Reply to Referee #1

by Johannes Hepp, Michael and Roland Zech & co-authors

We are grateful to anonymous Referee #1 for her/his constructive suggestions helping to improve our manuscript. Please find our replies to the individual comments below.

Major issues:

1) The brGDGT calibration presented here is of limited use, since the study uses an outdated method to measure brGDGTs and does not distinguish between the 5 methyl and 6 methyl compounds. Hepp et al. thus calibrate indices (CBT and MBT') that have fallen out of favor and been replaced by the more robust CBT' and MBT5Me indices. The new indices and new methods developed by De Jonge et al. (GCA, 2014, doi: 10.1016/j.gca.2014.06.013) and Hopmans et al. (Organic Geochemistry, 2016, doi: 10.1016/j.orggeochem.2015.12.006) are not even mentioned in the text, and the limitations of the brGDGT data presented here are not acknowledged. Without reanalyzing these samples with a method that resolves all isomers, I fear that the present sample set has limited value for the calibration of brGDGT-based proxies.

→ Referee #1 is right in his/her statement that the GDGT data presented in our manuscript were not acquired based on the up-to-date method. During revision, we will therefore explicitly emphasize that meanwhile new indices and methods were developed (including citations recommended by Referee #1). We would still see a high value of having our GDGT dataset published, because our results fit well to the calibrations done with the previous approach and this in turn allows evaluating GDGT proxy data published for Europe based on the previous approach.

2) There are some big assumptions in the proposed approach for reconstructing relative humidity using paired δ2H values of n-alkanes and δ18O values of sugars. In particular, the assumption that biosynthetic fractionation for these compounds is constant is contradicted by lots of existing work, which is briefly mentioned by the authors in their discussion. Figure 8 is not a very good advertisement for the utility of the paired δ2H-alkane/δ18O sugar approach, and the lack of correlation suggests that some of the many assumptions that go into this method are not valid. This paired approach has not caught on beyond the Zech group, and the data presented here suggests that it may not be useful as presently conceived. The authors state they have shown the "great potential" for this proxy. I remain unconvinced by the data and analysis shown here.

→ We accept that Referee #1 remains unconvinced by our coupled $\delta^2H_{n\text{-alkane}}$-$\delta^{18}O_{sugar}$ biomarker approach. We moreover (i) agree, (ii) are aware and (iii) explicitly state that the assumption of constant biosynthetic fraction is likely a major uncertainty of our approach. Still we are convinced that the **'opening of the second dimension'** by our group is a cutting-edge step forward and more promising than focusing on $\delta^2H_{n\text{-alkane}}$ alone. The reason for other working groups not having caught on the coupled approach might have to be seen, in our opinion, in the uniqueness of compound-specific $\delta^{18}O$ analyses: according to our knowledge, only 3 working groups world-wide have respective experience/publication records. Still, we

would be delighted to see the coupled approach being tested or applied by other groups, readily in cooperation with us. Please also note and be aware that any $\delta^2H_{n\text{-alkane}}$ paleoclimate study (without ^2H-^{18}O coupling!) could be rejected arguing with the uncertainty of biosynthetic fractionation, too.

Possibly, Referee #1 misunderstood Fig. 8. No correlation for the data points shown in Fig. 8 are to be expected. We clarified in our revision that Fig. 8 illustrates the 'concept of the coupled $\delta^2H_{n\text{-alkane}}$-$\delta^{18}O_{sugar}$ biomarker approach'. This conceptual figure illustrates (together with Fig. 9) that $\delta^2H/\delta^{18}O_{prec}$ values reconstructed by the coupled approach are more accurate than δ^2H_{prec} values reconstructed using $\delta^2H_{n\text{-alkane}}$ alone. Moreover, Fig. 10 illustrates that reconstructed RH values under deciduous forest sites and grassland sites are quite well in accordance with RH values of climate stations, thus indeed demonstrating the great potential of the coupled approach.

3) The writing is in places unclear and difficult to follow. I have noted a few of these instances in my technical corrections, but the manuscript would benefit from more careful editing.

→ We will insure a technical and grammatical improvement for the revised version of the manuscript.

