STATISTICAL PALEOCLIMATE RECONSTRUCTIONS VIA MARKOV RANDOM FIELDS

BY DOMINIQUE GUILLOT[*,1,2],
BALA RAJARATNAM[*,2] AND JULIEN EMILE-GEAY[†,3]

*Stanford University** and University of Southern California*[†]

Understanding centennial scale climate variability requires data sets that are accurate, long, continuous and of broad spatial coverage. Since instrumental measurements are generally only available after 1850, temperature fields must be reconstructed using paleoclimate archives, known as proxies. Various climate field reconstructions (CFR) methods have been proposed to relate past temperature to such proxy networks. In this work, we propose a new CFR method, called GraphEM, based on Gaussian Markov random fields embedded within an EM algorithm. Gaussian Markov random fields provide a natural and flexible framework for modeling high-dimensional spatial fields. At the same time, they provide the parameter reduction necessary for obtaining precise and well-conditioned estimates of the covariance structure, even in the sample-starved setting common in paleoclimate applications. In this paper, we propose and compare the performance of different methods to estimate the graphical structure of climate fields, and demonstrate how the GraphEM algorithm can be used to reconstruct past climate variations. The performance of GraphEM is compared to the widely used CFR method RegEM with regularization via truncated total least squares, using synthetic data. Our results show that GraphEM can yield significant improvements, with uniform gains over space, and far better risk properties. We demonstrate that the spatial structure of temperature fields can be well estimated by graphs where each neighbor is only connected to a few geographically close neighbors, and that the increase in performance is directly related to recovering the underlying sparsity in the covariance of the spatial field. Our work demonstrates how significant improvements can be made in climate reconstruction methods by better modeling the covariance structure of the climate field.

1. Introduction and preliminaries.

1.1. *Introduction.* Fundamental to an informed quantification of recent climate change is an accurate depiction of past climate variability [Masson-Delmotte et al. (2013)]. Since widespread instrumental observations of surface temperatures are only available after the mid-nineteenth century, climate scientists rely on proxy data (e.g., tree rings, ice cores, sediment cores, corals) to infer past temperatures via statistical modeling [Jones et al. (2009), National Research Council (2006)]— a task known as "paleoclimate reconstruction" in the climate literature. Given an instrumental temperature data set [see, e.g., Brohan et al. (2006)] and a global network of climate proxies [e.g., Mann et al. (2008), Figure 2], the temperature back in time can be estimated as a function of proxies.

Various CFR methods have been proposed to infer past climate [see Tingley et al. (2012)]. Here we adopt an approach based on multivariate linear regression as in the regularized EM algorithm [Schneider (2001)]. In that setting, the CFR problem is formalized as a missing data problem, which we now describe.

Consider a spatial grid and let p denote the number of temperature and proxy points. Let $n = n_a + n_m$ denote the sum of the number of years of available instrumental data, n_a, and missing data, n_m. In practice, $p \approx 3000$, $n \approx 2000$ and $n_a \approx 150$ (instrumental period). We model the temperature and proxy points as a multivariate random vector $(X_1, \ldots, X_p) \sim N_p(\mu, \Sigma)$ with missing values, where $\mu = (\mu_1, \ldots, \mu_p)$ is the mean vector and $\Sigma = (\sigma_{ij})_{p \times p}$ is the covariance matrix of the model. We denote by X the (incomplete) $n \times p$ data matrix where each row represents a year of observations containing r instrumental temperature observations and s proxy measurements. Hence, the rows represent time order and the columns represent different spatial locations of both instrumental temperature and proxy data (see Figure 1).
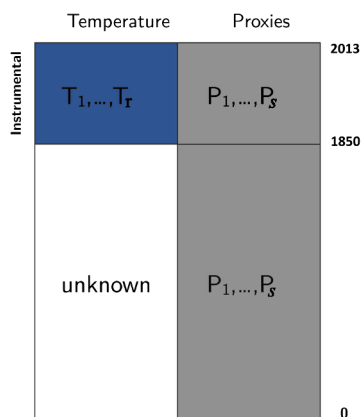

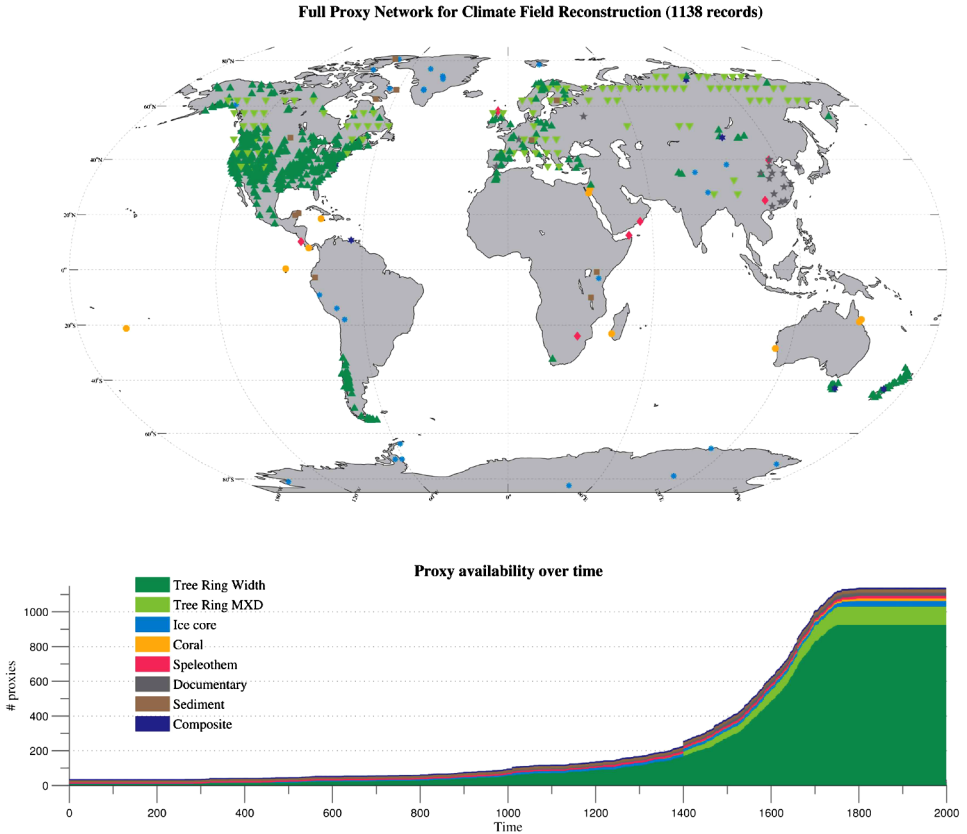FIG. 1. *Temperature/proxy matrix with missing values.*

Full Proxy Network for Climate Field Reconstruction (1138 records)

FIG. 2.   *Location, type and temporal availability of proxies in the* Mann et al. (2008) *database.*

Figure 2 shows that the availability of the proxy data from the network of Mann et al. (2008) decreases rapidly in time, and missing values constitute as much as 80% of the entries in the matrix. Reconstructing the pre-instrumental temperature field may be cast as a missing data problem, for which several strategies exist [Little and Rubin (2002)]. However, the high dimensionality of the problem ("large p, small n") makes it challenging to apply standard methods. For instance, it is well known that the sample covariance matrix is a poor estimator of Σ in that setting [Lin and Perlman (1985), Paul (2007), Stein (1986)]. In this paper, we explore the use of Gaussian Markov random fields (a.k.a. Gaussian graphical models) for estimating Σ. This approach provides flexibility in terms of modeling the inherent spatial heterogeneities of the field, but at the same time reduces the number of parameters that need to be estimated, thereby leading to improved reconstructions of past temperature. We start by recounting existing reconstruction strategies before introducing our new approach.

1.2. *The EM algorithm.* A popular method for the imputation of missing values is the EM algorithm [Dempster, Laird and Rubin (1977), Little and Rubin (2002)]. In the multivariate normal setting, given an estimate of μ and Σ, the EM algorithm reduces to regressing the missing values on the available ones, and thereafter updating the estimates of μ and Σ. This procedure is iterated until convergence. More precisely, let x denote the kth row of X, and let x_a and x_m denote the parts of x where data are available and missing, respectively. Let $\mu^{(0)}$ and $\Sigma^{(0)}$ be initial estimates of μ and Σ. For example, $\mu^{(0)}$ and $\Sigma^{(0)}$ could be the sample mean and sample covariance of the data set completed by replacing every missing value by the mean of the available values in the corresponding columns of X [Schneider (2001)]. The EM algorithm iteratively constructs a sequence $\mu^{(l)}$ and $\Sigma^{(l)}$ of estimates of μ and Σ. For every $l \geq 0$, the E-step consists of a linear regression

$$(1.1) \qquad \left(x_m - \mu_m^{(l)}\right)^\top = B^{(l)}\left(x_a - \mu_a^{(l)}\right)^\top,$$

where

$$(1.2) \qquad B^{(l)} = \Sigma_{ma}^{(l)}\left(\Sigma_{aa}^{(l)}\right)^{-1}, \qquad \Sigma^{(l)} = \begin{pmatrix} \Sigma_{aa}^{(l)} & \Sigma_{am}^{(l)} \\ \Sigma_{ma}^{(l)} & \Sigma_{mm}^{(l)} \end{pmatrix} \quad \text{and}$$

$$\mu^{(l)} = \left(\mu_a^{(l)}, \mu_m^{(l)}\right),$$

are the regression coefficients and the decompositions of $\Sigma^{(l)}$ and $\mu^{(l)}$ associated with the decomposition of x among its available and missing parts. Denote by $X^{(l+1)}$ the completed estimate of X, obtained after the regression (1.1) has been performed in order to impute the missing values in each row of X. In the M-step of the algorithm, the estimates of μ and Σ are updated by

$$(1.3) \qquad \mu_i^{(l+1)} = \frac{1}{n}\sum_{k=1}^{n} X_{ki}^{(l+1)},$$

$$\Sigma_{ij}^{(l+1)} = \frac{1}{n}\sum_{k=1}^{n}\left[(X_{ki}^{(l+1)} - \mu_i^{(l+1)})(X_{kj}^{(l+1)} - \mu_j^{(l+1)})\right] + C_{ij}^{(l+1)},$$

where $C_{ij}^{(l+1)}$ is the covariance of the residuals. Using the same block decomposition as in (1.2), we have

$$(1.4) \qquad C^{(l+1)} = \begin{pmatrix} 0 & 0 \\ 0 & \Sigma_{mm}^{(l)} - \Sigma_{ma}^{(l)}\left(\Sigma_{aa}^{(l)}\right)^{-1}\Sigma_{am}^{(l)} \end{pmatrix}.$$

The reader is referred to Little and Rubin (2002) and McLachlan and Krishnan (2008) for more details about the EM algorithm.

