UTRECHT UNIVERSITY
CLIMATE PHYSICS MASTER THESIS

Bayesian inference of plastic sources

by back-tracking virtual plastic particles in the Black-Sea

Supervisors:

Dr. Erik van Sebille
Dr. Delphine Lobelle

Second Examiner:

Dr. Henk Dijkstra

Utrecht University, July 2020

Utrecht University

Acknowledgements

I would like to extend my gratitude to the people that helped me all along this thesis. More specifically, to my supervisors, Dr. Erik van Sebille and Dr. Delphine Lobelle for their availability, support, and always insightful ideas and feedback. I also would like to thank the members of the *Ocean Parcels* team in a more general way, for the unbroken cheerful mood within the group but also for all the advises many of you gave me.

I would also like to thank Mattia Egloff for his tips on statistics, and in particular on the Bayes' theorem, but also Samuele Mercan for always encouraging me and helping me on many aspects throughout the last months.

Most importantly, I would like to express my gratefulness again to Dr. Erik van Sebille, for giving me the occasion to do my thesis reasearch on a subject I was really eager to work on and from which I learned a lot.

Utrecht, the 1st of Auguts 2020 D. B.

Abstract

In this thesis project, the baseline of new method has been developed to back-track the sources of plastic marine litter by coupling Bayesian inference with Lagrangian simulations of virtual particles tracking, for the Black Sea region. It has been concluded that the use of Bayesian statistics provides convincing results that can only be upgraded by the addition of new data. However, this approach still needs more thorough validation, through the use of observational data, to confirm its accuracy. In addition, the efficiency of this approach is limited by the quality of the prior knowledge and information about the studied domain.

Specifically to the Black Sea, when only considering the largest rivers of the basin as source of marine litter, it has been found that the Danube is the main contributor of plastic pollution in most of the zones of the Black Sea. In addition, the entropy of mixing has been calculated in order to understand over which timescales the sources of plastic could be inferred. For the open sea, the sources can be back-tracked over a timescale up to five years. After this period, all the particles are beached, and hence cannot be back-tracked anymore. This is mainly due to the northerly winds and the induced Stokes drift that drives the majority of the particles towards the Southern region of the Black Sea. Thus, if the plastic particle is located on the Southern beaches or along the corresponding coastal areas, its source can be inferred over a maximum timescale of 2 years.

Contents

List of Figures

List of Tables

I Introduction

Plastic accumulation is widely recognized as one of the largest pollutants of contemporary society. It can be found in any environmental reservoir: in the atmosphere, the biota, the terrestrial and marine habitat. Because it is everywhere and since plastic has a very slow degradation rate, it might be used, in the future, as a stratigraphic indicator of the Anthropocene (Zalasiewicz et al., 2016).

Despite this widespread in the environment, the global plastic production keeps rising: it passed from 348 to 360 millions tons from 2017 to 2018, which the largest contribution results from the production of packaging goods (PlasticsEurope, 2019). These types of products have a very short lifespan (<1 year) (Geyer et al., 2017) because they are mostly single-use items and thus almost directly disposed of (Jambeck et al., 2015). Hence, they considerably contribute to the circa 4 to 12 metric tons of plastic introduced in the oceans every year from terrestrial waste (Geyer et al., 2017; Hardesty et al., 2017).

In the oceans, plastic can be found everywhere, from the surface to the seafloor, the coastlines to the the open ocean (van Sebille et al., 2015). This extensive distribution is in part due to its longevity and its buoyancy (van Sebille et al., 2020). The longevity of plastic-based materials is associated to their principal component, fossil hydrocarbons (e.g. polyethylene or polypropylene), that makes them non-, or hardly, biodegradable (Geyer et al., 2017). In fact, fossil hydrocarbons are characterized by an high molecular weight and the presence of long and strong molecular bonds (Zheng et al., 2005), that lead to minimal biological degradation. Plastic preferentially shatters into smaller and smaller pieces, by diverse fragmentation mechanisms, such as the exposure to UV radiations, and it therefore can travel very long distances (van Sebille et al., 2015; O'Brine and Thompson, 2010). For example, plastic fragments have been sampled in some regions quite isolated from anthropogenic pressure, namely the Southern Ocean (Suaria et al., 2020) and the Arctic (Barrows et al., 2018).

Plastic can have different entry points into the marine environment, such as atmospheric fall out (Dris et al., 2016), beach littering, sewage outflows, shipping, fishing, rivers runoffs (Lebreton et al., 2017). Hence, it is necessary to understand the sources, the pathways, and the accumulation zones of plastic, in order to avoid its entry in the oceans (UNEP, 2016). Unfortunately, observations *in situ* are not always possible, since they require long periods of sampling and can be expensive (Hardesty et al., 2017). Furthermore, they are not always reliable, due to human error or biases caused by different instruments or methods used (Barrows et al., 2018). For example, in the case of the Black Sea region, a few observational studies have been made, the majority of which are beach surveys, and a couple of surface/seafloor/water column trawling experiments (Anton et al., 2013; Ioakeimidis et al., 2014; Topçu and Öztürk, 2010). Virtual particle modelling becomes then an important supplementary tool, to fill the gap of the general knowledge on their pathway.

The Black Sea is a semi-enclosed basin located in Eastern Europe and bordered by Turkey, Romania, Bulgaria, Ukraine, Russia, and Georgia (Fig.1.1). Because of the high anthropogenic pressure it is subject to, the Black Sea is considered as one of the most degraded ecosystem in the world (Ötzekin and Bat, 2017), degradation that is in part due to the presence of high levels of plastic. In Kershaw and Rochman (2015), it was estimated that the Black Sea contained between 425 to 900 gkm^{-2} of macro-plastics and between 20,000 to 93,000 items of micro-plastics per km^2, which is 400 times higher than in regions with the lowest value of micro-plastic concentration. Despite these estimations, as mentioned above, it is not a very well studied area. In fact, besides the small number of observational surveys, Stanev and Ricker (2019) is the only numerical study of plastic particle tracking for this region in the literature to date. In their study, a Lagrangian model is used with the purpose to detect accumulation patches of plastic, as well as to simulate its pathways when it originates from eight of the largest rivers of the region : the Danube, Dniester, Dniepr/Southern Bug, Rioni, Coruh, Sakarya, and Kizilirmak rivers.

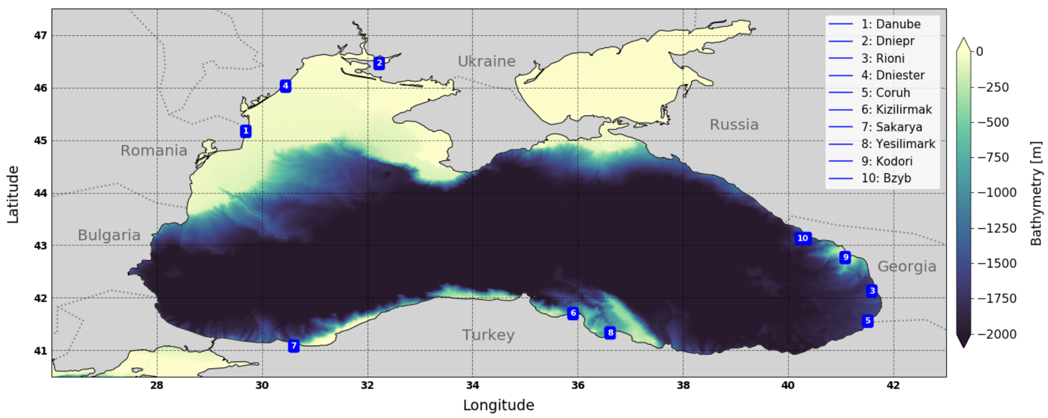


Fig. 1.1: Bathymetric map of the Black Sea. In blue, the main rivers studied in this project: Danube, Dniester, Dniepr, Rioni, Coruh, Sakarya, Kizilirmak, Kodori, Bzyb. (NB: in Stanev and Ricker (2019), cited in the introduction, the Kodori and the Bzyb are not taken in account, while for this project they are.)

Although virtual particle tracking is becoming more and more common in the marine litter context, there is not yet a fully reliable way to backtrack them to their source. The main limit for numerical simulations is that they can not simply be computed backward in time from the plastic sampling location to find their origin. In fact the propagation of particles in the ocean is largely influenced by random walking-like diffusion which, given its intrinsic stochasticity, cannot be reproduced. Moreover, once the mixing time ratio of the particles advected in the domain is reached, it becomes almost impossible to backtrack them (Wichmann et al., 2019). Nevertheless, the hindcast/forcast approach has showed some satisfying results (Carlson et al., 2017; Kako et al., 2011; Isobe et al., 2009). This method consists in advecting particles backward in time from their sampling location in order to single out a number of probable sources, and then advecting them forward in time from these selected entry-points, resulting in a list of more likely sources (Isobe et al., 2009). Although, the results obtained might be satisfactory, this approach encounters some restrictions: it seems not to work so well if extreme meteorological events, such as typhoons, took place (Kako et al., 2011); and the probable sources must be close (max two times the standard deviation of the particle distance from their averaged position in a certain direction) to the particle sampling location (Isobe et al., 2009). The latter limitation is the more constraining one, and a motivation to develop a new method.

The aim of this study is to focus on using Bayesian statistics in order to better understand the pathways of plastics in the ocean. More specifically, to understand the pathways from their sources to their sampling locations. Hence, the purpose of this project is to know if Bayesian inference can be used to back-track plastic sources. In addition, due to the limited observational studies that have been done in the Black Sea, this virtual simulation study would largely benefit the scientific community by having a first theoretical estimation of plastic distribution and their sources in the Black Sea.

II | Theory

Marine Litter

The UNEP (United Nations Environment Program) defines marine litter as "*any persistent, manufactured or processed solid material discarded, disposed of or abandoned in the marine and coastal environment*" (UNEP, 2020). Plastic has been estimated to be by far the most abundant marine litter (Pham et al., 2014). For instance, in the Black Sea, twelve studies out of thirteen, reported plastic to be the most abundant sampled material (Appendix A). Despite international laws on marine litter, such as MARPOL (International Convention for the Prevention of Pollution from Ships), up to 10% of the production of the worldwide manufactured plastic ends up in the oceans (Barnes et al., 2009; Cole et al., 2011).

Marine litter gets in the marine environment through several entry points: they can be sea-based or land-based, with the latter being reported as the largest contributor (Hardesty et al., 2017). However, before investigating these pathways, it is necessary to identify the size of the studied plastic fragments, since it influences their behaviour. Large fragments (i.e., mega-/macro-/meso-plastic), are most often directly discharged in the oceans, voluntarily or involuntarily, through shipping waste, navigation accidents or fishing activities. If these fragments are land-based, they are usually the result of abandoned objects, such as food packaging or cigarette butts on the beach, poorly managed landfills, improper management of city sewage plants, or they can be wash off from the streets by storms UNEP (2016). The smaller fragments, microplastics, are usually divided in two groups: primary and secondary. Secondary microplastics are the result of the fragmentation of larger pieces already present in the marine environment. Primary microplastics, in contrast, are discarded as such in the oceans. For instance, microplastics present in cosmetics, or generated by urban infrastructures, such as nurdles, are considered as primary plastics (UNEP, 2016). Another sub-group of microplastics are fibers, which can enter the marine environment via domestic water (e.g., washing machines) (Cesa et al., 2017) or from indoor cleaning water as a result of indoor fallout (Dris et al., 2016). Even though a large range of plastic litter size is found in the marine environment, Barnes et al. (2009) suggested that their size will tend to decrease over time and hence the total number of plastic concentration would increase (van Sebille et al., 2020).

