

Robertson, I.; Lucy, D.; Baxter, L.; Pollard, A.M.; Aykroyd, R.G.; Barker, A.C.; Carter, A.H.C.; Switsur, V.R.; Waterhouse, J.S.
Title: A kernel-based Bayesian approach to climatic reconstruction
Journal: The Holocene
Date: 1999
Volume: 9(4)
Pages: 525-530

A Kernel-based Bayesian Approach to Climatic Reconstruction

I. Robertson¹, D. Lucy², L. Baxter^{2,a}, A.M. Pollard², R.G. Aykroyd³, A.C. Barker⁴, A.H.C. Carter^{1,4}, V.R. Switsur^{1,4} and J.S. Waterhouse⁴

¹ Godwin Institute for Quaternary Research, University of Cambridge, Cambridge CB2 3SA, United Kingdom.

² Department of Archaeological Sciences, University of Bradford, Bradford BD7 1DP, United Kingdom.

³ Department of Statistics, School of Mathematics, University of Leeds, Leeds LS2 9JT, United Kingdom.

⁴ Environmental Science Research Centre, Anglia Polytechnic University, Cambridge CB1 1PT, United Kingdom.

Present address:

^a National Fingerprint Laboratory, New Scotland Yard, Broadway, London SW1H 0BG, United Kingdom.

Abstract

To understand recent climatic trends and possible future climatic change, it is necessary to examine the nature of past climatic variability (Nicholls *et al.*, 1996). Proxy measures of past climatic fluctuations can be used to extend this record beyond the limited period of instrumental measurements. Regression-based techniques are generally used to define transfer functions which describe the statistical relationship between these proxy estimates of past climate and measured climatic parameters. Although these regression-based techniques have been extremely successful, they can engender bias in the estimates if not used with care. More significantly, we also show that if regression errors are explicitly calculated they are often similar in magnitude to the total range of the parameter being estimated, implying that such reconstructions of past climate cannot be regarded as truly precise. A novel approach based upon Bayes' theorem is introduced which appears to increase the statistical veracity of such climatic reconstructions.

Key words: climatic reconstruction, regression analysis, Bayes' theorem, kernel-density estimation, carbon isotopes, dendroclimatology

Introduction:

Regression Procedures in Climatic Reconstruction

The objective of calibration in climatic reconstruction from proxy data is to derive a transfer function which expresses one or more environmental variables as a function of an indirect measure of past climate. To this end a wide range of regression techniques are routinely employed. Methods used for the reconstruction of past climate at a high-resolution include the measurement of the physical and chemical properties of ice-cores (Peel, 1995), coral records (Lough and Barnes, 1997), laminated sediments (Zolitschka, 1996), fossil beetle remains (Atkinson *et al.*, 1987) and tree-rings (Briffa *et al.*, 1995). The derived relationship is usually verified against independent data or data withheld from the fitting procedure as a test of the validity of the model (Fritts, 1976; Fritts and Guiot, 1990; Briffa, 1995). There are two conceptually different approaches to quantitative calibration - these are *classical* and *inverse calibration* (Osborne, 1991; Birks, 1995; ter Braak, 1995a).

Inverse calibration has been widely used for reconstructing past climate (Birks, 1995). Various approaches have been developed from simple linear regression (e.g. Williams and Johnson, 1975) to more complex techniques, such as those pioneered by Imbrie and Kipp (1971) who used factor analysis to derive assemblages for use in subsequent calibration. Past climates have been inferred using *inverse calibration* from foraminiferal data (Imbrie and Kipp, 1971; Williams and Johnson, 1975), pollen assemblages (Andrews *et al.*, 1980), limnological data (Whitmore, 1991) and measured tree-ring characteristics (Lofgren and Hunt, 1982; Robertson *et al.*, 1997). This approach is also known as *forward prediction* (Martens and Naes, 1989; Lucy, 1997). In inverse calibration the proxy measure of past climate is usually regarded as the 'controlling variable' (x axis) and regressed against the environmental variable of interest, the 'response variable' (y axis), using a 'training set' to define the nature of this relationship (ter Braak, 1995a, b; Birks, 1995). In its most basic form, the relationship is assumed to be linear. That is, it will have the form $y = ax + b$, where a represent the slope of the line and b its intercept with the y axis. Conventionally this relationship is found using a least squares criterion, where the values of the x variable are considered to be absolutely determined, and the values of y are considered to be subject to some random error, as illustrated in Figure 1. The parameters a and b are determined by the least squares criterion so that the sum of the squares of the distances e_1 , e_2 and e_3 (in the case illustrated) is minimised.

