PERGAMON

Computers & Geosciences 26 (2000) 293–307

COMPUTERS & GEOSCIENCES

Ramp function regression: a tool for quantifying climate transitions ☆

Manfred Mudelsee*

Institute of Mathematics and Statistics, University of Kent, Canterbury CT2 7NF, UK

Received 16 March 1999; received in revised form 15 September 1999; accepted 15 September 1999

Abstract

A method is proposed for fitting a 'ramp' to measured data. This is a continuous function, segmented in three parts: $x_{\mathrm{fit}}(t) = x1$ for $t \le t1$, $x2$ for $t \ge t2$, and linearly connected between $t1$ and $t2$. Its purpose is to measure transitions in the mean of time series as they occur, for example, in paleoclimatic records. The unknowns $x1$ and $x2$ are estimated by weighted least-squares regression, $t1$ and $t2$ by a brute-force search. Computing costs are reduced by several methods. The presented Fortran 77 program, RAMPFIT, includes analysis of weighted ordinary residuals for checking the validity of the ramp form and other assumptions. It fits an AR(1) model to the residuals to measure serial dependency; uneven time spacing is thereby allowed. Three bootstrap resampling schemes (nonparametric stationary, parametric, and wild) provide uncertainties for the estimated parameters. RAMPFIT works interactively (calculation/visualization). Example time series (one artificial, three measured) demonstrate that this approach is useful for practical applications in geosciences (n less than a few hundred, noise, unevenly spaced times), and that the ramp function may serve well to model climate transitions.

Keywords: Time series analysis; Three-phase regression; O-18/O-16; Sr-87/Sr-86; Cenozoic

1. Introduction

Accurately quantifying past transitions is one prerequisite for a better causal understanding of climatic change. Given a time series $x(i)$ measured at times $t(i)$, $i = 1, \ldots, n$, which documents a transition in the mean, it is important to know: when did the transition start? When did it end? What were the levels of the mean before and after the transition? These questions lead ultimately to the three-phase regression model which is shown in Fig. 1. We denote it a 'ramp function'.

As physical motivation for the ramp form, consider a (climate) system at equilibrium which is disturbed by some external action over a period of time, then attaining a new equilibrium state.

The mathematical problem is to fit the ramp function,

$$x_{\mathrm{fit}}(t) = \begin{cases} x1, & \text{for } t \le t1, \\ x1 + (t - t1)(x2 - x1)/(t2 - t1), & \text{for } t1 \le t \le t2, \\ x2, & \text{for } t \ge t2, \end{cases}$$

☆ Code available at http://www.iawg.org/CGEditor/index.htm

* Present address: Institute of Meteorology, University of Leipzig, Stephanstr. 3, D-04103 Leipzig, Germany. Fax: +49-341-97-32899.

E-mail address: mudelsee@tz.uni-leipzig-de (M. Mudelsee).

0098-3004/00/$ - see front matter
PII: S0098-3004(99)00141-7

to $x(i)$; that is, to estimate the parameters $t1$, $x1$, $t2$, and $x2$ for the regression model

$$x(i) = x_{\mathrm{fit}}(i) + \epsilon(i) \tag{1}$$

where $x_{\mathrm{fit}}(i) := x_{\mathrm{fit}}(t(i))$ is an abbreviation for the discrete form. We assume $\epsilon(i)$ to be distributed as $N(0, \sigma(i)^2)$, since for climate time series we anticipate heteroscedasticity. We further allow positive AR(1) dependency between the $\epsilon(i)$ since persistence is also typical for such time series. The two points of nondifferentiability (Fig. 1) cause the main difficulty, in that they prevent us from formulating the normal equations for a least-squares (LS) estimation of $t1$ and $t2$.

The proposed estimation method and the respective Fortran 77 program RAMPFIT are a pragmatic approach to that regression problem. By a brute-force search over a pre-defined grid, we find $\hat{t}1$ and $\hat{t}2$, whereas we calculate $\hat{x}1$ and $\hat{x}2$ directly by using a weighted LS criterion.

The computational burden for that approach is reduced by several approaches. First, interactive graphics (for smoothing and time interval evaluation) allow us to keep the boundaries of the search grid narrow. Second, an intelligent search path on the grid permits updating of values of constants rather than new evaluation. Third, the search grid is as coarse as the time values $t(i)$ are.

Section 2 explains the estimation method and also mentions possible extensions and alternatives. Section 3 demonstrates the analysis of the regression residuals which helps to assess whether the ramp function is indeed suited to the investigated data. Section 4 supplies three bootstrap resampling schemes for determining the uncertainties of the estimated parameters. Whereas an artificial time series serves to exemplify those theoretical Sections, measured data (Section 5) make clear that ramp function regression is of practical meaning for many types of time series encountered in the geosciences. The Appendix gives a short manual of RAMPFIT.

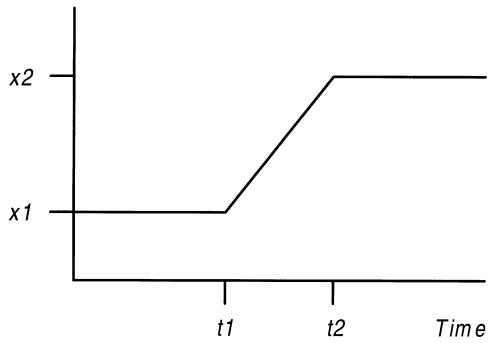

Fig. 1. Ramp function regression model.

2. Ramp function estimation

2.1. Weighted LS

Since the standard deviation $\sigma(i)$ can vary with time, we use a LS criterion which puts heavier weights on values with smaller $\sigma(i)$. The best ramp function fit to given data $x(i)$ thus minimizes

$$SSQW(t1, x1, t2, x2) = \sum_{i=1}^{n} [x(i) - x_{\mathrm{fit}}(i)]^2 / \sigma(i)^2.$$

We may either know $\sigma(i)$ a priori, or estimate it per eye with the help of smoothing (Subsections 2.3 and 3.3). We seek the solution $(\hat{t}1, \hat{x}1, \hat{t}2, \hat{x}2)$ for which $SSQW(\hat{t}1, \hat{x}1, \hat{t}2, \hat{x}2)$ is a minimum.

First, consider the case where $t1_{\mathrm{search}} = t(i1)$ and $t2_{\mathrm{search}} = t(i2)$ are fixed. Then we find $\hat{x}1$ and $\hat{x}2$ by putting

$$\partial / \partial x1 \, [SSQW(t1_{\mathrm{search}}, x1, t2_{\mathrm{search}}, x2)] \stackrel{!}{=} 0$$

and

$$\partial / \partial x2 \, [SSQW(t1_{\mathrm{search}}, x1, t2_{\mathrm{search}}, x2)] \stackrel{!}{=} 0.$$

This yields

$$\hat{x}2 = \left(\frac{K_3 K_4}{K_1} + K_6 \right) / \left(\frac{K_2 K_4}{K_1} + K_5 \right) \tag{2}$$

and

$$\hat{x}1 = (K_3 - \hat{x}2 K_2) / K_1. \tag{3}$$

The constants are:

$$K_1 = k_2 + (t1_{\mathrm{search}} k_4 - k_5) / (t2_{\mathrm{search}} - t1_{\mathrm{search}}),$$

$$K_2 = k_3 - (t1_{\mathrm{search}} k_4 - k_5) / (t2_{\mathrm{search}} - t1_{\mathrm{search}}),$$

$$K_3 = k_8,$$

$$K_4 = k_1 + [t2_{\mathrm{search}}(t1_{\mathrm{search}} + t2_{\mathrm{search}}) k_4 + 2k_6 - (t1_{\mathrm{search}} + 3t2_{\mathrm{search}}) k_5] / (t2_{\mathrm{search}} - t1_{\mathrm{search}})^2,$$

$$K_5 = k_3 + [t1_{\mathrm{search}}(t1_{\mathrm{search}} + t2_{\mathrm{search}}) k_4 + 2k_6 - (3t1_{\mathrm{search}} + t2_{\mathrm{search}}) k_5] / (t2_{\mathrm{search}} - t1_{\mathrm{search}})^2,$$

$$K_6 = k_9 - k_7 - 2[t1_{\mathrm{search}} k_{10} - k_{11}] / (t2_{\mathrm{search}} - t1_{\mathrm{search}}),$$

and

$$k_1 = \sum_{i=1}^{i1} 1/\sigma(i)^2, \quad k_2 = \sum_{i=1}^{i2-1} 1/\sigma(i)^2,$$

$$k_3 = \sum_{i=i2}^{n} 1/\sigma(i)^2, \quad k_4 = \sum_{i=i1+1}^{i2-1} 1/\sigma(i)^2,$$

