RESEARCH ARTICLE

10.1029/2019PA003576

Key Points:

- We develop Bayesian calibration models for planktic $\delta^{18}\text{O}_c$
- Accounting for seasonal abundance and species-specific sensitivities improves inference
- Our models produce realistic SST reconstructions for both recent and “deep-time” $\delta^{18}\text{O}_c$ data

Supporting Information:

- Supporting Information S1
- Data Set S1
- Data Set S2

Correspondence to:

S. B. Malevich,
malevich@email.arizona.edu

Citation:

Malevich, S. B., Vetter, L., & Tierney, J. E. (2019). Global core top calibration of $\delta^{18}\text{O}$ in planktic foraminifera to sea surface temperature. *Paleoceanography and Paleoclimatology*, 34, 1292–1315. <https://doi.org/10.1029/2019PA003576>

Received 29 JAN 2019

Accepted 17 MAY 2019

Accepted article online 29 MAY 2019

Published online 13 AUG 2019

Global Core Top Calibration of $\delta^{18}\text{O}$ in Planktic Foraminifera to Sea Surface Temperature

Steven B. Malevich¹ , Lael Vetter¹ , and Jessica E. Tierney¹

¹Department of Geosciences, The University of Arizona, Tucson, AZ, USA

Abstract The oxygen isotopic composition of planktic foraminiferal calcite ($\delta^{18}\text{O}_c$) is one of the most prevalent proxies used in the paleoceanographic community. The relationship between $\delta^{18}\text{O}_c$, temperature, and seawater oxygen isotopic composition ($\delta^{18}\text{O}_w$) is firmly rooted in thermodynamics, and experimental constraints are commonly used for sea surface temperature (SST) reconstructions. However, in marine sedimentary applications, additional sources of uncertainty emerge, and these uncertainty constraints have not as of yet been included in global calibration models. Here, we compile a global data set of over 2,600 marine sediment core top samples for five planktic species: *Globigerinoides ruber*, *Trilobatus sacculifer*, *Globigerina bulloides*, *Neoglobobulimina incompta*, and *Neoglobobulimina pachyderma*. We developed a suite of Bayesian regression models to calibrate the relationship between $\delta^{18}\text{O}_c$ and SST. Spanning SSTs from 0.0 to 29.5 °C, our annual model with species pooled together has a mean standard error of approximately 0.54‰. Accounting for seasonality and species-specific differences improves model validation, reducing the mean standard error to 0.47‰. Example applications spanning the Late Quaternary show good agreement with independent alkenone-based estimates. Our pooled calibration model may also be used for reconstruction in the deeper geological past, using modern planktic foraminifera as an analog for non-extant species. Our core top-based models provide a robust assessment of uncertainty in the $\delta^{18}\text{O}_c$ paleothermometer that can be used in statistical assessments of interproxy and model-proxy comparisons. The suite of models is publicly available as the Open Source software library bayfox, for Python, R, and MATLAB/Octave.

1. Introduction

The oxygen isotopic composition of foraminiferal shell calcite ($\delta^{18}\text{O}_c$ or δ_c) is the foundational proxy of paleoceanography (Emiliani, 1955). Foraminiferal $\delta^{18}\text{O}_c$ is a function of ambient $\delta^{18}\text{O}_w$ and temperature at the time of calcification, the latter of which determines the magnitude of equilibrium fractionation in line with thermodynamic expectations (Urey, 1947). The relationship of the isotopic difference between $\delta^{18}\text{O}_c$ and $\delta^{18}\text{O}_w$, and temperature is traditionally approximated as a second-order polynomial (e.g., Epstein et al., 1953; Erez & Luz, 1983; Shackleton, 1974)

$$T = \alpha + \beta_1 \cdot (\delta_c - \delta_w) + \beta_2 \cdot (\delta_c - \delta_w)^2$$

At environmental temperatures, the nonlinearity of the temperature sensitivity is slight, leading many to use a simplified linear form (e.g., Bemis et al., 1998)

$$T = \alpha + \beta \cdot (\delta_c - \delta_w)$$

Rearranging to the etiological form of the equation to express the dependence of $\delta^{18}\text{O}_c$ on temperature yields

$$\delta_c = \alpha + \beta \cdot T + (\delta_w - 0.27)$$

where -0.27 approximates the difference between the Vienna Pee Dee Belemnite (for $\delta^{18}\text{O}_c$) and Vienna Standard Mean Ocean Water (VSMOW; for $\delta^{18}\text{O}_w$) scales (Hut, 1987).

Synthetic calcite studies indicate that under equilibrium conditions, $\delta^{18}\text{O}_c$ sensitivity to temperature is approximately $-0.19\text{‰} \cdot ^\circ\text{C}^{-1}$ near 30 °C and approximately $-0.25\text{‰} \cdot ^\circ\text{C}^{-1}$ near 0 °C (Kim & O'Neil, 1997). Calibration studies using foraminifera from lab cultures and marine tows yield comparable temperature sensitivities (-0.20 to $-0.28\text{‰} \cdot ^\circ\text{C}^{-1}$; e.g., Bemis et al., 1998; Bouvier-Soumagnac & Duplessy, 1985; Mulitza et al., 2003a; Shackleton, 1974) although in some cases the foraminiferal calibrations are offset (0.2–0.8‰)

from inorganic calibrations (e.g., Caron et al., 1990; Duplessy & Blanc, 1981; Waelbroeck et al., 2005). The error in calibrations from culture or plankton tows is greater than the error of inorganic calibrations; for example, lab culture calibrations have standard errors of 0.15‰ (e.g., Bemis et al., 1998) and plankton tow calibrations of 0.26‰ (e.g., Mulitza et al., 2003a).

While there is general agreement of $\delta^{18}\text{O}_c$ sensitivity between synthetic calcite experiments and foraminiferal-based investigations, application of $\delta^{18}\text{O}_c$ paleothermometry to marine sediment records poses additional challenges and unavoidable sources of uncertainty. For one, $\delta^{18}\text{O}_w$ is not precisely known for past oceanographic conditions and must be estimated, potentially introducing a large source of uncertainty under both present and past oceanic conditions. To a leading order, $\delta^{18}\text{O}_w$ reflects the precipitation-evaporation balance over the open ocean, but it is also modified by local and regional processes such as ice formation, glacial meltwater, seasonal freshwater runoff, water mass advection, and mixing (Craig & Gordon, 1965). The climatological processes influencing $\delta^{18}\text{O}_w$, coupled with the scarcity of measurements in many regions of the modern ocean, can lead to large uncertainties in $\delta^{18}\text{O}_w$ for certain locations such as the high latitudes (LeGrande & Schmidt, 2006). Moving back through time, $\delta^{18}\text{O}_w$ distributions must either be estimated through a priori assumptions about oceanographic setting or predicted by isotope-enabled climate models. Alternatively, research questions about $\delta^{18}\text{O}_w$ can be addressed by reconstructing temperature with another independent proxy and then isolating $\delta^{18}\text{O}_w$ from $\delta^{18}\text{O}_c$.

In foraminiferal calcite, the uncertainty of shell $\delta^{18}\text{O}_c$ temperature calibrations is influenced by biological processes, such as photosynthesis in algal symbionts (e.g., Duplessy et al., 1970; Ravelo & Fairbanks, 1992; Spero & Lea, 1993; Spero et al., 1997) and biases in the formation of gametogenic and ontogenetic calcite (e.g., Hamilton et al., 2008; Spero & Lea, 1996; Williams et al., 1979). In addition, each species exhibits a distinct seasonality and depth habitat in the water column (e.g., Fairbanks & Wiebe, 1980; Fairbanks et al., 1982; Kohfeld et al., 1996; Sautter & Thunell, 1989, 1991; Žarić et al., 2005), and even within the morphospecies commonly used for classification of fossil foraminifera, there may be additional differences in life cycle and habitat preferences due to the genotypic diversity (Aurachs et al., 2011; Darling & Wade, 2008; Kucera & Darling, 2002). While such ecological relationships can be leveraged for season- and depth-specific climate reconstructions (e.g., Mulitza et al., 2003b; Spero et al., 2003; Williams et al., 1981), these relationships can change through time in response to environmental or biological variations, complicating paleoenvironmental interpretations (e.g., Mulitza et al., 1998). Changes in pH/carbonate ion concentration [CO_3^{2-}] during calcification also influence $\delta^{18}\text{O}_c$ (Bijma et al., 1999; Spero et al., 1997; Zeebe, 1999). Finally, the sedimentary environment influences the fidelity of $\delta^{18}\text{O}_c$ as it is preserved downcore. Shells deposited in bottom waters undersaturated in [CO_3^{2-}] may partly dissolve and recrystallize, a process that alters the original isotopic signature via exchange with pore water $\delta^{18}\text{O}$ (e.g., Schrag et al., 1995). Bioturbation can be an especially strong source of core top variability in areas with low sedimentation rates, where glacial age sediments may become mixed with relatively modern sediments in a core top sample (Waelbroeck et al., 2005).

Many paleoceanographic applications use laboratory calibrations to transform $\delta^{18}\text{O}_c$ data to sea surface temperatures (SST), but these calibrations do not capture the range of biological, chemical, and sedimentological uncertainties enumerated above. It is important to capture these uncertainties in order to realistically estimate paleo-SSTs from the marine sediment archive and critical for multiproxy or climate model-proxy comparisons. In this study, we develop a calibration for the $\delta^{18}\text{O}_c$ -SST relationship in planktic foraminifera using core top data and Bayesian regression. We use a Bayesian approach which explicitly models the uncertainty in calibration model parameters and then propagates this uncertainty into inferred SSTs, facilitating probabilistic estimates of past climate (Tierney & Tingley, 2014, 2018). In developing Bayesian regression for $\delta^{18}\text{O}_c$, we pay particular attention to species-specific differences and seasonal abundance. Depth habitat is also an important factor, but in this work we focus specifically on near-surface (mixed layer) dwelling species that are frequently used to reconstruct SST, for which slight differences in depth habitat are less likely to appreciably affect regression models. Calibration of deep-dwelling species is left to future efforts.

In what follows, we develop mean annual and seasonal calibration models for five planktic species commonly used in paleoceanography: *Globigerinoides ruber*, *Trilobatus sacculifer*, *Globigerina bulloides*, *Neoglobobulimina incompta*, and *Neoglobobulimina pachyderma*. We also develop a model that pools these species together for application to extinct species of planktic foraminifera commonly used for Cenozoic paleoceanographic reconstructions (e.g., Zachos et al., 1994). These models are freely available to researchers as a software library called `bayfox`. We give examples of how our calibrations can be applied for

Figure 1. Core top (a) and marine sediment trap (b) sites for each planktic species used in the calibration. We used the sediment trap data sets compiled by Žarić et al. (2005) to estimate seasonal sea surface temperatures for the seasonal calibration models.

paleotemperature reconstructions over the Late Quaternary as well as in deeper geologic time, comparing our calibration and its uncertainty with established Bayesian alkenone $U_{37}^{K'}$ and TEX_{86} reconstructions.

2. Methods and Data Selection

2.1. SST and Seawater $\delta^{18}O_w$

Our Bayesian models use modern SSTs and $\delta^{18}O_w$ as predictors. For SSTs, we used both monthly and annual fields from the World Ocean Atlas 2013 version 2 (Boyer et al., 2013). For $\delta^{18}O_w$, we used the top layer of the estimated annual fields from LeGrande and Schmidt (2006). Both of these products have $1^\circ \times 1^\circ$ spatial resolution. The LeGrande and Schmidt (2006) $\delta^{18}O_w$ field is based on $\delta^{18}O_w$ observations from the last half-century, and in areas with sparse isotope sampling coverage, it uses regional $\delta^{18}O_w$ -salinity relationships to estimate isotope values. The $\delta^{18}O_w$ field does not include uncertainty estimates for grid points, though LeGrande and Schmidt (2006) note that annual average values in regions near or under sea ice may be more uncertain due to a limited number of observations and large seasonal fluctuations in runoff and precipitation that induce high variance in $\delta^{18}O_w$.

2.2. Core Top Planktic Foraminiferal $\delta^{18}O_c$

We compiled planktic $\delta^{18}O_c$ sediment core records for five foraminiferal species. *G. ruber*, *G. bulloides*, *N. incompta*, and *N. pachyderma*. *G. ruber* (white), and *G. ruber* (pink) have some regional differences in seasonality and relative abundance (e.g., Bé, 1960; Williams et al., 1981). However, we opted to evaluate *G. ruber* white and pink together because our preliminary *G. ruber* (pink) calibration was strongly influenced by sites off the

northwest coast of Africa, leading to model parameters that we believe may reflect sampling and statistical artifacts more than differences in *G. ruber* (pink) calcification.

Core top and Late Holocene records were gathered from the Multiproxy Approach for the Reconstruction of the Glacial Ocean data set (Waelbroeck et al., 2005) which extends the collection of Schmidt and Mulitza (2002). We supplemented this collection with additional sources (Arbuszewski et al., 2010, 2013; Boussetta et al., 2012; Brown & Elderfield, 1996; Cléroux et al., 2008; Dahl & Oppo, 2006; Dekens et al., 2002; Dyez et al., 2014; Elderfield & Ganssen, 2000; Fallet et al., 2012; Farmer, 2005; Ganssen & Kroon, 2000; Garidel-Thoron et al., 2007; Gebregiorgis et al., 2016; Gibbons et al., 2014; Johnstone et al., 2011; Kozdon et al., 2009; Lea et al., 2006; Leduc et al., 2007; Linsley et al., 2010; Mashiotto et al., 1999; Mathien-Blard & Bassinot, 2009; Meland et al., 2006; Moffa-Sánchez et al., 2014; Mohtadi et al., 2010, 2011; Nürnberg et al., 2008; Oppo et al., 2009; Oppo & Sun, 2005; Pahnke et al., 2003; Palmer & Pearson, 2003; Parker et al., 2016; Regenberg et al., 2009; Richey et al., 2007; Riethdorf et al., 2013; Riveiros et al., 2016; Romahn et al., 2014; Rosenthal et al., 2003; Rustic et al., 2015; Sabbatini et al., 2011; Saraswat et al., 2013; Schmidt et al., 2012a, 2004, 2012b; Steinke et al., 2005, 2008; Steph et al., 2009; Stott et al., 2007; Sun et al., 2005; Thornalley et al., 2011; Tierney et al., 2016; Visser et al., 2003; Weldeab et al., 2005, 2006, 2007, 2014; Werner et al., 2013; Xu et al., 2010).