Plastic litter size, in addition, influences the type of environmental impacts it will have. For instance, microplastics can be ingested by some marine organisms as the particles have the same size range as plankton (UNEP, 2016) and thus bio-accumulate in the food chain, which may lead to the introduction of toxins at the base of the food chain (Maximenko et al., 2019; Barnes et al., 2009). They can also be a scattering vector of chemical additives or represent an habitat for other organisms which may result in the introduction of new species in other habitats, since plastic can travel over long distances (Suaria et al., 2016; Ioakeimidis et al., 2014; Pham et al., 2014). Larger plastic litter can lead to animals' entanglement (UNEP, 2016), be a threat to marine navigation (Hong et al., 2017), but also have a negative impact on the economy (Maximenko et al., 2019; UNEP, 2016). The leakage of large plastic fragments in touristic areas contributes to economic losses due to the landscape deterioration, leading to a diminution of the touristic activity and cleaning costs. For instance, the removal of 75% of marine debris in several Californian beaches in the United States resulted in a $40 million benefit, while an increase in marine litter on the beaches of Goeje Island in South Korea led to more than a $20 million loss (Maximenko et al., 2019).

The Black Sea

The Black Sea (Fig.1.1) is a semi-enclosed basin (Özsoy and Ünlüata, 1997), which makes it by definition more vulnerable to anthropogenic pressure. Indeed, because they are naturally favorable for human societies - because they generally provide water, food and economical expansion -, semi-enclosed basins are usually correlated with dense population centers, and are thus highly polluted (Healy and Harada, 1991). Furthermore, due to the limited outflow of these basins versus the large pollution inputs, this pollution tends to increase and accumulate. Following this general description, the Black Sea might be one of the most degraded ecosystems in the world, as a consequence of highly polluted rivers discharges, several industrial cities along its coastlines, intense shipping routes and fishing activities, all of which contribute to the release of large amounts of plastic litter and other pollutants (Ötzekin and Bat, 2017). Specifically to the Black Sea, the precipitation rate is almost equal to the evaporation rate, suggesting that the net water outflow, through the Bosphorus Strait, is balanced by the net inflow which comes largely from rivers discharge (Stanev and Ricker, 2019). The concentration of marine litter being larger in the rivers mouth than in the Bosphorus Strait, and since the water fluxes are equivalent, it implies that the marine litter concentration will, as expected, increase (Stanev and Ricker, 2019).

The main dynamical feature of the Black Sea its permanent cyclonic current, the so-called Rim current (Fig. 2.1), which flows much faster in the upper layer (max 100 cm/s) than in the sub-pycnocline layer (max 40 cm/s) (Oguz and Besiktepe, 1999). Another unique aspect of the basin, is its quasi-permanent halocline (i.e., the salinity vertical gradient) between 100-150m (Bat, 2017; Capet et al., 2012), separating the saltier, and warmer, deep waters from the less salty, colder surface waters (Özsoy and Ünlüata, 1997). This is due to the combination of two factors: the rapid and narrow Rim current which produces a dynamical barrier between the deep and surface waters, and the large freshwater inflow coming from the rivers that, as a consequence of the Rim current, remains at the surface (Kubryakov et al., 2016). The

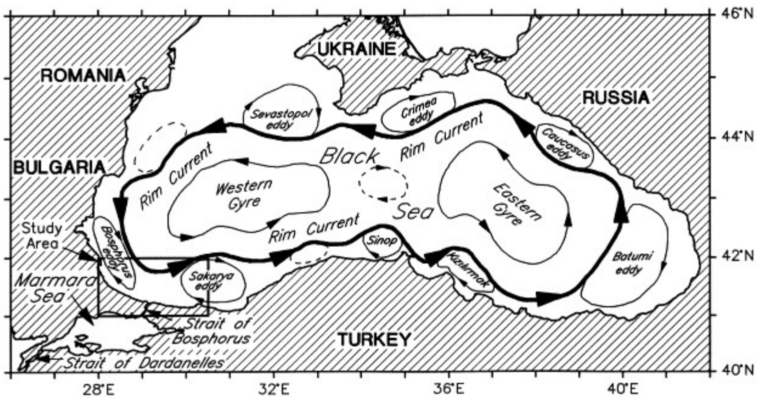


Fig. 2.1: Schematic of the main features of the Black Sea circulation, from Birkun et al. (2009).

basin interior, isolated by the Rim current, is the saltiest zone with 19 g/kg, while the North-Western part, where the largest freshwater discharges occur, is the least salty area of 14 g/kg (Bat, 2017; Sezgin et al., 2017). To illustrate this feature, it has been demonstrated that nutrients and plankton scatter over the whole basin surface with almost no sinking into deeper waters (Kubryakov et al., 2016). However, during a weaker Rim current period, which usually occurs in summer (Staneva et al., 2001), the dynamical barrier is less intense and baroclinic instabilities arise (Kubryakov et al., 2016). Those derive from the strong vertical stratification of the Black Sea basin. The combination of the steep bottom between the Rim current and the coasts and the weakening of the Rim current in summer favor vertical mixing which, due to the strong stratification, leads to baroclinic instabilities. Near the coast, these instabilities lead to mesoscale anticyclones (Fig. 2.1), and hence to even more mixing (Staneva et al., 2001). In short, the seasonal cycle of the Black Sea circulation is as follows: in winter the Rim current is strong and two quasi-persistent cyclonic eddies are noticeable (i.e., Sevastopol and Batumi in Fig.2.1), while during summer/autumn the Rim current is hardly present and several anticyclones appear (Staneva et al., 2001; Bat, 2017). Finally, the surface circulation is mainly influenced by the northerly winds and the resulting Stoke drift (Stanev and Ricker, 2019; Geyer et al., 2017).

The North-Western zone of the Black Sea is quite distinctive from the rest of the basin. It consists of the continental shelf and represents 25% of the total seafloor. The shallow water of that area is separated from the deep waters by a steep slope, passing from -10m to -1,000m quite abruptly (Fig.1.1) (Özsoy and Ünlüata, 1997). The North-West is also the region where most of the freshwater discharges. There, the Danube, Dniepr, and Dnister rivers flow into the Black Sea (Fig.1.1). Seven other relatively large rivers, located in other zones, flow in the basin: the Rioni, Çoruh, Kızılırmak, Sakarya, Yeşilırmak, Kodori, and the Bzyb, but also many smaller ones not cited here. All together, these rivers approximately count for 88% of the total freshwater flux into the basin (Jaoshvili, 2002). The Danube itself contribute to almost 60% (of the 88%), which makes it by far the largest water discharge into the basin. Due to the limited water outflow, exclusively through the Bosphorus Strait, the surface water is mainly influenced by the riverine freshwater inflows (Özsoy and Ünlüata, 1997). As already mentioned above, the evaporation rate is equivalent to the precipitation rate, such as the total rivers inflow ($300 km^3 y^{-1}$) is almost equivalent to the water outflow through the Bosphorus strait (Stanev and Ricker, 2019). In addition, these influxes could also influence the plastic pathways as it could accumulate in river plume fronts (van Sebille et al., 2020). Therefore, rivers act both as a barrier for marine litter from coastlines to the open oceans, since plume fronts can become accumulation zones, as well as a source of marine litter into the sea.

The transport of plastic litter into the Black Sea seems to be largely due to the coastal population density (Lebreton et al., 2019), and to shoreline activities (i.e., coastal tourism and recreational fishing) (UNEP, 2016). Such emissions are usually caused by landfills and wastewater mismanagement, or beach tourism. In fact, UNEP (2016) reported that on a national level, Ukraine and Turkey together had between 0.5 and 2 thousand tonnes per day of mismanaged plastic waste in 2010. In 2018, Bulgaria discarded 70% of its plastic waste in landfills (PlasticsEurope, 2019), making it more likely to end up in the Black Sea. Concerning the lack of wastewater management, Georgia treated less than 1% of its wastewater before 2015, while Turkey is the country within the entire Black Sea region which has the highest level of wastewater treatment, between 40-60 %, which is still relatively low (UNEP, 2016). Moreover, since the largest cities are situated along the coast, implying that the plastic released by mismanaged wastewater and landfills can result in the direct disposal into the Black Sea. Furthermore, rivers should also be considered as an additional key entry-point for plastics. It has been reported that at least 2,060 tonnes of plastic per year comes from rivers (Lebreton et al., 2017), of which a minimum of 530 tonnes are transported by the Danube only (van der Wal et al., 2015). Finally, Moncheva et al. (2016) reported that the North-Western part is the most polluted region. This is an expected outcome since it is the area where the largest river mouths are found and it is also an intense region for navigation (Fig.2.2).

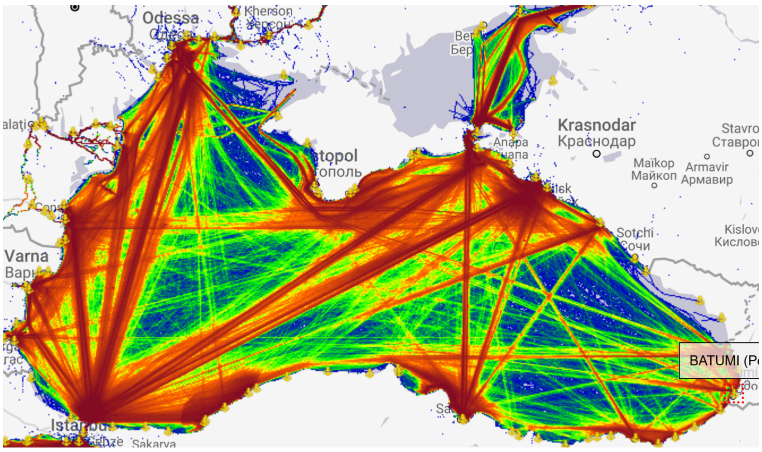


Fig. 2.2: Navigation routes in the Black Sea for the most recent year available (2017). Red to green lines represent the most to least used routes, the blue areas are where there is almost no navigation. [Source : www.marinetraffic.com]

Bayes' Theorem

Bayes' theorem is here used as a potential tool to infer the sources of plastic introduced in the marine environment. This theorem is widely used in our contemporary society: from online translation, to physics, or national security (McGrayne, 2011). It expresses the probability of an event, based on *prior* knowledge of conditions that are related to the event (McGrayne, 2011). In other words, this probability is used to quantify the uncertainty level of a statement (Martin, 2016). With a probability of 1, the statement is 100% certain, while a probability of 0 denotes a 0% certainty. Hence, to understand the theorem, one should be familiar with the notion of probability, in particular *conditional* and *conjoint* probability.

Conditional probability, expressed as $P(A|B)$ - read as the probability of A given B - is a probability based on some background information (Downey, 2013). By adding background information to a prediction, its probability changes. For instance, say that the probability that a random person gets sick after drinking milk is equal to 0.5, if we know that this person is lactose intolerant, then the probability goes up to 1 (i.e., the prediction is 100% certain). A conjoint probability expresses the fact that two statements are true, such as $P(A,B) = P(A)P(B)$ (Downey, 2013; Martin, 2016). However, this is valid if and only if A and B are independent: the outcome of the first event does not influence the probability of the second event (Downey, 2013). This means that A and B are commutative (i.e., interchangeable). Thus, in a more general sense, the conjoint probability can be equally expressed as $P(A \cap B) = P(B|A)P(A) = P(A|B)P(B)$ (Downey, 2013; McGrayne, 2011).

The definition of the conditional and the conjoint probability form the foundations of the Bayes' theorem. They are therefore used as follows :

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \tag{1}$$

with $P(A|B)$ and $P(B|A)$ conditional probabilities, while $P(A)$ and $P(B)$ are marginal probabilities, also called unconditional probabilities (i.e., the probability that event A/B occurs but without being conditioned by another event).

When speaking of Bayesian inference, the theorem is used to update the prior probability of an hypothesis H as new data D becomes available and so become a diachronic interpretation of Eq. 1 (Donovan and Mickey, 2019; Downey, 2013). Meaning that it is *" of, relating to, or dealing with phenomena [...] as they occur or change over a period of time"* (Merriam-Webster), with here, the change over time being due to the addition of new observational data. It is described as follows:

$$P(H|D) = \frac{P(D|H)P(H)}{P(D)}, \tag{2}$$

with $P(H|D)$ the posterior, $P(H)$ the prior, $P(D|H)$ the likelihood, and $P(D)$ the normalizing constant.

