



Effects of dating errors on nonparametric trend analyses of speleothem time series

M. Mudelsee^{1,2}, J. Fohlmeister³, and D. Scholz⁴

¹Climate Risk Analysis, Hannover, Germany

²Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, Germany

³Heidelberg Academy of Sciences, Heidelberg, Germany

⁴Johannes Gutenberg University Mainz, Institute for Geosciences, Mainz, Germany

Correspondence to: M. Mudelsee (mudelsee@climate-risk-analysis.com)

Received: 2 May 2012 – Published in *Clim. Past Discuss.*: 30 May 2012

Revised: 7 September 2012 – Accepted: 21 September 2012 – Published: 23 October 2012

Abstract. A fundamental problem in paleoclimatology is to take fully into account the various error sources when examining proxy records with quantitative methods of statistical time series analysis. Records from dated climate archives such as speleothems add extra uncertainty from the age determination to the other sources that consist in measurement and proxy errors. This paper examines three stalagmite time series of oxygen isotopic composition ($\delta^{18}\text{O}$) from two caves in western Germany, the series AH-1 from the Atta Cave and the series Bu1 and Bu4 from the Bunker Cave. These records carry regional information about past changes in winter precipitation and temperature. U/Th and radiocarbon dating reveals that they cover the later part of the Holocene, the past 8.6 thousand years (ka). We analyse centennial- to millennial-scale climate trends by means of nonparametric Gasser–Müller kernel regression. Error bands around fitted trend curves are determined by combining (1) block bootstrap resampling to preserve noise properties (shape, autocorrelation) of the $\delta^{18}\text{O}$ residuals and (2) timescale simulations (models StalAge and iscam). The timescale error influences on centennial- to millennial-scale trend estimation are not excessively large. We find a “mid-Holocene climate double-swing”, from warm to cold to warm winter conditions (6.5 ka to 6.0 ka to 5.1 ka), with warm–cold amplitudes of around 0.5 ‰ $\delta^{18}\text{O}$; this finding is documented by all three records with high confidence. We also quantify the Medieval Warm Period (MWP), the Little Ice Age (LIA) and the current warmth. Our analyses cannot unequivocally support the conclusion that current regional winter climate is warmer than that during the MWP.

1 Introduction

We can learn about the climatic variations over time during the past by examining natural archives. This task comprises (1) sampling an archive, (2) measuring on the samples a proxy variable indicative of climate, (3) dating the samples, and (4) statistically analysing the obtained proxy time series to infer the properties of the climatic process that generated the data. This paper uses the speleothem archive, $\delta^{18}\text{O}$ as proxy variable and U/Th as well as radiocarbon technology as dating tools. The statistical question we are elucidating regards centennial- to millennial-scale trends in the climate of western Germany during the later part of the Holocene.

It is well established that the measurement error of $\delta^{18}\text{O}$ in a speleothem is smaller than the “proxy error” because other considerable factors than climate influence $\delta^{18}\text{O}$ (Fairchild and Baker, 2012). These error sources and internal variability of the complex, nonlinear climate system (Lorenz, 1963) already make an inference (of trends) inexact and require us to report the uncertainty of a statistical analysis (Mudelsee, 2010). Estimates without error bars are useless. (With this statement we wish to say that one should give the measures of error, such as bias or standard deviation, when reporting the result of a statistical analysis. We do not mean that published records of paleoclimate are useless. In fact, the scientists measuring and establishing those records are mostly experts and do perform excellently when reporting the scope and limitations of the methods they develop and use.)

It is equally well established that also dating a sample has an error (Taylor, 1987; Ivanovich and Harmon, 1992).

In addition, construction of a curve of age against depth in a climate archive (Sect. 2) has to make assumptions, which may be violated, adding to the uncertainty of a time value in a proxy time series. It is fair to say that statistical theory and practice have largely ignored the effects of dating errors; notable exceptions are in the field of spectral analysis (Akaike, 1960; Moore and Thomson, 1991; Thomson and Robinson, 1996), see also Mudelsee et al. (2009). Climate time series analysis has only recently become aware of the need – and developed simulation-based capabilities – to fully include timescale errors for obtaining realistic uncertainty measures of an analysis (Mudelsee et al., 2009; Haam and Huybers, 2010; Mudelsee, 2010; Huybers, 2011; Rhines and Huybers, 2011; Stanford et al., 2011; Shakun et al., 2012).

In line with these recent advances, this paper aims to study the effects of dating errors on nonparametric trend analyses, which has been done neither in the work cited in the previous paragraph nor, to the best of our knowledge, in other work. Although our findings are obtained on speleothem $\delta^{18}\text{O}$ series, we expect them to be of relevance for analyses on other archives and other proxy variables as well.

A further purpose of this paper is to study the effects of using different age-modelling algorithms. On the other hand, we do not intend to discuss the merits of the algorithms themselves, and we do not make a recommendation which one to use.

The structure of this paper is as follows. Section 2 details sampling, measuring and dating and also gives an overview of construction of age–depth curves on the basis of absolutely dated depth points. Section 3 explains nonparametric regression for geoscientists with some background in statistical science. Section 4 presents results and discusses regional climate variations. Section 5 contains the conclusions.

2 Data

The speleothems analysed here are stalagmites stemming from two caves in the karstic Sauerland region, western Germany. Stalagmite AH-1 (Niggemann et al., 2003), from the Atta Cave ($51^{\circ}07'30''\text{N}$, $7^{\circ}54'56''\text{E}$), has a length of 61 cm. Stalagmites Bu1 and Bu4 (Fohlmeister et al., 2012), from the Bunker Cave ($51^{\circ}22'03''\text{N}$, $7^{\circ}39'53''\text{E}$), have lengths of 65 and 20 cm, respectively; these stalagmites grew separated by a distance of 12 m.

2.1 Dating points

AH-1 was dated using U/Th mass spectrometry at 13 depth points (measured as the distance from the top of the stalagmite). We here add information from radiocarbon dating of the topmost point (Table 1). Note that Table 1 shows slight deviations in U/Th ages and their errors from the original publication (Niggemann et al., 2003), which are due to a recalibration of the U and Th spike solutions used in the

Table 1. Dating points, stalagmite AH-1. Age errors are given conventionally as $2\text{-}\sigma$ standard deviation.

| Depth (cm) | Age (ka) | Age error (ka) |
|------------|----------|----------------|
| 0.1 | −0.03 | 0.02 |
| 2 | 0.819 | 0.067 |
| 4.5 | 1.61 | 0.1 |
| 10.2 | 1.666 | 0.063 |
| 14.5 | 2.37 | 0.11 |
| 19.3 | 2.511 | 0.078 |
| 25.7 | 3.47 | 0.17 |
| 30 | 4.17 | 0.23 |
| 34 | 4.25 | 0.12 |
| 39.5 | 4.58 | 0.38 |
| 43.7 | 5.5 | 0.18 |
| 48.2 | 5.44 | 0.1 |
| 50.2 | 6.87 | 0.16 |
| 55.3 | 8.51 | 0.087 |

Heidelberg U/Th laboratory (Hoffmann et al., 2007). All ages reported in this paper are relative to AD 1950. Bu1 and Bu4 were dated using U/Th mass spectrometry at 10 and 11 depth points, respectively, and using radiocarbon information on the topmost point (Fohlmeister et al., 2012). It was found that Bu1 shows a hiatus and problematic ages between 14.6 and 30 cm depth; hence, Bu1 is analysed separately in an early and a late part.

2.2 Age–depth curves

In general, construction of an age–depth curve utilizes information from the absolute dating points. It makes the additional assumption of a growth of the archive with time, that means, a strictly monotonically increasing age–depth curve. Still, the curve is not uniquely defined by the above, and additional assumptions are invoked, which ideally consider the physics of the archive's growth and, possibly, prior information (e.g. from other, neighbouring archives). It is therefore not astonishing that several approaches to the construction have been developed in different scientific communities (which often associate themselves with the type of archive typically employed). It is also reasonable to adopt Bayesian methods, which are designed to utilize prior information. The following is an overview of age–depth construction tools that can be used in a situation as here (dated depth points); it is based on the book by Mudelsee (2010, pp. 172–173 therein). We focus on tools that are also able to generate simulated age–depth curves and error bands that reflect the various sources of uncertainty (dating errors, violated assumptions).

