Corroborating Ecological Depth Preferences of
Planktonic Foraminifera in the Tropical Atlantic
With the Stable Oxygen Isotope Ratios of Coretop Specimens

E. Christa Farmer\textsuperscript{1,2}, Alexey Kaplan\textsuperscript{1}, Peter B. deMenocal\textsuperscript{1}, Jean Lynch-Stieglitz\textsuperscript{1,3}

\textsuperscript{1}Lamont–Doherty Earth Observatory of Columbia University, Palisades, NY 10964
\textsuperscript{2}Now at: Hofstra University, Geology Department, 145 Gittleson, Hempstead, NY 11549-1140 (geoecf@hofstra.edu)
\textsuperscript{3}Now at: School of Earth and Atmospheric Sciences, 311 Ferst Drive, Georgia Institute of Technology, Atlanta, GA 30332
Abstract. Past variability in upper ocean thermocline depth is commonly estimated from the abundance of different species of planktonic organisms or the difference in oxygen isotopic composition between two species of planktonic foraminifera, one that lives in the mixed layer and one that lives in or near the thermocline. To test the latter relationships, we measured the oxygen isotopic composition of eight species of planktonic foraminifera (pink and white varieties of *Globigerinoides ruber*, *Globigerinoides sacculifer* without the final chamber, *Orbulina universa*, *Pulleniatina obliquiloculata*, *Globoortalia menardii*, *Neogloborotalia dutertrei*, and *Globoortalia tumida*) in surface sediment samples from 31 tropical Atlantic deep-sea sediment cores. Bayesian analysis was used to compare measured oxygen isotopic compositions with their predictions based on modern data sets of annual temperatures and oxygen isotopic composition of ocean water in the upper 500m at the core sites. Posterior probability densities for predictive model parameters were computed. Probability distributions of calcification depth for analyzed species corroborated their ecological preferences inferred from net tow and sediment trap data. Robustness of the habitat signals in coretop specimens suggests that reconstructions of the entire upper ocean temperature profiles, not just their thermocline depth or temperature, might be possible.
1. Introduction

Study of the variation in water column depth preference among various species of planktonic foraminifera was pioneered by Bé and Tolderlund (1971), and furthered by Fairbanks et al. (1980) using MOCNESS tows in the western North Atlantic. *Globigerinoides ruber* (both pink and white varieties) and *Globigerinoides sacculifer* (without final chamber) species were most common in the surface mixed layer, whereas *Globorotalia menardii* and *Neogloborotalia dutertrei* were most abundant at the thermocline. *Orbulina universa* was often most abundant in the mixed layer but also sometimes most abundant in the thermocline (hence its “universal” name). Below the photic zone, at depths up to several hundred meters, the most abundant species were *Globorotalia truncatulinoides*, *Globorotalia crassiformis*, and *Globorotalia tumida*. Oxygen isotope ratios ($\delta^{18}O$) predicted from hydrography at the depth at which each species was found living matched well with isotopic values measured in coretop samples of each species (Fairbanks and Wiebe, 1980; Ravelo and Fairbanks, 1992).

Deuser and Ross (1989) collected planktonic foraminiferal samples from sediment traps in the Sargasso Sea and compared their $\delta^{18}O$ with hydrography as well. They estimated that *G.ruber* (pink variant) represented conditions in the surface waters, that *Pulleniatina obliquiloculata*, *Globorotalia inflata*, and *N. dutertrei* represented conditions in the winter mixed layer (down to 100m water depth), while *Globigerinoides conglobatus* represented the depth interval of 75–100m during the fall season. Mulitza et al. (1997) proposed that past changes in upper ocean thermal stratification could be estimated based on the calculated $\delta^{18}O$ gradient between surface dwelling species (*G.sacculifer* and *N.pachyderma,*...
right-coiling variant) and deeper-dwelling *G. truncatulinoides* which appeared to calcify near 250m. Based on these results and some further data for *G. ruber* (pink variant), Mulitza et al. (1998) proposed a theoretical model by which the temperature at various water depths can be “triangulated” from at least three species of planktonic foraminifera.

Dekens et al. (2002) looked at both calcite $^\delta$18O (hereinafter $^\delta$18Oc) and Mg/Ca values from coretop sediment samples in the tropics of all oceans, confirming that *G. ruber* record surface temperatures, *G. sacculifer* record temperatures at 20-30m, and the depth habitat of *N. dutertrei* is more broad but seems to be about 50m. Anand et al. (2003) also measured both $^\delta$18Oc and Mg/Ca from sediment trap samples, reaching similar conclusions. LeGrande et al. (2004) took a slightly different approach, measuring calcite $^\delta$18Oc of *Globorotalia truncatulinoides* and comparing it to upper ocean water $^\delta$18O (hereinafter $^\delta$18Ow) and temperature. The best match indicated that the species either calcified at 350m, or calcified at the surface and added post-gametogenic calcite at 800m. Schmidt and Mulitza (2002) used a Monte Carlo minimization procedure to match foraminiferal $^\delta$18Oc for six mixed layer species from a global data set of coretops with data sets of seawater $^\delta$18Ow and temperatures. They produced a description of temperature ranges and other ecological conditions preferred by these species.

Here we investigate whether the tropical Atlantic habitat depth preferences of eight species of planktonic foraminifera are consistent with the calcite $^\delta$18Oc of coretop specimens. Using a Bayesian analysis approach, measurements of $^\delta$18Oc in these species are compared with expected values based on published calibration equations and analyzed data sets of ocean water temperatures (Conkright et al., 2001) and $^\delta$18Ow (LeGrande and Schmidt, 2006). Our newly produced $^\delta$18Oc data set is smaller in size, but also smaller
in the geographical extent than that used by Schmidt and Mulitza (2002). Consequently, we can use calcification depth as a main factor defining the habitat of a given species, and allow, within some limits, species-dependent coefficients in their $\delta^{18}O_c$-to-temperature relationship.

