Biogeosciences Discussions

Biogeosciences Discuss., referee comment RC1
https://doi.org/10.5194/bg-2021-295-RC1, 2021

Comment on bg-2021-295

Xavier Dupla (Referee)

Referee comment on "Soil geochemistry as a driver of soil organic matter composition: insights from a soil chronosequence" by Moritz Mainka et al., Biogeosciences Discuss., https://doi.org/10.5194/bg-2021-295-RC1, 2021

An overall excellent paper. Informative, concise and very well written. I am confident that the scientific community will welcome it warmly. Please find, some minor comments below:

line 58: remove thereby which is misleading with the previous sentence

line 61: oxide do not "become" positively charged at low pH values, they are positively charged on the whole pH range of almost all soils (check pzc values). If you want to open the pandora box of variable charges, it is difficult to speak about the protonation of surfaces without saying that several OM functional groups too protonates when pH becomes acidic. Furthermore, if you want to maintain this sentence about acidic soils, then you could more explicitly mention that soil acidification is a key process behind soil weathering.

line 79: this section contradicts what you say from line 47 onwards. Rephrasing either the upper section (lines 47-52) of this one (lines 79-85) might help

Discussion section: overall excellent. However, you did not notice any significant decrease in base saturation along your chrono-sequence which contradicts general description of soil weathering sequences. This aspect is extremely interesting and should be discussed.

Non-binding suggestion : your discussion sticks very closely to the parameters. I was expecting your paper to zoom out at some point in order to 1) discuss how the climate and geomorphological changes that happened in your 3 million-year sequence may have impacted your results 2) discuss the general impact of your findings on our understanding of soil weathering, 2) outline the limits of your study and what should be done to go further.

Biogeosciences Discussions

Biogeosciences Discuss., referee comment RC2
https://doi.org/10.5194/bg-2021-295-RC2, 2022

Comment on bg-2021-295

Anonymous Referee #2

Referee comment on "Soil geochemistry as a driver of soil organic matter composition: insights from a soil chronosequence" by Moritz Mainka et al., Biogeosciences Discuss., https://doi.org/10.5194/bg-2021-295-RC2, 2022

General Appreciation / Overall Comment:

This is an interesting paper addressing a hot topic in the domain of SOC research (i.e. understanding long-term SOC stabilization / decomposition mechanisms). The paper is well written based on the innovative research idea of considering a soil chronosequence and reveals that way some new insights. However, the results section has been worked out rather weakly, in that sense that it is rather brief and descriptive, but more critically, there is a lack of making use of the quantitative measures / evidences (as being represented by the various figures in tables and figures) as well as associated statistical interpretation. Hence, in particular, this section requires extra attention when improving the manuscript based on following more specific comments:

Specific Comments:

L 36: You mention "responds to land use change" but I'm not too sure whether it is that relevant to mention it here, because land use change will have an impact on SOC stabilization mechanisms on a much shorter time period (scale of 10-100 years), whereas you are looking to a timescales exceeding 1000s of years.

L 40 – 65 I like the idea of the interaction between physical and chemical stabilization mechanisms and associated changes over time. But as regards the physical protection, I can see that the focus is mainly on sorption / binding of OM to mineral surfaces (and as such also the importance of clay % and type of clay in this context). However, I was wondering whether soil aggregate formation (macros and micros) shouldn't be considered as well / more explicitly here? (or is this not the right time scale. But if so, I guess you should neither mention 'land use change' – see my comment just above this one).

L66-69 & L 75-85: I think that the information given in these sections mainly belongs to the Material & Methods section. Hence, I would like to suggest to rewrite the end of the introduction section so that you end with making the problem statement clear followed by your main objective(s), i.e. the contribution of present research in addressing this particular problem statement.

L 75-85 I can see that in this section a kind of stepwise approach has been explained. Hence, in that respect I would like to suggest to consider making a methodological

flowchart and use that in the M&M section (see also my comment just above this one)

L 93 You mention climatic condition are homogeneous. Yes, that's true for this particular period in time, but not throughout time. So, as the alluvial deposits are differing in age they have being created under quite different climatic circumstances, and hence, I was wondering whether this fact could have influenced the relative importance of different stabilization mechanisms?

