

Dear Editor,

Please find submitted our rebuttal and the revised version of the paper "*Constraints on the applicability of the organic temperature proxies Uk'_{37} , TEX_{86} and LDI in the subpolar region around Iceland*" by Rodriguez –Gamiz et al. As outlined in our on-line responses to the referee's we have revised parts of our manuscript to reflect their comments. For a detailed rebuttal I refer you to our response (in bold) below to the comments of the reviewers (in italics).

We hope you find the revised version acceptable for publication.

On behalf of all co-authors,

Sincerely,

Sebastian Rampen

Response to reviewer Salcup:

In this work the authors investigated the applicability of Uk'_{37} , TEX_{86} , and LDI to paleotemperature reconstructions around Iceland using filtered water, sediment trap, and surface sediment samples. The results suggest that while there is good agreement between proxy derived temperatures in surface sediment samples and seasonally averaged temperatures, there are large discrepancies between the proxy inferred temperatures in suspended and sinking particulate matter and in situ temperature, as has been seen previously. The authors contribute these discrepancies to production seasonality, diagenetic alteration, and/or lack of producing species. This work represents an important study of three widely used paleotemperature proxies in a climatically important setting in which previous work has shown them to have difficulty reproducing meaningful temperature estimates. It warrants publication after significant editing for clarity and fluency.

We thank the reviewer for his kind words and for the comments and suggestions given below.

Specific comments:

In the Methods section, lines 14 to 24, the workup for the SPM filters is fairly complex and it is not clear to me why. Please expand slightly to explain your reasoning.

We agree that that this has been unclear. In the revised version this has been clarified and better explained (lines 144-161).

In figure two the authors show temperature reconstructions based on the different proxies. The concentration data for these samples is discussed in the results, but it would be good to see how the concentrations change between sampling stations and from year to year. Perhaps by adding a panel to figure two.

Regarding Fig. 2, this is an interesting point but unfortunately we do not have the concentrations of all the lipids used as proxies in the different stations for both cruises. Therefore, we cannot expand this figure as suggested by the referee.

As a general comment the authors should take care in comparing discrete suspended particulate matter results from summer, or sediment trap results from only one year, to results based on surface sediments which likely reflect a decade or more of sedimentation. That said I appreciate how expensive and time consuming water-column work can be.

We agree with the referee; surface sediments reflect the composition of descending particles over multiple years. In this sense, a sediment trap record of one year may still be seen as a “snapshot”, despite all the major efforts required to obtain such dataset. In our revised manuscript we have indicated this (Lines 410-415).

In their discussion of Uk'37, the authors suggest surface sediment based temperature estimates likely reflect summer temperatures as this is the time of maximum alkenone flux. This is supported by an XY plot of alkenone inferred temperature vs. summer mean temperature. However during summer, and time of maximum alkenone flux, Uk'37 inferred sea surface temperatures are up to 4_ colder than in situ temperature. This offset must complicate the interpretation of sedimentary Uk'37 inferred sea surface temperatures as a summertime temperature signal. The temperature offsets seen in their sediment trap time series, Uk'37 to warm in the winter and too cool in summer, has been seen previously [Harada et al., 2006; Lee et al., 2011; Prahl et al., 2001; Seki et al., 2007; Sikes et al., 2005; Yamamoto et al., 2007]. Please discuss within this context and reference appropriately.

The referee is correct with respect to the difficulties in the interpretation of the results of the $U^{K'}_{37}$ index. Several explanations for the discrepancies between $U^{K'}_{37}$ - derived SST from sedimenting particles and in-situ SST, or $U^{K'}_{37}$ -SST of surface sediments, have been discussed in the text but we will expand on this in the revised manuscript (lines 363-372 and 381-387).

Once the authors took into consideration an expanded depth habitat, and applied the proper calibration (TEX86L 0-200m), surface sediment derived temperature estimates were found to compare very well with mean annual or winter temperatures. This supports the interpretation that GDGT producers live throughout the water column and that TEX86 should be calibrated as such.

In all, the evidence presented by the authors supports a multi-proxy approach, particularly in troublesome environments like the high latitudes.

