

Revision article bg-2018-441 for Biogeosciences

We like to thank both reviewers for their constructive and very supportive comments. We have uploaded a pdf providing initial answers to the reviewers comments. Reviewer questions are indicated in italics and our answer below.

Reviewer 1:

1) My only wish is that this study would have also included another newly suggested proxy coral Sr-U for SST proposed by De Carlo et al., 2016, which could be performed simultaneously with their ICP-MS analysis, to assess this other suggested proxy along with the others. The method section (line 190) mentions the precision of U/Ca analysis so it seems U/Ca was measured. Why not add Sr-U to the proxy suite that was assessed?

We have tested Sr-U and found that it did not provide reliable SST reconstructions. Furthermore, our aim was to reconstruct bimonthly SST which is currently not possible with Sr-U as it requires averaging multiple data points to estimate temperature for one value and therefore makes it impossible to do direct comparisons of the methods. We have therefore not included such an assessment in the present paper and do not show U/Ca since it does not add to the interpretation. We add some figures here to show the results which we feel do not warrant inclusion, but the coral community might find it helpful since our comments will be published with the paper. Here, our focus is on testing Li/Mg vs. Sr/Ca and their combination for bimonthly time series.

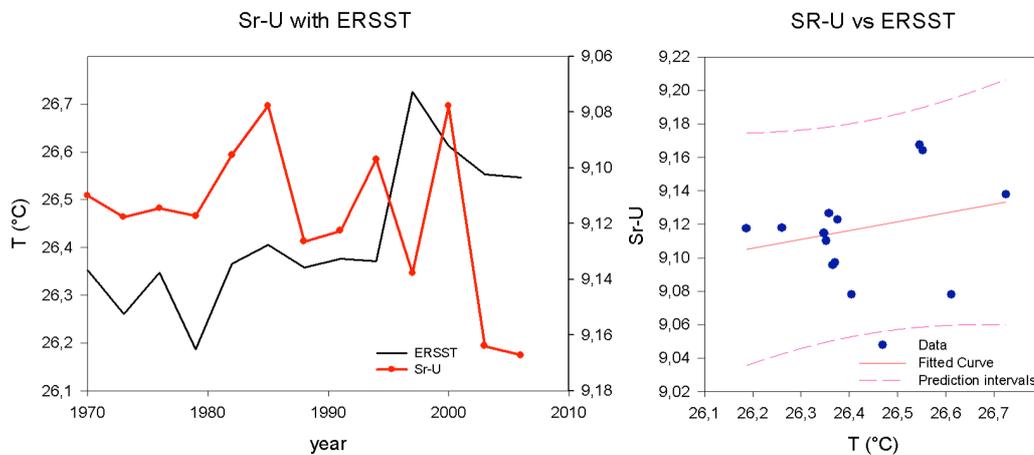


Figure - Left) 3 years averages of EU3 Sr-U (red) compared with NOAA ERSST4 (black). Right) Weighted regression of 3-year averaged Sr-U with ERSST (not significant). Sr-U was calculated following the method in De Carlo et al. (2016). No correlation was found between Sr-U and SST.

2) Move the paragraph in section 3.3 Coral growth parameters and SST to the first section in the Results.

Done.

3) In the methods section (lines 167–170) clarify if Coral XDS was performed on a single transect or replicate transects. Were these CoralXDS transects the same as your

micro-sampling for geochemical analysis?

Transects for CORALXDS were not the same. For geochemistry we have adopted a much more stringent approach to sample the main growth axes. Our age model confirmation with Sr/Ca seasonality was applied to cross-check CORALXDS results, as shown in Supplementary Figures S1 and S2.

We have added the CORALXDS sampling transects to the Supplementary figures S1 and S2 as dashed lines.

4) What were the lowest extension years? Any fall below the Porites threshold of 5 or 6 mm/year? If not, vital effects should not be a problem.

None of the years fall below 5mm/yr extension rates, ranging between 7-18mm/yr. Only one year in EU3 (1983) showed 5mm/yr in the CORALXDS data which was corrected with our Sr/Ca extension rate data. We now have included a Supplementary Table S2 listing all annual growth data.

