Dear editor,

in my opinion the revised manuscript gained a lot from the very constructive comments by two referees and the discussants Rodrigo Duran and Jens Meyerjürgens.

Addressing the concerns by Rodrigo Duran, the manuscript now much better deals with the distinction between convergence and divergence-free confluence. The theoretical basis is provided in Section 2.5, in the discussion a new figure (Fig. 6) and Table 1 have been added. Another major improvement is that now each example situation is supplemented with a wind rose that summarizes wind conditions during the trajectories' travel time. This helps to understand the role of strong winds for the generation of FTLE ridges and to identify memory effects during subsequent calm conditions.

Please find below my detailed responses (in black) to the referees' criticism and suggestions (in blue).

Referee #1:

First, I would like to thank the referee for the time spent on the thorough review of this paper.

The manuscript investigates Lagrangian metrics such as FTLE and absolute/relative dispersion using model results in the German Bight, showing the high variability of LCSs and enhanced space variability of trajectories close to ridges. The results are interesting but certainly not novel or unexpected, and the paper lacks in my opinion of clear focus and motivation.

I am not aware of any analysis of LCSs in the German Bight area. Results of such studies will always be site specific, therefore I do not agree that the shapes of the LCSs obtained are trivial or were foreseeable. In my opinion, the sometimes very sharp ridges in the FTLE field come as a surprise.

Fig. 3b (former Fig. 2a) provides the example of a situation in which FTLE ridges are obviously constrained by the specific German Bight bathymetry. To make this more clear, bathymetry is now explicitly shown in the new Fig. 1.

The general objective of the study is to make observers aware of the fact that beyond well known random dispersion, there are coherent structures along which the separation of simulated backward trajectories is shaped more systematically. This objective is now stated more clearly in the second paragraph of the introduction (lines 18-27).

The author mentions several motivating applications, such as characterization or guidance for the observing system, but it is unclear how this would be carry out.

'Guidance for the observing system' suggests that the distribution of LCSs would be used for a characterization of certain locations. I revised the discussion to make clear that this is not the intention of this study. The idea is rather that detailed simulations addressing a specific period of observations might support an analysis of these data, helping to better distinguish between real changes in a system and just a shift in the observed water bodies' origins. This point is now better addressed in the discussion (e.g. lines 302-312) and also clearly stated in the introduction (lines 58-59).

I think the paper needs an extensive revision or even better a re-submission, where the motivations and the elements of novelty are clearly indicated and developed. Also, there are several specific points that need further clarification, as detailed in the following.

Main points

1) As mentioned above, there is an extensive literature showing the high sensitivity of particle trajectories to their seeding and the use of LCS to characterize it, so the results presented here are not new.

The reviewer is right, a large body of literature on LCSs already exists. Nevertheless, the results from this study are new in the sense that so far no such study has been conducted for the German Bight area. The present study describes LCSs that may affect a proper interpretation of German Bight monitoring data. I do definitely not agree that the shapes of these large LCSs, partly constrained by the bathymetry in this specific region, are trivial or were foreseeable.

I think the paper needs a novel angle, and a more specific motivation to make the present results new and interesting. I thought that the angle mentioned by the author in the Introduction regarding the characterization of an observing system composed of fixed points is interesting. But it needs more focus and more practical applications.

With regard to the characterization of an observation system, please see my previous response. This study suggests using hydrodynamic simulations in support of data analysis. I hope that this point is now explained more clearly in both introduction (e.g. lines 55-63) and discussion (e.g. lines 307-312). Optimizing the site locations of an observation network would be another interesting task, challenging due to the high variability of LCSs.

The analysis of any specific data set would go far beyond the scope of the present paper. Such a study would have to address many data specific problems, most of them unrelated to marine transports.

For instance, could the results be used to quantify uncertainty at the stations using as proxies the distance from ridges? Or could they be used to indicate areas of influence of the stations, in terms of LCS patterns of dispersion properties? Investigating this type of questions would be very useful from the application point of view and could lead to new results.

I see the reviewer's point. However, in the light of high variability of the LCSs it seems difficult (if not impossible) to define such characteristic areas of influence for specific stations. The areas observations are representative for at a given time much depend on the recent history of winds. Instead of a static characterization of monitoring stations, a model based representation of wind driven, variable hydrodynamic conditions is supposed to be useful to support the interpretation of monitoring data.

In the revised manuscript, all figures now provide a summary of wind conditions over the last 10 days. An interesting result is that pronounced LCSs initiated by strong winds may survive and even sharpen under subsequent calm wind conditions.

2) Since the results are based on the BSH model outputs, it is very important that the model set up and its validation are adequately described, significantly improving Section 2.2. This is especially relevant since the model based results are envisioned to be used in support of the observing system, possibly also in real time.

