 112466, DOI: [10.1016/j.marpolbul.2021.112466](https://doi.org/10.1016/j.marpolbul.2021.112466).
- Song, Y. K., S. H. Hong, M. Jang, G. M. Han, S. W. Jung & W. J. Shim (2017), “Combined Effects of UV Exposure Duration and Mechanical Abrasion on Microplastic Fragmentation by Polymer Type”, *Environmental Science and Technology*, vol. 51, no. 8, pp. 4368–4376, DOI: [10.1021/acs.est.6b06155](https://doi.org/10.1021/acs.est.6b06155).
- Sorasan, C. et al. (2022), “Ageing and fragmentation of marine microplastics”, *Science of the Total Environment*, vol. 827, DOI: [10.1016/j.scitotenv.2022.154438](https://doi.org/10.1016/j.scitotenv.2022.154438).
- Stambler, N. (2014), “The Mediterranean Sea – Primary Productivity”, in: *The Mediterranean Sea: Its history and present challenges*, ed. by Goffredo, S. & Dubinsky, Z., Dordrecht: Springer Netherlands, pp. 113–121, ISBN: 978-94-007-6704-1, DOI: [10.1007/978-94-007-6704-1_7](https://doi.org/10.1007/978-94-007-6704-1_7).
- Sterl, M. F., P. Delandmeter & E. Sebille (2020), “Influence of Barotropic Tidal Currents on Transport and Accumulation of Floating Microplastics in the Global Open Ocean”, *Journal of Geophysical Research: Oceans*, vol. 125, no. 2, DOI: [10.1029/2019JC015583](https://doi.org/10.1029/2019JC015583).

- Suaria, G. et al. (2016), “The Mediterranean Plastic Soup: Synthetic polymers in Mediterranean surface waters”, *Scientific Reports*, vol. 6, pp. 1–10, DOI: [10.1038/srep37551](https://doi.org/10.1038/srep37551).
- Tarantola, A. (2005), *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM, ISBN: 0898715725.
- Thepwilai, S., K. Wangritthikraikul, S. Chawchai & R. Bissen (2021), “Testing the factors controlling the numbers of microplastics on beaches along the western Gulf of Thailand”, *Marine Pollution Bulletin*, vol. 168, p. 112467, DOI: [10.1016/j.marpolbul.2021.112467](https://doi.org/10.1016/j.marpolbul.2021.112467).
- Thompson, R. C. et al. (2004), “Lost at Sea: Where Is All the Plastic?”, *Science*, vol. 304, no. 5672, p. 838, DOI: [10.1126/science.1094559](https://doi.org/10.1126/science.1094559).
- Thorpe, S. A., T. R. Osborn, D. M. Farmer & S. Vagle (2003), “Bubble Clouds and Langmuir Circulation : Observations and Models”, *Journal of Physical Oceanography*, vol. 33, no. 9, pp. 2013–2031.
- Tikhonov, A. N. (1963), “On the solution of ill-posed problems and the method of regularization”, *Dokl. Akad. Nauk SSSR*, vol. 151, no. 3, pp. 501–504.
- Tokai, T., K. Uchida, M. Kuroda & A. Isobe (2021), “Mesh selectivity of neuston nets for microplastics”, *Marine Pollution Bulletin*, vol. 165, no. October 2020, p. 112111, DOI: [10.1016/j.marpolbul.2021.112111](https://doi.org/10.1016/j.marpolbul.2021.112111).
- Tolman, H. L. et al. (2014), “User manual and system documentation of WAVEWATCH III R version 4.18”.
- Turcotte, D. L. (1986), “Fractals and fragmentation”, *Journal of Geophysical Research*, vol. 91, no. B2, pp. 1921–1926, DOI: [10.1029/JB091iB02p01921](https://doi.org/10.1029/JB091iB02p01921).
- Turner, A. & L. Holmes (2011), “Occurrence, distribution and characteristics of beached plastic production pellets on the island of Malta (central Mediterranean)”, *Marine Pollution Bulletin*, vol. 62, no. 2, pp. 377–381, DOI: [10.1016/j.marpolbul.2010.09.027](https://doi.org/10.1016/j.marpolbul.2010.09.027).
- Vianello, A. et al. (2013), “Microplastic particles in sediments of Lagoon of Venice, Italy: First observations on occurrence, spatial patterns and identification”, *Estuarine, Coastal and Shelf Science*, vol. 130, pp. 54–61, DOI: [10.1016/j.ecss.2013.03.022](https://doi.org/10.1016/j.ecss.2013.03.022).
- Virkar, Y. & A. Clauset (2014), “Power-law distributions in binned empirical data”, *Annals of Applied Statistics*, vol. 8, no. 1, pp. 89–119, DOI: [10.1214/13-AOAS710](https://doi.org/10.1214/13-AOAS710), arXiv: [1208.3524](https://arxiv.org/abs/1208.3524).
- Vriend, P., C. van Calcar, M. Kooi, H. Landman, R. Pikaar & T. van Emmerik (2020), “Rapid Assessment of Floating Macroplastic Transport in the Rhine”, *Frontiers in Marine Science*, vol. 7, no. January, pp. 1–8, DOI: [10.3389/fmars.2020.00010](https://doi.org/10.3389/fmars.2020.00010).
- de Vries, R., M. Egger, T. Mani & L. Lebreton (2021), “Quantifying floating plastic debris at sea using vessel-based optical data and artificial intelligence”, *Remote Sensing*, vol. 13, no. 17, pp. 1–16, DOI: [10.3390/rs13173401](https://doi.org/10.3390/rs13173401).
- Waldschläger, K. & H. Schüttrumpf (2019), “Effects of Particle Properties on the Settling and Rise Velocities of Microplastics in Freshwater under Laboratory Con-

