Dear Luke,

Thank you for your insightful comments on our manuscript. In response to your and the reviewers’ suggestions we have made the following major changes to the manuscript:

- Changed the introduction to better position our paper in the context of what is known and has been done up to now about variable habitat bias in planktonic foraminifera. In this regard, please also see our response to a similar issue raised by reviewer 1.
- Included a comprehensive uncertainty assessment of the data.
- Assessed the sensitivity of our seasonality model to the parametrisation of the seasonal flux pattern.
- Added a section on how habitat tracking can be accounted/corrected for in paleoceanographic reconstructions.

We feel that thanks to these comments our study now has a clearer goal and is better positioned in the existing literature. The uncertainty analysis confirms our original conclusions, strengthening the main message that we wanted to convey with our paper.

We have uploaded point-by-point response to the reviews as separate files and append the revised version of the manuscript (with the changes highlighted) below and we hope that this version meets the criteria for publication in Climate of the Past.

Kind regards,

Lukas Jonkers and Michal Kucera

L. Skinner (Editor)

Dear Lukas and Michal,

I would like to invite you to prepare a full response to the review comments received so far, accompanied by a suitably revised manuscript if you so wish. These review comments cover a number of specific points, which you should strive to address. However they also dwell on two main themes: the need for more thorough statistical analyses; and the need to place your study more squarely in the context of previous work. I would add to this a further proposal to delve deeper into a discussion of what ultimately can be achieved in light of your results, and to add clear detail to your proposals regarding the possibility (if indeed this exists) of improving the accuracy of palaeoceanographic reconstructions and/or data-model comparisons. My impression is that the latter are currently only vaguely described.
The issue of relevant previous work is always a tricky issue to address, given the depth and interpretative nature of the literature; however I do tend to concur with the comment that the manuscript seems often to ‘preach to the converted’ regarding the importance of taking into account habitat bias in the interpretation of foraminifer proxy records, in particular for planktonic species. I also agree that for a paper on such a well-known (if also often completely ignored) issue, the literature review is rather sparse. I would therefore ask you to consider taking this opportunity in revising your manuscript to amplify on this aspect, e.g. for the benefit of readers who will be well aware of the issue, but perhaps not all of the key literature on the topic.

Of course, a full literature review of a topic is not a prerequisite for any given paper on that topic (and is not what I would propose you undertake); however, I suspect that in this case the issue comes to the fore in light of the apparent lack of progress in developing a ‘solution’ for dealing with the issue of habitat bias in planktonic foraminifera. In this regard, one might argue that there are two schools of thought: one that proposes to ‘fix’ foraminifer proxy records by correcting for habitat biases; and another that proposes to accept them (along with their ultimately cryptic nature) and to detect and/or ‘employ’ them as they arise, even if they cannot be known a priori. It seems that the manuscript currently makes a strong point that habitat biases are real (which arguably we already knew), as well as much weaker point regarding what is to be done (or can be done) in light of their existence. I would encourage you to consider engaging in a deeper discussion of the latter issue, as I think this would lend the manuscript greater force.

In order to illustrate my proposal, I would refer to the exclusion of G. bulloides from the analysis: this species does not show the ‘fingerprint’ of habitat bias as defined in the manuscript, and yet it does have a clear habitat bias. This may illustrate a shortcoming of the approach taken, with respect to accounting for habitat bias (or how it has been ‘modeled’); namely that it does not address habitat bias that results in fixed deviations from the annual mean for example. Further concerns might arise with respect to accounting for habitat bias when considering proxy reconstructions of past climatic change and the occurrence of ‘non-analogue’ situations (e.g. strong seasonal perturbations, or stratification changes). Can your analysis be expanded to consider in more detail ‘what can be done’, beyond recognizing that habitat bias is an issue? Would *parallel* multi-proxy and multi-species analyses be helpful for example (e.g. as a set of ‘parallel equations’ for habitat variability within a relatively constrained yet still unknown habitat range)?

I hope that you will find all of these comments helpful, and I look forward to receiving your revised manuscript and response.

Sincerely, Luke Skinner
Quantifying the effect of seasonal and vertical habitat tracking on planktonic foraminifera proxies

Lukas Jonkers* and Michal Kučera

MARUM | Universität Bremen, Leobenerstraße 8, Bremen, Germany.
* correspondence to ljonkers@marum.de

Keywords: planktonic foraminifera, seasonality, depth habitat, proxies

Key points:

- PF depth and seasonal habitat change with temperature
- PF proxy signals therefore underestimate gradients in space and time
- Depth and seasonal habitat variability can be predicted and thus accounted for
Abstract

The composition of planktonic foraminiferal (PF) calcite is routinely used to reconstruct climate change and variability. However, PF ecology leaves a large imprint on the proxy signal. The seasonal and vertical habitat of planktonic foraminifera (PF) species varies spatially, causing variable offsets from annual mean surface conditions recorded by sedimentary assemblages. PF seasonality changes with temperature in a way that minimises the environmental change that individual species experience. Although the controls on depth habitat variability are less well constrained, it is not unlikely that habitat tracking also affects PF depth habitat. While such habitat tracking could lead to an underestimation of spatial or temporal trends and variability in proxy records, most paleoceanographic studies are based on the assumption of a constant habitat. Despite the implications, the effect of this behaviour on foraminifera proxy records has not yet been formally quantified on a global scale. Here we attempt to characterise the effect of habitat tracking on the amplitude of environmental change recorded in sedimentary PF using core top δ¹⁸O data from six species, which we compare to predicted δ¹⁸O. We find that the offset from mean annual near-surface δ¹⁸O values varies with temperature, with PF δ¹⁸O indicating warmer than mean conditions in colder waters (on average by -0.1‰ or 0.4°C per °C), thus providing a first-order quantification of the degree of underestimation due to habitat tracking. We then use an empirical model to estimate the contribution of seasonality to the observed difference between PF and annual mean δ¹⁸O and use the residual Δδ¹⁸O to assess trends in calcification depth. Our analysis indicates that given an observation-based model parametrisation in all species calcification depth increases with temperature. A sensitivity analysis suggests that seasonal habitat adjustment is essential to explain the observed isotopic signal. Consistent with hydrographic conditions, vertical habitat adjustment is more prevalent in tropical species, whereas cold-water species mainly change their seasonality when tracking their ‘optimum’ habitat. Assumptions of constant PF depth or seasonal habitat made when interpreting proxy records are thus invalid. The approach outlined here can be used to account for these effects, enabling more accurate reconstructions and improved data-model comparison.
1. Introduction

