Review 2 of "Forcing-dependent submesoscale variability and subduction in the coastal sea area (Gulf of Finland, Baltic Sea)" by Salm K. et al., 2025

This paper aims to observe and understand variability in submesoscale processes within the Gulf of Finland, using a combination of in-situ observations from a glider platform taken in 2018 and 2019, and model simulations using GETM. They found that the depth at which submesoscale processes occurred at varied between spring and summer, and linked this with changes in stratification and atmospheric forcing. Salm et al also highlight the importance of a coastal current for stimulating submesoscale subduction and explore the connections with topography.

They predominantly use spice as a tracer to identify variability in the water column in both the model and observations but describe other parameters that can be used to identify submesoscale variability (such as horizontal buoyancy gradients) for the model results. The authors present a valuable dataset with glider observations and model simulations, and this is of importance to publish, especially for further understanding of submesoscales in the Baltic Sea. I feel that the presentation and discussion of the results is lacking in structure and some content, and my recommendation is for major revisions.

Thank you for the thorough review and comments. To better focus and structure the paper, we have defined the following hypotheses. 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.

General remarks

- Throughout, language and grammar need to be double checked. Specifically, the word "the" is often either missing, or unnecessarily added in to sentences.
We have thoroughly revised the manuscript for language and grammar. We have carefully reviewed the use of articles (e.g., "the") and corrected inconsistencies or errors throughout the text.

- The structure and motivation for the paper needs to be clarified. It feels that the glider observations have a minor (only validation) role to play and the bulk of the critical analysis is reliant purely on the model simulations.
We have substantially revised the Introduction to improve the presentation of the motivation and objectives of the study. Since we have published an earlier paper based solely on glider data from 2018, we see this paper as an extension of the study by incorporating two more glider missions and high-resolution model results. We have characterized SMS variability based on glider data in relation to forcing and mesoscale background. The model outputs are used to extrapolate (generalize) the findings over a larger spatial area and temporal extent. 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.

- Some of the analysis is very reliant on descriptive or more qualitative assessment of figures, or include reference to parameters not presented. The paper could benefit from identifying some more specific or quantitative metrics (such as those used in several of the papers cited and referred to).

We have added numerical values to clarify several previously qualitative assessments. Phrases that may have been misleading, such as those implying reference to parameters not shown, have been revised to more accurately reflect the intended interpretation of the figures and improve clarity.

We agree that incorporating more specific or quantitative metrics can enhance the analysis. Among the three glider missions, only one included a sufficiently long spatial transect to support meaningful calculation of horizontal gradients or derived diagnostics regarding background (mesoscale) conditions (partly presented by Salm et al. (2023)). Given these constraints, we used complementary metrics derived from the model, which provides spatially continuous coverage and spans over longer periods.

Specific comments

Abstract

- After all other corrections, I would review the abstract to tighten it up, be more specific when referring to dates, review language such as "suggest", "likely" – can you be more specific? What is the implication and impact from your findings?
We have reformulated the abstract to stress better the main findings and differentiate what is well founded by data and what can be suggested but must be studied in more detail in the future.

Introduction

- Throughout the introduction; double check grammar and sentence structures. Especially in Lines 20-40 there are lots of additional (or missing) "the"s which can make it a bit trickier to read through.
We have thoroughly revised the Introduction.

- Lines 20-25; I would include the importance and impact of SMS flows here (you talk about the links with carbon, heat etc later on).
We have reorganized the opening paragraphs to emphasize the importance and impact of SMS flows from the outset.

- Line 45; "In contrast to the open ocean…" ; this is not necessarily true, salinity has a significant stratifying role in many parts of the oean. Please rephrase.
The sentence was rephrased.

- Line 57: "captured" – remove this word, it doesn't add to the sentence.
The sentence was reworded.

- Line 57: "supports the prevalence" – the number of studies doesn't necessarily mean there are lots of SMS flows/variability, but they do highlight the importance. Maybe rephrase?
The sentence was reworded.

