Environ Ecol Stat (2015) 22:105–126
DOI 10.1007/s10651-014-0291-6

Specification and estimation of the transfer function in dendroclimatological reconstructions

Maximilian Auffhammer · Bo Li · Brian Wright · Seung-Jick Yoo

Received: 21 June 2013 / Revised: 10 April 2014 / Published online: 10 May 2014
© Springer Science+Business Media New York 2014

Abstract We identify two issues with the reverse regression approach as implemented in several classic reconstructions of past climate fluctuations from dendroclimatologcical data series. First, instead of estimating the causal relationship between the proxy, which is measured with significant error, as function of climate and formally inverting the relationship, most papers estimate the inverted relationship directly. This leads to biased coefficients and reconstructions with artificially low variance. Second, we show that inversion of the relationship is often done incorrectly when the underlying causal relationship is dynamic in nature. We show analytically as well as using Monte Carlo experiments and actual tree ring data, that the reverse regression method results in biased coefficients, reconstructions with artificially low variance and overly smooth reconstructions. We further demonstrate that correct application of the inverse regression method is preferred. However, if the measurement error in the tree ring index is significant, neither method provides reliable reconstructions.

Handling Editor: Ashis SenGupta.

We thank the Giannini Foundation for generous funding of this research. All remaining errors in the mansucript are soleley ours.

M. Auffhammer (✉) · B. Wright
University of California, 207 Giannini Hall, Berkeley, CA 94720-3310, USA
e-mail: auffhammer@berkeley.edu

B. Li (✉)
University of Illinois at Urbana-Champaign, 725 S. Wright St.,
Champaign, IL 61820, USA
e-mail: libo@illinois.edu

S.-J. Yoo
Greenhouse Gas Inventory & Research Center of Korea,
Saemunan-ro 92, Jongno-gu, Seoul 110-999, Korea

Keywords Climate change · Climate variability · Inversion · Paleoclimatic data · Reconstruction method

1 Introduction

Long run temperature records are of key importance in many fields, but maybe most significant in climate science. Unfortunately, the longest instrumentally-measured temperature series span just over three centuries (Jones and Hulme 1997; Jones et al. 1986; Hansen and Lebedeff 1987; National Research Council 1998). To supplement the brief observed temperature series, paleoclimatologists have reconstructed long histories of climate variables by using proxy data series including tree-ring indexes, ice cores, pollen series, coral, and faunal and floral abundance in deep-sea cores. They select these proxy data series for their length of sample span, sensitivity to climate variability, and relative lack of disturbance from non-climate factors. Examples of the vast number of such reconstructions include Briffa et al. (1990, 1992, 1998, 2001), Scuderi (1993), Hughes and Brown (1992), Bradley and Jones (1992), Mann (2002), Mann et al. (1998, 1999), Salzer et al. (2009) and others surveyed in Jones et al. (1998, 2001).

Recently, the attention to statistical issues in dendroclimatology has resulted in a rapidly growing literature examining the properties of reconstructions at single sites and the combination of these individual series to time series at larger spatial scales (Mann et al. 2008; National Research Council 2006; Jones et al. 2009). There is a vibrant literature which approaches the issue of proxy based reconstruction using a variety of approaches and methods. One commonly used approach is an application of the regularized EM algorithm (Schneider 2001; Mann et al. 2005). A more recent literature has focused on Bayesian approaches to reconstruction (Haslett et al. 2006; Lee et al. 2008; Tingley and Huybers 2010; Li et al. 2010; Brynjarsdóttir and Berliner 2011; McShane and Wyner 2011).

In this paper[1], we examine the statistical properties of a parametric method for proxy based temperature reconstructions at a single site, called *reverse regression*. This method is a straightforward application of the linear regression method. It is an appealing choice of technique, both because of its simplicity in execution and seemingly intuitive solution to the reconstruction problem. While it is just one of many possible techniques, it has been widely applied in a number of classic papers (see for example Graumlich and Brubaker 1986; Briffa et al. 1990; Till and Guiot 1990; Graumlich 1991; Scuderi 1993; Li et al. 2007). We examine two separate issues stemming from the application of the "reverse regression method". The first issue is a classic statistical issue related to the methodology of estimation, where measurement error in the proxy series leads to attenuation bias in the estimated parameter and suboptimal reconstructions. The second issue has to do with the specification of the estimated relationship between the climate and the paleoclimatic index if the underlying relationship is dynamic. We show using theory, Monte Carlo simulation and actual tree ring and climate data, that this reverse regression method will result in

[1] For information on the history of this paper see http://climateaudit.org/2009/12/16/climategatekeeping/#more-9569 Last accessed February 18, 2014.

reconstructions, relative to the true series of the climate index, with a smaller variance and very different time series properties. We examine an alternative reconstruction methodology called the inverse regression method, which is just as easily applied, and outperforms existing methods when the underlying tree ring data are relatively free of measurement error and strongly correlated with observed temperature. We also show, that in the absence of such quality data, neither method is capable of producing a satisfactory reconstruction.

Via a series of realistic pseudoproxy experiments, Christiansen (2011) also demonstrated that the inverse regression avoids the underestimation of low-frequency variability and suggested to use the inverse regression for the climate reconstruction. We share the same conclusion as Christiansen, but our paper focuses more on the statistical properties of the prediction in comparing two types of regression models. In addition to the mean and variance that are of Christiansen's interest as well, we also investigated the specification and estimation issues of a transfer function, and the dependence structure of the time series derived based on an estimated transfer function. Those two issues have often been neglected while applying the reverse regression to the paleoclimate reconstruction.

In the following section we describe the reverse regression method as well as the inverse regression method. Section 3 provides a theoretical examination of reconstructions from both methods. Section 4 contains the Monte Carlo evidence and describes the results from a reconstruction exercise using an actual sample of tree rings. Section 5 concludes.

2 Reconstruction of historical climate

Biological growth of trees, expressed in tree-ring width or latewood density is limited by operational environmental factors including temperature, precipitation and CO_2 content of the atmosphere (Fritts 1991). If climate has been limiting tree growth in a systematic way, and the relation between climate and tree growth is stationary through time, a very long history of tree growth can be employed to extend backward the relatively brief recorded history of climate.

The biological causal relation between tree-ring growth and climatic factors that are inputs to the growth process is known as the *response function*. This response function relates an index of tree ring growth, D_t, to a climate index, T_t. Both indexes are normalized to have mean zero and a unit variance. There are four potential sources of uncertainty when writing down this causal relationship. First, there is *model uncertainty* about how the two indexes are related. In practice there is some guidance based on the biological relationship between the climate index and the tree ring growth, yet there is significant uncertainty as to how many lags of the climate index causally affect D_t and whether they enter in a linearly additive or nonlinear fashion. Second, constructing a tree ring index is a complex and lengthy process (Fritts 1991). It is commonly assumed that D_t is *measured with error*. Third, the observed temperature index T_t may also be *measured with error* as the monitoring station is likely not located exactly at the location of the proxy sample. As is done in the literature, for the remainder of this paper we assume the most simple and best case scenario. Finally, the stochastic process ε_t follows is unknown (e.g. whether it is serially correlated or has constant variance). We assume that the true causal relationship between D_t and T_t is linear and additive and the number of lags of T_t affecting D_t is known. Further, we assume that T_t is measured without error and that ε_t is i.i.d. normal with mean zero and constant variance. The simplest possible response function under these assumptions is therefore given by:

$$D_t = \beta_o \cdot T_t + \varepsilon_t, \tag{1}$$

where ε_t is the measurement error of D_t, which is assumed to satisfy $E[\varepsilon_t|T_t] = 0$. If this strict exogeneity assumption holds, given an appropriate sample of T_t and D_t one can obtain a consistent and efficient estimate of β_o by ordinary least squares. Using this estimate for the population parameter, one can formally invert Eq. (1) and calculate predictions for T_t in years where only D_t is available. The inverse function of the response function, which has the climate indicator on the left hand side and is used to construct historical climate out of sample predictions, is called the *transfer function*. We refer to the approach of estimating the response function directly and formally inverting it to obtain the transfer function as "inverse regression".

