

**Lucy, D. Robertson, I. Aykroyd, R.G. & Pollard, A.M. (2008)
Estimates of uncertainty in the prediction of past climatic
variables *Applied Geochemistry* 23, 2961–2965.**

Estimates of uncertainty in the prediction of past climatic variables

D. Lucy¹, I. Robertson^{2*}, R.G. Aykroyd³ and A.M. Pollard⁴

¹ *Department of Mathematics and Statistics, Lancaster University, Lancaster LA1 4YF, UK.*

² *Department of Geography, University of Wales Swansea, Swansea SA2 8PP, UK.*

³ *Department of Statistics, University of Leeds, Leeds LS2 9JT, UK.*

⁴ *Research Laboratory for Archaeology and the History of Art, South Parks Road, Oxford OX1 3QY, UK.*

* corresponding author:

Tel + 44 1792 513693

Fax + 44 1792 295955

Email: i.robertson@swansea.ac.uk

Abstract

Fitting a regression line to a set of measurements to investigate the relationship between a proxy estimate of past climate and known climatic parameters is a routine procedure. It is generally accepted that the higher the correlation between parameters, the more reliable the reconstruction. However, there is a lack of published work upon what correlation is the minimum acceptable value. Simulated data was used to demonstrate that the relationship between proxy values and the climatic data are adversely affected by falling correlation, to the point where, in a training set consisting of 100 pairs of temperature and tree-ring proxies, the mean 95% confidence interval width for the

reconstructed temperature exceeds the total range of temperatures in the training set at or below $r = 0.65$. This correlation is typical of that used in many climate-proxy reconstructions, and it suggests that our understanding of past climate variability may be somewhat constrained.

Keywords

climate, correlation, calibration, proxy, simulation, carbon isotope, confidence interval

Introduction

With ever increasing interest, both political and scientific, in the variability of past climates, it is important to have confidence in the quality of the data and in the predictions made from the data. To produce absolute quantitative measures it is necessary to place considerable reliance on the accuracy and precision of climatic reconstructions (Briffa and Osborn, 1999). Proxy climate data have been obtained from many sources such as those climate sequences extracted from ice-cores (Siegenthaler *et al.*, 2005), tree-rings, coral reefs, speleothems, peat bogs, lake sediments, laminated sediments, loess deposits, marine sequences and many other natural assemblages obtained from the stratigraphic record (Mackay *et al.*, 2005). It has been demonstrated that the use of simple linear regression (inverse calibration) to quantify the relationship between a proxy and a climate variable can result in the introduction of bias, whereby extreme estimated values are drawn towards the mean, thus underestimating the frequency of extreme events (Robertson *et al.*, 1999). Moreover, the width of the 95% confidence intervals for the predicted values can be of the same order as the variation in the parameter being predicted. The degree to which each of these effects manifest themselves depends directly on the value of the correlation coefficient between the measured proxy and the climate variable in the training set. However, there is a lack of published work upon what correlation is the minimum acceptable value. In one of the few studies, McCarroll and Pawellek (2001) arbitrarily selected a minimum threshold value of $r = 0.71$, which explains half of the variance in the data. In this paper, the effect of variation in this coefficient is explored systematically using a series of simulated training sets exhibiting different but controlled

correlations between carbon isotope indices derived from oak cellulose and mean August temperature from the central England temperature (CET) record (Manley, 1974; Parker *et al.*, 1992). However, it is important to stress that any pair of proxy-environmental variables could have been used in the analyses.

Calibration and correlation

Although regression-based techniques are routinely applied to investigate the relationship between two variables; they should only be applied when the determination of one of the variables is without error. Unfortunately, this is rarely the case with environmental data. In such circumstances, alternative procedures are recommended to minimise bias. These procedures include the calculation of the reduced major axis (Till, 1973; Matthews, 1981), functional relationship estimation by maximum-likelihood (Ripley and Thompson, 1987) and Bayesian approaches to calibration (Robertson *et al.*, 1999). However, if regression-based techniques are still applied to environmental results, ordinary least squares regression may still be acceptable if the error on the independent variable (*x axis*) is less than one third of that of the dependent variable (*y axis*) (McArdle, 1988).