Specific comments:

Line 110: This adds up to more than 16, some sites were considered to be more than one of these categories? Would be good to rewrite to clarify

→ Following the recommendation of Referee #1 we will restructure this sentence. The revised version will read: "In November 2012, we collected 29 topsoil samples (0-5 cm depth) from 16 sites along a transect from Southern Germany to Southern Sweden (Fig. 1A). We distinguished between coniferous forest (con, n = 9), …".

Line 114: Was there a threshold for what was considered "close-by"?

→ We agree with Referee #1 that this was not obvious so far in the manuscript and especially not in the supplementary material where the longitude, latitude and altitude were provided for the climate stations (Tab. S2) but not for the locations/sites. In the revised manuscript, we will add the respective characteristics to Tab S1.

Line 133: Machine learning techniques like random forest aren't so commonly used in Biogeosciences and it would be helpful to provide more details here. How many trees did you use? How was data partitioned into training and testing sets? What metric was used to assess model performance? What was the minimum number of samples in the terminal nodes? What was the maximum number of terminal nodes? What variables ended up being ranked as most important (could be useful to show a plot of ranked variable importance in the supplemental materials)?

→ As suggested, we will add a supplementary method description part and refer to it in the text.

Line 136: Why wasn't it possible? Lack of measured data for a robust training data set? Please specify

→ Because no precipitation isotope data was available for the Danish and Swedish sites.

Lines 128-139: How did the calculated values you obtained for the German sites compare to OIPC? What is your evidence for your approach providing superior estimates of precip isotopes than OIPC? OIPC is obviously not perfect, but as written, we have no evidence to evaluate if your results are any more accurate. There is also no discussion of the implications of using one target for precip isotopes in the southern half of your transect and a different one in the northern half.

→ Please allow us to refer to the (cited) Diploma Thesis of Schlotter (2007): there are numerous reasons mentioned already in the introduction highlighting that OPIC is probably not the most robust estimator for middle and high latitudes. That's why we used our own regionalization where it was possible.

Section 2.3.1: No internal standard was added? How do you account for losses of brGDGTs during sample handling?

→ We used standard laboratory procedure for GDGT sample preparation. The internal standard was, as written, added before the measurements. A correction for GDGT losses during sample preparation is therefore not possible.

Lines 165-171: This is not the most current method used for robust brGDGT analysis (see Hopmans et al., Organic Geochemistry, 2016. DOI: 10.1016/j.orggeochem.2015.12.006). Does your method allow for 5' and 6' methyl brGDGTs to be distinguished from one another? If not, severely diminishes the accuracy of results. Based on the results that are shown, it seems like this method does not distinguish the different isomers.

→ That's correct. Please see our reply to major issue 1.

Lines 172-173: how was the pH measured?

→ We will include the information that a pH meter was used.

Section 2.3.2: Were the n-alkanes quantified prior to measuring their stable isotopes?

→ Yes, namely by Schäfer et al. (2016). We therefore added the following sentence in the section: "For more details about *n*-alkane quantification the reader is referred to Schäfer et al. (2016). ".

Also, please briefly describe the operating conditions of the GC-pyr-IRMS (or cite another publication that used an identical method and provides all the relevant details)

→ As suggested, we added now in the revised version of the manuscript a reference (Christoph et al., 2019), in which the method is described in more detail and we added that the ^{2}H pyrolysis reactor temperature was kept at 1420 °C.

Line 199: It is not clear how you had 29 samples from 16 sites. Were some of the sites sampled in duplicate?

→ We will clarify during revision that 29 samples were collected from 16 sites. These are, however, no duplicates, but rather different dominant vegetation types (see reply above).

Lines 211-221: The more robust indicator of soil pH is CBT' and the more robust indicator of soil temperature is MBT5me (De Jonge et al., GCA, 2014, DOI:10.1016/j.gca.2014.06.013).

→ See our reply to major issue 1.

Lines 227-229: A number of papers have shown that ebio is not constant and different among plant types and seasonally. See for example Feakins & Sessions 2010 (cited previously), Eley et al., GCA, 2014 (DOI: 10.1016/j.gca.2013.11.045), Cormier et al., New Phytologist, 2018 (DOI: 10.1111/nph.15016).

