Statistics for Strontium Isotope Stratigraphy: A Robust LOWESS Fit to the Marine Sr-Isotope Curve for 0 to 206 Ma, with Look-up Table for Derivation of Numeric Age[1]

Richard J. Howarth and John M. McArthur

Research School of Geological and Geophysical Sciences, Birkbeck and University College London, Gower Street, London WC1E 6BT, England

ABSTRACT

We provide a best-fit curve to 1849 strontium isotope data for the period 0 to 206 Ma using the LOcally-WEighted regression Scatterplot Smoother (LOWESS) method. This is a robust, nonparametric modern regression technique. Since it does not yield an explicit mathematical equation relating $^{87}Sr/^{86}Sr$ to time, a look-up table to determine numeric age has been generated in steps of 1×10^{-6} in $^{87}Sr/^{86}Sr$. The calibration uses the timescales of Shackleton and coworkers for 0–7 Ma; Cande and Kent for 7–72 Ma; Obradovich for 72–95 Ma and Gradstein and coworkers for >95 Ma. The look-up table includes 95% confidence intervals on the predictions of numeric age. When using this table, the uncertainty on the $^{87}Sr/^{86}Sr$ of the sample whose age is sought must be added to that inherent in the LOWESS regression. We show how to determine the uncertainty in $^{87}Sr/^{86}$, i.e., how best to obtain the 95% confidence bounds on a single measurement of $^{87}Sr/^{86}$ for a sample, and on the mean $^{87}Sr/^{86}$ value for 2 or more replicate measurements of $^{87}Sr/^{86}$ in the sample; these confidence intervals reflect analytical-system reproducibility for routine samples (as opposed to that of standard control materials, e.g., NIST 987) and are necessary to establish the final upper and lower bounds on predicted numeric age.

Introduction

While strontium isotope stratigraphy (SIS) is predominantly used for correlation and dating of marine sediments (Elderfield 1986; McArthur 1994; Veizer 1989) it can also distinguish between marine and freshwater environments (Schmitz et al. 1991) and identify reworking in sediments (MacLeod and Huber 1996). When SIS is used for dating, the quality of its numeric date depends upon several factors, including: the preservation quality of the original $^{87}Sr/^{86}Sr$ of a sample, whether artifacts are introduced during $^{87}Sr/^{86}Sr$ measurement, the slope of the curve of $^{87}Sr/^{86}Sr$ against numeric age, the accuracy of the age model used to assign a numerical calibration to the isotope curve, and finally, the way the best-fit curve and its associated uncertainty envelope are fitted to the Sr isotope data and used to derive a numeric age. The last of these factors is the focus of this paper.

$^{87}Sr/^{86}Sr$ data plotted against numeric age define a broad band of points. Isotope stratigraphers face the challenge of finding the best way to reduce that band to a best-fit curve, to derive uncertainty estimates on the curve's position, and to deduce the total uncertainty in a derived age by comparing to it an independent $^{87}Sr/^{86}Sr$ measurement.

This paper proposes a best way to do these things. We show that the statistical LOcally-WEighted regression Scatterplot Smoother (LOWESS) curve-fitting method of Cleveland (1979; Cleveland et al. 1992) provides an excellent way to calculate a best-fit model for the global standard strontium curve (GSSC hereafter) and how this best-fit curve can be used to derive numeric age from a measured $^{87}Sr/^{86}Sr$ value. We also provide a look-up table from which numeric age and the uncertainties can be derived from $^{87}Sr/^{86}Sr$. The table is large, so only an extract is reproduced in this paper: the full table, interpolated at intervals of 0.000001 in $^{87}Sr/^{86}Sr$ in order to permit precise determinations of numeric age, is available from the authors. In compiling the look-up table we have not used all the data available in the literature but have selected what we regard as the most appropriate.

[1] Manuscript received December 1, 1995; accepted November 19, 1996.

[The Journal of Geology, 1997, volume 105, p. 441–456] 0022-1376/97/10504-0008$01-00

As new data are released they will be evaluated and, where suitable, will be incorporated into the database and the GSSC recomputed. A note will be published showing the revisions, and a new look-up table will be made available to interested parties and sent automatically to recipients of the original. Readers are warned that the nonparametric fit, and so the resulting look-up table, is no better than the data used.

Best-Fit Functions for the Variation of $^{87}Sr/^{86}Sr$ through Time

Parametric Models. Most workers compute best-fit trends of $^{87}Sr/^{86}Sr$ with time using a variety of simple and familiar methods, such as linear or polynomial regression. These regressions are termed parametric because they are obtained by estimation of the parameters, i.e. the coefficients, of either linear ($y = a + bx$) or nth order polynomial ($y = a + bx + c^2 + dx^3$.....) equations (third-order in this example).

Users of linear regression must decide which of numeric age and $^{87}Sr/^{86}Sr$ is the dependent variable, and which the independent variable, as the former must be regressed on the latter. Miller et al. (1988) and McKenzie et al. (1988) treat age as dependent, whilst Hodell et al. (1991), Miller et al. (1991), Hodell and Woodruff (1994), and Oslick et al. (1994) treat $^{87}Sr/^{86}Sr$ as dependent. In practice, the choice makes little difference to the derived ages for timespans of a few million years (McArthur 1994). In some cases (e.g., figure 2 of Miller et al. 1991), separate regression lines fitted to adjacent segments of the GSSC may cross (i.e., they have discontinuous derivatives at the joint), but this may not always be the case (e.g., Oslick et al. 1994). Users of polynomial regression have the same decisions to make, in addition to deciding, arbitrarily, the order of the polynomial to use. It has been common to regress $^{87}Sr/^{86}Sr$ on age (Miller et al. 1991; Hodell and Woodruff 1994; Oslick et al. 1994; Farrell et al. 1995; Sugarman et al. 1995), and up to ninth-order equations have been fitted (e.g., Hodell and Woodruff 1994).

Although parametric regressions provide an apparently adequate fit over a few million years, there is no reason to suppose that the true GSSC conforms to a polynomial function. The use of such functions can distort the fit by, for example, forcing the number of inflection points by the number of terms (order) of the polynomial equation, rather than by the natural inflection points in the data trend. Use of such simplistic models can only be justified in terms of computational convenience. As a consequence, workers are increasingly turning to other methods of curve-fitting, particularly nonparametric methods (Ludwig et al. 1988; McArthur 1994; Smalley et al. 1994: Mc Laughlin et al. 1995).

Nonparametric Methods. Successful attempts have been made to get away from the purely mathematical constraints inherent in parametric methods by using nonparametric regression techniques. These methods require no assumptions regarding the underlying form of the relationship between $^{87}Sr/^{86}Sr$ and numeric age. Ludwig (1987, 1990) and Ludwig et al. (1988) used an approach based on smoothing splines (Wahba 1975), while the locally-weighted regression scatter-plot smoother, LOWESS (Cleveland 1979) has been used by RH in McArthur (1994), Smalley et al. (1994) and McLaughlin et al. (1995).

Comparison between the smoothing-spline and LOWESS approaches is difficult. Aspects of both are discussed in specialist texts such as Thisted (1988), Hastie and Tibshirani (1990), and Chambers and Hastie (1992). The choice of the number and placement of points of inflection (knots) in the smoothing-spline approach is both subjective and critical; problems with displacement of valleys and ridges in the fitted model can arise if the choice is poor. Although detailed comparative studies of the advantages of these two alternative approaches for GSSC calibration are desirable, our experience to date (McArthur 1994; Smalley et al. 1994; McLaughlin et al. 1955) suggests that LOWESS provides an excellent method for obtaining GSSC calibration, particularly over long time intervals within which the Sr-isotope curve exhibits many turning points. A particular value of the LOWESS method is that it is very resistant to bias in fitting caused by outliers (i.e., apparent aberrant values) in the data.

The LOWESS Regression Method

Introduction. In common with other workers (Hodell et al. 1991; Miller et al. 1991; Hodell and Woodruff 1994; Oslick et al. 1994; Farrell et al. 1995; Sugarman et al. 1995), we treat $^{87}Sr/^{86}Sr$ as the dependent variable and consequently assign numeric age to the x-axis and $^{87}Sr/^{86}Sr$ to the y-axis (the issue of uncertainty in the ages is discussed later). In the LOWESS method (Cleveland 1979; Chambers et al. 1983; Thisted 1988; Cleveland et al. 1992) a smooth curve is fitted to a set of n data points as follows:

A smoothing parameter (the *span, q*), defined as a fraction of the total number of data points (n), is first chosen. An appropriate span value is chosen

using trade-off between goodness-of-fit at local and global scales. The overall fitted curve must be faithful to the major turning points in the GSSC while minimizing small-scale crenulations about the main trend. A recent study of nonparametric regression by Marron and Tsybakov (1995) has shown that, because the usual goodness-of-fit criteria measure something different from what the eye can see in a graphic presentation, guided trial-and-error visual fitting (particularly by an experienced data-analyst) still remains a very effective way of choosing a smoothing parameter. In view of this, we used a combination of measures to determine the choice of span, such as the overall residual sum-of-squares, rate of change of slope and number of turning points (i.e. locations at which the slope of the GSSC changes sign).

Having set the span, a window parallel to the y ($^{87}Sr/^{86}Sr$) axis with a width equal to the span is then successively centered on each data point in turn, in order of increasing x (age) value. At each window position, the best-estimate of the mean value of y corresponding to the value x_0 lying at the centre of the window is determined using local regression, explained in more detail below. This yields a single predicted value of $^{87}Sr/^{86}Sr$ corresponding to the current window position. The overall LOWESS fit constitutes the set of calculated best-estimates of $^{87}Sr/^{86}Sr$ located at each age value, once the entire data set has been traversed.

