

Response to Reviewer 1

(line numbers refer to original manuscript/current manuscript)

line 3/3: I suggest reformulating the sentence “Boreal soils [...] organic soils. *I tried rewriting this sentence several times and have a hard time rewording it as suggested so I am leaving the text as it was.*

Line 4/36: Is there a difference between “layers” and “horizons”? You use both terms throughout the entire manuscript. I suggest you define that in the methods or stick to one of these expressions. *Layers and horizons are the same thing. This is now clarified in line 36. In addition, the text of the manuscript was changed so that horizon is consistently used throughout.*

Line 16/16: Please reformulate the sentence “These soils [...] organic soil layer. *Sentence simplified.*

Line 22/22: Please better link the last sentence of this paragraph to the rest of the paragraph or highlighting the importance of it. *A sentence has been added to link the last sentence to the rest of the paragraph: “Thus, both organic and mineral soil play an important role determining the amount of C stored in boreal ecosystems.”*

Line 24/24: Please add a source for the first sentence of the paragraph. *Reference added. In addition, the text of this entire paragraph has been strengthened and additional references have been added.*

Line 24/n.a.: Replace “is” with “are”: *This text no longer exists (see previous point).*

Line 30/36: layers or horizons? *This issue is now corrected (see line 4 response).*

Line 35/41: I suggest writing “C and N” instead of C/N. This might be misleading. C/N is often understood as C:N ratio. *This entire sentence has been rewritten so this request no long applies (see next response).*

Line 36-48/41-62: I suggest writings “Fires affect...”; I suggest replacing “several” with “multiple”; “First” but where is the second and third in this paragraph. Please restructure the paragraph and make it more clear, which are the several ways boreal soils are affect by fire; I suggest reformulating “the amount of which”. *This paragraph is now rewritten with these comments incorporated and now reads:*

The main disturbances that affect boreal soil properties are fire and permafrost thaw. Fires affect boreal soils through the combustion of litter and surface organic layers (as ground fuel; Harden et al., 2000), with the amount and depth of combustion regulated by fire severity (Turetsky et al., 2011). Fire directly effects surface organic soils, both in elemental composition and structure (Neff et al., 2005). In addition, there are indirect effects of fire on soil properties. The loss of insulating organic soil results in a darkened soil surface, which in turn warms post-fire soils, increasing decomposition rates from the surface downward (Genet et al., 2013; O'Neill et al., 2002). In addition, both fire return interval and fire severity influence post-fire vegetation and the re-accumulation of organic soil layers. As different tree and understory species have different amounts of C and N in their tissues (Van Cleve et al., 1983), changes in post-fire vegetation affect soil C and N accumulation rates and thus, the concentration of these elements in surface soil. Permafrost thaw also affects soil properties in several ways. By definition, thaw exposes older, previously sequestered C to warmer soil temperatures (Osterkamp et al., 2009),

increasing rates of decomposition (Mu et al., 2016; Schadel et al., 2016). In well drained sites post-thaw conditions usually result in water draining from the soil, resulting in oxic conditions (Estop-Aragonés et al., 2018). In lowlands, permafrost thaw often results in subsidence and inundation, changing the ecosystem from a forested permafrost plateau to a thermokarst wetland (Schoor et al., 2015). Fire can often be a trigger for this rapid permafrost thaw (Myers-Smith et al., 2008). Post fire vegetation changes affects both C and N inputs, again affecting the concentration of these elements within surface organic soil layers. As both fire frequency and permafrost thaw are expected to increase in the future (Hinzman et al., 2005), biogeochemical models have a need to characterize how these disturbances will impact C and N stocks. To accurately represent future scenarios, models need to include the distinct properties of organic soil horizons found in the boreal region (Flato et al., 2013).

Line 41/49: How does post-fire vegetation affect the chemistry of C and N inputs to the soil? Please add another sentence of give an example. *This paragraph is now rewritten, with this issue addressed. Please see above response.*

Line 47-48/58-60: This is a hard transition from these two paragraphs. Please try to better link these paragraphs: *This paragraph is now rewritten, with this issue addressed. Please see above response.*

Lines 52/71: I suggest writing “more than 3000 observations” instead of “> 3000 observations”. *Wording changed.*

Line 62/79-87: I suggest adding a sentence over which time period the samples were collected. Could you add information on the depth of sampling? How many locations and sites were sampled? In the data set I find 57 different locations with coordinates but in chapter 3.7 it is written that more than 290 soil profiles were sampled. Please state that in the text. *The beginning of the field methodology section now gives details on the number of sites, cores, and profiles taken as well as discusses the depth of sampling. This paragraph now reads*