5) Methods Line 190-191 The normalization of Sr/Ca with JCP-1, what about Li/Ca and Li/Mg? What is the precision for Li/Mg and Li/Ca?

All data is referenced to JCp-1, but we also include in coral *in-house* standards for long-term precision of TE/Ca and Li/Mg data. We have adjusted the methods section to align with the former statement and added Sr/Ca, Mg/Ca and Li/Mg values and RSDs in JCp-1 and the in house standard 'Davies Reef'.

It now reads:

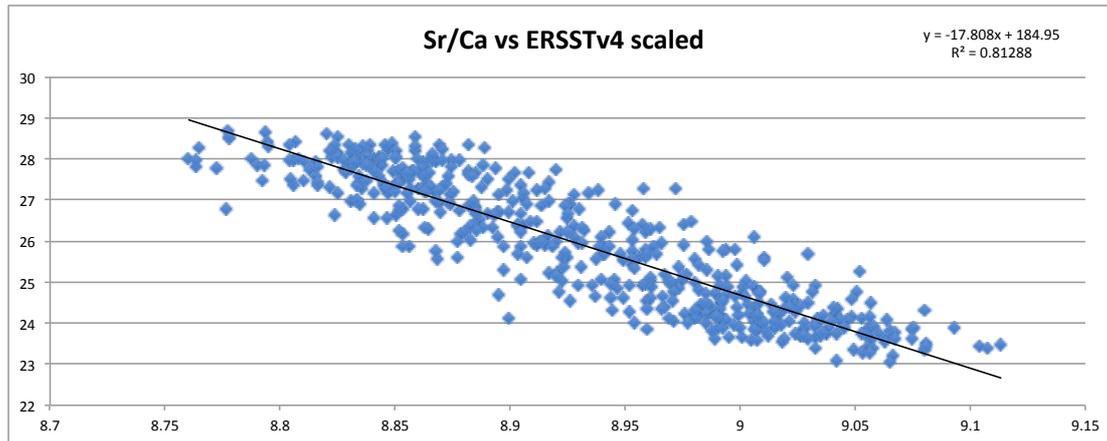
"The Sr/Ca, Mg/Ca and Li/Mg data reported here are normalized to the JCp-1 *Porites* sp. standard prepared by the Geological Survey of Japan (Okai et al., 2002) with Sr/Ca = 8.85 mmol/mol (2σ RSD = $\pm 0.41\%$), Mg/Ca = 4.20 mmol/mol (2σ RSD = $\pm 0.90\%$) and Li/Mg = 1.47 mmol/mol (2σ RSD = $\pm 1.04\%$) (N = 17). The Li/Ca data was estimated dividing the Li/Mg by the Mg/Ca data. Long-term reproducibility was determined using the UWA *in-house* Davies Reef coral standard solution with Mg/Ca = $\pm 6.24\%$, Sr/Ca = $\pm 0.45\%$ and Li/Mg = $\pm 1.39\%$ (2σ RSD; N = 139) (D'Olivo et al., 2018)."

6) Figure 4 and Figure S3 Time assignment. There seem to be three clusters of data in all the plots. This is odd, there should not be clusters of data. Double check your Analyseries time assignment, something is off, see next item.

We checked the linear regressions for monthly vs bimonthly interpolated data with regards to the clusters of data. We found that the clustering increases with decreasing resolution of the SST data, both in terms of spatial (from 0.25° to 2° gridded SST data) and temporal (from monthly to bimonthly) scales. ERSST4 ($2^\circ \times 2^\circ$ gridded) shows the most extreme clustering, while it is much less for AVHRR OISST. The reason for the clustering is the sudden jumps in SST between seasons at bimonthly time scales. Our monthly interpolated data don not show clustering (see figure below) confirming our

age model. We, however, prefer to work with bimonthly interpolated data since our sampling resolution (6-9 samples per year) is better suited for it.

We have now used 2 anchor points (summer and winter) for our age assignment which did not change the regressions or any conclusions drawn from our work. The methods section has been amended accordingly. Corals never grow constantly between summer and winter in subtropical sites like Europa, the SST seasonality varies between 4.5 and 6°C, so significant SST contrasts between seasons do prevail probably affecting density banding.



