Empirical estimate of the signal content of Holocene temperature proxy records

Maria Reschke[1,2], Kira Rehfeld[1,3], Thomas Laepple[1]

[1]Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Telegrafenberg A45, 14473 Potsdam, Germany
5 [2]Institute of Earth and Environmental Sciences, University of Potsdam, Karl-Liebknecht-Str. 24/25, 14476 Potsdam, Germany
[3]Institut für Umweltphysik, Ruprecht-Karls-Universität Heidelberg, Im Neuenheimer Feld 229, 69120 Heidelberg, Germany

Correspondence: M. Reschke (mreschke@awi.de)

10 **Abstract.** Proxy records from climate archives provide evidence about past climate changes, but the recorded signal is affected by non-climate related effects as well as time uncertainty. As proxy based climate reconstructions are frequently used to test climate models and to quantitatively infer past climate, we need to improve our understanding of the proxy records' signal content as well as the uncertainties involved.

In this study, we empirically estimate signal-to-noise ratios (SNRs) of temperature proxy records used in global compilations 15 of the mid to late Holocene. This is achieved through a comparison of proxy time series from close-by sites of three compilations and model time series data at the proxy sites from two transient Holocene climate model simulations.

In all comparisons, we found the mean correlations of the proxy time series on centennial to millennial time scales to be rather low (R < 0.2), even for nearby sites, which resulted in low SNR estimates. The estimated SNRs depend on the assumed time uncertainty of the proxy records, the time scale analysed, and the model simulation used. Using the correlation 20 structure of the ECHAM5/MPI-OM simulation, the estimated SNRs on centennial time scales the SNRs ranged from 0.05 - assuming no time uncertainty - to 0.5, for a time uncertainty of 400y. On millennial time scales, the estimated SNRs were generally higher. Use of the TraCE-21ka model simulation correlation structure resulted generally in lower SNR estimates than for ECHAM5/MPI-OM.

As the number of available high-resolution proxy records continues to grow, a more detailed analysis of the signal content of 25 specific proxy types should become feasible in the near future. The estimated low signal content of Holocene temperature records should caution against over-interpretation of these kinds of datasets until further studies are able to facilitate a better characterisation of the signal content in paleoclimate records.

1 Introduction

Improving our understanding of the climate systems and its variability requires knowledge about the climate of the pre-
30 instrumental period. Proxy records from different climate archives are available for determining past climate conditions (e.g.,

Bartlein et al., 2011; Huguet et al., 2006; Johnsen et al., 2001; Li et al., 2006; Luckman et al., 1997). However, as any observational estimate, paleoclimate proxies are affected by uncertainties (e.g., Breitenbach et al., 2012; Lohmann et al., 2013).

Paleoclimate proxies record very specific climate information that depends on temporal (seasonal recording, dating),
5 geological (mixing, transport, sorting), biological (life-time of organisms, habitat depth, bioturbation), and chemical (preservation and dissolution) processes (e.g., Bard, 2001; Berger and Heath, 1968; Goreau, 1980; Leduc et al., 2010; Lohmann et al., 2013; Mollenhauer et al., 2003; Ohkouchi et al., 2002; Rehfeld et al., 2016; Rosell-Melé and Prahl, 2013; Schneider et al., 2010; Telford et al., 2004; van Sebille et al., 2015).

In addition, the information of interest can be obscured by co-recorded non-climatic components which represents one
10 possible cause for misinterpretation of proxy signals, especially in systematic model-data comparisons and quantitative data synthesis efforts. Different approaches have been proposed in an effort to alleviate this problem and improve analyses:

    (i)   obtain a better statistical or mechanistic understanding of how and what a proxy actually records (e.g., Fisher et al., 1985; Grauel et al., 2013; Ho and Laepple, 2016; Münch et al., 2016, 2017; Richey et al., 2011; Rosén et al., 2003; Thirumalai et al., 2013);

15   (ii)   modelling of the proxy signal (e.g., Dee et al., 2011, 2015; Dolman and Laepple, 2018; Evans et al., 2013; Roche et al., 2018); and

    (iii)   through a detailed, expertise-driven analyses of single sites (e.g., Stebich et al., 2015).

In this study, we use a comparison of proxy records and model simulations to improve the characterisation of proxy uncertainties through empirical estimates of the signal-to-noise ratio (SNR) in temperature-related proxies. At present,
20 studies on SNRs in proxies are rare and mainly focussed on the instrumental period (e.g., Mann et al., 2007, 2008; Smerdon, 2012, Münch and Laepple, 2018). In contrast, the present study focusses on the pre-instrumental Holocene period which has received considerable attention in the community (e.g., Bakker et al., 2017; Gajewski, 2015; Mangerud and Svendsen, 2018; Marcott et al., 2013; Mischel et al., 2017; Moossen et al., 2015; Sejrup et al., 2016; Thibodeau et al., 2018; Wanner et al., 2015; Zhang et al., 2017). In particular, we focus on estimating SNRs in temperature-sensitive proxy records to improve
25 analyses of Holocene temperature evolution and variability. A better understanding of Holocene proxy time series SNRs will lead to improved and more reliable interpretations of proxy records and should raise awareness for the need of careful and critical evaluations of paleoclimate reconstructions.

2 Data

This study builds on existing compilations of recalibrated high-resolution Holocene temperature-sensitive proxy records to
30 facilitate intercomparisons of multiple time series. The analysis is based on three proxy datasets and two model simulations to test the robustness of our results and the sensitivity to the choice of a particular climate model.

2.1 Proxy records

We focus on globally distributed multi-archive and multi-proxy compilations of the Holocene temperature evolution from a wide variety of locations (Fig. 1a-c, Tab. 1; Tab. S1-S3), namely

    (1) M13: the compilation of Marcott et al. (2013) that was used to reconstruct the global and regional temperature

5   evolution of the past 11.3ky;

    (2) LH14: Uk37 and Mg/Ca proxy data compiled in the extended dataset of Laepple and Huybers (2014) that was used to reconstruct regional temperature variability and builds on the compilation of Leduc et al. (2010);

    (3) R18: the Holocene part of the compilation of Rehfeld et al. (2018) that was used to compare Glacial and Holocene temperature variability.

10 The datasets mostly originated from marine sediment cores and the proxy types include Uk37, planktonic foraminifera Mg/Ca, TEX_{86}, terrestrial bio-indicators (fossil pollen modern analogue technique, fossil chironomid transfer function), ice-core stable isotopes ($\delta^{18}O$, δ^2H) and several others (microfossil assemblages, MBT, tree ring width). As the early Holocene was influenced by deglaciation following the Last Glacial Maximum (e.g., Kaplan and Wolfe, 2006), we restricted the time series to the last 6ky (6ky BP to present day, where BP denotes years before 1950). We only analysed time series containing

15 climate information on at least centennial to millennial time scales (i.e., a mean inter-observation time step of $\Delta t < 500y$). Due to the limited number of available high-resolution time series, all datasets overlap (Tab. 1) to some degree and are thus not independent.

2.2 Climate model simulations

We analysed surface air temperature data from simulations of two coupled atmosphere-ocean general circulation models: A

20 6ky transient Holocene simulation from ECHAM5/MPI-OM (henceforth abbreviated as MPI6k) (Fischer and Jungclaus, 2011) and the TraCE-21ka (T21k) (Liu et al., 2009) simulation from the CCSM4 model, both of which have been used frequently in recent studies (e.g., Gregoire et al., 2016; Heinemann et al., 2009; Koldunov et al., 2010; Lu et al., 2018; Matei et al., 2012; May, 2008; Müller and Roeckner, 2008; Pausata and Löfverström, 2015; Werner et al., 2016). For the present study, annual means of temperatures of both model simulations were extracted at the nearest grid-box related to the proxy

25 record locations of M13, LH14 and R18.

MPI6k (Fischer and Jungclaus, 2011) is a 6ky transient run using ECHAM5/MPI-OM (Jungclaus et al., 2006) which consists of the atmosphere component ECHAM5 (Roeckner et al., 2003), the ocean component MPI-OM (Marsland et al., 2003), and the land surface model JSBACH (Raddatz et al., 2007) with dynamic vegetation module (Brovkin et al., 2009). The model outputs atmospheric variables on a regular longitude/latitude model grid with 96 by 48 horizontal grid-boxes (T31 resolution

30 corresponding to 3.75° in latitude/longitude). The simulation is forced only orbitally with greenhouse gas concentrations set to pre-industrial values. We extracted annual mean surface temperatures at an elevation of 2m from this model (temp2).

The TraCE-21ka dataset (Liu et al., 2009) originated from a simulation of the transient climate between 22ky BP and 1990 CE is based on a fully-coupled CCSM3 simulation with an atmospheric resolution of T31_gx3 (96 by 48 horizontal grid corresponding to 3.75° in latitude/longitude). Transient forcing factors included changes in the orbitally driven insolation, greenhouse gas concentrations, continental ice sheets and coastlines, and meltwater fluxes across the North Atlantic and Gulf

5 of Mexico. For the 6ky BP period, the changes in orbital insolation and changes in the greenhouse gas concentrations represent the most important forcings while changing continental ice sheets and meltwater fluxes are only of minor importance. Our analysis of the variability of annual surface temperatures is independent of the absolute changes and only relies on the simulated spatial correlation structure. For the time scales analysed and the proxy positions of our compilations, this correlation structure is not sensitive to the particular choice of temperature variable (sea surface temperature versus

10 surface temperature or near surface air temperature).