We excluded records from sites with annual SST $\leq 0^\circ\text{C}$ to reduce complications from local sea ice formation and poor $\delta^{18}O_w$ estimates. Waelbroeck et al. (2005) show that age and sedimentation rate filtering can help reduce uncertainty stemming from the ambiguous “modern” age constraints. However, we consider these to be important sources of uncertainty to include in our calibration, so we did not filter core top sites by age or sedimentation rates. This also means that we are calibrating ambiguously modern core top samples against SST and $\delta^{18}O_w$ fields influenced by anthropogenic climate change. This is an issue that affects all core top calibrations.

Our compilation consists of 2,636 observations (Figure 1a) with 1,002 for *G. ruber*, 635 for *G. bulloides*, 442 for *T. sacculifer*, 425 for *N. pachyderma*, and 132 for *N. incompta*. We then gridded the core top data to reduce the impact of spatial clustering by averaging samples for each species to the nearest $1^\circ \times 1^\circ$ grid point of our SST and $\delta^{18}O_w$ fields. After gridding, there were a total of 1,386 grid points, with 489 for *G. ruber*, 291 for *G.*

Table 1
Modern Annual SST Minimum, Maximum, Mean, and Standard Deviation (σ) for Each Core Top Species Group

Group	<i>n</i>	Modern annual SST (°C)			
		Min	Max	Mean	σ
<i>Globigerinoides ruber</i>	489	10.9	29.6	24.9	3.8
<i>Trilobatus sacculifer</i>	243	10.6	29.6	24.5	4.0
<i>Globigerina bulloides</i>	291	1.8	29.6	13.6	6.9
<i>Neogloboquadrina incompta</i>	90	2.6	19.6	11.5	4.7
<i>Neogloboquadrina pachyderma</i>	273	0.1	21.4	6.1	4.3
Pooled	1,386	0.1	29.6	17.8	9.0

Note. *n* values are sample size for gridded core top data.

bulloides, 273 for *N. pachyderma*, 243 for *T. sacculifer*, and 90 for *N. incompta*. References to core top data hereafter refer to the gridded core top data unless noted otherwise.

The core top data cover a wide range of modern SST values and reflect the general thermal preferences of each species (Table 1); for instance, *G. ruber* prefers relatively warmer waters, while *G. bulloides* is abundant across a wide range of temperatures, and *N. incompta* and *N. pachyderma* prefer cooler waters.

2.3. Estimation of Foraminiferal Seasonal Abundance

The abundance of individual planktic foraminiferal species varies seasonally in response to changes in temperature and nutrients, which affect food availability (e.g., Williams et al., 1979, 1981). This motivates the development of a seasonally adjusted calibration model, wherein foraminiferal seasonal expression is modeled as function of their preferred temperature range. To build such a model, we use sediment trap data compiled by Žarić et al. (2005) to identify temperature ranges that correspond to peak abundance for each foraminiferal species. The Žarić et al. (2005) data set pairs total foraminifera shell flux and local SSTs from 75 sites (Figure 1b). The data set contains a total of 5,548 observations with 1,807 for *G. ruber*, 1,034 for *T. sacculifer*, 1,255 for *G. bulloides*, 910 for *N. incompta*, and 542 for *N. pachyderma*.

To adjust for nonnormally distributed shell flux data, we applied a Box-Cox power transformation (Box & Cox, 1964) for comparison with SST (Figure 2). A Kernel Density Estimate was fit to the observations to estimate the SST interval that corresponds with the highest 10% (most abundant) flux observations, which is taken to represent their “ideal” thermal niche, similar to the approach in Žarić et al. (2005).

We then used these SST ranges to estimate the most likely seasons of peak shell flux for each core top foraminiferal observation, by averaging SSTs from all months at the location of the observation that fell within the SST range. If, for a given observation, no monthly temperatures fell within the peak abundance range, we took the average of the three monthly SSTs closest to the range. If less than 3 months fell within the range, we included the next closest months so as to guarantee a seasonal average of at least three months. The resulting seasonality estimates show that at core sites in the tropics, foraminiferal species abundances are typically annual or near annual, without a strong seasonal signal (Figure 3). For the warm-water species *G. ruber* and *T. sacculifer*, a stronger seasonal signal (typically Summer-Fall) appears outside of the tropics. Cold-water species are generally predicted to be annual within their expected extratropical ranges, although seasonal expressions do occur for locations with particularly broad annual temperature ranges (e.g., Figure S4 in the supporting information). Monthly maps of predicted niche SST ranges for each species are available in the supporting information (Figures S1–S5).

We recognize that this method of identifying seasonal signals in foraminiferal abundance is a simplification that does not consider factors such as light and nutrient availability. We chose this approach because it is easily replicated and used with a large global sample data set. This approach can be adapted more broadly for forward modeling experiments where seasonality or monthly SSTs change relative to modern conditions.

2.4. Bayesian Calibration Models

We designed and fit four linear Bayesian regression models with Markov chain Monte Carlo (MCMC) sampling (for review see Gelman, 2014; Kruschke, 2015; McElreath, 2016). With a Bayesian approach, we can explicitly estimate the uncertainty in calibration model parameters and produce a full prediction posterior

Figure 2. Peak abundance and SST ranges for each species. Points are observed foraminifera shell flux from the sediment trap network, plotted against SST measured at the sediment trap site (Žarić et al., 2005). Contours are the Kernel Density Estimate of observations. Horizontal lines are the highest 10% of observed shell flux. Vertical lines and annotations are the 90% range of SSTs within the highest 10% shell flux.

distribution of the predictant ($\delta^{18}\text{O}_c$ or SST), rather than a single-point estimate. We designed four models to uniquely consider seasonal and species-specific adjustments to calibration. We compared the performance of these four models with cross-validation statistics. These statistics were used to objectively assess whether considering species-specific variability and seasonality resulted in model improvement.

In the first model, we pooled all five species for a single calibration against annual SSTs. This model design is

$$\begin{aligned}\delta_c &\sim \mathcal{N}(\mu, \tau^2), \\ \mu &= \alpha + \beta \cdot T + (\delta_w - 0.27),\end{aligned}\tag{1}$$

where α , β , and τ^2 are the regression intercept, slope, and prediction variance, respectively. The unknown parameters, α , β , and τ were assigned weakly informed prior distributions. While we have discussed uncertainty in $\delta^{18}\text{O}_w$ fields, we did not explicitly model uncertainty in $\delta^{18}\text{O}_w$ because this is not constrained in the LeGrande and Schmidt (2006) data. We used a linear calibration model because the nonlinearity of temperature sensitivity is relatively small within the range of temperatures at which foraminifera calcify (e.g., Bemis et al., 1998), and this nonlinearity is apparently represented piecewise in species-specific calibrations (e.g., Mulitza et al., 2003a; Waelbroeck et al., 2005). To test the validity of our linear assumption, we developed a simple quadratic Bayesian model comparable to the temperature relationships for inorganic calcite from O'Neil et al. (1969) and Kim & O'Neil (1997; not shown). Cross validation showed that the quadratic model yielded no significant improvement over our linear regressions, suggesting that a linear assumption adequately explains the data.

Figure 3. Estimated seasonal preferences for core top foraminifera based on sediment flux and monthly sea surface temperature, for the tropics, northern extratropics (latitude $> 23^\circ$), and southern extratropics (latitude $< -23^\circ$). Numbers around the circle represent months. Line widths correspond to how frequently months are paired together for seasonal sea surface temperature estimates in each species. Stronger lines indicate common month pairings. n is the number of gridded core top samples available for the given region.

Our second model was also calibrated against annual SSTs but uses a hierarchical design to allow model parameters to vary for each of the five species. The hierarchical design imposes a “global” commonality for calcification in planktic foraminifera (supported by thermodynamics) but allows the calibration parameters for each species to vary from this global constraint—reflecting, for example, differences in depth of calcification, vital effects, or temperature sensitivity (e.g., Bemis et al., 1998; Mulitza et al., 2003a; Spero et al., 2003), as well as influence from spatial sampling bias. This model is

$$\delta_c \sim \mathcal{N}(\mu, \tau_i^2),$$

$$\mu = \alpha_i + \beta_i \cdot T + \delta_w - 0.27$$
(2)

where α_i , β_i , and τ_i are calibration parameters as described above but for given species, i . The hierarchical relationship is in the priors for species groups

$$\alpha_i \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2),$$

$$\beta_i \sim \mathcal{N}(\mu_\beta, \sigma_\beta^2),$$

$$\tau_i \sim \Gamma\left(\frac{\sigma_m^2}{\sigma_d^2}, \frac{\sigma_m^2}{\sigma_d^2}\right),$$
(3)

which have unknown parameters depending on hyperparameters shared across all species groups. The hyperparameters are described in the appendix.

Our third and fourth models use seasonal SSTs in the place of annual SSTs, while retaining the same pooled and hierarchical model designs described above. The seasonal SSTs are based on seasonal peak abundance estimated from a network of marine sediment traps, as described in section 2.3. These seasonal models use annual $\delta^{18}\text{O}_w$ estimates, as monthly fields are not available from LeGrande and Schmidt (2006).

All calibration models were cross validated using Pareto-Smoothed Importance Sampling leave-one-out cross validation (LOOCV; Vehtari et al., 2017). This statistic measures model predictive performance by approximating a LOOCV—estimating each core top sample as though it were left out of the calibration for validation purposes. A relatively low score is “better,” indicating improved performance.

Our four models are “forward models”—that is, they predict $\delta^{18}\text{O}_c$ given SST and $\delta^{18}\text{O}_w$ —but they can also be inverted, via Bayesian inference, to predict SST. To predict SST for a given $\delta^{18}\text{O}_c$ and $\delta^{18}\text{O}_w$, parameters are drawn from the full conditional posteriors of the calibration models for the likelihood and then combined with a prior distribution of SSTs to yield a posterior. The SSTs in Table 1 are a suggested starting point to develop a prior distribution, though users can specify values to fit different environmental settings.

Additional description of priors, hyperparameters, model inversion, and MCMC sampling is given in the appendix. We implemented the Bayesian models for this analysis with the `pymc3` library (Salvatier et al., 2016) on an Open Source Python software stack (Hunter, 2007; Hoyer & Hamman, 2017; McKinney, 2010; Met Office, 2010; Oliphant, 2015). Code implementing our analysis is available online (https://github.com/brews/d18oc_sst). The calibration models are available for broader use with the Open Source `bayfox` software library, described in section 4.7.

3. Results and Discussion

The four Bayesian calibration models differ in whether they account for species-specific differences (“pooled” vs. “hierarchical”) or the seasonality of foraminiferal abundance (“annual” vs. “seasonal”). Our

Figure 4. Scatter plots showing core top fractionation ($\delta^{18}\text{O}_c - \delta^{18}\text{O}_w$) and annual SST. (a) Pooled foraminifera species. (b) Individual species. Dotted line is the inorganic calcite temperature relationship of Kim and O'Neil (1997). Annotations are Pearson correlation (r) between core top fractionation and annual mean SST. All Pearson correlations are statistically significant ($p \ll 0.01$).

“pooled annual” model combines all five species together and calibrates core top $\delta^{18}\text{O}_c$ to annual mean SSTs. The “hierarchical annual” model also calibrates to annual mean SSTs but allows the calibration parameters to vary for each species. The “pooled seasonal” model and “hierarchical seasonal” model use the same pooled and hierarchical designs but are calibrated with our seasonal SST estimates—based on sediment trap fluxes (see section 2.3)—instead of annual SSTs.

3.1. The Core Top Relationship With SST

Within the core top data used for calibration, the relationship between annual SST and core top isotopic fractionation ($\delta^{18}\text{O}_c - \delta^{18}\text{O}_w$) is strongly negative for both the pooled data set ($r = -0.97$; $p \ll 0.01$) and individual species data sets. The relationship sits close to previous calibrations based on inorganic calcite precipitation, live cultures, and plankton tows (Bemis et al., 1998; Kim & O'Neil, 1997; Mulitza et al., 2003a; O'Neil et al., 1969; Figure 4). Despite this, the core top data have notable spread and deviations relative to previous calibrations, which can be understood as the expression of uncertainty related to sedimentological factors (to glacial sediment mixing from bioturbation, loss of core top material when coring, and low sedimentation rates), biological factors (seasonal abundance and vital effects), and uncertainties in $\delta^{18}\text{O}_w$.