- Line 60: this suggests that since observational data are limited, your glider data is of high importance? What are the limitations of relying on the model simulations? Can you include some review of literature/other modelling papers in the Baltic that support the usage of high-resolution models for this study.
Modelling advantage was stated more clearly. We do not present a literature review but refer to earlier modelling studies of SMS variability and processes in the Baltic (e.g. Väli et al., 2017, 2024; Chrysagi et al., 2021).

- Line 65: you introduce the gliders here; could this be mentioned earlier when talking about SMS observations, are there benefits to specifically using gliders when trying to observe SMS flows?
We have relocated the paragraph on glider studies to the third position and improved it.

- Line 65: what are the dates for the missions (what does spring-summer mean?)
The exact mission durations are included in Methods.

- Line 75: I'd be interested to see a bit more introduction of spice here for readers who aren't so familiar. Why use spice over other submesoscale parameters also reported in the papers you cite throughout (e.g. horizontal buoyancy gradients, parameterisations of SMS flows through Ekman and Mixed Layer Eddies)?
To improve clarity, the explanation of spice was relocated to immediately follow its initial mention.

- Line 86: you refer to a section after section 4/Discussion, but there is no final section.
We have added a Conclusion section that clearly summarizes the main findings and implications of the study.

Data and Methods

- Line 90-94: state explicitly the dates, was it the same transect (if not, why not, how are they different), water depths covered. Maybe a table could be useful for that?

We have included the exact dates of the missions and improved the paragraph. Although originally conducted for different research objectives, three glider missions in the GoF, Baltic Sea, collectively provide a data set for this study. 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.

- Line 94: “YOs” is a very specific term, can you find an alternative word to describe the half profile (upcast / downcast)?

We replaced “YOs” with “up- and downcasts”.

- Line 96: Is there a reason to interpolate on time rather than space?

We chose to interpolate the glider-derived parameters in time rather than in space because the glider observations are naturally indexed by time due to their profiling nature. Each measurement corresponds to a specific timestamp rather than a consistent spatial location. However, the choice depends on the purpose. In some cases, spatial interpolation – using distance, latitude, or longitude – can be more beneficial, e.g., when comparing repeated measurements or sections taken at the same location.

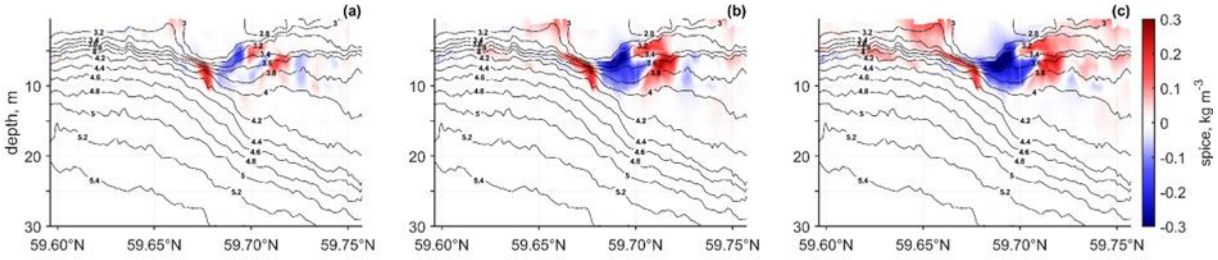
- Figure 1: very hard to see the magenta cross for the wind data
It was improved.

- Line 140: put the equation for spice into a separate line to accentuate it.
The equation was moved into a separate line.

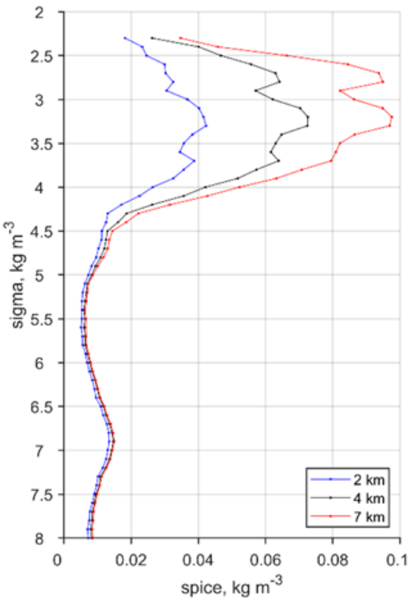
- Line 142: citation needed for TEOS-10?
The citation was added.

