The study outlines the outcomes of an inter-comparison experiment aimed at estimating glacier elevation changes using spaceborne optical stereo (ASTER) and synthetic aperture radar interferometry (TanDEM-X) data. The study also emphasizes the importance of accurate glacier mass change observations for understanding climate change impacts, such as regional runoff, ecosystem changes, and global sea-level rise.

I have some general comments to further improve this manuscript. Please specify the range of results obtained from the various studies, indicating both the differences and similarities. Quantifying these differences can provide a clearer picture of the challenges in current methodologies. Clearly articulate the significance of the findings in terms of their implications for understanding climate change impacts, emphasizing how the observed discrepancies impact broader applications such as regional runoff, ecosystem changes, and sea-level rise. Also, mention in one line about the details on the validation process with airborne data, including specific metrics used and the degree of accuracy achieved. Specify the level of community involvement in the experiment and how collaborative efforts contributed to the study. This can highlight the strength of community-based research. Encourage the adoption of transparent practices within the scientific community. Provide a concise summary of the major findings and contributions at the beginning or end of the abstract for better reader orientation.

- Elaborate on the methodologies used for bias corrections, co-registration, outlier filtering, void filling, radar signal penetration, and temporal corrections.
- Consider including a flowchart or a schematic representation of the data processing steps to enhance the clarity of the methodology section.
- Expand on the comparative analysis between optical and radar data, addressing the limitations and methodological differences. This will provide a more nuanced understanding of the challenges and strengths associated with each data type.

I believe by incorporating these suggestions and comments will refine the manuscript, making it clearer on the study's findings, their importance, and how collaboration shaped the research, thus making it suitable for publication.

Specific comments for each section are.

Introduction

- I will be better to begin by clearly stating the motivation behind the study and the specific research gap it aims to address. This could involve highlighting the importance of accurately assessing glacier mass balance for understanding climate change impacts and informing adaptation strategies.
- Provide a more detailed overview of the existing methods and techniques used to assess glacier elevation changes. This could involve discussing traditional mapping techniques, topographic surveys, and the role of digital elevation models (DEMs) derived from different sensor platforms.
- Emphasize the significance of spaceborne optical and radar sensors in revolutionizing glacier monitoring efforts. Discuss the advantages and limitations of optical and radar data for observing glacier elevation changes, including their spatial and temporal coverage, resolution, and sensitivity to environmental conditions.

- Introduce the ASTER and TanDEM-X satellite missions more prominently, including key characteristics such as sensor type, temporal coverage, and data processing capabilities. Briefly discuss the evolution of these missions over time and their contributions to glacier research.
- Clearly outline the objectives and methodology of the inter-comparison experiment described in the paper. Explain why comparing results from ASTER and TanDEM-X data is important, and how this contributes to improving our understanding of glacier elevation changes.
- Add some more relevant references such as Kääb et al., 2012; Bolch et al., 2011; Scherler et al., 2011; IPCC. (2021)

Data section

Ensure that the description of each component of the data is clear and concise. Provide sufficient details to allow readers to understand the study sites, airborne validation data, DEMs (ASTER, TanDEM-X, SRTM), and auxiliary data without overwhelming them with unnecessary information.

- Define any acronyms or specialized terms used in the description of the data to avoid confusion.
- Offer some context for why each type of data was selected for the study. Explain the rationale behind choosing specific study sites, DEMs, and auxiliary data, highlighting their relevance to the research objectives.
- If there are any unique aspects or novel approaches in the data collection or processing, highlight these to emphasize the original contributions of the study.
- Ensure that proper citations are provided for any previously published data sets or auxiliary data used in the study. This helps readers to trace the origins of the data and provides credit to the original sources.

Study Area

- Explain in more detail the criteria used for selecting the study sites, such as glacier size, topography, location, and availability of validation data. This will help readers understand why these specific sites were chosen for the inter-comparison experiments.
- Clearly articulate how each study site poses various data processing challenges for both optical and radar sensors. Provide specific examples of these challenges and explain their implications for DEM differencing.
- Where possible, provide supporting evidence or references to previous studies that have documented the challenges faced at each study site.
- Emphasize any unique characteristics or features of each study site that make it particularly interesting or challenging for the inter-comparison experiments. This could include factors such as glacier morphology, climate conditions, or geographical location.

