Mathematical Geology, Vol. 37, No. 2, February 2005 (© 2005)
DOI: 10.1007/s11004-005-1305-6

Estimating Climatic Timeseries From Multi-Site Data Afflicted With Dating Error[1]

Craig Loehle[2]

Timeseries of estimated temperature have been combined to create global or hemispheric climate series over periods exceeding 1000 yr. The data used in these studies, however, may be subject to dating errors. It is shown that when timeseries with dating error are combined, the noise in the data smoothes periodic signals but leaves linear trends intact. This means that the effect of dating error of sample data in a timeseries reconstruction is to smooth out any signals (waves, cycles) that may be present. The purpose of this study was to develop signal extraction methods that will work for this type of historical data. The method used was nonlinear estimation of sample series where dating error has been added by Monte Carlo sampling. Several algorithms were tested for handling the dating error problem. Results were that using nonlinear model fitting, the periods of signals can be identified even from the averaged data. In a second stage of the estimation procedure, the cycle magnitudes can be estimated. Very good fits were achieved for two example cases. Temperature estimation error (white noise due to the use of proxies) was also considered and the method was extended to cover this case with quite good results. Using the new estimation methods, the information inherent in multiple series can be used to overcome the problem of dating error.

KEY WORDS: Kalman filter, nonlinear estimation, proxy data, climate reconstruction.

INTRODUCTION

Historical climate series are of great interest for many reasons, including studies of the ENSO and other climate oscillations, solar-climate interactions (Hu *et al.*, 2003), and responses of ecosystems to past climates. To these ends, a number of global or hemispherical climate series have been constructed (Crowley, 2000; Crowley and Lowery, 2000; Jones, 1998; Mann, Bradley, and Hughes, 1998, 1999; Overpeck and others, 1997).

In a typical estimation problem more data improves the estimation. When reconstructing historical climate data, however, this usual result may not hold. This is because the critical independent variable, time, can only be estimated, and may be subject to large errors. It is suggested that these dating errors make problematic

[1]Received 14 January 2004; accepted 4 August 2004.
[2]National Council for Air and Stream Improvement, Inc. (NCASI), 552 S Washington St., Ste. 224, Naperville, Illinois 60540; e-mail: cloehle@ncasi.org

the averaging of multiple timeseries (e.g., to create a global series). In typical instrumental climate series, measurement error and local climate stochasticity constitute white noise, which may obscure the presence of underlying trends or cycles. Various filtering methods exist for removing this white noise and uncovering the fundamental system behaviors (Mann and Lees, 1996). In pre-instrumental climate reconstructions, however, dates are not known with precision and temperatures are estimated from some proxy such as glacier extent, aquatic organism abundance, or tree-ring width. We may consider sediment data as a typical example. ^{14}C measurements at control points are used as a basis for interpolating dates as a function of estimated sedimentation rates (Anderson, 2000; Pienitz, Smol, and MacDonald, 1999). Several sources of dating error are introduced during dating calibration. ^{14}C measurements are subject to laboratory errors. There may be varied and unknown mixing of sedimentary layers that introduces smearing of the time signal or the introduction of younger (by infusion) or older (by resuspension or erosion) carbon into the sample. Diatom-temperature calibration curves are subject to errors due to changing watershed geochemistry over time (Anderson, 2000). The sedimentation rate curves used for interpolation are not precise. The radiocarbon date conversion to calendar date is not precise because atmospheric ^{14}C has varied over time due to fluctuating cosmic ray intensities (Vogel, 2002), and the exact pattern of these fluctuations has not been worked out, and because of changes over time in ocean ventilation. When climate chronologies are determined from tree line, lake levels, or glacial activity, there is a problem of missing data (e.g., when a glacier destroys evidence of a previous glacier).

What is the magnitude of the dating error problem? The radiocarbon estimation error for recent samples ranges from 30 to 100 plus years (Kullman, 1998; Seppä and Weckström, 1999), with the conversion of radiocarbon years to calendar years adding more variability (Vogel, 2002). Radiocarbon dating errors tend to increase with the age of the sample (Seppä and Weckström, 1999; Vogel, 2002). For ice age period samples, dating can become quite difficult. For example, various studies using different methods estimated glacial stage Termination II at $142,000 \pm 3000$, $127,000 \pm 6000$, and $135,000 \pm 2500$ yr BP (Karner and Muller, 2000). Goslar and others (2000) showed that during the Late Glacial (15,000–11,000 BP) the dates for key events (e.g., the Younger Dryas) in various records differed by 100–400 yr. Mann, Park, and Bradley (1995) found 100-yr lags between estimated regional peaks and troughs in temperature, even over a five-century record.

