*Journal of Paleolimnology* **10**: 147–152, 1994.



*Note*

# WACALIB version 3.3 - a computer program to reconstruct environmental variables from fossil assemblages by weighted averaging and to derive sample-specific errors of prediction

J.M. Line[1], Cajo J.F. ter Braak[2] & H.J.B. Birks[3*]

[1]*University of Cambridge Computer Laboratory, Pembroke Street, Cambridge CB2 3QG, UK;*[2]*Agricultural Mathematics Group-DLO, Box 100, 6700 AC Wageningen, The Netherlands and DLO-Institute for Forestry and Nature Research, Box 23, 6700 AA Wageningen, The Netherlands;*[3]*Botanical Institute, University of Bergen, Allégaten 41, N-5007 Bergen, Norway and Environmental Research Centre, University College London, 26 Bedford Way, London WC1H OAP, UK;* Author for correspondence*

Received 30 September 1993; accepted 19 January 1994

*Key words:* weighted averaging, regression, calibration, bootstrapping, maximum likelihood, reconstruction, error estimation

## Abstract

A computer program for reconstructing environmental variables (e.g. lake-water pH) from fossil assemblages (e.g. diatoms) by weighted averaging regression and calibration is described. The estimation of sample-specific errors of prediction by bootstrapping is outlined. The program runs on IBM-compatible personal computers.

## Introduction

Weighted averaging (WA) regression and calibration (ter Braak & van Dam, 1989) are rapidly becoming standard techniques in palaeolimnology and other areas of palaeoecology for the quantitative reconstruction of past environmental variables (e.g. lake-water pH, salinity, total phosphorus, surface-water temperature) from fossil assemblages. Examples include Agbeti (1992), Anderson *et al.* (1993), Birks *et al.* (1990a, 1990b), Cumming *et al.* (1992a), Fritz *et al.* (1991), Gaillard *et al.* (1992), Hall & Smol (1992), Janssens *et al.* (1992), Juggins (1992), Kingston *et al.* (1992), Walker *et al.* (1991), and Wilson *et al.* (1993). WA is also beginning to be re-used in palaeo-oceanography to estimate sea-surface temperatures (e.g. Le, 1992; ter Braak *et al.* 1993). WA regression is being used to estimate ecological optima and tolerances of species in modern comparative autecological studies (Jonsgard & Birks, 1993; Hjertholm, 1992).

Although not a new technique (see ter Braak (1987) for a review), the increasing use of WA in palaeolimnology is, in part, due to the improved theoretical understanding of WA regression and calibration, as summarised by ter Braak & Prentice (1998) and ter Braak & van Dam (1989), and, in part, due to the consistently good or better performance of WA compared with the computationally more intensive and difficult, but theoretically more rigorous, approach of maximum likelihood (ML) regression and calibration (ter Braak & van Dam, 1989; Birks *et al.*, 1990a; Kingston & Birks,

1990; Cumming *et al.* 1992a, ter Braak *et al.*, 1993). In addition the availability of the computer program WACALIB version 2.1 (Line & Birks, 1990) may have contributed to the increased use of WA. WACALIB was designed to run interactively, to be easy to use, and to be totally compatible with the widely used CANOCO program of ter Braak (1990a) for indirect and direct gradient analyses.

Recent work with simulated data sets (ter Braak & Juggins, 1993; ter Braak *et al.* 1993) indicates that WA performs consistently well with noisy, taxon-rich compositional data containing many taxa that may be absent from a large proportion of the samples and extending over long (>3 standard deviation units; Hill & Gauch, 1980) ecological gradients. For less noisy data, an extension of WA called weighted averaging partial least squares (WA-PLS) (ter Braak & Juggins, 1993; ter Braak *et al.*, 1993) may outperform WA.

This note describes a substantially updated version of WACALIB 2.1 that extends WA regression and calibration to include estimation of sample-specific errors of prediction by bootstrapping, to provide improved ML calibration, to estimate WA tolerances of taxa more realistically, and to provide a range of editing facilities that are commonly required in the analysis of large modern training sets.

