



Mismatch between the depth habitat of planktonic foraminifera and the calibration depth of SST transfer functions may bias reconstructions

R. J. Telford^{1,2}, C. Li^{2,3}, and M. Kucera⁴

¹Department of Biology, University of Bergen, Thormøhlensgate 53A, 5006 Bergen, Norway

²Bjerknes Centre for Climate Research, Allégaten 55, 5007 Bergen, Norway

³Geophysical Institute, University of Bergen, Allégaten 70, 5008 Bergen, Norway

⁴MARUM & Fachbereich Geowissenschaften, Universität Bremen, Leobener Strasse, 28359 Bremen, Germany

Correspondence to: R. J. Telford (richard.telford@bio.uib.no)

Received: 31 July 2012 – Published in Clim. Past Discuss.: 24 August 2012

Revised: 31 January 2013 – Accepted: 27 February 2013 – Published: 22 March 2013

Abstract. We demonstrate that the temperature signal in the planktonic foraminifera assemblage data from the North Atlantic typically does not originate from near-surface waters and argue that this has the potential to bias sea surface temperature reconstructions using transfer functions calibrated against near-surface temperatures if the thermal structure of the upper few hundred metres of ocean changes over time. CMIP5 climate models indicate that ocean thermal structure in the North Atlantic changed between the Last Glacial Maximum (LGM) and the pre-industrial (PI), with some regions, mainly in the tropics, of the LGM ocean lacking good thermal analogues in the PI.

Transfer functions calibrated against different depths reconstruct a marked subsurface cooling in parts of the tropical North Atlantic during the last glacial, in contrast to previous studies that reconstruct only a modest cooling. These possible biases in temperature reconstructions may affect estimates of climate sensitivity based on the difference between LGM and pre-industrial climate. Quantifying these biases has the potential to alter our understanding of LGM climate and improve estimates of climate sensitivity.

1 Introduction

The composition of planktonic foraminifera assemblages in ocean surface sediments appears to be related to sea surface temperatures (SST) (Murray, 1897) and has been used to make quantitative SST reconstructions since the late 1960s, when transfer functions for estimating past climatic conditions from the modern relationship between species and the environment in a modern calibration set were developed (Sachs et al., 1977). These reconstructions constitute a large fraction of the available palaeoceanographic data, and have been synthesised into regional or global maps of past SST for key time periods. Compilations for the Last Glacial Maximum (LGM; 21 ka) by CLIMAP (CLIMAP Project Members, 1976) found colder conditions at high latitudes, but surprisingly reported warmer than modern conditions in the subtropical ocean. MARGO (MARGO Project Members, 2009) reported less extensive warming in the subtropics, and modest cooling in the tropics. Such estimates of LGM climate have been extensively used to validate climate models (Brannot et al., 2012) on the premise that if the models cannot reproduce past climate, they are unlikely to be useful for predicting future climate. Recently, foraminifera-derived SST estimates were central to an extensive network of LGM climate anomalies that Schmittner et al. (2011) used to estimate climate sensitivity, the amount that the global average climate will warm following a doubling of atmospheric CO₂ concentrations. Schmittner et al. (2011) estimated that

climate sensitivity is 2.2°C , considerably smaller than the IPCC estimate of about 3°C (Hegerl et al., 2007), and with a narrower uncertainty range. However, the Schmittner et al. (2011) estimate is sensitive to biases in oceanic LGM temperature anomalies, and these anomalies are dominated by foraminiferal records. As uncertainty in climate sensitivity constitutes the largest source of uncertainty in climate projections beyond a few decades (Knutti and Hegerl, 2008), potential sources of bias in planktonic foraminifera assemblages must be carefully evaluated.

Transfer functions for reconstructing past environmental conditions make a number of assumptions (Birks et al., 2010). If these assumptions are violated, reconstructions are potentially erroneous. The implications of some of these assumptions for planktonic foraminifera-derived SST have been explored, including the potential for spatial autocorrelation to make reconstructions appear more certain than justified by the data (Telford and Birks, 2005). In this paper we focus on the assumption that environmental variables other than the one of interest have negligible influence during the time window of interest or that the joint distribution of these variables of interest in the past is the same as today (Birks et al., 2010). If foraminifera assemblage composition is controlled by several environmental variables, and the correlation between these variables changes over time, this assumption will be violated.

The assemblage composition of planktonic foraminifera is usually calibrated against either seasonal or annual mean SST at a fixed depth in the upper ocean. Although it is well known from plankton tow studies (Fairbanks et al., 1980; Wilke et al., 2009) and from geochemical evidence of calcification depths using oxygen isotopes (Ganssen and Kroon, 2000) and magnesium/calcium ratios (Cl  roux et al., 2008) that planktonic foraminifera live at a broad range of depths in the upper ocean, Pflaumann et al. (1996) showed that, at least in the Atlantic Ocean, transfer function performance is best at 10 m or the mean of the top 75 m. If the temperature signal recorded in planktonic foraminifera assemblages integrates the influence of thermal structure of the whole upper ocean rather than the temperature at a fixed depth, the assumptions of transfer functions may be violated. Transfer function estimates of near-surface temperatures will then only be correct if the thermal structure of the upper few hundred metres of ocean has remained constant. This is unlikely given the reorganisation of ocean circulation during the LGM (e.g. Lynch-Stieglitz et al., 1999). Despite this potential bias, the effect of changing thermal structure on transfer function SST estimates has never been systematically explored. That it could be important has recently been demonstrated for the Mediterranean by Adloff et al. (2011).

We undertake several analyses to investigate the potential for bias in SST estimates should the foraminifera signal track subsurface temperature rather than near-surface conditions. First, we repeat the analyses of Pflaumann et al. (1996), both on an Atlantic-wide scale and on a regional scale, using

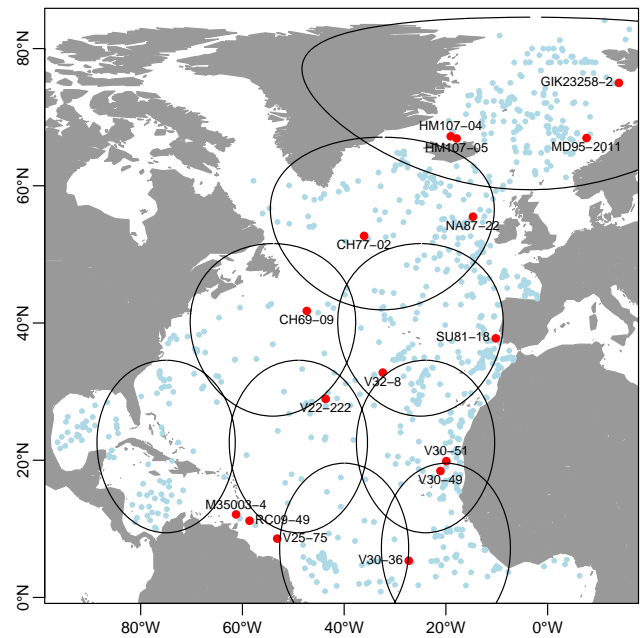


Fig. 1. Map of the North Atlantic Ocean showing calibration set sites (pale blue), core sites (red), and the 1400-km-radius circles for testing the regional performance of transfer functions.

updated foraminifera and ocean temperature data. We then make reconstructions of temperature at different water depths for a collection of North Atlantic cores, and argue that the reconstruction that best explains the variability in the faunal record since the last glaciation reflects the depth that most influenced faunal composition. We finish by using the output of CMIP5 climate models to identify parts of the ocean with thermal structures not present in the modern ocean; it is here that the potential for bias in reconstructions is greatest if foraminifera assemblages are tracking subsurface rather than near-surface temperatures. We conclude by discussing the likely sign of the potential, possibly regionally varying, biases in transfer function SST estimates using climate model output.