3 Method

3.1 Approach and assumptions

SNRs can be estimated by comparing proxy records that experienced the same or very similar climate signals, e.g., different proxies from the same site or the same proxy from different sites in close spatial proximity. If a pair of records contains the

15 same signal, an independent local noise component, and no time uncertainty, the SNR is given as $R/(1 - R)$ where R is the correlation between both time series (Fisher et al., 1985). Ideally, SNRs would be estimated from local replicates. This is often difficult, or impossible, due to the limited availability of replicated datasets. To increase the number of records and thus improve the robustness of estimates, we extended this approach to also include records from locations that are further apart. This increased spatial separation between sites requires knowledge of the signal covariance (as the climate signal will have

20 been slightly different at each location) and we rely on climate models to provide this information.

The underlying assumptions are thus: (1) when relying on model data, we must assume correctness of the model-based correlation structure; (2) when using different proxies, we must assume that all proxies recorded the same temporal (and spatial) variability of the climate signal (more specifically: annual mean surface temperature); (3) we must assume that differences in the spatial correlation structure between models and proxy observations are due solely to a site-independent

25 additive noise and time uncertainty.

Based on these assumptions, we can estimate the SNR by matching the spatial correlation of proxy records and model time series while accounting for time uncertainty and additive noise which can both lead to a deterioration in the spatial correlation. E.g., low correlations among time series can be caused by both: a low time uncertainty in combination with a high noise level and a high time uncertainty in combination with a high SNR (low noise level). Due to their relationship, we

30 quantify SNR estimates as a function of time uncertainty.

Sites that are very far apart only share a weak climate signal which does not represent any constraint on the SNR as both the climate and proxy correlations will be close to zero. For our SNR estimate, we therefore only included proxy pairs with

spatial separations of up to 5000 km, which we found to be a typical decorrelation distance on centennial time scales in the model simulations as we later show.

As climate variability is a function of time scale, we expect that both the spatial correlation structure and SNR will also be time scale dependent. However, the limited number of records and samples in each record prevents a more thorough time

5   scale dependent estimate which could be carried out using a spectral approach for instance (Münch and Laepple, 2018). In order to balance accounting for time scale and estimate robustness, we distinguish between a centennial time scale T_{cent} (with a cut-off frequency of 1/400y and by removing the linear trend of the time series) and a centennial to millennial time scale T_{mill} (using a cut-off frequency of 1/1000y and including the trend) (cf., Reschke et al., in review). To estimate T_{cent}, we only used records with a mean sampling interval of less than 200y while all records were included for estimating T_{mill}.

10   **3.2 Processing steps**

3.2.1 Estimation of the spatial correlation structure

From the MPI6k and T21k model time series we extracted annual mean temperatures at those grid cells that contain the location of the proxy record site. Through block averaging, data were then resampled at the same ages as in the corresponding proxy records. We chose to use averages rather than interpolation because sediment and ice samples, in

15   particular, often include adjacent depths or have a sample distance that is smaller than the typical mixing depth in the sediment or diffusion length in ice-cores. For each dataset (M13, LH14, R18), we estimated the time scale dependent (T_{cent}, T_{mill}) correlations between all possible proxy record pairs and between all model time series pairs. For this step, the irregularly sampled time series were linearly interpolated onto a more regular grid and subjected to a Gaussian filter. This approach has been shown to deliver good results for an irregularly sampled dataset with regard to time scale dependent

20   correlations (Reschke et al., in review).

The spatial separation between two sites was used to place the pair into 2000 km-sized bins (thus containing separations of 0-2000km, 2000-4000km, etc.) and averaging the correlations from proxy/model site pairs contained within the same bin. An overview of the processing steps is given in Fig. 2.

We performed a significance test of the spatial correlation structure based on spatially uncorrelated surrogate time series

25   with a temporal power-law scaling of $\beta = 1$, which is a typical value for Holocene sediment records (Laepple and Huybers, 2014). In a Monte Carlo procedure with 1000 repetitions, we generated annual surrogate records that were analysed using the same procedure as the true proxy observations, using the 90% quantile of the binned correlations of the surrogate time series as confidence intervals.

3.2.2 Estimation of the SNRs

30   The SNR estimate was obtained from a Monte Carlo simulation with 1000 repetitions. Through block averaging, we resampled the annual model data at the same resolution as the corresponding proxy records. We then added time

uncertainties (between 0 and 400y) and noise levels ($0.01 < SNR < 100$), before estimating the mean correlation using the interpolation method of Reschke et al. (in review). We estimated the SNRs as a function of the time uncertainty by minimising the absolute difference of the mean correlations of proxy records and modified model simulations.

We generated the modified model data by separately distorting the time axis and adding noise to the observations of the resampled model time series. As a simple heuristic to simulate time uncertainty, we defined four time control points at the ages of 1y, 2ky, 4ky, and 6ky and randomly shifted these points by adding a random value from a normal distribution (mean $\mu = 0$, standard deviation σ = time uncertainty [y]) except for the value of 1y. The new time axis was then created by linearly interpolating between the time control points. Noisy observations were generated by adding normally distributed noise, ε (with $\varepsilon \sim N$, mean $\mu = 0$ and variance σ^2 as $\frac{\sigma^2_{model,resampled}}{SNR}$), to the resampled model time series. Fig. 2 gives an overview of the processing steps.

4 Results

4.1 Spatial correlation structure and correlation decay length

The correlation analysis using all proxy types and locations yielded, unsurprisingly, a general decrease in correlation for larger spatial separations between proxy sites (Fig. 3). Both model simulations exhibit statistically significant spatial correlations at both analysed time scales (T_{cent} and T_{mill}) and for most inter-site separation distances. Throughout all datasets and separation distances, T21k yielded higher correlations than MPI6k, which is consistent with the generally higher correlation decay lengths l_d for T21k, estimated at grid cell level by fitting an exponential, $R = e^{-x/l_d}$, to the decay of correlations R as a function of site separation x (Fig. 1d, e).

While for T_{cent} the correlation of both model simulations decreases with increasing site separation (Fig. 3a-c), the T_{mill} estimate (Fig. 3d-f) shows a more complex pattern that includes a partial increase in correlation for separation distances larger than 8000 km. This is likely related to variations in orbital forcing affecting the temperature trend that is symmetric between both hemispheres. Especially for MPI6k, the correlation is weak for separation distances from 4000 to 6000 km.

The spatial correlations obtained from the proxy records differ systematically from those obtained from model simulation data. The mean correlation for close proxy site pairs (separation <5000 km) was 0.004 to 0.014 for T_{cent} and 0.101 to 0.186 for T_{mill} and thus lower than for model data (MPI6k: 0.303 to 0.338 for T_{cent}, 0.202 to 0.461 for T_{mill}; T21k: 0.634 to 0.719 for T_{cent}, 0.674 to 0.710 for T_{mill}). For T_{cent}, none of the proxy based correlations is statistically significant and no clear pattern emerges with regard to separation distance. All three datasets yielded a statistically significant correlation at T_{mill} for smaller separation distances, although visibly decreasing for longer separation distances (e.g., 6000-8000 km; cf. Fig. 3d, f).

Comparisons of temperature estimates from different proxy types are faced with the additional challenge that the actual recorded variable (e.g., summer atmospheric temperature vs. mixed layer winter temperature) may depend on the proxy type. We therefore also analysed the proxy-specific results (Fig. 4, Tab. 2). By performing separate analyses for each proxy type

(instead of analysing all proxies together) we obtained in all three datasets a higher mean correlation on the T_{mill} time scale for sites within a 5000 km range. For T_{mill}, the proxy-specific mean correlations across all datasets and proxies are between 0.149 and 0.357 compared to 0.101 to 0.186 when correlating sites across proxy types. For T_{cent}, most correlations are indistinguishable from zero and we observed no consistent increase when analysing proxy-specific correlations (Tab. 2).

5   Unfortunately, restricting the analysis to a single proxy type greatly reduces the number of available proxy pairs at any given distance and thus leads to less robust correlation estimates and rather large confidence intervals. We therefore only provide results for the most data-abundant proxy types Uk37 and Mg/Ca and one dataset (LH14) as example in the main manuscript (Fig. 4). The remaining data are shown in the Supplement (Fig. S1-S4). For LH14, both Mg/Ca and Uk37 show a decrease in correlation with increasing separation distance for both time scales. The correlations in this proxy-specific analysis are

10   stronger than the analysis across proxy types (Fig. 3). They are, however, only statistically significant for Uk37 on T_{mill} with separation distances smaller than 5000 km and for a single distance bin (2000-4000km) for Mg/Ca.

4.2 SNR estimates

The estimated SNRs of proxy records are a function of time uncertainty because correlations deteriorate due to both, time uncertainty and noise. In general, we found that low (high) SNRs were related to low (high) time uncertainties (Fig. 5). In

15   most cases, the estimated signal content for Holocene temperature-sensitive proxy records was quite low (<0.5).

By using the spatial correlation structure of MPI6k and assuming a time uncertainty (1 sd) of 220y (mean uncertainty in M13) we obtain an estimated SNR of between 0.05 and 0.2 for the T_{cent} time scale and 0.2 for the M13 and R18 datasets on the T_{mill} time scale. The LH14 dataset yielded a SNR of 10 at the T_{mill} time scale.

The M13, LH14 and R18 SNRs obtained for mixed proxy types differ for the simulation data. Generally, the SNRs were

20   more consistent with T21k compared to MPI6k, although lower due to higher correlations of spatially close (separation <5000 km) pairs (Fig. S5). In addition, the SNRs are lower for T_{cent} ($SNR_{T21k,T_{cent}}$ <0.05) compared to T_{mill} (0.05 < $SNR_{T21k,T_{mill}}$ < 0.2).