1.3. *The regularized EM algorithm.* Obtaining a precise estimate of Σ is a crucial step of the EM algorithm. In the sample-starved setting common to many paleoclimate problems, the sample covariance matrix is generally not invertible and can be a very poor estimator of Σ. This is a serious problem since parts of Σ need to be inverted to compute the regression coefficients B. Different ℓ_2-type methods to regularize the problem have been proposed in the literature. Among them are ridge regression [a.k.a. Tikhonov regularization, Hanke and Hansen (1993), Hastie, Tibshirani and Friedman (2008), Hoerl and Kennard (1970a, 1970b), Tikhonov and Arsenin (1977)] and truncated total least squares [TTLS, Fierro et al. (1997), Golub and Van Loan (1980), Van Huffel and Vandewalle (1991)] regression. These methods can be used to replace the regression matrix $B^{(l)}$ in equation (1.1) by a regularized estimate, and have been implemented within the EM algorithm. The resulting algorithm is known as RegEM [Schneider (2001)] and has been widely used in paleoclimate studies [Emile-Geay et al. (2013a, 2013b), Mann et al. (2005, 2007a, 2008, 2009), Riedwyl et al. (2009, 2005)]. For example, in RegEM-TTLS, the linear regressions in the EM algorithm are replaced by truncated total least squares (TTLS) regressions. The TTLS solution of a linear system $Ax = b$ is obtained by expressing the total least squares solution of the linear system as a function of the SVD of the matrix A, and then truncating all but a given number of eigenvalues. The number of retained eigenvalues corresponds to the *truncation parameter* of RegEM-TTLS [see Fierro et al. (1997) for more details].

To date, all direct regression methods have resulted in reconstructions that underestimate the amplitude of past climate variations to some extent [e.g., Smerdon et al. (2010, 2011), von Storch et al. (2004)]. This "regression dilution" [Frost and Thompson (2000)] is a direct consequence of modeling the temperature conditional on (noisy) proxy values [Christiansen (2010, 2013), Tingley and Li (2012), von Storch et al. (2004)]. Regularization may compound this problem, as with ridge regression the smoothness of the filter factors has been shown to leak energy from the leading SVD modes, resulting in overly damped estimates of past temperature [Smerdon and Kaplan (2007)]. This problem may be mitigated via TTLS [Mann et al. (2007b)], as it attempts to correct for regression dilution by steepening the regression slope; however, the solution is no longer guaranteed to be optimal even under broad assumptions [Carroll and Ruppert (1996)]. Furthermore, a major shortcoming of TTLS as currently used in climate applications is that the truncation parameter must be specified a priori, rather than being estimated adaptively. Given the applicability of the RegEM algorithm for missing data problems in the paleoclimate context (e.g., surface temperature reconstructions for the past 2000 years), we seek to develop an imputation method that rests on a more accurate and data-adaptive estimate of Σ itself.

1.4. *Gaussian Markov random fields.* A GMRF is a multivariate normal model which encodes conditional independence structure between variables [see

Lauritzen (1996), Whittaker (1990)]. More precisely, let $(X_1, \ldots, X_p)$ be a multivariate random vector with *inverse covariance matrix* (or *precision matrix*) $\Omega = (\omega_{ij}) = \Sigma^{-1}$. The *partial correlation coefficient* between X_i and X_j given the rest of the variables, denoted by $\rho_{ij|\text{rest}}$, can be obtained from the inverse covariance matrix [see Whittaker (1990), Corollary 5.8.2], and is given as follows:

$$(1.5) \qquad \rho_{ij|\text{rest}} = \frac{-\omega_{ij}}{\sqrt{\omega_{ii}\omega_{jj}}}.$$

In the case of multivariate normal data, one can show that $\rho_{ij|\text{rest}} = 0$ if and only if X_i is independent of X_j given the rest of the variables [Whittaker (1990), Corollary 6.3.4]. The zeros in the precision matrix therefore indicate conditional independence between the corresponding variables. The conditional independence relations in a distribution can be conveniently encoded using a *graph*. Recall that a graph $G = (V, E)$ is a pair of sets V and $E \subseteq V \times V$, where each element of V represents a vertex of the graph and each point of E is a pair of elements of V. We encode the conditional independence relations by adding an edge between i and j if and only if X_i is not conditionally independent of X_j given the rest of the variables. The random vector $(X_1, \ldots, X_p)$ is then said to satisfy the *pairwise Markov property* with respect to the graph G. For details on the *pairwise*, *local* and *global* Markov properties, we refer the reader to Lauritzen (1996) and Whittaker (1990).

Once the conditional independence structure (or graphical structure) of a Gaussian random vector is known, this information can be used for estimating its covariance matrix Σ. More specifically, given an i.i.d. sample $x_1, \ldots, x_n$ of $(X_1, \ldots, X_p)$ with mean $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, and a graph G, the *graphical maximum likelihood estimator* of Σ can be computed by solving

$$(1.6) \qquad \hat{\Sigma}_G = \underset{\substack{\Sigma = \Omega^{-1} > 0 \\ \omega_{ij} = 0, (i,j) \notin E}}{\operatorname{argmax}} \ \log \det \Omega - \operatorname{tr}(S\Omega),$$

where S is the sample covariance matrix of $x_1, \ldots, x_n$, given by

$$(1.7) \qquad S = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})(x_i - \overline{x})^{\top},$$

and $\log \det \Omega - \operatorname{tr}(S\Omega)$ is (up to a constant) the multivariate normal profile log-likelihood function. The problem (1.6) can be solved efficiently for up to a few thousand variables using, for example, regression-based algorithms [see Hastie, Tibshirani and Friedman (2008), Algorithm 17.1]. The resulting matrix $\hat{\Sigma}_G$ is generally a better estimate than the widely used sample covariance matrix, especially when the number of observations n is smaller than the number of variables p.

In this paper, we propose a methodology that combines graphical models with the EM algorithm for the purpose of reconstructing past temperature fields. In our approach, we first model the conditional independence structure of the target field based on structural assumptions or directly from the data. A sparse estimate of Σ

is then obtained in accordance with this graphical structure at every step of the EM algorithm. This approach greatly reduces the number of parameters to estimate, leads to better conditioned and more precise estimates of Σ, and also exploits the natural conditional independence structure of the spatial field. The regression step (1.1) can then be performed using any regularization method (or even no regularization at all). We call the resulting algorithm GraphEM (see Algorithm 1 in Appendix A; see also Appendix B for the derivation of the GraphEM algorithm within the EM framework).

The rest of the paper is structured as follows. In Section 2 we explore various methods to estimate the graphical structure of the joint temperature/proxy field. We then test the performance of GraphEM in a realistic geophysical context in Sections 3 and 4. The characteristics of the estimated conditional independence structures are then studied in Section 5. We conclude with a discussion section.

2. Methodology. Different methods have been proposed in the literature to discover the conditional independence relations (or *graphical structure*) of a data set, in either the Bayesian or frequentist framework [see e.g., Banerjee, El Ghaoui and d'Aspremont (2008), Dawid and Lauritzen (1993), Friedman, Hastie and Tibshirani (2008), Letac and Massam (2007), Rajaratnam, Massam and Carvalho (2008)]. In this work, we explore two different approaches: ℓ_1-penalized maximum likelihood [Banerjee, El Ghaoui and d'Aspremont (2008), Friedman, Hastie and Tibshirani (2008), Guillot et al. (2012), Hsieh et al. (2011)] and neighborhood graphs.

2.1. *ℓ_1-penalized maximum likelihood.* A flexible approach for obtaining a sparse estimate of the precision matrix Ω is to maximize the normal likelihood subject to an ℓ_1 penalty on its norm. More specifically, the ℓ_1-penalized maximum likelihood problem consists of solving

$$(2.1) \qquad \max_{\Omega > 0} l(\Omega) - \rho \|\Omega\|_1,$$

where $\Omega = \Sigma^{-1}$ denotes the precision matrix of the data, $l(\Omega)$ is the normal log-likelihood of Ω, $\rho > 0$ is a regularization parameter, and $\|\Omega\|_1$ is the 1-norm of Ω:

$$(2.2) \qquad \|\Omega\|_1 = \sum_{i=1}^{p} \sum_{j=1}^{p} |\omega_{ij}|.$$

The use of an ℓ_1 penalty, as first introduced in the context of the LASSO regression [Tibshirani (1996)], favors the introduction of zero elements and thus leads to sparse solutions [see Hastie, Tibshirani and Friedman (2008), Section 3.4.3]. At the same time, using an ℓ_1 penalty leads to a convex problem that can be solved efficiently using modern methods of convex optimization. Once an estimate of Ω is known, the associated graph can be inferred from the pattern of zeros in Ω. In

this work, we employ the *graphical lasso* (glasso) algorithm of Friedman, Hastie and Tibshirani (2008) to obtain a sparse estimate of Ω by solving an ℓ_1-penalized likelihood problem. As ρ varies, the matrix $\hat{\Omega}$ displays different sparsity patterns. When $\rho = 0$ and $n \geq p$, there is no penalty and $\hat{\Omega}$ is equal to the maximum likelihood estimate S^{-1} of Ω, where S denotes the sample covariance matrix of the data matrix. The estimate $\hat{\Omega}$ tends to a diagonal matrix as the regularization parameter ρ is increased. Problem (2.1) can also be easily modified to use a different penalty for different parts of the matrices. Consider, for example, the precision matrix of a temperature/proxies field. The matrix can be organized in block form

$$(2.3) \qquad \Omega = \begin{pmatrix} \Omega_{TT} & \Omega_{TP} \\ \Omega_{PT} & \Omega_{PP} \end{pmatrix},$$

where Ω_{TT}, Ω_{TP} and Ω_{PP} are block matrices corresponding to the temperature/temperature, temperature/proxy and proxy/proxy parts of the matrix. Since the signal contained in proxies is generally weaker than the temperature signal, it may be sensible to use different penalty parameters for different parts of the matrix when solving the ℓ_1-penalized maximum likelihood problem. Problem (2.1) can thus be replaced by

$$(2.4) \qquad \max_{\Omega > 0} l(\Omega) - \rho_{TT} \|\Omega_{TT}\|_1 - 2\rho_{TP} \|\Omega_{TP}\|_1 - \rho_{PP} \|\Omega_{PP}\|_1,$$

where ρ_{TT}, ρ_{TP}, $\rho_{PP} > 0$ are regularization parameters. This problem can also be solved efficiently by using a modified graphical lasso algorithm [see Friedman, Hastie and Tibshirani (2008), equation (15)]. Figure 3 displays the temperature neighbors of a few locations for a graph estimated using (2.4) (sparsity level = 1.4%) on a real temperature data set [Brohan et al. (2006)], and illustrates the potential of the ℓ_1 method to detect real geophysical structures. Note that the method correctly identifies anisotropic climate features like the equatorial Pacific cold tongue (left), the California current system (center) and east Atlantic structures related to the subtropical gyre circulation (right).