Annually dated series would seem to overcome these difficulties. Ice core data, for example, are highly resolved. However, ice does not leave such a clear record. Many years of ice can be melted and lost during a warm period. Ice layers can be compressed to the point where they can not be clearly identified. Rempel and others (2001) document anomalous diffusion in ice-sheet ice, which can displace impurities in the ice by 50 cm or more. These impurities (e.g., sulfuric acid spikes) are used to cross-date ice cores as well as to correlate ice dates with dates of outside events such as volcanic eruptions, although it must be remembered that the dates of such external events must be estimated also with some error. Tree rings would seem to be ideal since rings never vanish or melt. However, many trees produce false rings (multiple rings in a year) or have missing rings during bad years. This is particularly true of the trees that make the best climate indicators; those growing alone under harsh conditions. Going back a few hundred years, such problems are an annoyance but do not produce major dating errors, but over longer periods they accumulate. More seriously, since individual trees have limited lifespans (less than 1000 yr), longer series must be constructed by cross-dating multiple trees. This is done by matching up what appear to be similar major climate events such as very dry years. However, this is an art, and is subject to major errors. If cross-dating is done incorrectly, the series for that tree will be shifted by some unknown number of years. All of these errors accumulate as longer series are constructed.

The effect of these dating errors is to create a randomization in time of the data. Methods exist for estimating periodic components of a signal in a single detailed timeseries in which noise results from a first-order autoregressive process (Mann and Lees, 1996; Von Storch and Zwiers, 1999; Wilks, 1995). However, noise is not due to an autoregressive process when it is due to an dating estimation error. When timeseries afflicted by temporal noise are combined to obtain a regional or global average, unexpected problems occur which are not even mentioned in standard texts (e.g., Von Storch and Zwiers, 1999; Wilks, 1995). Consider a timeseries with measurement error (white noise) but no dating error. The more series are combined, the more clearly the signal (cycles, trends) can be seen, exactly according to the Central Limit Theorem. In contrast, the mean of n series with dating error is exactly equivalent to the convolution of the underlying signal with a time smoothing function (Grewal and Andrews, 1993). An example is when a Gaussian filter is applied to remove white noise from a timeseries (Fig. 1). But a Gaussian filter applied to the signal itself actually suppresses the signal. In the limit as $\sigma \rightarrow \infty$, a Gaussian filter completely eliminates a sinusoidal signal by averaging across the entire series. A Gaussian filter will leave a linear trend intact because the mean of random points to the left and right of a point average out to the true value at the point, in the limit. The result is that a timeseries constructed as the mean of n series with temporal noise will have a much lower amplitude than evidenced by the individual series (Fig. 2). For large dating error, only linear trends will be left in the averaged series. The averaged series may thus be giving a misleading impression of constancy that results purely from the averaging process for timeseries with temporal noise.

TWO-STAGE ESTIMATION PROCEDURE

If, as shown above, it is not valid to simply average across timeseries when dating errors exist, how do we obtain an overall estimated global record? Spectral

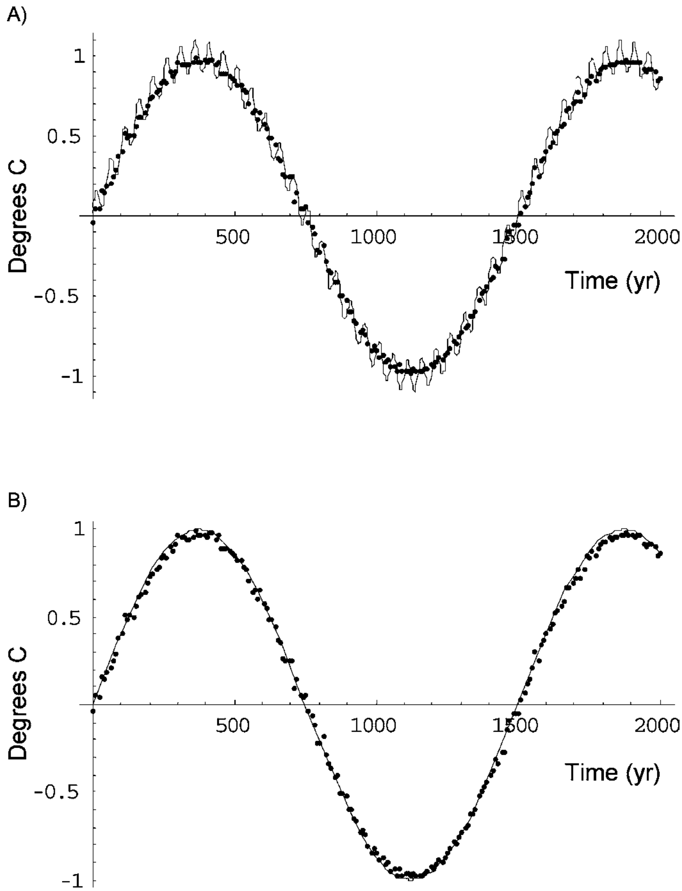

Figure 1. Example of Gaussian filter applied to a series with white noise. Points show result of applying a Gaussian filter to the signal. (A) Noisy signal with overlaid filtered points. (B) Smoothed points compared to true signal showing how true signal is recovered.

methods only give us an indication of the presence of cycles, with no way to recover the true cycle amplitudes. A new two-stage method is presented for detecting cyclic components of a signal from multiple series subject to dating error. The approach is developed sequentially, first addressing a simple one-cycle model with only dating error, then a two-cycle model with only dating error, then the approach is modified to account for temperature estimation error.