7) Looking at the residuals plot (Figures 5 and 6) and the methods lines 197–200, I see there is a sinusoidal character in the residuals.

The sinusoidal character of the residuals is most apparent between the 1980-1992 period (Figure 5d). For this particular period, the amplitudes in our TE/Ca SST reconstructions do not match those in instrumental AVHRR-OISST. The latter indicates very low seasonality in several years between 1980 and 1992, far lower than post-1992. It is beyond our capacity to check the quality of AVHRR-OISST in the oldest part of their record. We believe that the mismatch does not relate to the number of anchor points chosen (we have chosen 2 as requested by the reviewer) for calibration. The differences between coral-derived SST and instrumental record reflect either real differences in local SST or a combination of unknown vital effects and temperature. Figure 6 indicates SST anomalies (seasonality removed) and their residuals (difference between instrumental SST anomaly and coral-derived SST anomaly).

8) I do not see any major issues with your linear regression plots (Figure 4) but if you reverse the regressors, you get a different solution. OLS will either over fit or under fit the proxy data to SST, depending which regressor is the independent regressor. Reduced Major Axis regression is better and takes error in both regressors into account. Weighted least squares regression is even better (York and Evensen, 2004; Thirumalai et al., 2011) since you can vary the error for each data point. However, this switch in regression method should not change your slopes greatly but should reduce your RSME results. The bigger issue is the serial correlation. WLS can take this serial correlation into account in the regression and it increases your confidence intervals, but it would be better to just remove it.

Our regression analysis was done in Matlab using the Weighted Least Squares Regression, which we named 'robust regressions' in the figure captions. We apologize

for having missed to report this, and have now included it in the methods section. All absolute SST reconstructions are based on WLS regression, including the results for the residuals. Autocorrelation is accounted for in our WLS regressions.

We normally use OLS, and not RMA, because it is the method of choice for coral Sr/Ca-SST regressions (Solow & Huppert, 2004). This is because the relationship of coral Sr/Ca and SST is clearly asymmetric (SST determines coral Sr/Ca, it cannot be the other way round). This is the major criterion for the use of OLS, not the question whether there are errors in the SST observations or not (see e.g. Smith, 2009 “Use and Misuse of the RMA for Line-Fitting” for a discussion). (RMA should be used for symmetric relationships.)

The biggest problems with the application of RMA for coral-Sr/Ca calibrations are the unknown errors. This is nicely summarized in Solow and Huppert (2004). RMA assumes that the error variance in the SST observations equals the error variance of the Sr/Ca determinations. There is no reason to believe that this assumption is warranted. The RMA method can be extended to allow for differences in the error variances (see refs in Solow and Huppert, 2004). To do so, it is necessary to have an estimate of both the SST and Sr/Ca error variance. However, it is practically impossible to determine the error variance of coral Sr/Ca determinations, as these include not only the analytical error but also other factors such as vital effects or skeletal heterogeneities.

However, in our study we opted for weighted least squares regression in line with the suggestion of reviewer 1.

Sr/Ca and Li/Mg were used in age assignments since they are the most robust SST proxies to date. We have not opted for assigning the age model with other TE/Ca ratios. We agree with the reviewer that RMSE will therefore always be better for Sr/Ca and Li/Mg than other TE/Ca ratios.

9) Line 192 After doing your own very nice and comprehensive calibration study, why use Correge's -0.0607 ? Your slopes are lower than his and in my assessment, a better calibration study than Correge's that was just the average for all published studies to date (2006), regardless of the calibration method, SST data used, the possibility of using sub-optimally sampled corals for calibration, etc. Things have vastly improved from 2006 and I would not trust a community average value anymore. See recent paper by (Murty et al., 2018).

We have chosen the Correge slope of -0.061 mmol/mol $^{\circ}\text{C}^{-1}$ for the SST anomaly conversion only because this slope is better suited to look at interannual SST changes at regional space scales (not local) following Gagan et al., 2012. Gagan even suggested to use -0.084 mmol/mol $^{\circ}\text{C}^{-1}$. To our opinion, the -0.061 slope is best suited for SST anomaly conversions as shown in previous work (Pfeiffer et al., 2017). However, all absolute bimonthly SST reconstructions are based on our local regression slopes.

