Review of : soil-2019-41
Strong warming of subarctic forest soil deteriorated soil structure via carbon loss – Indications from organic matter fractionation
Christopher Poeplau[1], Páll Sigurðsson [2], Bjarni D Sigurðsson [2]

General comments
Long-term soil warming experiments are rare and always welcome. As so are organic matter fractionation studies that aim at comparing two ecosystem (grassland and forest). In large this is a good and interesting paper. However, there are some short-comings with regard to these types of studies. The authors themselves have discussed problems connected to interpretation of results from experiments using geothermically warmed soils in a global warming / climate change context, concerns that I also share. Though I am not convinced that the study is bringing us much forward in questions regarding the fate of carbon in subarctic forest soils in future warmer climates, I still see the relevance with regard to the fate of soil carbon in geothermically warmed Andesitic soil. I believe the most important results in this study are the relative changes in proportion of C in the different fraction with increasing temperature and this should be more in focus than the loss of C and deterioration of soil structure. I also think that the differences between the two ecosystems (grassland and forest) should be better communicated in the tittle.

More specific comments:
1. Does the paper address relevant scientific questions within the scope of SOIL?
   → yes
2. Does the paper present novel concepts, ideas, tools, or data?
   → the paper presents new date, but is not particularly novel in concept, ideas and tools.
3. Does the paper address soils within a multidisciplinary context?
   → Yes, in a using geothermic warming of soils as a proxy for warming of soils in climate change scenario. But does not address the ecosystem changes/ vegetation as much it ought
4. Is the paper of broad international interest?
   → Yes, but not as broad as the title suggests, these are Andesitic soil →and thermal warming og soils do have some limitations with regard to interpretation in a global change context.
5. Are clear objectives and/or hypotheses put forward?
   → Yes, three objectives are stated.
   1. advance our understanding of the temperature response of different SOC fractions representing kinetic pools
   2. assess the role of the ecosystem type in the temperature response of SOC
   3. investigate potential links between SOC loss and soil structure changes.
6. Are the scientific methods valid and clear outlined to be reproduced?
   →I have some questions with regard to sampling and interpretation of the term soil structure, see below.
7. Is the soil type/classification adequately described?
   → fairly, general information on soil type/classification at the experimental site is given. but I cannot see that the information that the soil

type/classification provides is actually used in the interpretation of the results. Though only the upper 30 cm is used in this study,- it would have given valuable information if this was related to soil horizons.

8. Are analyses and assumptions valid?
   → see comments below
9. Are the presented results sufficient to support the interpretations and associated discussion?
   → see comments below
10. Is the discussion relevant and backed up?
    → see comments below
11. Are accurate conclusions reached based on the presented results and discussion?
    → yes
12. Do the authors give proper credit to related and relevant work and clearly indicate their own original contribution?
    → Yes
13. Does the title clearly reflect the contents of the paper and is it informative?
    → See comments below
14. Does the abstract provide a concise and complete summary, including quantitative results?
    →Yes
15. Is the overall presentation well structured?
    → yes
16. Is the paper written concisely and to the point?
    1. → OK
17. Is the language fluent, precise, and grammatically correct?
    1. → long cumbersome sentences make particularly the discussion, but also elsewhere. You are really not making it easy for the reader when you write sentences like this e.g.: Page 9 lines 30 to 33 " Thereby, topsoil and subsoil samples scattered approximately around the same regression line, indicating that especially the amount of young SOC, may it be driven by warming intensity or the location within the soil profile, was a very good predictor for the amount of stable aggregates in the soil." Or page 11 line 24-26: "Together with the fact that also more labile SOC was present in the forest topsoil, which responded more sensitive to warming than the bulk soil, it seems likely that amount and fraction distribution of SOC drove the ecosystem specific warming response in the topsoil."
18. Are the figures and tables useful and all necessary?
    1. → yes (see point 20)
19. Are mathematical formulae, symbols, abbreviations, and units correctly defined and used according to the author guidelines?
    1. → I believe so
20. Should any parts of the paper (text, formulae, figures, tables) be clarified, reduced, combined, or eliminated?
    1. → most of the text is relevant and to the point, but I am unsure if the "poured bulk density - soil structure" part contributes in a meaningful way. Deleting it and giving more focus to the warming effects on fractionation - ecosystem comparisons – stable soil C would benefit the paper.

2. → Most of the figures are nice and useful. Figure 5 must have the wrong Y-axis tittle – subsoil not topsoil SOC. However, I wonder if it would be possible to add the warming (some sort of colour code) to Figure 6. It would be interesting to see if there were any systematic also with regard to temperature.
3. → I miss a table showing general soil properties such as pH, oxide extraction of some sort (Fe, Al, Si), sand silt clay%. I am not too keen on table 2 and 3 which only show summary of statistics, they are much more valuable when they are connected to measured properties, merge or delete?

