Interactive comment on "Intercomparison of snow density measurements: bias, precision and spatial resolution" by M. Proksch et al.

Anonymous Referee #1

Received and published: 23 July 2015

We appreciate the extensive review by Referee #1, in particular as he points towards several relevant studies which were not included in the paper. We also followed the advice of the Referee to better structure and increase readability of the paper by rephrasing "density per layer/ traditional stratigraphy" to "cylinder cutter" and referring in the results section better to the methods section. However, we want to clarify here that this paper focuses solely on snow, and not on firn or ice.

Please find our answers to the comments below in blue, *and the text changed in the manuscript in green.*

Comments on the discussion paper by M. Proksch et al "Intercomparison of snow density measurements: bias, precision and spatial resolution

General comments: The paper presents the results of an intercomparison of different density measurement methods. Compared methods are density cutters (3) with Micro- Computed-Tomography. Compared are 1. Box cutter to CT in the lab, 2. all cutters to CT in the field, 3. all methods to a layer mean (obtained by averaging over all methods). Statistical results of the comparison are provided.

A comparison of the different methods to obtain density of snow is a valuable approach to improve the quality of all applications where the measured density is used. The study compares three gravimetric methods (density cutters) with a high-resolution method (CT). The density cutter methods have been compared in earlier studies, the CT method, as it has a much higher resolution has been compared to other high resolution methods in earlier studies (which is not mentioned or discussed by the authors).

We thank the Referee for his valuable advice regarding additional earlier studies, which have been indeed not addressed in the current version. We included them in the manuscript, as detailed below.

In general the results presented here are interesting for a broader community. However, the study as its current stage lacks a larger scientific background. What is the advantage, the gain or benefit to recent scientific discussions? What is the take-home message for people using these measurement data in model applications or any other application? What is the implication of presented variations in density measurements (how much would that effect for example the calculation of swe, or the computation of metamorphism in models etc).

This study, as well as the whole "Intercomparison of snow grain size measurements methods workshop", which is an outcome of the IACS working group MicroSnow, aimed to provide an estimate how much the state of the art methods differ from each other. We see the advantage for the recent scientific discussion exactly there: to raise awareness about the differences between different measurement methods, and moreover to quantify these differences. With respect to the density paper, the take home message is that density derived from CT agrees in general within 9% with the density derived from density cutters.
As no "true" snow density measurement exists, we limited ourselves to an intercomparison of methods. Users, i.e. microwave modelers, which are driving their models with density measurements from one or the other instrument, now do have an estimate how much variation they have to expect solely from the differences in density measurement. Depending on the model, users are now able to estimate how this difference may propagate through their model and affect their results. It is, however, beyond our abilities to track the effect of density uncertainty in complex microwave models. Nevertheless, based on your suggestion, we provided estimates for thermal conductivity and the critical cut length (a measure of snow instability):

Assuming a density of 300 kg m^{-3} and a variation of 10% or 30 kg m^{-3}, the error in thermal conductivity based on Calonne2011 would equal 0.045 W K^{-1} m^{-1}, which results in an uncertainty of 21% . The critical cut length, a measure for snow instability, increases from 0.53 cm to 0.59 cm (increase of 9 \%), if the density of the snow slab on top of the weak layer is increased by 10 \% from 300 kmg-3 to 330 kgm3 following the procedure described in Reuter2015 (slab height 60 cm, weak layer fracture energy 0.5Jm-2, elastic modulus of the snow slab derived from Scapozza2004, slope angle 0°).

The data are presented, beyond that a profound presentation or discussion for example on the indeed interesting problem of different layering seen by different instruments and the intra-layer variability is missing.

We tried to address the issue of different layering (section 5.2.1 " Representation of the stratigraphy by the density measurements", Fig. 3 and 8) and intra layer variability (section 5.2.3 "Unresolved variation", Fig. 6 and 7):

- For the different layering, Fig 3 shows the all the density profiles of the different methods, and section 5.2.1 e.g. states on page 3594 ll. 16ff: "…the wedge cutter did not represent the variations measured by the box cutter, and the box cutter did not represent the variations measured by the CT." or l. 18ff: "on the one hand, layer boundaries which were defined following the traditional stratigraphic approach (Fierz et al., 2009) appeared less distinct in the CT, and on the other hand, the higher resolution methods resolved a high degree of variability within a layer." A detailed illustration of the stratigraphy for the lower part of the snow profile revealed by different methods is shown in Figure 8.
- The intra-layer variation is discussed in section 5.2.3. and quantified in section 4.3

It is obvious, that more instruments cannot be taken into account, since they were not involved in the measurement campaign, from which the presented data origin. However, it would have been much more interesting to include more high-resolution methods, since they represent the state of the art and are also widely used in the field. At least in the discussion part, published results of previous comparisons of high-resolution measurement methods (their known precision and bias) could be compared to the results obtained in this study. Overall, there have been earlier studies with a similar approach (comparing traditional gravimetric methods to high-resolution methods and stratigraphy). With regard to these publications the presented study does not provide new results.

