Where does plastic in the ocean come from?

Using the end location to find the likelihood of a possible start location of plastic particles in the ocean

L. Goudsmit

Delft University of Technology

TU Delft

Where does plastic in the ocean come from?

Using the end location to find the likelihood of a possible start location of plastic particles in the ocean

(Waar komt plastic in de oceaan vandaan?
De eindlocatie gebruiken om de waarschijnlijkheid van een mogelijke begin positie van plastic deeltjes in de oceaan te vinden)

by

L. Goudsmit

in partial fulfilment of the requirements
to obtain the degree of Bachelor of Science in

Applied Mathematics

at the Delft Institute of Applied Mathematics,
Faculty of Electrical Engineering, Mathematics & Computer Science,
Delft University of Technology.

Student number:	4701186	
Project duration:	April, 2020 – July, 2020	
Thesis committee:	Prof. dr. ir. M. Verlaan,	Supervisor
	Prof. dr. ir. C. Vuik,	
	Drs. E.M. van Elderen	

TU Delft

Abstract

In the ocean resides a large amount of plastic. This has severe environmental consequences for the oceans. To be able to clean up all this plastic the location of the plastic is needed. A lot of times plastic is cleaned up at the source, but to clean up here this location needs to be known. However, what if the start location is not known and only the end location is known.

This thesis looks at the following research question: *Can the method of particle filtering be used to find the start location of a plastic particle of which the end location is known?*. The method of particle filtering compares the end locations of the released particles and the end location of the end particle. (The end particle is the particle of which the end location is known and we want to know the start location.) The method calculates weights on the basis of this comparison for each of the released particles. Then the weights are used to calculate a probability density function that gives the likelihood of possible start locations being the right start location of the end particle.

The method is applied to the MSC Zoe case, which is a container accident, and a simplification of this case. The MSC Zoe lost in this accident hundreds of containers above the Wadden Islands in the Netherlands, in these containers were bags of plastic particles. A lot of these particles washed ashore on the Wadden Islands. However, it is not known exactly where all the containers fell overboard. So the start location of the plastic particles that washed ashore is not known. The method of particle filtering applied to the simplification showed that it can indeed be used to find the likelihood of possible start locations. However, it also showed that for some values of the variables it was possible to get multiple peaks with weights that are zero in between. This is not a desired outcome for a probability density function (pdf). So a kernel estimation is used to smooth out the pdf's. It uses the weights calculated in the method of particle filtering to estimate the pdf. This gives results that also in this way the likelihood of possible start locations can be found and now without multiple peaks. The kernel estimation was also applied to the MSC Zoe case and the results were calculated for multiple end particles. The locations that came out of this were very close to a few of the possible main locations were containers fell overboard. These possible main locations came from the international investigation into the MSC Zoe accident.

A simplification of this research is that the start location is found for one specific plastic particle. For future research it could be interesting to look if the method can be adapted to find the start location of concentrations of plastic particles.

L. Goudsmit
Delft, July 2020

Contents

Nomenclature

List of Abbreviations

Acronym	Description
AIS	Automatic Identification System
BSU	The German Federal Bureau of Maritime Casualty Investigation (Bundesstelle für Seeunfalluntersuchung)
DCSM-FM	Dutch Continental Shelf Model – Flexible Mesh
DSB	Dutch Safety Board (Onderzoeksraad Voor Veiligheid (OVV))
HDPE	High-Density-Polyethylene
LT	Local Time, UTC+1, (time in the Netherlands)
MSC	Mediterranean Shipping Company
PDE	Partial Differential Equation
pdf	Probability density function
PMA	Panama Marine Authority
SDE	Stochastic Differential Equation
TSS	Traffic Separation Scheme
UTC	Standard time
VDR	Voice Data Recorder

List of Symbols

Symbol	Description	Units
δ	Dirac delta function	–
Δt	Timestep	$seconds\ (s)$
ϵ	(ϵ_x, ϵ_y)	–
ϵ_x	Eddy diffusivity in the x- or long-direction	m^2/s
ϵ_y	Eddy diffusivity in the y- or lat-direction	m^2/s
$\hat{p_1}, \hat{p_2}$	The kernel density estimations	–
$\mathcal{N}(\mu, \sigma)$	Normal distribution with mean μ and standard deviation σ	–
ψ	State vector (A possible overboard location)	–
ψ_i	A model state that is followed (a released particle that is followed)	–
Σ	Covariance matrix	–
σ	The standard deviation that is used in the normal distribution in the kernel function	$meters\ (m)$

A	$(2\pi)^{\frac{-k}{2}} * det(R)^{\frac{-1}{2}}$	$-$
a_x	Wind velocity in the x- or long-direction	m/s
a_y	Wind velocity in the y- or lat-direction	m/s
B	constant	$-$
$c(t,x,y)$	concentration of particles at that time and place	$\#particles\ per\ x, y-coordinate$
d	Observations of the geophysical system (The position of the particle of which we only know the end position and want to find the start position)	$-$
f_1, f_2, g_1, g_2	Functions that are used to model the velocity of a particle	$-$
G_s	A stochastic process	$-$
$H(\psi_i)$	Measurement operator	$-$
K	Kernel function	$-$
k	Length of the vector d	$-$
l	The long- and lat-coordinates for which the likelihood of it being a start location is desired	$-$
L_i	The long- and lat-coordinates of the start location of particle i	$-$
lat	Latitude	$degrees, °$
lon	Longitude	$degrees, °$
N	The number of independent model states (the number of particles released)	$-$
n	Refers to the n^{th} timestep	$-$
N_t	A white noise process, a stochastic process	$-$
$p(\,)$	Probability density function	$-$
R	The error covariance of the observations	$-$
r	A value used in the formula of R	$-$
t	Time	$seconds\ (s)$
t_0	Start time	$seconds\ (s)$
t_{end}	End time	$seconds\ (s)$
u_x	Ocean velocity in the x- or long-direction	m/s
u_y	Ocean velocity in the y- or lat-direction	m/s
w_i	Weight of model state i (weight of particle i)	$-$
W_t	A Wiener process	$-$
x	Horizontal coordinate	$meters\ (m)$
X_t	Stochastic process for the x- or $long$-coordinate	$-$
y	Vertical coordinate	$meters\ (m)$
y_i	y-coordinate of the start location of particle i	$-$
Y_t	Stochastic process for the y- or lat-coordinate	$-$
z	Depth	$meters\ (m)$

1 Introduction

More than 5 trillion pieces of plastic can be found in the ocean. They mostly accumulate in the 5 big plastic garbage patches in the ocean, (The Ocean Cleanup, n.d.). Every year around 8 million ton of plastic ends up in the ocean, adding to this total. This has large environmental impacts on the oceans, (The Ocean Conservatory, n.d.). For example, animals in the sea can get entangled in pieces of plastic. Or they can swallow pieces of plastic which can lead to blockages or even death, (Wageningen University & Research, 2019a). In figure 1.1 is shown what could be the difference between cleaning up the plastic or letting it float away in the oceans according to The Ocean Cleanup, (The Ocean Cleanup, n.d.).

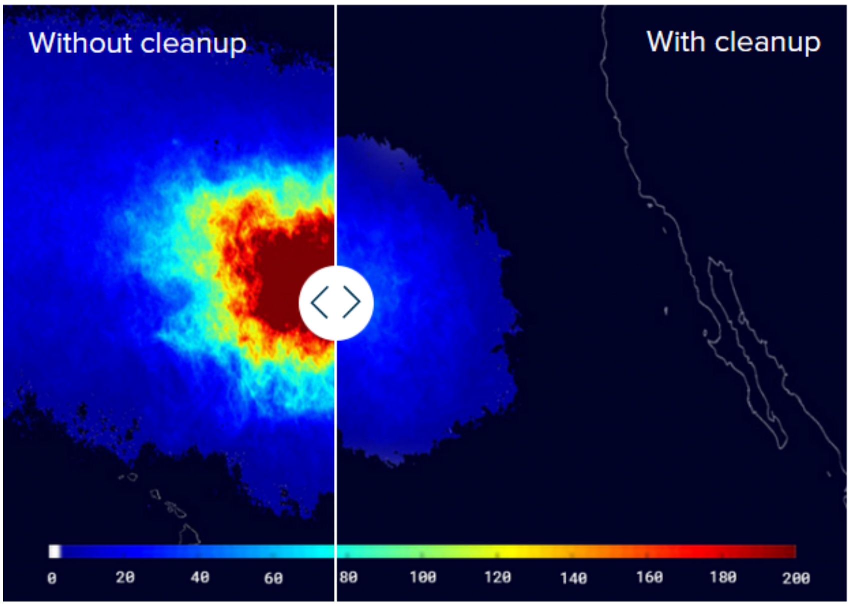


Figure 1.1: How the Great Pacific Garbage Patch could look like with and without cleanup in the year 2030. This Garbage Patch lies between Hawaii and California. The unit of the scale is kg/km^2, (The Ocean Cleanup, n.d.)

There are a few different ways to clean up the ocean. One of those is to clean near the source of the plastic or to make sure it never ends up in the ocean to begin with. To reduce the amount of plastic at the source, the location of the source has to be known. However, what if we do not know the start location and only the end location of some plastic and we would like to know where it originated from. To either clean up on this location or model the transport of plastic from there to see where it can end up.

This study will aim at an estimation of the initial position of plastic particles, if the end location is known. This is done by adapting the method of particle filtering to give weights to particles, that are released from different possible start locations, on the basis of their end location compared to the end location of the particle for which we want to know the the start location. For future reference, the particle for which only the end location is known and we want to find the start location, will be called an end particle. These weights are then used to calculate a probability density function of the likelihood of possible start locations of the end particle. This leads to the following research question:

Can the method of particle filtering be used to find the start location of a plastic particle of which the end location is known?

To test this method the specific case of the MSC Zoe disaster is looked at. Last year in January 2019 the container ship MSC Zoe had an accident above the Wadden Islands in the North Sea, here it lost 342 containers. Among the cargo in the containers that went overboard, were a lot of plastic products, and specifically micro plastics, (Panama Marine Authority et al., 2020). Which are plastics smaller than 5 mm, (Scheijen, 2018). Of these plastic particles only the end location is known. They washed ashore on for example the coast of the Wadden Islands of the Netherlands, a picture of this can be seen in figure 2.7. However the precise start location of the plastic particles is not known. So the method of particle filtering is applied to this specific case.

Figure 1.2: Plastic particles found on the beach after the MSC Zoe disaster, (Plastic Soup Foundation, 2019)

The structure of this report is as follows. In chapter 2 the MSC Zoe case is explained in a bit more detail, what happened that night. Additionally, some information is given about the plastic particles in question. In chapter 3 the particle filtering method is discussed. The method that is used to calculate a probability density function of the likelihood of possible start locations for the plastic particles of which we only know the end location. In the next chapter, chapter 4, the random walk model is explained. This model is used to model the transport of the plastic particles. In chapter 5 is an example to show how the method of particle filtering works and what the results are in a simplification of the MSC Zoe case. Furthermore, the effects of different values for a few variables are shown. Then in chapter 6 there is an overview of the data that is used to apply the method of particle filtering to the MSC Zoe case. Also the results are given from the application of the method of particle filtering to the MSC Zoe case. Lastly, in chapter 7 are the conclusion, discussion and recommendations for this report.

2
The MSC Zoe case

In this chapter is the MSC Zoe case described. Furthermore, some information is given about the plastic particles that were on board of the MSC Zoe and that fell overboard. The last part of this chapter contains a few consequences of the disaster pertaining to plastic particles.

2.1. What happened

In the night of 1 to 2 January 2019 the containership MSC Zoe lost 342 of the 8062 containers it had on board at the time. The ship belongs to the Mediterranean Shipping Company (MSC), and was at about 70% of its full capacity. At the time of the accident the ship was sailing in the North Sea, above the Wadden Islands of the Netherlands. It was coming from Sines, Portugal and was en route to Bremerhaven, Germany. Furthermore, it was sailing under Panamanian flag. This is why the investigation into the accident was done by the Panama Maritime Authority (PMA), the Federal Bureau of Maritime Casualty Investigation of Germany (BSU) and the Dutch Safety Board (DSB). The investigation reports of both the international investigation and another investigation of the DSB came out on June 25th 2020. Just in time to be included in this research, (Dutch Safety Board, 2020; Panama Marine Authority et al., 2020)

The first time that the crew of the ship detected that containers had fallen overboard, was around 01.00 hours local time (LT), which is UTC+1, on January 2. They noticed that not all containers were in their expected position. The one time that the crew saw containers falling overboard was at 01.30 hours LT, they informed German Bright Traffic of this. At this time they thought that only about 30 containers had fallen overboard. In reality, the first containers had fallen of shortly before 20.00 hours LT and a total of 342 containers had fallen of. At the time of the accident they were sailing on the southern route of the Traffic Separation Scheme (TSS) Terschelling - German Bright, this is the shortest route of the two routes and the closest to the coast. It had entered the southern track at around 18.50 hours LT. So at about 01.30 LT they changed the course of the ship to a northwesterly heading and slowed down. They headed towards the TSS German Bight Western Approach, were they arrived around 14.00 hours LT on 2 January. The two sailing routes can be seen in figure 2.2. Another difference between the two sailing routes, besides distance to shore, is that the northern route is slightly deeper, (Dutch Safety Board, 2020; Panama Marine Authority et al., 2020). Furthermore there is a more detailed timeline in figure 2.1. and in figure 2.3 there are pictures of the MSC Zoe in the daylight after the accident.

The weather conditions at the time of the accident were pretty heavy, but not extreme for this section of the North Sea. There was a north-northwesterly wind up to 8 Bft (Beaufort), between 16-18 m/s, and also gusts of wind. Additionally, there were high waves, the significant wave height was around 5 m and at the peak of the storm it went up to 6.5 m. There was also the possibility of individual waves of up to 11 meters. This was also not very extreme. Lastly, the ship was rolling the whole time, which is a motion around the longitudinal axis, it is tilting towards the port and starboard sides. The ship had a period of violent rolling around 23.00 hours, (Dutch Safety Board, 2020; Panama Marine Authority et al., 2020).

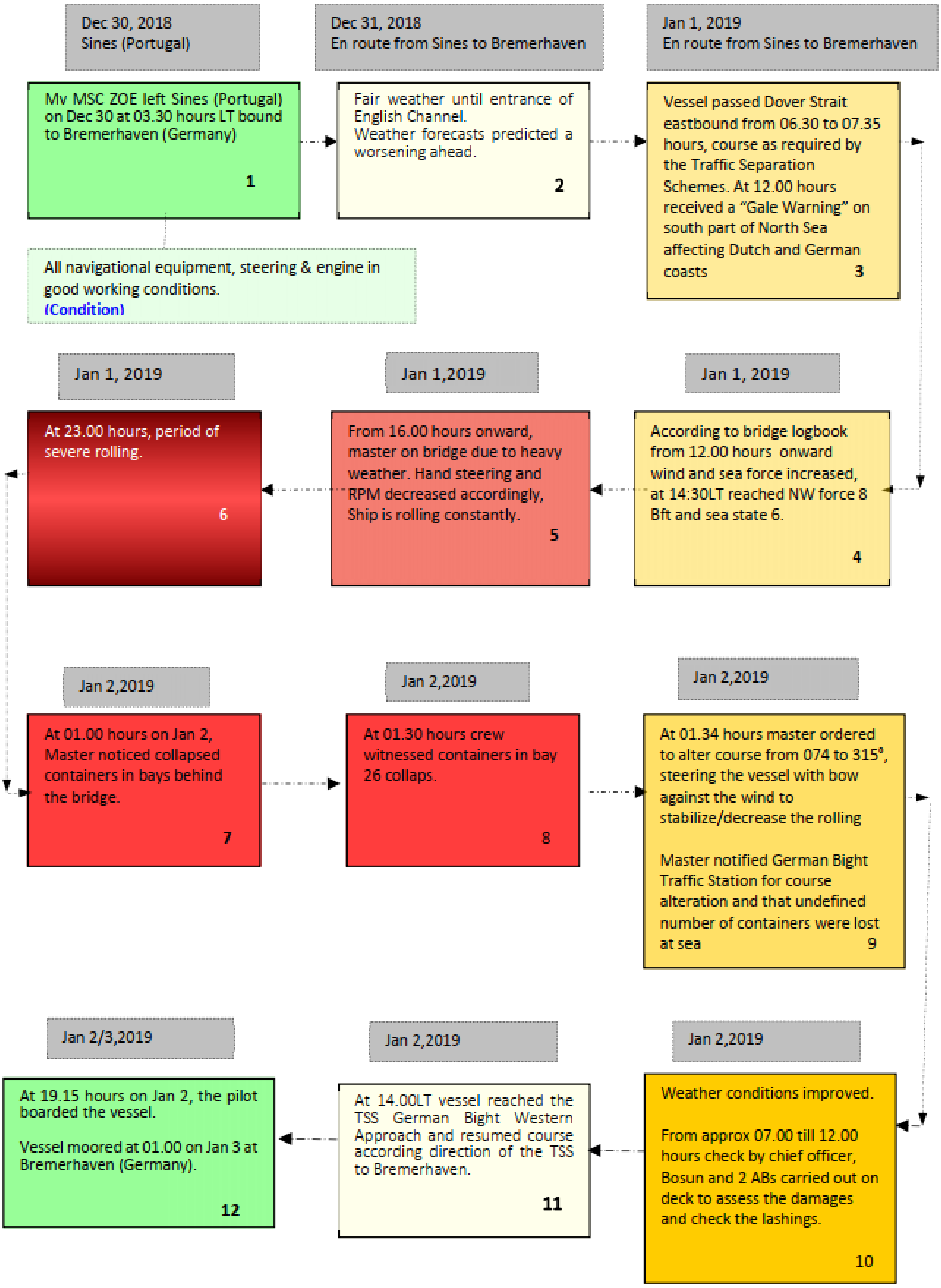


Figure 2.1: The timeline of the MSC Zoe accident according to the crew and the VDR (Voice Data Recorder). The VDR records for example the audio from the bridge of the ship, (Panama Marine Authority et al., 2020)

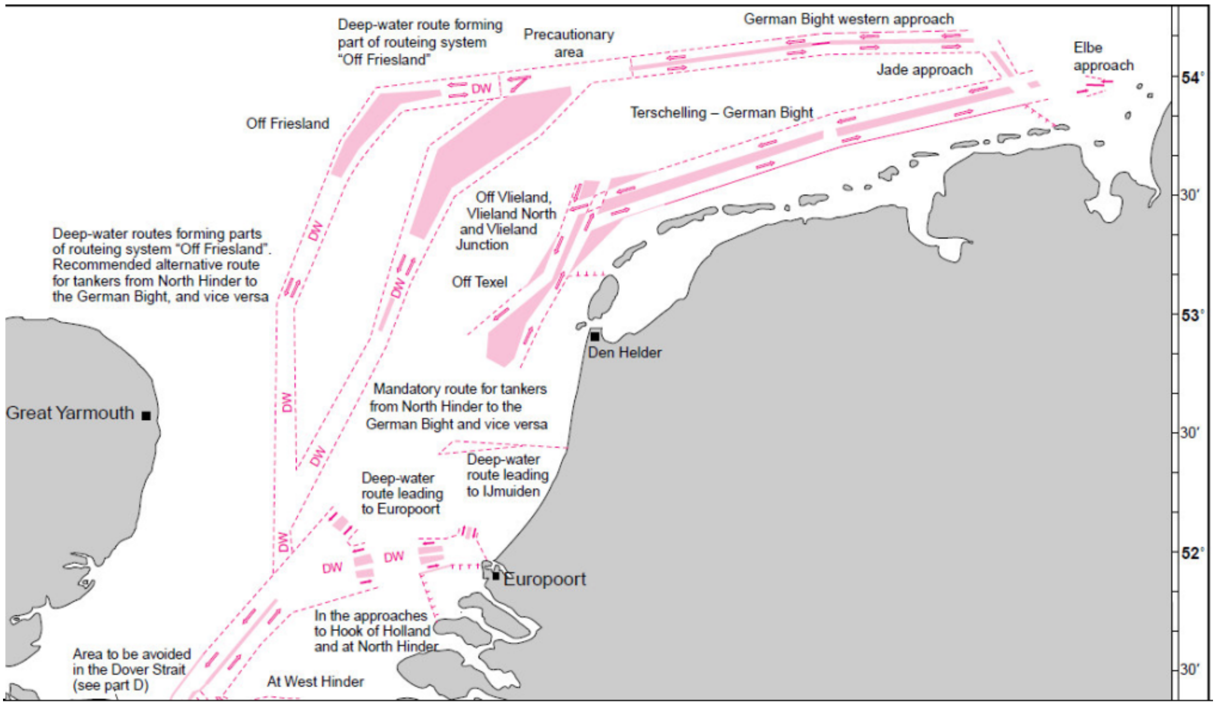


Figure 2.2: The shipping routes, including the TSS Terschelling-German Bight, (Panama Marine Authority et al., 2020)

Figure 2.3: Pictures of the MSC Zoe after the disaster, (Panama Marine Authority et al., 2020)

The recovery locations of the containers that were found have been plotted along the track of the MSC Zoe, this can be seen in figure 2.4. It shows that a lot of the containers were found at six main locations. In figure 2.5 the bays of the ship with lost containers are shown, with the same colours as in figure 2.4. These locations only pertain to the containers that were found, (Panama Marine Authority et al., 2020). In this research the origin location of the plastic particles found on the shore is looked at. It is not known where these fell overboard. This is partly due to the fact that it is not known in which containers the plastic particles were, because the shipping manifest is not public information. The plot of the locations of the found containers will be compared to the results in chapter 6.