Glacier outlines

- Clearly state the purpose of using glacier outlines from the Randolph Glacier Inventory (RGI) version 6.0 in the study. Explain how these outlines serve as a common reference

for comparison purposes and facilitate standardized analyses across different study sites.
- Acknowledge the limitations of using fixed glacier outlines, particularly in regions where glacier area changes significantly over time. Explain how this may introduce bias into the calculated specific elevation changes, especially in dynamic glacier environments such as the Alps.
- Maintain consistency in naming and abbreviations when referring to the RGI and its version number throughout the section.

Airborne validation DEMs

- Clarify why validation data from airborne lidar and aerial stereo images are important for the study. Explain how these data serve as ground truth measurements to assess the accuracy of spaceborne DEMs and validate their performance.
- Provide more detailed information about the acquisition dates, resolution, and sources of the airborne lidar and aerial stereo DEMs for each study site. This will help readers understand the characteristics and quality of the validation data used in the analysis.
- Describe the steps taken to process and prepare the airborne validation DEMs for comparison with spaceborne DEMs. This could include details on resampling, co-registration, and quality control procedures to ensure consistency and accuracy in the validation data.
- Discuss any limitations or challenges associated with the airborne validation DEMs, such as data gaps, varying acquisition dates, or potential artifacts. Address how these factors may impact the interpretation of validation results and the overall accuracy assessment.
- Include information on the accuracy of the validation data, particularly regarding the accuracy of elevation change measurements. Discuss how uncertainty was quantified and its implications for the interpretation of validation results.
- Acknowledge the absence of validation data for Baltoro and Northern Patagonian Icefield and discuss any potential implications for the study. Consider suggesting future research directions or alternative approaches for validating DEMs in these regions.

2.5 Spaceborne experiment DEMs

2.5.1 ASTER DEMs

- Clarify the significance of using ASTER DEMs for the study and explain how they contribute to the overall objectives. Discuss why ASTER data were chosen, emphasizing their global coverage and multi-temporal nature.
- Describe the process of acquiring and processing ASTER L1A images to generate DEMs for each study site. Provide information on the specific algorithms and parameters used for DEM generation, including any preprocessing steps or data projections applied.
- Discuss the quality of ASTER DEMs, including reported precision and potential limitations such as cloud cover, sensor saturation, and acquisition footprint. Address how these factors may affect the accuracy and spatial coverage of the DEMs, particularly in glacierized areas.

- Quantify the spatial coverage of ASTER DEMs for each study site, including the number of available DEMs and the variation in data points per glacier pixel. This will provide readers with a clear understanding of the spatial distribution and density of DEM data.
- Discuss the main challenges associated with using ASTER DEMs, such as incomplete spatial coverage and voids in the accumulation area. Provide insights into potential strategies or techniques used to address these challenges, such as interpolation or filtering methods.

2.5.2 TanDEM-X DEMs

- Clarify the significance of using TanDEM-X DEMs for the study and explain how they contribute to the overall objectives. Discuss why TanDEM-X data were chosen, emphasizing their high resolution and near-complete coverage of the glacier areas.
- Describe the process of acquiring and processing TanDEM-X data to generate DEMs for each study site. Provide information on the interferometric workflow, including steps for concatenating overlapping scenes, creating interferograms, and unwrapping phase data.
- Discuss the reported precision of TanDEM-X DEMs and any limitations or potential sources of error associated with their use. Address how accuracy may vary in mountainous terrain and how errors are mitigated during data processing.
- Quantify the spatial coverage of TanDEM-X DEMs for each study site, including the percentage of coverage on the glacier and the presence of voids or masked areas off-glacier. This will provide readers with a clear understanding of data availability and potential limitations.
- Include visual aids, such as figures or maps, to illustrate the spatial distribution of TanDEM-X DEMs over the study sites. This can help readers visualize the extent of data coverage and identify areas with potential data gaps or errors. Discuss the temporal coverage of TanDEM-X DEMs and how it varies between study sites due to the acquisition strategy of the mission. Address any challenges or limitations associated with the campaign-based acquisition mode and variable acquisition months.

