

Comments on “Reconstructing the NH Mean Temperature: Can Underestimation of Trends and Variability Be Avoided?”

MARTIN P. TINGLEY*

National Center for Atmospheric Research, Boulder, Colorado

BO LI

Department of Statistics, Purdue University, West Lafayette, Indiana

(Manuscript received 3 January 2011, in final form 26 July 2011)

1. Introduction

In a recent paper, Bo Christiansen presents and discusses “LOC,” a methodology for reconstructing past climate that is based on local regressions between climate proxy time series and instrumental time series (Christiansen 2011, hereafter C11). LOC respects two important scientific facts about proxy data that are often overlooked, namely that many proxies are likely influenced by strictly local temperature, and, to reflect causality, the proxies should be written as functions of climate, not vice versa. There are, however, several weaknesses to the LOC method: uncertainty is not propagated through the multiple stages of the analysis, the effects of observational errors in the instrumental record are not considered, and as the proxies become uninformative of climate, the variance of a reconstruction produced by LOC becomes unbounded—a result that is clearly unphysical. These shortcomings can be overcome by interpreting the LOC method in the context of recently proposed Bayesian hierarchical reconstruction methods.

Section 2 reviews the basic modeling assumptions underlying LOC and details the shortcomings of this approach. To illustrate one possible solution to the shortcomings of LOC, section 3 presents a Bayesian interpretation of LOC and briefly describes the connections

between LOC and two recently published Bayesian reconstruction methods. Section 4 discusses the variance of the reconstructed series in both the original and Bayesian LOC frameworks, and section 5 provides a few concluding remarks.

2. LOC: Description and shortcomings

The key *modeling assumption* subsumed under the LOC heading is

$$P_{ij} = \alpha_i + \lambda_i T_{ij} + \epsilon_{ij}, \quad (1)$$

where P_{ij} and T_{ij} are the proxy and true temperature values, respectively, at location i and time j , and the errors ϵ_{ij} are assumed to be independent and identically distributed (IID) normal for each location, with variance σ_i^2 . In other words, there is a linear relation between each proxy series and the local temperature series, with the latter treated as the independent variable. LOC makes the additional assumption that the instrumental series are free of observational errors.

In terms of *inference tools*, the parameters α_i and λ_i are inferred using ordinary least squares regression over a calibration interval. The resulting maximum likelihood estimates of the parameters are then used in the second stage of the analysis to impute past temperatures via [cf. Eq. (A6) from C11]

$$\hat{T}_{ij} = (P_{ij} - \hat{\alpha}_i) / \hat{\lambda}_i. \quad (2)$$

In the third and final stage of the analysis, the inferred local temperature series are combined into a regional

* Current affiliation: Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts.

Corresponding author address: Martin P. Tingley, Department of Earth and Planetary Sciences, Harvard University, 12 Oxford St., Cambridge, MA 02138.
E-mail: tingley@fas.harvard.edu

or global average by taking a weighted mean of the reconstructed time series, with weights proportional to the cosines of the latitudes of the proxies.

The most obvious shortcoming of the LOC method is the behavior of the local reconstructions in the limiting case of the proxy series being independent of the corresponding true temperature series. As the proxy series becomes less informative of the temperature series, $\hat{\lambda} \rightarrow 0$ and, via Eq. (2), $\text{Var}[\hat{T}_{ij}] \rightarrow \infty$. There is thus a singularity problem at $\hat{\lambda}_i = 0$, which C11 avoids by performing f tests to exclude proxies that are not informative of the corresponding local temperatures. In the case where $0 < |\hat{\lambda}| \ll 1$, prediction using Eq. (2) is well defined from a mathematical standpoint, but the variance of the predictions will still be inflated, and the temperature reconstructions unreliable. Bayesian inference (discussed below) provides a formalism for regularizing the inference and avoiding unphysical variance inflation in the presence of weakly informative predictors.