An analysis of proxy-specific SNRs yielded higher uncertainties due to the relatively small number of record pairs (see Fig. S6-S15 for the complete set of results). The dependence of SNR estimates on time uncertainty is very sensitive to how the

25   proxies are compiled and the type of model simulation. However, the overview of all proxy-specific SNR estimates (Fig. 6) suggests some proxy-specific tendencies. On T_{cent} ice-cores show the highest SNR. Mg/Ca shows a high SNR for the LH14 dataset but a low SNR in the two other compilations. Uk37 and terrestrial bio-indicators have the lowest SNR estimate on this time scale. In contrast analysing the T_{mill} time scale that also includes trends in the dataset leads to different results; Uk37 shows the highest SNRs whereas the other proxy types only show a small increase compared to the T_{cent} analysis.

5 Discussion

High-resolution temperature-sensitive proxy records for the Holocene are sparse, irregularly distributed, and rely on different proxy types. Thus, estimating the SNR in such datasets requires some simplifying assumptions. We assumed that: (1) the spatial correlation of the climate model simulations was realistic, (2) all proxy types were recording the same climate variable, and (3) any non-climatic components of the proxy signal can be fully accounted for through a combination of time uncertainty and additive noise.

The SNRs we estimated, based on these assumptions, generally suggest a low signal content of Holocene temperature records on centennial time scales (T_{cent}). We found a higher signal content on millennial time scales (T_{mill}), but the results were rather sensitive to the choice of the proxy compilation and model simulation. We now discuss the role of the different assumptions on the results.

5.1 Spatial correlation structure of model simulations

Our SNR estimates critically depend on the model-based temperature correlation structure as lower spatial temperature correlations in the models would lead to higher SNR estimates for the proxies and vice versa. In most regions, the model simulation MPI6k shows correlation decay lengths of 1295 to 6030 km (mean decay length: 3995 km) and the correlation decay length of T21k is generally in the range of 2130 and 8705 km (mean decay length: 5920 km) for the T_{cent} time scale (Fig. 1d, e). This is higher than previous estimates of correlation decay lengths from instrumental datasets in the range of 1000 to 3000 km (e.g., Hansen and Lebedeff, 1987; Jones et al., 1997; Madden et al., 1993), but plausible as an increase with time scale is to be expected. For example, Jones et al. (1997) found lower correlation decay lengths related to annual (2100 km) than to decadal (3800 km) time scales.

Nevertheless, spatial correlation could be overestimated in the model simulations for two reasons. Firstly, the spatial correlation of instrumental datasets always includes the anthropogenic forcing which strongly increases the correlation decay length. This effect is not or only weakly present in the 6ky time-period of our analysis. Instrumental records from the industrial period and pre-industrial model simulations might thus be in agreement for the wrong reasons. Secondly, the grid cell size of the models was on the order of several hundred kilometres whereas the records might be representative of a smaller spatial area. Hence, it is possible that proxy based correlations are lower compared to those obtained from the model due to the former being influenced by subgrid-scale temperature variations.

We also found that T21k yielded higher spatial correlations compared to MPI6k (Fig. 1d, e) which in turn resulted in lower SNR estimates if relying on this particular model simulation (Fig. S5). This difference might be related to the presence of transient greenhouse gas forcing in T21k (Timm and Timmermann, 2007), although the changes in forcing were small during the analysed time period.

Thus, there remain the possibility that the true temperature variations are more localised than suggested by the model simulations. In this case our estimates of the proxy signal content would be pessimistic. Ultimately, more replicate proxy records are needed to distinguish between these hypotheses.

5.2 Finite number of proxy records

5   Despite the strong overlap among records, we found our estimates of the spatial correlation structure and SNRs to be sensitive to the choice of proxy compilation (Tab. 2), which suggests that the number of records may have been limiting the robustness of our estimates. To test this hypothesis, we performed a sensitivity analysis using surrogate time series. N 6ky annual surrogate time series were generated from the sum of a common pseudo climate time series modelled as a random process that follows a power-law ($\beta = 1$) scaling and a separate non-climate component that is simulated as uncorrelated

10   white noise. The noise amplitude is chosen to yield SNR = 0.15. Irregular sampling times were used to mimic the observed sampling times of the M13 records. Surrogate inter-observation time steps were drawn from a gamma distribution (shape = rate = 2.25), rescaled with a mean inter-observation time step of 108.56y (cf., Reschke et al., in review). The final, pseudo proxy time series were then obtained by block averaging the annual time series to the irregular sampling times. The SNR of the surrogate time series were then calculated following the same method as the proxy records in the main study and

15   repeated for different sites using a Monte Carlo-based procedure with 2000 repetitions.

We found that the uncertainty of SNR estimates that are based on a small number of records can be high (Fig. 7). For a low number of only 15 records (105 correlation pairs), for instance, the uncertainty range of SNRs (90% quantiles of 0.08 to 0.26) is higher than the true SNR value of 0.15. Although we used more than 15 sites per compilation in our analysis (Fig. 7), there were often fewer than 15 time series per proxy type (Tab. 1) which might explain the strong scatter in the proxy-

20   type-specific SNR estimates.

To improve the robustness of SNR estimates, it is unavoidable to significantly increase the number of records that are collected not too far apart from one another (distances <5000 km). Additionally, a better global coverage of sites locations would likely lead to more robust results. Since we sampled the models at the locations of the proxy sites our results should be independent of the spatial sampling distribution if the models were perfect. In reality, however, spatial differences and

25   shifts in the simulated correlation structure are likely and can be overcome by sampling from a wide variety of sites from all over the globe.

5.3 Proxy-specific recording of climate variables

All proxy types used in this study have been reported in the literature as temperature-sensitive and are usually calibrated to the mean annual surface air or surface water temperature. However, this is a gross oversimplification as the true climate

30   variable influencing the recorded signal is proxy-specific and generally more complex. For example, signals reconstructed from marine organism-based proxies such as Mg/Ca, Uk37, and TEX_{86} are affected by the seasonal and depth-specific

preferred habitat of the organism (Ho and Laepple, 2016; Leduc et al., 2010; Lohmann et al., 2013). The consequence is a temporal and spatial averaging of the climate signal.

Thus, analysing different proxy types likely leads to an underestimation of the spatial temperature correlations. Indeed, in our study we found the spatial correlations related to records of the same proxy type for T_{mill} to be higher compared to those

5   for all types (Tab. 2). To gain a better understanding of proxies and their effect on the analyses, we suggested to use proxy-specific SNR estimates instead. However, this is hampered by the low number of records in close proximity to one another (Tab. 1; Tab. S4). For many proxy types, this leads to statistically non-significant correlations and unreliable SNR estimates (Fig. S6-S15).

On the other hand, analysing the spatial correlation among records of the same proxy type can also lead to overly optimistic

10   results as the correlation among records of the same proxy type could also stem from spatially correlated proxy-specific non-climatic components. A case in point would be the dissolution of foraminiferal shells (Lea, 2003) which could generate spatially correlated noise as the preferential dissolution of carbonate depends on the water depth (Brown and Elderfield, 1996; Dekens et al., 2002), the carbonate ion concentration, and the salinity of the surrounding seawater (Huguet et al., 2006; Lea, 2003; Spero et al., 1997).

15   **5.4 Time uncertainty and non-climatic components of the proxy signal**

Our SNR estimates depend on the assumed time uncertainty of the records. While we assumed a mean time uncertainty of 220y (as provided in the M13 dataset), the true time uncertainty for marine records might be considerably higher due to spatially varying reservoir effects (Ascough et al., 2005). This would imply that our SNR estimates are conservative, especially on centennial time scales. On the other hand, using one mean uncertainty value will clearly be too pessimistic for

20   ice-core data that is only subject to much smaller dating uncertainties. Using more sophisticated models to account for time uncertainty (e.g., Blaauw, 2010; Blockley et al., 2007; Telford et al., 2004) and the proxy- and site-specific information on the chronologies would allow to obtain more precise SNR estimates.

We modelled the transfer function between the temperature time series and the calibrated proxy records as a combination of time uncertainty and additive temporally uncorrelated noise. Our approach thus neglects other distortions of the signal and

25   non-additive parts of noise. Multiplicative noise can arise from aliasing due to subsampling that leads to errors that are proportional to short-term climate variability (Laepple and Huybers, 2013). Variable sedimentation rates, bioturbation and/or bioturbation depths varying over time have a low-pass filtering effect that is similar to irregular sampling. Proxy archive accumulation processes undergo temporal changes due to changes in bioturbation depths and advection (Mollenhauer et al., 2003) or spatial changes in ocean currents (van Sebille et al., 2015) that could introduce additional non-additive noise in the

30   obtained proxy records. Finally, even for a single proxy type, the data quality (i.e., the signal content) is site specific and will depend on the sampling and measurement protocol. For example, the SNR estimates using the LH14 dataset that is mainly based on very high-resolution records (mean sample distance <100y), are higher than estimates based on the two larger proxy compilations. Thirumalai et al. (2018) showed that foraminiferal records based on a large number (70-100) of

foraminiferal tests per sample were consistent between cores collected in close proximity to one another leading to much higher correlations compared to our study. As we rely on datasets of opportunity that consist of proxy records measured by various labs over a period of two decades, it seems conceivable that a small number of records could be of a relatively lower quality which would reduce our mean correlation and thus the SNR estimate. New studies, especially when based on a
5   careful design (Thirumalai et al., 2018), could help alleviate this situation.

6 Conclusion

In this study, we estimated SNRs of Holocene temperature-sensitive proxy records by comparing proxy- and model-based spatial correlations. We found that spatial correlations between proxy records were significantly lower than those computed for temperature time series extracted from climate models. Simply put, the proxy records varied more independently from
10   site to site, whereas the model simulations suggested spatially coherent temperature variations. This in turn led to low SNR estimates in multi-proxy-type analyses if we assume that the correlation structure that we obtained from the model simulations is reasonable.