2 Methods

We used the 862-site North Atlantic planktonic foraminifera calibration set compiled by Kucera et al. (2005) (Fig. 1). Ocean temperatures were extracted from the World Ocean Atlas (WOA, 1998), interpolated to the calibration set observations. We used the caloric warm and cold season, and mean annual temperatures at the 14 standard WOA depths between the surface and 500 m.

Sixteen planktonic foraminifera assemblage time series straddling the last termination from the North Atlantic were compiled and their taxonomy harmonised to match the calibration set (Table 1, Fig. 1). The following selection criteria were used: foraminifera counted in the greater than $150\ \mu\text{m}$

Table 1. Cores used in the analyses.

Core	Location	Number of observations	Timespan (ky BP)	Reference
GIK23258-2	14° E, 75° N	285	1–14	Sarnthein et al. (2003)
HM107-04	19.1° W, 67.2° N	92	0–13	Knudsen et al. (2004)
HM107-05	17.9° W, 66.9° N	56	0–16	Knudsen et al. (2004)
MD95-2011	7.6° E, 67° N	380	0–14	Risebrobakken et al. (2003)
NA87-22	14.7° W, 55.5° N	104	1–14	Duplessy et al. (1992)
CH77-02	36.1° W, 52.7° N	153	0–12	Marchal et al. (2002)
CH69-09	47.4° W, 41.8° N	58	1–17	Labeyrie et al. (1999)
SU81-18	10.2° W, 37.8° N	41	0–21	Duplessy et al. (1992)
V32-8	32.4° W, 32.8° N	25	2–23	Mix (2006a); Mix et al. (1986)
V22-222	43.6° W, 28.9° N	30	2–21	Mix (2006c); Mix et al. (1986)
V30-51	19.9° W, 19.9° N	18	2–32	Mix (2006d); Mix et al. (1986)
V30-49	21.1° W, 18.4° N	26	2–23	Mix (2006e); Mix et al. (1986)
M35003-4	61.2° W, 12.1° N	76	0–23	Hüls (1999)
RC09-49	58.6° W, 11.2° N	22	2–22	Mix (2006f); Mix et al. (1986)
V25-75	53.2° W, 8.6° N	35	2–28	Mix (2006g); Mix et al. (1986)
V30-36	27.3° W, 5.3° N	23	2–33	Mix (2006b); Mix et al. (1986)

fraction to match the calibration set, resulting in the rejection of some high-latitude sites; core location north of 5° N to avoid the edge of the calibration set; cores should span the last termination, with several observations from both the Holocene and the deglaciation with a temporal resolution of ~ one observation per millennium or better.

The SST reconstructions were derived for each depth and season using the modern analogue technique (MAT) with five analogues (as between four and six analogues, depending on depth and season, gave the lowest root mean square error of prediction) and squared chord distances (Prell, 1985). Performance of the transfer functions was estimated using leave-one-out cross-validation, and is reported both for the whole North Atlantic Ocean and for nine arbitrarily defined 1400-km-radius regions covering the North Atlantic.

Reconstructions of warm and cold season temperature at WOA standard depths from the surface down to 500 m were calculated using MAT. We attempted to determine the season and depth at which temperature variability appears to be most important by using a constrained ordination to find the proportion of the variance in the fossil data explained by each reconstruction (Telford and Birks, 2011). We assume that reconstructions that explain least variance in the fossil data are probably from depths that did not drive the variability in the fossil data. We used redundancy analysis (Rao, 1964) for the constrained ordination because the taxonomic turnover in the cores is relatively small. The statistical significance of the reconstructions was assessed by testing if they explained significantly more of the variance than a null model of 999 transfer functions trained on random data (Telford and Birks, 2011). We used a white noise null model because of the complexity of generating spatially autocorrelated random data with the correct spatial structure for each environmen-

tal variable. Our 95 % significance level will thus be liberal. We repeated this analysis on the deep- and shallow-dwelling species (Hemleben et al., 1989) in each core separately to try to determine which group carries a stronger temperature signal.

We used the output from time slice simulations of LGM and pre-industrial (PI) climate performed by four coupled climate models participating in the Coupled Model Intercomparison Project CMIP5 to identify areas of the North Atlantic LGM that have poor thermal analogues in the PI ocean. The models included in the analysis are listed in Table 2; the LGM and PI simulations were performed in compliance with PMIP3 protocol. Complete documentation of the models and experimental setup can be found on the PMIP3 website (<http://pmip3.lsce.ipsl.fr/>). We generated monthly climatologies for 50-yr periods (simulations years indicated in Table 2) of the LGM and PI simulations. For each model, we interpolated July temperatures in every grid box to 10-m intervals in the top 300 m of the ocean, and found the Euclidean distance from the simulated LGM ocean to the most similar grid box in the simulated PI ocean.

All analyses were run using the R statistical language version 2.14.1 (R Development Core Team, 2011). Reconstructions were generated with the rioja package version 0.7-3 (Juggins, 2012); the statistical significance and importance of reconstructions was tested with the palaeoSig package version 1.1-1 (Telford, 2012); redundancy analysis was fitted with the vegan package version 2.0-2 (Oksanen et al., 2011).

3 Results

Planktonic foraminifera-SST transfer function performance for the whole North Atlantic Ocean is best near the surface

Table 2. CMIP5 coupled models and periods of the time slice simulations used in this analysis. The sponsoring institutions of the models are GISS-E2-R, NASA/GISS Goddard Institute for Space Studies, USA; MIROC-ESM, Center for Climate System Research (University of Tokyo), National Institute for Environmental Studies, and Frontier Research Center for Global Change (JAMSTEC), Japan; IPSL-CM5A-LR, Institut Pierre Simon Laplace, France; and MPI-ESM-P, Max Planck Institute for Meteorology, Germany.

Model	Atmosphere	Ocean	Simulation period (years)
GISS-E2-R	$2^\circ \times 2.5^\circ \times \text{L40}$	$1^\circ \times 1.25^\circ \times \text{L32}$	PI: 4481–4530 LGM: 3050–3099
MIROC-ESM	$\text{T42} (\sim 2.8^\circ) \times \text{L80}$	$0.5^\circ - 1.7^\circ \times 1.4^\circ \times \text{L44}$	PI: 2280–2329 LGM: 4648–4699
IPSL-CM5A-LR	$3.75^\circ \times 1.9^\circ \times \text{L39}$	$2^\circ \times 2^\circ \times \text{L31}$	PI: 2750–2799 LGM: 2751–2800
MPI-ESM-P	$\text{T63} (\sim 1.9^\circ) \times \text{L47}$	$1.5^\circ \times 1.5^\circ \times \text{L40}$	PI: 2930–2979 LGM: 1850–1899