→ That's true and especially important when only $\delta^2H_{n\text{-alkane}}$ is used to reconstruct $\delta^2H_{\text{leaf-water}}$. Nevertheless, we emphasize in our manuscript that ε_{bio} is a major uncertainty in our coupled approach, too. At the same time, it's exactly such uncertainties why we need climate transect calibration studies as the one presented here for Europe.

Lines 383-385: are these concentrated weighted means? That is what is typically used to compare d2H values of n-alkanes where not all homologues are present in all samples

→ We used here mean values, because the areas and concentrations where not determined during isotope measurements.

Line 395: I think you mean "unenriched xylem water"?

→ Yes, changed.

Lines 431-432: This is not particularly convincing, the reconstructed precipitation isotopes are not correlated with the GIPR/OIPC precipitation isotopes. No evidence is provided to show that this approach is any better than the most up to date methods for obtaining precipitation isotopes from leaf wax n-alkane isotopes alone. For example, how do your results compare to the predictions from the proxy system model developed by Konecky et al. (JGR-Biogeoscience, 2019, DOI: 10.1029/2018JG004708)? Maybe your approach is better, but you need to prove this by providing a direct comparison, rather than just telling us

→ Please note that we do not necessarily expect a good correlation of our reconstructed $\delta^2H/\delta^{18}O_{prec}$ values with the GIPR/OIPC data, but rather a good (accurate) match on the 1:1 line. Nevertheless, many thanks for pointing us to the new publication by Konecky et al. (2019). While we will readily include a respective citation, we think that a direct comparison of our approach with the one suggested by Konecky et al. (2019) would be beyond the scope of our manuscript.

Lines 448-450: If this was the case, wouldn't you expect all the coniferous sites to be biased in the same direction? Instead, they are evenly distributed above and below the 1:1 line

→ No, please see Fig. 9: we do not see that the coniferous sites are evenly distributed around the 1:1 line. Except for one data point, they are clearly below the 1:1 line.

Line 454: Is this signal damping correction shown anywhere? How would this work practically in sediments?

→ No, sorry. This signal damping correction is not shown or quantified in this manuscript. This would require a quantitative estimation of the contribution of grass vegetation to the total biomass pool in the topsoil. For an example how such a correction can be applied to lake sediments please see e.g. Hepp et al. (2019, CP).

Lines 467-468: Actually, there are plenty of n-alkanes in roots and they have very different H isotopic composition than in leaves. See work from Guido Wiesenberg's group and Gamarra and Kahmen. I'm also confused about what you are referring to as "the discussion". There is not a separate discussion section to this manuscript.

→ Changed to "Zech et al. ,2012b and the discussion therein". We do not agree and we are not aware of any new studies showing that *n*-alkanes are produced in large amounts by roots in comparison to leaves. There is a clear respective dissense with the Wiesenberg group (see open discussion of Zech et al., 2012b, or Zech et al., 2013 - Response to the comments by G. Wiesenberg and M. Gocke. Quaternary Research 79(2), 306-307). Moreover, the work of Gamarra and Kahmen (2015) shows that root *n*-alkane concentration is always the lowest compared to the other plant tissues sampled.

Lines 489-494: Not stated here is that there is no correlation between the reconstructed and measured RH values. This suggests that this approach for reconstructing RH is not particularly useful Line 565: The data in the paper is not very convincing that there is great potential for the coupled d2H n-alkane d18O sugar approach

→ We don't agree. Given the low range of measured RH values along this European climate transect and the uncertainties of the coupled approach for reconstructing RH values, the lack of a respective correlation is not really surprising to us. Please compare a similar climate transect study by Tuthorn et al. (2015, BG) where the RH range is much larger and where indeed a significant correlation can be found. For this European transect study here, the usefulness of the coupled approach for reconstructing RH values should be rather inferred from the quite well 1:1 match for deciduous forest sites and grassland sites (cf. Fig. 9). The RH underestimation for coniferous forest sites can be easily explained with the extremely low n-alkane production of coniferous trees (see ll. 495-502).

Lines 566-567: I don't see evidence of this in your analysis, nor examples of how you would take vegetation into account when applying this proxy.