The low SNRs of Holocene proxy records are likely the result of processes occurring during the formation, preservation and measurement of the proxy signal. For Holocene, even small uncertainties in the process chain between the climate signal and
15   the climate reconstruction play an important role compared to the small temperature variations. In addition, as evidenced by the difference when comparing results between proxy types and within one proxy type, the proxy-specific recording of different temporal and spatial parts of the temperature (for example summer vs. winter) also affects the SNR of multi-proxy datasets. Nevertheless, our SNR estimates are still relevant for synthesis and model comparison efforts (e.g. Marcott et al., 2013), that usually interpret all proxy records together.

20   The precision of the SNR estimates is strongly dependent on the number of available proxy records. Due to the small number of spatially close records of the same proxy type, the uncertainty in our proxy-type-specific SNR estimates was very high.

Our SNR estimates implicitly depend on the expected time uncertainty, as well as on the model choice. However, for both tested models the multi-proxy-type estimates on centennial time scales (T_{cent}) were smaller ($SNR_{MPI6k,T_{cent}} < 0.5$; $SNR_{T21k,T_{cent}} < 0.05$) than on longer time scales T_{mill} ($SNR_{MPI6k,T_{mill}} \approx 0.2$; $0.05 < SNR_{T21k,T_{mill}} < 0.2$).

25   Our results of the low signal content of multi-proxy datasets, especially on centennial time scales, suggests that caution and a critical evaluation are in order when analysing and interpreting these datasets. Furthermore, optimising the sampling and measurement procedure is likely needed to faithfully reconstruct small climate variations over the Holocene. As the number of high-resolution proxy records continues to grow, a more detailed analysis of the signal content of specific proxy types and a model-independent estimate of the spatial correlation structure of climate variations will get feasible and will enable and
30   improve prospects for the interpretation and reconstruction of past climate changes.

Code availability. R scripts to estimate the spatial correlation structure and the signal-to-noise ratios are available on request from the authors and will be made available in an online repository after the manuscript is accepted.

Data availability. All datasets used in this study are available from the authors on request and will be made available in PANGAEA after the manuscript is accepted.

5 *Author contributions.* M.R. and T.L. designed the research. M.R. performed the analysis and wrote the first draft of the manuscript. All authors contributed to the interpretation and to the preparation of the final manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We would like to acknowledge Igor Kröner for fruitful discussions and comments on the manuscript. This study was supported by the Initiative and Networking Fund of the Helmholtz Association Grant VG-NH900 as well as
10 the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement no. 716092). It further contributes to the German BMBF project PALMOD. K.R. acknowledges funding by the German Research Foundation grants DFG RE3994-1/1 and DFG RE3994-2/1.

References

Ascough, P., Cook, G., and Dugmore, A.: Methodological approaches to determining the marine radiocarbon reservoir
15   effect, Prog. Phys. Geog., 29(4), 532-547, 2005.

Bakker, P., Clark, P. U., Golledge, N. R., Schmittner, A., and Weber, M. E.: Centennial-scale Holocene climate variations amplified by Antarctic Ice Sheet discharge, Nature, 541, 72-76, doi:10.1038/nature20582, 2017.

Bard, E.: Paleoceanographic implications of the difference in deep-sea sediment mixing between large and fine particles, Paleoceanography, 16(3), 235-239, 2001.

20 Bartlein, P. J., Harrison, S. P., Brewer, S., Connor, S., Davis, B. A. S., Gajewski, K., Guiot, J., Harrison-Prentice, T. I., Henderson, A., Peyron, O., Prentice, I. C., Scholze, M., Seppä, H., Shuman, B., Sugita, S., Thompson, R. S., Viau, A. E., Williams, J., and Wu, H.: Pollen-based continental climate reconstructions at 6 and 21 ka: a global synthesis, Clim. Dyn., 37, 775-802, doi:10.1007/s00382-010-0904-1, 2011.

Berger, W. H., and Heath, G. R.: Vertical mixing in pelagic sediments, J. mar. Res., 26, 134-143, 1968.

25 Blaauw, M.: Methods and code for 'classical' age-modelling of radiocarbon sequences, Quat. Geochronol., 5, 512-518, 2010.

Blockley, S. P. E., Blaauw, M., Bronk Ramsey, C., and van der Plicht, J.: Building and testing age models for radiocarbon dates in Lateglacial and Early Holocene sediments, Quaternary Sci. Rev., 26, 1915-1926, 2007.

Breitenbach, S. F. M., Rehfeld, K., Goswami, B., Baldini, J. U. L., Ridley, H. E., Kennett, D. J., Prufer, K. M., Aquino, V. V., Asmerom, Y., Polyak, V. J., Cheng, H., Kurths, J., and Marwan, N.: COnstructing Proxy Records from Age models (COPRA), Clim. Past, 8, 1765-1779, doi:10.5194/cp-8-1765-2012, 2012.

Brovkin, V., Raddatz, T., Reick, C. H., Claussen, M., and Gayler, V.: Global biogeophysical interactions between forest and climate, Geophys. Res. Lett., 36, L07405, doi:10.1029/2009GL037543, 2009.

Brown, S. J., and Elderfield, H.: Variations in Mg/Ca and Sr/Ca ratios of planktonic foraminifera caused by postdepositional dissolution: Evidence of shallow Mg-dependent dissolution, Paleoceanography, 11(5), 543-551, 1996.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. R. Meteorol. Soc., 137, 553-597, 2011.

Dee, S., Emile-Geay, J., Evans, M. N., Allam, A., Steig, E. J., and Thompson, D. M.: PRYSM: An open-source framework for PRoxY System Modeling, with applications to oxygen-isotope systems, J. Adv. Model. Earth Syst., 7, 1220-1247, doi:10.1002/2015MS000447, 2015.

Dekens, P. S., Lea, D. W., Pak, D. K., and Spero, H. J.: Core top calibration of Mg/Ca in tropical foraminifera: Refining paleotemperature estimation, Geochem. Geophys. Geosyst., 3(4), 1-29, doi:10.1029/2001GC000200, 2002.

Dolman, A. M., and Laepple, T.: Sedproxy: a forward model for sediment archived climate proxies, Clim. Past Discuss., 1-31, doi:10.5194/cp-2018-13, in review, 2018.

Evans, M. N., Tolwinski-Ward, S. E., Thompson, D. M., and Anchukaitis, K. J.: Applications of proxy system modeling in high resolution paleoclimatology, Quaternary Sci. Rev., 76, 16-28, 2013.

Fischer, N., and Jungclaus, J. H.: Evolution of the seasonal temperature cycle in a transient Holocene simulation: orbital forcing and sea-ice, Clim. Past, 7, 1139-1148, doi:10.5194/cp-7-1139-2011, 2011.

Fisher, D. A., Reeh, N., and Clausen, H. B.: Stratigraphic noise in time series derived from ice cores, Ann. Glaciol., 7, 76-83, 1985.

Gajewski, K.: Impact of Holocene climate variability on Arctic vegetation, Global Planet. Change, 133, 272-287, 2015.

Goreau, T. J.: Frequency sensitivity of the deep-sea climatic record, Nature, 287, 620-622, 1980.

Grauel, A.-L., Leider, A., Goudeau, M.-L. S., Müller, I. A., Bernasconi, S. M., Hinrichs, K.-U., de Lange, G. J., Zonneveld, K. A. F., and Versteegh, G. J. M.: What do SST proxies really tell us? A high-resolution multiproxy (UK'37, TEXH86 and foraminifera δ18O) study in the Gulf of Taranto, central Mediterranean Sea, Quaternary Sci. Rev., 73, 115–131, doi:10.1016/j.quascirev.2013.05.007, 2013.

Gregoire, L. J., Otto-Bliesner, B., Valdes, P. J., and Ivanovic, R.: Abrupt Bølling warming and ice saddle collapse contributions to the Meltwater Pulse 1a rapid sea level rise, Geophys. Res. Lett., 43, 9130-9137, doi:10.1002/2016GL070356, 2016.

Hansen, J., and Lebedeff, S.: Global Trends of Measured Surface Air Temperature, J. Geophys. Res., 92(D11), 13345-13372, 1987.

Heinemann, M., Jungclaus, J. H., and Marotzke, J.: Warm Paleocene/Eocene climate as simulated in ECHAM5/MPI-OM, Clim. Past, 5, 785-802, 2009.

Ho, S. L., and Laepple, T.: Flat meridional temperature gradient in the early Eocene in the subsurface rather than surface ocean, Nat. Geosci., 9(8), 606-610, doi:10.1038/NGEO2763, 2016.

Huguet, C., Kim, J.-H., Damsté, J. S. S., and Schouten, S.: Reconstruction of sea surface temperature variations in the Arabian Sea over the last 23 kyr using organic proxies (TEX86 and UK′37), Paleoceanography, 21, PA3003, doi:10.1029/2005PA001215, 2006.

Johnsen, S. J., Dahl-Jensen, D., Gundestrup, N., Steffensen, J. P., Clausen, H. B., Miller, H., Masson-Delmotte, V., Sveinbjörnsdottir, A. E., and White, J.: Oxygen isotope and palaeotemperature records from six Greenland ice-core stations: Camp Century, Dye-3, GRIP, GISP2, Renland and NorthGRIP, J. Quaternary Sci., 16(4), 299-307, doi:10.1002/jqs.622, 2001.

Jones, P. D., Osborn, T. J., and Briffa, K. R.: Estimating Sampling Errors in Large-Scale Temperature Averages, J. Climate, 10, 2548-2568, 1997.

Jungclaus, J. H., Keenlyside, N., Botzet, M., Haak, H., Luo, J.-J., Latif, M., Marotzke, J., Mikolajewicz, U., and Roeckner, E.: Ocean Circulation and Tropical Variability in the Coupled Model ECHAM5/MPI-OM, J. Climate, 19, 3952-3972, 2006.