10) Line 464–468, Figure 2 and Methods. It is not mentioned if you cleaned the corals to remove any endolithic algae, tissue layer or biological residue. Corals for Mg analysis are generally cleaned with bleach since Mg is present in biological tissues. In Figure 2, the Mg/Ca is high in the tissue layer, this section should be excluded from the

Li/Mg and Mg/Ca data and analysis. Is there any dirty or algae present in the coral skeletons that could be influencing the Mg?

We have already reported the cleaning method in the methods section of the original paper in lines 165-166. We have cleaned the corals following the protocol of Nagtegaal et al. (2012). We have excluded the top-most samples from the interpretation since certain TE/Ca ratios still showed an effect despite rigorous cleaning. The latter could well be an effect of the lower densities in the top mm where precipitation efficiency differs from underlying material.

11) For your regression analysis, why not combined the two corals to see if the results improve?

Corals EU2 and EU3 have offsets in their mean values, so its not easy to combine them. We have, however, combined proxies when assessing proxy and SST anomalies. Here, we found rather excellent agreement between cores and with SST anomalies.

12) Table 2 How were degrees of freedom determined? Was it adjusted for serial correlation? For EU3 it seems the Sr/Ca and Li/Mg are repeated in rows 9–12 or is this the average of EU2 and EU3 with the label missing in column 1?

The regressions applied assume a $DOF = N-2$, as indicated we used WLS as suggested by the reviewer.

Rows 9 and 11 show the regression parameters for the longer period 1981-2012 for EU3 Sr/Ca and Li/Mg, respectively, with AVHRR-OISSTv2. Rows 10 and 12 show the regression parameters for the period 1970-2012 for EU3 Sr/Ca and Li/Mg, respectively, in relation to ERSST4. This is indicated in the figure caption and periods in last column of Table 2.

13) Table 3 What is the second row for EU3 for? Is this EU2 and EU3 averaged together? Table 3 Why are you reporting a standard deviation for RSME? RSME does not have a standard deviation, it is a version of deviation itself between the calculated and measured. RSME is reported as since you take the square root of the sum of the deviations divided by the number of pairs.

The second row for EU3 in Table 3 is for the RMSE for the longer period 1981-2012. We have omitted the standard deviations, thanks for pointing this out. These numbers indicated the spread in RMSE for individual years over the period considered.

14) Table 4 and all correlations. Give the degrees of freedom (adjusted for serial correlation), # of pairs, and p-values for all correlations.

The number of pairs (period in years) and p-values (stars) are indicated in the legend of the table 3. DoF are now added.

15) Figure 2 Can you add the raw trace elemental ratios to the supplemental figures so

time assignment can be verified?

This will be done.

16) Line 59 Add a reference for trade route statement. I believe the southeastern coast of Africa was part of trade routes before the mid-19th century. ICOADS (V3.1) now goes back to the 1600s (Freeman et al 2017, doi:10.1002/joc.4775) but ERSST stops at 1854. Additionally, there is a newer version of ERSST (V5) that includes new ICOADS data (Huang et. al, 2017 doi: 10.1175/JCLI-D-16-0836.1). Does this version give you any better data coverage for your study site and improve the assessment?

ERSST5 does not provide improvements to ERSST4. In fact, at the coarse resolution of ERSST data (2°x2° gridded) no differences exist between the latest versions of ERSSTv3b, ERSST4 or ERSST5 at the coral reef scale. We have therefore opted for ERSST4 which has been tested more widely in several climatology papers.

17) Lines 87–94 There is another coral Li/Mg study (Fowell et al., 2016) that looked at intra-reef variability with a different coral species from the Atlantic and combine Sr/Ca and Li/Mg. That study should be included here but examine the coral sampling in their figure carefully, there were some sub-optimal sampling issues with their corals. The study Montagna et al., 2014 did look at several species of corals, but not the one used by Fowler 2016, so it worth including the Atlantic study in your discussion of previous Li/Mg studies.

Done. We added to discussion:”Applying the -0.097 mmol/mol per °C for Caribbean *Siderastra sidera* forereef corals would underestimate SST anomalies (Fowell et al., 2016).”