2. Methods

2.1. Study Area

In order to test the relationship between ocean water properties and the isotopic oxygen composition of multiple species of calciferous planktonic foraminifera, this study uses modern sea floor sediment samples from the tropical Atlantic. The oxygen isotope ratio of foraminiferal calcite, $\delta^{18}O_c$, is a function of temperature and the oxygen isotope ratio of seawater, $\delta^{18}O_w$ (Berger 1981; Craig and Gordon, 1965). Tropical Atlantic surface salinity variations have little effect on $\delta^{18}O_c$ of foraminifera formed in the mixed layer, because rainfall $\delta^{18}O$ is very close to that of the seawater. Changes in the water balance that affect salinity do not change the $\delta^{18}O_w$ of surface waters as much as in other parts of the Atlantic Ocean or in other ocean basins (Schmidt, 1999; Schmidt et al., 1999; LeGrande and Schmidt, 2006). Deeper waters, however, reflect the salinity and $\delta^{18}O_w$ of higher latitudes (Sarmiento et al. 2004), hence these parameters influence the $\delta^{18}O_c$ of foraminifera calcifying below the mixed layer.

Samples were collected from the top centimeter of 31 cores in a transect from the Caribbean to the Gulf of Guinea (Figure 1 and Table 1). The cores were selected from the Lamont-Doherty Earth Observatory Deep-Sea Sample Repository based on their full representation of the thermocline depth gradient in the tropical Atlantic (Figure 2a). Sediments from the tops of these cores are expected to be modern or of late Holocene
age based on stratigraphic information, like the presence or absence of faunal markers. Additionally, only cores collected at water depths less than 4150m were included (see Table 1), in order to avoid the influence of CaCO$_3$ dissolution below the tropical Atlantic lysocline (Broecker and Takahashi 1978; Lohmann 1995).

2.2. Sample Preparation

Sediment samples were shaken in a sodium metaphosphate surfactant solution for about two hours, washed through 150µm sieves with deionized water, and dried in a 50°C oven. The fine (<150µm) fraction was archived, and the coarse fraction was sieved in narrow size fractions to minimize the known ontogenetic fractionation effects of foraminiferal growth and size. The size ranges differed between species, but were selected to balance between the ideal size range where the growth effect is minimal and the size range in which the species is abundant (see Table 2). Between 10-20 individuals per species were picked from each sample with a red sable brush. The δ$^{18}$O$_c$ of all foraminiferal species from each core were measured using a Micromass Optima mass spectrometer with a Multiprep individual acid bath carbonate preparation device. Measurement precision is estimated from the 1-σ reproducibility of multiple measurements of a known standard in each run, averaged over all runs: 0.02 per mil for δ$^{13}$C and 0.06 per mil for δ$^{18}$O.

2.3. Instrumental data and calibration relationships

Foraminiferal calcite oxygen isotopic ratios (δ$^{18}$O$_c$) for each species from all cores were compared to the values predicted on the basis of existing analyses of mean water temperature and δ$^{18}$O$_w$. Temperature data was taken from the 2001 NOAA World Ocean Atlas (WOA2001, Conkright et al., 2001), accessed at
Mean temperatures, objectively analyzed, were used in the 14 top water depth intervals: 0, 10, 20, 30, 50, 75, 100, 125, 150, 200, 250, 300, 400, and 500m below the sea surface. Temperature values were interpolated bilinearly from a $1^\circ \times 1^\circ$ WOA2001 spatial grid to the core locations. Mean seawater oxygen isotope ratios ($\delta^{18}O_w$) were taken from a recently produced three-dimensional gridded data set by LeGrande and Schmidt (2006), accessed at http://data.giss.nasa.gov/o18data/grid.html. Since the grids of this data set match those of WOA2001, our processing of $\delta^{18}O_w$ data was identical to that of water temperatures. Vertical profiles of water temperatures and $\delta^{18}O_w$ at core locations are shown in Figure 2.

To date most foraminifera species used in our analysis do not have calibrations developed specifically for them. Therefore, we inform our interpretation of $\delta^{18}O_c$ measurements by commonly used calibration equations for the planktonic foraminifera: those developed by Bemis et al. (1998) for the culture experiments with $O. universa$ and $G. sacculifer$, and those reported by Mulitza et al. (2003) based on samples of living $G. ruber$ and $G. sacculifer$ pumped from surface waters of the Atlantic. These calibrations all express water temperature $T$ via a difference in oxygen isotope ratios between calcite and water, $\delta^{18}O_c - \delta^{18}O_w$:

$$T = 14.9 - 4.8 \times (\delta^{18}O_c - \delta^{18}O_w)$$  \hspace{1cm} (1)

was obtained in culture experiments with $O. universa$ under high light conditions (Bemis et al. 1998):

$$T = 16.5 - 4.8 \times (\delta^{18}O_c - \delta^{18}O_w)$$  \hspace{1cm} (2)
from culture experiments with *O. universa* under low light conditions (Bemis et al., 1998);

\[
T = 14.2 - 4.44 \times (\delta^{18}O_c - \delta^{18}O_w)
\]

(3)

from *G. ruber* specimens pumped from the surface ocean (Mulitza et al. 2003);

\[
T = 14.91 - 4.35 \times (\delta^{18}O_c - \delta^{18}O_w)
\]

(4)

from *G. sacculifer* specimens pumped from the surface ocean (Mulitza et al. 2003); and

\[
T = 13.2 - 4.89 \times (\delta^{18}O_c - \delta^{18}O_w)
\]

(5)

from culture experiments with *G. bulloides* (12-chambered) (Bemis et al. 1998).

2.4. Statistical analysis

Our goal is to analyze the measured values of \(\delta^{18}O_c\) to determine the likely depths in the upper water column at which each species of foraminifera was calcifying. All calibrations (1)-(5) specify the relationship between temperature \(T\), oxygen isotope ratio in calcite \(\delta^{18}O_c\), and that in the seawater \(\delta^{18}O_w\) in the form

\[
T = \alpha - \beta (\delta^{18}O_c - \delta^{18}O_w),
\]

for certain values of \(\alpha\) and \(\beta\). Temperatures \(T\) and seawater oxygen isotope ratios \(\delta^{18}O_w\) in this relationship should be taken at some species-specific calcification depth \(z\), to be determined. Importance of species-specific calibration equations for temperature reconstruction efforts having been demonstrated (Bemis et al. 2002), we allow coefficients \(\alpha\) and \(\beta\) to be different for different species. Therefore, we assume that the relationship

\[
T(z) = \alpha - \beta [\delta^{18}O_c - \delta^{18}O_w(z)]
\]

(6)

holds, within some error bounds, for each species for certain species-dependent values of
\(\alpha\), \(\beta\), and \(z\). With an additional assumption that mean vertical profiles of \(T\) and \(\delta^{18}O_w\) given by the present-time hydrographic data sets by Conkright et al. (2001) and LeGrande and Schmidt (2006), respectively, are applicable to the period when the foraminifera from the analyzed coretops were calcifying, we consider \(T(z)\) and \(\delta^{18}O_w(z)\) at all core locations to be known functions of \(z\). With \(\delta^{18}O_c\) measured, we can put a problem of estimating \(z\), \(\alpha\), and \(\beta\) into the Bayesian framework (Gelman et al., 2004).