In stage one, we assume a cyclic (harmonic) climate signal, use this to generate random series, and then fit this model directly to the timeseries using least-squares estimation. Harmonic analysis of climate data has been done previously

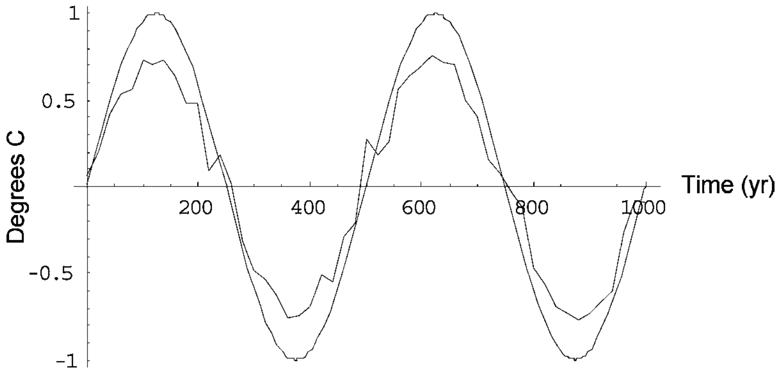

Figure 2. True signal compared to mean of 50 series sampled every 20 yr with $\sigma = 50$ showing damped peaks.

(e.g., Damon and Jirikowic, 1992; Loehle, 2004) This approach enables all parameters of climate cycles (offset, period, and amplitude) to be estimated without assuming evenly spaced data and without various other simplifying assumptions (e.g., Ferraz-Mello, 1981; Von Storch and Zwiers, 1999; Wilks, 1995). It is assumed here that there is in fact a periodic signal in the data. Given a candidate model of suitable form, we can form the sum of squares across all timeseries. For the first model considered here (Fig. 3A), we wish to find

$$\min(z) = \sum_i \sum_t ((c \sin[(a+t)2\pi/b]) - S_{it})^2 \tag{1}$$

by some suitable method, where S is the sample temperature for time t at location i, a is the date offset, b is the period, and c is the amplitude. The parameter set a,b,c was fit from Equation (1) using nonlinear minimization procedures available in *Mathematica*. Ten random series with dating error at each sample (with $\sigma = 60$ yr) were generated with samples every 20 yr over a 2000 yr horizon. An example plot of the best-fit model vs. the "true" model (Fig. 3A) shows that the proper period has been determined (498.8 yr vs. 500 yr) but the amplitude is much too small. The cycle offset (parameter c) is also very close to the true value (−0.02 yr vs. 0 yr). Fitting the model to the average of the 10 series at each time gives the same result because the methods are mathematically identical.

A bootstrap procedure was used (Efron and Tibshirani, 1993) to estimate method reliability. The model estimation procedure was repeated 400 times and summary statistics computed. The mean parameter values for offset and period with 95% confidence intervals (−0.58 ± 4.15 yr and 499.58 ± 0.94 yr) are very close to and bound the true values (0,500). However, the amplitude (0.75 deg. C) is much too low, due to the time averaging effect of dating error.