The fundamental difference between Eq.1 and Eq.2 is in the denominator $P(D)$: it should represent the probability of seeing the data under any hypothesis (Downey, 2013; Martin, 2016). However, in many cases, such as this project, it is difficult to narrow down an acceptable assumption for $P(D)$. Nevertheless, the law of total probability can be used to calculate it, such that :

$$P(D) = P(H|D)P(H) + P(H^c|D)P(H^c), \tag{3}$$

with the exponent c indicating the complementary of H. In the literature, it is often referred to as "does" (H) and "does not" (H^c). Hence, it can be assumed that H and H^c represent all the computable hypotheses of a specific case, allowing the assumption that $P(D) = \sum P(H_i|D)P(H_i)$.

Thus, for the purpose of this project, the Bayes' theorem is defined as follows:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum P(B|A_i)P(A_i)}, \tag{4}$$

with :

$\boldsymbol{P(B)}$ ≡ P(plastic being sampled in location B)

$\boldsymbol{P(A_i)}$ ≡ P(source i)

$\boldsymbol{P(B|A_i)}$ ≡ P(find sample from location B knowing the source A)

$\boldsymbol{P(A_i|B)}$ ≡ P(that source A is the origin of a known sample from B)

The probability of a source ($P(A_i)$) is based on the dataset found in the literature (Appendix A), while the probability of $P(B)$ and $P(B|A)$ are computed with the Lagrangian modelling of virtual particle trajectories. This allows to compute $P(A|B)$ and infer the source of a plastic sampled in a specific location.

As a final remark on the Bayes' theorem, as it can be seen in Eq.2 and Eq.3, the result of $P(A|B)$, equivalently $P(D|H)$, is optimal when combining all the possible hypotheses. In fact, the definition of $P(D)$ (Eq.3) involves the hypothesis (H) as well as its complementary (H^c), hence any other possible hypothesis within the sample space. Nevertheless, it is almost impossible to gather all the knowledge linked to an event, which should be done in order to carry out a full Bayesian analysis. As accurately formulated by Goldstein and Wooff (2007), "*Bayes falls victim to the ambition of its formulation*". In other words, there is no absolute answer to a probability question. Moreover, the fact that probabilities are numerical summaries of a person's epistemic state of knowledge about a subject, causes their inherent subjectivity (Rougier, 2007). Indeed, Bayesian statistics, and even more when considering Bayesian inferences, could seem to be subjective. Nevertheless, it is important to note that this subjectivity is more the expressions of the acknowledgment that our understanding of the world is conditioned by the data and models that have been made, and hence imperfect (Martin, 2016), than an actual real bias. However, these concepts are entering the philosophical sciences expertise and are therefore beyond the scope of this thesis. Thus, here, the assumptions and approximations have been made based on the data available.

Entropy of Mixing

As stated by Wichmann et al. (2019), plastic particles can only be back-tracked over a finite timescale t. Hence, for the purpose of this project, the entropy of mixing is calculated here in order to understand over what timescale t the source of plastic present in the Black Sea can be inferred.

To understand what the entropy of mixing consists of, it is easiest to start by defining *entropy*, or more specifically *Shannon entropy*, and *mixing*, in the context of this thesis. The Shannon entropy is a quantity that is used to estimate the uncertainty, or information, inherent to the outcome of random events - in probabilistic terms -, that can take any value between 0 - minimum entropy - and 1 - maximum entropy (Guida et al., 2010). Hence, here the entropy is not a property of the event itself but of the knowledge one has of the event. Here is a quick example to illustrate it : if we flip a coin that we know is loaded, says it gives only heads, the entropy of this event is equal to 0 because the outcome is known; if the coin is fair, then the entropy is equal to 1, since it is impossible to predict the outcome. Meanwhile, mixing represents the increase in homogeneity within a system (Guida et al., 2010; Camesasca et al., 2006). However, it can also be interpreted as a process by which particles of different species are exchanged between regions of the studied system (Guida et al., 2010), with the species being relative to the particles location of origin (Fig. 2.3). Thus, the entropy of mixing describes the efficiency of the mixing of these particles, or how they are exchanged, between the different regions of the system (Guida et al., 2010; Wichmann et al., 2019).

For the purpose of this project, the entropy of mixing S_k at time t is calculated for each particle i at each bin k (with a 0.2°x 0.2°resolution) after advection of particles initially homogeneously

(0.1°x 0.1°resolution) distributed over the Black Sea (Fig.2.3), which follows the Wichmann et al. (2019) approach:

$$S_k(t) = -\sum_i P_{i|k}(t) ln P_{i|k}(t), \tag{5}$$

where $P_{i|k}$ is the conditional probability to find a particle i in a bin k at time t and is defined as:

$$P_{i|k}(t) = \frac{\rho_{i,k}(t)}{\sum_i \rho_{i,k}(t)}, \tag{6}$$

where $\rho_{i,k}$ is the density of particle i in bin k.

Perfect mixing, and hence maximum entropy, arrives when the concentration of any particles in any bins is the same as the concentration of those particles in the whole basin (Camesasca et al., 2006). Thus, the maximum entropy is given by $S_k^{max} = lnM$, with M the number of particle species (Wichmann et al., 2019).

To summarize, here, the entropy of mixing is useful to understand on which timescales the information about the plastic fragment source is lost. This means that, once the timescale to arrive to the maximum mixing (i.e., $S_k/S_k^{max} = 1$) is reached, the plastic source can not be back-tracked anymore since the information containing it is lost (Wichmann et al., 2019). Therefore, calculating the entropy of mixing allows to better determine the timescale of the simulations, and this timescale gives also the range of time over which the source of a plastic particles can be inferred.

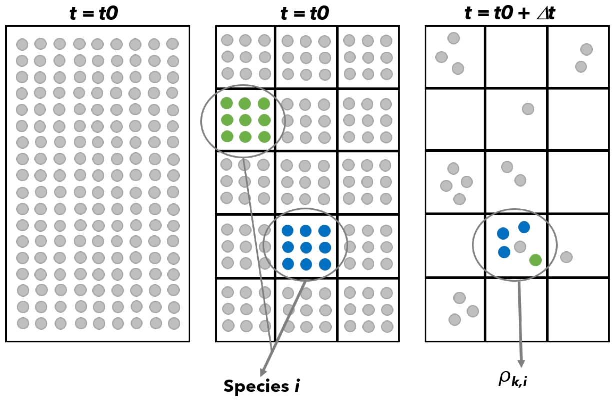


Fig. 2.3: Schematic view of how the entropy of mixing is calculated. At time $t = t0$, the particles are placed homogeneously over the studied domain. The domain is divided on several grid-cells k and all the particles located in the same grid-cell k is from the same species i. Hence, at time $t = t0$, any grid-cell contains particles from a unique species. After advection, at time $t = t0 + \triangle t$, the distribution is not homogeneous anymore and the density $\rho_{k,i}$ for each species i in each grid-cells k can be calculated.

III | Methods

Hydrodynamic and Observational Data

To apply Bayesian inference to the plastic pollution in the marine environment, virtual particle simulations are needed. Hence, to get results as close as possible to the realistic features of the Black Sea, accurate hydrodynamics' description is needed.

The surface currents velocity is obtained from 1/12°resolution (i.e., 9.3 km) global Copernicus Marine Environment Monitoring Service (CMEMS) reanalysis (i.e., GLORYS12V1)[1] provided by the EU Copernicus institute. The Black Sea is a eddy dominated system with a Rossby radius of 20-30 km (Staneva et al., 2001), so the CMEMS product resolution is high enough to resolve most of the dynamical features of the Black Sea. This dataset provides daily mean fields from 1993 to 2018 over 50 depth levels. In addition, it includes fields describing the horizontal velocities, the salinity, the temperature, and sea ice features (concentration, thickness and horizontal velocities). An detailed description of this product is given by Fernandez and Lellouche (2018) and a quality assessment is given by Drévillon and et al. (2018). For the purpose of this project, only the surface currents is used which corresponds to a depth of 0.49m in the CMEMS dataset.

To have an accurate simulation of the sea surface, the Stokes drift is usually a necessary field to add to the surface current. The Stokes drift data is taken from the 1/5 °resolution global WaveWatch3 model developed by the National Oceanic and Atmospheric Administration (NOAA). This dataset provides fields from 1999 to the present days. More details on the model can be found in WW3DG (2019).

Finally, the location of ten stations sampled along the Romanian coasts - from the river Danube mouth to the city of Constanta - during 'CoCoBLAS 2015' cruise held from the $26^t h$ to the $29^t h$ of May 2015 on board R/V Mare Nigrum is used. These locations are used as an example of how the results of the Bayesian inference can be applied to observational data.

[1] Product's name: GLOBAL_REANALYSIS_PHY_001_030

Lagrangian Modelling

All simulations were carried out using the Ocean Parcels framework, which is fully described in Lange and van Sebille (2017) and which is freely available at http://oceanparcels.org. In a nutshell, the particles are advected using a C-grid interpolation scheme and the trajectories are interpolated with a 4^{th}-order Runge-Kutta scheme. The trajectory of a displaced particle, $\overrightarrow{X}$ is computed as follows:

$$\overrightarrow{\mathbf{X}}(t+\Delta t) = \overrightarrow{\mathbf{X}}(t) + \int_{t}^{t+\Delta t} v(\vec{\mathbf{x}}(t), t)dt + \Delta\overrightarrow{\mathbf{X}}_b(t), \tag{7}$$

where $v(\vec{\mathbf{x}}(t), \tau)$ is the flow velocity from an Ocean General Circulation Model (OGCM) at the particle location $\vec{\mathbf{x}}(\mathbf{t})$ at time t, while $\Delta\mathbf{X}_b(t)$ is a change of position due to an additional process. Here, either $v = v_c$ or $v = v_c + v_{sto}$, with v_c the surface currents v_{sto} the stokes drift. As for $\Delta\mathbf{X}_b(t)$, it is an already built-in diffusion kernel of Ocean Parcels that computes the 2D-diffusion in the basin by using the 4^{th}-order Runge-Kutta scheme for the advection and the 1^{st}-order Milstein scheme to approximate the diffusion (c.f. AdvectionRK4DiffusionM1).

Finally, it is also necessary to precise that plastic particles are assumed to be fully immersed directly below the surface, hence the wind drag is neglected. Also, in all simulations, particles that reach the coastlines are assumed to accumulate at shore and stay there (i.e., they are beached) and therefore they are not re-suspended by waves dynamics.

Lagrangian Black Sea Circulation

In order to choose which parameters to implement to solve the Black Sea hydrodynamic features described in Chapter II, several combinations were tested : a) using the surface velocities only; b) the surface velocity plus the Stokes drift; c) the surface velocity plus the Stokes drift and diffusion (Table 1). The diffusion coefficients K_h zonal and K_h meridional are randomly determined, following the approach of Gräwe (2011). For those three sets of simulation, the particles were tracked for one year from the 1^{st} of January 2015. This year was chosen because it corresponds to the year when the plastic analysed fragments from the Black Sea were sampled (c.f. "Hydrodynamic and Observational Data" section of this Chapter). Their initial positions were distributed over an horizontal grid of 0.1°resolution and at each point twenty particles were released at the beginning of the simulation (38,260 particles in total).