A simple age–depth curve is a straight regression line fitted to the dating points. Also, polynomials of higher order can be used without technical effort. From the regression residuals follows an error measure (mean squared error), which can be compared with the typical size of the dating errors.

Agreement would then corroborate the validity of the estimated age–depth curve (e.g. its linear form). Papers on this approach include Bennett (1994) on the chronology of a lake sediment core dated with radiocarbon and Spötl et al. (2006) on a stalagmite dated with U/Th. Simulated curves can be generated by means of a random number generator using the dating errors for the individual depths and refitting the regression line. Drysdale et al. (2004) use a similar approach, also on a stalagmite, but make an additional, apparently ad-hoc assumption that stalagmite growth may vary by a factor of ten between the dating points, resulting in a wider error band. Bennett and Fuller (2002) study the influence of the age-model selection on the estimated date of the mid-Holocene decline of the hemlock tree in eastern North America. In the presence of hiatuses, it should be worth applying regression models with a jump in the mean (Mudelsee, 2010), additionally constrained to monotonic growth, to the construction of age–depth curves with error bands. Heegaard et al. (2005) offer an interesting extension to the case where the material in an archive at a depth is age-inhomogeneous. This can occur in sediment cores as a result of mixing processes. Bayesian tools for constructing chronologies are described in book length by Buck and Millard (2004). Some Bayesian examples are the papers by Agrinier et al. (1999) on a geomagnetic polarity record from the Cretaceous–Cenozoic, Blaauw and Christen (2005) on a proxy record from a climate archive in the form of a Holocene peat-bog core and Klauenberg et al. (2011) on glaciologically modelled chronologies for late Pleistocene ice cores. A new, continuous, piecewise linear, monotone stochastic process has been suggested (Haslett and Parnell, 2008) to model accumulation of a climate archive. These authors also present applications to radiocarbon-dated lake sedimentary records.

StalAge and iscam (intra-site correlation age modelling) are two recent approaches particularly designed to calculate age–depth curves and the corresponding uncertainty for speleothems. We also detail in the following the adaptations of StalAge and iscam for the purpose of generating simulated age–depth curves as an input to the statistical analysis.

2.2.1 StalAge

The StalAge algorithm (Scholz and Hoffmann, 2011) uses U-series ages and their corresponding age uncertainty for age–depth modelling and also includes stratigraphic information to constrain further and improve the age model. StalAge has been shown to be suitable for problematic data sets that include outliers, age inversions, hiatuses and large changes in growth rate. Potentially inaccurate ages are identified automatically, and manual selection of outliers prior to application is, thus, not required. To offer the highest degree of reproducibility and comparability for different studies, StalAge has – unlike many other methods – no adjustable free parameters. The algorithm consists of three major steps: first, major outliers are identified. Second, the age data are screened

for minor outliers and age inversions, and the uncertainty of potentially problematic data points is increased using an iterative procedure. Third, the age model and corresponding 95 % confidence limits are calculated by means of a Monte Carlo simulation fitting ensembles of straight lines to subsets of the age data. The algorithm is written in the open source statistical software R, and the code is available as an electronic supplement of the original publication (Scholz and Hoffmann, 2011), which also provides further details and examples.

Recently, StalAge has been applied to three speleothem records containing problematic sections, and its performance has been compared with four other methods (Scholz et al., 2012). 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, however, the algorithms provide clearly different age models as well as uncertainty ranges (Scholz et al., 2012). StalAge is capable of modelling hiatuses and accounts for problematic sections by increased uncertainty. Two spline-based age–depth construction models, in contrast, produced difficult-to-interpret results for the problematic record sections.

Although StalAge uses a Monte Carlo simulation to calculate the 95 % confidence limits, it does not directly provide simulated age–depth curves for all proxy depths as is required for the resampling methods applied in this study. Thus, the required 2000 age–depth curves are simulated as follows.

1. We first conventionally apply StalAge to the dating points obtaining a “best” age model (Fig. 1a, c, e) and the corresponding 95 % confidence limits for each data set. This provides a probability distribution for each sample depth. Note that the confidence limits obtained from StalAge are usually highly asymmetric. In rare cases, the resulting age model may still show age inversions (Scholz and Hoffmann, 2011), as is, for instance, the case for the age model calculated for stalagmite AH-1. StalAge thus performs a further test for monotonicity and, if necessary, a final correction of the “best” age model subsequent to this step (which has been omitted in the simulations for technical reasons).
2. In a second step, we simulate 2000 monotonically increasing age–depth curves (with a depth resolution of 2 mm) by randomly sampling from the probability distribution provided by StalAge. To simulate the curves within the provided probability distributions, we use a Markov chain Monte Carlo approach (the Gibbs sampler) similarly as described in previous publications (Spötl et al., 2008; Mudelsee et al., 2009). Since the final screening step performed by StalAge is omitted here (and the probability distributions for the AH-1 data thus include an age inversion), the Gibbs sampler does not find monotonically increasing curves for the AH-1 data set. To provide the 2000 simulated curves for AH-1, the

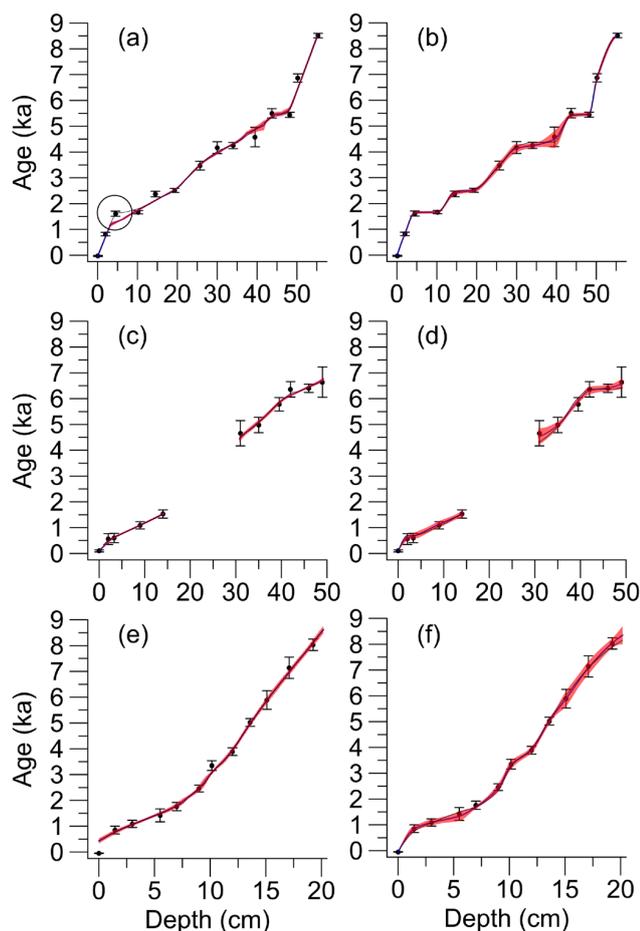


Fig. 1. Age–depth curves for stalagmites AH-1 (a, b), Bu1 (c, d) and Bu4 (e, f) constructed with algorithms StalAge (a, c, e; Scholz and Hoffmann, 2011) and iscam (b, d, f; Fohlmeister, 2012). Shown are the dating points (black dots) with the $2\text{-}\sigma$ dating errors (vertical black bars), the constructed age–depth curves (thin black lines) and the average from 2000 simulated age–depth curves (thick blue line – legible only in the zoomed electronic paper) with the corresponding $2\text{-}\sigma$ standard-error band (red shading). The circle in (a) indicates a systematic shift of simulated ages, which results from the methods used for obtaining the simulated curves (see Sects. 2.2.1 and 4.2).

region including the inversion (at around 5 cm depth, see Fig. 1a) has to be omitted from the simulation. This explains the systematic deviation between the original StalAge age model and the mean of the 2000 simulated curves for AH-1 (Fig. 1a, circle). The other data sets (Bu1 and Bu4) do not include problematic sections, and the original StalAge age model is within the range of the simulated age models (Fig. 1c, e).