All four Bayesian calibration models reasonably replicate core top data spread when we predict core top fractionation (Figure 5). Calibration model spread is measured as the mean sample standard deviation of the posteriors ($\bar{\sigma}$). The pooled models have larger spread (Figure 5; $\bar{\sigma} = 0.54\text{‰}$ for the pooled annual

Figure 5. Core top fractionation and SST compared with predictions from our (a) pooled annual model, (b) pooled seasonal model, (c) hierarchical annual model, and (d) hierarchical seasonal model. Our calibrations are compared with two inorganic calcite relationships (green lines), three culture calibrations (orange dots), and four species-specific calibrations from plankton tows (red dashes), over their respective calibration temperatures. The blue shading is the 95% credible interval (CI) of our calibration model prediction for the core top data (black circles). Annotated $\bar{\sigma}$ is mean sample standard deviation, or mean standard error, of the core top prediction posteriors. The inorganic relationships are from O'Neil et al. (1969) and Kim and O'Neil (1997). Culture calibrations are from Bemis et al. (1998), for *Orbulina universa* (high light and low light) and *Globigerina bulloides* with 12-chambered shells. The plankton tow calibrations are from Mulitza et al. (2003a) for *Globigerina bulloides*, *Globigerinoides ruber*, *Trilobatus sacculifer*, and *Neogloboquadrina pachyderma*.

model, $\bar{\sigma} = 0.51\text{‰}$ for the pooled seasonal model, 0.47‰ for the hierarchical annual model, and 0.49‰ for the hierarchical seasonal model). These $\bar{\sigma}$ values correspond to uncertainties in SST of approximately $2.5\text{--}2.8\text{ °C}$ (with $\beta = -0.19$) or $1.9\text{--}2.1\text{ °C}$ (with $\beta = -0.25$), depending on the regression slope (β) of the calibration model. For comparison, the *Orbulina universa* high-light and low-light culture calibrations of Bemis et al. (1998) have standard errors between 0.10‰ and 0.15‰ , which represents 0.5 to 0.7 °C using their calibration β of -0.21 . The wider standard errors of species-specific plankton tow-based calibration of Mulitza et al. (2003a) range from 0.21‰ to 0.32‰ (0.9 and 1.2 °C at their calibration $\beta = -0.23$ and $\beta = -0.21$, respectively) likely resulting from the broader range of depth habitats and thermocline structure captured during plankton tow collection. The prediction spread is larger in our calibration models because our compilation of global core top data contains a wider range of uncertainty (e.g., calcification depth, postdepositional bioturbation, and differences in depositional age).

The posterior distributions of the slope parameter (the sensitivity between seawater temperature and $\delta^{18}\text{O}_c$; β) for the pooled annual and pooled seasonal models likewise fall near the bounds of thermodynamic expectation (Figure 6), with nearly identical mean values of -0.23 . In the hierarchical models, β is allowed to deviate between species (β_i) resulting in greater variability (Figure 6). In four of the five species, the median value of β_i in the hierarchical annual model is lower than thermodynamic expectation (Figure 6). This tendency for lower sensitivity is also apparent in the posterior for the β hyperparameter—the slope parameter shared across all species—which is -0.18 ± 0.05 (95% CI) in the hierarchical annual model. In contrast, the hierarchical seasonal β hyperparameter more closely matches that of equilibrium fractionation (-0.21 ± 0.04 , 95% CI). However, species-specific offsets in sensitivity do persist; in particular, β_i is lower for *N. incompta* and *T. sacculifer*. These differences could simply be related to the narrower range of temperatures used for most of the species-specific calibrations (Bemis et al., 2002); for instance, β_i values in the seasonal model are close to the inorganic sensitivity, and *G. bulloides*, which has a wider temperature range than the other species, has a posterior β_i that is very similar to the pooled calibration values (Figure 6).

3.2. Cross-Validation Statistics and Model Comparison

Model cross-validation statistics (LOOCV; Figure 7) suggest that predictive performance of the model improves when we account for seasonality of foraminiferal abundance and species-specific differences. The pooled seasonal model (LOOCV = $2,063 \pm 65$) outperforms the pooled annual model (LOOCV = $2,248 \pm 74$), and predictive performance further improves when we incorporate species-specific differences in the hierarchical annual model (LOOCV = $1,783 \pm 68$). The hierarchical seasonal model shows a decrease in validation performance (LOOCV = $1,952 \pm 66$) compared to the hierarchical annual model, but this does not necessarily mean that seasonality is unimportant. The lack of improvement partly reflects the fact that under the hierarchical model design, seasonal abundance differences can be subsumed under species-specific differences. This said, patterns in the residuals suggest that data from the Mediterranean region may explain the increased LOOCV in the hierarchical seasonal model (see discussion below).

3.3. Model Residual Trends and Spatial Patterns

Taken as a whole, the residuals of the pooled annual calibration model are well behaved (Figure 8a). However, if we use the pooled annual model to predict $\delta^{18}\text{O}_c$ for individual species, strong trends emerge. At

Figure 6. Posterior distributions of the calibration slope parameter (β_i) for each species in the hierarchical annual model (blue) and hierarchical seasonal model (orange). The gray dotted vertical lines show the range of slopes from inorganic calcite temperature relationships (Kim & O'Neil, 1997). The green vertical line denotes β from the pooled annual and seasonal calibrations (both have a small standard deviation of ~ 0.0015).

G. bulloides in the Benguela Current, Peru Current, and other upwelling sites along the Southern Ocean boundary may reflect this habitat preference; these positive residuals persist even when accounting for seasonality and species-specific sensitivity (Figures 8–11). In the southwestern Atlantic Ocean, positive residuals (model predicts warmer temperatures) may reflect a response of *G. bulloides* to increased turbidity from seasonal river discharge from the Rio de la Plata estuary and/or variations in nutrient delivery related to wind direction and interactions with the Brazil Current (Piola et al., 2005). These factors might skew *G. bulloides* abundance to cooler seasons and/or a deeper depth habitat.

We also observe persistent positive mean $\delta^{18}\text{O}_c$ residual values for *G. ruber* in the Mediterranean region in both the pooled seasonal and hierarchical seasonal models (Figures 9b and 11b). This group of residuals from

the Mediterranean substantially contributes to the reduced performance of the seasonal hierarchical model ($\text{LOOCV} = 1,952 \pm 66$) over the annual hierarchical ($\text{LOOCV} = 1,783 \pm 68$) model. With Mediterranean core top data excluded, the two hierarchical models perform similarly ($1,716 \pm 66$ and $1,691 \pm 67$, respectively).

The Mediterranean bias in the seasonal models suggests that either our estimation of seasonality for this region is problematic or that there are other factors influencing foraminiferal calcification that are not accounted for in our basic model setup. Our seasonality estimates predict that *G. ruber* abundance and shell flux should be skewed toward boreal summer and fall, which agrees well with plankton tow surveys in the Mediterranean (e.g., Pujol & Grazzini, 1995). We therefore favor the second explanation; that is, that there is another mechanism that might explain the Mediterranean model residuals. We can rule out issues related to bioturbation because the majority of the core top samples for these sites have age control with dates placing them in the Holocene or Late

Figure 7. Comparison of model predictive performance based on cross validation (LOOCV). Lower LOOCV values indicate relatively better predictive performance, that is, lower expected deviance for predictions on new data. Circles are LOOCV mean; whiskers show ± 1 standard error.

Figure 8. Pooled annual calibration model core top residuals (observed - predicted mean) as scatter plots for (a) species pooled together, (b) species separated, and (c) maps of mean residuals for each core top grid point. Whiskers bars in (a) and (b) are ± 1 standard deviation.

Holocene (Sabbatini et al., 2011). Postdepositional overgrowth from the supersaturation of bottom water calcite is known to complicate the use of the foraminiferal Mg/Ca proxy in the Mediterranean (e.g., Kontakiotis et al., 2011) and could feasibly also account for the positive $\delta^{18}O_c$ residuals. However, if this were the case, we would expect to see strong positive residuals in other Mediterranean species and not just *G. ruber*. Another explanation for this apparent bias in the Mediterranean relates to habitat depth. During the summer-fall season, *G. ruber* expands its depth habitat within the water column and therefore may be recording cooler temperatures. This is supported by substantial abundance of *G. ruber* observed down to 100 m in the summer-fall season in the Mediterranean (Pujol & Grazzini, 1995). Yet another source of uncertainty is the seasonal changes in $\delta^{18}O_w$ and salinity (of 1PSU) observed in the Mediterranean (e.g., MEDAR Group, 2002), which are not captured in the LeGrande and Schmidt (2006) $\delta^{18}O_w$ data.

Figure 9. Pooled seasonal calibration model residuals (observed-predicted mean) as (a) scatter plots and (b) maps of mean residuals for each core top grid point. Whisker bars in (a) are ± 1 standard deviation.

We also observe a recurring pattern of positive $\delta^{18}\text{O}_c$ residuals (warmer predicted temperatures) near major oceanic frontal boundaries. This occurs for *G. bulloides* along the Southern Ocean front, for *N. incompta* near the confluence of the Agulhas and Benguela Currents, and for *N. pachyderma* near the boundary between the Labrador Current and Gulf Stream in the North Atlantic (Figure 11b). Such residual patterns could reflect advection of foraminifera shells which calcified along the colder side of these boundaries (Martínez-Méndez et al., 2010). However, sediment trap studies suggest that the rapid settling rate of foraminifera shells is unlikely to result in a strong advection bias (e.g., King & Howard, 2005), and differences in habitat related to water masses may be strongly dependent on species (e.g., Dyez et al., 2014). An alternate explanation is that sharp hydrological gradients in these frontal regions may bias the estimation of $\delta^{18}\text{O}_w$ from limited measurements (LeGrande & Schmidt, 2006; Waelbroeck et al., 2005). With a relatively coarse $1^\circ \times 1^\circ$ fields, our estimates of SST and especially $\delta^{18}\text{O}_w$ in boundary regions between water masses are likely poorly characterized.

A related issue is that assumptions about local $\delta^{18}\text{O}_w$ -salinity relationships may introduce a region-specific uncertainty in $\delta^{18}\text{O}_w$, which is translated into highly coherent spatial patterns in $\delta^{18}\text{O}_c$ residuals. For example, *G. bulloides*, *G. ruber*, and *T. sacculifer* all have negative $\delta^{18}\text{O}_c$ mean residuals (generally colder predicted temperatures or higher predicted $\delta^{18}\text{O}_c$ than observed) in the eastern boundary upwelling regions of the Atlantic Ocean (Figure 11b). These trends are present in both the annual and seasonal versions of the hierarchical model. It is unlikely that these negative residuals represent a biological bias. Sediment trap studies off the coast of Saharan Africa demonstrate that *G. ruber* is more abundant during the winter months (Abrantes et al., 2002), yet the negative model residuals in $\delta^{18}\text{O}_c$ suggest that *G. ruber* is calcifying in warmer SSTs

Figure 10. Hierarchical annual calibration model residuals (observed-predicted mean) as (a) scatter plots and (b) maps of mean residuals for each core top grid point. Whiskers bars in (a) are ± 1 standard deviation.

than predicted by the model. Here, a seasonal bias in the abundance of warm-water species is clearly not the cause of the geographic bias in model residuals, and a more likely candidate can be found in the predictor variables.

3.4. Summary of Calibration Model Performance

All four of our calibration models replicate the center and spread in the core top record (with model $\bar{\sigma}$ from 0.54‰ to 0.47‰) while reasonably reproducing the temperature- $\delta^{18}\text{O}_c$ relationship given in established equilibrium calcite relationships (e.g., Kim & O'Neil, 1997; O'Neil et al., 1969), as well as relationships determined for foraminiferal calcite from live culture and plankton tows (e.g., Bemis et al., 1998; Mulitza et al., 2003b). There are species-specific trends in the residuals for both pooled models; these are eliminated in the hierarchical models as the latter allow the regression parameters to vary by species. All models show some residual bias along oceanic fronts and upwelling zones, where dynamic hydrography may introduce complex seasonal patterns in abundance and habitat depth. Overall, model performance improves when accounting for foraminiferal seasonality and species-specific variability, with the exception of the hierarchical seasonal model versus the hierarchical annual model. As discussed, the reduced performance in the hierarchical seasonal model is related to the unusual behavior of *G. ruber* data in the Mediterranean, which may reflect depth habitat migration. Even though the hierarchical seasonal model objectively performs worse by the LOOCV metric, it produces posterior distributions of temperature sensitivity (β_i) in closer agreement with thermo-

Figure 11. Hierarchical seasonal calibration model residuals (observed-predicted mean) as (a) scatter plots and (b) maps of mean residual for each core top grid point. Whiskers bars in (a) are ± 1 standard deviation.

dynamic expectations (Figure 6). As we discuss below, the pooled annual model is a more appropriate choice for applications to extinct planktic species in the geologic past.

4. Examples and Applications

A key benefit of our Bayesian core top calibrations is that the models can propagate uncertainty from calibration into predictions about past climate conditions. We demonstrate this using several downcore examples from different oceanographic settings. For each example, we apply our calibration models to foraminiferal $\delta^{18}\text{O}_c$ and compare our results with SSTs inferred from independent organic geochemistry records from the same core sites (either $\text{U}_{37}^{K'}$ or TEX_{86} data) or sediment trap measurements. Inference of SST from $\delta^{18}\text{O}_c$ requires priors on SST means and standard deviations. We take these from Table 1, multiplying the standard deviation by 2. The $\delta^{18}\text{O}_c$ data are ice volume corrected before calibration to remove the changes in global $\delta^{18}\text{O}_w$ associated with ice sheets (see appendix). In all cases we use modern, annual $\delta^{18}\text{O}_w$ estimates from LeGrande and Schmidt (2006), unless noted otherwise.