## WACALIB 3.3

The original WACALIB version 2.1 (Line & Birks, 1990) was designed to make WA regression, WA calibration, and ML calibration available to palaeolimnologists. It was specifically written for the palaeolimnological projects within the Surface Waters Acidification Programme (SWAP) where it was used for all reconstructions of lake pH from fossil diatom assemblages (Battarbee *et al.*, 1990; Birks *et al.*, 1990a, 1990b). As part of the SWAP and the Paleoecological Investigations of Recent Lake Acidification II (PIRLA-II - see Charles & Smol, 1990) projects, WACALIB was extended to implement error estimation for individual fossil samples using bootstrapping (Birks *et al.*, 1990a) (see Birks *et al.* (1990b), Cumming *et al.* (1992a, 1992b), Kingston *et al.* (1992), Dixit *et al.* (1993), Cumming & Smol (1993a, 1993b), and Pienitz *et al.* (1994) for applications). Version 3.3 incorporates bootstrapping error estimation, as well as some additional theoretical improvements in WA regression and calibration, practical improvements in ML calibration, and extended editing facilities within WACALIB.

### *Editing facilities*

Besides the deletion of specific modern and fossil samples as in WACALIB 2.1, taxa can now be deleted if they occur in fewer than a specified number of samples, or if they occur only with abundance values smaller than a user-specified minimum. Either or both criteria can be applied. If both are used, they can be enforced jointly (taxa must occur at a specified abundance in the specified number of samples) or independently (taxa must occur in the specified number of samples and must also exceed the specified abundance level in at least *one* sample). In addition, particular taxa can be deleted explicitly by their taxon number. It is also possible to specify the minimum number of taxon occurrences within the modern training set for a taxon to be included in the WA regression and calibration (with or without bootstrap error estimation) on the basis of either the actual number of occurrences or the so-called "effective number" of occurrences, excluding those taxa already deleted. The effective number of occurrences for a taxon is estimated by N2, analogous to Hill's (1973) N2 diversity measure for samples which is the inverse of Simpson's diversity index (see ter Braak, 1990b). A taxon with 6 actual occurrences with values of say, 70%, 1%, 0.9%, 0.7%, 0.5%, and 0.1% will have its WA optimum effectively determined by the sample in which it occurs with an abundance of 70%. The N2 for this taxon is thus close to 1.

*Weighted averaging tolerance estimation*

The estimates of the WA tolerances for each taxon can be based, as in WACALIB 2.1, on the equation given in ter Braak (1987), ter Braak & van Dam (1989), and Birks *et al.* (1990a). However, this equation has no correction for bias resulting from the effective number of occurrences of the taxon. For an unbiased statistical comparison of tolerances (ter Braak, 1990b), the effective number of occurrences of the taxon in question, N2, should be taken into account (Hill, 1979). For presence-absence data, N2 is simply the actual number of occurrences. For quantitative data, N2 lies between 1 and the actual number of occurrences. In WACALIB 3.3 the tolerances for each taxon are corrected for bias by dividing by $(1-1/N2/)^{1/2}$. Uncorrected tolerances can also be calculated.

*Error estimation*

In WACALIB 3.3 the computer-intensive procedure of bootstrapping (Efron, 1982) is used to estimate the root mean square error (RMSE) of prediction for environmental inferences for all modern training samples, for the training set as a whole, and for all individual fossil samples.

The idea of bootstrap error estimation is to do many bootstrap cycles, say 1000. In each cycle, WACALIB selects at random but with replacement from the original training set a subset of training samples of the same size as the actual training set. As sampling is with replacement, some samples may be selected more than once in a cycle. Any samples not selected for the training set form a bootstrap *test* set for that cycle. WA regression and calibration are then used with the bootstrap training set to infer the environmental variable of interest for the modern samples (with known environmental variables) in the bootstrap *test* set. In each cycle, WA calibration is also used to infer the environmental value for each fossil sample. The standard deviation of the inferred values for both modern and fossil samples is calculated. This comprises one component ($s_1$) of the prediction error, namely that part due to estimation error in the optima and tolerances of the taxa. The second component ($s_2$), the error due to variations in the abundance of taxa at a given environmental value, is estimated from the training set by the root mean square (across all training samples) of the difference between the observed environmental value and the mean bootstrap estimate in all bootstrap cycles when that modern sample is in the bootstrap *test* set. The first component varies from fossil sample to fossil sample, the second component is constant. The estimated RMSE of prediction for a fossil sample is the square root of the sum of squares of these two components. The same procedure is applied to each modern training sample to derive sample-specific RMSE. The underlying theory is presented by Birks *et al.* (1990a).

