Reviewer #2

Dear reviewer,

We would like to show our sincere appreciation for your interest and deep analysis of our manuscript, entitled "A new Lagrangian based short term prediction methodology for HF radar currents". We would also like to thank the comments and suggestions you have proposed. The paper has been revised and carefully modified following them. They have undoubtedly helped to improve the quality of this manuscript. Our individualized response to your comments can be found below (in blue color).

You can find the new manuscript and the changes that we have done over it, in the final manuscript document that we will upload to the journal (both new and "track changes versions). Line references included in this document, are referred to the "track changes" version and they will be updated if any additional changes are requested by the editor before the final submission of the revised manuscript.

In the paper by Solabarrieta et al. a new short-term prediction method for surface marine transport is presented. The method is based on Lagrangian "analogues" calculated using velocity data from high-frequency coastal radars located in two different regions: the Bay of Biscay and the Red Sea. New-method errors and predictions are compared with those based on persistence. The performance is comparable to other methods reported in previous literature (e.g. Solabarrieta et al, 2016) as mean separation distances are shown to be similar. The new method can be more easily implemented operationally than the others due to its computational cost, which is allegedly low.

A process of major revisions is suggested to address the following concerns:

1) L123: "well demonstrated results". Please explain why OMA was chosen and quantify the OMA skills providing values and the advantages to other methodologies like DINEOF or SOM.

This paper is focused on the forecast of the surface currents and not on the gap filling techniques. This is why no more quantification values were included in the text. But we have now modified the text, to indicate that one of the main reasons to use OMAs is that it's well-functioning is demonstrated (Kaplan and Lekien, 2007, Hernández-Carrasco et al., 2018) but also because there are available codes in the HFR_progs package, that allow us not only to generate real time gap-filled fields but also to generate trajectories for our analysis (lines 128-130).

2) L138-146: not clear paragraph here and the concept may be missed. Are the authors trying to justify the choice of a Lagrangian vs Eulerian approach for the analogues? If so, wouldn't be enough to say that Lagrangian trajectories are direct measurements of transport of substances at sea? And also that they are more dependent on resolution as they are more keen on accumulating errors being integrals of the velocity fields?

We agree in this regard with referee. Accordingly, this paragraph has been rewritten in the manuscript (lines 148-153) as follows:

Lagrangian computations have proven to be robust in identifying dynamical flow structures and they are direct measurements of transport of substances at sea. Lagrangian diagnostic will capture dynamical features present in the flow that are not readily apparent in pure velocity. At this point we remark that they are more dependent on resolution since they are more keen on accumulating errors being integrals of the velocity fields

3) L151: uniqueness and originality of the work. Authors should clearly state whether or not this is the first application of the method of analogues in the ocean.

It has been clarified in the text that apart from the two-fold approach of the presented method, analogue finding to generate Short Term forecast has still not been applied to HF Radar ocean surface velocity fields (lines 181-182)

4) L156: numbers expressing a quantification of the computational costs for the different methods should be provided here. How long does it take to run this new method wrt the one in Solabarrieta et al (2016)? What about wrt other methods?

As it has been included in the text, this forecast can be done in seconds or few minutes (depending on the historical dataset size) (lines 184-185).

One of the main differences with the rest of the STP methods, is that this new method is not only fast but it can also modify (increase) the historical dataset (catalogue) with the last information as soon as new data are provided, without any requirement to re-analyze the whole catalogue. This clarification has been included later on in the text (lines 635-637)

5) L162-177: how do resolutions in the two regions compare with the Rossby radii? Are spatial resolutions of the HF radars fine enough to capture the marked seasonal variability of the mesoscale features in the whole year for both regions? Please provide number and quantify.

The Rossby first radius of deformation in the red Sea is around 30 km (Zhai and Bower, 2013) and between 20 and 50 km in the BoB (~ 3-8 km over the shelf (Charria et al., 2017)). Since the spatial resolutions of both systems are 3 and 5 km respectively they resolve adequately the mesoscale in both regions.