If values of \(\alpha\) and \(\beta\) were known, and as well as the calcification depth \(z\), we could invert (6) to predict measured values of \(\delta^{18}O_c\) as

\[
\delta^{18}O_c \approx \frac{\alpha}{\beta} - \frac{1}{\beta} T(z) + \delta^{18}O_w(z).
\] (7)

This model is approximate, because there are measurement errors, errors associated with a calibration relationship, and errors due to the difference between the actual values of \(T(z)\) and \(\delta^{18}O_w(z)\) during calcification and their values that were obtained from the modern data sets, as well as the error in the paradigm that foraminifera calcified at a single depth.

Assuming the total error in the relationship (7) to be normally distributed with zero mean and variance \(\sigma^2\), we can write down the sampling distribution, i.e. the distribution of a measured value \(\delta^{18}O_c\) conditional on all other parameters, as

\[
\delta^{18}O_c|\alpha, \beta, z, \sigma^2 \sim N\left(\frac{\alpha}{\beta} - \frac{1}{\beta} T(z) + \delta^{18}O_w(z), \sigma^2\right),
\]

or, in terms of its probability density function,

\[
p(\delta^{18}O_c|\alpha, \beta, z, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} \exp \left\{ -\frac{1}{2\sigma^2} \left[ \delta^{18}O_c - \delta^{18}O_w(z) + \frac{1}{\beta} T(z) - \frac{\alpha}{\beta} \right]^2 \right\}.
\]

For measurements \(\{\delta^{18}O_c\} = \{\delta^{18}O_{c1}, \delta^{18}O_{c2}, \ldots, \delta^{18}O_{cN}\}\) made on \(N\) cores, the joint sampling distribution will be the product of individual ones:

\[
p(\{\delta^{18}O_c\}|\alpha, \beta, z, \sigma^2) = \frac{1}{(2\pi)^{N/2}\sigma^N} \exp \left( -\frac{1}{2\sigma^2} \sum_{i=1}^N Y_i^2 \right),
\] (8)
where

\[ Y_i = \delta^{18}O_{ci} - \delta^{18}O_w(z) + \frac{1}{\beta}T_i(z) - \frac{\alpha}{\beta}, \quad i = 1, \ldots, N, \]  

(9)

are differences between measured and predicted values of $\delta^{18}O_c$ for individual cores.

Bayesian data analysis utilizes measurements by inverting predictive distribution (8) to obtain a “posterior” distribution for model parameters, i.e. their distribution conditional on all measured values:

\[
p(\alpha, \beta, z, \sigma^2|\{\delta^{18}O_c\}) = \frac{p(\{\delta^{18}O_c\}|\alpha, \beta, z, \sigma^2)p(\alpha, \beta, z, \sigma^2)}{p(\{\delta^{18}O_c\})}.
\]

(10)

Here $p(\alpha, \beta, z, \sigma^2)$ is a “prior” distribution of these parameters, i.e. our assumption about them, made before any measurements became available;

\[
p(\{\delta^{18}O_c\}) = \int \int \int p(\{\delta^{18}O_c\}|\alpha, \beta, z, \sigma^2)p(\alpha, \beta, z, \sigma^2)d\alpha d\beta dz d\sigma^2
\]

is just a normalizing factor, ensuring that $p(\alpha, \beta, z, \sigma^2|\{\delta^{18}O_c\})$ defined by (10) integrates to 1 over its joint domain of $\alpha$, $\beta$, $z$, and $\sigma^2$.

Before any information about measurements of $\delta^{18}O_c$ is given, we can expect $\alpha$ and $\beta$ to take with equal probability any values in their range known from published calibrations. We take these ranges of applicable values from equations (1)-(5) and then extend them by approximately two typical standard deviations of error in the published estimates of $\alpha$ and $\beta$, i.e. 0.3°C and 0.2°C/per mil, respectively (Bemis et al., 1998). Therefore, we take prior probability distributions for $\alpha$ and $\beta$ to be uniform in the following ranges:

\[
\alpha \sim U(12.5^\circ C, 17^\circ C), \quad \beta \sim U(3.9^\circ C/\text{per mil}, 5.3^\circ C/\text{per mil}).
\]

Similarly, a natural choice of the prior distribution for $z$ is a uniform distribution on an interval which is wide enough to contain the actual calcification depths of species that
we analyze. Since net tow data indicated that all species analyzed here live within the interval between surface and 500m, we use as a prior distribution

\[ z \sim U(0m, 500m). \]

A standard approach to selecting a non-informative prior distribution for a variance parameter like \( \sigma^2 \) is to use a uniform distribution for its logarithm, rather than the parameter itself (Gelman et al., 2004). The interval between 0.05 per mil and 1 per mil is wide enough to contain the standard deviation of the expected error in any useful prediction, thus we use a prior distribution \( U(\ln 0.05, \ln 1) \) for \( \ln \sigma \). Therefore,

\[
p(\alpha, \beta, z, \sigma^2) = U(12.5^\circ C, 17^\circ C)U(3.9^\circ C/\text{per mil}, 5.3^\circ C/\text{per mil}) \times \\
U(z|0\text{m}, 500\text{m})U(\ln \sigma|\ln 0.05, \ln 1)/2\sigma^2. \tag{11}
\]

(Note that variable changes in probability density functions between \( \sigma^2, \sigma \) and \( \ln \sigma \) are made according to the rules \( p(\cdots, \sigma^2) = p(\cdots, \ln \sigma)/2\sigma^2, p(\cdots, \sigma) = p(\cdots, \ln \sigma)/\sigma \).)