4.1. *G. bulloides* in the Mediterranean: A Cosmopolitan Foraminifera Species

G. bulloides has a wide temperature tolerance (2.8–29.6 °C, Bè, 1977; Hemleben et al., 1989, Figure 2), so the seasonal and annual models perform effectively the same for this particular species; this is evident in the similar posterior distributions of β and β_i (Figure 6). To investigate model performance on *G. bulloides*, we

Figure 12. SST reconstructions from sediment core records using our annual (a–c) and seasonal (d) hierarchical SST calibrations, compared with Bayesian alkenone $U_{37}^{K'}$ reconstructions using the BAYSPLINE calibration (Tierney & Tingley, 2018) at three core sites. (a) *Globigerina bulloides* in the Alboran Sea in the Mediterranean (Cacho et al., 1999). (b) *Neoglobobulimina pachyderma* in the Gulf of Alaska (Davies et al., 2011; Davies-Walczak et al., 2014; Praetorius et al., 2015). (c, d) *Globigerinoides ruber* in the eastern equatorial Pacific (Koutavas & Sachs, 2008; Koutavas & Joannides, 2012). The Alboran Sea and Gulf of Alaska alkenone reconstructions reflect SSTs from November to May and June to August, respectively (Tierney & Tingley, 2018). Solid lines are posterior distribution means with ± 1 standard deviation shading. Solid green bars are the site modern annual or seasonal SST means ± 1 standard deviation, using monthly (1981–2017) National Oceanic and Atmospheric Administration Optimum Interpolation Sea Surface Temperature V2 (Reynolds et al., 2002).

used our hierarchical annual calibration, though results are nearly identical for the seasonal calibration. As a test case, we applied the calibration to a *G. bulloides* $\delta^{18}O_c$ time series from core MD95-2043, spanning from 52 ka to present in the eastern Alboran Sea in the Mediterranean (Cacho et al., 1999). We used a modern $\delta^{18}O_w$ value of 1.05‰ (VSMOW; LeGrande & Schmidt, 2006). Resulting SSTs are compared to an alkenone $U_{37}^{K'}$ reconstruction from the same site, calibrated with BAYSPLINE (Tierney & Tingley, 2018).

Our model-based *G. bulloides* reconstruction matches the mean and variability of the alkenone-based reconstruction remarkably well (Figure 12a), though the alkenone posterior is noticeably tighter ($\sigma = 1.4^\circ\text{C}$) than for *G. bulloides* ($\sigma = 2.6^\circ\text{C}$). However, the BAYSPLINE calibration explicitly calibrates alkenone data in the Mediterranean to November–May SSTs (Tierney & Tingley, 2018), and indeed the late Holocene values approach the modern November–May value of 16.5°C . The median predicted values from *G. bulloides* are cooler than the present-day annual range and generally follow the alkenones. This may reflect a slight cool-seasonal bias for *G. bulloides*, which has peak abundance in the winter and spring in the Western Mediterranean (Bárcena et al., 2004; Rigual-Hernández et al., 2012). However, the uncertainty bounds are large and the latest Holocene estimates do encompass modern mean annual SSTs (18.8°C ; Boyer et al., 2013). Within calibration uncertainty, the model reconstruction using *G. bulloides* reasonably estimates annual SST.

4.2. High-Latitude Settings: Cool-Temperature Foraminifera

To investigate the performance of our model in a cold-water setting (Figure 12b), we apply our annual hierarchical calibration to a deglacial-to-Holocene *N. pachyderma* $\delta^{18}\text{O}_c$ record from core EW0408-85JC on the Gulf of Alaska (Davies et al., 2011; Davies-Walczak et al., 2014; Praetorius et al., 2015). The reconstruction uses modern $\delta^{18}\text{O}_w = -0.95\text{‰}$ (VSMOW; LeGrande & Schmidt, 2006). The results are similar if we used the seasonal hierarchical model (not shown). The *N. pachyderma* record shows similar changes through time as $U_{37}^{K'}$ but is offset from the latter by approximately 2–3 °C. At this location, the $U_{37}^{K'}$ calibration BAYSPLINE assumes a June–August bias and explicitly predicts summer temperatures; thus, the warm offset is expected and primarily reflects differences in the seasonal production. The *N. pachyderma* reconstruction is slightly cooler on average than the modern annual SST range. *N. pachyderma* occupies a broad depth habitat range in the North Pacific (Kuroyanagi et al., 2011), and so the slightly cooler SSTs predicted may reflect the offset between sea surface and actual *N. pachyderma* habitat depth, which we do not explicitly account for with our calibrations. Additionally, *N. pachyderma* tends to have peak abundance in the spring and late winter in the Gulf of Alaska (Sautter & Thunell, 1989), which has modern SSTs closer to the late-Holocene predictions. Regardless, the reconstructed Holocene SSTs fall within uncertainty of present-day mean annual values, suggesting that our model is reasonably accurate at estimating annual SST from *N. pachyderma* $\delta^{18}\text{O}_c$ at this location.

4.3. Annual Versus Seasonal Calibration

To explore the impact of accounting for seasonality, we apply both the annual and seasonal hierarchical models to a *G. ruber* record from core VM21-30 on the eastern equatorial Pacific spanning the last 30 ka (Koutavas & Sachs, 2008; Koutavas & Joanides, 2012). At this site, the effect of model selection is notable, with the annual model resulting in overall cooler temperatures and a larger range of variability (Figure 12c) than the seasonal model (Figure 12d). In both cases we used a modern $\delta^{18}\text{O}_w$ of 0.24‰ (VSMOW; LeGrande & Schmidt, 2006). The annual calibration and an alkenone-based reconstruction have similar uncertainties (both $\bar{\sigma} = 2.1$ °C; Figure 12c). The seasonal calibration results in temperatures that agree more closely with the alkenone-inferred SSTs, both in the mean and magnitude of reconstructed trends (Figure 12d), as well as producing a tighter reconstruction ($\bar{\sigma} = 1.9$ °C). Observations of *G. ruber* in the tropical Pacific show peak abundance for warm SSTs (Thunell et al., 1983), and our flux-based estimates suggest that *G. ruber* at this site should be seasonally biased toward December through May, the time of year when SSTs are at their warmest (Fiedler & Talley, 2006). The fact that the alkenone reconstruction best matches our seasonal *G. ruber* predictions (Figure 12d) suggests that the $U_{37}^{K'}$ record could also be biased toward warmer months rather than recording annual SSTs. However, there is no indication of warm bias in the eastern equatorial Pacific core top alkenone data (Kienast et al., 2012; Tierney & Tingley, 2018), and the $U_{37}^{K'}$ predictions during the Holocene still overlap with the range of annual SSTs at this site.

4.4. Influence of Freshwater Input and Changes in $\delta^{18}\text{O}_w$

To see the impact of freshwater inputs on a $\delta^{18}\text{O}_c$ -based temperature reconstruction, we apply our calibration to a *G. ruber* record from core GeoB 6518-1 (Schefuß et al., 2005) on the west coast of Africa in the Gulf of Guinea, near the mouth of the Congo River (Figure 13a). The core spans the last deglaciation (20 ka to present), which saw dramatic changes to the central African hydroclimate (Gasse, 2000; Schefuß et al., 2005). Alkenone $U_{37}^{K'}$ data, which are not affected by changing $\delta^{18}\text{O}_w$ or freshwater input, are available from this site for comparison with $\delta^{18}\text{O}_c$ (Schefuß et al., 2005). As in earlier examples, we apply our hierarchical seasonal calibration to the *G. ruber* $\delta^{18}\text{O}_c$ record and use BAYSPLINE (Tierney & Tingley, 2018) to reconstruct temperature from the $U_{37}^{K'}$ record. The estimated modern $\delta^{18}\text{O}_w$ for the site is 0.52‰ (VSMOW; LeGrande & Schmidt, 2006), and our estimated peak seasonal growth for *G. ruber* at this site is from September to June. The reconstructions from $\delta^{18}\text{O}_c$ and $U_{37}^{K'}$ are relatively similar in the Late Holocene, the Younger Dryas (~12 ka), and the Last Glacial Maximum. Outside of these periods, we see the foraminifera-based reconstruction diverge from the alkenone-based reconstruction, predicting warmer temperatures. These periods (the Early Holocene and Bølling-Allerød) correspond to times of larger freshwater inputs from increased precipitation across the extensive Congo River basin (Schefuß et al., 2005). This application demonstrates that, in coastal regions with large freshwater inputs, $\delta^{18}\text{O}_c$ will be biased toward more negative values, even when our new calibration models are applied.

4.5. Gulf of Mexico Sediment Traps: Predicting Monthly $\delta^{18}\text{O}_c$ of *G. ruber*

In this example we test how well our calibration can replicate the seasonality of monthly $\delta^{18}\text{O}_c$ measured from *G. ruber* (white and pink) at a sediment trap site, where SSTs and $\delta^{18}\text{O}_w$ values are well constrained

Figure 13. Other example applications. (a) The seasonal hierarchical calibration applied to *G. ruber* to reconstruct SST in the Gulf of Guinea (Schefuß et al., 2005). Separation between the foraminifera and alkenone-based reconstructions indicates periods with increased freshwater input. Solid lines are posterior distribution means with ± 1 standard deviation shading. Solid green bars are the site modern annual or seasonal SST means ± 1 standard deviation, using monthly (1981–2017) NOAA Optimum Interpolation Sea Surface Temperature V2 (Reynolds et al., 2002). (b) The annual hierarchical calibration used to predict $\delta^{18}\text{O}_c$ for *G. ruber* in the northern Gulf of Mexico, compared with averaged sediment trap observations (2010–2013) from Richey et al. (2019). The predictions were run with $\delta^{18}\text{O}_w$ of 1.66‰ and 0.86‰ (Vienna Standard Mean Ocean Water), based on the observed mean and range $\delta^{18}\text{O}_w$ for the site. These predictions used HadISST monthly average SSTs (Rayner et al., 2003), which reasonably replicated buoy-measured SST seasonality (Richey et al., 2019). All errors are ± 1 standard deviation. (c) The annual pooled calibration applied to *Morozovella* spp. $\delta^{18}\text{O}_c$ (John et al., 2008) analyzed in a Paleocene-Eocene Thermal Maximum section from Bass River, New Jersey. TEX_{86} (Sluijs et al., 2007) is calibrated using BAYSPAR (Tierney & Tingley, 2014). Samples are plotted against downcore depth. Solid and dotted lines show posterior means. Shading shows ± 1 standard deviation. Dotted $\delta^{18}\text{O}_c$ lines are the calibration run with 0.5‰ and -0.5 ‰ for $\delta^{18}\text{O}_w$, at around 0.0‰ modeled in Tindall et al. (2010).

(Figure 13b). The site is in the northern Gulf of Mexico and has repeated $\delta^{18}\text{O}_c$ measurements for each month from 2010 to 2013 (Richey et al., 2019). These years are pooled into monthly mean $\delta^{18}\text{O}_c$ values. We use the hierarchical annual calibration and not the seasonal variant because the seasonal calibrations favor SSTs associated with high foraminifera abundance and so tend to be better at predicting peak abundance seasons rather than growth for each month of the year. We use HadISST (Rayner et al., 2003) monthly mean SSTs from the nearest grid point to the sediment trap site, as Richey et al. (2019) found that HadISST reasonably replicated the seasonality of SSTs when compared with measurements from a local buoy. We create two predictions each with $\delta^{18}\text{O}_w$ values of 1.40‰ (VSMOW) and 0.86‰ (VSMOW), following the measured mean $\delta^{18}\text{O}_w$ of 1.13‰ and annual range of 0.53‰ between 0 and 50m depth (Richey et al., 2019). We lag our calibration predictions by 1 month to account for the time needed for foraminiferal tests to settle in the water column. Our calibration replicates the observed $\delta^{18}\text{O}_c$ seasonal pattern remarkably well (Figure 13b). The closest predictions are from the model using $\delta^{18}\text{O}_w = 0.86$ ‰. The absolute difference of its monthly prediction and observation means for *G. ruber* (pink) is between 0.02‰ and 0.24‰ and between 0.03‰ and 0.66‰ for *G. ruber* (white). This is an encouraging result given that the full prediction distribution has

$\bar{\sigma} = 0.40\text{‰}$ and the additional spread in the sediment trap observations can be as high as $\sigma = 0.47\text{‰}$ (see Figure 13b).

4.6. Application to Deep-Time Paleoclimate Reconstructions

Data from Late Quaternary sediment cores benefit from application of our hierarchical models, but these models cannot be reliably applied to non-extant planktic species in deeper geological time. Here, our annual pooled model is a more appropriate choice, and this was one of our primary motivations for developing this calibration model. In the pooled model, all species are assumed to calcify similarly, which enables us to approximate a general “planktic dependency” of $\delta^{18}\text{O}_c$ on annual SSTs. Arguably, this is the best first-order approach for applications to non-extant species, for which species-specific information such as seasonality, depth habitat, presence of symbionts, and crust formation are poorly constrained.

As a demonstration, we apply our pooled model to a $\delta^{18}\text{O}_c$ record of the planktic *Morozovella* spp. from the Paleocene-Eocene Thermal Maximum marine section from Bass River, New Jersey (John et al., 2008) (Figure 13). These specimens are “glassy” (well preserved) and thus are unlikely to show any of the isotopic overprinting common in “frosty” deep-sea specimens from this period, which can generate anomalously cold SST reconstructions from simplistic interpretation of $\delta^{18}\text{O}_c$ measurements (Kozdon et al., 2011). We show reconstructions for $\delta^{18}\text{O}_w = \pm 0.5\text{‰}$, which are plausible values for this site (John et al., 2008), bounding the modeled value of 0.0‰ from Tindall et al. (2010). We compare the results to a TEX_{86} -based SST reconstruction from the same section (Sluijs et al., 2007). The TEX_{86} data are calibrated to SSTs using the BAYSPAR analogue method as described in Tierney and Tingley (2014). We use a search tolerance of 0.15, resulting in 11 modern analog grid points. We use a wide, weakly informative prior for SST (Gaussian, 30.0°C mean and 20.0°C standard deviation) for both the TEX_{86} and foraminiferal data.

