General comments: This paper presents the snow pit and SnowMicroPen measurements over sea ice to recalibrate the SMP density model. The calibrated snow density and machine learning-based layer classification are combined to estimate density and length scale of variability differences in the composition of snow layers. Such density model and data are highly valuable in sea ice altimetry application as mentioned by authors. This in situ and model work are important in snowpack properties analysis and will draw wide interests from the community. This article is well-written and easy to follow.

Thank you for your review and helpful suggestions to improve the paper. We have made changes throughout section 3.2 to improve our description of how the SMP and density cutter measurements were compared. We have also revised some of the statistical descriptions as suggested. Inline responses to suggestions and questions are provided in bold below.

My major comments are as follows: Section 3.2 about how to estimate density form SMP profile is not quite clear to me in P6, L168 'Estimates of _smp were then extracted...'. From my understanding, what you are doing here is more like getting the original 5cm-thickness _smp profile scaled according to perturbed thickness inn individual layer. What do you mean by "average the scaled profiles within 3-cm height of cutter measurements"?

Once the SMP profiles are scaled we simply take the corresponding SMP values at the same height of each density cutter measurement. Because the density cutter is 3-cm in height we average the much higher resolution SMP estimates to make a 1:1 comparison. There is scope in the future to optimize how this comparison is made but we have not completed an extensive evaluation here. We modified the sentence on line L169 to make clear what is being averaged.

What does it mean by: "Another 6cm window moving averaging"?

We hope that the above response clarifies how the matching process was applied. However, we have not discussed a 6 cm moving average and are unsure which lines this comment is referring to.

P6, L180 and Figure 4, when you compare the density, do you compare each layer mean snow pit density and all SMP profiles estimation at that layer in one site?

Comparisons described here were between each density cutter measurement and the mean of the SMP estimates within their corresponding 3-cm height. Effectively each point in Figure 4 represents a single density cutter measurement. We've made small improvements to the text in an attempt to make this clear.

I noticed that in Eureka, one site has 2 or 3 pits (the distances between these pits are under 100m), how to divide the SMP measurements for these pits if SMP shave the same distance between two sites?

This is correct, not all sites have the same number of pits, and at times, they are unequally spaced. Placement of the snow pits was structured to characterize inter-site variability but the distance between each was not considered as part of our analysis. All analysis of spatial variability used the

distance between SMP profiles (GPS located). We relied on large SMP data volumes rather than strict spacing of profiles to understand scales between 0 and 100 m.

In section 3.4, when you use SVM to classify the snow layer's type, with 75% accuracy, have you tried other machine learning methods and have you tried other non-linear kernels except for the linear one? What is the accuracy in other methods, and what are the potential limitation of such methods in classifying snow properties?

Thank you for highlighting this important area of future work. We chose to apply a linear kernel with the SVM to limit complexity and focus on broader aspects of the density analysis. There are certainly non-linear divisions within the parameter space which the hyperplanes fail to delineate, limiting accuracy. To apply a non-linear kernel would require an extensive evaluation of the hyper-parameters which we feel is beyond the scope of this work. The work of Havens et al., (2012) stands as a strong example that enhanced SMP classification methods can be applied to improve accuracy. We hope to conduct an extensive assessment of other classification methods in future work.

P7, L219, what is the vertical resolution when snowpit and SMP measurements are both trained considering their vertical resolutions are different. Also, I am very curious about the results when further adding ice type information in the training.

Adding ice type information resulted in a small improvement of ~2% accuracy. The example we created will remain in our revised public code for reference. We chose not to use this configuration as ice type as ancillary information is not directly available from the SMP.

P9, L268, 'Profiles collected on FYI, and therefore exclusively near Eureka...'. Do you mean in Figure 7(a), over FYI, the distribution is negatively skewed? But from the figures, the density seems positively skewed over FYI. Also, the following sentence 'In contrast, densities on MYI were positively skewed...'. Please check it.

We have revised wording throughout the paper to use left- or right-skewed instead of negative or positive. The distribution in question is now described as left-skewed (Statistical skew of -0.41).

P9, L277, 'Measurements classified as faceted had on average a density...'.Figure 8c is over depth hoar not faceted and the distribution is not negatively skewed.

Thank you for noting the incorrect label, we have corrected this. See our previous comment regarding skew.

How to quantify the density uncertainty/error from the SMP density model in consideration of application on altimetry studies?

Errors quantified in the study showed the SMP to be comparable to those expected from manual density cutter measurements. We hope to use this information to build a more comprehensive

analysis of errors involved in altimetry of sea ice. However, we do not have any specific conclusions at this point on how best to address uncertainty when applying the SMP to altimetry studies.

Specific comments: P2, L50, 'Laxon et al. 2013' should be 'Laxon et al., 2013' P9,L271, 'However, these difference...' should be 'However, these differences' P13, L381,'however the errors appears' should be 'however the errors appear'

Thank you for noting these errors. Each has been revised as suggested.