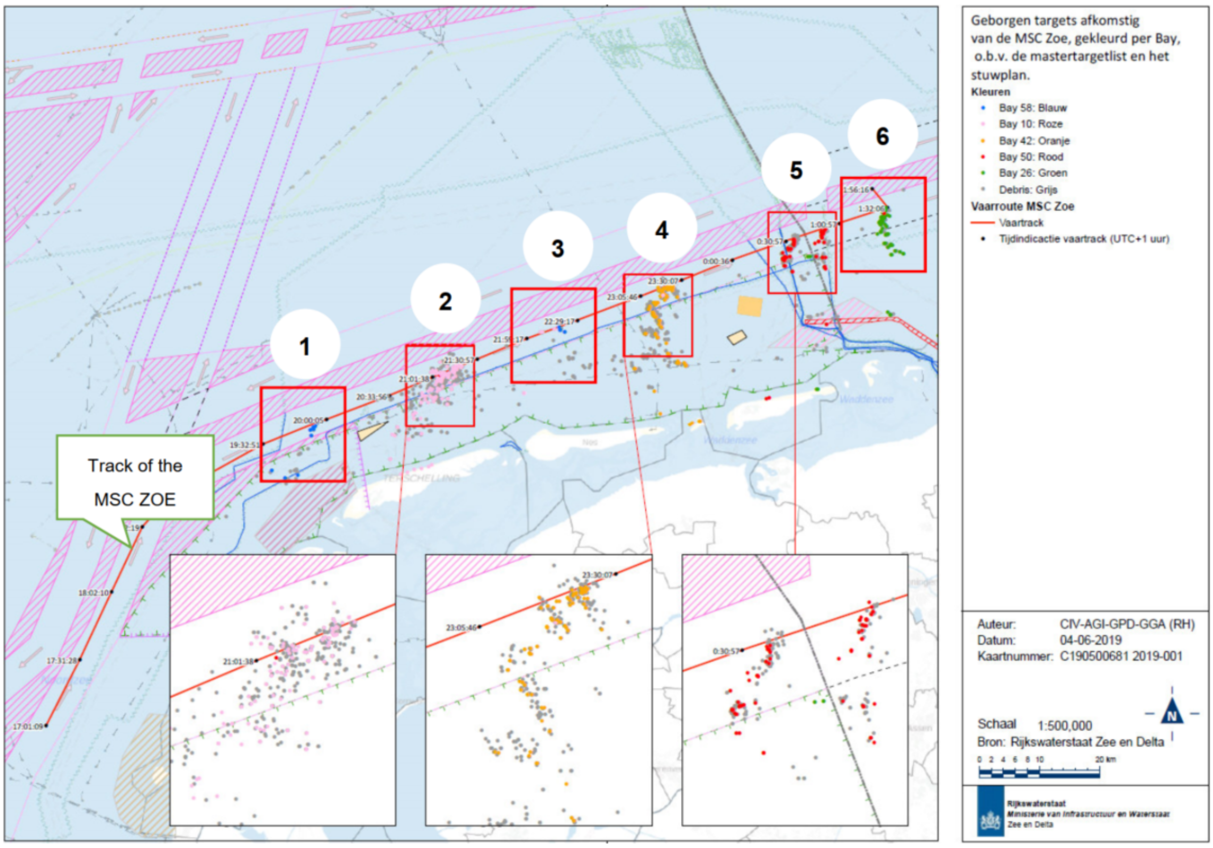


Figure 2.4: The six main locations that were identified for places that containers were found at in the ocean. The colours indicate bays which in turn indicate where on board the containers were, (Panama Marine Authority et al., 2020)

Figure 2.5: The MSC Zoe from above with all the bays and in colour where the containers that were lost came from, (Panama Marine Authority et al., 2020)

The investigation resulted in a few phenomena that they consider to have played a role in the accident and the probably six times that containers have fallen overboard. First there are the extreme motions and accelerations of the ship, next there is the point that the ship may have hit the bottom or has nearly hit the bottom. Furthermore, the investigators suspect that green water has also played a role in the accident. Green water means that water from the waves goes over the deck of the ship and/or slams against the cargo on the deck of the ship. The last reason is impulsive hits of the waves against the side of the ship. The results of the investigations says that one or more of these phenomena can be responsible for the accident of the MSC Zoe, (Dutch Safety Board, 2020; Panama Marine Authority et al., 2020).

2.2. Plastic particles

In this report the plastic particles that fell overboard in this disaster will be looked at. According to the investigation reports about the disaster there were different kinds of plastic particles on the ship. There were bags full of these plastic particles in the containers. One of the containers that had burst open contained 22.5 ton of polymeric beads. These are about 0.5 mm. Additionally, large amounts of HDPE, high-density-polyethylene, pellets were found. These pellets are about 4-5 mm, (Dutch Safety Board, 2020; Panama Marine Authority et al., 2020). A picture of these pellets and the bag that they were held in can be seen in figure 2.6. In this research the HDPE pellets are looked at, because there is data available on the end location of these plastic particles. The University of Groningen launched a website (Waddenplastic) on which people can indicate how many of these particles they found on the coast. They estimate that there are around 24 million of these plastic particles in the Wadden, (Dutch Safety Board, 2020; University of Groningen, 2019). More information on the end location of these HDPE pellets is given in chapter 6, where the data is used.

(a) HDPE pellets collected on the Wadden Islands, (Wageningen University & Research, 2019b)

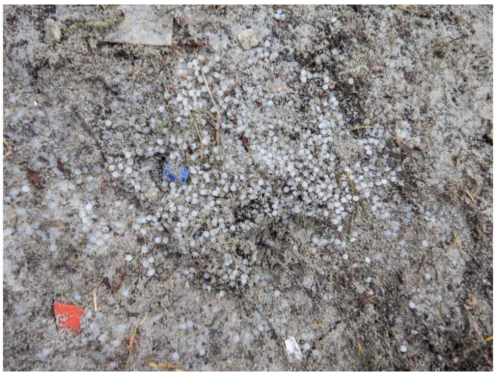
(b) A bag in which the HDPE pellets were contained, the bag is 25 kg, (Wageningen University & Research, 2019b)

Figure 2.6: Plastic particles - HDPE pellets

These plastic particles are called micro plastics, because they are smaller than 5 mm, (Scheijen, 2018). These HDPE pellets have a density of approximately 940 kg/m^3, (Ooms, 2020). Sea water has a density of around 1025 kg/m^3, which means these micro plastics have a floating character, (Scheijen, 2018).

2.3. Consequences of the disaster in regards to plastic particles

The immediate consequences of the disaster were already visible on January 2nd. A lot of the containers that fell overboard were damaged and smashed by the height of the fall and the waves. This meant that a lot of the cargo and residue of the cargo washed ashore on the coast. This happened on the coast of the Dutch and German Wadden Islands and also on the coast of the Dutch mainland. The Island that was hit the hardest was Schiermonikoog. The larger pieces of cargo are easy to clean up, but the plastic particles, HDPE pellets, are hard to remove because they are so small, see figure 2.7. Moreover, not only the fact that a lot of cargo fell overboard matters. Namely, the Wadden Sea is on the World Heritage List of the United Nations as a special nature reserve, it is a complex ecosystem with a lot of diversity, which makes it very vulnerable, (Dutch Safety Board, 2020; Nijpels, 2019a; Panama Marine Authority et al., 2020).



(a) HDPE pellets disappearing in the sand, (Wageningen University & Research, 2019b)

(b) HDPE pellets on the beach (Panama Marine Authority et al., 2020)

Figure 2.7: Plastic particles - HDPE pellets in the sand

If we look at the plastic particles in particular, then they have a lot of consequences for the nature, (Panama Marine Authority et al., 2020). Also the micro plastics on which this report focuses. These micro plastics can be ingested by sea animals. Which in turn can lead to a number of things, for example transferring these micro plastics through the food chain or the death of a sea animal. Furthermore, there are a lot of risks to the environment by these micro plastics, because they are difficult to degrade. The risks for humans are not yet very clear, but micro plastics can be ingested by eating for example shellfish. However for it to be harmful humans would need to ingest large amounts of micro plastics, (Dutch Safety Board, 2020; Wageningen University & Research, 2019a).

3

Particle Filtering Method

In this chapter the method of particle filtering will be discussed. First the theory will be given and after that the application of this method to the situation of finding the start location of a plastic particle when only the end location is known. At the end of the chapter will be a method to better estimate the result of the theory.

3.1. Theory of the Particle Filtering method

This section describes the parts of the theory of particle filtering that are relevant for this application. It is mainly based on the review article of Peter Jan van Leeuwen, (van Leeuwen, 2009). In the method they use the word particle for one run of a whole model, a model state. In this report it will instead be called a model state to avoid confusion with plastic particles on which the method will later be applied. The goal of the method of particle filtering is to calculate a probability density function that calculates the likelihood of a model run given the observations of a geophysical system for which the model is created, $p(\psi|d)$. d stands for the observations of the real system, the geophysical system for which a model is wanted. ψ is called the full-state or the state vector, in the vector are all the variables of the system. Furthermore, ψ is described by $p(\psi)$, this is the joint probability density of all variables. Moreover, $p(\psi)$ is also called the prior pdf, because it is the probability density of the model before the observations are taken into account.

To start, $p(\psi|d)$ can be rewritten with the help of Bayes rule, this can be seen in the following equation:

$$p(\psi|d) = \frac{p(d|\psi)p(\psi)}{p(d)} \tag{3.1}$$

In this equation, $p(d)$ is the probability density of the observations d and $p(d|\psi)$ is the probability density of the observations given ψ. The density $p(\psi)$ can be approximated by N random samples, random model states, which are drawn from the density $p(\psi)$. These random samples are the ψ_i's. The representation of $p(\psi)$ is then as follows:

$$p(\psi) = \frac{1}{N}\sum_{i=1}^{N}\delta(\psi - \psi_i) \tag{3.2}$$

The probability density function $p(d)$ can be written as the first line of equation 3.3. So we can see that $p(d)$ in equation 3.1 is the normalization of $p(\psi|d)$, $p(\psi|d)$ is called the posterior pdf. After that we fill in the definition of $p(\psi)$, this is also shown in equation 3.3.

$$\begin{aligned} p(d) &= \int p(d,\psi)d\psi = \int p(d|\psi)p(\psi)d\psi \\ &\overset{def\ p(\psi)}{=} \frac{1}{N}\sum_{j=1}^{N} p(d|\psi_j) \end{aligned} \tag{3.3}$$

Now the definitions of $p(\psi)$ and $p(d)$ are filled in into equation 3.1, the formula in equation 3.4 is the result. The weights w_i in this formula are given by equation 3.5. These weights are the relative importance of the model states in the probability density.

$$p(\psi|d) = \sum_{i=1}^{N} w_i \delta(\psi - \psi_i) \tag{3.4}$$

$$w_i = \frac{p(d|\psi_i)}{\sum_{j=1}^{N} p(d|\psi_j)} \tag{3.5}$$

In equation 3.5 we still have the probability density $p(d|\psi_i)$ that is unknown, so of the observation given we have the model state ψ_i. This is usually taken as a Gaussian distribution, this is given in equation 3.6.

$$p(d|\psi_i) = Ae^{-\frac{1}{2}[d-H(\psi_i)]^T R^{-1}[d-H(\psi_i)]} \tag{3.6}$$

In this equation are still a few unknown variables, these are defined as follows. R is defined as the error of the covariance of the observations and $H(\psi_i)$ is defined as the measurement operator, which is also said to be "the model equivalent of the observation d", (van Leeuwen, 2009).

If ψ has a subscript like ψ^n then the n stands for the n^{th} time step.

The procedure to calculate the weights (and implement this theory) goes as follows:

1. To start N model states are sampled from the density $p(\psi^0)$, this is the initial model probability density.
2. Then all the model states are integrated one step in time.
3. Now the weights are calculated for each model state and they are linked to each model state. Note that nothing of the model states is adapted, only the relative weight of the model state is changed.
4. Now take one more time step and repeat steps 2 and 3 until the end time is reached.

This procedure is also depicted in figure 3.1. In this figure can be seen that the model states move along the vertical axis, the dots. The weight is calculated at every time step. The time goes over the horizontal axis. The size of the weight is represented by the size of the point. At time $t = 0$ all variables have weight 0. Along the vertical lines at time $t = 10$ and $t = 20$ one can also see the likelihood calculated from the weights at that time.

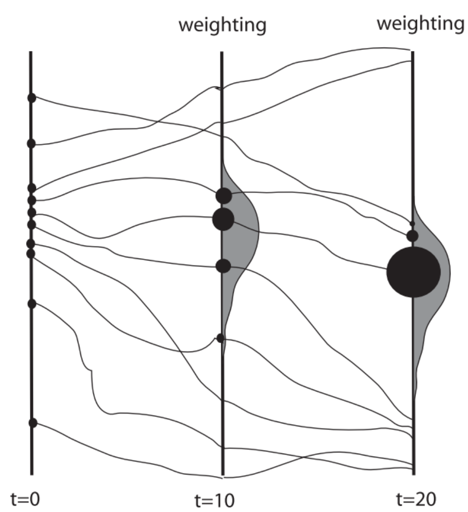


Figure 3.1: A representation of the particle filtering method/procedure, (van Leeuwen, 2009).

3.2. Application to the transport of plastic particles

In this project the method of particle filtering is not used to look for the most likely model run/state. It is used to find the most likely start location/overboard location of plastic particles, while only their end location is known. In this report the choice has been made to do a simplification of this case. Namely, the start location is modeled for one particle, so for one object. This is a simplification because for the MSC Zoe case we want to know the start location of multiple particles, a concentration of particles. This is explained a bit more at the start of chapter 5.

The symbols from the previous section will get slightly different definitions, which will be given now. The formulas to calculate the variables will stay the same unless stated otherwise. The variable d becomes the end location of the end particle, the particle from which we want the start location. ψ stands for the possible locations the particles fell overboard. So when ψ is filled in into the formula (equation 3.4) then a location is filled in for which it is desired to know how likely it is that this is the location that the particles fell overboard. Thus $p(\psi|d)$ is the probability density function of the possible overboard locations given the location of the end particle. When applying the method an assumption needs to be made for the pdf $p(\psi)$. A distribution needs to be chosen and the start location of all the ψ_i's need to abide by it. ψ_i is a plastic particle that is tracked from a possible overboard location to its end location to see if it might be a correct overboard location. In one run of the model multiple particles are tracked from their start location to their end location and weights w_i are attached to them. The symbol $H(\psi_i)$ will be taken as the information from ψ_i that can be compared to the information from d, so the current location (their end location) of the plastic particle ψ_i. R will still be seen as the error covariance of the observations and different values will be tested in the example in chapter 5.

Lastly, there is the symbol A, this is taken to be the formula of the coefficient in front of the exponent in the multivariate Gaussian distribution formula, (Do, 2008). This has the following formula in front of the exponent: $(2\pi)^{\frac{-k}{2}} * det(\Sigma)^{\frac{-1}{2}}$, adapting this to the formula in equation 3.6 will give equation 3.7. The variable Σ (the covariance matrix) is replaced by R, which has the place of Σ in equation 3.6.

$$A = (2\pi)^{\frac{-k}{2}} * det(R)^{\frac{-1}{2}} \tag{3.7}$$

With k being the length of the vector d and the length of the vector $H(\psi_i)$.

With all the formulas known, there are some assumptions that need to be made to choose values for R. Namely, the x- and y-coordinates are assumed to be independent, so R becomes a diagonal matrix, (the dimensions of R are $k \times k$). Moreover, it is assumed that x and y have the same error. So R becomes the following formula in equation 3.8. In this equation r is a number for which different values will be tested in the example in chapter 5.

$$R = r * \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{3.8}$$

In the MSC Zoe case is R a little bit different. This is because now there are longitude and latitude coordinates instead of the projected x- and y-coordinates from the example. This means that r which is in meters needs to be converted to degrees. This is done with a coordinate transformation. This does mean that the error is not exactly the same for both the long- and lat- coordinates. Which is due to the size of a circle around the earth, that is different in longitude and latitude direction in the North Sea. So, the two values in the matrix are slightly different.

An adaption to the method of particle filtering itself, is that the weights are not calculated at every time step, but only at the end. This is because we are only interested in the end location of the particles that are let loose in regards to their start location. So where they are in between does not matter, and only their end location has to be compared to the end location of the particle from which we want the start location.

The formula in equation 3.4 to calculate the probability density function $p(\psi|d)$ is interpreted as follows. The weights w_i are plotted against the start location of the particle to which it is attached. Then the surface under the graph is calculated numerically and the values of the weights are divided by the surface. The new values are then plotted against the start location of the particle, to which the weight was attached. This is done such that the surface under the plot is 1, so to make sure the plot gives a probability density function.

The method of particle filtering will be applied with these adapted definitions of the variables. First in a example in Chapter 5 and later in the case of the MSC Zoe disaster in Chapter 6.

3.3. Kernel Density Estimation

In this section Kernel Density Estimation will be explained. The kernel function can be used to estimate a probability density function, (Heemink, 2019). In this case it will be used to estimate the probability density function $p(\psi|d)$, from this chapter. The reason why this is needed will be shown in the example in chapter 5.