WACALIB 3.3 computes and outputs several bootstrap-derived statistics including $RMSE_{pred}$ (the RMSE of prediction for samples in the training set), $s_1$, $s_2$, and Est $SE_{pred}$ (the estimated standard error of prediction for individual samples, modern or fossil). In addition, $Var_{opt}$, the variance of the optima for the taxa contributing to the final non-bootstrapped WA reconstruction and Hill's (1973) N2, a diversity measure for each sample and an estimate of the effective number of abundant taxa in each modern and fossil sample, are calculated.

Bootstrap error estimation is only available for WA regression and calibration, with or without inverse weighting by squared tolerances of each taxon corrected or uncorrected for bias, and with classical or inverse deshrinking regression. Birks *et al.* (1980a) discuss the virtues of these two deshrinking methods. Bootstrap error estimation is not available if user-supplied values of optima and/or tolerances for selected taxa are supplied (e.g. from literature sources) or in ML regression and calibration.

*Maximum likelihood calibration*

Given the Gaussian logit regression coefficients for the taxa in the training set (estimated by ML using, for example, GLIM (Payne, 1986) or GLR (Juggins, 1993)), WACALIB will per-

form ML calibration using an iterative Gauss-Newton numerical optimisation procedure with Gallant's (1975) stopping rule for step-shortening. The WA estimates are used as initial estimates for the ML calibration. In WACALIB 2.1 some samples often failed to converge with this procedure. In WACALIB 3.3 if non-convergence occurs, a direct search algorithm is used. This simply evaluates the log-likelihood function at intervals throughout the likely range of the environmental variable being reconstructed. The point giving the largest value of the log-likelihood function is the result for that search. As long as the function does not have any really abrupt changes of slope, the direct search result is a reasonable approximation to the true result. This may actually fall between the points at which the function is evaluated in the direct search procedure.

The choice of step-size for the direct search is chosen by the user. The range that is searched is the same range as for normal ML calibration. It can, however, be selected by the user, but it defaults to the observed range of the environmental variable in the modern training set extended by half the range in either direction. The range searched is thus double the observed range in the training set, but shares the same central point. The direct search algorithm is inevitably rather computer-intensive and the computing time depends on how small a step-size is chosen. The default is 0.01 units of the environmental variable.

*Compatibility*

Data-input formats are identical to WACALIB 2.1. In addition in version 3.3, an explicitly defined zero value in so-called condensed format is now interpreted as a very small non-zero value. Data files readable by WACALIB 3.3 are fully compatible with CANOCO 3.1x (ter Braak, 1990a) and CEDIT 3.2 (van Tongeren, 1991). CANOCO 3.1x implements principal components, correspondence, detrended correspondence, canonical correspondence, detrended canonical correspondence, redundancy (= canonical principal components), and canonical variates analyses. CEDIT 3.2 is an extremely versatile data-manipulation program for editing, transforming, merging, etc. data files. Both are invaluable adjuncts to WACALIB. For much reconstruction work CANOCO provides an essential tool for exploring and assessing taxon/environment relationships prior to using WACALIB. CEDIT is invaluable for data-set manipulations prior to using WACALIB.

Ter Braak & Juggins (1993) and ter Braak *et al.* (1993) have extended WA to include partial least squares regression. Juggins & ter Braak (1993) have developed a $C^{++}$ program (CALIBRATE) for WA regression and calibration and WA-PLS. Data-input formats in CALIBRATE are not fully compatible with WACALIB or CANOCO, as modern and fossil samples have to be in separate files for CALIBRATE. CALIBRATE performs both WA and WA-PLS but cannot supply sample-specific error estimates.

*Availability*

WACALIB 3.3 is written in standard FORTRAN 77 and is available on 3.5 inch or 5.25 inch diskettes for IBM-compatible personal computers. On an IBM-compatible PC running DOS 3.2 or later with 640 KB RAM and 560 KB free memory, WACALIB 3.3 can analyse 400 samples and 400 taxa with up to 17500 non-zero taxon values. A math-coprocessor $80\times87$ is recommended, especially for bootstrapping and ML calibration, but is not essential.