3. In a final step the 2000 simulated age–depth curves are interpolated (within R) to the high-resolution proxy depths.

2.2.2 iscam

Another recently published approach to age–depth curve construction (Fohlmeister, 2012) is called intra-site correlation age modelling (iscam). It employs a random number generator to account for the dating error of depth points (Fig. 1). The additional constraints come from other contemporaneously grown, neighbouring stalagmites, which exhibit a similar oxygen isotope pattern preserved in the calcite to the original “reference” stalagmite. The assumption is that those age–depth model solutions in which the $\delta^{18}\text{O}$ records show higher correlation coefficients are more likely to represent the true age–depth relationship than random model solutions with lower correlations. This additional information is thought to lead to more accurate age–depths curves for periods of simultaneous growth.

To produce random age–depth simulations from iscam, we generate a very large number of realisations and take the first 2000 where the correlation coefficient exceeds the $2\text{-}\sigma$ significance level. The significance level is estimated by means of artificial time series prescribed to reflect similar statistical (i.e. autocorrelation) and sampling properties to the measured record (note that iscam assumes a normal distributional shape, which should not be strongly violated in case of the $\delta^{18}\text{O}$ records).

The age–depth model of Bu1 is calculated in relation to Bu4. However, Bu4 grew also during other intervals than Bu1 (Fig. 1). That means, the age–depth curve of Bu4 in such no-overlap intervals is unconstrained by Bu1 and may therefore show larger uncertainty than in the overlap intervals. On the other hand, since the growth periods of Bu4 and AH-1 are almost identical (Fig. 1), the individual age–depth relationships for both stalagmites can be calculated together via selecting the high-correlation age–depth realisations.

2.3 Proxy variable

Measurements of the $\delta^{18}\text{O}$ proxy variable were performed on the carbonate material (see Fairchild and Baker, 2012, for a general reference). The number, n , of data points is 380 for stalagmite AH-1 (Niggemann et al., 2003), 615 for the earlier part of stalagmite Bu1, 474 for the later part of stalagmite Bu1 and 1008 for stalagmite Bu4 (Fohlmeister et al., 2012). By means of the age–depth curves (Fig. 1), we obtain proxy time series. These records are unevenly spaced. (Even if the stalagmites had been sampled at even depth spacing (which was not the case), the resulting time series would have been unevenly spaced owing to a variable growth of the archive.) Cave monitoring and general physical–climatological considerations (Niggemann et al., 2003; Riechelmann et al., 2011; Fairchild and Baker, 2012; Fohlmeister et al., 2012) show that for the stalagmites analysed here, $\delta^{18}\text{O}$ variations are a proxy for changes in winter precipitation and temperature, with high (low) $\delta^{18}\text{O}$ values indicating dry/cold (wet/warm) conditions (“dry/warm” and

“wet/cold” combinations can be excluded, as demonstrated by Fohlmeister et al. (2012)). This certainly holds at the regional spatial scale (western Germany), but likely also beyond (Central and Northern Europe, North Atlantic). Note that “changes in precipitation” refers not only to the amount but also to the provenance of the precipitation and the trajectory of the air parcel that transported it (Fairchild and Baker, 2012, Ch. 3 therein).

3 Method

Trend is a property of genuine interest in climatology (Brückner, 1890; Hann, 1901; Köppen, 1923). It describes the mean state of the atmosphere or other compartments in the climate system on long timescales (conventionally 30 yr and longer). We adopt a simple decomposition of a stochastic climate process in discrete time (Mudelsee, 2010),

$$X(i) = X_{\text{trend}}(i) + S \cdot X_{\text{noise}}(i), \quad (1)$$

where $X(i)$ is the climate variable at time $T(i)$, i is an index, $X_{\text{trend}}(i)$ is the trend component and $X_{\text{noise}}(i)$ is the noise component with zero mean and unity standard deviation. Scaling the noise with a constant variability, S , is considered a reasonable assumption for our analysis of the (relatively stable, homoscedastic) Holocene regional climate. A continuous-time version of the “climate equation” (Eq. 1) is readily written as

$$X(T) = X_{\text{trend}}(T) + S \cdot X_{\text{noise}}(T). \quad (2)$$

The statistical task is to estimate $X_{\text{trend}}(T)$ on the basis of the time series data, $\{t(i), x(i)\}_{i=1}^n$, where $t(i)$ are the time and $x(i)$ the $\delta^{18}\text{O}$ values.

Instead of identifying $X_{\text{trend}}(T)$ with a specific linear or nonlinear function with parameters to be estimated, the nonparametric regression or smoothing method estimates the trend at a time point by, loosely speaking, averaging the data points $X(i)$ within a time neighbourhood around the time point. The idea is that the noise values cancel each other out to some degree and the trend can be extracted. A simple example is the running mean, where the points inside a window are averaged and the window runs along the time axis. Statistical science (Silverman, 1986; Härdle, 1990; Simonoff, 1996) recommends replacing the non-smooth weighting window (points inside receive constant, positive weight and points outside zero weight) by a smooth kernel function, K . A nonparametric kernel regression estimator of the trend is given by Gasser and Müller (1979, 1984),

$$\widehat{X}_{\text{trend}}(T) = h^{-1} \sum_{i=1}^n \left[\int_{s(i-1)}^{s(i)} K\left(\frac{T-y}{h}\right) dy \right] x(i), \quad (3)$$

where the “hat” indicates the estimation, T is the time for which the trend is calculated, h is the bandwidth, y is an

auxiliary variable and the sequence $s(i)$ fulfils the condition $t(i-1) \leq s(i-1) \leq t(i)$.

We use the following discretization: T is calculated over the sample interval $[t(1); t(n)]$ with an even, fine spacing of 1 a. We use the following sequence: $s(0) = 1.5t(1) - 0.5t(2)$, $s(i-1) = [t(i-1) + t(i)]/2$ for $i = 2, \dots, n$ and $s(n) = 1.5t(n) - 0.5t(n-1)$.

A measure of the estimation uncertainty, an error band, is essential for interpreting the results, for judging whether ups and downs of $\widehat{X}_{\text{trend}}(T)$ are statistically significant or not. Error-band construction in general has to take into account typical climate noise properties in the form of non-normal distributional shape and autocorrelation (also called serial dependence or persistence). Bootstrap resampling (Efron and Tibshirani, 1993; Davison and Hinkley, 1997) is a state-of-the-art approach to achieving this. By resampling from the regression residuals, the distributional shape is preserved. By resampling time blocks of residuals, denoted as moving overlapping block bootstrap or MBB resampling (Künsch, 1989), the persistence is preserved. Mudelsee (2010, Algorithm 3.1 therein) gives a description of the MBB accessible to geoscientists.

The bootstrap involves repeating the estimation (Eq. 3) on each of the B resampled time series (Fig. 2). With the adopted recommendation of $B = 2000$ (Efron and Tibshirani, 1993) and the typical data sizes of the stalagmite records (Sect. 2.3), the Fortran implementation of kernel estimation with MBB resampling means only minor computational costs (a few seconds on a modern PC).

Further technical details are as follows.

1. We adapt the source code from the University of Zurich, Division of Biostatistics (<http://www.biostat.unizh.ch/research/software/kernel.html>, last access: 3 May 2007). It utilizes a parabolic kernel function and lets the bandwidth shrink towards the interval boundaries: $h \rightarrow 0$ for $T \rightarrow t(1)$ or $T \rightarrow t(n)$; and $h = \text{const.} > 0$ elsewhere.
2. We use the block length selector via n and the parameter of a first-order autoregression for uneven time spacing (Mudelsee, 2002) fitted to the residuals (Mudelsee, 2010, Eq. 3.28 therein).

Timescale errors

Up to now in this section on nonparametric regression, we have assumed that in a stochastic, discrete-time process $\{T(i), X(i)\}_{i=1}^n$, the times $T(i)$ are exactly known, whereas we have conceded some noise from measurement, proxy and climate uncertainties, described by the time-dependent random variable $X_{\text{noise}}(i)$, to the climate variable $X(i)$. In the nonparametric regression problem, this assumption leads to Eq. (3) with the regressor “time” fixed and known. This model is called fixed-regressor model.