In the example, in chapter 5, only the y-coordinate of the plastic particles matters for their start location. This is because they all start on the same line, have the same x-coordinate. So the pdf $p(\psi|d)$ is only plotted against the start y-coordinate. The formula for the kernel density estimation for this pdf is given by equation 3.9, (Ruiter, 2020).

$$\hat{p_1}(y) = \frac{1}{N}\sum_{i=1}^{N} K(y - y_i) \tag{3.9}$$

In this equation N stands for the number of plastic particles that are released in the simulation. y stands for the vertical coordinate for which we want to know the likelihood of it being the start location of the end-particle. K is the Kernel function used. The Kernel function has the following property: $\int_{-\infty}^{\infty} K(x)dx = 1$. Lastly, y_i stands for the y-coordinate of the start location of particle i. However, for it to estimate $p(\psi|d)$ it needs to include the weights that are calculated in this pdf. This is added in equation 3.10, where w_i stands for the weight of particle i according to the formula in section 3.1.

$$\hat{p_1}(y) = \sum_{i=1}^{N} w_i * K(y - y_i) \tag{3.10}$$

The Gaussian distribution is chosen for the kernel function K in the example. With mean $\mu = y_i$ and in the example it will be tested for different values of the variance σ.

For the MSC Zoe case the bivariate kernel density estimation is used. It works in the same way, only now the long- and lat-coordinates of the start location matter. So we get equation 3.11, (Ruiter, 2020). Where l stands for the long- and lat-coordinates of the location for which we want to know the likelihood of it being the start location of the end particle. L_i stands for the long- and lat-coordinates of the start location of particle i.

$$\hat{p_2}(l) = \frac{1}{N}\sum_{i=1}^{N} K(l - L_i) \tag{3.11}$$

In the same way as in equation 3.9 weights (w_i) are added, s.t. it simulates the pdf $p(\psi|d)$. This resulted in the following equation:

$$\hat{p_2}(l) = \sum_{i=1}^{N} w_i * K(l - L_i) \tag{3.12}$$

For the MSC Zoe case the bivariate normal distribution is chosen for the kernel function K. With mean $\mu = L_i$ and standard deviation $\sigma = 100$ meters. The standard deviation is also here converted with coordinate transformation, same as with r in the previous section.

4

Random Walk Model & Eddy Diffusivity

In this chapter more mathematical theory will be discussed. To start the Random Walk model will be shown, a model used to model the transport of the particles in every time step. After that will be a more thorough explanation of the eddy diffusivity term used in the model.

4.1. Random Walk Model

In the Random Walk model stochastic differential equations (SDE) are used instead of partial differential equations (PDE) to describe the transport of the plastic particles. So the SDE's are used to model the spread of the plastic particles. This is done because now noise can be added to the model. Which is needed as a result of errors and uncertainties in the model of the flow. The model is based on information from the BEP report of Nelleke Scheijen and the course notes of TW3750TU from Arnold Heemink, (Heemink, 2019; Scheijen, 2018).

4.1.1. Stochastic Differential Equations

A SDE with only an advective term is given by formula 4.1. An advective term stands for the spread of particles from where they are dropped to a high concentration further down the flow, so in one direction.

$$\frac{dX_t}{dt} = f(X_t, t) \tag{4.1}$$

In this equation is X_t a stochastic process that represents the location of the variable. A stochastic process is something that evolves over time in a random process, (Spieksma, 2016). t stands for the time and f is a function for the velocity of the particle. Furthermore, there will also be a turbulent diffusivity term that is added to the model. Diffusion means a random spread of particles from areas with a lot of particles to areas with a low amount of particles, it happens because of molecular motion. This term is added to the model in equation 4.2. An initial condition is also added in this equation.

$$\begin{cases} \frac{dX_t}{dt} = f(X_t, t) + g(X_t, t)N_t \\ X_{t_0} = X_0 \end{cases} \tag{4.2}$$

The model that is given in equation 4.2 is called a random walk model. With this kind of model the paths of individual particles can be calculated. The function $g(X_t, t)$ in equation 4.2 has a relation to the turbulent diffusivity and eddy diffusivity. More explanation of turbulent and eddy diffusivity will be given in section 4.2. The other new term is N_t, which is a stochastic process. It is called a white noise process and it is there to model uncertainties in the displacement of the particle. There are some difficulties with calculating N_t, because it is not properly defined in mathematics. Before that will be discussed, a different notation of equation 4.2 will be given in equation 4.3. This equation will give the formula for the displacement of the particle, which is dX_t.

$$\begin{cases} dX_t = f(X_t, t)dt + g(X_t, t)N_t dt \\ X_{t_0} = X_0 \end{cases} \tag{4.3}$$

To give a solution for the variable N_t the definition of the Wiener process (it is also called the Brownian motion) is needed. This is because N_t models diffusion of a particle in the random walk model.

The definition of a standard Wiener process is as follows: a standard Wiener process is the process W_t defined as

$$\begin{cases} W_t = [W_t,\ t > 0]\ on\ [0, t_{end}] \\ W_0 = 0 \end{cases} \tag{4.4}$$

and $\forall\ 0 \leq s < t \leq t_{end}$ it holds that:

- $$\mathbb{E}(W_t - W_s) = 0 \tag{4.5}$$
- $$\mathrm{Var}(W_t - W_s) = t - s \tag{4.6}$$

From this definition it follows that dW_t is a Gaussian random variable with the following parameters: mean is 0 and variance is dt, so $dW_t \sim \mathcal{N}(0, \sqrt{dt})$. Moreover, it holds that $\frac{dW_t}{dt} = N_t$, so $dW_t = N_t dt$. This can be filled in into equation 4.3 and this results in equation 4.7. Additionally, the SDE of equation 4.3 can be turned into a two-dimensional model, namely an equation for X and Y, this can be seen in equations 4.7 and 4.8. X and Y specify the location of the particle.

$$dX_t = f_1(X_t, Y_t, t)dt + g_1(X_t, Y_t, t)dW_t \tag{4.7}$$

$$dY_t = f_2(X_t, Y_t, t)dt + g_2(X_t, Y_t, t)dW_t \tag{4.8}$$

The next step is to evaluate these SDE's, to get a formula for X_t and Y_t. First the integral is taken, this can be seen in equations 4.9 and 4.10.

$$X_t = X_{t_0} + \int_{t_0}^{t} f_1(X_s, Y_s, s)ds + \int_{t_0}^{t} g_1(X_s, Y_s, s)dW_s \tag{4.9}$$

$$Y_t = Y_{t_0} + \int_{t_0}^{t} f_2(X_s, Y_s, s)ds + \int_{t_0}^{t} g_2(X_s, Y_s, s)dW_s \tag{4.10}$$

To further evaluate these equations a definition is needed. Namely, the definition of the Îto integral, which is given by by equation 4.11.

$$\int_{t_0}^{t} G_s dW_s = \lim_{\Delta t \to 0} \sum G_{t_i}(W_{t+i} - W_{t_i}) \tag{4.11}$$

The variables G_s and W_s are stochastic processes. Using this definition, the Îto stochastic differential equations can be solved by using a numerical approximation. This is called the Euler scheme, and for a small time step Δt it results in equations 4.12 en 4.13.

$$\begin{aligned} X_{t+\Delta t} &= X_t + \int_{t}^{t+\Delta t} f_1(X_s, Y_s, s)ds + \int_{t}^{t+\Delta t} g_1(X_s, Y_s, s)dW_s \\ &\approx X_t + \int_{t}^{t+\Delta t} f_1(X_t, Y_t, t)ds + \int_{t}^{t+\Delta t} g_1(X_t, Y_t, t)dW_s \\ &= X_t + f_1(X_t, Y_t, t)\Delta t + g_1(X_t, Y_t, t)(W_{t+\Delta t} - W_t) \end{aligned} \tag{4.12}$$

$$\begin{aligned} Y_{t+\Delta t} &= Y_t + \int_{t}^{t+\Delta t} f_2(X_s, Y_s, s)ds + \int_{t}^{t+\Delta t} g_2(X_s, Y_s, s)dW_s \\ &\approx Y_t + \int_{t}^{t+\Delta t} f_2(X_t, Y_t, t)ds + \int_{t}^{t+\Delta t} g_2(X_t, Y_t, t)dW_s \\ &= Y_t + f_2(X_t, Y_t, t)\Delta t + g_2(X_t, Y_t, t)(W_{t+\Delta t} - W_t) \end{aligned} \tag{4.13}$$

4.1.2. Determining the functions f_1, f_2, g_1 and g_2

To use the model in equations 4.12 and 4.13, the functions f_1, f_2, g_1 and g_2 need to be determined. For this the Fokker-Planck equation is needed, from this equation the probability density functions of X_t and Y_t can be obtained. The equation from which the pdf for the two dimensional model can be obtained, is given in equation 4.14. The two dimensional model itself is given in equations 4.7 and 4.8.

$$\begin{aligned}\frac{\partial p(t,x,y)}{\partial t} = &- \frac{\partial(f_1(t,x,y)p(t,x,y))}{\partial x} + \frac{1}{2}\frac{\partial^2(g_1^2(t,x,y)p(t,x,y))}{\partial x^2} \\ &- \frac{\partial(f_2(t,x,y)p(t,x,y))}{\partial y} + \frac{1}{2}\frac{\partial^2(g_2^2(t,x,y)p(t,x,y))}{\partial y^2}\end{aligned} \tag{4.14}$$

In this equation is $p(t,x,y)$ the pdf of one particle. Moreover, there is a link between the pdf of the particles and the concentration of particles. If a large number of particles is simulated, then the random walk model can be used to describe the concentration. When this happens then the particle distribution is equivalent to pdf of one particle. This then results in that the pdf is the concentration divided by a constant. The division by the constant is to compensate for the mass difference. So if $c(t,x,y)$ is the formula for the concentration of particles, then we have $p(t,x,y) = \frac{c(t,x,y)}{B}$, where B is just a constant. This can be filled in into equation 4.14 and the following equation is obtained:

$$\begin{aligned}\frac{\partial(\frac{c(t,x,y)}{B})}{\partial t} = &- \frac{\partial(f_1(t,x,y)\frac{c(t,x,y)}{B})}{\partial x} + \frac{1}{2}\frac{\partial^2(g_1^2(t,x,y)\frac{c(t,x,y)}{B})}{\partial x^2} \\ &- \frac{\partial(f_2(t,x,y)\frac{c(t,x,y)}{B})}{\partial y} + \frac{1}{2}\frac{\partial^2(g_2^2(t,x,y)\frac{c(t,x,y)}{B})}{\partial y^2}\end{aligned} \tag{4.15}$$

Now this formula can be rewritten with the help of the product rule for differentiation and with the knowledge that B is a constant. This is given in equation 4.16.

$$\begin{aligned}\frac{\partial c(t,x,y)}{\partial t} = &- \frac{\partial(f_1(t,x,y)c(t,x,y))}{\partial x} + \frac{1}{2}\frac{\partial(g_1^2(t,x,y) + \frac{\partial c(t,x,y)}{\partial x})}{\partial x} + \frac{1}{2}\frac{\partial(c(t,x,y) + \frac{\partial g_1^2(t,x,y)}{\partial x})}{\partial x} \\ &- \frac{\partial(f_2(t,x,y)c(t,x,y))}{\partial y} + \frac{1}{2}\frac{\partial(g_2^2(t,x,y) + \frac{\partial c(t,x,y)}{\partial y})}{\partial y} + \frac{1}{2}\frac{\partial(c(t,x,y) + \frac{\partial g_2^2(t,x,y)}{\partial y})}{\partial y}\end{aligned} \tag{4.16}$$

Next, this equation will be set equal to the two-dimensional advection-diffusion equation, which has terms for turbulence in the water, the eddy diffusivity terms. More on eddy diffusivity can be found in the next section. This is done because the advection-diffusion equation can be used for the transport of particles in waters. Additionally, as mentioned above, the random walk model can describe a concentration when large amounts of particles are modelled. If this happens then the solution of the random walk model and advection-diffusion equation are similar. The advection-diffusion equation is given in equation 4.17.

$$\frac{\partial c(t,x,y)}{\partial t} = \frac{\partial(\epsilon_x \frac{\partial c(t,x,y)}{\partial x})}{\partial x} + \frac{\partial(\epsilon_y \frac{\partial c(t,x,y)}{\partial y})}{\partial y} - \frac{\partial(u_x(t,x,y)c(t,x,y))}{\partial x} - \frac{\partial(u_y(t,x,y)c(t,x,y))}{\partial y} \tag{4.17}$$

In this equation there are a few new variables. u_x and u_y stand for the velocity of the ocean/water in the x- and y-direction respectively. Additionally, ϵ_x and ϵ_y stand for the eddy diffusivity in the x- and y-direction respectively. More on the eddy diffusivity can be found in the next section. Equation 4.16 is equal to equation 4.17 if with the following formulas for the functions f_1, f_2, g_1 and g_2: $f_1 = u_x + \frac{\partial \epsilon_x}{\partial x}$, $f_2 = u_y + \frac{\partial \epsilon_y}{\partial y}$, $g_1 = \sqrt{2\epsilon_x}$ and $g_2 = \sqrt{2\epsilon_y}$. If this is filled in into equations 4.12 and 4.13, then it results in the following model:

$$X_{t+\Delta t} = X_t + (u_x + \frac{\partial \epsilon_x}{\partial x})\Delta t + \sqrt{2\epsilon_x}(W_{t+\Delta t} - W_t) \tag{4.18}$$

$$Y_{t+\Delta t} = Y_t + (u_y + \frac{\partial \epsilon_y}{\partial y})\Delta t + \sqrt{2\epsilon_y}(W_{t+\Delta t} - W_t) \tag{4.19}$$

Now there is only one last step to take to go from equations 4.18 and 4.19 to equations which can be used to program the model. $W_{t+\Delta t}$ and W_t need to be written as something that can be programmed. It holds that $\Delta W = W_{t+\Delta t} - W_t$, and because W_t is a Wiener process. It follows that $\mathbb{E}(\Delta W) = 0$ and $\mathrm{Var}(\Delta W) = \Delta t$. So $\Delta W = W_{t+\Delta t} - W_t$ can be modelled by taking a random number of the $\mathcal{N}(0, \sqrt{\Delta t})$ Gaussian distribution. If this is filled in into equations 4.18 and 4.19 it results in equation 4.20 and equation 4.21. Additionally, it is implemented in such a way that there is only eddy diffusivity when there is actually velocity of the water.

$$X_{t+\Delta t} = X_t + (u_x + \frac{\partial \epsilon_x}{\partial x})\Delta t + \sqrt{2\epsilon_x} * random(\mathcal{N}(0, \sqrt{\Delta t})) \tag{4.20}$$

$$Y_{t+\Delta t} = Y_t + (u_y + \frac{\partial \epsilon_y}{\partial y})\Delta t + \sqrt{2\epsilon_y} * random(\mathcal{N}(0, \sqrt{\Delta t})) \tag{4.21}$$

4.2. Eddy Diffusivity

The diffusion in the model is a result from turbulence in the ocean. A turbulent flow exists when not every pocket of fluid moves in the same direction and has the same amount of speed. This leads to eddies, which are swirls that emerge because of the changes in direction and velocities. In the ocean eddies can look like circular motions in the water, (Ruiter, 2020; Scheijen, 2018) In figure 4.1 an eddy is shown from above.

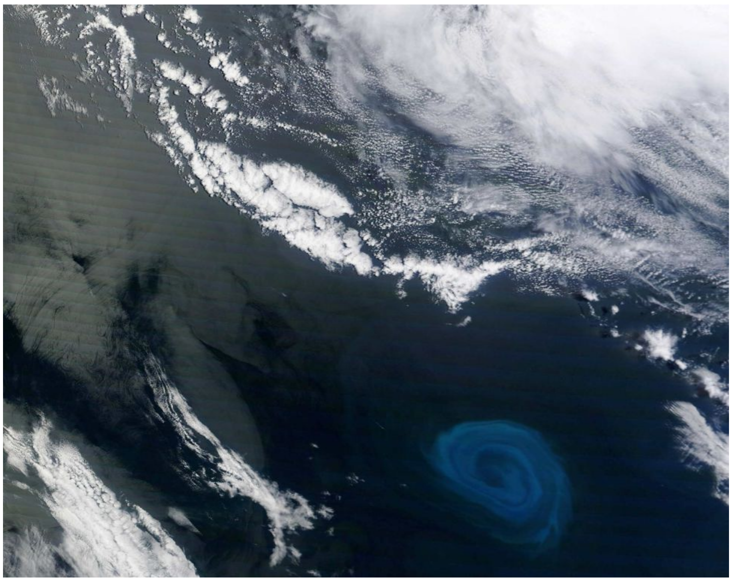


Figure 4.1: An eddy shown from above, (Allen, 2011)

These eddies can have all kinds of different sizes, from large ones that can be hundreds of kilometers big. To small ones of only a few centimeters. How long these eddies exist can also vary hugely. There are eddies that can exist for years, which are called long life time eddies. Additionally, there are eddies that exist for only a very short time and disappear quickly, they are called instantaneous eddies. (Ruiter, 2020; Scheijen, 2018).

The velocities of turbulence are random and turbulence itself is very chaotic. So it can be simulated using noise, which depends on how much turbulence there is, this is described by the eddy diffusivity variables ϵ_x and ϵ_y, (Ruiter, 2020; Scheijen, 2018). (There is no depth in the model, so the eddy diffusivity in that direction is not included). Due to the chaotic nature and randomness it is very hard to find values for the eddy diffusivity. Which is why different values for the eddy diffusivity will be tested in the example in chapter 5. These values will be between 0 and 2 m^2/s. These are rough estimations of common values of the eddy diffusivity. The values come from an article about the Gulf of Lion in the Mediterranean Sea, (Nencioli et al., 2013; Scheijen, 2018).

5

Example

In this chapter an example will be discussed, it is a simplified version of plastic particles falling of a ship and knowing the end location, but wanting to find the location where the particles fell overboard. This is because the bigger case, the MSC Zoe disaster, is about containers full of items that fell overboard, among these items were bags of plastic particles. We know the end location of a lot of these particles, but the location where the containers fell overboard is unknown. This is the location we want to know. For the explanation of this case see Chapter 2. One thing to explain before the example and before the MSC Zoe case is the assumption about what one particle means. As mentioned in chapter 3 one particle is literally one particle in this research. For this one particle the start location is being sought. One particle on the beach is not the same as another particle and it does not represent a concentration of particles. So it can be compared to a container that fell from the ship for which the overboard location is being sought. So the probability density function $p(\psi|d)$, represents the likelihood of different locations as start locations of the one particle. To get this pdf a lot of different particles are released, everyone of these represents one particle. So it can be compared to releasing the same amount of containers and calculating a pdf for this from for the start location of the one container. The setup is simplified because in reality there were huge concentrations of particles on the beaches and this is an easier first step. For the MSC Zoe case it would be interesting to find a way to apply the method to concentrations of particles. Another simplification is that for the example and the MSC Zoe case the time between containers falling overboard, containers breaking open and the bags with plastic particles rupturing is not taken into account. So it is assumed that as soon as the container hit the water it broke open and the bags ruptured. This is not realistic, but there is no way of knowing when exactly this all happened. On these assumptions is the most likely start location going to be found in this report.

5.1. The idea

So the idea for this example is that we have a ship that sails in a straight line along a coast that is also straight. We have 'found' one particle on the coast from which we want to know where it went overboard. So d is the location of this particle, it is the x and y-coordinates of the particle. In this report we do not look at the depth of the ocean, so no z-coordinate, because a 2D model is used. Now we are releasing plastic particles (ψ_i's) along the route of the ship, which is the line of $x = 4500$ m. We call N the number of particles released, and they are spread over this line with even space between them, between $y = 0$ m and $y = 10000$ m. As a result of this even spread of particles, the particles stick to the prior probability density function, $p(\psi)$, which in this case has been chosen as a Uniform distribution on the interval $[0, 10000]$ on which the particles are spread. So the samples ψ_i are being chosen from a Uniform distribution $p(\psi)$.

