**Climate
of the Past**

Open Access

Discussions

Multi-time scale data assimilation for atmosphere–ocean state estimates

N. Steiger and G. Hakim

Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA

Received: 22 July 2015 – Accepted: 30 July 2015 – Published: 19 August 2015

Correspondence to: N. Steiger (nathanjs@uw.edu)

3729

Abstract

Paleoclimate proxy data span seasonal to millennial time scales, and Earth's climate system has both high- and low-frequency components. Yet it is currently unclear how best to incorporate multiple time scales of proxy data into a single reconstruction frame-

5 work and to also capture both high- and low-frequency components of reconstructed variables. Here we present a data assimilation algorithm that can explicitly incorporate proxy data at arbitrary time scales. Through a series of pseudoproxy experiments, we find that atmosphere–ocean states are most skilfully reconstructed by incorporating proxies across multiple time scales compared to using proxies at short (annual) or long

10 (~ decadal) time scales alone. Additionally, reconstructions that incorporate long time-scale pseudoproxies improve the low-frequency components of the reconstructions relative to using only high-resolution pseudoproxies. We argue that this is because time averaging high-resolution observations improves their covariance relationship with the slowly-varying components of the coupled-climate system, which the data assimila-

15 tion algorithm can exploit. These results are insensitive to the choice of climate model, despite the model variables having very different spectral characteristics. Our results also suggest that it may be possible to reconstruct features of the oceanic meridional overturning circulation based solely on atmospheric surface temperature proxies.

1 Introduction

20 Paleoclimate proxies range across widely different time scales. High resolution paleoclimate proxies such as tree rings or corals have annual or seasonal resolution, whereas lower resolution proxies such as sediment cores can provide anywhere from annual to millennial scale information depending on the core and its location (Bradley, 2014). Additionally, high resolution proxies tend to have short records and are mostly limited

25 to the past two millennia, whereas some low resolution proxies can reach back across the Cenozoic (e.g., Zachos et al., 2008). In addition to the many time scales of proxies,

3730

the climate system itself varies across a large range of time scales: from atmospheric blocking to ocean overturning circulation to ice age cycles. Thus any faithful reconstruction of past climate must account for as many of these time scales, captured by both proxies and climate models, as possible.

Few paleoclimate reconstruction methods have been created that specifically incorporate multiple proxy time scales. Most reconstructions use either low or high resolution proxies alone. If multiple scales of proxy data are used together, researchers often resort to coarsening high resolution proxies (e.g., PAGES 2k Consortium, 2013) or linearly interpolating low resolution proxies to a "higher resolution" (e.g., Mann et al., 2008). One major reason for this is that many traditional multivariate regression methods are not easily constructed to calibrate in both low and high frequency domains. Nevertheless, a few methods have been modified for such purposes (e.g., Mann et al., 2005). Additionally, this is not entirely a methodological problem but partly a temporal sampling issue: given that instrumental temperature data only span the past ~ 150 years, low frequency reconstruction techniques have few degrees of freedom on which to be calibrated and validated if the time scale is longer than about a decade. Only Li et al. (2010) (using a Bayesian hierarchical model approach) and Hanhijärvi et al. (2013) (using an approach based on pairwise comparisons) present reconstruction methods that can specifically incorporate proxies at any time resolution without linear interpolation of coarse proxies or coarsening high resolution proxies to some uniform time scale. These methods have thus far been used only for time series reconstructions, not space-time reconstructions.

Data assimilation (DA) provides a flexible framework for combining information from paleoclimate proxies with the dynamical constraints of climate models. In principle, DA can provide reconstructions of any model variable, from surface temperature to sea water salinity to atmospheric geopotential height. Among DA techniques, we are unaware of any method that incorporates proxies across any arbitrary range of time scales. DA-based reconstructions have so far used only a single uniform time scale (e.g., Goosse et al., 2012) or have performed separate reconstructions at different

3731

uniform time scales (Mathiot et al., 2013). Here we develop a DA-based algorithm for space-time climate reconstructions that can assimilate proxies at any time resolution. Because of the limited time span of observational data sets, we explore the features and skill of this technique within a synthetic, pseudoproxy framework. This allows us to test the algorithm over long time spans, perform carefully controlled experiments, and unambiguously define errors.

Multiproxy reconstructions can potentially overcome some limitations of single proxy reconstructions, such as filling in for the missing frequency components of a particular proxy (Li et al., 2010). But besides this kind of benefit, it is possible that particular reconstruction methods could benefit from multi-scale proxy data. Within a coupled atmosphere–ocean DA framework, Tardif et al. (2014) suggest that assimilating time-averaged observations of atmospheric variables may improve present-day estimates of ocean circulation. They argue that these improvements arise from the fact that time averaging high-frequency observations improves the signal over noise in the covariance relationship between the atmosphere and the slowly-varying ocean overturning circulation. We test this hypothesis within a paleoclimate context and assesses whether or not atmosphere–ocean state estimates can be improved by including proxies and climate states at multiple time scales.

2 Assimilation technique

Data assimilation refers to a mathematical technique of optimally combining observations (or within this context, proxy data) with prior information, typically from a model. The model, in this case a climate model, provides an initial, or prior, state estimate that one can update in a Bayesian sense based on the observations and an estimate of the errors in both the observations and the prior. The prior contains any climate model variables of interest and the updated prior, called the posterior, is the best estimate of the climate state given the observations and the error estimates. The basic state

3732

update equations of DA (e.g., Kalnay, 2003) are given by

$$x_a = x_b + \mathbf{K}[y - \mathcal{H}(x_b)], \tag{1}$$

and where $\mathbf{K}$ can be written as

$$\mathbf{K} = \text{cov}(x_b, \mathcal{H}(x_b))[\text{cov}(\mathcal{H}(x_b), \mathcal{H}(x_b)) + \mathbf{R}]^{-1}, \tag{2}$$

with cov representing a covariance expectation. The prior (or "background") estimate of the state vector is x_b and x_a is the posterior (or "analysis") state vector. Observations (or proxies) are contained in vector y. The true value of the observations are estimated by the prior through $\mathcal{H}(x_b)$, which is, in general, a nonlinear vector-valued observation operator that maps x_b from the state space to the observation space. For example, tree-ring width may be estimated from grid-point values of temperature and moisture in the prior. Matrix $\mathbf{K}$, the Kalman gain, weights $y - \mathcal{H}(x_b)$, called the innovation, and transforms it into state space. Matrix $\mathbf{R}$ is the error covariance matrix for the observations. The DA update process involves computing Eqs. (1) and (2) to arrive at the posterior state; within the context of the climate reconstruction problem, the posterior state is the reconstructed state for a given time. Space-time reconstructions are obtained by iteratively estimating the posterior state for each year or time segment of the reconstruction.