The executable version along with some test data, test output, annotated input and output, relevant publications, and associated utility software are available from H.J.B. Birks for U.S. $50 (£35). Payment as a U.S. cheque or as a sterling cheque should be sent with the order. Please state what size diskette the software, etc. should be delivered on. Colleagues in countries with currency-exchange problems can request a free copy.

Any colleague who has purchased WACALIB 2.1 can request from H.J.B. Birks a free copy of WACALIB 3.3.. Please indicate what size diskette the upgrade should be sent on.

## Acknowledgements

We are grateful to SWAP, NAVF, and PIRLA for partial support, and to Rick Battarbee, Hilary Birks, John Boyle, Brian Cumming, and Steve Juggins for many helpful discussions.

## References

Agbeti, M. D., 1992. Relationship between diatom assemblages and trophic variables: a comparison of old and new approaches. Can. J. Fish. aquat. Sci. 49: 1171–1175.

Anderson, N. J., B. Rippey & C. E. Gibson, 1993 A comparison of sedimentary and diatom-inferred phosphorus profiles: implications for defining pre-disturbance nutrient conditions. Hydrobiologia 253: 357–366

Battarbee, R. W., J. Mason, I. Renberg & J. F. Talling, 1990. Paleolimnology and lake acidification. The Royal Society, London.

Birks, H. J. B., J. M. Line, S. Juggins, A. C. Stevenson & C. J. F. ter Braak, 1990a. Diatoms and pH reconstruction. Phil. Trans. Roy. Soc. Lond. B, 327: 263–278.

Birks, H. J. B., S. Juggins & J. M. Line, 1990b. Lake surface-water chemistry reconstructions from palaeolimnological data. In The Surface Waters Acidification Programme (ed. B. J. Mason), pp. 301–313. Cambridge University Press, Cambridge.

Charles, D. F. & J. P. Smol, 1990. The PIRLA II Project: Regional assessment of lake acidification trends. Verh. Int. Ver. Limnol. 24: 474–480.

Cumming, B. F. & J. P. Smol, 1993a. Development of diatom-based salinity models for paleoclimatic research from lakes in British Columbia (Canada). Hydrobiologia 269/270: 179–196.

Cumming, B. F. & J. P. Smol, 1993b. Scaled chrysophytes and pH-inference models: the effects of converting scale counts to cell counts and other species data transformations. J. Paleolimnol. 9: 147–153.

Cumming, B. F., J. P. Smol, J. C. Kingston, D. F. Charles, H. J. B. Birks, K. E. Camburn, S. S. Dixit, A. J. Uutala & A. R. Selle, 1992b. How much acidification has occurred in Adirondack region lakes (New York, USA) since preindustrial times? Can. J. Fish. aquat. Sci. 49: 128–141.

Dixit, S. S., B. F. Cumming, H. J. B. Birks, J. P. Smol, J. C. Kingston, A. J. Uutala, D. F. Charles & K. E. Camburn, 1993. Diatom assemblages from Adirondack lakes (New York, USA) and the development of inference models for retrospective environmental assessment. J. Paleolimnol. 8: 27–47.

Effon, B., 1982d. The jackknife, the bootstrap and other resampling plans. Society for Industrial and Applied Mathematics NSF-CBMS Monograph 38: 1-92.

Fritz, S. C., S. Juggins, R. W. Battarbee & D. R. Engstrom, 1991. Reconstruction of past changes in salinity and climate using a diatom-based transfer function. Nature 352: 706–708.

Gaillard, M.-J., H. J. B. Birks, U. Emanuelsson & B. E. Berglund, 1992. Modern pollen/land-use relationships as an aid in the reconstruction of past land-uses and cultural landscapes: an example from south Sweden. Veget. Hist. Archaeobot. 1: 3–17.

Gallant, A. R., 1975. Nonlinear regression. Am. Statist. 29: 73–81.

Hall, R. I. & J. P. Smol, 1992. A weighted-averaging regression and calibration model for inferring total phosphorus concentration from diatoms in British Columbia (Canada) lakes. Freshwat. Biol. 27: 417–434.

Hill, M. O., 1973. Diversity and evenness: a unifying notation and its consequences. Ecology 54: 427–432.

