Replies to reviewer 1 comments in calibri italic

General comments

I appreciate the effort that the authors have put into the revision of their manuscript, which has improved significantly. Many aspects are now more comprehensible, not least to the restructuring of the text and improved figures. Aside from minor and technical comments that were omitted during the first round of review, a few important points remain that need to be clarified for the reader.

Thank you for your review and input, we tried to address all issues in the revised version and will clarify the remaining points within this round.

Presentation of the instrumental setup

The presentation of the new cryo-LA-ICP-MS setup has been improved and includes now several important pieces of information. At least one optical mosaic that the authors show in the response should be included in the main text, e.g. adding to Figure 3. It is important to see the ablated laser path in this detail.

We already provided a close-up of the laser path in the revised Fig. 3. We added an overview mosaic picture in the appendix (Fig. A1) to not disproportionately extend the main manuscript.

To provide an example of why this is important: With regards to the difficulty in achieving ablation, the authors write in their response "This might be due to the finely microtomed and thus much smoother, more reflective and less uneven surface of the samples compared to other systems." First, the overview pictures do not allow to judge surface roughness adequately, e.g. in the comparison included in the response using the image taken from Bohleber et al. (2023). The latter was also prepared by a machined microtome prior to cryo-Raman analysis at which point the overview image was taken (Stoll et al., 2022).

We agree that this particular picture comparison might not have been the best choice to judge the surface roughness. However, over the course of three years of laser ablation measurements on many different samples we found some empiric evidence for the correlation between surface roughness and ablation efficiency. This is not yet quantifiable.

Second, there is no evidence for the surface roughness increasing ablation and, in a purely speculative way, one could also argue in the opposite direction: More roughness leading to the laser being out of focus more often, etc.

We agree that there is no theoretical evidence for the efficiency of ablation depending on the surface roughness. Roughness on a scale that would lead to the laser being out of focus would indeed result in lower ablation efficiency. But roughness on a scale below the laser spot size could have the opposite effect.

Ice is highly transparent at 193nm, the surface thus hardly reflective. At present the high fluence needed remains enigmatic but shows how little is known about the laser-ice interaction at UV-wavelengths.

Not just the laser-ice-interaction in general but also the apparent dependence on the different ice features (impurity concentration, opacity, crystal size and orientation (?), bubble content etc.). This also shows the need to compare more laser ablation measurements across different systems to better understand the mechanisms of ice ablation.

We can probably agree that this is interesting and hopefully it can be cleared up in the future with more dedicated experiments.

Agreed. Future experiments to investigate this phenomenon are planned.

The difficulty in coupling leads to the need to adjust the ablation settings, including fluence and repetition rate. The authors argue that this is caused by the eventual wear and tear of the laser optics. It should be clarified in the text if the fluence was actually measured at the sample level, where this degradation of the optics would manifest, or if the fluence reported corresponds to the nominal fluence settings in the software. Table 2 says "laser fluence on target" – was this measured at the target? Otherwise this should be "nominal laser fluence".

The fluence was not measured directly at the target but refers to the nominal laser fluence reported by the software. The term in Table 2 and throughout the manuscript was corrected accordingly.

The mosaic shows much better than the present close up in Figure 3 how deep the ablation craters are – and shows that they are consistent along the line as far as one can tell. It would be of interest to include the ablation settings used alongside this image, especially the repetition rate, which increases the overlap of spots fired and hence the ablated volume or crater depth.

We included the respective settings in the figure caption of Fig. A1. They were: Scanning speed: 50 µms^{-1}, nominal Laser fluence: 6.7 J cm^{-2}, Repetition rate: 120 Hz. Spot size: 150 µm round

The experiment for determining the washout time (or single pulse response, SPR) now nicely illustrates the performance of the system. However, as remarked in the first round, it is puzzling why the ablation settings were again different from the settings actually used in the experiments on ice?

We did not use the same settings as for ice on the NIST standard because this would have caused way too much material being ablated from the glass standard and therefore the ICP-MS signals going into saturation. The results would not have been meaningful.