Soil cores were sampled at 58 different sites located within several areas of Interior Alaska (Figure 1). Several different ecosystem types were sampled, including black spruce forests (~50%); wetlands (~26%); and deciduous and mixed forests (~16%). Between 1 and 14 soil profiles were sampled at each site, for a total of 292 soil profiles. Sampling took place over a 15-year period from 2000-2015. We examined the effect of fire or permafrost thaw disturbance on soil properties by categorizing each of the soil profiles in relation to time since the last disturbance, which we divided into three age classes: new (<5 yrs old), young (5 – 50 yrs old), and mature (> 50 yrs old). All new sites were recently burned and thus had lost some portion of their surface organic horizons (Harden et al., 2000), while young sites experienced either fire or permafrost thaw.

Line 77/130: I suggest naming this classification as Table 1. *This information is now included as a table. In addition, we added more text to Table 1 to expand upon these definitions. The manuscript gives very brief descriptions of the six main horizons: “A description of the horizons and the codes we used to represent them are found in Table 1, but in summary there are six main horizons: live moss (L), dead moss (D), fibric (mostly undecomposed; F), mesic (more decomposed; M), humic (very decomposed; H), and mineral soil (Min).”*

Line 93-94/132-134: The references here are basically the same, just different editions. I suggest choosing one and naming it Soil Survey Staff (1993) in the reference. *One of these references is for the US systems, while the other is for a Canadian manual so both references are still included. The references have been fixed in the citation program to provide the entire author name, making this fact clear.*

Line 98/94: I suggest replacing “in that” with “because”. *We prefer the original wording.*

Line 101/98: Redoximorphic – check the spelling. *Spelling corrected*

Line 136/175: “In the field the best call was made to if it was...” I suggest re-writing this sentence something like “in the field visual inspection of the soil samples gave a first indication.” *Rewritten as “Using visual and textural cues the field, horizons were categorized as either mineral (< 20 % C) or organic (≥ 20 % C).”*

Line 140/180: Please add “R core Team” to the source. Also, I could not find this reference in the reference list. *This reference was included, but the way in which it was entered into the citation program had it appearing as ‘Team, RC’. This citation is now fixed.*

Line 142/183: Please add “for significant differences among the different soil horizons” after “was tested”. *Verbiage added.*

Line 169/208: Please write “were not” instead of “weren’t”. *Sentence rewritten changed.*

Line 178/223: “likely due in large...” I suggest rephrasing this sentence. *Wording changed.*

Line 204/250: I suggest naming this sub-chapter “soil horizon thickness”. *Heading title changed.*

Line 215/262: Please write “was not” instead of “wasn’t”. *Wording changed.*

Line 216/263: Please write “it is” instead of “it’s”. *Wording changed.*

Line 222/270: I suggest adding another main chapter for the two subchapters 3.7 and 3.7 since it is more a discussion chapter. I suggest adding 4. Discussion of the data set” and then include 3.6 and 3.7 there. In addition, my question is, how does your data set relate to the soil pedon carbon and nitrogen data for Alaska by Michaelson et al. (2013). Maybe you can refer to that during the discussion and indicate how your data set adds or fits within this data set. Also, in a discussion chapter you could state again why your data set is so value. *The headings were changed as suggested. In addition, you ask how our data compares to Michaelson et al, 2013. This paper, as well as Ping et al, 2010, are now brought up in the Introduction (lines 66-69) as well as we now discuss how our data compare to Michaelson (2013) and Ping (2010) in section 4.1:*

Our data are the first of its kind to present organic horizon data across a range of Alaskan boreal ecosystems. Other studies have examined organic soil as a separate entity from mineral soil but with certain limitations. Michaelson et al. (2013) used Alaskan USDA-NRCS soil pedon data to examine soil properties of both organic and mineral soil but present these data for the organic portion as a whole. This study shows that there is significant variation in bulk density and C and N concentration across organic horizons, and therefore, one should not disregard these horizon-based variations. In a separate study, Ping et al. (2010) separated the organics into two horizons from boreal black spruce stands (Osurface, Oe/Oa). Our study supports the

results of Ping et al., (2010), which found a decrease in C:N ratios with increasing depth. Moreover, our study provides data from a fuller suite of soil horizons and includes data from bogs, fens, and deciduous forests.