- ditions”, *Environmental Science and Technology*, vol. 53, no. 4, pp. 1958–1966, DOI: [10.1021/acs.est.8b06794](https://doi.org/10.1021/acs.est.8b06794).
- Ward, C. P., C. J. Armstrong, A. N. Walsh, J. H. Jackson & C. M. Reddy (2019), “Sunlight Converts Polystyrene to Carbon Dioxide and Dissolved Organic Carbon”, *Environmental Science and Technology Letters*, vol. 6, pp. 669–674, DOI: [10.1021/acs.estlett.9b00532](https://doi.org/10.1021/acs.estlett.9b00532).
- Weideman, E. A., V. Perold, A. Ouardien, L. K. Smyth & P. G. Ryan (2020), “Quantifying temporal trends in anthropogenic litter in a rocky intertidal habitat”, *Marine Pollution Bulletin*, vol. 160, p. 111543, DOI: [10.1016/j.marpolbul.2020.111543](https://doi.org/10.1016/j.marpolbul.2020.111543).
- Weiss, L. et al. (2021), “The missing ocean plastic sink : Gone with the rivers”, *Science*, vol. 111, no. July, pp. 107–111.
- Wichmann, D. (2021), “Large-scale structure extraction in Lagrangian ocean trajectories”, PhD thesis, Utrecht University, ISBN: 9789463327367.
- Wichmann, D., P. Delandmeter & E. van Sebille (2019a), “Influence of Near-Surface Currents on the Global Dispersal of Marine Microplastic”, *Journal of Geophysical Research: Oceans*, vol. 124, no. 8, pp. 6086–6096, DOI: [10.1029/2019JC015328](https://doi.org/10.1029/2019JC015328).
- Wichmann, D., P. Delandmeter, H. A. Dijkstra & E. van Sebille (2019b), “Mixing of passive tracers at the ocean surface and its implications for plastic transport modelling”, *Environmental Research Communications*, vol. 1, no. 11, DOI: [10.1088/2515-7620/ab4e77](https://doi.org/10.1088/2515-7620/ab4e77).
- Wilcox, C., B. D. Hardesty & K. L. Law (2020), “Abundance of Floating Plastic Particles Is Increasing in the Western North Atlantic Ocean”, *Environmental Science and Technology*, vol. 54, no. 2, pp. 790–796, DOI: [10.1021/acs.est.9b04812](https://doi.org/10.1021/acs.est.9b04812).
- Williams, A. T. & D. T. Tudor (2001), “Litter burial and exhumation: Spatial and temporal distribution on a cobble pocket beach”, *Marine Pollution Bulletin*, vol. 42, no. 11, pp. 1031–1039, DOI: [10.1016/S0025-326X\(01\)00058-3](https://doi.org/10.1016/S0025-326X(01)00058-3).
- Wong, C. S., D. R. Green & W. J. Cretney (1974), “Quantitative Tar and Plastic Waste Distributions in the Pacific Ocean”, *Nature*, vol. 247, no. 5435, pp. 30–32, DOI: [10.1038/247030a0](https://doi.org/10.1038/247030a0).
- Woodall, L. C. et al. (2014), “The deep sea is a major sink for microplastic debris”, *Royal Society Open Science*, vol. 1, no. 140317, pp. 1–8.
- Ye, S. & A. L. Andrady (1991), “Fouling of floating plastic debris under Biscayne Bay exposure conditions”, *Marine Pollution Bulletin*, vol. 22, no. 12, pp. 608–613, DOI: [10.1016/0025-326X\(91\)90249-R](https://doi.org/10.1016/0025-326X(91)90249-R).
- Young, L. C., C. Vanderlip, D. C. Duffy, V. Afanasyev & S. A. Shaffer (2009), “Bringing home the trash: Do colony-based differences in foraging distribution lead to increased plastic ingestion in Laysan albatrosses?”, *PLoS ONE*, vol. 4, no. 10, pp. 11–13, DOI: [10.1371/journal.pone.0007623](https://doi.org/10.1371/journal.pone.0007623).
- Zeri, C. et al. (2018), “Floating plastics in Adriatic waters (Mediterranean Sea): From the macro- to the micro-scale”, *Marine Pollution Bulletin*, vol. 136, no. September, pp. 341–350, DOI: [10.1016/j.marpolbul.2018.09.016](https://doi.org/10.1016/j.marpolbul.2018.09.016).
- Zettler, E. R., H. Takada, B. Monteleone, N. Mallos, M. Eriksen & L. A. Amaral-Zettler (2017), “Incorporating citizen science to study plastics in the environment”, *Analytical Methods*, vol. 9, no. 9, pp. 1392–1403, DOI: [10.1039/c6ay02716d](https://doi.org/10.1039/c6ay02716d).

