REVIEW 1 :

Hi Phoebe,

You're not going to like what comes next. Sorry.

We thank the reviewer for their honest and constructive review and believe it has helped to notably strengthen this paper.

Major comments

I. The objective of this manuscript is unclear

Although both the abstract and the introduction are very long, they are also unfocused, so it is unclear whether the aim of this article is to:

- *combine satellite, reanalysis, and in-situ data to study the plume;*
- *demonstrate that satellite data can be used to look at sea surface salinity in the region;*
- *or determine which reanalysis product is most adapted for this study.*

The main objective of the paper: "combining satellite, reanalysis, and in-situ data to study the plume" has been highlighted more clearly at the end of the introduction. The introduction has also been shortened to help clarify the focus of the paper. In particular, the initial introduction section focusing on Arctic-wide processes, has been removed, and given the de-emphasis on satellite data, the section providing a broad overview of satellite SSS data has also been condensed.

The abstract has also been updated to reflect these changes and to better emphasize the main objective. The key points have also been re-arranged and altered to reflect this.

These different objectives would then result in different structures of the manuscript, with the majority of the product comparisons and validations going to the appendix. It would also affect the time period considered. The choice of in-situ data would then be affected as well; see my next point.

The components of the manuscript validating satellite data used with in-situ data have been moved to the appendix.

II. The choice of data, especially their resolution, was surprising

The method section did not clarify much, so I am not sure which time period you worked with. That is, you mention that you use UDASH, but that stopped in 2015 and would not really help with SMAP. The manuscript needs an overarching table that says for all types of products over which time period and at which spatio-temporal resolution they are available (not the resolution at which you use them; their native one).

An overarching table has now been added as requested for the satellite products to clarify all products used, their spatial resolutions and time-periods (new Table 1). All satellite products are now used/re-gridded onto a 0.25 degree grid for clarity and to aid comparison.

Coupled with the fact that the objective is unclear, I won't be able to give you a clear direction. But if you want to validate SSS, I would have looked for underway CTD data rather than CTD casts. It will be at approx. 10 m depth, but on most vessels (and especially so on Oden, i.e. for the SWERUS data) the upper 10 m of the CTD casts cannot be used anyway. If you want the upper water column, there might be ITPs nearby, and there should be at least one mooring, but I'm not sure of their time coverage.

The section detailing in-situ data used has been updated to more accurately reflect the range of in-situ data used for validation (not just CTD casts). The SWERUS (Oden) cruise CTD data is no longer used. NABOS underway data is used instead of CTD casts from the 2018 cruise. Table A1 (previously Table 1) has been updated to reflect these changes.

ITPs from the Mosaic cruise (from 2020) were considered for use but were >84N (and notably away from the shallow Laptev Shelf and the plume). The NABOS moorings (2013-2015, 2018-2021) were also considered for validation but are too far off the shelf and too deep (minimum depth 30m) for validation of the main plume.

The other thing that really surprised me is that you want to investigate a plume dynamics, have a 3-day product available… and downgrade it to monthly resolution. And then regularly show results in the manuscript that you explain with the poor resolution of your monthly product. Use the 3-day version.

And regardless of what you do, you need to say why you do so. Maybe you had a perfectly valid reason for using CTD casts and downgrading everything to one month, but you did not write it so the reader cannot know.

The shift to one main objective and the increased focus on reanalysis data will have helped to clarify the choice of data. In particular, much of the text describing comparison with in-situ data has been moved to the appendix, and therefore is a less central focus of the paper.

The reasoning behind the choice to use monthly data has also been clarified in the methods section: "Higher temporal resolution satellite products were considered for analysis but comparison with in-situ data suggested they did not notably improve correlations with in-situ data. Therefore, these results did not justify their use over monthly products."

III. Causation is not shown

This is my main issue with your manuscript. You do not demonstrate causation. You produce two composites, and declare that the variable you composited against explains the differences.

Let's start with the definition of the composites. I agree that on Figure 2, the circulation is different. However on Figure 4, the uncertainty is so large that some of the strongest years could in fact have a value with either sign. See for example 2012 and 2019. As rotation is involved, a metric based on the curl of the wind, or simply on the sea level pressure, may be more effective and robust.

To demonstrate causation, a lagged correlation analysis has now been run between ERA5 eastward turbulent surface stress in individual months and GLORYS SSS in September. The results of this analysis now establish that eastward wind stress is the dominant driver of September SSS. The correlation is strongest in June, July and August so we now use a 3-month metric (June-August) (rather than the 4-month metric previously being used).