The chemical composition of planktonic foraminifera shells reflects the environmental conditions in which they precipitate. Fossil shells therefore serve as the prime source of information about the past state of the oceans. Planktonic foraminifera species are non-uniformly distributed across the world ocean, indicating they inhabit distinct ecological niches (Bé and Tolderlund, 1971). Indeed, such habitat preferences allow paleoenvironmental reconstruction based on fossil foraminifera assemblages. However, habitat preferences also affect paleoceanographic reconstructions based on single species. Planktonic foraminifera inhabit a wide vertical range of the water column and often show distinct variability in their seasonal abundance (e.g. Field, 2004; Tolderlund and Bé, 1971; Fairbanks et al., 1980; Jonkers et al., 2010; Jonkers et al., 2013; Deuser et al., 1981). Hence, it is well-known that, rather than reflecting annual mean surface conditions, the average proxy signal in sedimentary planktonic foraminifera is weighted towards conditions at the depth and season of calcification (Mix, 1987). While species-specific seasonality and calcification depth are often taken into account, it is implicitly assumed that both remain constant in time and space. This assumption is at odds with observations from the present-day ocean and this could have important implications for reconstructions of climate change and proxy calibrations based on core top data.

Plankton tow and sediment trap studies have shown large variability in the phenology within individual planktonic foraminifera species (e.g. Tolderlund and Bé, 1971; Zarić et al., 2005). A recent review demonstrated that this variability is widespread and follows a predictable pattern consistent with the concept that foraminifera track their optimum habitat (Jonkers and Kučera, 2015). Two broad ecological groups with different seasonality patterns were found; it was shown that outside the tropics, warm-water species narrow their occurrence into the season of maximum temperatures. The seasonality in cold-water species also shows a clear relationship with near surface temperature as their flux peak generally occurs earlier in the year in warmer waters (Jonkers and Kučera, 2015). While the latter trend appears to be driven by the timing of food availability rather than by temperature, both patterns have the same effect on the fossil record since they reduce the amplitude of (temperature) change that bulk samples of their fossil shells will reflect (Kretschmer et al., 2016;
Fraile et al., 2009b). This indicates that changes in seasonality need to be taken into account when interpreting proxy records.

The depth habitat of planktonic foraminifera species also shows clear variability in space and time (Field, 2004; Fairbanks and Wiebe, 1980; Schiebel et al., 2001; Peeters and Brummer, 2002; Rebotim et al., 2017). Food and light availability, (thermal) stratification and temperature have all been suggested to be potential drivers of the vertical distribution of planktonic foraminifera (Schiebel et al., 2001; Fairbanks and Wiebe, 1980; Ortiz et al., 1995; Salmon et al., 2015; Fairbanks et al., 1982). In contrast to seasonality a global overview is lacking and the exact controls on depth habitat variability within species remain poorly constrained. The issue of changing depth habitat is further complicated by the tendency that many foraminifera likely migrate in the water column during their life and add proportionally more calcite at later stages in their life, potentially resulting in a mismatch between depth habitat and calcification depth (e.g. Duplessy et al., 1981). Whereas depth habitat can be directly observed, calcification depth is generally estimated from Mg/Ca or stable oxygen isotope data and hence more uncertain. Nevertheless, depth habitat and calcification depth are related, as deeper dwelling species will also have a greater calcification depth.

Here we hypothesise that, similar to seasonality, the depth habitat and therefore calcification depth is related to temperature and that changes in temperature will lead to adjustments in depth habitat such that the environmental changes planktonic foraminifera experiences during their life cycle are minimised. The combined effect of seasonal and depth habitat tracking would be that temporal and spatial gradients in planktonic foraminiferal proxy records are reduced compared to the gradients in the mean annual value of the reconstructed parameter. For instance, in the case of temperature, it would lead to positive offsets from annual mean near-surface temperatures at times of cooling as planktonic foraminifera shift their seasonal and depth habitats to the warmer season and/or to shallower depths. The partitioning of this effect into seasonality and depth habitat likely varies by region, depending on the ratio of seasonal over vertical temperature variability in the upper water column (Fig. 1). This implies that for tropical species constraining the depth habitat will be more important than seasonality, whereas the opposite is true for species living in mid- and high latitudes.
Variability in seasonal and vertical habitat within individual species is well known and the dependency of foraminifera habitat on climate has been implicated before (Ganssen and Kroon, 2000; Mix, 1987; Mulitza et al., 1998; Jonkers and Kučera, 2015; Skinner and Elderfield, 2005). In addition, several modelling studies have investigated the potential dampening effect of seasonality (Fraile et al., 2009a; Fraile et al., 2009b; Kretschmer et al., 2016). Surprisingly for an effect that is so obvious and may strongly affect proxy records, studies attempting to demonstrate the effect of habitat tracking, or the non-passive recording by foraminifera (and how to deal with the problem) are scarce.

In addition, while habitat tracking behaviour can be expected, a systematic quantification of the effect on planktonic foraminifera proxies based on observational evidence, as well as an assessment of the respective roles of seasonality and depth habitat, is lacking. Essentially, the existence of a habitat tracking effect on proxy signals in planktonic foraminifera has either been treated by attempting to derive a “correction factor” which was applied more or less uniformly (e.g. Barker et al., 2005), or by trying to detect the effect by multi-species analyses (Skinner and Elderfield, 2005). In both approaches, the most difficult aspect was to deal with the possibility of the effect of habitat tracking being variable.

Here we aim to bring the issue (back) to the attention of the community, stimulate discussion and suggest some ways forward. We use core top stable isotope data to first demonstrate that foraminifera proxies are indeed affected by habitat tracking. We then show that this effect can be parametrised and assess the relative importance of variable seasonality and depth habitat. Our findings have important implications for the interpretation of paleoceanographic records and could help to bridge the gap between paleoceanographic data and model simulations.

2. Data and approach

If shifts in depth and seasonal habitat would act to minimise the change in the ambient environment of the planktonic foraminifera, then the proxy signal preserved in their shells should show an offset from mean annual values that varies with temperature. To test this conjecture we compare core top stable oxygen isotope ($\delta^{18}O$) values from different species with seasonally and vertically resolved
estimates of equilibrium $\delta^{18}O$. We use quality controlled data with strict age control (chronozone 1-4) from the MARGO core top dataset (Waelbroeck et al., 2005). This dataset contains data for six morpho-species of planktonic foraminifera: *Trilobatus sacculifer* (n = 38), *Globigerinoides ruber* (white (n = 131) and pink (n = 20) varieties), *Globigerina bulloides* (n = 131), *Neogloboquadrina incompta* (n = 46) and *Neogloboquadrina pachyderma* (n = 253); the existence of different genotypes within these species was not taken into account. We exclude samples from the Mediterranean from our analysis because of difficulties of estimating seawater $\delta^{18}O$ in this evaporative basin, and removed *T. sacculifer* data from Pacific Ocean sites deeper than 3 km as these are thought to be affected by dissolution (Wu and Berger, 1989).