- Zhao, S., E. R. Zettler, R. P. Bos, P. Lin, L. A. Amaral-Zettler & T. J. Mincer **(2022)**, “Large quantities of small microplastics permeate the surface ocean to abyssal depths in the South Atlantic Gyre”, *Global Change Biology*, vol. 28, no. 9, pp. 2991–3006, DOI: [10.1111/gcb.16089](https://doi.org/10.1111/gcb.16089).
- Zhu, L., S. Zhao, T. B. Bittar, A. Stubbins & D. Li **(2020)**, “Photochemical dissolution of buoyant microplastics to dissolved organic carbon: Rates and microbial impacts”, *Journal of Hazardous Materials*, vol. 383, no. August 2019, p. 121065, DOI: [10.1016/j.jhazmat.2019.121065](https://doi.org/10.1016/j.jhazmat.2019.121065).

Acknowledgments

Looking back at the past four years, so much happened and yet the time flew by. It seems like it was only yesterday that I came to the IMAU to do the interview for the PhD position with Erik and Henk. I would like to thank both for their encouragement and positive attitudes these past years. I really appreciate how Erik created and managed such a stimulating environment to work in, with so many amazing colleagues. Erik, your guidance has been wonderful, with lots of freedom to explore topics I was particularly interested in, and trusting me with so many great opportunities besides just the ‘main’ research. Already at the start of my PhD you gave me the opportunity to take part in workshops on plastic pollution both in the Netherlands and abroad, encouraged me to take part in programs such as *Faces of Science*, and let me join a scientific cruise to the middle of the Atlantic Ocean. Henk, I would like to thank you for showing so much interest in my research these past years, and your valuable feedback to the chapters of this thesis.

