Replies to referees
Referee 2

Ref2: The manuscript presents a series of detailed investigations of cm-scaled sections from the EGRIP ice core, mostly from the Holocene ice core section. The sections are analyzed for cryo-Raman spectra, microstructure, microscopy, and compared to high-resolution dust records. Based on this, the temporal development of the ice core mineralogy is discussed.
The manuscript (MS) is very detailed almost providing a review in some section, figures are generally good, and referencing is satisfactory. The analytical work presented in the paper is impressive and of relevance to the community. I have some comments and concerns in the following.

General comments
Ref2: Whereas the MS discusses to a great detail how the impurity composition results may relate to the long-term climatic context, I think another important property/sample characteristics is somewhat overlooked, namely the sample seasonality**.** Because almost all of the impurities in the Greenland ice sheet show some kind of seasonal pattern/variation, I think the season from where your samples are taken may influence their composition just as much as the longer term climate context (eg where in the Holocene the sample is taken). For example concerning dust, we know that today we have a large Asian-derived dust spike in Greenland in spring/summer, whereas the dust in other seasons could be of different origin(s). The detailed sample composition may therefore depend strongly of which season the sample is take from. In Fig. 1 (b)-(l), we see that all of the samples appear to be associated with a dust spike. Does that mean that all samples are from spring/summer? I think that with all of the high resolution profiles that are available for EGRIP, it should be possible to determine the approximate seasonality of the samples? For the Holocene, you may for example compare the CFA profiles to those of (Gfeller et al., 2014).
Likewise, it may be of use to make more comparison to the CFA / DEP / ECM record or to the line scan profile across the sections you have sampled. Are samples with high sulfate associated with high DEP/ECM/Conductivity? Are you in a winter layer with high NaCl? Are the dust concentrations typical? Is here a possibility that one or several samples coincidence with forest fires (NH4), a volcanic eruption (DEP/ECM/Conductivity) or some other atypical feature?
One sample is from a cloudy band. Is this a 'typical' cloudy band for the period or is it somehow exceptional? It is stated that in this MS you 'focus on the chemistry', but there is very little chemistry data shown and the comparison between 'chemistry' and the Raman and other results is sparse.
Reply: We thank the referee for a detailed review and several good ideas to put our work in context and to enhance readability. The majority of the suggestions were adopted and are discussed below. You present several good ideas, but not all of them are feasible within this study and would require dedicated personnel and specific studies (e.g., on the seasonal changes of certain parameters). The general idea of our study is to give an overview of the mineralogy of inclusions in EGRIP Holocene ice, tackling very specific features of more records is thus not expedient.