$$k_5 = \sum_{i=i1+1}^{i2-1} t(i)/\sigma(i)^2, \quad k_6 = \sum_{i=i1+1}^{i2-1} t(i)^2/\sigma(i)^2,$$

$$k_7 = \sum_{i=1}^{i2-1} x(i)/\sigma(i)^2, \quad k_8 = \sum_{i=1}^{n} x(i)/\sigma(i)^2,$$

$$k_9 = \sum_{i=i2}^{n} x(i)/\sigma(i)^2, \quad k_{10} = \sum_{i=i1+1}^{i2-1} x(i)/\sigma(i)^2,$$

$$k_{11} = \sum_{i=i1+1}^{i2-1} x(i)t(i)/\sigma(i)^2.$$

2.2. Brute-force search

Since the ramp function (Fig. 1) is not differentiable with respect to time at $t1$ and $t2$, we cannot derive $\hat{t}1$ and $\hat{t}2$ in the same manner as $\hat{x}1$ and $\hat{x}2$. Instead, we search over all points of a $t1$–$t2$ grid (Fig. 2). The knots of this grid are determined by the time values $t(i)$ and the restriction $t1_{\text{search}} < t2_{\text{search}}$. For every search point, we calculate $\hat{x}1$, $\hat{x}2$, and $SSQW$.

The search starts in the upper left corner of the grid (Fig. 2) where constants $k_1 - k_{11}$ have to be calculated explicitly. An intelligent path then permits updating of the values of the constants rather than new evaluation (Table 1 in Mudelsee, 1999, Technical Report UKC/IMS/99/10). This reduces the computing costs by a factor of about 2.

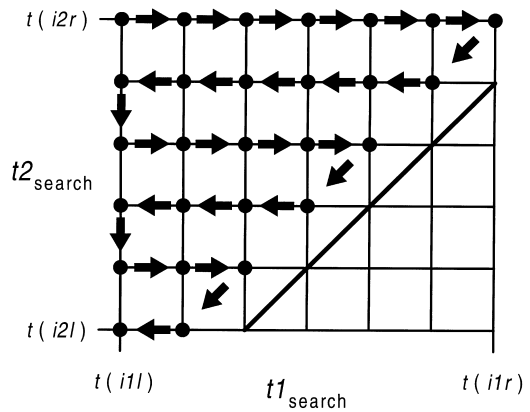

Fig. 2. Search grid for $\hat{t}1$ and $\hat{t}2$: search points (●), search steps (arrows), and 1:1 line (heavy).

We prescribe the boundaries of the $t1$–$t2$ search grid. They have to be broad enough to reduce the possibility of a solution on the grid boundary. In such a case it would be likely that a better fit solution exists outside the grid. On the other hand, the grid boundaries have to be narrow enough to keep computing costs small. The computer program RAMPFIT helps to suit the search boundaries by means of interactive graphics (smoothing).

The coarseness of the search grid, that is, the spacing of the $t(i)$, limits the precision of $\hat{t}1$ and $\hat{t}2$. However, when the uncertainties of $\hat{t}1$ and $\hat{t}2$ (which are determined by bootstrap resampling, see Section 4) are larger, that limitation becomes unimportant. The brute-force search, despite its computational burden and its tedious appearance, has the advantage of providing a *global* optimal solution within the search grid.

2.3. Standard deviation

The time-dependent standard deviation provides the weights for the LS regression. If no prior knowledge (e.g., measurement error) is available, we have to estimate $\sigma(i)$. We assume also $\sigma(i)$ to have a ramp form. This is motivated by the definition of climate not only in terms of mean value but also in terms of variability. A climate change therefore may consist of a ramp function change of standard deviation.

RAMPFIT allows us to fit $\sigma(i)$ per eye. Within a running window, the standard deviation of the data is calculated. Interactive graphics and an adjustable degree of smoothing facilitate the estimation (Subsection 3.3 and Appendix A.1).

In principle, it would be possible to carry out a LS regression of the standard deviation or to use more complicated functions for $\sigma(i)$. Since the major interest in the present context is an appropriate distribution of the weights, we avoid such extensions.

2.4. Extensions and alternative approaches

My approach of brute-force search is similar to Williams' (1970) adaptation of Hudson's search to a three-phase linear regression. He used a search path which made re-calculation of constants necessary, something our path (Fig. 2) completely avoids. He further conducted a 'fine estimation' to overcome the coarseness of the grid.

One possible alternative to brute-force search is numerical minimization of $SSQW(t1, x1, t2, x2)$ which also would provide a 'fine estimation'. A numerical technique without derivative calculation, such as Brent's search in two dimensions, would be necessary. The solution, however, would only be a local minimum. Typical questions such as choice of starting point and stopping rule would arise. Furthermore, updating of constants which reduces computing costs would not be possible.

Since we allow AR(1) dependency between the errors $\epsilon(i)$, generalized LS estimation could be a useful extension, especially for a strong dependency. In addition to $\sigma(i)$, the serial correlation would have to be determined. Note that because of the bootstrap resampling, correlated errors do not distort evaluated uncertainty ranges of estimated parameters.

We consider robust regression to be a less helpful extension. The simple calculation of $\hat{x}1$ and $\hat{x}2$ by LS minimization (Eqs. 2 and 3) would not be possible. Most climate time series seem not to depart so strongly from the Gaussian shape that would justify such a complication.

The statistical literature to 'change-point detection' indicates alternative approaches to ramp function regression. For example, we could use tests for an abrupt change in the mean (see, e.g., Basseville and Nikiforov, 1993) to estimate $t1$ 'from the left' (Fig. 1) and $t2$ 'from the right'. However, when a significant number of time series points lie *between* $t1$ and $t2$, that approach looses considerable power.

It might be interesting to adapt the approach of Tishler and Zang (1981) and approximate $x_{\mathrm{fit}}(t)$ by a continuously differentiable function in some arbitrarily small neighborhood of $t1$ and $t2$. That smooth regression model could then be estimated by numerical techniques with derivative calculation. Similar questions to the above would arise, and updating would not be possible.

We finally mention change-point detection by nonparametric regression (Müller, 1992).

2.5. Computing costs

Costs not only depend on the number of data points, n, and the number of grid search points (Fig. 2), say n_{search}, but also on the proportion of data points between $t1$ and $t2$ because in that interval evaluation of constants is more expensive. The costs to calculate k_1–k_{11} for one search point are roughly proportional to n whereas updating costs for all search points are roughly proportional to n_{search}. Thus, for smaller values of n/n_{search}, numerical minimization which requires calculation only at a small number of search points, is an interesting alternative.

Costs further are determined by the type of machine and compiler. The following numbers are based on RAMPFIT test calculations with a PC with 80486-50 DX processor and the Microsoft Fortran PowerStation 1.0 compiler, and a Sunsparc workstation with 143 MHz clock and the EPCF77 2.7 compiler. They should be seen as a rough guideline.

For $n = 500$, the CPU time per search point is about 150 μs on the work-station, for $n = 300$ about 50 μs, and for $n = 100$ about 18 μs. The PC was about 100 times slower.

3. Residual diagnostics

We have made the following assumptions for the ramp function regression:

1. The errors $\epsilon(i)$ are Gaussian distributed.
2. The errors have time-dependent standard deviation $\sigma(i)$, that is, the weighting is appropriate.

Table 1
Analyzed time series[a]

Name	Description	Time interval	n	V
Artificial	$t1 = 200$, $x1 = 2.0$ $t2 = 300$, $x2 = 4.0$ $\sigma(t) = 1.0$ ($t \leq 250$) $\sigma(t) = 0.5$ ($t > 250$) $a = 0.40$	[0; 499]	500	0.21
ODP 925	δ^{18}O [per mil vs PDB], benthic foraminifera	[500 ka; 1400 ka]	429	0.79
87-Sr/86-Sr	87-Sr/86-Sr, various cores, planktonic foraminifera	[2.0 Ma; 8.8 Ma]	65	1.94
ODP 929 y	δ^{18}O [per mil vs PDB], benthic foraminifera	[23.4 Ma; 23.8 Ma]	70	0.17
ODP 929 o	δ^{18}O [per mil vs PDB], benthic foraminifera	[23.7 Ma; 24.6 Ma]	150	0.16

[a] Time units refer to years before present. V is the coefficient of variation of the time spacing, that is, the standard deviation divided by the average. The ODP 929 time series has been split into a young ('y') and an old ('o') part.