I would like to thank the parcels group for being such an amazing team, and for all the great discussions. When I started in 2018 it was still a relatively small group. Thank you Philippe, Peter, and David for welcoming me with open arms on my first day, it was always a joy to step into your office. Thank you Philippe for your help with technical questions and your enthusiasm in explaining and improving the parcels code. Thank you Peter for a great EGU, and your infectious enthusiasm for anything sports-related. Victor and Cleo, it was great having discussions with you both, I really enjoyed attending my first conference ever with you on Lanzarote. Thank you Delphine, it was a joy to work together with you on the global project in chapter 5, I admire your ambition to have a positive impact on the ocean. Additionally, thank you Delphine and Cleo for the great trip after OSM in San Diego, it was an awesome way to end the pre-COVID era. Thank you Claudio, I really enjoyed our gym sessions, intense table tennis matches, and teaching together. Darshika, thank you for teaching me the ways of Indian cooking, it was great fun. Thank you Mike for the fun conversations on floating stuff and coffee. Thank you Christian, Daan, Laura, Steffie, Reint, Siren, and all other parcels team members for everything, a special thanks for being there during the strange COVID times and keeping some structure in daily life.

Thank you Floor, Sandra, Clara, Judith, Mariken for your organisational support.

Marcel and Michael, thank you for your support regarding everything IT related. Thank you André and Claudia for being a great officemates. I would like to thank all colleagues at the IMAU and in the ocean group for making it such a great place to work, including Michiel, René, Janneke, Arthur, Amber, Anneke, Getachew, Leo, Hossein, Jinyang, Tjebbe, Daniele, Maurice, Julia, Hanne, Sacha, Valérian, Daniel, Simon.

Thank you Aaron for giving me the opportunity to spend a very memorable month on the Atlantic. Jannes, Vincent, Gabriella, Louise, Stephan, Sarah: I couldn't have wished for better company while there. Femke, thank you so much for adopting me into your data assimilation group, and a thank you to all its members for the interesting presentations and discussions that we had.

Thank you to my parents and sister, Jaap, Sarita, and Sara for always being there. Thank you Truus, Jaap, Ingeborg, and all the family in Finland. A big thanks to my Casticum friends for the good times: Egbert, Luc, Dennis, Bastiaan, Coen, Gijs, Rick. A special thanks to Egbert and Dennis for the great relaxing surf sessions: spending time in the sea is even better than just studying it. Thomas and Frederik, thanks for the awesome Teksel and Christmas market trips, and for the thesis cover tips.

Finally, I would like to thank the mud from the bottom of the Atlantic: it takes a long time to sieve you for plastics and other hidden treasures. Enough time to get to know everything about the amazing Rebecka. Rebecka, I am so glad the universe brought us together, you are the best and I can't wait for all the fun times we will have together.