As opposed to the inverse regression, the reverse regression method does not estimate Eq. (1), but rather the transfer function directly:

$$T_t = \gamma_o \cdot D_t + \eta_t.$$

Since D_t is measured with error, this creates a variety of estimation issues, which go back in the statistics literature to Eisenhart (1939). As we show in the next section, direct estimation of the transfer function causes the least squares estimator to be biased in large as well as small samples. Bias of the estimator even in a large sample poses a critical limitation in prediction or in reconstruction of temperature series in dendroclimatology. In addition to the biased coefficient estimates, the reverse regression also results in a downward bias in the variance of the reconstructed historical climate in large and small samples, even when assuming a valid underlying response function (Storch et al. 2004).

The second set of issues relates to the specification of the transfer function, when the underlying response function is thought to be dynamic. Scuderi (1993), for example, uses current temperature and two lagged years of temperature in the response function, in which case the specification is given by

$$D_t = \beta_o \cdot T_t + \beta_1 \cdot T_{t-1} + \beta_2 \cdot T_{t-2} + \varepsilon_t. \tag{2}$$

The inclusion of the lagged temperature variable is interpreted as reflecting year-to-year persistence of the effect of climate variables upon the response of tree growth (LaMarche 1974; Yoo and Wright 2000). In the standard application of the reverse regression approach, the number of lags for T_t in the response function determines the number of leads of D_t in the transfer function. This transfer function is then estimated directly using ordinary least squares. Scuderi (1993), based on the response function specified in Eq. (2) above, choses the transfer function

$$T_t = \gamma_o \cdot D_t + \gamma_2 \cdot D_{t+1} + \gamma_3 \cdot D_{t+2} + \eta_t.$$

It is standard practice in papers using this reverse regression method to use the directly estimated transfer function coefficients ($\hat{\gamma}_1,\ \hat{\gamma}_2,\ \hat{\gamma}_3$) to reconstruct T_t back beyond the horizon of observation. This reconstructed climate record can then be used to infer the time series properties of climate, including the dynamics of baseline variation, counts, and duration of climatic extremes such as severe droughts and heat waves.

This approach introduces two additional issues to the estimation stemming from measurement error in D_t discussed above. First, the fact that current and lagged observations of the climate variable causally affect tree growth does not imply that current and future tree growth should appear in the formal inversion of this relation. As we show below, whether the leads or the lags of the index of tree ring growth should be included as predictors in the transfer function for a given response function depends upon the relative magnitude of the parameters in the response function. Even when leads should be included in the transfer function, the number of leads is always greater than the number of lags in the response function. Furthermore we show that when the inclusion of leads is justified, current tree growth should be excluded from the transfer function.

Second, the misspecification of the transfer function results in an overestimate of the order of the autoregressive process of the reconstructed temperature series. This is consistent with the observation of very smooth reconstructions [e.g. Seater (1993)]. While empirically the estimated order of the AR process also depends on the selection criterion used, we argue based on theory and Monte Carlo evidence that direct estimation of the misspecified transfer function tends to produce reconstruction with artificially high autoregressive processes. This aspect of reconstruction is of crucial importance in the analysis of climate change, since overestimation of the order of the AR process of temperature underestimates the variability of the true temperature movement so that inferences about the significance of observed temperature deviations in a particular period prior to the availability of direct climate measurement are unreliable.

3 Comparing inverse and reverse regression: theory

3.1 Bias and variance of temperature estimates

We have discussed two basic types of underlying models that have been typically considered for estimating the temperatures from the proxy. One is directly regressing the temperatures on proxies (reverse regression) and the other is regressing the proxies on temperatures (inverse regression). The former corresponds to directly estimating the transfer function and the latter to the inversion of the estimated response function. In addition to reverse and inverse regression, these two methods have carried a variety of names in the statistics literature. For example, they sometimes are called the direct and indirect regression (Christiansen 2011), while at other times may have been called the inverse and classical calibration (Krutchkoff 1967). The comparison between these two statistical estimation methods has long been discussed going back to Eisenhart (1939) and later reviewed in Brown (1993). Here we compare these two approaches in the context of estimating temperatures using proxy data.

Both temperatures and proxies are random variables and they are jointly distributed. In principle we can construct the conditional distribution of temperature given proxies in order to estimate unknown temperatures at given proxies. However, from a modeling perspective the inverse regression is preferred over reverse regression for two reasons. First, the transfer function may not be the most appropriate model for describing the relationship between temperatures and proxies. It is commonly agreed that the proxy is noisy with the signal to noise ratio to be roughly 0.25 by variance (Smerdon 2012), while the temperature measurements, though contain noise, are relatively more accurate estimates. In this sense, it might be more appropriate to model the linear relationship by having proxies as the response variable with temperature as the predictor and to actually represent the source of errors. This seems to suggest that estimation of the response function is more natural and sensible than the direct estimation of the transfer function. Estimates of out of sample temperatures should be obtained through the inversion of the estimated response function. Another disadvantage of the reverse regression method is that it is unclear whether the past temperatures are from the same population as the observed temperatures. With this uncertainty it may be less risky to model temperatures as predictors.

In terms of statistical properties of the estimated temperatures, the two estimation methods each have their own advantages and disadvantages. Direct estimation of the transfer function encounters the classical problem of errors in variables (Klepper and Leamer 1984). Ordinary Least Squares (OLS) estimators are biased and inconsistent if the tree-ring index series used as a predictor in the transfer function is subject to measurement error due to disturbances (ε_t) in the response function. Therefore, the estimates of the parameters of this direct estimation of the transfer function are biased downward. However, compared to the inversion of response function, the direct estimation reduces the variability. To demonstrate this, consider the simplest static case below, which omits the issue of lags and leads discussed earlier, but can be easily extended to the general case.

Specify the response, D_t, as determined by the contemporaneous impact of T_t with measurement error $\varepsilon_t | Temp_t \sim N(0, \sigma^2)$:

$$D_t = \beta T_t + \varepsilon_t. \tag{3}$$

As is standard in the literature, both D_t and T_t are standardized to have mean 0 and standard deviation 1. In Eq. (3) this implies a coefficient of determination (R^2) of β^2. Given the standardization the OLS estimate of β based on a sample of size n is:

$$\hat{\beta} = \frac{\sum_{t=1}^{n} D_t \cdot T_t}{\sum_{t=1}^{n} T_t^2} = \mathrm{cov}(D_t, T_t).$$

The specification of the reverse regression used to reconstruct T_t's from D_t's is

$$T_t = \gamma D_t + \eta_t. \tag{4}$$

Given the standardization, the OLS estimate of γ is given by:

$$\hat{\gamma} = \frac{\sum_{t=1}^{n} D_t \cdot T_t}{\sum_{t=1}^{n} D_t^2} = \text{cov}(D_t, T_t).$$

The OLS estimate, $\hat{\gamma}$, of the reverse equation equals the OLS estimate of the response function parameter $\hat{\beta}$, because of the standardization of D_t's and T_t's. Given the causal relation in (3), the matching transfer function to be used for reconstruction is given by:

$$T_t = \frac{1}{\beta}(D_t - \varepsilon_t) = \frac{1}{\beta} D_t + \eta_t,$$

noting that $\eta_t = -\frac{1}{\beta}\varepsilon_t$. If we take the ratio of the coefficients on the tree ring index from the reverse regression in (4) and the inverse regression in (3) and notice that for $|\beta| < 1$, we get:

$$\frac{\hat{\gamma}}{\frac{1}{\hat{\beta}}} = \hat{\gamma}^2 \ll 1.$$

Thus the OLS estimate of the coefficient of D_t in the reverse regression is biased toward zero by a proportion $(1 - \beta^2)$. Since β^2 is the coefficient of determination of the OLS estimate of the response function, the bias grows as the fit of the response function deteriorates. This corroborates the discoveries in Ammann et al. (2010).