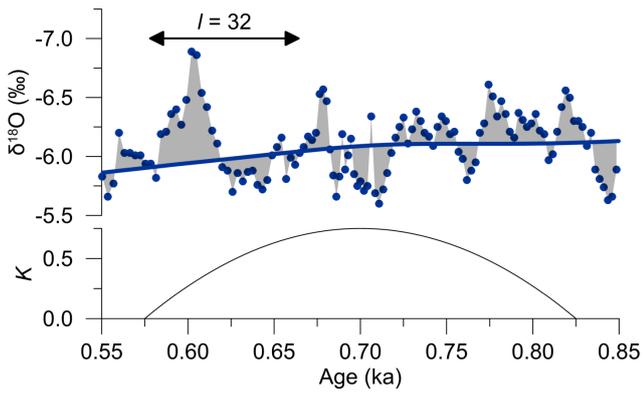


Fig. 2. Kernel trend estimation and MBB resampling illustrated by means of the Bu1 $\delta^{18}\text{O}$ record (blue dots, upper panel) during the interval from 0.55 to 0.85 ka on the StalAge-derived timescale. Also shown are trend estimate (blue line), regression residuals (indicated by the shading), parabolic kernel function K with a bandwidth of $h = 250$ a (black line, lower panel) and one block of length $l = 32$ for MBB resampling.

For several types of climate time series with exactly known $T(i)$ (e.g. climate model output or instrumental observations), the fixed-regressor model is adequate. On the other hand, for the climate archives studied here, the assumption $T(i) = T_{\text{true}}(i)$ (true time value) cannot be maintained (Sect. 2.2). In principle, we have to write the measured times as

$$T(i) = T_{\text{true}}(i) + T_{\text{noise}}(i), \quad (4)$$

$i = 1, \dots, n$, where $T_{\text{noise}}(i)$ is the error owing to age uncertainties, and we have to insert this expression into the calculation of the $s(i)$ sequence for the estimation (Eq. 3). The resulting regression model is called errors-in-variables model.

Since the typical size of the dating errors is rather small against the spread of the time points, the effects on the estimation itself (e.g. bias) are likely negligible (Mudelsee, 2010). For example, the AH-1 chronology has an average $1-\sigma$ dating error of 67 a (Table 1), and the time points for AH-1 constructed with StalAge have a $1-\sigma$ standard deviation or spread of $\left\{n^{-1} \sum_{i=1}^n [t(i) - n^{-1} \sum_{i=1}^n t(i)]^2\right\}^{1/2} \approx 1838$ a.

However, it is the objective of this paper to take into account timescale errors for the construction of error bands. This can be achieved by augmenting MBB resampling with parametric timescale simulation. This leads to the algorithm used for the calculations (Fig. 3). The idea behind this algorithm is roughly as follows. Whereas the usual MBB (Mudelsee, 2010) consists in doing steps 1–5 and then using in step 7 the original times $t(i)$ instead of the simulated times $t^*(i)$, our adaptation includes timescale simulation (step 6) and using in step 7 the $t^*(i)$. We show (Sect. 4.2) that ignoring timescale errors usually leads to error bands that are too narrow.

| | | |
|----|---|---|
| 1 | Data | $\{t(i), x(i)\}_{i=1}^n$ |
| 2 | Nonparametric regression (Eq. 3), inserted trend value for time $t(i)$ | $\hat{x}_{\text{trend}}(i) = h^{-1} \sum_{s(i-1)}^n \left[\int_{s(i-1)}^{t(i)-h} K\left(\frac{t(i)-y}{h}\right) dy \right] x(i),$ $i = 1, \dots, n$ |
| 3 | Regression residuals | $e(i) = x(i) - \hat{x}_{\text{trend}}(i), \quad i = 1, \dots, n$ |
| 4 | Apply MBB (Fig. 2) to $\{e(i)\}_{i=1}^n$ | $\{e^*(i)\}_{i=1}^n$ |
| 5 | Resampled $x(i)$ | $x^*(i) = \hat{x}_{\text{trend}}(i) + e^*(i), \quad i = 1, \dots, n$ |
| 6 | Simulated timescale (Section 2.2) | $t^*(i), \quad i = 1, \dots, n$ |
| 7 | Resample (b , counter) | $\{t^{*b}(i), x^{*b}(i)\}_{i=1}^n$ |
| 8 | Replication (continuous time, T) | $\hat{x}_{\text{trend}}^{*b}(T)$ |
| 9 | Go to 4 until $b = B$ replications exist | $\{\hat{x}_{\text{trend}}^{*b}(T)\}_{b=1}^B$ |
| 10 | Standard error given by | |
| | standard deviation over B replications | $\left\{B^{-1} \sum_{b=1}^B \left[\hat{x}_{\text{trend}}^{*b}(T) - B^{-1} \sum_{b'=1}^B \hat{x}_{\text{trend}}^{*b'}(T) \right]^2\right\}^{1/2}$ |
| 11 | Standard error-band given by | |
| | trend estimate (continuous time, T) | |
| | \pm standard error | |

Fig. 3. Algorithm for construction of standard-error bands for nonparametric regression estimates under consideration of timescale errors. The MBB (step 4) draws with replacement random blocks of length l from the residuals until n values have been resampled.

4 Results and discussion

We first look in a more qualitative manner at the overall results (Sect. 4.1). The effects of the dating errors are then studied separately for each of the two age–depth curve construction tools (Sect. 4.2). The next, critical point is how results depend on the type of age–depth model construction (Sect. 4.3). To study this dependence in detail, we perform a new simulation experiment. With the knowledge generated by this procedure, we look again at the overall results, the regional climate during the Holocene, in a more quantitative manner (Sect. 4.4).

4.1 Qualitative inspection

At first sight, the agreement of resulting trends among the three stalagmite $\delta^{18}\text{O}$ records AH-1, Bu1 and Bu4 is remarkable for each of the two age–depth construction tools (StalAge, Fig. 4a; iscam, Fig. 4b). We emphasize that no shift or correction has been applied to the $\delta^{18}\text{O}$ values – nevertheless their overall levels and the amplitudes of the trends are rather similar. It appears that the deviations between the AH-1 and each of the Bunker Cave records (Bu1 and Bu4) are larger than those between Bu1 and Bu4.

Note that the time series plots (Figs. 2, 4, 5, 6, 7 and 8) conventionally show time t on the horizontal and time series value x on the vertical axis.

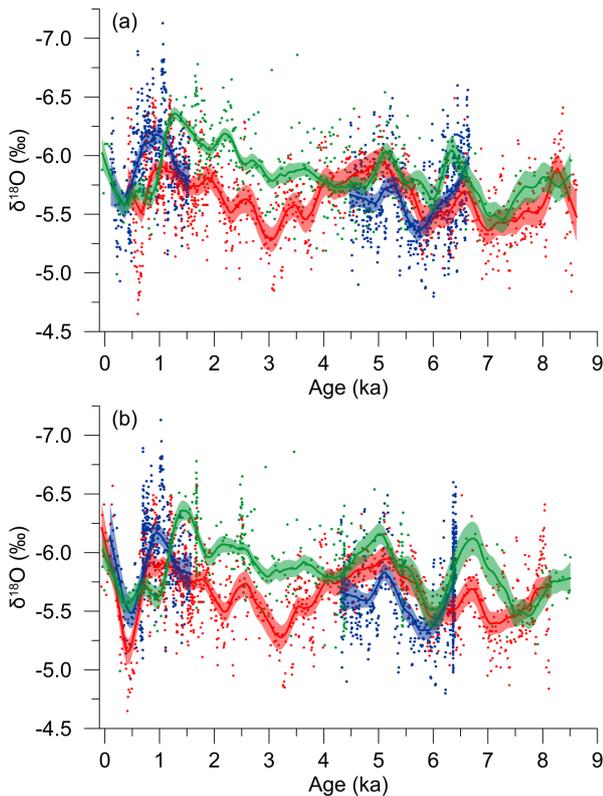


Fig. 4. Results. Trend estimation of stalagmite time series for age–depth construction algorithms StalAge (a) and iscam (b). Data points (dots), trends (thick solid lines) and $1-\sigma$ standard-error bands (shaded) are shown for stalagmites AH-1 (green), Bu1 (blue) and Bu4 (red). The early and late parts of the Bu1 series are analysed separately. The trends are calculated with a kernel bandwidth of $h = 250$ a and the standard-error bands with a number of MBB resamplings of $B = 2000$ and taking dating errors into account. The block lengths for MBB resampling are for (a): $l = 1$ (AH-1), 33 (Bu1, late part), 32 (Bu1, early part) and 22 (Bu4); and for (b): $l = 1$ (AH-1), 32 (Bu1, late part), 33 (Bu1, early part) and 22 (Bu4).