Local Fitting Within Each Window. The k = integer$(qn + 0.5)$ nearest-neighbors on either side of x_0 (x_i; $i = -k$ to k) are used as a basis for estimation of the mean value of y which corresponds to x_0. Since the local density in x varies along the data sequence, the width of each window will vary correspondingly. All the data points that lie within the current window are given an initial local neighborhood weight (w_d) using the symmetric bicubic function:

$$w_d = \left\{ \begin{array}{c} (1 - |d|^3)^3 \quad \text{for } 0 \le |d| < 1 \\ 0 \text{ otherwise} \end{array} \right\} \qquad (1)$$

where

$$d_i = |x_0 - x_i| / \max(|x_0 - x_i|). \qquad (2)$$

$|d|$ denotes the absolute value of d, i.e. the magnitude of the quantity without regard to its sign, and the subscript refers to the ith data point within the window. These local weights will have their maximum at (x_0) and will decrease to zero at its furthest neighbors (x_{-k} and x_k), which lie at opposite edges of the window. Note that these local weights are in addition to any *global* pre-weighting assigned to the raw data, on input to the LOWESS smoothing process, to reflect relative data quality. (We return to this aspect later.) The y_i points within the window are then fitted by a quadratic function:

$$\hat{y}_i = b_0 + b_1 x_i + b_2 x_i^2 \qquad (3)$$

where $\hat{y}_i$ is the estimated value of y at location x_i. The values of the parameters $\{b_0, b_1, b_2\}$ of the fitted equation are those for which the weighted sum-of-squares $\sum_{i=-k,k} w_{d_i} (y_i - \hat{y}_i)$ becomes a minimum. The residuals from the regression are the differences between the observed and fitted $^{87}Sr/^{86}Sr$ values:

$$r_i = y_i - \hat{y}_i. \qquad (4)$$

In order to down-weight any data points that have large residual values, the local fit is repeated, but this time using a bisquare weighting function based on the magnitude of the residuals (Huber 1981; Hastie and Tibshirani 1990; Hastie 1992):

$$w_r = \left\{ \begin{array}{c} (1 - |u|^2)^2 \quad \text{for } 0 \le |u| < 1 \\ 0 \text{ otherwise} \end{array} \right\} \qquad (5)$$

where

$$u_i = r_i / 6(\text{median}|r_i|);\ i = j = -k,k \qquad (6)$$

and the local quadratic function (equation 3) is fitted (with w_r now replacing w_d), and a corresponding new set of residuals (equation 4) is computed. This robust-fitting cycle is repeated until the sum of the squared weighted-residuals reaches an overall minimum. The outcome is the final estimate of the mean value of $^{87}Sr/^{86}Sr$ ($\hat{y}_0$) for the current window together with its associated pointwise upper and lower 95% confidence bounds (see Hastie and Tibshirani [1990] and Cleveland et al. [1992]), corresponding to the age at the center of the window (x_0) based on a local fit that is robust against any bias caused by the presence of outliers in the data.

Obtaining the Global LOWESS Fit. Once a local-fitting operation is completed, the window is then re-centered on the next x value, which then becomes x_0, and a new local fit begins. This process is continued until all the data points in the current data set have been visited in turn. The LOWESS curve is then given by the set of discrete $\hat{y}_0$ values, located at $x_i (i = 1,n)$, once the entire data set has been traversed.

Because the set of $\hat{y}_0$ values has been derived from a composite of overlapping local fits, its over-

all fit is smoothed with respect to the original data. However, the resultant ensemble of points is not, in itself, an analytical function, consequently, it cannot provide a unique formula with which to calculate either numeric age from $^{87}Sr/^{86}Sr$, or vice versa (as would be possible with, say, a globally fitted polynomial equation), nor does it explicitly define values of either variable in the intervals between the discrete data points. Estimation of these intermediate values therefore has to be carried out using interpolation.

Inverting the LOWESS Curve. To predict a sample's age based on its measured $^{87}Sr/^{86}Sr$ value, the LOWESS curve has to be inverted. To accomplish this, we first fitted separate conventional piecewise cubic splines through (i) the final LOWESS point estimates of $^{87}Sr/^{86}Sr$ values ($\hat{y}_0$); (ii) the pointwise lower 95% confidence bound; and (iii) the pointwise upper 95% confidence bound. In each case these were computed as functions of numeric age. In keeping with the classical regression approach to instrumental calibration (Draper and Smith 1981; Miller et al. 1991), we then inverted these spline-interpolated curves, in order to obtain numeric age as a function of $^{87}Sr/^{86}Sr$. This enabled the numeric age corresponding to the best $^{87}Sr/^{86}Sr$ estimate in a sample to be determined, either from a plot of the fitted curve (figure 1*a*), which is easy to approximate if the curve is overlain by a grid (figure 2) or, for more accurate work, by look-up table.

Our interpolation of the final LOWESS GSSC and its associated confidence bounds using conventional splines (numerically equivalent to using a flexible draftsman's curve to draw a line through the points defining the smooth curves) simply to derive the look-up table values should not be confused with direct estimation of a smoothed GSSC from raw data using *B*-splines (Ludwig 1987, 1990; Ludwig et al. 1988). Fitting smoothing *B*-splines similarly results in pointwise estimates of a smoothed curve (for detailed discussion see Hastie and Tibshirani 1990; Hastie 1992). Subsequent interpolation of this GSSC for prediction purposes would also be required, for similar reasons. In this case, we found that smoothing-splines showed no particular advantage as their cross-validation solutions were oversmoothed, and one was again forced to use a relatively subjective choice of a smoothing parameter.

Combining LOWESS Fits To Subsets. Because of the great variation with age of the local data density, and also because of the number of major turning points in the curve, we found that the best overall GSSC fit was obtained by splitting the dataset into six overlapped subsets: 0–1.40; 1.30–29.0; 28.0–63.0; 62.0–84.0; 83.0–118.0; and 116.0–206.0 Ma. A separate LOWESS fit was made to each. *B*-splines proved extremely useful for combining the smooth LOWESS curves (and their associated upper and lower confidence bounds) for the six overlapping subsets into a final composite GSSC, which enabled optimum local fits to be obtained throughout its entire range.

Result. The final smoothed combined LOWESS GSSC for predicted mean $^{87}Sr/^{86}Sr$, together with the original data, is shown in figure 1*a,* and the half-width of its 95% confidence interval is plotted in figure 1*b.* Enlarged views of overlapped sections of the final GSSC and their associated 95% confidence intervals are shown in figure 2*a–f.*

The computed GSSC embodied in the full table and illustrated in figure 2 requires some comment. First, data density between 0 and 28 Ma, and between 65 and 83 Ma is great. Consequently, the position of the mean $^{87}Sr/^{86}Sr$ is known with great certainty, the regression standard error is correspondingly small, and so the confidence limits on the mean are not resolved at the scale of the figure. Second, a distinct local maximum in $^{87}Sr/^{86}Sr$ with a magnitude of about 36×10^{-6} occurs in the Middle Eocene, where the curve has previously been thought to be flat. If confirmed by more data, this maximum may provide resolution for dating with SIS in the Eocene. Third, the rate of change in $^{87}Sr/^{86}Sr$ between the local minimum at 52.7 Ma and the K/T boundary at 65 Ma averages about 9×10^{-6}. With a resolution of 10×10^{-6} in measurement of $^{87}Sr/^{86}Sr$, this gives a potential mean temporal resolution over this interval of about 0.9 Ma with a 95% confidence interval half-width of 0.8 Ma. Fourth, a sharp and deep minimum occurs at 110.8 Ma, close to the Albian/Aptian boundary (112.2 Ma; Gradstein et al. 1995). The position of this minimum is constrained largely by the data of Jenkyns et al. (1995) which is, as these authors acknowledge, not well-constrained biostratigraphically. We have imposed a numerical calibration on their data that may be optimistic, and revision may be required in the future. Fifth, there is only one datum for the Berriasian (137.0 to 144.2 Ma), which may introduce artifact into the fitting of the GSSC to this interval, and to the curve near the Jurassic/Cretaceous boundary at 144.2 Ma. Sixth, the turning point at 156.0 Ma arises from the data of Podlaha (pers. comm. 1996) and it introduces structure into this part of the curve (Kimmeridgian, 150.7 to 154.1; Oxfordian, 154.1 to 159.4 Ma) where none was present in the data of Jones et al. (1994a). Last, the fit rejects a number of Toarcian data that plot between 186.0 and 187.5 Ma (around the *falciferum*