The calculated values of a and b are then used to provide estimates for unknown values of the y variable for particular values of the x variable, providing that x lies within the range of x values in the training set. However, if this process is carried out for a set of known x - y co-ordinates, and the

differences between the observed and the estimated values of y plotted against the observed y values, then it is found that there is a systematic error in these estimates (Figure 2). This systematic error makes low values of the estimated variable too high and high values of the estimated variable too low. In other words, a bias is present. From least squares criterion, it can be shown that the magnitude of the slope of this bias is $1-r^2$, where r is the Pearson correlation coefficient between the x and y values in the training set (Draper and Smith, 1981; Aykroyd *et al.*, 1997). If, as in most analyses, the y variable is the one being reconstructed (e.g., temperature or rainfall from tree-ring characteristics), then the estimates made from the proxy measure will always show a bias which is the complement of the square of the observed correlation coefficient. As most proxy measures of past climate do not demonstrate strong correlations with measured simple climatic parameters, there is always likely to be a significant bias in such estimates. One manifestation of this bias is that extreme climatic events are likely to be under-represented in the reconstructed data.

A possible solution to the problem of bias lies in a standard variant on regression analysis known as *classical calibration* (Birks, 1995). In classical calibration, the regression model is calculated in exactly the same way as before, except that the variable from which estimates are to be made is regarded as the x value, and its proxy data are y . The regression model, $y = ax + b$, then has to be rearranged before estimates of x can be made from y . As before, if the differences between the observed and estimated x values are plotted against the observed x values for a known dataset, the slope of a least squares line through them will be zero and will in fact lie along the x axis (Figure 2). This is the expected relationship from estimates that are unbiased.

Unfortunately, classical calibration is not a panacea for the reconstruction of climatic data from poorly correlated variables and, therefore, the approach has not been used widely in palaeoclimatology, except for inferring sea-surface temperatures from the mean size of foraminifera (Malmgren and Kennett, 1976; 1978). There are two major limitations to this approach. Firstly, the error estimates for any reconstructed point are always larger than those obtained by inverse calibration and can, in the case of climatic reconstruction, be as large as the total range of the reconstructed variable. Secondly, no version of the classical calibration procedure currently exists for situations where multiple proxy measures are employed to give estimates.

The selection of a suitable calibration technique is not just a modern problem: Eisenhart (1939) first demonstrated the differences between *inverse calibration* and *classical calibration*. In the years that followed, the approach of classical calibration was generally adopted until the techniques were re-investigated by Krutchkoff (1967) and Tallis (1969). Using Monte Carlo simulations, Krutchkoff (1967) demonstrated the superiority of inverse calibration over the classical calibration,

with the former approach giving smaller mean squared errors. It is now accepted that the inverse approach performs slightly better when the unknown samples are from a large central part of the distribution of the training set, even though there will be a bias present (Aykroyd *et al.*, 1997). However, the classical approach may perform better at extremes and under slight extrapolation (Birks *et al.*, 1990; Birks, 1995; ter Braak, 1995a).

Bayesian Approaches to Calibration

Another approach is to use Bayesian methods, which are now applied to a wide range of archaeological and environmental problems (Buck *et al.*, 1996). The Bayesian approach does not rely upon the explicit modelling of the relationship between two variables, but the modification of some prior belief about the specific value of a variable on the basis of some additional information. If a sequence of tree-ring characteristics of *Quercus robur* L. were to be used as a proxy for a climatic parameter such as mean July-August temperature, then we can initially assume that the mean July-August temperature through the period could not have been outside the range 0-50°C. We know this because the species has adapted to a particular ecological niche and will not grow outside these limits. In fact, it is possible to be more precise. It is unlikely that the mean July-August temperature was as low as 10°C or as high as 40°C, but quite likely that it was about 20°C. Working through a sequence of steps like this it is possible to define, based upon observation and other knowledge, a probability distribution for the temperature range experienced by the particular species of tree. In Bayesian terms, this is known as *prior* information, or *prior probability*.

This type of prior information on its own is not a very useful estimate of temperature as it is constrained only by the fact that species distribution exists over a restricted temperature range. If some property of tree rings is being measured, such as ring width, or carbon isotope values, then further information is available to refine this initial prior estimate. This further information is termed *conditional probability* or *likelihood*, and can be obtained by examination of a '*training set*' of controlled data similar to those used in more conventional calibrations, such as the measured tree ring response to known mean summer temperatures. In geometric terms, the conditional probability is derived from a larger dataset called the joint probability which, is visualised as a surface, rather than a line. The observed data points define the shape of the surface. A complicating factor is that a full joint probability surface for any climatic and response variables requires an extremely large dataset to build up the surface, and no such large datasets exist. Instead a smoothing function is used to interpolate values. The function used here is called a kernel-density function; whereby a kernel, which in geometric representation looks like a small bump, is added to a two dimensional surface, one kernel

for every (x, y) pair of points in the known dataset (Silverman, 1986). The result is a ‘sum of bumps’ and can be interpreted as a smoothed joint probability density function. For example, the resulting joint probability density function for mean July-August temperature and carbon isotope index is shown in Figure 3.