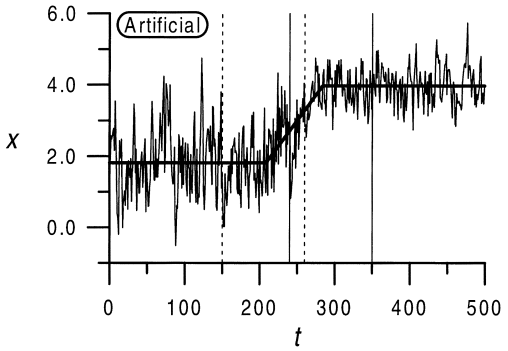

Fig. 3. Ramp function fit (heavy line) to artificial time series (Tables 1 and 2). Grid boundaries for $t1$-search (dashed vertical lines) and $t2$-search (solid).

3. The relationship between x and t is indeed a ramp function.
4. The errors show AR(1) dependency (that is not too strong).

Analysis of regression residuals is an important tool for assessing how well these assumptions are fulfilled. For an excellent treatment of that subject refer to Montgomery and Peck (1992, Ch. 3). In the present study, an artificial time series with distinct transition in the mean (Fig. 3 and Table 1) illustrates residual analysis.

3.1. Weighted ordinary residuals

The weighted ordinary residuals are calculated as

$$e(i) = [x(i) - \hat{x}_{\mathrm{fit}}(i)]/\hat{\sigma}(i)$$

where $\hat{x}_{\mathrm{fit}}(i)$ is the ramp function fitted by brute-force search and LS minimization, and $\hat{\sigma}(i)$ is the standard deviation fitted per eye. Given the uncertainty of the latter fit, we regard it unnecessary to furthermore studentize $e(i)$.

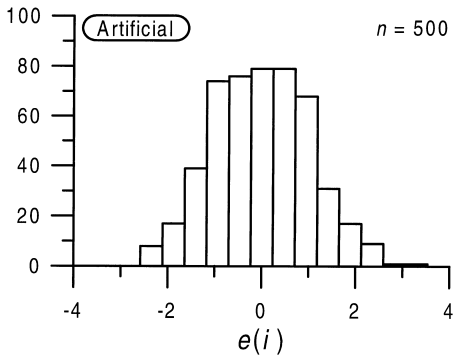

Fig. 4. Histogram of weighted ordinary regression residuals, artificial time series.

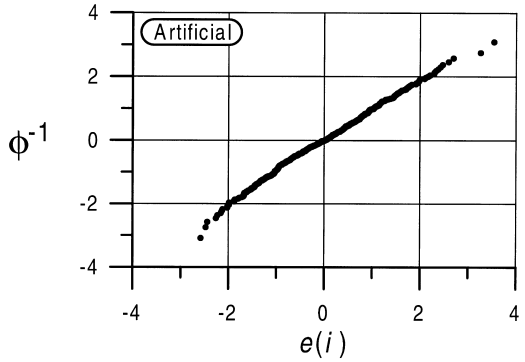

Fig. 5. Normal probability plot of weighted ordinary regression residuals, artificial time series.

3.2. Gaussian shape

A histogram of the $e(i)$ (Fig. 4) is the first visual tool for examining a Gaussian shape. We follow Scott (1979) and use the number of classes

$$NINT[(e_{\max} - e_{\min})n^{1/3}/3.49e_{\mathrm{std}})]$$

where NINT( ) is the nearest integer function; and $e_{\max}$, $e_{\min}$, and e_{std} are, respectively, maximum, minimum, and standard deviation of $e(i)$.

Fig. 5 shows the normal probability plot, $\Phi^{-1}[(i - 1/2)/n]$ vs $e_{[i]}$ (where Φ^{-1} is the inverse cumulative normal function and $e_{[i]}$ are the ranked residuals); the Gaussian shape appears as a straight line, especially in the crucial central part of the plot. RAMPFIT uses the approximation of Hamaker (1978) to Φ^{-1}.

A stronger violation of the Gaussian assumption would necessitate more robust measures than LS regression.

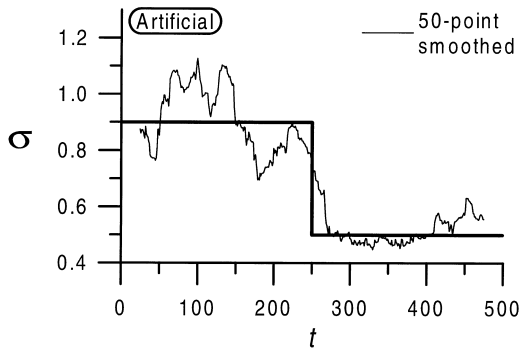

Fig. 6. Per-eye fit (heavy line) of $\sigma(i)$, and standard deviation (smoothed) of detrended values (light line), artificial time series. Smoothing window contains 50 points.

Table 2
Results: ramp function fits[a]

Name	$\hat{t}1$	$\hat{x}1$	$\hat{t}2$	$\hat{x}2$
Artificial	208.1 ± 10.1	1.81 ± 0.08	283.8 ± 6.9	3.98 ± 0.04
ODP 925	918.7 ± 13.2	3.40 ± 0.03	931.8 ± 14.1	3.03 ± 0.03
87-Sr/86-Sr	4.39 ± 0.12	7090.30 ± 0.06	5.66 ± 0.18	7089.27 ± 0.06
ODP 929 y	23.615 ± 0.024	1.28 ± 0.03	23.726 ± 0.027	1.82 ± 0.07
ODP 929 o	23.756 ± 0.027	1.88 ± 0.08	24.010 ± 0.018	0.89 ± 0.02

[a] Given errors are *mad'*, using the nonparametric stationary bootstrap scheme (Subsection 4.1). In case of the 87-Sr/86-Sr time series, the x-values have been multiplied by 10^4.

3.3. Appropriate weighting

In Fig. 6, $\hat{\sigma}(i)$ (fitted per eye) is plotted together with the standard deviation of the detrended values, $[x(i) - \hat{x}_{\mathrm{fit}}(i)]$. Their standard deviation is calculated within a window which contains a fixed number of points and which is shifted on the time axis (k-nearest-neighbor smoothing). The assigned time value is the average of time values contained in the window (Tukey, 1977). That is more precise since, for climate time series, we expect uneven time spacing. RAMPFIT facilitates comparison with $\hat{\sigma}(i)$ by permitting the choice of the degree of smoothing. RAMPFIT also gives the values of reduced chi-square,

$$SSQWN = SSQW(\hat{t}1, \hat{x}1, \hat{t}2, \hat{x}2)/(n-4),$$

which — in the case of independent $\epsilon(i)$ — should be around unity for an adequate fit to $x(i)$ and a proper guess of $\sigma(i)$ (see, for example, Bevington and Robinson, 1992, Ch. 11). In the situation of strong dependency, $SSQWN$ is no longer useful, and, also, the smoothed standard deviation is then lower than the true. (The example of the artificial time series (Table 2 and Fig. 6) indicates that an autocorrelation coefficient of 0.4 is not a strong dependency in that respect.) Note that this underestimation would be a roughly *constant percentage* over time, provided the coefficient of variation of the time spacing is small. Thus, the underestimation would not corrupt the appropriate distribution of the weights.

3.4. Suitable ramp model

The plots of $e(i)$ vs $t(i)$ (Fig. 7) and $e(i)$ vs $x_{\mathrm{fit}}(i)$ (Fig. 8), together with the fit (Fig. 3), help to identify outliers or systematic deviations from the assumed ramp form. In the case of the artificial time series with a pre-defined ramp transition, as expected, we observe no such behavior: the points are homogeneously distributed.

3.5. Weak AR(1) dependency

We inspect AR(1) dependency between errors by means of a lag-1 scatterplot of $e(i)$ (Fig. 9). The orientation of the cloud of points along the 1:1 line indicates a positive correlation. We quantify that dependency by fitting the stationary time series model

$$e(i) = a^{[t(i)-t(i-1)]}e(i-1) + [1 - a^{2[t(i)-t(i-1)]}]^{0.5}u(i) \qquad (4)$$

to the residuals where the $u(i)$ are assumed to be independent and identically distributed as $N(0,1)$. This model is a generalization of an AR(1) model to discrete data with an uneven time spacing. Estimation ($\hat{a}$) is carried out using a LS criterion and numerical techniques. RAMPFIT also gives the estimated value of $\tau = -1/\ln(a)$ which is the decay period of the autocorrelation function of the AR(1) model. My approach is similar to, amongst others, Robinson's (1977).