Another shortcoming of the LOC method concerns the treatment of measurement errors. It is well known that when both the predictor and response variables in a regression are subject to measurement errors, inference on the model parameters is ill defined without additional information on the variance of the error terms (e.g., Fuller 1987). If such information is unavailable, it is possible to establish bounds on model parameters, but precise parameter estimation is not possible. The indirect regression underlying LOC bypasses the estimation of the measurement errors in the proxy and instrumental observations, and the resulting $1/\hat{\lambda}_i$ is in fact simply an estimate of the upper bound of β_i , rather than of β_i itself, in the model $T_{ij} = \gamma_i + \beta_i P_{ij} + \epsilon_{ij}$ with P_{ij} contaminated by errors of unknown variance. This upper bound corresponds to the case when the P_{ij} are observed subject to measurement errors, but the observations of T_{ij} are free of such errors (Fuller 1987, chapter 1.1.3). If the T_{ij} also contain measurement errors, then the magnitude of $1/\hat{\lambda}_i$ estimated from Eq. (1) will be biased high, and the variance of the reconstruction artificially inflated. Given that the instrumental temperature observations are typically contaminated with substantial noise (Brohan et al. 2006), inverse regression must be used cautiously to avoid overcorrecting the attenuation effect in the estimate of $\hat{\lambda}_i$ caused by measurement errors in T_{ij} . These possible biases can be mitigated by specifying a model that explicitly takes into account measurement errors in both P_{ij} and T_{ij} , or by specifying a parametric model for the target process within a hierarchical model; both of these possibilities are discussed below.

The methodology used to combine the local reconstructions into a regional or hemispheric average

is likewise unsatisfying. The fact that certain proxy series may have weaker relationships with local temperatures than others, and thus that the inferred temperature at those locations are less reliable, is not taken into account. In addition, the spatial covariance of the underlying temperature process is not exploited (cf. Tingley and Huybers 2010a). These two issues can both be resolved via hierarchical modeling, as discussed below. Given that C11 does not consider the spatial covariance between the local temperature reconstructions, and that the goal of the analysis is to infer Northern Hemisphere (NH) mean temperature, the advantage of first reconstructing local temperatures is unclear. As an alternative, C11 could use each proxy series to directly reconstruct the hemispheric average temperature, and then take the average over these reconstructions (cf. Li et al. 2010).

A final concern with LOC is that the uncertainty introduced at the various stages of the analysis, such as in the estimation of $\hat{\alpha}_i$ and $\hat{\lambda}_i$, is not propagated through to the final estimate of the NH mean temperature. Indeed, C11 provides no estimates of the uncertainty in the reconstructed temperature series (regional or local) and the word ‘‘uncertainty’’ does not appear anywhere in the text. Given the increased focus on uncertainty quantification throughout the climate sciences (e.g., Solomon et al. 2007), a method that does not produce internally consistent uncertainty estimates is not satisfying. Bayesian inference (discussed below) is the obvious tool to rectify this lack of uncertainty quantification.

3. LOC in a Bayesian framework

Many of the weaknesses in the LOC method can be avoided by placing the key underlying assumptions into a Bayesian framework. Indeed, the indirect regression that forms the basis of the LOC method [Eq. (1)] has the form of a linear data-level model in a Bayesian hierarchical model. From a Bayesian perspective, the local temperature prediction, conditional on the proxies and model parameters, takes the following form:

$$[T_{ij} | P_{ij}, \lambda_i, \alpha_i, \sigma_i^2] \propto [P_{ij} | T_{ij}, \lambda_i, \alpha_i, \sigma_i^2][T_{ij}] \\ P_{ij} | T_{ij}, \lambda_i, \alpha_i, \sigma_i^2 \sim N(\alpha_i + \lambda_i T_{ij}, \sigma_i^2), \quad (3)$$

where $[A]$ denotes the probability density function (PDF) of the random variable A and $[A | B]$ denote the distribution of the random variable A conditional on the random variable B . Note that the distribution for $P_{ij} | T_{ij}, \lambda_i, \alpha_i, \sigma_i^2$ is equivalent to the form of Eq. (1): the indirect regression is the natural linear data-level model within the Bayesian framework. As discussed below, even

a weakly informative prior for $[T_{ij}]$ regularizes the inference and thus avoids the problem of unbounded variance in the presence of uninformative proxies. It is worth emphasizing that Bayesian models treat temperature as a random variable rather than an unknown, but fixed, quantity.