The reverse regression is subject not only to bias of the estimated coefficient but also to underestimation of the variance of the reconstructions, $\widehat{T}_t^r$:

$$\text{var}\left(\widehat{T}_t^r\right) = \gamma^2 \text{var}(D_t) = \gamma^2\beta^2 \text{var}(T_t) + \gamma^2 \text{var}(\varepsilon_t) = \gamma^2,$$

since $\text{var}(\varepsilon_t) = 1 - R^2 = 1 - \gamma^2$. The variability of the reconstruction by reverse regression is determined by the estimate of the coefficient, γ , which equals β in expectation. Hence the variability of the reconstruction is also biased downward and decreases as the fit of the response function declines. This is consistent with the observation made in Storch et al. (2004). On the other hand, if the transfer function is obtained by inversion of the estimated response function, the variance of the reconstruction, $\widehat{T}_t^i$, in expectation is overestimated relative to the true (unobserved) climate:

$$\text{var}\left(\widehat{T}_t^i\right) = \text{var}\left(\frac{1}{\beta} D_t\right) = \frac{1}{\beta^2}\text{var}(\beta T_t + \varepsilon_t) = \text{var}(T_t) + \frac{1}{\beta^2}\text{var}(\varepsilon_t) = \frac{1}{\beta^2} > 1.$$

The last equal sign in the above equation holds because $var(\epsilon) = 1 - \beta^2$. The overestimation of the true variance of the climate series grows as the fit of the response function declines. A noisy proxy therefore leads to a noisy reconstruction.

In summary, the direct estimate of transfer function will yield biased parameter estimates and thus bias predictions. However, these predictions have smaller variance than those obtained from the inversion of response function.

Let $\sigma_T^2 = \sum(T_t - \bar{T}_t)^2/n$. Asymptotically, Brown (1993, p 32) gives

$$E(\widehat{T}^i) = T,\ \text{var}(\widehat{T}^i) = \sigma^2/\beta^2;$$
$$E(\widehat{T}^r) = \rho^2 T + (1-\rho^2)\bar{T},\ \text{var}(\widehat{T}^r) = \rho^4\sigma^2/\beta^2,$$

where $\rho^2 = \beta^2\sigma_T^2/(\sigma^2 + \beta^2\sigma_T^2), 0 \le \rho^2 \le 1$.

The bias of $\widehat{T}^r$ is small if the new T is close to $\bar{T}$, the mean of the temperature in the training data. In general, $\widehat{T}^r$ has smaller mean-squared error than $\widehat{T}^i$ when T is close to $\bar{T}$, particularly so if ρ^2 or signal to noise ratio $\beta^2\sigma_T^2/\sigma^2$ is small (Krutchkoff 1967), otherwise $\widehat{T}^i$ performs better by this measure. With T centered, $E(\widehat{T}^r) = \rho^2 T$, and

$$\text{MSE}(\widehat{T}^i) - \text{MSE}(\widehat{T}^r) = \sigma^2/\beta^2 - \rho^4\sigma^2/\beta^2 - (1-\rho^2)^2T^2.$$

Then $\widehat{T}^r$ is better than $\widehat{T}^i$ if and only if $T^2 < \sigma^2/\beta^2 + 2\sigma_T^2 = \sigma_T^2(1/\rho^2 + 1)$. In climatology, we mainly focus on the long term climate variability, so we would prefer the unbiased estimation method regardless of the comparison in MSE. Tingley and Li (2012) point out that in some settings a Bayesian framework can regularize the instability problem of the inverse regression.

3.2 Specification of the transfer function

In this section we illustrate the specification and estimation issues in a simple case in which the response function contains only the current value and one lag of the climate variable (e.g. Briffa et al. 1992). The issue in this section generalizes to situations with more complicated lag structures and is only applicable to settings where the response function is not static. Let D_t be the annual increment of tree-ring width and T_t temperature at date t. The regression equation of the response function of tree growth on the temperature index is given by

$$D_t = \beta_o \cdot T_t - \beta_1 \cdot T_{t-1} + \varepsilon_t = \beta_o\,(1 - \phi L)\,T_t + \varepsilon_t, \tag{5}$$

where $\phi = \frac{\beta_1}{\beta_o}$ and the lag operator L is defined as $L^p T_t = T_{t-p},\ p\epsilon\ N$. The measurement error ε_t is assumed to be i.i.d. and $\varepsilon_t|Temp_t \sim N(0, \sigma^2)$. Inversion of Eq. (5) will vary depending upon the ratio of the parameters, β_o and β_1. First consider the case where $|\phi| < 1$. Assuming that T_t is a bounded sequence, we can solve Eq. (5) for T_t as follows (Hamilton 1994, page 19):

$$\begin{aligned} T_t &= (\beta_o\,(1-\phi L))^{-1}\,(D_t - \varepsilon_t) \\ &= \frac{1}{\beta_o}\left(D_t + \phi D_{t-1} + \phi^2 D_{t-1} + \phi^3 D_{t-3}\ldots\right) + \nu_t \\ &\cong \frac{1}{\beta_o}\left(D_t + \phi D_{t-1} + \phi^2 D_{t-1} + \phi^3 D_{t-3}\ldots + \phi^m D_{t-m}\right) + \nu_t^m \end{aligned} \tag{6}$$

where $\nu_t = -\frac{1}{\beta_o}\left(1+\phi L+\phi^2L^2+\phi^3L^3\ldots\right)\varepsilon_t$ and $\nu_t^m = -\frac{1}{\beta_o}\left(1+\phi L+\phi^2L^2 +\phi^3L^3\ldots+\phi^mL^m\right)\varepsilon_t$. If D_t is determined by the current and the first lagged value of T_t in the response function, the corresponding transfer function should be defined as the weighted sum of the lags of the tree growth measures and of the errors in the estimated response function back to the distant past with geometrically decreasing weights. Only then will it reflect the biological relation that specifies the direction of causality in the response function. If $|\phi| < 1$, the marginal effect of current temperature is larger than that of the previous year. The above inversion shows that given response relation (3), T_t in the transfer function should be regressed on the current observation and previous observations of tree-ring growth, rather than on D_t and its first lead, D_{t+1}, as is the standard approach.

Now consider the case where $|\phi| > 1$. We can invert Eq. (5) using the lead operator, which is the inverse of the lag operator $L^{-p}T_t = T_{t+p}$ where $p > 0$, to get the transfer function (Hamilton 1994, page 41):

$$\begin{aligned} T_t &= \left(\beta_o\left(1-\phi L\right)\right)^{-1}\left(D_t-\varepsilon_t\right)\\ &= -\frac{1}{\beta_o}\left(\phi^{-1}D_{t+1}+\phi^{-2}D_{t+2}+\phi^{-3}D_{t+3}\ldots\right)+\nu_t\\ &\cong -\frac{1}{\beta_o}\left(\phi^{-1}D_{t+1}+\phi^{-2}D_{t+2}+\phi^{-3}D_{t+3}\ldots+\phi^{-(m+1)}D_{t+m+1}\right)+\nu_t^m, \end{aligned}$$

where $\nu_t = -\frac{1}{\beta_o}\left(\phi^{-1}L^{-1}+\phi^{-2}L^{-2}+\phi^{-3}L^{-3}\ldots\right)\varepsilon_t$ and $\nu_t^m = -\frac{1}{\beta_o}\left(\phi^{-1}L^{-1} +\phi^{-2}L^{-2}+\phi^{-3}L^{-3}\ldots+\phi^{-(m+1)}L^{-(m+1)}\right)\varepsilon_t$. For this parameterization of the response function (5), where lagged climate has a larger impact on tree ring growth than current climate, the matching transfer function is the one where current temperature is not a function of current tree growth, but is a weighted sum of leads of tree growth and the errors ε_t with geometrically declining weights.

It is common (e.g. Briffa et al. 1990, 1992) to estimate the transfer function directly, after estimating the response function as a first step to determine the relevant number of lags. This approach ignores the fact that a transfer function including one or two leads of D_t is misspecified if the underlying response function that is correctly specified has one or two lags. Assuming that climate is accurately measured by the climate index in the sample period, the transfer function relating a climate variable to tree growth should be generated by inverting the estimated response function. If tree growth is generally limited by the current and previous years' climatic conditions (LaMarche 1974) the climate measure in a given year should be calculated as a function of either current and past tree growth if $|\phi| < 1$, or of future tree growth excluding current year tree growth if $|\phi| > 1$. The relevant specification of the transfer function therefore depends upon the relative size of the marginal effect of the current (β_o) and the previous year's (β_1) climate on current tree growth. Empirically, the weights of the lags/leads should be calculated from the estimated response function parameters.