The inferred SSTs from the *Morozovella* spp. and TEX_{86} data generally agree with one another and overlap within uncertainties, though there is increased separation just below 355 m, near the onset of the Paleocene-Eocene Thermal Maximum (Figure 13c). It is tempting to interpret this as temporary differentiation of depth habitat (e.g., *Morozovella* spp. migrating deeper in the water column and/or TEX_{86} producers forced to the surface), but the large uncertainties on both SST reconstructions must also be considered. We note that the average standard deviation of the foraminiferal reconstruction ($\bar{\sigma} = 2.4^\circ\text{C}$) is much smaller than the uncertainty on the TEX_{86} estimates ($\bar{\sigma} = 5.0^\circ\text{C}$). This can be attributed to the fact that the TEX_{86} values for this sequence (which approach 0.9) are far outside the range of the calibration data set—which contains only a few values above 0.75—requiring heavy extrapolation. Another important factor is that overall, the temperature sensitivity of TEX_{86} is poorly constrained relative to β in our $\delta^{18}\text{O}_c$ calibrations. Indeed, β is well estimated by our pooled model and remarkably similar to thermodynamic expectation (annual pooled β posterior 95% CI is between -0.234 and -0.228) for modern species. That said, the sensitivity of the $\delta^{18}\text{O}_c$ proxy to $\delta^{18}\text{O}_w$ must be considered, where a difference of 1.0‰ corresponds to approximately 4.1°C (Figure 13c). In an unglaciated Eocene world at 55 Ma, global $\delta^{18}\text{O}_w$ can be fairly well estimated, but poorly constrained local variations in $\delta^{18}\text{O}_w$ are another potential source of uncertainty to interpreting $\delta^{18}\text{O}_c$. This deep-time example underlines the importance of constraining variability and quantifying uncertainty in $\delta^{18}\text{O}_w$ when using $\delta^{18}\text{O}_c$ to reconstruct SST.

Overall, these examples show that our calibration models produce reasonable results when applied to foraminiferal records of $\delta^{18}\text{O}_c$. Our SST reconstructions compare favorably with independent SST reconstructions using biomarker-based proxies in a variety of paleoceanographic settings and using different planktic species. We are also able to predict subannual $\delta^{18}\text{O}_c$ seasonality found in sediment trap measurements. Importantly, these examples demonstrate how our models incorporate multiple sources of uncertainty in reconstructing SST, which is often not reported or considered in published downcore records and subsequent interpretations. This fundamental uncertainty in SST reconstructions, which is inherent in any paleo-proxy, needs to be considered when comparing records, integrating SST reconstructions across geographic regions or comparing data to climate model output.

4.7. bayfox: Bayesian Foraminifera Calibration Software

Our calibration models are available to users as a software library called `bayfox`. This library is packaged and available in both Python (<https://github.com/brews/bayfox>) and R (<https://github.com/brews/bayfoxr>). Scripts are also available for MATLAB/Octave (<https://github.com/brews/bayfoxm>). These packages include both forward and reverse calibration models so that users can infer $\delta^{18}\text{O}_c$ from SSTs and infer SSTs from $\delta^{18}\text{O}_c$. The software is available under an Open Source license.

5. Conclusions

Our Bayesian calibration models enhance the widespread use of planktic $\delta^{18}\text{O}_c$ in paleoceanography by providing a realistic representation of proxy uncertainty based on core top variability. We find that, in spite of the many biological and environmental factors that can influence planktic $\delta^{18}\text{O}_c$, the inferred sensitivity to SST is remarkably similar to established inorganic calcite calibration curves, attesting to the fidelity of $\delta^{18}\text{O}_c$. Our annual, seasonal, and species-specific calibration exercises demonstrate that model performance is improved by accounting for foraminiferal seasonal abundance and species-specific variability in calibration parameters. However, some residual patterns remain and can be difficult to diagnose due to the complex set of environmental factors which control the abundance of each foraminiferal species, as well as uncertainty in observed variables—specifically $\delta^{18}\text{O}_w$. We demonstrate how the calibration can be used to reconstruct SST in the Late Quaternary, where generally speaking, the most applicable model is the hierarchical seasonal model. We demonstrate how the calibration can replicate the seasonal signal of $\delta^{18}\text{O}_c$ observed in Gulf of Mexico sediment traps. We also demonstrate how our pooled annual model can be used to infer SSTs from $\delta^{18}\text{O}_c$ of non-extant species of planktic foraminifera in deeper geological time. We have made the calibration models available in Open Source software libraries (bayfox), so that users can apply these calibrations to both forward and inverse $\delta^{18}\text{O}_c$ modeling problems.

Appendix A: Bayesian Regression Model Design and Priors

As described in section 2.4, we trained four Bayesian calibration models using two model designs: one that pools all species together and one that uses a hierarchical structure to account for variations between species. The pooled model likelihood is described in equation (1), where α , β , and τ are the regression intercept, coefficient, and standard error, respectively. α , β and τ were given weakly informed prior distributions

$$\begin{aligned}\alpha &\sim \mathcal{N}(3, 4) \\ \beta &\sim \mathcal{N}(-0.2, 1) \\ \tau &\sim \text{HalfCauchy} \quad (1)\end{aligned}$$

The normal priors around α and β loosely reflect existing inorganic precipitation and culture-derived calibrations from the literature (see Bemis et al., 1998). However, we found that less informed α and β prior distributions, for example, $\mathcal{N}(0, 10)$, produced comparable results.

The hierarchical model (equation (2)) and its priors (equation (3)) depend on shared hyperparameters

$$\begin{aligned}\mu_\alpha &\sim \mathcal{N}(3, 4) \\ \sigma_\alpha &\sim \text{HalfCauchy} \quad (0.5) \\ \mu_\beta &\sim \mathcal{N}(-0.2, 1) \\ \sigma_\beta &\sim \text{HalfCauchy} \quad (0.25) \\ \sigma_m &\sim \text{HalfCauchy} \quad (1) \\ \sigma_d &\sim \text{HalfCauchy} \quad (1)\end{aligned}$$

With this design, updates to individual species parameters also inform shared hyperparameters. At the same time, hyperparameters influence the individual species parameters.

We infer the posterior distribution of the models with a No-U-Turn Sampler (Hoffman & Gelman, 2014)—an MCMC sampler variant using Hamiltonian mechanics. We initialized No-U-Turn Sampler using an identity mass matrix with a diagonal adapted to the variance of the sampler tuning steps. The start of each sampling chain was the prior mean with added uniform noise (between -1 and 1). The sampler was run in two chains each with 1,000 tuning draws followed by 5,000 draws. We compared the chains for convergence and autocorrelation. Tuning steps were removed from the draws, and the chains were combined for analysis.

The forward models described above can be “reversed” to infer SSTs. Inference of SSTs, $\mathbf{C}$, requires a vector of observed δ_c and δ_w , and Gaussian priors for SST. The full conditional posterior for $\mathbf{C}$, given all other variables, is then a multivariate normal distribution

$$\mathbf{C} \sim \mathcal{N}(\Psi_C \cdot \mathbf{V}_C, \mathbf{V}_C)$$

where

$$\Psi_C = \Sigma_C^{-1} \cdot \mu_C + \tau^{-2} \cdot \beta(\delta_c - \alpha - [\delta_w - 0.27])$$

$$\mathbf{V}_C = (\Sigma_C^{-1} + \tau^{-2} \cdot \beta^2)^{-1}$$

where Σ_C^{-1} is the prior covariance matrix and μ_C is the prior mean vector. This inversion is very close to that described in Tierney and Tingley (2014) but includes the $\delta^{18}\text{O}_w$ offset. Other parameters (τ , α , β) are drawn from the posterior of the forward calibration models. This applies to the pooled calibration model. For the hierarchical calibrations, τ , α , and β are replaced with their species-specific parameters τ_i , α_i , and β_i . Given these parameters, no additional steps are needed for the hierarchical calibrations.

Appendix B: Ice Volume Correction

We applied a simple ice volume correction to the $\delta^{18}\text{O}_c$ proxy records used in section 4. We assumed that the global $\delta^{18}\text{O}_w$ change since the Last Glacial Maximum is 1.0‰, following the pore water estimates of Schrag et al. (1996). We applied this to scale the LR04 benthic stack (Lisiecki & Raymo, 2005)—a proxy for global ice volume. This scaled time series is an estimate of global ice volume change in $\delta^{18}\text{O}$ for seawater ($\delta^{18}\text{O}_{\text{ice}}$). We removed the ice volume change from a $\delta^{18}\text{O}_c$ proxy sample with

$$\delta_{\text{cor}} = \frac{1,000 + \delta_c}{1 + \frac{\delta_{\text{ice}}}{1,000}} - 1,000$$

where δ_{cor} is the ice volume corrected $\delta^{18}\text{O}_c$ proxy sample and δ_{ice} is $\delta^{18}\text{O}_{\text{ice}}$ for the corresponding date of δ_c . If $\delta^{18}\text{O}_{\text{ice}}$ had no exact corresponding date, δ_c was linearly interpolated from $\delta^{18}\text{O}_{\text{ice}}$ values.

Software implementing this ice volume correction is available for broader use in the Open Source Python package `erebusfall` (<https://github.com/brews/erebusfall>).

References

Acknowledgments

This research was funded by the Heising-Simons Foundation (2016-015) and the National Science Foundation (AGS-1602156). We thank our Editors and anonymous reviewers for their time and thoughtful comments. Thanks to Kaustubh Thirumalai for help collecting sediment trap data. Core top data used for this analysis are available as supporting information. Open Source Software packages implementing these calibrations are available for Python (<https://github.com/brews/bayfox>) and the R statistical environment (<https://github.com/brews/bayfoxr>). Scripts are also available for MATLAB/Octave (<https://github.com/brews/bayfoxm>).

- Abrantes, F., Meggers, H., Nave, S., Bollman, J., Palma, S., Sprengel, C., et al. (2002). Fluxes of micro-organisms along a productivity gradient in the Canary Islands region (29°N): Implications for paleoreconstructions. *Deep Sea Research Part II: Topical Studies in Oceanography*, 49(17), 3599–3629. [https://doi.org/10.1016/S0967-0645\(02\)00100-5](https://doi.org/10.1016/S0967-0645(02)00100-5)
- Arbuszewski, J. A., deMenocal, P. B., Cléroux, C., Bradtmiller, L., & Mix, A. (2013). Meridional shifts of the Atlantic intertropical convergence zone since the Last Glacial Maximum. *Nature Geoscience*, 6(11), 959–962. <https://doi.org/10.1038/ngeo1961>
- Arbuszewski, J., deMenocal, P., Kaplan, A., & Farmer, E. C. (2010). On the fidelity of shell-derived $\delta^{18}\text{O}_{\text{seawater}}$ estimates. *Earth and Planetary Science Letters*, 300(3), 185–196. <https://doi.org/10.1016/j.epsl.2010.10.035>
- Aurahs, R., Treis, Y., Darling, K., & Kucera, M. (2011). A revised taxonomic and phylogenetic concept for the planktonic foraminifer species *Globigerinoides ruber* based on molecular and morphometric evidence. *Marine Micropaleontology*, 79(1), 1–14. <https://doi.org/10.1016/j.marmicro.2010.12.001>
- Bárcena, M. A., Flores, J. A., Sierro, F. J., Pérez-Folgado, M., Fabres, J., Calafat, A., & Canals, M. (2004). Planktonic response to main oceanographic changes in the Alboran Sea (Western Mediterranean) as documented in sediment traps and surface sediments. *Marine Micropaleontology*, 53(3), 423–445. <https://doi.org/10.1016/j.marmicro.2004.09.009>
- Bé, A. W. H. (1960). Ecology of Recent planktonic foraminifera—Part 2. Bathymetric and seasonal distributions in the Sargasso Sea off Bermuda. *Micropaleontology*, 6(4), 373–392.
- Bé, A. W. H. (1977). An ecological zoogeographic and taxonomic review of Recent planktonic foraminifera. In A. T. S. Ramsay (Ed.), *Oceanic Micropaleontology* (pp. 1–100). London, UK: Academic Press.
- Bemis, B. E., Spero, H. J., Bijma, J., & Lea, D. W. (1998). Reevaluation of the oxygen isotopic composition of planktonic foraminifera: Experimental results and revised paleotemperature equations. *Paleoceanography*, 13(2), 150–160. <https://doi.org/10.1029/98PA00070>
- Bemis, B. E., Spero, H. J., & Thunell, R. C. (2002). Using species-specific paleotemperature equations with foraminifera: A case study in the Southern California Bight. *Marine Micropaleontology*, 46(3), 405–430. [https://doi.org/10.1016/S0377-8398\(02\)00083-X](https://doi.org/10.1016/S0377-8398(02)00083-X)
- Bijma, J., Spero, H. J., & Lea, D. W. (1999). Reassessing foraminiferal stable isotope geochemistry: Impact of the oceanic carbonate system (experimental results). In G. Fischer, & G. Wefer (Eds.), *Use of proxies in paleoceanography* (pp. 489–512). Berlin, Heidelberg: Springer
- Berlin Heidelberg. https://doi.org/10.1007/978-3-642-58646-0_20
- Boussetta, S., Kallel, N., Bassinot, F., Labeyrie, L., Duplessy, J.-C., Caillon, N., et al. (2012). Mg/Ca-paleothermometry in the western Mediterranean Sea on planktonic foraminifer species *Globigerina bulloides*: Constraints and implications. *Comptes Rendus Geoscience*, 344(5), 267–276. <https://doi.org/10.1016/j.crte.2012.02.001>
- Bouvier-Soumagnac, Y., & Duplessy, J.-C. (1985). Carbon and oxygen isotopic composition of planktonic foraminifera from laboratory culture, plankton tows and Recent sediment: Implications for the reconstruction of paleoclimatic conditions and of the global carbon cycle. *The Journal of Foraminiferal Research*, 15(4), 302–320. <https://doi.org/10.2113/gsjfr.15.4.302>
- Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 26(2), 211–252.
- Boyer, T. P., Antonov, J. I., Baranova, O. K., Garcia, H. E., Johnson, D. R., Mishonov, A. V., et al. (2013). World ocean database 2013. <https://doi.org/10.7289/v5nz85mt>
- Brown, S. J., & Elderfield, H. (1996). Variations in Mg/Ca and Sr/Ca ratios of planktonic foraminifera caused by postdepositional dissolution: Evidence of shallow Mg-dependent dissolution. *Paleoceanography*, 11(5), 543–551. <https://doi.org/10.1029/96PA01491>