Compared to ice analysis in the SPR experiment fluence was lower, 4 Jcm-2 vs 6 – 6.7 Jcm-2 , and spot size smaller, 30 µm vs 150 µm. It has been shown that both parameters matter for determining the SPR (e.g. Jerše et al. 2022). More material will generally need more time for transport, and a 30µm spot ablates only a small fraction of the area of a 150µm spot.

We agree that there might be a general relation between peak shape and beam size. However, Jerše et al. 2022 state that the appearance of double peaks is additionally strongly dependent on the material matrix and the generation of large particles. The matrix of the NIST standard is very different to glacier ice. The effects on the SPR are not quantified in Jerše et al. 2022. Since in this study it was not possible to perform the SPR experiment on a homogeneous ice matrix standard, we argue that the NIST experiment with the reported settings gives the currently best estimate for the SPR of our setup. We added a paragraph (L 161 ff revised manuscript) weakening the statement about the determined washout time.

The SPR value is further on used to determine the spatial resolution. This is an important number that appears in several central places throughout the manuscript. The authors also write in line 177 "Therefore the acquisition time is the dominating factor regarding depth resolution compared to the much smaller washout time." – this can only be stated if we know the SPR for 150µm.

As stated above, we can only use the results SPR determined with the 30 µm spot size and 4 Jcm^{-2} as a best estimate. The acquisition time of the ICP-MS is with 500 ms almost double the measured SPR times (198-270 ms). Even if the peak shape would change for the ice ablation, we doubt that this would cause an increase of the SPR by a factor of two or more.

This is an important issue that needs to be addressed, either via a clear explanation to the reader that the SPR for 150µm remains unknown and likely underestimated by the 30µm experiment or preferably by determining this value experimentally.

We recognize that it cannot be proved entirely that the acquisition time is larger than the washout time, but we still consider it very likely (see comment above). We changed the sentence in line 177 (L 185 ff revised manuscript) accordingly.

Please also adjust the statement in the abstract which reports "depth resolution of down to 80 µm".

Done. Changed to 182 µm

The spikes in the background signal and their subsequent removal remains an important issue to clarify further, especially with respect to the discussion about a potential imprint of grain boundaries (see below). The authors write in line 209: "The origin of the sharp isolated, high peaks is not entirely clear." And then "Therefore, we conclude that the very high frequency signals and sharp peaks in intensity are most likely caused by electronic interference and should be regarded as analytical noise." – There is no evidence provided for thsis effect being electronic interference to support this conclusion. More importantly, based on the statement in lines 215 – 217, it is not fully clear if peaks were removed from the actual data acquired on ice. If so, how would this affect a localized high intensity spike potentially caused by a grain boundary?

We reformulated the paragraph (L 216-223 rev. manuscript) and clarified the statements. The sharp, high, single point excursions in the data were regarded as noise and removed from the data during background correction (L 221f rev. manuscript). We found that the crossing of grain boundaries causes peaks much wider than just one datapoint (see discussion Fig. 11). The identification of these features should therefore not be affected by the background noise removal.

Layer detection in the Skytrain ice core

If a seasonality is established for LA-ICP-MS Na it makes sense to look for such periodicities in deeper ice, using spectral analysis if not evident otherwise. However, the imprint of grain boundaries needs to be demonstrated more clearly than what is presently discussed around the new Figure 11. It is generally important to clarify how the authors regard the high frequency signal components in their data. Are they considered noise, like for the blank ice, or are they a signal that has a physical origin in the ice, e.g. from grain boundaries?

As stated above, only the very high, single point excursions were regarded as noise and removed from the data. Still, some high frequency variations in the depth domain remain after this outlier removal.

Figure 11 relates to the aforementioned issue of outlier removal. In the data shown here, were high peaks removed as outliers as for the ice blank data?

Yes. As described above, the very high single point outliers were removed from the data for all samples.

In line 304 the authors write "The main challenge of identifying small scale features in the LA-ICP-MS data, which can be interpreted as layers, is to separate the signal of interest from the high frequency, noisy background." – this statement suggests that the noisy background illustrated for the blank ice still exists?