We agree partially with the reviewer and tried to add the results of the previous studies, see comment below". However, we would like to point out that our study includes for the first time a detailed analysis of the vertical spatial resolution and the effects caused by different sampling volume of the instruments.

Specific comments:

Title: Spatial resolution is not addressed at all in the paper
We agree to the reviewer and changed the wording in the title to: *"vertical resolution"*

Abstract Line 4-5: This is not true. Examples: Freitag et al. 2004, comparing CT density measurements with gamma-absorption method Kawamura, 1990, comparing CT to hydrostatical method (which is comparable gravimetric), ice Lundy et al., 2002, comparing CT to traditional method, snow
We agree with the reviewer, however, Lundy et al, 2002, reported only qualitative results (derived from weighing a snow sample, but without using a density cutter), and the other studies did not focus on alpine snow. Therefore we changed to: *"No study has yet quantitatively considered the recent advances in snow measurement methods such as micro-computed tomography (CT) in alpine snow"*

Line 18: what is meant by “introduced by the observer” ? Check Phrasing

The observer decides on the layering of the snowpack, and as such "introduces" layers into the snowpack, as described in section 2.1. Layers are not an a-priori property of the snowpack, but layers represent a stratigraphic classification of the snowpack. Different observers will judge the stratigraphy differently, and consequently result in different layers. The code of practice for classifying the layers in an alpine snowpack is given in Fierz et al., 2009.

Introduction Page 3583, Line: 20ff: This part could be improved a lot: Take for example Hawley et al., 2008, or Harper and Bradford, 2003, who take very different methods and compare them and already discuss the issue of different resolutions; Kendra et al., 1994, comparing gravimetric methods with snow probe

We agree and included the relevant studies:

Page 3583, l.20: Despite its relevance, few studies quantified so far the differences between the methods to measure snow density. Indeed several studies focused on firn and ice density, but those were mostly limited to firn and ice, i.e. density ranges (>500) larger than the one typically found in alpine snow (50 - 400).

Page 3583, l.28:

Page 3584, l.3: …micro-computed tomography (CT, Schneebeli and Sokratov, 2004, Lundy, 2002)…

Page 3584, l.10: Dielectric devices… (Denoth et al., 1984; Tiuri and Sihvola, 1986; Kendra, 1994; Mätzler, 1996). Neutron absorption (Kane 1969,Morris2003) was in particular used to measure density inside a firn or ice bore hole.

Page 3584, l.17ff: The impact of measurement resolution was in particular demonstrated by Harper and Bradford 2003, who showed that the identification of stratigraphy is a function of a tool’s sensitivity to vertical contrast. Hawley 2008 in addition highlighted the smoothing of the density profile of an ice core for instruments with larger measurement/sensor length. In terms of measurement time, the SMP is more time efficient, as excavation of a snow pit is not necessary. Vertical profiles of snow density through repeated measurements with the SMP allow the investigation of the spatial variability of snow density.

Page 3583,Line 24 ff: What is then the new approach of your study, the three cutter methods already compared elsewhere?

All three cutter methods were never compared to the CT (and Lundy reports only qualitative results and did not use a density cutter.)

Page 3584, Lines 1-13: there is more around, not all need to be included, but currently the presented selection is quite narrow (DEP: S. Fujita or F. Wilhelms; Neutron-scattering, R. Hawley and Liz Morris etc).

See comment Introduction Page 3583, Line: 20ff.

Page 3584, Line 23: why keeping ’spatial variability’ in the title?

The title does not contain "spatial variability". However, we changed the expression to *"vertical resolution"*, see comment above.

General comment to that section: Introducing other methods and comparisons is good, unfortunately the presented study only uses gravimetric methods (3) and CT – What is the sense of the introduced methods here, if not picking up on them in the discussion? Maybe the authors could use the results of previous publications and comparisons by discussing their findings with results found by others (i.e. is the difference between the gravimetric methods and CT comparable to the difference between CT and other high-resolution methods etc – to get a feeling, were the main uncertainties are).

