We thank the reviewer for spending time and effort reviewing our work. Below the reviewer's comments are in **black** and the replies in **blue.**

In the manuscript "Evaluation of SWOT HR PIXC version D water level time series of small lakes", the authors validate the SWOT HR PIXC product for constructing water surface elevation (WSE) time series over 37 Danish lakes. The study proposes a filtering and aggregation workflow for HR PIXC point-cloud data and evaluates the resulting lake-level estimates against in situ gauge observations.

The topic is relevant to the hydrology and remote-sensing communities, and the technical processing appears generally sound. However, in its current form the paper reads largely as a product-validation and methodological benchmarking study, and it feels borderline relative to the scope typically expected for a full Research Article in HESS. The manuscript could potentially fit better as a Technical Note in HESS unless the broader hydrological implications and generalizability are strengthened.

Reply: We thank the reviewer for their careful reading and thoughtful suggestions, which we believe will help us strengthen and clarify the manuscript. We address each of the points raised below. We hope that these comments and revisions make the manuscript's contributions more compelling and clearer for a broader readership.

Major comments

Native PIXC spatial sampling and implications for robustness

A key advantage of SWOT is its dense, two-dimensional spatial sampling of water surfaces. While the validation ultimately compares a single aggregated WSE value per lake per overpass, the reliability of this estimate depends strongly on the spatial sampling and point density of the HR PIXC data retained after filtering.

Reply: We thank you for your comment. We agree with you that SWOT's strength is its dense 2D sampling of WSE and WSS. We plan to strengthen this argument in the manuscript. We furthermore agree that the quality of the aggregated single WSE measurement per lake and SWOT acquisition depends on the number of points retained after filtering. We address this issue by requiring the minimum number of points ('N' – fx. table 1). We will make sure to make the argument clearer.

Please

(i) clearly state the native spatial sampling characteristics of the HR PIXC dataset (e.g., effective point spacing/footprint and any relevant along-/across-track considerations), and

Reply: We thank the reviewer for the helpful feedback regarding the lack of discussion of the SWOT sampling geometry. We understand that adding these will improve the understandability of out Manuscript. We plan to add a paragraph on it in the revised version.

(ii) discuss how these characteristics interact with the selected filtering/selection criteria (e.g., water fraction thresholds, quality flags, distance-to-shore constraints, etc.). In particular, it would be helpful to explain how the retained point-cloud geometry affects representativeness for small lakes, near-shore zones, and narrow/irregular lake shapes.

Reply: That is a good point. For example, it will be harder to reach the required threshold of 200 retained points in the near range of the swath given SWOT's sampling geometry. Yet the individual point quality should likely be higher in the close range. An adaptive threshold based on lake position in SWOT's swath may improve temporal resolution. However, we think a uniform filter is more tangible. Especially given that the temporal resolution of our approach is higher than filtered LakeSP already. The 2nd assumption we take – using only open water – may particularly affect small lakes & narrow elongated lakes (expecting a higher fraction of water near land for such geometries). However, we believe that we can see that – for scales we investigated here – this poses no significant disadvantage. The smallest lake (Tivolisoen) shown in Figure 5A and the 'ill shaped' Valsølille Sø in Figure 2 display good metrics en par with the median of all 37 lakes. We will make sure to include this discussion in the revised manuscript.

Hydrological relevance and limitations of validating only relative WSE changes with a static mask

The Introduction motivates applications such as flood detection and small-lake monitoring, but the methodology primarily validates relative WSE changes and relies on a static lake mask for extracting PIXC data. This choice has important implications: a static mask can include land pixels during low-water conditions and exclude wetted areas during high-water conditions, potentially introducing systematic errors—especially when hydrological dynamics cause appreciable changes in lake extent.

Reply: thank you for raising this concern. We tried to address this concern in line 200-206 but see that it should be refined in a revised version. We think adding an additional analysis where we increase the buffer size to deliberately include dry land for some lakes will show what to expect from our proposed method in dynamic situations. Analogously for a buffer which is too small for the lake. We hope that this will improve generalizability and relevance of our study.

Moreover, the study does not attempt to infer changes in surface area, storage, or volume—metrics that are often more directly relevant for water resources management and flood/drought assessment than WSE alone. As written, this constraint represents the main methodological weakness, because it limits interpretability in dynamic conditions and may lead to systematic volumetric biases even if WSE variability is captured reasonably well.

Reply: We agree with you, that measuring volumetric changes would be an interesting and relevant study. However, we think this is not the purpose of our study. We hope that the above proposed artificially extended / narrowed buffers will indeed show what to expect from our approach in more varied and dynamic conditions than they are present in Denmark. Visualizing the WSE recovered from these 3 buffer sizes would provide the reader with an estimate of how a static mask affects WSE recovery.

I recommend that the authors more rigorously justify this design choice and explicitly discuss the hydrological implications and potential biases. At minimum, the manuscript should clarify the intended use-case (e.g., lake-level anomaly tracking vs. hydrologic accounting) and delineate when the approach is expected to be reliable or unreliable.