- Line 143: how does the choice of 4km impact results? Did you test with higher/lower scales?
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. We examined smaller/larger length scales when considering the scale for the glider data.

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).

- Line 146: You refer to spice variance throughout the paper, but it is predominantly spice that is plotted in the figures. Is it a qualitative assessment of the figures that results in your analysis of the variance, or do you numerically calculate the variance? If so, can you state that here?
Spice variance is not calculated in this study. The phrase was reworded.

- Line 149/150 (and fig 2): Is it possible to plot spice on the same y-axis? Hard to compare between the panels when T and S are plotted against depth and spice is plotted against sigma.
We have harmonized axes across related plots where relevant.

- Line 155: This reads as a result? Maybe rephrase this to introduce this paragraph.
We agree. It was moved to Results.

- Line 161: you calculate several SMS characteristics from the model data. Is it possible to calculate some of these from the glider data too (as is done in papers you cite, such as Thompson, du Plessis et al)
Yes, most of the parameters are, in principle, possible to estimate from glider sections (with certain limitations). We did so in an earlier paper (Salm et al., 2023). In this study, we preferred to calculate them (except for spice and characteristics of vertical stratification) from model data, which provide spatially continuous coverage and span longer periods.

- Line 170: Not sure what you mean by "Central scheme…"
The centred finite-difference scheme estimates spatial derivatives by combining forward and backward differences, which correspond to differences taken with respect to neighbouring points ahead and behind in space.

- Line 171: Were these gradients used for N2 as well?
Vertical buoyancy gradient was calculated using a 2 m vertical interval. It is now clarified in the text.

- Line 174: Why do you smooth over 6 hours? Is there a motivation for choosing this timescale?
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.

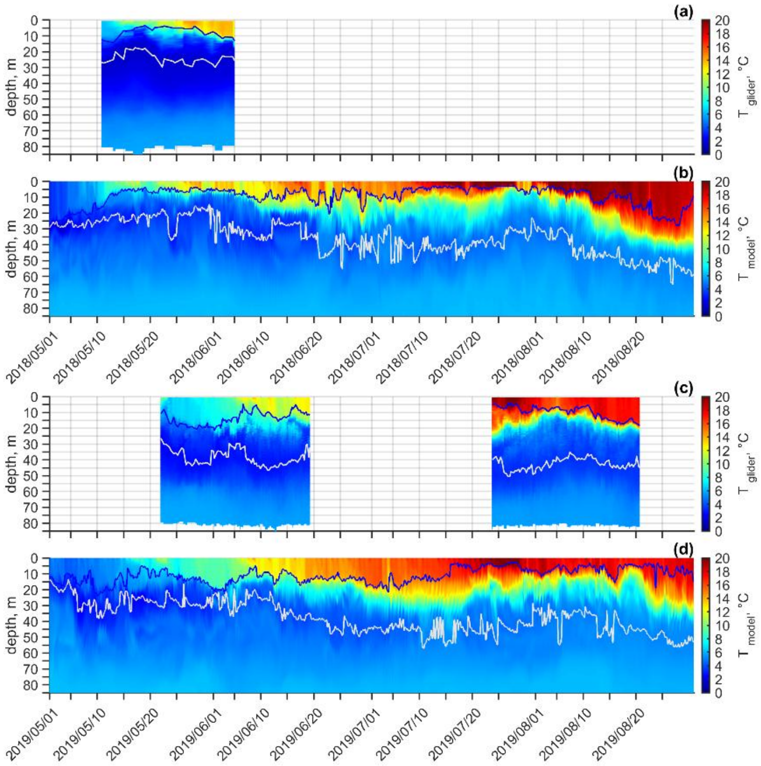
<u>Results</u>

- Line 183: "over which ten profiles were gathered": is that per transect of the glider, so in total you have 10x the number of repeats the glider did?