In practice, choosing suitable penalty parameters in (2.1) or (2.4) can be difficult. A high penalty forces many zero entries in the precision matrix, while a low penalty adds some edges that make little geophysical sense. An optimal choice should strike a balance between those extremes. If $\rho > \rho_{max} := \max_{i \neq j} |S_{ij}|$, it can be shown [see, e.g., Witten, Friedman and Simon (2011), Theorem 2] that the
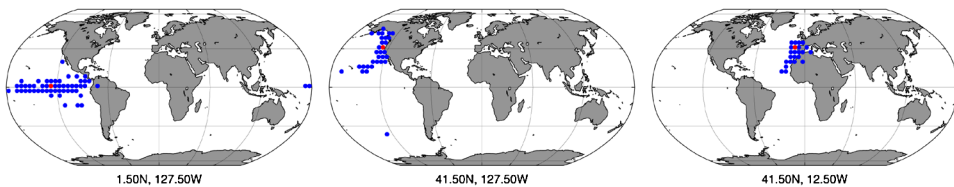

FIG. 3. *Example of estimated graphical structure of a temperature field (HadCRUT3v).*

resulting glasso estimate of Σ is a diagonal matrix. A relevant finite number of regularization parameter values can therefore be obtained by dividing the interval between some small value $\rho_{\min}$ and the biggest relevant value $\rho_{\max}$. In our numerical work, we have chosen $\rho_{\min} = 0.1 \cdot \rho_{\max}$ and have divided the interval $[\rho_{\min}, \rho_{\max}]$ into 10 values. Problem (2.1) can then be solved for each of these penalty parameters to obtain estimates $\hat{\Omega}$ of the precision matrix Ω. To each estimate corresponds a graph based on the structure of zeros in $\hat{\Omega}$. When the dimension of the problem to solve is small (e.g., in regional reconstructions) or a single penalty parameter is used for the whole precision matrix [as in equation (2.1)], an optimal parameter can be chosen using k-fold cross-validation. However, when a different penalty parameter is used for each part of the precision matrix, performing cross-validation for an array of regularization parameters (e.g., a $10 \times 10 \times 10$ grid of penalty parameters) incurs a prohibitive computational cost. A possible solution consists of searching for a graph that is (a) dense enough to capture the salient spatial dependences, and (b) sparse enough to make the reconstruction possible and stable (by reducing the dimension of the problem to a size comparable to the sample size). A triple $(\rho_{\mathrm{TT}}, \rho_{\mathrm{TP}}, \rho_{\mathrm{PP}})$ of regularization parameters with the desired sparsity can be chosen by starting with large values of the three penalty parameters, and progressively reducing the value of each penalty parameter until a given target sparsity is obtained for each part of the precision matrix. This technique requires computing the solution of problem (2.4) at only a few points of the grid. This sparsity approach is implemented in our proposed version of GraphEM, and is compared to the neighborhood approach described below in Section 4. In this paper, we have chosen fixed sparsity levels when performing large reconstruction ensembles, after verifying via targeted experiments that the specified sparsity levels were close to those deemed optimal by 5-fold cross-validation.

2.2. *Neighborhood graphs.* Since temperature variations at a given point are to a large extent explained by temperature of surrounding points, it is natural to use a *neighborhood graph* (i.e., a graph where two vertices are connected if and only if they are within a specified radius R) to approximate the true graphical structure of the joint temperature/proxy field; see, for example, Cook et al. (1999) where a similar assumption was made. The radius can be either specified or chosen from the data. As we illustrate in Figure 4, the choice of an optimal radius can be made by performing cross-validation over the instrumental period and choosing the radius that minimizes the MSE of the reconstructed values. Besides this natural and meaningful Markov random field structure in spatial temperature fields, a neighborhood graph approach has the distinct advantage that the underlying graph does not have to be estimated from sample-deficient high-dimensional data, and that the procedure does not require solving computationally intensive optimization problems. Dimensionality reduction is achieved with great ease and at the same time has an intuitive geophysical interpretation; sparsity is entirely governed by the
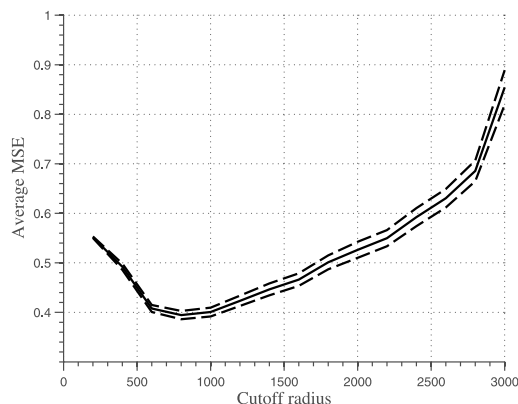
FIG. 4. *Cross-validation scores for choosing a neighborhood graph radius* (5-*fold cross-valida-tion*).

neighborhood radius R. On the other hand, neighborhood graphs are less flexible and cannot model in an adaptive way (1) conditional independence relations result-ing from anisotropic structures present in the data (such as land/ocean boundaries, mountain ranges, atmospheric flow patterns, etc.), and (2) long range dependen-cies that arise due to teleconnections. However, when the noise level is too high, a simple model such as a neighborhood graph may be preferable to the ℓ_1-penalized covariance estimation method.

As an illustration, Figure 5 displays the neighborhood of size 800 km, 1000 km and 1200 km at different locations with the same latitude on a $5° \times 5°$ grid. The average number of neighbors (and their standard deviation) are 8.42 (2.08), 10.24 (3.37) and 16.61 (4.92), respectively.

We also consider sparser variants of the neighborhood graph model based on the structure of the paleoclimate reconstruction problem. First, since climate is the
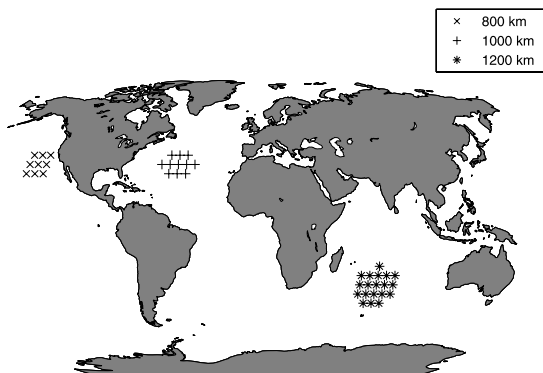

FIG. 5. *Illustration of neighborhood graphs.*

TABLE 1
Neighborhood graph variants for the joint temperature/proxy field

Name	TT	TP	PP
Neigh	Neighborhood	Neighborhood	Neighborhood
Ind$_{PP}$	Neighborhood	Neighborhood	Diagonal
CAR$_{TP}$	Neighborhood	CAR	Diagonal
CAR$_{TT}$	CAR	Neighborhood	Diagonal
CAR$_{TT+TP}$	CAR	CAR	Diagonal

signal shared between proxies, it is natural to assume that the proxies are independent of each other conditional on the temperature data (i.e., to assume that Ω_{PP} is diagonal). We thus explore a simpler model where the temperature/temperature (TT) and the temperature/proxy (TP) parts of the graph are constructed as above with a neighborhood graph, but where Ω_{PP} is diagonal. Further, since temperature proxies are reflective of local temperature only, it is natural to impose a local structure in Ω_{PT} as well (i.e., $\Omega_{PT} = 0$ except for the columns corresponding to each proxy's closest temperature grid point). Finally, given that the optimal neighborhood graphs chosen by cross-validation tend to feature only the immediate neighbors of each temperature gridpoint, it is natural to impose such constraints on the TT part of the graph a priori. Note that such a model is equivalent to a spatial conditionally auto-regressive (CAR) model [Besag (1974)]. The variants considered in the paper are summarized in Table 1, and their performance in modeling the conditional independence structure of the temperature/proxy field is studied in Section 4.

3. Validation via pseudoproxy experiments.

3.1. *Background.* In the climate literature, pseudoproxy experiments have become the method of choice to objectively evaluate the performance of CFR techniques against a geophysically-relevant target [see Smerdon (2011) for a recent review]. This target temperature field is often the output of coupled general circulation model (GCM) simulations for the past 1000 years or so, sampled at a fixed spatiotemporal resolution. Although GCM-simulated temperature fields do not exactly match the characteristics of observed temperature fields, they are generated in accordance with physical laws embedded in such models, and thus provide a controlled, realistic framework to test reconstruction methods.

In practice, a pseudoproxy is obtained by adding noise to a GCM-simulated temperature field at locations where proxy observations are available in the real world. Because such observations are sparse, the pseudoproxy network therefore comprises a small collection of time series. Given only knowledge of the temperature field over a 150-year calibration interval, the CFR method is then used to

backcast a thousand-year long temperature field based on this relatively small sample of noisy temperature time series. Given a simulated temperature field $T(l, t)$ at location l and time t (standardized to have mean 0 and variance 1 over time) from a GCM model, the pseudoproxies $P(l, t)$ are constructed as follows:

$$(3.1) \qquad P(l, t) = T(l, t) + \frac{1}{\text{SNR}} \cdot \xi(l, t),$$

where $\xi(l, t)$ are independent realizations of a Gaussian white noise process, and the (scalar) signal-to-noise ratio SNR controls the amount of noise in the pseudoproxy. Although pseudoproxies constitute an oversimplification of reality, they have been used extensively in the climate literature [Annan and Hargreaves (2012), Bradley (1996), Christiansen, Schmith and Thejll (2009), Li and Smerdon (2012), Mann and Rutherford (2002), Smerdon (2011), Smerdon et al. (2011), Tingley and Huybers (2010a)] to provide a numerical laboratory to test the performance of CFR methods.

In our simulations, we used the NCAR CSM 1.4 model experiment [Ammann et al. (2007)], which simulates the climate of the last millennium (850–1980 AD) on a $5° \times 5°$ grid. As per previous work [Li, Nychka and Ammann (2010), Mann et al. (2007a), Smerdon et al. (2011), Wang et al. (2014)], the locations of the pseudoproxies were chosen in accordance with Mann, Bradley and Hughes (1998) (MBH98, Figure 6) and the value of SNR has been fixed to 0.5. Other SNR values have also been investigated but, for the sake of brevity, are not presented here. The last 150 years of data have been used as a calibration period, and the remaining 981 years of temperature data have been reconstructed using GraphEM. As a benchmark, we follow recent work [Steiger et al. (2013), Tingley and Huybers (2010a)] and use RegEM-TTLS, which was widely used in high-profile climate reconstructions [Mann et al. (2008, 2009)].