A very simplified flow field has been used for the ocean flow field. Namely only a flow in the x-direction and not in the y-direction (and not in the z-direction). The speed in the x-direction is 0.2 m/s, this is the average speed of the surface current in the North Sea, (Lee et al., 1968; Winther et al., 2006). The coast starts on $x = 23000$, this is a distance of 18500 m from the sailing route to the coast. This distance has been chosen because the MSC Zoe sailed on the southern sailing route above the Wadden Islands, which is where the disaster happened, and this route is 18500 m from the coast, (Smid, 2019). To see the coast clearly in the example, it will have a width of 3500 m. So the total length in the x-direction of the example is 26.500 m. In the example the flow stops when it reaches the coast. So the flow is only a straight line to the coast and stops there, it does not go back (no ebb and flow). This is signified by the following formulas:

$$u_x = \begin{cases} 0.2 & if\ x < 23000 \\ 0.0 & otherwise \end{cases} \tag{5.1}$$

$$u_y = 0.0 \tag{5.2}$$

The model used for calculating the position each times step is the Random Walk model explained in section 4. The values of u_x and u_y in the formulas above are filled in in equations 4.20 and 4.21. Furthermore the time goes from $t_{start} = 0$ to $t_{end} = 108000$ seconds (it is 30 hours) with time step $\Delta t = 1800\ s$. This end time has been chosen because after this amount of time all the released particles have ended up on the coast. The reason for this time step is that it is a balance between the accuracy of the model and the runtime of the simulation. In this example it is not necessary to add a time delay because the flow does not change over time. Now the method from chapter 3 is applied. The results from this and from the movement of particles will be shown in the next section. In the results section will the values of the 3 variables that are not yet determined be given, (N, r and ϵ).

5.2. Results & Analysis of the Results

First the plots of the movement of the particles will be shown and after that multiple plots of the probability density function $p(\psi|d)$ for different values of N, r and $\epsilon = (\epsilon_x, \epsilon_y)$.

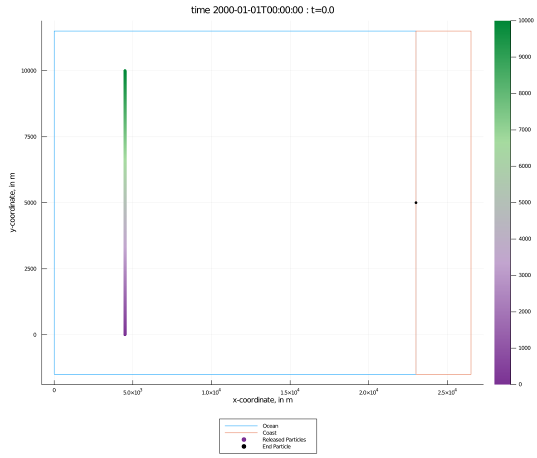
5.2.1. The movement of particles

In the plots of the movement the blue lines signify the edges of the ocean in this example and the red lines signify the edges of the coast in this example. Moreover, some particles are coloured according to their start location, such that it is easy to see at the end where they originated from. The particles that are released along the route of the ship are the coloured particles. The particle from which we only know the end location and want to find the start location is coloured black, such that the difference between the two kinds of particles can be seen. The end particle (of which we only know the end location), is already on the coast at the start of the simulation. It has the following coordinates: $(x, y) = (23000, 5000)$. In figure 5.1, on the next page, can the movement from the particles to the coast be seen. For these pictures the following values of the still undetermined variables are used: $N = 5000$ particles, $\epsilon_x = 1.0$ m^2/s and $\epsilon_y = 1.0$ m^2/s. At the top of the plot it says that the date is in the year 2000, this is a default time setting.

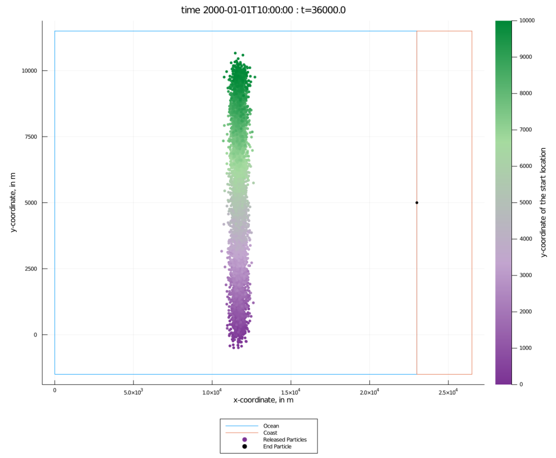
In figure 5.1 it can be seen that the released particles flow from the left to the right, which follows the flow given in equations 5.1 and 5.2. The time of the last picture in figure 5.1 is 29 hours, and not 30 hours because at 29 hours all particles are on the coast. So after this the particles do not move anymore. Furthermore, it can be seen that the particles spread out a little, but that the order does not change that much. This can be seen when comparing the picture in figure 5.1d to the gradient bar next to it, which is the color gradient on the basis of the y-coordinate of the start location of a particle.

5.2.2. Analysis of the movement of particles

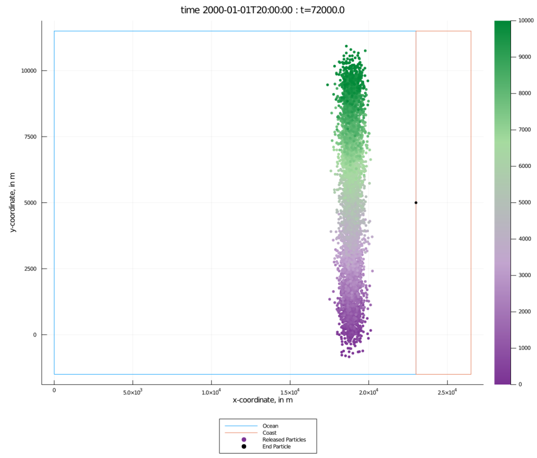
Now will the results in the previous subsection be discussed. It can be seen that the particles spread out once they are moving, this is due to the eddy diffusivity ($\epsilon_x = 1.0$ and $\epsilon_y = 1.0$). This means that random numbers from the Gaussian distribution are added to the model, see chapter 4. So the transport is more random, and this means that not every particle moves at the same speed. If the eddy diffusivity would be 0 in both directions, then the plots would only show the straight line from figure 5.1a moving to the right. Such a transport can be seen in figure A.1 in Appendix A.



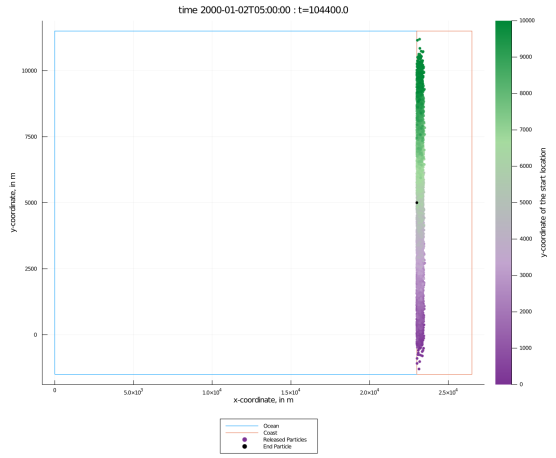
(a) $t = 0$ seconds (0 hours)



(b) $t = 36000$ seconds (10 hours)



(c) $t = 72000$ seconds (20 hours)

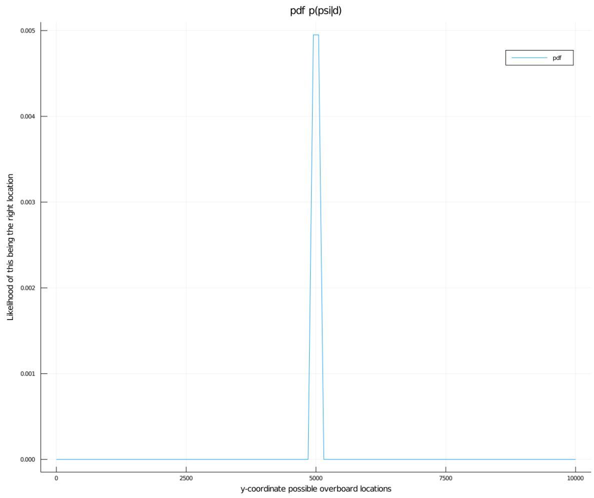


(d) $t = 104400$ seconds (29 hours)

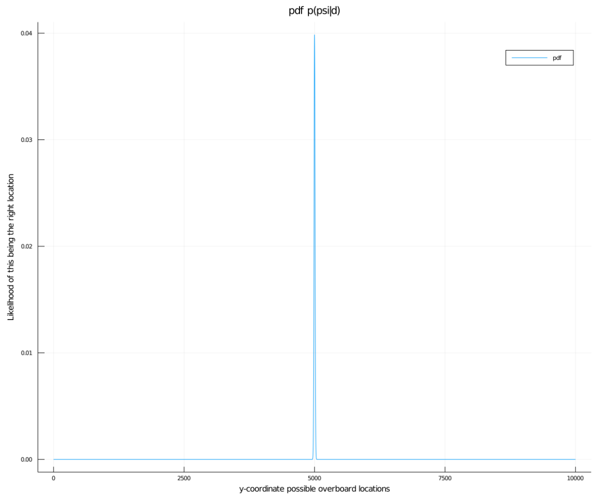
Figure 5.1: The movement of the plastic particles, with $N = 5000$, $\epsilon_x = 1.0$ and $\epsilon_y = 1.0$. With a color gradient on the basis of the y-coordinate of the start position

5.2.3. Results of the method of particle filtering

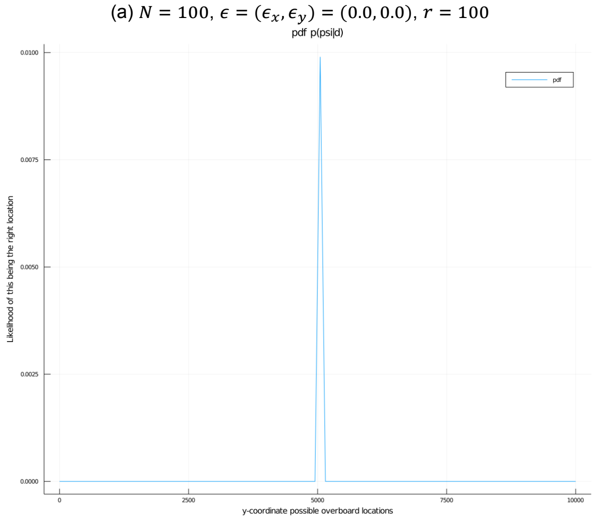
Now the results from the method of particle filtering will be discussed. The results are given for different values of the three variables for which we can vary the values. These variables are N, the number of particles released, r, used in the calculation of the error covariance of the observations R, and ϵ, the eddy diffusivity. N is a number, r is in meters and ϵ_x and ϵ_y are in m^2/s. The following values are used for N, r and ϵ: $N \in \{100, 10000\}$, $r \in \{0, 1000000\}$ and $\epsilon = (\epsilon_x, \epsilon_y) \in \{(0.0, 0.0), (1.0, 1.0)\}$. The terms for the eddy diffusivity ϵ_x and ϵ_y do not have to have the same value, but for this example it is chosen that they do have the same value. The values for the variables are very far apart, such that the difference that it makes can be easily seen. The results, the probability density functions $p(\psi|d)$ can be seen in figure 5.2. As can be seen in figure 5.2 the different values of N, ϵ and r give very different results. This will be discussed in the next subsection.



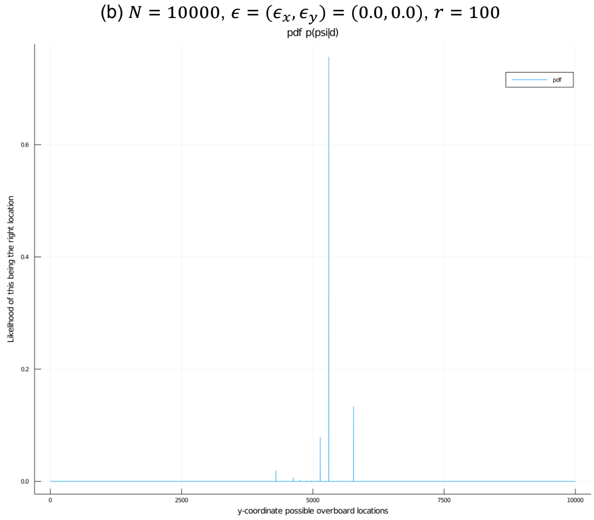
(a) $N = 100$, $\epsilon = (\epsilon_x, \epsilon_y) = (0.0, 0.0)$, $r = 100$



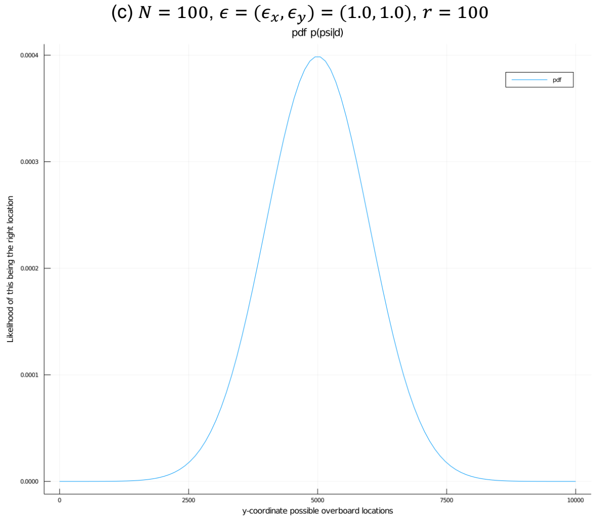
(b) $N = 10000$, $\epsilon = (\epsilon_x, \epsilon_y) = (0.0, 0.0)$, $r = 100$



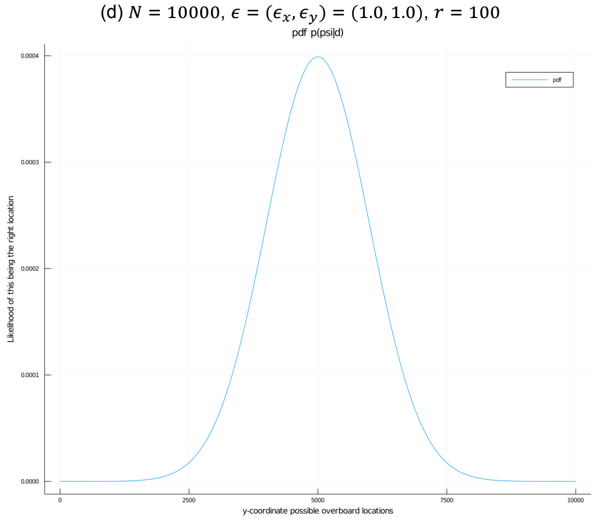
(c) $N = 100$, $\epsilon = (\epsilon_x, \epsilon_y) = (1.0, 1.0)$, $r = 100$



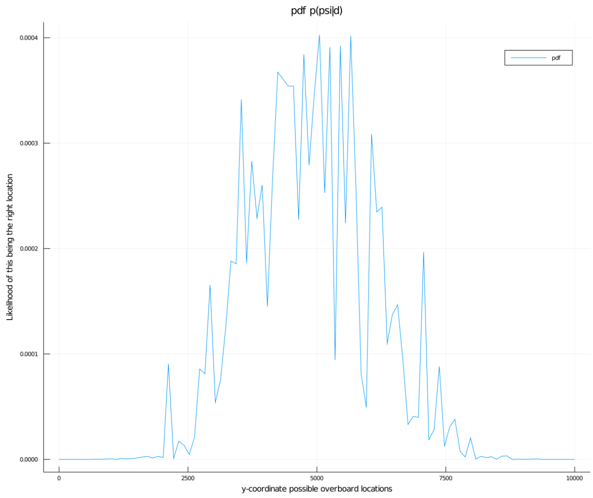
(d) $N = 10000$, $\epsilon = (\epsilon_x, \epsilon_y) = (1.0, 1.0)$, $r = 100$



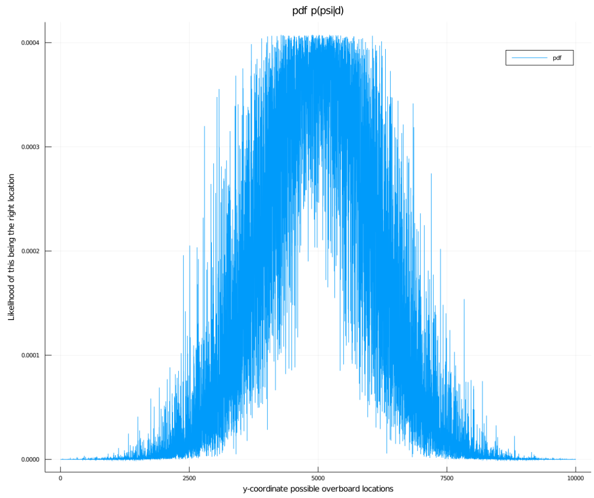
(e) $N = 100$, $\epsilon = (\epsilon_x, \epsilon_y) = (0.0, 0.0)$, $r = 1000000$



(f) $N = 10000$, $\epsilon = (\epsilon_x, \epsilon_y) = (0.0, 0.0)$, $r = 1000000$



(g) $N = 100$, $\epsilon = (\epsilon_x, \epsilon_y) = (1.0, 1.0)$, $r = 1000000$



(h) $N = 10000$, $\epsilon = (\epsilon_x, \epsilon_y) = (1.0, 1.0)$, $r = 1000000$

Figure 5.2: The probability density function $p(\psi|d)$

5.2.4. Analysis of the results of the method of particle filtering

As said above, the different values give very different results. For example, if ϵ is $(1.0, 1.0)$, it means that there is diffusion and randomness in the end location of the particles. It allows for more than one peak in the plot, when N or r increases. N says how many particles are released. The larger N is the less space there is between the released particles, so the more accurate the results. However, another side is that the larger N is the longer the runtime becomes. r, used in the calculation of the error covariance of the observations R, indicates how correct the information that is given is. Namely, the larger r is, the more error one perceives in the calculations and the wider the peak of the plot becomes. So the pdf then gives more possibilities to other options to be the right start location, and more uncertainty is added. Also, for a small r it can be seen that N determines the width of the peak. Additionally, the larger r gets the more the different peaks (if these exist) flow into each other. The peaks don't go all the way back to 0. So the difficult part is to find a balance between these values, between what is realistic and what gives good results.

Another thing to look at is if the method of particle filtering works, does it give a logical result. As can be seen in the plots it does give a logical result. The end particle had the location of $(x, y) = (23000, 5000)$. So with the flow of the water only going towards the coast and depending on the value of ϵ a little diffusivity. It would be expected that the start location/the overboard location would be around $y = 5000$ on the line $x = 4500$. The different plots in figure 5.2 all show the highest likelihood around this point. So this shows that the method indeed works for this kind of problem.

In figures 5.2g and 5.2h it can be seen that the pdf has a lot of different peaks that flow into one another. This does not represent a very nice probability density function. We do not suspect that there are different peaks with weight is 0 between it. So another method is used to estimate the pdf $p(\psi|d)$, this is the kernel estimation method explained in section 3.3. The results of this will be given in the next subsection.

There is a big disadvantage to this calculation of the weights. Namely, the weights are the relative importance of a particle. So what happens is that when no particles get close, the particle that gets the 'closest' still gets the most weight. Its start location is still identified as the most likely. However, it is not very likely at all because the particle did not get very close to the end particle. So to see if the method really works the end positions of the released particles still need to be close to the location of the end particle for the result to be reliable. In this example that is the case, as can be seen in figure 5.1 the released particles get very close to the end particle.

