

Developments in Age-Depth Modelling of Holocene Stratigraphical Sequences

H.J.B. BIRKS AND EINAR HEEGAARD

Botanical Institute, University of Bergen, Bergen, Norway; John.Birks@bot.uib.no; Einar.Heegaard@bot.uib.no

Absolute chronology is the basis for all comparisons and correlations of Holocene stratigraphical proxy records. It is essential to develop independent absolute chronologies for such stratigraphic sequences before attempting comparisons, correlations, and syntheses, so as to avoid the well-known problems of visual "curve-matching" and associated dangers of the "reinforcement syndrome" and the "suck-in" phenomenon (Bennett, 2002a). A statistical approach for developing age-depth models is desirable so that the uncertainties in the radiometric age determinations can be taken into account and confidence intervals for the resulting models estimated and displayed.

As almost all Holocene stratigraphic sequences (e.g. peats, lake sediments) are from non-laminated sediments, an absolute chronology must be based on radiometric dating (^{14}C , also ^{210}Pb for the very recent past) and, if available, volcanic tephra of known age. Radiocarbon years do not equal calendar years. It is, however, possible to calibrate radiocarbon dates (with their associated uncertainties) for the Holocene into calibrated ages, thanks to the development of the internationally adopted INTCAL98 radiocarbon calibration data-set and the availability of calibration software such as CALIB, OXCAL, and BCAL. Age-depth modelling using calibrated ages is essential if correlations and comparisons are to be made with proxy records based on an absolute chronology, such as ice-cores or tree-rings. Such modelling using calibrated ages has, however, several statistical problems.

Developing a Chronology

The stages in developing an absolute chronology for a Holocene sequence based on a series of ^{14}C dates are (1) calibration of the ^{14}C dates, (2) statistical age-depth modelling using calibrated dates, and

(3) model selection and evaluation. Assuming that the ^{14}C dates are reliable, are not subject to problems of hard-water effects and incorporation of "old" carbon from other sources and, in the case of marine sequences, that the marine reservoir effect is known, the first step is calibration of the radiocarbon dates. Different calibration procedures can give different calibrated ages and very different ranges and probability distributions, especially if Bayesian approaches are adopted. In the Bayesian approach, it is possible to take account of other dates from the same sequence during the calibration of individual radiocarbon dates. If it is known that one date is younger than another because of their known stratigraphical relationship, it is possible to use this information to reduce the range of uncertainties for the calibrated ages of both dates (Bennett, 2002b). The irregular probability distribution functions of calibrated dates can be summarized in many ways (e.g. intercept, mode, median, weighted average) and the resulting summaries can have important influences on the final age-depth models. An additional potentially critical question is whether the radiocarbon calibration curve should be smoothed to match the potential resolution of the stratigraphical sequence of interest. The effect of such smoothing is usually to alter the uncertainties in the resulting calibrated ages.

Given a set of calibrated ages (and associated uncertainties) for a stratigraphical sequence, the next step is to convert the sample depths into estimated calibrated ages prior to comparison with other proxy records. This can readily be done by simple linear interpolation with confidence intervals estimated for the resulting age estimates (Bennett, 1994, 2002b). However, it is often also necessary to estimate sediment deposition times, fossil accumulation rates ("influx"), or

rates of biotic change. Here one needs to develop a statistical model for estimating calibrated ages for each sample based on a geologically realistic model of sediment accumulation. Linear interpolation between dated horizons is not appropriate here. Age-depth models are required that (1) are statistically reliable and robust, (2) provide estimates of the uncertainties of the calibrated age estimates for each sample, (3) take account of the uncertainties in the datings, radiocarbon calibrations, and sampling, (4) are geologically realistic, and (5) are easy to implement. Various methods such as linear interpolation, cubic splines, and regression models (e.g. polynomial least-squares regression) in the framework of generalized linear models (GLM) have been used (see Bennett, 1994, 2002b). The approach we have adopted (Heegaard and Birks, 2003) is weighted, non-parametric regression in the framework of generalized additive models (GAM) (Yee and Mitchell, 1991). In contrast to GLM, GAM models are driven primarily by the data themselves rather than by fitting prespecified GLM regression models to the data. Smoother functions are fitted to the data with appropriate prespecified link and error functions. Weightings can be introduced for all dated samples, with weights inversely proportional to the uncertainties in the calibrated ages. Well-dated horizons (e.g. tephra layers) or the lake-sediment core-top receive high weights in the regression, whereas calibrated ages within a radiocarbon plateau, and hence with a large calibration uncertainty, receive low weights.