A general term for the inference of one quantity, for which measurements cannot be directly made, from another quantity which can be observed, is calibration. It is in many instances the only way in which quantitative knowledge can be generated about past events. For example, a direct record of monthly temperature observations exists only for relatively short time spans in geographically

discrete locations. However, both theory and empirical data show that temperature has an effect, directly or indirectly, upon several growth parameters in tree-rings. Knowledge of the way in which the tree-ring record is influenced by temperature can then be used to extend the temperature record back to times, and locations, for which direct observations are not available. This requires that a training set (or reference dataset) be available, in which both the proxy measurements and the climatic variable(s) are simultaneously recorded, in order to define the relationship between the two. It is normal practice to use only part of this paired dataset to generate the relationship (transfer function) between climate and proxy – the remainder is used to verify the relationship by comparing known with predicted values for the climatic variable (Briffa, 1995). Generally speaking, however, this is done by comparing point values, neglecting the uncertainty in the predicted values, and this can be somewhat misleading.

Correlation expresses the way in which two or more quantities are systematically associated throughout their range of variation. Mathematically, in the case of two continuous variables, the Pearson Correlation Coefficient (r) is given by:

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2 \Sigma y^2}} \quad (1)$$

where x is the deviation of x_i from the mean of all the x 's, and y is the deviation of y_i from the mean of all the y 's. The denominator in Equation 1 treats x and y separately, and itself is unaffected by the covariance between x and y , being a scaling factor. The numerator is far more important. The

magnitude of the numerator will be maximised by having all large deviations of x and y occurring together, and minimised by having large deviations in one variable paired with small deviations in the other.

The magnitude of the correlation coefficient can vary between zero and unity; zero indicating no relationship, unity indicating a fully determined relationship. Intuition suggests that the higher the correlation coefficient between two variables, the better any estimate of the target value made from the proxy should be. Any assessment of quality of estimation has to include some criteria for accuracy, such as how far estimates differ from true values in a known dataset, and must have some measure of precision insofar as ascribing a confidence interval to any estimate. A better estimate would, on the whole, have a lower deviation of estimate from true, and a narrower confidence interval.

Correlation within training set

The success of the regression-based techniques is illustrated by the many subject areas that routinely use regression such as zoology (Slabbekoorn and Peet, 2003) and phenology (Warren *et al.*, 2001). If there is thought to be a strong association between variables, the correlation usually lies between 0.6 and 0.9. For example, in the literature associated with climatic reconstruction from tree-ring isotopic proxies, correlations with environmental variables range usually from 0.4 to 0.8

(Heaton, 1999; Schleser *et al.*, 1999) with a reported value as high as $r = 0.94$ (Freyer and Belacy, 1983).

Proxy-climate correlation simulation

In order to explore the relationship between the accuracy of any reconstructed climatic value, that is the closeness of the reconstructed value to the true value, the estimated uncertainty of the reconstructed variable, and the correlation observed between the selected proxy and climate variable in the training set, a large number of simulated datasets were constructed and used to calculate simulated climate sequences. Two thousand datasets were generated for carbon isotope indices and temperature. To simulate the typical reference datasets n random ordinates were generated from a Gaussian distribution with parameters taken from the marginal parameters for the CET with mean value of 15.91°C and a standard deviation 1.005°C (Manley, 1974; Parker *et al.*, 1992). Response values were generated from the simulated temperature series by a linear transformation using the parameters from the fit of the isotopic sequence to the CET (Robertson *et al.*, 1997). Those parameters were: gradient -003959 and intercept 19.874425 . Some random noise generated from a Gaussian distribution with mean zero and standard deviation sampled from a random uniform with range 10 to 50 and added to the simulated isotopic measurements. This produced simulated temperature/isotopic measurement series with uniformly varying correlation coefficient. The simulations were run three times, with 15, 100 and 1000 temperature/carbon

isotope index pairs selected, the 15 and 100 pairs are considered typical of datasets encountered in paleoclimatic research.

A calibration for each simulated dataset was performed using linear regression (with y as isotopic measurement and x as temperature), and the reconstructed temperature was then calculated for each simulated tree-ring response. Since, for each dataset the ‘true’ and the reconstructed values are known, the mean absolute deviation (MAD; the mean of the difference in magnitude between the estimated temperature and the temperature in the simulated dataset) and mean confidence interval (95% CI) calculated from the following approximation (Draper and Smith, 1981, p.31).

EQUATION IN HERE

where:

EQUATION TERMS HERE

where:

MORE EQUATION TERMS

These performance estimation parameters were then plotted against the correlation for all 2000 datasets (Figures 1-2).