Fig. 3.1 shows the results of the three experiments as density maps. The densities were calculated for each bin as an average over the full simulation (i.e., over one year) since the idea here is to find a model that best represents the dynamical features of the Black Sea. It can be seen that the addition of the Stokes drift to the surface currents leads to a depletion of

	Surface currents ($v = v_c$)	**Stokes drift** ($v = v_c + v_{sto}$)	**Diffusion** ($\Delta \overrightarrow{X}_b(t) \neq 0$)
Experiment A	×		
Experiment B	×	×	
Experiment C	×	×	×

Table 1: Summary of the three experiments (A, B, C) tested to setup the model. The mathematical expressions in italic refer to eq.7. For experiment A, only the surface currents were used; for B the surface currents and the Stokes drift; and for C, the surface currents, the Stokes drift and the diffusion. Experiment C provided the more realistic results, and hence is the setup used for the later simulations.

particles in the Northern region of the Black sea (Fig. 3.1a vs. Fig. 3.2b) and to more stranded particles along the Southern coasts. This is probably due to the northerly winds present in this region (Stanev and Ricker, 2019). Also, more particles stay trapped in the Western gyre (Fig.2.1) compared to the situation in which they are advected only with the surface currents. Finally, the addition of diffusion, leads to more diffusive density patterns (Fig. 3.1c) compared to the case in which only the Stokes drift is added (Fig. 3.1b). Experiment C (Fig. 3.1c), capture the Western gyre as well as the Batumi eddy. Since it is one of the permanent features of the Black Sea, it has been decided that this model setup best represents the Black Sea's hydrodynamic feature and therefore it has been used for all the further simulations. Note that the addition of diffusion to the model does not smooth the flow as one could expect. Instead, the addition of random scattering here tend to amplify the quantity of particles into the center of the basin and hence towards the permanent gyres which are isolated from the rest of the basin by the Rim current. Thus, more particles are trapped and those features, in particular the Batumi gyre (Fig.2.1), are enhanced (Fig. 3.1c).

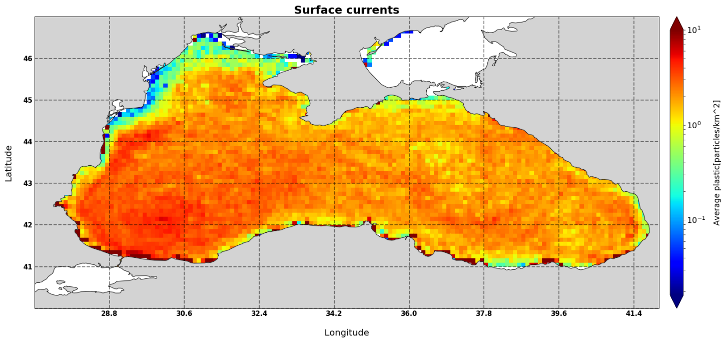
Probabilistic information

The probabilistic information was partially derived with the simulations, and partially deducted from the data available in the literature. As specified in Chapter II, the Bayes' theorem is defined by Eq.4, and is applied to this project as follows :

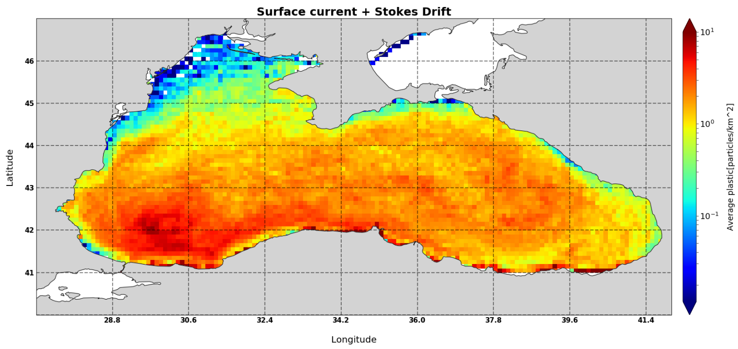
$$P(River_i|Location_j) = \frac{P(Location_j|Source_i)P(Source_i)}{\sum P(Location_j|Source_i)P(Source_i)} \tag{8}$$

In order to compute Eq.8, the first step is to determine $P(Source)$ from the literature. The Marine Litter Report of the Black Sea Commission (Birkun et al., 2009) presented a ranked list of the primary sources of marine litter in the Black Sea for each bordering country (Annex B). The ranks range from 1 - least likely - to 5 - most likely - and were assessed by the own appraisal of national experts. Hence, from the mean score results, in descending order, the sources likelihood are : sewage plants, ship waste/port and rivers, tourism, industry, and fishery (NB: *tourism* includes any recreational activities in coastal area). This estimation is quite consistent with the studies mentioned in Appendix A. Subsequently, and because it is the best data in term of quality and quantity that was found, this work focuses exclusively on rivers as possible sources of plastic.

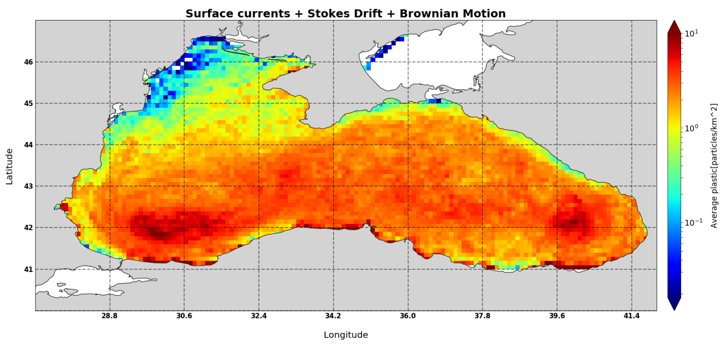
Previous studies that have estimated the water discharge of the main rivers of the Black are



(a)



(b)



(c)

Fig. 3.1: Results of the three experiments made to determine which components are required to best represent the physical features of the Black Sea: the particles are advected a) only with the surface currents; b) with the surface currents and the Stokes drift; c) with the surface currents, the Stokes drift, and diffusion.

	Danube	**Dniepr**	**Dniester**	**Rioni**	
Obs. years	1960-2000	1960-2006	1960-1984	1960-1984	*Ludwig et al. (2009)*
Mean flux [m^3/s]	6573.8	1488.1	376.6	110.5	408.5
Obs. years	bf. 1997	bf. 1997	bf. 1997	bf. 1997	*Bat et al. (2018)*
Mean flux [m^3/s]	6595.6	1623.5	323.4	405.9	
Obs. years	-	-	-	-	*Jaoshvili (2002)*
Mean flux [m^3/s]	6300	1375	288	31.6	
Mean [m^3/s]	**6489.8**	**1495.5**	**329.3**	**407.2**	

	Çoruh	**Kodori**	**Bzyb**	**Yeşilırmak**	**Kızılırmak**	**Sakarya**	
Obs. years	bf. 1997	bf. 1997	bf. 1997	bf. 1997	bf. 1997	bf. 1997	*Bat et al. (2018)*
Mean Flux [m^3/s]	275.6	129.4	97.3	156.3	159.2	202.3	
Obs. years	-	-	-	-	-	-	2*Jaoshvili (2002)*
Mean Flux [m^3/s]	12.5	132	120	168.1	187.1	177.6	
Mean [m^3/s]	**275.6**	**130.7**	**108.7**	**162.2**	**173.2**	**190**	

Table 2: Main rivers and their water discharge in the Black Sea. (NB : the red value for the Rioni has been neglected since it was excessively small compared to the other reported values.)

presented in Table 2. From this information, the mean freshwater runoff in the Black Sea is 9762.2 m^3s^{-1} and the contribution of each rivers is as follows : Danube: 65.9%, Dniepr: 15.2%, Rioni: 4.1%, Dniester: 3.3%, Çoruh: 2.8%, Kızılırmak: 1.8%, Sakarya: 1.9%, Yeşilırmak: 1.6%, Kodori: 1.3%, Bzyb: 1.1%; which is consistent with the results of Jaoshvili (2002). There are obviously many smaller rivers flowing into the Black Sea but since their water discharge is so little (< 1%), they have been neglected.

Since the concentration of plastic in a river is unknown, one of the main assumption of this study is that the plastic concentration stays constant over time and corresponds to the river's water discharge. Thus, in this context, Eq.8 can be rewritten as :

$$P(River_i|Location_j) = \frac{P(Location_j|River_i)P(River_i)}{\sum_{i,j=1}^{i,j=n} P(Location_j|River_i)P(River_i)}, \tag{9}$$

with i being the rivers listed above, and their probability is their corresponding contribution (e.g., $P(River_{Danube}) = 0.659$).

The probability to find a plastic fragment in a specific location knowing its source is determined by tracking virtual Lagrangian particles. Thus, ten distinct simulations, corresponding to the number of possible sources, were carried out using the model setup determined earlier (Fig. 3.1c). For these simulations, particles were released and tracked everyday during one year, from 1^{st} of January 2015 to the 31^{st} December 2015, from the geographical position of the source. Afterwards, density maps were constructed by dividing the Black Sea in bins of 0.1°resolution (i.e., the $Location_i$) in which the average density over the timescale of the event (e.g., one year) was calculated. Here, the average density is taken because nor the particle age nor the sampling date is known. Hence, it is assumed that this information would have no effect on the resulting density (i.e., sampling a particle in January is equivalent to sampling it in September). This way, $P(Location_j|River_i)$ is obtained for each grid-cell ($Location_j$) and for each river.

Finally, by combining the probabilistic information deduced from the literature ($P(River_i)$) and from the simulations ($P(Location_j|River_i)$), the origins of plastic could be inferred by using Bayes' theorem (Eq.9).

Inference Timescales

After time t, the origin of a particle cannot be back-tracked anymore and so by extension the origin of a plastic fragment cannot be inferred anymore. To determine the timescale for which the Bayesian inference can be applied, two methods can be used : a) the Shannon entropy; b) the Bayes' theorem itself. For method b), Eq.4 can be rearranged as :

$$\frac{P(A_i|B_j)}{P(A_i)} = \frac{P(B_j|A_i)}{\sum P(B_j|A_i)P(A_i)}, \tag{10}$$

with $P(B_j|A_i)$ is the density of plastic from the source A_i at location B_j and $\sum P(B_j|A_i)P(A_i)$ is the maximum density over the basin.

Hence, if Eq.10 equals 1, then the mixing is maximum, which mean that the source of a plastic fragment cannot be inferred. For Eq.10 to be equal to 1, $P(B|A)$ must be equivalent to $P(A)$. Thus, when the probability that a plastic from the source A_i is the same as the probability that a plastic comes from the source A_i knowing its sampling location B_j, then its source cannot be inferred anymore.

Secondly, the entropy of mixing, derived from the Shannon entropy, was also calculated to determine the mixing timescales. The Shannon entropy is highly sensitive to the amount of information we have about an event, or an item in this case. More information known about an event leads to longer timescales to reach maximum entropy. A common example is the marble bag game: the player has a marble bag with both blue and red marbles and they must guess which color marble they will randomly draw. If the player does not have the information that the marbles can be either blue or red, the probability that they guess the color is close to 0 and the entropy is maximum (i.e., close to 1). However, if they know the colors and if, for instance, they also know how many marbles from each color are in the bag, then the entropy is smaller than 1. Here, the entropy of mixing was calculated for the following situation: very little is known about the plastic particles (i.e., only their origin and final position is known). This is mainly for consistency: to compute the Bayesian inference, time averaged densities have been used, implying that the particle age and its sampling date is unknown. If these two parameters were known, the quantity of information about the plastic fragments would be more, and so the timescales to reach the maximum entropy would be longer, meaning that the plastic fragment could be back-tracked over a longer time.

Finally, to compute the entropy of mixing, twenty particles were released at each point of a homogeneous 0.1 degree resolution horizontal-grid and tracked for one year. The mixing entropy was calculated by comparing the initial position of a particle with its final position, for different timescales, up to one year (Fig. 4.1a-4.1d). It is important to note that the entropy of mixing is quite sensitive to the number of released particles, if they are too little the entropy values could be equal to zero everywhere. Here, 38,260 particles in total were released at each point of the initial grid. In addition, the entropy analysis was carried out over a coarser grid than the initial grid (0.2 degree and 0.1 degree resolution, respectively).