18) Line 154-156 Please give the interval for the air temperature records

Done.

19) Line 161 Please clarify, is the depth of 12 and 13 m the water depth of the top of the coral colony where the cores were removed, or the water depth of the base of the coral? Did you sample the whole colony? Just trying to see how big these colonies were. Table 1 Add units for Length

The depth of 12 and 13 m corresponds to the water depth at the base of the megacolonyes. Approximately less than half of the colony EU3 was sampled as we estimate its height to ~3.50 m and its width to 6m. The core top is therefore at 8.5-9.5 m.

Reviewer 2:

1) *In my opinion, it is circular to reconstruct SST from trace element-to-Ca data of the same time interval and specimens that were used for proxy calibration. The calibration (+ verification) and reconstruction intervals must be separate, or at least this approach would require different colonies (of both species) for calibration/verification and reconstruction, respectively. One possibility is to limit the calibration interval to 2003-2012 and apply the resulting model then just to the pre-2003 section of core EU3. You could also limit the reconstruction interval to 1970-1980 of EU3, because this time interval was – according to Figure 4 – not used for calibration; both options would still miss a verification interval though. Another possibility is to just present regression models spanning the entire lengths of the cores. Alternatively, the authors use the calibration model by D’Olivio et al. (2018) to estimate SST and compare those data to instrumental SST.*

The calibration periods were all indicated in the figure captions. We have now also included it in methods. We have opted for using the regression model that spans the period of overlap with satellite SST between 1981-2012 for the long core EU3 and 2003-2012 for the short core EU2. Table 2 shows the regressions for different calibration periods. However, our main aim is not the quantification of absolute SST, rather relative changes in SST (anomalies) and how they compare to SST records. We have added a paragraph to the discussion in which we assess the multiproxy method from D’Olivio et al. (2018) applied to our corals.

2) *I am missing details on how exactly the Sr/Ca and Li/Mg data were combined. Is the composite Sr/Ca-Li/Mg SST proxy based on a multi-regression model? The equation for this regression model should be provided.*

We calculated the arithmetic mean between Sr/Ca-SST and Li/Mg-SST.

3) *There is some inconsistency regarding the statement on proxy reliability of Sr/Ca and combined Sr/Ca-Li/Mg (compare, e.g., L461-463a and L508/9). Authors should say more clearly (in the main text and in the Abstract) that relative SST changes are reflected equally well by both proxies, but Sr/Ca is still superior to quantify SST. The current Abstract is too vague as the main findings are concerned. It also remained unclear how a combination of Sr/Ca and Li/Mg can “improve SST reconstructions”. Table 3 does not support this claim, because RMSE errors are, on average, lower for Sr/Ca-based SST estimates than for those computed from combined Sr/Ca-Li/Mg.*

We have modified the abstract. It now reads: “In our study, Sr/Ca is still superior to Li/Mg to quantify absolute SST and relative changes in SST.”

Regarding the combined Sr/Ca-Li/Mg-SST, Table 3 does show improvement relative to Sr/Ca or Li/Mg only for shorter core record EU2. RMSEs are lower for the combined proxy. The latter hints at the possibility that combined Sr/Ca-Li/Mg can improve SST reconstructions when both proxies work equally well.

We have modified the section in lines 508 onwards and it now reads:

“Overall the results from this study indicated that Sr/Ca is still superior to Li/Mg and was the most reliable SST proxy when applied to a longer time series. However, the

excellent agreement between Sr/Ca and Li/Mg and their combination in core EU2 demonstrated that both SST proxies and their combination can provide with greater confidence, more reliable SST reconstructions with lower RMSEs (D’Olivo et al., 2018).”

4) *The authors need to better describe the innovative aspect of the study. In particular, they need to highlight (in the Introduction) how their work differs from D’Olivo et al. (2018) and Montagna et al. (2014) (e.g., how much longer were the new coral records in comparison to previous studies; which species were used here and in previous works; how compare the results?).*

In the Introduction we stated that our record is the first application of the Li/Mg proxy on a multi-decadal scale for *Porites* corals.