If the order of lags of T_t on D_t is not known *a priori*, the order of lags of the temperature series can be estimated *ex post* and utilized in inverting the response function into the transfer function, as illustrated above for the one-lag case for $|\phi| < 1$

or $|\phi| > 1$. In the following section we discuss an additional statistical problem that arises if the transfer function is misspecified as described above.

3.3 Time series properties of climate reconstructions

The autocorrelation order of long run history of the reconstructed climate series is another important characteristic in making inferences about its baseline fluctuation. When the dependent variable (D_t) is caused by the lagged and the current realizations of T_t, the dependent variable exhibits time dependent behavior (even without serial correlation of ε_t). Let us consider the implications for the autoregressive order of the estimated climate series from the two estimation approaches using the following simple response function equivalent to (5) as an example:

$$D_t = \beta_o T_t + \beta_1 T_{t-1} + \varepsilon_t, \tag{7}$$

where both D_t and T_t are standardized to have mean zero and unit variance and $\varepsilon_t \sim N(0, \sigma^2)$. Applying reverse regression the following transfer function is estimated directly:

$$T_t = \gamma_o D_t + \gamma_1 D_{t+1} + \eta_t. \tag{8}$$

Since the variables are standardized we can represent the OLS estimates in (7) as a function of the OLS estimates in (8):

$$\hat{\gamma}_o = \frac{\hat{\beta}_o\left(1-\hat{\beta}_1^2\right)}{1-\hat{\beta}_o^2\hat{\beta}_1^2} \text{ and } \hat{\gamma}_1 = \frac{\hat{\beta}_1\left(1-\hat{\beta}_o^2\right)}{1-\hat{\beta}_o^2\hat{\beta}_1^2}.$$

Then the reconstruction of the climate series using reverse regression, $\widehat{T}_t^r$, is given by $\widehat{T}_t^r = \hat{\gamma}_o D_t + \hat{\gamma}_1 D_{t+1}$. To focus on the problem caused by estimating a misspecified transfer function, abstracting from the errors-in-variables problem discussed above, we remove the unobserved factor, ε_t, from D_t, resulting in a perfect fit for the response function. Recognizing that $\hat{\beta}_i$ is an unbiased estimate of β_i for $i = 0, 1$, and using the true relation between β's and γ's, we obtain:

$$\begin{aligned}\widehat{T}_t^r &= \frac{1}{1-\beta_o^2\beta_1^2}\left[\beta_o^2\left(1-\beta_1^2\right)+\beta_1^2\left(1-\beta_1^2\right)\right]T_t\\ &+\frac{1}{1-\beta_o^2\beta_1^2}\left[\beta_o\beta_1\left(1-\beta_1^2\right)T_{t-1}+\beta_o\beta_1\left(1-\beta_o^2\right)T_{t+1}\right].\end{aligned}$$

The reconstructed $\widehat{T}_t^r$ is the weighted sum of the lag and lead of T_t as well as the current value. Reconstructions show dependency across time even if they are reconstructed from error-free variables. This is due to the misspecified transfer function, which does not reflect the physical causal relationship given by the response function. On the other hand the reconstruction by inversion of the response function will only exactly match the true T_t if D_t is free from errors, ε_t, as assumed in this illustration. When there are errors in the response function, its inversion also introduces time dependent behavior in the reconstructed series—even if the errors, ε_t, are i.i.d.

In summary, a noisier response function will result in noisier reconstructions, with potential time dependence, even if the response function is correctly inverted. But the estimation of a misspecified transfer function induces time dependency due not only to the errors in the response relation but also to the misspecification of the transfer function. The limitations of this method of reconstruction do not disappear if the fit of the response relation is perfect. For high-quality data, for which the response relation has high explanatory power, the inversion method is superior. For low-quality data, little can be expected from either method.

4 Performance comparison

In order to more fully investigate the issues of specification estimation addressed above, we conduct several Monte Carlo experiments as well as a reconstruction using an actual tree ring index.

4.1 Monte Carlo experiment

In the first experiment, we choose β_o and β_1 so that their ratio is equal to the ratio of the values found by estimating the response function from the data by Briffa et al. (1992). A set of independent T_t's is generated from the standard normal distribution and held constant for the experiment. The i.i.d. disturbances, ε_t, were repeatedly drawn from a mean zero normal with the variance chosen so that the variance of the D_t's is 1 given the values of β_o and β_1. For each experiment we chose the sample size $T = 1,000$, and we replicated the sample 1,000 times.

To apply the inversion method, we estimated the response function (5) for the last 100-year subsample, and reconstructed the series T_t for the whole sample from the series for D_t via Eq. (6) using $m = 10$. The reverse regression method was implemented by estimating the transfer function (8) from the last 100-year subsample, and reconstructing the whole sample from this estimated regression relation. As an initial comparison of the two methods of reconstruction we take an average of the reconstructions for each year across the 1,000 experiments. If a reconstruction method produces unbiased reconstructions, the mean reconstruction should be very close to the true value for any given year. Hence, we plot the difference of the *mean reconstruction* from the true T_t *for each year* in Fig. 1 over the whole sample period for both methods.

Figure 1 shows that the deviations of the *averaged* reconstructions for each year produced via the inversion method from the true generated time series of T_t's are dense around 0, indicating that the mean reconstruction is quite accurate. On the other hand, the averaged reverse regression method result in a much noisier estimate of the true mean reflected in the larger variance of the series plotted in Fig. 1. It is important to note the order of magnitude difference in the vertical scale, and it is important to differentiate the variance of the series plotted here form the variance of a single reconstruction. What this figure shows is the difference between the average reconstruction (across

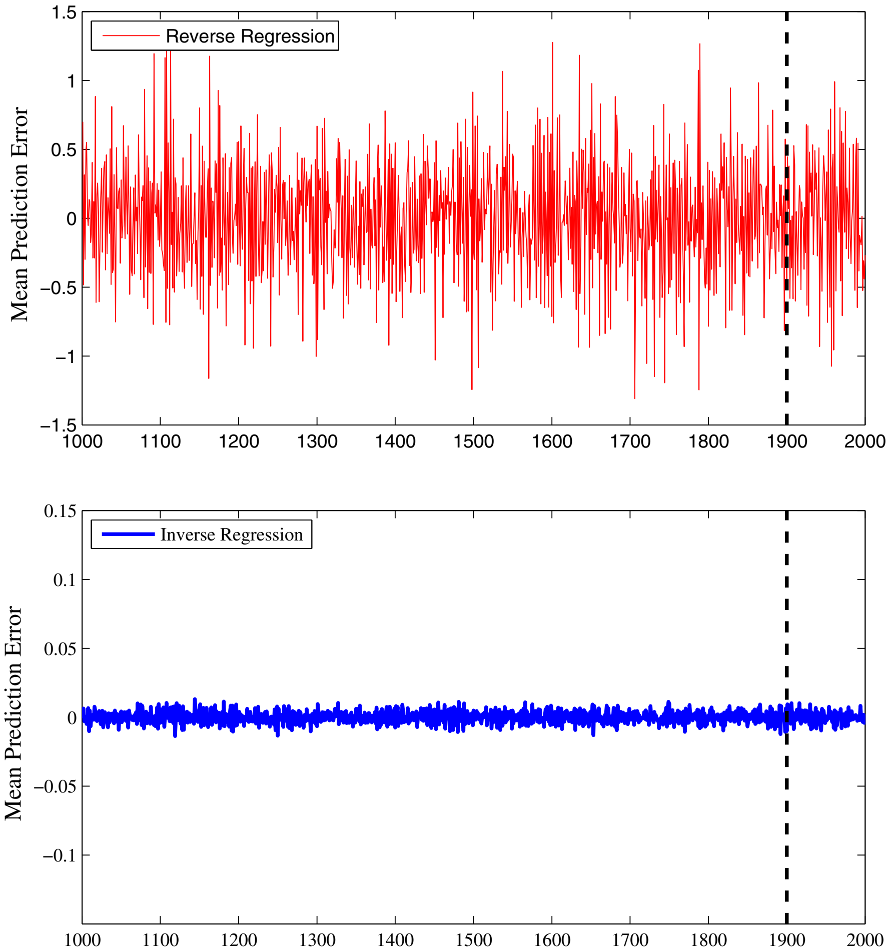

Fig. 1 Difference of the Mean Annual Reconstruction from the True Temperature ($\beta_o = 0.68$; $\beta_1 = 0.10$). The *top panel* plots the difference of the mean reconstruction from the true T_t for each year in the sample period using the reverse regression method. The *bottom panel* plots the difference of the mean reconstruction from the true T_t for each year in the sample period using the inverse regression method. The scale for the bottom panel is compressed