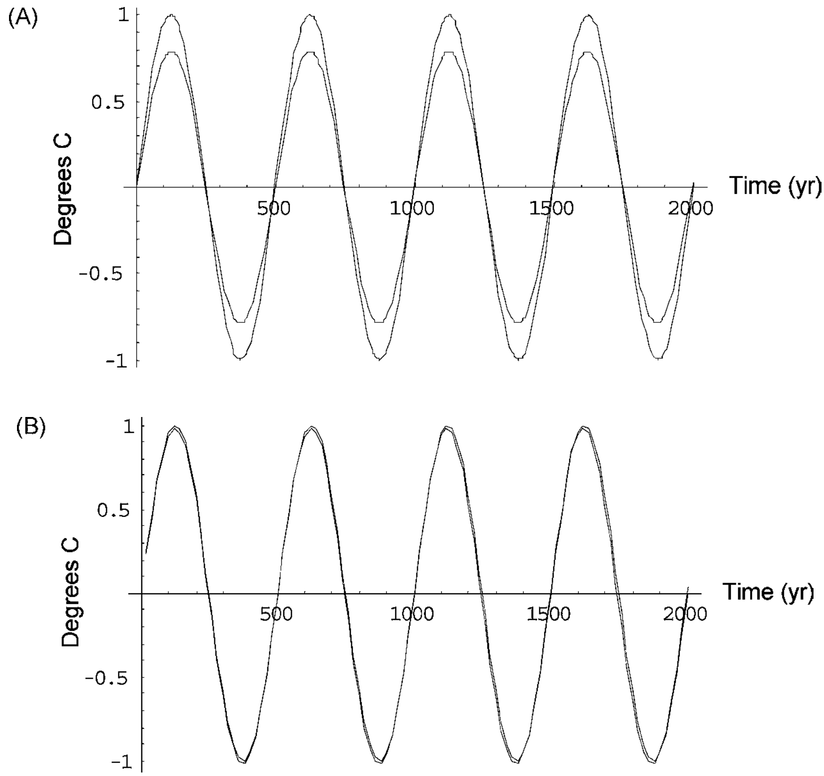

Figure 3. (A) True signal compared to best-fit curve from mean of 10 sampled series at 20 yr intervals with $\sigma = 60$ yr. Best-fit curve is damped compared to true signal. (B) Second stage estimation of sinusoidal signal showing nearly perfect fit.

The second stage of the estimation procedure recovers the amplitudes. The offset and period (a,b) from stage 1 are retained for each run. For each of the 10 series, the range is determined by the difference between the maximum and minimum temperature over the series. The mean range (1.996) over all series is used to adjust the fitted model. The rationale for the adjustment is that when there is just dating error the law of large numbers dictates that the true max and min will be found when many samples are taken. The result (Fig. 3B) is an almost perfect fit.

The bootstrap was again applied after the stage 2 adjustment of amplitude, with 400 replicates. The mean amplitude is now 0.999, with 95% confidence interval ± 0.00775 and $R^2 = 0.998$. Thus in the presence of only dating error, the true series can be recovered almost perfectly with this technique.

It could be argued that the above example is too simplistic. A second case was tested with two component waves, periods 300 yr with 80 yr offset and 800 yr with no offset, both with amplitude 1.0 deg. C (Fig. 4A). Ten series at 20 yr intervals were again generated, again with $\sigma = 60$ yr (Fig. 4B). The first stage estimation gives a good fit over the 400 replicates, with average offsets (79.85 $\pm$ 0.77 and -0.74 ± 1.0 yr) close to and bracketing the true values (80 and 0 yr), and periods (299.98 $\pm$ 0.18 and 799.56 $\pm$ 0.63 yr) close to and bracketing the true values (300 and 800 yr), but the amplitudes are again too small (Fig. 4C), with a mean over the 400 replications of 0.455 and 0.897 vs 1.0 and 1.0, respectively, and an R^2 of 0.84.

Correction of the amplitudes in stage two is done as follows. Peaks and troughs are found by numerically computing turning points of the estimated model. At each peak of the fitted curve, the largest value across all 10 series at that date is found. At each trough, the lowest value over all 10 series at that date is found. This forms a set of data of length equal to the number of peaks and troughs. The two-cycle model is then fit with only the amplitudes free to vary and the other parameters taken from stage one, with the goal being a least-square fit across the values at the peaks and troughs. This forces the model to approximate the underlying series, with an almost perfect result (example fit in Fig. 4D). This works because on average some of the sample series will have been sampled with a correct date, and have a correct magnitude at that date. By selecting the peaks and troughs, the procedure corrects for the underestimation resulting from fitting the mean series. Over the 400 replications of the bootstrap, the average R^2 value increased to 0.986, and the two peaks increased to 1.095 $\pm$ 0.003 and 1.007 $\pm$ 0.003. There is a slight positive bias but overall the result is surprisingly good.

The cases considered above assumed that measurement error is zero, which of course is never true. Measurement error is thus considered next. While measurement error does not affect the mean of n series at any given date, it does expand the range of values over the n series likely to be observed at any given date. In the above algorithm, the maximum or minimum found at a peak or trough, respectively, was used to adjust the cycle amplitudes in a second stage fitting. If this approach is tried in the presence of measurement error, the amplitudes are significantly overestimated (not shown) and the R^2 values sometimes actually get worse. Since dating error alone causes an underestimate of amplitudes and measurement error causes an overestimate, it seems logical that a combined estimate might prove useful. If d_1 is the mean data value at a peak time, and d_2 is the mean over the n series of the maximum of the value at that date and of the two neighboring dates, the following weighting was found to produce very good results (Table 1; Figs. 5 and 6) though it is not claimed to be universal:

$$d = (d_1 + 3d_2)/4$$

where the d values at each peak and trough become the fit criteria for stage 2 with offset and period held constant, as before. Two levels of dating error were tested,