In the example of the artificial time series, we estimate $\hat{a} = 0.41$, that is, approximately the true serial correlation. It would be difficult to evaluate the bias and the variance of estimator $\hat{a}$ analytically. Despite that, $\hat{a}$ serves well for judging whether the dependency between the errors is weak enough to circumvent a

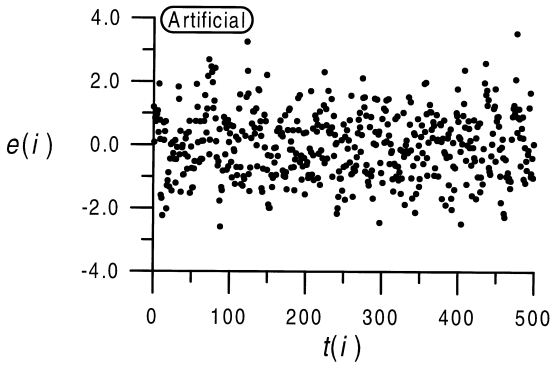

Fig. 7. Weighted ordinary regression residuals vs time, artificial time series.

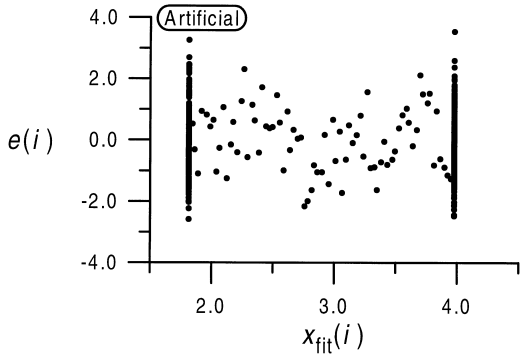

Fig. 8. Weighted ordinary regression residuals vs x_{fit}, artificial time series. (Because of ramp form, clusters of points at $\hat{x}1$ and $\hat{x}2$ emerge.)

generalized LS estimation, and whether $SSQWN$ is a useful quantity (see Subsection 3.3).

Strictly speaking, we do not test whether the dependency is indeed first-order autoregressive which, for unevenly spaced time series, would be a formidable task. This omission should be justified in case of climate time series ('natural persistence').

4. Bootstrap resampling

After we have fitted a ramp function and — assisted by the residual plots — accepted that model for a particular climate time series, geological interpretation of the fit result requires knowledge about how close the estimated parameters are to the true parameters. Those biases and uncertainties of $\hat{t}1$, $\hat{x}1$, $\hat{t}2$, and $\hat{x}2$, we estimate by bootstrap resampling: from the original time series $x(i)$, we draw a first sample $x(i)^*(j=1)$ that is intended to have the same statistical properties; three methods of doing so are described in this Section. A ramp function is fitted to $x(i)^*(j=1)$, resulting in

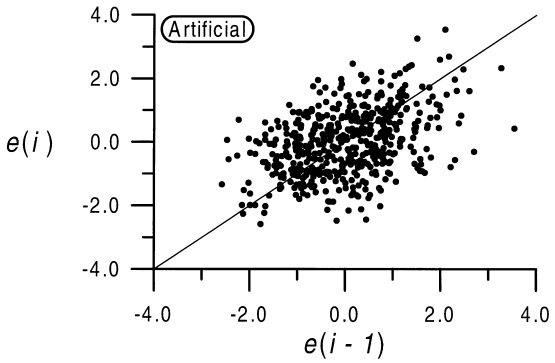

Fig. 9. Lag-1 scatterplot of weighted ordinary regression residuals (artificial time series), 1:1 line.

$\hat{t}1^*(j=1)$, $\hat{x}1^*(j=1)$, $\hat{t}2^*(j=1)$, and $\hat{x}2^*(j=1)$. Resampling and fitting is repeated, B times in total. The distributions of $\hat{t}1^*(j)$, $\hat{x}1^*(j)$, $\hat{t}2^*(j)$, and $\hat{x}2^*(j)$ then inform about biases and uncertainties.

Following Efron and Tibshirani (1993) in their excellent monograph on the bootstrap, we use $B = 200$. With regard to CPU time and coarseness of estimation ($t1$–$t2$ grid), a higher number of replications would make no sense. The search grid (Fig. 2) remains unchanged.

The more traditional approach for estimating uncertainties, via the second derivative of $SSQW$, is not applicable since the ramp function is not continuously differentiable. The numerical minimization technique (Subsection 2.4), however, can use that approach.

4.1. Nonparametric stationary bootstrap

The nonparametric stationary bootstrap (Politis and Romano, 1994), adapted to heteroscedasticity, uses the set $\{e(k); k = 1, \ldots, n\}$ of regression residuals for resampling:

$$x(i)^* = \hat{x}_{\mathrm{fit}}(i) + \hat{\sigma}(i)e(i)^*, \quad i = 1, \ldots, n,$$

where $e(1)^*$ is picked at random from $\{e(k)\}$. The other $e(i)^*$ are derived as follows. Let $e(i)^* = e(l)$, say. Then:

$$e(i+1)^* = \begin{cases} \text{with probability } p\text{:} \\ e(l+1) \text{ or, if } l = n, e(1), \\ \text{with probability } 1-p\text{:} \\ \text{picked at random from } \{e(k)\}. \end{cases}$$

The $e(i)^*$ consist of blocks of variable length which preserve a sub-chain of the $e(i)$. That recognizes eventual autocorrelation of the $\epsilon(i)$. The probability p has to be prescribed. RAMPFIT suggests a value for

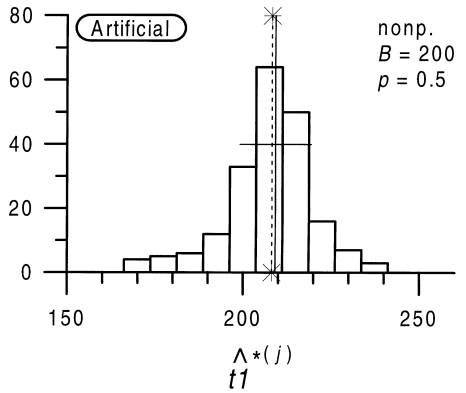

Fig. 10. Histogram of $\hat{t}1^*(j)$, nonparametric stationary bootstrap, artificial time series. Vertical, dashed line (with stars) marks $\hat{t}1$. Crossed, solid lines mark $med[\hat{t}1^*(j)] \pm mad'[\hat{t}1^*(j)]$.

which the average block length becomes equal to the estimated autocorrelation decay period $\hat{\tau}$.

In the histogram of $\hat{t}1^*(j = 1, \ldots, B = 200)$ (Fig. 10), the fit value, $\hat{t}1$, is marked. Also the median, $med[\hat{t}1^*(j)]$, is drawn with error bars $\pm mad'[\hat{t}1^*(j)]$ where $mad' := 1.4826\, mad$. The *mad* is the median of absolute distances to *med*; a normal distribution has standard deviation $\simeq mad'$. For comparison with $\hat{t}1$, we prefer those robust measures which are also meaningful for heavier skewed distributions. Fig. 10 reveals that $\hat{t}1$ has negligible bias. Furthermore, $mad'[\hat{t}1^*(j)] = 10.1$ is clearly larger than the average spacing (1.0) of the time series within the $t1$ search region, justifying the coarse $t1$–$t2$ grid. In addition to *med* and *mad'*, RAMPFIT also calculates the average, standard deviation, minimum, and maximum; the same also applies for $\hat{x}1^*(j)$, $\hat{t}2^*(j)$, and $\hat{x}2^*(j)$. The two other bootstrap schemes are dealt with similarly.

4.2. Parametric bootstrap

The parametric bootstrap resamples $x(i)^*(j)$ after Eq. (1):

$$x(i)^* = \hat{x}_{\mathrm{fit}}(i) + \hat{\sigma}(i)v(i), \quad i = 1, \ldots, n,$$

where $\hat{x}_{\mathrm{fit}}(i)$ is the estimated fit function, $\hat{\sigma}(i)$ is the estimated standard deviation, and the $v(i)$ are distributed as $N(0, 1)$ with the AR(1) correlation in Eq. (4).

For generating the random numbers, RAMPFIT uses the algorithm of Park and Miller (1988) in combination with routine Gasdev from Press et al. (1992).

4.3. Wild bootstrap

The wild bootstrap of Härdle and Marron (1991), adapted to heteroscedasticity, resamples as

$$x(i)^* = \hat{x}_{\mathrm{fit}}(i) + \hat{\sigma}(i)e(i)w(i), \quad i = 1, \ldots, n,$$

where $w(i)$ is the two-point distribution,

$$w(i) = \begin{cases} \text{with probability } r: & w_1, \\ \text{with probability } 1 - r: & w_2, \end{cases}$$

with $r = (\sqrt{5} + 1)/(2\sqrt{5})$, $w_1 = (1 - \sqrt{5})/2$, and $w_2 = (1 + \sqrt{5})/2$. The wild bootstrap attempts "to reconstruct the distribution of each residual through the use of one single observation" (Härdle and Marron, 1991). This resampling scheme, in contrast to the two others, preserves no dependency between the residuals.

