Review 1 of "Forcing-dependent submesoscale variability and subduction in the coastal sea area (Gulf of Finland, Baltic Sea)" by Salm K. et al., 2025

General Comments

• Clarity and Structure: The manuscript requires significant reorganization and revision for clarity. The writing style is often unclear, with missing or incorrect articles, tenses, and sentence structures throughout.
Thank you for the thorough review. We have carefully revised the manuscript to improve grammar, clarity, and overall flow.

• The Introduction lacks a clear narrative and needs to better articulate the study's motivation, background, and novelty relative to existing work in the GoF.
We have substantially revised the Introduction to clarify the narrative and strengthen the study's motivation and context. The updated text provides a clearer explanation of the relevance of submesoscale processes and the use of spice to characterise them in stratified estuarine systems, such as the Gulf of Finland. We also highlight one aspect of the novelty of our study in terms of combining glider-based observations with a regional high-resolution model to investigate the vertical structure and evolution of spice anomalies, which has not been previously addressed in the GoF context or similar basins.

• Consistency and Rigor: A number of concepts, such as "spice", "tracer variance", and various metrics (e.g., Rossby number, buoyancy gradients), are introduced too late, without sufficient explanation or justification. Terminology should be defined clearly and used consistently.
We have revised the manuscript to introduce key physical and dynamical concepts, including spice, more clearly and earlier. Definitions are now provided at their first mention in both the Introduction and Methods sections, with consistent terminology used throughout.

• Use of Observations vs. Model: The manuscript leans heavily on the model results, and the glider observations appear to be used mostly for validation. This raises concerns about how well the observational data actually constrain or inform the main findings.
We have published an earlier paper based solely on glider data from 2018. We incorporated two more glider missions from the study area to characterize SMS variability in relation to forcing and mesoscale background and used model data to extrapolate (generalize) the findings over a larger spatial area and temporal extent. While the numerical model provides the spatial and temporal coverage necessary to explain the development of submesoscale processes, we have revised the manuscript to better emphasize the role of glider observations beyond validation. Specifically, we state that glider observations revealed maxima of spice above the maximum vertical density gradient during spring missions and below it during the late summer mission. Also, glider data indicated high spice in the sub-surface layer in the case of forcing conditions favorable for coastal upwelling and formation of a long-coastal baroclinic current. Both of these findings were generalized, and potential mechanisms were suggested using model data from a larger area and extended periods.

• Figures and Analysis: Several figures are difficult to interpret due to inconsistent domains, axes, or color scales. Some key diagnostics are not shown (e.g., vertical velocity, currents (depth vs time)), limiting the reader's ability to assess the conclusions about subduction and mixing.
We have reworked several figures to improve clarity and comparability – for example, by harmonizing axes across related plots and adding subplots to include vertical velocity and horizontal currents. New figures were introduced where necessary due to the large number of subplots. These enhancements allow for a more transparent evaluation of subduction and help readers more effectively assess the physical mechanisms discussed in the manuscript.

• A Conclusion section is missing and the Discussion lacks direct references to the figures and results.
We have added a Conclusion section that clearly summarizes the main findings and implications of the study. The Discussion section has been reorganized to directly reference relevant figures and results and includes a more focused comparison with previous studies. These changes improve the cohesion of the manuscript and highlight how our findings contribute to understanding submesoscale dynamics and vertical structure in the GoF.

Specific Comments:

Abstract: the abstract has to be improved so it matches text body and the analysis and it is self-explanatory so the reader knows what exactly was done.

• L2: Rephrase the second sentence—it's unclear and does not reflect the analysis accurately.
• L4: Specify what "tracer variance" refers to (i.e., spice variance).
• L6: Clarify "around UML" – upper mixed layer? Provide depth range.
• L10–13: Specify what atmospheric forcing is meant; consider merging sentences.
• L14: "likely SMS flows" – was this demonstrated? If so, how?
• L15: Be specific—"high tracer variance" of what exactly?
• L16: The final sentence is vague—was this shown, or is it speculative?

Thank you for your suggestions. We have reformulated the abstract to stress better what was done and what are the main findings.

Introduction (L20–85): this needs to be seriously improved as the writing style is not good, the paragraphs need to be reorganized so the background, motivation and what will be done and why is clear to a reader.