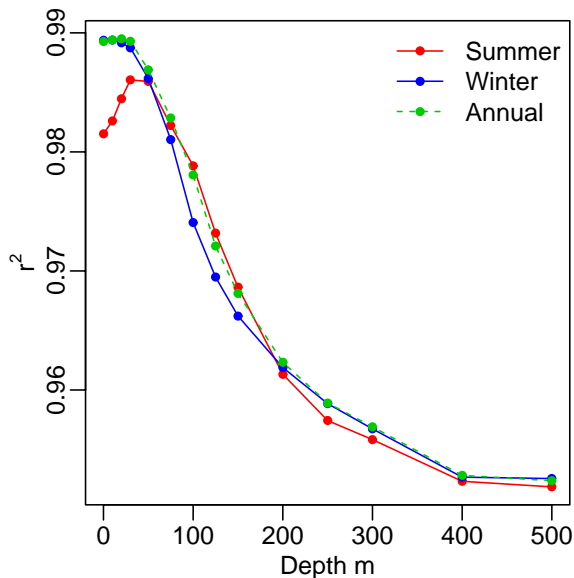


Fig. 2. Transfer function performance for the whole North Atlantic calibration set, shown as the r^2 between measured and predicted temperatures for different depths and seasons.

for the cold season and the annual mean temperature (Fig. 2). The warm season has a maximum r^2 between 30 and 50 m. This result is similar to that of Pflaumann et al. (1996), who only considered the top 75 m, confirming that their result was not an artefact of the SIMMAX method (Telford et al., 2004) or the older, less precise SST data they used (Levitus, 1982). This would seem to support the current practice of attributing transfer function results to a fixed depth representing the mixed layer. However, different patterns emerge when performance of the transfer functions in different regions is estimated (Fig. 3). In several regions there is a pronounced drop in r^2 near the surface during the warm season, and in the tropics this is also found for the cold season and the annual mean temperature. This result suggests that the near monotonic decline in r^2 with depth for the cold season and annual mean temperature in the whole North Atlantic is, at least in part, an artefact of mixing together different regions with different depth sensitivities.

The regional differences in the performance of transfer functions with depth are reflected in the time series of reconstructions at different depths. For cores north of 25°N , the reconstructions from different depths and seasons resemble one another, with an offset (Fig. 4a). Tropical cores have very different reconstructions for different depths. For example, near-surface temperature reconstructions from site V30-36 (Fig. 4b) indicate little variability over the last 30 000 yr, whereas reconstructions from 75- and 100-m depth suggest a warming from the glacial into the Holocene of over 5°C and reconstructions from greater depths have less variability, with a cooling of up to 2°C .

The difference between reconstructions at different depths in the time series is manifested in the proportion of the variance in the fossil data that they explain (Fig. 5). The profile of variance in the fossil data explained by temperature reconstructions at different water depths and seasons varies geographically (Fig. 5).

The four sites in the Nordic Seas have similar shaped profiles (Fig. 5a–d). The amount of variance explained is high and statistically significant for the top 200–300 m, and declines steeply below this. Towards the surface, the variance explained by warm-season reconstructions declines.

The three sites in the North Atlantic Drift (Fig. 5e–g) have broad peaks in the amount of variance explained by warm-season reconstructions, with a maximum at about 200 m. The patterns for the cold season are flatter, and explain less than the warm season at its peak. At these sites all reconstructions are statistically significant.

The single site off the Portuguese margin (Fig. 5h), which is possibly influenced by seasonal upwelling, has a unique profile, with the amount of variance explained by the reconstructions slowly increasing with depth.

The two sites in the North Atlantic subtropical gyre have idiosyncratic profiles (Fig. 5i–j). The amount of variance explained by reconstructions at V32-8 is greatest in the warm season at the surface, and declines with depth until reconstructions are not statistically significant below about 200 m. At V22-222 the pattern is more complex, with the warm-season reconstruction explaining more at about 75 m than at the surface, below this the amount explained first falls then rises, but the magnitude of the changes is small. The

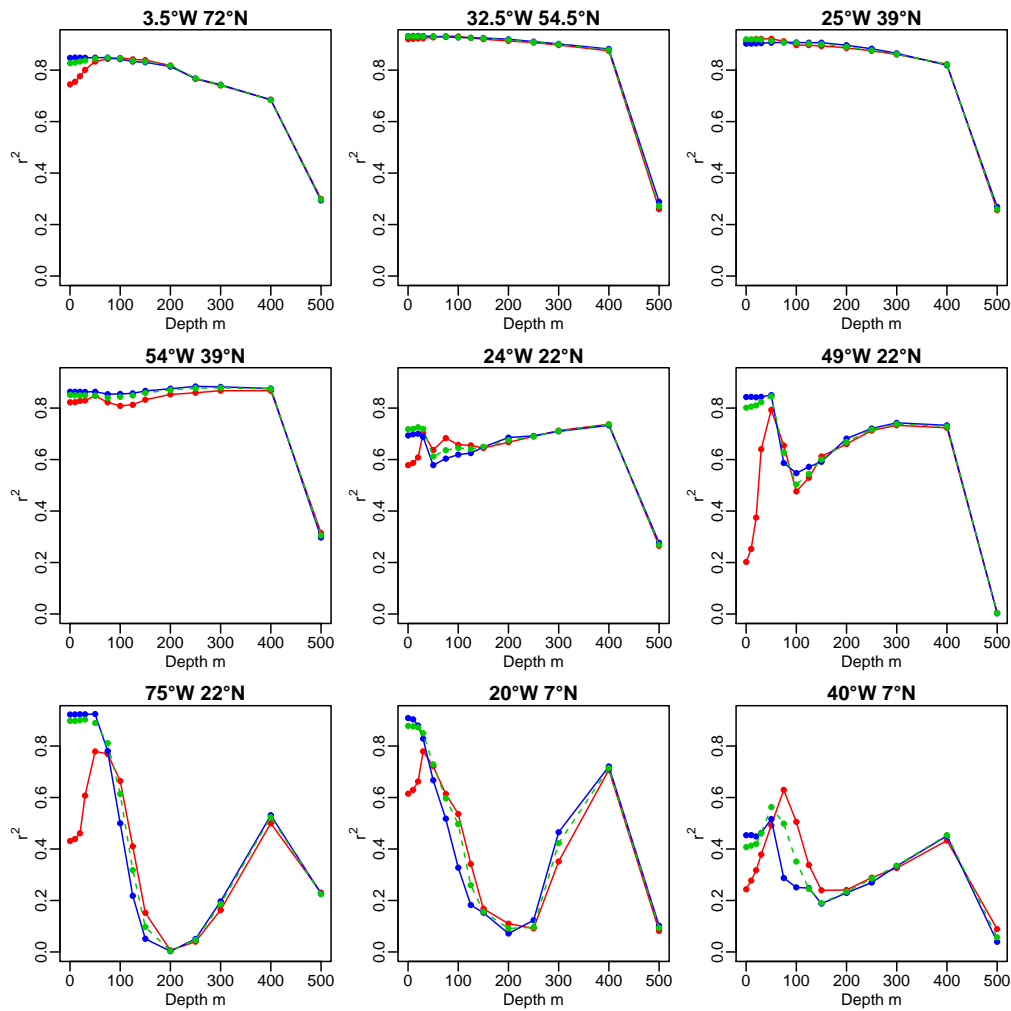


Fig. 3. Transfer function performance, shown as the r^2 between measured and predicted temperatures for different depths and seasons, for sites in different 1400 km radius regions (see Fig. 1) of the North Atlantic calibration set. Legend as Fig. 2.