The question of which of the resampling schemes may be most adequate for ramp function regression has to be pursued further, for example, by means of Monte Carlo simulations. At present, I conjecture that eventual deviations from Gaussianity or AR(1) dependency are best preserved by the nonparametric stationary bootstrap.

An eventual important addition to RAMPFIT would be the incorporation of a scheme which also resamples the time $t(i)$. An uncertain timescale is most common for paleoclimatic time series. It may, however, be necessary to develop a statistical model for the $t(i)$.

Table 3
Results: standard deviation and AR(1) fits[a]

Name	$\hat{\sigma}$					AR(1) fit	
	$\hat{t}1_s$	$\hat{x}1_s$	$\hat{t}2_s$	$\hat{x}2_s$	$SSQWN$	$\hat{a}$	$\hat{\tau}$
Artificial	250.0	0.90	250.0	0.50	1.08	0.41	1.1
ODP 925	650.0	0.45	1000.0	0.35	1.02	0.81	4.6
87-Sr/86-Sr	(measurement error)				(1.72)	0.00	
ODP 929 y	23.5	0.16	23.6	0.20	1.18	0.00	
ODP 929 o	0.19 (constant)				1.06	0.00	

[a] $\hat{t}1_s$, $\hat{x}1_s$, $\hat{t}2_s$, and $\hat{x}2_s$ are the ramp function parameters for $\hat{\sigma}$ (cf Fig. 1).

5. Examples

After having tested ramp function regression, residual analysis, and bootstrap resampling for an artificial time series, we investigate three measured marine, paleoclimatic time series. Table 1 describes the data sets and Tables 2 and 3 present the results. As error bars, we use $\pm mad'$ from the nonparametric stationary bootstrap. The results from the other schemes, as well as the bootstrap histograms and scatterplots, (as in Fig. 11) can be found in Mudelsee (1999, Technical Report UKC/IMS/99/10).

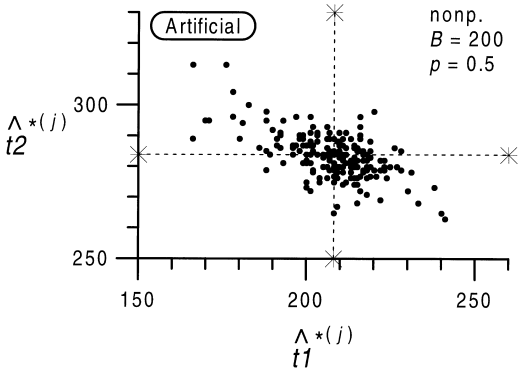

Fig. 11. Scatterplot (correlation) between $\hat{t}1^*(j)$ and $\hat{t}2^*(j)$, nonparametric stationary bootstrap, artificial time series. Crossed, dashed lines (with stars) mark $\hat{t}1$ and $\hat{t}2$, respectively.

5.1. ODP 925

The ODP 925 δ^{18}O record (Bickert et al., 1997) documents growth and decay of Northern Hemisphere ice mass. High δ^{18}O values mean large ice mass. It contains the Mid-Pleistocene Transition (MPT), a relatively sudden increase in mean ice mass at around 920 ka before present (BP) which led to Late Pleistocene ice ages (see Mudelsee and Stattegger, 1997, Table 1 for further references). For fit region, we choose the same time interval, [500 ka; 1400 ka], as Mudelsee and Stattegger (1997) for other records.

The ramp function fit (Fig. 12, Table 2) confirms timing and abruptness of the MPT. The estimated δ^{18}O step agrees with results from other records. The location of the ODP 925 drill hole (tropical W-Atlantic) adds further evidence that the MPT was a global climate phenomenon. Residual analysis (Fig. 13) attests that made assumptions are valid and the ramp form is suited. Dependency between the $e(i)$ is considerably strong ($\hat{a}=0.81$) which may be due to harmonic components (Milankovitch frequencies of Earth orbital parameter variations). That deviation from AR(1) dependency should make the nonparametric bootstrap (we take $p = 0.7$) the most reliable scheme. However, all three resampling methods reveal nearly zero bias of estimated parameters and no correlation between them (exception: $\hat{t}1^*(j) < \hat{t}2^*(j)$). Estimated parameter uncertainties (Table 2) roughly agree with those of previous quantifications of the MPT on other records (Mudelsee and Schulz, 1997; Mudelsee and Stattegger, 1997). The estimated uncertainties of $\hat{t}1$ and $\hat{t}2$ are about 7–10 times as large as the average spacing, justifying the coarse search grid.

5.2. 87-Sr/86-Sr

The 87-Sr/86-Sr record (Hodell et al., 1989), constructed from various marine sedimentary cores, documents the geochemical cycling of strontium in the late Neogene ocean which is related to processes such as chemical weathering of the continents, hydrothermal circulation at mid-ocean ridges, and carbonate dissolution on the seafloor. Hodell et al. fitted a ramp function per eye ($\hat{t}1=4.5$ Ma, $\hat{x}1=0.709025$, $\hat{t}2=5.5$ Ma, $\hat{x}2=0.708925$) and concluded that the 87-Sr/86-Sr record additionally offers considerable potential as a stratigraphic tool for the transition interval. At $\sigma(t)$ we overtake the reported measurement errors. (Per-eye fit of the standard deviation gave similar results.)

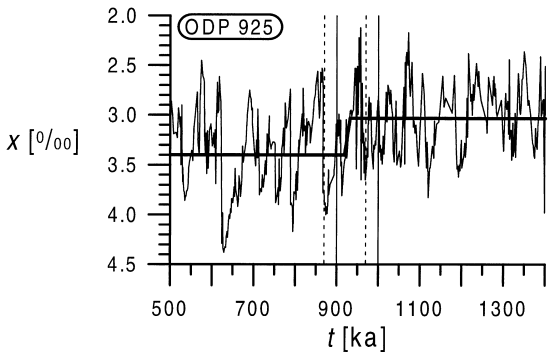

Fig. 12. Ramp function fit (heavy line) to ODP 925 time series (Tables 1 and 2). Grid boundaries for $t1$-search (dashed vertical lines) and $t2$-search (solid).

The ramp function regression carried out with RAMPFIT (Fig. 14, Table 2) confirms the findings of Hodell et al. Residual analysis (Fig. 15) attests that made assumptions are valid and the ramp form is suited. There is no dependency between the $e(i)$, and all three resampling schemes reveal the following similar results: negligible biases for the estimated parameters and minor correlation between them. The estimated uncertainties of $\hat{t}1$ and $\hat{t}2$ (Table 2) are about as large as the average spacing, which seems to be still acceptable.

5.3. ODP 929

The ODP 929 δ^{18}O time series (Zachos et al., 1997) from the tropical W-Atlantic is a high-resolution record of Global Change over the Oligocene/Miocene boundary. Climate was moderately warm but punctuated by episodes of global cooling. The first and largest such episode is Mi-1, at around 23.8 Ma BP. The ODP 929 record mainly reflects changes in bottom water temperature (signal proportion > 2/3 (J.C. Zachos, pers. comm., 1998)), which enables us to quantify the Mi-1 event. High δ^{18}O values mean low temperature. We use the same time interval, [23.4 Ma; 24.6 Ma], as Zachos et al. (1997: Fig. 1B) for analyzing Mi-1.

A *single* ramp function would not provide an adequate model for this data. This can be seen directly from Fig. 16 and would also be revealed by residual analysis. Instead we fit *two* ramps, the older transition corresponding to initial, strong cooling, the younger to the following, minor warming (Fig. 16, Table 2). Residual analysis for the old part (Fig. 17) generally confirms the ramp form. However, the Gaussian assumption is eventually violated. Likewise, the regression model cannot be rejected for the young part (Fig. 18). But, a good alternative might be a harmonic model for the mean which would describe the Mi-1 event as part of a 400 ka Milankovitch cyle, see also Zachos et al. (1997). Bootstrap resampling does not indicate estimation bias, except perhaps of minor degree of $\hat{x}1$ and $\hat{x}2$ in the old part. For both parts the esti-