This surface shows the probability of any x, y pair occurring together (e.g., in Figure 3 there is a high probability that a mean summer temperature of 16.2°C will result in a carbon isotope index of 1.14). The conditional probability function is a slice through this surface parallel to the axis for which estimates are to be made (in this case, mean July-August temperature) and scaled in such a way that it has a total area below the surface of unity. Once the conditional probability density function has been obtained it can be combined with the prior probability density function to produce what is called a posterior probability density function using an appropriate form of Bayes’ theorem, the mathematics of which are set out in the appendix.

The posterior probability distribution can be regarded as the final probability distribution based on the knowledge of how one variable responds to another, as captured in the training set. For example, where the variables involved are the mean July-August temperature and the carbon isotope index, the posterior probability distribution is a statement of the probability distribution of mean July-August temperature given the particular observation of carbon isotope index. This statement is based upon previous observations of known carbon isotope indices from *Quercus robur* L. and mean July-August temperature. Figure 4 represents the probability distribution for mean July-August temperature predicted for a carbon isotope index value of 1.14.

A posterior probability distribution and the more conventional point estimate and associated confidence interval of that estimate are in certain respects equivalent, and it may be useful to consider the similarities when attempting to evaluate the efficacy of both approaches. A conventional estimate is a Gaussian distribution with the mean defined by the point estimate, and a confidence interval defining the ‘spread’ of the distribution. Seen in this way, a conventional estimate can be regarded as a probability distribution function for the parameter being estimated, albeit one with a Gaussian distribution. The posterior probability distribution will have a central value which, can be used as a point estimate if required, and a confidence interval that can be defined by truncating the distribution at a pre-defined value. These values can be calculated in a number of ways - for instance the modal value can be used as a point estimate and the points on the temperature axis which enclose 95% of the area of the posterior probability density function can be regarded as a confidence interval. However, to take the analogue between the conventional approach and a probability density approach much further than this negates some of the power of the latter technique - in particular the fact that the

posterior probability function is an accurate reflection of the possible outcomes, whereas the conventional approach imposes a Gaussian distribution, which may be unrealistic. A commonly occurring example is the case where the posterior probability density is more complex than a unimodal distribution. Here the conventional estimate should be regarded as unreliable, and the actual shape of the posterior distribution may help to refine ideas about how the response variable changes with the controlling variable.

A further advantage of the probability density approach is that there is no requirement for the response variables to be continuous, or even that they are ordinal. The calibration method is just as applicable to species counts as it is to carbon isotope indices, and can be used when the indicator variables contain a mixture of both types of observation. The only condition for these techniques is that a value in an observable must have a distribution for the variable of interest associated with it. The final advantage is that the posterior probability distribution is in itself a conditional probability distribution. In the example above, it is the probability of temperature given the information that the tree is capable of growth, and given the particular carbon isotope index. Were some new information to become available, for example tree-ring density, then the first posterior probability distribution could itself be regarded as a prior probability distribution and the new information used in exactly the same way to produce a new posterior probability distribution which would reflect all available information about the climatic variable of interest. Thus the method can be used to combine a wide range of differing sources of information to yield the best estimate of the climate variable given the available data.

Examples of Comparative Applications

We present a set of preliminary comparative results based on the reconstruction of summer temperature from stable isotopes in oak tree rings which indicates that a Bayesian approach can improve the estimates of climatic parameters. Perhaps more importantly, it serves to demonstrate that, in poorly correlated data of the type often encountered in many forms of climatic reconstruction from proxy data, the errors produced during the calibration (inverse or classical or Bayesian) procedure, if correctly estimated, are likely to be of a magnitude similar to, or greater than, variation in the parameter being reconstructed.

The example involves the reconstruction of mean July and August temperatures between AD 1895 and 1994 from high-frequency carbon stable isotope indices, calculated from five oaks (*Quercus robur* L.) growing at Sandringham Park (52°50'N, 0°30'E) in east England (Robertson *et al.*, 1997).

The high-frequency carbon stable isotope indices were obtained by dividing the $\delta^{13}\text{C}$ values of each tree by a fitted value obtained from a 60-year Gaussian filter to remove the very low frequency variance. The individual indices were then transformed into standardised indices by calculating standard anomalies. A 10-year high-pass Gaussian filter was applied to these standardised indices to define the high-frequency variance. Standard anomalies were again calculated and the mean site chronology represented here is the average of these values. The correlation between the mean July and August temperature and the high-frequency carbon stable isotope indices is 0.65, which is typical of the correlations encountered in such proxy palaeoclimatic reconstruction. Figure 5 shows a comparison between the observed mean July and August central England temperature and the temperature reconstructions from high-frequency carbon isotope indices, using in turn inverse calibration, classical calibration and Bayesian calibration. All estimates of temperature were made using a jackknife resampling strategy (Efron, 1982), which is essentially a technique operating on a 'leave one out' basis for removing the effect of including the specific values for the value to be estimated in the training set.