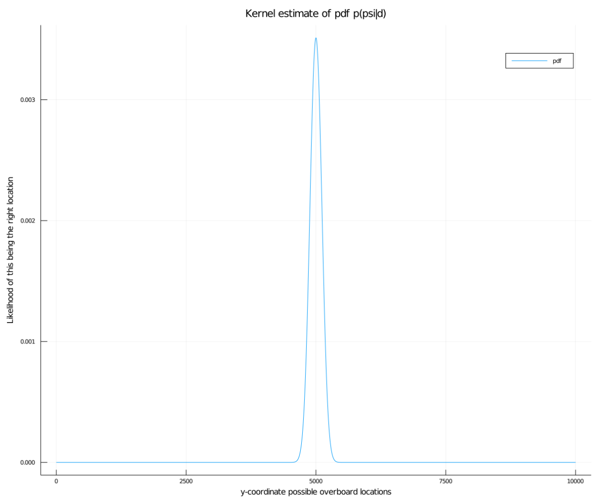
5.2.5. Results of the kernel estimation

The kernel estimation is given for the same values of N, r and ϵ as used in section 5.2.3. In this method there is another variable that can be varied, namely the σ. In figures 5.3 and 5.4 the kernel estimation of $p(\psi|d)$ is shown for $\sigma = 100$ m and $\sigma = 300$ m, respectively. It can be seen that the different values of σ give different results.

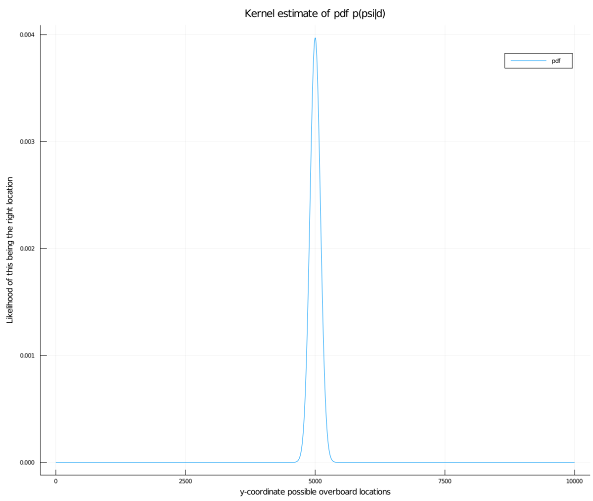
5.2.6. Analysis of the results of the kernel estimation

The results of the kernel estimation show that it gives a more probable pdf. This is due to the fact that there are no more multiple peaks in the plots. It smooths out the probability density functions from figure 5.2. The value of σ matters in how much the functions are smoothed out. The larger σ is the more the functions are smoothed out and the wider the peak of the pdf becomes. So also here the value of σ needs to be chosen such that it smooths out the function, but does not give too much error/that the standard deviation too big is. So an average between multiple peaks and smoothness of the plot needs to be chosen. How large N is also matters for the kernel. The larger N the longer the runtime to calculate the kernel estimation, because a lot more calculations need to be done.

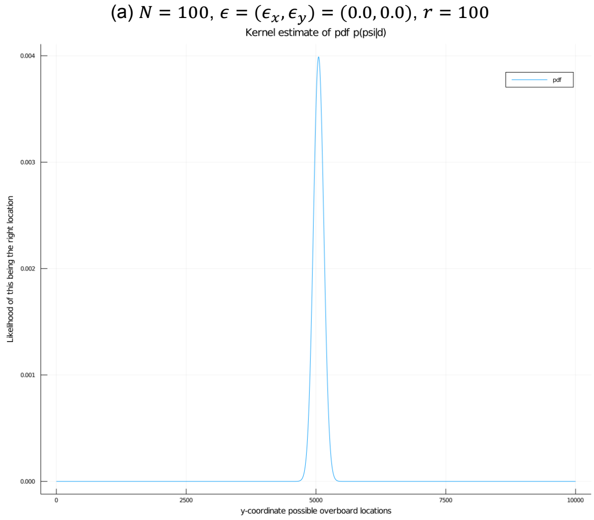
As can be seen in figures 5.3 and 5.4 that the method still works. The highest likelihood is still around $y = 5000$. Which is to be expected with the location of the end particle being $(x, y) = (23000, 5000)$.



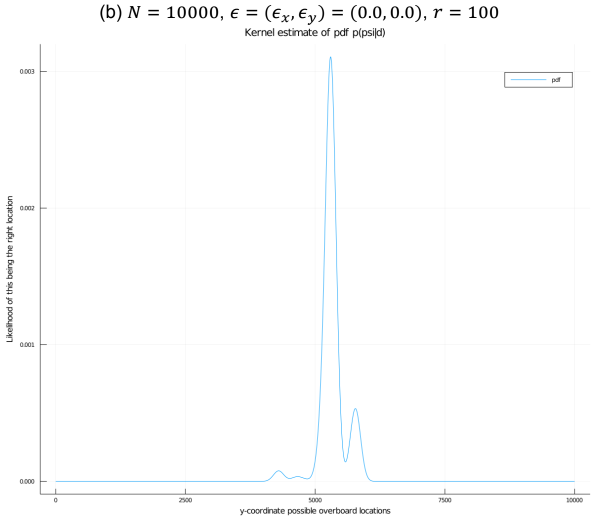
(a) $N = 100$, $\epsilon = (\epsilon_x, \epsilon_y) = (0.0, 0.0)$, $r = 100$



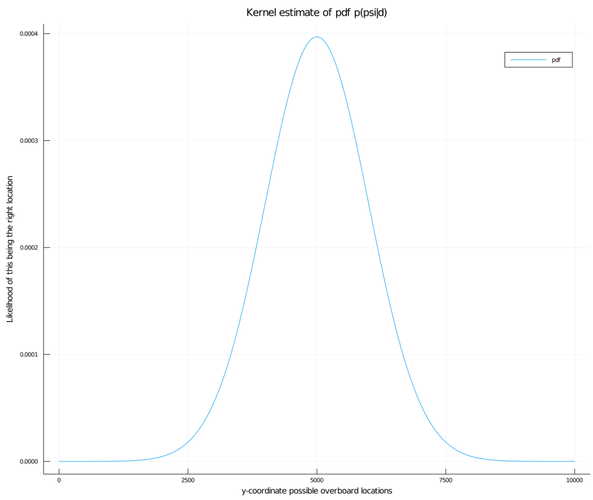
(b) $N = 10000$, $\epsilon = (\epsilon_x, \epsilon_y) = (0.0, 0.0)$, $r = 100$



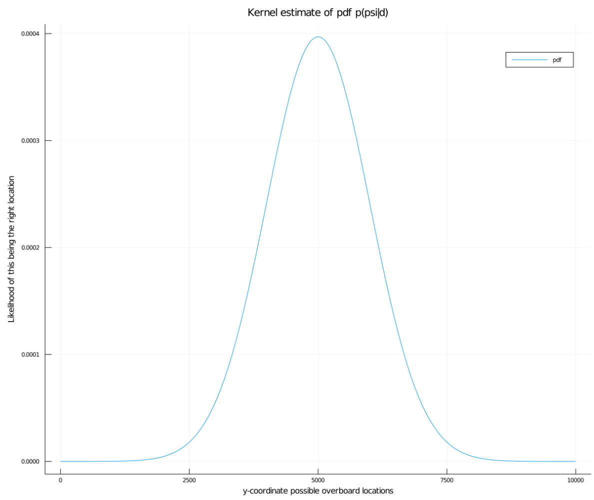
(c) $N = 100$, $\epsilon = (\epsilon_x, \epsilon_y) = (1.0, 1.0)$, $r = 100$



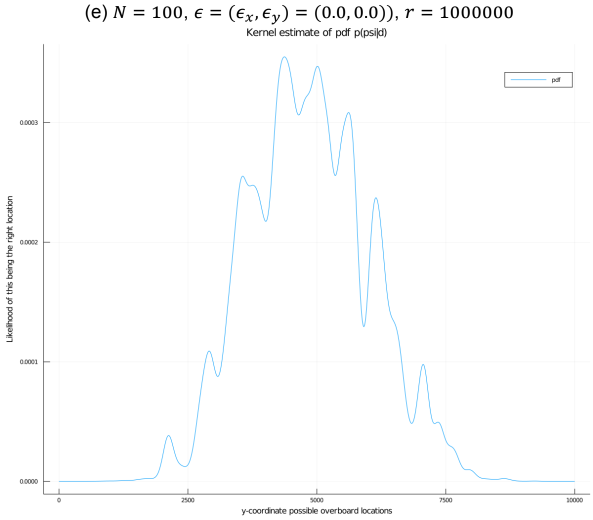
(d) $N = 10000$, $\epsilon = (\epsilon_x, \epsilon_y) = (1.0, 1.0)$, $r = 100$



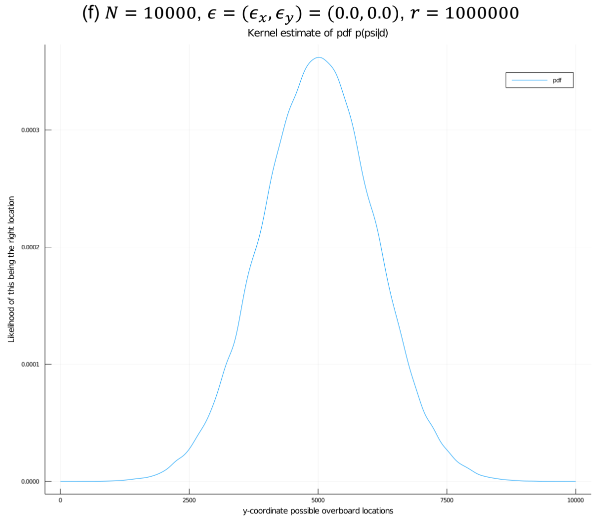
(e) $N = 100$, $\epsilon = (\epsilon_x, \epsilon_y) = (0.0, 0.0))$, $r = 1000000$



(f) $N = 10000$, $\epsilon = (\epsilon_x, \epsilon_y) = (0.0, 0.0)$, $r = 1000000$

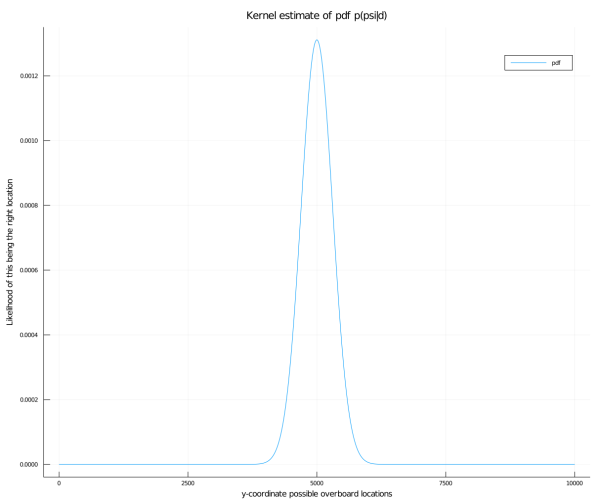


(g) $N = 100$, $\epsilon = (\epsilon_x, \epsilon_y) = (1.0, 1.0)$, $r = 1000000$

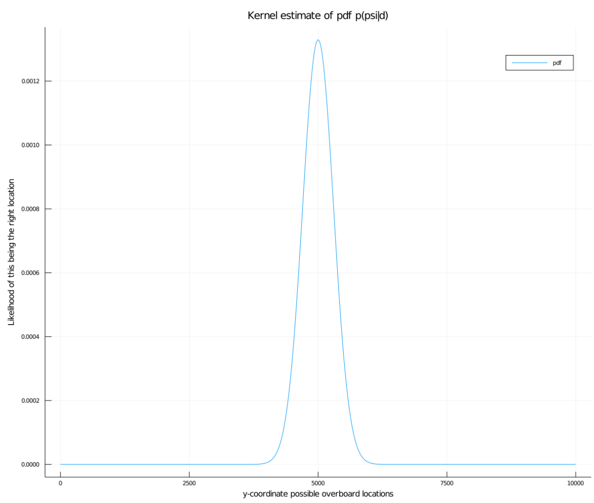


(h) $N = 10000$, $\epsilon = (\epsilon_x, \epsilon_y) = (1.0, 1.0)$, $r = 1000000$

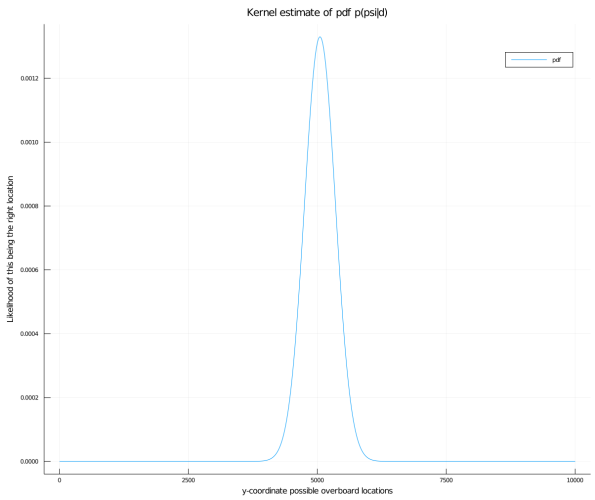
Figure 5.3: The kernel estimate of the probability density function $p(\psi|d)$, with $\sigma = 100\ m$



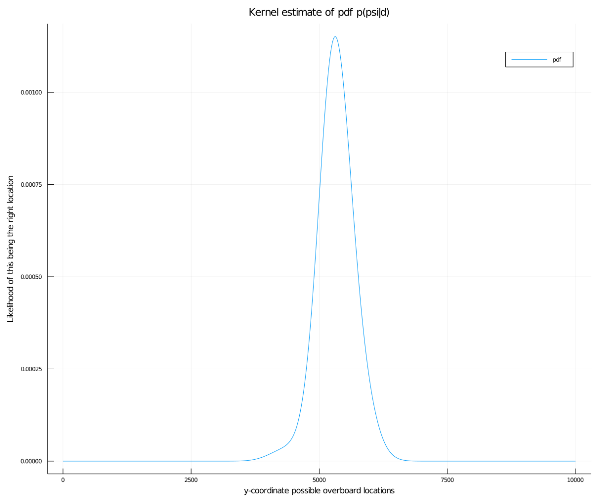
(a) $N = 100$, $\epsilon = (\epsilon_x, \epsilon_y) = (0.0, 0.0)$, $r = 100$



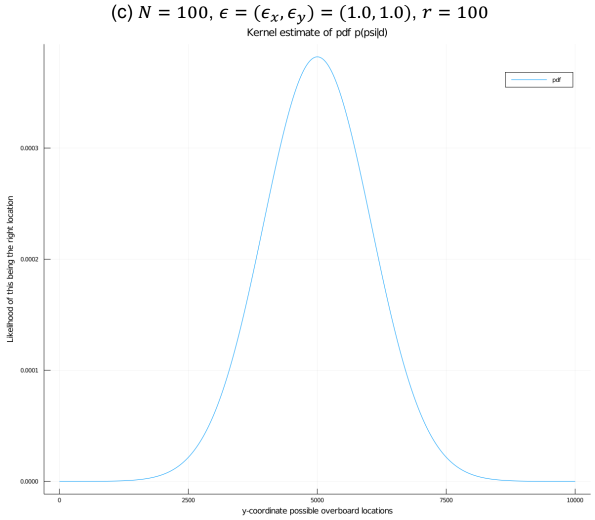
(b) $N = 10000$, $\epsilon = (\epsilon_x, \epsilon_y) = (0.0, 0.0)$, $r = 100$



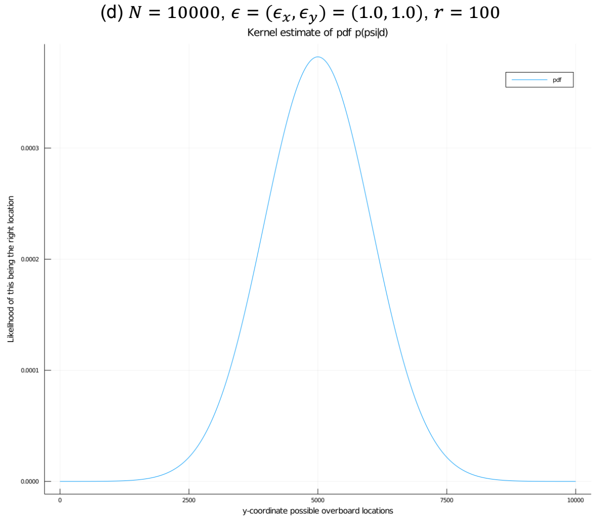
(c) $N = 100$, $\epsilon = (\epsilon_x, \epsilon_y) = (1.0, 1.0)$, $r = 100$



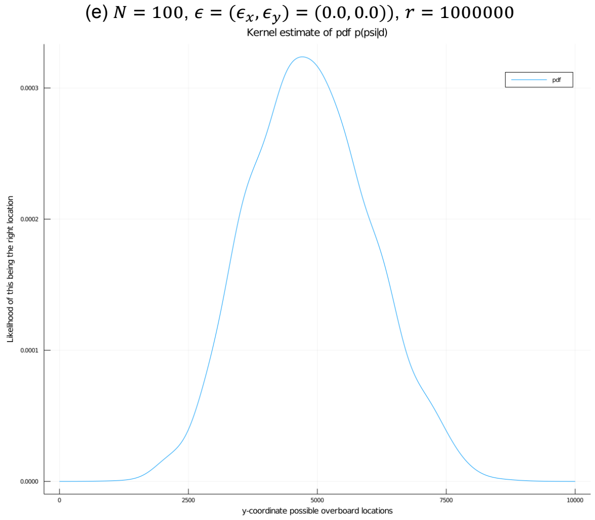
(d) $N = 10000$, $\epsilon = (\epsilon_x, \epsilon_y) = (1.0, 1.0)$, $r = 100$



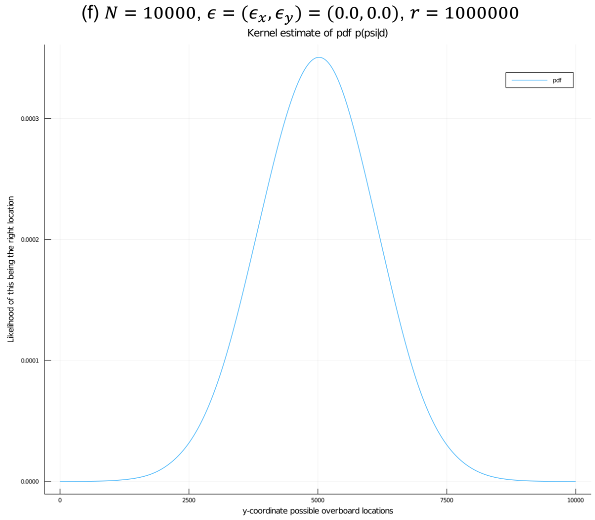
(e) $N = 100$, $\epsilon = (\epsilon_x, \epsilon_y) = (0.0, 0.0))$, $r = 1000000$



(f) $N = 10000$, $\epsilon = (\epsilon_x, \epsilon_y) = (0.0, 0.0)$, $r = 1000000$



(g) $N = 100$, $\epsilon = (\epsilon_x, \epsilon_y) = (1.0, 1.0)$, $r = 1000000$



(h) $N = 10000$, $\epsilon = (\epsilon_x, \epsilon_y) = (1.0, 1.0)$, $r = 1000000$

Figure 5.4: The kernel estimate of the probability density function $p(\psi|d)$, with $\sigma = 300\ m$

6
The MSC Zoe Case - Results

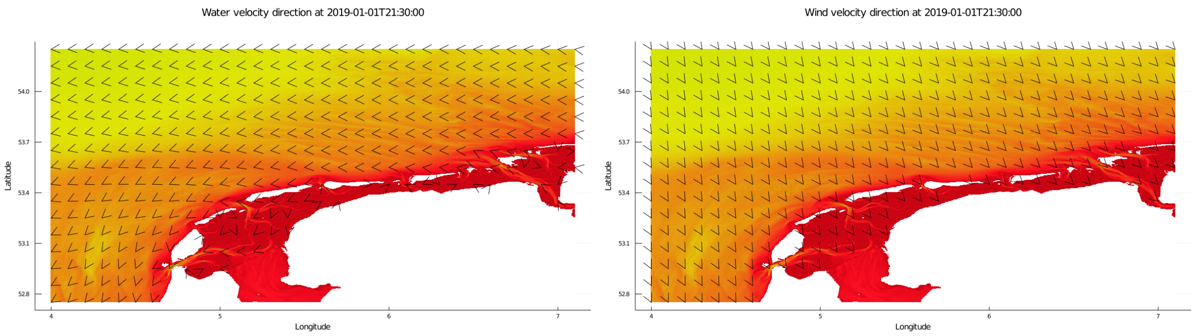
This chapter is about the theory applied to the MSC Zoe disaster in 2019. To start, the data that is used will be explained and how it is used. After that the chapter will end with the results of the method of particle filtering that is applied. In this chapter are two different time zones used to indicate what happened at which time. The data is in UTC time, so this is seen in all the plots. The text is mostly in local time (LT), UTC+1. This is indicated everywhere. There are some simplifications made in this research, these are explained at the beginning of chapter 5.