The description in Section 2.2 has been extended, it now provides more detailed information about the model.

BSHcmod has been in operational use for about two decades, with revisions being applied where needed and possible. Practical experiences were gained also in the context of search and rescue. The operational model was chosen exactly for the reason the referee mentions: BSHcmod results are reliably available on an everyday basis. Of course, higher spatial resolution would be useful, but for the moment such more expensive simulations are not available operationally.

More specifically, has Lagrangian validation ever been performed using drifters?

Yes, the same model setup plus a small leeway (0.6% of 10m winds) has successfully been used for the evaluation of drifter experiments. Corresponsing references (Callies et al. 2017b, 2019) have been added at the end of Section 2.2 (line 119).

It is also important to be up front regarding model limitations. For instance, given the 1 km resolution we can expect that coastal submesoscale is only partially resolved at best.

Thanks, this is indeed an important aspect, now addressed in the discussion (lines 326-330): "*Computationally more demanding FTLE analyses on a finer grid would have enabled identification of structures even smaller than resolution of the Eulerian hydrodynamic model arising, however, from tracer simulations over longer distances (Huhn et al., 2012). This shows that a classification of kinematic LCSs in terms of mesoscale or submesoscale features and processes may be difficult. Longer integration periods underlying the LCS analysis may filter more short-term features (Serra and Haller, 2016).*"

Also, if the model is hydrostatic, we cannot expect that near surface divergence processes are correctly described.

The model is hydrostatic, but in the absence of significant bathymetric variations implying large vertical velocities, this shouldn't be a major restriction. The present analysis focusses on open sea conditions, where currents proved to provide a reliable basis for drift simulations (Callies et al., 2017b, 2019).

3) The description of the used techniques in Section 2.4- 2.5 should be improved, indicating also possible limitations and clarifying definitions. For instance, is the definition of FTLE in eq (3) valid in the case of 2-dimensional flows (as the text at line 120 seems to imply)?

The theoretical concepts are now explained giving more details.

Section 2.4: The formalism deals with a purely kinematic description of flow motions (according to Haller (2015): "…*mimicking experimental flow visualization by tracers*") so that there are no limitations with regard to the underlying physical dynamics. The whole study deals with two-dimensional surface currents, which is now more obvious from the 2D deformation gradient being explicitly specified in Eq. (2).

Also, what is the difference between eq (4) and (5) for dilation? From the text (line 144-145), they seem to indicate the same thing, but it is unclear. Indeed the results in Fig.1b,c are quite different.

The reviewer is right that in the original paper the values in Fig. 1b and Fig.1c differed by a factor of 2. This factor was due to a mistake. Erroneously, dilation rate was calculated from the eigenvalues of the Cauchy-Green strain tensor rather than the singular values of the deformation tensor. As these eigenvalues equal the squared singular values (line 146 in the revised manuscript), this error resulted in a factor of two when the logarithm was taken.

Having corrected this error, the FDLD and dilation rate differ just due to different numerical discretization. While the dilation rate is calculated from the deformation gradient defined on the 1km FTLE grid, calculation of the Eulerian divergences uses auxiliary points at a 250m distance (see explanation in lines 184-186).

Because FDLE and dilation rate are so similar, I removed dilation rate from the figure (now Fig. 2) and replaced it by the temperature plot (formerly Fig. 4a, now Fig. 2c). The colour maps of both Fig. 2b and Fig. 2c have been modified to make related structures better visible.

4) In general, I think that the text commenting the results should be more realistic throughout the paper. For instance the comparison between LCS ridges, and the 2 forms of dilation in Fig.1 (lines 188-195 and lines 312-13) is very positive, while I fail to see a good comparison between the figures. I do not see a "striking similarity" between Fig.1a and 1b, where the main North-South ridge is absent.

I do not understand why the reviewer claims that he does not see the North-South ridge in Fig. 1b (now Fig. 2b). Actually, the FTLE ridge from Fig. 1a (now Fig. 2a) is reproduced very clearly in terms of a line of convergence. To improve visibility, the colour scale of Fig. 1b (now Fig. 2b) has been changed, replacing green by brown.

The author acknowledges the clear difference between Fig.1b and 1c, but I do not understand the point of the comparison, given that the model itself is not well suited for this diagnostics.

Why should the model not be suited? I guess that the reviewer refers to the factor of 2, which (see my above explanation) was an error that has been corrected. Nevertheless, Fig. 1c (now Fig. 2c) showing the dilation rate has been removed and replaced by the temperature distribution (formerly Fig. 4a, see above). In this way, all relevant information for the particular example (12 June 2015) is now concentrated in one figure and can more easily be assessed.

