We greatly thank the referee for the effort he applied on his review and for his helpful comments.

In the following, the referee's comments are shown in blue.

In this paper, an investigation of the properties of relative dispersion, structure functions and spectra is presented, from drifters released in the German Bight. The paper is written in a rather clear and competent way, but the results are in my opinion insufficiently robust and inconclusive.

I think the paper is not publishable in its present form, and it should go through a major revision or a resubmission.

MAIN COMMENT

The data set is relatively small (a total of 19 drifter pairs), and the authors choose to present dispersion properties for each pair independently, attempting to discuss their individual characteristics and statistics. They justify this approach in terms of coastal inhomogeneity which would prevent a global statistical approach. This hypothesis, though, is not sufficiently substantiated by the data as discussed in the following, and the end result is that the statistics of each pair (with duration of 1-4 days) is too poor to reach robust conclusions.

We fully agree (and state that in the paper) that, due to the low number of drifters, our findings are not robust in a statistical sense. Fig. 8 shows very clearly how differently drifter pairs with relatively large initial separation (> 9 km) behave. Also for smaller distances (< 1 km) a comparison of Figs. 5 and 7, for instance, suggests that averaging over different drifters would not generate useful information. We agree with Referee #1 that only future experiments could improve the situation. For the time being, we believe that the best that can be done is to summarize all (admittedly weak) indications available.

My suggestion is the following. I think that the authors could indeed start with a description of the individual launches, in terms of geographical positions and wind and tidal forcing, without though going in the details of the individual dispersion plots and fits. After the general presentation, I think the authors should present some clear working hypotheses on parameters that could influence the statistics, that will then be consistently tested throughout the paper. The parameters could be related to topography, forcing or distance from offshore wind farms (OWF). These hypotheses will be tested though conditional statistics, using selected sub ensemble of data. Given the small number of data, the conditional sub ensembles should be as broad as possible, based on the chosen parameter.

The results from these conditional statistics will then be compared with the total statistics obtained from all the pairs, in order to verify whether or not significant differences emerge.

The referee asks for a formalized statistical analysis, testing well-defined hypotheses. We agree that a number of different parameters could influence drifter behaviour. Forcing (weather conditions) undoubtedly is among these important factors. Our analysis combines three experiments at different times. We do not see, however, how different weather conditions at these times could be formally described or characterized. Weather conditions cannot be characterised in terms of just one parameter. During experiment HE496 wind speeds tended to be smaller than during the other two experiments. Does that already mean that environmental conditions during HE445 and HE490 fall into one class (regarding weather) while conditions during HE496 establish a second class? Experiment HE496 also happens to be the experiment in which drifters travelled at larger distances from the wind farm. How could the impacts of these two factors be separated from each other?

We are afraid that formalizing the study in terms of conditional statistics would generate a substantial formal overhead without promising a clear benefit. Note, however, that in Figs. 11, 12 and 13 we already did some conditioning, showing distributions and structure functions for different groups of drifters, roughly defined in terms of distance from wind farms. This grouping is necessarily qualitative, considering also the fact that these distances change when drifters move. It is in HE496 (drifter set C, Fig. 6) that larger distances from the wind farm occur.

We appreciate the referee's intention to improve the common thread of the discussion. To clarify the general structure of the analysis we included the following introductory paragraph at the very beginning of Section 3 (Results): "*Section 3.1 presents details of all drift trajectories analysed in this study. Plots show how drifters are located relative to wind farms and which winds they are exposed to. In Section 3.2 kinetic energy spectra are studied to assess the possible relevance of tidal movements as a source of turbulent energy. Section 3.3 then presents probabilities of separation velocities and velocity structure functions. To check the hypothesis that drifter separation might be influenced by wind farm related turbulence, these functions are shown for different groups of drifters, separating in particular those drifters that are far enough to presumably not experience wind farm effects. The section concludes with some results of simulated drifter dispersion (Section 3.4).*"

This will provide a logical structure to the paper, and a setting that will allow testing working hypothesis. It might be that the data set is too small and the errors are too big to actually differentiate between conditional statistics, but at least this will be shown in a quantitative way. In the present version of the paper, the authors actually take a similar approach for the discussion of the spectra and structure functions, but the hypotheses are not presented in a clear fashion and are not consistent throughout the paper.

