

Author Response

We thank the both the reviewer for spending their time and effort reviewing our work. We have carefully considered all points and revised our manuscript. Below, we list the major changes we have done in this review.

Detailed responses to all points raised by the reviewers follow afterwards.

Major points:

- We re-ran our analysis with data up to date. Now all data is process baseline version D. Additionally, now, LakeSP data was available up to date, which made the comparison fairer, more streamlined and expressive.
- We completely re-wrote the results and discussion section. We removed all discussing elements in the results and broadened the discussion. We introduced several dedicated points regarding transferability and hydrological impact.
- We have made sure to highlight the advantage our approach brings. Namely retaining double, the observations compared to filtered LakeSP throughout the paper.
- We have added a new figure to show how performance metrics compare from our approach to LakeSP, visualizing lake surface area and max. WSE amplitude.
- We re-ran our analysis, artificially altering the lake mask by +20m / -20m to simulate changing lake extents to address concerns regarding transferability of the used static mask

REV 1

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

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 his careful reading and thoughtful suggestions, which we believe strengthened and clarified the manuscript. We address each of the points raised below.

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: Thank you. We added a sentence to the introduction, highlighting SWOTs advantage (Line 30f). "SWOT fundamentally transforms surface water hydrology by providing dense spatial sampling and temporally coherent observations of water surface slopes (WSS)."

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: Thank you. We added a paragraph on this in the data section in Line 70ff.: "SWOT sampling geometry is not uniform. The change of incidence angle across track is large, in relative terms, causing effective pixel spacings of ca. 60m to 10m from close to far range (Fjørtoft et al., 2014). However, as noise levels increase toward the far-range, near-range observations provide higher data reliability. The along track resolution is fixed at 22m (JPL, 2023b)."

- (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: Thank you. We added a new section to the discussion addressing the interactions of our method with the SWOT particularities (Section 5.1).

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. We have addressed your concern in a new section of our Discussion (5.3) where we show a new analysis extending / decreasing the lake buffers artificially by 20m. This mimics changing lake levels and uncertainties attached to the lake mask. We see that the metrics do not change considerably.

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 added discussion of Section 5.3 satisfies your concerns regarding transferability to other hydrological

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 added a cleared purpose statement in the introduction (line 50ff.) and hope that the added discussion points outline possible limitations of our approach.

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 for your point. We hope that varying the lake mask has sufficiently addressed your concerns.

(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: thank you for your suggestion. We have added a discussion on implementation of a dynamic lake mask in section 5.3.

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: thank you for your input. We have completely revised our Results and Discussion sections. We have particularly removed discussing elements from our Results and tried to have a better story throughout the Results by removing subsections. We broadened the subsections in the discussion and expanded particularly in the discussion of our filter choices and application to other hydrological domains.

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: Thank you. We have added a dedicated section in the Discussion to address this whole topic. We believe this improved the manuscript considerably. (Section 5.5)

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 have incorporated your suggestions and thoroughly reviewed our manuscript. We want to thank you for your input, we think it substantially improved the paper.

REV 2

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

This manuscript evaluates the use of SWOT L2 HR PIXC data to derive relative water surface elevation (WSE) time series for small Danish lakes and compares the results with the LakeSP product. The topic is timely and the validation dataset is potentially valuable. However, in its current form, the paper reads more like a processing/engineering report than a scientific study, and the broader significance of the proposed workflow is not yet convincingly demonstrated. Several methodological choices also require clearer definition and stronger justification. Below I outline major issues that should be addressed before the manuscript can be considered for publication.

Reply: The authors would like to thank the anonymous reviewer for their time and effort in providing their expert views and comments on the manuscript during the review process.

Major comments

1) Novelty and scientific contribution relative to LakeSP are not sufficiently established.

At present, the manuscript mainly documents a processing workflow and reports summary metrics for Danish lakes. The improvement over the filtered LakeSP product appears modest (median RMSE 5.34 cm vs. 6.05 cm; median PCC 0.91 vs. 0.89), which raises the question of whether the contribution is a conceptual advance or primarily an alternative filtering/aggregation strategy. The authors should explicitly clarify:

Reply: thank you for pointing this out. We have made sure to highlight the distinct advantage we see in our approach throughout the paper. Namely, that we retain double the amount of timeseries observations vs. LakeSP.

What is fundamentally new in the proposed approach (beyond parameter choices and filtering rules)?