21. Are the number and quality of references appropriate?
    1. →Yes
22. Is the amount and quality of supplementary material appropriate and of added value?
    1. → as far as I can see yes

I have a couple of more specific comments that I think the authors need to address to improve the paper

1) I think the use of the term "soil structure" is used wrong. Soil structure has not been investigated in this study. What has been studied is stable aggregate (SA) (63-2000µm) from the fractionation procedure and the carbon (C) connected to this fraction. The natural soil bulk density (BD) was not measured – which could have given some indication of soil structure and the "poured bulk density" does not replace the measurement of the natural BD, though it does give a relative difference between dried soil material in the fraction less than 2 mm. I therefore suggest a change in tittle. e.g.: «Strong warming of subarctic forest soil reduces stabilisation of carbon in soil aggregates – Indications from organic matter fractionation". I would also suggest including "subarctic andesitic forest soil" in the tittle as these soils normally show both chemical and physical properties that are markedly different from "subarctic forest soils" formed on other parent materials.

2) Sampling method and comparisons, sampling should be better explained, and it would also be nice to know how many samples (N=) were behind each temperature/location?

- In studies like this sampling method and a good description of these are crucial. Many sophisticated analyses in the laboratory will never compensate for errors, flaws and inaccuracy in sampling – or description of sampling. I am afraid that the lack attention to the sampling procedure may make the results of this study none reproducible. Personal experience with similar sampling schemes suggests that a soil of this nature (Silandic Andosol with a silt loam texture, 60 % silt) is easily compressed during sampling. How were the depth intervals determined? If sampling was done as I anticipate by extracting a 30 cm cylindrical core and then splitting it into 0-10, 10-20 and 20-30 samples this could be a real challenge. The comparisons between the different layers could be based on pure artefacts - please convince me of the opposite.

- This brings me to my next point – if the warming of the soil has caused changes in the soil density particularly in the "top soil" this would cause the sampling at the warmer place to go deeper into the subsoil extracting soils that naturally (before the shift in geothermal flow) had a lower content of C. This would then be compared to the lower layer of the "none" warmed soil and we would wrongly conclude that the warming has caused loss of carbon?
- Bringing me to my third point –as one of the main objectives of the study clearly is to assess losses of C and also quantify the losses, we need to be sure we that what we compare are comparable. Studies like this should be done by comparing equivalent mass (See Ellert, B. H. and J. R. Bettany (1995). "Calculation of organic matter and nutrients stored in soils under contrasting management regimes." Canadian Journal of Soil Science 75(4): 529-538 or others more recent paper). Also the 10-20 cm was not analysed, understandable from a resource point of view (many time-consuming and expensive analysis), but a simple analysis of SOC + weighing of the total dry sample would have added valuable information particularly for interpretation in a climate change context.
- The most important results in this study is the relative change in proportion of C in the different fraction with increasing temperature and this should be more in focus than the loss of C.
- All analysis were done on the fraction < 2mm, but it would have been interesting to know the proportion of the coarse fraction.

3) The study appears to focus on loss of C from the soil with warming, however there is little information in change in input of C to the system. Warming of the soil may have had an influence on the forest growth/productivity and litter input. Some more information on the vegetation would have been appreciated. In this study the litter /O horizon is removed – it would have been nice to at least know how thick it was at the different locations? Several papers have been written with data from this experimental area– surly some information could have been extracted from these not just giving references and leaving the reader to find out for themselves. Additional information on mineralogy/alternatively selective extraction of different oxides would also aid the discussion on C stabilization mechanism. Also, DOC and pH normally are correlated – it would have been nice to have some pH measurements to go with the top and subsoil samples. As this warming is by geothermal heating, I am naturally curious to how this is affects the the top- and subsoil, are there any effects on soil moisture, any gradient between the two layers.

4) The use of the term "topsoil" and "sub soil" when this refers to 0-10cm (topsoil) and 20-30 (sub soil) is ill-conceived. I normal soil terminology both these layers refer to topsoil. Readers with an interest in C subsoil – non surface layers will perhaps be misled. Why not simply us "Upper" and "Lower" or even better were there any genetic differences A horizon – B horizon?

5) Ecosystem comparison: I believe there should have been more focus on differences in input of C. You observe differences between the ecosystems only in the topsoil – ascribing this to the fact that the forest was planted on former unmanaged grassland. However, you also find that the forest soil has a more pronounced depletion of C in the subsoil. Could also part of the explanation also

be due to the fact that warming was geothermal – from beneath. In a situation where global warming (air warming) is the case the differences in the topsoil would be equally reflected in the subsoil.