We appreciate the suggestion of the reviewer to discuss results of other studies:

Several studies have compared different methods to measure density, but were mostly limited on firn and ice, i.e. density ranges (>500) larger than the one typically found in alpine snow (50 - 350). Freitag 2004 compared firn densities measured by CT with those measured by gamma-absorption for three sections of a firn core, each approx. 60 cm long. The authors report an deviations of less than 1 \% for both the methods in the density range from 640 - 733 kgm-3, but also qualitatively higher values for the CT in the range 460 - 550 kgm-3 and lower values for the CT in for densities above 733 kgm-3. However, no results are reported for densities below 460 kgm-3, which were in evidence at the workshop. Kawamura reported good agreement between CT and the hydrostatical method to determine the density of ice cores. Hawley 2008 compared neutron probing, dielectric profiling, optical stratigraphy of the core and gravimetric measurements on an 11 m firn/ice core from Kongsvegen, Svalbard. The authors reported a smoothing of thin ice layer in particular for the neutron probe due too its large detector size of 13.5 cm, but also for the dielectric device due too its finite sampling volume, where the authors estimated a sensing length of approx 4 cm. Other problems related to the gravimetric and dielectric measurements were mentioned with respect to collecting cores (accurate measurement of borehole diameter, depth registration, core breaks, poor core quality or melting of cores during shipping) or loose snow at the surface of the bore hole.

Studies which quantitatively focus on snow rather than firn or ice are rarely available. A study which compared density measured by CT and by weighing samples of sieved snow was presented by Lundy 2002. The authors reported qualitatively a good agreement between both methods for their 4 investigated samples, however, none of the three density cutter used in this study was used. Dielectric devices were also compared to gravimetric measurements. Kendra 1994 found an RMSE of their snow probe of \pm 50 kgm-3 compared to gravimetric measurements only in a qualitative way, whereas a RMSE of maximum 9 \% (29.7 kgm-3; lab data, box cutter to CT) was found for our data.

Methods General: The definition, use and presentation of the 'stratigraphic method' is unclear. Reading the first part of the method section one expects later two profiles of each method – the continuous profile and the profile with samples from each layer. However, this is not the case. Later the authors refer to the stratigraphy method, but it is not clear, how it is determined and which line in the graphs actually shows this method.
We agree that the term " stratigraphic method" or "density per layer" is misleading and changed it to "cylinder cutter", to be in line with the legend of the figures.

The stratigraphic method is the only method that determines the density per layer, and was solely performed using the cylinder cutter. All other methods were used to determine the density per sample, and not per layer (section 2.1.). We clarified this in the methods section and added the following phrase:

Page 3585 l. 26:"In this study, the cylinder cutter was used to measure the density per layer, after the layers were determined following Fierz 2009. All other methods were used to measure the density per sample. As such, the cylinder cutter provided a density profile with varying vertical resolution, based on the thickness of the layers, contrasted by box and wedge cutter, as well as CT, which were operated with constant vertical resolution."

Page 3587 l. 28: "The density per layer or traditional stratigraphy is termed "cylinder cutter" hereafter, as only by the cylinder cutter was used in this study to determine the density per layer. All other devices (box and wedge cutter, CT) were operated without considering layering or stratigraphy of the snowpack, i.e. with constant vertical resolution (see also Section 2.1)."

Page 3586, Lines 8-15: sample size and resolution are missing here (there are included in the sections dealing with the other methods below) General: CT samples are extremely small compared to the others. A discussion on the difference of the samples size and its effect on the comparison is missing. As the snow is not homogeneous in space and over different scales some words or even numbers / references need to be included.
We agree and added the following sentence in the discussions:

Page 3594, l.26.: The fact that the higher resolution methods resolved a higher degree of density variability is closely related to the measurement volume of the different instruments. For instance, the measurement volume of the CT (15^3 mm-3 = 3375 mm^3 = 3.375 cm^3) is around 3\% the measurement volume of the 100 cm^3 box density cutter. A larger measurement volume is immutably connected to a smoothing of the measured density profile, as very thin layers are averaged within the measurement volume. This explains the lower *variability of the box cutter density profile, compared to the high frequency density variations resolved by the CT. As the measurement volume of the CT was sufficiently large to be representative (1.25^3 mm^3 = 1.95 mm^3 found by Kaepfer2005, section 3.1.), these high frequency density fluctuations are not an artifact of a small measurement volume.*

(Is there a possibility to have many samples from the same layer /depth interval etc to look at the variability of a number of CT samples within the same 'bulk sample' captured by the other methods, and compare this variability then to the variability of different methods?) At least this issue needs to be discussed and an estimation of the value of the variability introduced by this compared to the method-induced variability should be given.