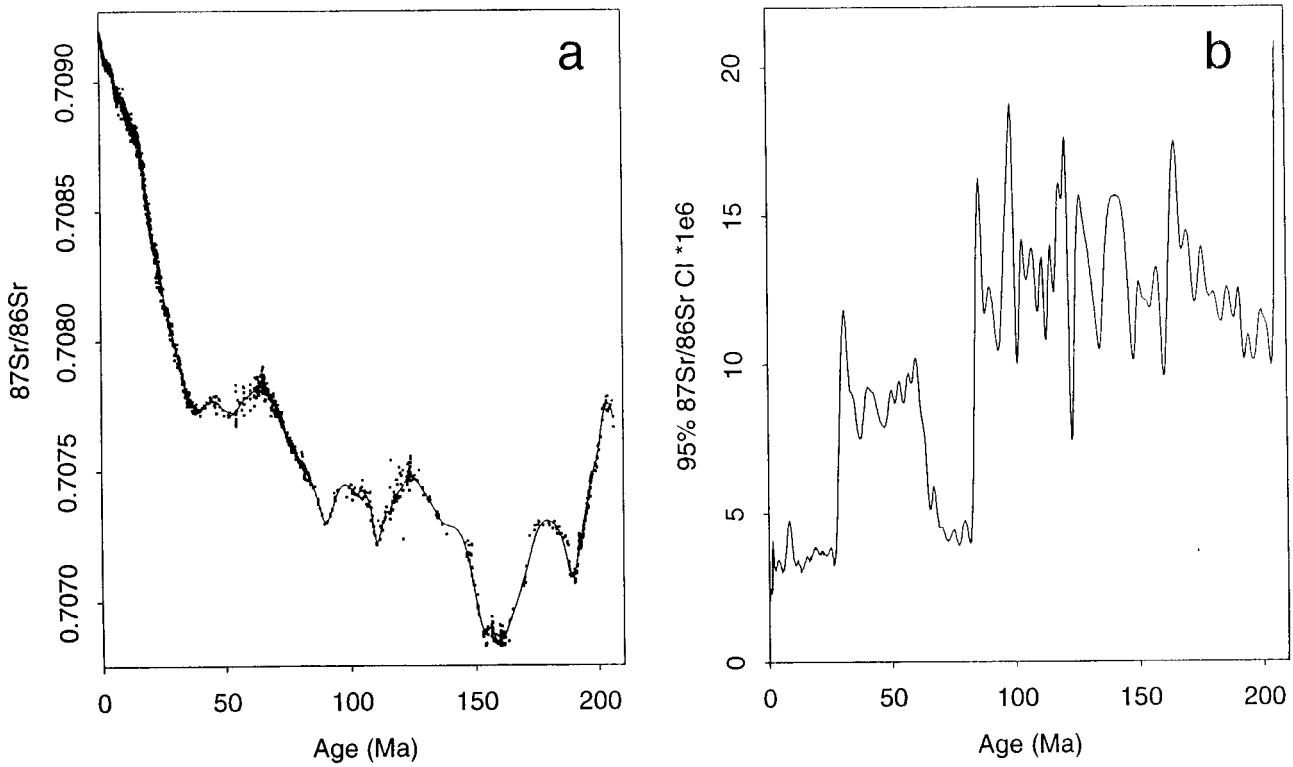

Figure 1. (*a*) Robust LOWESS fit to 1849 $^{87}Sr/^{86}Sr$ data for the period 0 to 206 Ma; curve shows mean $^{87}Sr/^{86}Sr$ corresponding to a given age (Ma) of a sample; solid dots show data points. *(b)* Half-width of the two-sided 95% confidence interval on mean $^{87}Sr/^{86}Sr$ ($\times 10^6$) as a function of age (Ma).

Zone; Jones et al. 1994*a*) where a discontinuity in the curve may reveal a hiatus in the sedimentation at this time in Yorkshire. We are analyzing more samples from this interval to resolve this problem.

Data Selection for the GSSC

Sources. Our LOWESS model for $^{87}Sr/^{86}Sr$ data for the period 0 to 206 Ma is based on 1849 data points from the literature (table 1). Rather than "average" all data available for any period we have generally selected the most recently published data because analytical techniques and supporting biostratigraphy have improved with time (McArthur 1994). For example, we use only Farrell et al. (1995) for the period 0–7 Ma. Where data sets join, they have been overlapped slightly in time to ensure statistical continuity; e.g., we have overlapped Hodell et al. (1989, 1991) and Farrell et al. (1995) from 5.5 Ma to 7 Ma.

Although a LOWESS fit for the entire Phanerozoic was provided by Smalley et al. (1994), those authors did not use a substantial body of new data we have incorporated here (Hess et al. 1989; Montanari et al. 1991; Denison et al. 1993; Paytan et al. 1993; Hodell and Woodruff 1994; McArthur 1994; McArthur et al. 1994; Farrell et al. 1995; Mead and Hodell 1995; Sugarman et al. 1995) and McArthur et al. (unpub. data), nor did they provide a look-up table.

Bias. As shown in McArthur (1994) and Hodell and Woodruff (1994), substantial interlaboratory differences exist for measured values of standard reference material NIST 987 (previously named SRM 987) and modern seawater strontium (MSS). Such bias can be monitored only by: (i) regular in-house use of standard control materials (Analytical Methods Committee 1995); and (ii) regular participation in interlaboratory cooperative trials using a variety of control materials (Pierson and Fay 1959; Analytical Methods Committee 1987; Mandel 1991). The latter should enable recognition of bias owing to the different physical nature of routine samples from that of control materials. Such procedures should enable a bias correction to be applied to all results from a given laboratory. Current practice in this respect is ad hoc and far from ideal.

We have corrected for interlaboratory bias by adjusting all data to a value of 0.709175 for MSS. In addition, 20×10^{-6} has been added to all data from the University of Florida to correct for an apparent interlaboratory bias that remains after correction to

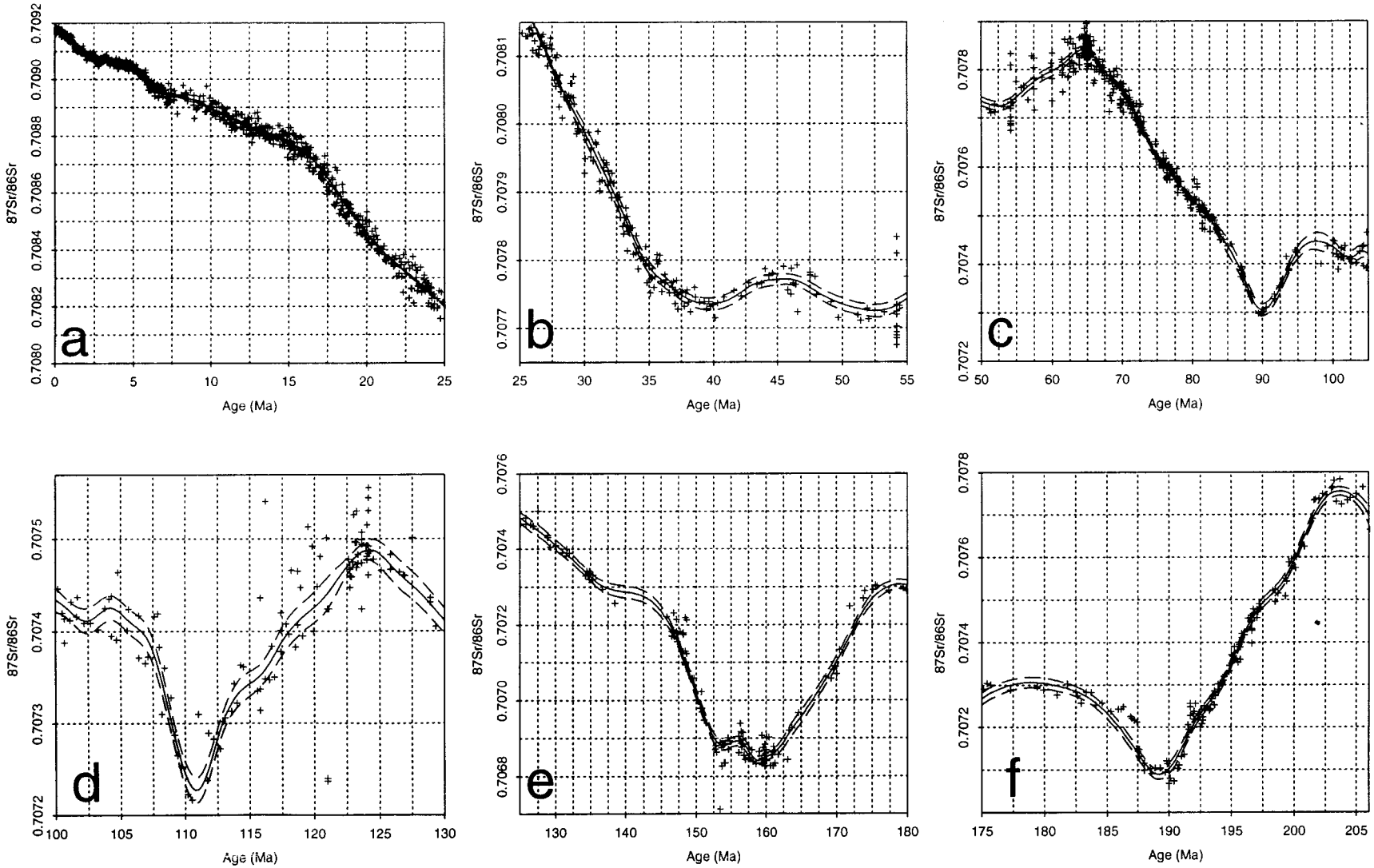

Figure 2. Details of robust LOWESS fit to $^{87}Sr/^{86}SR$. (*a*) 0–25 Ma; (*b*) 25–55 Ma; (*c*) 50–105 Ma; (*d*) 100–130 Ma; (*e*) 125–180 Ma; (*f*) 175–206 Ma. In each case the bold line defines the mean $^{87}Sr/^{86}Sr$, the dashed lines define the 95% confidence interval on the mean predicted $^{87}Sr/^{86}Sr$ value (these lines merge in figure 2*a* where uncertainties are small at the scale of the figure). The background grid is drawn at intervals of 2.5 Ma and 0.0001 $^{87}Sr/^{86}Sr$.

Table 1. Data Sources for Robust LOWESS Fit and Associated Uncertainty Based on Quoted Long-term Standard Deviation for Control Materials

Data source[a]	Code[b]	Ma	$\pm 1\sigma \times 10^{-6}$
Denison et al., 1993	15	47–65	15.0
DePaolo and Ingram 1985	11	39–65	9.0
Farrell et al., 1995	10	0–7	10.5
Hess et al., 1989	1	48–68	10.0
Hodell et al., 1989, 1991	5	6–24	11.0
Hodell and Woodruff, 1994	6	11–24	11.0
Mead and Hodell, 1995	16	18–46	11.0
Jenkyns et al., 1995	13	99–121	10.5
Jones et al., 1994a, 1994b	12	100–206	11.5
Martin and Macdougall, 1991	9	61–68	11.0
McArthur et al., 1993, 1994	7	69–98	9.0
McArthur et al., submitted	7	64–67	9.0
Miller et al., 1988	2	24–35	15.0
Miller et al., 1988, 1991	2	9–25	15.0
Montanari et al., 1991	14	29–36	9.0
Oslick et al., 1994	3	10–26	10.0
Payton et al., 1993	8	10–33	11.0
Podlaha, pers. comm., 1996	17	120–162	10.0
Sugarman et al., 1995	4	65–73	10.0

[a] Not all data from each author has been included; age ranges denote data used.
[b] See figure 4.