1,000 runs) and the truth (which is the same in each run) for each year. This number should be close to zero if the estimator performs well on average. Without surprse, it shows that the reverse regression reconstruction performs very poorly so the bias of this reconstruction is very noisy, but the inverse regression is very close to the target value so the variance of the difference from the inversion regression reconstruction is very small. We have verified this by looking at the absolute value of the deviation for each run from the truth for both methods and averaged that across runs. For the inverse regression for the average year this deviation is about 70 % smaller than that of the reverse regression.

To further compare the performance of the two reconstruction methods, we perform additional Monte Carlo experiments for different sets of parameter values for β_o and β_1. For the first set of experiments we hold β_o fixed and explore the performance of the two reconstruction methods for a range of β_1 such that $|\phi| < 1$, given that the T_t's and D_t's are again normalized. As a first measure of performance, we use the mean absolute deviation of reconstructions from the generated "true" value for each of the two reconstruction methods, which is defined as

$$\text{MAD} = \frac{\sum_{t=1}^{T} |f_t - T_t|}{T},$$

where f_t is the average of the reconstructions from 1,000 replications. The quality of the reconstruction of T_t depends on two factors, the explanatory power of T_t in the response function and the ratio of the current impact of climate to its lagged impact. Figure 2 shows the MAD performance of the two methods for three different levels of β_o across the valid range of β_1. Reconstruction either by inversion of the estimated response function or by estimation of the transfer function improves as the explanatory power of the climate variable in the response function increases. If the response variable is severely contaminated by measurement error, high quality reconstructions of climate variables from the respondent variable are not possible with either method.

Irrespective of the explanatory power of climate variables in the response function, the inversion method proves superior in reproducing the true T_t's. As the ratio of β_o to β_1 of the response function decreases, the relative role of the lagged explanatory variables increases and the accuracy of the reconstructions by inversion deteriorates. The average of the mean absolute deviation from the true values of reconstructions by reverse regression is almost an order of magnitude higher than that of the inversion method even when the current effect of T_t on D_t compared to the lagged impact is dominant, for example $\beta_o = 0.9$, $\beta_1 = 0.01$. Reverse regression bias makes the reconstruction less accurate.

Table 1 shows results from Monte Carlo experiments of reconstructions for seven combinations of (β_o, β_1). For each experiment, the true standard deviation of T_t is exactly 1 for the sample period. The top row in the table indicates the magnitude of the coefficients of the response function. The share of the variance in temperature coming from a signal instead of noise (ε_t) increases as we go to the right in the table, as indicated by the implicit variance of ε_t.

It should be emphasized that the deviations of the variance of reconstructions in the inversion approach are not due to a problem with the method, but are an inevitable consequence of the low explanatory power of climate in the determination of tree growth. Consistent with Fig. 2, the table indicates that the MAD of the reconstructions using the inverse regression methods is consistently smaller than that of the reverse regression method and converges towards zero as the response function becomes less noisy. The reverse regression reconstructions never become unbiased. The mean of the estimated standard deviation (SD) of reconstruction is biased upwards for the inversion approach and biased downwards for the reverse regression approach. When the explanatory power of T_t is 0.99 as in the case with (β_o, β_1) equal to (0.8, 0.59) or (0.9, 0.42), the means of the estimated standard deviation for the reconstructions from

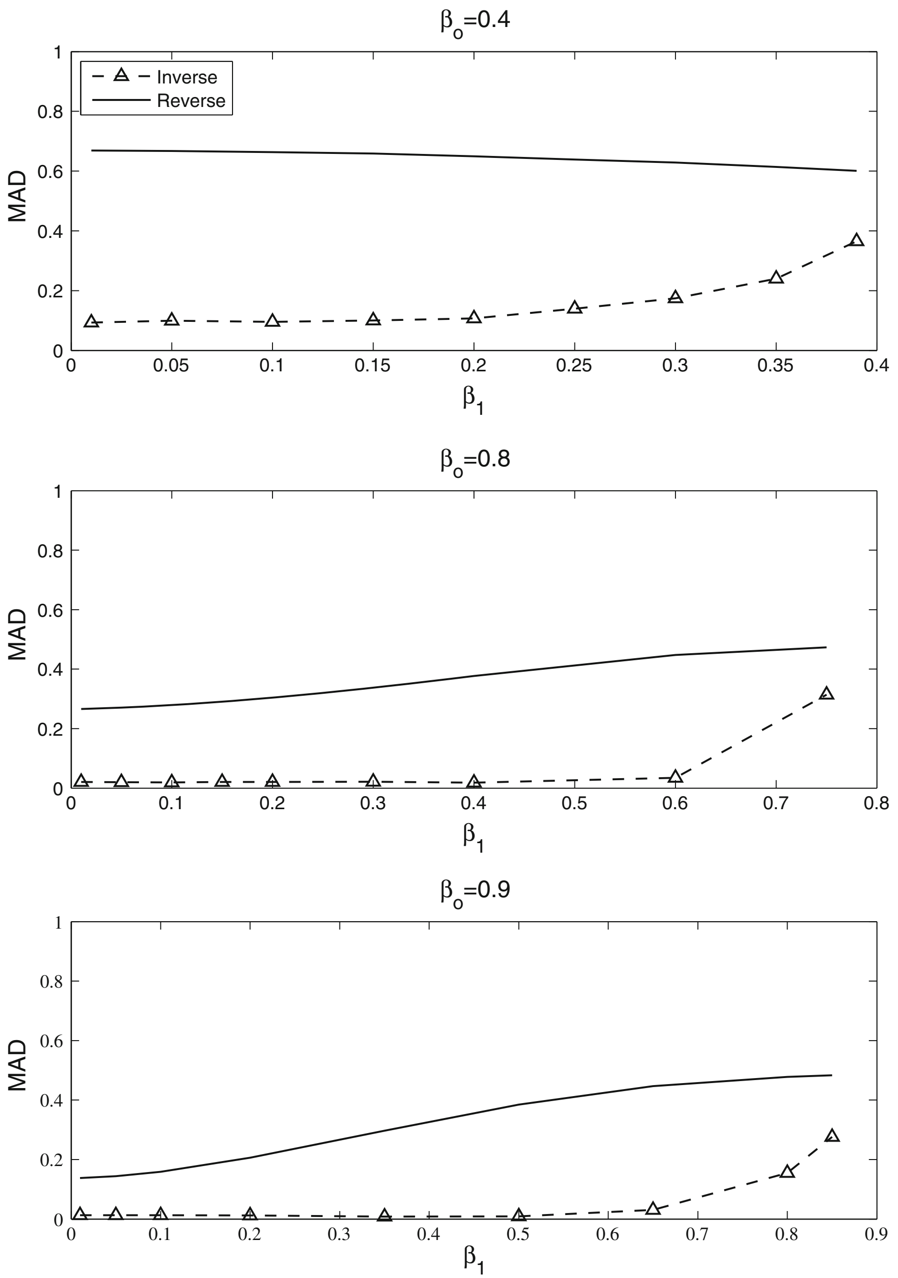

Fig. 2 Mean Absolute Deviations of Reconstructions. The three panels above show the MAD performance of the two methods for three different levels of β_o across the valid range of β_1. Reconstruction either by inversion of the estimated response function or by estimation of the transfer function improves as the explanatory power of the climate variable in the response function increases