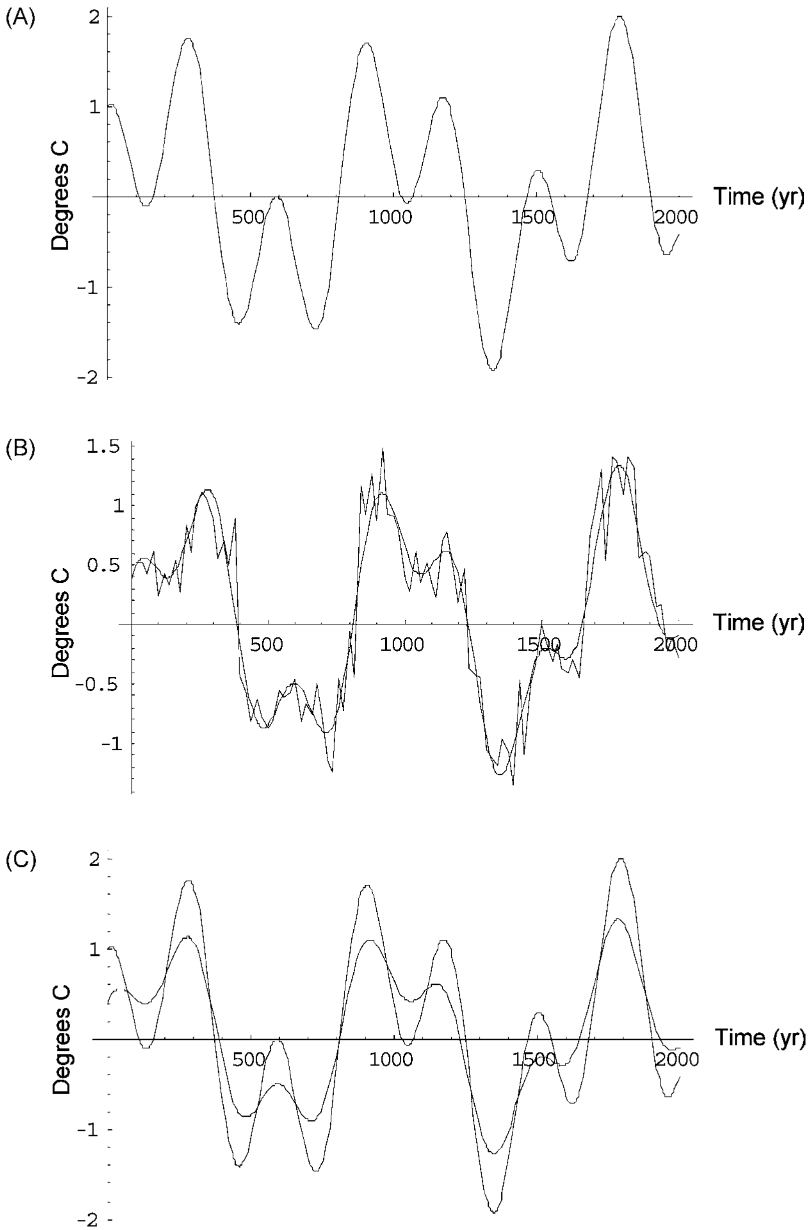

Figure 4. (A) Two component signal for testing. (B) Best-fit two-component model compared to the mean of 10 randomly sampled series with dating error. (C) True signal vs. best-fit (damped curve) from stage one. (D) Best-fit model vs. true signal after second stage estimation.

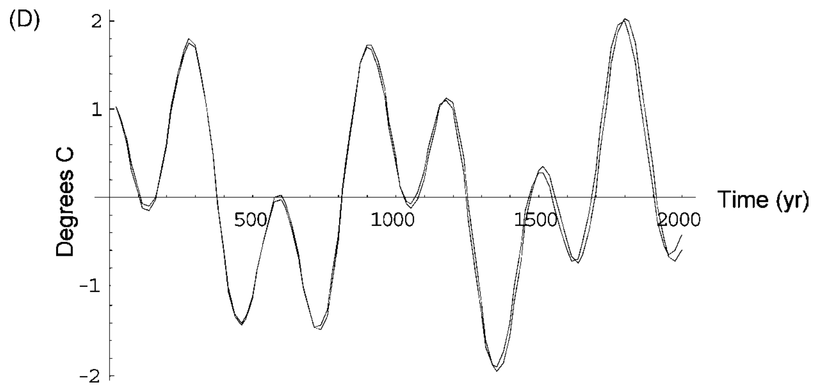

Figure 4. Continued.

with three levels of measurement error. Bootstrap estimation of accuracy and bias of accuracy and bias were performed with 400 replications for each parameter combination, with 10 series generated for each trial, as before.

For all cases, the offset and period were estimated without bias and with high precision (Table 1). As before, stage 1 underestimated the amplitudes of the climate cycles. After the stage 2 adjustment, all amplitudes were either very close to 1 or nearly identical to 1, and R^2 values were exceptionally good (Table 1).