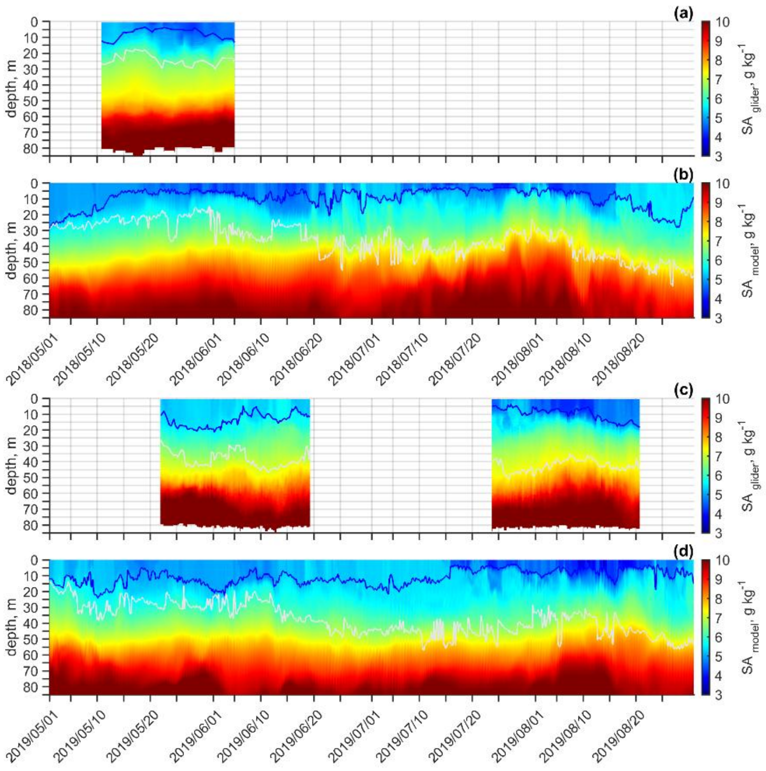
We have rephrased the text for clarity. This common area corresponds to approximately 1 km of glider track per section. Therefore, a profile is obtained per section, and the total number corresponds to the number of repeated sections.

- By averaging the profiles you lose a lot of the benefits and advantages of the glider data

We now present temperature and salinity as we previously did for the squared Brent-Väisälä frequency. See the figure below.



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.

– it would be interesting to see the variability in spice (from glider data) over the transect/mission timeline, and to compare that variability to the seasonal variability that the paper is focused on (i.e. is the shorter timescale comparable to the seasonal?).

- Line 187-193: I struggle to understand what magnitude of difference between model and observations is significant. How much can it impact the spice calculations and the final results? From the figures it is clear that the model does not perfectly replicate the observations, what level is acceptable? (e.g. "slight differences" – what does that. Mean?)
The paragraph was revised and the term "slight" has been replaced with specific numerical differences to improve clarity and avoid ambiguity.

- For quite a few of the figures (4, 8, 9, 10 , 11), it could be beneficial to point towards or highlight features that you discuss in the text (e.g. with a small triangle/arrow/line).
We agree. We have marked referred periods and features to improve the readability of the figures (and text).

- Line 199: "UP" – if you have enough space, I recommend typing this out in full, "upper pycnocline". Try to remove unnecessary acronyms as much as possible.
We removed this acronym.

- Figure 4: can you interpolate the glider data across time? It is hard to pick up the features you discuss due to the individual profiles.
The gaps have been removed and discussed periods/features marked.

- Missing more quantified discussion of submesoscale characteristics (e.g. line 245: can you calculate how much the wind changes could impact wind induced SMS flows?). I think this links to the somewhat qualitative or descriptive discussion of the spice tracer; adding quantification where possible will strengthen your conclusions.