Also the comments on Fig.4 do not seem very grounded to me. In a case with very little gradients, except for the obvious coastal ones, as in Fig.4a and at some extent 4c, it is impossible to draw any meaningful conclusion.

I agree with the reviewer that the display of temperatures in Fig. 4 is problematic indeed, the challenge being the resolution of small temperature differences in the presence of large gradients towards the coast. Therefore, in the revise manuscript Fig. 4 was removed, keeping only Fig. 4a which now occurs as Fig. 2c. The colour scale was changed in such a way that now features corresponding with the FTLE field in Fig. 2a should be well recognizable.

5) Finally, and very importantly in my opinion, new diagnostics and metrics should be investigated, related to the observing system as mentioned in point 1). How can LCS be used to evaluate the observing system? How do LCSs vary on time? At which scales? Which proxies can we use to quantify these changes?

The first question (*How do LCSs vary on time?*) can hardly be answered based on a limited number of example situations shown in static figures of the manuscript. This is why a video was provided in the supplement, showing the evolution of FTLE fields over the full year 2016. This video also answers your second question (*At which scales?*). Features at very different scales can be distinguished, arising from an interaction between variable wind conditions and topographic constraints. In the revised manuscript, a summary of wind conditions during the last ten days has been added in each example and also in the supplementary video.

The referee's third question (*Which proxies can we use to quantify these changes?*) is the most difficult one. It is addressed in the last paragraph of the conclusions (lines 414-419). It would be very useful if LCS could be predicted as a simple function of atmospheric forcing. However, the LCSs depend on detailed past wind histories over extended periods of time, so that such relationship cannot easily be established. The display of recent wind history in each example illustrates the difficulty.

More specific points

The Introduction (and possibly also the title) should be re-written with more focus to-ward point 1) above.

The introduction has been revised. Parts of the discussion were moved to the introduction and shortened. A special paragraph was added (starting at line 55), focussing on the use of drift simulations for an improved interpretation of observational data. It is now clearly stated that the study does not aim at the optimization of a monitoring network.

Title: Slightly changed, now explicitly mentioning 'backward simulations'.

More in details, many phrases are unclear. Some examples are listed below - below line 20: "deficiencies of the underlying hydrodynamics...". Is this phrase indicating subgrid uncertainties or what? Deficiencies is certainly not the right word - around line 35.

I agree, the whole sentence has been deleted.

Discussion on local-versus nonlocal is not very precise. Indeed, local relative dispersion has been shown to be much faster at small scales and initial times than non local (Poje et al., 2014).

I agree that a definition of local relative dispersion was missing. This has been added.

It should also be clarified throughout the text whether the emphasis is on mesoscale or submesoscale dynamics

In the discussion it is now mentioned (lines 326-330) that the scale of ridges in the FTLE field can be smaller than grid resolution in the underlying field. Separation of tracer particles released very close together may evolve along drift paths that much exceed numerical grid resolution. This fact makes it very difficult to classify FTLE ridges in terms of either mesoscale or submesoscale dynamics. These ridges provide a merely kinematic description.

Section 2.1. It would be useful to mention from the beginning (lines 80-85) the geographical extension of the German Bight (lat/long are now mentioned at line 128 in Section 2.4), and clarify that the area is depicted in all the figures.

This information has been added in Section 2.1 (line 88).

Section 3, on results. The author shows 3 examples of LCS (Fig1, 2) for three different flow realizations and dates, 1 example of particle stats (Fig.3) for another realization, and finally SST (Fig.4) for a mix of realizations. It would be better to focus on 3 cases only, and compare LCS with particle stats, as well as SST.

Thanks for this suggestion. In the revised manuscript, the temperature field in the former Fig. 4a has been moved to Fig. 2 to be combined with the corresponding FTLE and FDLD fields shown in panels (a) and (b). The two other panels in former Fig. 4 have been removed, together with section 3.2 on surface temperatures. To complement the analysis, in Fig. 3 (now Fig. 4), a third panel has been added, showing the corresponding FTLE field (previously provided only in the supplement). The FTLE field in Fig. 3a is new, it nicely illustrates the origin of FTLE ridges in Fig. 3b about 8 days later, demonstrating the relevance of changing wind conditions.

Section 4, provides a broad discussion on FTLEs and their applications, but there is no clear connection with the present results. Indeed, most of the information are more suitable for the introduction, and in any case should be trimmed and focused on the paper's goals.