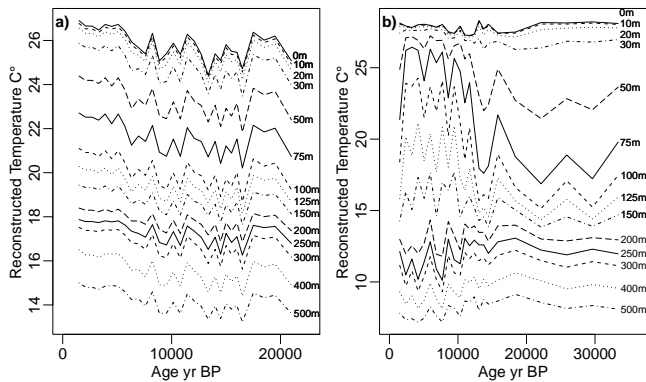


Fig. 4. Warm-season reconstructions for different water depths from (a) V22-222 and (b) V30-36.

pattern for the cold season is similar, except it does not show the near-surface decline. All reconstructions at this site are statistically significant.

The patterns in variance explained and the significance levels change south of 25° N. The two sites near the West African upwelling cells (Fig. 5k–l) both have near-surface reconstructions that explain the most variance, with the warm season explaining more than the cold. Reconstructions of subsurface temperatures explain very little variance, but reconstructions below 200 m explain more, especially at V30-49.

At the tropical sites (Fig. 5m–p), near-surface reconstructions explain little variance and are typically not statistically significant. Subsurface reconstructions explain about twice as much of the variance as near-surface reconstruction, and are statistically significant, or almost so. Reconstructions from below 150 m also explain little variance and

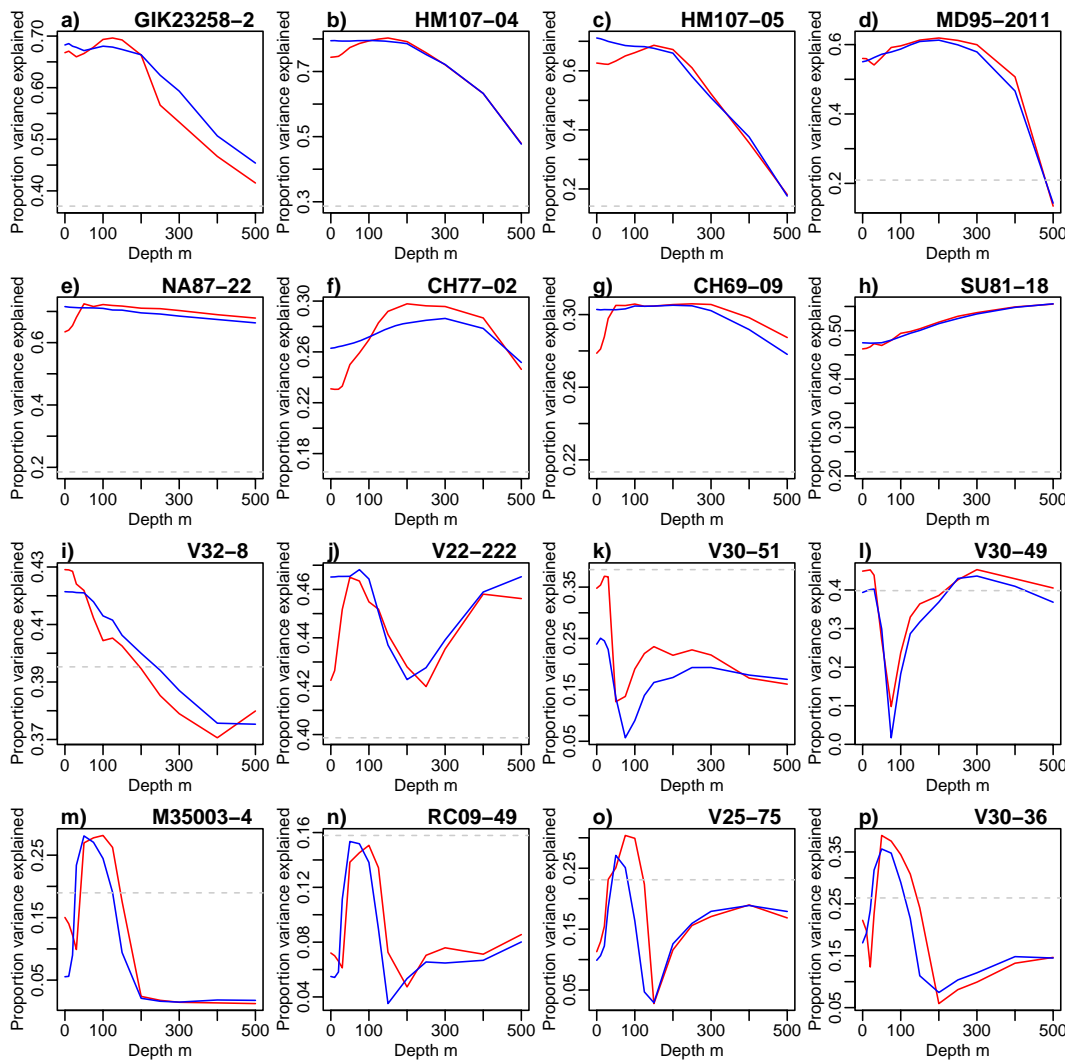


Fig. 5. The proportion of variance in the fossil data explained by reconstructions of warm (red) and cold (blue) season temperatures at different depths. The horizontal dashed grey line represents the 95 % significance level, assessed by finding the proportion of variance explained by reconstructions of random environmental data. Note that the y-axis scale changes between plots and cannot be directly compared as the diversity is different at each site.

are not statistically significant. Both seasons explain a similar amount of variance.

When the foraminifera assemblages are divided into deep- and shallow-dwelling taxa, the similarity between the variance explained by the reconstructions in these groups and the whole assemblage varies geographically (Fig. S1 in the Supplement). In the northern North Atlantic (Fig. S1a–h in the Supplement), the sites are dominated by deep-dwelling taxa and the pattern of variance explained in the deep-dwelling taxa is similar to that explained in the whole community, whereas the shallow-dwelling taxa have varied responses that are often not statistically significant. The pattern is more varied in the tropical and subtropical North Atlantic (Fig. 1i–p in the Supplement). Reconstructions from the deep-dwelling taxa, which constitute about 10 % of the tropical assem-

blages, are not statistically significant in most sites, whereas the shallow-dwelling taxa give statistically significant reconstructions in half these sites, and the pattern resembles that of the whole assemblage in most sites. These results show that the subsurface SST signal is not solely carried by deep-dwelling species. This may reflect the ecology of foraminifera, with species migrating through the water column during their life cycle. Alternatively, this result could be because the properties of the upper water column depend on the temperature gradient in the subsurface – a measure for mixing/stratification.