Combining formulas (8),(10), and (11), we obtain

\[
p(\alpha, \beta, z, \sigma^2|\{\delta^{18}O_c\}) = C \times \begin{cases} 
\exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{N} Y_i^2\right)/\sigma^{N+2}, & \text{if } 12.5 < \alpha < 17, \\
3.9 < \beta < 5.3, \\
0 < z < 500, \\
0.05 < \sigma < 1, \\
0, & \text{otherwise},
\end{cases} \tag{12}
\]

where \( Y_i \) are defined by (9), and a normalizing factor \( C \) is determined by the condition

\[
\int \int \int \int p(\alpha, \beta, z, \sigma^2|\{\delta^{18}O_c\})d\alpha d\beta dz d\sigma^2 = 1. \tag{13}
\]

The calculations were performed in Matlab. Values of the function

\[
\exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{N} Y_i^2\right)/\sigma^{N+2}
\]
were tabulated on a grid covering the four-dimensional domain of $\alpha$, $\beta$, $z$, and $\ln \sigma$ in which the probability density $p(\alpha, \beta, z, \sigma^2|\delta^{18}O_{c1}, \delta^{18}O_{c2}, \cdots, \delta^{18}O_{cN})$ is non-zero, as defined by (12). Uniform grids of 40 points were used for $\alpha$, $\beta$, and $\ln \sigma$, and the 14 grid points from WOA2001 were used for $z$. Normalizing factor $C$ was computed from (13) using numerical integration, thus the full joint posterior probability density function (12) for all parameters became available. Marginal distributions were computed by further numerical integration:

$$p(\alpha, \beta) = \int \int p(\alpha, \beta, z, \sigma^2|\delta^{18}O_c)dzd\sigma^2, \quad p(\alpha) = \int p(\alpha, \beta)d\beta, \quad p(\beta) = \int p(\alpha, \beta)d\alpha,$$

$$p(z) = \int \int p(\alpha, \beta, z, \sigma^2|\delta^{18}O_c)d\alpha d\beta d\sigma^2,$$

$$p(\sigma) = 2\sigma \int \int \int p(\alpha, \beta, z, \sigma^2|\delta^{18}O_c)d\alpha d\beta dz.$$  (14)

These distributions for parameter values were then used to compute their means and confidence intervals.

Prediction of the values that depend on the estimated parameters are done by the integration over the entire parameter space, i.e. we compute

$$\langle \delta^{18}O_c \rangle = \int \int \left[ \frac{\alpha}{\beta} \frac{1}{\beta}T(z) + \delta^{18}O_w(z) \right] p(\alpha, \beta, z)d\alpha d\beta dz,$$  (15)

$$\langle T \rangle = \int \int \left\{ \alpha - \beta[\delta^{18}O_c - \delta^{18}O_w(z)] \right\} p(\alpha, \beta, z)d\alpha d\beta dz$$  (16)

for predictions of $\delta^{18}O_c$ and ocean temperature at the calcification depth. Uncertainties in these predictions are computed using these integrals and similar ones but with the predictive function squared:

$$\sigma_T^2 = \langle T^2 \rangle - \langle T \rangle^2$$

for the standard deviation $\sigma_T$ of error in predictions of $T$, etc.
3. Results

Eight planktonic foraminifer species were found in more than 20 cores and thus were sufficiently abundant for the statistical analysis. These were: *G. ruber* (both pink and white variants) and *G. sacculifer* (without a final sac-like chamber) from the 355-425µm size fraction; *O. universa*, *P. obliquiloculata*, and *G. menardii* from the 600-710µm size fraction; and *N. dutertrei*, and *G. tumida* from the 500-600µm size fraction. Data are presented in Table 1, and size fractions are listed in Table 2.

3.1. Parameter Estimates

The analysis method described above was applied to all measurements presented in Table 1, to each species separately. Figure 3 illustrates the analysis results for *G. ruber* (white) by presenting marginal posterior probability density functions (PDF) of all parameters, computed by equations (14). For easier interpretation, isolines of joint PDF for \((\alpha, \beta)\) pairs (Figure 3a) are labeled not by the probability density value but by the probability with which this value is exceeded. For example, the contour marked “95%” surrounds the area of \((\alpha, \beta)\) values with cumulative probability of 0.95. Areas of high probability reach limits of the ranges set by prior distributions for both \(\alpha\) and \(\beta\) values, suggesting that these parameters are not well-constrained by the data in this analysis. Indeed, the individual marginal PDF for \(\beta\) (Figure 3b) reaches its maximum at the high limit of the range set by the prior distribution and remains high in the most of the range. The marginal PDF for \(\alpha\) (Figure 3c) shows that it is only slightly better constrained than \(\beta\). However, PDFs for the calcification depth \(z\) and calibration error \(\sigma\) (Figures 3de) have large probability density values concentrated on relatively small segments of the prior range, hence these parameters are well-constrained.
The tendency of $z$ and $\sigma$ to be better constrained than $\alpha$ and $\beta$ parameters holds for all species. This can be concluded from inspection of Table 3, which provides expected values (means of posterior distributions) and 95% confidence intervals of parameters for analyses of all species. This outcome is particularly important for our analysis of the calcification depth of different species. Figure 4 puts together their posterior PDFs for $z$, creating a striking pattern of species’ inferred distribution in the water column which is quite consistent with their known ecological preferences.

Fairbanks et al. (1980) and Dekens et al. (2002) reported that $G.\text{ruber}$ (both pink and white variants) and $G.\text{sacculifer}$ (without final chamber) were most common in the mixed layer. Consistent with these results, our analysis suggests that $G.\text{ruber}$ calcified in the top 30m and 40m for white and pink variants, respectively, their expected depths being 16m and 21m. The expected calcification depth of $G.\text{sacculifer}$ (without final chamber), from our analysis, is near 30m, with a rather narrow 95% confidence interval between 18m and 40m.

Our results for $O.\text{universa}$ are also consistent with the Fairbanks et al. (1980) data: the 95% confidence interval, covering the top 60m, encompasses, therefore, both the mixed layer and upper thermocline. Ravelo and Fairbanks (1992) found that $P.\text{obliquiloculata}$ was most abundant at 60m, near the base of the seasonal thermocline: this agrees well with our estimate of 50m for its calcification depth. Our results also include an expected calcification depth near 100m for $N.\text{dutertrei}$ (albeit with a wide confidence interval 64m–169m), and near 75m for $G.\text{menardii}$ (with a narrow confidence interval between 63m and 87m): both these depths are consistent with conclusions of Fairbanks et al. (1980) and Dekens et al. (2002). The depth range of 176–273m obtained in our analysis for
G. tumida, suggesting that this species calcifies below the thermocline, also matches well with the Fairbanks et al. (1980) net tow results.

