Cite as: Lucy,D. Aykroyd,R.G. Pollard,A.M. and Robertson,I. (1999)
Estimates of Uncertainty in the Prediction of Past Climatic Variables,
unpublished report

# Estimates of Uncertainty in the Prediction of Past Climatic Variables

D. Lucy[1], R.G. Aykroyd[3], A.M. Pollard[2] and I. Robertson[4]

[1] Department of Mathematics and Statistics, Edinburgh University, Edinburgh, EH9 3JZ

[2] Department of Archaeological Sciences, University of Bradford, Bradford, BD7 1DP

[3] Department of Statistics, University of Leeds, Leeds, LS2 9JT

[4] Quaternary Dating Research Unit, CISR, Environmentek, South Africa.

## Abstract

We discuss two important considerations when reconstructing climate data from proxy information, such as the carbon isotope index in tree-rings. The first is the relationship between the correlations observed in a training set, which consists of pairs of isotopic and climate data, and the quality of the resulting reconstructed climate data. Quality in this case is measured in terms of the mean absolute deviation between reconstructed point and true value, and the mean 95% confidence interval width for each reconstructed point. We show by simulation that both parameters are adversely affected by falling correlation, to the point where, in a training set consisting of 100 pairs of temperature and carbon isotope index points, the mean 95% confidence interval width for the reconstructed temperature exceeds the total range of temperatures in the training set at $r = 0.67$. This correlation is typical of many climate-proxy reconstructions, and it suggests that our understanding of past climate variability may be severely restricted. Our second point expands on this observed sensitivity to variations in correlation. We have studied time-variant trends in climate-proxy correlations by applying a 15 year moving window to a 100 year training set, and show that the correlation between proxy and any particular climate variable can vary significantly over this period. The implications of this are that the quality of any reconstructed climate parameter will vary with time, in a manner which we cannot as yet predict outside the

confines of the training set.

## Keywords

**Climate, Correlation, Calibration, Time-series**

# 1 Introduction

With ever increasing interest, 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. For the period before instrumental records the longest climate sequences come from ice-cores (Petit *et al.*, 1999), however, these are mostly used as relative measures. To produce absolute quantitative measures it is necessary to place considerable reliance on the accuracy and precision of climatic reconstructions such as temperature (Lara and Villalba, 1993) and rainfall (Bednarz and Ptak, 1990), from proxy variables like tree ring widths. As individual records may contain bias and errors (Lough and Barnes, 1997), a multi-proxy approach to high resolution climatic reconstruction is often adopted (Bradley and Jones, 1993; Overpeck *et al.*, 1997; Mann *et al.*, 1998, 1999). We have shown elsewhere (Robertson *et al.*, 1999) that the use of simple linear regression to calibrate the relationship between a proxy and a climate variable can result in the introduction of bias (whereby extreme values are drawn towards the mean), thus underestimating the frequency of extreme events. 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 proxy and the climate variable in the training set. Here we explore systematically the effect of variation in this coefficient using a series of simulated training sets exhibiting different but controlled correlations between carbon isotope index as derived from oak cellulose and mean August temperature.

We have also considered the time-dependent nature of the relationship between a climate variable and the measured proxy, in complex situations where more than one climate variable might have an influence on the value of the proxy. Using a moving window technique on carbon isotope data on oaks from two sites in eastern England, we show that the value of the correlation calculated over a 15 year period can vary from 0.05 to 0.84. Combining this with the previous observation, we conclude that the magnitude of the error on the reconstructed climate variable might be expected to vary significantly with time, reflecting the variation in correlation. Finally, we examine possible ways in which this behaviour might be detected, and suggest ways in which it might be used to give usefully improved estimates over limited ranges of the climatic-proxy time series.

## 2 Calibration and correlation

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 knowledge can be generated about past events. For example, a direct record of monthly temperature observations exist only for relatively short time spans in geographically discrete locations. However, both theory and empirical data show that, for example, temperature has an effect upon several growth parameters in trees which are recorded in the annual growth increments, and which subsequently can be measured. Knowledge of the way in which the tree-ring record is influenced by temperature can then be used to extend the temperature record 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 record! ed, in order to define the relationship between the two. In other words a proxy can be used in conjunction with a training set to produce a series of estimates for the target variable of interest.

A general concept of correlation is widely understood, and is 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 $x$'s, and $y$ is the deviation of $y_i$ from the mean of all $y$'s.