IV Results and Discussion

Shannon Entropy

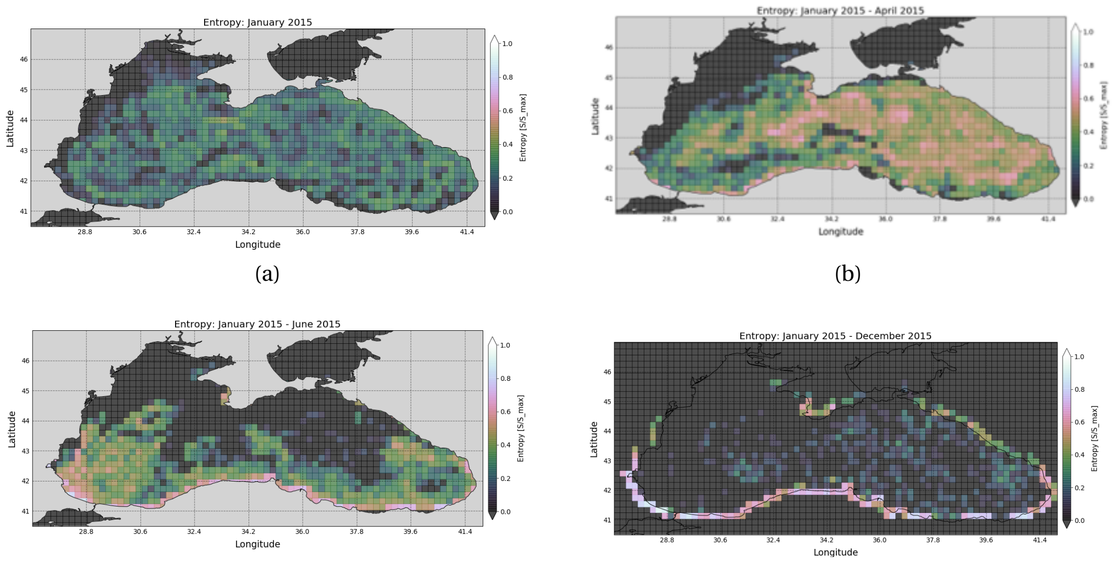
The main result obtained was that in the open ocean, the highest level of entropy reached is of 0.657 after two months of simulation, while near the coasts, in particular in the South-Western area, values close to the maximum entropy emerge after 6 months. The decrease of the entropy in the open ocean is due to the model assumption that the particles are stranded when they reach the coast. Hence, the density of particles in the open ocean lessen while they are cast ashore the Southern coasts of the Black Sea, leading to an increasing density - and entropy - on the beaches and in the coastal zones. Therefore, the maximum entropy of mixing is never reached in the Black Sea. Thus, in theory, the source of plastic fragments sampled in the open ocean can be back-tracked for any timescales, up to five years. In fact, after five years, for this model setup, all the released particles have stranded (i.e., there is no particles left in the open sea). The plastic particles sampled near the Southern coasts of the Black Sea reaches values close to maximum entropy (0.92) in some regions after 1 year (Fig.4.1d). This suggests that the inference of the plastic that is found on the beaches and the coastal zonal is time-limited. This result is not surprising since most of the particles are transported towards the Southern coasts by the northerly winds.

To summarize, the results of the entropy of mixing is convenient because it means that for most of the regions, in the open sea, the sources of plastic can be inferred after any time t smaller than five years after their release. However, stranded particles can be back-tracked only over a limited timescale, smaller than two years.

Plastic Source Back-tracking

The results of the Bayesian inference of particle source are presented as probability maps of the Black Sea for each potential source: the Danube, Dniepr, Dniester, Coruh, Kilizirmak, Rioni, Sekarya, Yesilimark, Kodori, and Bzyb rivers (Fig.4.4).

Because the simulations setup assumes that the plastic concentration is proportional to the water discharge of the rivers, plastic most likely originates from the Danube (Fig.4.5a). Since the Danube contributes to 66.5% of the total freshwater inflow in the Black Sea, the other



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

Fig. 4.1: Entropy mixing for particles advected from a homogeneous horizontal grid, after: a) 1 month; b) 4 months; c) 6 months; d) 1 year. An entropy mixing value of 1 represents the maximum entropy, while the value 0 indicates that a region is completely depleted of particles (i.e., does not mean in *per se* that there is no entropy). Note that after 1 year, the majority of the released particles have already beached, hence the vast majority of depleted grid-cells.

rivers have a minimal role on the input of plastic litter into the Black Sea, compared to the Danube. Focusing on some of the other rivers, in the far North-Western area, plastic comes almost exclusively from the Dniepr (Fig. 4.5b). This can be explained by the general circulation of the Black Sea: the combination of the northerly winds and the strong Rim current lead to the accumulation of particles southward and limits the influx of particles of other zones into the North-Western part. Hence, the particles found in this region almost exclusively originates from the Dniepr.

Regarding the Bzyb, is that even though it contributes only to 1.1% of the total freshwater discharge into the Black Sea, its particles largely spread over the Eastern region of the basin and seems to be the only possible source of plastic in some specific locations close to the North-Eastern coasts around 45°N 36°E (Fig. 4.5j). Again, this can be explained by the general circulation: the Bzyb mouth is located near an area (i.e., North-East) in which many small eddies appear between the coast and the Rim current, and it is also close to the Batumi gyre. Thus, plastic from the Bzyb can easily get trapped in these eddies and gyres, which is why those scatter so much over the Eastern part of the basin. As a consequence, in this regions, the Bzyb is more likely to be the source of plastic than, for instance, the Coruh, which has a larger contribution to the total freshwater (i.e., 2.8%) input than the Bzyb. For the other rivers, generally, the probability that a plastic fragment originates from them is relatively high when it is sampled near the river mouth, and becomes quite low when the distance to the source increases (Fig. 4.5c-4.5i). Note that for this project, the rivers contributing to less than 1% of

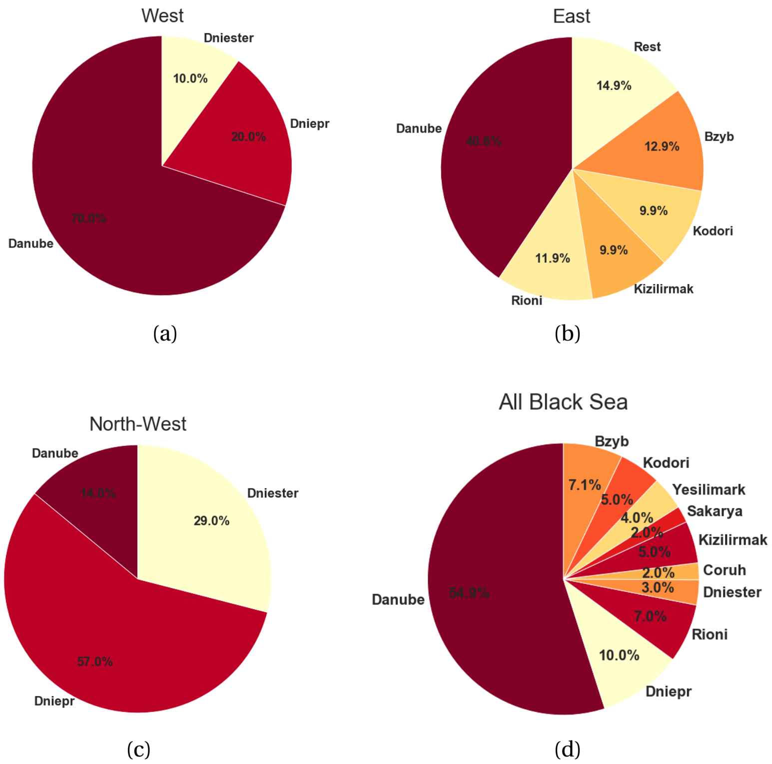
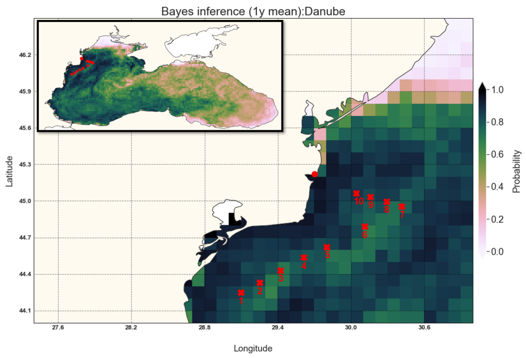


Fig. 4.2: Pie charts displaying the probabilities that sampled plastics comes from a specific river, at a regional scale: a) West; b) East; c) North-West; and d) for the full basin. The delimitation of the regions are not political but in function of the results of the Bayesian inference. A map of this delimitation is presented in Appendix C.

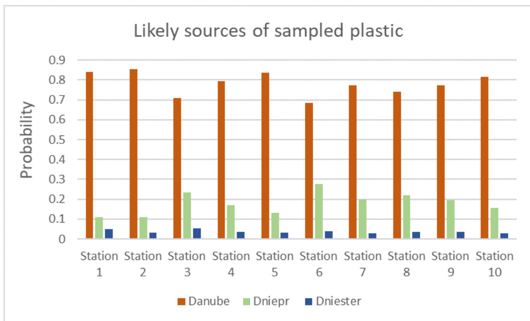
the total freshwater discharge have been neglected. By adding these rivers to the analysis, the final probability maps might be different. However, since the actual plastic concentration in the rivers is unknown, and hence based on the water discharge alone, these differences would be here, quite small.

These results can also be visualized as a pie chart (such as Fig.4.2). As expected, the Danube is the main source of plastic in all parts of the Black Sea besides the North-West (West: 70%, East: 40.6%, North-West: 14%). Considering the whole basin, after the Danube, the Dniepr (10%) and the Bzyb (7.1%) are the two most important sources of plastic of the basin. The pie chart for the whole Black Sea (Fig. 4.2d) is important because it allows us to see if $P(A|B)$ is equal to $P(A)$ or not, and by extension understand if the mixing of particles is at its maximum or not (cf. Chapter III). As a reminder, the $P(A)$ of each river here is proportional to their water discharge and is as such: Danube: 65.9%, Dniepr: 15.2%, Rioni: 4.1%, Dniester: 3.3%, Çoruh: 2.8%, Kızılırmak: 1.8%, Sakarya: 1.9%, Yeşilırmak: 1.6%, Kodori: 1.3% , and Bzyb: 1.1%. The comparison between these contributions and Fig.4.2d, shows that $P(A|B)$ and $P(A)$ are not equal. In other words, the $P(A)$ of each rivers, mentioned above, is not the same as the results for the full basin (Fig.4.2d), which induces that the the mixing is not at its maximum. This implies that the plastic source could be inferred over an even longer time, which follows the result given earlier by the Shannon entropy (c.f. "Shannon Entropy" section of this Chapter).

Finally, Fig.4.3 shows a possible application of the probability results to real sampled data. The ten sampling stations (red crosses) are interpolated over the probability map (Fig. 4.3a), of the Danube since they are located close to its mouth and also because it is the most likely source of plastics over the Black Sea (Fig.4.2d, Fig.4.44). This way, it can be seen that the most probable origin of the sampled plastics is, all stations included, the Danube (Fig. 4.3b) with an average probability of 0.78. However, there is also a small chance that they provide from the Dniepr or the Dniester, with an averaged likelihood over all the stations of 0.18 and 0.03, respectively. Note that the sum of the averaged probability is not equal to 1. This is because, there is very small likelihood that these plastic fragments comes from the other analysed rivers (c.f Annex E).



(a)



(b)

Fig. 4.3: a) Interpolation of sampling locations (red crosses) and the result of the probability map for the Danube as a source of plastics. The red dot shows the location of the Danube mouth; b) Bar chart showing the most likely source for each station. For all the stations, the Danube (orange) is the most probable source, followed by the Dniepr (green) and the Dniester (blue).