6.1. The Data & How it is implemented

In this section an overview of the data used to model the MSC Zoe case will be given. It will start with an explanation of the velocities of the ocean and wind that are used to model the transport of the particles. Next, is the AIC data of the MSC Zoe during that night. Finally, there are the locations of the plastic particles that are found on the coast of the Netherlands.

6.1.1. Ocean & Wind velocity

For the ocean and wind velocities the 2D DCSM-FM, Dutch Continental Shelf Model – Flexible Mesh, model is used, (Zijl et al., 2019). From this model the velocities in the x and y- direction are taken for both the ocean/water and wind velocities. The data was taken from the time of the accident, namely during the start of January of 2019. At first the model for the transport of the plastic particles, chapter 4, was only modelled with the water velocity. However, the particles did not come close to coast. In figure 6.1a is a vector field of the water velocity at a moment during the night of 1 to 2 January 2019. It can be seen that particles released a little bit from the shore will not flow to the shore. The picture that is the background of the vector field is a possible bathymetry of the North Sea, (Ruiter, 2020; The European Marine Observation and Data Network, n.d.). A bathymetry visualizes the depths and shapes of terrain that lies underwater, (US National Ocean Service, n.d.). This background image is used in multiple images in this chapter to provide a map of the sea and the coast as a background.

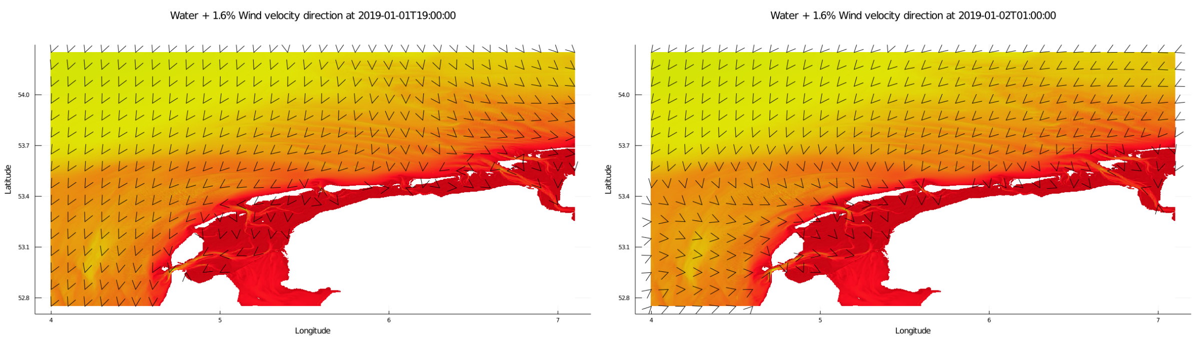


(a) Vector field of the water velocity

(b) Vector field of the wind velocity

Figure 6.1: Velocity vector fields of the water and wind velocities at 01-01-2019 22:30 LT (21:30 UTC time)

However, at the time of the accident a strong wind was coming from the north-northwesterly direction, so going in the direction of the coast, (Panama Marine Authority et al., 2020). This can also be seen in figure 6.1b, which is a vector field of the wind velocity at a time during that night. So it is needed to incorporate the wind velocity into the model to get the particles to the coast. This is done by adding 1.6% of the wind velocity measured at 10 meter height to the water velocity. Since the velocity of wind is usually measured at a height of 10 meters, (Ooms, 2020). So u_x and u_y from equations 4.20 and 4.21 in chapter 4 will be replaced by $u_x + 0.016 * a_x$ and $u_y + 0.016 * a_y$. Where a_x and a_y are the wind velocities in the x- and y-direction. Additionally, these velocities are all in m/s, so they are converted to °/s with coordinate transformation. This happens after the application of the random walk model, so all values including the eddy diffusivities are converted. This is needed because in this case is the step form projected to spherical coordinates. There are now longitude and latitude coordinates (long and lat) instead of x- and y-coordinates. So all values given in this chapter that are in meters are converted to degrees (r and σ). Vector fields of these velocities are shown in figure 6.2, two vector fields 6 hours apart. It can be seen that the flow now goes more towards the coast. This is in line with the findings that a lot of the cargo from the MSC Zoe that went overboard was found on the coast of the Wadden Islands, (University of Groningen, 2019).



(a) Date: 01-01-2019, Time: 20:00 LT (19:00 UTC time)

(b) Date: 02-01-2019, Time: 02:00 LT (01:00 UTC time)

Figure 6.2: Velocity vector fields of ocean velocity + 1.6% wind velocity

6.1.2. AIS Data MSC Zoe

The next piece of data that is used is the AIS data, Automatic Identification System, of the MSC Zoe. The AIS is a transponder on the ship. This data contains for example, the coordinates and speed of the ship at specific times, (Made Smart Group, n.d.). In figure 6.3 is the route of the MSC Zoe plotted. With the time and date at specific moments, the start and end point of the data. Additionally, the time at which the ship's course was fully turned. Lastly, there were a few small gaps in the data, the times just before and after these gaps are also given. The particles in the MSC Zoe case are released along the route. However, it does hold that the initial position of the particles needs to meet the prior distribution $p(\psi)$. So the long (longitude) coordinates are a range from 4.781112 to 6.557998 and the lat (latitude) coordinates are a range from 53.409382 to 53.786342. So the particles are evenly spread on the line from $(4.781112, 53.409382)$ to $(6.557998, 53.786342)$, this is sometimes just below the route of the MSC Zoe. This results in the following distributions, the long-coordinate has a Uniform distribution on the interval $[4.781112, 6.557998]$ and the lat-coordinate has a Uniform distribution on the interval $[53.409382, 53.786342]$. So $p(\psi)$ is the Uniform distribution on the rectangle made by these coordinates. The long- and lat-coordinates still hold to this distribution because they are chosen on a regular interval.

In figure 6.4 the speed of the MSC Zoe is plotted against the time. This is to show how fast the MSC Zoe moved forward along it's route plotted in figure 6.3. The speed is indirectly used to put a time delay on the release of the particles along the route. The MSC Zoe is not everywhere on the route at the same time, so not all the particles should be released at the same time. The time delay is implemented as follows: At every time step the UTC time at for the next time step is compared to all the times in the AIS data, the closest time is found. Then the coordinates of the MSC Zoe at that time are saved. After that all the particles with coordinates lower than the one of the MSC Zoe are released. This means that all particles that are on the route where the MSC Zoe sails in the next time step are released. So, the particles are now released a bit earlier than they should be. The choice was between releasing the particles a bit too early or releasing them a bit too late. In this research the choice has been made to release them a bit too early. The method will become more and more accurate the smaller the time step is. One thing that can impact the time delay a little bit are the gaps in the data. This is due to fact that in the time delay is sought for a closest time. This time may not be very close because of the small gaps in the data. In the next section some plots of the movement of the particles will be shown. All the particles are always plotted, but they are only moving when they are allowed to by this time delay.

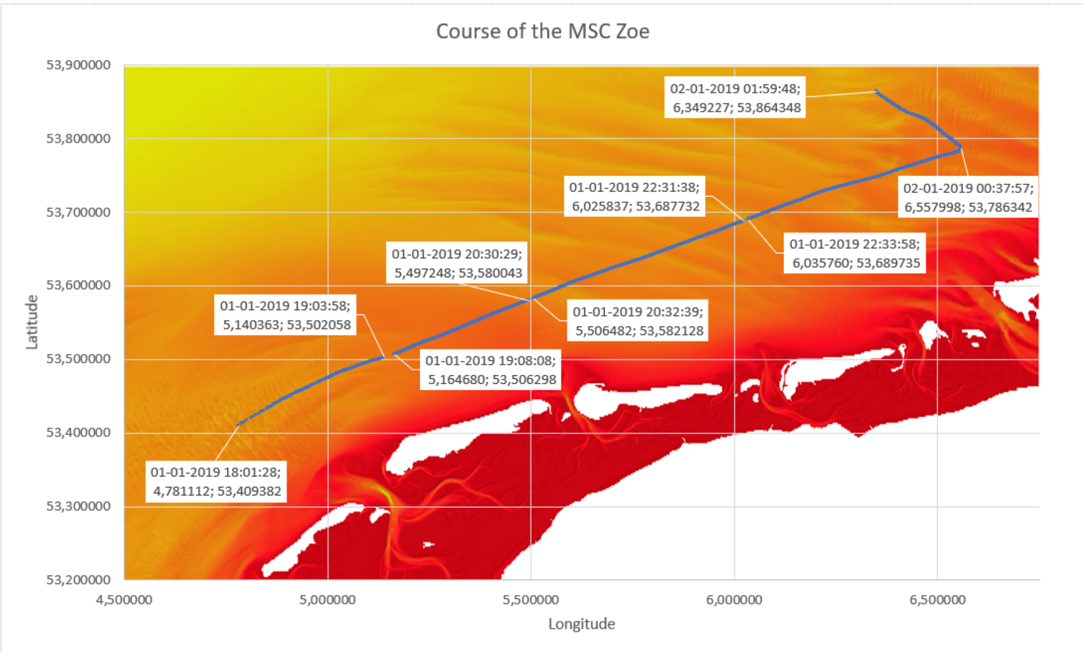


Figure 6.3: The route/course of the MSC Zoe (from the AIS data of the MSC Zoe) plotted on a map of the Wadden Sea

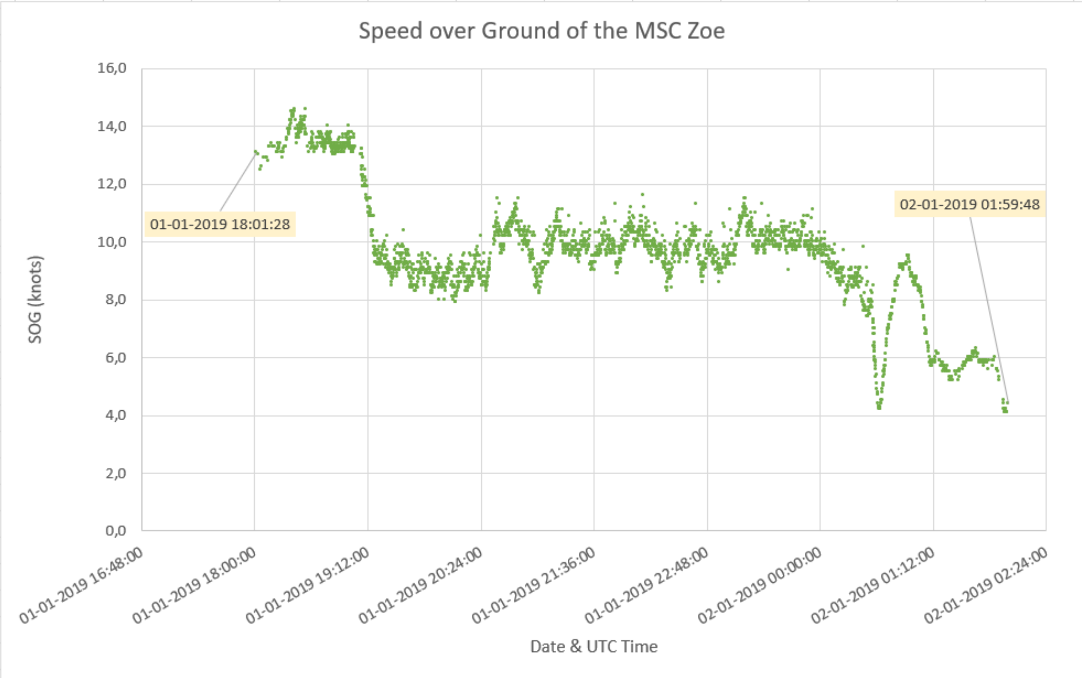
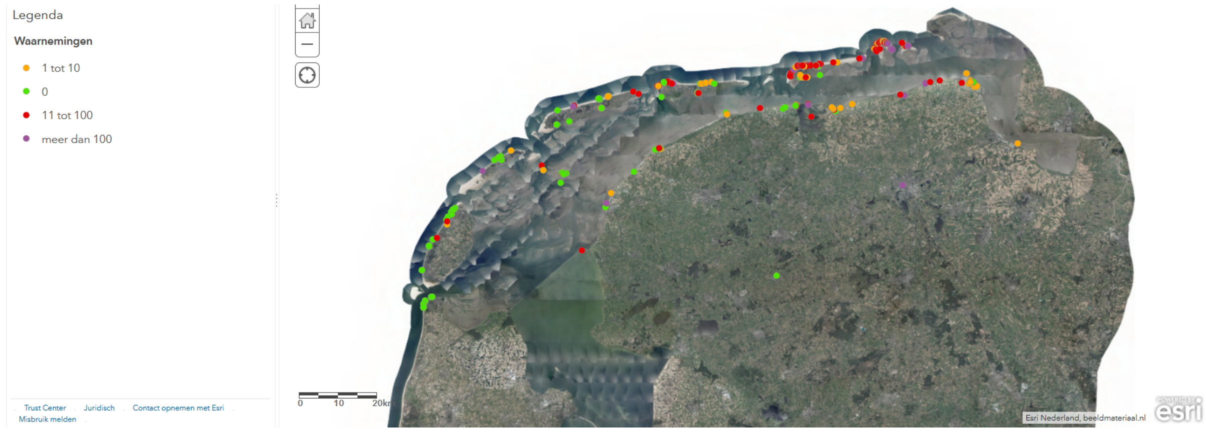


Figure 6.4: The speed over ground of the MSC Zoe (from the AIS data of the MSC Zoe) against the time and data

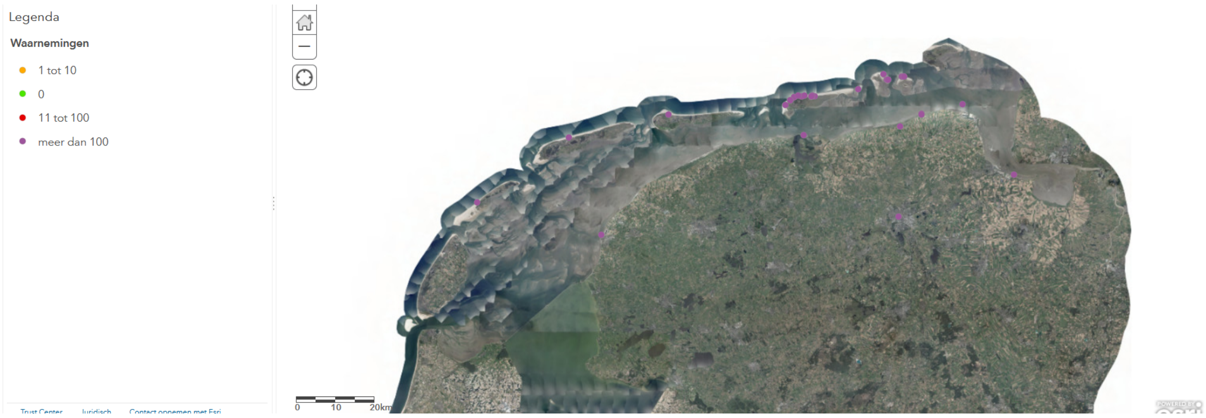
6.1.3. End locations of plastic particles

As mentioned in chapter 2 the University of Groningen set up a website to keep track of where plastic particles were found after the MSC Zoe accident. The counts are from how many particles the person saw in a grid of 40 cm by 40 cm, (University of Groningen, n.d.; 2019). All the counts that have been made can be seen in figure 6.5a. However, in this research only the purple dots are used. These are the 40x40 cm squares where the largest amounts of plastic particles are found. Only the purple dots can be seen in figure 6.5b. Furthermore, not all the purple dots are used, but a few are chosen. The few that are chosen are all modelled as 1 end particle, this is explained at the start of chapter 5. For each one will be calculated what the most likely overboard location is, which can be seen in the next section.

There are a few side notes to this data. For one, everyone can fill in one of these dots, so it is not known if every dot is accurate. However, the overall picture is mostly accurate because several sources state that the plastic particles washed ashore on the Wadden Islands and that Schiermonikoog was one of the Islands that was hit the hardest, (Dutch Safety Board, 2020; Nijpels, 2019a). Another side note is is that probably not all plastic particles are on this map. It is estimated that over 24 million plastic particles are in the Wadden, (University of Groningen, 2019). Not all of these plastic particles are on this map.



(a) The locations of all the dots



(b) The locations of only the purple dots

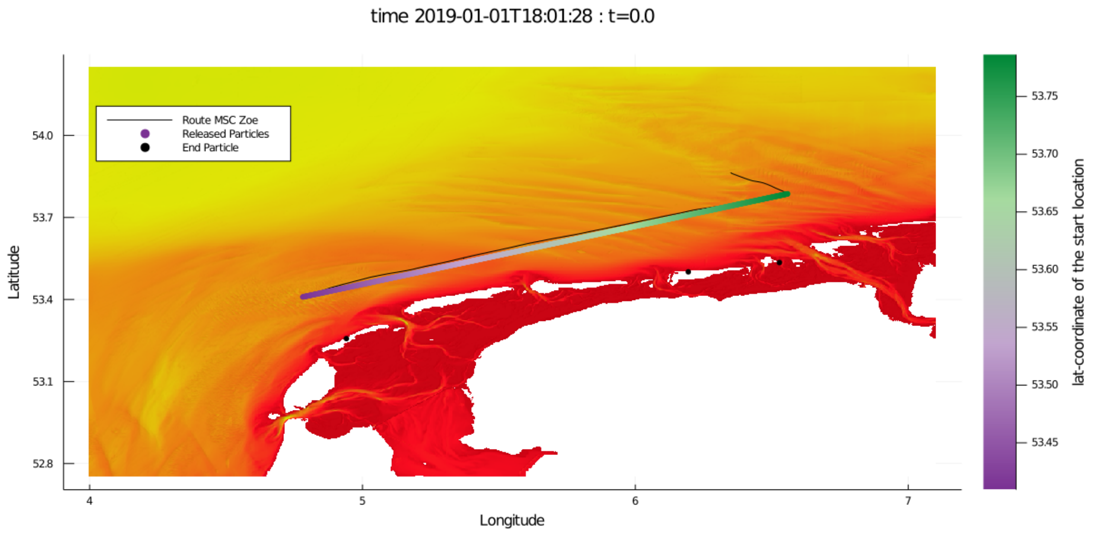
Figure 6.5: The locations of the found plastic particles, (University of Groningen, n.d.)