Section 4 (Discussion) has been thoroughly revised. Following the reviewer's advice, some paragraphs were moved from the discussion to the introduction. A new key issue in the discussion (raised by Rodrigo Duran) is the question to which extent FTLE ridges may be seen as lines of convergence rather than convergence-free confluence (lines 331-363). In the context of this discussion, Fig. 6 and Table 1 have been added. Another important aspect being addressed is the role of strong winds for the generation of FTLE ridges and memory effects during subsequent calm conditions (lines 313-325). The problem of dealing with possible simulation errors is addressed at lines 377-385.

Referee #2:

Once more I would like to thank the referee for the time spent on reviewing my paper.

Here, I try to provide my comments about this manuscript.

In general, the topic is appropriate for this journal, and I'm not aware of a similar study for the German Bight, which it makes it worth of publishing.

The used tools (PELETS in combination with surface current velocities obtained from BSH) are state-of-the-art and appropriate. Furthermore, the different quantitative measures to identify Lagrangian Coherent Structures (LCS) or objects with similar meaning are appropriate.

I must admit I'm not an expert for LCS and similar structures. However, reading the article and the theory included was very informative and helpful for me. Furthermore, using more the one quantitative measure is quite illustrative.

Major comments and questions:

- Tides are very important for this region. Might it be possible to give some idea how and if the FTLE structures will change when tides are not subtracted from the currents. So, what might happen when not using residual currents.

The referee is right that tidal currents are very important in the German Bight area. Accordingly, all drift simulations were conducted using the full hydrodynamic fields stored on a 15min basis. Looking at the trajectories in Fig. 2a (revised manuscript), for instance, tidal movements can easily be identified. In addition, these trajectories show the more long-term movements related to residual currents. However, tidal currents were never separated from the total flow fields when conducting numerical simulations.

- It was very illustrative to see how variable the FTLE structures are in time. What would happen, if one considers time averaged FTLE-fields. Perhaps averaged over a season or half a year. Will there be stationary or more robust FTLE ridges visible?

I wasn't able to find any meaningful non-trivial mean fields. The reason for this is that even when similar structures occur, the sharp FTLE ridges usually reside in different locations, so that any averaging results in extremely smooth, non-informative fields.

Due to the strong influence of wind forcing, a classification with regard to wind directions would probably be the most promising approach. In the revised manuscript, each figure now contains a wind rose that summarizes the recent history of wind conditions. Winds are never constant and the recent history of wind forcing differs for each specific date. This means that resulting FTLE structures can hardly be classified in a simple way. The only chance might be some more fuzzy classification of FTLE fields.

- Instead of showing the temperature and salinity fields, perhaps one could use the density distribution. Would this do any difference to the interpretation of the relationship between FTLE ridges and T/S fields.

Temperature and salinity were used just as simple indicators of the accumulated effects of surface divergences over some period of time. In the revised manuscript, the whole section on surface temperatures has been removed together with the temperature fields in Fig. 4. Just panel Fig. 4a was kept, being shown as Fig. 2c in the revised manuscript. The problem with finding a proper colour map could be solved after the exclusion of very nearshore regions. Salinity fields are no longer considered.

Surface temperatures have the advantage that they can be observed by remote sensing. Density fields would be interesting to look at, but unfortunately these fields were not available from the archived data underlying this study.

Some minor comments and questions:

- How do you calculate the residual currents from the BSH data? The current velocities of the BSH include tides, which are important for this region. Are the results dependend on the way these residual currents are calculated?

All simulations were based on hydrodynamic fields including tidal effects. See my comments above.

- Do you think it might be helpful to plot the absolute value of the gradient of temperature or salinity instead of the pure fields. So, one could see the sharp fronts more clearly.

This might indeed be one option to solve the problem of showing small local differences in the presence of substantial changes on a larger scale. I decided, however, to remove the former Fig. 4, keeping just Fig. 4a, which is now presented as Fig. 2c. The colour scale was changed in such a way that now the structure corresponding with the FTLE and FDLD fields in Fig. 2a and 2b can well be recognized.

Some direct comments to the text:

- Line 13: Gap at the end of sentence; (2017). Stanev . . .

Corrected

- Line 143: The year of the Huntley citation is missing

Corrected

- Equation (5): It is divergence of the 2d surface velocity field, isn't it?

Yes, thanks for the hint. I changed symbol ∇ into ∇_H and mentioned 'horizontal divergence' in the text.

- Figure 2a: The endpoint of the red drifter starting at 54 30' is missing?

Has been corrected, thanks for checking the plot so carefully!

- Perhaps one could also include the river Ems to the plots.

Location of river Ems is now indicated in all figures.

- Figure 4, caption: . . . times of the FTLE fields . . . ?

Has been changed.

- Additional figure S3: The unit of salinity should be different from PSU. Perhaps no unit or g/kg

Thanks for the advice, but the consideration of salinity was removed from the paper.