Table 1 Summary of Monte Carlo experiment

var(ε_t)	β_0	β_1	β_0	β_1	β_0	β_1	β_0	β_1
	0.4	0.1	0.4	0.2	0.6	0.2	0.8	0.2
	0.83		0.8		0.6		0.32	
	Inv.	Rev.	Inv.	Rev.	Inv.	Rev.	Inv.	Rev.
MAD	0.10	0.66	0.12	0.65	0.04	0.50	0.02	0.30
SD	2.84	0.47	3.13	0.49	1.75	0.67	1.26	0.85
RE	0.15	−3.39	0.12	−2.79	0.35	−0.45	0.66	0.50
CE	0.15	−3.42	0.12	−2.81	0.35	−0.46	0.66	0.49
AR(0)	19 %	18 %	5 %	2 %	15 %	1 %	53 %	0 %
AR(1)	80 %	46 %	87 %	23 %	78 %	59 %	47 %	95 %
AR(2)	1 %	20 %	9 %	31 %	7 %	30 %	1%	5 %
AR(3)	0 %	9 %	0 %	16 %	0 %	7 %	0 %	0 %
AR(4)	0 %	4 %	0 %	14 %	0 %	3 %	0 %	0 %
AR(5+)	0 %	3 %	0 %	14 %	0 %	0 %	0 %	0 %
var(e)	β_0	β_1	β_0	β_1	β_0	β_1		
	0.9	0.1	0.8	0.59	0.9	0.42		
	0.18		0.0119		0.0136			
	Inv.	Rev.	Inv.	Rev.	Inv.	Rev.		
MAD	0.01	0.16	0.03	0.44	0.00	0.34		
SD	1.13	0.94	1.04	0.85	1.03	0.92		
RE	0.82	0.78	0.96	0.57	0.98	0.78		
CE	0.82	0.78	0.96	0.56	0.98	0.78		
AR(0)	93 %	28 %	100 %	0 %	100 %	0 %		
AR(1)	7 %	72 %	0 %	0 %	0 %	0 %		
AR(2)	0 %	0 %	0 %	0 %	0 %	0 %		
AR(3)	0 %	0 %	0 %	0 %	0 %	20 %		
AR(4)	0 %	0 %	0 %	0 %	0 %	80 %		
AR(5+)	0 %	0 %	0 %	100 %	0 %	0 %		

This table shows the results from the Monte Carlo experiment discussed in Sect. 4.1 We use a time series of length T = 1, 000 with 1,000 iterations. *Inv.* stands for the Inverse Regression Method and *Rev* stands for the Reverse Regression Method. The coefficients β_0 and β_1 correspond to the estimated parameters in Briffa et al. (1992). The ε_t is chosen to give the temperature series a variance of 1. MAD the mean absolute devation as defined in Sect. 4.1. SD is the standard deviation of the reconstructions, which if unbiased should be equal to 1. RE and CE are measure of fit as given in Mann and Rutherford (2002). The AR( ) rows indicate which order autoregressive process is chosen for the reconstructions based on the Schwarz Information Criterion

reverse regression ($\widehat{T}_t^r$) are 0.85 and 0.92 while the reconstructions from the inversion method ($\widehat{T}_t^i$) have an estimated standard deviation of 1.04 and 1.03 respectively, which is close to the true value of 1. The bias in variance reduces dramatically going from left to right as the explanatory power of the T_t's increases (as var(ε_t) goes to 0.01) and almost disappears for the inverse regression method, but not so for the reverse regression method. The latter technique's bias, given a valid response function with high explanatory power, is predominantly due to misspecification of the transfer function - not to the error in D_t.

Since MAD is only one measure of performance, we calculate the mean "Reduction of Error" (RE) and the "Coefficient of Efficiency", which range from $-\infty$ to 1. RE is zero if the reconstruction is set equal to the mean of the target series over the calibration interval, while the CE is zero if the reconstruction is set equal to the mean of the target series over the validation interval. Both of the measures are standard in the dendroclimatology literature and described in detail in Cook et al. (1994) and National Research Council (2006). For each statistic, a value of zero indicates that the reconstruction method is equivalent in performance to using the sample mean. Values less than zero indicate that the mean outperforms the reconstruction. Values greater than zero and less than one indicate the superior predictive ability of the model relative to simply using the mean. As Table 1 indicates, the inverse regression method for both statistics always outperforms the mean. For inverse regression, both RE and CE come close to the theoretical maximum of 1 for the last two low noise scenarios. The reverse regression method has negative values for both statistics for the first two scenarios. Further, for the reverse regression method neither the RE nor the CE is greater than those of the inverse regression method for any of the considered coefficients.

The bottom six rows of Table 1 show supporting Monte Carlo evidence of our theoretical claim from the previous section, that the reverse regression method will overestimate the order of the autoregressive process. In each experiment above we have generated T_t as an independent process over time, which implies that T_t is AR(0). For each iteration and reconstruction method, we use the Schwarz criterion to render a consistent estimate of the order of the process. (Note that using the inconsistent AIC would further increase the estimated order of the autoregressive process). As above, the success of each method in correctly identifying the order of the autoregressive process depends on the explanatory power of the series T_t in determining D_t. When explanatory power is very low, neither method is successful at correctly identifying the correct order. When the variances of errors, ε_t, are less than 0.2, the reconstructions by inversion of the response function are successfully identified as independent processes over time in most of the cases. For the response function, which contains the least degree of noise ($\beta_0 = 0.8$; $\beta_1 = 0.59$), the inverse regression method based reconstructions are correctly identified as following an AR(0) process for 100 % of the reconstructions, while the autoregressive order of the reverse regression method based reconstructions is identified as higher than AR(5) 100 % of the time.

The reconstructions by reverse regression yield consistently erroneous, positive orders of the autoregressive process. This is one likely reason that Seater (1993) identified the underlying process of the reconstructed temperature series of Briffa et al. (1992) as AR(17). The second Monte Carlo (II) uses parameters whose ratio is equal to that of those estimated in Briffa et al. (1992) and shows that the reconstructions from reverse regression indicate an AR(5) or higher order AR process in 14 % of the cases even though the true process is AR(0). Storch et al. (2004) have pointed out this underestimation of the variability of the climate reconstructions, looking at studies using principal components of the proxy data series as predictors of the instrumental temperature series. Above we formalize this notion and show an additional methodological issue contributing to the underestimate of variability in the reconstruction.

4.2 Estimation and reconstruction: Briffa et al. (1992)

To provide a concrete example of the consequences from reconstructing a climate index using a proxy index, we estimate the response function using the 1876–1974 temperature and latewood density index used by Briffa et al. (1992), and reconstruct the climate series for a 1,524 year time span. For the reverse regression we estimate the first model given in Table 2 in Briffa et al. (1992) using one lead of the tree ring and maximum latewood density index via least squares, which corresponds to Eq. (7) above. For the inverse regression method we only use the maximum latewood density index with one lag, which is significant in the response function using Newey and West (1994) standard errors. The fit of the response function in this example is very good for this given data set (similar to the 4th and 5th example in Table 1). The results from the two reconstruction methods are summarized in Table 2. Figure 3 plots the actual temperature and reconstructions using both methods.

Matching the results from the Monte Carlo experiment, the estimated standard deviation of the reconstructions from the inversion method is 1.45 times the true standard deviation, thus overestimating the climate variability. The reconstruction via the inversion method results in a reconstruction with an autoregressive process matching the order of original series—AR(0). On the other hand, estimation of the reverse regression underestimates the standard deviation by 29 %, resulting in an overly smooth series. It follows an AR(3) process, not matching the order of the actual temperature series. These results provide supporting evidence of what Storch et al. (2004) noted—namely that traditional reconstruction methods result in an overly smooth historical climate record. Here we show that this excessive smoothing is a consequence of the statistical estimation issues pointed out above as well as the misspecification of the transfer function. We further conducted calibration/verification exercise using the year 1925–1976 as a calibration period and the year 1876–1924 as a verification period. As Table 2 shows, the inverse regression method outperforms the reverse regression method using the commonly used RE and CE measures of fit that we have also employed earlier.