The median standard deviation of replicate measurements within the dataset is 0.12 ‰. This value reflects analytical uncertainty as well as uncertainty associated with different integration time of each sample and selective sampling due to low and variable numbers of shells used for analysis.

We compare the planktonic foraminifera $\delta^{18}O$ ($\delta^{18}O_{foram}$) to predicted $\delta^{18}O$ ($\delta^{18}O_{eq}$) calculated using $\delta^{18}O$-temperature equation by Kim and O’Neil (1997). Calibration uncertainty is estimated from measurements on planktonic foraminifera shells from sediment traps from a period of complete mixing of the upper water column (Jonkers et al., 2013). This value (0.2 ‰) is larger than the calibration error based on laboratory cultures (Bemis et al., 1998). Following the approach of LeGrande and Schmidt (2006) we estimate seawater $\delta^{18}O$ using regionally defined salinity-$\delta^{18}O_{sw}$ relationships for the upper 200 m using the Global Seawater Oxygen-18 Database (Schmidt et al., 1999). Standard errors of prediction vary regionally and reach a maximum of 0.91 ‰ in the Arctic. Conversion from the SMOW to PDB scale was done by subtracting 0.27 ‰ (Hut, 1987). Temperature and salinity data were taken from the World Ocean Atlas 2001 (Boyer et al., 2002; Stephens et al., 2002) and area weighted averages were obtained from the four 1 degree areas surrounding each core top position.

We start with comparing the observed $\delta^{18}O_{foram}$ to annual mean $\delta^{18}O_{sw}$ for the upper 50 m as this is the depth interval where these species are most likely to calcify. We use a Monte Carlo approach to estimate the uncertainty on the relationships between predicted and observed values. We assume a normal distribution of the uncertainty and perform regressions on 10,000 representations of the data.
sampled from/modified within the combined range of uncertainty based on the square root of the sum of the squared errors mentioned above.

3. Results and discussion

3.1 Habitat tracking in planktonic foraminifera

The observed $\delta^{18}$O of all species show deviations from expected mean annual $\delta^{18}$O$_{eq}$ by up to 3‰. If our hypothesis of habitat tracking holds, the $\Delta\delta^{18}$O ($\delta^{18}$O$_{foram}$ - $\delta^{18}$O$_{eq}$) should show a positive relationship with temperature. Indeed, the $\Delta\delta^{18}$O$_{annual.mean}$ values of the three tropical species (G. ruber (white and pink) and T. sacculifer) show a significant positive relationship with mean annual temperature (Fig. 2). The slopes vary between 0.04 and 0.14 ‰ °C$^{-1}$. In general, $\Delta\delta^{18}$O$_{annual.mean}$ values are close to 0 at high mean annual temperatures and negative $\Delta\delta^{18}$O$_{annual.mean}$ values, indicating calcification temperatures higher than annual mean, occur in colder waters. Among the cold-water species, N. incompta also displays a consistent positive relationship between temperatures and $\Delta\delta^{18}$O$_{annual.mean}$ above ~5°C, whereas below this temperature the relationship has the opposite sign (Fig. 2). These observations at low temperature stem from samples in the Nordic Seas outside the direct path of the North Atlantic Drift. We suspect that these (partly) reflect right-coiling variants of N. pachyderma (Bauch et al., 2003) and we have therefore excluded them from further analysis.

$\Delta\delta^{18}$O$_{annual.mean}$ values of N. pachyderma are generally positive and show an increased spread towards higher values in warmer waters (Fig. 2). G. bulloides is the only species that does not show any trend in $\Delta\delta^{18}$O$_{annual.mean}$ modal values are close to 0, but the distribution is skewed towards positive offsets (Fig. 2). This species was therefore excluded from further analyses. In summary, five of the six species analysed display a pattern in their $\Delta\delta^{18}$O$_{annual.mean}$ that is consistent with the hypothesis of habitat tracking acting to minimise the temperature change they experience. In all cases, the relationships are associated with substantial noise, but they are statistically significant and have similar slopes with the same sign. Such coincidence is unlikely to have arisen by chance and we conclude that the data...
demonstrate the existence of a variable habitat bias in foraminifera proxies, a part of which is predictable as a function of temperature.

3.2 Seasonality

Next, using simple empirical models for seasonality we assess how much of the trend in $\Delta^{18}O_{\text{annual}}$ could be due to changes in seasonality alone. To this end we calculate a flux-weighted $\delta^{18}O_{\text{eq}}$ for the upper 50 m of the water column using a simple seasonality model and compute $\Delta\delta^{18}O_{\text{season}}$. Based on previous work (Jonkers and Kučera, 2015) we describe the log-transformed flux pattern as a sine wave of which we change the amplitude and phasing as a function of mean annual temperature. For tropical species we fix the peak in the flux in September (March in the southern hemisphere), which is generally the warmest month, and increase the amplitude linearly with decreasing temperature with a species-specific slope derived from sediment trap data (Fig. 3; Table 1). While this model does not account for the random peak flux timing at high temperatures (Jonkers and Kučera, 2015), the seasonal amplitude of the shell flux and of $\delta^{18}O_{\text{eq}}$ are very small at these temperatures, and the model serves as a reasonable approximation of the seasonality pattern that characterises this species group. For cold-water species we fix the amplitude at the average value for this group (0.66) and vary the timing of the peak flux as a function of temperature (Fig. 3). Below a critical low temperature we set the peak timing to September and above a critical high temperature to March (reversed for Southern hemisphere); between these temperatures, the modelled flux pattern has two peaks a year that linearly shift towards earlier in the year in colder waters (table 2). While simple, this model represents a realistic scenario, derived from observations and can thus be applied to the studied species. We gauge the effect of the flux weighting by determining i) the change in the spread of the $\Delta^{18}O$ values and ii) the slope of the $\Delta^{18}O$-temperature relationship.

The uncertainty of the $\Delta^{18}O_{\text{season}}$ values derived when using this model are initially estimated using the same Monte Carlo approach with the same error estimates as outlined above. We later test the sensitivity of the result to the parametrisation by repeating the analyses with formulations assuming stronger/weaker flux seasonality.
Accounting for seasonality using this model reduces the root means square error (RMSE) in the Δδ¹⁸O values of G. ruber (pink) by 21% and the slope by 37% (Fig. 4). For G. ruber (white) the values are 12 and 77 % respectively (Fig. 4). Due to large positive Δδ¹⁸O values for T. sacculifer at high temperatures, flux-weighting has a negligible effect on the spread in the values (1 %), but it reduces the slope by 22 % (Fig. 4). The values for N. incompta are 47 and 57 % and for N. pachyderma 16 and 51 % (Fig. 4). For none of the species this adjustment for seasonality leads to an increase in the spread of the Δδ¹⁸O values, on the contrary, in most cases the predicted δ¹⁸Oeq are closer to the δ¹⁸Oforam. This indicates that even by using a simple empirical model for seasonality, predictions of the fossil signal can be improved, as long as the mean annual temperature is constrained from independent data.