- Cacho, I., Grimalt, J. O., Pelejero, C., Canals, M., Sierro, F. J., Flores, J. A., & Shackleton, N. (1999). Dansgaard-Oeschger and Heinrich event imprints in Alboran Sea paleotemperatures. *Paleoceanography*, 14(6), 698–705. <https://doi.org/10.1029/1999PA900044>
- Caron, D. A., Anderson, O. R., Lindsey, J. L., Faber, W. W., & Lin Lim, E. E. (1990). Effects of gametogenesis on test structure and dissolution of some spinose planktonic foraminifera and implications for test preservation. *Marine Micropaleontology*, 16(1), 93–116. [https://doi.org/10.1016/0377-8398\(90\)90031-G](https://doi.org/10.1016/0377-8398(90)90031-G)
- Cléroux, C., Cortijo, E., Anand, P., Labeyrie, L., Bassinot, F., Caillon, N., & Duplessy, J.-C. (2008). Mg/Ca and Sr/Ca ratios in planktonic foraminifera: Proxies for upper water column temperature reconstruction. *Paleoceanography*, 23, PA3214. <https://doi.org/10.1029/2007PA001505>
- Craig, H., & Gordon, L. I. (1965). Isotope oceanography: Deuterium and oxygen 18 variation in the ocean and marine atmosphere. In E. Tongiorgi (Ed.), *Stable isotopes in oceanographic studies and paleotemperatures* (pp. 9–130). Italy: Cons. Nax. di Rech., Spoleto.
- Dahl, K. A., & Oppo, D. W. (2006). Sea surface temperature pattern reconstructions in the Arabian Sea. *Paleoceanography*, 21, PA1014. <https://doi.org/10.1029/2005PA001162>
- Darling, K. F., & Wade, C. M. (2008). The genetic diversity of planktic foraminifera and the global distribution of ribosomal RNA genotypes. *Marine Micropaleontology*, 67(3), 216–238. <https://doi.org/10.1016/j.marmicro.2008.01.009>
- Davies, M. H., Mix, A. C., Stoner, J. S., Addison, J. A., Jaeger, J., Finney, B., & Wiest, J. (2011). The deglacial transition on the southeastern Alaska Margin: Meltwater input, sea level rise, marine productivity, and sedimentary anoxia. *Paleoceanography*, 26, PA2223. <https://doi.org/10.1029/2010PA002051>
- Davies-Walczak, M., Mix, A., Stoner, J., Southon, J., Cheseby, M., & Xuan, C. (2014). Late Glacial to Holocene radiocarbon constraints on North Pacific Intermediate Water ventilation and deglacial atmospheric CO₂ sources. *Earth and Planetary Science Letters*, 397, 57–66. <https://doi.org/10.1016/j.epsl.2014.04.004>
- Dekens, P. S., Lea, D. W., Pak, D. K., & Spero, H. J. (2002). Core top calibration of Mg/Ca in tropical foraminifera: Refining paleotemperature estimation. *Geochemistry, Geophysics, Geosystems*, 3(4), 1022. <https://doi.org/10.1029/2001GC000200>
- Duplessy, J., & Blanc, P. (1981). Oxygen-18 enrichment of planktonic foraminifera due to gametogenic calcification below the euphotic zone. *Science*, 213(4513), 1247–1250. <https://doi.org/10.1126/science.213.4513.1247>
- Duplessy, J. C., Lalou, C., & Vinot, A. C. (1970). Differential isotopic fractionation in benthic foraminifera and paleotemperatures reassessed. *Science*, 168(3928), 250–251. <https://doi.org/10.1126/science.168.3928.250>
- Dyez, K. A., Zahn, R., & Hall, I. R. (2014). Multicentennial Agulhas leakage variability and links to North Atlantic climate during the past 80,000 years. *Paleoceanography*, 29, 1238–1248. <https://doi.org/10.1002/2014PA002698>
- Elderfield, H., & Ganssen, G. (2000). Past temperature and d18o of surface ocean waters inferred from foraminiferal Mg/Ca ratios. *Nature*, 405(6785), 442–445. <https://doi.org/10.1038/35013033>
- Emiliani, C. (1955). Pleistocene temperatures. *The Journal of Geology*, 63(6), 538–578. <https://doi.org/10.1086/626295>
- Epstein, S., Buchsbaum, R., Lowenstam, H. A., & Urey, H. C. (1953). Revised carbonate-water isotopic temperature scale. *Geological Society of America Bulletin*, 64(11), 1315. [https://doi.org/10.1130/0016-7606\(1953\)64\[1315:RCITS\]2.0.CO;2](https://doi.org/10.1130/0016-7606(1953)64[1315:RCITS]2.0.CO;2)
- Erez, J., & Luz, B. (1983). Experimental paleotemperature equation for planktonic foraminifera. *Geochimica et Cosmochimica Acta*, 47(6), 1025–1031. [https://doi.org/10.1016/0016-7037\(83\)90232-6](https://doi.org/10.1016/0016-7037(83)90232-6)
- Fairbanks, R. G., Sverdrlove, M., Free, R., Wiebe, P. H., & Bé, A. W. H. (1982). Vertical distribution and isotopic fractionation of living planktonic foraminifera from the Panama Basin. *Nature*, 298(5877), 841–844. <https://doi.org/10.1038/298841a0>
- Fairbanks, R. G., & Wiebe, P. H. (1980). Foraminifera and CHLOROPHYLL MAXimum: Vertical distribution, seasonal succession, and paleoceanographic significance. *Science*, 209(4464), 1524–1526.
- Fallet, U., Castañeda, I. S., Henry-Edwards, A., Richter, T. O., Boer, W., Schouten, S., & Brummer, G.-J. (2012). Sedimentation and burial of organic and inorganic temperature proxies in the Mozambique Channel, SW Indian Ocean. *Deep Sea Research Part I: Oceanographic Research Papers*, 59, 37–53. <https://doi.org/10.1016/j.dsr.2011.10.002>
- Farmer, E. C. (2005). Reconstructing tropical Atlantic climate from magnesium/calcium and oxygen isotopes of planktonic foraminifera (Ph.D thesis), Columbia University, New York.
- Fiedler, P. C., & Talley, L. D. (2006). Hydrography of the eastern tropical Pacific: A review. *Progress in Oceanography*, 69(2), 143–180. <https://doi.org/10.1016/j.pcean.2006.03.008>
- Ganssen, G. M., & Kroon, D. (2000). The isotopic signature of planktonic foraminifera from NE Atlantic surface sediments: Implications for the reconstruction of past oceanic conditions. *Journal of the Geological Society*, 157(3), 693–699. <https://doi.org/10.1144/jgs.157.3.693>
- Garidel-Thoron, T. d., Rosenthal, Y., Beaufort, L., Bard, E., Sonzogni, C., & Mix, A. C. (2007). A multiproxy assessment of the western equatorial Pacific hydrography during the last 30 kyr. *Paleoceanography*, 22, PA3204. <https://doi.org/10.1029/2006PA001269>
- Gasse, F. (2000). Hydrological changes in the African tropics since the Last Glacial Maximum. *Quaternary Science Reviews*, 19(1), 189–211. [https://doi.org/10.1016/S0277-3791\(99\)00061-X](https://doi.org/10.1016/S0277-3791(99)00061-X)
- Gebregiorgis, D., Hathorne, E. C., Sijinkumar, A. V., Nath, B. N., Nürnberg, D., & Frank, M. (2016). South Asian summer monsoon variability during the last ~54kyrs inferred from surface water salinity and river runoff proxies. *Quaternary Science Reviews*, 138, 6–15. <https://doi.org/10.1016/j.quascirev.2016.02.012>
- Gelman, A. (2014). *Bayesian data analysis, Chapman CRC texts in statistical science* (3rd ed.). Boca Raton: CRC Press.
- Gibbons, F. T., Oppo, D. W., Mohtadi, M., Rosenthal, Y., Cheng, J., Liu, Z., & Linsley, B. K. (2014). Deglacial $\delta^{18}\text{O}$ and hydrologic variability in the tropical Pacific and Indian Oceans. *Earth and Planetary Science Letters*, 387, 240–251. <https://doi.org/10.1016/j.epsl.2013.11.032>
- Gibson, K. A., Thunell, R. C., Machain-Castillo, M. L., Fehrenbacher, J., Spero, H. J., Wejnert, K., et al. (2016). Evaluating controls on planktonic foraminiferal geochemistry in the Eastern Tropical North Pacific. *Earth and Planetary Science Letters*, 452, 90–103. <https://doi.org/10.1016/j.epsl.2016.07.039>
- Hamilton, C. P., Spero, H. J., Bijma, J., & Lea, D. W. (2008). Geochemical investigation of gametogenic calcite addition in the planktonic foraminifera *Orbulina universa*. *Marine Micropaleontology*, 68(3), 256–267. <https://doi.org/10.1016/j.marmicro.2008.04.003>
- Hemleben, C., Spindler, M., & Anderson, O. R. (1989). *Modern planktonic foraminifera*. New York, NY: Springer New York. <https://doi.org/10.1007/978-1-4612-3544-6>
- Hoffman, M. D., & Gelman, A. (2014). The No-U-turn Sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15(1), 1593–1623.
- Hoyer, S., & Hamman, J. J. (2017). xarray: N-D labeled arrays and datasets in Python. *Journal of Open Research Software*, 5, 10. <https://doi.org/10.5334/jors.148>
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 9(3), 90–95. <https://doi.org/10.1109/MCSE.2007.55>
- Hut, G. (1987). Consultants' group meeting on stable isotope reference samples for geochemical and hydrological investigations.