From the numerator of Eq. (2), we can interpret $\mathbf{K}$ as "spreading" the information contained in the observations through the covariance between the prior and the prior-estimated observations. This implies that, other things being equal, larger values of $\text{cov}(x_b, \mathcal{H}(x_b))$ will weight the innovation more heavily; thus this new information not contained in the prior has a bigger influence. One way to improve this covariance relationship may be to use time-averaged observations, particularly if the model or climate system has better covariance relationships at longer time scales.

For the update calculations we employ an ensemble square-root Kalman filter, applied to time averages (see Steiger et al., 2014 for a detailed DA algorithm and fuller discussion of DA terminology). We extend the technique of Dirren and Hakim (2005);

Huntley and Hakim (2010), and Steiger et al. (2014) by iteratively applying the state-update equations across multiple time scales by leveraging the serial observation processing approach to the Kalman filter (Houtekamer and Mitchell, 2001). The following general algorithm allows one to assimilate any collection of observation or proxy data, including irregular time averages:

1. Construct a prior ("background") ensemble $\mathbf{x}_b$ at the highest temporal resolution of interest (e.g., monthly or annual), or a collection of them with one for each time step (e.g., monthly or annual ensembles assigned to particular months or years).

2. Loop over observations and assimilate each at their own time scale:

   a. Decompose the prior ensemble(s) that overlap in time with the observation y into time averages (overbar) and deviations from this average (prime) via $\mathbf{x}_b = \overline{\mathbf{x}}_b + \mathbf{x}'_b$, such that the time-average of $\overline{\mathbf{x}}_b$ matches the time scale of y.

   b. Estimate the observation via the forward model, or "proxy system model" (Evans et al., 2013) $\mathcal{H}(\overline{\mathbf{x}}_b)$, and update (compute Eqs. 1 and 2) the time averaged ensemble(s) $\overline{\mathbf{x}}_b \xrightarrow{\text{DA}} \overline{\mathbf{x}}_a$, with $\overline{\mathbf{x}}_a$ as the posterior (or "analysis") time-mean ensemble(s).

   c. Add back the time-deviations, $\mathbf{x}_b = \overline{\mathbf{x}}_a + \mathbf{x}'_b$, which can serve as the prior(s) for another observation.

   Note that if y shares the same time scale as $\mathbf{x}_b$, then the method is the same, but with $\mathbf{x}'_b = 0$.

3. After all observations have been assimilated, the ensemble mean of $\mathbf{x}_b$ provides the best estimate of the state for each analysis time.

We now discuss an illustrative implementation of this general algorithm that we employ for the experiments in this paper. Consider a paleoclimatic situation where the observations are a collection of annual proxy data and also proxy data representing irregularly averaged climatic information. Prior ensembles can be composed of annually

averaged climate states that have been randomly drawn from a long climate simulation and initially assigned to each year of a reconstruction, Fig. 1. Following the steps outlined above, proxies representing differently averaged time intervals can be assimilated by averaging over the prior ensembles for the time intervals defined by the proxy. For example, a proxy value representing information over the years 1700–1720 would update the prior ensembles averaged in time over that same interval. Annual proxies can simply be assimilated by updating the ensembles for each year of available proxy data, Fig. 1. This approach proceeds by assimilating each proxy over its full time extent and after every proxy is assimilated, one is left with an updated version of what one started with: a time sequence of ensemble state estimates at annual resolution.

In all of the experiments shown here we use a "no-cycling" or "off-line" DA approach, where the prior ensembles are drawn from existing climate model simulations. This approach has vast computational benefits over a "cycling" or "on-line" approach where one must integrate an ensemble of climate model simulations forward in time after each DA update step. Indeed, for the paleoclimate reconstruction problem, it is infeasible to cycle an ensemble of tens to hundreds of CMIP5-class coupled climate models (as used here) for hundreds or thousands of years. Moreover, in the off-line case one may use hundreds to thousands of ensemble members from multiple models and simulations, reducing the potential for model bias and sampling error. It is also advantageous to use an off-line approach when the predictability time limit of the model is shorter than the time scale of the observations: for example, if observations are only available at annual resolution yet the model cannot skilfully forecast the climate state a year into the future, then no useful information is gained by cycling the model. Matsikaris et al. (2015) recently compared on-line and off-line approaches to paleoclimate DA with a fully-coupled earth system model and found no improvement with the on-line method, suggesting that the model was unable to provide useful information at analysis times. Nevertheless, one way the approach outlined here can generalize to the on-line approach is by cycling on the shortest time scale (e.g., annual or seasonal) and updating longer time scales at the end of the appropriate interval without cycling.

If the reconstructions use the off-line approach together with multiple time scales, temporal consistency of the priors will need to be maintained in order to have coherent long-time-scale covariance relationships. Thus portions of these priors can be randomly drawn in blocks of consecutive years from the employed climate model simulation, see Fig. 1. The length of these blocks can be determined based on the needs of the specific reconstruction problem (e.g., the length of the longest proxy time scale) and the length of available model simulations. If multiple long simulations are available (they need not be from the same model), different rows in Fig. 1 could be different model simulations and the block length could be the length of the reconstruction; this option avoids any discontinuities in time that result from small block lengths.

Also note that for the sake of simplicity in the above example and throughout the paper, we are assuming that an irregular, long-time-scale proxy is just an average of some climate variable over a given time interval. Real proxies are nearly always more complex than this and would necessitate a more sophisticated proxy system model ($\mathcal{H}(\overline{x}_b)$ in Eqs. 1 and 2); however, the algorithm described above is general and covers the case when such models are available.