Results and discussion

Figures 1a and 2a depict, for each of the 2000 simulated datasets, the mean absolute deviation (MAD) as a function of correlation in the training set (containing 100 and 15 pairs of data respectively), which allows us to estimate the expected difference between the predicted and true value, for any particular correlation. Figures 1b and 2b illustrate the corresponding width of the mean 95% confidence interval plotted as a function of correlation. The case for 1000 pairs of data in the training set is not illustrated.

The results of Figure 1 are discussed further as it is the simulation most representative of training sets likely to be encountered in practice. The solid horizontal line marked on Figure 1b represents the maximum temperature range in the simulated datasets (approximately 5°C, compared to 5.8°C in the mean August temperature data for central England). This can be taken as a baseline for evaluating the performance of reconstructed temperature estimates, since estimates which have an associated error term greater than the total range of the parameter they are attempting to reconstruct must be considered to be of limited value. The vertical line on Figure 1b indicates the correlation corresponding to the point at which the average 95% confidence interval equals the average total range of simulated temperature values - in this case a correlation of $r = 0.65$ (or $r^2 = 0.42$). From Figure 2a it can be seen that at this level of correlation, each estimated point would have, on average, an absolute error of 1.2°C.

Figure 2 shows the same information for training sets with only 15 pairs of points, with the superimposed lines marking the same values discussed above. Figure 2b shows that the mean 95% confidence interval width reaches the same value as the mean range of the simulated temperatures (4°C) at a correlation coefficient of $r = 0.85$ ($r^2 = 0.72$), and at this value the mean absolute deviation (Fig. 2a) is around 0.7°C. As might be expected, a reduction in the *quantity* of data in the training sets leads to the need for a higher *quality* of data.

The equivalent values for 1000 pairs of data in the training set (not illustrated) are $r = 0.53$ ($r^2 = 0.28$) and 1.6°C, respectively, showing the opposite trend. Again, as might be expected, with 1000 pairs of data in the training set, $r = 0.53$ ($r^2 = 0.28$) is the minimum correlation necessary to give a 95% confidence interval less than the total range of temperatures in the training sets. The mean absolute deviation at this correlation is 1.6°C.

In the normal course of climatic reconstruction it would not be possible to access this information. The value of these simulations is that the results can be used to estimate important relationships between observations and proxy data, given only the correlation within the training set, providing the training sets contain around 15, 100 or 1000 points respectively. Parameters for other sizes of training set could, of course, be produced by re-running the simulations or (approximately) by interpolation. It was assumed for the sake of this argument that the temperature data being reconstructed have a distribution similar to the CET data for mean August temperature.

Investigations of climatic data which show higher (or lower) variation would require new simulations to elucidate the relationship.

Conclusions

This study shows that the strength of the correlation between two variables directly affects the ability to predict values of one given the other. Generally, and intuitively, the higher the correlation, the better estimates for the variable of interest. There is a fundamental need to consider the magnitude of the expected error when converting proxy data into climate data. Through simulated temperature and carbon isotope index data, for 100 pairs of variables which are correlated at about $r = 0.65$, the 95% confidence interval width for any estimate will be similar to the total range of variation for temperature. This threshold value for correlation can be in some ways regarded as a minimum correlation required for climate-related data (given 100 pairs of observations in the training set), where verification procedures are not employed. This result underscores the need to use highly correlated data for palaeoclimatic work. As most biological systems are not usually highly correlated, an alternative approach is required. The simplest solution may be to increase the sample size of the training set of climate parameter and proxy, and/or, combine several independent measures of past climate using multi-proxy techniques (Bradley and Jones, 1993; Overpeck *et al.*, 1997; Mann *et al.*, 1998, 1999; McCarroll *et al.*, 2003). If these proxies have different sources of noise, the climatic signal will be enhanced (McCarroll and Loader, 2004). The work shown here, whilst only illustrative, suggests that for many palaeoclimatic reconstructions, the average 95%

confidence interval for any estimated point is no better than 5°C for those sites whose marginal temperature distribution is similar to the CET. Although it is tempting to generalise these results the variances of the Northern Hemisphere temperatures datasets frequently used in climate reconstruction are usually lower than the CET, this would mean that the 95% CI is likely to be smaller than 5°C. Current climatic proxies indicate that any perceived anthropogenic effect on temperature would have to be identifiable above this level.