The denominator in Equation 1 treats $x$ and $y$ separately, and itself is unaffected by the $xy$ correlation, being a scaling factor. The numerator is far more important in that it 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 between the two variables, unity indicating a fully determined relationship. Intuition would suggest 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 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.

## 2.1 Correlation within training sets

In most natural complex systems where there is thought to be a strong association between variables, the correlation usually lies between 0.6 and 0.8. Briffa and Osborn (1999) found the correlation between ring widths, and April to September temperatures during the period 1881-1960 for five Northern Eurasian tree-ring series, in the range 0.60 to 0.71. Robertson *et al* (1997) found a correlation of 0.64 and 0.53 between temperature and the high frequency carbon isotopic composition for a series of tree rings from Sandringham and Babingley in East Anglia over the period 1895-1994. They also found correlations of -0.67 and -0.57 between July relative humidity and high frequency carbon isotope index for the period 1921-1944 for each site respectively. Heaton (1999) gives a Table (their Table 1), summarising a range of recent climatic studies showing correlation coefficients ranging from 0.44 to 0.81 between $\delta^{13}$C and temperature, and 0.52 and 0.90 for $\delta^{13}$C and water availability.

## 2.2 Proxy-climate correlation simulation

In order to better understand the relationship between the accuracy and estimated uncertainty of the reconstructed variable and the correlation observed in the training set, we have carried out a large number of simulations. Two thousand datasets were generated for carbon isotope index and temperature, each consisting of 100 temperature/carbon isotope index pairs to simulate the size of a typical reference dataset. Each temperature value was sampled from a Gaussian distribution of parameters having a mean value of 15.91 and a standard deviation 1.005, which closely reflects the measured mean August Central England temperature distribution (Manley, 1974; Parker *et al.*, 1992). High frequency carbon isotope data obtained from absolutely-dated chronologies derived from the tree rings of five trees growing on each of two distinct sites in eastern England (Sandringham Park and Babingley Osier Carr ($52^o50'N$, $0^o30'E$) near King's Lynn (Robertson *et al.*, 1997)) were used to establish the parameters for a suitable tree ring response. The actual values being sampled from another Gaussian distribution using a linear transformation of the simulated temperature to give the mean, with variance being controlled by adjusting the standard deviation of the second random distribution. This gives a dataset consisting of two variables with an adjustable correlation, and which obeys the assumption of least squares regression for an absolutely determined $x$, $y$ being known with some degree of normally distributed error.

A calibration for each dataset was performed using linear regression (with $y$ as dependent and $x$ as independent variable), and the reconstructed temperature was then calculated for each simulated tree-ring response. For each dataset the mean absolute deviation (MAD) and mean confidence interval width were calculated. These were then plotted against the correlation of the original dataset for all 2000 datasets (Figure 1).

Although most training sets for environmental reconstruction seem to contain around 100 pairs of observations, we have explored the effect of reducing the number of pairs in order to determine the sensitivity of the analysis to this number. As an example, Figure 2 shows the results of an identical simulation experiment, with the exception that only 15 pairs of data were included in each simulated training set.

Figure 1a depicts the mean absolute deviation as a function of correlation for training sets containing 100 pairs. This allows us to estimate the expected difference between predicted and true value, for any particular correlation. Figure 1b shows the mean 95% confidence interval width again plotted as a function of correlation.

The horizontal line marked on Figure 1b represents the mean of the total range of temperature over all of the simulated datasets (approximately $5^o$C). 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 $r = 0.67$ ($r^2 = 0.45$). From Figure 1a it is can be seen that at this level of correlation ($r = 0.67$), each estimated point would have, on average, an absolute error of 1.2$^o$C.

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

In the normal course of dendroclimatological reconstruction it would not be possible to access this information. The value of these simulations is that Figures 1 and 2 can be used to estimate these rather important parameters, given only the correlation within the training set, providing the training sets contain around 100 or 15 points respectively. Parameters for other sizes of training set could, of course, be produced by re-running the simulations.