3   Experimental framework

3.1   Models and variable characterizations

For the experiments presented here, we are interested in (1) how the reconstruction methodology proposed in Sect. 2 performs in both the atmosphere and ocean, (2) how the differing time scales of the atmosphere and ocean may be leveraged in the reconstruction process, and (3) how these results vary with two different models having very different spectral characteristics in their coupled-climate systems. To this end we choose two long pre-industrial control simulations (part of the Coupled Model Intercomparison Project Phase 5 available for download at http://www.earthsystemgrid.org/), one from the climate model GFDL-CM3 and the other from CCSM4. We also choose

two illustrative reconstruction variables, global mean 2 m air temperature and the Atlantic meridional overturning circulation (AMOC). Figures 2 and 3 characterize the global mean temperature and an AMOC index for each simulation (defined here as the maximum value of the overturning streamfunction in the North Atlantic between 25 and 70° N and between depths of 500 and 2000 m), respectively. Note that even though these are only time series variables, the DA framework proposed here can trivially reconstruct spatial variables as well (Steiger et al., 2014). From Figs. 2 and 3 we see that despite both being control simulations, these two models display very different spectral characteristics for both global mean temperature and the AMOC index.

We next assess whether there are strong covariance relationships between the observation variables and the reconstruction variables at different time averages. Recall that the key covariance relationship in the DA update equations is between the prior variables and the prior estimate of the observations, Eq. (2). For an experiment where the pseudoproxies, y, are composed of atmospheric surface temperature (and thus what $\mathcal{H}(x_b)$ estimates), we would need to know the covariance between these and the state variables of global mean temperature and the AMOC index, contained in x_b. A simple assessment of this is shown in Fig. 4 with panels a and b showing box plot summaries of the correlations between the global mean temperature time series and all the surface temperature time series at every grid point for both climate simulations; these correlations are then computed over a range of averaging times. Panels c and d in Fig. 4 are the same calculation but for the AMOC index instead of the global mean temperature. (Note that the correlation of two time series is simply the covariance normalized by the product of the standard deviations of the two time series.) Figure 4 indicates that there is increased covariance information (or more locations with higher correlations) between surface temperature and the reconstruction variables at longer time scales. This information is leveraged by the equations of DA to potentially improve the low-frequency components of the reconstructed variables. An important point about computing correlations at increasing time averages is that the number of degrees of freedom in the time series are also reduced, potentially spuriously inflating the correla-

tions. A test of statistical significance accounting for the reduced degrees of freedom is, however, not particularly germane: the DA equations do not "know" about 95 % confidence intervals, just the covariance information. If, after performing the reconstructions and computing several different skill metrics, we see an increase in reconstruction skill, then we can infer that the information was in fact useful for the reconstructions.

3.2 Pseudoproxy construction

The pseudoproxy experiments employed here follow the general framework of many previous studies (see Smerdon, 2012 for a summary and review) but with some important modifications. Generally, after one or more climate model simulations are chosen to represent nature, a pseudoproxy network is chosen that mimics real world proxy availability, similar to the network chosen here and shown in Fig. 5a; this particular network is composed of a spatially thinned version of the proxy collection of PAGES 2k Consortium (2013) (thinned over Asia and North America where the proxy density is high) and all of the proxy locations in Shakun et al. (2012) and Marcott et al. (2013). Pseudoproxies are typically generated by adding random white noise to the chosen network of climate model temperature series. The added noise is usually assumed to be the same value for all proxy locations, with a common signal-to-noise ratio (SNR) being 0.5 (where $\mathrm{SNR} \equiv \sqrt{\mathrm{var}(X)/\mathrm{var}(N)}$, and where X is a grid-point temperature series drawn from the true state and N is an additive noise series, and var is the variance). Following recent work by Wang et al. (2014), we instead randomly draw SNR values from a distribution characteristic of real proxy networks, Fig. 5b. This distribution is a shifted Gamma distribution (shape parameter = 1.667, scale parameter = 0.18, shifted by 0.15) with a mean SNR of 0.45 and is modelled after Fig. 3 from Wang et al. (2014).

Also in contrast to nearly all pseudoproxy experiments, we use pseudoproxies at two different time scales for each model. Importantly for the comparison of results, we use the same SNR distribution for both time scales and add the noise to the time

series after averaging. Within the DA framework, the additive error for each proxy is accounted for in the entries of the diagonal matrix **R**. The SNR equation above is related to **R** in that that each of these entries is equal to var(N) for a given proxy. The process of adding the noise after averaging ensures that **R** is statistically identical for each reconstruction. This process isolates the role of the covariance relationships in Eq. (2). By drawing from the same SNR distribution for all pseudoproxy time scales we are also assuming that the distribution is an appropriate characterization of the error in long time scale proxies; we assume this for simplicity and also because we are not aware of a systematic assessment of SNR values for low-resolution proxies as Wang et al. (2014) have done for annual-resolution proxies.

We also note an important idealization of the present pseudoproxy experiments, which we share with all pseudoproxy experiments heretofore published, is that we use a perfect model approximation where the pseudoproxies from one model simulation are used to reconstruct that same simulation (e.g., pseudoproxies from the CCSM4 simulation are used to reconstruct the CCSM4 simulation, not the GFDL-CM3 simulation). In a real DA-based reconstruction, the climate model will never be a perfect description of the real climate system, from which the assimilated observations are derived. Since the purpose of the present work is to illustrate a novel algorithm, we therefore have not considered this additional layer of complexity, which can only be fully assessed within a study of real proxy climate reconstructions: using one simulation to reconstruct another can assess inter-model differences, but it is unclear how these results would relate to model–nature differences.

3.3 Pseudoproxy experiments

The primary results of this paper are presented with a series of 12 experiments using only atmospheric surface temperature pseudoproxies to reconstruct the global mean temperature and AMOC index of the two climate model simulations discussed previously. For each variable, and each model, three experiments are performed: (1) short (annual) pseudoproxies only, (2) long (5 or 20 year time averages) pseudoproxies only,

and (3) both short and long time averaged pseudoproxies. We have chosen the long time-scale for the CCSM4 simulation to be 20 years, and we note that an alternative choice of one to several decades gives similar results (not shown). The situation is more complex with the GFDL-CM3 simulation because of the presence of an approximate 22 year periodic signal in the AMOC, Fig. 3a and c. A choice of 20 years for GFDL-CM3 would effectively undersample the AMOC variability and so we have chosen a long time scale of 5 years for GFDL-CM3. Unfortunately, a long time scale of 5 years for CCSM4 shows little difference in the results over the annual time scale reconstructions (not shown), as would be suggested by the small difference in correlation (covariance) between 1 and 5 years, Fig. 4b.