3.2. Model Consistency Checks

As a summary measure of the model fit, we perform the omnibus $\chi^2$ discrepancy test (Gelman et al. 2004, Sec. 6.5, Eq. (6.4)), which in our case amounts to the statistic

\[ \chi^2 = \sum_{i=1}^{N} \frac{Y_i^2}{\sigma^2}, \]  

(17)

where $Y_i$ are model discrepancies computed by (9) for observed values of $\delta^{18}O_{ci}$ and a certain set of model parameters. In general, statistic (17) is a function of the parameters $\alpha$, $\beta$, $z$, and $\sigma$. We take these parameters at their posterior means, and report sample values of (17) in Table 4. Under a null hypothesis that model errors are normal with zero mean distribution, these sample values can be interpreted as coming from the $\chi^2$ distribution with $N - 4$ degrees of freedom (4 parameters were selected on the basis of $N$ data points). Corresponding $p$-values of the two-sided test to the null hypothesis, i.e. the theoretical probability that the $\chi^2$-distributed random variable with $N - 4$ degrees of freedom would reach further towards the tails of its distribution than the sample value of this statistic (DeGroot and Schervish, 2002), are all quite large, no less than 30%. Therefore the null hypothesis is accepted for the analysis of all species: model fit is generally consistent with its assumptions. This is not surprising, because the model error, $\sigma$, is one of the estimated model parameters.

To inspect model residuals in greater detail, we present in Figure 5 scatterplots of measured versus predicted $\delta^{18}O_c$. Predictions were computed using equation (15). Their estimated theoretical errors (two standard deviations, shown in Figure 5 by horizontal
lines) vary strongly between species, and sometimes even between cores for individual species. The largest Bayesian prediction errors are for *N. dutertrei* and *O. universa*, because the analyses of these species produced larger posterior uncertainty in their calcification depth, particularly in the depth intervals with large temperature variability. Large theoretical prediction errors for δ\(^{18}\)O\(_c\) of *G. ruber* (white) and *G. sacculifer*, which stand out visually in Figures 5a and 5c, are expected for the core VM019-284, in which location large vertical temperature gradients occur, according to the WOA2001 data set (*G. ruber* (pink) δ\(^{18}\)O\(_c\) measurement is not available for this core, thus a point with the large uncertainty is missing in Figure 5b).

An important characteristic of prediction is its bias. Do we have evidence of prediction bias in the scatterplots of Figure 5? Mean error of the prediction (ME) is a sample mean of the actual differences between measured and predicted δ\(^{18}\)O\(_c\). To evaluate their significance, they need to be compared to the error standard deviations (STDE), also estimated from the sample. These parameters are reported in the panels of Figure 5 and Table 4. Under normality assumption, the null hypothesis that the errors have zero mean can be tested using Student’s *t*-statistics (DeGroot and Schervish, 2002),

\[ t = \sqrt{N} \times \frac{\text{ME}}{\text{STDE}}. \]

Corresponding *p*-values (Table 4) are large for all species, except *P. obliquiloculata*, but even for them the *p*-value is 0.06. Therefore the null hypothesis of unbiased prediction for each species successfully passes the *t*-test with 5% significance.

Lack of bias having been established, one can ask if the predictions in Figure 5 are optimal in terms of the slope of a predictive line. Could we improve the skill by rescaling the predictions up or down? A formal way to address this question is to assume that measurements (M) can be presented as scaled predictions (P) with some offset (C) and
random error ($\varepsilon$):

$$M = r \times P + C + \varepsilon.$$  \hfill (18)

For predictions shown in Figure 5, we effectively use $r = 1$ and $C$ equal zero. To check if $r = 1$ is consistent with the data, we test the null hypothesis of $r$ being equal to one in equation (18) against its two-sided alternative. This is effectively a $t$-test for the slope of a univariate linear regression. If the null hypothesis is true, the test statistic

$$t = (r - 1)/\sigma_r,$$

where $r$ and $\sigma_r$ are respectively sample regression coefficient and its standard error estimate, is distributed as Student’s $t$ with $N - 2$ degrees of freedom (DeGroot and Schervish, 2002, Sec. 10.3, Eq. (10.3.20)). Results presented in Table 4 show that the null hypothesis of $r = 1$ passes the test with 5% significance for both variants of $G. ruber$, $G. sacculifer$, and $N. tumida$, but has to be rejected for the four other species. By equation (7), the slope of the measured-to-predicted $\delta^{18}O_c$ relationship is, to a large degree, controlled by the model parameter $\beta$. Failure of the model, at least for some species, to select parameters that fit data best are discussed below along with other model limitations.

4. Discussion

4.1. Intercomparison of calibration relationships

The model (7) used here can be viewed as a constrained linear regression with some additional parametric dependence of predictors on $z$. The constraint on the regression coefficients came in a form of the uniform prior distribution for $\alpha$ and $\beta$ defined by (11). Figure 6a shows a box corresponding to our prior distribution on the $(\alpha, \beta)$ plane and indicates parameter positions for published calibration equations (1)-(5) (which informed
our prior constraints) together with the means of posterior distributions that were obtained in our analyses of different species. All solutions are concentrated in the part of the prior domain corresponding to larger values of $\beta$. The 95% confidence intervals for our parameter estimates for most species are quite wide, so they generally cover a large part of the prior domain. Nevertheless we note that the solution for *N.tumida* is especially close to the relationship obtained by Bemis et al. (1998) from culture experiments with 12-chambered *G.bulloides*, while solutions for both variants of *G.ruber* are closest to the calibration from culture experiments with *O.universa* under high light conditions (Bemis et al. 1998). Our solution for *O.universa* has within its 95% confidence area the parameters of the Bemis et al. (1998) calibration for *O.universa* under low light conditions, although our solution for *G.sacculifer* is even closer to it.