This paper has focussed on the relationship between carbon isotope indices and some estimate of average summer temperature. It appears to be the case, however, that the simulations shown here have very general application to the reconstruction of any environmental variable from any proxy measurement; differing only in so far as the distribution of the variable being reconstructed varies. As such, it suggests that the conclusions relating to the minimum value of r required to give definition to the variation have broad relevance in all environmental studies, and this requires a much more explicit consideration of uncertainty than has typically been the case.

Acknowledgements

We thank A.C. Barker, K.R. Briffa, A.H.C. Carter, M.A. Hall, J. Rolfe, V.R. Switsur and J.S.

Waterhouse for help with the construction of the stable isotope time-series and P.D. Jones and B.

Horton for updating the central England temperature series. This research was partially supported

by the European Union (EVK2-CT-2002-00136) and NERC (GR3/11395).

References

Bradley, R.S. and Jones, P.D. 1993. 'Little Ice Age' summer temperature variations: their nature and relevance to recent global warming trends. *The Holocene* 3, 387-396.

Briffa, K.R., 1995. Interpreting high-resolution proxy climate data – the example of dendrochronology. In von Storch, H. and Navarra, A. (eds.), *Analysis of Climate Variability: Applications of Statistical Techniques*. New York, Springer-Verlag, 77-94.

Briffa, K.R. and Osborn, T.J., 1999. Seeing the wood from the trees. *Science*, 284, 926-927.

Draper, N. and Smith, H. 1981. *Applied Regression Analysis: second edition* John Wiley and Sons Inc.

Freyer, H.D. and Belacy, N., 1983. 13C/12C records in northern hemispheric trees during the past 500 years - anthropogenic impact and climatic super-positions. *Journal of Geophysical Research - Oceans and Atmospheres* 88, 6844-6852.

Heaton, T.H.E., 1999. Spatial, species, and temporal variations in the 13C/12C ratios of C3 plants: implications for palaeodiet studies. *Journal of Archaeological Science* 26, 637-649.

Mackay, A., Battarbee, R., Birks, J., Oldfield, F., 2005. *Global Change in the Holocene*. London, Arnold.

Manley, G., 1974. Central England temperatures: monthly means 1659 to 1973. *Quarterly Journal of the Royal Meteorological Society* 100, 389-405.

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.* 26, 759-762.

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

Matthews, J.A., 1981. The form of relationships and prediction by regression. In *Quantitative and statistical approaches to geography*. Oxford, Pergamon Press Ltd, pp. 155-170.

McArdle, B.H., 1988. The structural relationship: regression in biology. *Canadian Journal of Zoology* 66, 2329-2339.

McCarroll D. and Loader N.J., 2004. Stable isotopes in tree rings. *Quaternary Science Reviews* 23, 771-801.

McCarroll, D., Jalkanen, R., Hicks, S., Tuovinen, M., Pawellek, F., Gagen, M., Eckstein, D., Schmitt, U., Autio, J. and Heikkinen, O. 2003. Multi-proxy dendroclimatology: a pilot study in northern Finland. *The Holocene* 13, 829-838.

McCarroll, D. and Pawellek, F. 2001. Stable carbon isotope ratios of *Pinus sylvestris* from northern Finland and the potential for extracting a climate signal from long Fennoscandian chronologies. *The Holocene* 11, 517-526.

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*, 278, 1251-1256.

Parker, D.E., Legg, T.P. and Folland, C.K., 1992. A new daily central England temperature series, 1772-1991. *International Journal of Climatology* 12, 317-342.

Ripley, B.D. and Thompson, M., 1987. Regression techniques for the detection of analytical bias. *Analyst* 112, 377-383.

Robertson, I., Switsur, V.R., Carter, A.H.C., Barker, A.C., Waterhouse, J.S., Briffa, K.R. and Jones, P.D., 1997. Signal strength and climate relationships in the $^{13}\text{C}/^{12}\text{C}$ ratios of tree ring cellulose from oak in east England. *Journal of Geophysical Research* 102, 19507-19516.

Robertson, I., Lucy, D., Baxter, L., Pollard, A.M., Aykroyd, R.G., Barker, A.C., Carter, A.H.C., Switsur, V.R., and Waterhouse, J.S., 1999. A kernel-based Bayesian approach to climatic reconstruction. *The Holocene* 9(4), 495-500.

Schleser, G.H., Helle, G., Lucke, A. and Vos, H., 1999. Isotope signals as climate proxies: the role of transfer functions in the study of terrestrial archives. *Quaternary Science Reviews* 18(7), 927-943.