If we understand the question correctly, we answered this question by calculating the "unresolved variation".

Page 3587, Lines 5-28: Where in paper are the profiles shown? Might be overseen, but a plot, where the profile (layer) is compared to the continuously sampled profile is missing.
Figure 3.

Page 3588, Lines 1 - 23: General Structure: Three different methods of comparison are introduced here (a-c). For the reader it is hard to find them in the following text. It would be more convenient to structure the results in the same way. Another option could be to add a link/reference to the sub chapter/ figure/table, where the relevant results of this method is described/shown here, so that the reader can find the results of this method (i.e. method a (see chapter x.y and figure x.y)
We agree and structured the results section better:
Page 3590 l.24: Three types of comparisons (Sect. 2.3) were performed, all excluding ice layer. For comparison a), the bulk densities derived from each method were compared. In addition, a cylinder of inner diameter 9.44cm and length 55 cm (Sect 2.2.3. was used.

Page 3591 l.6: "For comparison b), all methods were compared to the CT density profile. For this reason, the high resolution CT profile was averaged...."

Page 3591 l.16: "For comparison c) and to facilitate a more objective comparison, all measurements were averaged to the same vertical resolution, i.e. to match the traditional stratigraphic layers. The mean density per layer of all instruments was then set as reference. With respect to this reference, the different methods agreed within 2 to 5% (Fig. 5, Table 6), the bias was between -1 and 1 % and R2 = 0.99 for all instruments, significant at the 1% level. When ice layers..."

Meaning of approach a: Reading section on page 3588 lines 4-7 (not quite straightforward to understand – What is meant by 'it' in line 5?) one understands the following: The whole measured profile is taken and converted into one swe value - that gives one value for each profile. This value obtained by one method is compared with the value obtained by another method. What is the meaning of such an approach? Where are the results of this comparison?
This approach shows whether the means of all methods were consistent with each other. By this check we determine if there is a systematic over- or undersampling of the methods agree or not (if they are biased towards lower or higher values).

Approach c: One has to search and read twice in order to find the results of this comparison. It would be interesting to have a sketch of the observed traditional stratigraphy and the measured densities together.
See comment above on Page 3588, Lines 1 - 23. We also modified Figure 3 for better illustration of the different profiles:

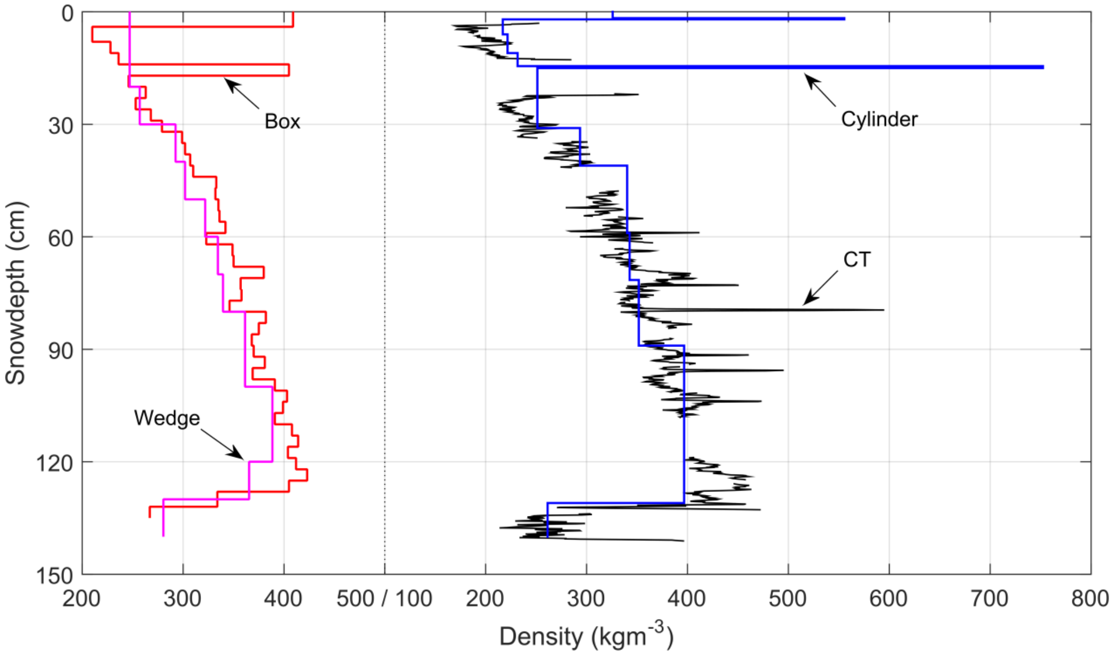

Figure 1: Density profile measured by different methods. Two methods each are displayed separately for better visibility. Note that the cylinder profile shows the density with respect to the stratigraphic layers.