3.3 Calcification depth

In none of the species investigated here, the adjustment for seasonality completely removes the relationship between Δδ¹⁸Oseason and temperature, although in the case of G. ruber seasonality adjustment could remove the Δδ¹⁸O-MAT slope when the full range of uncertainties is considered (confidence intervals on slopes as shown in Figure 3 contain zero). Therefore, one may assume that at least a part of the relationship could reflect an adjustment of calcification depth. To investigate if the trends in the Δδ¹⁸Oseason reflect an increase of calcification depth towards the tropics, we determine the depth at which Δδ¹⁸Oseason is smallest and assess if there is a relationship between this apparent calcification depth and mean annual temperature. This analysis reveals that of the tropical species G. ruber (pink) shows the shallowest (apparent) calcification depth, followed by G. ruber (white) and T. sacculifer (Fig. 5). All species show an increase in calcification depth with temperature (even though the scatter is large). This rules out that the relationships between Δδ¹⁸O and temperature (Fig. 2 and 4) reflect calcification at a constant, but greater depth than in the near surface layer. Rather, this correlation is consistent with the hypothesis that planktonic foraminifera (passively) track an optimum vertical habitat. N. incompta has variable calcification depths that show a steep slope with temperature (Fig. 5). The positive Δδ¹⁸Oseason values of N. pachydermo indicate a calcification depth consistently below 50 m (Fig. 5).
Next we use the linear relationships between apparent calcification depth and temperature (Fig. 5) to explain the fossil signal. We thus adjust the δ²⁸O$_{season}$ to a depth-specific signal, using the depth-temperature relationship identified earlier (Fig. 5) to calculate Δδ²⁸O$_{season, depth}$. In G. ruber (pink) this leads to a further 50 % reduction in the RMSE and a Δδ²⁸O$_{season, depth}$-temperature slope that is close to 0 (Fig. 6). In G. ruber (white) the reduction in the spread in the data is more modest (12 %) and so is the reduction in slope (Fig. 6). For T. sacculifer also only modest additional reductions are achieved: 8 and 11 % for RMSE and slope, respectively (Fig. 6). Among the cold-water species N. incompta shows the clearest relationship between Δδ²⁸O$_{season, depth}$ and temperature and adjustment for calcification depth yields a reduction of the RMSE of 8 and of the slope of 46 %, which is now within the uncertainty range of 0 (Fig. 6). In N. pachyderma no further reduction in the slope is achieved and the RMSE decreases by 22 %.

### 3.4 Seasonality vs. depth habitat

Our analysis allows assessing the relative contribution of seasonality and calcification depth change in explaining the variable Δδ¹⁸O$_{season, mean}$ for species where temperature seems important for determining their habitat. In general, the improvement of the prediction of the δ¹⁸O$_{season}$ is larger for the slope of the Δδ¹⁸O-temperature relationship than for the spread in the Δδ¹⁸O values (Fig. 7). This may point to some degree of inherent noise in the observations (e.g. related to different size fractions used for the measurements (Friedrich et al., 2012)), or it could also be due to uncertainty in the δ¹⁸O$_{eq}$ values, which are based on climatology and salinity-based estimates of δ¹⁸O$_{eq}$. Moreover, the noise may also reflect the simplicity of the seasonality model we have used. Nevertheless, G. ruber (pink) and N. incompta show coherent behaviour with respect to both parameters (Fig. 7). For N. incompta seasonality explains most of the trend in Δδ¹⁸O$_{season, mean}$ whereas for G. ruber (pink) depth habitat appears more important. This is consistent with their distribution: N. incompta predominantly inhabits high and mid latitudes where seasonal temperature change is larger than vertical temperature gradients and G. ruber (pink) is restricted to the tropics where the opposite situation prevails (Fig. 1). This pattern provides support for our approach and suggests that both seasonality and depth habitat variability are important for interpretation of the proxy signal preserved in the
sediment. The picture is less clear for *G. ruber* (white) and *T. sacculifer*. For the latter species the improvement in the prediction of their $\delta^{18}O$ is generally smaller, which may be due to a remnant dissolution signal at the high temperature end of the species distribution in the Pacific. For *G. ruber* (white), the signal/noise ratio in the data appears lower than in the other species, which may reflect a disproportionate effect of secondary variables, such as changing proportionality and inconsistent recognition of the ecologically distinct morphotypes (Steinke et al., 2005) that are now assigned to different taxa (Aurahs et al., 2011).

An important caveat in the attribution of the improvement in the prediction of the fossil proxy signal to either seasonality or calcification depth is the form and parametrisation of the seasonality model used. We have explored the sensitivity of our model to changes in the slope and intercept of the flux amplitude-temperature relationship (Sfig. 2). This suggests that the formulation of seasonality in our model is conservative: weaker seasonality parametrisation leaves much larger residuals and a slope that cannot be accounted for by depth habitat adjustment. However, we note that in the case of *G. ruber* pink there exists a parametrisation of flux seasonality that leads to a greater improvement in the prediction of $\delta^{18}O_{\text{foram}}$ and implies a constant habitat depth adjustment. However, we feel that the parametrisation based on actual data (Jonkers and Kučera, 2015), even if conservative, is the most realistic. It is also important to realise that the sine wave as a template for the seasonal flux pattern is only an approximation and seasonal flux pulses are often narrower and more focussed, leading to flux-weighting to a shorter period within the year. This too implies that the model used here is therefore a conservative estimate of the importance of seasonality.

Next, implicit in our approach is the assumption that planktonic foraminifera form their skeleton in accordance with inorganic calcite precipitation and that their $\delta^{18}O$ can be described using the equation by Kim and O’Neil (1997). While this appears to be the case for some species (Jonkers et al., 2010; Jonkers et al., 2013; Loncaric et al., 2006), there are also indications that, in particular for tropical species, different equations are more appropriate (Mulitza et al., 2003; Spero et al., 2003). Species-specific paleotemperature equations proposed by the latter authors have a non-quadratic
form, but almost identical slopes as the Kim and O’Neil (1997) equation yet are offset by 0.3-0.6 ‰, with the offset increasing with temperature. For instance, using the Mulitza et al. (2003) equation for T. sacculifer would lead to more positive $\Delta\delta^{18}O_{annual\ mean}$ values and slightly steeper $\Delta\delta^{18}O$-temperature relationships (Fig. 8). This suggests a generally greater calcification depth and would change the attribution of depth habitat and seasonality influence, rendering depth habitat more important (Fig. 7). However, it would not affect our main conclusion that the proxy signal of planktonic foraminifera is affected by habitat tracking.