Figure 6a also demonstrates how prior constraints for $\beta$ affected solutions for four species, *P.obliquiloculata*, *G.menardii*, *O.universa*, and *N.dutertrei*: their $(\alpha, \beta)$ points are particularly close to the highest allowable value in $\beta$. Incidentally, these are precisely the species for which the prediction slope was inconsistent with the data best fit (Table 4). In other words, the prior constraint on the allowable range in $\beta$ prevented the model from selecting higher $\beta$ values for these species, even though higher values would fit the available data better.

Are larger values of $\beta$, implied by these analyses, real or an artefact of relatively small sample sizes (around 30) and perhaps some systematic error? Figure 6b plots calibration lines corresponding to all $(\alpha, \beta)$ points in Figure 6a, together with data points of temperature versus oxygen isotope ratio difference, both evaluated at the species' calcification
depths by

\[ T_{calc} = \int T(z)p(z)dz, \quad \left[ \delta^{18}O_c - \delta^{18}O_w \right]_{calc} = \delta^{18}O_c - \int \delta^{18}O_w(z)p(z)dz, \quad (19) \]

where \( p(z) \) is the species-dependent posterior marginal density function for the calcification depth \( z \). In the context of the general scatter of observational points and the spread of calibration lines (reaching almost 4°C in temperature and 0.8 per mil in \( \delta^{18}O \)), different calibration lines look almost parallel; their slopes appear consistent with the large-scale cross-species arrangement of the data. In fact, a standard unconstrained linear best fit to this entire multi-species data set produces a line (red in Figure 6) with parameters \( \alpha = 15.3 \) and \( \beta = 5.11 \), which are within the range of published calibrations. Note that the slope of this line is controlled not by the scatter of points for individual species, but by the general shifts of individual species subsets with regards to each other.

The latter conclusion is consistent with our earlier observation of no significant bias in \( \delta^{18}O_c \) predictions for individual species (Table 4). Since mean shifts in measured \( \delta^{18}O_c \) and calcification interval temperatures between individual species might be as large as the data point scatter within species, these data sets of individual species, when pooled together, extend along their assumed (and almost parallel) calibration lines. This phenomenon is illustrated further by Figure 7, where the scatterplots from different panels of Figure 5 are presented in the same panel. Even though the standard deviation of error (STDE) in Figure 7 (0.33 per mil) is in the range of STDE values in panels of Figure 5 (0.17–0.49 per mil), the normalized prediction skill, which can be measured by the signal to noise ratio, is much higher in Figure 7 than in any of Figure 5 panels. The reason for this is a larger “signal” variability: the range of measured \( \delta^{18}O_c \) exceeds 3 per mil in Figure 7, while these ranges are smaller at least by a factor of two in most panels of Figure 5.
A contrast in prediction skill between data sets of individual species in Figure 5 and the multi-species data set in Figure 7 emphasizes the fact that measured $\delta^{18}O_c$ have a better consistency with published calibration relationships in the “vertical direction”, i.e. between species, from hydrographic conditions of one calcification depth interval to another, than it has in the “horizontal directions”, i.e. between the same species but for different geographic locations. Obviously, something hurts the prediction skill for individual species in data sets which include multiple locations.

Figure 7 provides an opportunity for putting our analysis results in the context of that by Schmidt and Mulitza (2002, their Figure 8, right panel). Their root mean square analysis error is 0.53 per mil. It was obtained for a global data set, with $\delta^{18}O_c$ varying from -3 to 4 per mil. Their selection of species also differed from ours: it included three species which we had as well ($G.ruber$, white and pink variants, and $G.sacculifer$) and three species which we did not use ($Neogloboquadrina pachyderma$, left- and right-coiling variants, and $G.bulloides$).

4.2. Caveats of the analysis

One particular weakness of this data set is the lack of precise age control on the coretop sediment samples. Using a foraminiferal sample from an older coretop in locations of slower sedimentation rate can bias the results of our comparison with modern temperatures because the sampled foraminifera have grown in systematically different temperature regimes. In an attempt to identify such cores, we plot their temperature profiles in Figure 8, along with $\delta^{18}O_c$-based temperature predictions from our analysis. Visual inspection does not identify any cores which are consistently offset. Temperature estimates are generally within $3^\circ$C of co-located WOA2001 profiles, most outliers being associated with
O.universa, N.dutertrei, and N.tumida, the species with large uncertainty in their calcification depth estimates. This conclusion is also confirmed by Figure 6b, which shows some increase in the scatter near ocean surface (at high temperatures). Part of this increase is due to the diminished representativeness of WOA2001 temperatures near the surface, where trends and decadal variability are larger than at depth, but high δ18Oc – high temperature O.universa outliers contributed to it as well.

The O. universa example highlights the obvious weakness in our analysis: an assumption that each species calcifies at a single depth. This assumption is known to be wrong for O. universa and likely has resulted in large depth uncertainty and the largest prediction error for this species in our analysis. Yet, because of the assumption of a single depth, we might still have underestimated the uncertainty in its depth interval. Relaxing the assumption of a single depth in our analysis, along with the introduction of other controlling parameters (light, density, nutrients, seasonality) that the species might optimize in choosing their habitat, can improve the model performance and help with better interpretations for species able to live under a variety of environmental conditions.

The possibility that some species form a secondary calcite layer can also be addressed by allowing more than one calcification depth interval in the model, at least theoretically. For example, G.sacculifer is known to form 20-30% of secondary calcite, by weight (Lohmann, 1995). This estimate was reconfirmed in the global analysis by Schmidt and Mulitza (2002). In our analysis, however, considering the narrowness of the calcification depth interval that we obtained for G.sacculifer with the present model, it is unlikely that we could derive reliably an additional, much deeper, calcification depth interval for
G. sacculifer from the small size data set that we are analyzing here, even if our model were given this flexibility.

Since we allow species-dependent adjustments of calibration coefficients in the model, mean impacts of the “vital” effect (species-specific fractionation due to biophysical processes) over the entire data set can be absorbed by some change in $\alpha$. Similarly, the mean impact of carbonate ion concentration $[CO_3^{2-}]$ on $\delta^{18}O_c$ (Zeebe, 1999; Schmidt and Mulitza, 2002) can be compensated for by changes in both $\alpha$ and $\beta$. To identify effects like that in an unambiguous way, they not only need to be introduced into the model, but will also require a larger coretop data set, featuring significant variations of these effects from one location to another.