The summary results are presented in Table 1. In this example, the overall inaccuracy (defined as the mean absolute difference between the observed and estimated temperature) is lowest for inverse calibration (0.62°C). However, the bias is also highest (0.56) which translates to a systematic discrepancy of about 1.5°C at the extreme ends of the temperature range, and is clearly unacceptable for the believability of any reconstructed temperature sequence. Both classical and kernel density based estimates are less accurate (0.93°C and 0.81°C respectively), but have negligible bias (-0.02 and 0.10). The best 95% confidence interval is from inverse calibration (3.07°C), closely followed by kernel density estimation (4.25°C). Classical calibration has an associated confidence interval of 11.43°C ¹, which again is unacceptable when the total temperature range is 4.70°C .

¹ Classical calibration seems to have been unduly affected by the jackknife process, giving very large confidence intervals for the extremes of the temperature range. Removing these outliers leaves a confidence interval of about 6°C which is probably more representative of the true confidence interval.

Discussion

Although this case is not optimal, in that it does not display particularly high correlations between proxy data and climate variables, the results clearly demonstrate the general nature and magnitude of the problem we perceive in current calibration methodologies. We believe that there are two principal difficulties - the presence of a systematic discrepancy when using inverse calibration techniques, which we term bias, and the fact that the average 95% confidence width of the estimate can in some cases exceed the variability of the parameter being estimated, if the correlation between proxy and climate variable is relatively low. Bias will have the important effect of causing under-representation of extreme conditions in the reconstructed data set, i.e., a tendency to underestimate the frequency of very low or very high values.

Bayesian methods have been attempted previously for climatic reconstruction, but have proved disappointing (e.g., Van Deusen and Reams, 1996). However, the results presented here demonstrate clearly the potential for kernel-based Bayesian calibration to be a useful statistical tool for climatic reconstruction. Only some of the mathematical properties of kernel-density calibrations are known to the authors at present, although simulation of the mathematical properties of kernel-density calibration has shown that regression-based techniques are superior when the observed correlation is greater than ca. 0.9 (Lucy, 1997). However, for biological systems such as those used in climatic reconstruction, correlations are typically between 0.5 and 0.8, with figures greater than 0.9 being extremely rare. Our results indicate that in the majority of cases there use of kernel density based calibration techniques may offset the undesirable properties of bias seen in inverse calibration, and large confidence intervals and inaccuracy seen in classical calibration.

Conclusions

At a time when knowledge of past climatic variability is of increasing political and economic importance in world affairs, it is crucial that the statistical techniques employed by climatologists do not give misleading impressions of past climatic change, and give estimates for reconstructed climate parameters with the lowest possible errors. Conventional least squares regression-based techniques may fail in both these respects. According to Cook *et al.* (1994), even much more complex regression models cannot extract weak climatic signals from tree-rings. We suggest that closer scrutiny be applied to not just bias and accuracy, but the estimated confidence with which any technique can give a predicted value, as all the statistical techniques described here give 95% confidence intervals which are similar in magnitude, or greater than, the total variation in the climatic parameter of interest. However, there may be some potential in investigating further the properties and applications of

kernel-density estimates for climatic reconstruction from a wide range of proxy data, as part of a wider review of the performance of such techniques.

Acknowledgements

The authors thank our colleagues, Simon Crowhurst, Mike Hall, Debbie Hemming, Monni Kotilainen, Neil Loader and James Rolfe for their constant help and encouragement. This work was supported in part by the European Commission, the Leverhulme Trust and the British Academy. One of the authors (IR) was supported by a Henry Giles Fellowship. The work was also supported in part by a grant from the Natural Environment Research Council (NERC GR3/11395).

References

NEED IMBRIE AND KIPP 1971

Atkinson, T.C., Briffa, K.R. and Coope, G.R. 1987: Seasonal temperatures in Britain during the past 22,000 years, reconstructed using beetle remains. *Nature* 325, 587-592.

Andrews, J.T., Mode, W.N. and Davis, P.T. 1980: Holocene climate based on pollen transfer functions, eastern Canadian Arctic. *Arctic and Alpine Research* 12, 41-64.

Aykroyd, R.G., Lucy, D., Pollard, A.M. and Solheim, T. 1997: Technical note: Regression analysis in adult age estimation. *American Journal of Physical Anthropology* 104, 259-265.

Birks, H.J.B., Line, J.M., Juggins, S., Stevenson, A.C., ter Braak, C.F.J. 1990: Diatoms and pH reconstruction. *Philosophical Transactions of the Royal Society of London B*, 327, 263-278.

Birks, H.J.B. 1995: Quantitative palaeoenvironmental reconstructions. In Maddy, D. and Brew, J.S., editors, *Statistical Modelling of Quaternary Science Data, Technical Guide 5*, Cambridge, Quaternary Research Association, 161-254.

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

Briffa, K.R., Jones, P.D., Schweingruber, F.H., Shiyatov, S.H. and Cook, E.R. 1995: Unusual twentieth-century summer warmth in a 1,000-year temperature record from Siberia. *Nature* 376, 156-158.