This statement was misleading, the background removal of high single, spikes was done for all samples before further (e.g. spectral) analysis. We reformulated the sentence (L 309 rev. manuscript).

Then, in line 307: "When the laser ablation path crosses a grain boundary, the enhanced concentrations can cause sharp peaks in the LA-ICP-MS data which could then lead to misinterpretation of grain boundaries as larger scale features." – this statement argues that there is a physical origin of these signals.

As stated above, after background removal of the single datapoint excursions, some sharp peaks and steep excursions remain. The sentence was reformulated (L 310 rev. manuscript).

And in line 328: "Superimposed to the large trend the smoothed signal still shows high frequency variability. These small scale variations are most likely caused by either (i) particles that show up as sharp peaks in the ICP-MS signal (ii) actual small scale intensity variations e.g. caused by accumulation of Na along crystal grain boundaries."
The latter statement sounds similar to line 307 but is speculative, because there is no direct evidence for (i) or (ii). Figure 11 is interpreted as little evidence for Na at grain boundaries (line 370, 399).

We reformulated the respective paragraph and emphasized the hypothetic nature of the statement (L 333 ff rev. manuscript)

I find the hypothesis that localization of elements at grain boundaries – if present – produce detectable signals in the frequency domain still not convincing. Grains would need to be highly uniform in size to generate periodic signals, which is not the case – all visual images included in the manuscript show highly variable grain sizes. This needs to be addressed better and rephrased (e.g. line 430).

We agree that there is some variation and range in the grain size. And that it is unlikely that a certain grain size will show up as frequency in the PSD. However, we know about the effect of impurity concentrations in the boundaries and that this will have subsequent impacts on the overall impurity distribution of the sample. This will in return influence the periodicity. We therefore still regard the ranges of the PSD that cover the span of the grain sizes as influenced by these effects. We added a table documenting the average grain sizes and ranges for all samples in the Appendix (Table C1). We reformulated the statement in L 434 rev. manuscript

I suggest removing indicating the grain size in the PSD plots, as this suggests that such periodicities could be expected.

We would not like to remove the grain size ranges from the PSD plots. They function as a guidance for the reader indicating the mean and range of the grain sizes and therefore highlighting the period lengths that might be influenced by these effects. We do not discuss direct connections between the periodicities and the grain sizes, only ranges.

The parallel line approach is much more convincing for reducing grain boundary related signals. (Figure 9). This is why I find it confusing that in Figure 11, three lines were stacked specifically to discuss grain boundary imprints. Stacking would make the grain boundary signals much weaker, not stronger. How does this comparison look for the single line profiles?

We used the stacking method because the three lines were very close together, crossing the same grain boundaries at almost the same positions. We hereby expected the stacking to enhance the common features, but that might have been misleading. We changed Fig. 11 and now only show a single line comparison and additionally the two other lines in the Appendix (Fig. B2).

In their response the authors write that grain boundaries are observed not wider than 5µm. At a resolution of more than 180µm (but see comment on SPR above), the signal contrast caused by a grain boundary could be too weak to detect, depending on the actual impurity concentration in the boundary. This point could be worth mentioning.

We agree in principle. Nevertheless, we observed the same phenomenon (some grain boundaries showing a very distinct signal, some showing none) during 2D mapping (data not shown in the paper) with much smaller spot sizes (50 µm) and larger scanning speeds (100 µm/s). Therefore we do not believe that the larger spot size combined with a slower speed and thus longer dwell time on the grain boundary would smooth out the signal completely. It would be interesting to analyse ice samples from different origins, impurity contents, microstructure on the same laser ablation system, to investigate if this is an observation unique to Skytrain ice samples. Such analyses are planned for the future.

Line 235 and Figure 6: Please explain how the grey vertical lines for the annual layers were identified in the CFA. The text mentions CFA-Ca, but in Figure 6 only Na shows a clear oscillating pattern. Interestingly the correlation with Ca is weak at best, e.g. the big peak around 83.8 m was skipped in counting and at 83.95m a Ca peak is missing but a grey line is set?