Tool Validity

The main purpose of this project is to develop a tool to back-track plastic fragments to their source. Hence, a fundamental point is to demonstrate that this method is in fact reproducible. This could be done by comparing the density maps of each source (Annex D) to the probability map of the same source. In fact, the probability should be high where the particle density is large. However, a more solid statistical proof is obtained by adding up the probability of all the sources at each grid-cells (i.e., summing all the probability maps together). The results should be that the total probability is equal to 1 everywhere. Thus, since this criterion is attained (Annex F), the inference of plastic source with Bayesian statistics is, statistically speaking, a reliable tool. However, it still would need a more comprehensive verification to completely prove its reliability.

A thorough validation of an oceanographic or statistical model should be validated by observations on site. For oceanographic circulation models, drifter measurements are usually adequate and in the context of plastic pollution they could also work. However, in this project, plastic fragments were assumed to flow just below the water surface which it also implies that their size is really small (i.e. micro- or nano-plastics). Thus, drifter measurements might not be enough, since the pathways of these kinds of particles are believed to be influenced by other factors than the ocean (surface) dynamics such as bio-fouling (Kooi et al., 2017; Fazey and Ryan, 2016). Another option could be to compare the sources using the Scoring Matrix Technique (MTS) (Tudor and Williams, 2004), that is applied directly on sampled marine litter, with the results obtained by Bayesian inference. The baseline of the MTS approach is to give a score based on the character and the aspect of the litter. For instance, an oil drum will be considered to be more likely sea-based or land-based if it is marked for ship use or domestic use, respectively (Tudor and Williams, 2004). Finally, a last alternative, would be to voluntarily release labelled plastic tracers. However, this would obviously bring up an ethical problem: trying to find answers to marine pollution by adding pollution ?

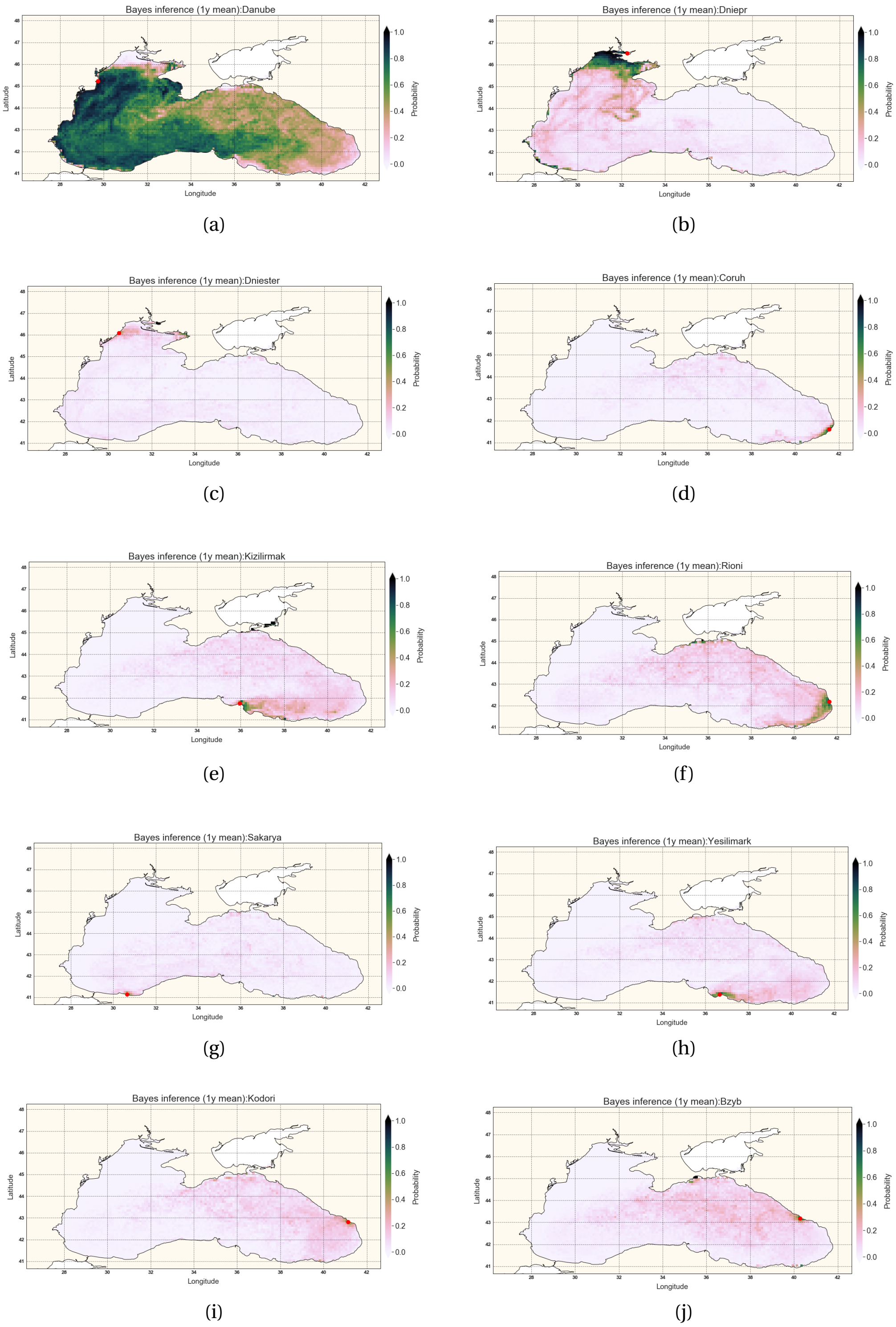


Fig. 4.4: Probability maps that a plastic fragment comes from a specific source with the source being: a) the Danube; b) the Dniepr; c) the Dniester; d) the Coruh; e) the Kilizirmak; f) the Rioni; g) the Sakarya; h) the Yesilimark; i) the Kodori; j) the Bzyb. These probabilities were inferred after tracking particles daily for one year.

V Conclusion

For this thesis, the baseline of a new method to assess the entry-points of sampled plastic in the oceans was developed by using Bayesian statistics. This method uses Lagrangian simulations of virtual particles performed via the Ocean Parcels framework from which the sources are inferred by applying Bayes' theorem. On a wider scale, this method could be applied to generate a map for each possible source of plastic sampled in a specific location in the ocean, with its associated probability values (such as in Fig.4.4).

This project focuses on the Black Sea exclusively with only the largest rivers being considered as possible sources of plastic (i.e., Danube, Dniepr, Dnister, Rioni, Coruh, Sakarya, Kilizirmak, Yesilimark, Kodori, Bzyb). The concentration of plastic in each river has been assumed to be constant through time and relative to the river's water discharge. Hence, the results might be biased: there is no observational study to date that report which of these rivers in fact transports the most quantity of plastic. The water discharge of the Danube (66.5% of the total freshwater discharge) being seventeen times larger than the mean of all the other nine rivers (3.72% of the total freshwater inflow), it is most likely that in fact it transports the largest quantity of plastic litter into the Black Sea. However, for the rest of the rivers, the assumption that their plastic concentration is relative to their water discharge is not so straightforward. In addition, if a plastic particle reaches the coast, it has been assumed to be stranded (i.e., re-suspension not possible). In this scenario, it has been found that the main source of plastic in the Black-Sea in absolute is the Danube (54.9%), followed by the Dniepr (10%) and the Bzyb (7.1%). The large difference between these probabilities is due to the assumption that the plastic concentration is relative to the river's runoff: the Danube's water discharge is by far the largest (6,489 m^3/s and the others are <1,496 m^3/s). However, when looking at the results region-wise, the outcome varies (although the Danube is always dominant). For instance, in the North-West region, the Danube, Dniepr, and Dniester are probably the only sources.

In addition, by calculating the Shannon entropy, it as been found that for simulations in which particles are allowed to strand, the plastic particle source can be back-tracked over a time t of five years in the open Black Sea. After five years, all the plastic has beached and cannot be back-tracked anymore, at least along the Southern coast. This is because particles tend to beach along the Southern shores, as a consequence of the northerly winds and the induced Stokes drift. Thus, the source of beached particles in that region and in the corresponding coastal zone can only be back-tracked over a maximum time t of two years.

The main benefit of using Bayesian inference for plastic pollution is that it can potentially be applied to any region, globally or locally, and to any plastic fragment size-range. In addition, because of the definition of the Bayes' theorem, the results can be updated and improved as soon as new observational data is reported. Nevertheless, using Bayesian statistics has also some down sides. It is mainly based on the previous knowledge of a specific system, including a prior idea of where the plastic could come from. For instance, in the case of the Black Sea for which the prior information is limited, it would be difficult, presently, to generate satisfying results for sources other than rivers. Plus, it largely depends on what the user considers as a source of plastic or not. Hence, before applying this method, it is necessary to clearly define what can be or cannot be a source and to assess the quality of the prior information related to that source.

Overall, and in conclusion, using Bayesian inference is a promising tool to be used to back-track the source of floating marine debris and to better understand the pathways of plastic in the marine environment. However, an imperative aspect missing here is the validation of this approach by observational data. Without such validation, there is no solid evidence that the results obtain via Bayesian inference reflect the reality. It is necessary to develop and improve such tools in order to switch from a bottom-up approach (i.e., cleaning the ocean after release) to a top-down approach (i.e., reducing the release directly from the source) concerning ocean plastic pollution. In fact, it would more efficient to focus on the sources of plastic to avoid its entry into the marine environment than removing it afterwards. Therefore, identifying plastic litter sources would as formulated by Carlson et al. (2017) "*maximize the effectiveness of prevention and response efforts by providing scientific support to the implementation of public policies*", and ideally by reducing the release of marine debris directly from their sources.

Bibliography

E. Anton, G. Radu, G. Ţiganov, M. Cristea, and M. Nenciu. The situation of marine litter collected during demersal surveys in 2012 in the romanian black sea area. *Cercetări Marine*, 43:350–357, 2013.

D. K. Barnes, F. Galgani, R. C. Thompson, and M. Barlaz. Accumulation and fragmentation of plastic debris in global environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1526):1985–1998, 2009.

A. Barrows, S. E. Cathey, and C. W. Petersen. Marine environment microfiber contamination: Global patterns and the diversity of microparticle origins. *Environmental Pollution*, 237: 275–284, 2018.

L. Bat. Black sea marine environment : The turkish shelf. 2017.

L. Bat, A. Öztekin, F. Şahin, E. Arıcı, and U. Özsandıkçı. An overview of the black sea pollution in turkey. *Mediterranean Fisheries and Aquaculture Research*, 1(2):66–86, 2018.

A. Birkun, A. Atudorei, T. Gamgebeli, S. Dedeoglu, N. Movchan, A. Nikolova, and Y. Yurenko. Marine litter in the black sea region, 2009.

M. Camesasca, M. Kaufman, and I. Manas-Zloczower. Quantifying fluid mixing with the shannon entropy. *Macromolecular theory and simulations*, 15(8):595–607, 2006.

A. Capet, A. Barth, J.-M. Beckers, and G. Marilaure. Interannual variability of black sea's hydrodynamics and connection to atmospheric patterns. *Deep Sea Research Part II: Topical Studies in Oceanography*, 77:128–142, 2012.

D. F. Carlson, G. Suaria, S. Aliani, E. Fredj, T. Fortibuoni, A. Griffa, A. Russo, and V. Melli. Combining litter observations with a regional ocean model to identify sources and sinks of floating debris in a semi-enclosed basin: the adriatic sea. *Frontiers in Marine Science*, 4:78, 2017.

F. S. Cesa, A. Turra, and J. Baruque-Ramos. Synthetic fibers as microplastics in the marine environment: A review from textile perspective with a focus on domestic washings. *Science of The Total Environment*, 598:1116–1129, 2017.

M. Cole, P. Lindeque, C. Halsband, and T. S. Galloway. Microplastics as contaminants in the marine environment: a review. *Marine pollution bulletin*, 62(12):2588–2597, 2011.

T. M. Donovan and R. M. Mickey. *Bayesian Statistics for Beginners: A Step-by-step Approach*. Oxford University Press, 2019.