Estimation of model parameters in LOC is conducted using collocated instrumental and proxy observations during a calibration interval. The calibration stage can be expressed as a Bayesian regression:

$$[\lambda_i, \alpha_i, \sigma_i^2 | P_i, T_i] \propto [P_i | T_i, \lambda_i, \alpha_i, \sigma_i^2][\lambda_i, \alpha_i, \sigma_i^2], \quad (4)$$

where a subscripted dot refers to all time points in the calibration interval. The likelihood, $[P_i | T_i, \lambda_i, \alpha_i, \sigma_i^2]$, is a product of normals of the form in Eq. (3), while the prior for the model parameters, $[\lambda_i, \alpha_i, \sigma_i^2]$, may be more or less informative. See standard texts (e.g., Gelman et al. 2003) for details on Bayesian regression models.

The key advantage of putting LOC into Bayesian framework is that a draw from the posterior of the parameters in Eq. (4) can be plugged into Eq. (3), which results in a draw of the temperature at each time and location for which there is no instrumental observation. The end result is an estimate of the *posterior distribution* of the temperature at these times and locations. Likewise, draws from various locations can be averaged at each year to produce a posterior distribution of the global or regional mean time series. Posterior draws of the local temperature time series are less variable at locations where the corresponding proxy series have strong relationships with local temperature (small σ_i^2). As a result, the mean of the posterior distribution of the global or regional time series is more influenced (regardless of the weights used to calculate the regional mean) by those proxies having strong relationships with local temperature. The Bayesian inference scheme provides a framework for propagating uncertainty through the analysis in an internally consistent manner, as uncertainty in the estimation of regression parameters is included in the final estimate of uncertainty in the global or regional mean time series.

In contrast, uncertainty introduced at numerous stages of the LOC method is not propagated, and the resulting reconstructions (see Figs. 3, 5, 7, 9, 10, and 11 from C11) include no estimates of uncertainty. As the calibration model [Eq. (1)] that underpins LOC attributes all errors to the proxies and treats the target climate as a fixed quantity, estimates of residual variance after fitting the model in Eq. (1) cannot be readily used to estimate the uncertainty in predictions of T . In

contrast, methods based on direct regression, such as the various implementations of the regularized EM algorithm (RegEM; Schneider 2001), attribute the errors to the target climate quantity, and uncertainty estimates can then be obtained from the residual analysis.

The Bayesian interpretation of LOC is closely related to two other published methods for reconstructing past climate.¹ Li et al. (2010) describe a method for reconstructing the NH mean temperature time series using proxy time series that reflect temperatures at different temporal scales, and estimates of solar, volcanic, and greenhouse gas forcings. Forward models for each type of proxy link the true temperatures to the proxy observations, and a second-order autoregressive [AR(2)] model is assumed for the errors. The indirect regression model in LOC is essentially a special case of the data models used in Li et al. (2010). In contrast to LOC, the Bayesian model of Li et al. (2010) propagates uncertainty introduced at each level of the hierarchy and quantifies the uncertainty in the final estimate of the NH mean temperature time series. In addition, the process-level model used in Li et al. (2010), as well as the vague priors on parameters, provide regularization when the proxy signal-to-noise ratio is low.