## 3 Time dependent correlations

The above simulations have shown that the correlation coefficient between proxy and climate data is a very important parameter when considering the quality of the reconstructed variable. Small changes in this coefficient could lead to rather large changes in the quality of the reconstruction. In the situation where more than one climatic variable might be expected to have an influence on the proxy, it is conceivable that the strength of the correlation between any single climate parameter and the proxy might vary with time. The situation might be visualised as one in which the value of a proxy for a particular year is most strongly influenced by whichever climatic parameter is most limiting during the growing season (Fritts, 1976) - perhaps rainfall in times of drought, in which case temperature may have little influence, and vice versa in a wet summer, in which temperature exerts more influence. The following season may, of course, be completely different.

In this situation, the derived value of the correlation coefficient must be seen as the average value for the strength of the relationship between the climate parameter and the proxy, over the period for which the two variables are mutually defined. In a typical training set there might be 100 pairs of points, thus the correlation produced is an average value over that time period. It would be possible to divide the initial 100 pairs into smaller subsets, each subset will have a correlation coefficient describing the local strength of relationship between the two variables, and the overall correlation can be seen as a function of all the correlations of all subsets.

Using the above approach, the change in relationship between two variables in a time series (such as tree-ring response to temperature) can be examined. Since the overall correlation can be considered as a function of the correlations in the subsets, some of which will be smaller, and some larger, than the overall correlation, the subsets can in their turn be thought of as mapping out the change with time in the relationship between the two variables.

## 3.1 The Moving Window Method

Using the dataset mentioned previously (Robertson, 1997) the correlation between the carbon isotopic index and mean August temperatures was calculated using a moving window of 15 years, and for display purposes the correlation was attributed to the centre year of the moving window. For example, the first window considered the temperature/carbon isotope index subset for the period 1895-1910, and the correlation was associated for display purposes with the year 1902. The significance of any correlation can be tested using student's t-test, allowing for the fact that only 15 pairs of data are included in the correlation. If the overall correlation is significant then there must be a value for the running window width, the 'saturation point', above which the correlation for each and every subset is significant (for instance where the window width covers all pairs). Equally, there will be a value of the window width below which all correlations are non-significant. In this case the saturation point is a window width of 20 years, so a width of 15 years has been arbitrarily selected as a value just before this saturation point (see Figure 3). However, any other value between the two extremes could equally well be adopted. The 15 year window width was cross-validated by repeated sampling of random $xy$ pairs correlated at the same level. It was found that the probability of getting the distribution of correlations seen in the real data was extremely low, indicating their non-random nature. Figure 4 shows the results of this procedure for trees from Sandringham and Babingley.

We have shown elsewhere (Aykroyd *et al.*, In Preparation) using this technique that there are significant variations in the correlations between monthly averaged climatic data and monthly averaged relative humidity data and the carbon isotope index for trees growing in Sandringham. We suggested there that this is direct evidence for changes to the growing season of oaks in a northerly latitude, providing evidence for an extension of this period, consistent with other observations. Here we focus on just one correlation, between mean August temperature and carbon isotope index, but present data from two sites which a geographically separated by 3 kilometres. Figure 4 shows the general similarity in pattern between the response of trees growing at Sandringham and Babingley. For both sites there are periods of significant correlation between 1906 and 1923, and in the post war period for 1947 to 1980 at Sandringham, and 1958 to 1984 at Babingley. The period 1940 to 1950 shows low correlations at both sites, even as low as zero. One is drawn to wonder whether the observation that the correlation time series shows a significant drop during the Second World War may suggest misrecording of climatic data during that period. However this is unlikely as the homogeneous Central England Temperature Record is a 'three' station-average that has undergone extensive quality control and has been correlated with twenty individual stations over the British Isles (Jones and Hulme, 1997). To alter the record for reasons of national security during this time would have taken considerable coordination by those responsible, and would have only been undertaken from 1939 to 1944. It would not explain the more progressive and extended low correlation period seen at Babingley (Figure 2b).

To what extent it is valid to regard the time series functions represented in Figure 4 as being real is of some debate. The technique used above may be regarded as a smoothing method, which will of necessity produce smoothed functions, the degree of similarity expected between any two points in the time series being proportional to the number of shared elements and could give sufficient smoothing to impose the observed structure onto any dataset. However the fact that the two sites, which are independent apart from experiencing a broadly similar climate, have such similar peaks and troughs throughout the period shown here, suggests that there is a deeper structure (of climatic relevance) being revealed by this technique, and that the high frequency $\delta^{13}$C values are being climatically *forced* in the same way observed by Anderson (1998). The discrepancy as to which climatic parameter is forcing $\delta^{13}$C between Anderson's (1998) work, and that presented here, can be accounted for by the regional differences in conditions between central England and central Switzerland.