Kaplan, M. R., and Wolfe, A. P.: Spatial and temporal variability of Holocene temperature in the North Atlantic region, Quaternary Res., 65, 223-231, doi:10.1016/j.yqres.2005.08.020, 2006.

Koldunov, N. V., Stammer, D., and Marotzke, J.: Present-Day Arctic Sea Ice Variability in the Coupled ECHAM5/MPI-OM Model, J. Climate, 23, 2520-2543, doi:10.1175/2009JCLI3065.1, 2010.

Laepple, T., and Huybers, P.: Reconciling discrepancies between Uk37 and Mg/Ca reconstructions of Holocene marine temperature variability, Earth Planet. Sc. Lett., 375, 418-429, doi:10.1016/j.epsl.2013.06.006, 2013.

Laepple, T., and Huybers, P.: Ocean surface temperature variability: Large model–data differences at decadal and longer periods, P. Natl. Acad. Sci. USA, 111(47), 16682-16687, doi:10.1073/pnas.1412077111, 2014.

Lea, D. W.: Elemental and Isotopic Proxies of Past Ocean Temperatures, In: Treatise on Geochemistry, Vol. 6, Editor: Elderfield, H., Executive Editors: Holland, H. D., and Turekian, K. K., Elsevier, pp. 365-390, doi:10.1016/B0-08-043751-6/06114-4, 2003.

Leduc, G., Schneider, R., Kim, J. H., and Lohmann, G.: Holocene and Eemian sea surface temperature trends as revealed by alkenone and Mg/Ca paleothermometry, Quaternary Sci. Rev., 29, 989-1004, doi:10.1016/j.quascirev.2010.01.004, 2010.

Li, J., Gou, X., Cook, E. R., and Chen, F.: Tree-ring based drought reconstruction for the central Tien Shan area in northwest
5   China, Geophys. Res. Lett., 33, L07715, doi:10.1029/2006GL025803, 2006.

Liu, Z., Otto-Bliesner, B. L., He, F., Brady, E. C., Tomas, R., Clark, P. U., Carlson, A. E., Lynch-Stieglitz, J., Curry, W., Brook, E., Erickson, D., Jacob, R., Kutzbach, J., and Cheng, J.: Transient Simulation of Last Deglaciation with a New Mechanism for Bølling-Allerød Warming, Science, 325(5938), 310-314, doi:10.1126/science.1171041, 2009.

Lohmann, G., Pfeiffer, M., Laepple, T., Leduc, G., and Kim, J.-H.: A model-data comparison of the Holocene global sea
10   surface temperature evolution, Clim. Past, 9, 1807-1839, doi:10.5194/cp-9-1807-2013, 2013.

Lu, F., Ma, C., Zhu, C., Lu, H., Zhang, X., Huang, K., Guo, T., Li, K., Li, L., Li, B., and Zhang, W.: Variability of East Asian summer monsoon precipitation during the Holocene and possible forcing mechanisms, Clim. Dynam., 1-21, doi:10.1007/s00382-018-4175-6, 2018.

Luckman, B. H., Briffa, K. R., Jones, P. D., and Schweingruber, F. H.: Tree-ring based reconstruction of summer
15   temperatures at the Columbia Icefield, Alberta, Canada, AD 1073-1983, The Holocene, 7(4), 375-389, 1997.

Madden, R. A., Shea, D. J., Branstator, G. W., Tribbia, J. J., and Weber, R. O.: The Effects of Imperfect Spatial and Temporal Sampling on Estimates of the Global Mean Temperature: Experiments with Model Data, J. Climate, 6, 1057-1066, 1993.

Mangerud, J., and Svendsen, J. I.: The Holocene Thermal Maximum around Svalbard, Arctic North Atlantic; molluscs show
20   early and exceptional warmth, The Holocene, 28(1), 65-83, doi:10.1177/0959683617715701, 2018.

Mann, M. E., Rutherford, S., Wahl, E., and Ammann, C.: Robustness of proxy-based climate field reconstruction methods, J. Geophys. Res., 112, D12109, doi:10.1029/2006JD008272, 2007.

Mann, M. E., Zhang, Z., Hughes, M. K., Bradley, R. S., Miller, S. K., Rutherford, S., and Ni, F.: Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia, P. Natl. Acad. Sci. USA, 105(36),
25   13252-13257, doi:10.1073/pnas.0805721105, 2008.

Marcott, S. A., Shakun, J. D., Clark, P. U., and Mix, A. C.: A Reconstruction of Regional and Global Temperature for the Past 11,300 Years, Science, 339(6124), 1198-1201, doi:10.1126/science.1228026, 2013.

Marsland, S. J., Haak, H., Jungclaus, J. H., Latif, M., and Röske, F.: The Max-Planck-Institute global ocean/sea ice model with orthogonal curvilinear coordinates, Ocean Model., 5(2), 91-127, doi:10.1016/S1463-5003(02)00015-X, 2003.

30 Matei, D., Pohlmann, H., Jungclaus, J., Müller, W., Haak, H., and Marotzke, J.: Two Tales of Initializing Decadal Climate Prediction Experiments with the ECHAM5/MPI-OM Model, J. Climate, 25, 8502-8523, doi:10.1175/JCLI-D-11-00633.1, 2012.

May, W.: Climatic changes associated with a global "2°C-stabilization" scenario simulated by the ECHAM5/MPI-OM coupled climate model, Clim. Dynam., 31, 283-313, doi:10.1007/s00382-007-0352-8, 2008.

Mischel, S. A., Scholz, D., Spötl, C., Jochum, K. P., Schröder-Ritzrau, A., and Fiedler, S.: Holocene climate variability in Central Germany and a potential link to the polar North Atlantic: A replicated record from three coeval speleothems, The Holocene, 27(4), 509-525, doi:10.1177/0959683616670246, 2017.

Mollenhauer, G., Eglinton, T. I., Ohkouchi, N., Schneider, R. R., Müller, P. J., Grootes, P. M., and Rullkötter, J.: Asynchronous alkenone and foraminifera records from the Benguela Upwelling System, Geochim. Cosmochim. Ac., 67(12), 2157-2171, doi:10.1016/S0016-7037(00)00168-6, 2003.

Moossen, H., Bendle, J., Seki, O., Quillmann, U., and Kawamura, K.: North Atlantic Holocene climate evolution recorded by high-resolution terrestrial and marine biomarker records, Quaternary Sci. Rev., 129, 111-127, doi:10.1016/j.quascirev.2015.10.013, 2015.

Müller, W. A., and Roeckner, E.: ENSO teleconnections in projections of future climate in ECHAM5/MPI-OM, Clim. Dynam., 31, 533-549, doi:10.1007/s00382-007-0357-3, 2008.

Münch, T., and Laepple T.: What climate signal is contained in decadal to centennial scale isotope variations from Antarctic ice cores?, Clim. Past Discuss., 1-25, doi:10.5194/cp-2018-112, in review, 2018.

Münch, T., Kipfstuhl, S., Freitag, J., Meyer, H., and Laepple, T.: Regional climate signal vs. local noise: a two-dimensional view of water isotopes in Antarctic firn at Kohnen Station, Dronning Maud Land, Clim. Past, 12, 1565-1581, doi:10.5194/cp-12-1565-2016, 2016.

Münch, T., Kipfstuhl, S., Freitag, J., Meyer, H., and Laepple, T.: Constraints on post-depositional isotope modifications in East Antarctic firn from analysing temporal changes of isotope profiles, The Cryosphere, 11, 2175-2188, doi:10.5194/tc-11-2175-2017, 2017.

Ohkouchi, N., Eglinton, T. I., Keigwin, L. D., and Hayes, J. M.: Spatial and Temporal Offsets Between Proxy Records in a Sediment Drift, Science, 298(5596), 1224-1227, doi:10.1126/science.1075287, 2002.

Pausata, F. S. R., and Löfverström, M.: On the enigmatic similarity in Greenland δ18O between the Oldest and Younger Dryas, Geophys. Res. Lett., 42, 10470-10477, doi:10.1002/2015GL066042, 2015.

Raddatz, T. J., Reick, C. H., Knorr, W., Kattge, J., Roeckner, E., Schnur, R., Schnitzler, K.-G., Wetzel, P., and Jungclaus, J.: Will the tropical land biosphere dominate the climate–carbon cycle feedback during the twenty-first century?, Clim. Dynam., 29, 565-574, doi:10.1007/s00382-007-0247-8, 2007.

Rehfeld, K., Trachsel, M., Telford, R. J., and Laepple, T.: Assessing performance and seasonal bias of pollen-based climate reconstructions in a perfect model world, Clim. Past, 12, 2255-2270, doi:10.5194/cp-12-2255-2016, 2016.

Rehfeld, K., Münch, T., Ho, S. L., and Laepple, T.: Global patterns of declining temperature variability from the Last Glacial Maximum to the Holocene, Nature, 554, 356-359, doi:10.1038/nature25454, 2018.

Reschke, M., Kunz, T., and Laepple, T.: Comparing methods for analysing time scale dependent correlations in irregularly sampled time series data, Comp. Geosci., in review.

Richey, J. N., Hollander, D. J., Flower, B. P., and Eglinton, T. I.: Merging late Holocene molecular organic and foraminiferal-based geochemical records of sea surface temperature in the Gulf of Mexico, Paleoceanography, 26, PA1209, doi:10.1029/2010PA002000, 2011.

Roche, D. M., Waelbroeck, C., Metcalfe, B., and Caley, T.: FAME (v1.0): a simple module to simulate the effect of planktonic foraminifer species-specific habitat on their oxygen isotopic content, Geosci. Model Dev., 11, 3587-3603, doi:10.5194/gmd-11-3587-2018, 2018.

Roeckner, E., Bäuml, G., and Bonaventura, L.: The atmospheric general circulation model, 2003.