Buck, C.E., Cavanagh, W.G. and Litton, C.D. 1996: *Bayesian Approach to Interpreting Archaeological Data*, Chichester: John Wiley and Sons.

Cook, E.R., Briffa, K.R., and Jones, P.D. 1994: Spatial regression methods in dendroclimatology: a review and comparison of two techniques. *International Journal of Climatology* 14, 379-402.

Draper, N.R. and Smith, H. 1981: *Applied Regression Analysis: Second Edition*, New York, John Wiley and Sons Inc.

Efron, B 1982: The jackknife, the bootstrap, and other resampling plans. Philadelphia: Society for Industrial and Applied Mathematics.

Eisenhart, C. 1939: The interpretation of certain regression methods and their use in biological and industrial research. *Annals of Mathematical Statistics* 10, 162-186.

Fritts, H.C. 1976: *Tree Rings and Climate*, London, Academic Press.

Fritts, H.C. and Guiot, J. 1990: Methods of calibration, verification, and reconstruction. In Cook, E.R. and Kairiukstis, L.A., editors, *Methods of Dendrochronology: Applications in the Environmental Sciences*, London, Kluwer Academic Publishers, 163-217.

Krutchkoff R.G. 1967: Classical and inverse regression methods of calibration. *Technometrics* 9(3), 425-439.

Lofgren, G.R. and Hunt, J.H. 1982: Transfer functions. In Hughes, M.K., Kelly, P.M., Pilcher, J.R. and LaMarche Jr, V.C., editors, *Climate from Tree Rings*, Cambridge, Cambridge University Press, 50-56.

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.

Lucy, D. 1997: *Human Age Estimation from Skeletal and Dental Evidence*. Unpublished PhD Thesis, University of Bradford.

Malmgren, B. and Kennett, J.P. 1976: Biometric analysis of phenotypic variation in recent *Globigerina bulloides* D'Orbigny in the southern Indian Ocean. *Marine Micropaleontology* 1, 3-25.

Malmgren, B.A. and Kennett, J.P. 1978: Late Quaternary paleoclimatic applications of mean size variations in *Globigerina bulloides* D'Orbigny in the southern Indian Ocean. *Journal of Paleontology* 52(6), 1195-1207.

Martens, H. and Næs, H. 1989: *Multivariate calibration*, Chichester, J. Wiley and Sons.

Nicholls, N., Gruza, G.V., Jouzel, J., Karl, T.R., Ogallo, L.A. and Parker, D.E. 1996: Observed climate variability and change. In Houghton, J.T., Meira Filho, L.G., Callander, B.A., Harris, N., Kattenberg, A. and Maskell, K., editors, *Climate Change 1995 - The Science of Climate Change*, Intergovernmental Panel on Climate Change, Cambridge, Cambridge University Press, 133-192.

Osborne, C. 1991: Statistical calibration: a review. *International Statistical Review* 59, 309-336.

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.

Peel, D.A. 1995: Ice core evidence from the Antarctic Peninsula region. In Bradley, R.S. and Jones, P.D., editors, *Climate Since A.D. 1500: Revised Edition*, London, Routledge, 549-571.

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 $^{13}\text{C}/^{12}\text{C}$ ratios of tree ring cellulose from oak in east England, *Journal of Geophysical Research* 102(D16), 19,507-19,516.

Silverman, B.W. 1986: *Density Estimation for Statistics and Data Analysis*, London, Chapman and Hall.

Tallis G.M. 1969 Note on a calibration problem. *Biometrika* 56, 505-508.

ter Braak, C.J.F. 1995a: Non-linear methods for multivariate statistical calibration and their use in palaeoecology: a comparison of inverse (k -nearest neighbours, partial least squares and weighted averaging partial least squares) and classical approaches. *Chemometrics and Intelligent Laboratory Systems* 28, 165-180.

ter Braak, C.J.F. 1995b: Calibration. In Jongman, R.H.G., ter Braak, C.J.F. and Van Tongeren, O.F.R., editors, *Data Analysis in Community and Landscape Ecology: Second Edition*, Cambridge, Cambridge University Press, 78-90.

Van Deusen, P.C. and Reams, G.A. 1996: Bayesian procedures for reconstructing past climate. In Dean, J.S., Meko, D. M. and Swetnam, T. W., editors, *Tree Rings, Environment and Humanity*, Arizona, Radiocarbon, 335-339.

Whitmore, T.J. 1991: Sedimentary diatom concentrations and accumulation rates as predictors of lake trophic state. *Hydrobiologia* 214, 163-169.

Williams, D.F. and Johnson II, W.C. 1975: Diversity and recent planktonic foraminifera in the southern Indian Ocean and Late Pleistocene paleotemperatures. *Quaternary Research* 5, 237-250.