A. Downey. *Think Bayes: Bayesian statistics in python*. " O'Reilly Media, Inc.", 2013.

R. Dris, J. Gasperi, M. Saad, C. Mirande, and B. Tassin. Synthetic fibers in atmospheric fallout: a source of microplastics in the environment? *Marine pollution bulletin*, 104(1-2):290–293, 2016.

M. Drévillon and et al. Quality information document for the global ocean physical reanalysis products global$_r eanalysis_p hy_0 01_0 30.2018.$

F. M. Fazey and P. G. Ryan. Biofouling on buoyant marine plastics: An experimental study into the effect of size on surface longevity. *Environmental Pollution*, 210:354–360, 2016.

E. Fernandez and J. Lellouche. Product user manual for the global ocean physical reanalysis product global_reanalysis_ phy_001_030. *Report CMEMS-GLO-PUM*, (001-030):15, 2018.

R. Geyer, J. R. Jambeck, and K. L. Law. Production, use, and fate of all plastics ever made. *Science advances*, 3(7):e1700782, 2017.

M. Goldstein and D. Wooff. *Bayes linear statistics: Theory and methods*, volume 716. John Wiley & Sons, 2007.

U. Gräwe. Implementation of high-order particle-tracking schemes in a water column model. *Ocean Modelling*, 36(1-2):80–89, 2011.

A. Guida, A. W. Nienow, and M. Barigou. Shannon entropy for local and global description of mixing by lagrangian particle tracking. *Chemical engineering science*, 65(10):2865–2883, 2010.

B. D. Hardesty, T. Lawson, T. van der Velde, M. Lansdell, and C. Wilcox. Estimating quantities and sources of marine debris at a continental scale. *Frontiers in Ecology and the Environment*, 15 (1):18–25, 2017.

T. Healy and K. Harada. Enclosed and semi-enclosed coastal seas. *Journal of Coastal Research*, pages i–v, 1991.

S. Hong, J. Lee, and S. Lim. Navigational threats by derelict fishing gear to navy ships in the korean seas. *Marine Pollution Bulletin*, 119(2):100–105, 2017.

C. Ioakeimidis, C. Zeri, H. Kaberi, M. Galatchi, K. Antoniadis, N. Streftaris, F. Galgani, E. Papathanassiou, and G. Papatheodorou. A comparative study of marine litter on the seafloor of coastal areas in the eastern mediterranean and black seas. *Marine pollution bulletin*, 89(1-2): 296–304, 2014.

A. Isobe, S. Kako, P.-H. Chang, and T. Matsuno. Two-way particle-tracking model for specifying sources of drifting objects: application to the east china sea shelf. *Journal of Atmospheric and oceanic technology*, 26(8):1672–1682, 2009.

J. R. Jambeck, R. Geyer, C. Wilcox, T. R. Siegler, M. Perryman, A. Andrady, R. Narayan, and K. L. Law. Plastic waste inputs from land into the ocean. *Science*, 347(6223):768–771, 2015.

S. Jaoshvili. The rivers of the black sea. 2002.

S. Kako, A. Isobe, S. Magome, H. Hinata, S. Seino, and A. Kojima. Establishment of numerical beach-litter hindcast/forecast models: An application to goto islands, japan. *Marine pollution bulletin*, 62(2):293–302, 2011.

P. Kershaw and C. Rochman. Sources, fate and effects of microplastics in the marine environment: part 2 of a global assessment. *Reports and Studies-IMO/FAO/Unesco-IOC/WMO/IAEA/UN/UNEP Joint Group of Experts on the Scientific Aspects of Marine Environmental Protection (GESAMP) Eng No. 93*, 2015.

M. Kooi, E. H. v. Nes, M. Scheffer, and A. A. Koelmans. Ups and downs in the ocean: effects of biofouling on vertical transport of microplastics. *Environmental science & technology*, 51(14): 7963–7971, 2017.

A. A. Kubryakov, S. V. Stanichny, A. G. Zatsepin, and V. V. Kremenetskiy. Long-term variations of the black sea dynamics and their impact on the marine ecosystem. *Journal of Marine Systems*, 163:80–94, 2016.

M. Lange and E. van Sebille. Parcels v0. 9: prototyping a lagrangian ocean analysis framework for the petascale age. *arXiv preprint arXiv:1707.05163*, 2017.

L. Lebreton, M. Egger, and B. Slat. A global mass budget for positively buoyant macroplastic debris in the ocean. *Scientific reports*, 9(1):1–10, 2019.

L. C. Lebreton, J. Van Der Zwet, J.-W. Damsteeg, B. Slat, A. Andrady, and J. Reisser. River plastic emissions to the world's oceans. *Nature communications*, 8:15611, 2017.

W. Ludwig, E. Dumont, M. Meybeck, and S. Heussner. River discharges of water and nutrients to the mediterranean and black sea: major drivers for ecosystem changes during past and future decades? *Progress in oceanography*, 80(3-4):199–217, 2009.

O. Martin. *Bayesian analysis with python*. Packt Publishing Ltd, 2016.

N. Maximenko, P. Corradi, K. L. Law, E. Van Sebille, S. P. Garaba, R. S. Lampitt, F. Galgani, V. Martinez-Vicente, L. Goddijn-Murphy, J. M. Veiga, et al. Towards the integrated marine debris observing system. *Frontiers in marine science*, 6:447, 2019.

S. B. McGrayne. *The theory that would not die: how Bayes' rule cracked the enigma code, hunted down Russian submarines, & emerged triumphant from two centuries of controversy*. Yale University Press, 2011.

Merriam-Webster. Diachronic. URL https://www.merriam-webster.com/dictionary/diachronic. [Online accessed 5-July-2020].

S. Moncheva, K. Stefanova, A. Krastev, A. Apostolov, L. Bat, M. Sezgin, F. Sahin, and F. Timofte. Marine litter quantification in the black sea: A pilot assessment. *Turkish Journal of Fisheries and Aquatic Sciences*, 16(1):213–218, 2016.

T. Oguz and S. Besiktepe. Observations on the rim current structure, ciw formation and transport in the western black sea. *Deep Sea Research Part I: Oceanographic Research Papers*, 46(10): 1733–1753, 1999.

A. Ötzekin and L. Bat. Microlitter pollution in sea water: A preliminary study from sinop sarikum coast of the southern black sea. *Turkish Journal of Fisheries and Aquatic Sciences*, 17(6), 2017.

E. Özsoy and Ü. Ünlüata. Oceanography of the black sea: a review of some recent results. *Earth-Science Reviews*, 42(4):231–272, 1997.

T. O'Brine and R. C. Thompson. Degradation of plastic carrier bags in the marine environment. *Marine pollution bulletin*, 60(12):2279–2283, 2010.

C. K. Pham, E. Ramirez-Llodra, C. H. Alt, T. Amaro, M. Bergmann, M. Canals, J. Davies, G. Duineveld, F. Galgani, K. L. Howell, et al. Marine litter distribution and density in european seas, from the shelves to deep basins. *PloS one*, 9(4):e95839, 2014.

PlasticsEurope. Plastics—the facts 2019. *An Analysis of European Latest Plastics Production, Demand and Waste Data.*, 2019. URL https://www.plasticseurope.org/application/files/4315/1310/4805/plastic-the-fact-2016. Online accessed on 25-March-2020.

J. Rougier. Probabilistic inference for future climate using an ensemble of climate model evaluations. *Climatic Change*, 81(3-4):247–264, 2007.

M. Sezgin, L. Bat, D. Ürkmez, and B. Öztürk. Black sea marine environment : The turkish shelf. *Turkish Marine Research Foundation (TUDAV)*, 46, 2017.

E. V. Stanev and M. Ricker. The fate of marine litter in semi-enclosed seas. case of the black sea. *Frontiers in Marine Science*, 6:660, 2019.

J. V. Staneva, D. E. Dietrich, E. V. Stanev, and M. J. Bowman. Rim current and coastal eddy mechanisms in an eddy-resolving black sea general circulation model. *Journal of Marine Systems*, 31(1-3):137–157, 2001.

G. Suaria, C. G. Avio, A. Mineo, G. L. Lattin, M. G. Magaldi, G. Belmonte, C. J. Moore, F. Regoli, and S. Aliani. The mediterranean plastic soup: synthetic polymers in mediterranean surface waters. *Scientific Reports*, 6, 2016.

G. Suaria, V. Perold, J. R. Lee, F. Lebouard, S. Aliani, and P. G. Ryan. Floating macro-and microplastics around the southern ocean: Results from the antarctic circumnavigation expedition. *Environment International*, 136:105494, 2020.

E. N. Topçu and B. Öztürk. Abundance and composition of solid waste materials on the western part of the turkish black sea seabed. *Aquatic Ecosystem Health & Management*, 13(3):301–306, 2010.

D. T. Tudor and A. T. Williams. Development of a 'matrix scoring technique'to determine litter sources at a bristol channel beach. *Journal of Coastal Conservation*, 10(1):119–127, 2004.

UNEP, 2020. URL https://www.unenvironment.org/explore-topics/oceans-seas/what-we-do/working-regional-seas/marine-litter. [Online accessed 5-April-2020].

G. UNEP. Arendal: Marine litter vital graphics. *Nairobi and Arendal: United Nations Environment Programme and GRID-Arendal*, 2016.

M. van der Wal, M. Van der Meulen, G. Tweehuijsen, M. Peterlin, A. Palatinus, M. Viršek, L. Coscia, and A. Kržan. Identification and assessment of riverine input of (marine) litter final report for the european commission dg environment under framework contract no env. 2015.

E. van Sebille, C. Wilcox, L. Lebreton, N. Maximenko, B. D. Hardesty, J. A. Van Franeker, M. Eriksen, D. Siegel, F. Galgani, and K. L. Law. A global inventory of small floating plastic debris. *Environmental Research Letters*, 10(12):124006, 2015.

E. van Sebille, S. Aliani, K. L. Law, N. Maximenko, J. M. Alsina, A. Bagaev, M. Bergmann, B. Chapron, I. Chubarenko, A. Cózar, et al. The physical oceanography of the transport of floating marine debris. *Environmental Research Letters*, 15(2):023003, 2020.

D. Wichmann, P. Delandmeter, H. A. Dijkstra, and E. van Sebille. Mixing of passive tracers at the ocean surface and its implications for plastic transport modelling. *Environmental Research Communications*, 1(11):115001, 2019.

WW3DG. User Manual and System Documentation of WAVEWATCH III version 6.07, The WAVEWATCH III Development Group. *Tech. Note 326 pp. + Appendices, NOAA/NWS/NCEP/MMAB*, 2019.

J. Zalasiewicz, C. N. Waters, J. A. I. do Sul, P. L. Corcoran, A. D. Barnosky, A. Cearreta, M. Edgeworth, A. Gałuszka, C. Jeandel, R. Leinfelder, et al. The geological cycle of plastics and their use as a stratigraphic indicator of the anthropocene. *Anthropocene*, 13:4–17, 2016.

Y. Zheng, E. K. Yanful, and A. S. Bassi. A review of plastic waste biodegradation. *Critical reviews in biotechnology*, 25(4):243–250, 2005.