# 4 Conclusions

We have demonstrated that the strength of the correlation between two variables directly effects our ability to predict values of one given the other. Generally the higher the correlation the better we are able to make estimates for the variable of interest. However, there is a fundamental need to consider the magnitude of the expected error when converting proxy data into climate data, and we have shown that for variables which are correlated at about the 0.67 level the 95% confidence interval width for any estimate will be similar to the total range of variation for the estimated variable. This puts $r = 0.67$ as an absolute minimum required correlation for climate-related data (given 100 pairs of observations in the training set). This is possible using the most highly correlated proxy and climate variables, but the implications of this relationship even with higher correlations need further investigation. An obvious way forward is to use only highly correlated data ($r > 0.9$) for all such work. Sadly, few climate proxies show such high correlation with simple climatic parameters, and perhaps the next step is to look more closely at models for relating proxy to climate variable so that the predictive power can be improved.

The search for natural climatic variability revolves around examining climatic parameters over recent history, and comparing them to the variability experienced during the rest of the Holocene when it is thought that industrial activities had no significant effect on the natural environment. The work above shows that on average the confidence interval for any estimated point is about $5^o$C. This means that using current climatic proxies any perceived anthropogenic effect on temperature would have to be identifiable above this level, probably having to be in the order of $2.5^o$C greater than any extended periods of Holocene temperature anomalies.

It is possible that the variation of correlation over time could be a useful means for increasing the predictive power of any climate and proxy variables. At those points in Figure 4 where the correlation between carbon isotope index and August temperature is significant the correlation varies around the 0.8 level at Sandringham and between 0.6 and 0.7 at Babingley, and would lead us to expect a mean deviation of around $0.75^o$C, and a mean 95% confidence interval width of about $4.5^o$C, which is an improvement in predictive power, even though the same accuracy and precision would not be available across the entire time series. What needs to be found is some feature of the data which would be an indication of how well the tree ring system was responding to the given climatic parameter. This might be available through modelling the factors which control isotopic parameters, or by considering the combined correlations of more than one pair of variables, and by using non-linear models examined recently by Schleser *et al.* (1999).


## 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 Commission and NERC.


## References


Anderson W.T., Bernasconi S.M., McKenzie J.A. and Saurer M. (1998) Oxygen and carbon isotopic record of climatic variability in tree ring cellulose (Picea abies): An example from central Switzerland (1913-1995). *Journal of Geophysical Research*, **103**(D24), 31625-31636.

Aykroyd R.G., Lucy D., Pollard A.M. and Robertson I. (In Preparation) Extended vegetation growth: influence on proxy-climate correlations: submitted to *Geophysical Research Letters.*

Bednarz Z. and Ptak J. (1990) The influence of temperature and precipitation on ring widths of oak (*Quercus robur* L.) in the Niepolomice forest near Cracow, southern Poland. *Tree-Ring Bull*, **50**, 1-10.

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**, 367-376.

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

Fritts H.C. (1976) *Tree Rings and Climate.* Academic.

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**(6), 637-649.

Jones P.D. and Hulme M. (1997) The changing temperature of 'Central England'. In *Climate of the British Isles: Present, Past and Future* (eds Hulme, M. and Barrow, E.) pp. 173-196, Routledge.

Lara A. and Villalba R. (1993) A 3620-year temperature record from Fitzroya cupressoides tree rings in southern South America. *Science*, **260**, 1104-1106.

Lough J.M. and Barnes D.J. (1997) Several centuries of variation in skeletal extension, density and calcification in massive Porites colonies from the Great Barrier Reef: a proxy for seawater temperature and a background of variability against which to identify unnatural change. *Journal of Experimental Marine Biology and Ecology*, **211**, 29-67.

Manley G., (1974) Central England temperatures: monthly means 1659 to 1973. *Quart. J. Royal Met. Soc.*, **100**, 389-405.

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**(6678), 779-787.

Mann M.E., Bradley R.S., and Hughes M.K. (1999) Northern hemisphere temperatures during the past millennium: Inferences, uncertainties, and limitations. *Geophysical Research Letters*, **26**6, 759-762.

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. *Int. J. Climatol.* **12**, 317-342.

