Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2020-428-RC2, 2021

Biogeosciences
Discussions

Interactive comment on "Development of global temperature and pH calibrations based on bacterial 3-hydroxy fatty acids in soils" *by* Pierre Véquaud et al.

Anonymous Referee #2

Received and published: 3 February 2021

I have finished the review of the manuscript "development of global temperature and p calibrations based on bacterial 3-hydroxy fatty acids in soils", submitted to biogeosciences by Pierre Vequaud and co-authors.

In general, the authors present a good contribution to the developing field of 3-OH fatty acids, by compiling several published and novel altitudinal gradients. The performance of the local and compiled linear regressions are compared with that of brGDGT lipids, for those sites where 5- and 6-methyl compounds were separated (and where the MBT'5ME could thus be calculated). This is an interesting dataset, and a necessary next step in the development of 3-OH fatty acids as temperature proxies.

However, the performance of the three additional models (multiple linear regression, random forest and k-nearest neighbor), that move away from the previously established RAIN, RAN15 and RAN17 ratios, are not evaluated thoroughly. There is no explanation of what the models represent, and why they allow a better correlation between the 3-OH fatty acids and MAAT or pH, compared to the simple linear regression. There is for instance no analysis of residuals, and no indication of maximum reconstructed temperatures. I would recommend the authors to look at a recent calibration paper (fi Dearing-Crampton Flood, 2020), to see what the state-of-the-art in lipid proxy calibration is. As it stands, the authors will have a difficult time convincing readers that the more complex models are the preferred choice when doing 3-OH FA-based climate reconstructions.

When developing a MAAT-calibration, the authors should also mention the following weakness: there is currently only 3 samples with MAAT > 15 âĎČ in the dataset (maximum temperature: +- 20 âĎČ). Do these samples influence the MAAT calibration disproportionally?

Minor comments: L 91. Include 'bacterial' before 'lipids'. L 129. 'more developed statistical approach', perhaps rephrase as 'further development of the statistical approaches'? L 148. For easy comparison: include the altitudinal, pH and MAAT ranges of the previously published transects. L 200-201. Soil sensors have been used, but the calibration is done with mean annual air temperature (MAAT) instead. As the soil sensor data is not used or discussed further, I'd remove their mention here, and use the MAAT for this site as well. L 322 (and further in the manuscript). As far as I can see, the supp. tables referred to are not present. Please include these tables in a revised version.

L 541. Perhaps the authors can refer here to De Jonge et al., 2019, who argue that temperature can modulate the soil bacterial composition, and the dependency between MBT'5ME and soil temperature? L551. Here the authors should explain better how 'more complex models' allow to take the complexity of each site into account. At this

point, in the manuscript, we have observed that the linear dependency between the 3-OH fatty acids is better on a local scale, than on a global scale. What mechanism do the authors propose for this, and how can a more complex model correct for it? L 567. Have the authors done a selection of the fatty acids that are necessary for the model? For instance, forward selection or reverse selection? Does the model not suffer from overfitting? Is there any fatty acids that are generally present in low abundance (and can thus be absent in geological archives) that are important for the regression? If yes, is it prudent to include these low abundance compounds in the model as well? L 587. The random forest model and k-NN model need to be explained much better before the results are presented. What are they based on and how do they compare variability in the lipid distribution with the MAAT? L 593. Same comments as at L 567. L 651. Can the authors comment on the probable source of the 3-OH fatty acids in the speleothem? Can we assume that all lipids are derived from the soil, or is there a (variable) proportion produced in the cave environment as well?

Refs: Dearing Crampton-Flood E., Tierney J. E., Peterse F., Kirkels F. M. S. A. and Sinninghe Damsté J. S. (2020) BayMBT: A Bayesian calibration model for branched glycerol dialkyl glycerol tetraethers in soils and peats. Geochimica et Cosmochimica Acta 268, 142–159. De Jonge C., Radujković D., Sigurdsson B. D., Weedon J. T., Janssens I. and Peterse F. (2019) Lipid biomarker temperature proxy responds to abrupt shift in the bacterial community composition in geothermally heated soils. Organic Geochemistry, S0146638019301275.

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2020-428, 2020.