Reply: thank you for your question. We have expanded on it in the Discussion (Section 5.2). But as stated in our initial reply, fundamentally we are using the same data as LakeSP. The main novelty is, that we can produce timeseries of WSE from SWOT with on average double the observations per lake.

Why this constitutes a meaningful advancement over LakeSP in terms of methodological design, robustness, or interpretability?

Reply: thank you. We have made sure to highlight the distinct advantage we see in our approach throughout the paper. Namely, that we retain double the amount of timeseries observations vs. LakeSP. We also discussed the added robustness of our approach as we do not use "dark water" (line 229ff.)

Under what conditions the proposed method is expected to outperform LakeSP (e.g., lakes with small seasonal amplitudes? high-latitude multi-pass coverage?), and why.

Reply: thank you for your comment. We have expanded on this in the discussion. Fx. in *“While the operational LakeSP product provides accurate WSE estimates with slightly better metrics, our filtering approach for the HR PIXC data preserves substantially more valid timestamps, thereby maximizing the temporal resolution necessary for robust local lake monitoring and other hydrologic applications. Especially for lakes with highly variable WSE patterns: even in Tivolisøen (Figure 3A) peaks in WSE (11.23-02.24; 03.24-05.24; 01.26-03.26) are completely missed by filtered LakeSP, but retained by our approach. This demonstrates that, despite similar metrics, our approach displays a clear advantage over LakeSP, given its higher temporal resolution.”*

Without a clearer positioning, the contribution currently appears incremental.

Reply: thank you. We hope that we adequately addressed your concerns above and throughout the complete rewrite process of the result and discussion sections.

2) Limited geographic scope and weak discussion of transferability/scalability.

The study focuses exclusively on Danish lakes. While this is a reasonable testbed, the manuscript does not clearly explain how the findings generalize to other hydroclimatic settings (e.g., arid regions, vegetated shorelines, tropical systems)

Reply: Thank you. We have added a dedicated section in the Discussion to address this whole topic. We believe this improved the manuscript considerably. (section 5.5)

or to regions with different viewing geometries and sampling densities.

Reply: Thank you. We have added a sentence in the new section of the Discussion “our Danish lake dataset samples across the entire swath, hence cross-track positioning does not limit transferability; instead, we see steep terrain remaining as the primary operational bottleneck due to topographic observation loss. “

In particular, it is unclear whether key selection criteria (e.g., ≥ 200 points per acquisition, strict geo-flag handling) are transferable to areas with fewer valid observations per pass, smaller lakes, or more heterogeneous shorelines.

Reply: Thank you. We have added a paragraph in the new section of the Discussion (Section 5.1 and 5.5).

3) Methodological description needs clearer formalization and stronger justification.

Despite the technical focus, several core steps are not described with sufficient precision to ensure reproducibility or to interpret why performance improves:

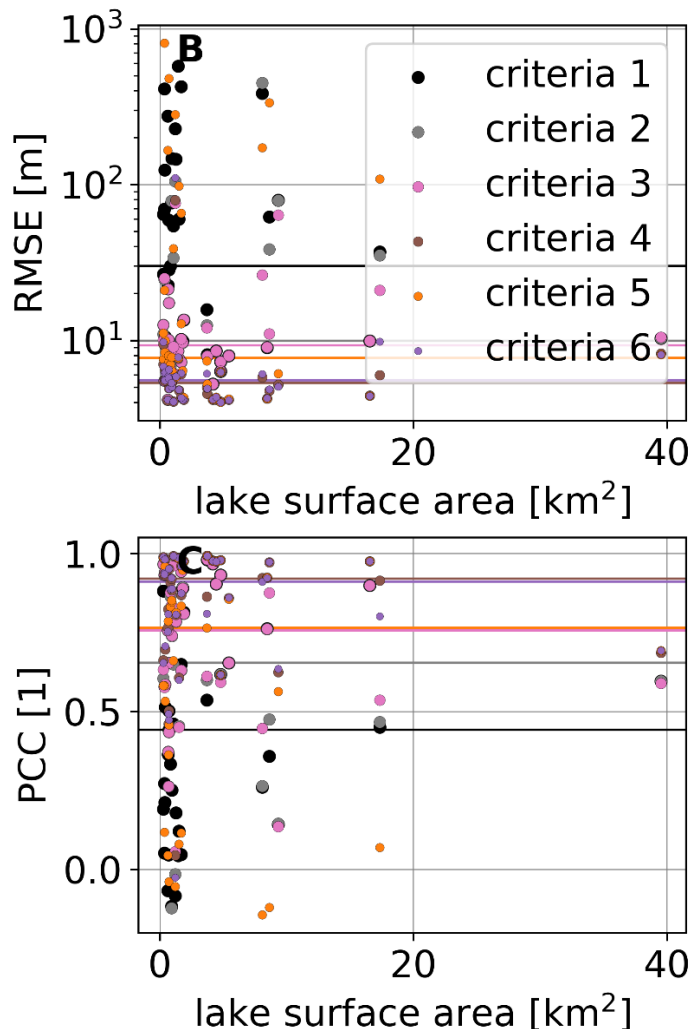
Reply: thank you for your comment. We aimed to improve the clarity of our statements in the revised manuscript.