Line 238/297 & 312: I wonder whether the accuracy of the bulk density measurements could be a reason for the differences, too. Often it is difficult to accurately measure bulk density, therefore I could think that maybe the accuracy of the bulk density measurements in both, the reference and your data set might be a reason for the differences. *Yes, bulk density measurements are hard to get. This idea is now acknowledged in both paragraphs of section 4.2: "In addition, accurate bulk density measurements is time consuming to do correctly (Nalder and Wein, 1998) and could also play a role." And "Thus, bulk density measurements play a role in these differences."*

Line 239/299: I do not understand the first sentence. What do you mean with "previous results"? Please, also consider restructuring this sentence. *This wording is now changed to be clearer and reads "To determine if the above findings..."*

Lines 246, 261, 264, 265, 274, and 282: Please insert a "the" between "than" and "measured"; Please insert a "the" between "into" and "mineral"; Please add Boreal in this sentence; I suggest writing "drained sites" instead of "drainages"; Please change "is" to "are"; I suggest writing "lack input data for" instead of "do not do a good job". *Wording changed.*

Line 264/330: I suggest moving the first sentence of this paragraph to the conclusions. *This paragraph has been rewritten.*

Figure 1: I suggest writing "included" instead of "used" in the figure caption. I suggest to improve the map because the scale bar is hardly visible and the map looks a bit blurred in general. Maybe add more names in the maps for the regions or locations where the sites are. *Region names and additional city added to the figure. Scale text enlarged and made white to improve readability. Resolution of image made the maximum size possible. Figure caption modified accordingly.*

Figure 2. I like this figure. However, I suggest to add photographs with a higher resolution. *The size and resolution of the photographs have been increased. In addition, if helpful, we can provide the original photographs to the journal.*

Figure 3. This figure shows that humic soils are most important in C and N storage. Maybe you can mention this in the text as well. *The role of lower organic horizons in C storage is now called out in section 3.2: "Therefore, even though the deeper organic horizons (M and H) have slightly lower C concentrations than the shallow horizons, their high bulk densities result in large amounts of C at depth. In fact, given average thickness, bulk density, and % C (Table 2), approximately 75% of the soil C is stored in the mesic and humic soil horizons."*

Table 1 & 2 (now Table 2 & 3) plus Table S2-S5: I have some troubles understanding the tables and whether it is significant different or not. There are a lot of superscript letters, sometimes the same, sometimes two or three. While I acknowledge the effort in putting everything into one table, I would suggest to make a separate cross-table for the p values and whether it is significant different or not. The same for the tables S2-S5 in the supplementary material. *The formatting of these tables has been redone to help make comparison of the superscript letters easier.*

Thank you for publishing this very valuable data set. *Thank you for your suggestions on how to make this manuscript better.*

Generalized_models_for_C&N_Alaska.csv: You write in the methods section that four different methods were used in collecting the soil cores. I suggest adding a column to the data set indicating which method was used for the collection of the samples. *This information is now a part of the dataset (see Table 5).*

Site_GPS_coordinates.csv: Could you add the key for the abbreviations of the regions and sites. I could not find it in the metadata what e.g. HCCS or BZ means. Also, it would be nice if the sites could be found in Figure 1, e.g. by adding the region names to the map. *This information is now added to the file (see Table 6) and mentioned in Figure 1.*

Response to Reviewer #2

Line 19/Figure 1: Add some more town names to Figure 1. Add footnote that several profiles were sampled at a given site. Include more visible North arrow and scale bar (in SI units rather than miles). Region names and additional city added to the figure. *This information is now included.*

Line 20/Figure 2: Picture bottom right. Legend includes two drainage classes? How have such cases, when arising, been processed in the dataset itself? *The caption for Figure 2 is now modified to make it clear that gleyed soil can be found in both very poorly and poorly drained soil. The determining factor between the two classes is the length of time of saturated soils, as mentioned in the table.*

Line 35/41: Change C/N to C and N dynamics to avoid possible confusion with C:N ratio. *Sentence rewritten*

Line 37/41-62: There is no second or third, please rephrase this paragraph. *Paragraph rephrased. Please see the revised text in response to Reviewer 1 above.*

Line 52/71: over 3000 thousand. *Text changed.*

Line 77/Table 1: Possibly present this as a table (new Table 1): code (X), name, description. *This information is now presented as Table 1.*

Line 156/195: one progresses. *Sentence rewritten*

Line 157/196: express bulk density as g cm⁻³. *Text changed.*

Line 187/36: Both horizon and layer are used. Are they used as synonyms or were (thicker) horizons divided into separate layers for sampling? Please clarify this. *Layers and horizons are the same thing. This is now clarified in line 36, but the text of the manuscript was changed so that horizon is consistently used throughout.*

Line 190/235: than then → (change to) than the. *Text changed.*

Line 204/250: Change Thickness to Horizon thickness. *This header has been changed to "Soil horizon thickness".*