### 3.5 Additional factors affecting foraminifera proxies

Five out of six species analysed show a temperature dependency of the offset between $\delta^{18}O$ of the foraminiferal shells and the annual mean $\delta^{18}O$ of the upper water column (Fig. 2). In addition, these species show a positive relation between apparent calcification depth and temperature (Fig. 5). Together, these observations provide a strong indication that temperature, either directly or by acting on other temperature-related variables, causes changes in the habitat of foraminifera. Such an important role for temperature in predicting the vertical and seasonal habitat is not unexpected given that temperature appears to be strongly correlated to the spatial distribution of species (Morey et al., 2005; Bé and Hutson, 1977), their flux (Zaric et al., 2005) and seasonality (Jonkers and Kučera, 2015) and appears important for test growth (Lombard et al., 2009).

Several studies have shown that formation of secondary calcite layers (e.g. gametogenic calcite or a crust) at the end of the life of a specimen, presumably deep in the water column could be responsible for higher $\delta^{18}O$ of sedimentary foraminifera compared to those collected in the upper water column (Duplessy et al., 1981; Bé, 1980). To the best of our knowledge there is no evidence that such secondary calcite is formed with a different isotopic (dis)equilibrium than the lamellar calcite. We therefore assume that our inferences are not affected by differences in calcification during ontogeny. Nevertheless, the addition of such a crust in deeper (colder) waters could in principle lead to the observed increase in apparent calcification depth with temperature because of steeper vertical temperature gradients in the tropics. However, foraminifera shell mass increases exponentially during growth, and the last chambers that make up most of the test mass are formed in the last few days of
their life, presumably close to the time of the secondary calcite formation (Bé, 1980). The compositional contrast between the bulk of the lamellar calcite and the crust calcite is thus likely to be smaller than estimated from the comparison of surface taws and sediment (cf. Jonkers et al., 2016). Consequently, the apparent calcification depth we infer here likely incorporates this effect and the increase in apparent calcification depth that we observe most likely reflects habitat adjustment.

Next to temperature and δ¹⁸O seawater, the δ¹⁸O of foraminiferal calcite is to a lesser degree also influenced by the CO₃⁻ concentration in seawater (Spero et al., 1997). Because of the generally positive correlation between temperature and [CO₃⁻] in seawater, the trends we observe in Δδ¹⁸Oannual values could be dampened by a CO₃⁻ influence. However, the CO₃⁻ effect is only modest (0.002 ‰ µmol⁻¹ kg⁻¹) and to fully account for the on average 1‰ difference we observe over the temperature range in our dataset, unrealistically large gradients in [CO₃⁻] would be required. The trends thus most likely dominantly reflect real changes in the habitat of planktonic foraminifera.

While the majority of the species investigated here show clear indications of temperature-dependent depth and seasonal habitat variability, the picture for N. pachyderma is less clear. In the species most of the trend in Δδ¹⁸Oannual values appears driven by an increased spread in Δδ¹⁸O at higher temperatures (Fig. 2). Some of these values are unrealistically large and stem from observations in the northern North Atlantic south of 50°N, thus outside the general distribution range of the species. This suggests that these observations may reflect expatriated specimens that calcified in colder regions or may point to inaccuracies in the chronological control and reflect (partly) shells of glacial age. Alternatively, these samples could be affected by admixture of sinistrally coiled N. incompta (Darling et al., 2006). It is puzzling though that the effect of seasonality is not larger since the species shows a clear latitudinal shift in the timing of the peak flux (Jonkers et al., 2010; Jonkers et al., 2013; Jensen, 1998; Wolfteich, 1994; Kohfeld et al., 1996). However, the species is also known to inhabit a broad, but generally deeper, zone of the upper water column (Carstens et al., 1997; Pados and Spielhagen, 2014) where seasonal temperature is smaller than in the near surface layer, possibly rendering a seasonality effect difficult to detect.
Even though the absence of a Δδ¹⁸O_annual-mean temperature trend in *G. bulloides* may suggest that this species holds the best promise of providing reconstructions of mean annual near surface conditions (Fig. 2), its Δδ¹⁸O_annual-mean is noisy, suggesting that caution is required to interpret the species proxy signal. Similar to *N. pachyderma* this species also shows clear latitudinal changes in seasonality (Jonkers and Kučera, 2015; Tolderlund and Bé, 1971). However, *G. bulloides* is characterised by considerable cryptic diversity (Darling and Wade, 2008). Possible genotypic ecological differences could therefore obscure ecological patterns at the morphospecies level. Alternatively, being an opportunistic species, depth and seasonal habitat variability of *G. bulloides* may be driven by other parameters than temperature. Indeed, previous studies have shown that the distribution of this species is driven by food availability (Schiebel et al., 1997; Jonkers and Kučera, 2015). Whether or not the species shows habitat tracking and how this would affect its fossil record remains unclear, but we caution that the result of our study cannot be taken to indicate that proxy records from this species record the actual magnitude of environmental change.

Since planktonic foraminifera seasonality and calcification depth appear to be affected by habitat tracking, our ability to accurately reconstruct past ocean properties would benefit from improved understanding of the drivers of their habitat variability. In particular, the controls on depth (and calcification) habitat remain poorly constrained. Due to logistical challenges, very few studies exist that have attempted to systematically investigate depth habitat variability. On the other hand, the realisation of the importance of habitat tracking in planktonic foraminifera could help to formulate more realistic, mechanistic models of planktonic foraminiferal distribution in time and space (e.g. Lombard et al., 2011) and further improve our capabilities of forward proxy modelling. Even though our analysis reconfirms that an observed change in a proxy value reflects a change in the climate state as well as a change in the species habitat, very few studies explicitly acknowledge the possibility and effects of temporally variable habitat (e.g. Antonarakou et al., 2015). This likely reflects a combination of ignorance of the problem and the lack of clear solutions to account for variable habitat bias. At any rate, our observations and the simple conceptual modelling exercise shown here serve as reminder that assumptions of constant seasonality and depth habitat are not universally valid and the implications thereof are likely to be substantial.
3.6 Implications and outlook

Habitat tracking behaviour of planktonic foraminifera has important implications for paleoceanographic reconstructions. For example, it implies that the temperature niche of planktonic foraminifera inferred from their abundance in the sediment (e.g., Kucera, 2007) may be overestimated since their occurrence is not related to mean annual sea surface temperature, but rather by whether their temperature niche is realised at any depth or season. It should thus be possible to define planktonic foraminifera temperature ranges (sensitivity) more precisely, which may help to improve transfer functions and is important for understanding of their ecology.