Both the short only and long only reconstructions use 100 pseudoproxies randomly drawn from the network of 274 proxy locations shown in Fig. 5a. For the mixed resolution reconstructions, 100 pseudoproxies are randomly drawn from the network for each time scale, giving a total of 200. This is an approximation of the real world setting where one usually has proxies at multiple time scales and would like to use all of them. Following the algorithm outlined in Sect. 2, for the multi-scale reconstructions, we assimilate the long time-scale pseudoproxies first, followed by the annual time-scale pseudoproxies; we also performed these reconstructions by swapping which time scale was assimilated first and found statistically identical results (not shown), as would be expected from the linearity of this approach. For these mixed resolution reconstructions, we have also ensured that there is no overlap between locations associated with the two time scales.

We have reconstructed the first 500 years of each simulation while randomly drawing the prior, of size 500, in 20 year continuous blocks from the entire length of the simulations (800 years for GFDL-CM3 and 1051 years for CCSM4); this uniform block length was chosen because it was the longest time scale of the pseudoproxies and because the pseudoproxies were constructed over regular long intervals and thus discontinuities at block edges were not a concern (see Fig. 1 and the discussion in Sect. 2). Because of the large prior ensemble size, we did not employ covariance localization,

a common DA practice for controlling sampling error. Each of the 12 reconstructions are repeated 100 times in a Monte Carlo fashion where new SNR values are randomly drawn from the distribution, Fig. 5b, for each pseudoproxy location and a new pseudoproxy network is also randomly drawn for each time scale from the original network. All the reconstruction figures show the mean of 100 of these Monte Carlo reconstruction iterations.

4   Reconstruction results

The reconstructions of global mean temperature are shown in Fig. 6. The top panels of a and b show the reconstructions with the annual pseudoproxies, the middle panels show the reconstructions with the long time-scale proxies, and the bottom panels show the reconstructions for both time scales. Skill metrics, computed at annual resolution, are shown for each reconstruction: correlation (r) and coefficient of efficiency (ce), which for a data series comparison of length N is defined as

$$ce = 1 - \frac{\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}{\sum_{i=1}^{N}(x_i - \overline{x})^2},$$

where x is the "true" time series, $\overline{x}$ is the true time series mean, and $\hat{x}$ is the reconstructed time series. The metric ce has the range $-\infty < ce \leq 1$, where $ce = 1$ corresponds to a perfect match and $ce < 0$ indicates that the error variance is greater than the true time series variance. Comparing the reconstructions in Fig. 6 we see that the bottom panels with both time scales have the highest skill.

Note that the long-time scale reconstructions in the middle panels of Fig. 6a and b have sharp edges at 5 (for GFDL-CM3) or 20 year (for CCSM4) intervals. This is due to the simplified experimental design we have employed where all the long-time scale pseudoproxies are averages over a given 5 or 20 year period. As discussed in Sect. 2, this experimental design is only a single illustrative example of the general

3741

algorithm. The data from real proxies are not apportioned into specific time frames, but are scattered irregularly in time. Using many irregular proxies will act to smooth the long time-scale reconstructions. As long as the time scales can be estimated and an appropriate proxy system model is used, the algorithm of Sect. 2 can handle any real proxy data. While not dealt with explicitly here, real long time-scale (low resolution) proxies also have dating uncertainty, which will also tend to smooth the reconstructions. The algorithm can account for dating uncertainty through the Monte Carlo framework by repeating the reconstructions many times and sampling from an ensemble of ages dates for a given proxy.

One assessment of skill as a function of time scale is to compute the cross spectra of the reconstructed time series with the true time series (Fig. 7). The cross spectra in this case reveals the relationship between the two time series as a function of frequency or period. As a point of reference, the dashed gray lines in this figure indicate the cross spectra of the true time series with itself, which is the same as its own power spectrum.[1] Considering Fig. 7b we see that the annual-only reconstruction does a better job of matching the power at short periods than the 20 year-only reconstruction; however, the 20 year-only reconstruction performs better at longer periods. The mixed time-scale reconstruction, 20 + 1, does better or just as well as the single time-scale reconstructions at both short and long periods. This same general result holds for Fig. 7a, though it is more difficult to see because of the much larger power at longer periods in the GFDL-CM3 simulation.

Figure 8 shows three time scale reconstructions of the AMOC index for the two model simulations, similar to Fig. 6. In these AMOC index reconstructions, we see the same general patterns as with the global mean temperature reconstructions, where the

---

[1]Following a common technique to reduce noise in the cross spectra, they are computed using Welch's averaged periodogram method, which samples segments of the time series and averages the power spectra of these samples to arrive at the cross power spectral densities. As a result, the gray line spectra in Figs. 7 and 9 should not be expected to precisely match up with Figs. 2 and 3.

3742

multi-scale reconstructions provide the most skill (r and ce). Given that the pseudoproxies are of surface air temperature, it is not surprising that the absolute skill values of the AMOC reconstructions are reduced relative to the reconstructions of global mean temperature. Figure 9 shows the corresponding cross spectra for the reconstructions shown in Fig. 8. The most robust result from this figure is the improved low-frequency components of the AMOC reconstructions when time-averaged surface temperature pseudoproxies are used. We argue that this result follows from the fact that the annual observations of atmospheric surface temperature are essentially noise to the slowly varying ocean. One may improve the information content relevant to the ocean by averaging over the atmospheric noise. This interpretation may also be seen in Fig. 4, where the correlation (covariance) information between the atmosphere and the ocean is particularly low at annual averages but improves at longer time averages.

We note that all the cross spectra of the reconstructions shown in Figs. 7 and 9 show a decrease in power relative to the true state, though this need not always be the case. In additional experiments we performed using global ocean heat content, we found that this reconstructed variable tended to have more power than the true state and was thus higher than the respective gray lines (not shown). Therefore the reduced power relative to the true state in Figs. 7 and 9 should not be interpreted as saying something general about the nature of DA-based reconstructions or the particular approach employed here.

