Comment on tc-2022-79 (RC1)
Mikhail Kanevskiy (Referee)

Computed tomography of frozen sediments is a relatively new method of permafrost studies. This non-destructive method allows to obtain high-resolution 3D images of frozen soils and estimate volumes of solids, gas, and ice inclusions in permafrost samples. We believe that in the future CT will become a main method of field permafrost studies of frozen soils, and this study makes a significant contribution that may help to improve this method. This manuscript provides unique information on application of CT imaging and image processing methods to quantify properties of the upper permafrost in Yedoma region of northern Yakutia. The authors present a systematic quantitative comparison between the laboratory-measured and the CT-derived composition of permafrost cores, which is extremely important for future development of the CT method. I strongly support publication of this manuscript and recommend to accept it with minor revision. My comments and suggestions are attached.

Dear Mikhail Kanevskiy,

We thank you for taking the time to thoroughly review our manusricpt and appreciate the overall positive evaluation. We agree with your view that applying CT imaging to study permafrost soils bears great potential for future research and are happy if our study can contribute towards this strand of research.

We are very thankful for your numerous suggestions to improve the text, and will consider all of them for the revised manuscript. In particular, we appreciated your comments on improving the terminology to refer to different types of excess ice and massive ice contained in the core. Moreover, we will replace the term „sediment phases“ by „sediment types“ which is hopefully less confusing, as the term „phase“ was also criticized by the second reviewer.

We are confident that our revised manuscript will be significantly improved thanks to your valuable comments and suggestions.

Kind regards,

Jan Nitzbon (on behalf of all authors)

Comment on tc-2022-79 (RC2)
Philip Pika (Referee)

This paper presents a thorough analysis and comparison of a permafrost core using both an innovative non-destructive computed tomography scan and a destructive laboratory based sampling method to address the composition of permafrost soils including gas, ice, and solids. The three-dimensional and high resolution CT scanning technique is a great tool for advancing our understanding of the state and history of permafrost soils. This study qualitatively showcases the strengths and weaknesses of CT scans and laboratory based analysis and their combination, and will be a milestone for future studies in this field. However, while I definitely support the publication of this study, I must also ask the authors to address several major points and comments made in the manuscript. Please see the attached document.

Dear Philip Pika,

We thank you for taking the time to review our manuscript and for the overall positive evaluation. We are happy about your appreciation of the importance of our study and are thankful for your further suggestions on how to improve it.

Regarding your numerous comments on the use of British vs American English, we would like to note that we decided to use British English with Oxford spelling (i.e. -ize instead of -ise) and indicated this during manuscript submission with Copernicus. Otherwise, we thank you for pointing out instances of using AE and will replace these with BE in the revised manuscript text.

Subsequently, we respond directly to those of your comments which require more explanation or justification (referee comments are in *italics*, changes to the manuscript text in **bold**).

We are confident that our revised manuscript will considerably be improved by considering your suggestions and are looking forward to its potential publication in TC.

Kind regards,
Jan Nitzbon (on behalf of all co-authors)

Point-by-point reply to RC2 (reviewer comment in *italics*, changes to the manuscript text in **bold**):

L. 74: during samping? Or when did these cracks occur?

The core cracked at different locations during the drilling processes (see Fig. 1a), resulting in some parts being not suitable for the CT analysis.

L. 76: Detecting edges between grey tones in CT scans is one of the most important and difficult steps in during image processing, so does your flitering have any effect on the analysis of the CT scans? Why did you opt for three steps, not two or four? And how did you assess if the filtering was sufficient?

A 3D non-local means filter is currently not implemented in Fiji/ImageJ. Three sweeps of 2D non-local means filters with a noise standard deviation of 6, each in another direction, simply replaces a 3D non-local means filter with a noise standard deviation of 18. Different noise parameters were tested beforehand and evaluated visually. The non-local means filter is known to be superior to more conventional denoising filters, in that it preserved edges while efficiently removing noise in homogeneous areas (Schlüter et al. 2014: 10.1002/2014WR015256). We added the following explaining sentence:

This denoising procedure which is equivalent to a single 3D Non-Local Means filtering step with a noise standard deviation of 18, is known to be superior to more conventional denoising filters, in that it preserves edges while efficiently removing noise in homogeneous areas (Schlüter et al, 2014).

L. 81: I would use another word. "class" would be more appropriate, for instance.

We understand that the term „phases“ in this context might cause confusion, which was also pointed out by reviewer 1. Thus, we will replace the term by a less confuing term in the revised mansucript.

L. 82: I would really appreciate the threshold values (HU values) to be made available somewhere in the main body of the manuscript. I have not seen them. Adding these values helps, to understand where and how you made the cut.

We would like to point out that we decided not to use Hunsfield units (HU) to normalize the greyscale values. Instead, we decided to normalize each CT scan individually using the percentile method described in L. 68ff. Thus, the thresholds used for the segmentation have been determined individually for each CT scan and cannot be generalized. Therefore, reporting the thresholds for the individual CT scans would not be of any use for the study of other CT scans.

L. 141: Great to see that the CT scans are able to identify new sections of the core previously invisible to the eye. The sections are crucial to the rest of the text, and therefore mark these sections also in figures 1 and 2 to better guide the reader.

Thank you for this suggestion. We will mark the sections in Fig. 1 in the revised manuscript.

Figure 1 Caption: What is the blue background in the photography? It looks like a mask from the image processing tool? Remove it if it has no meaning, or refer to it in the main text.

The blue background is actually a foil on which the core was put directly after drilling in the field. As it provides a good contrast to the core, we see no need to remove it from the photography.

Figure 1 Caption: Mention these are density gradients, not just brightness values. The gradient has a physical meaning and can interpreted as such.

We agree that the brightness values are indicative of the density of the scanned material and thus added this information.