Figure 6 shows, for each grid box, the Euclidean distance between the vertical temperature profiles in the top 300 m of the simulated LGM ocean and the most similar profile in the PI. Large values indicate grid boxes that lack good analogues

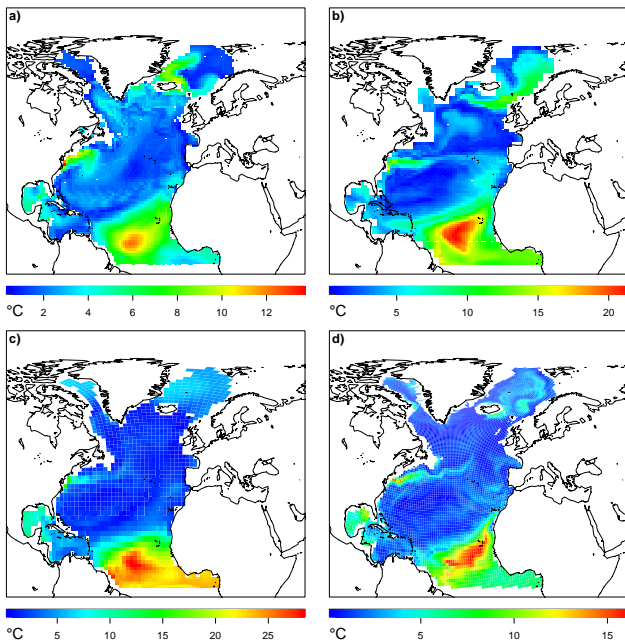


Fig. 6. Euclidean distance between LGM July temperatures over the top 300 m of the water column and the nearest analogues in the PI ocean for four CMIP5 models: (a) GISS, (b) MIROC, (c) IPSL, (d) MPI.

for the LGM thermal structure in the modern ocean. The spatial pattern is similar for all four CMIP5 models (Fig. 6), with less good analogues in the Nordic Seas and around the Gulf Stream separation, and the worst analogues in the area south of the subtropical gyre. This inter-model consistency suggests that these are robust features.

4 Discussion

Planktonic foraminifera assemblages in the sediment integrate live communities from different seasons and water depths. Since their composition reflects the thermal structure of the entire upper ocean, it would be surprising if reconstructions of temperature from one depth in one season perfectly captured SST changes through space and time. However, as long as the thermal structure of the upper ocean remains the same, transfer functions based on a fixed calibration depth should not be biased. Our analysis of North Atlantic foraminifera time series across the last termination (Figs. 4, 5) as well as model simulations of the LGM and PI (Fig. 6) indicate that ocean thermal structure has changed and that transfer-function-based reconstructions are affected. We find that the depth at which planktonic foraminifera assemblages are usually calibrated to – 10 m – is rarely the depth that explains the most variance in the fossil data in the North Atlantic. Given our ecological knowledge on the vertical and seasonal abundances of planktonic foraminifera (Chapman, 2010; Fairbanks et al., 1980; Wilke et al., 2009), these re-

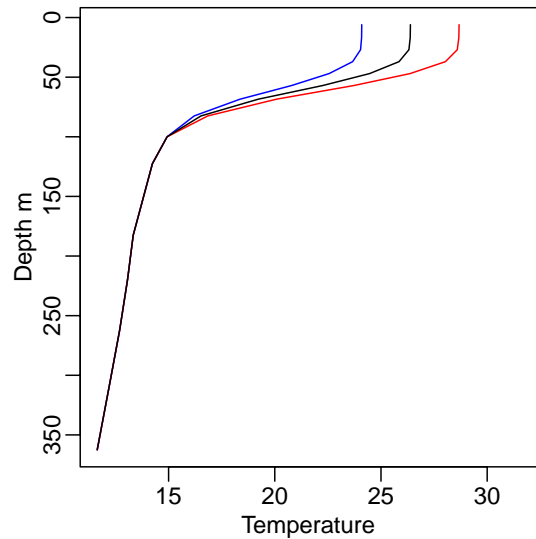


Fig. 7. Schematic plot showing a modern temperature profile (black) and two possible past temperature profiles that are the same at depth, but have either stronger (red) or weaker (blue) stratification. If planktonic foraminifera responded to temperature at 100 m, a transfer function calibrated against 10 m will reconstruct the same SST in all three cases. The reconstruction for the case with the weaker stratification will be biased towards warm temperatures; conversely, the case with stronger stratification will have a cold bias to the reconstruction.

sults are not surprising. They pose both opportunities and challenges for palaeoceanography, which we discuss below.

4.1 Opportunities

The greatest opportunity is the potential for more meaningful reconstructions, leading to improved understanding of past climate. For example, the temperature reconstructions from 10 and 75 m at V30-36 (Fig. 4) are very different, suggesting either only minor temperature changes over the last 30 000 yr or pronounced variability. If strong subsurface cooling at the LGM can be reconstructed at other sites across the tropics, it will challenge the consensus that the tropical oceans only experienced moderate cooling at the LGM (MARGO Project Members, 2009).

Acknowledging that changes in foraminifera assemblages across the last termination, in most of the North Atlantic, were more sensitive to subsurface conditions than surface conditions will allow greater insight into palaeoceanographic processes by exploiting information on the seasonal and depth sensitivity of proxies. This has already been attempted, for example by Jansen et al. (2008) who examined the contrasting SST reconstructions from the Vøring Plateau in the Nordic seas: diatoms and alkenones reconstruct a warm early Holocene, whereas planktonic foraminifera SST reconstructions, both isotopic and faunal, are warmest in the late Holocene. Jansen et al. (2008) argue

that alkenone and diatoms represent summer SST, while planktonic foraminifera represent subsurface conditions and are sensitive to the temperature set during winter ventilation. They conclude that the contrasting reconstructions indicate that direct enhanced insolation rather than advection of warm water was responsible for the early Holocene thermal optima. Another study by Adloff et al. (2011) finds a discrepancy between climate model output and reconstructed SST for the Eastern Mediterranean during the early Holocene thermal optimum. They find that these data can be reconciled by considering the foraminifera-derived reconstruction to represent the upper water column rather than just the surface.

4.2 Challenges

The first challenge is to interpret the ambiguity in having multiple reconstructions – one for each depth and season examined. There may be a desire to try to attach palaeoclimate meaning to several, or indeed all of these. For example, the record from V30-36 (Fig. 4b) could be interpreted as having little temperature change at the surface, but pronounced temperature changes subsurface. However, it is unlikely that there is sufficient information in the fossil data to reconstruct several independent variables simultaneously (Telford and Birks, 2011). Therefore, at any given time, it is likely that only one of the multiple reconstructions can be considered. The statistical significance of each reconstruction may help guide the choice of which should be considered.

The second challenge is that the depth and season that explains the most variance may vary not only geographically, but also with time at one site. Here we consider assemblage changes over the last termination. Different patterns may have been found if we had considered only interglacial or only glacial periods. Theoretically, the framework we develop here could be used to explore this problem using a moving-window analysis, but there may be problems with obtaining adequate faunistic analogues.

Other challenges include problems using and displaying this information. There is a long tradition of drawing maps of SST anomalies (CLIMAP Project Members, 1976; MARGO Project Members, 2009). If reconstructions have to be made for different depths in different regions, producing such maps will be problematic. This does not reduce the utility of the reconstructions for comparing with climate model output.

Planktonic foraminifera assemblage-based SST reconstructions have traditionally been assigned a priori to a fixed depth, typically 10 m. This contrasts with reconstructions based on foraminifera geochemistry, where the temperature change in their habitat may be possible to reconstruct accurately, but the depth and season which this represents is unclear a priori and needs to be estimated post hoc. Our results suggest that assignment of the signal in assemblage-based reconstructions to a particular depth and season needs to be treated in a similar manner to the geochemistry data.