Data collection Page 3589, Lines 1-18: Maybe this part should be moved to the methods-part, where the CT is introduced, the same for the following lines on page 3590.
We agree and moved this section to the methods part.

Results Page 3590, Line25 ff: General: Suggestion to structure this chapter according to the methods – that makes it a lot easier for the reader to follow Line 25:Is this the result belonging to method a (Page 3588) ?
See comment above on Page 3588, Lines 1 - 23.

Probably this is a problem of wording: What is meant by reference value? How is the swe calculated – some words on that could be included in the methods part. How does the ratio of swe to snow depth look like? What can one learn from that? Where are the values presented? As mentioned above, this needs some explanation.
We agree with the reviewer and changed the wording of reference value to: "An additional density measurement…." The calculation of this value is explained in the same paragraph (Page 3591, l.1-3) and not in the methods, as this is the only time were this method is used, solely to increase the amount of mean density values and to better illustrate the comparison of these values.
The comparison of the mean values (reported on Page 3591 l.3-5) shows that wedge cutter gives the lowest mean density and box cutter the highest density. In terms of SWE, box cutter would give the highest SWE. We could not, however, perform the calculation of SWE, as the CT profile was sampled without overlap, i.e. no continuous measurement needed to calculate SWE could be achieved.
However, we believe the mean density of the entire snow profile is valuable information (i.e. for single layer microwave models) and its comparison necessary.

Page 3591, Line 6: Again, taking it right – here starts the method b part? Or is this the comparison to the layer density (as it reads "density per layer")? It reads like the mixture of both... Same Page, Line 13: Why these thresholds?
We changed the term density per layer to *cylinder cutter*.
The thresholds are for over/undersampling of the cutters with respect to the Ct density (Fig. 4). Their calculation is explained in the methods section 2.3, Page 3588, l. 12-15.

Same Page, Line 16: Is this method c? What is the difference to the lines above (10 – 13).

Method c allows a more objective comparison without setting one instrument as reference (see section 2.3), as done in lines 10-13.

Same Page, Lines 23 – 28: A definition /introduction of the phrasing 'unresolved variation' is missing and how it is estimated and why. For here and the discussion part it would be interesting to discuss this with regard to the variability within a 'layer', variability due to thin layers, which are not considered or the variability 'lost' by merging layers (section 2.2.3) with adjacent layers or variability due to ice crusts. Otherwise these values do hardly have any meaning.
We agree that the definition of the unresolved variation was too short and included the following sentence: *"The unresolved variation is the density variation within a layer. It was calculated as standard deviation of the CT density within a certain vertical distance. For instance, for the 100cm3 box cutter, which had a vertical resolution of 3 cm, the CT profile was averaged to 3 cm vertical resolution and the standard deviation for each 3 cm window was derived. The mean of all these standard deviations was then defined as unresolved variance (in this case for the 100 cm3 box cutter with respect to the CT density)."*
As such, the unresolved variance also included the effect of merging layers, as the CT profile was subsequently averaged to coarser resolution, i.e. subsequently two adjacent layers were merged. The unresolved variance is discussed in section 5.2.3.
Thin layers are discussed also within the unresolved variance section 5.2.3. on Page 3596 l.10ff, where thin density variations are mentioned to be critical for several properties. Thin layers are illustrated in Fig. 8, where a close-up of the high resolution CT profile (as well as a 3D reconstruction) is compared to the methods with coarser resolution.
The variability due to ice crusts is discussed in section 5.2.2., statistics are given in Table 6.

Discussion Page 3592, Lines 6-11: What is the authors conclusion from the results then?
Conclusion are given in section 6. For the lab data, the conclusion is that it resembles the field data, i.e. that all methods agree within 5 to 9%. However, for the discussion here we wanted to point out that our results (oversampling of the box cutter) are in agreement with Carroll 1977, even so neither of the snow blocks used in the lab consisted of the snow type for which Carroll1977 in particular reported the oversampling (light snow, depth hoar).