Note: In Figure 7 (previously Figure 4), the grey overlay is the maximum and minimum wind stress over the 4 month period (not the uncertainty). As eastward wind stress in September is highly variable, the range in surface stress over this three month period is still intentionally left on Figure 7 to help interpretation of individual years.

Anyway, the outcome of the composites is that the SSS looks different. But so do the sea ice and the SST, which could both explain the SSS pattern, and even be responsible for the wind differences. Or maybe wind and SSS are both the result of another variable that is not included in your analysis.

What you are really showing is that the hypothesis that wind drives SSS is not incompatible with the observations. But as your analyses currently are, you do not demonstrate the causality. One option is therefore to just rephrase everything, removing all mentions of "the wind drives" and saying what I just wrote. But that's rather underwhelming a result.

Instead, you could utilise the 3-day product in its full 3-day glory. For each point, or for the overall region, do a lagged (temporal) correlation analysis of the relationship between wind (available at daily resolution; downgrade to 3-day) and the 3-day SSS from BEC. Then see which variable drives the other, based on these values. I personally would push the analysis and perform the same calculation also with the SST and the sea ice, and have an overall result matrix that shows for each pair of variables for which lag the correlation is maximal and what that correlation value is.

A lagged-temporal correlation analysis has now been conducted of both wind stress and runoff with SSS using monthly rather than 3-daily data (see previous response on the 3 day product). A section has been added to the paper to discuss the results of this correlation analysis (depicted in Figures 4, 5 and 6). This shows that eastward turbulent surface stress in June-August is particularly strongly correlated with September SSS.

The same temporal correlation analysis was also conducted of both wind stress and runoff with SST and SIC. The results of this analysis have also been included in the paper. The differing spatial patterns highlighted by the correlation maps highlight the different processes controlling SSS and SST/SIC.

Whilst sea ice formation does play a strong role in controlling salinity in winter, sea ice melt (and precipitation) contribute several orders of magnitude too little freshwater to be dominant drivers of SSS variability in summer. The minimal direct role of sea ice melt has been clarified earlier in the text with reference to prior literature in the introduction.

Less major comments, in order of appearance

Salinity unit: You meant "psu"or really"pss"? Oceanographers use the absolute salinity in g/kg now.

We tried altering the paper to not use any salinity unit but found this to be unclear. Whilst the difference between absolute and practical salinity is negligible, especially compared to the uncertainty in salinity in this region, both satellite and GLORYS12V1 SSS are provided as practical salinities. Therefore, we choose to use units for practical salinity rather than absolute salinity. We avoided the use of "psu" following the guidelines of the TEOS manual "Note that Practical Salinity is a unit-less quantity. Though sometimes convenient, it is technically incorrect to quote Practical Salinity in "psu"; rather it should be quoted as a certain Practical Salinity "on the Practical Salinity Scale PSS-78"." (Intergovernmental Oceanographic Commission et al., 2015).

Subsection 2.1.1: not detailed enough. Either there or in the introduction, you need to be more specific about which satellite measures at which band, has which footprint, repeat time, etc. Something similar to the first paragraph of section 2.1 of these people: https://tc.copernicus.org/articles/12/921/2018/tc-12-921-2018.pdf , but for all sensors (at least SMOS and SMAP, and then explanations about how the different products combine them).

As noted previously, our main objective is to study the plume rather than validate the satellite SSS data. Having made this clear, we do not feel it is necessary to include specific details about the satellite measurements.

Subsection 2.2.1: also not detailed enough; what do you mean by "correlation" and "RMSD"? I assume that you took all points available, regardless of location and time, and basically did a regression? Given that the plume is both time and space dependent, as shown on your figure, I would recommend you verify the temporal and spatial accuracy separately. You may need more points for this, agreed, but see Major comment II.

The details of the methodology are now included in the appendix. By "correlation" and "RMSD" we mean the Pearson correlation coefficient and root mean square difference between all in situ data that has a corresponding satellite SSS value across the entire region and time period. The reviewer is correct that we took all points available irrespective of space and time. This has been more clearly explained within the appropriate section of the methodology in the appendix.

Section 3.1: You do not show the RMSD. See my previous comment anyway, but that could be added to the table -which could be shortened once the manuscript is more focused, see Major comment I.