3.2. *Performance metrics.*   Various metrics have been used in the literature to measure the quality of CFR methods and reconstructed temperature fields [Bürger (2007), Cook, Briffa and Jones (1994)]. Let $T(l, t)$ denote the temperature at a location l and at time t, and denote by $\hat{T}(l, t)$ a reconstruction of $T(l, t)$. The
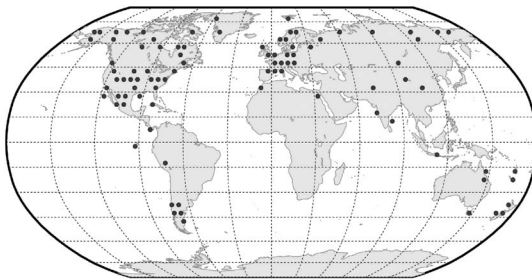

FIG. 6.   *Geographic location of the pseudoproxies in the MBH98 database.*

mean squared error (MSE) measures the mean difference between the two fields at a given location l:

$$(3.2) \qquad \text{MSE}(\hat{T})(l) = \frac{1}{N} \sum_t (T(l, t) - \hat{T}(l, t))^2,$$

where N is the number of time points. To measure the improvement made by our proposed graphical method, we define the *relative MSE difference* at a location l by

$$\text{relative MSE difference } (l) = \frac{\text{MSE}_{\text{RegEM-TTLS}}(l) - \text{MSE}_{\text{GraphEM}}(l)}{\text{MSE}_{\text{RegEM-TTLS}}(l)}.$$

Although a small MSE indicates a good reconstruction, it is not immediately clear how small the MSE has to be for the reconstruction to be considered a "good reconstruction." A useful approach is to compare the MSE of a given reconstruction to that of a reconstruction that is equal to a constant value over time (a "constant reconstruction"). The *reduction of error* (RE) compares the MSE of a given reconstruction to a constant reconstruction equal to the mean temperature of the field $\overline{T}_c(l)$ over the calibration period:

$$(3.3) \qquad \text{RE}(l) = 1 - \frac{\text{MSE}(\hat{T})(l)}{\text{MSE}(\overline{\overline{T}}_c)(l)}.$$

Similarly, the *coefficient of efficiency* (CE) compares the MSE of the reconstruction to a constant reconstruction equal to the mean of the temperature field $\overline{T}_v(l)$ over the validation interval:

$$(3.4) \qquad \text{CE}(l) = 1 - \frac{\text{MSE}(\hat{T})(l)}{\text{MSE}(\overline{\overline{T}}_v)(l)}.$$

Finally, the bias at point l is the difference between $\hat{T}(l, \cdot)$ and $T(l, \cdot)$ averaged over time. A perfect reconstruction would have a MSE of 0, a CE and a RE of 1 and a bias of 0. The closer to those values, the better the reconstruction.

4. Results. In order to test the performance and the sensitivity of GraphEM to reconstruct temperature over the whole globe, we performed 50 reconstructions, each corresponding to a different noise realization $\xi(l, t)$. The performance of GraphEM is then compared to the performance of RegEM-TTLS. The truncation parameter was set to 5, but the results show little sensitivity to this choice.

To study the performance of GraphEM, reconstructions were performed using both the neighborhood graph methods and the ℓ_1 method (Section 2). For illustration purposes, in Sections 4.1 and 4.2, we present detailed results for the neighborhood graph method with a cutoff radius of 800 km, as suggested by cross-validation (see Figure 4). Verification statistics for other cutoff radii, for the neighborhood graph variants and for the ℓ_1 method are also provided in Tables 2 and 3.
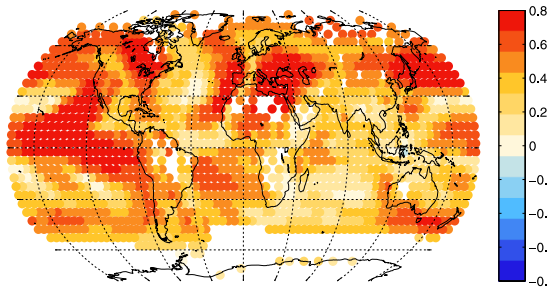
FIG. 7. *Relative MSE improvement for SNR = 0.5 (neighborhood graph, cutoff radius = 800 km).*

4.1. *Spatial reconstructions.* We begin by studying the performance of GraphEM in space. Figure 7 displays the average relative MSE improvement for the 50 reconstructions, and shows that the improvement can be substantial when using GraphEM. The improvement is positive for almost every location. The average improvement is about 43%, whereas improvements as large as 80% are recorded in certain regions. Figure 7 also provides some compelling evidence that the magnitude of the percentage improvement appears to be even greater at some locations that are distant from proxy sites. In particular, vast swathes of the entire central and northern Pacific stretching from East Asia to North and central America display significantly higher improvements in MSE. The same appears to be true for parts of the southern Atlantic. This is remarkable given the high degree of locality of the chosen graph. Hence, a local graph does not translate into short-range correlations; on the contrary, it can actually improve the representation of long-range dependencies. Improvements over the Indian ocean, however, tend to be modest perhaps because of the paucity of data.

Figure 8(a) and (b) display the CE statistics (averaged over the 50 noise realizations) for RegEM-TTLS and GraphEM, respectively. Again, in many regions, GraphEM leads to substantial improvements, particularly where the skill was very poor with RegEM-TTLS. The different precision metrics averaged over space (for
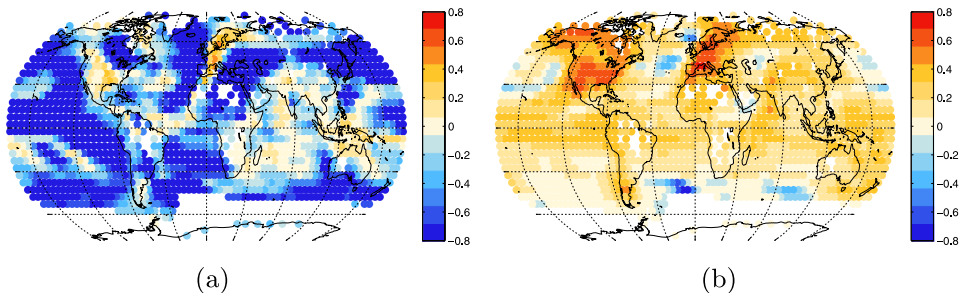

(a)   (b)

FIG. 8. *CE map for the* (a) *RegEM-TTLS and* (b) *GraphEM reconstructions for SNR = 0.5 (neighborhood graph, cutoff radius = 800 km).*

Mean (and standard deviation) of the performance metrics averaged over space for the global reconstructions

Method	MSE	RE	CE	Bias
ℓ_1 method				
GraphEM (0.3% target sparsity)	0.44 (0.01)	0.33 (0.01)	0.11 (0.02)	0.09 (0.01)
GraphEM (0.5% target sparsity)	0.42 (0.01)	0.36 (0.01)	0.15 (0.01)	0.08 (0.01)
GraphEM (0.7% target sparsity)	0.41 (0.01)	0.36 (0.01)	0.16 (0.01)	0.08 (0.01)
GraphEM (0.9% target sparsity)	0.41 (0.01)	0.36 (0.01)	0.15 (0.01)	0.08 (0.01)
Neigh				
GraphEM (600 km radius)	0.42 (0.01)	0.35 (0.01)	0.14 (0.01)	0.06 (0.01)
GraphEM (800 km radius)	0.39 (0.01)	0.39 (0.01)	0.19 (0.01)	0.06 (0.01)
GraphEM (1000 km radius)	0.40 (0.01)	0.38 (0.01)	0.18 (0.01)	0.06 (0.01)
GraphEM (1200 km radius)	0.41 (0.01)	0.36 (0.01)	0.16 (0.01)	0.06 (0.01)
Ind$_{PP}$				
GraphEM (600 km radius)	0.42 (0.01)	0.35 (0.01)	0.13 (0.01)	0.06 (0.01)
GraphEM (800 km radius)	0.39 (0.01)	0.39 (0.01)	0.19 (0.01)	0.06 (0.01)
GraphEM (1000 km radius)	0.40 (0.01)	0.38 (0.01)	0.19 (0.01)	0.06 (0.01)
GraphEM (1200 km radius)	0.41 (0.01)	0.37 (0.01)	0.16 (0.01)	0.06 (0.01)
CAR$_{TP}$				
GraphEM (600 km radius)	0.42 (0.01)	0.35 (0.01)	0.14 (0.01)	0.06 (0.01)
GraphEM (800 km radius)	0.39 (0.01)	0.39 (0.01)	0.19 (0.01)	0.06 (0.01)
GraphEM (1000 km radius)	0.39 (0.01)	0.39 (0.01)	0.19 (0.01)	0.06 (0.01)
GraphEM (1200 km radius)	0.40 (0.01)	0.38 (0.01)	0.18 (0.01)	0.06 (0.01)
CAR$_{TT}$				
GraphEM (600 km radius)	0.39 (0.01)	0.39 (0.01)	0.20 (0.01)	0.06 (0.01)
GraphEM (800 km radius)	0.39 (0.01)	0.39 (0.01)	0.19 (0.01)	0.06 (0.01)
GraphEM (1000 km radius)	0.40 (0.01)	0.39 (0.01)	0.19 (0.01)	0.06 (0.01)
GraphEM (1200 km radius)	0.40 (0.01)	0.38 (0.01)	0.18 (0.01)	0.06 (0.01)
CAR$_{TT+TP}$				
GraphEM (600 km radius)	0.39 (0.01)	0.39 (0.01)	0.19 (0.01)	0.06 (0.01)
GraphEM (800 km radius)	0.39 (0.01)	0.39 (0.01)	0.19 (0.01)	0.06 (0.01)
GraphEM (1000 km radius)	0.39 (0.01)	0.39 (0.01)	0.19 (0.01)	0.06 (0.01)
GraphEM (1200 km radius)	0.39 (0.01)	0.39 (0.01)	0.19 (0.01)	0.06 (0.01)
RegEM-TTLS				
RegEM-TTLS	0.84 (0.10)	-0.24 (0.14)	-0.61 (0.19)	0.01 (0.02)

the unsmoothed reconstruction) are presented in Table 2 along with their standard deviation computed using the 50 reconstructions. This table confirms once more that GraphEM performs better spatially and is more stable than RegEM-TTLS.

Although the results presented in Table 2 are quite similar for the different GraphEM methods, the neighborhood graphs seem to perform slightly better than the ℓ_1 method. They could therefore be useful in noisy cases for which discovering the structure of the field from the data is difficult. Another advantage of the neighborhood method is that the cutoff radius is easy to choose by cross-validation.

In comparison, choosing appropriate regularization parameters to use with the ℓ_1 method is computationally intensive.

We also observe that the four neighborhood graph variants produce very similar results. In particular, the simplest graph CAR_{TT} performs quite well, underlying the importance of locality in the temperature/proxy field. In the pseudo proxy experiment, the better validation metrics also correspond to the models that best reflect the data generating mechanism, that is, the models where Ω_{PP} is diagonal. Climate fields found in nature may display a more complex structure, but to the extent that it can be reasonably approximated by a neighborhood graph, our results suggest that GraphEM could produce very skillful reconstructions.