Another consequence of habitat tracking is that spatial and temporal differences reflected in the sedimentary foraminifera represent an underestimation of the actual gradients in the mean conditions, because temperature change forces the foraminifera to live in a seasonal or vertical 'window' where conditions are closest to optimal (cf. Jonkers and Kučera, 2015). We observe considerable variability in the slope of the $\Delta^{18}O_{\text{annual}}$-temperature relationships, but the average for the four species that show the clearest signal (G. ruber (pink and white), T. sacculifer and N. incompta) is 0.1% °C$^{-1}$ (Fig. 2). This is equivalent to a 40% (0.4 °C °C$^{-1}$) underestimation of reconstructed temperature change.

The existence of such underestimation can be observed through comparison of time series of different temperature proxies. Previous studies have shown that Holocene temperature trends and temperature variability inferred from foraminiferal Mg/Ca ratios are generally of lower magnitude than those derived from alkenone unsaturation indices (Gill et al., 2016; Leduc et al., 2010). While it is not a priori clear that the alkenone signal is unaffected by seasonal habitat variability of cocolithophores (Rosell-Melé and Prahl, 2013), this comparatively low variability inferred from planktonic foraminifera proxies provides support that habitat tracking minimises amplitude of the recorded environmental change. Comparison of Mg/Ca-derived and transfer function based temperature evolution across the deglaciation provides further indications that habitat tracking dampens the foraminifera proxy signal (Fig. 9). While both proxies indicate a clear warming step during the deglaciation, the amplitude of the Mg/Ca-based estimate is significantly lower. In addition, the single species Mg/Ca-temperature estimate lacks the smaller cooling and warming trends seen in
the transfer function-based estimate during the glacial and Holocene respectively. Using the linear 
\[ \Delta \delta^{18}O_{\text{annual mean}} - \] temperature relationships (Fig. 2) we also predicted the G. ruber (pink) temperature 
signal assuming that the assemblage-based temperatures represent an accurate estimate of mean 
annual temperature and using a conversion from \( \delta^{18}O \)-temperature sensitivity of 0.25 \( \% \ ^{\circ} \text{C}^{-1} \) (Fig. 9).

The high degree of agreement between the predicted and observed temperature evolution provides 
qualitative support for the idea that habitat tracking reduces the amplitude of the foraminifera 
proxy signal.

Accounting for the dampening effect due to habitat tracking would likely increase the magnitude of 
reconstructed climate change as well as estimates of climate variability on longer time scales. This 
could have profound implications for inferred climate dynamics; it may mean, for instance, that 
estimates of climate sensitivity (e.g. Snyder, 2016) may be too low (or at least that the uncertainty of 
the estimate can be reduced). In addition, model-data comparison indicates that climate models 
 systematically underestimate temperature variability (Laepple and Huybers, 2014), which has 
implications for both attribution of ongoing climate change as well as for climate predictions. Since 
habitat tracking dampens variability in the foraminifera proxy record, the mismatch between 
modelled and reconstructed climate variability may be even larger.

It is important to note that habitat tracking would not only affect stable isotope records and Mg/Ca-
based temperature estimates, but any geochemical proxy based on planktonic foraminifera. However, 
the size of the effect will depend on the magnitude of the seasonal and vertical gradients in the 
parameters that are inferred. Recognising habitat tracking and deconvolving the effects of seasonality 
and calcification depth in data from the fossil record is however not straightforward. For instance, 
minor changes in mean temperature may be accommodated by changes in the habitat of foraminifera 
and remain invisible.

The need to take habitat tracking into account is clear (and reinforced by this study). Correcting for 
the effect on paleoceanographic records will be an iterative approach. Ideally, several approaches 
should be combined – and their consistency checked - to derive at the most accurate reconstruction 
of past climate associated with the most meaningful estimate of the uncertainty. Several lines of
approach are possible to account for underestimation due to habitat tracking. They range from very simple and widely applicable to more sophisticated, but harder to apply as they require more input parameters to model the foraminifera habitat.

The simplest approach would be to use the observed slopes of Δ818Oannual-MAT relationships (Fig. 2) to correct the dampening effect. This approach can be applied to any single-species planktonic foraminifera proxy time series without the need for additional data. However, it is a ‘black box’ method that relies on parametrisation derived from core top observations, and it may not work under oceanographic conditions, which are not represented today. Nevertheless, the similarity between the Mg/Ca signal of G. ruber (pink) and the prediction based on the annual mean reconstructed temperature in the example in (Fig. 9) suggest that this method provides a useful first-order approximation of the actual amplitude of temperature change.

Alternatively, multi-proxy and/or multi-species approaches can be applied to observe offsets between species and proxies and use this information to directly constrain the variable habitat bias (e.g. Skinner and Elderfield, 2005). To use this method to correct for the variable bias, better constraints than currently available on the calibration of the individual proxies are needed. This is because most proxies are affected by multiple parameters, rendering a multi-proxy (or multi-species) dataset underdetermined. In this context, a potential shortcut could be to build on the central premise of assemblage transfer functions, i.e. that a species becomes rarer further away from its optimum habitat, and investigate whether relative abundance can be used to correct for habitat tracking. This approach integrates all the environmental forcings contained in fossil assemblages and it would provide estimates of habitat tracking independent of the reconstructed parameter.

Finally, habitat tracking could be accounted for using direct proxy modelling. This approach requires a priori knowledge of the habitat predictor(s) and their change through time. This renders the approach suitable for comparison of climate model simulations with proxy data. If the predictor can be reconstructed independently, the approach could, in theory, also be used in purely proxy-based studies. The simplest strategy is represented by the conceptual model of Mix (1987). In this study, we build on an empirical parametrisation of flux seasonality using temperature as the single predictor and derive a likely habitat depth parametrisation. This is of course a simplification that could be
outcompeted by more complex ecological models that take multiple factors affecting foraminifera

habitats into account (Lombard et al., 2011; Fraile et al., 2008).