0.709175 for MSS (Hodell and Woodruff 1994). The value of this correction is derived from the results of an interlaboratory comparison between the University of Florida and Rutgers University (Hodell and Woodruff 1994). Furthermore, compared to the data of McArthur et al. (unpub. data) and Sugarman et al. (1995), the data of Martin and Macdougall (1991) appear high, so an additional 17×10^{-6} has been subtracted from their data to correct for this apparent interlaboratory bias. We acknowledge that this correction is subjective, in that it normalizes their results, alone of the data we use, to a value of 0.710248 for NIST 987 rather than 0.709175 for MSS.

Assignment of Numeric Age

Age Models. We have used the timescales of Shackleton et al. (1994) for the period 0–7 Ma; Cande and Kent (1995) for 7–72 Ma, Obradovich (1993) for 72–95 Ma, and Gradstein et al. (1995) for ages >95 Ma. Where original sources used other timescales, numeric ages have been converted using the table in Wei (1994) or by interpolation between tie-points using the formulae of Wei (1994). For the data from DePaolo and Ingram (1985), conversion has not been possible owing to their use of non-standard timescales, so we use their given ages. Direct calibration to radiometric ages of Obradovich (1993) has been used for the data of McArthur et al. (1993, 1994) for the interval 64–95 Ma. We have used the data of Jones et al. (1994a, 1994b), Jenkyns et al. (1995) and Podlaha (pers. comm., 1996) for the interval 95–206 Ma and have converted the numeric dates from their original timescales to that of Gradstein et al. (1995). Data for samples in Jones et al. (1994a, 1994b) and Podlaha (pers. comm. 1996) with Fe concentrations >150 ppm have not been used in this study, as such concentrations may indicate that alteration has occurred.

Uncertainties in Age Estimates. The LOWESS regression assumes that the numeric ages assigned to $^{87}Sr/^{86}Sr$ data are accurate. This is clearly not the case, but how such uncertainties can be quantified presents considerable difficulty. Using chronograms, statistical estimates of uncertainty on stage boundaries have been attempted by Harland et al. (1990), Agterberg (1994), and Gradstein et al. (1995), with good agreement. At present, however, we do not believe it is feasible to assign age uncertainties to each of the *individual* datum in our database because they cannot be rigorously quantified. It seems highly desirable for the rigorous maximum-likelihood approach to chronogram construction, used by Agterberg (1994), to be extended throughout the geological timescale.

Numerous possibilities for uncertainty should be noted: e.g., inaccuracies of up to 4 myr from uncertainties in biostratigraphically based age models used to calibrate Cenozoic curves (Miller et al. 1991) and inaccuracy increases, although in a rather unpredictable way, with numeric age. Biostratigraphic age models require correct recognition of boundaries and their defining taxa, interpolation, extrapolation, high-order correlations, and assumptions about sedimentation rate, all of which introduce non-systematic and unquantifiable uncertainty into the GSSC. Further uncertainty is introduced by diachronicity (Hess et al. 1989; Miller et al. 1991; Hodell and Woodruff 1994; McArthur et al. 1994) and the restricted geographical range of taxa. Biostratigraphic data in DSDP/ODP reports are subject to all these problems, and magnetostratigraphic data has its own shortcomings, but both are essential as a basis for age assignments to $^{87}Sr/^{86}Sr$ data. Whilst in the long run it ought to be possible (given reasonable quantitative data on uncertainty in age values as the explanatory variable) to compute a GSSC taking uncertainty in both age and $^{87}Sr/^{86}Sr$ values into account, we have no option here but to follow earlier investigators in treating numeric age as free from uncertainty. This should not significantly affect the estimated mean age for a given $^{87}Sr/^{86}Sr$ value, but does imply that the confidence bounds on mean predicted ages

should be used with caution as they are smaller than would be the case were it possible to include age uncertainties in the calculation.

Calculation of the LOWESS-Smoothed GSSC

Assignment of Initial Weights. Analytical precision for published $^{87}Sr/^{86}Sr$ data, reported as one standard deviation, varies from $\pm 5\times10^{-6}$ (Montanari et al. 1991) to $\pm 15\times10^{-6}$ (Denison et al. 1993) and is usually based on replication of a laboratory standard such as NIST 987. We assigned initial weights (w_A) to each data set, on input to the LOWESS regression, to reflect this variation on the basis

$$w_A = \begin{cases} 1.0 & \text{if } \sigma_A \le 9 \times 10^{-6} \\ & \text{else } e^{-0.20(\sigma_A - \sigma_t)} \end{cases} \quad (7)$$

where σ_A is the reported analytical standard deviation of each investigator's data set and $\sigma_T = 9 \times 10^{-6}$ is that reported for recent, very well-controlled data (table 5 of Thirlwall 1991); see also McArthur (1994). These global weights are simply combined with the LOWESS neighborhood weights in the initial stage of computing the local LOWESS fit (equations 1–3), which then continues with the robust fit using weights based on the local regression residuals (equations 4–6).

Obtaining the Global LOWESS Model. Although the LOWESS fitting method is locally resistant to the effects of outliers, magnitudes of some of the residuals from the final piecewise six-subset global model proved so large that to obtain a further improvement in the overall fit, the data were globally re-weighted using a scheme

$$w_r = \begin{cases} e^{-2.5(|\lambda| - 5.0)} & \text{if } \lambda > 5.0 \\ & \text{else } 1.0 \end{cases} \quad (8)$$

where λ is the ratio of the residuals from the unified spline-interpolated LOWESS curve-fit to the half-width of its associated confidence interval. The *entire* LOWESS curve-fitting process was then repeated, using the same six subsets as before but now pre-weighted using a new residual-based weight, w_r rather than w_A. This was again followed by smoothing spline interpolation of these submodels to obtain the GSSC. As explained above, choice of the span is not too critical (cf. Marron and Tsybakov 1995). We found visually-guided choice of the span preferable to a purely algorithmic approach, which could easily lead to oversmoothing, particularly at sharp inflection points. The spans used in fitting the final submodels were 0.30, 0.15, 0.25, 0.20, 0.19, and 0.12, respectively.

The final unified LOWESS fit (figure 1*a*) contains 19 turning-points (at which the first derivative changes sign) on the curve for predicted mean $^{87}Sr/^{86}Sr$ at c. 0.13, 0.38, 0.42, 39.5, 45.5, 52.5, 64.5, 90.0, 97.4, 102.3, 104.2, 110.8, 124.2, 153.8, 156.0, 159.0, 179.0, 189.2, and 203.7 Ma), the majority of which are visible in figures 1*a* and 2.

Checking the Regression Model. The appropriateness of a regression model can be tested by inspection of the frequency distribution of the $^{87}Sr/^{86}Sr$ residuals. Figure 3*a* shows a histogram of the residuals from the final global LOWESS model. Sorting these into ascending order of magnitude allows comparison with the equivalent expected quantiles (percentiles) of the standardized normal distribution, with a mean of zero and unit standard deviation, $N(0,1)$. If a residual set conforms to a normal distribution, then a plot of its ordered values as a function of the normal quantiles would be linear. In this case, the curvature in the plot (figure 3*b*) as the tails of the distribution are approached indicates both nonnormality as well as the presence of some extremely large and small residuals. This is to be expected, since we are using a robust-fitting approach, and outliers have been downweighted in the fitting process. There is no systematic relationship (correlation) between the regression-residuals and their predicted values (figure 3*c*). These facts suggest that the LOWESS model is an appropriate one. In addition, the residuals grouped by data source (figure 4) show no major bias.

Uncertainty in the LOWESS Regression Model. Owing to random uncertainties in sample-preservational factors, subsampling, and analytical uncertainty in measured $^{87}Sr/^{86}Sr$, there is uncertainty in the estimation of the parameters of the local regression fits, and so in the overall smoothed global curve. The computed GSSC gives the best estimate of the *mean* value of $^{87}Sr/^{86}Sr$ which corresponds to a given numeric age. This relationship, once inverted, forms the basis of an age prediction based on the experimental measurement of $^{87}Sr/^{86}Sr$ in a sample.

Uncertainty in the fitted LOWESS model is given by the upper and lower $100(1 - \alpha)\%$ confidence bounds on the expected mean $^{87}Sr/^{86}Sr$ value for a given age, where α is the chosen level of risk (0.05 has been used here). Our choice therefore yields a 95% confidence interval on the predicted mean $^{87}Sr/^{86}Sr$ value. The half-width of the pointwise 95% LOWESS confidence interval for a given age is taken to be twice the LOWESS regression standard error (Hastie and Tibshirani 1990; Cham-