6) L209: a conceptual question that should be addressed. It is my understanding that the OMA method is based on finding the best combination of geometrical modes in a specific region able to maximize the fit with the observations at a specific time. In a way, isn't the combination and gap-filling technique already based on "analogues" modes? Isn't this procedure already creating analogue situations from a dynamical perspective, introducing a bias when epsilon_ANL is calculated? I guess that the other way to pose the same question is: how sensitive are results to the use of OMA? How much do they change if a simple linear interpolation technique is used instead of OMA?

As pointed by the reviewer, the OMA method finds the best combination of geometrical modes in a specific region to maximize the fitting to the radar surface velocity observations. But it is not “based” in temporal analogues as this fitting is applied independently to each specific hour field, not related to the previous and later fields. Indeed, the OMA method is applied to radial velocities and it can be applied to spatial gaps (due to range fails for example) where linear interpolation technique could not be applied.

7) L213: clearly say here that the "most similar" concept will be defined later in the paper.

Included in the text (lines 245)

8) L212-218 and L220-226: more concepts are repeated in both paragraphs. Please combine them and shorten accordingly

The text has been reorganized and double concepts have been removed (lines 243-253) to make it clearer for the reader.

9) L228-230: where is this shown? I have the impression that a section has been completely cut off from the paper. This is also related to point 23 below

It has been clarified in the text that those results were done during the analysis for this work but that those results are not shown in this paper.

We want to maintain it there, as the reader may think that the direct application of the methodology to the Eulerian fields could be a better approach but we saw that it is not.

10) L237: is conceptually correct to use the whole period as a test period and a Lagrangian catalogue at the same time for the Red Sea? How do results change if the first year is used as catalogue and the second year as test period?

In the Red Sea case, it was indicated in line 236, that the data availability is from July 2017 to October 2018 (2 years). This is just 1 year and 4 months and it has been corrected in the text (Line 264 in the “track control” version).

Ideally, it would be better to use past data as a training period, like the Lagrangian catalogue used for the Bay of Biscay data (because this is the situation that we will have once this method is applied in real time). But taking in account that we know (from previously published works; not HF Radar data) that there is a clear seasonality in the Red Sea study area, and the HF Radar data availability was short, we have used the whole year as a training and test period, but we have removed the previous 2.5 days and the next 2.5 days to avoid the overlapping.

11) L244: I would suggest swapping Fig.2 and Fig.3 positions as this latter is introduced in the text before.

Figures have been swapped and the references corrected accordingly in the text.

12) L269: please remove not needed.

It has been removed and the magnitude of δ_t has been indicated in line 309

13) L326-330 and Fig.4: contradictions and big confusion here. Not easy to understand whether or not black dots show periods when epsilon_STP is either larger or smaller than epsilon_PRS. My guess is that dots are when errors in the predictions are larger than in the persistence. Please double-check and rephrase the whole paragraph

Your guess is correct. Black dots are plotted for the periods when ε_{PRS} is lower than the ε_{STP}. It has been corrected in the text (line 421) and it is consistent now.

14) L331: what is the time-scale of the persistence of these currents during winter months?

Rubio et al. (2018, 2019) and Solabarrieta et al. (2014) show that currents during winter months show an eastward flow than can least for several weeks during winter and that these currents are higher than eastward flow present during summer season.

It has been completed in the manuscript, in the first paragraph of section 3.1.

15) L343: indicatES

Corrected in the text (line 445)

16) L349-357, Fig.6 and throughout the manuscript: please use the already introduced notation for the mean separation distance like, for example, \delta^STP_6h (\delta^PRS_6h) and not STP_dist (PRS_dist).

δ_{STP} or δ_{STP} has been used for the previos STP_{dist} and PRS_{dist}.

It has been modified throughout the whole manuscript and the figures.

17) L356: not sure what "especially after 12 hours mean"? Maximum values are at 36h. Do the authors want to say that larger values are reached and remain almost constant after 24h? Please rephrase.