- John, C. M., Bohaty, S. M., Zachos, J. C., Slujs, A., Gibbs, S., Brinkhuis, H., & Bralower, T. J. (2008). North American continental margin records of the Paleocene-Eocene thermal maximum: Implications for global carbon and hydrological cycling. *Paleoceanography*, 23, PA2217. <https://doi.org/10.1029/2007PA001465>
- Johnstone, H. J. H., Yu, J., Elderfield, H., & Schulz, M. (2011). Improving temperature estimates derived from Mg/Ca of planktonic foraminifera using X-ray computed tomography-based dissolution index, XDX. *Paleoceanography*, 26, PA1215. <https://doi.org/10.1029/2009PA001902>
- Kienast, M., MacIntyre, G., Dubois, N., Higginson, S., Normandeau, C., Chazen, C., & Herbert, T. (2012). Alkenone unsaturation in surface sediments from the eastern equatorial pacific: Implications for sst reconstructions. *Paleoceanography and Paleoclimatology*, 27, PA1210. <https://doi.org/10.1029/2011PA002254>
- Kim, S.-T., & O'Neil, J. R. (1997). Equilibrium and nonequilibrium oxygen isotope effects in synthetic carbonates. *Geochimica et Cosmochimica Acta*, 61(16), 3461–3475. [https://doi.org/10.1016/S0016-7037\(97\)00169-5](https://doi.org/10.1016/S0016-7037(97)00169-5)
- King, A. L., & Howard, W. R. (2005). $\delta^{18}\text{O}$ seasonality of planktonic foraminifera from Southern Ocean sediment traps: Latitudinal gradients and implications for paleoclimate reconstructions. *Marine Micropaleontology*, 56(1), 1–24. <https://doi.org/10.1016/j.marmicro.2005.02.008>
- Kohfeld, K. E., Fairbanks, R. G., Smith, S. L., & Walsh, I. D. (1996). Neoglobobulimina pachyderma (sinistral coiling) as paleoceanographic tracers in polar oceans: Evidence from northeast water polynya plankton tows, sediment traps, and surface sediments. *Paleoceanography*, 11(6), 679–699. <https://doi.org/10.1029/96PA02617>
- Kontakiotis, G., Mortyn, P. G., Antonarakou, A., Martinez-Boti, M. A., & Triantaphyllou, M. V. (2011). Field-based validation of a diagenetic effect on *G. ruber* Mg/Ca paleothermometry: Core top results from the Aegean Sea (eastern Mediterranean). *Geochemistry, Geophysics, Geosystems*, 12, Q09004. <https://doi.org/10.1029/2011GC003692>
- Koutavas, A., & Joannides, S. (2012). El Niño-Southern Oscillation extrema in the Holocene and Last Glacial Maximum. *Paleoceanography*, 27, PA4208. <https://doi.org/10.1029/2012PA002378>
- Koutavas, A., & Sachs, J. P. (2008). Northern timing of deglaciation in the eastern equatorial Pacific from alkenone paleothermometry. *Paleoceanography*, 23, PA4205. <https://doi.org/10.1029/2008PA001593>
- Kozdon, R., Eisenhauer, A., Weinelt, M., Meland, M. Y., & Nürnberg, D. (2009). Reassessing Mg/Ca temperature calibrations of *Neoglobobulimina pachyderma* (sinistral) using paired $\delta^{44}\text{Ca}/^{40}\text{Ca}$ and Mg/Ca measurements. *Geochemistry, Geophysics, Geosystems*, 10, Q03005. <https://doi.org/10.1029/2008GC002169>
- Kozdon, R., Kelly, D. C., Kita, N. T., Fournelle, J. H., & Valley, J. W. (2011). Planktonic foraminiferal oxygen isotope analysis by ion microprobe technique suggests warm tropical sea surface temperatures during the Early Paleogene. *Paleoceanography*, 26, PA3206. <https://doi.org/10.1029/2010PA002056>
- Kruschke, J. K. (2015). *Doing Bayesian data analysis: a tutorial with R, JAGS, and Stan edition 2 ed*. Boston: Academic Press.
- Kucera, M., & Darling, K. F. (2002). Cryptic species of planktonic foraminifera: Their effect on paleoceanographic reconstructions, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 360(1793), 695–718. <https://doi.org/10.1098/rsta.2001.0962>
- Kuroyanagi, A., Kawahata, H., & Nishi, H. (2011). Seasonal variation in the oxygen isotopic composition of different-sized planktonic foraminifera *Neoglobobulimina pachyderma* (sinistral) in the northwestern North Pacific and implications for reconstruction of the paleoenvironment. *Paleoceanography*, 26, PA4215. <https://doi.org/10.1029/2011PA002153>
- LeGrande, A. N., & Schmidt, G. A. (2006). Global gridded data set of the oxygen isotopic composition in seawater. *Geophysical Research Letters*, 33, L12604. <https://doi.org/10.1029/2006GL026011>
- Lea, D. W., Pak, D. K., Belanger, C. L., Spero, H. J., Hall, M. A., & Shackleton, N. J. (2006). Paleoclimate history of Galápagos surface waters over the last 135,000yr. *Quaternary Science Reviews*, 25(11), 1152–1167. <https://doi.org/10.1016/j.quascirev.2005.11.010>
- Leduc, G., Vidal, L., Tachikawa, K., Rostek, F., Sonzogni, C., Beaufort, L., & Bard, E. (2007). Moisture transport across Central America as a positive feedback on abrupt climatic changes. *Nature*, 445(7130), 908–911. <https://doi.org/10.1038/nature05578>
- Linsley, B. K., Rosenthal, Y., & Oppo, D. W. (2010). Holocene evolution of the Indonesian throughflow and the western Pacific warm pool. *Nature Geoscience*, 3(8), 578–583. <https://doi.org/10.1038/ngeo920>
- Lisiecki, L. E., & Raymo, M. E. (2005). A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}\text{O}$ records: PLIOCENE-PLEISTOCENE BENTHIC STACK. *Paleoceanography*, 20, PA1003. <https://doi.org/10.1029/2004PA001071>
- MEDAR Group (2002). MEDATLAS/2002 database: Mediterranean and Black Sea database of temperature, salinity and bio-chemical parameters in climatological atlas [CD-ROM] (Tech. rep.) France: Fr. Res. Inst. for Exploit of Sea, Issy-les-Moulineaux.
- Martínez-Méndez, G., Zahn, R., Hall, I. R., Peeters, F. J. C., Pena, L. D., Cacho, I., & Negre, C. (2010). Contrasting multiproxy reconstructions of surface ocean hydrography in the Agulhas Corridor and implications for the Agulhas Leakage during the last 345,000 years. *Paleoceanography*, 25, PA4227. <https://doi.org/10.1029/2009PA001879>
- Mashiotta, T. A., Lea, D. W., & Spero, H. J. (1999). Glacial-interglacial changes in Subantarctic sea surface temperature and $\delta^{18}\text{O}$ -water using foraminiferal Mg. *Earth and Planetary Science Letters*, 170(4), 417–432. [https://doi.org/10.1016/S0012-821X\(99\)00116-8](https://doi.org/10.1016/S0012-821X(99)00116-8)
- Mathien-Blard, E., & Bassinot, F. (2009). Salinity bias on the foraminifera Mg/Ca thermometry: Correction procedure and implications for past ocean hydrographic reconstructions. *Geochemistry, Geophysics, Geosystems*, 10, Q12011. <https://doi.org/10.1029/2008GC002353>
- McElreath, R. (2016). *Statistical rethinking: A Bayesian course with examples in R and Stan*, no 122 in Chapman CRC texts in statistical science series. Boca Raton: CRC Press/Taylor & Francis Group.
- McKinney, W. J. (2010). Data structures for statistical computing in python. In *Proceedings of the 9th Python in Science Conference* Walt, S. v. d., Millman, pp. 51–56.
- Meland, M. Y., Jansen, E., Elderfield, H., Dokken, T. M., Olsen, A., & Bellerby, R. G. J. (2006). Mg/Ca ratios in the planktonic foraminifera *Neoglobobulimina pachyderma* (sinistral) in the northern North Atlantic/Nordic Seas. *Geosystems, Geochemistry, Geophysics*, 7, Q06P14. <https://doi.org/10.1029/2005GC001078>
- Met Office (2010). *Cartopy: A cartographic python library with a matplotlib interface*. Devon: Exeter.
- Moffa-Sánchez, P., Hall, I. R., Barker, S., Thornalley, D. J. R., & Yashayaev, I. (2014). Surface changes in the eastern Labrador Sea around the onset of the Little Ice Age. *Paleoceanography*, 29, 160–175. <https://doi.org/10.1002/2013PA002523>
- Mohtadi, M., Oppo, D. W., Lückge, A., DePol-Holz, R., Steinke, S., Groeneveld, J., et al. (2011). Reconstructing the thermal structure of the upper ocean: Insights from planktic foraminifera shell chemistry and alkenones in modern sediments of the tropical eastern Indian Ocean. *Paleoceanography*, 26, PA3219. <https://doi.org/10.1029/2011PA002132>
- Mohtadi, M., Steinke, S., Lückge, A., Groeneveld, J., & Hathorne, E. C. (2010). Glacial to Holocene surface hydrography of the tropical eastern Indian Ocean. *Earth and Planetary Science Letters*, 292(1), 89–97. <https://doi.org/10.1016/j.epsl.2010.01.024>

- Mulitza, S., Boltovskoy, D., Donner, B., Meggers, H., Paul, A., & Wefer, G. (2003a). Temperature- $\delta^{18}\text{O}$ relationships of planktonic foraminifera collected from surface waters. *Paleogeography, Paleoclimatology, Paleoecology*, 202(1), 143–152. [https://doi.org/10.1016/S0031-0182\(03\)00633-3](https://doi.org/10.1016/S0031-0182(03)00633-3)
- Mulitza, S., Donner, B., Fischer, G., Paul, A., Pätzold, J., Rühlemann, C., & Segl, M. (2003b). The South Atlantic oxygen isotope record of planktic foraminifera. In G. Wefer, S. Mulitza, & V. Ratmeyer (Eds.), *The South Atlantic in the Late Quaternary: Reconstruction of material budgets and current systems* (pp. 121–142). Berlin: Springer-Verlag.
- Mulitza, S., Wolff, T., Pätzold, J., Hale, W., & Wefer, G. (1998). Temperature sensitivity of planktic foraminifera and its influence on the oxygen isotope record. *Marine Micropaleontology*, 33(3), 223–240. [https://doi.org/10.1016/S0377-8398\(97\)00040-6](https://doi.org/10.1016/S0377-8398(97)00040-6)
- Nürnberg, D., Ziegler, M., Karas, C., Tiedemann, R., & Schmidt, M. W. (2008). Interacting loop current variability and Mississippi River discharge over the past 400kyr. *Earth and Planetary Science Letters*, 272(1), 278–289. <https://doi.org/10.1016/j.epsl.2008.04.051>
- O'Neil, J. R., Clayton, R. N., & Mayeda, T. K. (1969). Oxygen isotope fractionation in divalent metal carbonates. *The Journal of Chemical Physics*, 51(12), 5547–5558. <https://doi.org/10.1063/1.1671982>
- Oliphant, T. E. (2015). *Guide to NumPy*. Austin, Tex: Continuum Press.
- Oppo, D. W., Rosenthal, Y., & Linsley, B. K. (2009). 2,000-year-long temperature and hydrology reconstructions from the Indo-Pacific warm pool. *Nature*, 460(7259), 1113–1116. <https://doi.org/10.1038/nature08233>
- Oppo, D. W., & Sun, Y. (2005). Amplitude and timing of sea-surface temperature change in the northern South China Sea: Dynamic link to the East Asian monsoon. *Geology*, 33(10), 785–788. <https://doi.org/10.1130/G21867.1>
- Pahneke, K., Zahn, R., Elderfield, H., & Schulz, M. (2003). 340,000-year centennial-scale marine record of Southern Hemisphere climatic oscillation. *Science*, 301(5635), 948–952. <https://doi.org/10.1126/science.1084451>
- Palmer, M. R., & Pearson, P. N. (2003). A 23,000-year record of surface water pH and PCO_2 in the western equatorial Pacific Ocean. *Science*, 300(5618), 480–482. <https://doi.org/10.1126/science.1080796>
- Parker, A. O., Schmidt, M. W., Jobe, Z. R., & Slowey, N. C. (2016). A new perspective on West African hydroclimate during the last deglaciation. *Earth and Planetary Science Letters*, 449, 79–88. <https://doi.org/10.1016/j.epsl.2016.05.038>
- Piola, A. R., Matano, R. P., Palma, E. D., Möller, O. O., & Campos, E. J. D. (2005). The influence of the Plata River discharge on the western South Atlantic shelf. *Geophysical Research Letters*, 32, L01603. <https://doi.org/10.1029/2004GL021638>
- Praetorius, S. K., Mix, A. C., Walczak, M. H., Wolhowe, M. D., Addison, J. A., & Prahl, F. G. (2015). North Pacific deglacial hypoxic events linked to abrupt ocean warming. *Nature*, 527(7578), 362–366. <https://doi.org/10.1038/nature15753>
- Pujol, C., & Grazzini, C. V. (1995). Distribution patterns of live planktic foraminifera as related to regional hydrography and productive systems of the Mediterranean Sea. *Marine Micropaleontology*, 25(2), 187–217. [https://doi.org/10.1016/0377-8398\(95\)00002-I](https://doi.org/10.1016/0377-8398(95)00002-I)
- Ravelo, A., & Fairbanks, R. (1992). Oxygen Isotopic composition of multiple species of planktonic foraminifera: Recorders of the modern photic zone temperature gradient. *Paleoceanography*, 7(6), 815–831. <https://doi.org/10.1029/92PA02092>
- Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., et al. (2003). Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research*, 108(D14), 4407. <https://doi.org/10.1029/2002JD002670>
- Regenberg, M., Steph, S., Nürnberg, D., Tiedemann, R., & Garbe-Schönberg, D. (2009). Calibrating Mg/Ca ratios of multiple planktonic foraminiferal species with $\delta^{18}\text{O}$ -calcification temperatures: Paleothermometry for the upper water column. *Earth and Planetary Science Letters*, 278(3), 324–336. <https://doi.org/10.1016/j.epsl.2008.12.019>
- Reynolds, R. W., Rayner, N. A., Smith, T. M., Stokes, D. C., & Wang, W. (2002). An improved in situ and satellite SST analysis for climate. *Journal of Climate*, 15(13), 1609–1625. [https://doi.org/10.1175/1520-0442\(2002\)015<1609:AIHSAS>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<1609:AIHSAS>2.0.CO;2)
- Richey, J. N., Poore, R. Z., Flower, B. P., & Quinn, T. M. (2007). 1400 yr multiproxy record of climate variability from the northern Gulf of Mexico. *Geology*, 35(5), 423–426. <https://doi.org/10.1130/G23507A.1>
- Richey, J. N., Thirumalai, K., Khider, D., Reynolds, C. E., Partin, J. W., & Quinn, T. M. (2019). Considerations for *Globigerinoides ruber* (White and Pink) paleoceanography: Comprehensive insights from a long-running sediment trap. *Paleoceanography and Paleoclimatology*, 34(3), 353–373. <https://doi.org/10.1029/2018PA003417>
- Riethdorf, J.-R., Max, L., Nürnberg, D., Lembke-Jene, L., & Tiedemann, R. (2013). Deglacial development of (sub) sea surface temperature and salinity in the subarctic northwest Pacific: Implications for upper-ocean stratification. *Paleoceanography*, 28, 91–104. <https://doi.org/10.1002/palo.20014>
- Rigual-Hernández, A. S., Sierro, F. J., Bárcena, M. A., Flores, J. A., & Heussner, S. (2012). Seasonal and interannual changes of planktic foraminiferal fluxes in the Gulf of Lions (NW Mediterranean) and their implications for paleoceanographic studies: Two 12-year sediment trap records. *Deep Sea Research Part I: Oceanographic Research Papers*, 66, 26–40. <https://doi.org/10.1016/j.dsr.2012.03.011>
- Riveiros, N. V., Govin, A., Waelbroeck, C., Mackensen, A., Michel, E., Moreira, S., et al. (2016). Mg/Ca thermometry in planktic foraminifera: Improving paleotemperature estimations for *G. bulloides* and *N. pachyderma* left. *Geochemistry, Geophysics, Geosystems*, 17, 1249–1264. <https://doi.org/10.1002/2015GC006234>
- Romahn, S., Mackensen, A., Groeneveld, J., & Pätzold, J. (2014). Deglacial intermediate water reorganization: New evidence from the Indian Ocean. *Climate of the Past*, 10(1), 293–303. <https://doi.org/10.5194/cp-10-293-2014>
- Rosenthal, Y., Oppo, D. W., & Linsley, B. K. (2003). The amplitude and phasing of climate change during the last deglaciation in the Sulu Sea, western equatorial Pacific. *Geophysical Research Letters*, 30(8), 1428. <https://doi.org/10.1029/2002GL016612>
- Rustic, G. T., Koutavas, A., Marchitto, T. M., & Linsley, B. K. (2015). Dynamical excitation of the tropical Pacific Ocean and ENSO variability by Little Ice Age cooling. *Science*, 350(6267), 1537–1541. <https://doi.org/10.1126/science.aac9937>
- Sabbatini, A., Bassinot, F., Boussetta, S., Negri, A., Rebaubier, H., Dewilde, F., et al. (2011). Further constraints on the diagenetic influences and salinity effect on *Globigerinoides ruber* (white) Mg/Ca thermometry: Implications in the Mediterranean Sea. *Geochemistry, Geophysics, Geosystems*, 12, Q10005. <https://doi.org/10.1029/2011GC003675>
- Salvatier, J., Wiecki, T. V., & Fonnesebeck, C. (2016). Probabilistic programming in Python using PyMC3. *PeerJ Computer Science*, 2, e55. <https://doi.org/10.7717/peerj-cs.55>
- Saraswat, R., Lea, D. W., Nigam, R., Mackensen, A., & Naik, D. K. (2013). Deglaciation in the tropical Indian Ocean driven by interplay between the regional monsoon and global teleconnections. *Earth and Planetary Science Letters*, 375, 166–175. <https://doi.org/10.1016/j.epsl.2013.05.022>
- Sautter, L. R., & Thunell, R. C. (1989). Seasonal succession of planktonic foraminifera: Results from a four-year time-series sediment trap experiment in the Northeast Pacific. *Journal of Foraminiferal Research*, 19(4), 253–267. <https://doi.org/10.2113/gsjfr.19.4.253>
- Sautter, L. R., & Thunell, R. C. (1991). Planktonic foraminiferal response to upwelling and seasonal hydrographic conditions: Sediment trap results from San Pedro Basin, Southern California Bight. *Journal of Foraminiferal Research*, 21(4), 347–363. <https://doi.org/10.2113/gsjfr.21.4.347>