4.3 Implications

If palaeoceanographic change can be thought of as the redistribution, expansion and contraction of water masses with fixed thermal properties, our results are but a curiosity. In this case, in a given water mass, the surface temperature is tied to the subsurface temperature which the foraminifera are responding to, so a surface reconstruction will be valid because of this correlation. Our comparison of CMIP5 LGM and PI ocean temperature data demonstrates that this simplistic notion of climate change is incorrect. In the model output, much of the tropical North Atlantic has a thermal structure in the LGM that is not found in the PI ocean. It is here that the consequences of foraminifera being sensitive to subsurface rather than surface temperatures are likely to be most severe, and it is here where reconstructions from different depths differ most. Perhaps not coincidentally, the tropical ocean is where LGM faunas have poor modern analogues (Mix et al., 1999) and where there is a large mismatch between climate model output and proxy data (Otto-Bliesner et al., 2009). For example, proxy data show an east–west gradient in the size of the LGM temperature anomaly in the tropical Atlantic which is not replicated in the model output (Otto-Bliesner et al., 2009). Our results suggesting large subsurface cooling in the western tropical Atlantic imply that this result may need revisiting.

The non-analogue ocean thermal structure has two consequences for planktonic foraminifera assemblage-based reconstructions: first that the uncertainty in the tropical reconstructions is likely to be underestimated, and second that reconstructions are likely to be biased. Therefore, we need to consider the sign and likely magnitude of this bias.

If the thermal gradient was weaker in the past, subsurface temperatures now associated with warm surface conditions would have been associated with cooler SST in the past (Fig. 7). Consequently, foraminifera assemblages now associated with warm SSTs because of their relationship with cooler subsurface conditions would in the past have been associated with cooler SSTs, and reconstructions will have a warm bias. Conversely, if the thermal gradient was steeper, SST reconstructions will have a cold bias.

Our reconstructions using transfer functions calibrated against different depths show marked subsurface cooling at some tropical sites, in contrast to previous reconstructions of a modest cooling. If this bias is widespread in tropical sites, it may be sufficient to produce a bias in estimates of climate sensitivity based on the difference between LGM and modern climate. A sensitivity study undertaken by Schmittner et al. (2011) finds that a global 0.5 °C bias in LGM ocean anomalies gives a 1 °C change in climate sensitivity. Hargreaves et al. (2012) find that there is a 1 : 1.2 relationship between modelled tropical LGM temperatures anomalies and climate sensitivity in an ensemble of PMIP2 models, and use this relationship to estimate climate sensitivity from LGM proxy data; if the LGM cooling has been underestimated by

1 °C, their estimate of climate sensitivity would be 1.2 °C too low. Both these results suggest that estimates of climate sensitivity are sensitive to biases in the proxy data, but because the sign of the bias in the foraminifera-SST reconstruction may vary by region, we cannot estimate the sign or magnitude of the global mean bias.

There has been some debate about which season planktonic foraminifera assemblages can be used to reconstruct (Kucera et al., 2005). Our results render this debate largely moot as they show that the signal in the foraminifera assemblage data is usually from the subsurface where seasonality is subdued relative to the surface.

4.4 Solutions

The obvious solution to the problem of planktonic foraminifera being sensitive to subsurface rather than near-surface temperatures in much of the North Atlantic would seem to be to calibrate planktonic foraminifera assemblages against a different, more ecologically relevant depth. This has been done, for example by Andersson et al. (2010) who reconstructed 100-m SST. Reconstructions of subsurface conditions could be interpreted in much the same way as surface reconstructions, and could be used as a target for estimating climate sensitivity. However, the most ecologically relevant depth varies in space and time, and the assemblages will probably integrate the communities from several depths and seasons, so selecting a more appropriate fixed depth for temperature reconstructions for each location is probably not trivial and does not completely circumvent the problem. The method we develop here can only be used to identify the most relevant depth from time series. It cannot be used for single assemblages.

Since the most appropriate depth probably changes with time, it might be useful to try to identify the most appropriate depth for each time period in each assemblage time series with a moving window analysis rather than using a single fixed depth. Assigning the signal in the assemblage data to a dynamic rather than fixed depth is probably not tractable. It will be difficult to demarcate periods where specific depths are optimal, and assemblages with poor analogues in the modern ocean may generate spurious results.

An alternative solution is to use forward modelling – ecological models that predict the planktonic foraminifera assemblages given the output of a climate model. The match between the fossil assemblage and the model assemblage could then be assessed directly rather than having to estimate temperatures from the assemblages to compare with model temperatures. Ideally, forward models of foraminifera assemblages would be run in conjunction with forward modelling of the geochemistry of planktonic foraminiferal tests (Schmidt and Mulitza, 2002) for a comprehensive solution. Forward modelling of planktonic foraminifera assemblages has been developed by Fraile et al. (2008) and Lombard et al. (2011), but these forward models are not yet sufficiently

comprehensive for assessing the fit between observed and modelled assemblages; for example, they include only a subset of taxa and are not resolved in depth. A more achievable short-term goal is to use forward models of foraminifera assemblages to help constrain the sign and likely magnitude of biases in SST reconstructions. Transfer functions, calibrated against 10-m SST, could be generated for simulated foraminifera assemblages forced by PI climate model output. These transfer functions could be used to reconstruct SST from simulated foraminifera assemblages forced by LGM conditions, and the spatial extent, sign and magnitude of any bias in the reconstructions determined.

4.5 Other proxies

A potential solution would seem to be to try to constrain foraminifera assemblage-based estimates of SST with other proxies. If proxies that represent different depth are compared, any discrepancies may indicate periods when the thermal structure of the upper ocean changed. However, discrepancies could also be due to changing seasonality of the production of the other proxies (Chapman et al., 1996).

The potential biases we discuss in this paper for planktonic foraminifera assemblage-based SST reconstructions in the North Atlantic will apply to other oceans and some other proxies. Assemblage or geochemistry data from taxa that are constrained to live in the photic zone because they, or their symbionts, are photosynthetic – for example diatoms (Koç Karpuz and Schrader, 1990) – can be used to calculate surface temperatures, but may generate biased results if the seasonality of the proxy production changes. Other micropalaeontological proxies that are not constrained to live in the photic zone, for example radiolarians (Pisias et al., 1997), risk the same biases as planktonic foraminifera transfer functions if the ocean thermal structure changes.

Estimates of SST from geochemical proxies on subsurface micropalaeontological proxies may also be biased. Mg/Ca ratios and $\delta^{18}\text{O}$ of foraminiferal tests may accurately record temperature at the time and depth of calcification, but if the seasonality or depth of calcification has changed (Cléroux et al., 2009), or the ocean thermal structure has changed, estimates of SST will be biased. Therefore, constraining the apparent recording depth of foraminifera transfer function SST reconstructions by comparison with other proxies is unlikely to yield sufficient confidence.

5 Conclusions

We present evidence that planktonic foraminifera assemblages can be more sensitive to subsurface temperatures than the 10-m SST they are usually calibrated against. Consequently, reconstructions of 10-m SST are likely to be biased, especially in the tropics where non-analogue ocean thermal structures occurred in the LGM. The sign and magnitude of

the bias is likely to vary regionally, probably with a warm bias in the tropical North Atlantic. Foraminifera-based reconstructions for other ocean basins remain to be assessed.

This problem exposes the limitations of using transfer functions to reconstruct past climate to compare with model output. The most promising solution is to use forward models of planktonic foraminifera assemblages which can be directly compared with observed fossil foraminifera assemblages; however, considerable work is needed to develop these models.

Supplementary material related to this article is available online at: <http://www.clim-past.net/9/859/2013/cp-9-859-2013-supplement.pdf>.

Acknowledgements. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modelling groups listed in Table 2 for producing and making available their model output. For CMIP the US Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. We thank three anonymous reviewers for their comments on a previous version of this paper.