Zolitschka, B. 1996: High resolution Lacustrine sediments and their potential for palaeoclimatic reconstruction. In Jones, P.D., Bradley, R.S. and Jouzel, J., editors, *Climatic Variability and Forcing Mechanisms of the Last 2000 Years*, Berlin, Springer-Verlag, 453-478.

TABLE 1. July and August temperature reconstruction for the period AD 1895-1994 from stable carbon isotopes in oak cellulose from east England, UK ($r=0.65$)

	Mean inaccuracy (°C)	Bias	Percentage estimates within 95% CI	Mean 95% confidence interval width (°C)
Inverse calibration	0.62	0.5938	94	3.07
Classical calibration	0.93	-0.0202	96	11.43
Bayesian	0.81	0.1033	94	4.25

The mean inaccuracy is the magnitude of the mean difference between the reconstructed temperature and the observed temperature. Mean 95% confidence intervals were calculated from the reconstructed temperature time-series.

Figures

Figure 1. The sample x - y pairs here are illustrated with their least squares line. The least squares parameters are calculated by assuming that all x values are known absolutely and that the y values are subject to some random error. Hence, the least squares method selects parameters for the model which minimise this random error shown as e_1 , e_2 and e_3 .

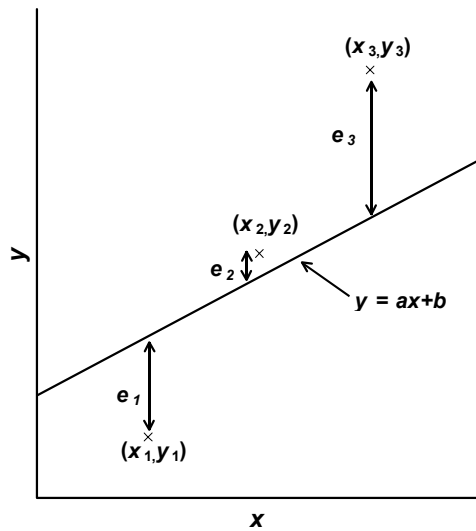


Figure 2. Plot of error for six estimates against their observed values, based upon the sample data represented in Figure 1. Both inverse and classical calibration techniques have been employed on these data which clearly illustrate the systematic error in estimates made by inverse calibration.

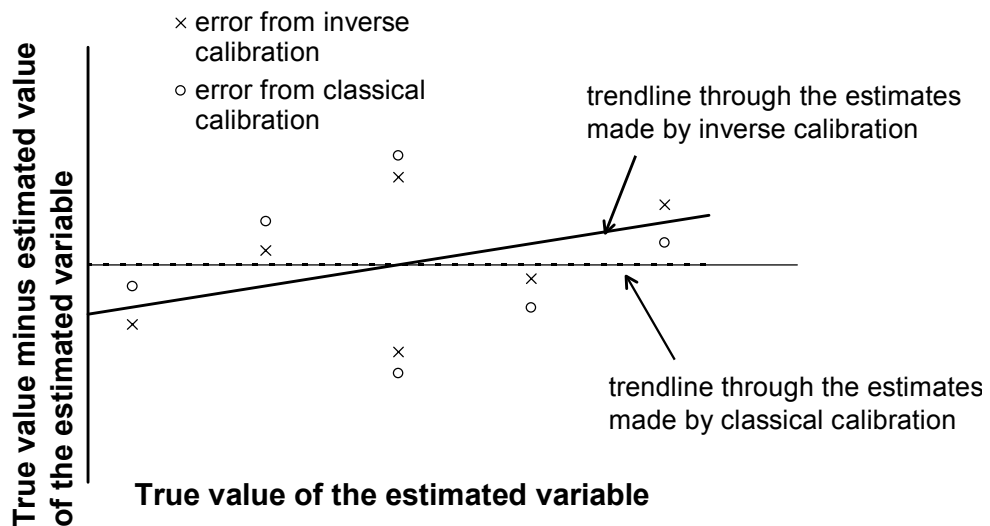


Figure 3. Joint density surface for Sandringham high-frequency carbon isotope indices and July-August temperature. Here, x is temperature, y is the high-frequency carbon isotope index and z is the probability for the joint event. A conditional probability for temperature, given a particular carbon isotope index, can be obtained by taking a slice through this surface at the appropriate value of y and dividing by $f(x)$.