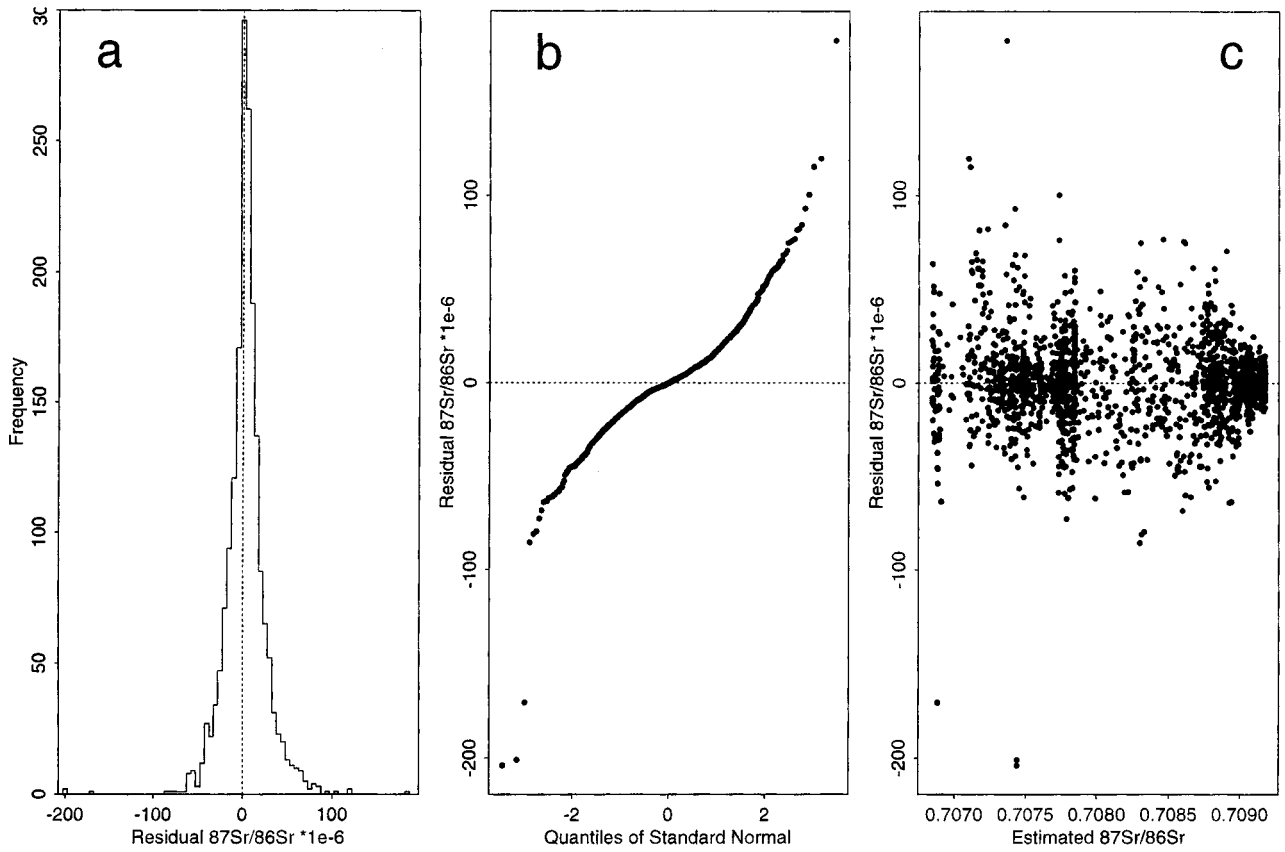

Figure 3. Diagnostic plots for residuals (i.e. observed–fitted) $^{87}Sr/^{86}Sr \times 10^6$) from the fitted values of the robust LOWESS model: (*a*) Histogram. (*b*) Plot of the ordered residuals as a function of the expected values of the standard normal distribution $N(0,1)$; plot would be linear if the residuals were themselves normally distributed. (*c*) Plot of residuals as a function of predicted $^{87}Sr/^{86}Sr$; there is no sign of correlation between the residuals and predicted values.

bers and Hastie 1992). However, the global LOWESS curve fit, and hence its associated confidence interval, is defined only at the *x*-coordinates of the data set. It therefore has to be interpolated to obtain sufficient intermediate values to yield a "continuous" prediction function and its associated confidence limits (dashed lines in figure 2*a–f*, based on intervals of 0.25 myr). These values are later inverted to enable age estimations from an observed $^{87}Sr/^{86}Sr$ best-estimate in a sample of unknown age.

Confidence and Prediction Intervals. The spread of the data used in the GSSC reflects gross sampling, subsampling and analytical errors, interlaboratory variation, etc. This is why the GSSC has been computed using a robust statistical method. A common misconception appears to be that 95% of the individual data points on which a regression model is based should fall within the 95% confidence interval of the fitted regression function. This is not the case: the *confidence interval* is the interval within which one has $100(1 - \alpha)\%$ confidence that the true (but unknown) *mean* $^{87}Sr/^{86}Sr$ corresponding to a given age-value (estimated by the fitted regression function) will lie. The *prediction interval* is the interval within which one has $100(1 - \alpha)\%$ confidence that an *individual* $^{87}Sr/^{86}Sr$ value might lie. The uncertainty on this single value is considerably broader than the uncertainty on the mean, since it is inflated because of the dispersion of the individual *y*-values about the mean, which still represents the best estimate of an individual $^{87}Sr/^{86}Sr$ value. In this application, we are principally concerned with the uncertainty in the predicted mean value of $^{87}Sr/^{86}Sr$ corresponding to a given age.

Deriving the Look-up Table. The ages in the final look-up table are derived from interpolation of the inverted LOWESS-derived curves as a function of regularly-spaced $^{87}Sr/^{86}Sr$ values. Separate inversions have been made between all 19 turning-points on each of the three curves, viz. the mean and its upper and lower 95% confidence bounds, as the locations of their turning points are not exactly

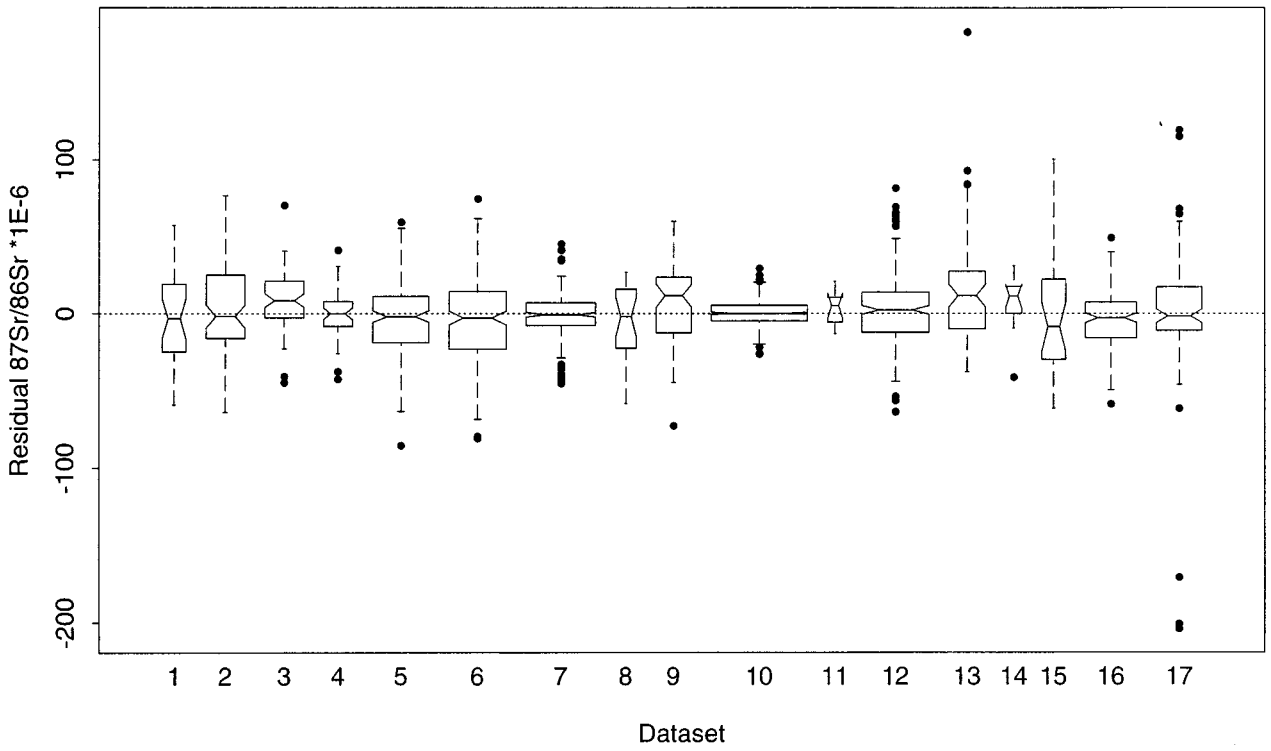

Figure 4. Boxplots of $^{87}Sr/^{86}Sr$ residuals from global robust LOWESS fit ($\times 10^{-6}$) by principal author of data source (see table 1 for author codes). Central dotted straight line represents the mean of all data. Width of the boxes is proportional to the number of data in each subset. Height of the boxes represents the inter-quartile range (IQR), i.e. the central 50% of the data; vertical dashed lines extend from boxes to most outlying data $\leq 1.5 \times$ (IQR); individual dots show outliers. Vertical width of the notches in the sides of the boxes shows approximate 95% confidence intervals on the medians; three subsets show slight upward bias.

coincidental. The example in table 2 illustrates the style of the full look-up table and is interpolated in steps of $5 \times 10-6$ $^{87}Sr/^{86}Sr$ in order not to take up undue space.

Uncertainty on Numeric Age Estimates Using the LOWESS GSSC

Contributing Factors. The uncertainty in any numeric age from the look-up table should combine uncertainties from: (i) the appropriateness of the regression model used, i.e., LOWESS as opposed to an alternative regression technique; (ii) the uncertainty in the LOWESS fit, which will depend on the uncertainty in the $^{87}Sr/^{86}Sr$ values, and in the ages assigned to the data points; and (iii) the uncertainty in the $^{87}Sr/^{86}Sr$ value of the new sample whose age is to be estimated.

The LOWESS model appears to be well behaved; in common with earlier GSSC investigators, we are unable to quantify age uncertainties used to calibrate the GSSC. We have shown how the uncertainty in the GSSC has been derived and therefore concentrate ensuing discussion on how best to determine the true uncertainty associated with the measurement of $^{87}Sr/^{86}Sr$ in a sample.

The uncertainty in $^{87}Sr/^{86}Sr$ of the sample to be dated must be added to that inherent in the LOWESS fit. The uncertainty in measured $^{87}Sr/^{86}Sr$ of the sample may be estimated in several ways. For a single $^{87}Sr/^{86}Sr$ measurement, the uncertainty may be derived from the long-term variance of the analytical system as measured by long-term repeated measurements of a standard reference material (NIST 987, E & A, or MSS) or using pooled duplicates, triplicates, etc. of different run-of-the-mill samples analyzed over an extended period (it is important to note that the best uncertainty estimate is not given by $\pm$ twice the standard deviation, as will be seen later). In contrast, where a critical sample is replicated, estimates of uncertainty may be derived from the results. We here use the term "analytical system" in the sense of Thompson and Howarth (1973, 1976) to mean all procedures from subsampling to delivery of the result, considered as an ensemble.