• L20: Replace "forces" with a clearer term
The sentence was made more precise.

• L21: Expand the second sentence to explain the background of SMS dynamics.
The sentence was revised, including the SMS intermediate horizontal scale.

• L23: Explain why SMS features are important—link to physical or biogeochemical processes.
We have reorganized the opening paragraphs to emphasize the importance and impact of SMS flows from the outset.

• L25: Add citations to definitions of SMS.
Added.

• L28: Rephrase the final sentence and state the region explicitly.
We have relocated the paragraph on glider studies to the third position, where it now naturally leads into the introduction of the study region.

• L31–42: Revise for English grammar and completeness. "The" and "of" are frequently missing.
Revisions were made to the regional description paragraphs to improve coherence and readability.

• L36–38: Add "semi-enclosed" to describe the Baltic Sea.
Added.

• L41: Specify freshwater input sources—mention the main rivers.
We now mention the Neva River, the largest freshwater source and most relevant in the context of the GoF. We believe listing all major rivers is not necessary in this context.

• L45: "In contrast to the open ocean" implies salinity is unimportant there—rephrase.
The sentence was revised and moved.

• L46–47: Sentence unclear—needs rewording.
The sentence was reworded.

• L48: Clarify whether Lips (2009) and Vali (2017) estimated mixing or just described it.
The sentence was revised for clarity. These works address different aspects of the topic. While Lips et al. (2009) emphasized the role of coastal upwelling in enhancing vertical mixing and nutrient transport, Väli et al. (2017) focused on the SMS structures that arise during such events. Lips et al. (2009) estimated via TS-analysis that the upwelled water comprised approximately 85% intermediate layer water and 15% surface mixed layer water.

• L57–58: Replace "captured" and "prevalence" with more precise terms.
The sentence was reworded.

• L60–61: What is the key modeling advantage? State it up front.

Modelling advantage was stated more clearly (including, models help to extend the findings over space and time).

• L62–64: Introduce the glider earlier; explain why the upper half of the water column is the focus.
We have relocated the paragraph on glider studies to the third position. We focus on the upper half of the water column, where SMS activity is most prominent, and the glider data are more densely sampled, allowing more robust analysis.

• L65: Provide the exact mission durations.
The exact mission durations were included in Methods.

• L66: Rephrase "favored".
The sentence was revised ("associated").

• L68: Define "tracer variability" – first mention needs explanation.
To improve clarity, the explanation of spice was relocated to immediately follow its initial mention.

• L70–72: Why is only spring-summer analysed? given the model covers a longer period. The two summers you analyze are similar or different?
Our study is limited to the spring–summer period, when seasonal stratification develops and becomes well established. SMS processes are known to exhibit seasonal variability (e.g., Wang et al., 2018; Yu et al., 2021), and this time window allows us to isolate conditions, such as upwelling, frontal activity, and spice gradients, when SMS dynamics is most active and relevant to our objective of understanding SMS generation under stratified, wind-driven conditions. Furthermore, spice, as defined in the present study, is best applicable during seasons when both temperature and salinity have comparable contributions to density variations. It is not the case in brackish waters at low temperatures when salinity mostly defines the density variations. For instance, the temperature of maximum density in the Gulf of Finland is about 2.5 °C.

Although the model simulations cover a longer time period, our analysis remains focused on spring–summer to ensure consistency with the observational data and to isolate seasonal conditions favourable for SMS generation. Notably, while both summers analysed share common features – such as surface stratification and wind-driven upwelling events – there are also interannual differences in the timing and intensity of these processes, which are examined in the context of their influence on SMS variability in the Discussion section.

• L72–75: Clarify the relation to Salm (2018)—same dataset?
The explanation was added in the text. This study builds on earlier observations of SMS features in the GoF (Salm et al., 2023), extending the analysis to include all three missions conducted in the same area.

• L77: Explain why spice is used and what it captures. This is not a common term.
The explanation of spice and the motivation for using it was expanded. Among other arguments, spice reveals anomalies and spatial gradients along isopycnals, where SMS processes often act. It serves as a proxy for SMS intensity capturing variability without being masked by vertical excursions of isopycnal surfaces, i.e., internal waves.