As the referee already states, strict hypothesis testing will not be possible given the small number of drifters and the variety of uncontrolled influencing factors. We cannot see how a statistical formalism could help overcome this very obvious fact. A major problem is also that distances between drifters and wind farms are ill-defined parameters. Given the size of the wind farm, it is not clear how such distances should be measured. Effective distances might also depend on wind direction relative to a drifter's location. It is possibly not very beneficial to apply formalized statistics to a small number of values that are just vaguely defined.

DETAILED COMMENTS

- Section 1
    - Lines 1-5. There are a number of recent papers that investigate "local" initial conditions (e.g. Ohlman et al, 2017; Berta et al., 2016; Poje et al., 2014)

      These three references were already referred to later in the paper. But we agree that they should be mentioned already here in the introduction. We changed the passage accordingly: "*The sub-mesoscale we focus on has also been addressed by numerous other studies (e.g. Berta et al. 2016; Ohlmann et al. 2017; Poje et al. 2014). A key observation is that spreading* *rates may be much higher than those observed on the large-scale (Corrado et al. 2017).*"

- Line 20 Please expand on the mechanisms through which OWF are expected to impact on surface dispersion
  The third paragraph of Section 1 (Introduction) has been revised, extending the already existing summary of relevant processes (wakes, vertical mixing, atmospheric or marine turbulence).

- Section 2
  - Lines 10-20. Please discuss expected slippage errors of the MDO3 drifters. Have they been quantitatively tested? and compared with other types of drifters such as the classic CODE? Please provide references
    Slippage errors are now addressed in a new paragraph (second paragraph of Section 2.1). "*Although Albatros MD03 drifters have been widely used during the last years (e.g. Lana et al. (2016), Callies et al. (2017), Sentchev et al. (2017), Onken et al. (2018)), to our knowledge slippage of this drifter type has never been quantified. However, considering the drag ratio of 33.2, the parametrization exposed in Niiler et al. (1995) would predict a slippage of 1.1 to 1.6 cm/s, for 10 m/s wind speed and a velocity difference across the vertical extent of the drogue of roughly 0.1 cm/s. Quantification of a drifter's slip is not trivial due to an influence of sea-state. For another type of drifter, the {CODE} drifter, Poulain et al. (2009) estimated slippage to be 1 % of wind speed. By contrast, according to Poulain and Gerin (2019) slippage was estimated to be 0.1 % of wind speed. Fortunately, specification of slippage effects is of minor importance for the present study. First, it can be expected that slippage effects affecting two drifters of the same type will not dominate separation of these drifters. Second, when comparing observations with corresponding simulations, the additional wind drag tuned for successful simulations will cover also slippage effects. Therefore, for the present study slippage effects were neglected.*"
    Unfortunately drifter specific estimates are not available. However, it seems plausible that slippage effects will not dominate separation of identical drifters exposed to the same forcing. The wind drag assumed for numerical simulations will implicitly cover also slippage effects without, however, distinguishing them from effects of Stokes drift, for instance.
  - Table 1. It should be improved or complemented by an other table. Initial distances between pairs and distances from OWFs should be included.
    We see the point that spatial scales of drifter separation should be indicated more clearly. To solve this problem, we added in each plot of drift trajectories (panels in Figs. 2, 4 and 6) an explicit length scale, which in particular emphasizes the small initial distances between drifters (< 100m).