Some extensions of the method were tested. It was found that if there is an underlying trend, the method can be applied to the detrended data or by directly adding a trend (e.g., linear trend) coefficient to the model. In either case, the method worked with the same accuracy as above. A second case considered was lognormal error distribution for sampling error. Using this distribution, a positive and skewed error term is added to all samples, in addition to the temporal estimation error. Although the coefficient of a lognormal can not be compared directly to those used in Table 1, with roughly similar error terms the algorithm used here was able to cut bias in half or better from stage 1 to stage 2 and improve R^2 from an average of 0.90–0.95 over the cases considered (not shown). The approach should work equally well for sawtooth or other periodic models without modification, although testing all possible model types is outside the scope of this study.

Examining Figure 5A, we see that noise has created a very irregular signal that looks nearly random. However, it is clear from Figure 5B that the true signal can be recovered even in this extreme case. It is important to note that signal smoothing procedures will not duplicate the results obtained here. It is only by utilizing the information inherent in multiple series that the problems of dating error and sampling error can be overcome. The new methods presented here are not applicable to real data sets without some further work. In particular, it is necessary to develop a greater understanding of the nature of dating error in constructed

Table 1. Model Estimation with a Two-Stage Algorithm When Both Dating Error and Sampling Error are Present

S.D. of	S.D. of	Stage 1							Stage 2		
age (yrs)	sample (°C)	Cycle 1 offset	Cycle 1 period	Cycle 1 amplitude	Cycle 2 offset	Cycle 2 period	Cycle 2 amplitude	Average R^2	Cycle 1 amplitude	Cycle 2 amplitude	Average R^2
40	0.1	80.03 ± 0.42	300.01 ± 0.10	0.71 ± 0.002	0.31 ± 0.73	800.09 ± 0.47	0.953 ± 0.0024	0.953	0.952 ± 0.003	0.98 ± 0.003	0.996
40	0.2	79.91 ± 0.43	300.0 ± 0.11	0.70 ± 0.0024	0.28 ± 0.81	800.21 ± 0.51	0.952 ± 0.0025	0.952	0.990 ± 0.003	0.978 ± 0.003	0.996
40	0.3	80.20 ± 0.45	300.05 ± 0.10	0.71 ± 0.0024	0.21 ± 0.80	800.19 ± 0.51	0.952 ± 0.0025	0.953	1.04 ± 0.0036	0.982 ± 0.0034	0.995
60	0.1	79.64 ± 0.77	299.99 ± 0.19	0.457 ± 0.0031	0.21 ± 1.06	800.3 ± 0.67	0.894 ± 0.0032	0.841	0.83 ± 0.004	0.964 ± 0.0037	0.977
60	0.2	79.73 ± 0.78	300.00 ± 0.19	0.456 ± 0.0033	0.49 ± 1.07	800.26 ± 0.69	0.890 ± 0.0039	0.841	0.86 ± 0.0046	0.96 ± 0.0039	0.981
60	0.3	80.05 ± 0.82	300.03 ± 0.21	0.46 ± 0.0033	0.18 ± 1.08	800.21 ± 0.69	0.896 ± 0.0032	0.84	0.894 ± 0.0048	0.964 ± 0.0045	0.984

Note. The parameters offset, period, and amplitude are (80, 300, 1) for cycle 1 and (0, 800, 1) for cycle 2 in the model. Ten series were generated for each of 400 replicates at each combination of error terms. Values are shown ± 95% bootstrap confidence intervals.

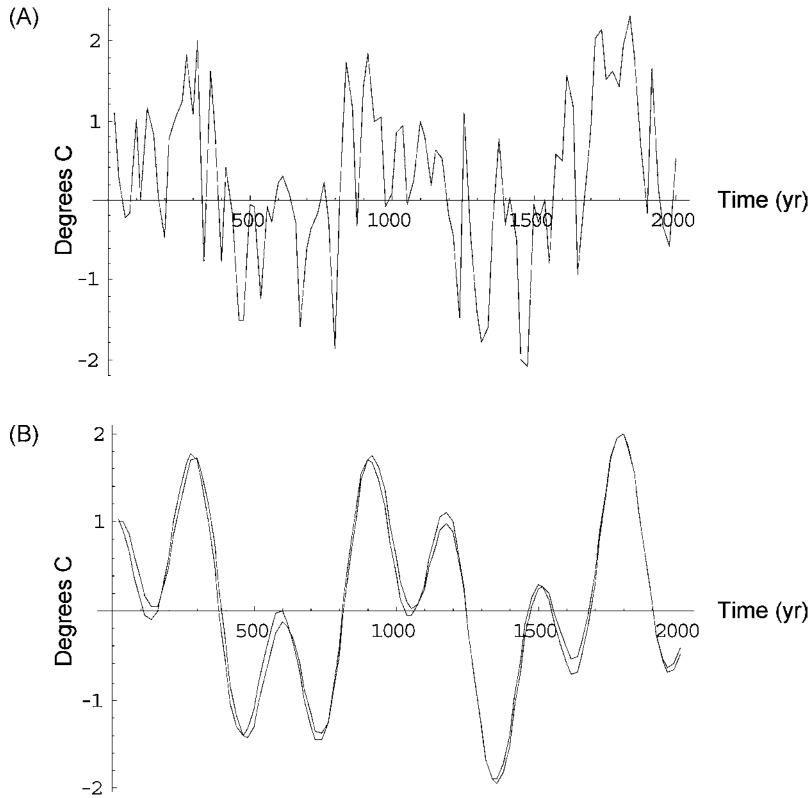

Figure 5. (A) Example data series with both dating error ($\sigma_d = 60$ yr) and sampling error ($\sigma_s = 0.3$). (B) After two-stage estimation, showing extremely good fit of the model.