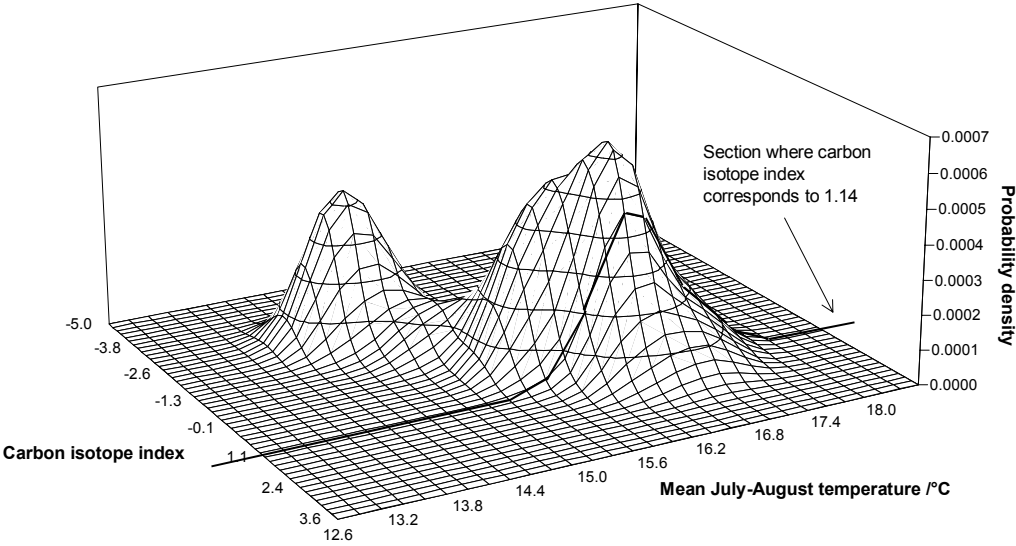


Figure 4. Conditional probability density for mean July-August temperature given a carbon isotope index value of 1.14 (the probability density is scaled so that the total area under the surface is unity).

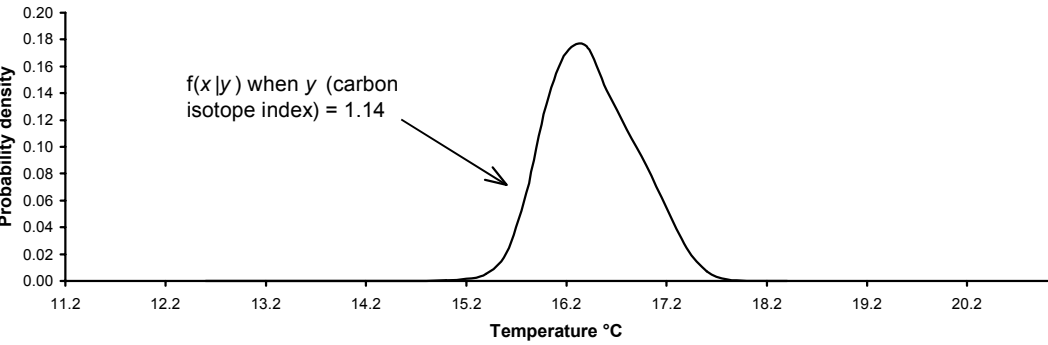
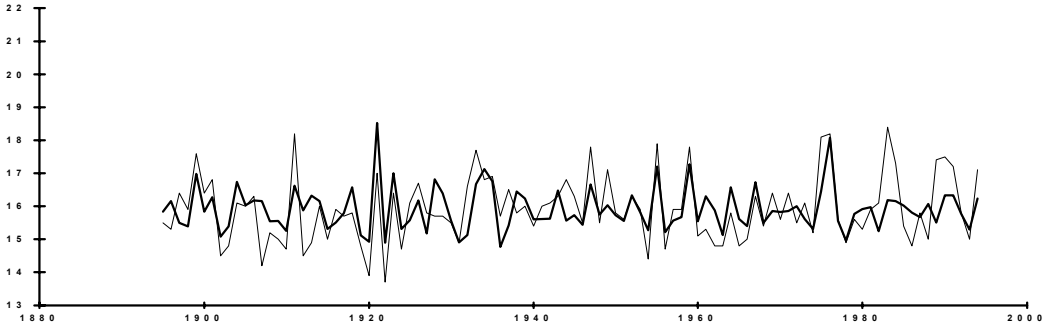


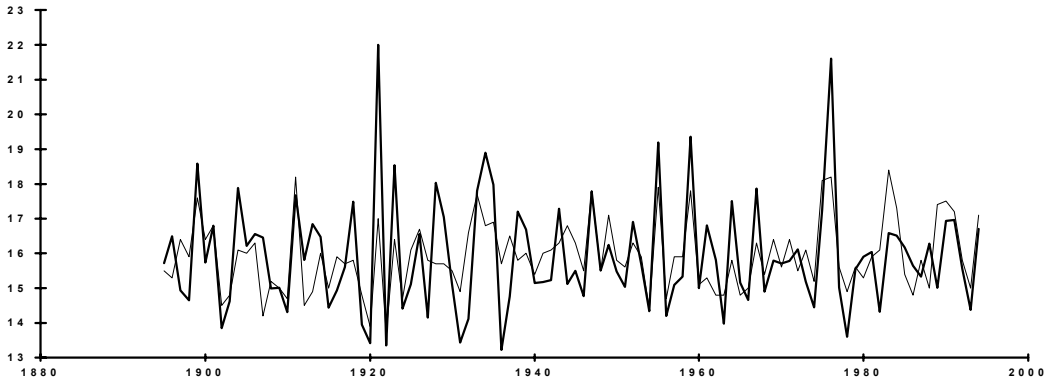
Figure 5. Observed (a) and reconstructed (b-d) July and August temperature (°C) for the period 1895-1994. The high-frequency carbon isotope indices used in the analyses were calculated from carbon isotope values using a 60-year Gaussian filter to remove the very low frequency variance. A 10-year high-pass Gaussian filter was then applied to the residuals to define the high-frequency variance. (a) Observed central England temperatures were calculated using the mean 24h maxima and minima daily values read at 0900h GMT (Parker *et al.*, 1992). (B) Reconstructed central England temperatures using inverse calibration ($r = 0.65$). (c) Reconstructed central England temperatures using classical calibration. (d) Reconstructed central England temperatures reconstructed using kernel-density techniques.