3 Methods

3.1 Inter-comparison experiment description

- Describe in more detail the different strategies, post-processing steps, and corrections applied by participants to calculate their spaceborne estimates. This could include specific methods used for DEM selection, filtering, co-registration, and uncertainty estimation.
- Provide clearer explanations of the two experiments conducted as part of the inter-comparison study. Clearly state the objectives, target periods, and study sites involved in each experiment. This will help readers understand the specific goals and scope of each experiment.
- Discuss the rationale and methodology behind the temporal corrections applied in the first experiment, particularly for aligning spaceborne estimates with airborne validation

data. Explain how temporal discrepancies were addressed and why they are important for the accuracy of elevation-change estimates.

- Provide more details about the sensitivity study conducted in the second experiment, including the specific processing steps evaluated and their potential impact on glacier elevation change estimates. Discuss the significance of this sensitivity analysis for understanding the robustness of results.

3.2 Participants and spaceborne results

- Discuss the variability in the number of submissions for each study site. Explore potential reasons why some sites received more submissions than others, such as site characteristics, data availability, or participant preferences.
- Provide more detail about the different DEM sources and approaches employed by participants for each study site. Discuss any notable differences or trends in the types of data sources and methodologies used.
- Highlight the prevalence of ASTER DEMs over TanDEM-X DEMs in the submissions. Discuss potential reasons for this discrepancy and any implications it may have for the interpretation of results.
- Provide context for why Hintereis and Baltoro were chosen to illustrate the experiments in the main body of the manuscript. Discuss any unique characteristics or significance of these sites compared to the other study sites.

3 General workflow

- Expand on the description of each step in the workflow to provide more detail about the specific processes involved. For example, explain the methods used for bias correction, co-registration, noise filtering, and void-filling, and discuss the rationale behind each process.
- Highlight any key differences in approach or methodology used by different groups.
- Provide examples or case studies to illustrate how the workflow steps were applied in practice for specific study sites.
- Discuss the various types of corrections that can be applied in the workflow, such as temporal corrections to match the target period or corrections for radar signal penetration. Explain the purpose of each correction and how it contributes to improving the accuracy of the elevation change estimates.
- Encourage readers to refer to supplementary tables for more detailed descriptions of the DEM selection and processing strategies used by each group for each site. Provide a brief overview of the information available in these tables and its relevance to understanding the workflow.

3.3.2 Spatial bias correction

- Provide a clearer definition and explanation of the spatial trends and vertical deformation mentioned. Explain how these biases occur due to sensor behaviour, data acquisition, and processing. Specify the magnitude and extent of the biases more precisely, including the range of horizontal scales affect. Clarify the concept of "height of ambiguity" and its role in generating the observed linear trends in TanDEM-X DEMs..

- Provide more detail on how each correction technique works and why it is effective. Explain the rationale behind fitting planes, polynomial functions, or sine functions, and how they address specific biases.
- Include references to relevant literature to support the choice of correction techniques and their effectiveness. Provide proper attribution for the correction techniques mentioned, citing the original sources or key references where these methods were introduced or validated.
- Discuss any quantitative assessments or validations of the correction techniques used. Provide details on how the effectiveness of each technique was evaluated and any metrics used to measure improvement in DEM quality. include statistical analysis or comparisons to demonstrate the impact of the corrections on reducing biases. Discuss the broader context of spatial bias correction in DEMs, including common challenges, emerging trends, and areas for future research.

3.3.3 DEM co-registration

- Start by clearly stating the importance of DEM co-registration in minimizing systematic errors before performing DEM differencing. Explain how shifts and rotations resulting from georeferencing techniques and processing distortion can impact the accuracy of DEMs.
- Discuss the variability in co-registration approaches adopted by different groups in the experiment. Explain the reasons behind the selection of specific stable terrain areas and reference DEMs.
- Describe the co-registration algorithms used by the groups, including specific references to equations or methodologies.Discuss the strengths and limitations of each algorithm in addressing shifts, rotations, and scaling between DEMs.
- Describe how remaining biases after co-registration were addressed, including any criteria used to determine whether biases were significant enough to warrant correction.