We thank the referee for the positive appraisal of our work.

Response to reviewer Mollenhauer:

General comments: In this manuscript, Rodrigo-Gámiz and co-authors examine the applicability of three organic biomarker based proxies for sea surface temperature (SST) using samples of suspended matter collected from near-surface waters, sinking particles collected using a sediment trap, and surface sediments. All samples are analyzed for the three biomarker SST proxies, and the results are compared with insitu SST, satellite SST estimates, and World Ocean Atlas temperatures. The data presented are interesting and have the potential to lead to a better understanding of that require consideration and likely will result in a much different interpretation of the results and different conclusions:

We thank the reviewer for her comments and suggestions.

1) *The core piece of data stems from samples collected using a sediment trap moored for one year at a water depth of 1850 m. There is abundant literature discussing sinking rates of particles based on data obtained on samples collected with sediment traps, all indicating that considerable time elapses between the formation of a biogenic particle and its settling to deeper water depths (e.g., Muñ' Ller and Fischer, 2001, DSR, Fischer and Karakas, 2009, Biogeosciences, Yamamoto et al., 2007, DSR, 2012, OG, etc.). Some of these papers include the finding that settling rates might be different for different types of particles. Sinking rates are often calculated from the phase shift between proxy records and satellite observations, which in turn means that the time the sinking*

requires needs to be considered when comparing proxy data and observed SST. This is completely ignored when discussing differences in temperature estimated using the lipid biomarker proxies and satellite derived temperatures.

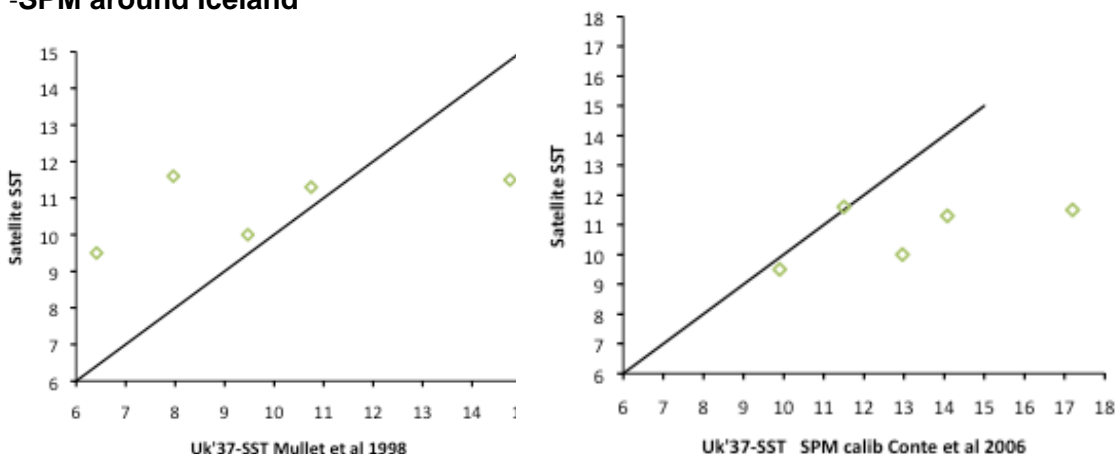
We agree with the idea that usually there is a time lag between the production of the biomarker and its sinking to deeper water depths, arriving at the sediment trap. However, in our case, almost synchronous variations in net primary production and the fluxes of the various lipids at 1850 m are observed (Fig. 3). This suggests that there is no major difference, within the resolution of the sediment trap, between signals generated in the upper part of the water column and those received in the sediment trap at 1850 m. A rough estimation of the sinking velocity based on the resolution of the trap suggests that this would correspond to a sinking rate of up to ca. 230 m per day, which is quite fast. Reported lipid fluxes for GDGTs are ca. 260 m per day in the NW Pacific (Yamamoto et al., 2012) and for alkenones are ca. 280 m per day in the filamentous upwelling region off Cape Blanc (Müller and Fischer, 2001). These fluxes compare well with our estimated lipid fluxes. This is now indicated in lines 392-395 and 497-499.