• L81–85: The hypotheses should be clearly formulated and tested in the Results.
We refined the hypotheses to enhance their clarity and alignment with the study objectives. First, we suggest that SMS variability is modulated by both atmospheric forcing, particularly surface heat flux and wind stress, and the background (larger-scale) hydrographic structures, including mesoscale frontal gradients. Second, we propose that topographically induced instabilities of baroclinic coastal currents create favorable conditions for SMS subduction, enabling offshore and downward transport of tracers.

2. Materials and Methods

2.1 Glider Observations

• L90: Grammar issues—"Three missions were performed…"
The beginning of this section was reworded, including the exact dates of the missions.

• Specify exact dates, transect directions, and water depth coverage.

These transects were oriented across the southern coast of the GoF, capturing cross-shore variability. The glider profiled the water column from the surface down to depths of 80–100 meters, depending on the position. While under the surface, the glider started to turn around either 4 m before the surface or 5–6 m before the seafloor.

• How was data quality-controlled? Cite appropriate methods.
The raw data were quality controlled following procedures adapted from Argo quality control protocols (Wong et al., 2025).

• Why were transects oriented differently between missions?
Although originally conducted for different research objectives, three glider missions in the GoF, Baltic Sea (Fig. 1) collectively provide a data set for this study.

• L98: Describe how the data were interpolated (vertical/horizontal resolution).
Explanations added. The interpolated data fields had a vertical resolution of 0.5 dbar and a horizontal resolution of 10 min. Interpolation was performed on a two-dimensional dataset using pressure as the vertical coordinate and time as the horizontal coordinate.

• Fig. 1b: Include a broader regional map with coastline for context.
The map is slightly extended.

2.2 Model

• Clarify whether "adaptive vertical coordinates" refers to sigma or z-coordinates.
We have specified the description of adaptive vertical coordinates. Such a grid is a generalization of sigma layers with the potential to enhance vertical resolution near boundaries and in layers with strong stratification and shear (Klingbeil et al., 2018).

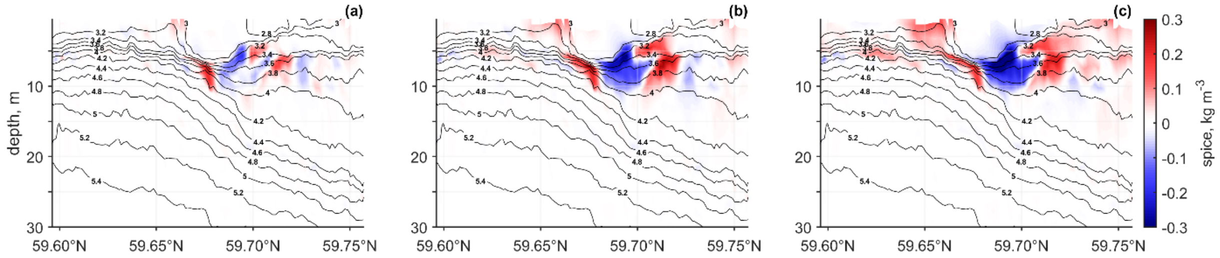
• L110: Why was the model vertically interpolated? Was this to match glider data?
It is not clear which line/sentence is referred to here. In this section, interpolation is discussed in relation to generating boundary data for the high-resolution model from the coarse-resolution model. The output of the coarse-resolution model at boundaries was vertically interpolated to fixed z-levels for input to a high-resolution model, as there are spatial differences between these models.

• Use consistent tenses (past for methods).
The tenses were checked.