The WSE aggregation procedure should be formally defined (e.g., summary statistic used, whether any weighting is applied, outlier handling rules, spatial screening, and how multiple water bodies within a mask are treated).

Reply: We have added a flow chart to the appendix. We hope that these aids understanding of the information that is present in the text. Additionally, we have tweaked our formulations in Section 3.3 of our method section.

The sensitivity of results to the ≥ 200 -point threshold should be assessed more explicitly (e.g., dependence on lake area, swath edge vs. near-nadir sampling, and seasonal/ice conditions).

Reply: Thank you for pointing this out. We added the new figure to the appendix and referenced it in the results section. We hope this addresses your point. Furthermore, we expanded on your point in the discussion in the new section 5.1.



The rule to exclude all geo-flagged observations (unless the point count drops below 200) needs a clearer statistical rationale. Why is this threshold optimal or conservative, and how robust is it?

Reply: Thank you. We rephrased your text slightly to avoid the impression that we “exclude geo-flagged observations (unless the point count drops below 200) “. New text: “The median is applied to all points as filtered by the selection criteria (either all available points or points without a geo-flag) “. Otherwise we would like to point to Table 1.

We addressed your point by adding “With this threshold, we deliberately accept less retained timestamps to trade off for ‘no’ outliers in our timeseries.” in section 3.3.

In addition, the roll-error discussion remains largely qualitative. If residual roll error is emphasized as affecting $\sim 18\%$ of acquisitions, the manuscript should provide a more systematic quantification approach (e.g., uncertainty envelopes, swath-position dependence, cycle-to-cycle variability), and discuss the implications for inland-water monitoring beyond Denmark.

Reply: thank you. We have added to the discussion on roll error that it is not our purpose to quantify but only to highlight its existence: “We are not trying to correct or quantify the residual roll error but only want to investigate its existence. Our set-up is not capable to robustly quantify it and perform an in-depth quantitative analysis of its amplitude. This is out of scope for our study and necessitates a dedicated research paper. However, our results clearly demonstrate its presence.”

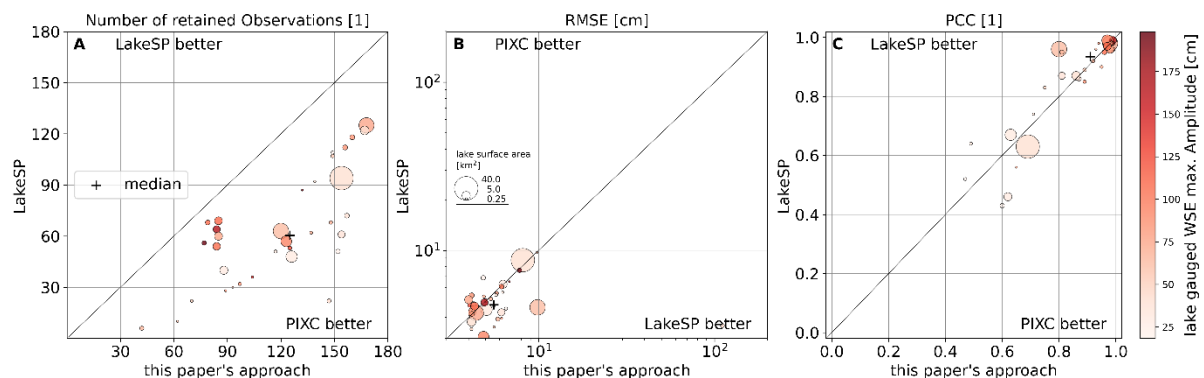
Additionally, we have expanded on the discussion of impact of the residual roll error: “Consequently, while investigators can likely ignore this effect for water bodies with large WSE variations, the tilt may influence the accuracy of WSE time series for lakes with small seasonal signals. More significantly, measurements of WSS could be superimposed with the residual roll error. Rivers perpendicular to a SWOT ground track would be particularly affected. We think, future SWOT users should maintain awareness of this potential error component.”