Based on these estimations, the differences observed between proxy derived SST and satellite derived SST are probably not related to delayed signals. We have provided alternative explanations in the manuscript to explain the differences between proxy derived SSTs and satellite-derived SSTs (lines 363-372 and 381-387 for alkenones, and lines 473-478 for GDGTs).

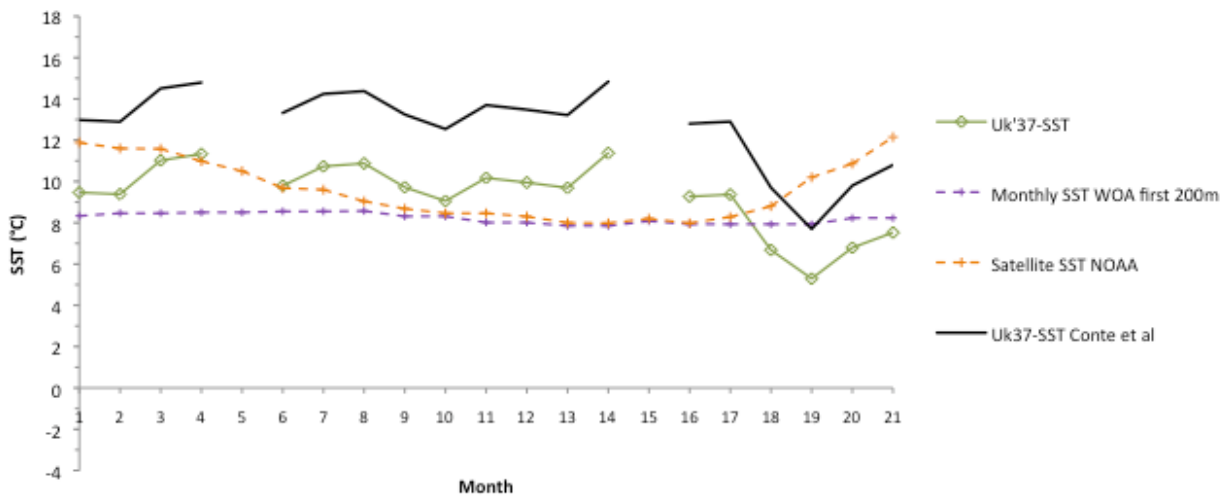
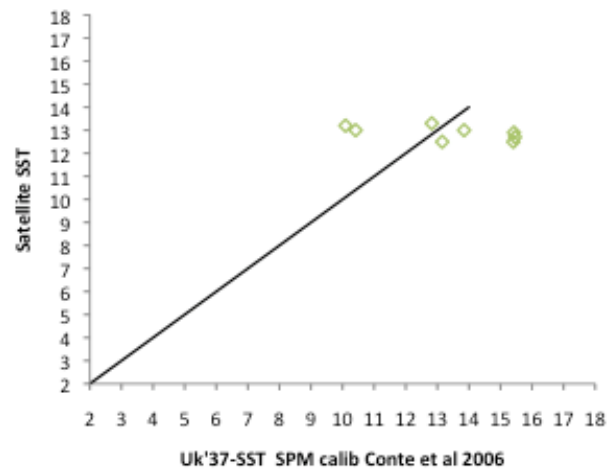
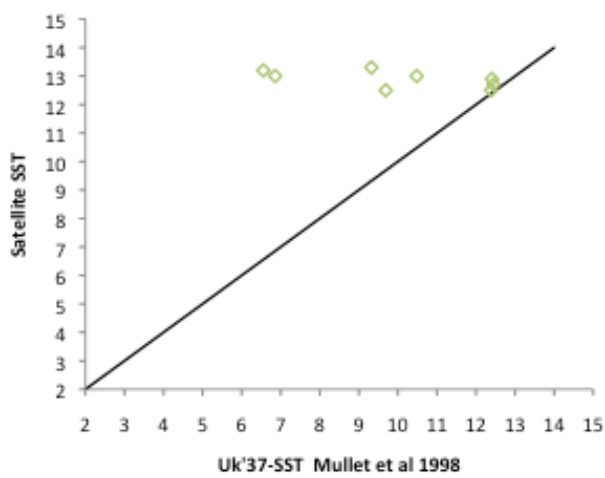
2) All UK'37-based SST estimates are based on the core-top calibration by Müller et al. (1998), even though this calibration is explicitly derived for sediments. Since the publication of this seminal paper, however, more efforts have been undertaken to refine the calibration of the UK'37 proxy, in particular for samples of suspended matter. In their paper published in 2006 in GC, Conte and co-authors compile a large data set obtained on SPM and compare it with core-top data. Their calibration for SPM is polynomial, while the best fit for core-top sediments is linear. The largest discrepancy between the two is approximately between UK'37 values of 0.15 and 0.45 and can amount to up to >3.5 °C. This is a) exactly the range of UK'37 values observed in this study and b) very similar to the temperature offsets between UK'37-SST and in-situ values. Moreover, the authors provide an explanation for the discrepancy between the two, which should be considered in this manuscript as well. It has furthermore been previously observed that the polynomial calibration also results in better agreement between observations and reconstructions from samples collected by sediment traps (Mollenhauer et al., 2015, DSR). I thus suggest that the authors re-calculate their temperature estimates using the polynomial regression for SPM and sediment trap samples and the core-top calibration for core-top sediments. I expect that the agreement between observations and reconstructions will be much improved and, as a result, the conclusions will be substantially different.