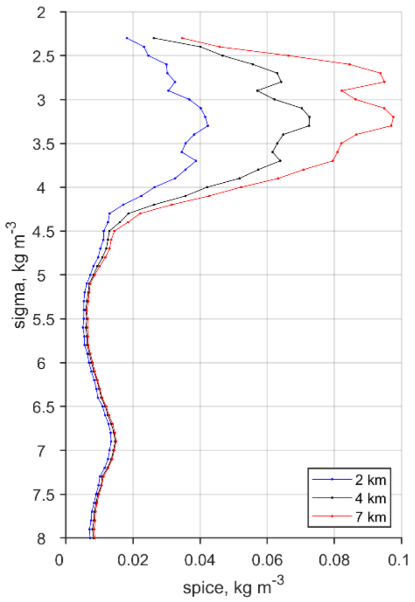
2.3 Analysis

• L122: Why use 4 km filtering? What would 2 or 7 km yield? Discuss sensitivity.
• L128–130: Refer to Alenius et al. (2003); discuss how scale selection affects variance.
The chosen length scale is consistent with the internal Rossby deformation radius in the GoF, which is typically 2–4 km (Alenius et al., 2003). The choice of a 4 km averaging scale offers a practical balance between resolving SMS structures and suppressing high-frequency noise. Smaller scales may exaggerate variability and obscure persistent features, while larger scales risk smoothing out key SMS signals. Thus, 4 km averaging preserves the essential gradients and anomalies linked to SMS dynamics without compromising interpretability.

To illustrate sensitivity to the choice of horizontal scale, we present an example using the same section shown in Figure 2. Despite variations in horizontal scale, the structure and location of spice anomalies remain consistent across all three estimates, supporting the robustness of the observed frontal features. However, the magnitude of anomalies increases with smoother (larger) scales, likely due to spatial averaging. This is further supported by the depth-resolved standard deviation profiles, which show systematically lower spice variability at 2 km and higher values at 7 km, especially in the upper layers. See the referred figures below.



The figure above shows an example of spice distribution using horizontal averaging scales of 2 km (a), 4 km (b), and 7 km (c). Each panel displays the same section, overlaid with density contours at 0.2 kg m^{-3} intervals. The data are based on glider observations from 24–25 May 2018.



The figure above shows the standard deviation of spice calculated from the glider mission conducted between 9 May and 6 June 2018. The colours indicate the horizontal averaging scales used for spice: 2 km (blue), 4 km (black), and 7 km (red).

• L140: Define how spice variance is calculated—add the equation.
Upon review, we realized that the sentence regarding spice variance was inaccurate and, therefore, removed. Spice variance is not calculated in this study.

• L148: "For it" – unclear, remove.
Removed.

• L150–155: How was N^2 calculated? What vertical spacing and smoothing were used?
Vertical buoyancy gradient was calculated using a 2 m vertical interval.

• L155- this is a Result or it could be also in the Introduction- move it
We agree. It was moved to Results.

• Why the features can be displaced in the model- this should be covered in the introduction
This content has been repositioned and is now covered in the Introduction.

• Parameters such as Rossby number, Ri , buoyancy gradients should be at least mentioned in the Introduction, what the analysis will be performed and why?
Parameters and their relevance are mentioned in the Introduction.

• L170: central 'difference' ? 'above' missing words
Centred finite-difference scheme was meant here.

• Why the wind components were smoothed?
The wind components in the analysis were smoothed by a Gaussian low-pass filter for 6 h to reduce high-frequency noise and highlight relevant forcing scales.

• Fig 2. This should be presented in the separate section e.g model validation or section 3.1 , the spice is shown in density domain, but the rest in time-depth domain, why? What about the currents and vertical velocities? They are important for SMS, maybe not in the observations but the dynamics can be shown in the model.
Spice is derived in density space, where isopycnal-referenced anomalies yield the most meaningful physical insights. We ensure that spice captures the variability relevant for detecting SMS dynamics, even if it means sacrificing easy visualization in fixed-depth space. However, to make the presentation clearer, we revised the figures where relevant to display spice in depth coordinates. We also now show vertical velocities where relevant, but state that vertical velocities in the frequency domain of interest are largely contaminated by the vertical movements of isopycnals (internal waves). Thus, their more detailed presentation is not relevant here.

<u>3. Results</u>: While the section presents several relevant observations and model outputs, it lacks key quantitative metrics to support the conclusions. If the main findings rely primarily on model-derived interpretations of physical processes, this should be substantiated with appropriate statistics and objective measures. For example, the influence of topographic steering is not sufficiently demonstrated, and vertical velocities—crucial for diagnosing subduction or vertical exchange—are not shown. Including such diagnostics would significantly strengthen the analysis and the credibility of the inferred processes.

3.1 Validation:

• Correct the language, there are some missing words e.g in the title etc.
We have carefully reviewed and revised the manuscript to correct grammatical and typographical issues, including missing words.

• L181–185: Provide not only mean differences but also standard deviations.
We have updated the text about the UML, maximum vertical density gradient, and CIL depths, providing both mean values and standard deviations when comparing observations and simulations. This offers a more comprehensive representation of variability and model performance.