The idea that authors want to show with the combination of figure 6 and table 2 is that there is no correlation between ε_{ANL} (used to find the analogue in the catalogue) and PRS_{dist} (distance between real and PRS simulated trajectories); while there is higher correlation between ε_{ANL} and $STP_{dist,}$ specially after 12 hours of simulation ($R2(\varepsilon_{ANL}$ vs $STP_{dist})$ increases rapidly after 12 hours, from 0.37 to 0.54) as indicated in table 2)

It has been clarified in the text (lines 460-461)

18) L357: it should be also mentioned that at t=6h PRS is always better that STP (Fig.6). However we have a problem here: at t=6h R^2 for PRS is is lower than for STP

We have mentioned in the text that PRS at 6 hours is always better than STP (lines 454)

Regarding the correlation, there is no any problem. From our understanding, it means that the ε_{ANL} is correlated with the STP error (bigger $\varepsilon_{ANL\ will}$ have bigger ε_{STP} or STP_{dist}) but it is no correlated with PRS error, even when persistence is better than the STP. The point here is that during the first 6 hours, it is better to use persistency than the STP. But it is worth it to use STP for longer time forecasts (for example, to predict where a possible oil spill could move).

19) L364: isn't this choice unfair wrt persistence? Shouldn't we consider all of them for a fair comparison?

With this comparison, we want to show the capabilities of the methodology for the times when we consider that the STP will be better than the Persistence ($\varepsilon_{ANL} < 853km^{2,}$ for BoB case). When $\varepsilon_{ANL} > 853km^2$, we suggest to use persistent currents

ε_{ANL} can be considered as a real-time skill-score metric for the L-STP. In fact, this value has been investigated and presented to be able to tell to the final user if our forecast is good enough or not.

20) L371: correct, it should be indeed added that persistence during the first hours is actually slightly better

It has been included in the text (line 510)

21) L380-381: why does the mean drift follow more the persistence curve in the Red Sea case?

It is probably related to temporal size of the HF Radar data availability in the Red Sea case. Longer the dataset, better results will be obtained using the presented L-STP method.

22) L390: the advantage is not clear as this is the difference between the two, does not necessarily mean that one is better than the other. Please modify Figs.9 and 10 as suggested in point 37 below

Figures 9 and 10 have been converted to figures 10 and 11, as we have included a new figure.

This point has been replied in point 37 below.

23) L404-407: what does this mean? Only Lagrangian analogues are shown in the manuscript. Has a section been cut off from the paper? This is also related to point 9 above.

As in the point 9 above, it has been again clarified in the text that those results were done during the analysis for this work but that those results are not shown in this paper.

24) L417: contradiction with L327-328

Corrected in the text (line 421)

25) L423: "first and only the first". Not really but please quantify as it looks that for BoB is at least during the first 6h and for the Red Sea at least for the first 15h!

Quantified in the text (line 578-579)

26) L429: not sure about this value as it was reported 853 km2 before (e.g. at L342 and L364)

It was a typo mistake and it has been corrected in the text (line 585)

27) L441: Fig.7 not Fig.4, correct?

Figure 7, correct. It has been corrected in the text (line 597)

28) L447-453: these lines belong more to the introduction. They are also qualitative while differences and comparisons between methods should really be quantified.

They are qualitative but we would prefer to maintain them there, as it is a comparison between both methodologies.

29) L463-472 and in general for the whole section: discussion is poor. Why aren't HF radars able to capture currents if they are persistent? I would expect radars not to be able to resolve highly-variable small-scale structures, not persistent features! Not (0.07 vs 0.19). How is this possible? getting (or buying) the idea that something persistent cannot be seen by analogues. A better dynamical insight is needed and expected in the discussion of the results.

Since temporal resolution of HF-Radars is hourly, they capture well all scales of interest above hours. This includes persistent currents. The comparison in the discussion is made between the STP system based on radars in front of a prediction made with persistence (in an abuse of language since persistence here means that the prediction for the next hour is simply the velocity measured in the last observation).