Hill, M. O., 1979. DECORANA- a FORTRAN program for detrended correspondence analysis and reciprocal averaging. Cornell University, Ithaca, New York, USA.

Hill, M. O. & H. G. Gauch, 1980. Detrended correspondence analysis - an improved ordination technique. Vegetatio 42: 47–58.

Hjertholm, E., 1992. The autecology of four species of *Sedum* along a west-east climatic gradient, Sognefjorden, western Norway. Cand. Scient. thesis, University of Bergen, 83 pp.

Janssens, J. A., B. C. S. Hansen, P. H. Glaser & C. Whitlock, 1992. Development of a raised-bog complex. In The Patterned Peatlands of Minnesota (eds. H. E. Wright, B. A. Coffin & N. E. Aaseng), pp 189–221. University of Minnesota Press, Minneapolis.

Jonsgard, B. & H. J. B. Birks, 1993. Quantitative studies on saxicolous bryophyte-environment relationships in western Norway. J. Bryology 17: 579–611.

Juggins, S., 1993. GLR - a program for Gaussian logit regression. Unpublished program.

Juggins, S. & C. J. F. ter Braak, 1993. CALIBRATE - a program for species-environment calibration by [weighted-averaging] partial least squares regression. Unpublished program.

Kingston, J. C. & H. J. B. Birks, 1990. Dissolved organic carbon reconstructions from diatom assemblages in PIRLA project lakes, North America. Phil. Trans. Roy. Soc. Lond. B, 327: 279–288.

Kingston, J. C., H. J. B. Birks, A. J. Uutala, B. F. Cumming & J. P. Smol, 1992. Assessing trends in fishery resources and lake water aluminum from paleolimnological analyses of siliceous algae. Can. J. Fish. aquat. Sci. 49: 116–127.

Le, J., 1992. Palaeotemperature estimation methods: sensitivity test on two western equatorial Pacific cores. Quat. Sci. Rev. 11: 801–820.

Line, J. M. & H. J. B. Birks, 1990. WACALIB version 2.1 - a computer program to reconstruct environmental variables from fossil assemblages by weighted averaging. J. Paleolimnol. 3: 170–173.

Pienitz, R., J. P. Smol & H. J. B. Birks, 1994. Assessment of freshwater diatoms as quantitative indicators of past climatic change in the Yukon and Northwest Territories, Canada. Arctic and Alpine Research (submitted).

ter Braak, C. J. F., 1987. Unimodal models to relate species to environment. Doctoral thesis, University of Wageningen, 152 p.

ter Braak, C. J. F., 1990a. CANOCO - a FORTRAN program for canonical community ordination. Microcomputer Power, Ithaca, New York, USA.

ter Braak, C. J. F., 1990b. Update notes: CANOCO version 3.10. Agricultural Mathematics Group, Wageningen, 35 pp.

ter Braak, C. J. F. & S. Juggins, 1993. Weighted averaging partial least squares regression (WA-PLS): an improved method for reconstructing environmental variables from species assemblages. Hydrobiologia 269/270: 485–502.

ter Braak, C. J. F., S. Juggins, H. J. B. Birks & H. van der Voet, 1993. Weighted averaging partial least squares (WA-PLS): definition and comparison with other methods for species - environmental calibration. In Multivariate Environmental Statistics (eds. G. P. Patil & C. R. Rao), North Holland, Amsterdam: 525–560.

ter Braak, C. J. F. & I. C. Prentice, 1988. A theory of gradient analysis. Adv. Ecol. Res. 18: 271–317.

ter Braak, C. J. F. & H. van Dam, 1989. Inferring pH from diatoms: a comparison of old and new calibration methods. Hydrobiologia 178: 209–223.

Walker, I. R., R. J. Mott & J. P. Smol, 1991. Allerød-Younger Dryas lake temperatures from midge fossils in Atlantic Canada. Science 253: 1010–1012.

Wilson, S. E., I. R. Walker, R. J. Mott & J. P. Smol, 1993. Climatic and limnological changes associated with the Younger Dryas in Atlantic Canada. Climate Dynamics 8: 177–187.

van Tongeren, O. F. R., 1991. CEDIT program. Limnological Institute, Nieuwersluis, The Netherlands.