Table 2 Comparison of the reconstructions for the in sample period using the Fennoscandia sample used by Briffa et al. 1992

	Actual temperature	Reconstruction by inversion method	Reconstruction by reverse regression
Average	0	0	0
Standard deviation	1.01	1.46	0.71
Order of AR process	0	0	3
Correlation with temperature	1	0.71	0.72
RE**	–	0.52	0.2
CE**	–	0.52	0.21

The series were normalized to have mean zero and standard deviation 1 for the entire sample. We used the years 1925–1976 as a calibration period and 1876–1924 as the verification period. RE and CE are calculated using the formulas given in Mann and Rutherford (2002) with N = 101

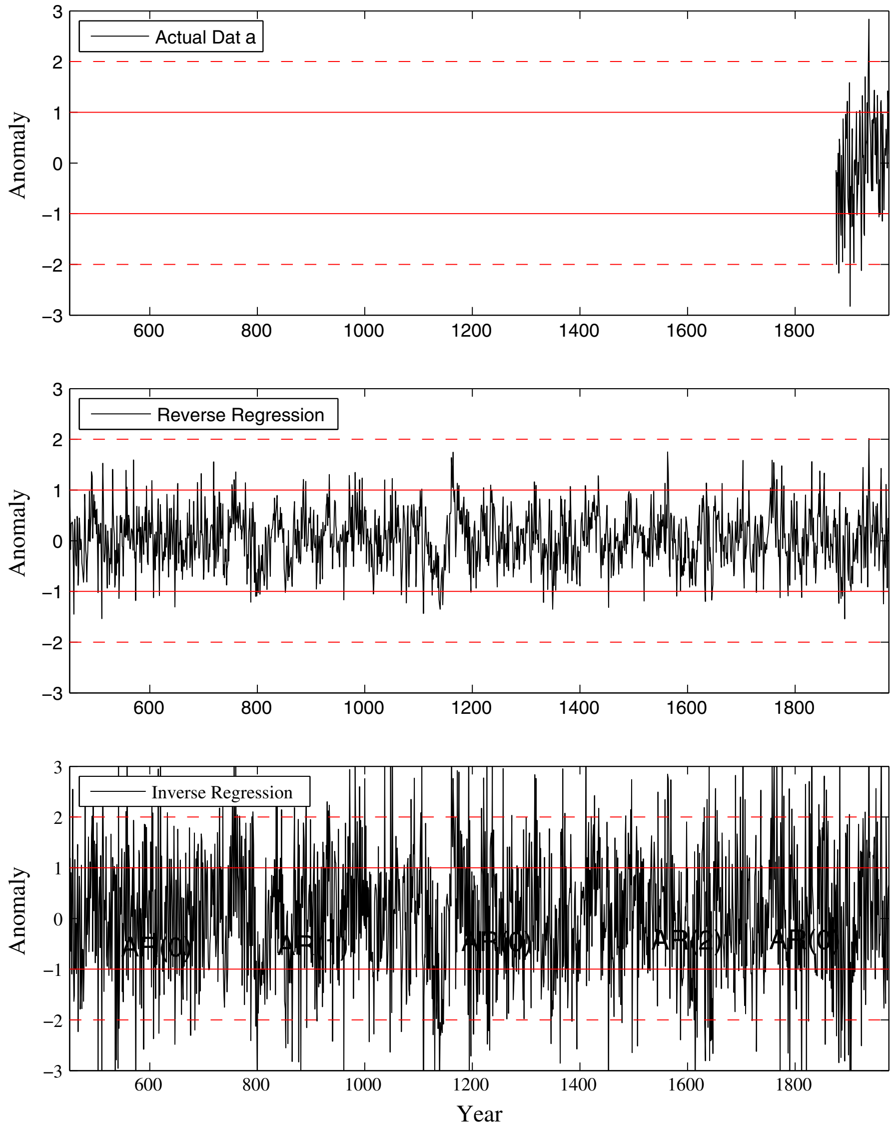

Fig. 3 Reconstruction of the Observed Temperature Record in Briffa et al. (1992). The *top panel* plots the the actual temperature data from Briffa et al. (1992), which are normalized to mean 0 and standard deviation 1. The *middle panel* displays the reconstructions from the reverse regression method. The *bottom panel* displays the reconstructions from the inverse regression method

5 Conclusion

In this paper we investigate the adequacy of a simple, intuitive but traditional reconstruction method of past climate fluctuations from paleoclimatic data series such as tree-ring index series, ice cores, pollen series, assuming that the response function relating the chosen paleoclimatic index to the climate measure is correctly specified and the temperature index is free of measurement error. We identify two problems with the traditional approaches, one with the estimation methodology, the other with the specification of the estimated relationship between climate and the paleoclimatic index. We demonstrate that reconstruction from the correctly specified inversion of the estimated response function is preferred, from a statistical point of view, to direct estimation of a transfer function relating the climate index to a paleoclimatic index.

We show that the specification of the transfer function should be determined by the specification of the response function. Whether leads and/or lags of the environmental factors should be included as predictors in the transfer function depends on the relative size of the parameters of the response function. Whether or not the transfer function is correctly specified, direct estimation of the transfer function is a classic example of the reverse regression problem, which causes the estimators to be biased in large as well as small samples.

Further, we have explored the underlying causes of the issue raised by Storch et al. (2004): The estimated degree of variation of reconstructed climate, which is central to the climate change debate, is underestimated by the standard approach as measured by the standard deviation of the reconstructions. Further, the method proposed in this study improves the accuracy of a climate reconstruction found in the literature (Briffa et al. 1992) on average by an order of 3 or 4, measured by mean absolute deviations, in our Monte Carlo experiment.

In addition, the misspecification of the transfer function results in overestimation of the order of the autoregressive process of the reconstructed series. The reconstructions by the direct estimation of a misspecified transfer function yield consistently erroneous and positive orders of the autoregressive process in Monte Carlo experiments and the empirical example. As a result of misspecification of the transfer function, and the bias induced by the reverse regression estimation procedure, fluctuations of the reconstructed climate series are underestimated. Inferences from such series regarding the existence of abnormalities in particular periods, including the most recent period, are unreliable.

The quality of the paleoclimatic data series, the specification of the response relation, and its explanatory power, are of course crucial issues in reconstruction of the history of climate change. What we show here is that even if the data are excellent, and the response relation is correctly specified and has high explanatory power, direct estimation of a correctly or incorrectly specified transfer function can produce highly unreliable information about the history of climate. The inversion of the estimated response function is the preferred method for reconstruction of climate history. It generally generates more reliable information, given the quality of the available data and the specification of the response relation. It will inevitably tend to overstate the variation of climate due to the errors in the noise component in the underlying response function, but the bias is small when the underlying response function has high explanatory power. If the underlying response function is noisy, neither method provides reconstructions of sufficient quality.

This article mainly attempts to illustrate the problem with reverse regression in the paleoclimate reconstruction, but it by no means indicates that inverse regression is the optimal method. Indeed, the Bayesian hierarchical modeling (BHM) framework has many advantages except that it often requires heavy computation. The BHM typically sets the response function at one level, and then it can include other information from other sources, such as the dynamics of the physical process, climate models or energy balance equations (Li et al. 2010; Tingley and Huybers 2010). In addition, Tingley and Li (2012) showed that the inverse regression becomes unstable when the response function is noisy, while the BHM can potentially ameliorate this problem by introducing only a weakly informative prior. Recently, Tingley et al. (2012) provides a detailed review of the paleoclimate reconstruction and discusses the modeling of each hierarchy and challenges when using a BHM model for the reconstruction.

Acknowledgments Auffhammer thanks the Giannini Foundation for support. Lis research was partially supported by the NSF Grant DMS-1007686. The authors thank the Guest Editor and two anonymous referees for their helpful comments and suggestions.