6.2. Results & Analysis of the Results

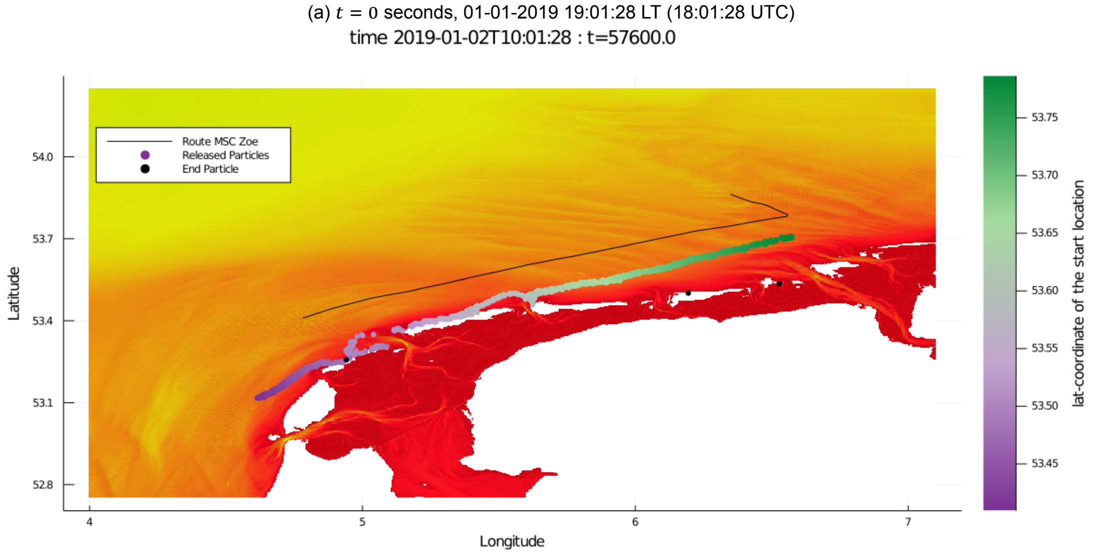
In this section are the results of the MSC Zoe case. It will start with the plots of the movement of particles. After that will come the results of the method of particle filtering. So what is the most likely overboard location of a specific end particle. The section will end with a comparison between the results of the method and the conclusions from the investigation about possible main overboard locations for containers.

6.2.1. The movement of particles

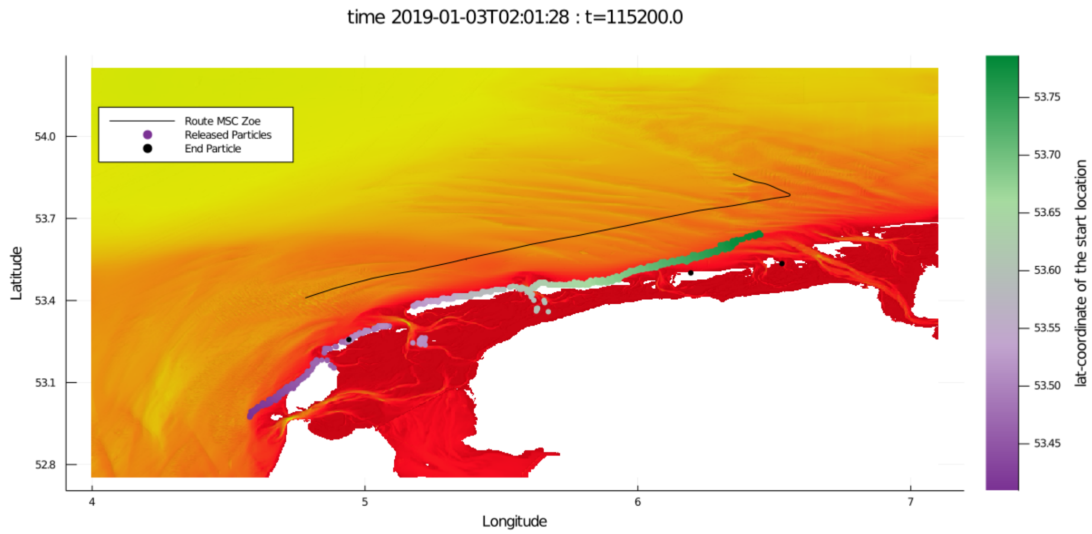
In the previous section is already explained what the flow is and what the start locations of the particles are. However, there are a few variables for which we still need a value before the simulation can be run. To start, the time variables. The investigation into the accident reported that it is very likely that the first containers fell overboard a little before 20:00 LT on January 1$^{\text{th}}$. So the $t_0 = 0.0$ seconds will be about an hour before this time, at 19:01:28 LT. This time is chosen, because the AIS data starts here and it is enough time before they suspected that the first containers fell overboard. The timestep will be $\Delta t = 1800.0$ seconds and the simulation will be run for two days, so $t_{end} = 172800.0$ seconds. Which is at 03-01-2019 19:01:28 LT. The time step is chosen such that there is a balance between how long the runtime is and accuracy (the smaller the timestep the more accurate the model). The end time is chosen such that the end position of the released particles is around/on the coast. Furthermore, the number of particles that will be released will be $N = 500$. Next, there will be eddy diffusivity, $\epsilon = (\epsilon_x, \epsilon_y) = (1.0, 1.0)$. The choice has been made to add eddy diffusivity, because of the turbulence in the ocean. Finally, the method of particle filtering will be applied for 3 end particles in the next section, so the method is applied 3 times. The end particles are chosen from the purple dots in figure 6.5b. These have the following coordinates: $(4.941, 53.257)$, $(6.194, 53.501)$ and $(6.528, 53.535)$. The end particles can also be seen in the plots of the movement of particles, these plots can be seen in figure 6.6. The figure shows 4 plots and between every plot is 16 hours. In this figure it can be seen that the particles move towards different points of the coast, depending on where they were released. It can also be seen that for every end particle there are released particles that ended up nearby.



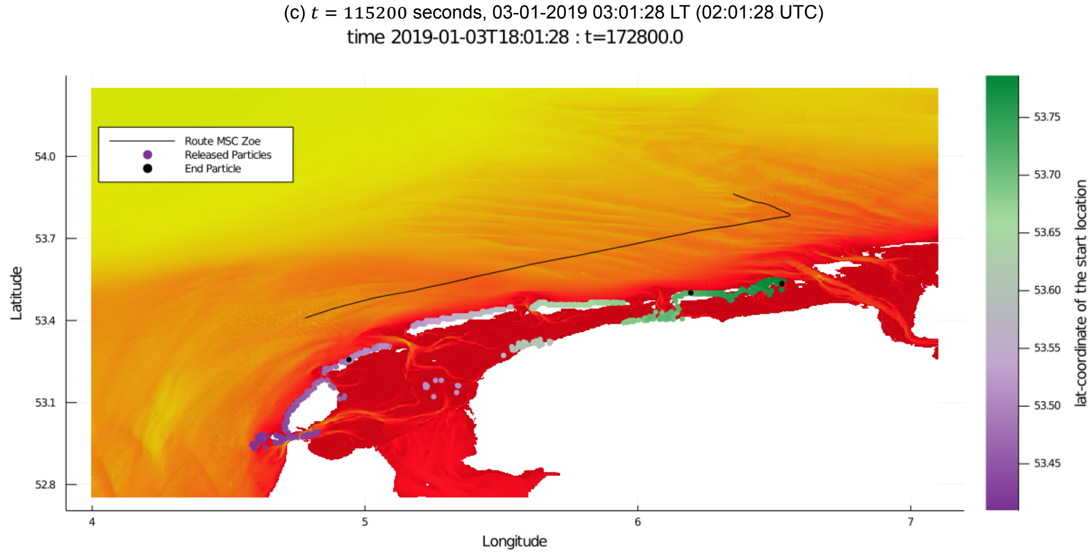
(a) $t = 0$ seconds, 01-01-2019 19:01:28 LT (18:01:28 UTC)



(b) $t = 57600$ seconds, 02-01-2019 11:01:28 LT (10:01:28 UTC)



(c) t = 115200 seconds, 03-01-2019 03:01:28 LT (02:01:28 UTC)

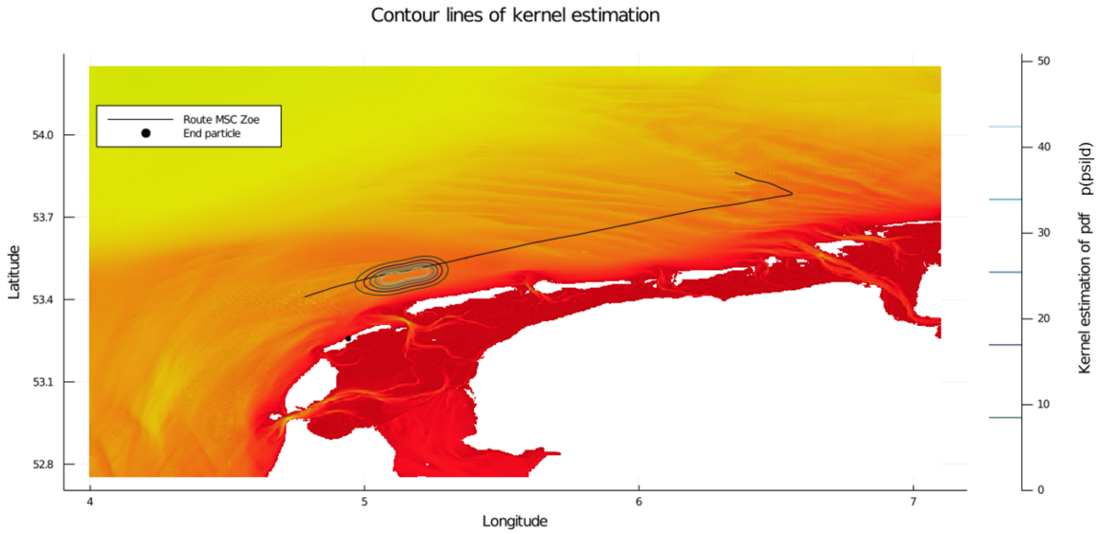


(d) t = 172800 seconds, 03-01-2019 19:01:28 LT (18:01:28 UTC)

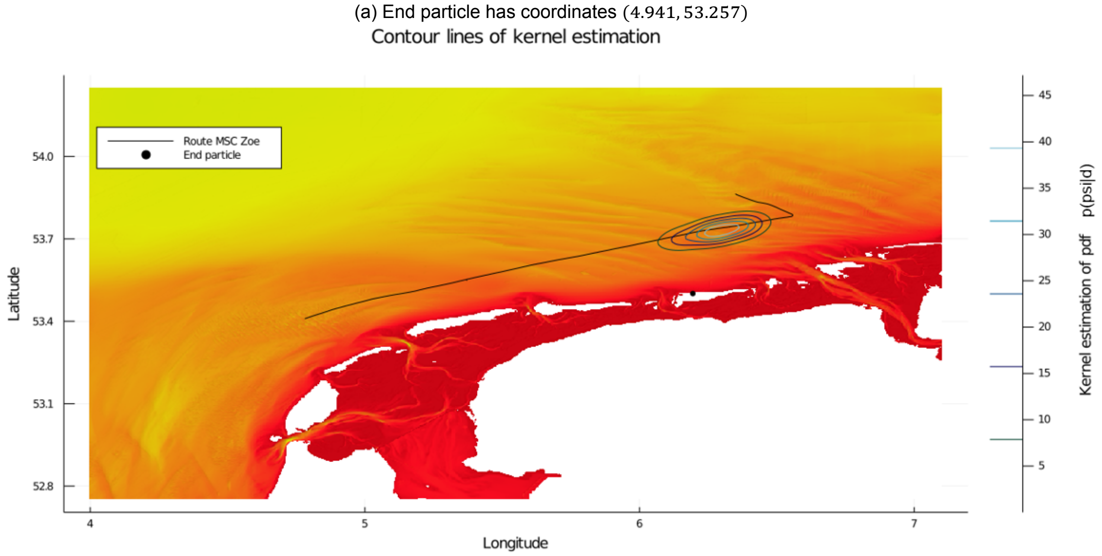
Figure 6.6: The movement of the plastic particles, with $N = 500$, $\epsilon_x = 1.0$ and $\epsilon_y = 1.0$. With a color gradient on the basis of the lat-coordinate of the start position

6.2.2. Results of the particle filtering method

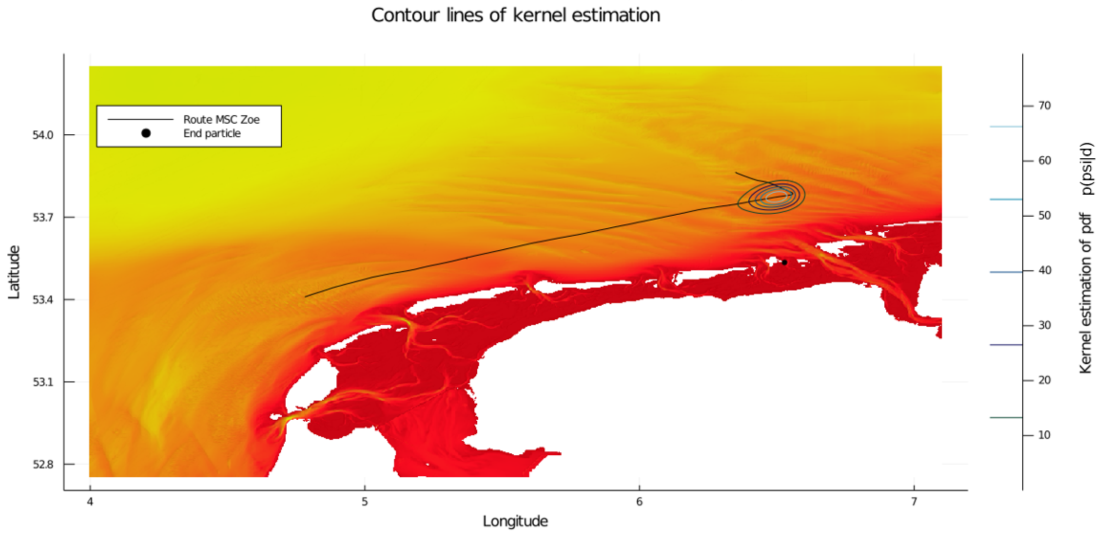
Next are the results of the particle filtering method. The choice has been made to directly use the kernel estimation, because it gives better results for the probability density function $p(\psi|d)$. In the calculation of the weights the value $r = 1000$ meters is used and in the calculation of the kernel estimation is the value $\sigma = 100$ meters used. For these values a balance is made between how much error/standard deviation is put into the formulas and what gives better results. As mentioned in the previous subsection, the method is done 3 times. For each end particle once. The results can be seen in figure 6.7. It can be seen that the likely overboard location of the particles really depends on its end position.



(a) End particle has coordinates $(4.941, 53.257)$



(b) End particle has coordinates $(6.194, 53.501)$



(c) End particle has coordinates $(6.528, 53.535)$

Figure 6.7: The results of the kernel estimation of the pdf $p(\psi|d)$, with $\sigma = 100$ m. The kernel estimation is drawn with contour lines.

6.2.3. Analysis of the results of the particle filtering method

A side note to the kernel estimation of the method of particle filtering is that the results really depend on the values of r and σ. This is also already mentioned in chapter 5. The larger r and σ are the larger the area that contour lines indicate is. This is because they are both indicators of how large the error, standard deviation, is in the values that are filled in.

Another thing to keep in mind is that an end particle can only be linked to an end location. There is no way in knowing how long it has already been there. This can have effect on the results, because of for example ebb and flow. Due to ebb and flow particles can come to the coast but later flow away again. They may end up on another part of the coast. So the end location of the released particles really depends on how long the simulation is run. They could, depending on ebb and flow, return to the ocean once they were on the coast and end up somewhere else. So then other particles may be closer to the end particle and the results will be different. So how long the simulation is run may influence the results in this way.

6.3. Comparison of the results with the investigation report

In figure 6.8 are the kernel estimations, given in the previous section, all in one plot. This is done to compare it more easily to the results of the investigation into the MSC Zoe accident. The plot of the six suspected main locations where containers fell overboard is repeated in figure 6.9 for easy reference.

It can be seen that the results of the kernel estimation of the pdf $p(\psi|d)$ are rougly on the same place as some of the 6 main overboard locations of containers. Namely, location 1, 5 and 6 of figure 6.9 are roughly on the same place as the 3 contour lines. So it depends on where the plastic particles ended up, on where they probably went overboard. This would also mean that the HDPE plastic particles were in more than one container, unfortunately this is not something that can be checked because the shipping manifest is not public information.

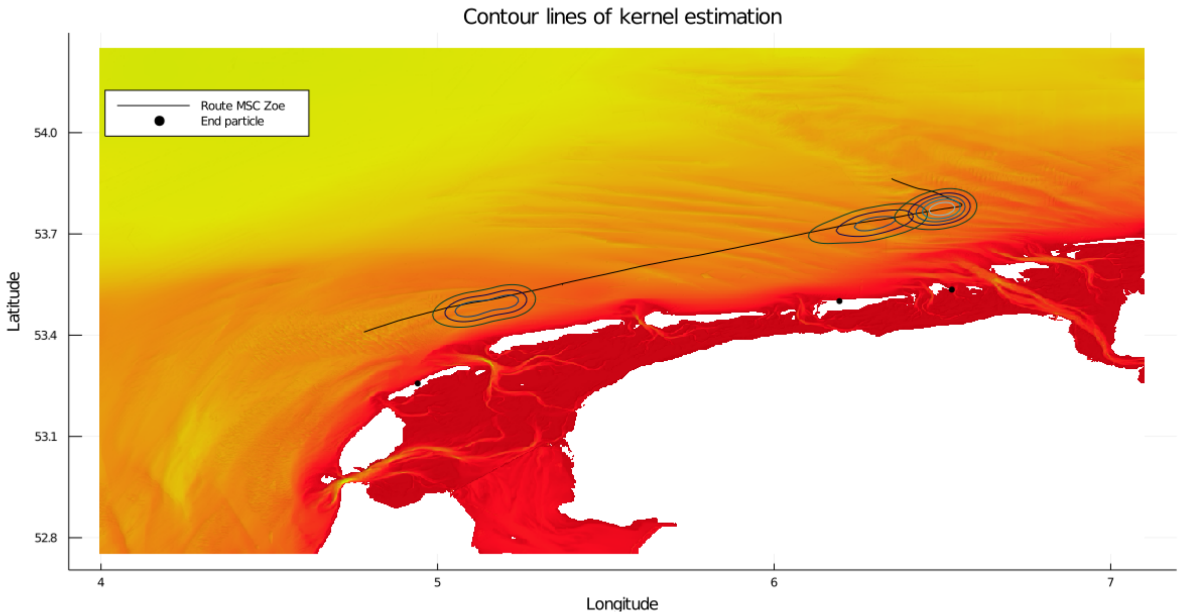


Figure 6.8: The results of the kernel estimation of the pdf $p(\psi|d)$ for all 3 end particles together, with $\sigma = 100$ m. The kernel estimation is drawn with contour lines.