Tingley and Huybers (2010a) propose a Bayesian Algorithm for Reconstructing Climate Anomalies in Space and Time (BARCAST), which is, like LOC, based on the data-level assumption that each proxy observation reflects strictly local (in space and time) information about the target process. Each type of proxy observation is assumed to share a common, linear relation of the form in Eq. (1) with the target climate field, while the instrumental observations are modeled as the true underlying value of the field perturbed by additive white noise with constant variance.

Temporally averaged climate processes, such as annual mean surface temperature anomalies, generally display persistence in both space and time, and one of the key features that differentiates BARCAST from LOC is the inclusion of a parametric space–time model. At the process level, the target field is assumed to be a multivariate AR(1) process, with a common AR(1) parameter for all locations. Spatial structure enters through the innovations that drive the AR(1) process, which are assumed to be draws from a mean-zero multivariate normal distribution, with an exponential spatial covariance function (see, e.g., Banerjee et al. 2004). This process-level model allows information to be shared between locations and through time, and the temperature inferred for a given

¹ For more detail on the links between various reconstruction methods see Tingley et al. (2012).

location and year reflects information from neighboring times and places. In particular, explicitly modeling the space–time covariance structure allows for imputation of the temperature field at locations where there are no observations. Inference can then be made on a regular latitude–longitude grid, in which case weighting the inferred temperatures by the cosine of latitude, as done in C11, is a defensible spatial averaging procedure. Finally, Bayesian inference insures that uncertainty in the estimation of model parameters is included in the final estimate of uncertainty in the spatial mean time series.

An important feature of Bayesian hierarchical models is the simultaneous modeling of measurement errors in both the proxy and the instrumental observations (Tingley and Huybers 2010a,b; Li et al. 2010). The process level of a hierarchical model specifies the statistical structure of the target quantity, be it temporal (Li et al. 2010) or spatiotemporal (Tingley and Huybers 2010a), while the data level describes the errors in each type of observation of the target field. In contrast, LOC only models the errors in the proxies [Eq. (1)]. A measurement error model that explicitly takes into account the errors in the observations of both P_{ij} and T_{ij} seems more appropriate, as measurement error models allow for the estimation of the true relationship between the temperature process and the proxies. While appendix B of C11 notes that such models suffers from identifiability issues without additional information or constraints on the parameters,² Ammann et al. (2010) construct an identifiable measurement error model by imposing the constraint of minimizing the bias of the reconstruction. Other methods that account for measurement errors in both predictor and response variables are discussed in Hegerl et al. (2007) and Mann et al. (2007).

4. Variance of the reconstruction

C11 argues that one key advantage of LOC is that the variability of the reconstruction accurately reflects the variability of the target process. While C11 states and illustrates that the variability of LOC reconstructions is larger than that for other methods (see in particular Fig. 10 of C11), no validation results are presented that speak to the method’s ability to capture the *true* variability of the inferred mean time series. Given Eq. (A6) from C11, LOC will drastically overestimate the variability if a proxy series has a weak relationship with local temperature (λ is close to zero) or the instrumental observations are contaminated by a substantial amount of noise. In addition, C11 does not investigate what must

be an inevitable trade-off between minimizing bias and preserving variance, and doing so would lend insight into the properties of the resulting reconstructions. See Tingley and Huybers (2010b) for a comparison of the variability in various (pseudoproxy based) reconstructions to the target variability.

There is strong prior scientific knowledge about the plausible amplitude of temperature variability; for example, Jansen et al. (2007) estimate that, during the Last Glacial Maximum, the mean global surface temperature was about 4°–7°C colder than at present. Using such scientific knowledge to specify a weakly informative prior for the unknown temperature T_{ij} regularizes the inference and avoids the problem of unbounded variance in the presence of noninformative proxies and error-prone instrumental temperatures.