Petit J.R., Jouzel J., Raynaud D., Barkov N.I., Barnola J.-M., Bender M., Chappellaz J., Davis M., Delaygue G., Delmotte M., Kotlyakov V.M., Legrand M., Lipenkeo V.Y., Lorius C., Pepin L., Ritz C., Saltzman E., and Stievenard M. (1999) Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. *Nature*, **399**, 429-436.

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.

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 $^{13}C/^{12}C$ ratios of tree ring cellulose from oak in east England. *J. Geophys. Res.* **102**, 19507-19516.

Schleser G.H., Helle G., Lücke 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.

Figure 1: Accuracy and precision against correlation. Each of the 2000 points represent *(a)* the mean average deviation ($^oC$ vertical axis) and *(b)* mean 95% confidence interval width ($^oC$ vertical axis) for a given correlation (horizontal axis). These are found by generating a random dataset comprising 100 $xy$ pairs with parameters closely mirroring a real proxy/temperature dataset. Then calibration was used to calculate a best estimated value and estimate of error for each temperature from it's corresponding proxy value. The mean absolute deviation was then calculated as the mean of the difference in magnitude between the estimated temperature and the temperture in the simulated dataset, the mean 95% confidence interval width being the mean of the calculated error ascribed to each estimate. The horizontal line on *(b)* is the mean of the range of temperatures in the training sets, and is taken to be a minimum requirement for precision. The vertical line is the minimum correlation required to be able to achieve that precision. The same level of correlation on *(a)* gives a measure of the expected inaccuracy.

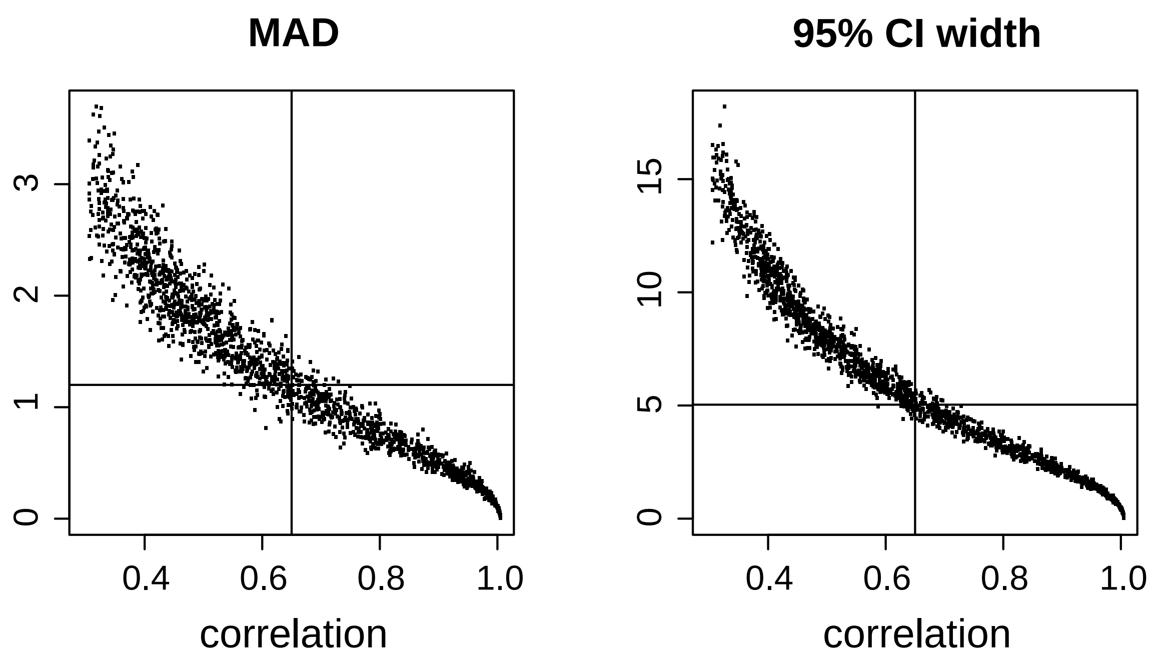

Figure 2: Plot of correlation against accuracy and precision for 2000 simulated datasets: These are the same as for Figure 1, but to illustrate the effect of sample size each of the 2000 points is based on a training set size of 15 points