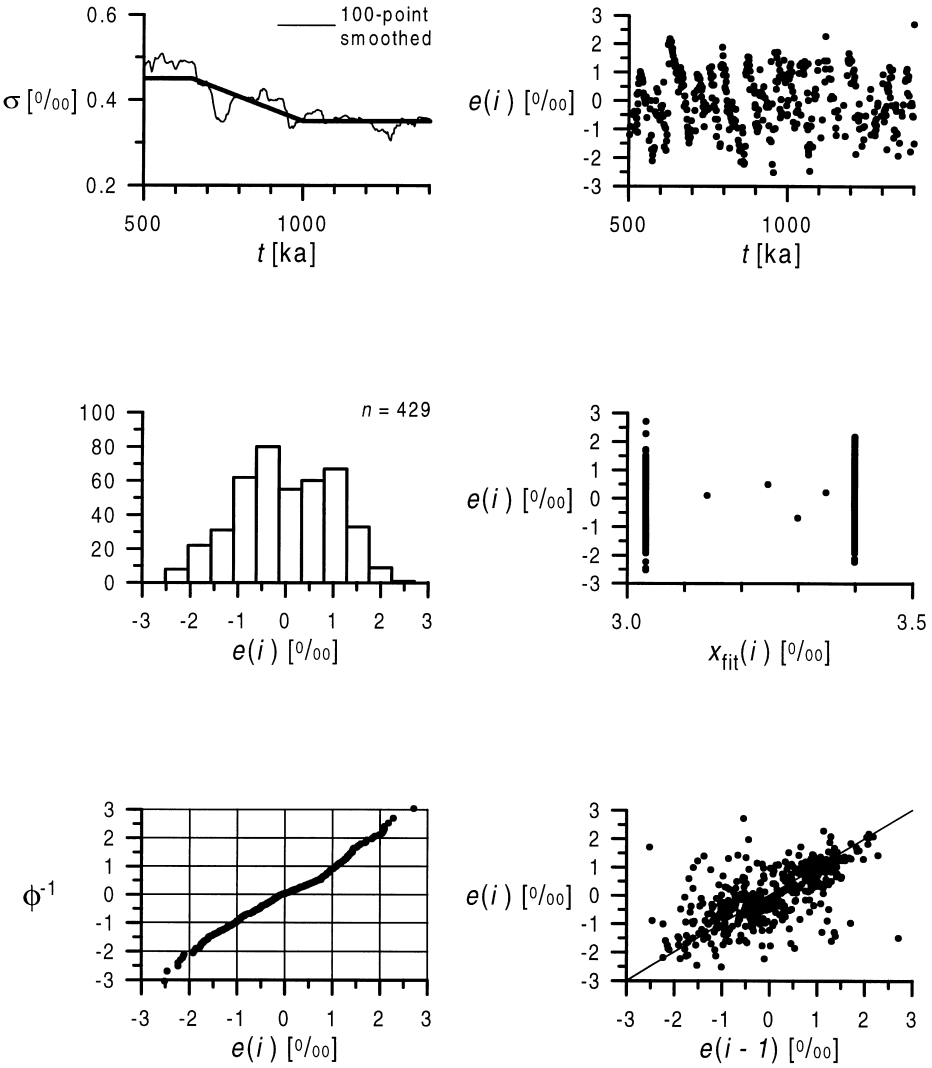

Fig. 13. Standard deviation fit and residual analysis for ODP 925 time series (Tables 1 and 3).

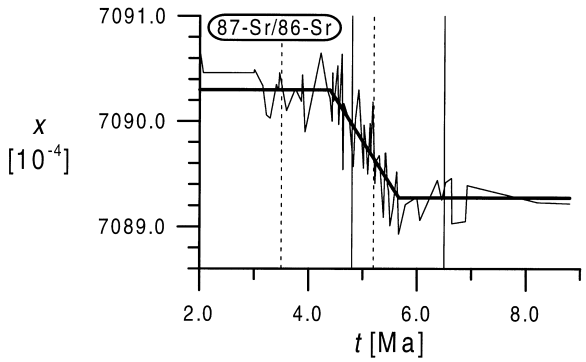

Fig. 14. Ramp function fit (heavy line) to 87-Sr/86-Sr time series (Tables 1 and 2). Grid boundaries for $t1$-search (dashed vertical lines) and $t2$-search (solid).

mated uncertainties of $\hat{t}1$ and $\hat{t}2$ are about 3 times as large as the average spacing.

6. Conclusions and summary

A method is presented accompanied by a Fortran 77 computer program (RAMPFIT), for fitting a ramp function (Fig. 1) to measured paleoclimatic time series.

- A brute-force search over a grid (Fig. 2) estimates $t1$ and $t2$, whereas a weighted LS estimation provides $\hat{x}1$ and $\hat{x}2$. An intelligent search path reduces costs. Further, RAMPFIT's interactive graphics facilitate keeping the grid boundaries narrow. The example time series (one artificial, three measured; $n \leq 500$)

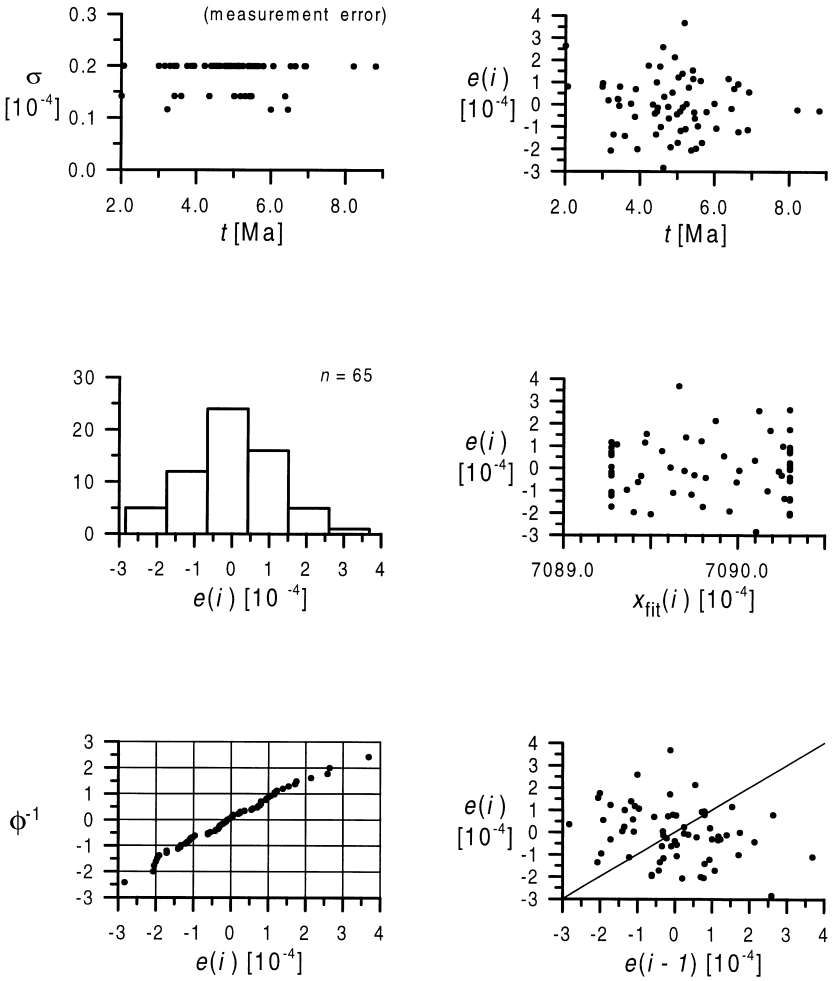

Fig. 15. Standard deviation and residual analysis for 87-Sr/86-Sr time series (Tables 1 and 3).

show the practicability of this approach in terms of computing costs and precision of the solution: estimated parameter uncertainties were greater than the grids were coarse.

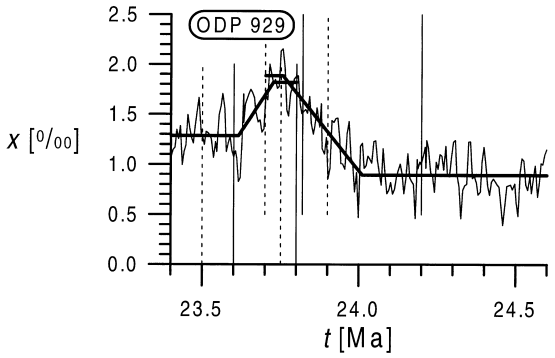

Fig. 16. Two-fold ramp function fit (heavy line) to ODP 929 time series (Tables 1 and 2). Grid boundaries for $t1$-searches (dashed vertical lines, those for young part border on t-axis) and $t2$-searches (solid).

- $\sigma(t)$ is fitted per eye, also as a ramp function. For the example time series, this seems to be sufficient for providing approximate weights, including when dependency is present.
- Interesting extensions/alternative approaches would be: (1) generalized LS; (2) numerical minimization; (3) a continuously differentiable approximation of the ramp function, and (4) LS fitting of $\sigma(t)$.