The results also demonstrate that a larger graph (e.g., neighborhood 1200 km vs CAR) can still lead to a very good reconstruction. This is to be expected since an edge between two vertices does not prohibit the corresponding entry in Ω from being very small. Thus, a graph containing a certain number of spurious edges (such as the graphs obtained from the ℓ_1 method) may still perform well, which means that results are broadly insensitive to the graph density. Finally, we note that although the ℓ_1 method performs slightly worse in our experiments, it has the potential to detect real geophysical structures, and could lead to improvements when working with data sets with a stronger signal.

4.2. *Spatial average.* The spatial reconstructions given by RegEM-TTLS and GraphEM can also be averaged over space to obtain (area-weighted) spatial averages. Figure 9 displays a 95% deviation band (constructed using the 50 reconstructions) for the mean temperature series reconstructed with RegEM-TTLS and GraphEM. The instrumental period is also reconstructed via the pseudoproxies
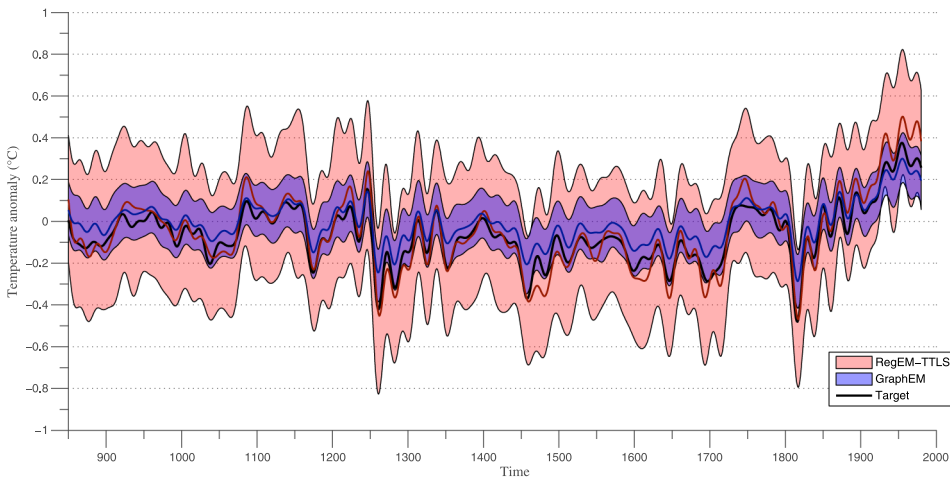

FIG. 9. *Global spatial averages (multiple noise realizations, 95% deviation interval) for SNR = 0.5 (neighborhood graph, cutoff radius = 800 km). The thick lines represent the median of each ensemble.*

using the estimated mean and covariance matrix obtained from GraphEM. The uncertainty bands have been obtained by computing the (weighted) average temperature at each time for each reconstruction, and then constructing a confidence interval containing 95% of the 50 simulated values. A 20 year low-pass filter has been applied after computing the quantiles for illustration and interpretation purposes. The mean width of the deviation interval for GraphEM and RegEM-TTLS are 0.25 and 0.66, respectively. The associated reconstruction statistics are provided in Table 3. Note that the CE scores for GraphEM are significantly larger

TABLE 3

Mean (and standard deviation) of the performance metrics for the spatial average reconstructions

Method	MSE	RE	CE	Bias
ℓ_1 method				
GraphEM (0.3% target sparsity)	0.12 (0.01)	0.75 (0.02)	0.25 (0.08)	0.09 (0.01)
GraphEM (0.5% target sparsity)	0.11 (0.01)	0.79 (0.02)	0.36 (0.05)	0.08 (0.01)
GraphEM (0.7% target sparsity)	0.11 (0.01)	0.79 (0.02)	0.37 (0.05)	0.08 (0.01)
GraphEM (0.9% target sparsity)	0.11 (0.01)	0.79 (0.02)	0.36 (0.05)	0.08 (0.01)
Neigh				
GraphEM (600 km)	0.12 (0.01)	0.82 (0.01)	0.46 (0.04)	0.06 (0.01)
GraphEM (800 km)	0.11 (0.01)	0.83 (0.01)	0.50 (0.04)	0.06 (0.01)
GraphEM (1000 km)	0.10 (0.01)	0.83 (0.01)	0.50 (0.04)	0.06 (0.01)
GraphEM (1200 km)	0.10 (0.01)	0.83 (0.01)	0.48 (0.04)	0.06 (0.01)
Ind$_{PP}$				
GraphEM (600 km radius)	0.12 (0.01)	0.82 (0.01)	0.46 (0.04)	0.06 (0.01)
GraphEM (800 km radius)	0.11 (0.01)	0.83 (0.01)	0.50 (0.04)	0.06 (0.01)
GraphEM (1000 km radius)	0.11 (0.01)	0.83 (0.01)	0.50 (0.04)	0.06 (0.01)
GraphEM (1200 km radius)	0.11 (0.01)	0.83 (0.01)	0.48 (0.04)	0.06 (0.01)
CAR$_{TP}$				
GraphEM (600 km radius)	0.12 (0.01)	0.83 (0.01)	0.47 (0.04)	0.06 (0.01)
GraphEM (800 km radius)	0.11 (0.01)	0.83 (0.01)	0.49 (0.04)	0.06 (0.01)
GraphEM (1000 km radius)	0.11 (0.01)	0.83 (0.01)	0.50 (0.04)	0.06 (0.01)
GraphEM (1200 km radius)	0.11 (0.01)	0.83 (0.01)	0.49 (0.04)	0.06 (0.01)
CAR$_{TT}$				
GraphEM (600 km radius)	0.11 (0.01)	0.83 (0.01)	0.49 (0.04)	0.06 (0.01)
GraphEM (800 km radius)	0.11 (0.01)	0.83 (0.01)	0.49 (0.04)	0.06 (0.01)
GraphEM (1000 km radius)	0.11 (0.01)	0.83 (0.01)	0.49 (0.04)	0.06 (0.01)
GraphEM (1200 km radius)	0.11 (0.01)	0.84 (0.01)	0.50 (0.04)	0.06 (0.01)
CAR$_{TT+TP}$				
GraphEM (600 km radius)	0.11 (0.01)	0.83 (0.01)	0.47 (0.04)	0.06 (0.01)
GraphEM (800 km radius)	0.11 (0.01)	0.83 (0.01)	0.47 (0.04)	0.06 (0.01)
GraphEM (1000 km radius)	0.11 (0.01)	0.83 (0.01)	0.47 (0.04)	0.06 (0.01)
GraphEM (1200 km radius)	0.11 (0.01)	0.83 (0.01)	0.47 (0.04)	0.06 (0.01)
RegEM-TTLS				
RegEM-TTLS	0.15 (0.03)	0.63 (0.18)	-0.12 (0.56)	0.01 (0.02)

than the corresponding scores for RegEM-TTLS. Moreover, the standard deviations of the CE scores are significantly smaller for GraphEM. The results thus demonstrate that GraphEM can also be useful for reconstructing indices such as the mean temperature, with better risk properties than RegEM-TTLS.

4.3. *Uncertainty quantification.* Section 4.1 demonstrates the ability of GraphEM to reduce the uncertainties in paleoclimate reconstructions via an ensemble of pseudoproxies. In practice, it is necessary to obtain an estimate of the uncertainties internally [see, e.g., Li, Nychka and Ammann (2010)]. We therefore produce prediction intervals for both RegEM and GraphEM using a nonparametric block bootstrap method [Liu (1988)]. The technique is described in Appendix C, and is illustrated for the global reconstruction of Section 4.1. Using the reconstruction $\hat{X}_1, \ldots, \hat{X}_N$ provided by the nonparametric bootstrap, we estimate a 95% prediction interval for each reconstructed mean by computing the 2.5th and 97.5th percentiles of the empirical distribution. The mean width of the uncertainty bands for GraphEM and RegEM-TTLS are 0.35 and 0.45, respectively. Comparing Figures 9 and 10, we observe that the uncertainties of GraphEM seem slightly overestimated, whereas the uncertainties of RegEM-TTLS seem underestimated by the bootstrap.

The coverage rates over the validation period for GraphEM and RegEM-TTLS are, respectively, 92.3% and 91.4%. The coverage rates of our method thus appear reasonable. Two natural techniques can be used if a given coverage rate needs to be obtained: (1) modify the band width to obtain the right coverage, or (2) inflate the variance of the reconstructed values in the bootstrap [see Janson and Rajaratnam
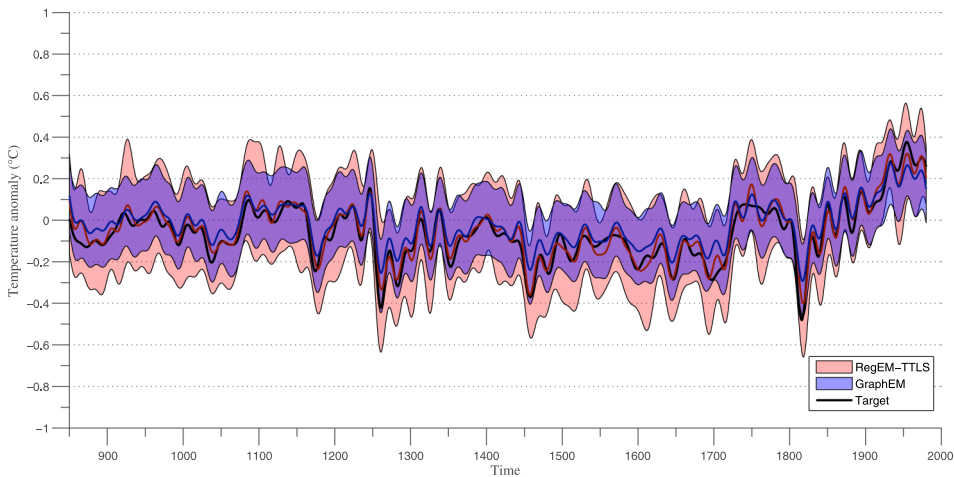

FIG. 10.   *Inflated spatial average uncertainty estimated by nonparametric bootstrap, blocksize = 2 (neighborhood graph, cutoff radius = 800 km). The thick lines represent the median of each ensemble.*

(2014), Li, Nychka and Ammann (2010) for details]. Recall that in our reconstructions, the instrumental period is also reconstructed using the pseudoproxies. The reconstructed values over the instrumental period can thus provide guidance about how much to inflate the uncertainty bands to obtain a given coverage rate. In Figure 10, the coverage rates over the instrumental period for GraphEM and RegEM-TTLS are 91.3% and 90.7%, respectively. In order to obtain a coverage rate of, say, 95% over the instrumental period, the GraphEM and RegEM-TTLS bands must be inflated by a factor of 1.15 and 1.42, respectively. Inflating the bands by these factors yields coverage rates of 94.2% and 97.2% on the validation period, respectively. Inflation factors can also be computed in a more principled way by using k-fold cross-validation over the instrumental period. In our simulations, we split the instrumental period into 5 blocks and used the bootstrap to reconstruct each block using the other 4 blocks. In each case, an inflation factor can be computed so that the uncertainty bands cover 95% of the targeted mean over the instrumental period. Using this technique, we obtained an average inflation factor of 1.10 with GraphEM (similar to the inflation factor obtained without cross-validation).