CFA and Raman spectroscopy data were both categorized as chemistry data in this context. To clarify this we changed the text to "we here focus on the mineralogy and insoluble particle content." To clarify this throughout the text we specified the terms of data we use throughout the manuscript. We included the DEP and ECM records (new Fig. 1 below) which provide a good overview since water isotope data are not available yet. We indicate the depths of the analysed samples in the dust record and also show an age scale.
We did a high-resolution analysis of conductivity and acidity at the analysed depths, but no prominent features were found (Figure below). The most prominent one is a small wavy increase of acidity from 757.18 to 757.24 m (sample 6). This is also related to the fact that sample 6 is the longest sample and a thus higher chance of a change in acidity or conductivity. The more acidic part is characterized by different sulphates as displayed in Fig. 7a. No clear correlations are visible from the plots and we thus did not include them. If the editor thinks this figure helps the manuscript, we will include it in the supplements. This is potentially due to different reasons. The main reason is probably the difference in resolution of the applied methods (DEP every 5 mm, ECM every 1 mm, inclusions for Raman a few micrometer) again demonstrating the challenge in combining different scales and methodological approaches (continuous vs. discrete). Another aspect is the slight difference in the measured planes of the samples. The same applies to Visual Stratigraphy data, deciphering clear signals is difficult on such a small scale.
We did not find any indications for an "exceptional cloudy band", but you raise an interesting point. There is work in progress regarding the cloudy bands in EGRIP ice and Raman spectroscopy might play a role in this.
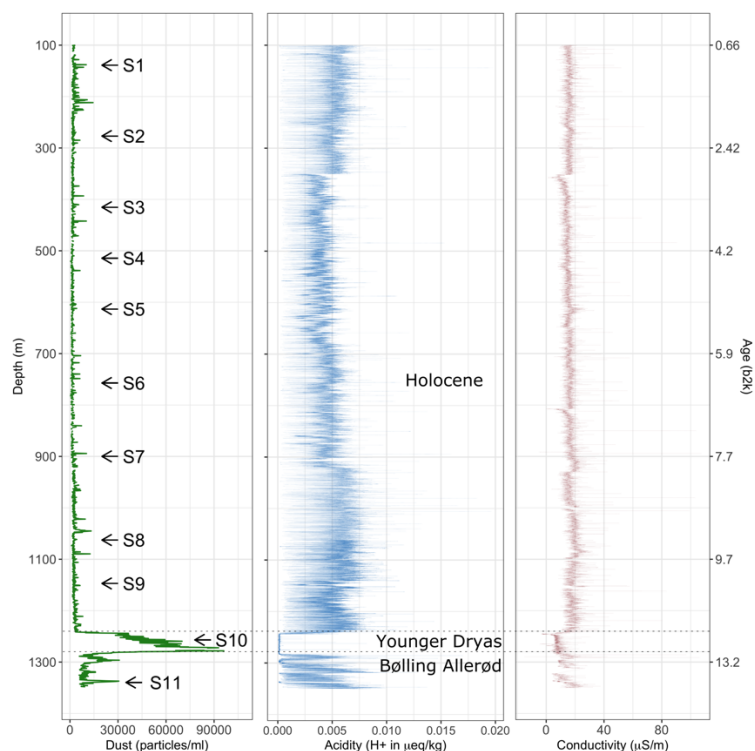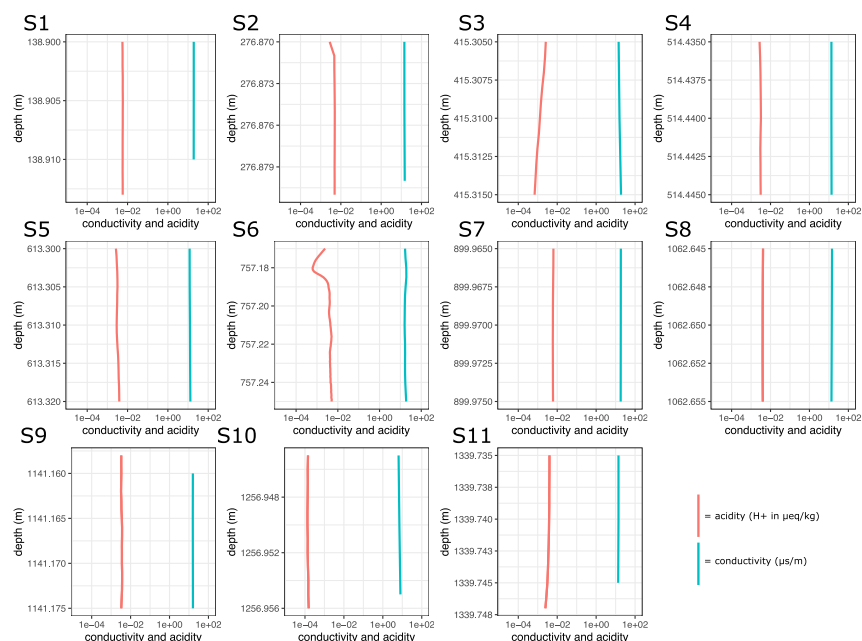
Figure 1 Number counts of particles larger than 1μm per ml of melt water (dust) derived via CFA from the upper 1350 m of the EGRIP ice core. Samples analysed with Raman spectroscopy are indicated with arrows. Acidity data from Mojtabavi et al. (2020b), conductivity data from Mojtabavi et al. (2020c). Age from Mojtabavi et al. (2020a), b2k = before 2000 C.

Ref2: I think the authors need to spend a little more time working with the wording of the text. In every other line, I think there are imprecise statements or the wording is not concise. I gave up making a list of specific places where I think the text could be improved, as I think this is a task of the authors.

Reply: We enhanced the wording throughout the manuscript, it is now more concise and precise. Examples are found throughout the text, this reply ,and in the reply to referee 1.

Specific comments

Ref2: Throughout the MS there is reference to 'bags'. Whereas this may be a meaningful notation for those working on ice cores on an everyday basis it may not be the most obvious notation for the reader. I would suggest to replace with '55 cm sample' or similar throughout the MS.

Reply: We agree and changed "bag" to "55 cm sample" throughout the text.