A | Observational Surveys

Region	Year	Method	Type	Nb of locations	Total ML [items]	Total ML [kg]	ML Density [item/m^2]	ML Density [g/m^2]	Plastic ratio [%]	Sources sampled ML[a]	ML Seasonal variation	Authors
S (Turkish Coast)	2015-2016	beach survey	macro	1	20132	451.933	1.512 ±0.578	31.875 ±10.684	95.61	Packaging, domestic, fishing, recreational, construction, medical	low	*Öztekin et al. (2020)*
S (Turkish Coast)	2014	beach survey	macro	13	17024	n/a	0.92 ±0.36	7.43 ±2.68	70.3	Land-based local as fishing, packaging, medical/personal hygiene, industrial, personal use, rapid consumption, recreation, smoking.	n/a	*Aydın et al. (2016)*
SW (Turkish Coast)	2008-2009	beach survey	macro	10	18597	n/a	0.8 ±0.94	n/a	91	Land-based local (recreational activities)/ neighboring countries[g] (rivers), sea-based (international shipping[g], Rim current)	high (autumn)	*Topçu et al. (2013)*
SE (Turkish Coast)	2016-2017	beach survey	macro	1	17015	168.9	2.10 ±1.38	21.11 ±11.35	92	**Land-based (62%):** rivers (22%), landfill/dumping (21%), tourism (15%), sewage (4%); **Sea-based (38%):** fishing (18%), shipphing (13%), offshore activities (7%)	high (summer)	*Aytan et al. (2019)*
SE (Turkish Coast)	2009-2018	beach survey	macro	11	4138	108.75	0.09–3.24	3.90–32.65	79.69	Rivers (21.96%), landfill/ improper disposal (21.18%), recreational activities (16.93%)	n/a	*Terzi et al. (2020)*
SE (Turkish Coast)	2012-2013	beach survey	macro	9	5690	108.28	0.16 ±0.02	3.35 ±1.63	61.65	Land-based (73.5%), sea-based (26.5%)	low	*Terzi and Seyhan (2017)*
NW (Bulgarian Coast)	2015-2016	beach survey	macro	4	16690[e]	n/a	0.0587 ±0.005[b]	n/a	84.3	Tourism, fishing, wild camping, recreational activities, rivers	high (summer)	*Simeonova et al. (2017)*
SW (Bulgarian Coast)				4			0.1343 ±0.008[b]	n/a				
NW (Bulgarian Coast)	2015-2016	beach survey	macro	8	19805	n/a	n/a	n/a	67.6	Tourism, fishing, wild camping, recreational activities, rivers	high (summer/winter)	*Simeonova and Chuturkova (2019)*
NW (Romanian Coast)	2014-2017	beach survey	macro	6	13150	n/a	n/a	n/a	80[f]	Hospital (sewage), tourism, urban devlopment, maritime traffic, fishing activities, harbours pollution	n/a	*Muresan et al. (2017)*
NW (Romanian Coast)[d]	2012	trawling net at different depths	macro	69	n/a	554.53	n/a	n/a	2	Shipphing, fishing	n/a	*Anton et al. (2013)*
NW (Romanian Coast)	2014	visual survey at sea	macro	30	225	n/a	3.9*10e-5	n/a	89.1	n/a	n/a	*Suaria et al. (2015)*
NW (Romanian Coast)	2014-2017	beach survey	macro	8	46 024	n/a	n/a	n/a	80.6	n/a	n/a	*Paiu et al. (2017)*
NW (Romanian Coast)	2013	trawling net	macro	16	210	87	n/a	n/a	45.2	Land-based (28.1), vessel (27%), fishing(13.3%)	n/a	*Ioakeimidis et al. (2014)*
NW (Romanian Coast)	2015	trawling net	micro	10	3262	0.000178	2.35	0.00012	n/a	n/a	n/a	*Suaria and Bassotto*

Fig. A: Summary of the observational survey carried out in the Black Sea.

B Marine Litter Source

Country/Sources	Bulgaria	Georgia	Romania	Russia	Turkey	Ukraine	Mean Score
Garbage/Sewage	5	5	5	5	3	4	4.5
Ship waste/ports	4	3	2	3	5	5	3.66
Tourism	3	4	3	4	1	3	3
Rivers	1	-	-	5	5	-	3.66
Industry	1	-	1	-	4	-	2
Fishery	1	1	4	-	-	-	2

Table 3: Likelihood that marine litter comes from a source, 5 is the most likely, specific for the Black Sea. From Birkun et al. (2009).

C | Regional Delimitation

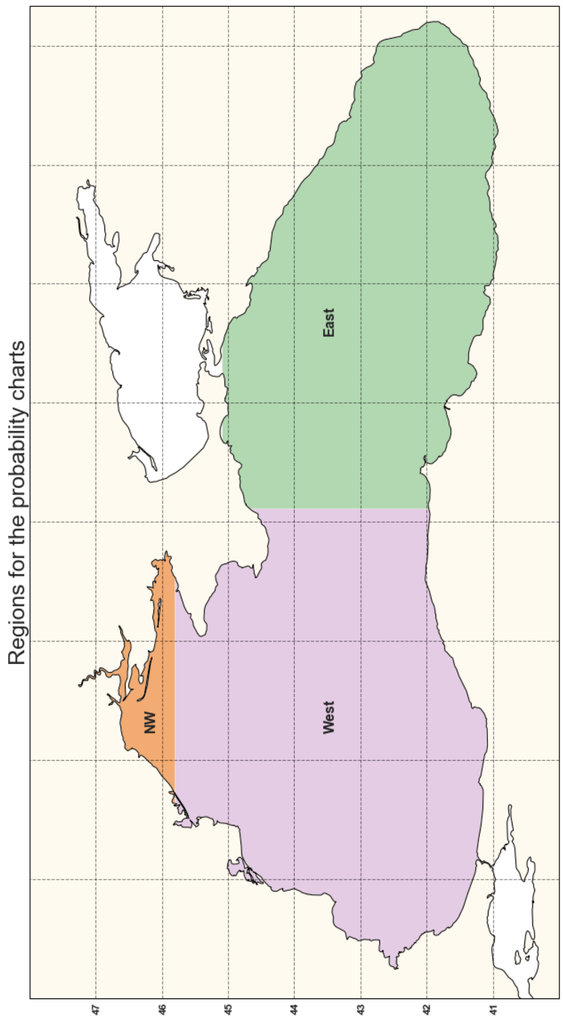
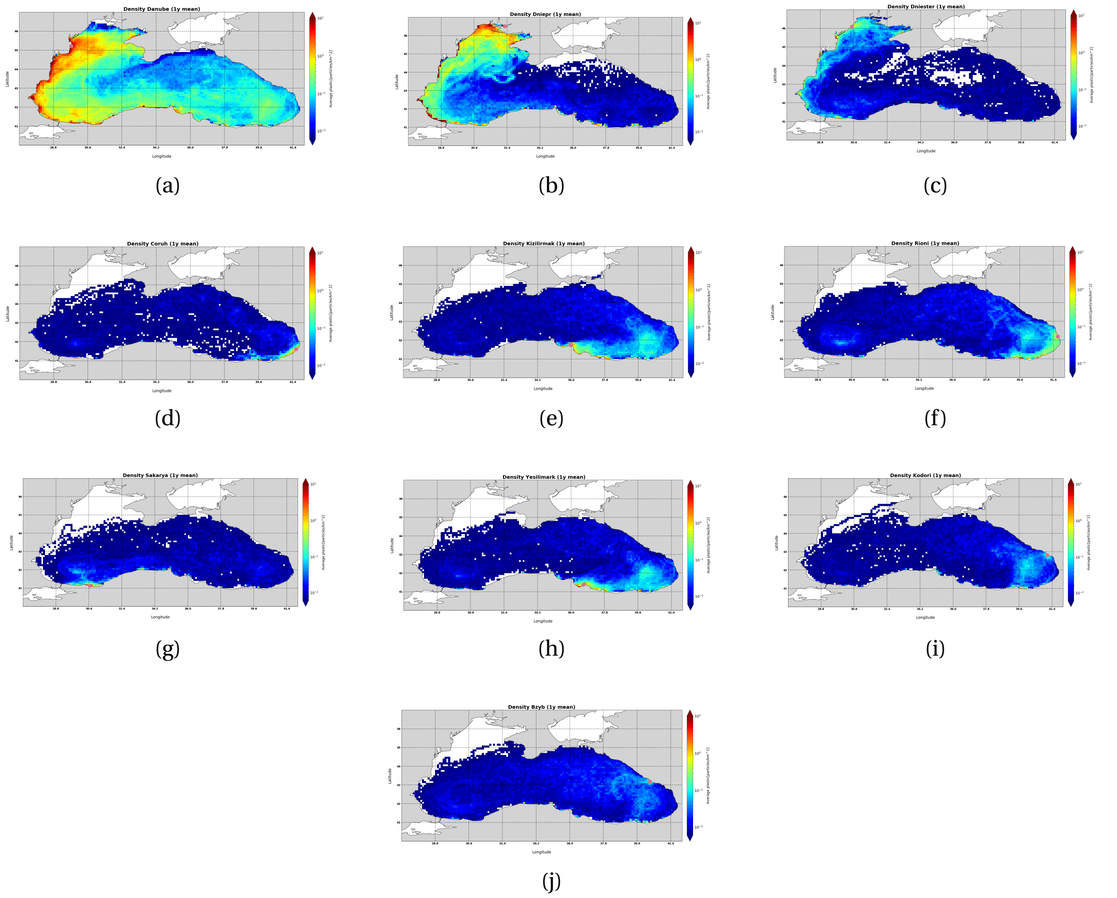


Fig. C: Illustration of the regions used to determine the pie chart of the sources probabilities (Fig.4.2)

D | Density Maps



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

(j)

Fig. D: 1-year average density of particles advected from each source: a) Danube; b) Dniepr; c) Dniester; d) Coruh; e) Kizilirmak; f) Rioni; g) Sakarya; h) Yesilimark; i) Kodori, j) Bzyb

E Source Probability Sampled Data

River/Station	1	2	3	4	5	6	7	8	9	10
Danube	0.84	0.85	0.71	0.79	0.83	0.68	0.77	0.74	0.77	0.81
Dniepr	0.11	0.11	0.23	0.17	0.13	0.27	0.19	0.22	0.19	0.15
Dniester	0.04	0.03	0.05	0.03	0.03	0.03	0.02	0.03	0.03	0.02
Rioni	0.0005	0.0008	0.0017	0	0	0	0	0	0	0
Coruh	0.0007	0	0	0	0	0	0	0	0	0
Kizilirmak	0	0	0	0	0	0	0	0	0	0
Sakarya	0	0.0007	0.0005	0	0	0	0	0	0	0
Yesilimark	0	0.0003	0.0009	0	0	0	0	0	0	0
Kodori	0.0005	0	0	0	0	0.0009	0	0.001	0	0
Bzyb	0	0.0004	0.0006	0	0	0	0	0	0	0

Table 4: Detailed probability values for the plastic sampled in the Black Sea, used to illustrate a possible application of the Bayesian inference results, originates from a specific river at each sampling station

F | Statistical Validation

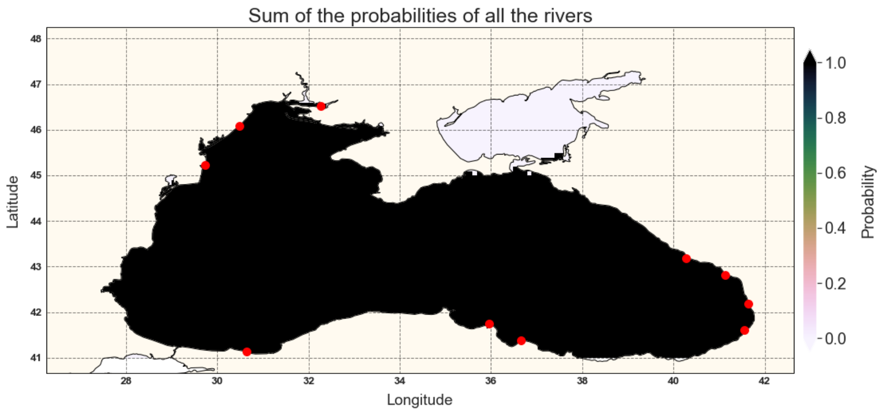


Fig. F: Sum of the probability of all the source at each grid-cells. The red dots represent the location of the source (i.e., the river mouth). Nb: the sum is not precisely 1 for all the cells, it actually ranges from 0.0 to 1.0000000000000004, consequence of some rounding that had to be done to compute the Bayesian inference. However, this divergence is small enough to be tolerated.