4.2 Effects of dating errors I

The nonparametric regression residuals (Fig. 2) are realisations of one class of noise: uncertainties of the x values in the form of measurement noise, proxy error and unknown/unresolved climate variability. The MBB algorithm (Fig. 3) takes this noise class fully into account (i.e. it preserves distributional shape and persistence) by means of the resample values $x^*(i)$ for calculating the error band around the nonparametric trend estimate. The second class of noise is on the t values in the form of dating errors at the construction of the age–depth curve. The error-band construction algorithm (Fig. 3) takes this into account by means of the resample values $t^*(i)$ generated by the timescale simulations (Sect. 2.2). We now study the pure effects of timescale errors by comparing two versions of error-band construction: one with timescale errors, the other without (i.e. setting $t^*(i) = t(i)$ for $i = 1, \dots, n$ in Fig. 3 at step 6).

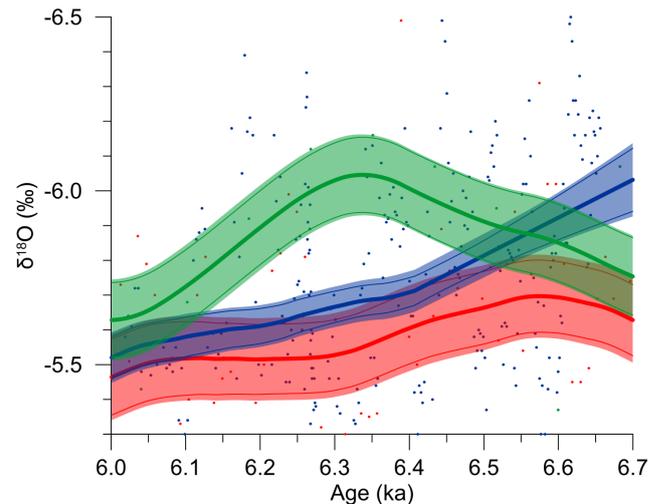


Fig. 5. Results. Effects of timescale errors on centennial- to millennial-scale trend estimation of stalagmite time series, interval from 6.0 to 6.7 ka, for age–depth construction algorithm StalAge. Cf. Fig. 4a; $1-\sigma$ standard-error bands resulting from taking timescale errors into account (shaded) and standard-error bands from ignoring timescale errors (thin solid lines).

The effects of timescale errors, exemplified by a zoom of the interval from 6.0 to 6.7 ka (Fig. 5), are not large. The conclusion for this time interval, namely that trends for the two records Bu1 and Bu4 from the same cave are rather close to each other while the record AH-1 from the other cave exhibits a trend with about 0.4 ‰ lighter $\delta^{18}\text{O}$ values at around 6.3 ka, is almost unaffected. The average contribution over the 6.0–6.7 ka interval from dating errors to the overall error band for algorithm StalAge is only 20.5 % (Bu1), 14.6 % (Bu4) and 8.2 % (AH-1). The average contribution over the same time interval for algorithm iscam is 2.4 % (Bu1), 11.3 % (Bu4) and 0.4 % (AH-1). Although the scatter among those percentages is considerable, it appears clear that timescale errors do not contribute much to the overall error of the estimation.

It is interesting to note that timescale errors on nonparametric trend estimates do not always widen the error bands. The error band for the AH-1 record on the StalAge-derived timescale (Fig. 6) shows for the interval around 1.3 ka a clearly narrower band than from ignoring timescale errors. The explanation of this effect is as follows. Trend estimates are more accurate for “high-density” data regions (Silverman, 1986; Härdle, 1990; Simonoff, 1996), where the time-dependent temporal spacing is small. The AH-1 series on the StalAge-derived timescale has such a high-density region at around 1.7 ka (Fig. 6). The average StalAge simulation shifts the region at around 1.7 ka systematically to smaller ages, by an amount of approximately 0.4 ka (Fig. 1a, circle). This means a shift of the high-accuracy region from 1.7 to 1.3 ka. The resulting narrowing effect on error bands at around 1.3 ka is stronger than the widening effect of the

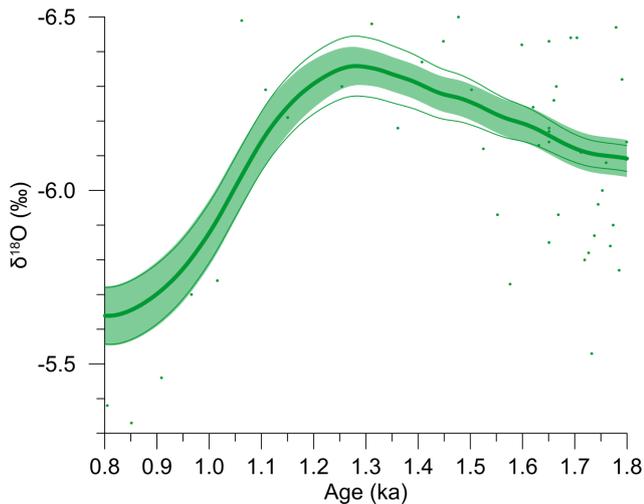


Fig. 6. Results. Effects of timescale errors on trend estimation of stalagmite time series, interval from 0.8 to 1.8 ka, for age–depth construction algorithm StalAge and stalagmite AH-1. Cf. Fig. 4a.

inclusion of timescale errors. We emphasize that this may in general only happen when the timescale simulation produces systematic shifts.

The main climatological objective of this paper is to study centennial- to millennial-scale trends in regional climate. For this purpose we have selected a bandwidth of $h = 250$ a for the trend estimation (Figs. 2 and 4). These estimations yield error bands that are rather robust against the presence of timescale errors; see Figs. 5 and 6 as well as the second paragraph in this Sect. 4.2. Let us look at decadal-scale trends by means of selecting a bandwidth of $h = 30$ a. Several new “wiggles” appear (Fig. 7) compared to using $h = 250$ a (Fig. 5). The immediate question relates to the statistical significance of these additional wiggles. The expectation of a reduced robustness against the presence of timescale errors is quantitatively confirmed (Fig. 7). For the 6.0–6.7 ka interval, the average contribution from dating errors to the overall error band for algorithm StalAge is as large as 39.4 % (Bu1), 55.8 % (Bu4) and 22.7 % (AH-1). The average contribution over the same time interval for algorithm iscam is 20.9 % (Bu1), 56.8 % (Bu4) and 6.6 % (AH-1). It is difficult to see significant decadal-scale trends when timescale errors are of similar (decadal) size (Fig. 1). Notwithstanding this verdict, there may be agreement on decadal-scale trends between the two Bunker Cave records Bu1 and Bu4; we note the wet/warm peak in both records at around 6.25 ka, preceded by a dry/cold peak at around 6.30 ka (Fig. 7). The difficulty to find significant peaks in decadal-scale climate is likely exacerbated when also considering “age–depth model uncertainty” (Sect. 4.3).

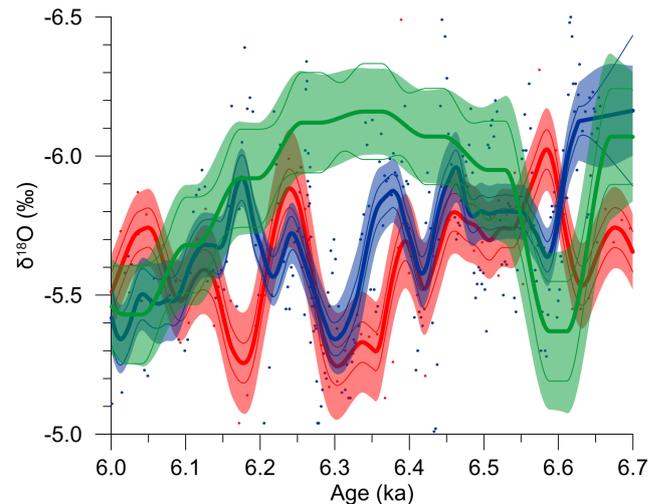


Fig. 7. Results. Effects of timescale errors on decadal-scale trend estimation of stalagmite time series, interval from 6.0 to 6.7 ka, for age–depth construction algorithm StalAge. The trends are calculated with a kernel bandwidth of $h = 30$ a. Cf. Figs. 4a and 5; block lengths: $l = 1$ (AH-1), 12 (Bu1, early part) and 6 (Bu4).