List of Publications

Published

1. Kaandorp, M. L., S. L. Ypma, M. Boonstra, H. A. Dijkstra & E. Van Sebille (2022), “Using machine learning and beach cleanup data to explain litter quantities along the Dutch North Sea coast”, *Ocean Science*, vol. 18, no. 1, pp. 269–293, DOI: [10.5194/os-18-269-2022](https://doi.org/10.5194/os-18-269-2022)
2. van Duinen, B., M. L. Kaandorp & E. van Sebille (2022), “Identifying Marine Sources of Beached Plastics Through a Bayesian Framework: Application to Southwest Netherlands”, *Geophysical Research Letters*, vol. 49, no. 4, pp. 1–9, DOI: [10.1029/2021GL097214](https://doi.org/10.1029/2021GL097214)
3. Onink, V., M. L. A. Kaandorp, E. van Sebille & C. Laufkötter (2022b), “Influence of Particle Size and Fragmentation on Large-Scale Microplastic Transport in the Mediterranean Sea”, *Environmental Science & Technology*, vol. 56, no. 22, pp. 15528–15540, DOI: [10.1021/acs.est.2c03363](https://doi.org/10.1021/acs.est.2c03363)
4. Kaandorp, M. L. A., H. A. Dijkstra & E. van Sebille (2021), “Modelling size distributions of marine plastics under the influence of continuous cascading fragmentation”, *Environmental Research Letters*, vol. 16, no. 5, p. 54075, DOI: [10.1088/1748-9326/abe9ea](https://doi.org/10.1088/1748-9326/abe9ea), arXiv: [2011.01775](https://arxiv.org/abs/2011.01775)
5. Kaandorp, M. L. A., H. A. Dijkstra & E. van Sebille (2020), “Closing the Mediterranean Marine Floating Plastic Mass Budget: Inverse Modeling of Sources and Sinks”, *Environmental Science & Technology*, vol. 54, no. 19, pp. 11980–11989, DOI: [10.1021/acs.est.0c01984](https://doi.org/10.1021/acs.est.0c01984)
6. van Sebille, E., S. Aliani, K. L. Law, N. Maximenko, J. M. Alsina, A. Bagaev, et al. (2020), “The physical oceanography of the transport of floating marine debris”, *Environmental Research Letters*, vol. 15, no. 2, DOI: [10.1088/1748-9326/ab6d7d](https://doi.org/10.1088/1748-9326/ab6d7d)
7. Kaandorp, M. L. & R. P. Dwight (2020), “Data-driven modelling of the Reynolds stress tensor using random forests with invariance”, *Computers & Fluids*, vol. 202, p. 104497, DOI: [10.1016/j.compfluid.2020.104497](https://doi.org/10.1016/j.compfluid.2020.104497)
8. Kaandorp, M. L., S. Menzel & S. Schmitt (2017), “An Aerodynamic Perspective on Shape Deformation Methods”, *18th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, DOI: [10.2514/6.2017-3145](https://doi.org/10.2514/6.2017-3145)

In Preparation

1. Kaandorp, M. L., D. Lobelle, C. Kehl, H. A. Dijkstra & E. van Sebille, “The global mass of buoyant marine plastics is dominated by large long-lived debris”
2. Beck, A., M. L. Kaandorp, T. Hamm, B. Bogner, E. Kossel, M. Lenz, et al., “Rapid shipboard measurement of net-collected marine microplastic polymer types using near-infrared hyperspectral imaging”
3. Kehl, C., P. D. Nooteboom, M. L. A. Kaandorp & E. V. Sebille, “Efficiently Simulating Lagrangian Particles in Large-Scale Ocean Flows – Data Structures and their Impact on Geophysical Applications”
4. Lobelle, D., L. Shen, B. van Huet, T. van Emmerik, M. Kaandorp, G. Iattoni, et al., “Mapping Managed and Mismanaged Dutch Plastic Waste Flows”

About the Author

Mikael Kaandorp was born on the 24th of April 1993, in Amsterdam. He grew up in Castricum, close to the North Sea. After finishing high-school in 2011 at the Bonhoeffer College Castricum, he did his Bachelor and Master studies in aerospace engineering at the Delft University of Technology. For his Master's he specialized in aerodynamics, graduating cum laude. As part of his Master's he did an internship in 2016 at the Honda Research Insitute Europe in Germany on the optimization of aerodynamic shapes, resulting in his first scientific publication. This internship sparked his interest in optimization and machine learning methods. This resulted in a Master thesis and scientific publication on data-driven turbulence models for aerodynamic flows.

His interest in combining numerical models for fluid dynamics with data-driven techniques continued at the Utrecht University, where he started his PhD at the physical oceanography group in 2018 under supervision of Prof. dr. ir. Henk Dijkstra and Prof. dr. Erik van Sebille. During this PhD, Mikael used data assimilation and machine learning methods to estimate the sources, pathways, and sinks of marine plastic pollution. He was selected as one of the *Faces of Science* (an initiative of KNAW, De Jonge Akademie, and NEMO Kennislink) in 2021, and his research on plastic pollution on Dutch beaches was covered by various media such as *het Parool* and the radio program *Vroege Vogels*. Mikael currently continues to work on his fascination for data-driven numerical models, applying data assimilation to coupled atmospheric/land surface/subsurface models at the Forschungszentrum Jülich, in Germany.