There is a reason why persistence is better during persistent periods than STP and it is not that STP does not capture persistence. It is mainly because in both cases (BoB and the Red Sea) the persistent periods show high surface velocities and the persistent structures take place in similar longitude and latitude but not exactly the same positions. A small separation distance between real and analogue fields generate high separation distances between real and simulated trajectories. But it does not happen when the real current field is used as persistent current, as it is located exactly in the place where the persistent structure is located in the study time and it will remain there at least during the first few hours.

This paragraph has been rewritted/completed in the manuscript in order to clarify and provide more dynamical insight of the presented results.

30) Fig.1: can we have GDOP maps in the two regions? Can they help discussion? Asking for more reasons: a) obtained ranges look large compared to the radar system positions and distances between them; b) it would be important to visualize in which areas OMA operations are more to be carried out; c) it would be nice to compare/discuss GDOP maps wrt to the error distributions of the new Figs.9 and 10 (see point 37 below)

Figures 9 and 10 have been converted to figures 10 and 11, as we have included a new figure.

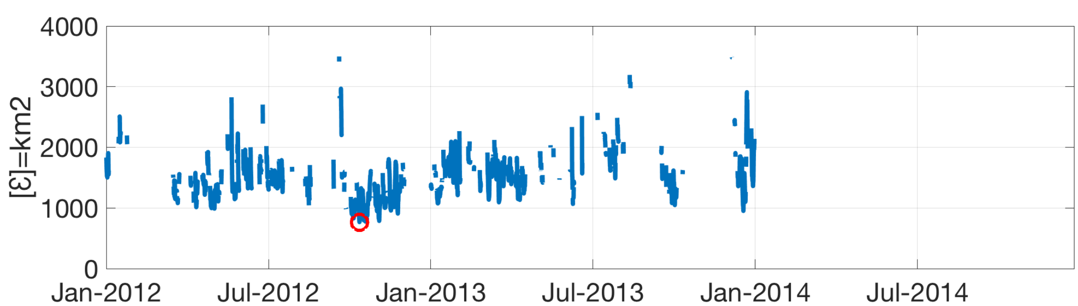
This point has been replied in point 37 below.

31) Fig.2: resolution is really terrible, please increase it. Line should be thicker as in Figs. 7 and 8. Why are there gaps in the blue line? Really confused by the fact that caption is reporting Nov 17 2015 instead of April 13-15, 2015 as in Fig.3.

We used different examples during the writing of the manuscript and we finally did not change the date of the caption. But it is corrected now with the correct date: April 15, 2015.

There are gaps in the blue line because the methodology doesn't calculate the errors when the δ_cg > 10 km, as indicated in the text and in the caption of this figure.

We have tried to make the line thicker but we loss the details of the times when the error is not calculated because of the δ_cg > 10 km condition, as you can see in the next figure:



Regarding the resolution of the figure, we hope that it is just a problem with the revision version of the manuscript. We will submit a high resolution independent file to the journal for the final publication.