References

Ammann C, Genton M, Li B (2010) Technical note: correcting for signal attenuation from noisy proxy data in climate reconstructions. Clim Past 6:273–279

Bradley R, Jones P (1992) Climate since AD 1500. Routledge, London

Briffa K, Bartholin T, Eckstein D, Jones P, Karlén W, Schweingruber F, Zetterberg P (1990) A 1,400-year tree-ring record of summer temperatures in Fennoscandia. Nature 346(6283):434–439

Briffa K, Jones P, Bartholin T, Eckstein D, Schweingruber F, Karlén W, Zetterberg P, Eronen M (1992) Fennoscandian summers from ad 500: temperature changes on short and long timescales. Clim Dyn 7(3):111–119

Briffa K, Osborn T, Schweingruber F, Harris I, Jones P, Shiyatov S, Vaganov E (2001) Low-frequency temperature variations from a northern tree ring density network. J Geophys Res 106(D3):2929–2942

Briffa K, Schweingruber F, Jones P, Osborn T, Harris I, Shiyatov S, Vaganov E, Grudd H (1998) Trees tell of past climates: but are they speaking less clearly today. Philos Trans R Soc Lond B 353:65–73

Brown PJ (1993) Measurement, regression and calibration. Oxford University Press, New York

Brynjarsdóttir J, Berliner LM (2011) Bayesian hierarchical modeling for temperature reconstruction from geothermal data. Ann. Appl. Stat. 5(2B):1328–1359

Christiansen B (2011) Reconstructing the nh mean temperature: Can underestimation of trends and variability be avoided? J Clim 24:674–692

Cook ER, Briffa KR, Jones PD (1994) Spatial regression methods in dendroclimatology: a review and comparison of two techniques. Int J Climatol 14(4):379–402

Eisenhart C (1939) The interpretation of certain regression methods and their use in biological and industrial research. Ann Math Stat 10:162–186

Fritts H (1991) Reconstructing large-scale climatic patterns from tree-ring data: a diagnostic analysis. University of Arizona Press, Tuscon

Graumlich L (1991) Subalpine tree growth, climate, and increasing CO_2: an assessment of recent growth trends. Ecology 72(1):1–11

Graumlich L, Brubaker L (1986) Reconstruction of annual temperature(1590–1979) for Longmire, Washington, derived from tree rings. Quat Res 25(2):223–234

Hamilton J (1994) Time series analysis. Princeton University Press, Princeton

Hansen J, Lebedeff S (1987) Global trends of measured surface air temperature. J Geophys Res 92(D11):13345–13372

Haslett J, Whiley M, Bhattacharya S, Salter-Townshend M, Wilson S, Allen J, Huntley B, Mitchell F (2006) Bayesian palaeoclimate reconstruction. J R Stat Soc Ser A (Stat Soc) 169(3):395–438

Hughes M, Brown P (1992) Drought frequency in central California since 101 BC recorded in giant sequoia tree rings. Clim Dyn 6(3):161–167

Jones P, Briffa K, Barnett T, Tett S (1998) High-resolution palaeoclimatic records for the last millennium: interpretation, integration and comparison with general circulation model control-run temperatures. The Holocene 8(4):455

Jones P, Briffa K, Osborn T, Lough J, Van Ommen T, Vinther B, Luterbacher J, Wahl E, Zwiers F, Mann M et al (2009) High-resolution palaeoclimatology of the last millennium: a review of current status and future prospects. The Holocene 19(1):3–49

Jones P, Hulme M (1997) The changing temperature of Central England. In: Hulme M, Barrow E (eds) Climate of the British Isles: past, present and future. Routledge, London, pp 173–196

Jones PD, Osborn TJ, Briffa KR (2001) The evolution of climate over the last millennium. Science 292(5517):662–667

Jones P, Wigley T, Wright P (1986) Global temperature variations between 1861 and 1984. Nature 322(6078):430–434

Klepper S, Leamer E (1984) Consistent sets of estimates for regressions with errors in all variables. Econometrica 52(1):163–184

Krutchkoff RG (1967) Classical and inverse regression methods of calibration. Technometrics 9:425–439

LaMarche V Jr (1974) Paleoclimatic Inferences from long tree-ring records: intersite comparison shows climatic anomalies that may be linked to features of the general circulation. Science 183(4129):1043

Lee T, Zwiers F, Tsao M (2008) Evaluation of proxy-based millennial reconstruction methods. Clim Dyn 31(2):263–281

Li B, Nychka D, Ammann C (2007) The 'hockey stick' and the 1990s: a statistical perspective on reconstructing hemispheric temperatures. Tellus A 59(5):591–598

Li B, Nychka D, Ammann C (2010) The value of multi-proxy reconstruction of past climate. J Am Stat Assoc 105(491):883–911

Mann M, Bradley R, Hughes M (1998) Global-scale temperature patterns and climate forcing over the past six centuries. Nature 392:779–787

Mann M, Bradley R, Hughes M (1999) Northern hemisphere temperatures during the past millennium: inferences, uncertainties, and limitations. Geophys Res Lett 26(6):759–762

Mann ME (2002) The value of multiple proxies. Science 297(5586):1481–1482

Mann ME, Rutherford S, Wahl E, Ammann C (2005) Testing the fidelity of methods used in proxy-based reconstructions of past climate. Journal of Climate 18:4097–4107

Mann ME, Zhang Z, Hughes MK, Bradley RS, Miller SK, Rutherford S, Ni F (2008) Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. Proc Natl Acad Sci 105(36):13252–13257

Mann M, Rutherford S (2002) Climate reconstruction using 'pseudoproxies'. Geophys Res Lett 29(10):139–141

McShane BB, Wyner AJ (2011) A statistical analysis of multiple temperature proxies: are reconstructions of surface temperatures over the last 1000 years reliable? Ann Appl Stat 5(1):5–44

National Research Council (1998) Decade-to-Century-Scale climate variability and change: a science strategy. National Academies Press, Washington

National Research Council (2006) Surface Temperature Reconstructions for the last 2000 years. National Academies Press, Washington, DC

Newey W, West K (1994) Automatic lag selection in covariance matrix estimation. Rev Econ Stud 61(4):631–653

Salzer MW, Hughes MK, Bunn AG, Kipfmueller KF (2009) Recent unprecedented tree-ring growth in bristlecone pine at the highest elevations and possible causes. PNAS 106(48):2034820353

Schneider T (2001) Analysis of incomplete climate data: estimation of mean values and covariance matrices and imputation of missing values. J Clim 14:853–871

Scuderi LA (1993) A 2000-year tree ring record of annual temperatures in the Sierra Nevada mountains. Science 259(5100):1433–1436

Seater J (1993) World temperature-trend uncertainties and their implications for economic policy. J Bus Econ Stat 11(3):265–277

Smerdon JE (2012) Climate models as a test bed for climate reconstruction methods: pseudoproxy experiments. WIREs Clim Change 3:63–77

Till C, Guiot J (1990) Reconstruction of precipitation in Morocco since 1100 AD based on Cedrus atlantica tree-ring widths. Quat Res 33:337–351

Tingley M, Huybers P (2010) A Bayesian algorithm for reconstructing climate anomalies in space and time. Part 1: development and applications to paleoclimate reconstruction problems. J Clim 23(10):2759–2781

Tingley MP, Craigmile PF, Haran M, Li B, Mannshardt-Shamseldin E, Rajaratnam B (2012) Piecing together the past: statistical insights into paleoclimatic reconstructions. Quat Sci Rev 35:1–22

Tingley MP, Li B (2012) Comments on “reconstructing the nh mean temperature: Can underestimation of trends and variability be avoided? by Bo Chirstiansen. J Clim 25:3441–3446

von Storch H, Zorita E, Jones JM, Dimitriev Y, Gonzalez-Rouco F, Tett SFB (2004) Reconstructing past climate from noisy data. Science 306(5696):679–682

Yoo SJ, Wright BD (2000) Persistence of growth variation in tree-ring chronologies. For Sci 46(4):507–520

Maximilian Auffhammer is Associate Professor of Department of Agricultural and Resource Economics, George M. Pardee, Jr. Family Chair in International Sustainable Development and Director of International Area Studies Teaching Program at University of California at Berkeley.

Bo Li is Associate Professor of Department of Statistics at University of Illinois at Urbana-Champaign.

Brian Wright is Professor and Chair of Department of Agricultural and Resource Economics at University of California at Berkeley.

Seung-Jick Yoo is the President of Greenhouse Gas Inventory & Research Center of Korea.