4.3 Effects of dating errors II

Let us take a rational, critical, scientific standpoint (Kant, 1781; Popper, 1935; Einstein, 1949) and consider that a true but unknown age–depth curve exists and that the task of curve construction (Sect. 2.2) is to infer the true age–depth curve. The problem then is that different algorithms for achieving this task can be used, which lead to different inference results and also to different timescale uncertainties. The question which model to use (for age–depth curve construction) is referred to in the statistical literature as model uncertainty; see Chatfield (1995), Draper (1995), Candolo et al. (2003) and Chatfield (2004, Sect. 13.5 therein). It appears to us that it should be solved not entirely by means of statistical model-selection criteria but rather be considered at a deeper level, touching areas of philosophy of science. In fact, an irrational, “post-normal” standpoint (Ravetz, n.d.; Fleck, 1980), which holds that truth is a sociological construct, seems not to lead to a way of advancing our understanding. These remarks do of course apply not only to age–depth construction for stalagmite records but to a wider class of inferential problems in climatology (“unknown unknowns”), for example the question of how to deal with uncertainties in greenhouse-gas emissions for climate forecasting.

The stalagmite problem is that the two different models of age–depth curve construction lead to two different Holocene climate trend estimates and error bands: model StalAge (Fig. 4a) and model iscam (Fig. 4b). The two schools of statistical thinking may inspire following practical approaches to tackle this model-uncertainty problem. Bayesian model averaging attaches prior probabilities to the age-model candidates. Chatfield (1995, p. 48) reports: “*The*

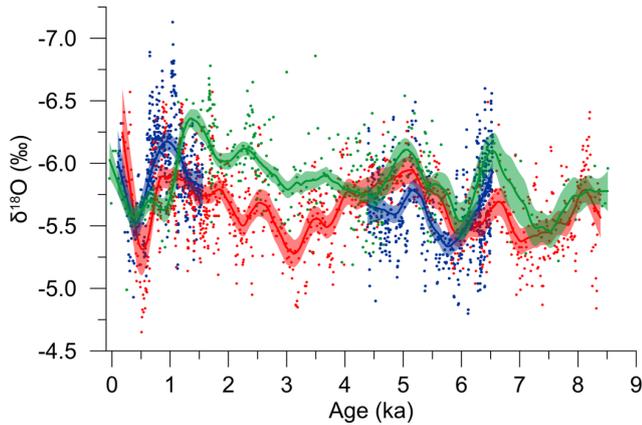


Fig. 8. Results. Trend estimation of stalagmite time series for a combination of age–depth construction algorithms StalAge and iscam. Cf. Fig. 4; block lengths: $l = 1$ (AH-1), 32 (Bu1, late part), 33 (Bu1, early part) and 22 (Bu4).

data are then used to evaluate posterior probabilities for the various models. Models with “low” posterior probabilities may be discarded to keep the problem manageable, and then a weighted sum of the remaining competing models is taken”. Since we have not taken the Bayesian approach but one based on resampling (Fig. 3), we prefer to bootstrap the modelling process; see, for example, Efron and Gong (1983) or Roberts and Martin (2010). In detail (cf. the algorithm in Fig. 3): first, at step 1, we average the time values, $t(i) = 0.5[t_{\text{StalAge}}(i) + t_{\text{iscam}}(i)]$, $i = 1, \dots, n$, where the subscript denotes the timescale model; second, at step 6, we take the first $B/2$ timescale resamples $t^*(i)$ from StalAge and the second $B/2$ from iscam. Thus, our procedure corresponds to a uniform prior or an unweighted resampling.

The resulting nonparametric trend estimates (Fig. 8) show some mild widening (relative to results shown in Fig. 4) of the error bands owing to model uncertainty. A larger effect is on the result for the late part of the $\delta^{18}\text{O}$ record from Bu4, on which individual StalAge and iscam results display considerable age deviations (Fig. 4).

4.4 Quantitative interpretation

On the basis of the trend estimation with a full consideration of the various error sources (measurement, proxy, timescale, model uncertainty), we infer the significant features of the climate evolution in western Germany over the past 8.6 ka (Fig. 8).

The early part of the analysed time interval does not show signs of a “recovery” from the previous glacial climatic state (Davis et al., 2003; Davis and Brewer, 2009) in the form of a long-term warming trend. This is likely due only to the relatively late start of our analysed records; older speleothems from Bunker Cave do show an early-Holocene warming trend (Fohlmeister et al., 2012).

Around the middle of the Holocene, all three records go into a swing from warm ($T \approx 6.5$ ka) to cold ($T \approx 6.0$ ka) and again to warm ($T \approx 5.1$ ka), with warm–cold amplitudes of around 0.5 ‰. Strictly speaking, we should write “wet/warm” instead of “warm” and “dry/cold” instead of “cold”. We should also bear in mind that kernel smoothing is not without bias (Härdle, 1990) and that the amplitude estimate of 0.5 ‰ may be too small. It therefore seems premature to assign certain temperature amplitudes to the observed swings. Nevertheless, this is a significant feature of winter-climate evolution in Europe. Note that the “Holocene climate optimum” between 5 and 9 ka regards another climate element, namely Northern Hemisphere summer (Jansen et al., 2007). It is an interesting question whether records from other locations in Europe do reveal a similar double swing in winter climate when examined with advanced methods of time series analysis.

The interval from approximately 5 to 1 ka, covered by only one stalagmite from Bunker Cave (Bu4) and the other from Atta Cave (AH-1), experienced “wiggly” trends of longer wetting/warming interrupted by phases of reduced drying/cooling (Fig. 8). Evidence of a quantitative agreement between Bu4 and AH-1 is not strong; AH-1 has generally lighter $\delta^{18}\text{O}$ values than Bu4, and Bu4 shows, unlike AH-1, a clear dry/cold peak at around 3 ka. A considerable amount of local climate influences and even inter-cave variability seems to prevail.

The warming trends described in the previous paragraph reach a maximum; for AH-1 this maximum is at $T \approx 1.4$ ka or AD 550 and $\delta^{18}\text{O} \approx -6.4$ ‰, for Bu4 it is at $T \approx 0.85$ ka or AD 1100 and $\delta^{18}\text{O} \approx -5.9$ ‰ and for Bu1, which commenced to grow again at around 1.5 ka, this wet/warm maximum is at $T \approx 0.95$ ka or AD 1000 and $\delta^{18}\text{O} \approx -6.2$ ‰. The scatter among these peak values again reflects local climatic influences; it also illustrates “taxonomic difficulties” in defining an MWP (Medieval Warm Period) or Medieval Climate Anomaly (Jansen et al., 2007, Box 6.4 therein).

The inter- and intra-record spread of values is rather small for the most recent centennial-scale cold, the LIA (Little Ice Age). It peaked at $T \approx$ AD 1550 and $\delta^{18}\text{O} \approx -5.6$ ‰ (AH-1), $T \approx$ AD 1550 and $\delta^{18}\text{O} \approx -5.5$ ‰ (Bu1) and $T \approx$ AD 1450 and $\delta^{18}\text{O} \approx -5.3$ ‰ (Bu4). Since the LIA, a warming trend, enhanced by anthropogenic contributions (Solomon et al., 2007), has persisted to the present. The warming is rather fast, although it is not straightforward to place realistic error bars around the rate of warming. However, such an estimation target (first derivative) is accessible to a quantitative analysis by means of kernel smoothing (Gasser and Müller, 1984). As regards the often discussed levels of the current warmth in relation to MWP levels, the results (Fig. 8) allow us only to conclude that current warmth is stronger (lighter $\delta^{18}\text{O}$) than MWP for stalagmite Bu4, about the same for Bu1 and weaker for stalagmite AH-1.