Comparing the predicted temperatures from LOC to those from a similar Bayesian model demonstrate the effects of weak prior information in the Bayesian context. For simplicity, we assume that the intercept parameters [α_i from Eq. (1)] are zero, and we condition on the true values of β_i and σ_i^2 . We consider a single proxy series at a single time point, and drop the subscripts i and j ; we can then write the data-level model as

$$P | T, \lambda, \sigma_e^2 \sim N(\lambda T, \sigma_e^2), \quad (5)$$

where the subscript e indicates that σ_e^2 corresponds to the data-level error. We specify a mean-zero normal prior for the unknown temperature, $T \sim N(0, \sigma_T^2)$. Following Eq. (3), the posterior distribution is then

$$T | P, \lambda, \sigma_e^2 \sim N \left[\frac{\lambda P}{\lambda^2 + \sigma_e^2/\sigma_T^2}, \left(\frac{\lambda^2}{\sigma_e^2} + \frac{1}{\sigma_T^2} \right)^{-1} \right]. \quad (6)$$

In the limit as the prior becomes uninformative (i.e., $\sigma_T^2 \rightarrow \infty$), the posterior converges to $N(P/\lambda, \sigma_e^2/\lambda^2)$. If the prior is uninformative, then in the limit as the regression coefficient $\lambda \rightarrow 0$, the variance becomes unbounded. In the case of finite but large σ_T^2 , which corresponds to a weakly informative prior distribution on T , the variance remains bounded even as $\lambda \rightarrow 0$. Indeed, if the proxies contain no information, then the posterior is equivalent to the prior—which is the only source of information.

Now suppose we estimate T using the posterior mean, $\hat{T} = \lambda P / (\lambda^2 + \sigma_e^2/\sigma_T^2)$, and substitute for P with the data-level relation $P = \lambda T + \sigma_e z$, where z is a draw from $Z \sim N(0, 1)$. We then have

$$\hat{T} = \frac{\lambda^2 T}{\lambda^2 + \sigma_e^2/\sigma_T^2} + \frac{\lambda \sigma_e}{\lambda^2 + \sigma_e^2/\sigma_T^2} z. \quad (7)$$

² For a more thorough discussion of the identifiability issues brought up in appendix B of C11, see Fuller (1987, chapter 1.1.3).

In the case of an uninformative prior ($\sigma_t^2 \rightarrow \infty$), Eq. (7) reduces to Eq. (A6) in C11. If the error structure is white (as is assumed here), then all frequencies of variability are exaggerated, especially as $\lambda \rightarrow 0$. In contrast to LOC, if the prior variance σ_t^2 is finite, then the expected value \hat{T} remains bounded even as $\lambda \rightarrow 0$. Equation (7) indicates that the predicted temperature is the sum of a term proportional to the true temperature T and an error term. As the scaling factor for T is less than one, the amplitude of the signal in the reconstruction is biased toward zero, which is a direct result of the prior mean for T being set to zero. However, the informative prior stabilizes the error component of the predicted temperature, and prevents both the expected value [Eq. (7)] and the variance thereabout [Eq. (6)] from growing unbounded as $\lambda \rightarrow 0$.

Finally, we emphasize that within the hierarchical framework, it is possible to incorporate prior information about the space–time behavior of the target process. Li et al. (2010) incorporate time series of climate forcings and specify a parametric temporal process for the NH mean time series based on a simple energy balance model, while Tingley and Huybers (2010a) specify a spatiotemporal process for the surface temperature anomalies. In each case, weakly informative prior distributions are then placed on the parameters of the (spatio) temporal process, so that the estimates of the parameters are dominated by the data. Conditional on the process level being correctly specified, the posterior draws of the temperature time series or space–time process will have, on average, a reasonable variability—regardless of the signal-to-noise ratio for the proxies. If the proxies are uninformative, however, the variability of the *mean* across these draws will be attenuated—see Fig. 2 of Tingley and Huybers (2010b).

