Quaternary Geochronology 14 (2012) 94–104

Contents lists available at SciVerse ScienceDirect

Quaternary Geochronology

journal homepage: www.elsevier.com/locate/quageo

Research paper

A comparison of different methods for speleothem age modelling

Denis Scholz [a,*], Dirk L. Hoffmann [b,c], John Hellstrom [d], Christopher Bronk Ramsey [e]

[a] Institute for Geosciences, University of Mainz, Johann-Joachim-Becher-Weg 21, 55128 Mainz, Germany
[b] CENIEH, Paseo Sierra de Atapuerca s/n, 09002-Burgos, Spain
[c] Bristol Isotope Group, School of Geographical Sciences, University of Bristol, University Road, BS8 1SS Bristol, United Kingdom
[d] School of Earth Sciences, The University of Melbourne, VIC 3010, Australia
[e] Oxford Radiocarbon Accelerator Unit, University of Oxford, Dyson Perrins Building, South Parks Road, Oxford OX1 3QY, UK

ARTICLE INFO

Article history:
Received 27 November 2011
Received in revised form
24 February 2012
Accepted 21 March 2012
Available online 9 April 2012

Keywords:
Speleothem
Chronology
Age modelling
Software
U-series dating

ABSTRACT

Speleothems, such as stalagmites and flowstones, can be dated with unprecedented precision in the range of the last 650,000 a by the ^{230}Th/U-method, which is considered as one of their major advantages as climate archives. However, a standard approach for the construction of speleothem age models and the estimation of the corresponding uncertainty has not been established yet.

Here we apply five age modelling approaches (`StalAge`, OxCal, a finite positive growth rate model and two spline-based models) to a synthetic speleothem growth model and two natural samples. All data sets contain problematic features such as outliers, age inversions, large and abrupt changes in growth rate as well as hiatuses.

For data sets constrained by a large number of ages and not including problematic sections, all age models provide similar results. In case of problematic sections, the algorithms provide significantly different age models and uncertainty ranges.

`StalAge`, OxCal and the finite positive growth rate model are, in general, more flexible since they are capable of modelling hiatuses and account for problematic sections by increased uncertainty. The spline-based age models, in contrast, reveal problems in modelling problematic sections.

Application to the synthetic data set allows testing the performance of the algorithms because the 'true' age model is available and can be compared with the age models. OxCal and `StalAge` generally show a good performance for this example, even if they are inaccurate for a short section in the area of a hiatus. The two spline-based models and the finite positive growth rate model show larger inaccurately modelled sections.

1. Introduction

Speleothems such as stalagmites and flowstones are increasingly used as palaeoclimate archives due to their almost worldwide occurrence and because they offer long, often continuous, high-resolution climate records. Their greatest advantage in comparison to other climate archives, which are often difficult to date beyond the ^{14}C-dating limit, is that ^{230}Th/U-dating enables age determination with unprecedented precision in the range of the last 650,000 a (Richards and Dorale, 2003; Scholz and Hoffmann, 2008; Cheng et al., 2009).

Since the spatial resolution of determined ages is usually lower than that of the proxy analyses, the age between two adjacent U-series ages needs to be estimated. This is usually done by calculating a relationship between ages determined at certain distances and the distance of the proxy analyses along the growth axis of the speleothem. This relationship is usually referred to as the *age-depth model*.

Age modelling poses a challenge for all climate archives, and a variety of approaches have been used. Interestingly, the details of the methods applied for age modelling and in particular for the estimation of age uncertainty are often insufficiently described. A recent literature review, for instance, showed that 65 out of 93 publications relying on some kind of age model do not discuss the uncertainties associated with the age models (Blaauw, 2010).

For speleothems, a variety of different methods have been applied. The most frequently used method is linear point-to-point interpolation between dated sub-sample distances (e.g., McDermott et al., 1999; Wang et al., 2005), but least-squares polynomial fits (Spötl and Mangini, 2002) and various kinds of splines (Beck et al., 2001; Spötl et al., 2006; Vollweiler et al., 2006; Hodge et al., 2008) have also been used. Furthermore, methods

* Corresponding author. Tel.: +49 6131 3924767; fax: +49 6131 3924769.
E-mail address: scholzd@uni-mainz.de (D. Scholz).

1871-1014/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.quageo.2012.03.015

based on the general growth mechanisms of speleothems (Drysdale et al., 2004; Genty et al., 2006) were applied. These methods are precursors of the finite positive growth rate model presented here (compare Section 2.4). Recently, Scholz and Hoffmann (2011) have presented an algorithm particularly developed for speleothem age modelling. However, a *standard* approach for construction of speleothem age models or to estimate the corresponding uncertainty has not been established yet. This is particularly surprising because the precise determination of the timing and duration of climatic events is considered as the major advantage of speleothems as climate archives.

Here we compare five different age modelling approaches by applying them to synthetic as well as natural speleothem age data. All data sets are non-trivial and represent specific challenges because they include potential outliers, age inversions, hiatuses and sections with substantial changes in growth rate. We compare the resulting age models and the corresponding uncertainties and discuss the specific advantages and disadvantages of the different methods.