A – Sorry Can't find plot A

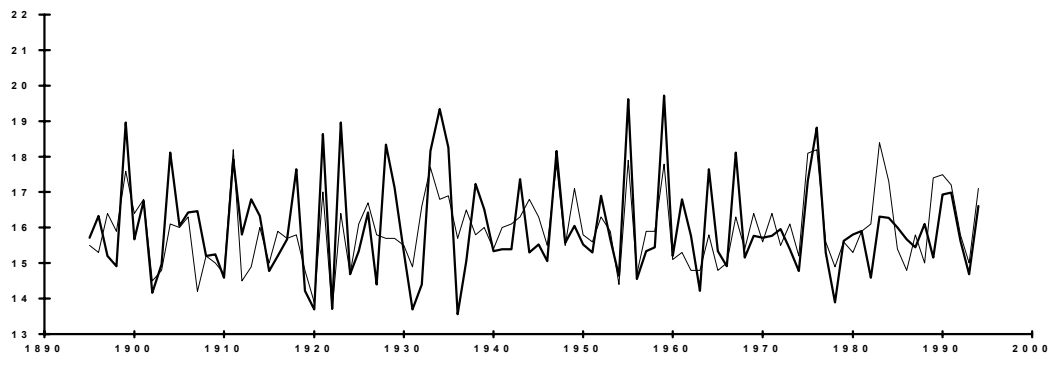
B



C



D



Appendices

Bayes' theorem

Suppose we are interested in calibrating variable X given the values of m indicator variables Y_1, Y_2, \dots, Y_m then a suitable form of Bayes' theorem (assuming conditional independence of the Y 's given X) that can be used when all variables are continuous is:

$$f(x | y_1, y_2, \dots, y_m) = \frac{f(y_1 | x)f(y_2 | x) \dots f(y_m | x)f(x)}{f(y_1, y_2, \dots, y_m)}$$

where: $f(x | y_1, y_2, \dots, y_m)$ is the posterior distribution of X given all m indicator variables,

$f(y_i | x)$ ($i=1, 2, \dots, m$) is the likelihood of indicator Y_i given X , and

$f(x)$ is the prior distribution of X .

The probability density $f(y_1, y_2, \dots, y_m)$ ensures proper normalisation and can be found by integrating the numerator; in practice this is usually done by numerical approximation.

Kernel density estimation

To find the posterior density function we must first have estimates of all required likelihood and prior densities. The likelihood is found using $f(y_i | x) = f(x, y_i) / f(x)$, $i=1, 2, \dots, m$, where $f(x, y_i)$ is the joint probability distribution of x and y_i . The kernel density estimate of the joint probability function based on a training data set of n observations $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ is:

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n k(x - x_i)$$

with a univariate Gaussian kernel and window width of h_x , that is

$$k(x) = \frac{1}{\sqrt{2\pi}h_x} \exp\left\{-\frac{1}{2} \frac{x^2}{h_x^2}\right\}, -\infty < x < +\infty$$

In practice, the ranges of x is truncated to make implementation easier, for example $-5h_x < x < 5h_x$. There are many approaches to estimation of the window width parameter h_x (see Silverman (1986) for examples). The simplest says that $\hat{h}_x = n^{-1/5} s_x$, where s_x is the sample standard deviation of the x values.

The kernel density estimate of the joint probability function based on the training data set is:

$$\hat{f}(x, y) = \frac{1}{n} \sum_{i=1}^n k(x - x_i, y - y_i)$$

where the kernel function chosen in this example is the general bi-variate Gaussian, window widths h_x, h_y and correlation ρ , with form:

$$k(x, y) = \frac{1}{2\pi h_x h_y \sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2\sqrt{1-\rho^2}} \left(\frac{x^2}{h_x^2} - \frac{2\rho xy}{h_x h_y} + \frac{y^2}{h_y^2} \right) \right\}, \quad -\infty < x, y < +\infty$$

Again the ranges of x and y are truncated. The window width parameters h_x , h_y and ρ are selected as functions of the sample variances and correlations from the training data.