Ref2: Throughout the MS there is reference to 'a companion paper'. Rather than making the reader start guessing about what paper that might be, I suggestion to cite to the full reference.
Reply: As already stated for Referee 1 at the time of the submission there was no proper way to cite the companion paper, because of the submission at the same time. Since part 1 and part 2 now have a citable doi etc. we changed "companion paper" to "Stoll et al. (2021a)".

Ref2: In several places there is mentioning of the upstream effects at EGRIP. There is now a paper discussing those effects at EGRIP and it may be discussed how important upstream effects may be for the sampled intervals (Gerber et al., 2021).
Reply: We are happy that this paper is published and refer to it in the methods and the discussion:
"Accumulation rates were highest 7.8 kyr ago (0.249 ma−1) and decreased towards the Last Glacial Period with a peak during the Bølling Allerød. Due to the flow of NEGIS ice from the last 8 kyr was deposited under increasingly higher accumulation rates with increasing age caused by higher precipitation closer to the ice divide (Gerber et al., 2021)."
"Gerber et al. (2021) propose that Last Glacial Period ice was deposited 197 to 332 km upstream from EastGRIP."
"Our analysed samples were deposited within 197 km upstream from EGRIP and thus at slightly higher surface elevations (2993±7 m a.s.l. at 1400 m depth) (Gerber et al., 2021), which limits the impact on the aerosol input. Accumulation rates for ice from depths of 900 to 1400 m were low, except the peak during the Bølling Allerød. This peak coincidences with the high I$_{var}$ value in S11 displaying that a high accumulation rate enhances mineral diversity in this climatic period, contrary to the Holocene. However, it is difficult to compare the Holocene samples to the two samples from the Younger Dryas and Bølling Allerød. A systematic follow-up study on EGRIP Glacial ice is needed to investigate if the observed trends, e.g., of mineral diversity, continue with depth."

Ref2: For the discussion of the extend of the Greenland ice sheet in earlier periods and possible costal dust sources (l. 422-432), you may refer to (Simonsen et al., 2019).
Reply: We added Simonsen et al. (2019) and extended the discussion to: "The RECAP dust record shows the exposition of local dust sources, e.g., in King Christian X land, between 12.1 ± 0.1 to 9.0 ± 0.1 ka b2k (Simonsen et al., 2019).".

Ref2: Figure 1: This is the figure where you put your samples into a climatic context and I have a number of comment/suggestions:
- It is very important that overall climatic context is clear, therefore, it would be very helpful to include to water isotopic profile (d18O) in the figure, eg we need to know exactly where the YD onset and terminations are in relation to the samples and which part of BA your sample are taken from. If the EGRIP isotopes are not released you can show DEP or ECM or transfer the isotopes from another deep ice core.
Reply: This is a valuable comment to increase the accessibility for a broader community. Unfortunately, the EGRIP d18O record is not released yet, we thus plot the ECM and DEP profiles next to the dust record as explained above (see Fig. 1 above).

Ref2: - In (a) I do not understand what the black dots represents. How are the depths chosen? Isn't there a continuous dust profile for the Holocene? It seems like the dot density is very irregular? Considering the abrupt change in dust concentration at the YD boundaries by an order of magnitude or more, the smoothing of the dust profile appears somewhat unjustified.
Reply: The black dots were chosen to represent the 55 cm bag values which were analysed for physical properties, and thus available for Raman measurements. We now show a continuous dust profile of all 55 cm sample means together with continuous DEP conductivity and ECM acidity data.

Ref2: - In all figures it says the x-axis shows dust particles per ml. Are those the >1 micron particles only or what size fractions are included. This should be specified.
Reply: This is addressed in section 2.2 Continuous Flow Analysis: "Micro-particle concentrations where determined using an Abakus (Fa Klotz) Laser Particle Sizer (e.g., Ruth et al., 2003) operating in the size range between 1-15 µm, which covers the size range of optical microscopy."
To clarify this in the figure we changed the figure caption to: "Number counts of particles larger than 1 µm per ml of melt water (dust) derived via CFA from the upper…".

Ref2: - In all of the Greenland dust profiles I know of (Eg (Ruth et al., 2002; Schüpbach et al., 2018)), the Younger Dryas is characterized by a much high (order of magnitude) dust concentration than the Holocene, whereas the BA period has intermediate dust levels. This pattern is not at all reflected in the dust curve shown in (a). The indication of the YD interval appears inconsistent with the depths provided lines 100-103.
Reply: We mixed something up, the YD is deeper in the core than shown in the plot – thanks for noticing! We corrected this and the dust profile now aligns with the mentioned literature and is cleary visible in Fig. 1.