Reply: we agree that our manuscript will benefit from a clearer justification of design choices and a discussion of their implications. We plan to include some of the points we commented here above in the revised form of the paper and expand on these. Additionally, we agree, that we should make our intended use case clearer. Our intend is to look at lake-level anomaly tracking, for small lakes in a context like that of Denmark. We will reformulate our motivation to reflect this more clearly.

To address the reliability of out approach we plan to expand on the discussion and to include the additional analysis using varying lake mask buffer sizes.

If feasible, the paper would be substantially strengthened by either (i) a sensitivity analysis demonstrating the impact of mask mismatch across hydrologic states, or

Reply: Thank you, this is a good point. We hope, that by including the new analysis using varying buffer sizes we address you concerns about varying lake extents and different hydrological states. While not a complete sensitivity analysis, we will expand on the new results in the discussion section.

(ii) a discussion (and possibly a proof-of-concept) of strategies to incorporate dynamic water extent (e.g., time-varying masks or classification) and how this would affect WSE and downstream storage estimates.

Reply: We agree, this would greatly improve impact. Thank you for pointing this out. We plan to expand on out discussion section to include strategies how to incorporate dynamic water extent in out approach. And its downstream implications.

Structure and separation of Results vs. Discussion (Sections 4 and 5)

Sections 4 and 5 would benefit from clearer separation between Results and Discussion. Currently, interpretive statements appear interwoven with results, and the subdivision into multiple short subsections makes the narrative feel fragmented. I suggest consolidating the results into a more coherent sequence (e.g., overall performance metrics → sensitivity to filtering choices → lake-size/geometry effects → outlier/failure cases), followed by a Discussion that synthesizes the key findings, explains mechanisms behind observed patterns, and articulates limitations and practical guidance.

Reply: Agreed, we will restructure and rewrite the Results and Discussion sections. We hope that this will improve readability and ease of understanding.

The reasoning of fragmenting both sections into multiple small subsections was to separate the different points we show in the manuscript (filtering choices, result metrics, error sources, etc.) in a logical way. We can see though, that this fragments the story and hinders the reading flow. We will revert this. Thank you for pointing this out.

In addition, a concise discussion of the broader applicability of the proposed workflow beyond the Danish lake set (e.g., expected behavior in different geomorphologies, vegetation, turbidity, wind conditions, or hydrological regimes) would improve the manuscript's relevance to the wider HESS readership.

Reply: Agreed, we will expand the manuscript to include a broader discussion.

- the **geomorphology** can impact our approach twofold. (i) tall flanking topography to the sides of the lake can cause layover → this should be masked by the SWOT intrinsic DEM. It will affect the number of observations per lake and time stamp which inturn may cause our filter to reject more observations as the threshold might not be reached. (ii) significant elevation differences along the lake extent will cause problems for using a singular WSE aggregate for lake representation. Depending on where retained SWOT observations are sampled on the lake, a bias will be introduced. The impact on small lakes should be small however, but extremely elongated lakes in mountainous terrain should likely be excluded from out approach.
- For **ice cover** we expect a degradation of coherence leading to fewer reliable observation points. The general point observation quality should also degrade compared to open water. We don't know the scattering mechanisms for Ka band SAR over ice and snow. More research is required. (https://doi.org/10.21203/rs.3.rs-8454201/v1 - accepted and soon to be published in final version)

- For **submerged vegetation** and **turbid** water, we expect no impact on the WSE measurements. Ka band SAR is expected to not penetrate the water surface.
- **Surrounding vegetation** will hinder Ka band penetration to the water surface, but its impact should remain small as it is limited to the immediate shoreline. We would expect its impact to be analogous to a slightly too small lake polygon.
- **Wind** conditions will impact backscatter intensity. Initially we expect an increase in backscatter intensity for winds up to ca. 3 m/s which will fall again for wind speeds exceeding this threshold (doi: 10.1109/TGRS.2013.2258402) (interplay of surface roughness, Ka band wavelength and satellite look angle). Hence in high wind regime environments we expect worse data quality.
- Changing **hydrological regimes** may impact SWOT measurement. The static lake mask could cause error by including land. However, we expect the impact to be small as we (i) use SWOT's intrinsic water classification by using 'open water' only and (ii) using the median to aggregate a unified WSE measure per lake and time stamp – dampening the impact of outliers.
  At the same time, if seasonal WSE signal is highly variable, higher temporal resolution increases in importance, which is where our approach retains more time stamps than filtered Lake SP.

Overall recommendation

The study is technically robust, but it remains primarily a validation/benchmarking effort. I therefore consider it borderline for a full Research Article in HESS in its current form. It may be better positioned as a Technical Note in HESS, or alternatively as a Research Article if the authors (i) clarify and justify key methodological choices—especially the static mask and the focus on relative WSE—and (ii) strengthen the discussion of hydrological implications and generalizability.

Reply: thank you for your thorough critique of our manuscript. We agree with your general sentiment and hope that the comments we provided throughout this document will improve our study once implemented in the revised version.