As an approximation of a real reconstruction scenario, the experiments shown in Figs. 6 and 8 with two time scales use twice as many pseudoproxies as the single time scale experiments (200 vs. 100). Therefore the improved skill might simply be a consequence of having more observation information. We tested this idea by repeating all the experiments shown here but instead increasing the number of observations to 200 for each experiment: the single time scale reconstructions used 200 randomly drawn pseudoproxies and the multi-scale reconstructions used 100 randomly drawn pseudoproxies each for the two time scales (the same as in the previous multi-scale reconstructions). Figure 10 is a characteristic example of the results of these additional

tests. Figure 10a shows the reconstructions of the AMOC index with the CCSM4 model output and Fig. 10b shows the respective cross spectra. In a, the skill is best for the multi-scale reconstructions and in b the cross spectra shows the same general result of improved low-frequency power for the time-averaged pseudoproxies. However, the cross spectra for the 20 + 1 reconstruction is not always closest to the true spectrum, suggesting that the number of pseudoproxies does play a role in improving the spectrum of the reconstructions. However, we note that the r and ce skill metrics for the single time scale reconstructions in Fig. 10a are hardly changed relative to those in Fig. 8b even though the number of observations are doubled.

5 Conclusions

This paper presents a data assimilation algorithm that can explicitly incorporate proxy data on arbitrary time scales. This approach generalizes previous data assimilation techniques in the sense that many scales of both proxies and climate states can be included explicitly in a single reconstruction framework. This may be particularly useful given the many inherent time scales of the climate system, such as the fast time scales of the atmosphere and the slow time scales of the ocean. We performed three types of realistic atmosphere–ocean pseudoproxy reconstructions to assess the impact of using observations at multiple time scales: (1) short (annual) pseudoproxies only, (2) long ($\sim$ decadal) pseudoproxies only, and (3) both short and long time-averaged pseudoproxies. We found that for both global mean temperature and an index of the Atlantic meridional overturning circulation, the reconstructions that incorporated proxies across both short and long time scales were more skilful than reconstructions that used short or long time scales alone (Figs. 6 and 8). This result holds even when the number of pseudoproxies for the single time-scale reconstructions are doubled, Fig. 10a. Multi-scale reconstructions would be expected to perform better than single-scale reconstructions because they can include information at multiple time scales and because the prior can be better conditioned as it's used from one time scale to the next.

We found that reconstructions incorporating long time scale pseudoproxies improve the low-frequency components of the reconstructions over reconstructions that only use annual time-scale pseudoproxies, Figs. 7, 9, and 10b. This result may at first seem surprising because the annual pseudoproxies should contain the low-frequency infor-
5 mation. It is helpful to recall that the data assimilation algorithm outlined here proceeds by sequentially finding the optimal state at each time segment of interest given the prior, the observations, and their respective errors. This state update critically relies on the covariance between the prior and the model estimate of the observations, Eq. (2). If, for example, surface temperature observations do not covary well with the AMOC at
10 annual resolution, then the posterior AMOC estimate will be little changed compared to the prior (Tardif et al., 2014). But if the time average of surface temperatures has a large covariance with the AMOC, the posterior will be more influenced by the observations. This result is not controlled by the noise added to the pseudoproxies because, as noted in Sect. 3.2, we ensured that **R** from Eq. (2) remains fixed for both time scales.
15 These results indicate that data assimilation-based atmosphere–ocean state estimates may be improved by including proxies and climate states from multiple time scales. The general results outlined above appear to be insensitive to the choice of climate model simulation. These results also show, as suggested by Kurahashi-Nakamura et al. (2014), that given a representative prior ensemble, features of the
20 Atlantic meridional overturning circulation may be reconstructed using observations of surface variables alone.

3745

References

Bradley, R. S.: Paleoclimatology, 3rd edn., Academic Press, Oxford, UK, 2014. 3730
Dirren, S. and Hakim, G. J.: Toward the assimilation of time-averaged observations, Geophys. Res. Lett., 32, L04804, doi:10.1029/2004GL021444, 2005. 3733
5 Evans, M. N., Tolwinski-Ward, S., Thompson, D., and Anchukaitis, K. J.: Applications of proxy system modeling in high resolution paleoclimatology, Quaternary Sci. Rev., 76, 16–28, 2013. 3734
Goosse, H., Crespin, E., Dubinkina, S., Loutre, M.-F., Mann, M. E., Renssen, H., Sallaz-Damaz, Y., and Shindell, D.: The role of forcing and internal dynamics in explaining the
10 "Medieval Climate Anomaly", Clim. Dynam., 39, 2847–2866, 2012. 3731
Hanhijärvi, S., Tingley, M., and Korhola, A.: Pairwise comparisons to reconstruct mean temperature in the Arctic Atlantic Region over the last 2,000 years, Clim. Dynam., 41, 2039–2060, doi:10.1007/s00382-013-1701-4, 2013. 3731
Houtekamer, P. L. and Mitchell, H. L.: A sequential ensemble Kalman filter for atmospheric data
15 assimilation, Mon. Weather Rev., 129, 123–137, 2001. 3734
Huntley, H. S. and Hakim, G. J.: Assimilation of time-averaged observations in a quasi-geostrophic atmospheric jet model, Clim. Dynam., 35, 995–1009, doi:10.1007/s00382-009-0714-5, 2010. 3734
Kalnay, E.: Atmospheric Modeling, Data Assimilation and Predictability, Cambridge, Cambridge,
20 UK, 2003. 3733
Kurahashi-Nakamura, T., Losch, M., and Paul, A.: Can sparse proxy data constrain the strength of the Atlantic meridional overturning circulation?, Geosci. Model Dev., 7, 419–432, doi:10.5194/gmd-7-419-2014, 2014. 3745
Li, B., Nychka, D. W., and Ammann, C. M.: The value of multiproxy reconstruction of past
25 climate, J. Am. Stat. Assoc., 105, 883–895, 2010. 3731, 3732
Mann, M. E., Rutherford, S., Wahl, E., and Ammann, C.: Testing the fidelity of methods used in proxy-based reconstructions of past climate, J. Climate, 18, 4097–4107, 2005. 3731
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
30 variations over the past two millennia, P. Natl. Acad. Sci. USA, 105, 13252–13257, doi:10.1073/pnas.0805721105, 2008. 3731

3746

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, 1198–1201, doi:10.1126/science.1228026, 2013. 3738, 3752