We appreciate the comment to use the specific SPM calibration proposed by Conte et al. (2006). We have therefore plotted a comparison between Müller et al. (1998; left) and Conte et al. (2006 right) calibrations for:

-SPM around Iceland



-SPM transect south of Iceland



-Sediment trap

It is clear from these figures that we do not obtain substantially reduced offsets between $U^{K'}_{37}$ -SST and satellite SST for the SPM around Iceland; the average offset

remains 2.7°C. For the SPM transect south of Iceland, we do observe a slightly reduced offset ranging from 0.5 to up to 2.5°C.'

For the sediment trap, the $U^{K'_{37}}$ -SSTs (black line) using the SPM calibration of Conte et al. (2006) overestimates satellite SST by about 3°C on average (dashed orange and purple lines), whereas the $U^{K'_{37}}$ -SSTs (green line) using the calibration by Müller et al. (1998) overestimate satellite SSTs on average by only 0.3°C.

Based on this, we prefer to keep using the calibration by Müller et al. (1998). In the revised manuscript it is now noted that we have tested the calibration by Conte et al. (2006) but that this didn't provide better temperature estimates (lines 358-361).

3) *In the abstract and conclusions, it is fairly strongly stated that a good agreement is observed between TEXL86 0-200 m temperatures and WOA observations of annual mean and winter depth-integrated temperatures. This is, however, only based on data from the core-top sediments (n=10), while the entire data set on settling particles (n=21) does not support this conclusion. In contrast, the TEXL86-temperatures for 0-200 m are substantially overestimated with respect to the WOA data. In my view, this discrepancy mandates further investigation and does not allow to draw the conclusion presented in the manuscript.*

We agree and we have rephrased these sentences in both the abstract and conclusions (lines 597-600). However, the restricted number of words for the abstract prevented us to present a more detailed statement.

4) *The language requires improvement. There are several errors in grammar and a number of awkward expressions and overly long sentences.*

In the revised version we have tried to modify the text and to improve our phrasing.

5) *The data obtained within this study are not completely presented in the tables. In Table 1, information on sampling stations is given, for all samples. In Table 2, however, where proxy data are presented, only the samples from the sediment trap are listed. In situ temperatures used to compare the proxy results with are missing entirely, as well as total fluxes. Please add missing information.*

We thank the reviewer for this suggestion. We have added the bulk fluxes and main lipid fluxes to Table 2. Furthermore, a new table, as supplementary material, will be included with the index values and estimates SSTs of each proxy and temperatures derived from WOA and NOAA.

Below I list a number of specific comments:

Page 1115, lines 25 and following: It is a bit too simplistic to state that soil-derived contributions of isoGDGTs can be neglected at $BIT < 0.3$; please reword.