Page 3592,Lines 13: Again, it would be very helpful for the reader , if this chapter was structured according to the methods(a-c). It is done bit in the following lines, by adding the related method in brackets, but this could be made much clearer, by having separated paragraphs and the first sentence related to each method.
We excluded an extra sentence repeating the method a, b or c in the beginning of each paragraph for better readability and in order to keep the paper short. Summarizing each method in one sentence is a bit too short, and we therefore preferred referring to the methods in brackets, where the precise description can be found.

Page 3592, Line 17: In the method/results part – it is explained, that the measurement methods (3 cutter plus CT) are compared to the mean value. What is meant here by "traditional stratigraphy" then? Why is this profile not shown somewhere?
We changed the term "traditional stratigraphy" to *"cylinder cutter"* in the entire manuscript. This profile is shown in Figure 3.

What can one learn from this approach (related to the question concerning the results part) and what is the authors conclusion from the results?
Please refer to the answer to comment Page 3590, Line25 ff, second part.

Page 3593. Line 12: Repetition of lines 8-10 of previous page?
Yes, as it applies for both field and lab results.

Page 3594 Lines 4 – 10: What is meant by 'traditional stratigraphy' here and where can one see it in figure 8 (as the reference is give to figure 8)?

See answer on comment on Page 3592, Line 17.

No line shown in figure 8 follows the description given here ('highly detailed representation of specific types of density variations....contrasted by a very coarse representation in the lower part..'). What is meant by 'specific types of density variations'?
In the manuscript it is written " Traditional stratigraphy showed a highly detailed representation of specific types of density variations such as ice layers….". As such, the specific density variations are ice layers.

Line 14ff: same problem as above, what is meant by 'traditional stratigraphy' and where is it shown? One could assume at one point, that the box cutter measurements are named as 'traditional stratigraphy', however in lines 14-15 the box cutter is compared to 'traditional stratigraphy'... Because of this, it is hard to follow the argumentation given in this chapter.
See answer on comment Page 3592, Line 17.

Line 21ff: What is meant by 'introduced by the observer'?
See answer on comment on Abstract Line 18.

Line 24: Why not? At least this would improve the study and add some new aspects to this topic.
We agree and discuss the results of previous studies:

Page 3594 l.24: The effect of different stratigraphic representations on microwave emission modeling was unambiguously demonstrated. Durand 2011 estimated the error in retrieved snow depth from PM simulations up to 50\% due to neglecting stratigraphy. Rutter2014 showed that the bias of a three-layer representation of a tundra snowpack with respect to microwave emission was half of the bias for a single layer representation. For the validation of snow cover models, Monti et al.(2012) mentioned the higher number of simulated layers than observed ones to be critical.

Page 3595: Lines 5-7: strange sentence Lines 8-10: unclear sentence Line 17: What about sample 9 shown in figure 8; at 104 cm depth there seems to be an ice crust? With the resolution of the CT an ice layer should be detectable and with some image processing the density of this layer should be possible to estimate.
Unfortunately, the CT samples did not contain any ice crusts, as mentioned in line 17. The 3D reconstruction shows that the layer at 104 cm depth is not an ice crust but only a layer with higher density.

Page 3596 General: same comment on this issue as given above: "Please define /introduce somewhere your meaning of 'unresolved' variation and how it is estimated why. For here and the discussion part it would be interesting to discuss this with regard to the variability within a 'layer', variability due to thin layers you do not consider or the variability you 'loose' by merging layers (section 2.2.3) with adjacent layers or variability due to ice crusts."
See answer on comment Page 3591 Lines 23 – 28.

Table 1: chose a more common currency, why is this value added anyway, as it is not discussed in the paper?
We changed the currency to Euro. The paper focused only on the performance of the instruments independent of their costs. However, the cost can become a major practical limitation when it comes to buy instrumentation, so that we decided to include it for informational reasons in the table.

Table 2: What is the depth of 'bottom'
"Bottom" refers to the snow-ground interface, the bottom of the snowpit.

Figure 3: add what is called 'stratigraghic method' , 'traditional stratigraphy' , and/or show the layers and the related 'mean densities' at least as a scetch
See answer on comment Page 3592, Line 17.

Technical corrections: Page 3582, Line 23-25: redundant Page 3584, Line 21: n missing (known)
Changed

Page 3589, Lines 1-18: you have an extra cubic over each number
15^3 refers to 15x15x15, here 15 denotes the length of a square cube.