5. Characteristics of paleoclimate Markov random fields. Our results demonstrate that the GraphEM approach produces substantial improvements in comparison to RegEM-TTLS almost uniformly over space. This section examines the characteristics of paleoclimatic Markov random fields. More precisely, we study the properties of the joint temperature/proxy graph estimated using the ℓ_1 method, with the goal of understanding (a) whether the GraphEM approach is indeed achieving its original aim of parameter reduction, and (b) what are the important features of estimated temperature/proxy fields. In particular, we examine the difference between the graphical structures estimated from the data using the ℓ_1 method (Section 2.1) and the neighborhood structures described in Section 2.2.

We first illustrate the achieved parameter reduction when the graph is estimated from the data. Figure 11 displays the distribution of the average distance from
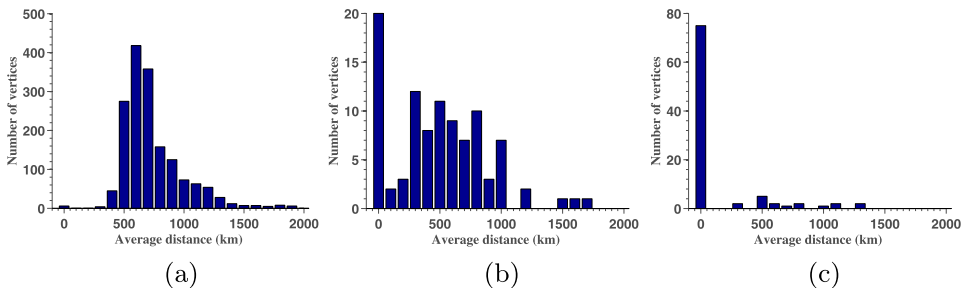

FIG. 11. *Distribution of the average distance to each vertex in the different part of the graph (sparsity level = 0.5%, SNR = 0.5). (a) TT, (b) TP, (c) PP.*

each vertex to its neighbors in the TT, TP and PP part of the temperature/proxy graph estimated with the ℓ_1 method with a sparsity level of 0.5% in each part of the precision matrix. We observe that each point is generally only connected to geographically close neighbors, although the graph can display some far away connections (which may or may not represent geophysical relations). The average number of neighbors in the TT, TP and PP parts of the graph are 10.5, 9.4 and 0.42, respectively. The graph therefore displays a neighborhood structure in the TT and TP part of the graph, with a cutoff radius of roughly 800 km. Note also that the absence of connections in the PP part of the graph suggests that proxies are conditionally independent given the temperature data, and that the estimated graph is very similar to the 2 families of graphs described at the end of Section 2.2. The main message is that the number of neighbors is relatively few compared to one that would be present with a full precision matrix, and that locality seems to be an important characteristic of paleoclimate Markov random fields.

Figure 12 displays the distribution of the number of temperature neighbors of each proxy (in the graph estimated with the ℓ_1 method) when no noise has been added to the temperature time series when generating pseudoproxies (SNR $= \infty$), as compared to the typical noise case that has been studied thus far (SNR $= 0.5$). Both graphs have been obtained using the ℓ_1 method with a sparsity level of 0.5%. This comparison shows that many proxies do not have any temperature neighbors in the SNR $= 0.5$ case. In comparison, a relation between each proxy and some temperature locations has been detected in the SNR $= \infty$ case. Detecting temperature/proxy relations from the data can thus be an issue when the level of noise is high. The potential for the ℓ_1 method to detect spurious relations in the presence of noise is also to be expected [Banerjee, El Ghaoui and d'Aspremont (2008)]. This problem may be mitigated by adding further constraints on the estimated graph. Neighborhood graphs offer a natural solution and provide a good graphical structure independently of the level of noise.

We now examine how sparsity translates to improvements in paleoclimate reconstructions. Figure 13 displays the improvements given by GraphEM (as com-
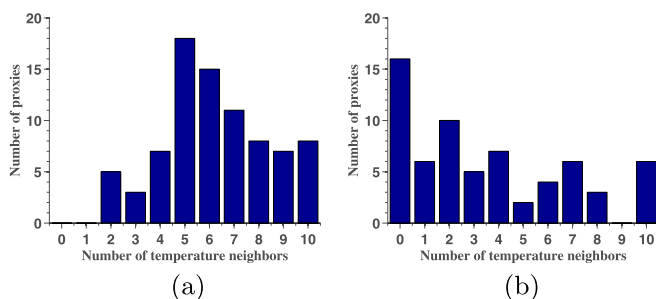

FIG. 12.   *Distribution of the number of temperature neighbors of proxy points* (*sparsity level* $= 0.5$). (a) *SNR* $= \infty$, (b) *SNR* $= 0.5$.
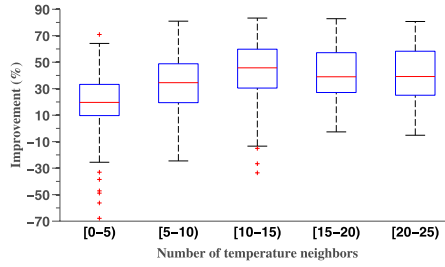
FIG. 13.   *Boxplot of the % improvement as a function of the number of temperature neighbors* (*sparsity level* $= 0.5\%$, *SNR* $= 0.5$).

pared to RegEM-TTLS) at different temperature points vs. their connectivity (number of temperature neighbors) in the corresponding graph.

The figure indicates that percentage improvement is smaller for temperature points with very few neighbors. The improvement is maximal when the number of neighbors roughly corresponds to the number of immediate geographical neighbors of the vertex. Once again, this demonstrates the importance of locality in paleoclimate Markov random field structures. We note, however, that large improvements are still recorded at locations with a larger number of neighbors. The larger neighborhoods may represent real geophysical structures, in which case the reconstruction may benefit from the flexibility of the model. These edges may also be spurious. We note, however, that the presence of an edge (i, j) in the graph simply does not force the corresponding entry ω_{ij} in the precision matrix to be zero. When Ω is estimated in accordance with the graph, ω_{ij} can still be very small. Large improvements are therefore possible when spurious edges are present in the graph.

6. Concluding remarks.   The main objective of the paper was to explore the efficacy of recent advances in the theory of graphical models and high-dimensional inference for statistical paleoclimate reconstructions. Markov random fields provide a sparse representation of the precision matrix of spatial fields, and thus achieve the dimension reduction that is often necessary in high-dimensional settings.

We explored two families of methods to estimate the graphical structure of climate fields: a neighborhood approach and a ℓ_1 penalized inverse covariance estimation approach. In neighborhood graphs, each vertex is only connected to its immediate neighbors, reflecting the fact that variables at two locations are expected to be independent given the temperature in a geographical neighborhood. The size of the neighborhoods can be chosen from the data by cross-validation. The ℓ_1 method, in contrast, provides more flexibility to represent the spatial heterogeneities of geophysical fields (e.g., land/ocean contrasts, topographical boundaries, teleconnection patterns), which would in general be difficult with parametric

(e.g., Matérn family) covariance functions. The GraphEM algorithm was subsequently tested on pseudoproxy data. We also proposed a block bootstrap method to internally estimate the uncertainties in the reconstructions performed using GraphEM and RegEM-TTLS.

Our experiments show that the GraphEM approach gives consistently better reconstructions than the frequently used RegEM-TTLS [see, e.g., Mann et al. (2008, 2009)] almost uniformly over space. We show that Gaussian Markov random fields yield demonstrably improved estimates of the underlying spatio-temporal process, which we tied to the sparsity of the estimated covariance model. A caveat of the ℓ_1 method is the tendency to sometimes detect spurious edges in the graph, that is, to detect relationships that arise from the presence of noise, instead of physical links between the temperature field and the proxies (or pseudoproxies) that derive from it. This is to be expected due to the signal to noise relationship in the data, and is inherent in all statistical and signal processing recovery techniques. Further constraints on the graph can naturally be added to ensure that the graphs selected by the *graphical lasso* retain a high degree of locality. In contrast, neighborhood graphs seem to provide an adequate approximation to the conditional structure of the temperature/proxy field, independently of the level of noise present in the data. The size of the neighborhoods can also be chosen from the data so as to minimize the prediction error. As we demonstrate in our simulations, neighborhood graphs perform well and can be used in situations where there is less hope of discovering the graphical structure of the field from the data. We also observed that most locations in the graphs estimated using the ℓ_1 method are connected to geographically close locations. Locality is therefore an important feature in paleoclimate graphs. We also note that most proxies have no proxy neighbors in graphs estimated from the data, suggesting that proxies are independent of each other given the temperature data.

Finally, and although we were primarily motivated by paleoclimate applications and the use of the EM algorithm in this context, it is worth pointing out that graphical models are also applicable within Bayesian CFR methods [e.g., Tingley and Huybers (2010b, 2010a)] and beyond the confines of climate science. GraphEM as described here provides a useful addition to the RegEM framework, one that will be applicable to any high-dimensional imputation problem, and one that can be used in tandem with other ℓ_2 regularization approaches, especially data-adaptive ones. Future work will extend the use of Gaussian Markov random fields as process models for geophysical fields, in tandem with hierarchical models.

APPENDIX A: DESCRIPTION OF THE GRAPHEM ALGORITHM

Algorithm 1 The graphical EM algorithm (GraphEM)

Input: Incomplete $n \times p$ matrix X, graph G.

1: Initialize $X^{(0)}$ by replacing the missing values in X by the sample mean of each variable over the instrumental period;

2: Compute initial estimates $\mu^{(0)}$ and $\Sigma^{(0)}$ of μ and Σ by computing the sample mean and sample covariance of $X^{(0)}$;

3: Initialize $i \leftarrow 0$;

4: Initialize $\Sigma_G^{(0)} = \Sigma^{(0)}$;

5: **repeat**

6:     Compute $X^{(i+1)}$ by performing a linear regression of the missing values on the available ones for each row of X, using the current estimate $\mu^{(i)}$ of μ and the current graphical estimate $\Sigma_G^{(i)}$ of Σ [see (1.1)];

7:     Compute $\mu^{(i+1)}$ by computing the sample mean of $X^{(i+1)}$;

8:     Compute $\Sigma^{(i+1)}$ as in (1.3);

9:     Compute the new graphical estimate $\Sigma_G^{(i+1)}$ by solving (1.6) with $S = \Sigma^{(i+1)}$, that is,

(A.1)
$$\Sigma_G^{(i+1)} = \operatorname*{argmax}_{\substack{\Sigma=\Omega^{-1}>0 \\ \Omega_{ij}=0,(i,j)\notin E}} \log\det\Omega - \operatorname{tr}(\Sigma^{(i+1)}\Omega);$$

10:    $i \leftarrow i + 1$;

11: **until** convergence

Output: Completed matrix $\hat{X}$, estimate $\hat{\mu}$ of μ, estimate $\hat{\Sigma}$ of Σ.