Norwegian Research Council FriMedBio project palaeoDrivers (213607) helped support this work. CL acknowledges the support of the Centre for Climate Dynamics (SKD) at the Bjerknes Centre. This is publication no. A402 from the Bjerknes Centre for Climate Research.

Edited by: L. Beaufort

References

- Adloff, F., Mikolajewicz, U., Kučera, M., Grimm, R., Maier-Reimer, E., Schmiedl, G., and Emeis, K.-C.: Upper ocean climate of the Eastern Mediterranean Sea during the Holocene Insolation Maximum – a model study, *Clim. Past*, 7, 1103–1122, doi:10.5194/cp-7-1103-2011, 2011.
- Andersson, C., Pausata, F. S. R., Jansen, E., Risebrobakken, B., and Telford, R. J.: Holocene trends in the foraminifer record from the Norwegian Sea and the North Atlantic Ocean, *Clim. Past*, 6, 179–193, doi:10.5194/cp-6-179-2010, 2010.
- Birks, H. J. B., Heiri, O., Seppä, H., and Bjune, A.: Strengths and weaknesses of quantitative climate reconstructions based on Late-Quaternary biological proxies, *Open Ecol. J.*, 3, 68–110, 2010.
- Braconnot, P., Harrison, S. P., Kageyama, M., Bartlein, P. J., Abe-Ouchi, V. M.-D. A., Otto-Bliessner, B., and Zhao, Y.: Evaluation of climate models using palaeoclimatic data, *Nat. Clim. Change*, 2, 417–424, doi:10.1038/nclimate1456, 2012.
- Chapman, M. R.: Seasonal production patterns of planktonic foraminifera in the NE Atlantic Ocean: Implications for paleotemperature and hydrographic reconstructions, *Paleoceanography*, 25, PA1101, doi:10.1029/2008pa001708, 2010.
- Chapman, M. R., Shackleton, N. J., Zhao, M., and Eglinton, G.: Faunal and alkenone reconstructions of subtropical North Atlantic surface hydrography and paleotemperature over the last 28 kyr, *Paleoceanography*, 11, 343–357, 1996.
- Cléroux, C., Cortijo, E., Anand, P., Labeyrie, L., Bassinot, F., Caillon, N., and Duplessy, J. C.: Mg/Ca and Sr/Ca ratios in planktonic foraminifera: proxies for upper water column temperature reconstruction, *Paleoceanography*, 23, PA3214, doi:10.1029/2007PA001505, 2008.
- Cléroux, C., Lynch-Stieglitz, J., Schmidt, M. W., Cortijo, E., and Duplessy, J.-C.: Evidence for calcification depth change of *Globorotalia truncatulinoides* between deglaciation and Holocene in the Western Atlantic Ocean, *Mar. Micropaleontol.*, 73, 57–61, 2009.
- CLIMAP Project Members: The surface of the ice-age Earth, *Science*, 191, 1131–1137, 1976.
- Duplessy, J.-C., Labeyrie, L., Arnold, M., Paterne, M., Duprat, J., and van Weering, T. C. E.: Changes in surface salinity of the North Atlantic Ocean during the last deglaciation, *Nature*, 358, 485–488, 1992.
- Fairbanks, R. G., Wiere, P. H., and Be, A. W. H.: Vertical distribution and isotopic composition of living planktonic foraminifera in the Western North Atlantic, *Science*, 207, 61–63, 1980.
- Fraille, I., Schulz, M., Mulitza, S., and Kucera, M.: Predicting the global distribution of planktonic foraminifera using a dynamic ecosystem model, *Biogeosciences*, 5, 891–911, doi:10.5194/bg-5-891-2008, 2008.
- Ganssen, G. M. and Kroon, D.: The isotopic signature of planktonic Foraminifera from NE Atlantic surface sediments; implications for the reconstruction of past oceanic conditions, *J. Geol. Soc. Lond.*, 157, 693–699, 2000.
- Hargreaves, J. C., Annan, J. D., Yoshimori, M., and Abe-Ouchi, A.: Can the Last Glacial Maximum constrain climate sensitivity?, *Geophys. Res. Lett.*, 39, L24702, doi:10.1029/2012gl053872, 2012.
- Hegerl, G. C., Zwiers, F. W., Braconnot, P., Gillett, N. P., Luo, Y., Marengo Orsini, J. A., Nicholls, N., Penner, J. E., and Stott, P. A.: Understanding and Attributing Climate Change, in: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.
- Hemleben, C., Spindler, M., and Anderson, O. R.: *Modern planktonic foraminifera*, Springer, New York, 363 pp., 1989.
- Hüls, M.: Distribution of planktic foraminifera of sediment core M35003-4, doi:10.1594/PANGAEA.55756, 1999.
- Jansen, E., Andersson, C., Moros, M., Nisancioglu, K. H., Nyland, B. F., and Telford, R. J.: The early to mid-Holocene thermal optimum in the North Atlantic, in: *Natural Climate Variability and Global Warming – A Holocene Perspective*, edited by: Battarbee, R. W. and Binney, H. A., Wiley-Blackwell, Chichester, 123–137, 2008.
- Juggins, S.: *Rioja: An R Package for the Analysis of Quaternary Science Data*, Version 0.7-3, 2012.
- Knudsen, K. L., Jiang, H., Jansen, E., Eiriksson, J., Heinemeier, J., and Seidenkrantz, M. S.: Environmental changes off North Iceland during the deglaciation and the Holocene: foraminifera,