Line 222/270: I would suggest you move subsection 3.6 and 3.7 to a new Discussion section (4). This section could also include a comparison with results derived from other studies for boreal regions. *The headings were moved as suggested. In addition, we have added a comparison of our study to the two other studies we know of that discuss organic soil properties in section 4.1. (See response to Reviewer #1, line 222 above.)*

Line 274 – 278/336-343: The authors should at least indicate that the units of measurement, respectively domains for observations, for the properties under consideration in the two csv-files can be found in file Mega-AK metadata.xml. However, I would recommend this information is also summarised in an Appendix. As indicated, it would be a ‘plus’ for this data paper if the underpinning ‘raw’ profiles could also be made available as supplemental information, as they are not presented in <https://doi.org/10.5066/P960N1F9>, rather than just pointing at several open file reports (pdf’s) as is now the case. *There is now a short description of the data found within the ScienceBase publication (“This publication includes both .csv data files as well as metadata. A short description of these files and the data found within them can be found in Tables 5 and 6.”) as well as two Tables that list the column names, units (if applicable), and column descriptions. I would also like to clear that all raw data used in this study are within this ScienceBase publication. The reason to point the reader to the Open-File Reports is because there is additional information (such as Von Post descriptions and ²¹⁰Pb data) that may be of interest that are not in the ScienceBase publication.*

Figure 3: What depth interval is considered in this figure? The unit of g/cm² is rather confusing (a typo?). *This graph is of C and N density and was incorrectly labeled. Therefore, depth or thickness is not included in the calculation. Thank you very much for spotting this labeling error. The figure caption and graph axes have been fixed.*

Table 1: Bulk density given as g/cm², this should be g/cm³. *Units fixed.*

Tables 1-4: For legibility, and future typesetting by ESSD, it would be better to create three columns for each row (e.g. bulk density): n, mean, sd. The symbols for statistical significance would then also become more ‘legible’. *The formatting has been redone for these tables to help make comparison of the superscript letters easier.*

Table 3 (now Table 4): Bulk density, should be g/cm³. *Units fixed.*

Under ‘Data access’ briefly describe the content of the zip file (csv and xml). Further, please provide an Appendix that describes the content of the csv files. *There are now two tables (5 & 6) describing that data available in the two .csv files.*

Site_GPS_coordinates.csv: There are 57 sites, yet the paper refers to over 289 profiles. It would be useful to know how many profiles were sampled at each site without readers having to digest this from file (Generalized_models_for_C&N_Alaska.csv). Further, the abbreviations for regions and sites should be provided, preferably in a look-up table (i.e. as a separate csv file). Please note that data in row 57 have ‘shifted’ to the right; this should be corrected. *We do not see a good place to put the number of cores per site in the text and feel that, if needed, this information is accessible using the data file. Region abbreviations in addition to names are now included in this file (see Table 6). Site names do not have much meaning outside of the research group, so we did not create a lookup table for this information.*

Generalized_models_for_C&N_Alaska.csv. Specify units of measurement (depth (cm), bulk density (g/cm³), 13C etc.); explain all codes/abbreviations used in the file, as a 'look up' table (i.e. as a separate csv file). *This information is now included in Table 5. Tables 1, 5, and 6 have been added to the ScienceBase publication site as .txt files.*

Supplemental information S2: Please add units for bd table. *Units fixed.*

Additional improvements

Since our initial submittal of this manuscript we have come to realize that the original test used to determine differences among drainage types or age class, *diffsmeans*, does not correct for multiple comparisons. Therefore, we redid our analyses using estimated marginal means (*emmeans*; line 185), which does this correction, the result of which made some of significant differences originally presented in Tables 2-3 and Tables S2-S5 no longer significant. This change did not alter our conclusions in any way.

To aid those interested in better understanding the predicted versus measured relationships discussed in section 4.2 we have added graphs showing those results to the Supplemental Information as Figures S1 and S2.

We have added a paragraph (line 318) suggesting that, due to the inherent variability of thickness measurements, that we recommend that researchers continue to measure thickness at their sites and only use our bulk density and concentration data. This combination minimizes errors while allowing researchers, if needed, to bypass soil sampling and processing, both of which are quite labor intensive, and, thus, not always possible. The new text is below:

One of the important uses of this dataset is the potential for estimating C and N stocks based on simple field characterizations of organic soil horizons of North American boreal forests and wetlands. Because soil sampling and processing is quite time intensive, researchers may decide to measure thicknesses of the various soil horizons within their sites, using the descriptors in Table 1, and then calculate C and N stocks using the average values presented in Tables 2, S2, S3, or S4. This approach minimizes errors associated with the high variability found for horizon thicknesses, due to variable site histories.