APPENDIX B: DERIVATION OF THE GRAPH-EM ALGORITHM

We follow the notation in Little and Rubin [Little and Rubin (2002)]. The complete data belongs to a regular exponential family given by a Gaussian Markov random field with graph $G = (V, E)$ (as compared to a complete model in the classical EM algorithm). The sufficient statistics are given by

(B.1)
$$S = \left(\sum_{i=1}^{n} y_{ij}, j = 1, \ldots, k; \sum_{i=1}^{n} y_{ij} y_{ik} \text{ with } (j, k) \in E \right).$$

Let $\theta^{(t)} = (\mu^{(t)}, \Sigma^{(t)})$ denote the current estimate of the parameters. The E-step is given as follows:

(B.2)
$$E\left[\sum_{i=1}^{n} y_{ij} \,\middle|\, Y_{\text{obs}}, \theta^{(t)} \right] = \sum_{i=1}^{n} y_{ij}^{(t)}, \qquad j = 1, \ldots, k$$

and

$$
\text{(B.3)} \qquad E\left[\sum_{i=1}^{n} y_{ij} y_{ik} \,\middle|\, Y_{\text{obs}}, \theta^{(t)}\right] = \sum_{i=1}^{n}(y_{ij}^{(t)} y_{ik}^{(t)} + c_{jki}^{(t)}), \qquad (j,k) \in E
$$

with

$$
\text{(B.4)} \qquad y_{ij}^{(t)} = \begin{cases} y_{ij}, & \text{when } y_{ij} \text{ is observed}, \\ E[y_{ij}|y_{\text{obs},i}, \theta^{(t)}], & \text{when } y_{ij} \text{ is missing} \end{cases}
$$

and

$$
\text{(B.5)} \qquad c_{jki}^{(t)} = \begin{cases} 0, & \text{if at least one of the } y_{ij} \text{ or } y_{ik} \text{ is observed}, \\ 0, & \text{if } j \perp_G k|\text{obs}, i, \\ \text{Cov}[y_{ij}, y_{ik}|y_{\text{obs},i}, \theta^{(t)}], & \\ & \text{if both } y_{ij} \text{ and } y_{ik} \text{ are missing and } j \not\perp_G k|\text{obs}, i, \end{cases}
$$

where $j \perp_G k$ means that j and k are separated in the graph G [see, e.g., Lauritzen (1996), Example 3.2]. At a first glance, it would appear as if there is little difference between the treatment in the graphical vs. the complete case. A closer look reveals that there are some notable differences, the first being in the calculation of the sufficient statistics. Second, note that the definition of $y_{ij}^{(t)}$ and $c_{jki}^{(t)}$ below are different: $y_{ij}^{(t)}$ when y_{ij} is missing is given as follows:

$$
\text{(B.6)} \qquad \begin{aligned} &E[y_{ij}|y_{\text{obs},i}, \theta^{(t)}] \\ &\quad = \mu_j^{(t)} + (\Sigma_{j,\text{obs}}^{G})^{(t)}[(\Sigma_{\text{obs},\text{obs}}^{G})^{(t)}]^{-1}(y_{\text{obs},i} - \mu_{\text{obs}}^{(t)}), \end{aligned}
$$

where $(\Sigma^G)^{(t)}$ corresponds to a graphical covariance matrix Σ. When both y_{ij} and y_{ik} are missing and $(j,k) \in E$,

$$
\text{(B.7)} \qquad \begin{aligned} &\text{Cov}[y_{ij} y_{ik}|y_{\text{obs},i}, \theta^{(t)}] \\ &\quad = (\Sigma_{jk}^{G})^{(t)} - (\Sigma_{\{j,k\},\text{obs}}^{G})^{(t)}[(\Sigma_{\text{obs},\text{obs}}^{G})^{(t)}]^{-1}(\Sigma_{\text{obs},\{j,k\}}^{G})^{(t)}. \end{aligned}
$$

Note, however, that $\text{Cov}[y_{ij} y_{ik}|y_{\text{obs},i}, \theta^{(t)}] = \Sigma_{jk|\text{obs},i}^{(t)}$. Thus,

$$
\text{(B.8)} \qquad c_{jki}^{(t)} = \Sigma_{jk|\text{obs},i}^{(t)} = 0 \qquad \text{if } j \perp_G k|\text{obs}, i.
$$

The M-step in the GraphEM algorithm therefore consists of using the sufficient statistics for the complete data derived in (B.2) and (B.3) to determine the graphical mle. In particular, the estimate of the mean parameter is given by the sample mean and the estimate of the graphical covariance is given in equation (A.1).

APPENDIX C: NONPARAMETRIC BOOTSTRAP

Algorithm 2 RegEM/GraphEM Uncertainty quantification (nonparametric bootstrap)

Input: Incomplete $n \times p$ matrix X containing n_i years of instrumental data, number of bootstrap samples $N > 1$, blocksize b.

1: **for** $i = 1, \ldots, N$ **do**
2:    Construct a bootstrap sample $X_{\text{boot},i}$ by sampling with replacement $\lceil n_i/b \rceil$ blocks of size b from the lines of X in the instrumental period, and $\lceil (n - n_i)/b \rceil$ blocks of size b from lines in the rest of the matrix;
3:    Reconstruct the missing values in $X_{\text{boot},i}$ using RegEM/GraphEM. The algorithm outputs estimates $\mu_{\text{boot},i}$, $\Sigma_{\text{boot},i}$ of the mean and covariance matrix of the field;
4:    Obtain $\hat{X}_i$ by reconstructing the missing values in X by performing the regression step of RegEM/GraphEM starting with $\mu_{\text{boot},i}$ and $\Sigma_{\text{boot},i}$;
5:    For each line (x_m, x_a) of $\hat{X}_i$ that originally contained missing values, add a noise realization to x_m from the conditional distribution of $x_m|x_a$, where we assume $(x_m, x_a) \sim N(\mu_{\text{boot},i}, \Sigma_{\text{boot},i})$.
6: **end for**

Output: Ensemble of N reconstructions $\hat{X}_1, \ldots, \hat{X}_N$ of the incomplete field X.

REFERENCES

AMMANN, C. M., JOOS, F., SCHIMEL, D. S., OTTO-BLIESNER, B. L. and TOMAS, R. A. (2007). Solar influence on climate during the past millennium: Results from transient simulations with the NCAR climate system model. *Proc. Natl. Acad. Sci. USA* **104** 3713–3718.

ANNAN, J. D. and HARGREAVES, J. C. (2012). Identification of climatic state with limited proxy data. *Clim. Past* **8** 1141–1151.

BANERJEE, O., EL GHAOUI, L. and D'ASPREMONT, A. (2008). Model selection through sparse maximum likelihood estimation for multivariate Gaussian or binary data. *J. Mach. Learn. Res.* **9** 485–516. MR2417243

BESAG, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *J. Roy. Statist. Soc. Ser. B* **36** 192–236. MR0373208

BRADLEY, R. S. (1996). Are there optimum sites for global paleotemperature reconstruction? In *Climate Variations and Forcing Mechanisms of the Last* 2000 *Years. NATO ASI* **41** 603–624. Springer, Berlin.

BROHAN, P., KENNEDY, J. J., HARRIS, I., TETT, S. F. B. and JONES, P. D. (2006). Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850. *Journal of Geophysical Research*: *Atmospheres* (1984–2012) **111** D12106, 1–21.

BÜRGER, G. (2007). On the verification of climate reconstructions. *Clim. Past* **3** 397–409.

CARROLL, R. J. and RUPPERT, D. (1996). The use and misuse of orthogonal regression in linear errors-in-variables models. *Amer. Statist.* **50** 1–6.

CHRISTIANSEN, B. (2010). Reconstructing the NH mean temperature: Can underestimation of trends and variability be avoided? *J. Climate* **24** 674–692.

CHRISTIANSEN, B. (2013). Straight line fitting and predictions: On a marginal likelihood approach to linear regression and errors-in-variables models. *J. Climate* **27** 2014–2031.

CHRISTIANSEN, B., SCHMITH, T. and THEJLL, P. (2009). A surrogate ensemble study of climate reconstruction methods: Stochasticity and robustness. *J. Climate* **22** 951–976.

COOK, E. R., BRIFFA, K. R. and JONES, P. D. (1994). Spatial regression methods in dendroclimatology: A review and comparison of two techniques. *Intern. J. of Clim.* **14** 379–402.

COOK, E. R., MEKO, D. M., STAHLE, D. W. and CLEAVELAND, M. K. (1999). Drought reconstructions for the continental United States. *J. Climate* **12** 1145–1162.

DAWID, A. P. and LAURITZEN, S. L. (1993). Hyper-Markov laws in the statistical analysis of decomposable graphical models. *Ann. Statist.* **21** 1272–1317. MR1241267

DEMPSTER, A. P., LAIRD, N. M. and RUBIN, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc. Ser. B* **39** 1–38. With discussion. MR0501537

EMILE-GEAY, J., COBB, K. M., MANN, M. E. and WITTENBERG, A. T. (2013a). Estimating central equatorial pacific SST variability over the past millennium. Part 1: Methodology and validation. *J. Climate* **26** 2302–2328.

EMILE-GEAY, J., COBB, K. M., MANN, M. E. and WITTENBERG, A. T. (2013b). Estimating central equatorial pacific SST variability over the past millennium. Part 2: Reconstructions and implications. *J. Climate* **26** 2329–2352.

FIERRO, R. D., GOLUB, G. H., HANSEN, P. C. and O'LEARY, D. P. (1997). Regularization by truncated total least squares. *SIAM J. Sci. Comput.* **18** 1223–1241. MR1453566

FRIEDMAN, J., HASTIE, T. and TIBSHIRANI, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostat.* **9** 432–441.

FROST, C. and THOMPSON, S. G. (2000). Correcting for regression dilution bias: Comparison of methods for a single predictor variable. *J. Roy. Statist. Soc. Ser. A* **163** 173–189.

GOLUB, G. H. and VAN LOAN, C. F. (1980). An analysis of the total least squares problem. *SIAM J. Numer. Anal.* **17** 883–893. MR0595451

GUILLOT, D., RAJARATNAM, B., ROLFS, B., WONG, I. and MALEKI, A. (2012). Iterative thresholding algorithm for sparse inverse covariance estimation. In *Advances in Neural Information Processing Systems* **25** 1583–1591. Curran Associates, Red Hook.

HANKE, M. and HANSEN, P. C. (1993). Regularization methods for large-scale problems. *Surveys Math. Indust.* **3** 253–315. MR1253521

HASTIE, T., TIBSHIRANI, R. and FRIEDMAN, J. (2008). *The Elements of Statistical Learning*: *Data Mining*, *Inference*, *and Prediction*, 2nd ed. Springer, New York.