[Siegenthaler U.](#), [Stocker T.F.](#), [Monnin E.](#), [Luthi D.](#), [Schwander J.](#), [Stauffer B.](#), [Raynaud D.](#), [Barnola J.M.](#), [Fischer H.](#), [Masson-Delmotte V.](#), [Jouzel J.](#), 2005. Stable carbon cycle-climate relationship during the late Pleistocene. *Science* 310 (5752), 1313-1317.

Slabbekoorn, H. and Peet, M., 2003. Birds sing at a higher pitch in urban noise, *Nature* 424, 267.

Till, R., 1973. The use of regression in geomorphology. *Area* 5, 303-308.

Warren, M.S., Hill, J.K., Thomas, J.A., Asher, J., Fox, R., Huntley, B., Roy, D.B., Telfer, M.G.,
Jeffcoate, S., Harding, P., Jeffcoate, G., Willis, S.G., Greatorex-Davies, J.N., Moss, D. and
Thomas, C.D., 2001. Rapid response of British butterflies to opposing forces of climate and habitat
change. *Nature* 414, 65-69.

Figure Captions

Figure 1. Plot of (a) mean absolute deviation (MAD) and (b) 95% confidence interval (CI) width from 2000 simulated datasets of 100 pairs of carbon isotope indices and mean August temperature showing varying correlation. The solid horizontal line in (b) shows the maximum temperature range observed in the simulated datasets. The vertical line in (b) denotes the value of the correlation coefficient at which the 95% confidence interval width exceeds the maximum range in the training set. The vertical line in (a) shows the same value of correlation coefficient, and the horizontal line allows the corresponding mean absolute deviation to be estimated.

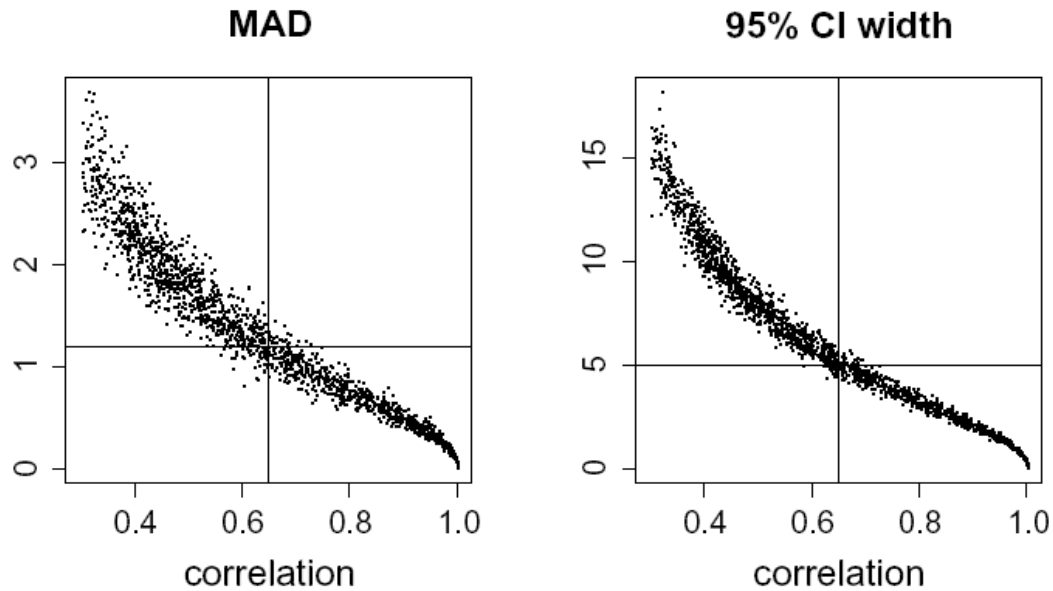


Figure 2. Plot of (a) mean absolute deviation (MAD) and (b) 95% confidence interval (CI) width from 2000 simulated datasets of 15 pairs of carbon isotope indices and mean August temperature showing varying correlation. The solid horizontal line in (b) shows the maximum temperature range observed in the simulated datasets. The vertical line in (b) denotes the value of the correlation coefficient at which the 95% confidence interval width exceeds the maximum range in the training set. The vertical line in (a) shows the same value of correlation coefficient, and the horizontal line allows the corresponding mean absolute deviation to be estimated.