RAMPFIT includes an analysis of weighted ordinary residuals for testing whether made assumptions (namely Gaussian distribution, appropriate weighting, model suitability, and only weak AR(1) dependency) are valid. In the example of the measured climate time series (Table 1), we found that:

- The deviations from Gaussianity are of only minor

degree.

- The ramp function proved to be a suitable model. In one case (ODP 929) its form was too simple and had to be changed for a two-fold ramp.
- Serial dependency between the $e(i)$ might be strong in the presence of harmonic components (ODP 925).

RAMPFIT supplies three bootstrap resampling schemes (nonparametric stationary, parametric, and wild) for determining the accuracy of estimated parameters.

- For the measured time series, the three schemes produced similar results. It remains to be investigated further which resampling scheme is most suitable. Using one scheme alone should be sufficient. The nonparametric stationary bootstrap might be favorable because it might best tolerate deviations from Gaussianity or AR(1) dependency.
- An eventual important addition would be the incorporation of a scheme which also resamples the time.
- It is recommended that 200 bootstrap replications (Efron and Tibshirani, 1993) be used. For typical paleoclimatic time series, this demands computing power at the Pentium/workstation level.

To summarize, ramp function regression seems to be a valuable tool for quantifying climate transitions on measured time series. Such a functional form is not unusual in the geologic past. Modern computer power allows practical determination of parameter uncertainties.

Acknowledgements

Thanks for discussions during these five years of development go to: Karl Stattegger and Michael Sarnthein (both at GPI, Kiel, Germany), Pieter

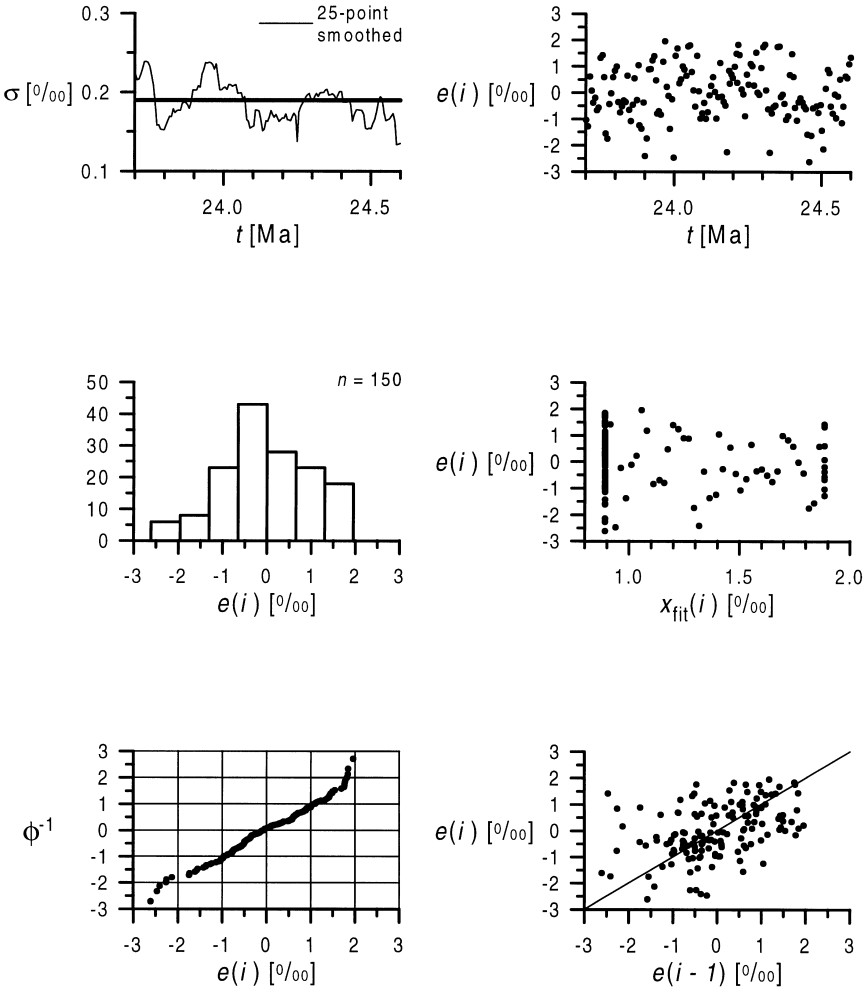

Fig. 17. Standard deviation fit and residual analysis for ODP 929 time series, old part (Tables 1 and 3).

Grootes (Leibniz Laboratory, Kiel, Germany), Michael Schulz (SFB 313, Kiel, Germany), Martin H. Trauth (Geosciences, Potsdam, Germany), Olaf Müller (SAP, Walldorf, Germany), Qiwei Yao (IMS, Canterbury, UK), and Cees G. H. Diks (CeNDEF, Amsterdam, The Netherlands). The term 'ramp function' I owe to Wolfgang H. Berger (Scripps, San Diego, USA). Torsten Bickert (GeoB, Bremen, Germany) and James Zachos (Earth Sciences, Santa Cruz, USA) kindly provided their data. Walter Schwarzacher and Karl Stattegger gave constructive review comments. Alistair Duncombe (IMS, Canterbury, UK) helped refining the English. Financial support by the Deutsche Forschungsgemeinschaft (GK Kiel) and the European Union (Marie Curie Research Grant ERBFM-BICT971919) is acknowledged.

Appendix A. RAMPFIT manual

RAMPFIT, written in Fortran 77, guarantees high portability. RAMPFIT was tested on a PC with 80486-50 DX processor under MS-DOS 6.2, and on a Sun Ultra SPARC 5.5 workstation with 143 MHz clock under Unix V 4.0, but it is expected to run on other systems as well.

During execution, RAMPFIT regularly calls the gnuplot program for plotting results of current calculations on the screen; thus a Fortran compiler is needed which allows external calls. (Microsoft Fortran Power-Station 1.0 compiler was used for the PC, and the EPCF77 2.7 compiler for the workstation.)

A copy of the Fortran code and hints for installa-

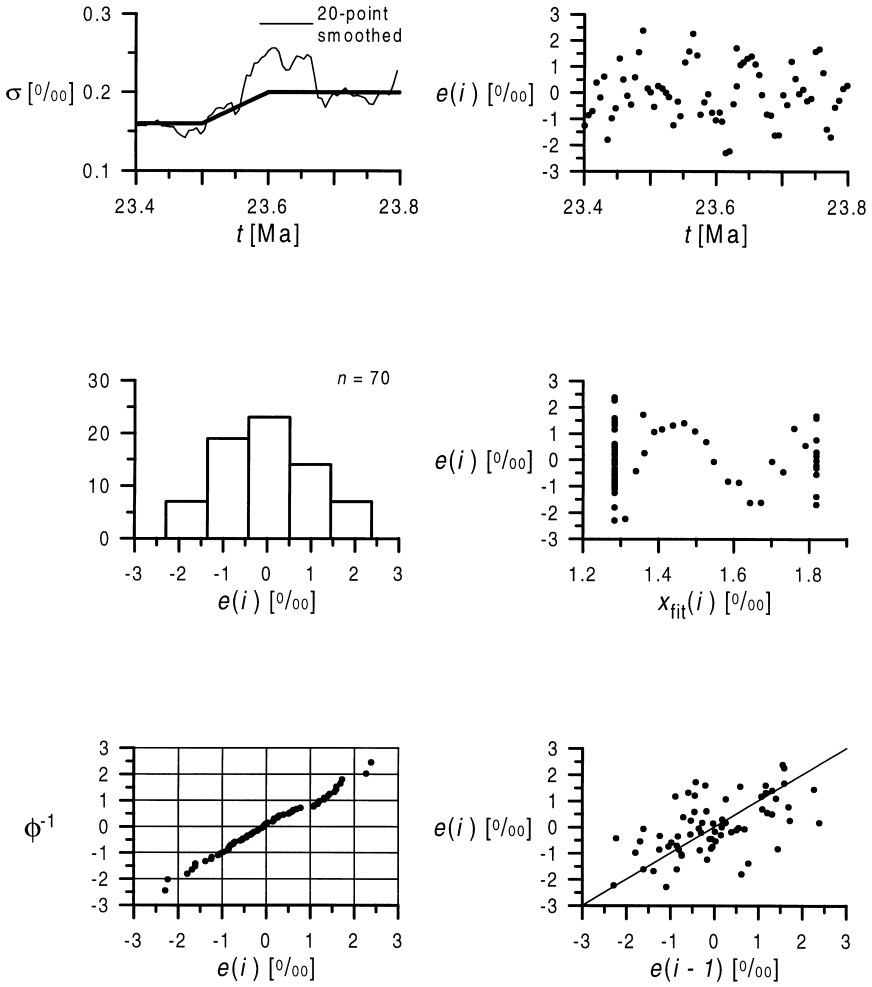

Fig. 18. Standard deviation fit and residual analysis for ODP 929 time series, young part (Tables 1 and 3).

tion may be requested from the author. Copies of gnuplot (version 3.6 or higher) can be obtained from the internet (for example, FTP directory, 1999).