Choosing the Model

After fitting a series of GAM models with different error specifications and smoothers using different degrees of freedom to the available calibrated ages and their associ-

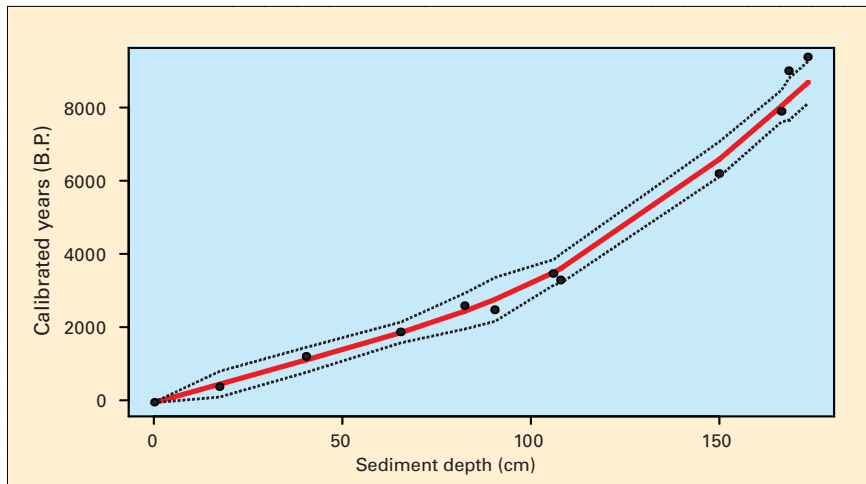


Figure 1: The relationship between sediment depth and calibrated years B.P. for Gåvåli lake, Dovre, Norway (data are kindly lent to us by Wenche Eide). The red line is the expected calibrated year for a particular sediment layer. The dotted lines represent the 95% confidence interval based on the insecurity of the observed calibrated dates and the regression line in combination. Here we used a constant variance in the regression.

ated uncertainties, the next step is to select the simplest parsimonious model (called the "minimal adequate model" in statistical model building), namely the simplest statistically significant solution that uses the fewest terms in the model and the fewest numbers of degrees of freedom in the fitted smoother. Model selection is done by examining regression diagnostic plots and PRESS statistics. The point-wise standard errors for the mean regression response are added to the standard error of the individual age estimates to calculate 95% confidence intervals for the entire age-depth model. All these analyses are implemented as a S-PLUS macro (Heegaard and Birks, 2003). The resulting age-depth model and associated age estimates and 95% confidence intervals should be critically evaluated by comparison with independently derived age estimates from the same geographic region. This GAM approach allows the use of calibrated ^{14}C dates, ^{210}Pb dates, tephra, and other independently dated horizons in one age-depth model and has been found to produce realistic age-depth relationships for a range of lake-sediment sequences in Norway, Sweden, Finland, Austria, Switzerland, Greenland and the UK.

An age-depth model is statistically based and as such includes various assumptions: (1) the datings used are reliable, (2) the radiocarbon calibrations are reliable and

the weighting functions used are realistic, (3) sample age is known with a larger error than sample depth, so we are justified in regressing age on depth rather than depth on age, (4) the principle of parsimony, (5) there are no sedimentary disturbances in the sequence and there are no gaps due to coring artifacts or errors, and (6) the mean radiocarbon date is independent of its variance. The final choice is the simplest adequate model possible statistically and has the largest degrees of freedom. It should therefore be relatively robust and have good predictive power. However, the model is a working hypothesis to be tested by independent criteria. Being a statistical model, it has uncertainties and all Holocene age estimates generally have associated 95% confidence intervals of between 50-200 calibrated years. Such intervals can easily encompass rapid events of short duration (Bennett, 2002a). If age-depth modelling is applied consistently to different proxy records at different sites, this should provide a robust basis for comparing and correlating many Holocene proxy records. Future developments should be directed at trying to reduce the uncertainties of the resulting age estimates. Decomposition of the total model variance suggests that to reduce the model uncertainties from 100-200 to 25-50 years, uncertainties in the radiocarbon calibration need to be reduced by 75%, assuming

that other sources of uncertainty (sampling, radiocarbon assay, etc.) remain the same.

Conclusion and Outlook

In comparing and interpreting Holocene stratigraphical records, more attention needs to be paid to the uncertainties associated with the age-depth models. The uncertainties provide a guide for the basis of any correlation (Bennett, 2002a). In the correlation of an "event" between sequences, the 95% confidence intervals of the age estimates should consistently overlap but as Bennett (2002a) notes, the correlation may still be valid even if a 1 in 20 confidence interval fails to overlap. However, when most or all of the confidence intervals do not overlap, the proposed correlation should be rejected. Comparison and correlation of stratigraphical sequences are, in reality, hypotheses that should be tested and falsified using statistical criteria (Bennett, 2002a). Statistically based age-depth models are valuable because they force the researcher to be explicit about the assumptions of the model and the quality of the data used. How much confidence do we really have in the basic radiocarbon age determinations and in the subsequent calibrations? Age-depth modelling is perhaps one of the weakest areas in Holocene research at present and it is an area that urgently requires further attention. Such work is currently in progress at the Bjerknes Centre for Climate Research in Bergen.

REFERENCES

- Bennett, K.D., 1994: Confidence intervals for age estimates and deposition times in late-Quaternary sediment sequences. *The Holocene* **4**: 337-348.
- Bennett, K.D., 2002a: Comment: the Greenland 8200 cal. yr BP event detected in loss-on-ignition profiles in Norwegian lacustrine sediment sequences. *Journal of Quaternary Science* **17**: 97-99.
- Bennett, K.D., 2002b: Documentation for psimpoll 4.10 and pscomb 1.03. (<http://www.kv.geo.uu.se> under link to "Software"). Uppsala University.
- Heegaard, E. and Birks, H.J.B., 2003: Age-depth modelling of late-Quaternary sediment sequences (in preparation).
- Yee, T.W. and Mitchell, N.D., 1991: Generalized additive models in plant ecology. *Journal of Vegetation Science* **2**: 587-602.