RMSD values have now been included in Table A2 in the appendix.

Sea ice: Line 377 onwards you give statistics of the sea ice area, without specifying over which region. Overall, and based on the figures shown in the manuscript, the area does not seem to matter as much as the southern / eastern extent. I would rather use such extent, if doing the correlations suggested above. If not, then do not even quantify it; your maps are very clear.

The sea ice area metric is no longer included in the paper, as it was deemed to no longer be needed with the inclusion of the sea ice concentration correlation plots.

Runoff: Line 405, the discussion starts with an analysis of the correlation between runoff and SSS. It is not specified, but I assume that the runoff values are published elsewhere? If so, reproduce them here, and do a proper (lagged) correlation analysis.

Interannual variability in GRO runoff is now considered and correlated to variability in the spatial pattern of GLORYS12V1 SSS, SST and SIC. A section has been added to the paper to discuss the results of this correlation analysis (depicted in Figures 4, 5 and 6).

Arctic Oscillation: Same comment, no information about where the AOI comes from and the correlation analysis is not shown.

The source of Arctic Oscillation data has been clarified in the methods section.

REVIEW 2 *:*

Dear coauthors,

Your paper offers valuable insights into the drivers of Laptev Sea dynamics and interannual variability in salinity and temperature. It provides evidence that the salinity and temperature signatures agree in different reanalyses and satellite products under different wind regimes. Addressing the major comments proposed will improve the clarity and robustness of the research, enhancing its contribution to the scientific community.

We thank you for your positive review and agree that implementing the suggested changes will strengthen the paper.

Major Comments:

The focus of the paper:

The stated objective of the paper is to determine the drivers of the interannual variability of the Laptev Sea dynamics. However, a significant portion of the paper is devoted to the validation/intercomparison of satellite Sea Surface Salinity (SSS) products. While the effort to validate and compare these products is commendable, it appears to dominate the narrative, diverting attention from the primary objective of identifying the drivers of Laptev Sea dynamics.

To address this concern, I see two options:

a) Dividing the current study into two separate papers, each with a distinct focus and stating clearly the objectives (Validation and Intercomparison of Satellite SSS Products / Drivers of Laptev Sea Interannual Variability in Salinity and Temperature).

b) Focus on the primary objective during the narrative and if you feel that the validation/intercomparison of satellite SSS products is an essential part of the methodology, it would be beneficial to include at least one example figure showcasing the different products. Additionally, providing information on the p-value of your correlations, bias, and spectral analysis to assess the effective resolution of the products will improve the methodological rigor and transparency of the study.

We agree that the validation distracts from the main objective. The abstract and introduction have been shortened to help more clearly highlight the focus of the paper. The section validating in-situ data has been moved to the appendix to help better emphasize the main objective.

Methods:

1. *In subsection 2.2.2, it is not clear why the analysis is not performed with all four satellite products. To ensure the robustness of the study, it would be beneficial to explain the reasons for the exclusion, if any, of certain products in the analysis.*

All four products have now been included in Figure 2. As the validation section has been moved out of the results section and into the appendix, discussion of correlation and RMSD values now occurs alongside this figure which helps to clarify the choice to exclude BEC and RSS.

2. *The temporal resolution of satellite SSS products is a critical factor when studying the dynamics of a region like the Laptev Sea, where rapid changes can occur over short time scales. If you choose not to use 3-day or available 8-day satellite Sea Surface Salinity (SSS) products, providing a clear and well-justified argument for this decision is crucial.*

The reasoning behind the choice of data has been clarified in the methods section: "Higher temporal resolution satellite products were considered for analysis but comparison with in-situ data suggested they did not notably improve correlations with in-situ data. Therefore, these results did not justify their use over monthly products.

3. *The paper mentions a validation/intercomparison of satellite SSS products. To strengthen this aspect, I suggest including an example figure showing the different products for comparison. Additionally, information on the p-values of your correlations, bias, and spectral analysis to determine the effective resolution of the different products should be included.*

All four products have now been included in Figure 2. P values for each correlation have not been included in Table A1 as they are all << 0.01 but the title has been updated to reflect this.

4. *The decision to use the median of the products for analysis should be justified. It might be more appropriate to use the product that best aligns with in-situ information, has a higher spatiotemporal resolution, or realistically agrees with the expected dynamics of the area. If the median approach is retained, the reasoning behind this choice should be elaborated.*

The median product is no longer used in analysis as we feel it is more useful to include each of the four products to more clearly illustrate their similarities/differences.