Figure 2 Caption: Directly and indirectly are confusing terms. Use lab-derived and CT scan derived? I guess this needs to be adjusted across the manuscript.
The legend is confusing.
The CT scan is the indirect method, but you write gas and excess ice to be estimated directly. In the case of gas you even write that it was not possible to measure in the lab!
And the organic and mineral phases are also estimated by Ct scans, no? But called phases A and B?
Correct the legend and see other comments on the figrue.

From these and other comments we recognized, that our use of the terms „direct“ and „indirect“ for the derivation of the volumetric content profiles might require more explanation. We would like to emphasize that our terms direct/indirect are not simply synonmys for laboratorey-measured and CT-derived.
The volumetric content profiles of gas, excess ice and the two sediment types A/B were derived *directly* from the CT, as they correspond to a certain range of grayscale

values determined during image segmentation.
Importantly, the volumetric contents of sediment types A/B do not correspond to the volumetric contents of organic an mineral material. Both sediment types are composed of pore ice, mineral, and organic material, and their composition has been determined via a least squares regression aginast the laboratory measurements (see Section 2.4 and Table 1).
The results of the regression analysis have then been used to calculate the volumetric content profiles of total ice, organic, and mineral material at the high vertical resolution of the CT scan (using Eq. (5)). We refer to these volumetric content profiles as being determined *indirectly* as they are based on both the results of the CT scanning and the results of the laboratory measurements.
In order to make this more clear, we added the following sentence at the end of Section 2.4:

Note that we refer to the profiles $\theta_{A/B/ei/a}(z)$ as being *directly* derived from the CT images, while the profiles $\theta_{o/m/i}(z)$ were *indirectly* derived form the CT images as they require the composition of the sediment types to be estimated through the statistical regression against laboratory measurements.

Figure 2 Caption: mark that the x-axis is not linear, but has a jump between 24/40. And why is that so?

In fact, the axis is linear. There is just a tick label missing at „32“ which will be added in the revised manuscript.

Figure 2 Caption: Do these include all data points? More data points seem available from panels ab-d. And I would expect point between 80-100%, which would represent the subsections below 85cm.

For the comparison of CT-derived and lab-measured volumetric ice, organic and mineral contents (Fig. 2 e,f,g), we only considered values from the „upper“ and „middle“ part of the core, but not from the „lower“ part (depth >85cm) consisting of massive wedge ice.

Figure 2 Caption: Add the statistical values of the fit in the figure. They get lost in the text below.

We added a reference to Table 1 in the caption of Figure 2 which contains all the evaluation metrics of the linear fits.

Figure 2 Caption: Axis label need adjusting:
CT is the indirect method, correct? Then add "(direct)" to "Lab. measured".

See comment on indirect/direct above.

L. 151: How did you assess the active layer depth?

We measured the thaw depth before drilling the core as mentioned in L. 53.

L. 171: I would avoid using the term 'peat' and suggest the use of "organic material". This is more consistent with the rest of the text, prevents misinterpretation as peat per se cannot be "seen" with the used method. Finally, PF core can may contain woody bits/sticks, which are organic material, but not peat.

We agree that using the term peat in this context might be misleading and replaced it with „organic-rich soil“. Please note that here we refer to the composition of the entire sediment phase A and not just the organic matter content.

L. 180: These are the total ice contents calculated using equation 5, using the excess ice from the CT scan?

Yes. The high-resolution profiles of total ice, mineral, and organic contents have been determined *indirectly* using equations (5). For this, the high-resolution profiles of excess ice, and sediment types A and B which were determined *directly* from the CT images, have been used in combination with the fitted parameters ($\gamma_{A/B,m/o/pi}$). This information might be helpful for understanding our distinction between *directly* and *indirectly* determined volumetric content profiles.

L. 202: What do yo mean? You would just scan an inner section of the core? Also this whole sentence is not correct, because the resolution issue could be circumvented by using a micro-CT scanner, with a higher resolution. See Cnudde and Boone, 2013, where they show the benefits of microCT.

Please note that the CT scans in this study had been acquired with a micro-CT scanner (Xtek-XTH225, Nikon Metrology). There is a fixed ratio between sample diameter and voxel size, which is somewhere between 1000-3000 for modern scanners. For an ice core with a diameter of 75 mm this resulted in a voxel size of

50µm. Higher resolution would have been possible with region-of-interest scans of the whole core, which in our experience are afflicted with radial intensity drifts, or by drilling smaller cores out of the original cores and scanning these small cores. Irrespective of the chosen strategy, the analyzed volume would have been smaller and therefore less representative of the heterogeneity encompassed in the entire core. Moreover, partial volume effects would still be encountered at a smaller voxel size of e.g. 10-30µm.

L. 215: I would not go that far. This will certainly change in the near future. Crabeck et al. 2016 used an innovative way to get a higher resolution in their Ice core, with the current "normal" CT scanner.

We agree that limited resolution is less of a problem for air entrapment in ice cores. However, the problem of partial volume voxels will remain for fine-textured soil with varying organic matter content and small ice-filled voids no matter how far the resolution limit is pushed (at the same time sacrificing representativeness). We therefore have rephrases sentence accordingly:
"...because the resolution of the CT is often insufficient to resolve the phase or object of interest completely."

L. 219: Could you elaborate on this? What would be the benefit? Would across study comparisons be affected in any way?

The benefit would be that the same set of thresholds could be applied to all CT scans. Across-study comparisons are typically done with morphological properties based on segmented images. These would not be affected by gray value normalization. Only the average gray value within material classes would be affected. We have rephases the sentence now:
"Homogeneous reference materials could be used to standardize the gray values prior to segmentation similar to Hounsfield units so that the same set of thresholds could be applied to all CT images"