Rosell-Melé, A., and Prahl, F. G.: Seasonality of UK′ 37 temperature estimates as inferred from sediment trap data, Quaternary Sci. Rev., 72, 128-136, doi:10.1016/j.quascirev.2013.04.017, 2013.

Rosén, P., Segerström, U., Eriksson, L., and Renberg, I.: Do Diatom, Chironomid, and Pollen Records Consistently Infer Holocene July Air Temperature? A Comparison Using Sediment Cores from Four Alpine Lakes in Northern Sweden, Arct. Antarct. Alp. Res., 35(3), 279–290, 2003.

Schneider, B., Leduc, G., and Park, W.: Disentangling seasonal signals in Holocene climate trends by satellite-model-proxy integration, Paleoceanography, 25, PA4217, doi:10.1029/2009PA001893, 2010.

Sejrup, H. P., Seppä, H., McKay, N. P., Kaufman, D. S., Geirsdóttir, Á., de Vernal, A., Renssen, H., Husum, K., Jennings, A., and Andrews, J. T.: North Atlantic-Fennoscandian Holocene climate trends and mechanisms, Quaternary Sci. Rev., 147, 365-378, doi:10.1016/j.quascirev.2016.06.005, 2016.

Smerdon, J. E.: Climate models as a test bed for climate reconstruction methods: pseudoproxy experiments, WIREs Clim. Change, 3, 63-77, doi:10.1002/wcc.149, 2012.

Spero, H. J., Bijma, J., Lea, D. W., and Bemis, B. E.: Effect of seawater carbonate concentration on foraminiferal carbon and oxygen isotopes, Nature, 390, 497-500, 1997.

Stebich, M., Rehfeld, K., Schlütz, F., Tarasov, P. E., Liu, J., and Mingram, J.: Holocene vegetation and climate dynamics of NE China based on the pollen record from Sihailongwan Maar Lake, Quaternary Sci. Rev., 124, 275-289, doi:10.1016/j.quascirev.2015.07.021, 2015.

Telford, R. J., Heegaard, E., and Birks, H. J. B.: All age–depth models are wrong: but how badly?, Quaternary Sci. Rev., 23, 1-5, doi:10.1016/j.quascirev.2003.11.003, 2004.

Thibodeau, B., Bauch, H. A., and Knies, J.: Impact of Arctic shelf summer stratification on Holocene climate variability, Quaternary Sci. Rev., 191, 229-237, doi:10.1016/j.quascirev.2018.05.017, 2018.

Thirumalai, K., Partin, J. W., Jackson, C. S., and Quinn, T. M.: Statistical constraints on El Niño Southern Oscillation reconstructions using individual foraminifera: A sensitivity analysis, Paleoceanography, 28, 401–412, doi:10.1002/palo.20037, 2013.

Thirumalai, K., Quinn, T. M., Okumura, Y., Richey, J. N., Partin, J. W., Poore, R. Z., and Moreno-Chamarro, E.: Pronounced centennial-scale Atlantic Ocean climate variability correlated with Western Hemisphere hydroclimate, Nat. Commun., 9, 392, doi:10.1038/s41467-018-02846-4, 2018.

Timm, O., and Timmermann, A.: Simulation of the Last 21 000 Years Using Accelerated Transient Boundary Conditions, J. Climate, 20, 4377-4401, doi:10.1175/JCLI4237.1, 2007.

van Sebille, E., Scussolini, P., Durgadoo, J. V., Peeters, F. J. C., Biastoch, A., Weijer, W., Turney, C., Paris, C. B., and Zahn, R.: Ocean currents generate large footprints in marine palaeoclimate proxies, Nat. Commun., 6, 6521, doi:10.1038/ncomms7521, 2015.

Wanner, H., Mercolli, L., Grosjean, M., and Ritz, S. P.: Holocene climate variability and change; a data-based review, J. Geol. Soc. London, 172, 254-263, doi:10.1144/jgs2013-101, 2015.

Werner, M., Haese, B., Xu, X., Zhang, X., Butzin, M., and Lohmann, G.: Glacial–interglacial changes in H218O, HDO and deuterium excess – results from the fully coupled ECHAM5/MPI-OM Earth system model, Geosci. Model Dev., 9, 647-670, doi:10.5194/gmd-9-647-2016, 2016.