A1. Part 1: detection

After you have started the program and supplied path and name of the data file (x vs t), and the total number of points, RAMPFIT first helps to detect a transition. It plots the original time series (continue with 'Enter'), and then it displays the Part 1 decision tree and the Part 1 information. That field informs about: current time interval and number of points, and (separately for x and the time spacing) the maximum, minimum, mean, and standard deviation. The decision tree allows you to choose a time interval and perform k-nearest-neighbor smoothing (in x and t together) for the mean and the standard deviation. Repeat this until you have a first guess about where a transition is located and how the time-dependent mean and standard deviation look. You proceed then to Part 2 where you take the current time interval for the regression.

A2. Part 2: regression

To provide the weighting for the regression, you first prescribe time-dependent standard deviation, either as a constant or as a ramp. Then you adjust the boundaries of the $t1$–$t2$ search grid. RAMPFIT calculates the best fit and subsequently plots it (see Fig. 3). The next plot (see Fig. 6) shows $\hat{\sigma}(t)$ (that is, the prescribed standard deviation) compared against a k-nearest-neighbor smoothing of $[x(i) - \hat{x}_{\mathrm{fit}}(i)]$. Then the Part 2 decision tree appears with Part 2 information. That field displays: current time interval, prescribed standard deviation, $t1$–$t2$ search boundaries, and the fit result. The latter comprises $\hat{t}1$, $\hat{x}1$, $\hat{t}2$, $\hat{x}2$, $SSQWN$, and $SSQWNS := SSQWN/[\text{average of } \hat{\sigma}(t)]$. This quantity helps to compare two fits that differ in $\hat{\sigma}(t)$. The decision tree allows new per-eye fits of $\sigma(t)$ and evaluation of new search boundaries; a new fit follows each adjustment. Repeat this until you are convinced of $\hat{\sigma}(t)$ and the $t1$–$t2$ search grid. You proceed to Part 3 to analyze the regression residuals. (Alternatively you may go back to Part 1 to evaluate a new time interval.)

A3. Part 3: residual diagnostics

The residual plots appear in succession: histogram of $e(i)$ (see Fig. 4), normal probability plot (Fig. 5), $e(i)$ against time (Fig. 7), $e(i)$ against $x_{\mathrm{fit}}(i)$ (Fig. 8), and $e(i)$ against $e(i-1)$ (Fig. 9). Then the Part 3 decision tree appears, with the Part 2 information field (time interval, fit result) and the Part 3 information field. The latter displays the mean, standard deviation, and standard error or $e(i)$, and the result of the AR(1) fit, namely the values of $\hat{\tau}$ (same units as $t(i)$) and $\hat{a}$. The decision tree allows you to go back to Part 1 or 2. This may be necessary for achieving a properly selected time interval, an accurate fit to the standard deviation, and a well-adjusted $t1$–$t2$ grid. Part 4, bootstrap resampling, costs the most CPU time. It is recommended that the grid boundaries for the $t1$-search ($t2$-search) be centered about $\hat{t}1$ ($\hat{t}2$) and not be adjusted too marrow in order to keep the number of grid boundary solutions small.

A4. Part 4: bootstrap resampling

The Part 4 decision tree appears with Part 3 information (residuals). This field helps to take into account serial correlation between the $e(i)$. The decision tree allows you to select among nonparametric stationary, parametric, and wild bootstrap. Each scheme demands a choice for the value of B and a seed for the random number generator. The nonparametric scheme additionally needs a value for p as input. RAMPFIT suggests a value which is determined by $\hat{\tau}$ (see Subsection 3.5). You may, however, wish to take a smaller value which would increase the average block length. The parametric bootstrap requires input of a value for the AR(1) correlation a; you may simply use $\hat{a}$.

After calculation with a particular resampling scheme, RAMPFIT gives: $\hat{t}1$, average of $\hat{t}1^*(j)$, standard deviation of $\hat{t}1^*(j)$, median, mad', minimum, maximum, number nl of solutions on the left (lower) search boundary, and number nr of those on the right boundary; analogously for $x1$, $t2$, and $x2$. RAMPFIT also gives the results from other schemes, if such exist. Note that only the most recent results are stored for each scheme; re-running a scheme will overwrite its previous results. Now you can visualize the existing bootstrap results by selecting a scheme, a first variable ($t1$, $x1$, $t2$, $x2$), and a second variable ($t1$, $x1$, $t2$, $x2$, 'no'). The value 'no' leads to the bootstrap histogram for the first variable (see Fig. 10), the other values to bootstrap scatterplots (as in Fig. 11). You can repeat visualization or go back to the Part 4 decision tree. From there, you may re-run the same scheme with a new value for B, select one other scheme, go back to Part 1, 2, or 3, or exit RAMPFIT. The latter choice first allows you to change the mean smoothing and the standard deviation smoothings (for graphical reasons). Finally, RAMPFIT prints the output file names and their respective content on the screen.

References

Basseville, M., Nikiforov, I.V., 1993. Detection of Abrupt Changes: Theory and Application. Prentice-Hall, Englewood Cliffs, NJ [447 pp. (http://www.irisa.fr/sigma2/kniga/, 2 November 1998)].

Bevington, P.R., Robinson, D.K., 1992. Data Reduction and Error Analysis for the Physical Sciences, 2nd ed. McGraw-Hill, New York 328 pp.

Bickert, T., Curry, W.B., Wefer, G., 1997. Late Pliocene to Holocene (2.6–0 Ma) western equatorial Atlantic deep-water circulation: inferences from benthic stable isotopes. In: Shackleton, N.J., Curry, W.B., Richter, C., Bralower, T.J. (Eds.). Proceedings of the ODP, Scientific Results, 154. Ocean Drilling Program, College Station, TX, pp. 239–254.

Efron, B., Tibshirani, R.J., 1993. An Introduction to the Bootstrap. Chapman & Hall, London 436 pp.

FTP directory., 1999. ftp://cmpc1.phys.soton.ac.uk/incoming/, 6 March 1999.

Hamaker, H.C., 1978. Approximating the cumulative normal distribution and its inverse. Applied Statistics 27 (1), 76–77.

Härdle, W., Marron, J.S., 1991. Bootstrap simultaneous error bars for non-parametric regression. The Annals of Statistics 19 (2), 778–796.

Hodell, D.A., Mueller, P.A., McKenzie, J.A., Mead, G.A., 1989. Strontium isotope stratigraphy and geochemistry of the late Neogene ocean. Earth and Planetary Science Letters 92 (2), 165–178.

Montgomery, D.C., Peck, E.A., 1992. Introduction to Linear Regression Analysis, 2nd ed. Wiley, New York 527 pp.

Mudelsee, M., Schulz, M., 1997. The Mid-Pleistocene climate transition: onset of 100 ka cycle lags ice volume build-up by 280 ka. Earth and Planetary Science Letters 151 (12), 117–123.

Mudelsee, M., Stattegger, K., 1997. Exploring the structure of the mid-Pleistocene revolution with advanced methods of time-series analysis. Geologische Rundschau 86 (2), 499–511.

Müller, H.-G., 1992. Change-points in nonparametric regression analysis. The Annals of Statistics 20 (2), 737–761.

Park, S.K., Miller, K.W., 1988. Random number generators: good ones are hard to find. Communications of the ACM 31 (10), 1192–1201.

Politis, D.N., Romano, J.P., 1994. The stationary bootstrap. Journal of the American Statistical Association 89 (428), 1303–1313.

Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T., 1992. Numerical Recipes in Fortran, 2nd ed. Cambridge University Press, Cambridge 963 pp.

Robinson, P.M., 1977. Estimation of a time series model from unequally spaced data. Stochastic Processes and their Applications 6, 9–24.

Scott, D.W., 1979. On optimal and data-based histograms. Biometrika 66 (3), 605–610.

Tishler, A., Zang, I., 1981. A new maximum likelihood algorithm for piece-wise regression. Journal of the American Statistical Association 76 (376), 980–987.

Tukey, J.W., 1977. Exploratory Data Analysis. Addison-Wesley, Reading, MA 688 pp.

Williams, D.A., 1970. Discrimination between regression models to determine the pattern of enzyme synthesis in synchronous cell cultures. Biometrics 26 (1), 23–32.

Zachos, J.C., Flower, B.P., Paul, H., 1997. Orbitally paced climate oscillations across the Oligocene/Miocene boundary. Nature 388 (6642), 567–570.