Finally, we want to emphasize that at least parts of the $\delta^{18}\text{O}$ signal could have been altered by changes in the

strength of kinetic isotope fractionation (Fairchild and Baker, 2012) between the two caves. This may be the reason for differences in the trends between AH-1 and both Bunker Cave stalagmites. Even when we use *iscam* for the age modelling, there are periods where differences in the timing of the trends occur (e.g. from 2.5 to 2 ka). As shown for Bunker Cave, the artificial cave opening in the late 19th century had no influence on the oxygen isotopic composition of the calcite (Fohlmeister et al., 2012). Therefore, a rapid opening or closing of the cave does not seem to have had an influence on the degree of kinetic fractionation, whereas stalagmite-growth modelling studies (Mühlinghaus et al., 2009; Scholz et al., 2009; Dreybrodt and Scholz, 2011) suggest that changes in drip rate may indeed have an influence.

5 Conclusions

1. To assess the significance of climate trends requires in general the construction of error bands that capture fully the error sources. For proxy records from speleothems, the error sources regard measurement, proxy quality, dating and age–depth curve construction.
2. The three analysed stalagmite $\delta^{18}\text{O}$ time series reflect regional, western-Germany winter climate over the past 8.6 ka. The error influences on centennial- to millennial-scale trend estimation are not excessively large, allowing us to infer real climatic features. One such apparently previously less well recognized feature is the “mid-Holocene climate double swing” from warm to cold to warm winter conditions (6.5 ka to 6.0 ka to 5.1 ka). We also quantify the MWP and LIA as expressed in the stalagmite records.
3. One should be aware of a general potential for circular reasoning regarding results from age–depth construction algorithm *iscam*. Since *iscam* employs two series and assumes a correlation between them to construct an age–depth curve, a good agreement between trends may also result from the stretching and shifting of timescales imposed by *iscam*. In case of the results shown in the present paper, however, the danger of circular reasoning is rather small owing to the fact that also the other age–depth construction algorithm (StalAge) yields trends that exhibit a similar amount of agreement among different stalagmites.
4. Differences between trends shown by Atta Cave (AH-1) on the one hand and Bunker Cave (Bu1 and Bu4) on the other likely reflect varying small-scale climatic and hydrological influences and possibly also effects of kinetic fractionation. Variations in terms of provenance or trajectories (Sect. 2.3) are an unlikely explanation since both caves are as close as about 39 km to each other (Sect. 2).
5. Construction of age–depth curves for climate archives on the basis of dating points, constraints (e.g. strictly monotonically increasing curves) and the physics relevant for describing the archive’s growth is a challenging task. It is currently being tackled by means of Bayesian and other simulation-based approaches. Construction of these curves is, however, not a means in itself. Age–depth modelling must also provide simulated curves, which can then be fed into modern resampling methods of climate time series analysis, resulting in realistic measures of uncertainty in our knowledge about the climate.

Supplementary material related to this article is available online at: <http://www.clim-past.net/8/1637/2012/cp-8-1637-2012-supplement.zip>.

Acknowledgements. Troy Rasbury and Nereo Prato provided reviews, and Silvia Frisia gave editorial comments, all of which rich in detail, carefully prepared and constructive in tone. We hope to have adequately dealt with these contributions and produced a paper that is easier to access for readers than the previous version in *Climate of the Past Discussions*. We thank also Eric Wolff for helpful editorial comments. We prepared the manuscript bearing in mind the electronic version, where figures can be zoomed. The Supplement is a compressed archive of high-resolution versions of the figures that contain shading. We acknowledge gratefully financial support by the Deutsche Forschungsgemeinschaft (Bonn, Germany) via Research Unit FOR 668 “Dated Speleothems: Archives of the Paleoenvironment (DAPHNE)”.

Edited by: S. Frisia

References

- Agrinier, P., Gallet, Y., and Lewin, E.: On the age calibration of the geomagnetic polarity timescale, *Geophys. J. Int.*, 137, 81–90, 1999.
- Akaike, H.: Effect of timing-error on the power spectrum of sampled-data, *Ann. I. Stat. Math.*, 11, 145–165, 1960.
- Bennett, K. D.: Confidence intervals for age estimates and deposition times in late-Quaternary sediment sequences, *Holocene*, 4, 337–348, 1994.
- Bennett, K. D. and Fuller, J. L.: Determining the age of the mid-Holocene *Tsuga canadensis* (hemlock) decline, Eastern North America, *Holocene*, 12, 421–429, 2002.
- Blaauw, M. and Christen, J. A.: Radiocarbon peat chronologies and environmental change, *Appl. Stat.-J. Roy. St. C.*, 54, 805–816, 2005.
- Brückner, E.: Klimaschwankungen seit 1700 nebst Bemerkungen über die Klimaschwankungen der Diluvialzeit, *Geogr. Abh.*, 4, 153–484, 1890.
- Buck, C. E. and Millard, A. R. (Eds.): *Tools for Constructing Chronologies: Crossing Disciplinary Boundaries*, Springer, London, 2004.