Variance Estimation Using Standard Reference Materials. Long-term analytical-system variance can

Table 2. Example of Look-up Table

$^{87}Sr/^{86}Sr$	Best estimate of mean age and 95% confidence intervals (Ma)																				
.707700	>72.03	72.16	<72.29	>201.48	201.68	<201.89	>205.61	205.95	ND												
.707705	>71.86	72.00	<72.13	>201.56	201.76	<201.99	>205.52	205.86	ND												
.707710	>71.69	71.83	<71.97	>201.65	201.86	<202.1	>205.41	205.76	ND												
.707715	>71.53	71.67	<71.80	>201.74	201.96	<202.21	>205.29	205.65	ND												
.707720	>51.26		<53.71	>71.36	71.51	<71.64	>201.84	202.07	<202.34	>205.16	205.53	<205.96									
.707725	>50.21		<54.23	>71.19	71.33	<71.48	>201.94	202.18	<202.46	>205.01	205.41	<205.84									
.707730	>38.68		<40.82	>49.49	52.36	<54.67	>71.00	71.16	<71.31	>202.04	202.31	<202.63	>204.83	205.27	>205.71						
.707735	>38.03		<41.42	>48.94	50.06	<51.93	>53.01	54.33	<55.05	>70.83	70.98	<71.13	>202.16	202.44	<202.81	>204.62	205.10	<205.56			
.707740	>37.60	39.66	<41.85	>48.48	49.36	<50.70	>53.89	54.77	<55.45	>70.62	70.80	<70.97	>202.27	202.60	<203.06	>204.34	204.91	<205.41			
.707745	>37.22	37.82	<39.05	>39.90	41.50	<42.25	>48.04	48.79	<49.90	>54.44	55.16	<55.82	>70.41	70.60	<70.77	>202.41	202.79	<203.75	>203.75	204.68	<205.23
.707750	>36.88	37.41	<38.08	>41.06	41.93	<42.63	>47.63	48.32	<49.23	>54.87	55.53	<56.19	>70.18	70.39	<70.58	>202.57	203.70	<205.02			
.707755	>36.53	37.04	<37.61	>41.58	42.31	<43.10	>47.20	47.88	<48.65	>55.24	55.89	<56.58	>69.92	70.15	<70.37	>202.75		<204.76			
.707760	>36.18	36.71	<37.22	>41.99	42.71	<43.83	>46.69	47.46	<48.15	>55.60	56.24	<56.99	>69.63	69.90	<70.14	>202.99		<204.40			
.707765	>35.83	36.38	<36.90	>42.39	45.12	<47.72	>55.93	56.61	<57.39	>69.30	69.60	<69.87									
.707770	>35.50	36.03	<36.56	>42.80	45.31	<47.28	>56.28	57.00	<57.85	>68.93	69.28	<69.57									
.707775	>35.24	35.70	<36.24	>43.38		<46.79	>56.64	57.45	<58.36	>68.56	68.91	<69.24									
.707780	>35.01	35.43	<35.91	>45.25		<45.60	>57.04	57.90	<59.01	>68.18	68.54	<68.87									
.707785	>34.82	35.18	<35.62	>57.47	58.45	<59.72	>67.85	68.19	<68.52												
.707790	>34.66	34.97	<35.35	>57.97	59.05	<60.34	>67.52	67.86	<68.19												
.707795	>34.51	34.79	<35.12	>58.49	59.69	<60.90	>67.23	67.55	<67.88												
.707800	>34.37	34.63	<34.93	>59.07	60.36	<61.34	>66.96	67.27	<67.58												

Notes. (1) Where estimated mean age is blank, $^{87}Sr/^{86}Sr$ corresponds to confidence region only; ND, not determined.
(2) A copy of the full table (interpolated in steps of 0.000001 $^{87}Sr/^{86}Sr$) may be obtained, either on a 3.5″ disk or via e-mail, by contacting the second author at j.mcarthur@ucl.ac.uk.

be calculated from replicate determinations of $^{87}Sr/^{86}Sr$ in standard reference materials. In theory, it provides a false view of the uncertainty in $^{87}Sr/^{86}Sr$ measurement in real samples, since standards and samples may differ in Sr concentration and matrix composition; matrix differences may cause mass fractionation during mass spectrometry if Sr is not cleanly separated by ion-exchange chromatography. $^{87}Sr/^{86}Sr$ in a real sample will, unlike a reference material, differ between subsamples. In practice, however, the reproducibility of $^{87}Sr/^{86}Sr$ determinations on standards and samples are close (Hodell et al. 1990) but not identical. Never the less, estimates of variance based on replicated long-term measurements of samples is sounder.

Variance Estimation Using Routine Samples. Subsamples are best analyzed in random order, and replicate determinations are made on a proportion of them. Typically, two to five replicate analyses should be made per sample.

In evaluating the variance of the analytical system, the first step is to collate the results for all samples of a similar matrix type for which replicate determinations (N_r) exist (e.g., results for biogenic calcite should be treated separately from those for biogenic phosphate). The range, R, of each set of replicate determinations for individual samples can be thought of as

$$R = \max \{x_i - \bar{x}\} - \min \{x_i - \bar{x}\} \qquad (9)$$

where x_1 are the N_r individual $^{87}Sr/^{86}Sr$ determinations and $\bar{x}$ = mean ($^{87}Sr/^{86}Sr$). A check for constancy of the analytical system variance is made by plotting R as a function of $\bar{x}$.

Figure 5 show such a plot based on routine replicate $^{87}Sr/^{86}Sr$ determinations at Royal Holloway and Bedford New College, Surrey, England (RHBNC), made in the course of analysis of 104 routine samples of biogenic carbonate, with N_r equal to either two determinations, or three in 33% of the samples. The slope of the fitted least-squares regression line of R on mean ($^{87}Sr/^{86}Sr$) (dashed line in figure 5), is not significantly different from zero. A constant-variance statistical model is therefore acceptable. If this is not the case, special methods are required to estimate the analytical system standard deviation as a function of the mean. (See Thompson and Howarth [1973, 1976] and Howarth and Thompson [1976] for a method applicable to duplicate analyses of routine samples.)

Samples with the same N_r should now be grouped together and, assuming that the errors inherent in the analytical system conform to the normal distribution (an assumption that ought to be checked), the standard deviation (s) can either be estimated directly or from the average range (ASQC Statistics Division 1983). The standard deviation estimates based on each set of replicates should then be combined to give the *pooled* estimate of the long-term analytical system standard deviation ($\hat{\sigma}_A$) using a weighted average of the variances:

$$\hat{\sigma}_A = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2 + n_3 s_3^2 + \cdots}{n_1 + n_2 + n_3 + \cdots}} \qquad (10)$$

Table 3 gives the mean range and estimated standard deviation for the two sets of biogenic calcite

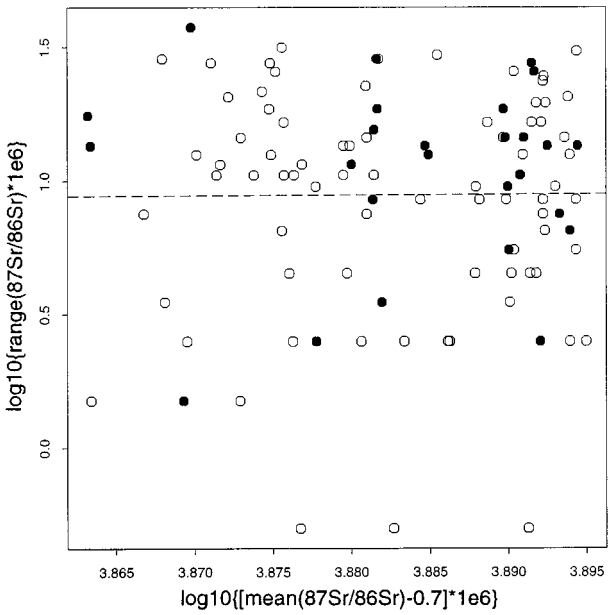

Figure 5. Log-log plot of range as a function of mean $^{87}Sr/^{86}Sr$ ($\times 10^6$) for biogenic calcite in 104 biogenic carbonate samples for which replicate determinations are available; open circles, $N_r = 2$; solid circles, $N_r = 3$. Fitted linear regression of range on mean (dashed line) is virtually horizontal, indicating complete absence of systematic dependency of range on the mean.

Table 3. Computation of Analytical System Variance for 104 Biogenic Calcite Samples Analyzed with Replicate Determinations of $^{87}Sr/^{86}Sr$

No. of. replicates (N_r)	No. of replicated samples (n)	Mean range ($\bar{R} \times 10^{-6}$)	c (ASQC 1983)	Estimated std. dev. ($[\bar{R}/c] \times 10^{-6}$)
2	78	11.3	1.128	10.05
3	26	13.1	1.693	7.75

samples with different values of N_r plotted in figure 5. The pooled estimate of the analytical system standard deviation for these data is therefore:

$$\hat{\sigma}_A = [\sqrt{(78 \times 10.05^2 + 26 \times 7.75^2)/(78 + 26)}] \times 10^{-6}$$

$$= 9.53 \times 10^{-6}$$

If the interval over which the replicate data has been collected represents instrumental usage under varying conditions of sample preparation and instrumental measurement (different operators, etc.), then this may be a worst-case estimate.

Confidence Interval on Mean $^{87}Sr/^{86}Sr$ in the Sample

Single Measurements In the case of a *single* determination of $^{87}Sr/^{86}Sr(x)$, the inherent combined sampling and analytical variability is unknown, and one must assume that the magnitude of the analytical system variance is as it has been in the past. The best-estimate of the true (but unknown) mean $^{87}Sr/^{86}Sr$ of the sample is the single observed value x, and the prior estimated value of the standard deviation $\hat{\sigma}_A$ (equation 10) is assumed to apply. A two-sided $100(1 - \alpha)\%$ confidence interval on this estimate, $\{\mu_l, \mu_u\}$, is obtained by inverting the equivalent prediction interval for a future single observation from a normal distribution. This yields:

$$\{\mu_l, \mu_u\} = x \pm z_{1-\alpha/2}\left(1 + \frac{1}{n}\right)^{0.5} \hat{\sigma}_A \qquad (11)$$

where μ_l and μ_u are the lower and upper $100(1 - \alpha)\%$ confidence limits; n is the number of sets of replicates the prior estimate $\hat{\sigma}_A$ is based; and $z_{1-\alpha/2} = 1.96$ is the $100(1 - \alpha/2)$th percentile of the standardized normal distribution, $N(0,1)$, corresponding to a probability $\alpha = 0.05$, i.e. $\{\mu_l, \mu_u\}$ is a two-sided 95% confidence interval.