Zhang, Y., Renssen, H., Seppä, H., and Valdes, P. J.: Holocene temperature evolution in the Northern Hemisphere high latitudes – Model-data comparisons, Quaternary Sci. Rev., 173, 101-113, doi:10.1016/j.quascirev.2017.07.018, 2017.
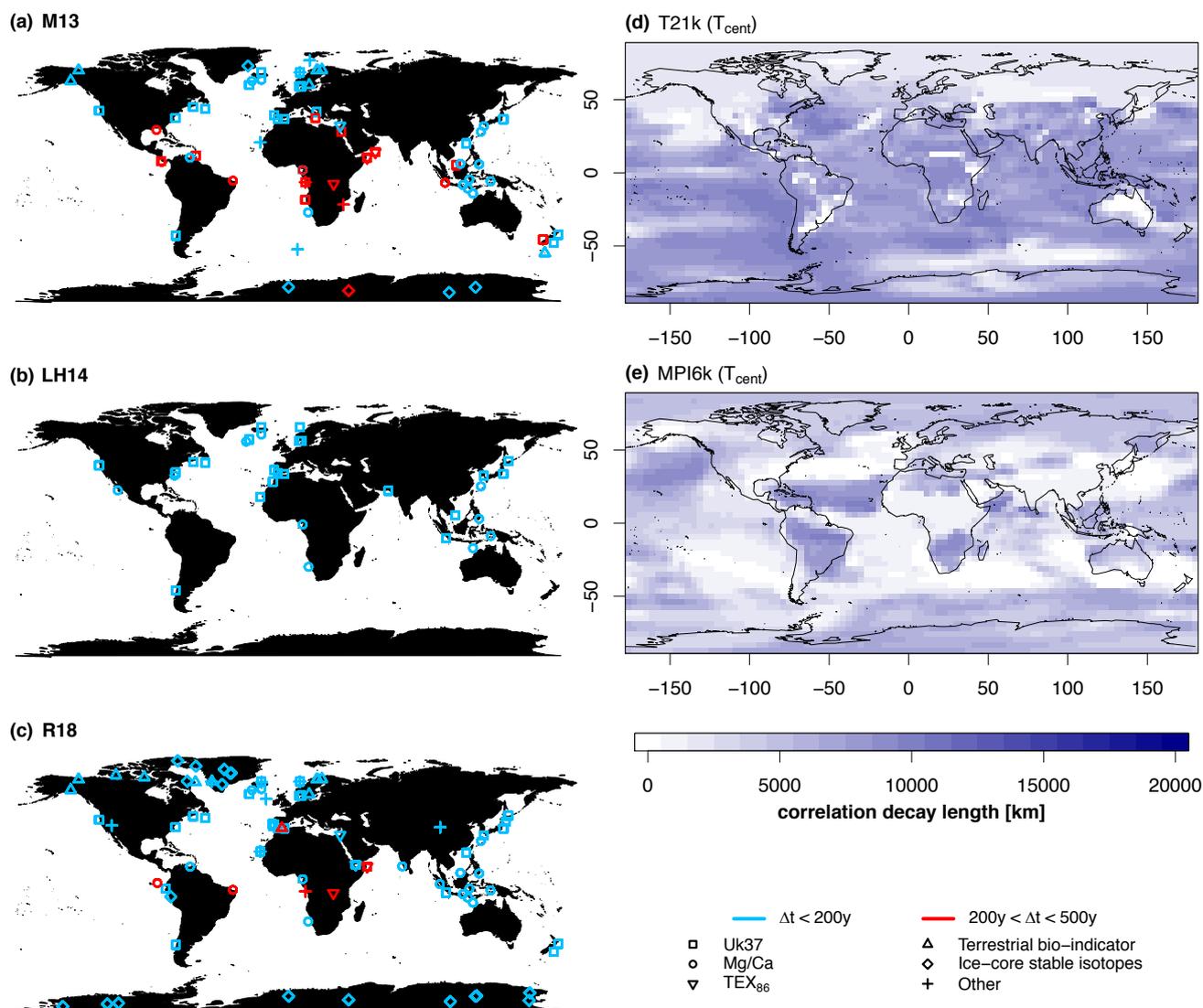
Figure 1: Overview of proxy and model datasets. Site locations of the proxy compilations (a) M13: Marcott et al. (2013), (b) LH14: Laepple and Huybers (2014), and (c) R18: Rehfeld et al. (2018) used in this study. Proxy types are indicated by symbols and the mean inter-observation time step by colours. Correlation decay length of (d) the T21k and (e) the MPI6k simulations estimated on centennial time scales (T_{cent}; not detrended). The spatial correlation decay length is generally higher for T21k than for MPI6k.
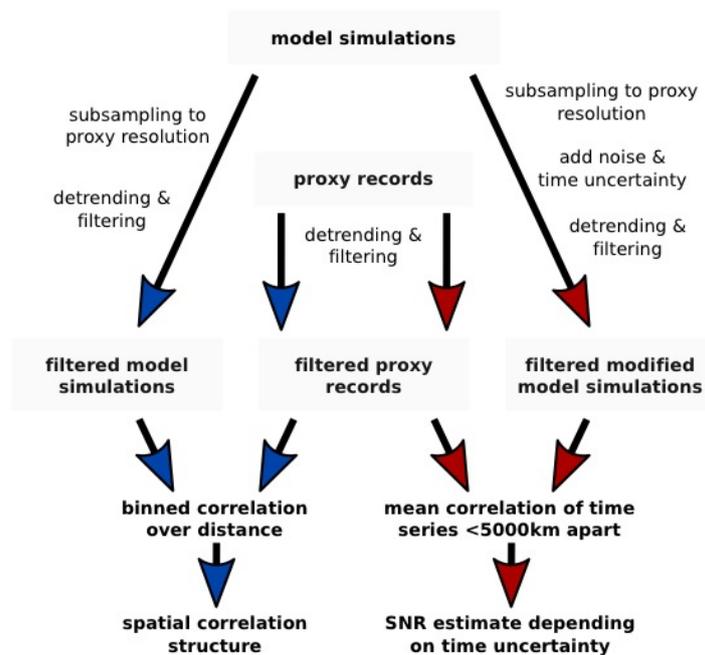
Figure 2: Processing steps for the proxy and model time series. Blue paths illustrate the analysis of the spatial correlation structure. Red paths represent the estimation of SNRs of proxy records as a function of time uncertainty.
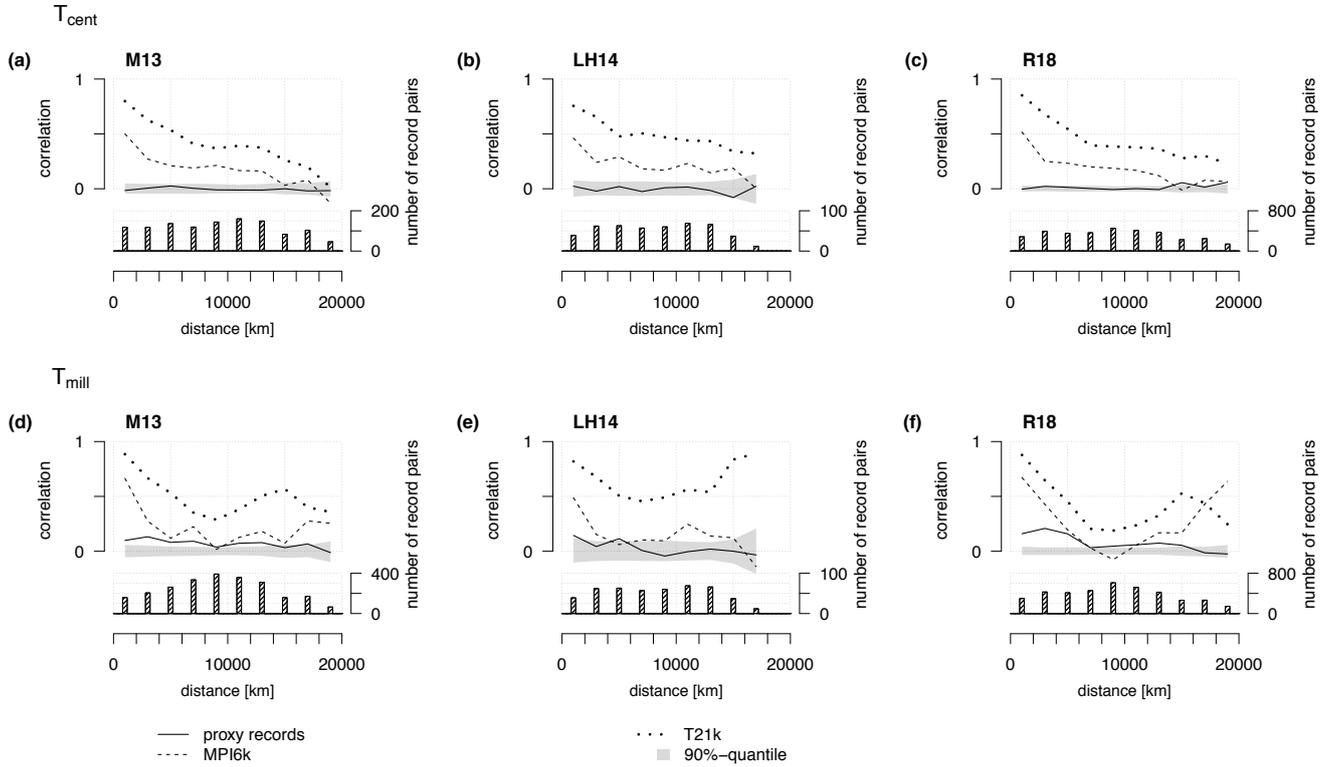
Figure 3: Spatial correlation structure of Holocene temperature proxy records and simulated surface temperatures based on three multi-proxy datasets and related to (a-c) centennial T_{cent} and (d-f) centennial to millennial time scales T_{mill}. In each panel, the upper part shows the mean correlation of the model simulation (for 2000 km sized bins as a function of the separation distance between record pairs) evaluated at the proxy locations (dotted/dashed line) and the proxy dataset (continuous line). The grey polygon represents the 90%-quantile of mean correlations of uncorrelated surrogate time series with a power-law scaling of $\beta = 1$. The lower parts of the panels show the number of record pairs used in each estimate. The spatial correlation structure of the model time series is generally higher than of proxy records which are only statistically significant on T_{mill} at neighboured sites. The highest correlations belong to sites with separation distances less than 4000-6000 km.
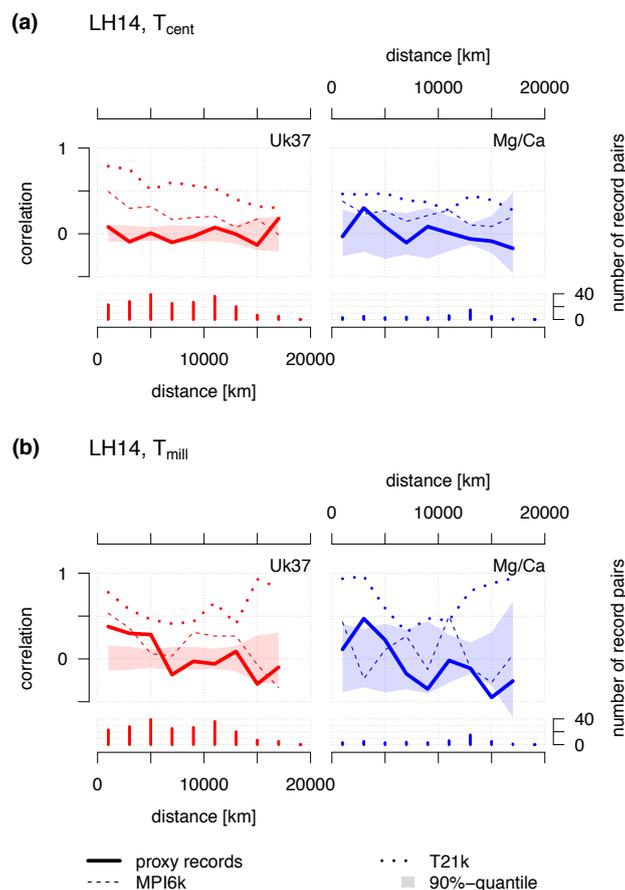
Figure 4: Proxy-type-specific (Uk37, Mg/Ca) spatial correlation structure related to (a) centennial T_{cent} and (b) centennial to millennial time scales T_{mill} based on the LH14 dataset. The upper parts of the panels show mean correlations of 2000 km sized bins as a function of the separation distance between record pairs in the proxy dataset (continuous line) and model simulations evaluated at proxy locations (dotted/dashed line). Polygons represent the 90%-quantiles of mean correlations of uncorrelated surrogate time series with a power-law scaling of $\beta = 1$. The lower parts of the panels show the number of record pairs used for each estimate. The spatial correlation structure of proxy records is non-significant for individual proxy types, except for close (separation <6000 km) sites of Uk37 temperature records at T_{mill}.
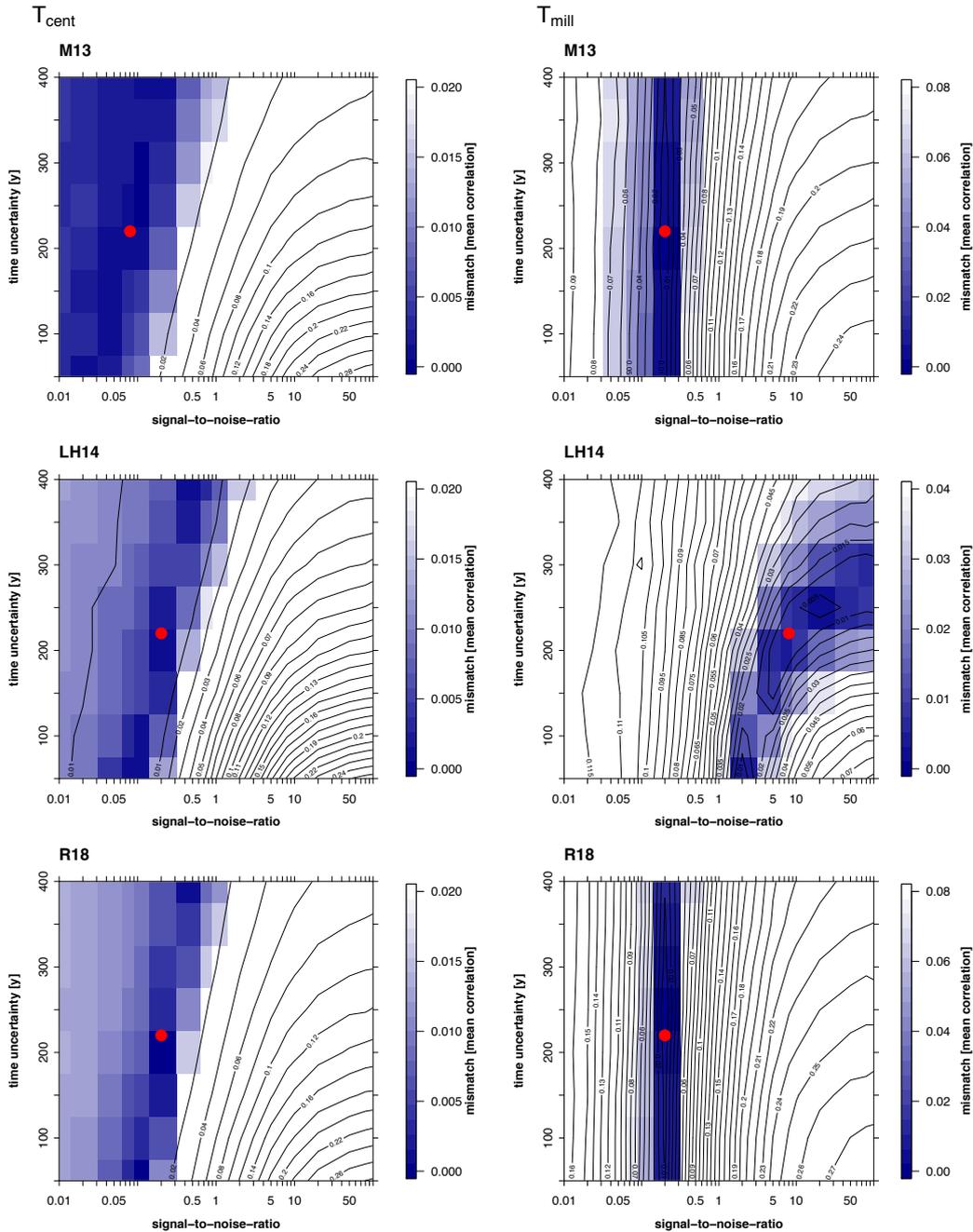
Figure 5: SNR$_{MPI6k}$ estimates of proxy records as a function of time uncertainty related to centennial T$_{cent}$ and millennial time scales T$_{mill}$. Colour coating and contour lines in each panel show the mismatch between mean correlations of close-by (separation <5000 km) proxy records and time series extracted from the MPI6k simulation at proxy locations as a function of time uncertainty (vertical axis) and SNR (horizontal axis). Areas with the lowest mismatch are represented by the darkest colours and mark suitable combinations of SNR$_{MPI6k}$ estimates and time uncertainties. The red dots illustrate SNR estimates for a time uncertainty of 220y.
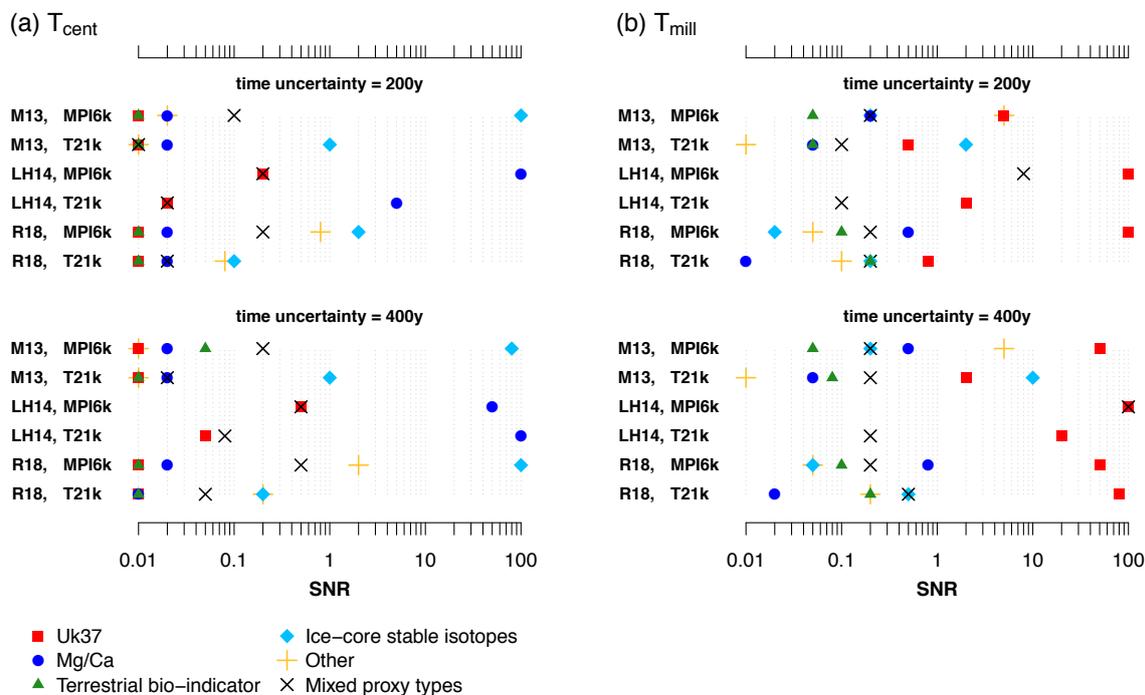
Figure 6: Overview of proxy-specific SNR estimates on (a) centennial T_cent and (b) millennial time scales T_mill. The symbols represent the SNRs estimated from the different proxy compilations using the simulations of MPI6k and T21k. Upper panels show the results for an assumed time uncertainty of 200y and lower panels for 400y. The SNRs are proxy-type-specific different, but generally higher on the T_mill than on the T_cent time scale.
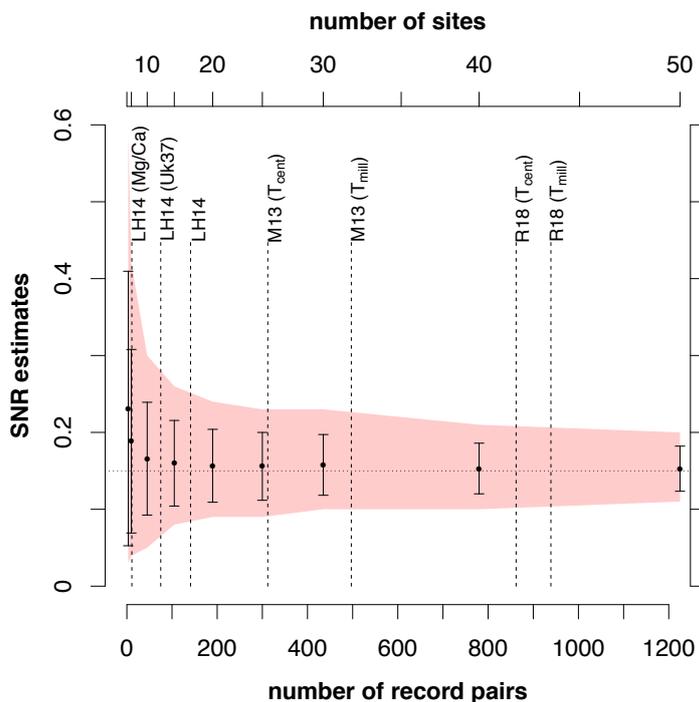
Figure 7: Sensitivity of the SNR estimates on the number of sites/record pairs based on surrogate time series. The time series were generated with a predefined SNR = 0.15 (horizontal line). SNR estimates with standard deviations based on 2000 repetitions are shown as dots with error bars. The uncertainties in the SNR are illustrated as polygons showing the 90%-quantiles of the estimates. The uncertainty of SNR estimates is high when only considering a small number of sites. Vertical lines show the numbers of selected sites/record pairs contained in each data compilation. This indicates that for single proxy type analysis the uncertainties in the SNR estimates are high.