L 99 – 100 I think that this sentence needs some rephrasing: what about somethings as follows: "Soil samples were collected in December 2013 from 1 m3 soil pits located within a circular area with a diameter of approximately 40 km in the North of Merced County."

L 100 you make reference to the fact that this are "glacial periods" but I can see that the considered time-window also includes interglacial periods, and as such I would rephrase this as "quaternary".

L125 Can you specify these correction factors?

L 185-221 As mentioned in my overall comment (see above), I think that this sections requires considerable rewriting, including a much clearer engagement with the quantitative information (as being presented in various figures and tables). Moreover, in some cases a bigger effort could be made to assess the statistical value of a given statement. For example, you say in L189-190 "a strong increase" but it would have much more value to indicate whether this strong increase is significant (and at what level, e.g. $p < 0.05$, $p < 0.01$, ect…). In that respect, I would also like to encourage the authors to undertake a much bigger effort in terms of providing more statistical based evidences as regards the values presented in table 2. More precisely as "n = 30" I guess you could have also added standard deviations and / or standard error values, which on its turn could be of use for the statistical interpretation of the results.

L 195 You make reference to table 3, but I think that isn't the correct, because that's the table representing the output of the regressions. Anyway, please check your in-text references to figures / tables throughout the entire text, because I think they aren't always correct.

L 196-198 I think that this information comes from table 2? If, so please make reference to the corresponding table. Please also check in other parts of the "Results" and "Discussion" sections whether you always make in-text references to the corresponding tables / figures, because in many cases these seem lacking.

L 210 Did you perform a multicollinearity analysis? (if so can you give the correlation coefficient matrix and explain which kind of correlation coefficient threshold you did consider in order to say variables were too strongly correlated?).

L 214 I guess this should be "table 3" instead of "table 4"? Please check the entire text (see also my related comments above (L195, L196-198)).

L214-216 In line of my overall comment (and specific comment related to L185-221) I like to iterate that it is important to mention values as being given in various tables and figures, because the engagement with the quantitative measures (and its statistical interpretation) is very important.

L219-221. I agree that the R2 values can be of use when comparing the models. However, I have my doubts about the usefulness of the RMSE values, because the RMSE values aren't dimensionless and are depending on the value range of the considered variable. Hence, I think that a relatively scaled variant of RMSE could be more useful, e.g. relRMSE

or RPD? In addition, I was wondering whether it could make sense to give (beside a measure for random error) also a measure for the (relative) bias (e.g. %BIAS)?

L234-236 Is this significant?

L 248-250 I think this interpretation should be made with more care, because as the vegetation type did vary over time (as a function of climate variations), the quantity of C-input as well as associated origin and stability will have been different along you chronosequence (see also my comment related to L39). Hence, I think some more critical reflection is required here when interpreting the results.

L 274 – 276 When looking to the coefficients of variable importance in table 3, I can see that Fetotal:Sitotal has a value of 3.2 whereas that of FeDCB is 1.8. Hence, following your statement, I understand that 1.8 is considered as not significant? I that correct? Can you please explain to the readers somewhere (e.g. in the "material and methods" section) how to interpret these values in terms of their statistical meaning?

Figure 1: Can you add an extra subplot to this figure indicating where this study site is located within the western part if the USA / California?

Table 2: This table has multiple layout issues. First of all, I think "30" is missing in the last 5 rows of the first column (so "10-" should be "10-30"). Moreover, I think that it could be a better idea to switch rows with columns because 17 columns re fare too much (I can't read the headings properly) and/or this figure should be made in landscape format. Another suggestion could be to split the table in 2 sub-tables (or 2 separate tables) one considering 0-10 and the other considering 10-30. Finally, I was also wondering why you couldn't present this information in graphs (with "soil age" on the x-ax) just as you did in figure 3 and 4? And last but not least, I like to reiterate my comment about providing more statistical measures, because as "n = 30", I guess you could have also added standard deviations and / or standard error values. (which you actually do in figures 3 and 4).

Figure 3 & 4: I would like to suggest to use a logarithmic x-ax in order to place the soil ages accordingly on it (instead of using just equal distances between the different soil ages, and hence, not having 'a real numerical x-axe').