Ref2: - Figures (b)-(l) nicely show the position of the sample in context of the continuous dust profile, but they do not give an impression of the absolute dust level in each sample, which vary by orders of magnitude and may have important implications for the interpretation. I would suggest to either use a common log scale for all the figures or to keep-as-is but then add another column of figures that shows the absolute level at the same detailed depth resolution. The dust level at the sample resolution is a basic parameter that that may fundamentally impact the sample composition, and it is not deducible from (a).
Reply: We appreciate this detailed feedback and decided to split up both figures. A new figure now shows the insoluble particle number with depth (log-scale) together with acidity and conductivity data as described above (now Fig. 1). The detailed sections are now shown in a second figure (Fig. 2 below). We changed the labeling of (now) Fig. 2 to "…Please note the different scales in dust concentration on the abscissa. S10 is within a cloudy band, a horizontal layer of much higher dust content than the area above and below."
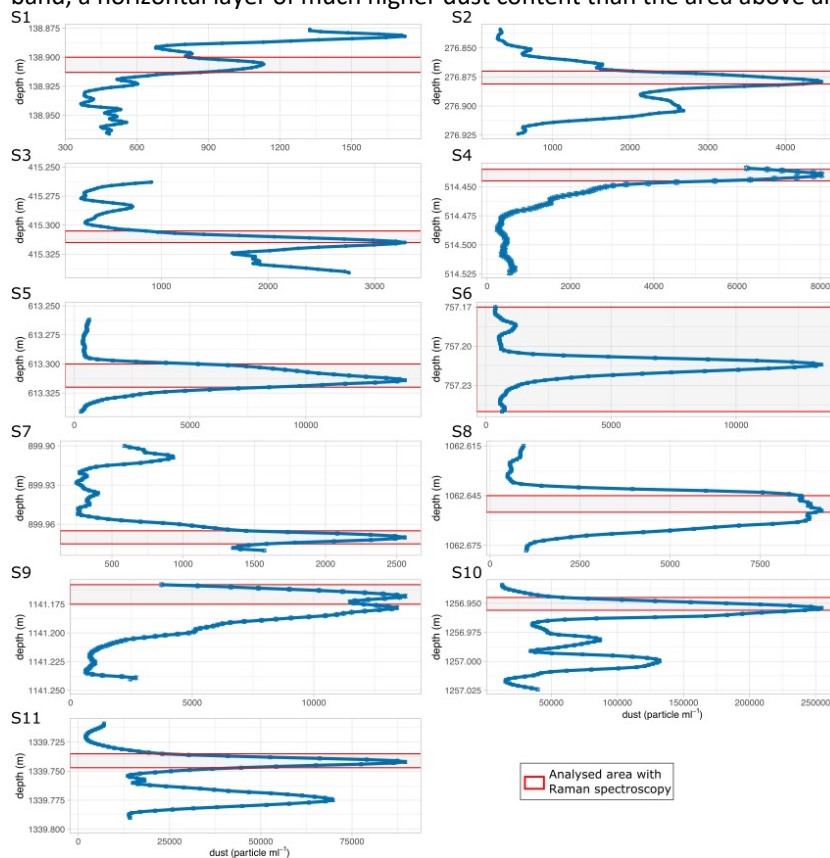

Figure 2 Number counts of insoluble particles larger than 1 µm per ml (dust) derived via CFA from the chosen depths analysed with Raman spectroscopy. Please note the different scales in dust concentration on the abscissa. S10 is within a cloudy band, a horizontal layer of much higher dust content than the area above and below.

Ref2: - Caption: please be somewhat more precise in this and other captions: Eg 'Dust data' potentially means 'Number counts of particles larger than 1 micron per ml of melt water'? '55 cm bags' may not make sense to the reader. 'Cloudy band' may not make sense to the reader. Refer to main text if explained elsewhere.
Reply: We edited all captions and applied the suggested changes.