HOERL, A. E. and KENNARD, R. W. (1970a). Ridge regression: Biased estimation for non-orthogonal problems. *Technometrics* **12** 55–67.

HOERL, A. E. and KENNARD, R. W. (1970b). Ridge regression: Applications to non-orthogonal problems. *Technometrics* **12** 69–82.

HSIEH, C.-J., SUSTIK, M. A., DHILLON, I. S. and RAVIKUMAR, P. K. (2011). Sparse inverse covariance matrix estimation using quadratic approximation. In *Advances in Neural Information Processing Systems* 24 (J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira and K. Q. Weinberger, eds.) 2330–2338. Curran Associates, Red Hook.

JANSON, L. and RAJARATNAM, B. (2014). A methodology for robust multiproxy paleoclimate reconstructions and modeling of temperature conditional quantiles. *J. Amer. Statist. Assoc.* **109** 63–77. MR3180547

JONES, P. D., BRIFFA, K. R., OSBORN, T. J., LOUGH, J. M., VAN OMMEN, T. D., VINTHER, B. M., LUTERBACHER, J., WAHL, E. R., ZWIERS, F. W., MANN, M. E.,

SCHMIDT, G. A., AMMANN, C. M., BUCKLEY, B. M., COBB, K. M., ESPER, J., GOOSSE, H., GRAHAM, N., JANSEN, E., KIEFER, T., KULL, C., KUTTEL, M., MOSLEY-THOMPSON, E., OVERPECK, J. T., RIEDWYL, N., SCHULZ, M., TUDHOPE, A. W., VILLALBA, R., WANNER, H., WOLFF, E. and XOPLAKI, E. (2009). High-resolution palaeoclimatology of the last millennium: A review of current status and future prospects. *The Holocene* **19** 3–49.

LAURITZEN, S. L. (1996). *Graphical Models*. *Oxford Statistical Science Series* **17**. Oxford Univ. Press, New York. MR1419991

LETAC, G. and MASSAM, H. (2007). Wishart distributions for decomposable graphs. *Ann. Statist.* **35** 1278–1323. MR2341706

LI, B., NYCHKA, D. W. and AMMANN, C. M. (2010). The value of multiproxy reconstruction of past climate. *J. Amer. Statist. Assoc.* **105** 883–895. MR2752583

LI, B. and SMERDON, J. E. (2012). Defining spatial comparison metrics for evaluation of paleoclimatic field reconstructions of the common era. *Environmetrics* **23** 394–406. MR2958920

LIN, S. P. and PERLMAN, M. D. (1985). A Monte Carlo comparison of four estimators of a covariance matrix. In *Multivariate Analysis VI* (*Pittsburgh*, *PA*, 1983) (P. R. Krishnaiah, ed.) 411–429. North-Holland, Amsterdam. MR0822310

LITTLE, R. J. A. and RUBIN, D. B. (2002). *Statistical Analysis with Missing Data*, 2nd ed. Wiley, New York. MR1925014

LIU, R. Y. (1988). Bootstrap procedures under some non-i.i.d. models. *Ann. Statist.* **16** 1696–1708. MR0964947

MANN, M. E., BRADLEY, R. S. and HUGHES, M. K. (1998). Global-scale temperature patterns and climate forcing over the past six centuries. *Nature* **392** 779–787.

MANN, M. E. and RUTHERFORD, S. (2002). Climate reconstruction using 'pseudoproxies.' *Geophys. Res. Lett.* **29** 139-1–139-4.

MANN, M. E., RUTHERFORD, S., WAHL, E. and AMMANN, C. (2005). Testing the fidelity of methods used in proxy-based reconstructions of past climate. *J. Climate* **18** 4097–4107.

MANN, M. E., RUTHERFORD, S., WAHL, E. and AMMANN, C. (2007a). Robustness of proxy-based climate field reconstruction methods. *Journal of Geophysical Research* (*Atmospheres*) **112** D12109.

MANN, M. E., RUTHERFORD, S., WAHL, E. and AMMANN, C. (2007b). Reply to comments on "Testing the fidelity of methods used in proxy-based reconstructions of past climate" by Smerdon and Kaplan. *J. Climate* **20** 5671–5674.

MANN, M. E., ZHANG, Z., HUGHES, M. K., BRADLEY, R. S., MILLER, S. K., RUTHERFORD, S. and NI, F. (2008). Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. *Proc. Natl. Acad. Sci. USA* **105** 13252–13257.

MANN, M. E., ZHANG, Z., RUTHERFORD, S., BRADLEY, R. S., HUGHES, M. K., SHINDELL, D., AMMANN, C., FALUVEGI, G. and NI, F. (2009). Global signatures and dynamical origins of the little ice age and medieval climate anomaly. *Science* **326** 1256–1260.

MASSON-DELMOTTE, V., SCHULZ, M., ABE-OUCHI, A., BEER, J., GANOPOLSKI, A., ROUCO, J. F. G., JANSEN, E., LAMBECK, K., LUTERBACHER, J., NAISH, T., OSBORN, T., OTTO-BLIESNER, B., QUINN, T., RAMESH, R., ROJAS, M., SHAO, X. and TIMMERMANN, A. (2013). Information from Paleoclimate Archives. In *Climate Change* 2013: *The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (T. F. Stocker, D. Qin, G. K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P. M. Midgley, eds.) 383–464. Cambridge Univ. Press, Cambridge.

MCLACHLAN, G. J. and KRISHNAN, T. (2008). *The EM Algorithm and Extensions*, 2nd ed. Wiley, Hoboken, NJ. MR2392878

NATIONAL RESEARCH COUNCIL (2006). *Surface Temperature Reconstructions for the Last* 2000 *Years*. The National Academies Press, Washington, DC.

PAUL, D. (2007). Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statist. Sinica* **17** 1617–1642. MR2399865

RAJARATNAM, B., MASSAM, H. and CARVALHO, C. M. (2008). Flexible covariance estimation in graphical Gaussian models. *Ann. Statist.* **36** 2818–2849. MR2485014

RIEDWYL, N., KÜTTEL, M., LUTERBACHER, J. and WANNER, H. (2009). Comparison of climate field reconstruction techniques: Application to Europe. *Clim. Dyn.* **32** 381–395.

RUTHERFORD, S., MANN, M. E., OSBORN, T. J., BRADLEY, R. S., BRIFFA, K. R., HUGHES, M. K. and JONES, P. D. (2005). Proxy-based northern hemisphere surface temperature reconstructions: Sensitivity to method, predictor network, target season, and target domain. *J. Climate* **18** 2308–2329.

SCHNEIDER, T. (2001). Analysis of incomplete climate data: Estimation of mean values and covariance matrices and imputation of missing values. *J. Climate* **14** 853–871.

SMERDON, J. E. (2011). Climate models as a test bed for climate reconstruction methods: Pseudoproxy experiments. *WIREs Clim Change* **3** 67–77.

SMERDON, J. E. and KAPLAN, A. (2007). Comments on "Testing the fidelity of methods used in proxy-based reconstructions of past climate:" The role of the standardization interval. *J. Climate* **20** 5666–5670.

SMERDON, J. E., KAPLAN, A., CHANG, D. and EVANS, M. N. (2010). A pseudoproxy evaluation of the CCA and RegEM methods for reconstructing climate fields of the last millennium. *J. Climate* **23** 4856–4880.

SMERDON, J. E., KAPLAN, A., ZORITA, E., GONZÁLEZ-ROUCO, J. F. and EVANS, M. N. (2011). Spatial performance of four climate field reconstruction methods targeting the Common Era. *Geophys. Res. Lett.* **38** L11705.

STEIGER, N. J., HAKIM, G. J., STEIG, E. J., BATTISTI, D. S. and ROE, G. H. (2013). Assimilation of time-averaged pseudoproxies for climate reconstruction. *J. Climate* **27** 426–441.

STEIN, C. (1986). Lectures on the theory of estimation of many parameters. *J. Math. Sci.* **34** 1373–1403.

TIBSHIRANI, R. (1996). Regression shrinkage and selection via the lasso. *J. Roy. Statist. Soc. Ser. B* **58** 267–288. MR1379242

TIKHONOV, A. N. and ARSENIN, V. Y. (1977). *Solution of Ill-Posed Problems*. V. H. Winston and Sons, Washington.

TINGLEY, M. P. and HUYBERS, P. (2010a). A Bayesian algorithm for reconstructing climate anomalies in space and time. Part II: Comparison with the regularized expectation-maximization algorithm. *J. Climate* **23** 2782–2800.

TINGLEY, M. P. and HUYBERS, P. (2010b). A Bayesian algorithm for reconstructing climate anomalies in space and time. Part I: Development and applications to paleoclimate reconstruction problems. *J. Climate* **23** 2759–2781.

TINGLEY, M. P. and LI, B. (2012). Comments on "Reconstructing the NH mean temperature: Can underestimation of trends and variability be avoided?" *J. Climate* **25** 3441–3446.

TINGLEY, M. P., CRAIGMILE, P. F., HARAN, M., LI, B., MANNSHARDT, E. and RAJARATNAM, B. (2012). Piecing together the past: Statistical insights into paleoclimatic reconstructions. *Quaternary Science Reviews* **35** 1–22.

VAN HUFFEL, S. and VANDEWALLE, J. (1991). *The Total Least Squares Problem*: *Computational Aspects and Analysis. Frontiers in Applied Mathematics* **9**. SIAM, Philadelphia, PA. MR1118607

VON STORCH, H., ZORITA, E., JONES, J. M., DIMITRIEV, Y., GONZÁLEZ-ROUCO, F. and TETT, S. F. B. (2004). Reconstructing past climate from noisy data. *Science* **306** 679–682.

WANG, J., EMILE-GEAY, J., GUILLOT, D., SMERDON, J. E. and RAJARATNAM, B. (2014). Evaluating climate field reconstruction techniques using improved emulations of real-world conditions. *Climate of the Past* **10** 1–19.

WHITTAKER, J. (1990). *Graphical Models in Applied Multivariate Statistics*. Wiley, Chichester. MR1112133

WITTEN, D. M., FRIEDMAN, J. H. and SIMON, N. (2011). New insights and faster computations
    for the graphical lasso. *J. Comput. Graph. Statist.* **20** 892–900. MR2878953

D. GUILLOT   J. EMILE-GEAY
B. RAJARATNAM   DEPARTMENT OF EARTH SCIENCES
DEPARTMENT OF STATISTICS   AND CENTER FOR APPLIED MATHEMATICAL SCIENCES
STANFORD UNIVERSITY   UNIVERSITY OF SOUTHERN CALIFORNIA
SEQUOIA HALL   3651 TROUSDALE PARKWAY, ZHS 275
390 SERRA MALL   LOS ANGELES, CALIFORNIA 90089-0740
STANFORD, CALIFORNIA 94305-4065   USA
USA   E-MAIL: julieneg@usc.edu
E-MAIL: dguillot@stanford.edu
        brajarat@stanford.edu