    In Section 2.1 (after the description of the three drifter sets) we clearly state that initial drifter separations shown refer to the time at which the first signals were received from the positioning system. That means that initial separations are even smaller than shown, unfortunately the precise values cannot be specified.
    The referee would like to see information on initial distances from wind farms being included in a table. We thought about this idea but came to the conclusion that such information cannot be given in a meaningful way. Figures like Fig. 1a, for instance, show that the distance in question is much smaller than the size of the wind farm. This means that it would rather arbitrary choice how to define the reference location of the wind

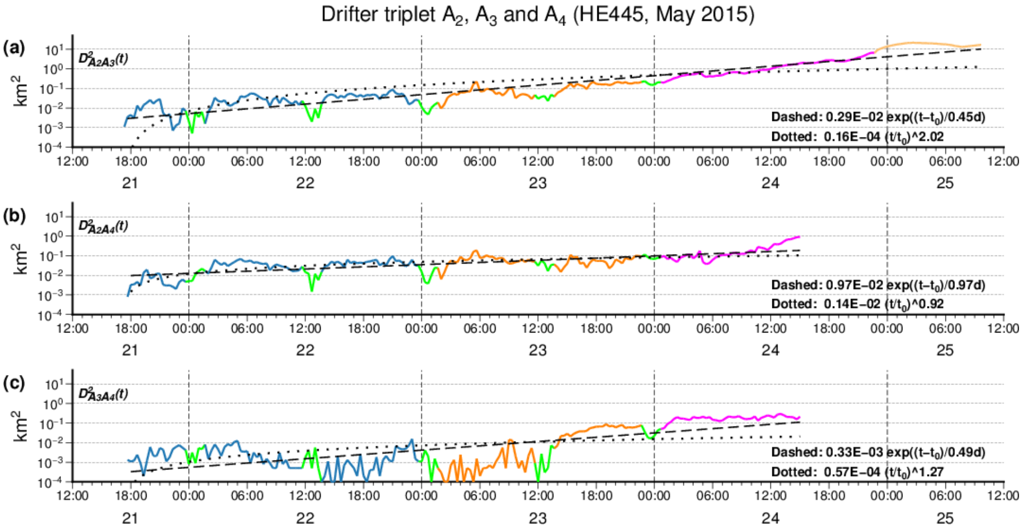
farm. Should it be the location of the nearest engine or instead the centre of the wind farm? This choice would dominate the value one obtains. Therefore we came to the conclusion that a pure listing of such fuzzy numerical values would not be helpful for the reader, given the fact that the information the referee asks for is easily accessible from the trajectory plots in Figs. 2, 4 and 6.

  - Also in the text, in Section 2 and 3, please be more quantitative, avoid mentioning that pair are “close” or far, and refer to the i.c. in Table 1.
    We presume that this remark addresses mainly the discussion of Figs. 11-13 in Section 3.3 where we classified drifters with regard to their location relative to wind parks. As already mentioned, giving absolute distances is difficult as these are time dependent and wind farms cover large areas. However, the group of drifters being close to wind farms can also be described as those that even entered the wind farm area. Throughout the paper we now use this more precise wording.
  - Section 2.4. Please specify model initial distances between pairs and comment on the fact that given a model resolution of 900 m, local structures beyond 2-4 km are not correctly resolved.
    To simulate drifter dispersion, all particles are started at exactly the same location. This is said in the caption of Fig. 14 (“…*100 trajectories initialized at the same location…*”) and also the first paragraph of Section 3.4 (“…*spreading from a common source point…*”). In the revised manuscript we now also included in Section 2.4 the following sentences, which explicitly address the problem of lacking grid resolution and stress the point that no initial particle separation is needed for simulating dispersion: “*Grid resolution limits the scale of flow features that can be resolved. Drifter separations of less than 1~km are clearly beyond the resolution of BSHcmod. The general approach to overcome such problem is to include sub-grid scale turbulent processes via a scale-dependent random diffusion term. With such approach being implemented, even particles released at the same initial location will start separating.*”
  - Fig. 1. It should be improved, showing the deployment design and the topography
    Thank you for giving this hint: Although in Fig. 1 the bathymetry was already shown, the numeric scale corresponding with the different colours was missing. In the revised manuscript, a corresponding legend has been added to the figure. We also found that in the horizontal length scale an error had slipped in. This has been corrected.