As elaborated in Hoffmann et al. 2022 the annual layer identification was done using a combination of the Na and Ca CFA signals together with absolute time markers (e. g. volcanic eruptions). The variations in the higher resolution Ca data were used to complement the Na data. More details can be found in the respective paper.

I suggest changing the statement of line 238 ff. “This indicates that the LA-ICP-MS technique is capable of identifying annual signals in this depth of the core, which would not be visible in the Na signal of the CFA alone.” Based on Figure 6, Na is the most plausible indicator of annual signals, and the LA-ICP-MS agrees well with CFA.

Yes, in general it agrees well but still shows some higher frequency variations compared to the CFA data (e. g. between 83.45 and 83.6 m). We will therefore refrain from changing that statement.

Minor and some technical comments that were skipped in the first round of review

Line 51: Reinhardt et al. (2003) used an IR laser, not UV.

Added IR in Line 52 rev. manuscript

Line 53: The main difference among the systems is also the design of the ablation chamber: Two volume vs. single volume, fast vs. slow washout, etc.

Added a respective sentence L 55f revised manuscript.

Line 55: It is unclear what is meant by “laser cell” – ablation chamber?

Changed to ablation chamber.

Table 2: Helium flow: Does an inner and an outer volume exist that has separate flow rates?

No, only one helium flow is adjustable.

Line 117: Here it is “ablation chamber” – please check for consistency

Done.

Line 129: “no significant sublimation” – this is another reason why the optical mosaics are important. The fact that the grain boundaries are visible indicates some sublimation has happened (after surface decontamination). I would suggest phrasing this as “no additional increase in sublimation”, although I am not 100% sure if this is what the authors are trying to say, and how this would be determined – are we talking about visual observations?

The grain boundaries are already faintly visible immediately after microtoming. They will constantly deepen over time, which is the fundamental physical process and can only be slowed, not stopped. We did however not observe a deepening of the grain boundary grooves that would have caused the laser to go out of focus, which would then lead to insufficient ablation. “No significant sublimation” refers to the sublimation of the main ice sample body, which would for example be indicated by a rounding of the outer edges. This was not observed during the measurement and on the sample pictures. We added a short remark in L 130f rev. manuscript.

Line 155: Was the dwell time really 1s for 238U?

Typo, corrected to 1 ms

Line 166: Does “laser power” refer to nominal fluence settings?

Changed to nominal laser fluence, L 171 rev. manuscript

Line 167 – 169: this statement is unclear. What is a “drift in laser system sensitivity”?

This was related to the need of readjustment of the fluence settings and was misleading. We removed the sentence and reformulated it (L 170-174 rev. manuscript)

Line 177: “Therefore the acquisition time is the dominating factor regarding depth resolution compared to the much smaller washout time.” See comment above. The washout time for 150µm spot size is unknown.

See comment above, paragraph changed.

Line 184: Remove “therefore” in this sentence.

Done.

Line 223: I think there is a “for” missing here: “… for which … “

For added.

Table 4 is a very good overview. Was the rectangular spot not used? If it was used, it should be included here.

It was only used for the ice section from the Last Interglacial, which is no longer discussed in the revised version of the paper.

Line 315: Is “500 year old” correct?

Yes.

Hoffmann, H. M., Grieman, M. M., King, A. C. F., Epifanio, J. A., Martin, K., Vladimirova, D., Pryer, H. V., Doyle, E., Schmidt, A., Humby, J. D., Rowell, I. F., Nehrbass-Ahles, C., Thomas, E. R., Mulvaney, R., and Wolff, E. W.: The ST22 chronology for the Skytrain Ice Rise ice core – Part 1: A stratigraphic chronology of the last 2000 years, Clim. Past, 18, 1831–1847, https://doi.org/10.5194/cp-18-1831-2022, 2022.

Jerše, A., Mervič, K., van Elteren, J. T., Šelih, V. S., & Šala, M. (2022). Quantification anomalies in single pulse LA-ICP-MS analysis associated with laser fluence and beam size. Analyst, 147(23), 5293-5299. https://doi.org/10.1039/D2AN01172G