Ref2: Table 1: Please specify what the depth refers to: top, middle or bottom of sample. If someone wants to compare your results to other records it is important to know the exact sample position. You may consider naming your samples, eg S01, S02, ... S11 rather than referring to the sample depth in the text. This may improve the readability of Figure 4 and others. If you do that, this table should include the sample names. You

may also include information about what criteria the individual samples are selected from. Furthermore, information about the mean crystal size, fraction of sample covered by crystal boundaries, the mean dust and salt concentrations and the sample average conductivity/DEP/ECM level(s). If possible, information about the season from where the sample is taken could be included as well.

Reply: We changed the table in several ways. We happily adopted the suggestion to use S1-S11 and changed it in the text and Table 1. Furthermore, we added the exact depths intervals of the Raman samples. The exact depths were shown in Fig. 1, adding them to Table 1 should make it easier to compare them to other records. For these exact depths we added the mean acidity and conductivity values to Table 1. Conductivity and acidity are now also displayed in Fig. 1 (see longer explanation above).

Including much more information in one table is not possible for various reasons. For example, information about the crystal size would not make sense due to poor statistics at the usual Raman sample size of ~1 x 1 cm. This could only be shown with a good statistic for the entire 9 cm sample (thin section) which would give a wrong impression of the specific area analysed with Raman spectroscopy. The "bulk" information is however included in part 1 and thus easily available (Stoll et al., 2021a).

We appreciate the other ideas, but it is not possible to do all this within this study. For example, investigating the seasonality along the core is a major task as explained above. More details regarding the DEP/ECM record are mentioned above.

Ref2: Table 2: You investigate the fraction of impurities that are found in grain boundaries. I think it is relevant in this context also to state the fraction of each sample that is covered by grain boundaries according to your 300 micron definition? For some samples it appears that a quite a large fraction of the analyzed area is covered by boundaries. If you then subtract the area covered by air bubbles, it could be that in the end there is no preference of the impurities to be located in a boundary or not?

Reply: We did not include the area occupied by the grain boundary as this is shown (and thus referred to) in Stoll et al. (2021a) and the table is already quite broad. It would be necessary to include each sample, the area occupied by inclusions and by grain boundaries, ergo a new table (which would be similar to Table 1 in Stoll et al., 2021a. We thus added a link to Stoll et al. (2021a): "Details on grain boundaries are shown in Table 1 in Stoll et al. (2021a). In the updated version of Stoll et al. (2021) we also discuss the impact of different grain boundary thicknesses (100, 200, and 300 micron).

Investigating the area covered by air bubbles is an interesting approach, but is not suitable for this study. The volume of bubbles is magnitudes larger than grain boundaries and the insecurity with depth (sample surface and inclusions 500 um below) would be large. Thus, a different methodological approach might be needed. It is definitely an interesting approach, which could be investigated in a different study.

Ref2: Figure 4: Based on the sulfate diversity presented in Fig. 4(b) you conclude that there is a general decrease in the sulfate diversity with depth and in the abstract you mention that there is a change at around 900 m depth that is also discussed at length in the discussion. I think this conclusion is poorly supported by the data. Indeed, the sulfate diversity of the four deepest samples is low, but it is also low for two other samples from above 900 me depth. The two deepest samples are from the last glacial period where many climatic conditions were quite different from the Holocene conditions, so I am not sure those two deeper samples are directly comparable to the younger ones. An alternative interpretation of the figure would be to say that the sample from 1062.65 m depth looks unusual in terms of sulfates, but that all of the other samples are similar, leaving out the two deepest samples that are from a different climatic period. In other words, I think the statistics may not allow for the conclusion you make.

Reply: We agree that it is difficult to draw definite conclusion from our statistics even though numbers and spatial-resolution are comparably large. We would like to address the two mentioned aspects separately:

1) Decrease in sulphate diversity

We mainly focus on the (almost) non-abundance of other sulphates than gypsum below 900 m. Sulphates are not consistently diverse throughout the upper 900 m, but there is usually more than one sulphate at each depth (6/7 samples have at least 2 sulphates in the upper 900 m).

You are right, there are two samples within the upper 900 m of similar properties. However, higher insoluble particle (dust) content mainly correlates with the dominance of gypsum in comparison to other sulphates (see answer to Reviewer 1) and thus supports our interpretation regarding the deeper samples.

2) Change in mineralogy around 900 m

In the abstract we say "A variety of sulphates dominate the upper 900 m while gypsum is the only sulphate in deeper samples, which however contain more mineral dust, nitrates and dolomite." We thus not only refer to (the debatable distribution of) sulphates, but also to other minerals (e.g., nitrates and dolomite). These minerals only occur below 900 m while other minerals, such as hematite and titanite, only occur above 900 m.