→ See for example Hepp et al. (2019).

Technical corrections and typing errors:

Lines 54-56: The way this sentence is written is confusing. Suggest rewriting as "Climate proxies based on molecular fossils, also known as biomarkers, have great potential...

→ Changed.

Line 56: don't need the comma after "particular"

→ Changed.

Line 59: "need to be known"?

→ Changed.

Line 61: It would be better to start this paragraph with a clear link back to the previous one

→ We now start the paragraph with "One famous and widely applied lipid biomarker group are terrestrial branched glycerol dialkyl glycerol tetraethers (brGDGTs). They are synthesized… and…"

Line 74: don't need commas before and after "it is known"

→ Changed.

Line 79: Again, some sort of transition would be helpful to begin this paragraph

→ We now start the paragraph with "Concerning paleohydrology proxies, compound-specific…"

Line 82: "all along the way" too wordy

→ Changed

Lines 93-94: "as well as concerning possible effects related to" awkward phrasing

→ Changed.

Figure 1: would be nice to have a legend on panel B or have the axis colors match the variable colors. At the moment we are left to guess that blue bars are precip and the red dots are temp, since this is not stated in the figure caption or the legend. Also would be nice to offset the panel letters with a () or . to break them apart from the title of the panel

→ Changed.

Line 180: No "the" needed in front of ETH

→ Changed.

Line 225: the n at the beginning of n-alkane should be italicized. Check throughout

→ Changed.

Line 234: Generally, figures should be numbered in the same order that they are referenced in the text

→ Checked and changed if necessary.

References

Christoph, H., Eglinton, T. I., Zech, W., Sosin, P. and Zech, R.: A 250 ka leaf-wax δD record from a loess section in Darai Kalon , Southern Tajikistan, Quaternary Science Reviews, 208, 118–128, doi:10.1016/j.quascirev.2019.01.019, 2019.

Gamarra, B. and Kahmen, A.: Concentrations and δ^2H values of cuticular *n*-alkanes vary significantly among plant organs, species and habitats in grasses from an alpine and a temperate European grassland, Oecologia, 178, 981–998, doi:10.1007/s00442-015-3278-6, 2015.

Hepp, J., Wüthrich, L., Bromm, T., Bliedtner, M., Schäfer, I. K., Glaser, B., Rozanski, K., Sirocko, F., Zech, R. and Zech, M.: How dry was the Younger Dryas? Evidence from a coupled δ^2H–δ^{18}O biomarker paleohygrometer applied to the Gemündener Maar sediments, Western Eifel, Germany, Climate of the Past, 15, 713–733, doi:10.5194/cp-15-713-2019, 2019.

Konecky, B., Dee, S. G. and Noone, D. C.: WaxPSM: A Forward Model of Leaf Wax Hydrogen Isotope Ratios to Bridge Proxy and Model Estimates of Past Climate, Journal of Geophysical Research: Biogeosciences, 124, 2107–2125, doi:10.1029/2018JG004708, 2019.

Schlotter, D.: The spatio-temporal distribution of δ^{18}O and δ^2H of precipitation in Germany - an evaluation of regionalization methods, Albert-Ludwigs-Universität Freiburg im Breisgau. [online] Available from: http://www.hydrology.uni-freiburg.de/abschluss/Schlotter_D_2007_DA.pdf, 2007.

Tuthorn, M., Zech, R., Ruppenthal, M., Oelmann, Y., Kahmen, A., del Valle, H. F., Eglinton, T., Rozanski, K. and Zech, M.: Coupling δ^2H and δ^{18}O biomarker results yields information on relative humidity and isotopic composition of precipitation - a climate transect validation study, Biogeosciences, 12, 3913–3924, doi:10.5194/bg-12-3913-2015, 2015.

Zech, M., Kreutzer, S., Goslar, T., Meszner, S., Krause, T., Faust, D. and Fuchs, M.: Technical Note: *n*-Alkane lipid biomarkers in loess: post-sedimentary or syn-sedimentary?, Discussions, Biogeosciences, 9, 9875–9896, doi:10.5194/bgd-9-9875-2012, 2012.