- Candolo, C., Davison, A. C., and Demétrio, C. G. B.: A note on model uncertainty in linear regression, *Statistician*, 52, 165–177, 2003.
- Chatfield, C.: Model uncertainty, data mining and statistical inference (with discussion), *J. Roy. Stat. Soc. A Sta.*, 158, 419–466, 1995.
- Chatfield, C.: *The Analysis of Time Series: An Introduction*, 6th Edn., Chapman and Hall, Boca Raton, FL, 2004.
- Davis, B. A. S. and Brewer, S.: Orbital forcing and role of the latitudinal insolation/temperature gradient, *Clim. Dynam.*, 32, 143–165, 2009.
- Davis, B. A. S., Brewer, S., Stevenson, A. C., Guiot, J., and Data Contributors: The temperature of Europe during the Holocene reconstructed from pollen data, *Quaternary Sci. Rev.*, 22, 1701–1716, 2003.
- Davison, A. C. and Hinkley, D. V.: *Bootstrap methods and their application*, Cambridge University Press, Cambridge, 1997.
- Draper, D.: Assessment and propagation of model uncertainty (with discussion), *J. Roy. Stat. Soc. B*, 57, 45–97, 1995.
- Dreybrodt, W. and Scholz, D.: Climatic dependence of stable carbon and oxygen isotope signals recorded in speleothems: From soil water to speleothem calcite, *Geochim. Cosmochim. Ac.*, 75, 734–752, 2011.
- Drysdale, R. N., Zanchetta, G., Hellstrom, J. C., Fallick, A. E., Zhao, J., Isola, I., and Bruschi, G.: Palaeoclimatic implications of the growth history and stable isotope ($\delta^{18}\text{O}$ and $\delta^{13}\text{C}$) geochemistry of a middle to late Pleistocene stalagmite from Central–Western Italy, *Earth Planet. Sc. Lett.*, 227, 215–229, 2004.
- Efron, B. and Gong, G.: A leisurely look at the bootstrap, the jackknife, and cross-validation, *Am. Stat.*, 37, 36–48, 1983.
- Efron, B. and Tibshirani, R. J.: *An Introduction to the Bootstrap*, Chapman and Hall, London, 1993.
- Einstein, A.: Autobiographisches – autobiographical notes, in: *Albert Einstein: Philosopher – Scientist*, edited by: Schilpp, P., Library of Living Philosophers, Evanston, IL, 1–95, 1949.
- Fairchild, I. J. and Baker, A.: *Speleothem Science: from Process to Past Environments*, Wiley-Blackwell, Chichester, 2012.
- Fleck, L.: *Entstehung und Entwicklung einer wissenschaftlichen Tatsache*, Suhrkamp, Frankfurt am Main, reprint of the 1st Edn. (1935), 1980.
- Fohlmeister, J.: A statistical approach to construct composite climate records of dated archives, *Quat. Geochronol.*, in press, 2012.
- Fohlmeister, J., Schröder-Ritzrau, A., Scholz, D., Spötl, C., Riechelmann, D. F. C., Mudelsee, M., Wackerbarth, A., Gerdes, A., Riechelmann, S., Immenhauser, A., Richter, D. K., and Mangini, A.: Bunker Cave stalagmites: an archive for central European Holocene climate variability, *Clim. Past Discuss.*, 8, 1687–1720, doi:10.5194/cpd-8-1687-2012, 2012.
- Gasser, T. and Müller, H.-G.: Kernel estimation of regression functions, in: *Smoothing Techniques for Curve Estimation*, edited by: Gasser, T. and Rosenblatt, M., Springer, Berlin, 23–68, 1979.
- Gasser, T. and Müller, H.-G.: Estimating regression functions and their derivatives by the kernel method, *Scand. J. Stat.*, 11, 171–185, 1984.
- Haam, E. and Huybers, P.: A test for the presence of covariance between time-uncertain series of data with application to the Dongge cave speleothem and atmospheric radiocarbon records, *Paleoceanography*, 25, PA2209, doi:10.1029/2008PA001713, 2010.
- Härdle, W.: *Applied nonparametric regression*, Cambridge University Press, Cambridge, 1990.
- Hann, J.: *Lehrbuch der Meteorologie*, Tauchnitz, Leipzig, 1901.
- Haslett, J. and Parnell, A.: A simple monotone process with application to radiocarbon-dated depth chronologies, *Appl. Stat.-J. Roy. St. C*, 57, 399–418, 2008.
- Heegaard, E., Birks, H. J. B., and Telford, R. J.: Relationships between calibrated ages and depth in stratigraphical sequences: an estimation procedure by mixed-effect regression, *Holocene*, 15, 612–618, 2005.
- Hoffmann, D. L., Prytulak, J., Richards, D. A., Elliott, T., Coath, C. D., Smart, P. L., and Scholz, D.: Procedures for accurate U and Th isotope measurements by high precision MC-ICPMS, *Int. J. Mass Spectrom.*, 264, 97–109, 2007.
- Huybers, P.: Combined obliquity and precession pacing of late Pleistocene deglaciations, *Nature*, 480, 229–232, 2011.
- Ivanovich, M. and Harmon, R. S. (Eds.): *Uranium-series Disequilibrium: Applications to Earth, Marine, and Environmental Sciences*, 2nd Edn., Clarendon Press, Oxford, 1992.
- Jansen, E., Overpeck, J., Briffa, K. R., Duplessy, J.-C., Joos, F., Masson-Delmotte, V., Olago, D., Otto-Bliesner, B., Peltier, R. W., Rahmstorf, S., Ramesh, R., Raynaud, D., Rind, D., Solomina, O., Villalba, R., and Zhang, D.: Palaeoclimate, in: *Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by: Solomon, S., Qin, D., Manning, M., Marquis, M., Averyt, K., Tignor, M. M. B., Miller Jr., H. L., and Chen, Z., Cambridge University Press, Cambridge, 433–497, 2007.
- Kant, I.: *Critik der reinen Vernunft*, Hartknoch, Riga, 1781.
- Klaunberg, K., Blackwell, P. G., Buck, C. E., Mulvaney, R., Röthlisberger, R., and Wolff, E. W.: Bayesian glaciological modelling to quantify uncertainties in ice core chronologies, *Quaternary Sci. Rev.*, 30, 2961–2975, 2011.
- Köppen, W.: *Die Klimate der Erde: Grundriss der Klimakunde*, de Gruyter, Berlin, 1923.
- Künsch, H.: The jackknife and the bootstrap for general stationary observations, *Ann. Stat.*, 17, 1217–1241, 1989.
- Lorenz, E.: Deterministic nonperiodic flow, *J. Atmos. Sci.*, 20, 130–141, 1963.
- Moore, M. I. and Thomson, P. J.: Impact of jittered sampling on conventional spectral estimates, *J. Geophys. Res.*, 96, 18519–18526, 1991.
- Mudelsee, M.: TAUEST: a computer program for estimating persistence in unevenly spaced weather/climate time series, *Comput. Geosci.*, 28, 69–72, 2002.
- Mudelsee, M.: *Climate Time Series Analysis: Classical Statistical and Bootstrap Methods*, Springer, Dordrecht, 2010.
- Mudelsee, M., Scholz, D., Röthlisberger, R., Fleitmann, D., Mangini, A., and Wolff, E. W.: Climate spectrum estimation in the presence of timescale errors, *Nonlin. Processes Geophys.*, 16, 43–56, doi:10.5194/npg-16-43-2009, 2009.
- Mühlinghaus, C., Scholz, D., and Mangini, A.: Modelling fractionation of stable isotopes in stalagmites, *Geochim. Cosmochim. Ac.*, 71, 2780–2790, 2009.
- Niggemann, S., Mangini, A., Mudelsee, M., Richter, D. K., and Würth, G.: Sub-Milankovitch climatic cycles in Holocene stalagmites from Sauerland, Germany, *Earth Planet. Sc. Lett.*, 216,

- 539–547, 2003.
- Popper, K.: *Logik der Forschung: Zur Erkenntnistheorie der modernen Naturwissenschaft*, Julius Springer, Wien, 1935.
- Ravetz, J.: *The No-Nonsense Guide to Science*, New Internationalist, Oxford, n.d.
- Rhines, A. and Huybers, P.: Estimation of spectral power laws in time uncertain series of data with application to the Greenland Ice Sheet Project 2 $\delta^{18}\text{O}$ record, *J. Geophys. Res.*, 116, D01103, doi:10.1029/2010JD014764, 2011.
- Riechelmann, D. F. C., Schröder-Ritzrau, A., Scholz, D., Fohlmeister, J., Spötl, C., Richter, D., and Mangini, A.: Monitoring Bunker cave (NW Germany): a prerequisite to interpret geochemical proxy data of speleothems from this site, *J. Hydrol.*, 409, 682–695, 2011.
- Roberts, S. and Martin, M.: Bootstrap-after-bootstrap model averaging for reducing model uncertainty in model selection for air pollution mortality studies, *Environ. Health Persp.*, 118, 131–136, 2010.
- Scholz, D. and Hoffmann, D.: StalAge – an algorithm designed for construction of speleothem age models, *Quat. Geochronol.*, 6, 369–382, 2011.
- Scholz, D., Mühlinghaus, C., and Mangini, A.: Modelling $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ in the solution layer on stalagmite surfaces, *Geochim. Cosmochim. Ac.*, 73, 2592–2602, 2009.
- Scholz, D., Hoffmann, D. L., Hellstrom, J., and Ramsey, C. B.: A comparison of different methods for speleothem age modelling, *Quat. Geochronol.*, in press, 2012.
- Shakun, J. D., Clark, P. U., He, F., Marcott, S. A., Mix, A. C., Liu, Z., Otto-Bliesner, B., Schmittner, A., and Bard, E.: Global warming preceded by increasing carbon dioxide concentrations during the last deglaciation, *Nature*, 484, 49–54, 2012.
- Silverman, B. W.: *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London, 1986.
- Simonoff, J. S.: *Smoothing Methods in Statistics*, Springer, New York, 1996.
- Solomon, S., Qin, D., Manning, M., Marquis, M., Averyt, K., Tignor, M. M. B., Miller Jr., H. L., and Chen, Z. (Eds.): *Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, 2007.
- Spötl, C., Mangini, A., and Richards, D. A.: Chronology and paleoenvironment of Marine Isotope Stage 3 from two high-elevation speleothems, *Austrian Alps, Quaternary Sci. Rev.*, 25, 1127–1136, 2006.
- Spötl, C., Scholz, D., and Mangini, A.: A terrestrial U/Th-dated stable isotope record of the penultimate interglacial, *Earth Planet. Sc. Lett.*, 276, 283–292, 2008.
- Stanford, J. D., Hemingway, R., Rohling, E. J., Challenor, P. G., Medina-Elizalde, M., and Lester, A. J.: Sea-level probability for the last deglaciation: a statistical analysis of far-field records, *Global Planet. Change*, 79, 193–203, 2011.
- Taylor, R. E.: *Radiocarbon Dating: an Archaeological Perspective*, Academic Press, Orlando, FL, 1987.
- Thomson, P. J. and Robinson, P. M.: Estimation of second-order properties from jittered time series, *Ann. I. Stat. Math.*, 48, 29–48, 1996.