All these results support our interpretation of a change at 900 m.

We weaken our statement throughout the text and e.g., delete "considerable change" in the abstract: "Inclusions of the same composition tend to cluster, but clustering frequency and mineralogy changes with depth."

We also mention that the deepest two samples are difficult to compare with the samples above: "However, it is difficult to compare the Holocene samples to the two samples from the Younger Dryas and Bølling Allerød. A systematic follow-up study on EGRIP Glacial ice is needed to investigate if the observed trends, e.g., of mineral diversity, continue with depth."

Ref2: l. 37: 'CFA' is unexplained at this point.
Reply: We added the explanation for CFA.

Ref2: l. 111: 'thin sections' is unexplained at this point.
Reply: We changed it to: "Depth co-registration to the samples analysed with Raman spectroscopy is limited by…".

Ref2: l. 290-294: This section appears to belong in the conclusions?
Reply: We partly agree with this statement. Referring to Eichler et al. (2019) and to the mentioned figures helps to convey the concept. We add a summary of this paragraph to the conclusions: "Combining these methods, and thus covering different scales, provides a good basis for a systematic analysis of different depth regimes while ensuring a sufficient number of micro-inclusions."

Ref2: l. 344: Does 'the stadial' refer to the Younger Dryas interval in this case?
Reply: This refers to the Greenland Stadial 1, i.e. the Younger Dryas. To clarify this we changed the text to:" However, it is interesting that the transition from the Glacial to the Holocene is not represented by a major change in mineralogy.".

Ref2: l. 445: What is Dome Fuji Interstadial ice?
Reply: This refers to the Dome Fuji ice core (introduced in l. 340) and the discussed differences between Holocene and interstadial ice by Ohno et al. (2005) on p. 176:"There are also small differences between Holocene and interstadial ice (1351 m)."

Ref2: l. 458-465: It seems unnecessary to repeat part of the introduction here.
Reply: We deleted l. 457-463.

Additional note: We found a mistake in our diversity index calculation, the numerator and denominator were exchanged. After correcting this a value of 1 means every inclusion has a different mineralogy, while small values indicate a low diversity. This led to changes in the figure (see figure below) and in the text:
"To compare our samples despite the varying amount of total identified Raman spectra per sample we calculated the ratio of the amount of different minerals per sample (n_m) to the total amount of identified micro-inclusions per sample (n_i) resulting in the diversity index Ivar with a maximum of 1. Ivar of 1 indicates that every inclusion is of different mineralogy while values close to 0 indicate a low diversity.
$$Ivar = \frac{n_m}{n_i}$$
Ivar varies between 0.099 and 0.308, the mean value is 0.158. Mineralogy diversity decreases slightly with depth (Fig. 6), the large diversity of sulphates is only found in the upper 900 m (Fig. 5B)."
"The lack of a variety of sulphate minerals below 900 m is shown in our diversity index, which decreases with depth even though other minerals occur at these depths, such as dolomite."
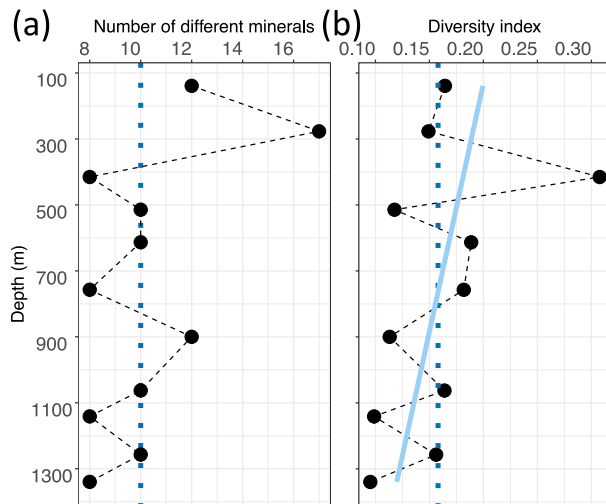
Figure 3 Mineral number and diversity with depth in EGRIP ice. A) Absolute numbers of different minerals per sample. The dotted blueline is the median value (10). B) Diversity index values calculated after Eq. (1). The light blue line is a linear regression, the dotted blue line is the mean value (0.158). Higher values indicate a larger mineral diversity in relationship to the amount of identified Raman spectra per sample.