Ocean Science Discussions

Ocean Sci. Discuss., author comment AC1
https://doi.org/10.5194/os-2021-83-AC1, 2021

Reply on RC3

Mikael L. A. Kaandorp et al.

Author comment on "Using machine learning and beach cleanup data to explain litter quantities along the Dutch North Sea coast" by Mikael L. A. Kaandorp et al., Ocean Sci. Discuss., https://doi.org/10.5194/os-2021-83-AC1, 2021

Summary and scientific relevance:

The purpose of this manuscript is to investigate the physical processes leading to the accumulation of litter on the beaches of the Dutch North Sea coast. Data from six years of beach cleanups were used to fit a variety of environmental parameters with a random forest model to identify possible correlating variables affecting litter accumulation on the coast. Tidal height and variability were found to be the strongest explanatory variables, leading to a decrease in litter accumulation at the coast with increasing tidal height and variability. In addition, shoreline geometry was found to have explanatory power for litter accumulation on Dutch beaches. Based on the best explanatory variables, the authors extrapolated the distribution and abundance of beached litter on the Dutch North Sea coast, which may contribute to effective cleanup strategies by identifying hotspots of litter accumulation along the coast.

General comments:

The manuscript is well organized and generally easy to follow. The presentation of data and methods is very well structured and the results are clearly illustrated. The method presented for studying coastal litter accumulation is very innovative and can help provide valuable insights into the processes governing the beaching of marine litter in shelf seas. I wonder about two points that I think are particularly important for coastal shelf seas. I would like to recommend this manuscript for publication after moderate revision according to the following points.

We would like to thank the reviewer for these kind words, and the very useful comments below.

- **As far as I understand it correctly, the authors of this study used the AMM7 model of the Copernicus Marine Service for the advection of the virtual particles. One key point I wonder about is how the authors analyzed the**

influence of tides on marine litter washing ashore on Dutch beaches. The authors found that tidal height is the most important influencing variable, resulting in a decrease in the amount of litter as the variability and height of the tides increases. Numerical studies (e.g., Stanev and Ricker, 2020) and observational studies (e.g., Meyerjürgens et al., 2020) have found that tidal forces (including overtides) have a significant influence on the length of Lagrangian trajectories and particle residence times, in addition to affecting particle dispersion at different spatial scales. Since the authors used daily average values of ocean current fields, the effect of tides is suppressed in the analyzed Lagrangian model, which is very important for the reader to classify the results of this study. Please clarify and discuss the important effects of tidal currents on the results of this study.

We use the CMEMS North West Shelf Reanalysis product in our analysis for the ocean surface currents. This product uses 15 tidal constituents, so their net effect is contained in the Lagrangian particle simulations (i.e. the long-term effect on time scales above a day). In our first submission we included additional tidal forcing in the Lagrangian particle simulations as well, using the FES2014 data, at a high temporal resolution (the integration time step in the Lagrangian model is 20 minutes). But, as reviewers noted correctly in a related paper we submitted (van Duinen et al. (2021), 10.1002/essoar.10508992.1) this means the effect of tides is overpredicted. We therefore re-did our simulations without this extra tidal forcing.

To still include the high frequency effect of tides on our analysis, we included the tidal height as a separate feature. Since this is obtained from the FES2014 data (which is spectral data), we can calculate the tidal components at any frequency we want. This allows us to calculate the tidal height and currents at any given lead time, or what the maximum tidal height was during the cleanup stage. Given the feedback below, we have now extended the previously included components (M2, S2, K1, and O1) with the M4 and M6 components.

We have added some comments on the different effects that tides might have on litter concentrations in the main text (line 314 track changes):

In general, a higher tidal maximum and variability lead to less litter measured on the coastline (see the Appendix B5 for further details). A higher tide during or preceding the cleanup could re-suspend some of the litter from the beach. Furthermore, a higher tide encountered during the cleanup stage reduces the beach width that can be sampled. Perhaps a stronger variability in the tidal height leads to less persistent high strandlines where the highest litter concentrations are normally found (Heo et al., 2013). It has been shown in numerical studies that residual tidal currents can lead to a net transport of both suspended and floating matter (Gräwe et al., 2014; Børve et al., 2021; Schulz and Umlauf, 2016). While the regression model indicates that tides play an important role, it is difficult to separate the causal relations between all these different effects and the litter quantities found on beaches. To quantify this in more detail, further experimental and numerical studies are required.

Since this is an important aspect, we have added this as a recommendation for future research to the conclusions and discussions (line 475 track changes):

Future studies could further investigate the causal relations between the variables seen as important predictors by the regression model and the litter concentrations found on beaches. This is especially the case for tides, which constitute the two most important features in the regression model (see Figure 5). Experimental studies could further determine whether lower litter concentrations at locations with higher tidal variability are mainly caused by litter re-suspending back into the sea, or for example due to the fact *that less area of the beach is sampled during high tide. It should additionally be investigated how these effects compare to the role of (residual) tidal currents, as it has been shown that this can play an important role in transporting suspended matter towards the shore (Schulz and Umlauf, 2016). Experimental investigations can be done in combination with numerical studies of the nearshore marine environment, to capture the interactions between processes such as tides, waves, and particle sizes (Alsina et al., 2020).*

- **Another point I wonder about is why the authors of the study use anthropogenic variables such as fishing intensity as possible sources (for the Lagrangian model), but do not include anthropogenic factors (other than population density) in the statistical analysis. I would recommend that the authors include anthropogenic parameters such as "ship density" and "fishing activities" (which can be taken from EMODNET, for example) as possible explanatory variables in their analysis to get a sense of how anthropogenic factors affect coastal litter accumulation compared to hydrodynamic and atmospheric parameters.**

We have added more anthropogenic factors to the set of features to make this consistent. The population within a given radius was already included ($n_{pop.}$), we have added the fishing effort in a nearby radius as well now ($n_{fis.}$). Furthermore, we have included salinity (sal.), a parameter which can be used as a proxy to see how close a given beach is located near upstream river mouths. The entire analysis was redone including these features.