    Fig. 1 is meant to give an overview of the larger region where wind farms and corresponding drifter experiments are situated. At the spatial scale of Fig. 1 it is impossible to display the deployment design of the small scale drifter experiments. However, Fig. 1 clearly indicates the locations of the two wind farms within the German Bight region. Throughout the paper, each plot of drifter trajectories (such as Fig. 2a, for instance) shows these farms in much larger resolution. In our opinion each of these detailed plots, resolving even individual wind engines, displays very clearly how the respective drifters were deployed relative to the wind farm.
- Section 3
  - Fig.3 5,7 and related text. The exponential fit seems very arbitrary to me. Were other fits tested as well? The initial distances from which the fit start should be mentioned. Please discuss errors and confidence limits. In order to compare results, the initial distance should be comparable. See also the

point on model pairs above. In general, please see General Comment above.

In an earlier version of the manuscript we also provided a fitted power law. However, these fits are very sensitive and obviously do not provide better results than the exponential fit. The following figure shows this for the example of Fig. 3 (see additional dotted lines). We therefore decided to not include this in the paper.



Panel (c) of the above figure also provides an example of how the data the exponential law is fitted to do not just correspond to a signal superimposed by some (e.g. Gaussian) noise. Roughly between 22 May 12:00 and 23 May 12:00 the fluctuations of squared drifter distance do not seem to be purely random. Therefore the exponential model is just a possibly weak indicator that underlying processes are not too far from theoretical expectations. Specification of uncertainties is not really meaningful or even possible in such context. We now comment on this problem at the end of Section 3.1.1: "*In sum, the exponential model should be seen as just an indicator of what could be expected theoretically. Specification of uncertainty bounds of the fitted model does not seem reasonable in this context.*"

- Section 3.2. The computed spectra are in time, while the general discussion in 2.2 is in terms of wavenumbers. Please discuss the hypotheses used to link the two types of spectra. The drifter spectra (except for one case) are obtained from time series of 1-3 days. Can they effectively resolve tidal frequency, even using MMT? Please discuss errors and confidence limits.

  In Fig. 9 we considered energy as a function of frequency as this is the natural approach for the analysis of local time series. A transformation into the domain of wave numbers would have to be based on the assumption of some transport velocity. Panels in Fig. 10 are thought to be directly contrasted with Fig. 9 so that changing the independent variable would not make sense. An important aspect in the section is to identify the relevance of tidal motions. The most straightforward approach for doing that is an analysis in terms of frequencies, needing no further assumptions.

  As suggested by the referee, we checked statistical significance of spectral peaks, a corresponding paragraph added at the end of Section 2.2 mentions the methods applied and gives all relevant references: "*Besides* *all mentioned advantages, a drawback of the MEM method is that the statistical significance of the spectral peaks is difficult to assess. Nevertheless, to estimate the statistical significance of spectral peaks the method applying a permutation test (Good, 2000) as proposed and exemplified by Pardo-Igúzquiza and Rodríguez-Tovar (2005, 2006) has been followed. Identified spectral peaks referred to in the discussion section show high statistical confidence levels with values between 95% and 99% based on the permutation test (10,000 spectra) using an underlying red noise spectrum.*" However, it is also to be noted that the tidal constituents indicated in Figs. 9 and 10 (magenta coloured lines) were not analysed from the data. They rather represent the values that are expected according to physics.

- Section 3.3. What do the authors mean by "Eulerian and Lagrangian" separation?

  The corresponding explanation has probably been a bit too short. We added (third paragraph of Section 3.3) the exact definition of the Lagrangian velocity increments: "*Increments $\delta v^{(L)}(t)$ were obtained as differences between velocities of the same drifter at times t and t+τ, where τ=20 min corresponds with the time resolution of drifter observations.*" Regarding Eulerian velocity increments we now explicitly refer to the definitions given in Eqs. (3) and (4).

- Section 3.4. What are the initial distances of the model pairs? Given the model resolution, the dynamics is not expected to be local beyond 2-4 km, so that the exponential behavior is simply a consequence of the setting.

  The referee is absolutely right, the exponential growth of distances is to be expected when this kind of parameterization is used in numeral modelling. Fig. 14 was included to demonstrate that. Initial distances between particles were assumed to be zero, stated in the figure caption: (" … *100 trajectories initialized at the same location …*") and at the beginning of Section 3.4 *("… spreading from a common source point* …").