If one has no information on the long-term performance of the analytical system based on comparable routine samples, then the only recourse is to use an estimate based on a reference material such as NIST 987. Although the analytical system performance may be reasonably stable with time (figure 6), because of differences of sampling, subsampling, and preparation technique, and long-term variation in instrumental operating conditions, an estimate $\hat{\sigma}_A$ based on a reference material cannot be truly applicable to short-term analysis of routine samples. For these reasons, it is preferable that the $\hat{\sigma}_A$ used to establish confidence bounds is based on long-term replicate analyses of similar run-of-the-mill samples rather than a reference material.

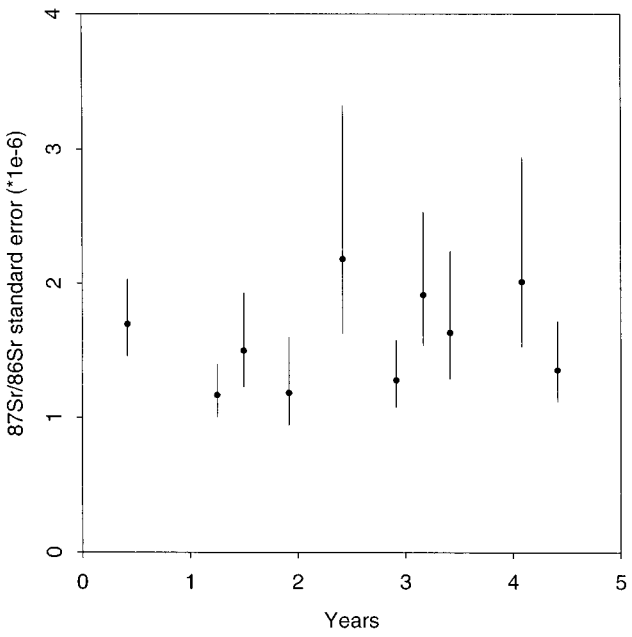

Figure 6. Plot of standard error ($\hat{\sigma}_s/\sqrt{n}$) and 95% confidence intervals of $^{87}Sr/^{86}Sr$ ($\times 10^{-6}$) in long-term replicate analyses of NIST 987 in a single laboratory over 4.5 years. Number of replicates in each sub-period (n) varies between 15 and 73. Calculated from data in table 5 of Thirlwall (1991), with addition of 95% confidence intervals after Hahn and Meeker (1991).

For example, if $\hat{\sigma}_A$ is taken to be 9.53×10^{-6} as estimated from the 104 biogenic carbonate samples (above), the two-sided 95% confidence interval on the mean $^{87}Sr/^{86}Sr$ in a new sample, estimated by a *single* measurement of $^{87}Sr/^{86}Sr$ (x), becomes (equation 11):

$$\{\mu_l, \mu_u\} = x \pm 1.96\left(1 + \frac{1}{104}\right)^{0.5} (9.53 \times 10^{-6})$$

$$= x \pm 18.77 \times 10^{-6} \approx x \pm 19 \times 10^{-6}$$

In comparison, the value of $\hat{\sigma}_A$ based on the most recent medium-term estimate using NIST 987 in the same laboratory (44 measurements over a 6 month period: Thirlwall 1991) is 9.0×10^{-6}, yielding a 95% confidence interval of $x \pm 17.84 \times 10^{-6} \approx 18 \times 10^{-6}$.

The bias incurred by an estimate of $\hat{\sigma}_A$ based on a reference material rather than on comparable routine samples could easily become considerably larger.

Two or More Replicated Measurements. To reduce the uncertainty on the $^{87}Sr/^{86}Sr$ estimate. N_r replicate determinations yield an observed mean $\bar{x}$. If, as before, we assume that $\hat{\sigma}_A$ is a prior estimate of the standard deviation applicable to the analytical system, then the two-sided 100(1-α)% confidence interval for the true (but unknown) mean is given by:

$$\{\mu_l, \mu_u\} = \bar{x} \pm z_{1-\alpha/2} (\hat{\sigma}_A/\sqrt{N_r}) \quad (12)$$

Taking $z_{1-\alpha/2} = 1.96$, as before, again yields a two-sided 95% confidence interval on the mean. The ratio $(\hat{\sigma}_A/\sqrt{N_r})$ is often referred to as the standard error of the mean.

Suppose we had measured the $^{87}Sr/^{86}Sr$ value of five subsamples and obtained the results: $^{87}Sr/^{86}Sr = 0.7077 + \{25, 23, 17, 31, 19\} \times 10^{-6}$. Their mean is 0.707723. If, as before, we assume that $\hat{\sigma}_A$ is equal to 9.53×10^{-6} based on the results of long-term replication of similar biogenic carbonate samples, then the 95% confidence bounds on the mean $^{87}Sr/^{86}Sr$ value are given by:

$$\{\mu_l, \mu_u\} = 0.7077 + \{23 \pm 1.96(9.53/\sqrt{5})\} \times 10^{-6}$$

$$= 0.707723 \pm 8.3 \times 10^{-6}$$

$$= \{0.707715, 0.707731\}$$

An alternative is to base the confidence internal on the standard deviation $(\hat{\sigma}_S)$ estimated from the $^{87}Sr/^{86}Sr$ values of the current subsample set, as opposed to using a prior estimate, $\hat{\sigma}_A$. In this case, we risk assuming that a confidence interval solely based on these analyses adequately reflects the true variance of the analytical system (contributed to by sampling/subsampling variation, current instrumental operating conditions, and human factors). Since $\hat{\sigma}_S$ is calculated using the mean $^{87}Sr/^{86}Sr$ estimate $(\bar{x})$ from the *same* suite of measurements, the *t*-distribution must be used to calculate the two-sided confidence bounds and not the normal distribution. The confidence bounds are consequently given by:

$$\{\mu_l, \mu_u\} = \bar{x} \pm t_{(1-\alpha/2),\nu} (\hat{\sigma}_s/\sqrt{N_r}) \quad (13)$$

where $t_{(1-\alpha/2),\mu}$ is the $100(1 - \alpha/2)$th percentile of the *t*-distribution corresponding to a probability $\alpha = 0.05$ and $\nu = N_r - 1$ degrees of freedom (df; this is one less than N_r because 1 df has already been used in the calculation of $\bar{x}$ from which $\hat{\sigma}_S$ is subsequently obtained). The *t*-distribution has longer tails than $N(0,1)$, so that even if $\hat{\sigma}_S = \hat{\sigma}_A$, the confidence interval in equation 13 is wider than that of equation 12.

Taking the same data as in the previous example, the sample standard deviation is 5.48×10^{-6} and $t_{0.975,4} = 2.776$; hence (equation 13) the 95% confidence bounds become:

$$\{\mu_l, \mu_u\} = 0.707723 \pm 2.776(5.48/\sqrt{5}) \times 10^{-6}$$

$$= 0.707723 \pm 6.80 \times 10^{-6}$$

$$= \{0.707716, 0.707730\}$$

Suppose, however, that $^{87}Sr/^{86}Sr$ had been determined on only the first three subsamples, with the results: $0.7077 + \{25, 23, 17\} \times 10^{-6}$. The mean would now be 0.707722 and the standard deviation 4.16×10^{-6}. Because the number of determinations is smaller, $t_{0.975.4} = 4.303$ and hence:

$$\{\mu_l, \mu_u\} = 0.707722 \pm 4.303(4.16/\sqrt{3}) \times 10^{-6}$$

$$= 0.707722 \pm 10.34 \times 10^{-6}$$

$$= \{0.707712, 0.707732\}$$

Thus, unless a large number of replicate determinations are made, an estimate of the standard deviation based on the long-term behavior of replicate analyses of routine samples can yield a confidence-interval half-width smaller than that based simply on replicate determinations of a current sample and considerably smaller than if only a single $^{87}Sr/^{86}Sr$ determination is made.

Look-up Table

The full table, interpolated in steps of 1×10^{-6} $^{87}Sr/^{86}Sr$, may be obtained on disk from the authors (see table 2, note, for details). An accompanying document describes recommended procedures for (i) determination of mean $^{87}Sr/^{86}Sr$ in the sample(s) to be dated; and (ii) obtaining an overall best-estimate of age and its accompanying 95% confidence interval from the mean $^{87}Sr/^{86}Sr$ values of a set of independent samples and/or subsamples: together with worked examples.

ACKNOWLEDGEMENTS

We thank O.G. Podlaha (Ruhr University, Bochum, Germany) for providing data for Jurassic and Creta-

ceous samples; the referees, Jack Schuenemeyer (University of Delaware) and John Compton (University of Cape Town); Managing Editor Barbara Sivertsen and Co-Editor Robert Newton; and Wuchang Wei (Scripps Institution of Oceanography) for their thoughtful comments on the manuscript.

REFERENCES CITED

Agterberg, F. P., 1994, Estimation of the Mesozoic Geological Time Scale: Math. Geol., v. 26, p. 857–876.

Analytical Methods Committee. 1987, Recommendations for the Conduct and Interpretation of Co-operative trials: Analyst, v. 112, p. 679–686.

———, 1995, Internal Quality Control of Analytical Data: Analyst, v. 120, p. 29–34.

ASQC Statistics Division, 1983, Glossary and Tables for Statistical Quality Control: Milwaukee, Wisconsin, American Society for Quality Control, 160 p.

Cande, S. C., and Kent, D. V., 1995, Revised calibration of the geomagnetic polarity timescale for the Late Cretaceous and Cenozoic: Jour. Geophys. Res., v. 100, p. 6093–6095.