4. Conclusions

Through comparison of observed and predicted δ¹⁸O data of six common planktonic foraminifera we have demonstrated that the average geochemical signal preserved in a population of fossil shells shows a temperature-dependent offset from mean annual sea surface conditions. This most likely reflects shifts in the seasonal and depth habitat in response to temperature, or temperature-related environmental changes (Fig. 9). As a consequence of this behaviour, the fossil record of these species, and likely also others, does not reflect the full range of climate variability. Our analysis indicates that spatial and temporal gradients in temperature may be underestimated by 40%, clearly highlighting the need to account for climate-dependent habitat variability in the interpretation of paleoceanographic records based on planktonic foraminifera. Using a simple empirical model we attempted to assess the relative influence of seasonality and depth habitat variability. While improvements to this empirical approach are possible, we observe species-specific partitioning of depth habitat versus seasonality that appears consistent with oceanographic conditions within their areal distribution. In the tropical species G. ruber (pink) we find that habitat tracking is primarily due to adjustments in the calcification depth. This is in agreement with the larger vertical than seasonal temperature gradients in the tropics. The offsets from annual mean surface conditions in N. incompta, on the other hand, appear dominantly driven by changes in the seasonality, consistent with the dominance of seasonal over vertical temperature variability in the regions where it occurs. Our data underscore the importance of ecology in setting the climate signal preserved in fossil foraminifera. The recognition of predictable habitat tracking will help to improve the accuracy of paleoceanographic reconstructions and aid model-data comparison.

Moved up [2]: The need to take habitat tracking into account is clear (and reinforced by this study). Correcting for the effect on paleoceanographic records will be an iterative approach. Ideally, several approaches should be combined – and their consistency checked – to derive at the most accurate reconstruction of past climate associated with the most meaningful estimate of the uncertainty.

Deleted: homeostatic
Acknowledgments

We thank Stefan Mulitza and Thomas Laepple for valuable discussions and constructive comments from two reviewers and the editor Luke Skinner, which helped to improve this manuscript. LJ was supported by the German climate modelling initiative PalMod funded by the Federal Ministry of Education and Research (BMBF). The data used in this study are all in the public domain, but a data sheet is provided as electronic supplement. R (R core team, 2016) code is available upon request from LJ.

Tables

Table 1: temperature-amplitude relationships for the modelled flux pattern of tropical species based on Jonkers and Kucera (2015).

<table>
<thead>
<tr>
<th>Species</th>
<th>intercept</th>
<th>slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>G. ruber (pink)</td>
<td>2.16</td>
<td>-0.07</td>
</tr>
<tr>
<td>G. ruber (white)</td>
<td>0.99</td>
<td>-0.02</td>
</tr>
<tr>
<td>T. sacculifer</td>
<td>0.85</td>
<td>-0.02</td>
</tr>
</tbody>
</table>

Table 2: critical temperatures (°C) that determine the phasing of the shell flux of cold-water species. Between these two temperatures the flux pattern is characterised by two peaks a year that shift as a function of temperature to earlier in the year in warmer water (Jonkers and Kucera, 2015).

<table>
<thead>
<tr>
<th>Species</th>
<th>T.crit.lo</th>
<th>T.crit.hi</th>
</tr>
</thead>
<tbody>
<tr>
<td>N. incompta</td>
<td>9</td>
<td>15</td>
</tr>
<tr>
<td>N. pachyderma</td>
<td>-5</td>
<td>7</td>
</tr>
</tbody>
</table>

Figure captions

Fig. 1: Distribution of core top δ¹⁸O data used in this study. Background colours represent the log₁₀ ratio of the temperature range at the surface to the temperature range in the annual mean values.
between 0 and 200 m depth. Blue colours thus indicate areas where seasonal temperature gradients are larger than vertical gradients and red colours indicate the opposite. The thin black contour line shows the zero level of this ratio.

Fig. 2: Temperature dependent offsets between predicted annual mean near surface and observed δ¹⁸O (Δδ¹⁸Oannual−mean), suggesting habitat tracking. All species except G. bulloides show a trend in Δδ¹⁸Oannual−mean values with mean annual temperature in the upper 50 m (MAT₀−50m) of the water column suggesting that planktonic foraminifera adjust their habitat to minimise temperature change in their environment. Histograms show the spread in the Δδ¹⁸Oannual−mean values. The root mean square error (RMSE) and the linear slope (m ± 90% confidence interval) of the Δδ¹⁸O-temperature relationship are indicated in the upper left corner of each panel. The uncertainty envelopes depict the 5 to 95 percentile of the uncertainty on the MAT-Δδ¹⁸Oannual−mean relationships, which are indicated by the solid black line. The grey dots in the panel for N. incompta show the data that are excluded from further analyses as they most likely stem from right-coiling morphotypes of N. pachyderma.

Fig. 3: Schematic representation of the seasonality model. Upper panels show the annual flux patterns; colours indicate temperature, where blue is cold and red is warm. Lower panels show the timing of the peak in the year. For a more detailed explanation of the model see section 3.2 and tables 1 and 2.

Fig. 4: Offset between flux-weighted predicted using the seasonality model depicted in Fig. 3 and observed δ¹⁸O (Δδ¹⁸Oannual−mean). Grey symbols represent Δδ¹⁸Oannual−mean. Note the general reduction in the spread of the data (RMSE) and slopes of the Δδ¹⁸Oannual−mean-temperature relationship (m) approaching 0, compared to Δδ¹⁸Oannual−mean (Fig. 2), suggesting that seasonal habitat tracking partly explains the trends shown in Fig. 2. Error envelope as in Fig. 2.

Fig. 5: Relationship between apparent calcification depth (ACD) and temperature. Data are summarised in 2-degree bins and error bars represent standard deviations within each bin. The data
points at the cold temperature end of *G. ruber* (white) are excluded since these are more likely to reflect outliers or advected specimens.

Fig. 6: Offset between seasonality and depth adjusted $\delta^{18}O_{foram}$ and $\delta^{18}O_{calc}$. In most cases both the RMSEs and the $\Delta\delta^{18}O$-temperature slopes are reduced indicating further improvement in the prediction of foraminifera $\delta^{18}O$ and suggesting that the simple empirical parametrisation of habitat variability can be used to correct for habitat tracking. Grey symbols represent $\Delta\delta^{18}O_{\text{error}}$; error envelopes as in Fig. 2.

Fig. 7: Partitioning of the improvement in the prediction of the fossil $\delta^{18}O$ signal into seasonality and depth habitat for both RMSE of $\Delta\delta^{18}O$ and the slope between $\Delta\delta^{18}O$ and MAT$_{\text{0-50m}}$. Colours denote species and the size of each dot is proportional to the total improvement achieved. The open circles illustrate the partitioning for *T. sacculifer* using the paleotemperature equation of Mulitza et al. (2003).

Fig. 8: Assessing the effect of the use of a different paleotemperature equation. The panels show the same as Fig. 2, 4, 5 and 6 respectively, but for *T. sacculifer* and using the equation of Mulitza et al. (2003). Note that the basic patterns indicative of habitat tracking remain, but that the general calcification depth appears greater, also at lower temperatures.