Our results now show that fisheries likely play a large role in the amount of litter that is found on the Dutch coastlines. The Lagrangian model run with particles released at fishing locations is now one of the most important explanatory features, ranked 3rd. In general, the Lagrangian model simulations now play a more important role compared to our first submission. This can be explained due to the fact that the new Lagrangian model runs are likely more accurate now, as in the first submission the effect of tides on the particle trajectories was overestimated. Additionally, nearby fishing intensity ($n_{fis.}$) is ranked as the 10th most important feature.

We have not included ship density as a separate variable, as a lot of the litter on beaches is estimated to originate from the fishing industry specifically (see .e.g van Duinen et al. (2021), 10.1002/essoar.10508992.1). Littering from ships in the North Sea is highly regulated, and likely more of a very local nature in space and time (e.g. container accidents doi.org/10.3389/fmars.2021.607203). This is difficult to include in the model without having a priori knowledge of when and where this happened, which is out-of-scope for this project.

Specific comments:

Line 110: The authors have considered the most important tidal components in their calculations. In the North Sea, shallow-water tides (M6 and M4) play an important role in the currents and advection of particles. Please keep this in mind in your analysis and consider it when discussing your results.

We thank the reviewer for this comment. We have added the M4 and M6 tides to our analysis, as we indeed found in the literature that these components play an important role in the transport of suspended particles. We have added to the manuscript (line 113 track changes):

Tidal currents (U_{tides}) and heights (h_{tide}) are calculated, taking the M_2, S_2, K_1, and O_1 constituents into account (Sterl et al., 2020), as well as the M_4 and M_6 components which have been shown to play an important role in transport of suspended particles in the North Sea (Grawe et al, 2014).

Line 143-144: Unclear to me. Please rephrase this.

We have rephrased this part (line 153 track changes):

Each virtual particle starts with a unit mass. Each time step that a virtual particle spends near the coast, a fraction of its mass is lost due to the beaching process. This means that as t_{coast} increases for a virtual particle, a fraction of its mass is lost, which is calculated using (1). For each virtual particle, we calculate where and when it loses mass due to the beaching process. These masses lost to beaching are binned in a 1/9° x 1/15° beaching flux histogram for each day.

Line 174-175: I can't see a dashed brown line in Fig. 3a. Do you mean the brown arrow in this context?

Yes, this should indeed be an arrow, we have adjusted the text.

Line 213: Please clarify why Utide is not considered as a scalar feature.

We have added the tidal currents as scalar and directional features now, as this was indeed not consistent. At first, they were left out since the tidal height was already included, but we added them now since it might be important to consider the currents to include the effect of residual tidal transport. As can be seen in the main text, one of the important explanatory variables is now the dot product of the tidal velocity with respect to the coastline.

Line 344: This sentence seems incomplete.

We have modified the sentence a bit for more clarity (line 395 track changes):

For the smallest lag distance (h = 5 +- 5 km), we find γ = 0.08. This variance estimate was also used to create the error bars in Figure 4.

Line 352: Where does this grid size come from? AMM7 should have a resolution of 7 km x 7 km. Perhaps it arises from the inclusion of lower resolution Stokes drift data in the numerical grid? This is not clear to me. Please clarify and add a section in Section 3.1.1 on how your grid size is defined by merging the different data sources.

The different sources of data have different grids. We have clarified this in the text now (line 404 track changes):

The grid sizes used for our numerical data ranges from about 7 km (the surface current data), to about 20 km (the wind data). This means that the variance at and below these length scales is not captured by the numerical data. The variance calculated for lag distances up to 20 km is quite substantial (γ = 0.05-0.12).

Technical corrections:

Line 13: … the Dutch coastline.

We consider the Dutch North Sea coastline here, inland waters and shorelines around the Wadden Sea are not taken into account

Line 15: … the need for…

adjusted

Line 75: …end of the stage…

adjusted

Line 76: Most litter found… or Most of the litter found…

adjusted

Line 152: … on how…

adjusted

Line 184: …that are…

we are not sure what is meant here, the sentence seems fine to us

Line 218: …are calculated…

adjusted

Line 225: … is used…

adjusted

Line 227: … is available.

adjusted

Line 236: …from the coastal population.

adjusted

Line 323: "not to have" not well formulated, please rephrase.

Changed to

It is estimated that the number of participants taking part in the tour does not have a large influence on the amount of litter that is found

Line 363: … coastline into…

adjusted

Line 381: …play an important role…

adjusted

Line 386: …is seen…

adjusted

Line 408: …the importance…

adjusted

Line 420: … can be taken into account.…

adjusted