Mathiot, P., Goosse, H., Crosta, X., Stenni, B., Braida, M., Renssen, H., Van Meerbeeck, C. J., Masson-Delmotte, V., Mairesse, A., and Dubinkina, S.: Using data assimilation to investigate the causes of Southern Hemisphere high latitude cooling from 10 to 8 ka BP, Clim. Past, 9, 887–901, doi:10.5194/cp-9-887-2013, 2013. 3732

Matsikaris, A., Widmann, M., and Jungclaus, J.: On-line and off-line data assimilation in palaeo-climatology: a case study, Clim. Past, 11, 81–93, doi:10.5194/cp-11-81-2015, 2015. 3735

PAGES 2k Consortium: Continental-scale temperature variability during the past two millennia, Nature Geosci., 6, 339–346, doi:10.1038/ngeo1797, 2013. 3731, 3738, 3752

Schneider, T. and Neumaier, A.: Algorithm 808: ARfit a Matlab package for the estimation of parameters and eigenmodes of multivariate autoregressive models, ACM T. Math. Software, 27, 58–65, 2001. 3749, 3750

Shakun, J. D., Clark, P. U., He, F., Marcott, S. A., Mix, A. C., Liu, Z., Otto-Bliesner, B., Schmittner, A., and Bard, E.: Global warming preceded by increasing carbon dioxide concentrations during the last deglaciation, Nature, 484, 49–54, 2012. 3738, 3752

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. 3738

Steiger, N. J., Hakim, G. J., Steig, E. J., Battisti, D. S., and Roe, G. H.: Assimilation of time-averaged pseudoproxies for climate reconstruction, J. Climate, 27, 426–441, doi:10.1175/JCLI-D-12-00693.1, 2014. 3733, 3734, 3737

Tardif, R., Hakim, G., and Snyder, C.: Coupled atmosphere–ocean data assimilation experiments with a low-order model and CMIP5 model data, Clim. Dynam., doi:10.1007/s00382-014-2390-3, 2014. 3732, 3745

Wang, J., Emile-Geay, J., Guillot, D., Smerdon, J. E., and Rajaratnam, B.: Evaluating climate field reconstruction techniques using improved emulations of real-world conditions, Clim. Past, 10, 1–19, doi:10.5194/cp-10-1-2014, 2014. 3738, 3739, 3752

Zachos, J. C., Dickens, G. R., and Zeebe, R. E.: An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics, Nature, 451, 279–283, 2008. 3730
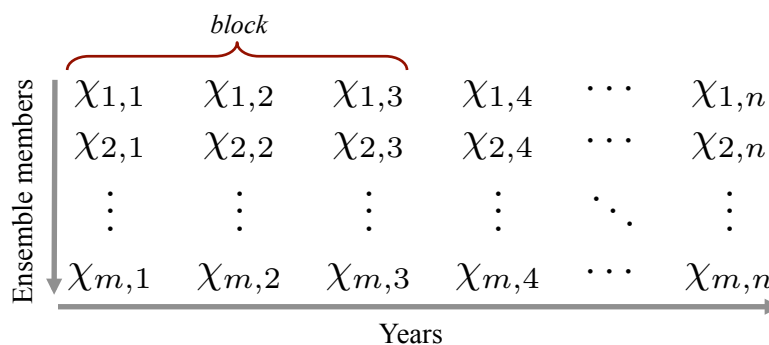
Figure 1. Schematic of the general reconstruction method using an off-line approach. Prior ensembles of m state vectors, χ, are assigned to each of the n years. To retain some temporal coherency, the rows are composed of time-coherent blocks drawn from a climate model simulation (arbitrarily illustrated here as a 3 year block, or 3 consecutive annual states). The method updates prior ensembles for specific years corresponding to annual proxy data points, while for long time-scale proxies prior ensembles are computed by time averaging across the rows corresponding to the years of a proxy data point.
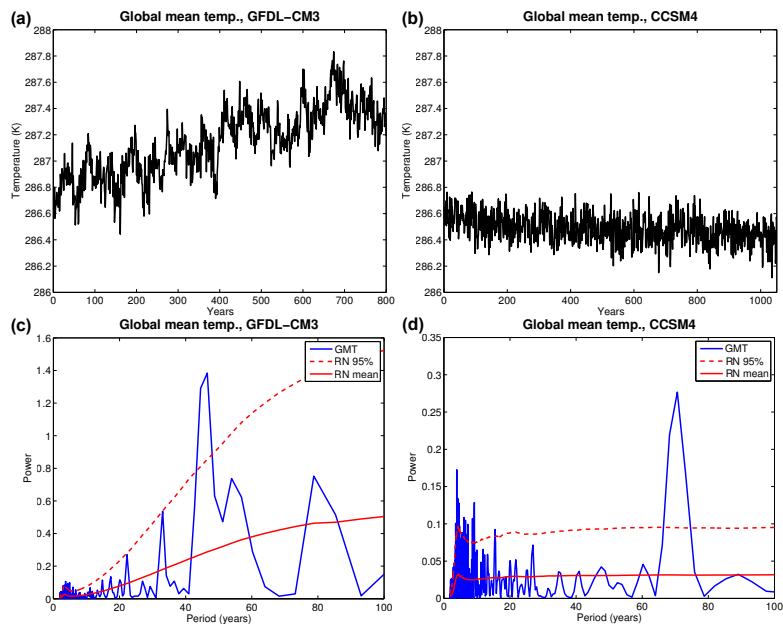
Figure 2. Characterization of the global mean 2 m air temperature variables used in this paper. Panels **(a, b)** show the global mean temperature time series for the pre-industrial control simulations of GFDL-CM3 and CCSM4, respectively. Panels **(c, d)** show their respective power spectra (GMT) with a best-fit red noise (RN) spectrum (computed as in Schneider and Neumaier, 2001) and an estimated 95 % confidence interval.
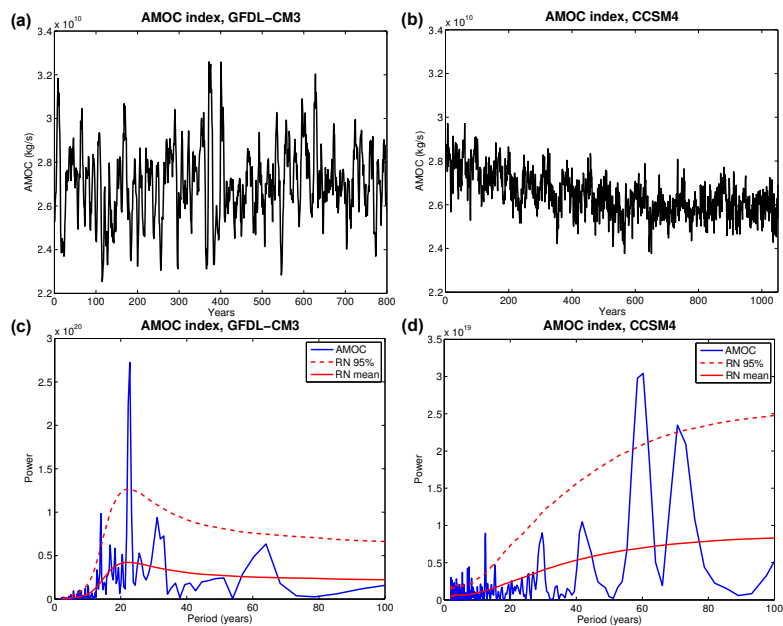
3749