4) LakeSP comparison is not sufficiently diagnostic to support strong conclusions.

The manuscript argues that the proposed approach performs slightly better while retaining more timestamps, but the comparison does not yet isolate *why* differences occur. In particular:

The evaluation should include a more systematic statistical comparison (e.g., lake-by-lake paired differences in RMSE/PCC, confidence intervals/bootstrapping, performance stratified by lake size, WSE amplitude, swath position, and presence/absence of roll error).

Reply: Thank you. We have added a new figure to the results section comparing the metrics and retained timestamps for our approach and filtered LakeSP vs lake surface area and max WSE amplitude.



The paired differences are shown in the Table in the appendix. And we added on the discussion as to why differences occur (section 5.2).

“While both our approach and the baseline LakeSP product aggregate WSE values from the SWOT PIXC product, two methodological choices separate the pipelines. First, LakeSP includes “dark water” and “water near land” pixels, whereas we restrict extraction strictly to the “open water” class. Second, LakeSP dynamically adapts the initial Prior Lake Database mask instead of utilizing a static mask. Specifically, “dark water” observations often carry higher noise levels, and the dynamic mask can inadvertently expand over shoreline features – such as wet soil – misclassified as water (Figure A1F). By relying on a static mask and open-water pixels, our method circumvents these noise and misclassification pathways, reducing non-water artifacts and thus preserving a greater number of valid timestamps”.

The trade-off between operational simplicity (LakeSP) and configurability (proposed workflow) should be framed explicitly: in what use cases does the proposed method justify additional complexity?

Reply: thank you. We have improved upon our discussion in section 5.2:

“While the operational LakeSP product provides accurate WSE estimates with slightly better metrics, our filtering approach for the HR PIXC data preserves substantially more valid timestamps, thereby maximizing the temporal resolution necessary for robust local lake monitoring and other hydrologic applications. Especially for lakes with highly variable WSE patterns: even in Tivolisøen (Figure 3A) peaks in WSE (11.23-02.24; 03.24-05.24; 01.26-03.26) are completely missed by filtered LakeSP, but retained by our approach. This demonstrates that, despite similar metrics, our approach displays a clear advantage over LakeSP, given its higher temporal resolution. [...] For global use cases or applications, where the temporal resolution is not of the highest priority, but ease of use and computational resources, we see LakeSP as preferable. Quarrying LakeSP from Hydrocron takes seconds per lake. Quarrying commuted lake level values from our approach from a local database is comparable to that. However, computing the PIXC data takes about one to two minutes per SWOT pass over Denmark -- this has to be done only once.”

As written, the marginal performance gains make it difficult to justify the proposed workflow as a replacement or substantive supplement to LakeSP.

Reply: We have tried to workout the key advantage of our approach throughout the revised manuscript: that we have double the retained observations. We hope that this addresses your concern adequately.

5) Structure and framing should be revised toward a hypothesis-driven scientific narrative.

The manuscript would benefit from substantial restructuring to move from a workflow description to a clearer scientific contribution. Specifically:

Reply: we agree. We have revised our manuscript substantially.

State explicit research questions/hypotheses early (e.g., “Can PIXC-derived relative WSE time series for small lakes achieve cm-level accuracy, and what selection criteria are required?”).

Reply: we agree. Thank you for pointing this out. We incorporated this in the revised version. We are sure it improved the manuscript.

Separate methodology development, validation results, and implications for broader SWOT inland-water applications.

Reply: we agree. We have rewritten result and discussion section completely.

Strengthen the discussion to emphasize what the Danish case study reveals about SWOT capabilities/limitations for small lakes globally.

Reply: We hope that incorporating the comments we made here and in reply to reviewer 1 addressed your concern adequately.

Summary recommendation

The dataset and topic are valuable, and applying SWOT to small lakes is an important research direction. However, the current manuscript does not yet demonstrate a sufficiently strong methodological innovation or convincingly argue global relevance. I recommend **Major Revision**, focusing on (i) clarifying novelty relative to LakeSP, (ii) formalizing and justifying methodological choices, and (iii) providing a clearer framework for transferability and scalability beyond Denmark.

Reply: Thank you for your comments. We have tried to address them throughout the review of our manuscript and addressing them here. We think that your feedback considerably improved our work.