The reviewer is correct in that it will depend on the location whether a BIT index threshold of 0.3 is sufficient to exclude an impact of terrestrial input (cf. Schouten et al., 2013). More clues may be obtained by correlating the BIT index with TEX_{86} values where a significant correlation indicates an impact of terrestrial input. In our case the correlation is negligible ($R^2 < 0.08$) and thus the TEX_{86} is unlikely to be affected by a terrestrial influx of GDGTs. We have rewritten the text about the BIT index in the revised manuscript (lines 29-33).

Page 1116, line 28: there is only one paper by Rodrigo-Gámiz listed, so please omit the "b" after 2014.

This has been corrected.

Page 1117: Please add information on the productivity regime and the timing of phytoplankton blooms, in particular on coccolithophorid blooms, to the description of the

study area. Nutrient regimes might also be interesting.

We have added additional information about the productivity regime in the revised text (lines 90-99).

Page 1118: The fact that the cruises during which the samples were collected are specifically named suggests that additional information on these cruises (e.g., a cruise report) is available. However, no reference is made to such information. Please clarify.

There are two relevant cruise reports to which we refer in the revised manuscript (lines 103-104).

Page 1119, lines 14 and following: At which temperature was the saponification carried out? The method description for the extraction of the filters is not clear: How can you extract water with a mixture of water and methanol? It seems to me that there is an error. Please clarify.

With respect to the saponification method in the SPM, we understand that this section was unclear. In the revised version we have clarified the method and explained it better, as also requested by the other reviewer (lines 144-161). Concerning the temperature at which the saponification was carried out, it is indicated that this is under reflux, hence, at boiling temperature.

Page 1120, line 1-2: Why was there only the diol standard added to “some” samples? On which grounds was decided which of the sediment trap samples were treated with copper to remove sulphur?

The section concerning the standard has been removed, as lipids in sediment samples and SPM samples have not been quantified.

With respect to the treatment to remove elemental sulfur; we treated those samples in which we detected elemental sulfur by GC-MS analysis in order to improve that analysis (lines 140-142).

Page 1124, line 15: Please provide total fluxes also in the table.

We have added this to Table 2.

Page 1129, line 18 and following: This line of arguments is not convincing: Usually, the TEX₈₆ paleothermometer is determined on core lipids, not on IPLs. Therefore, a mismatch between core-lipid SST estimates and observations is a relevant signal. Including IPL-TEX₈₆ does not help in resolving the discrepancies.

We agree that TEX₈₆-derived SST is commonly based on core lipids, but IPLs might better reflect *in situ* conditions. We have rewritten parts of this section although we have already previously provided some explanations about the mismatch between proxy estimated SST and satellite derived SST in the manuscript (lines 455-465).

Page 1131, line 4 and following: The fact that fluxes are highest in the summer does not necessarily mean that GDGTs were produced during this time and represent summer SST. Considering that the TEX-based SST estimates are lowest during the high-flux periods, as can be seen in Figure 4, this is a rather unlikely scenario.

We agree and we addressed this observation in the revised manuscript (lines 497-516). Indeed, based on previous studies (e.g., Wuchter et al., 2006; Huguet et al., 2007; Mollenhauer et al., 2015), high GDGT flux periods in summer are explained by a preferential transport of GDGTs due to high phytoplankton bloom.

Table 1, caption: The caption is incomplete: What does “Flow meter (l)” and “Cross cut (l)” mean? Column headers should be “core length” and “volume” instead of “Long” and “Flow meter” and “Cross cut”.

We agree, “cross cut” has been deleted as it gives similar information to “flow meter”, which has been replaced by “volume pumped”.

Figure 2: Cross plots of the in-situ and satellite temperatures versus the reconstructed SST might be more revealing.

We appreciate this comment. However, the cross-plots suggested (such as shown in the above comment about $U^{K'_{37}}$ calibrations) are, in our opinion, not adding additional information. Adding these plots to the current Fig. 2 would result in many more panels in Figure 2 (i.e. cross-plots of proxy values against satellite SST, in-situ SST and summer mean temp. at 50 m). Thus, we prefer to keep the original Figure 2.

Figure 3: Please consider adding vertical lines to help guide the eye.

We have added a number of relevant vertical lines in this figure.