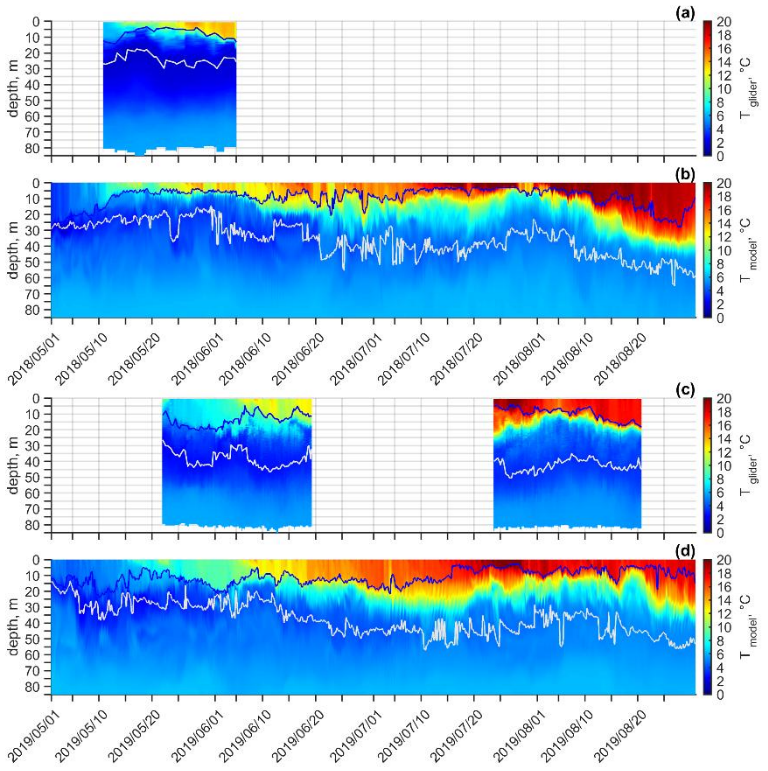
• L190: Define “slight” differences numerically.
The term “slight” has been replaced with specific numerical differences to improve clarity and avoid ambiguity.

• L192–196: Indicate which depths the model fails to resolve secondary maxima.
We now explicitly indicate the depths at which the model fails to capture the observed secondary maxima in the vertical density gradient (observed at depths of 19-20 m), allowing readers to better assess the limitations of the model.

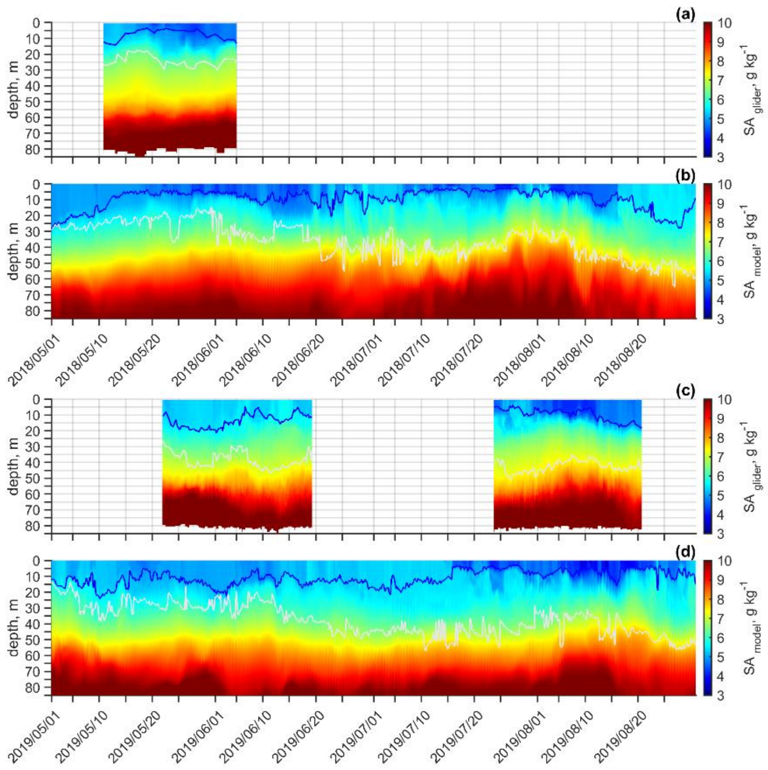
• L197: If stratification was weaker, state it clearly.
The revised sentence now clearly states that stratification was weaker.