Annual mean temperature and $\delta^{18}O_w$ values of the modern period used in this work are subject to measurement, sampling, and analysis error: gridded WOA2001 and LeGrande and Schmidt (2006) data sets were produced by interpolation and smoothing of sparse observations. The smoothing may have introduced some false consistency into the gridded data, masking partly its natural variability. The data sets’ own errors, however, are most likely dwarfed by the errors introduced in our usage of these data: a lack of control for the species’ seasonality or precise time period of the coretop sample.

The main remaining problem in the interpretation of this data set is evident in the clustered, and thus possibly non-random, deviations from one-to-one line in the panels of Figure 5. Some effects which change from place to place and from species to species must be at work there, creating these deviations from our estimates of calibration line at calcification depths. What is their nature: age variations? additional controlling factors? non-constant vital effects? seasonality? The data set in its present form is
too small to identify these unknown effects; a larger database could produce a clearer picture. Radiocarbon age control of coretop samples as well as the development of species-specific calibration equations are other directions where future progress can help reduce uncertainties in the analyses of this and similar data sets.

5. Conclusions

Measured $\delta^{18}O_c$ of eight species of planktonic foraminifera from 31 tropical Atlantic sediment coretops is consistent with upper ocean temperatures and $\delta^{18}O_w$ at various water depths at the core locations and with ecological water depth preferences of these species known from plankton tow and sediment trap abundance data. Bayesian statistical analysis highlighted this correspondence.

Continuous distribution of apparent calcification depths for these species throughout the top 250m (Figure 4) suggests that a reconstruction of thermocline profiles based on a multi-species approach may yield higher-resolution thermocline reconstructions than have been possible previously. We envision multiple regression using a matrix inversion approach, in which a multi-point thermocline profile may be reconstructed rather than a simple thermocline depth.

Some of the species studied here have not been subject to rigorous attempts to calibrate their $\delta^{18}O_c$ as paleotemperature proxies. New data relating $\delta^{18}O_c$ of deep-dwelling species $G. \text{menardii}$ and $N. \text{tumida}$ to ocean temperatures at roughly 75m and 175–275m respectively highlight the potential contributions these species can make to further development of paleotemperature reconstructions.
Acknowledgments. This research was supported by grants from the Climate Center of LDEO and a Graduate Research Environmental Fellowship to ECF from the Global Change Education Program, which is administered by the Oak Ridge Institute for Science and Education for the U.S. Department of Energy’s Office of Biological and Environmental Research. AK was supported by the NSF grants ATM04-17909 and OCE 03-17941. Samples were supplied by the LDEO Deep-Sea Sample Repository, and Martha Dees and Linda Baker provided expert guidance in sample preparation and analysis. Early versions benefited greatly from discussion with Yair Rosenthal, and constructive criticism from two anonymous reviewers greatly improved the manuscript. LDEO contribution number 7025.

References


56-60.


Table 1. Data produced in this analysis listed with Core ID, core’s latitude, longitude, and water depth. Cores locations are arranged from west to east. Measured $\delta^{18}O$, is given for eight species of planktonic foraminifera: WRU = *G. ruber* (white), PRU = *G. ruber* (pink), SAC = *G. sacculifer* (without final chamber), UNI = *O. universa*, OBL = *P. obliquiloculata*, MEN = *G. menardii*, DUT = *N. dutertrei*, and TUM = *N. tumida*. Starred core IDs indicate cores for which all eight species were measured. Bottom line gives the number of cores in which each species was measured.
<table>
<thead>
<tr>
<th>Species</th>
<th>Size fraction</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>G. ruber (white)</td>
<td>355 425</td>
<td>Mulitza (pers. comm.)</td>
</tr>
<tr>
<td>G. ruber (pink)</td>
<td>355 425</td>
<td>Mulitza (pers. comm.)</td>
</tr>
<tr>
<td>G. sacculifer</td>
<td>355 425</td>
<td>Mulitza (pers. comm.)</td>
</tr>
<tr>
<td>O. universa</td>
<td>500 600</td>
<td>Ravelo and Fairbanks (1992)</td>
</tr>
<tr>
<td>P. obliquiloculata</td>
<td>500 600</td>
<td>Ravelo and Fairbanks (1992)</td>
</tr>
<tr>
<td>G. menardii</td>
<td>600 710</td>
<td>Ravelo and Fairbanks (1992)</td>
</tr>
<tr>
<td>N. dutertrei</td>
<td>500 600</td>
<td>Ravelo and Fairbanks (1992)</td>
</tr>
<tr>
<td>N. tumida</td>
<td>500 600</td>
<td>Ravelo and Fairbanks (1992)</td>
</tr>
</tbody>
</table>

**Table 2.** Species and size fractions of planktonic foraminifera used in this analysis, with references used to determine size fraction for each species.
<table>
<thead>
<tr>
<th>Species</th>
<th>$\alpha$, °C</th>
<th>$\beta$, °C/per mil</th>
<th>$z$, m</th>
<th>$\sigma$, per mil</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G.\text{ruber}$ (white)</td>
<td>15.4</td>
<td>14.1 – 16.8</td>
<td>4.78</td>
<td>4.18 – 5.26</td>
</tr>
<tr>
<td>$G.\text{ruber}$ (pink)</td>
<td>14.7</td>
<td>13.4 – 16.6</td>
<td>4.86</td>
<td>4.12 – 5.27</td>
</tr>
<tr>
<td>$G.\text{sacculifer}$</td>
<td>16.2</td>
<td>15.4 – 16.9</td>
<td>4.94</td>
<td>4.57 – 5.26</td>
</tr>
<tr>
<td>$O.\text{universa}$</td>
<td>16.5</td>
<td>15.4 – 17.0</td>
<td>5.11</td>
<td>4.68 – 5.29</td>
</tr>
<tr>
<td>$P.\text{obliquiloculata}$</td>
<td>16.8</td>
<td>16.5 – 17.0</td>
<td>5.22</td>
<td>5.06 – 5.30</td>
</tr>
<tr>
<td>$G.\text{menardii}$</td>
<td>16.6</td>
<td>16.0 – 17.0</td>
<td>5.20</td>
<td>5.00 – 5.30</td>
</tr>
<tr>
<td>$N.\text{dutertrei}$</td>
<td>14.6</td>
<td>12.5 – 16.9</td>
<td>5.09</td>
<td>4.55 – 5.29</td>
</tr>
<tr>
<td>$N.\text{tumida}$</td>
<td>13.1</td>
<td>12.5 – 14.1</td>
<td>4.95</td>
<td>4.31 – 5.28</td>
</tr>
</tbody>
</table>