5) *Linear regression equations may have been the most obvious to use in a mathematical sense, but the actual relationships between trace element-to-Ca ratios (or TE/TE) and SST are nonlinear (compare Gaetani & Cohen 2006; Montagna et al. 2014 etc.). Would it not be better to work with non-linear equations?*

All regression models for the proxies used in this paper are linear models, which hold for the narrow SST range at our coral site (22-29 °C). This is in line with all tropical corals shown in Montagna et al. 2014 and D’Olivo et al. 2018, which fall on a linear regression line. The exponential relationship for Li/Mg vs SST applies to the entire suite of carbonate secreting archives that grow at vastly different SST.

6) *I may be worth mentioning somewhere in the Discussion that the Li/Mg serves a more robust thermometer at low temperatures because of the non-linearity of Li/Mg vs. SST relationship.*

This will be done.

Other issues: L22: It should read “temperature proxy”

Done.

L25-26: Coral = archive, Sr/Ca etc. = proxy. Rephrase sentence accordingly.

Done.

L51-53: It should read “The oceans respond”. If two full sentences are linked by conjunction, a comma must precede “and”.

Done.

L54: replace semicolon by comma

Done.

L55-58: This sentence needs to be rephrased. Three times “climate” in one sentence. Complicated phrasing.

Word ‘climate’ removed in one instance

L76-78: *Here you list the slope ranges, but not provide the values revealing the different strengths of the correlations.*

The strength of correlations is not assessed here; it is merely a listing of slopes from the literature.

L81: *“bio-smoothing effects”*: Explain. Has this expression been used in cited papers? This expression has been used by Gagan et al. (2012) and is cited.

L201: *Tell the reader how long were the calibration intervals in EU3 and EU2.*
Calibration periods don't apply here. They are reported in section 3.3 where it is required.

L226: *I would not call 9 years “long-term”.*
It refers to the longer period used for anomalies, it reads “Longer time series anomalies were calculated relative to the 1981 to 2010 period.”

L228/9: *“record’s lowest ratios” or just “lowest ratios”*
Not changed.

L232: *I do not understand what you mean: “showed higher values between 2003 and 2012”? Offset relative to what? Why “absolute”?*
added “relative to EU3..”

L233: *“Li/Mg between EU2 and EU3: : :”*: Replace “between” by “of”
Done.

L237: *“EU3 showed larger amplitude seasonal variations”*: Delete “amplitude”
Done.

L262/3: *“The Sr/Ca and Li/Mg time series of cores EU3 and EU2 were highly consistent in the period of overlap: : :”*. What does “consistent” mean here? Rephrase.
We deleted this sentence.

L264 (and other occasions in ms): *“Li/Mg performed equally well”*. Odd phrasing; unclear what you mean exactly.
We deleted this sentence.

L285: *“AVHRR-OISST2 display a more limited seasonality: : :”*: Delete one “S” in “OISST”; replace “more limited” by “truncated” or “attenuated”.
Done.

L291: *“reference period” should read “reference periods”*
Done.

L295: *Following Table 2, Sr/Ca-SST slopes are lower than the average reported in Corrège (2006); so this sentence needs to be rephrased.*
Latter part of sentence was removed.

L321: *replace “cooler temperature anomalies” by “smaller anomalies”*
We replaced it with ‘colder anomalies’ because that is what we are trying to convey here.

Table 2 and elsewhere: Use consistent number of decimals throughout ms (including Figures and Tables), i.e., instead of a slope of "-0.04", write "-0.040".

We adjusted all slope and p-values to three numbers of decimals.

Figure 2: Comparison would be facilitated if scaling of the y-axes was the same.

We deliberately did not scale the y-axes the same because absolute proxy values have an offset. Plotting all on the same y-axes would make the seasonality harder to spot. We therefore did not re-scale y-axes in Figures 2 and 4. Figure 3 illustrates all proxy versus proxy relationships on the same axes for both cores, complementing Figure 2.

Figure 4: ditto for x and y-axes. Also check all other figures.

We deliberately did not scale the y-axes the same because absolute proxy values have an offset. Here, we show the regression plots for both cores reflecting their individual relationship with SST. We therefore did not re-scale this Figure 4.

Best regards,
Zinke and co-authors