Fig. 9: Effect of habitat tracking: reduced magnitude of deglacial temperature change estimated from Mg/Ca of *G. ruber* (pink) (Elderfield and Ganssen, 2000) compared to faunal assemblage based seasonal temperature estimates (Chapman et al., 1996) in the subtropical North Atlantic. The predicted *G. ruber* (pink) temperature, which is similar to the Mg/Ca temperature, is based on the relationship identified in Figure 2 and the assemblage-derived temperatures. Values are anomalies with respect to the 0-10,000 years BP average.
References


Barker, S., Cacho, I., Benway, H., and Tachikawa, K.: Planktonic foraminiferal Mg/Ca as a proxy for past oceanic temperatures: a methodological overview and data compilation for the Last Glacial Maximum, Quaternary Science Reviews, 24, 821-834, 2005.


Carstens, J., Hebbeln, D., and Wefer, G.: Distribution of planktic foraminifera at the ice margin in the Arctic (Fram Strait), Marine Micropaleontology, 29, 257-269, 1997.


Ortiz, J. D., Mix, A. C., and Collier, R. W.: Environmental Control of Living Symbiotic and Asymbiotic Foraminifera of the California Current, Paleoceanography, 10, 987–1009, 1995.


Skinner, L. C., and Elderfield, H.: Constraining ecological and biological bias in planktonic foraminiferal Mg/Ca and δ18Occ: A multispecies approach to proxy calibration testing, Paleoceanography, 20, n/a-n/a, 10.1029/2004PA001058, 2005.

Snyder, C. W.: Evolution of global temperature over the past two million years, Nature, advance online publication, 10.1038/nature19798


Fig. 1
Fig. 3

WARM-WATER SPECIES

COLD-WATER SPECIES

temperature

peak time
Fig 4

G. ruber (white)

T. sacculifer

N. incompta

N. pachyderma

M: 0.05
RMSE: 0.67

M: 0.07

M: 0.09
RMSE: 0.53

M: 0.01

T. sacculifer (pink)
Fig 5

G. ruber (white)

G. ruber (pink)

T. sacculifer

N. incompta

N. pachyderma

ACD [m]

−5 0 5 10 15 20 25 30

50 40 30 20 10 0

MMT [℃]

0−50 m
Fig. 6

G. rubra (white)

\[ \text{MAT}_{0-50m} \]

\[ \Delta \theta \]

G. rubra (pink)

\[ \text{MAT}_{0-50m} \]

\[ \Delta \theta \]

T. sextovera

\[ \text{MAT}_{0-50m} \]

\[ \Delta \theta \]

N. incompta

\[ \text{MAT}_{0-50m} \]

\[ \Delta \theta \]

N. pachyderma

\[ \text{MAT}_{0-50m} \]

\[ \Delta \theta \]

RMSE: 0.47

RMSE: 0.34

RMSE: 0.38
Fig. 7
Fig 8

- Annual mean
- Seasonality adjusted
- Apparent calcification depth
- Seasonality and depth adjusted

RMSE: 0.72
m: 0.11 ± 0.04

RMSE: 0.77
m: 0.09 ± 0.04

RMSE: 0.49
m: −0.01 ± 0.04
Fig. 9

Temperature anomaly (°C)

- mean annual temperature
- G. oeder (pink) Mg/Ca temperature
- predicted G. oeder (gray) temperature

Age [ka BP]
WARM-WATER SPECIES

COLD-WATER SPECIES

Fig. 10
Supplementary figures to Jonkers and Kucera: ‘Quantifying the effect of seasonal and vertical habitat tracking on planktonic foraminifera proxies’.

SFig. 1: observed $\delta^{18}$O$_{foram}$ vs predicted $\delta^{18}$O$_{eq}$; error bars on observations are median standard deviation form repeat measurements and include the uncertainty on the salinity-$\delta^{18}$O$_{sw}$ conversion and an estimate of calibration uncertainty on the predictions. The grey line denotes the 1:1 relationship between observations and predictions and the black line the observed relationship, including the 5-95 percentile of the estimated uncertainty in dark grey. Note that the large offsets between the observations and the predictions and that none of the relationships follows the 1:1 line.
SFig. 2: sensitivity test of the seasonality model parametrisation for *G. ruber* (pink). Top panels show on the left the improvement in the prediction of the foraminifera $\delta^{18}$O based on the seasonality model described in section 3.2 of the main text, and on the right the resulting relationship between apparent calcification depth (ACD) and mean annual temperature in the upper water column (MAT). Both graphs are also shown in Fig. 4 and 5 of the manuscript. RMSE and slope values in the right panels pertain to the $\Delta\delta^{18}$O-MAT relationships after correction for seasonality and calcification depth.

We have increased and decreased the seasonality by doubling and halving the slope and intercept of the MAT-flux relationship with respect to the empirical values obtained from (Jonkers and Kučera, 2015). Increasing the seasonality reduces the RMSE and the dependency of $\Delta\delta^{18}$O on MAT in the seasonally weighted $\Delta\delta^{18}$O estimates. It yields estimates of ACD that appear not or positively correlated with MAT and leads to (seasonality and depth weighted $\Delta\delta^{18}$O) RMSE and $\Delta\delta^{18}$O-MAT slopes similar to the observation-based model.

The reverse holds true for a reduction in seasonality, which yields RMSE larger and $\Delta\delta^{18}$O-MAT slopes steeper than when using mean annual values and implies $\Delta\delta^{18}$O-MAT relationships similar to our original seasonality-only $\Delta\delta^{18}$O estimates and RMSE close to the $\Delta\delta^{18}$O based on annual mean values. This suggests that the formulation of seasonality in our model is conservative: weaker seasonality parametrisation leaves much larger residuals and a slope that cannot be accounted for by depth habitat adjustment. However, we note that in the case of *G. ruber* (pink) there exists a parametrisation of flux seasonality that leads to a greater improvement in the $\delta^{18}$O prediction and implies a constant habitat depth adjustment.
Seasonality stronger

- \( \delta^{18}O \) [% PDB] vs. MAT [°C]
  - **Empirical seasonality model**
  - RMSE: 0.27
  - \( m = 0.09 \pm 0.1 \)

- \( \delta^{18}O \) [% PDB] vs. ACD [cm]
  - RMSE: 0.27
  - \( m = 0.03 \pm 0.1 \)

Seasonality weaker

- \( \delta^{18}O \) [% PDB] vs. MAT [°C]
  - **Slope * 0.5**
  - RMSE: 0.64
  - \( m = 0.10 \pm 0.1 \)

- \( \delta^{18}O \) [% PDB] vs. ACD [cm]
  - RMSE: 0.55
  - \( m = 0.09 \pm 0.1 \)

- **Intercept * 2**
  - RMSE: 0.31
  - \( m = 0.03 \pm 0.1 \)

- **Intercept * 0.5**
  - RMSE: 0.27
  - \( m = -0.02 \pm 0.1 \)