• L205: Are the largest spice values associated with maximum T/S gradients? Clarify.
Yes, largest spice values are close to maximum T/S gradients, but above the pycnocline during spring missions and below it during the late summer mission.

• Fig. 3: Instead of mean-removed slices, show actual matched cross-sections (as in Fig. 4).
• Fig. 4: Difficult to compare due to gaps in glider profiles. Consider interpolating data.
Both figures are revised. An example of sections is shown below (new Fig. 3), where glider data are interpolated to match the model output (although there is no complete coincidence of the data in space from the model and observations). These will be shown instead of average profiles from the missions).



The figure above shows temperature variability based on glider data (panels a, c) and model output (panels b, d) for May–August 2018 (a, b) and 2019 (c, d). The glider data represent average profiles for each section within the selected area, forming a composite data field. The model data correspond to average profiles within a 1×1 km window, providing a profile at each model timestep. The blue and white lines show the UML and CIL depth, respectively.



The figure above shows salinity variability based on glider data (panels a, c) and model output (panels b, d) for May–August 2018 (a, b) and 2019 (c, d). The glider data represent average profiles for each section within the selected area, forming a composite data field. The model data correspond to average profiles within a 1×1 km window, providing a profile at each model timestep. The blue and white lines show the UML and CIL depth, respectively.

• L220: This sentence is confusing—aims should be explained in the Introduction.
The beginning of the paragraph was revised.

• L225–235: Clarify whether text refers to model or observations.
We have clarified in the text that the analysis in this section is based on model-derived spice fields. The paragraph describes seasonal variability using standard deviations of modelled spice.

• Fig. 5: Clearly label data sources (glider vs. model).
This figure shows the standard deviations of modelled spice, which is now explicitly noted in the figure caption.

3.2 Wind Forcing

• Are wind differences between years sufficient to explain SMS variability?
We used both wind data and heat flux estimates to explain temporal variability, including differences between the years. Wind is most crucial for the development of mesoscale features (e.g. baroclinic current), while heat flux may define the potential of SMS activity at the base of the upper mixed layer (if it is large enough compared to the wind mixing).

• Fig. 7: Add mean wind arrows for reference.
We prefer to keep this information in the text since it is not obvious which period should be used to derive a mean vector that is relevant for each situation.

3.3 Submesoscale Indicators

• L279: 'Maxima of relative vorticity'—show this in a figure.
This phrase was potentially misleading and has been revised for clarity to more accurately reflect the intended meaning to describe Fig. 8.

• L300: "Changing wind forcing" – specify exact changes.
We have revised the sentence to specify the observed transition in wind direction.

• Fig. 8: Why show minimum temperature instead of UML?
The depth of the minimum temperature was chosen as a boundary in Fig. 8 because it effectively captures the upper part of the water column, encompassing both the UML and the thermocline. The UML is typically only a few meters thick, and a significant portion of the spice variability is observed not just within the UML but within the thermocline. Therefore, using the minimum temperature depth provides a more representative boundary for analysing upper-layer variability.

• Relationship between surface Ro and spice variance should be discussed quantitatively. E.g correlations etc.
We chose not to emphasize a direct correlation analysis between surface Ro and spice variance. As shown in Fig. 8, there are situations with coinciding high Ro and high spice, high Ro and low spice, and low Ro and high spice. We chose the case with the highest spice (but low surface Ro) and tried to explain the mechanism of its creation.

• Discuss limitations of model—e.g., can EO data be used for validation? In the summer this should be possible
It is not the topic of the Results section and is discussed in the Discussion section with reference to other studies.

• Fig. 10: Use consistent axes (depth vs. density space); add vertical velocity if available.
We have revised Fig. 10 to use consistent axes across all subplots and included vertical and horizontal velocity fields.

• L345: "Probable" – vague. Can this be quantified?
It is difficult to quantify it. We explained this probable link between the coastal current, topography, configuration of the coastline and SMS variability in a descriptive way. In the revised manuscript, an analysis is

added to incorporate wind forcing into this topic in more detail. The conclusions will remain as suggestions (not quantitatively approved). We clarified the text in the Results and Discussion section accordingly.