- Schefuß, E., Schouten, S., & Schneider, R. R. (2005). Climatic controls on central African hydrology during the past 20,000 years. *Nature*, 437(7061), 1003–1006. <https://doi.org/10.1038/nature03945>
- Schmidt, M. W., Chang, P., Hertzberg, J. E., Them, T. R., Ji, L., & Otto-Bliesner, B. L. (2012a). Impact of abrupt deglacial climate change on tropical Atlantic subsurface temperatures. *Proceedings of the National Academy of Sciences*, 109(36), 14,348–14,352. <https://doi.org/10.1073/pnas.1207806109>
- Schmidt, G. A., & Mulitza, S. (2002). Global calibration of ecological models for planktic foraminifera from coretop carbonate oxygen-18. *Marine Micropaleontology*, 44(3), 125–140. [https://doi.org/10.1016/S0377-8398\(01\)00041-X](https://doi.org/10.1016/S0377-8398(01)00041-X)
- Schmidt, M. W., Spero, H. J., & Lea, D. W. (2004). Links between salinity variation in the Caribbean and North Atlantic thermohaline circulation. *Nature*, 428(6979), 160–163. <https://doi.org/10.1038/nature02346>
- Schmidt, M. W., Weinlein, W. A., Marcantonio, F., & Lynch-Stieglitz, J. (2012b). Solar forcing of Florida Straits surface salinity during the early Holocene. *Paleoceanography*, 27, PA3204. <https://doi.org/10.1029/2012PA002284>
- Schrag, D. P., DePaolo, D. J., & Richter, F. M. (1995). Reconstructing past sea surface temperatures: Correcting for diagenesis of bulk marine carbonate. *Geochimica et Cosmochimica Acta*, 59(11), 2265–2278. [https://doi.org/10.1016/0016-7037\(95\)00105-9](https://doi.org/10.1016/0016-7037(95)00105-9)
- Schrag, D. P., Hampt, G., & Murray, D. W. (1996). Pore fluid constraints on the temperature and oxygen isotopic composition of the glacial ocean. *Science*, 272(5270), 1930–1932. <https://doi.org/10.1126/science.272.5270.1930>
- Shackleton, N. J. (1974). Attainment of isotopic equilibrium between ocean water and the benthonic foraminifera genus *Uvigerina*: Isotopic changes in the ocean during the last glacial. *Colloque international du CNRS*, 219, 203–209.
- Sluijs, A., Brinkhuis, H., Schouten, S., Bohaty, S. M., John, C. M., Zachos, J. C., et al. (2007). Environmental precursors to rapid light carbon injection at the Palaeocene/Eocene boundary. *Nature*, 450(7173), 1218–1221. <https://doi.org/10.1038/nature06400>
- Spero, H. J., Bijma, J., Lea, D. W., & Bemis, B. E. (1997). Effect of seawater carbonate concentration on foraminiferal carbon and oxygen isotopes. *Nature*, 390(6659), 497–500. <https://doi.org/10.1038/37333>
- Spero, H. J., & Lea, D. W. (1993). Intraspecific stable isotope variability in the planktic foraminifera *Globigerinoides sacculifer*: Results from laboratory experiments. *Marine Micropaleontology*, 22(3), 221–234. [https://doi.org/10.1016/0377-8398\(93\)90045-Y](https://doi.org/10.1016/0377-8398(93)90045-Y)
- Spero, H. J., & Lea, D. W. (1996). Experimental determination of stable isotope variability in *Globigerina bulloides*: Implications for paleoceanographic reconstructions. *Marine Micropaleontology*, 28(3), 231–246. [https://doi.org/10.1016/0377-8398\(96\)00003-5](https://doi.org/10.1016/0377-8398(96)00003-5)
- Spero, H. J., Mielke, K. M., Kalve, E. M., Lea, D. W., & Pak, D. K. (2003). Multispecies approach to reconstructing eastern equatorial Pacific thermocline hydrography during the past 360 kyr. *Paleoceanography*, 18(1), 1022. <https://doi.org/10.1029/2002PA000814>
- Steinke, S., Chiu, H.-Y., Yu, P.-S., Shen, C.-C., Löwemark, L., Mii, H.-S., & Chen, M.-T. (2005). Mg/Ca ratios of two *Globigerinoides ruber* (white) morphotypes: Implications for reconstructing past tropical/subtropical surface water conditions. *Geochemistry, Geophysics, Geosystems*, 6, Q11005. <https://doi.org/10.1029/2005GC000926>
- Steinke, S., Kienast, M., Groeneveld, J., Lin, L.-C., Chen, M.-T., & Rendle-Bühning, R. (2008). Proxy dependence of the temporal pattern of deglacial warming in the tropical South China Sea: Toward resolving seasonality. *Quaternary Science Reviews*, 27(7), 688–700. <https://doi.org/10.1016/j.quascirev.2007.12.003>
- Steph, S., Regenberg, M., Tiedemann, R., Mulitza, S., & Nürnberg, D. (2009). Stable isotopes of planktonic foraminifera from tropical Atlantic/Caribbean core-tops: Implications for reconstructing upper ocean stratification. *Marine Micropaleontology*, 71(1), 1–19. <https://doi.org/10.1016/j.marmicro.2008.12.004>
- Stott, L., Timmermann, A., & Thunell, R. (2007). Southern Hemisphere and deep-sea warming led deglacial atmospheric CO₂ rise and tropical warming. *Science*, 318(5849), 435–438. <https://doi.org/10.1126/science.1143791>
- Sun, Y., Oppo, D. W., Xiang, R., Liu, W., & Gao, S. (2005). Last deglaciation in the Okinawa Trough: Subtropical northwest Pacific link to Northern Hemisphere and tropical climate. *Paleoceanography*, 20, PA4005. <https://doi.org/10.1029/2004PA001061>
- Thornalley, D. J. R., Elderfield, H., & McCave, I. N. (2011). Reconstructing North Atlantic deglacial surface hydrography and its link to the Atlantic overturning circulation. *Global and Planetary Change*, 79(3), 163–175. <https://doi.org/10.1016/j.gloplacha.2010.06.003>
- Thunell, R. C., Curry, W. B., & Honjo, S. (1983). Seasonal variation in the flux of planktonic foraminifera: Time series sediment trap results from the Panama Basin. *Earth and Planetary Science Letters*, 64(1), 44–55. [https://doi.org/10.1016/0012-821X\(83\)90051-1](https://doi.org/10.1016/0012-821X(83)90051-1)
- Tierney, J. E., Pausata, F. S. R., & deMenocal, P. (2016). Deglacial Indian monsoon failure and North Atlantic stadials linked by Indian Ocean surface cooling. *Nature Geoscience*, 9(1), 46–50. <https://doi.org/10.1038/ngeo2603>
- Tierney, J. E., & Tingley, M. P. (2014). A Bayesian, spatially-varying calibration model for the TEX86 proxy. *Geochimica et Cosmochimica Acta*, 127, 83–106. <https://doi.org/10.1016/j.gca.2013.11.026>
- Tierney, J. E., & Tingley, M. P. (2018). BAYSPLINE: A new calibration for the alkenone paleothermometer. *Paleoceanography and Paleoclimatology*, 33(3), 281–301. <https://doi.org/10.1002/2017PA003201>
- Tindall, J., Flecker, R., Valdes, P., Schmidt, D. N., Markwick, P., & Harris, J. (2010). Modelling the oxygen isotope distribution of ancient seawater using a coupled ocean-atmosphere GCM: Implications for reconstructing early Eocene climate. *Earth and Planetary Science Letters*, 292(3), 265–273. <https://doi.org/10.1016/j.epsl.2009.12.049>
- Urey, H. C. (1947). The thermodynamic properties of isotopic substances. *Journal of the Chemical Society (Resumed)*, 0(0), 562–581. <https://doi.org/10.1039/JR9470000562>
- Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5), 1413–1432. <https://doi.org/10.1007/s11222-016-9696-4>
- Visser, K., Thunell, R., & Stott, L. (2003). Magnitude and timing of temperature change in the Indo-Pacific warm pool during deglaciation. *Nature*, 421(6919), 152–155. <https://doi.org/10.1038/nature01297>
- Waelbroeck, C., Mulitza, S., Spero, H., Dokken, T., Kiefer, T., & Cortijo, E. (2005). A global compilation of late Holocene planktonic foraminiferal $\delta^{18}\text{O}$: Relationship between surface water temperature and $\delta^{18}\text{O}$. *Quaternary Science Reviews*, 24(7–9), 853–868. <https://doi.org/10.1016/j.quascirev.2003.10.014>
- Weldeab, S., Lea, D. W., Oberhänsli, H., & Schneider, R. R. (2014). Links between southwestern tropical Indian Ocean SST and precipitation over southeastern Africa over the last 17kyr. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 410, 200–212. <https://doi.org/10.1016/j.palaeo.2014.06.001>
- Weldeab, S., Lea, D. W., Schneider, R. R., & Andersen, N. (2007). 155,000 Years of West African monsoon and ocean thermal evolution. *Science*, 316(5829), 1303–1307. <https://doi.org/10.1126/science.1140461>
- Weldeab, S., Schneider, R. R., & Kölling, M. (2006). Deglacial sea surface temperature and salinity increase in the western tropical Atlantic in synchrony with high latitude climate instabilities. *Earth and Planetary Science Letters*, 241(3), 699–706. <https://doi.org/10.1016/j.epsl.2005.11.012>
- Weldeab, S., Schneider, R. R., Kölling, M., & Wefer, G. (2005). Holocene African droughts relate to eastern equatorial Atlantic cooling. *Geology*, 33(12), 981–984. <https://doi.org/10.1130/G21874.1>

- Werner, K., Spielhagen, R. F., Bauch, D., Hass, H. C., & Kandiano, E. (2013). Atlantic water advection versus sea-ice advances in the eastern Fram Strait during the last 9 ka: Multiproxy evidence for a two-phase Holocene. *Paleoceanography*, 28, 283–295. <https://doi.org/10.1002/palo.20028>
- Williams, D. F., Bé, A. W. H., & Fairbanks, R. G. (1979). Seasonal oxygen isotopic variations in living planktonic foraminifera off Bermuda. *Science*, 206(4417), 447–449. <https://doi.org/10.1126/science.206.4417.447>
- Williams, D. F., Bé, A. W. H., & Fairbanks, R. G. (1981). Seasonal stable isotopic variations in living planktonic foraminifera from Bermuda plankton tows. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 33(1), 71–102. [https://doi.org/10.1016/0031-0182\(81\)90033-X](https://doi.org/10.1016/0031-0182(81)90033-X)
- Xu, J., Kuhnt, W., Holbourn, A., Regenberg, M., & Andersen, N. (2010). Indo-Pacific warm pool variability during the Holocene and Last Glacial Maximum. *Paleoceanography*, 25, PA4230. <https://doi.org/10.1029/2010PA001934>
- Zachos, J., Stott, L., & Lohmann, K. (1994). Evolution of Early Cenozoic marine temperatures. *Paleoceanography*, 9(2), 353–387. <https://doi.org/10.1029/93PA03266>
- Zeebe, R. E. (1999). An explanation of the effect of seawater carbonate concentration on foraminiferal oxygen isotopes. *Geochimica et Cosmochimica Acta*, 63(13–14), 2001–2007. [https://doi.org/10.1016/S0016-7037\(99\)00091-5](https://doi.org/10.1016/S0016-7037(99)00091-5)
- Žarić, S., Donner, B., Fischer, G., Mulitza, S., & Wefer, G. (2005). Sensitivity of planktic foraminifera to sea surface temperature and export production as derived from sediment trap data. *Marine Micropaleontology*, 55(1), 75–105. <https://doi.org/10.1016/j.marmicro.2005.01.002>