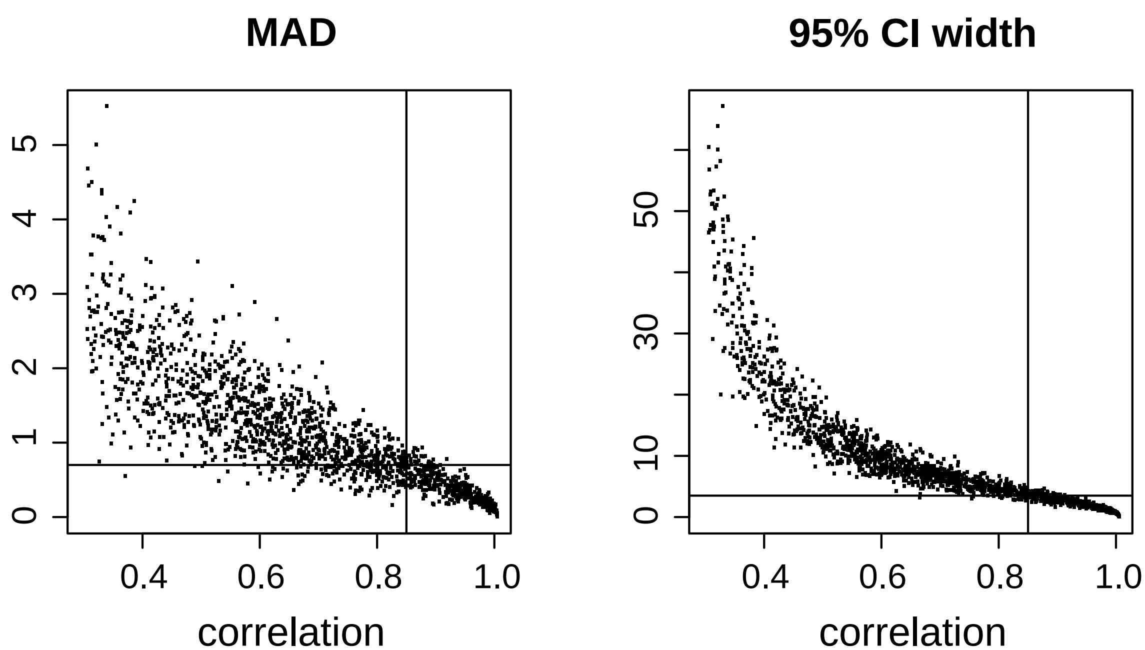

Figure 3: Effect of window width on the proportion of significant correlations seen between carbon isotope index and mean July/August temperature. The 15 year window selected gives good properties before the saturation point is reached.

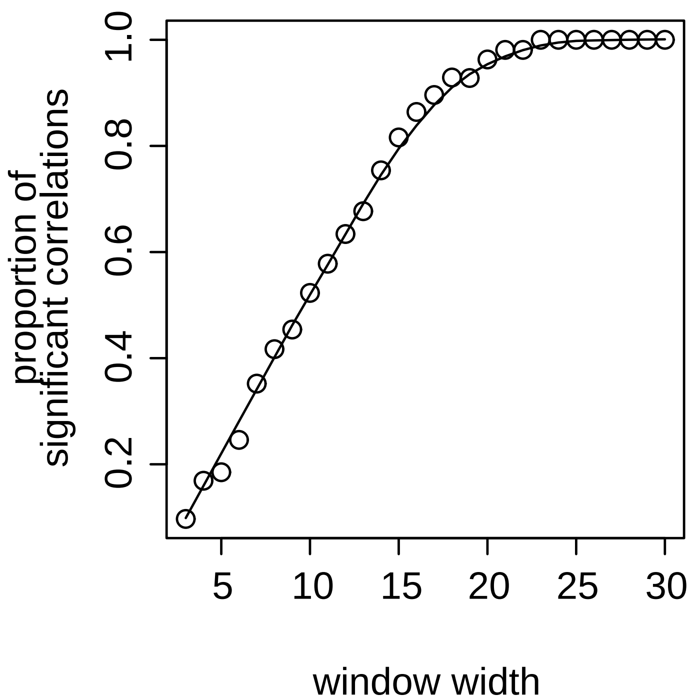

Figure 4: Time series correlation: the dashed horizontal lines indicate the overall correlation between temperature and carbon isotope index, the lines with shorter dashes are the level at which the correlation for each of the windows represented by the points in the time series is significantly greater than zero.