Table 1: Numbers of records and their overlap in the proxy compilations used in this study. The total number of time series is separated by proxy type for each proxy compilation (upper part). T_{mill} refers to the number of time series with a mean inter-observation time step of Δt <500y and T_{cent} counts time series with Δt <200y. The overlap is shown for each pair and for all compilations (lower part).

	Sum of records	Uk37	Mg/Ca	TEX$_{86}$	Terrestrial bio-indicator	Ice-core stable isotopes	other
$M13 - T_{mill}$	70	28	19	4	8	5	6
$M13 - T_{cent}$	49	18	14	1	8	4	4
$LH14 - T_{mill}$	31	21	10	-	-	-	-
$LH14 - T_{cent}$	31	21	10	-	-	-	-
$R18 - T_{mill}$	88	27	19	4	11	18	9
$R18 - T_{cent}$	81	26	17	2	10	18	8
M13 ∩ LH14	20	13	7	-	-	-	-
M13 ∩ R18	45	16	14	3	6	4	2
LH14 ∩ R18	22	16	6	-	-	-	-
M13 ∩ LH14 ∩ R18	19	13	6	-	-	-	-

Table 2: Mean correlations of proxy time series with separation distances <5000 km for different proxy types. For each dataset, the mean correlation was estimated for millennial time scales T_{mill} and proxy time series with a mean inter-observation time step of Δt <500y and related to centennial time scales T_{cent} for proxy time series with Δt <200y. Mixed proxy types contain all combinations of time series pairs independent of the proxy type. The mean of single proxy types summarises the proxy-type-specific mean correlations weighted by the number of record pairs of each proxy type. Correlations in brackets are not statistically significant (p = 0.1).

	Mixed proxy types	Mean of single proxy types	Uk37	Mg/Ca	TEX_{86}	Terrestrial bio-indicator	Ice-core stable isotopes	other
$M13 - T_{mill}$	0.101	0.149	0.211	(0.068)	(-0.105)	(0.095)	0.414	(-0.14)
$M13 - T_{cent}$	0.004	-0.009	(-0.003)	(-0.006)	-	(-0.09)	(0.244)	(-0.388)
$LH14 - T_{mill}$	0.12	0.357	0.365	0.304	-	-	-	-
$LH14 - T_{cent}$	0.012	0.031	(0.013)	0.151	-	-	-	-
$R18 - T_{mill}$	0.184	0.208	0.347	(0.034)	(-0.188)	0.17	0.23	(0.107)
$R18 - T_{cent}$	0.014	0.009	(-0.017)	(-0.024)	(0.148)	(-0.057)	0.1	(0.05)