2. Description of the different age modelling techniques

2.1. `StalAge` *(Scholz and Hoffmann, 2011)*

`StalAge` (Scholz and Hoffmann, 2011) is written in the open source statistical software `R` (R Development Core Team, 2011) and has been particularly designed for construction of speleothem age models. In addition to the individual ages and the corresponding uncertainties, `StalAge` includes stratigraphic information in order to further constrain and improve the age model. It is applicable to problematic data sets including outliers, age inversions, hiatuses and large changes in growth rate, and potentially inaccurate ages are automatically identified. The age model and the corresponding 95%-confidence limits are calculated by a Monte-Carlo simulation fitting ensembles of straight lines to sub-sets of the age data.

The algorithm has no adjustable free parameters and, thus, offers the highest degree of reproducibility and comparability.

2.2. *Mixed effect regression model of Heegaard et al. (2005)*

The Heegaard et al. (2005) age model utilises a generalized mixed effect regression combining the uncertainty of the individual ages (within-object variance) with an estimation of the uncertainty of the age distribution as a whole (between-object variance). The code used for age modelling is, as for `StalAge`, available as a function for R (R Development Core Team, 2011). The algorithm uses a cubic smooth spline function, and the ‘best fit’ is determined using the Generalized Cross-Validation provided by Wood (2001) in the `mgcv` library of R.

Two parameters can be adjusted by the user in order to increase and reduce the roughness of the fit, respectively. vspan is a parameter of the local regression used in the diagnostic plot, and *k* denotes the number of splines used in the cubic smooth spline regression (Heegaard et al., 2005). This enables the user, at least to some extent, to adjust the age model according to his/her point of view.

For speleothems, the Heegaard et al. (2005) algorithm was applied to an MIS 7 speleothem from the Austrian Alps in combination with Bayesian statistics for chronological ordering (Spötl et al., 2008).

2.3. *Smoothing cubic spline*

An age model can also be generated using a smoothing cubic spline following the procedures outlined in Beck et al. (2001) or Spötl et al. (2006). The approach here is identical to the one applied in Hoffmann et al. (2010) and uses the `smooth.spline` function of the `‘stats’` package of R (R Development Core Team, 2011) in order to derive a spline fit and associated 95%-confidence limits. A Generalized Cross-Validation approach is used to determine the best fit based on paired distance-age data and their associated uncertainty. A weight for each predictor value is derived from the inverse of the associated variance. Lower and upper confidence limits are calculated using the smoothing spline results object and the uncertainties of the age data.

The `smooth.spline` function uses a variety of adjustable parameters including equivalent degree of freedom (df) or smoothing parameter (spar). Parameters can, thus, be manually manipulated to ‘improve’ the fit. Here, automatic derivation of the parameters is used to achieve a maximum of reproducibility and exclude personal biases. The result of the smoothing spline is not necessarily monotonic as required for a stalagmite age model, especially where significant ‘outliers’ are observed in the ^{230}Th–U-ages. In cases where the automatic spline yields sections with inversions of the age model, user dependent modifications may be necessary. In this case, either a parameter like equivalent degree of freedom (df) is restricted to the maximum value that maintains monotonicity or, where unambiguous ‘outliers’ cause age inversions, the input data set needs to be modified accordingly. Where there are significant cessations in growth, the spline should ideally be developed for individual growth sections since this approach cannot identify and account for hiatuses.

2.4. *Finite positive growth rate model (Drysdale et al., 2004, 2005; Genty et al., 2006; Hendy et al., in press)*

This model has evolved from a Bayesian technique, based on an implementation in *Isoplot* (Ludwig, 2003; Drysdale et al., 2004), in which all ages are randomised within their uncertainties and only the sub-population in which all ages are in stratigraphic order is retained for age model calculation. This model was found to unreasonably minimise periods of rapid growth rate and was replaced with a Monte-Carlo approach that randomises all dates within their uncertainty and for each iteration conducts a least-squares fit of a sequence of line segments for which growth rate must be positive but less than an arbitrary threshold (usually 1 mm/yr, Drysdale et al., 2005; Genty et al., 2006).

In the most recent version of this model, which is used here, this arbitrary constraint has been replaced with a requirement that relative change in growth rate is minimised between age determinations according to a user-defined parameter (Hendy et al., in press), similarly as in a model recently presented by Blaauw and Christen (2011). In all three versions of this model, growth rate is allowed to vary randomly between age determinations with a user-defined intensity in each iteration in order to allow for interpolation uncertainty. In a final step, all iterations are ranked to determine the median model and its uncertainty.

2.5. *Deposition models implemented in OxCal (Bronk Ramsey, 2008)*

The P_Sequence model from OxCal (Bronk Ramsey, 2008) assumes monotonic growth and is based on an underlying random Poisson process. The models may also include outlier analysis (Bronk Ramsey, 2009), and this has been applied in all cases here.

For the P_Sequence model there is one user option, the *k* parameter, which defines the rigidity of the model: an infinite *k* implies a uniform deposition rate, and a *k* of zero assumes no more than that the deposition occurs in order. Since there is not much experience of which values of *k* are suitable for speleothems, in this