3.3.4 Noise filtering and void-filling

- Start by providing a brief overview of the types of artefacts and noise commonly found in ASTER and TanDEM-X DEMs, along with their causes. This will help readers understand the significance of noise filtering and void-filling.
- Clearly explain the sources of artefacts in ASTER and TanDEM-X DEMs, including sensor characteristics, surface properties, and topographic factors. Provide examples of artefacts such as sinks, phase unwrapping errors, shadowing, and layover.
- Clearly explain the sources of artefacts in ASTER and TanDEM-X DEMs, including sensor characteristics, surface properties, and topographic factors. Provide examples of artefacts such as sinks, phase unwrapping errors, shadowing, and layover.
- Provide a detailed explanation of the noise filtering approaches adopted by the participants, including the use of statistical parameters like standard deviation, NMAD, and threshold values. Clarify how these approaches help remove gross errors and artefacts from the DEMs.
- Highlight any novel or innovative approaches used by certain groups for noise filtering and void-filling. Discuss the advantages and limitations of these approaches compared

to more traditional methods. Ensure that relevant references and citations are provided to support the discussion of noise filtering and void-filling techniques.

3.3.5 Radar penetration correction

- Clarify the significance and implications of radar penetration correction in the context of glacier studies. Elaborate on the methods employed by different groups to address radar penetration. Provide a clearer explanation of how radar penetration correction was implemented in the study.
- Explain how inaccurate correction can lead to biased results and affect the reliability of glacier mass balance assessments. Emphasize the role of penetration depth variability in influencing the accuracy. Provide a detailed explanation of the probabilistic framework employed by the UST group and how it accounts for elevation-dependent penetration using a Gaussian probability distribution. Similarly, explain the specifics of the elevation-dependent C-band penetration model applied by the GAC group, referencing the findings of Kumar et al. 2019 of InSAR measurements.

3.3.6 Temporal corrections

- Begin by clearly stating why temporal corrections are necessary when comparing different datasets, especially when spaceborne observation dates differ from the airborne validation period.
- Provide a more detailed explanation of the various strategies employed by participants for temporal corrections. Describe each strategy (no temporal correction, linear scaling, annual corrections using glaciological observations, non-linear regressions) and clarify the assumptions underlying each approach.
- Discuss the significance of the remaining temporal differences after applying corrections. Explain why these differences occurred despite the correction efforts and discuss their potential impact on the accuracy of the elevation change estimates. This discussion will provide insights into the limitations of the correction methods and the challenges involved in achieving precise temporal alignment.
- Provide a clear explanation of the approach used by Zemp and Welty (2023) to correct for the remaining temporal differences. Describe how seasonal observations of glaciological mass balance were temporally downscaled using sine functions and how this adjustment was applied to spaceborne elevation change results.
- Quantify the effects of the corrections on the temporal differences between spaceborne elevation changes and validation dates. Provide specific examples or statistics to illustrate the magnitude of the corrections and their impact on the elevation change estimates.

3.4 Uncertainty assessment

- Provide a clearer explanation of how the overall uncertainty was calculated by the participants. Describe the process of quadratic summation of different error sources and how the 95% confidence interval was derived.

- Expand on the different error sources considered by participants and the methods employed to quantify them. Provide additional context and explanation for each error source, including pixel elevation change error, errors in glacier outline and area mapping, errors due to missing observations, and errors related to temporal mismatch or radar penetration correction. This will help readers understand the specific challenges addressed by each group and the methodologies used to mitigate them.
- Discuss the variation in error quantification methods employed by different groups. Highlight any differences in approaches, assumptions, or model parameters used for error estimation.

4 Results

- Provide additional clarification on why participants preferred using the provided DEMs rather than self-processed DEMs. Elaborate on any advantages or limitations of using provided DEMs and discuss how this choice may have influenced the results.
- Expand on the correction of radar signal penetration by two groups for the Baltoro site and the absence of such correction for the Northern Patagonian Icefield. Discuss the implications of this difference in correction approaches for the accuracy and reliability of the results. Additionally, provide context on the challenges or limitations of radar signal penetration correction and how participants addressed them.
- Discuss the limited application of temporal corrections by participants and the relatively low participation in sensitivity studies. Provide insights into the reasons for these trends and discuss their implications for the reliability and robustness of the results. Additionally, highlight any key findings or observations from the sensitivity studies that were conducted.
- Provide clarity on the availability of spaceborne results and related processing steps for each study site. Ensure that readers understand where they can access this information and how it can contribute to their understanding of the study findings. Consider providing direct links or references to the tables containing this information for each site.