• Fig 11. Why you not use UML or vertical currents? For comparison with spice?
Thank you. We stress more clearly that high spice below the maximum vertical density gradient occurs when the upper mixed layer is shallower (upwelling-favourable winds prevail). See Fig. 11. As stated above, vertical currents are contaminated by vertical movements of isopycnals (internal waves) and are difficult to directly link with SMS activity.

4. Discussion

• L367–374: General background not linked to results—consider trimming.
We omit general statements and link the discussion to the results of the present study.

• L385–395: Was vertical velocity shown anywhere? If not, speculative statements must be softened.
It is a referred statement. We did not deal with vertical velocities as indicators of SMS activity. The sentences are reworded accordingly to clarify it.

• Discuss how "elongated regions" of spice relate to fronts or subduction processes.
Thank you. We added relevant discussion when analyzing the appearance of high spice in relation to baroclinic currents, their instability and the reasons behind (topography and/or changes in wind forcing).

• Consider plotting spice vs. Ro for correlation.
As we explained above, we do not consider it worthwhile to show (there are situations when the correlation exists and when it does not).

• L400: Discussion on winter SMS processes seems out of place—study only covers summer.
We left it out.

• L402: "SMS only visible with T and S gradients"—is this a limitation of the method?
Yes, it is discussed as a limitation of the method. It is mentioned now explicitly.

• L409–411: Clarify whether this is your result or literature-based.
It is a literature-based statement. We link it with the results of the present study, indicating that our results reveal an opposite tendency (as do some other studies) towards enhanced SMS activity at the base of the UML when surface heating prevails over wind mixing.

• L423–428: Cite figures for all claimed results.
We checked it and cited the figures where relevant.

• L437–443: Strong claims based on limited evidence—can they be supported by broader statistics?
As we explained, adding statistics is challenging, and we base our suggestions on the analysis of specific situations. We agree that the statements need to be softened (as we did in the revised text).

• L444–452: This paragraph appears unrelated to the study and could be removed.

We removed most of it but retained a suggestion that such SMS subduction could also influence the creation of subsurface phytoplankton maxima, as observed earlier in the Gulf of Finland during summer.

Missing Conclusion

• A summary of key findings and a clear answer to the hypotheses are needed.
• Clarify how glider data contributed—was it only for validation?
• Highlight the study's novelty and limitations clearly.

We formulated concluding remarks as follows: This study demonstrates that submesoscale variability in the Gulf of Finland is strongly modulated by both atmospheric forcing – particularly surface heat flux and wind stress – and background hydrographic structures such as mesoscale frontal gradients. Glider observations, supported by high-resolution modelling, revealed consistent spatial patterns of SMS activity, with spice anomalies concentrated near the UML base in spring and within the thermocline in late summer, demonstrating the vertical sensitivity of SMS features to seasonal stratification. While seasonal stratification played a key role in shaping SMS structure, wind forcing became dominant under weaker surface buoyancy input. High spice variability and subduction signatures were consistently found on the offshore side of a baroclinic coastal current, where sloped isopycnals aligned with velocity and spice gradients indicated downward and lateral transport of surface-layer water masses. The integration of observations and model output allowed for extrapolation beyond individual glider transects, confirming that SMS processes in this coastal sea are both dynamically active and responsive to variations in external forcing. Together, these results clarify the physical mechanisms driving SMS variability and subduction in stratified coastal environments.

Recommendation: Major Revisions

The manuscript addresses a relevant topic and includes valuable datasets. However, substantial improvements in structure, clarity, and analysis are needed. The roles of observations and model outputs must be better defined, and the analysis should be aligned more closely with the stated objectives.

References

Klingbeil, K., Lemarié, F., Debreu, L., and Burchard, H.: The numerics of hydrostatic structured-grid coastal ocean models: State of the art and future perspectives, Ocean Modelling, 125, 80–105, doi:10.1016/j.ocemod.2018.01.007, 2018.

Wong, A., Keeley, R., Carval, T., and the Argo Data Management Team (2025). Argo Quality Control Manual for CTD and Trajectory Data. http://dx.doi.org/10.13155/33951