- diatoms and stable isotopes, *Mar. Micropaleontol.*, 50, 273–305, 2004.
- Knutti, R. and Hegerl, G. C.: The equilibrium sensitivity of the Earth's temperature to radiation changes, *Nat. Geosci.*, 1, 735–743, doi:10.1038/ngeo337, 2008.
- Koç Karpuz, N. and Schrader, H.: Surface sediment diatom distribution and Holocene paleotemperature variations in the Greenland, Iceland and Norwegian Sea, *Paleoceanography*, 5, 557–580, doi:10.1029/PA005i004p00557, 1990.
- Kucera, M., Weinelt, M., Kiefer, T., Pflaumann, U., Hayes, A., Weinelt, M., Chen, M.-T., Mix, A. C., Barrows, T. T., Cortijo, E., Duprat, J., Juggins, S., and Waelbroeck, C.: Reconstruction of the glacial Atlantic and Pacific sea-surface temperatures from assemblages of planktonic foraminifera: multi-technique approach based on geographically constrained calibration datasets, *Quaternary Sci. Rev.*, 24, 951–998, doi:10.1016/j.quascirev.2004.07.017, 2005.
- Labeyrie, L., Leclaire, H., Waelbroeck, C., Cortijo, E., Duplessy, J.-C., Vidal, L., Elliot, M., LeCoat, B., and Auffret, G.: Temporal variability of the surface and deep waters of the North West Atlantic Ocean at orbital and millennial scales, in: Mechanisms of global climate change at millennial time scales, edited by: Webb, R., Clark, P. U., and Keigwin, L. D., Geophysical Monograph, American Geophysical Union, 77–98, 1999.
- Levitus, S.: Climatological atlas of the world ocean, NOAA Professional Paper, 13, 1–173, 1982.
- Lombard, F., Labeyrie, L., Michel, E., Bopp, L., Cortijo, E., Retailleau, S., Howa, H., and Jorissen, F.: Modelling planktonic foraminifer growth and distribution using an ecophysiological multi-species approach, *Biogeosciences*, 8, 853–873, doi:10.5194/bg-8-853-2011, 2011.
- Lynch-Stieglitz, J., Curry, W. B., and Slowey, N.: Weaker Gulf Stream in the Florida Straits during the Last Glacial Maximum, *Nature*, 402, 644–648, 1999.
- Marchal, O., Cacho, I., Stocker, T. F., Grimalt, J. O., Calvo, E., Martrat, B., Shackleton, N., Vautravers, M., Cortijo, E., van Kreveld, S., Andersson, C., Koç, N., Chapman, M., Saffi, L., Duplessy, J.-C., Sarnthein, M., Turon, J.-L., Duprat, J., and Jansen, E.: Apparent long-term cooling of the sea surface in the northeast Atlantic and Mediterranean during the Holocene, *Quaternary Sci. Rev.*, 21, 455–483, 2002.
- MARGO Project Members: Constraints on the magnitude and patterns of ocean cooling at the Last Glacial Maximum, *Nat. Geosci.*, 2, 127–132, doi:10.1038/ngeo411, 2009.
- Mix, A. C.: Percentages of planktonic foraminifera species in sediment core V32-8, doi:10.1594/PANGAEA.355352, 2006a.
- Mix, A. C.: Percentages of planktonic foraminifera species in sediment core V30-36, doi:10.1594/PANGAEA.355347, 2006b.
- Mix, A. C.: Percentages of planktonic foraminifera species in sediment core V22-222, doi:10.1594/PANGAEA.355340, 2006c.
- Mix, A. C.: Percentages of planktonic foraminifera species in sediment core V30-51, doi:10.1594/PANGAEA.355351, 2006d.
- Mix, A. C.: Percentages of planktonic foraminifera species in sediment core V30-49, doi:10.1594/PANGAEA.355350, 2006e.
- Mix, A. C.: Percentages of planktonic foraminifera species in sediment core RC09-49, doi:10.1594/PANGAEA.355331, 2006f.
- Mix, A. C.: Percentages of planktonic foraminifera species in sediment core V25-75, doi:10.1594/PANGAEA.355345, 2006g.
- Mix, A. C., Ruddiman, W. F., and McIntyre, A.: Late Quaternary paleoceanography of the tropical Atlantic, I: Spatial variability of annual mean sea-surface temperatures, 0–20,000 years B.P., *Paleoceanography*, 1, 43–66, 1986.
- Mix, A. C., Morey, A. E., Pisias, N. G., and Hostetler, S. W.: Foraminifera faunal estimates of paleotemperature: circumventing the no-analog problem yields cool ice age tropics, *Paleoceanography*, 14, 350–359, 1999.
- Murray, J.: On the distribution of the pelagic foraminifera at the surface and on the floor of the ocean, *National Sci.*, 11, 17–27, 1897.
- Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., and Wagner, H.: *vegan: Community Ecology Package*, R package version 2.0-2, 2011.
- Otto-Bliesner, B. L., Schneider, R., Brady, E. C., Kucera, M., Abe-Ouchi, A., Bard, E., Braconnot, P., Crucifix, M., Hewitt, C. D., Kageyama, M., Marti, O., Paul, A., Rosell-Melé, A., Waelbroeck, C., Weber, S. L., Weinelt, M., and Yu, Y.: A comparison of PMIP2 model simulations and the MARGO proxy reconstruction for tropical sea surface temperatures at last glacial maximum, *Clim. Dynam.*, 32, 799–815, doi:10.1007/s00382-008-0509-0, 2009.
- Pflaumann, U., Duprat, J., Pujol, C., and Labeyrie, L. D.: SIMMAX: A modern analog technique to deduce Atlantic sea surface temperatures from planktonic foraminifera in deep-sea sediments, *Paleoceanography*, 11, 15–35, doi:10.1029/95pa01743, 1996.
- Pisias, N. G., Roelofs, A., and Weber, M.: Radiolarian-based transfer functions for estimating mean surface ocean temperatures and seasonal range, *Paleoceanography*, 12, 365–379, doi:10.1029/97PA00582, 1997.
- Prell, W. L.: The stability of low-latitude sea-surface temperatures: An evaluation of the CLIMAP reconstruction with emphasis on the positive SST anomalies, Department of Energy, Washington, DC, 60 pp., 1985.
- R Development Core Team: *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, 2011.
- Rao, C. R.: The use and interpretation of principal component analysis in applied research, *Sankhyā: The Indian Journal of Statistics, Series A*, 26, 329–358, 1964.
- Risebrobakken, B., Jansen, E., Andersson, C., Mjelde, E., and Hevroy, K.: Planktonic foraminiferal and stable oxygen isotope record of Holocene sediments from the Norwegian Sea, doi:10.1594/PANGAEA.760166, Supplement to: Risebrobakken et al., 2003: A high-resolution study of Holocene paleoclimatic and paleoceanographic changes in the Nordic Seas, *Paleoceanography*, 18, 1017, doi:10.1029/2002PA000764, 2003.
- Sachs, H. M., Webb, T., and Clark, D. R.: Paleocological Transfer Functions, *Ann. Rev. Earth Planet. Sci.*, 5, 159–178, doi:10.1146/annurev.ea.05.050177.001111, 1977.
- Sarnthein, M., van Kreveld, S. A., Erlenkeuser, H., Grootes, P. M., Kucera, M., Pflaumann, U., and Schulz, M.: Distribution of foraminifera of sediment core GIK23258-2. doi:10.1594/PANGAEA.114682, In Supplement to: Sarnthein et al., 2003: Centennial-to-millennial-scale periodicities of Holocene climate and sediment injections off western Barents shelf, 75° N, *Boreas*, 32, 447–461, doi:10.1111/j.1502-3885.2003.tb01227.x, 2003.

- Schmidt, G. A. and Mulitza, S.: Global calibration of ecological models for planktic foraminifera from coretop carbonate oxygen-18, *Mar. Micropaleontol.*, 44, 125–140, 2002.
- Schmittner, A., Urban, N. M., Shakun, J. D., Mahowald, N. M., Clark, P. U., Bartlein, P. J., Mix, A. C., and Rosell-Mele, A.: Climate Sensitivity Estimated from Temperature Reconstructions of the Last Glacial Maximum, *Science*, 334, 1385–1388, doi:10.1126/science.1203513, 2011.
- Telford, R. J.: palaeoSig: Significance Tests of Quantitative Palaeoenvironmental Reconstructions, R package version 1.1, 2012.
- Telford, R. J. and Birks, H. J. B.: The secret assumption of transfer functions: problems with spatial autocorrelation in evaluating model performance, *Quaternary Sci. Rev.*, 24, 2173–2179, doi:10.1016/j.quascirev.2005.05.001, 2005.
- Telford, R. J. and Birks, H. J. B.: A novel method for assessing the statistical significance of quantitative reconstructions inferred from biotic assemblages, *Quaternary Sci. Rev.*, 30, 1272–1278, 2011.
- Telford, R. J., Andersson, C., Birks, H. J. B., and Juggins, S.: Biases in the estimation of transfer function prediction errors, *Paleoceanography*, 19, Pa4014, doi:10.1029/2004pa001072, 2004.
- Wilke, I., Meggers, H., and Bickert, T.: Depth habitats and seasonal distributions of recent planktic foraminifers in the Canary Islands region (29° N) based on oxygen isotopes, *Deep Sea Res. Part I*, 56, 89–106, 2009.
- WOA: World Ocean Atlas 1998 Version 2, National Oceanographic Data Center, Silver Spring, Maryland, 1998.