Results:

1. *The results section lacks concrete analysis and tangible results to support the discussed relationships with the Arctic Oscillation Index, and river runoff.*

Lagged correlation analysis has been conducted with GRO runoff and SSS, SST and SIC and has been included as an additional results section. A section has been added to the paper to discuss the results of this correlation analysis (depicted in Figures 4, 5 and 6).

The AOI timeseries has been added to the eastward turbulent surface stress timeseries to enable visual comparison and better support the discussion in text.

Discussion:

1. *The discussion/conclusion emphasizes that wind is the dominant driver of offshore or onshore Lena River plume transport. To strengthen this claim, the study should include additional evidence from the analysis correlating composites and other drivers, for example, the river runoff, ice melting, etc.*

A lagged correlation analysis has been conducted between eastward wind stress, river runoff, and SSS, SST and SIC. The discussion section has been updated to discuss the results of this correlation analysis (depicted in Figures 4, 5 and 6).

2. *Line 415: your claim that because GLORYS12V1, which doesn't include interannually varying river runoff, replicates the SSS pattern well as compared to satellite SSS, suggests that variability in river runoff is not a significant contributor to the interannual variability in GLORYS12V1 SSS. However, GLORYS12V1 utilizes in-situ SSS data, which is how it reproduces the SSS pattern. This does not negate the potential influence of river runoff on interannual SSS variability. The absence of river runoff does not imply that it has no impact on SSS dynamics. Moreover, the correlation between GLORYS12V1 and satellite SSS patterns does not necessarily indicate causation, river runoff could influence Laptev SSS variability, if you make this argument, the paper should conduct a more comprehensive analysis that explicitly investigates the impact of river runoff on interannual SSS variability.*

The section described has now been altered to no longer make this claim. This section has also been expanded to include the results of the lagged correlation analysis.

3. *In Section 4.1, there is a discussion about correlating composites to river runoff, but no results are shown to support the argument. The analysis should be included to provide tangible evidence for the discussion. Additionally, the use of the BEC SSS should be addressed if you want to compare your results to the study of the product as in Umbert et al. 2021, who uses this product. The absence of BEC SSS figures and the correlation against river runoff data should be explained to ensure a coherent argument.*

Correlations with river runoff have now been included in results and discussion sections.

Figure 2 now includes a comparison of all four satellite products which helps to explain the reasoning behind the choice to exclude BEC. The choice to exclude BEC from analysis makes comparison with the Umbert et al. 2021 paper more challenging. However, the discussion and comparison with Umbert et al. (2021) has been expanded upon based on the lagged-correlation analysis between runoff and GLORYS SSS variability.

Figures:

1. *Figure 1, it is unclear why the wind over plots are not the mean for September, similar to the salinity over plots. Providing an explanation for this difference would enhance the figure's clarity and interpretation.*

The choice of wind metric was derived from the lagged-correlation analysis between wind stress and September SSS. June, July and August were found to have the strongest correlations with September SSS so a mean of these months was chosen as the wind metric. The strong correlations with June-August can be seen in Figure 7. This has also been more clearly explained in text in the methods section.

2. *In Figure 3, I strongly suggest to include the other two satellite SSS products*

All four products have now been included in Figure 2.

Minor comments:

Introduction

I suggest including a reference to Umbert et al. 2021 as it also uses SMOS SSS to characterize the Lena River plume in the Laptev Sea, which could provide valuable context and potential links between the two studies.

Reference to the Umbert et al., 2021 paper has now been made in the introduction as well as in the discussion. The comparison in the discussion section has also been expanded to incorporate correlation results.

Methods

In line 280, it is mentioned that the median product is generated using GLORYS12V1, LOCEAN SMOS, and both SMAP products. However, it seems there might be a discrepancy, as it was previously stated that there were four satellite products. This inconsistency needs clarification.

The median product is no longer used. This sentence has been changed to reflect this.

Results

Section 3.3 is missing, but it is referred to in the text as "3.2 3". The authors should correct this discrepancy and make sure the section numbers are accurate.

The section has been correctly renamed to resolve this.

In Table 1, it is puzzling that the median product has more observations than any of the individual products. The authors should address this discrepancy and provide an explanation for the data variations to ensure the table's accuracy and consistency.

The median product is no longer included in the paper so this discrepancy is no longer present.