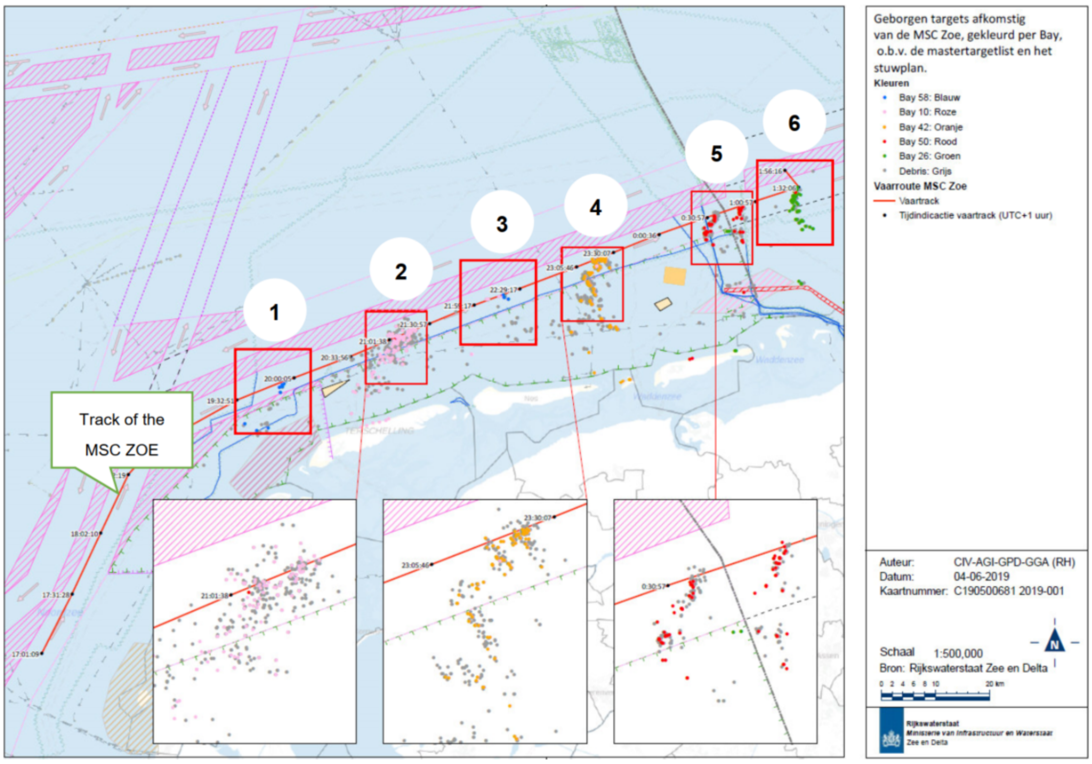


Figure 6.9: The six main locations that were identified for places that containers were found at in the ocean. The colours indicate bays which in turn indicate where on board the containers were, (Panama Marine Authority et al., 2020)

7

Conclusion, Discussion & Recommendations

7.1. Conclusion

In this report the following research question is investigated:

Can the method of particle filtering be used to find the start location of a plastic particles of which the end location is known?

In the results of both the example and the MSC Zoe case it can be seen that the method can indeed be adapted and used to find the start location of a specific plastic particle. In this report a simplification has been made, namely to find the start location of one specific particle instead of a concentration of particles. The last option would be more interesting for the MSC Zoe case, because huge amounts of particles were found and the particles were counted in concentrations on a grid of 40 cm by 40 cm.

As mentioned above the method can indeed be used to find a start location of one specific particle. In the example it gives the logical start position as the most likely. Additionally, in the MSC Zoe case it gives, for the end particles that are modelled, a location that corresponds to the results from the international investigation into the accident. However, the better results do come from a kernel estimation of the probability density function (pdf) that is calculated with the weights of the method of particle filtering. This gives a more smoother result of the probability density function $p(\psi|d)$, the pdf that we want to calculate. The problem of there being multiple peaks and that there are points with weight equal to 0 between these peaks have mostly disappeared.

The method of particle filtering has its advantages and disadvantages. An advantage is that it is very adaptable. The method of particle filtering could be changed from finding a most likely model run to finding the start location of one particle. A disadvantage is that the more particles released, the longer it takes to do the calculations. In other words the runtime gets longer. In the end it would be a good method to use for finding the start location, but the kernel estimation is needed to achieve good results. However, if a very large amount of particles is needed in the simulation then it might be better to look at a method that is faster.

Besides the conclusion, the answer to the research question, there are some other points of the research that need to be discussed. This takes place in the next section, the discussion.

7.2. Discussion

As mentioned above there are a few additional topics that need to be discussed about this research. First of all, the method of particle filtering can only work if the model for transport is correct. If this model is not correct then the released particles will not end up in the correct place. Then the weights are calculated according to these incorrect locations and the probability density function will not be correct.

Another side note is that the method can also assign high weights to particles that are not very close to the end particle. This would happen in the event that none of the released particles come close to the end particle. Then the particle that is 'closest' to the end particle will receive the most weight. However, it is not very close and the results will not be very accurate. So, to conclude if the results are accurate the plot showing the locations of the released particles needs to be looked at, to see if the particles have come or have not come close to the end particle.

Furthermore, the time delay that is added in chapter 6 depends on the time step. So the smaller the time step is, the more accurate the time delay. However, the smaller the time step is, the longer it takes to run the simulation. So a balance needs to be chosen for the time step, between accuracy and runtime.

The results in the MSC Zoe case say that the particles fell overboard on multiple locations. This is possible, because the investigation said that there were six main locations were containers fell overboard. To check if this is correct, we would need to know in which and in how many containers the HDPE pellets were. However, the shipping manifest is not public information so this can not be easily checked. Additionally, it is hard to check if the results are completely correct because it is not known when the containers with HDPE pellets broke open and when the bags containing the particles broke open.

Moreover, the weights are calculated on the basis of the location of the released particles at the end of the simulation. This location depends on how long the simulation is run. So how long the simulation runs influences the weights and in turn influences the probability density function that is calculated, the end result.

Lastly, the results also depend on the variables that do not have a set value. For example, N, r and σ. Also in this case a balance needs to be found between good results, accuracy, how much error there is and how long the runtime is.

7.3. Recommendations

There are multiple things that can be looked at for further research. Firstly, the model that is used to model the transport of the particles can be improved. If this model is improved than the locations of the particles that are released will become more accurate and in turn the results of the method of particle filtering will become more accurate. S. C. Ruiter researched if a 3D model is more accurate than a 2D model in the German Bright. This means that a depth component is added to the model. He showed that the 3D model gave improvements in the results, (Ruiter, 2020). This was done in a different part of the North Sea. So an option for future research could be too look if a 3D model gives better results in the part of the North Sea where the MSC Zoe accident happened.

Another way to improve the model is to improve the values from some variables in the model. In this report the eddy diffusivity is taken as a constant. However it would be more realistic that it does not have the same value at every location. So an adaption would be to find a formula to calculate the eddy diffusivity. The turbulence and eddy diffusivity do not matter for the method of particle filtering itself. However, as mentioned above for the method to work, the model used to simulate the transport of the particles needs to be correct. In the model are the terms for eddy diffusivity so if these terms are improved then the model improves. The influence of the eddy diffusivity is seen in where the released particles end up. So with a more accurate representation of the eddy diffusivity comes a more accurate representation of the end location of the released particles.

Additionally, the method of particle filtering can be adapted. One can look if the method can be adapted such that more than one end particle can be taken into account for the likelihood of where they went overboard. Now the method is applied in the situation where the start location is found for one specific end particle. However, in reality there were millions of particles that washed ashore in the MSC Zoe case. So an adaptation to the model could be that the observations describe, for example, concentrations of particles at multiple locations. Then it calculates the weights on the basis of this. So to find the start locations of the concentrations of particles counted in the 40 cm by 40 cm grids in the case of the MSC Zoe.

Another possibility for further research could be that the method of particle filtering is applied to find objects. For example, in the case of things falling of a ship. What if multiple objects fell overboard (it is not known where) and a few are found, and someone wants to find the other items. Then the method could be used to find the location that the items fell overboard and from there a simulation can be run to see where objects can end up if that place is used as a start location.

There is also the possibility to look at another way to find the start location from the end location of a particle. So to use another method than particle filtering. For example, another method to look at could be a backwards in time simulation of the particles. This method simulates the transport from the particles backwards in time from their end location. S. C. Ruiter has looked into this option in his research. He concluded that the results of this method are not very reliable, because the outcomes varied a lot, (Ruiter, 2020). However, future research in this method could maybe result in a more accurate method that can be used instead of or next to the particle filtering method.

Bibliography

Allen, J. (2011). NASA Earth Observatory image [Image only]. https://www.livescience.com/18445-eddy-ocean-nasa-satellite-image.html (Accessed: 07-07-2020)

Copernicus. (2020). OBSERVER: Cleaning our oceans with observations and models - Copernicus keeps track of plastic pollution in oceans. https://www.copernicus.eu/en/news/news/observer-cleaning-our-oceans-observations-and-models-copernicus-keeps-track-plastic?utm_source=Service+Desk+-+Subscribers+CMEMS&utm_campaign=5db8bf3762-EMAIL_CAMPAIGN_2020_04_20_08_41_COPY_01&utm_medium=email&utm_term=0_789b491dbc-5db8bf3762-205999393 (Accessed: 29-04-2020)

Do, C. B. (2008). The Multivariate Gaussian Distribution. http://cs229.stanford.edu/section/gaussians.pdf

Dutch Safety Board. (2020). *Veilig containertransport ten noorden van de Waddeneilanden: Lessen na het containerverlies van de MSC ZOE*. https://www.onderzoeksraad.nl/nl/page/13223/veilig-containertransport-ten-noorden-van-de-waddeneilanden.-lessen

Heemink, A. (2019). Numerical Methods for Stochastic Differential Equations [Course Notes TW3750TU].

Lee, A., & Ramster, J. (1968). The hydrography of the North Sea. A review of our knowledge in relation to pollution problems. *Helgoläinder wiss. Meeresunters*, *17*, 44–63. https://doi.org/10.1007/BF01611211

Made Smart Group. (n.d.). The World's Largest AIS Data Store By Made Smart Group [The AIS data of the MSC Zoe]. https://www.madesmart.nl/worlds-largest-ais-data-store/

Nencioli, F., d'Ovidio, F., Doglioli, A. M., & Petrenko, A. A. (2013). In situ estimates of submesoscale horizontal eddy diffusivity across an ocean front. *Journal of Geophysical Research: Oceans*, *118*, 7066–7080. https://doi.org/10.1002/2013JC009252

Nijpels, B. (2019a). *Zembla: De ramp op het wad (1/2)* [Documentary]. https://www.npostart.nl/zembla/17-10-2019/BV_101395849 (Accessed: 04-05-2020)

Nijpels, B. (2019b). *Zembla: De ramp op het wad (2/2)* [Documentary]. https://www.npostart.nl/zembla/24-10-2019/BV_101393876 (Accessed: 04-05-2020

Ooms, E. (2020). *Implementation of Stokes Drift into a Particle Model: Internship Deltares*. Delft University of Technology.

Panama Marine Authority, Dutch Safety Board, & Bundesstelle für Seeunfalluntersuchung. (2020). *Loss of containers overboard from MSC Zoe: 1-2 January 2019*. https://www.bsu-bund.de/EN/Publications/Unfallberichte/Unfallberichte_node.html

Plastic Soup Foundation. (2019). Extensive loss of pellets at sea remains without sanctions [Image only]. https://www.plasticsoupfoundation.org/en/2019/01/extensive-loss-of-pellets-at-sea-remains-without-sanctions/ (Accessed: 09-07-2020)

Ruiter, S. (2020). *Internship Deltares: Modeling particle drifters using a 3D velocity field and Stokes drift in the German Bight*. Delft University of Technology.

Scheijen, N. (2018). *Plastic litter in the ocean: Modeling of the vertical transport of micro plastics in the ocean*. Delft University of Technology.

Smid, R. (2019). 'Nederland en Duitsland kunnen aanvraag voor vaarverbod zuidelijke vaarroute Waddeneilanden doen'. https://www.lc.nl/friesland/Nederland-en-Duitsland-kunnen-aanvraag-voor-vaarverbod-zuidelijke-vaarroute-Waddeneilanden-doen-25075439.html (Accessed: 04-06-2020)

Spieksma, F. (2016). An Introduction to Stochastic Processes in Continuous Time. http://pub.math.leidenuniv.nl/~spieksmafm/colleges/sp-master/sp-hvz1.pdf

The European Marine Observation and Data Network. (n.d.). Bathymetry: Understanding the topography of the european seas [(The bathymetry image of the North Sea)]. https://www.emodnet-bathymetry.eu/

The Ocean Cleanup. (n.d.). Cleaning up the garbage patches. https://theoceancleanup.com/oceans/ (Accessed: 09-07-2020)

The Ocean Conservatory. (n.d.). The Problem with Plastics. https://oceanconservancy.org/trash-free-seas/plastics-in-the-ocean/ (Accessed: 09-07-2020)

University of Groningen. (n.d.). ArcGIS Wadden Plastic. https://www.arcgis.com/home/webmap/viewer.html?webmap=3eb9b39fda224e1086ce8b8ffdacf053&extent=6.0562,53.4204,6.2258,53.4812 (Accessed: 26-06-2020)

University of Groningen. (2019). Eerste onderzoeksresultaat Waddenplastic.nl: Schiermonnikoog hotspot van aangespoelde plastic korrels. https://www.rug.nl/news/2019/03/eerste-onderzoeksresultaat-waddenplastic.nl_-schiermonnikoog-hotspot-van-aangespoelde-plastic-ko (Accessed: 08-07-2020)

US National Ocean Service. (n.d.). What is bathymetry? https://oceanservice.noaa.gov/facts/bathymetry.html (Accessed: 12-07-2020)

van Leeuwen, P. J. (2009). REVIEW Particle Filtering in Geophysical Systems. *Monthly Weather Review*, *137*, 4089–4114. https://doi.org/10.1175/2009MWR2835.1

van Sebille, E., Griffies, S. M., Abernathey, R., Adams, T. P., Berloff, P., Biastoch, A., Blanke, B., Chassignet, E. P., Cheng, Y., Cotter, C. J., Deleersnijderk, E., Döös, K., Drake, H. F., Drijfhout, S., Gary, S. F., Heemink, A. W., Kjellsson, J., Koszalka, I. M., Lange, M., … Zika, J. D. (2018). Lagrangian ocean analysis: Fundamentals and practices. *Ocean Modelling*, *121*, 49–75. https://doi.org/10.1016/j.ocemod.2017.11.008

Verstraten, P. (n.d.). *Internship report: Deltares*. Delft University of Technology.

Wageningen University & Research. (2019a). Vijf vragen over de impact van de containerramp met de MSC Zoë. https://www.wur.nl/nl/show/Vijf-vragen-over-de-impact-van-de-containerramp-met-de-MSC-Zoe.htm (Accessed: 08-07-2020)

Wageningen University & Research. (2019b). Wadden Sea island Schiermonnikoog two weeks after the container incident with MSC Zoe. https://www.wur.nl/en/blogpost/Wadden-Sea-island-Schiermonnikoog-two-weeks-after-the-container-incident-with-MSC-Zoe.htm (Accessed: 08-07-2020)

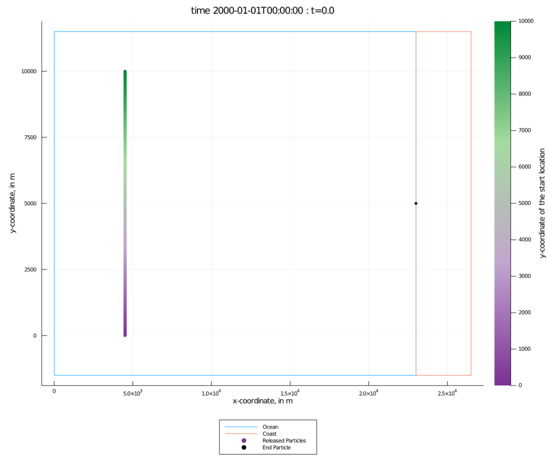
Winther, N. G., & Johannessen, J. A. (2006). North Sea circulation: Atlantic inflow and its destination. *Journal of Geophysical Research*, *111*, 1–12. https://doi.org/10.1029/2005JC003310

Zijl, F., & Groenenboom, J. (2019). *Development of a sixth generation model for the NW European Shelf (DCSM-FM 0.5nm): Model setup, calibration and validation*. Deltares.

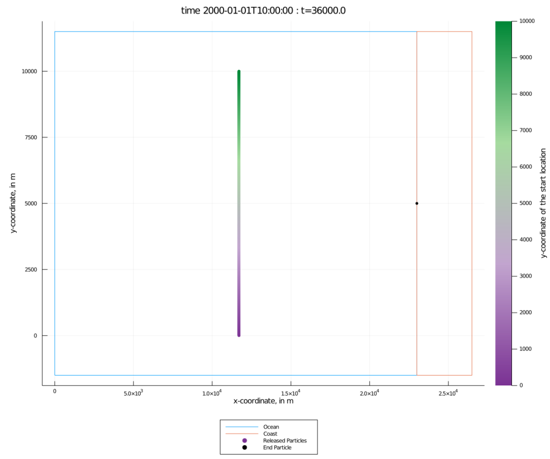
A

Extra Plots

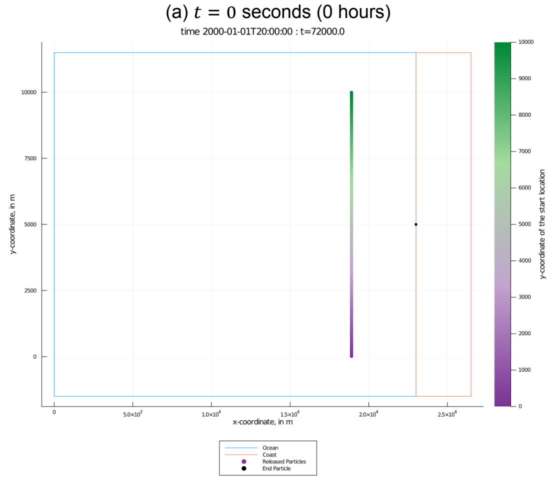
Movement of particles from the example without eddy diffusivity



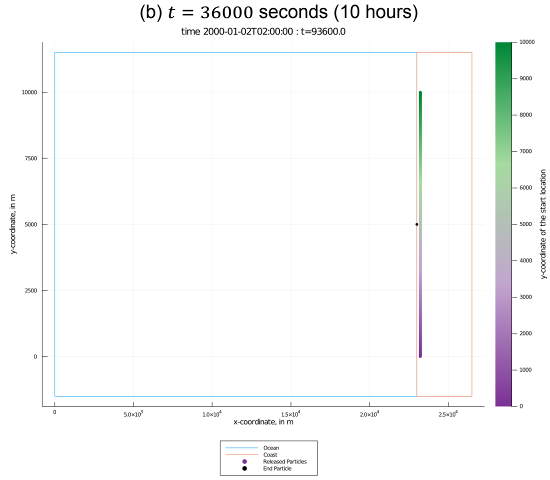
(a) $t = 0$ seconds (0 hours)



(b) $t = 36000$ seconds (10 hours)



(c) $t = 72000$ seconds (20 hours)



(d) $t = 93600$ seconds (26 hours)

Figure A.1: The movement of the plastic particles, with $N = 5000$, $\epsilon_x = 0.0$ and $\epsilon_y = 0.0$. With a color gradient on the basis of the y-coordinate of the start position