Figure 3. Characterization of the Atlantic meridional overturning circulation (AMOC) index variables used in this paper. Panels **(a, b)** show the AMOC index time series (defined in the text) for the pre-industrial control simulations of GFDL-CM3 and CCSM4, respectively. Panels **(c, d)** show their respective power spectra with a best-fit red noise (RN) spectrum (computed as in Schneider and Neumaier, 2001) and an estimated 95 % confidence interval.
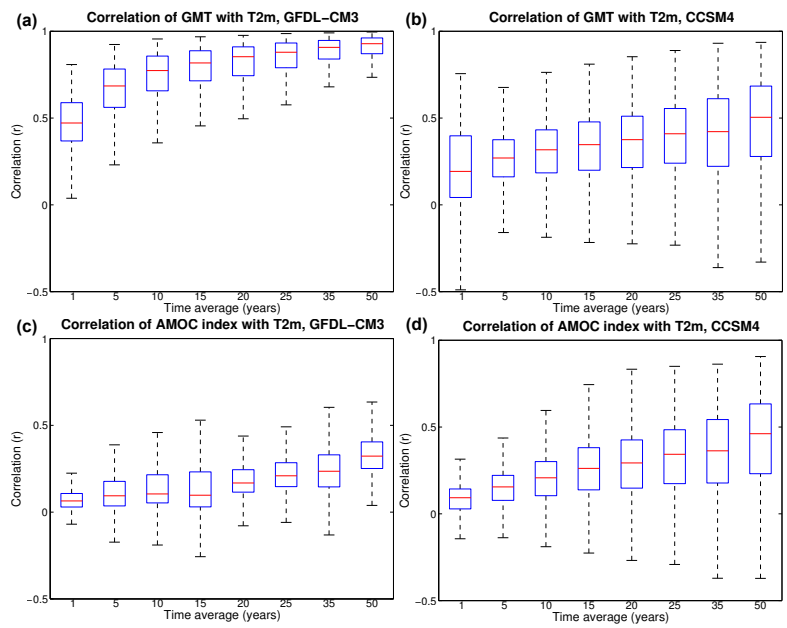
Figure 4. Correlation of global mean 2 m air temperature (GMT) with the spatial 2 m air surface temperatures (T2m), panels **(a, b)**, and correlation of the Atlantic meridional overturning circulation (AMOC) index with the spatial 2 m air surface temperatures, panels **(c, d)**, at a range of time scales. The correlations are computed for each spatial grid point at a given averaging time scale, with the spatial correlation information summarized with these box plots (outliers have been omitted for clarity).
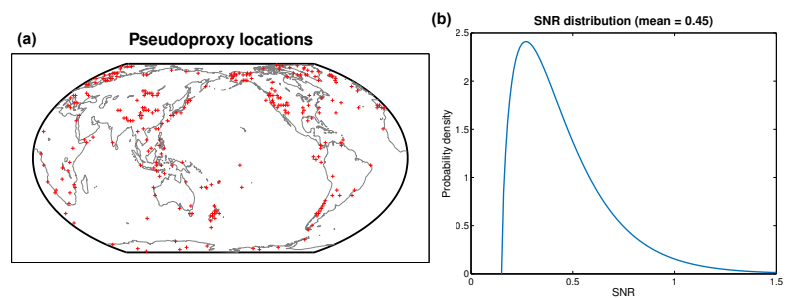
Figure 5. (a) Pseudoproxy locations used in this study ($n = 274$), drawn from the predominantly high resolution (annual) proxy collection of PAGES 2k Consortium (2013) and all the comparatively low resolution (decadal to centennial) proxy locations in Shakun et al. (2012) and Marcott et al. (2013). **(b)** The signal to noise ratio (SNR) distribution for the pseudoproxies, based on a real-world estimate of Wang et al. (2014). For a given Monte Carlo experiment, the SNR for each pseudoproxy was randomly drawn from this distribution.
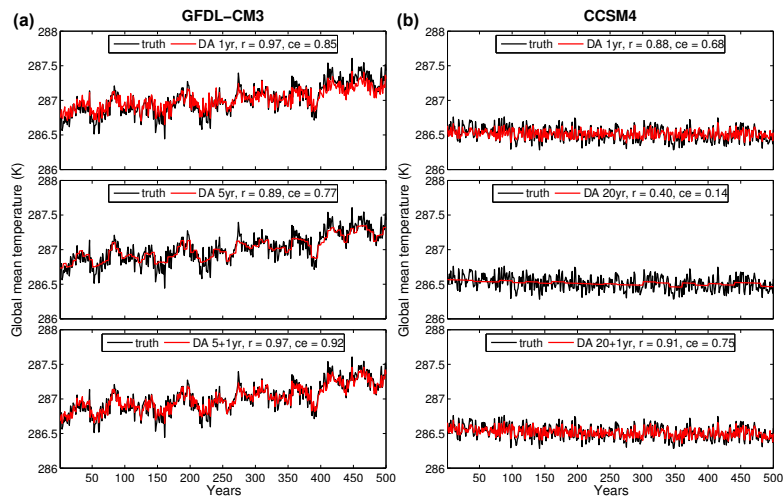
Figure 6. Global mean temperature reconstructions (mean of 100 Monte Carlo iterations) for the three types of experiments discussed in the text and for each climate model simulation. Black lines indicate the true time series while red lines indicate the reconstructed time series for only short time scale (annual) pseudoproxies, only long time scale (5 or 20 years) pseudoproxies, and both long and short time scale pseudoproxies. Skill metrics of the reconstructions, correlation (*r*) and coefficient of efficiency (ce), are shown at the top of each subpanel.