5. Conclusions

Hierarchical statistical modeling combined with Bayesian inference provides a flexible framework for analyzing data subject to numerous sources of uncertainty. In the paleoclimate context, hierarchical modeling allows the proxies to be linked to the target climate process via a probabilistic data-level model, and a parametric model of the target climate process to be incorporated at the process level. The indirect regression basis of LOC, which specifies the proxy observations as the dependent variable, can be understood as the data-level specification of a hierarchical model. By casting LOC into a Bayesian hierarchical modeling framework, weakly informative priors for the target temperature process eliminate the variance inflation that occurs in LOC if the proxies are

uninformative of the climate process or if the observations of the target climate are subject to errors. Furthermore, since Bayesian hierarchical models allow for uncertainty propagation through multiple steps of an analysis, framing LOC as a Bayesian model thus allows for internally consistent uncertainty quantification—a key omission in the original description of LOC.

LOC is closely related to the data level of the model presented in either Li et al. (2010) or Tingley and Huybers (2010a). In turn, both of these Bayesian models make unrealistic but simplifying assumptions in order to facilitate the inference. The model of Li et al. (2010) allows the proxies to reflect the target climate process at different levels of temporal averaging, but is non-spatial, while that of Tingley and Huybers (2010a) includes a spatial component but assumes that each proxy observation reflects spatially and temporally local information about the climate field. Despite the simplicity of current implementations, Bayesian hierarchical models are more flexible and hold the potential to be much more comprehensive than LOC. Bayesian hierarchical modeling is a promising approach to inferring past climate, and the specification of more realistic data-level and process-level models remains an area of active research.

Acknowledgments. This work was conducted while MPT was part of The Institute for Mathematics Applied to Geosciences at The National Center for Atmospheric Research. MPT is supported in part by the NSF under Grants ATM-0902374 and ATM-0724828, and BL by the NSF under Grant DMS-1007686.

REFERENCES

- Ammann, C., M. Genton, and B. Li, 2010: Technical note: Correcting for signal attenuation from noisy proxy data in climate reconstructions. *Climate Past*, **6**, 273–279.
- Banerjee, S., B. P. Carlin, and A. E. Gelfand, 2004: *Hierarchical Modeling and Analysis for Spatial Data*. Chapman & Hall/CRC, 472 pp.
- Brohan, P., J. J. Kennedy, I. Harris, S. F. B. Tett, and P. D. Jones, 2006: Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850. *J. Geophys. Res.*, **111**, D12106, doi:10.1029/2005JD006548.
- Christiansen, B., 2011: Reconstructing the NH mean temperature: Can underestimation of trends and variability be avoided? *J. Climate*, **24**, 674–692.
- Fuller, W. A., 1987: *Measurement Error Models*. Wiley, 440 pp.
- Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin, 2003: *Bayesian Data Analysis*. 2nd ed. Chapman & Hall/CRC, 696 pp.
- Hegerl, G., T. Crowley, M. Allen, W. Hyde, H. Pollack, J. Smerdon, and E. Zorita, 2007: Detection of human influence on a new,

- validated 1500-year temperature reconstruction. *J. Climate*, **20**, 650–666.
- Jansen, E., and Coauthors, 2007: Palaeoclimate. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 433–497.
- Li, B., D. Nychka, and C. Ammann, 2010: The value of multi-proxy reconstruction of past climate. *J. Amer. Stat. Assoc.*, **105** (491), 883–911.
- Mann, M., S. Rutherford, E. Wahl, and C. Ammann, 2007: Robustness of proxy-based climate field reconstruction methods. *J. Geophys. Res.*, **112**, D12109, doi:10.1029/2006JD008272.
- 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.
- Solomon, S., and Coauthors, 2007: Technical summary. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 19–91.
- Tingley, M., and P. Huybers, 2010a: 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.
- , and —, 2010b: 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.
- , P. Craigmile, M. Haran, B. Li, E. Mannshardt-Shamseldin, and B. Rajaratnam, 2012: Piecing together the past: Statistical insights into paleoclimatic reconstructions. *Quat. Sci. Rev.*, **35**, 1–22.