4.2 Elevation change assessment for glaciers with airborne validation data

- Provide a brief overview or introduction to the first experiment, outlining its objectives and the glaciers studied (Hintereis, Aletsch, and Vestisen). This will help orient readers and provide context for the subsequent discussion of the results.
- Offer a more detailed explanation of the observed elevation change patterns for each glacier. Describe the main findings regarding glacier-wide mean elevation changes, trends in ice loss, changes in accumulation areas, and any notable spatial variations. Providing this information will help readers understand the key findings and interpretations derived from the comparison of spaceborne results with airborne validation data.
- Discuss the accuracy and reliability of the spaceborne results compared to the airborne validation data. Address any discrepancies or differences observed between the two datasets, including variations in spatial coverage, noise levels, and patterns of elevation change. Consider discussing potential sources of error or uncertainty in the spaceborne results and their implications for the interpretation of the findings.

- Discuss any insights gained from comparing spaceborne results with airborne validation data and highlight areas for future research or improvement in methodology. This will provide closure to the discussion and guide readers on the significance of the results presented.

4.3 Regional-scale elevation-change assessment and sensitivity study

- Provide a brief introduction or context for the second experiment, outlining its objectives and focus on regional-scale elevation change assessment for Baltoro and the Northern Patagonian Icefield.
- Clarify the main focus of the second experiment, which was to analyse the effect of various processing chain steps rather than comparing spaceborne results with airborne validation data. Explain why this approach was chosen and how it contributes to understanding the sensitivity of elevation change estimates to different processing methods.
- Offer a detailed description of the observed elevation change patterns for both Baltoro and the Northern Patagonian Icefield. Discuss any notable features or trends identified in the elevation change rate maps, such as surge-type patterns, spatial variability, and differences between observation periods. Providing this information will help readers visualize the results and understand the variability in elevation change estimates.
- Discuss the spread of calculated elevation differences and mean elevation change rates for both study regions and observation periods. Explain the significance of the variability in results and any trends observed between the two periods. Address any factors contributing to the spread in results, such as differences in input data sources and processing techniques.

Discussion

- Address any additional considerations or issues that were not covered in the experiments but may have a significant impact on geodetic glacier mass change assessments, such as changes in glacier area and density conversion factors. Discuss the implications of these factors for interpreting elevation change estimates and comparing them to other sources of data.
- Provide clear and concise explanations of the findings presented in the results section, highlighting the main conclusions drawn from the analysis. Use specific examples and references to support your arguments and illustrate key points.
- Offer thoughtful interpretations of the findings and discuss their implications for glacier elevation change assessment using spaceborne data. Consider addressing questions such as the reliability of spaceborne estimates, the significance of systematic differences between approaches, and the potential impact of processing steps on the accuracy of elevation change estimates.
- Compare and evaluate the different processing steps and approaches used by the participating groups in the experiments. Discuss the strengths and limitations of each approach, as well as any observed trends or patterns in the results.
- Address any additional considerations or issues that were not covered in the experiments but may have a significant impact on geodetic glacier mass change assessments, such as changes in glacier area and density conversion factors. Discuss

the implications of these factors for interpreting elevation change estimates and comparing them to other sources of data.

- Highlight areas for future research and methodological development based on the gaps and limitations identified in the experiments. Discuss potential avenues for further investigation, such as the development of new processing techniques, the integration of additional data sources, and the exploration of advanced uncertainty quantification methods.
- By summarizing the main insights gained from the inter-comparison experiments and discussing their broader implications for the field of glacier remote sensing. Offer final reflections on the significance of the findings and their relevance for advancing our understanding of glacier dynamics and climate change.

Conclusion

- Begin the conclusions section by summarizing the main findings of the study, focusing on aspects such as the large spread in elevation change estimates, the challenges associated with individual results from spaceborne surveys, and the potential of ensemble approaches to reduce random errors.
- Discuss the challenges and limitations associated with uncertainty assessment and validation of elevation change estimates. Address the variability in reported uncertainties and the importance of using high-quality DEMs for validation purposes. Highlight the need for a coordinated community effort to compile a benchmark dataset of validation DEMs for selected glaciers worldwide.
- Provide recommendations for future research based on the findings of the study. Discuss potential avenues for further investigation, such as conducting more extensive inter-comparison experiments, exploring the impact of different processing steps at the glacier scale, and developing common good practices for uncertainty assessment and material density assignment.