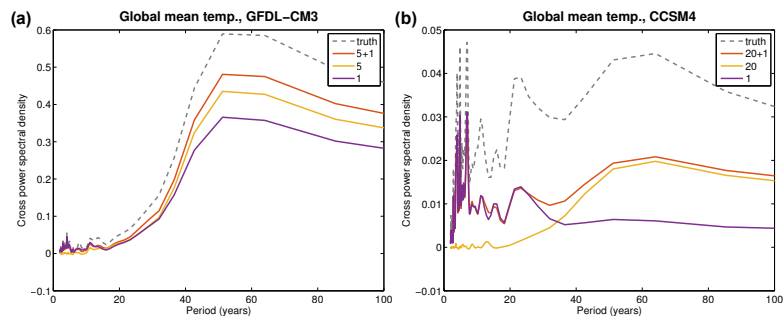
Figure 7. Cross spectra of the reconstructed global mean temperature time series with the true global mean temperature time series, for the reconstructions shown in Fig. 6. For reference, the dotted gray line indicates the cross spectra of the true time series with itself, or equivalently its own power spectrum.
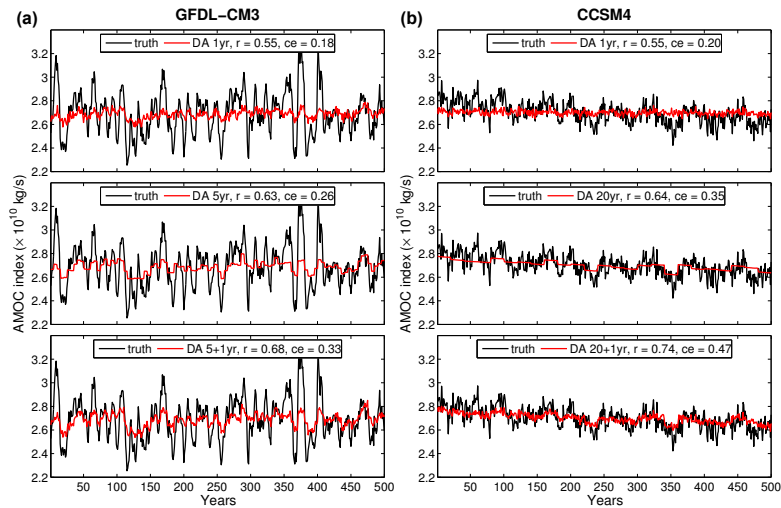
Figure 8. AMOC index reconstructions (mean of 100 Monte Carlo iterations) for the three types of experiments discussed in the text and for each climate model simulation. Black lines indicate the true time series while red lines indicate the reconstructed time series for only short time scale (annual) pseudoproxies, only long time scale (5 or 20 years) pseudoproxies, and both long and short time scale pseudoproxies. Skill metrics of the reconstructions, correlation (r) and coefficient of efficiency (ce), are shown at the top of each subpanel.
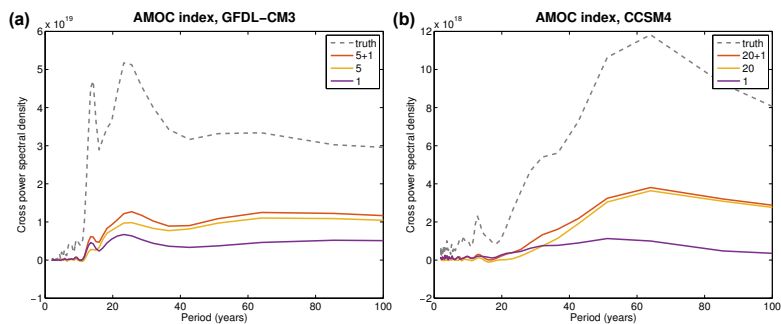
Figure 9. Cross spectra of the reconstructed AMOC index time series with the true AMOC index time series, for the reconstructions shown in Fig. 8. For reference, the dotted gray line indicates the cross spectra of the true time series with itself, or equivalently its own power spectrum.
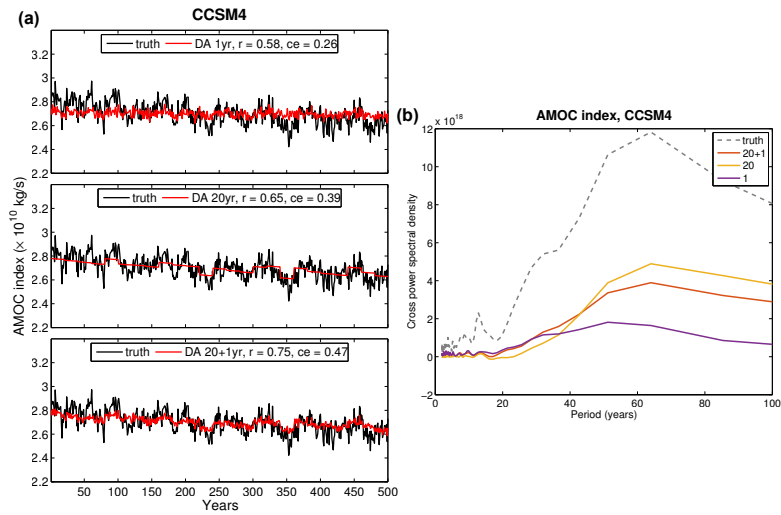
Figure 10. AMOC index reconstructions (mean of 100 Monte Carlo iterations) and corresponding cross-spectra similar to those shown in Figs. 8b and 9b but for the case where each experiment uses 200 pseudoproxies: the single time scale reconstructions use 200 pseudoproxies each, while the multi-time scale reconstructions use 100 pseudoproxies for the short time scale and 100 pseudoproxies for the long time scale.