series, because this topic has been largely ignored in studies to date. There may be local temporal correlation in errors, for example, that is not considered here, and dating error could differ between series. Characteristics of sampling error also need to be explored. Sampling errors could be different with time or across samples, or be non-normally distributed. In addition to characterizing error in field data, the effect of these various complications on the methods presented here need exploration. If cycles are only quasi-periodic, an enhanced model may be necessary. Nevertheless, the methods presented show great promise for revealing patterns in historical data in spite of the complications caused by dating error.

CONCLUSIONS

The results of this analysis are relevant to questions concerning the climate at particular periods such as the Medieval Warm Period (MWP). The first

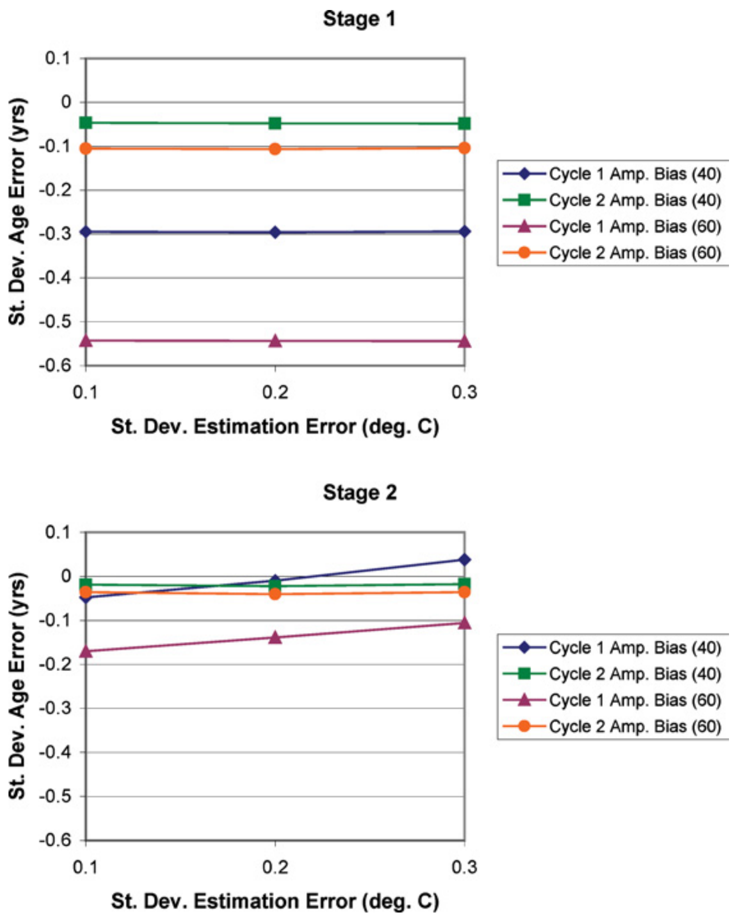

Figure 6. Bias reduction from stage 1 to stage 2 estimation for 2-cycle model at two levels of dating error. Shown is the mean result of 400 Monte Carlo simulations of sampling 10 time series with both dating error and estimation error with the given variances.

implication is that if there is dating error in samples, then the peaks indicating the MWP in series from different parts of the world would be expected to not line up, as has been observed (Bradley, Hughes, and Diaz, 2003). In contrast, if a looser definition of peak timing is used, evidence for the MWP is strong (Soon and Baliunas, 2003). Dating error would also lead to an underestimate of the height of the peak at the MWP (and other periods) for any regional to global estimate. This bears substantively on whether recent climate is anomalously warm. The extent to which past reconstructions are affected by dating error deserves examination because of the potential major impact on the conclusions from these studies.

It is important to characterize long-term global temperature trends so that recent climate changes can be put in the proper context. However, it is not sufficient to simply average historical data to produce a global or hemispheric timeseries, because dating error afflicts virtually all extant pre-instrumental reconstructions. There are two options available for obtaining statistically valid global or hemispheric timeseries. If better dating methods can be developed to reduce dating errors in proxy records, then simple averaging of series is valid. Given the many sources of dating error, this is a challenge. If, on the other hand, multiple series can be estimated, the estimation methods developed in this paper can be applied to identify trends and cycles in the historical record, even with dating errors. In either case, the analysis is important and work on climate trend reconstructions should continue.