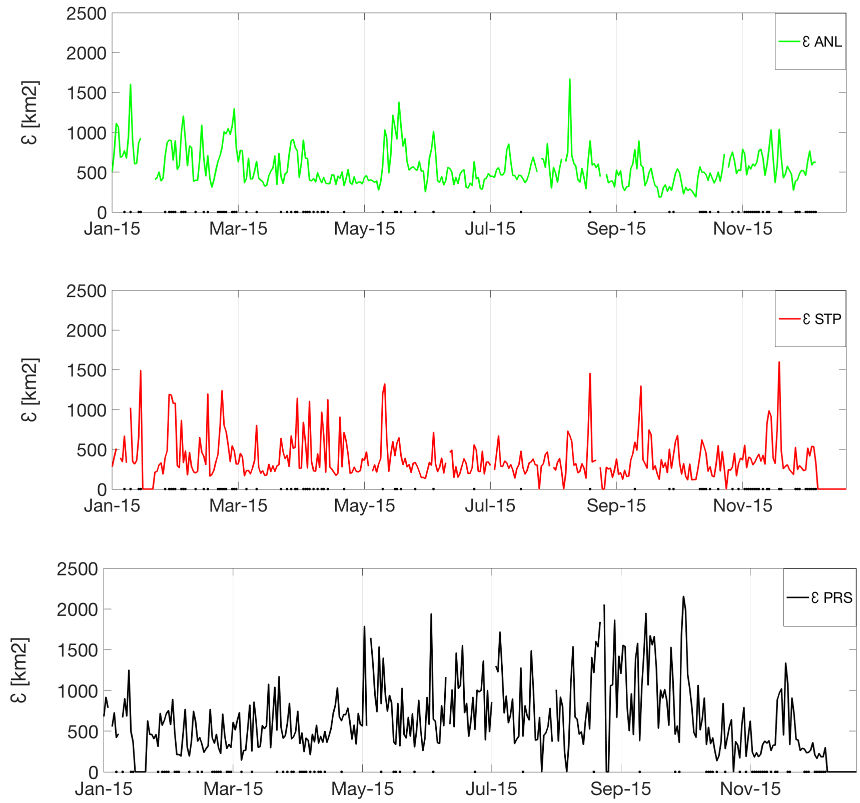
32) Fig.3: why is this time chosen? Is this a good or bad example?

This figure has become figure 2, following your advice.

It has been selected as an example of the good functioning of the methodology. There are better and worse examples and we wanted to show something intermediate.

33) Fig.4: resolution is really terrible, please increase it. Lines should be thicker as in Figs. 7 and 8. I would suggest to put them in three different panels as they mostly overlap. Double-check figure and text for black dots meaning.

We have modified the figure increasing the thickness of the lines. We want to maintain the three lines in just one panel to be able to see the comparison of the values. It is too complicated if we separate it into 3 panels, as you can see in the next plots:



Regarding the resolution, we will proceed in the same way as with figure 2, to submit the figures with high resolution.

34) Figs.5 and 6: resolution is really terrible, please increase it. Lines should be thicker as in Figs. 7 and 8.

Modified as requested.

35) Figs.7 and 8: rearrange x-axis labels to have 6-h intervals ending at 48h.

The figure has been corrected.

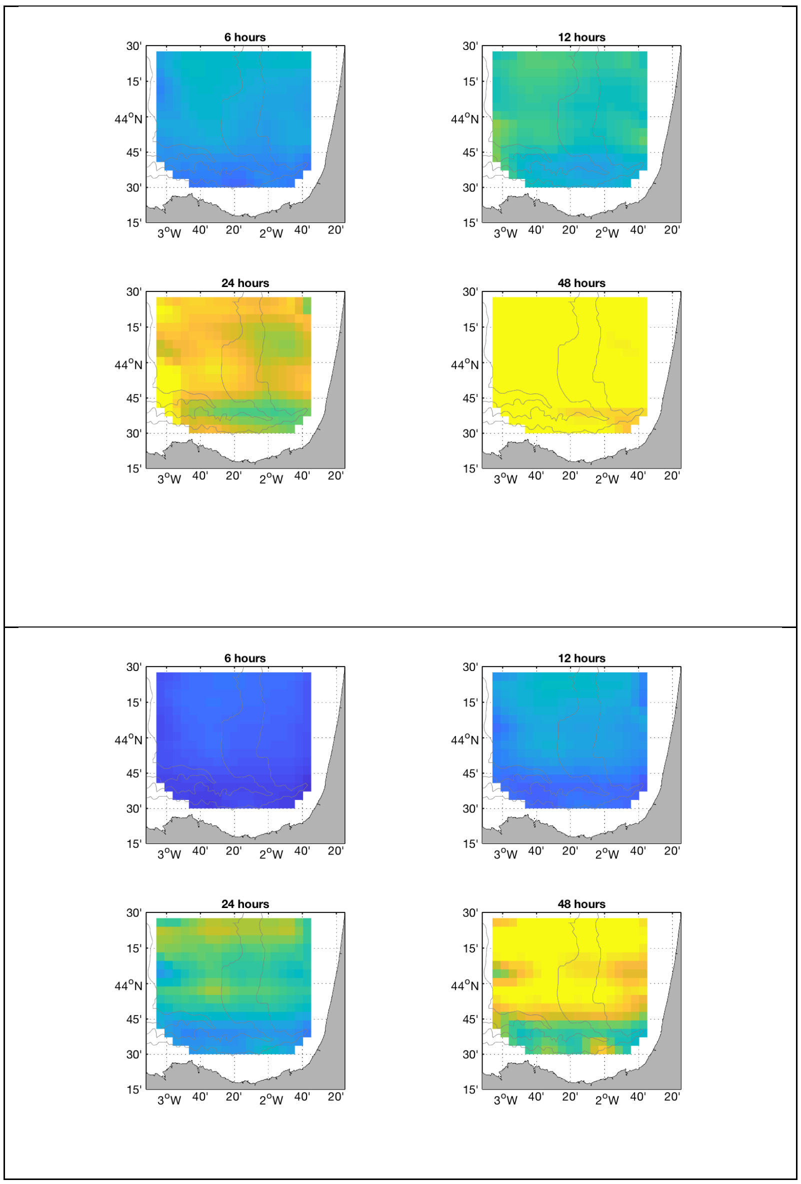
36) Fig.8 caption: remove (UP)