**Table 3.** Posterior estimates for parameters obtained from Bayesian analysis. Expected values (means of the posterior probability distribution) and 95% confidence intervals are presented for the calibration intercept $\alpha$, slope $\beta$, calcification depth $z$, and the standard deviation $\sigma$ of calibration error.
<table>
<thead>
<tr>
<th>Species</th>
<th>N</th>
<th>Omnibus $\chi^2$ test $p$-value</th>
<th>Student’s $t$-test for the prediction bias</th>
<th>Student’s $t$-test for the prediction slope</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\chi^2$</td>
<td>ME, STDE, $t$ $p$-value</td>
<td>$r$, $\sigma_r$, $t$ $p$-value</td>
</tr>
<tr>
<td></td>
<td></td>
<td>per mil</td>
<td>per mil</td>
<td></td>
</tr>
<tr>
<td>G.ruber (white)</td>
<td>31</td>
<td>29.9 0.64</td>
<td>0.002 0.26 0.05 0.96</td>
<td>0.97 0.30 -0.1 0.91</td>
</tr>
<tr>
<td>G.ruber (pink)</td>
<td>28</td>
<td>26.8 0.63</td>
<td>0.000 0.24 0.01 0.99</td>
<td>0.47 0.27 -2.0 0.06</td>
</tr>
<tr>
<td>G.sacculifer</td>
<td>31</td>
<td>30.1 0.62</td>
<td>0.003 0.17 0.09 0.93</td>
<td>1.00 0.19 0.0 0.99</td>
</tr>
<tr>
<td>O.universa</td>
<td>27</td>
<td>25.8 0.62</td>
<td>0.089 0.49 0.95 0.35</td>
<td>-0.02 0.44 -2.3 0.03</td>
</tr>
<tr>
<td>P.obliquiloculata</td>
<td>24</td>
<td>24.3 0.46</td>
<td>0.109 0.27 1.99 0.06</td>
<td>0.34 0.13 -5.0 0.00</td>
</tr>
<tr>
<td>G.menardii</td>
<td>30</td>
<td>29.8 0.55</td>
<td>0.045 0.33 0.75 0.46</td>
<td>0.50 0.08 -6.1 0.00</td>
</tr>
<tr>
<td>N.dutertrei</td>
<td>26</td>
<td>28.8 0.30</td>
<td>-0.040 0.46 -0.45 0.66</td>
<td>0.24 0.13 -6.1 0.00</td>
</tr>
<tr>
<td>N.tumida</td>
<td>31</td>
<td>30.0 0.63</td>
<td>-0.029 0.32 -0.49 0.63</td>
<td>0.77 0.17 -1.3 0.20</td>
</tr>
</tbody>
</table>

**Table 4.** Consistency tests for Bayesian analysis and predictions (see text for explanations).
Figure 1. Map showing location of all cores (all depths below 4150m isobath are shaded gray). Dots show locations of 23 cores for which oxygen isotope data were obtained for all 8 species used in the analysis; stars show the locations of 8 cores for which data were obtained for only some of the species. Note that though VM19-297 at 2.617°N, 12.000°W appears below the 4150m isobath, it is located on a local high point at 4122m (see Table 1).
Figure 2. Mean annual ocean water (a) temperatures from NOAA World Ocean Atlas (Conkright et al 2001) and (b) stable oxygen isotope ratios $\delta^{18}O_w$ from LeGrande and Schmidt (2006) analysis at all 31 core locations used in this study.
Bayesian analysis for *G. ruber* (white)

(a) $p(\alpha, \beta)$

(b) $p(\beta)$

(c) $p(\alpha)$

(d) $p(z)$

(e) $p(\sigma)$

**Figure 3.** Results of Bayesian analysis for *G. ruber* (white). Shown are (a) joint posterior probability density function (PDF) for calibration coefficients $\alpha$ and $\beta$ (the star indicates the mean of the posterior distribution, with thin lines showing 95% confidence intervals for individual marginal distributions of $\alpha$ and $\beta$), and marginal PDFs for (b) $\beta$, (c) $\alpha$, (d) calcification depth $z$, and (e) calibration error standard deviation $\sigma$. 
Figure 4. Probability density functions for the calcification depth. Color scale saturates at 0.03 m$^{-1}$. Stars indicate expected values (posterior distribution means). Vertical black lines show 95% confidence intervals. Positions on the horizontal axis correspond to different species: WRU = *G. ruber* (white), PRU = *G. ruber* (pink), SAC = *G. sacculifer* (without final chamber), UNI = *O. universa*, OBL = *P. obliquiloculata*, MEN = *G. menardii*, DUT = *N. dutertrei*, and TUM = *N. tumida*. 

February 2007, 5:51pm
Figure 5. Measured versus predicted (i.e., expected) calcite $\delta^{18}O_c$. Predictions are made using equation (15). Horizontal lines drawn through individual markers show two standard deviations of Bayesian prediction error. Sample means (ME) and standard deviations (STDE) of the actual measured–predicted $\delta^{18}O_c$ errors are indicated in each panel.
Figure 6. Published and estimated calibration relationships. (a) Open circles mark \((\alpha, \beta)\) pairs for published calibration equations (1)-(5); crosses indicate posterior means obtained in the present analysis for different species; lines drawn through crosses show 95% confidence intervals for \(\alpha\) and \(\beta\); parameters of the simultaneous linear best fit to all data in the bottom panel and their two standard errors are shown in red. (b) A scatterplot of calcification depth temperatures versus \(\delta^{18}O_c - \delta^{18}O_w\), estimated by equation (19), is shown along with the calibration lines defined by equations (1)-(5) and by \((\alpha, \beta)\) posterior means from present analyses (identified by species symbols); red line is the linear best fit to all data in this plot.
Figure 7. Comparison of measured and predicted $\delta^{18}O_c$ values: data from all panels of Figure 5, pooled together.
Figure 8. Comparison of observed and predicted by equation (16) temperature values. Markers identify different species by the legend from Figure 7. Vertical lines through markers are 95% confidence intervals for the calcification depth; horizontal lines are two standard errors for temperature predictions.