ACKNOWLEDGMENT

I am grateful for helpful comments provided by Jeffrey Gove.

REFERENCES

Anderson, N. J., 2000, Diatoms, temperature and climate change: Eur. J. Phycol., v. 35, p. 307–314.

Bradley, R. S., Hughes, M. K., and Diaz, H. F., 2003, Climate in medieval time: Science, v. 302, p. 404–405.

Crowley, T. J., 2000, Causes of climate change over the past 1000 years: Science, v. 289, p. 270–277.

Crowley, T. J., and Lowery, T. S., 2000, How warm was the Medieval warm period?: Ambio, v. 29, p. 51.

Damon, P. E., and Jirikowic, J. L., 1992, The sun as a low-frequency harmonic oscillator: Radiocarbon 34, p. 199–205.

Efrom, B., and Tibshirani, R. J., 1993, An introduction to the bootstrap: Chapman & Hall, New York, 436 p.

Ferraz-Mello, S., 1981. Estimation of periods from unequally spaced observations: Astronom. J., v. 86 p. 619–624.

Goslar, T., Arnold, M., Tisnerat-Laborde, N., Czernik, J., and Wickowski, K., 2000, Variations of Younger Dryas atmospheric radiocarbon explicable without ocean circulation changes: Nature, v. 403, p. 877–880.

Grewal, M. S., and Andrews, A. P., 1993, Kalman filtering: Theory and practice: Prentice-Hall, Englewood Cliffs, New Jersey. 381 p.

Hu, F. S., Kaufman, D., Yoneji, S., Nelson, D., Shemesh, A., Huang, Y., Tian, J., Bond, G., Clegg, B., and Brown, T., 2003, Cyclic variation and solar forcing of holocene climate in the Alaskan subarctic: Science, v. 301, p. 1890–1893.

Jones, P., 1998, It was the best of times, it was the worst of times: Science, v. 280, p. 544–545.

Karner, D. B., and Muller, R. A., 2000, A causality problem for Milankovitch: Science, v. 288, p. 2143–2144.

Kullman, L., 1998, Holocene history of the forest-alpine tundra ecotone in the Scandes Mountains (central Sweden): The New Phytologist, v. 108, p. 101–110.

Loehle, C., 2004, Climate change: detection and attribution of trends from long-term geologic data: Ecol. Model., v. 171, p. 433–450.

Mann, M. E., and Lees, J. M., 1996, Robust estimation of background noise and signal detection in climatic timeseries: Clim. Change, v. 33, p. 409–445.

Mann, M. E., Bradley, R. S., and Hughes, M. K., 1998, Global-scale temperature patterns and climate forcing over the past six centuries: Nature, v. 392, p. 779–787.

Mann, M. E., Bradley, R. S., and Hughes, M. K., 1999, Northern hemisphere temperatures during the past millennium: Inferences, uncertainties, and limitations: Geophys. Res. Lett., v. 26, p. 759–762.

Mann, M. E., Park, J., and Bradley, R. S., 1995, Global interdecadal and century-scale climate oscillations during the past five centuries: Nature, v. 378, p. 266–270.

Overpeck, J., Hughen, K., Hardy, D., Bradley, R., Case, R., Douglas, M., Finney, B., Gajewski, K., Jacoby, G., Jennings, A., Lamoureux, S., Lasca, A., MacDonald, G., Moore, J., Retelle, M., Smith, S., Wolfe, A., and Zielinski, G., 1997, Arctic environmental change of the last four centuries: Science, v. 278, p. 1251–1256.

Pienitz, R., Smol, J. P., and MacDonald, G. M., 1999, Paleolimnological reconstruction of Holocene climatic trends from two boreal tree-line lakes, Northwest Territories, Canada: Arctic, Antarctic Alpine Res., v. 31, p. 82–93.

Rempel, A. W., Waddington, E. D., Wettlaufer, J. S., and Worster, M. G., 2001, Possible displacement of the climate signal in ancient ice by premelting and anomalous diffusion: Nature, v. 411, p. 568–571.

Seppä, H., and Weckström, J., 1999, Holocene vegetational and limnological changes in the Fennoscandian tree-line area as documented by pollen and diatom records from Lake Tsulbmajavri, Finland: Ecoscience, v. 6, p. 621–635.

Soon, W. H., and Baliunas, S., 2003, Proxy climatic and environmental changes of the past 1000 years: Clim. Res. v. 23, p. 89–110.

Vogel, J. C., 2002, Secular variations in carbon-14 and their geophysical implications: South African J. Sci., v. 98, p. 154–160.

Von Storch, H., and Zwiers, F. W., 1999, Statistical analysis in climate research: Cambridge University Press, Cambridge, 484 p.

Wilks, D. S., 1995, Statistical methods in the atmospheric sciences: Academic Press, New York. 467 p.