It has been corrected in the caption.

37) Figs.9 and 10: both figures need improvements to show the errors and not only their differences. Suggestion is to have a total of 12 panels in each region and show for each time three panels, one with \delta^STP, the second with \delta^PRS and the third one with their difference.

We generated those figures before the submission of the paper and we decided to show just the difference between δ_PRS and δ_STP, as the purpose of this figures is to show the advantage (when exists) of the L-STP methodology vs the usage of persistent fields. But it may help to the reader to have them, so we could include the δ_PRS and δ_STP panels for each study area, as you suggest, as supplementary material for the paper.

As an example, we show here the results of the Bay of Biscay System:



38) Figs.9 and 10: put labels indicating times either on top of each panel or in the right bottom corners, on land

Times have been included on top of each panel.

REFERENCES:

Charria G., Theetten S., Vandermeirsch F., Yelekçi Ö., and Audiffren N. Interannual evolution of (sub)mesoscale dynamics in the Bay of Biscay. Ocean Sci., 13, 777–797, https://doi.org/10.5194/os-13-777-2017, 2017

Hernández-Carrasco, I., Solabarrieta, L., Rubio, A., Esnaola, G., Reyes, E., and Orfila, A.: Impact of HF radar current gap-filling methodologies on the Lagrangian assessment of coastal dynamics, Ocean Sci., 14, 827-847, https://doi.org/10.5194/os-14-827-2018, 2018.

Rubio A., Caballero A., Orfila A., Hernández-Carrasco I., Ferrer L., González M., Solabarrieta L., Mader J. Eddy-induced cross-shelf export of high Chl-a coastal waters in the SE Bay of Biscay. Remote Sensing of Environment 205, pp. 290–304, 2018.

Rubio, A., Manso-Narvarte, I., Caballero, A., Corgnati, L., Mantovani, C., Reyes, E., Griffa, A., and Mader, J.: The seasonal intensification of the slope Iberian Poleward Current, in: Copernicus Marine Service Ocean State Report, J. Oper. Oceanogr., Issue 3, 13-18, doi: 10.1080/1755876X.2019.1633075, 2019.

Solabarrieta, L., Rubio, A., Castanedo, S., Medina, R., Charria, G., Hernández, C.: Surface water circulation patterns in the southeastern Bay of Biscay: new evidences from HF radar data. Cont Shelf Res 74:60–76 doi:10.1016/j.csr.2013.11.022, 2014.

Zhai P. and Bower A. The response of the Red Sea to a strong wind jet near the Tokar Gap in summer. Journal of Geophysical Research: Oceans, Vol. 118, 422–434 https://doi.org/10.1029/2012JC008444