- Line 278-285: can you put lines/mark out these events on the figure?
Yes, we did so in the revised manuscript.

- Line 290: highlight on the figure when these three events occur, help the reader out as much as possible by making it easy to follow your arguments and analysis.
Done.

- Line 294: you talk about horizontal buoyancy gradients, but there are none shown in Figures 7 or 8) that you refer to. In addition, you are talking about an event in the second half of May 2018, but you refer to figure 7j,m which shows an event in June.
We checked the text and cited figures. There was a mistake, and we corrected it.

- From line 335; is this intended to be a new subsection? It is somewhat detached from the previous paragraph (maybe this is just the formatting in the draft version though).
We ensured that all the text for subsection 3.3 appears before the figures.

- Paragraph around line 345: Are there any calculations or analysis that can be done to make this more certain? Or rephrasing of your results (a lot of "probable", "suggests" etc).
We have made an additional thorough analysis of the selected situation with high spice, which was detected offshore from the coastal baroclinic current. We explained high spice by the presence of the coastal current and its instabilities. It is evident that such situations emerge when upwelling-favorable conditions prevail and the coastal baroclinic current develops. The conclusions will remain as suggestions (not quantitatively approved) since it is difficult to identify a single reason for instabilities and the observed high spice immediately below the strongest vertical buoyancy gradient. We added/refined figures. We clarified the text in the Results and Discussion section accordingly.

- Figure 11: can you colour in the bathymetry in black or similar? It took me a while to spot the features that you refer to in the text. And is there space along the bottom of the figure to add a panel showing the overall topography so we have context of what the rest of the bathymetry is like away from the <40 m peaks?

The figure was redrawn to also show current vectors in the subsurface layer, and the topography was added as a layer on it.

Discussion

- Paragraph 367-374: Feels more like a paragraph for the introduction? Or link it to your discussion/results a lot earlier on.

We omit general statements and link the discussion to the results of the present study.

- Line 391: what do you mean by elongated regions? Vertically or horizontally?

They are elongated horizontally. We added a reference to the figure in the Results section.

- Lines 396-403: the discussion here is comparing to studies that include winter SMS flows / full annual cycles of submesoscale variability (seasonality). The paper presents itself as looking at seasonal variability, but you only look at spring-summer and not the full seasonal cycle. It seems that this is done to only use the model for periods when the observational data is available to validate it, but this results in a compromise on both of the data sources: we lose the high resolution variability that could be interesting to look at in the glider data, and also lose the full seasonal cycle that the model could present us. Do you trust the model simulations enough to gain a small insight into other seasons?

We do not discuss winter conditions in the revised manuscript.

- Line 404: Can you give some insight into how much this could impact your conclusions? Is this a significant limitation in using only spice as a tracer for submesoscale flows?

It is stated now explicitly that the analysis based on spice (as it is 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.

I feel that the paper is missing a final conclusion statement or section, to give a summary of the main results and findings. Also a clear comment on the impact/implications of the study, and any limitations or future aspects to explore.

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.

References

Chrysagi, E., Umlauf, L., Holtermann, P., Klingbeil, K., and Burchard, H.: High-resolution simulations of submesoscale processes in the Baltic Sea: The role of storm events, J. Geophys. Res.: Oceans, 126(3), doi:10.1029/2020JC016411, 2021.

Väli, G., Meier, H.E.M., Liblik, T., Radtke, H., Klingbeil, K., Gräwe, U., and Lips, U.: Submesoscale processes in the surface layer of the central Baltic Sea: A high-resolution modelling study, Oceanologia, 66 (1), 78−90, doi: 10.1016/j.oceano.2023.11.002, 2024.

Väli, G., Zhurbas, V., Lips, U., and Laanemets, J.: Submesoscale structures related to upwelling events in the Gulf of Finland, Baltic Sea (numerical experiments), J. Mar. Syst., 171, 31–42, doi:10.1016/j.jmarsys.2016.06.010, 2017.