Chambers, J. M.; Cleveland, W. S.; Kleiner, B.; and Tukey, P. A., 1983, Graphical Methods for Data Analysis: Belmont, CA, Wadsworth, 395 p.

———, and Hastie, T. I. 1992, Statistical Models in S, Pacific Grove, California; Belmont, CA, Wadsworth and Brooks/Cole, 608 p.

Cleveland, W. S., 1979, Robust Locally Weighted Regression and Smoothing Scatterplots: Jour. Am. Statist. Assoc., v. 74, p. 829–836.

———; Grosse, E.; and Shyu, W. M., 1992, Local regression models, *in* Chambers, J. M., and Hastie, T., eds., Statistical Models in S, Pacific Grove, California: Belmont, CA, Wadsworth and Brooks/Cole, p. 309–376.

Denison, R. E.; Koepnick, R. B.; Fletcher, A.; Dahl, D. A.; and Baker, M. C., 1993, Reevaluation of Early Oligocene, Eocene, and Paleocene seawater strontium isotope ratios using outcrop samples from the U.S. Gulf Coast: Paleoceanography, v. 8, p. 101–126.

DePaolo, D. J., and Ingram, B., 1985, High-resolution stratigraphy with strontium isotopes: Science, v. 227, p. 938–941.

Draper, N. R., and Smith, H., 1981, Applied Regression Analysis: New York, Wiley, 709 p.

Elderfield, H.; 1986, Strontium isotope stratigraphy: Palaeogeog. Palaeoclimatol. Palaeoecol., v. 57, p. 71–90.

Farrell, J. W.; Clemens, S. C.; and Gromet, L. P., 1995, Improved chronostratigraphic reference curve of late Neogene seawater $^{87}Sr/^{86}Sr$: Geology, v. 23, p. 403–406.

Gradstein, F. M.; Agterberg, F. P.; Ogg, J. G.; Hardenbol, J.; Van Veen, P.; and Huang, Z., 1995, A Mesozoic Time Scale: Jour. Geophys. Res., v. 99, p. 24,051–24,074.

Hahn, G. J., and Meeker, W. Q., 1991, Statistical Intervals. A Guide for Practitioners: New York, Wiley, 392 p.

Harland, W. B.; Armstrong, R. L.; Cox, A. V.; Craig, L. E.; Smith, A. G.; and Smith, D. G., 1990, A Geologic Time Scale: Cambridge, Cambridge University Press, 263 p.

Hastie, T. J., 1992, Generalized Additive Models, *in* Chambers, J. M., and Hastie, T. J., eds., Statistical Models in S, Pacific Grove, California; Belmont, CA, Wadsworth and Brooks/Cole, p. 249–307.

———, and Tibshirani, R. J., 1990, Generalized Additive Models: London, Chapman and Hall, 335 p.

Hess, J.; Stott, L. D.; Bender, M. L.; Kennet, J. P.; and Schilling, J.-G., 1989, The Oligocene marine microfossil record: Age assessments using strontium isotopes: Paleoceanography, v. 4, p. 655–679.

Hodell, D. A.; Mead, G. A.; and Mueller, P. A., 1990, Variation in the strontium isotopic composition of seawater (8 Ma to present): Implications for chemical weathering rates and dissolved fluxes to the oceans: Chem. Geol. (Isotope Geosci. Sect.), v. 80, p. 291–307.

———; Mueller, P. A.; and Garrido, J. R., 1991, Variations in the strontium isotopic composition of seawater during the Neogene: Geology, v. 19, p. 24–27.

———; ———; McKenzie, J. A.; and Mead, G. A., 1989, Strontium isotope stratigraphy and geochemistry of the late Neogene ocean: Earth Planet. Sci. Lett., v. 92, p. 165–178.

———, and Woodruff, F., 1994, Variations in the strontium isotopic ratio of seawater during the Miocene: stratigraphic and geochemical implications: Paleoceanography, v. 9, p. 405–426.

Howarth, R. J., and Thompson, M., 1976, Duplicate analysis in geochemical practice. Part II. Examination of the proposed method and examples of its use: Analyst, v. 101, p. 699–709.

Huber, P. J., 1981, Robust Statistics: New York, Wiley, 308 p.

Jenkyns, H. C., Paull, K.; Cummins, D. I.; and Fullagar, P. D., 1995, Strontium-isotope stratigraphy of Lower Cretaceous atoll carbonates in the mid-Pacific Mountains, *in* Proc. Ocean Drilling Program, Scientific Results, College Station, Texas, p. 89–97.

Jones, C. E.; Jenkyns, H. C.; Coe, A. L.; and Hesselbo, S. P., 1994*a*, Strontium isotope variations Jurassic and Cretaceous seawater: Geochim. Cosmochim. Acta, v. 58, p. 3061–3074.

———; ———; and Hesselblo, S. P., 1994b, Strontium isotopes in Early Jurassic seawater: Geochim. Cosmochim. Acta, v. 58, p. 1285–1301.

Ludwig, K. R., 1987, ISOPLOT VERSION 2, a plotting and regression program for isotope geochemists for use with the HP Series 200/300 computers: U.S. Geol. Survey, Open File Rept., 49 p.

———, 1990, ISOPLOT—a plotting and regression program for radiogenic-isotope data, for IBM-PC compatible computers: U.S. Geol. Survey, Open File Rept. 44 p.

———; Halley, R. B.; Simmons, K. R.; and Peterman, Z. E., 1988, Strontium-isotope stratigraphy of Enewetak Atoll: Geology, v. 16, p. 173–177.

MacLeod, K. G., and Huber, B. T., 1996, Strontium isotopic evidence for extensive reworking in sediments spanning the Cretaceous-Tertiary boundary at ODP Site 738: Geology, v. 24, p. 463–466.

Mandel, J., 1991, The validation of measurement through interlaboratory studies: Chemomet. Intell. Lab. Syst., v. 11, p. 109–119.

Marron, J. S., and Tsybakov, A. B., 1995, Visual error criteria for qualitative smoothing: Jour. Am. Statist. Assoc., v. 90, p. 499–507.

Martin, E. E., and Macdougall, J. D., 1991, Seawater Sr isotopes at the Cretaceous/Tertiary boundary: Earth Planet. Sci. Lett., v. 104, p. 166–180.

McArthur, J. M., 1994, Recent trends in strontium isotope stratigraphy: Terra Nova, v. 6, p. 331–358.

———; Chen, M.; Gale, A. S.; Thirlwall, M. F.; and Kennedy, W. J., 1993, Strontium isotope stratigraphy for the Late Cretaceous: Age models and intercontinental correlations for the Campanian: Paleoceanography, v. 8, p. 859–873.

———; Kennedy, W. J.; Chen, M.; Thirlwall, M. F.; and Gale, A. S., 1994, Strontium isotope stratigraphy for the Late Cretaceous: Direct numerical age calibration of the Sr-isotope curve for the U.S. Western Interior Seaway: Palaeogeog. Palaeoclimatol. Palaeocol., v. 108, p. 95–119.

———; Sahami, A. R.; Thirlwall, M.; Hamilton, P. J.; and Osborn, A. O., 1990, Dating phosphogenesis with strontium isotopes: Geochim. Cosmochim. Acta, v. 54, p. 1343–1351.

McKenzie, J. A.; Hodell, D. A.; Mueller, P. A.; and Muller, D. W., 1988, Application of strontium isotopes to late Miocene–early Pliocene stratigraphy: Geology, v. 16, p. 1022–1025.

Mc Laughlin, O. M.; McArthur, J. M.; Thirlwall, M. F.; Howarth, R. J.; Burnett, J.; Gale, A. S.; and Kennedy, W. J., 1995, Sr isotope evolution of Maastrichtian seawater determined from the Chalk of Hemmoor, Germany: Terra Nova, v. 7, p. 491–499.

Mead, G. A., and Hodell, D. A., 1995, Controls on the $^{87}Sr/^{86}Sr$ composition of seawater from the middle Eocene to Oligocene: Hole 689B, Maud Rise, Antarctica: Paleogeography, v. 10, p. 327–346.

Microsoft Corporation, 1992, Microsoft EXCEL Spreadsheet with Business Graphics and Database. Function Reference: Redmond, WA, Microsoft Corporation, 536 p.

Miller, K. G.; Feigenson, M. D.; Kent, D. V.; and Olson, R. K., 1988, Upper Eocene to Oligocene isotope ($^{87}Sr/^{86}Sr$, ^{118}O, ^{113}C) standard section, Deep Sea Drilling Project Site 522: Paleoceanography, v. 3, p. 223–233.

———; ———; Wright, J. D.; and Clement, B. M., 1991, Miocene isotope reference section. Deep Sea Drilling Project Site 608: An evaluation of isotope and biostratigraphic resolution: Paleoceanography, v. 6, p. 33–52.

Montanari, A.; Deino, A.; Coccioni, R.; Langenheim, V. E.; Capo, R.; and Monechi, S., 1991, Sr isotope analysis, magnetostratigraphy and plankton stratigraphy across the Oligocene-Miocene boundary in the Contessa section (Gubbio, Italy): Newsl. Stratigr., v. 23, p. 151–180.

Obradovich, J. D., 1993, A Cretaceous timescale, *in* Caldwell, W. G. E., ed., Evolution of the western interior foreland basin: Geol. Assoc. Canada Spec. Paper, p. 379–396.

Oslick, J. S.; Miller, K. G.; Feigenson, M. D.; and Wright, J. D., 1994, Oligocene-Miocene strontium isotopes: Stratigraphic revisions and correlations to an inferred glacioeustatic record: Paleoceanography, v. 9, p. 427–443.

Paytan, A.; Kastner, M.; Martin, E. E.; Macdougal, J. D.; and Herbert, T., 1993, Marine barite as a monitor of seawater strontium isotope composition: Nature, v. 366, p. 445–449.

Pierson, R. H., and Fay, E. A., 1959, Guidelines for Interlaboratory Testing Programs [reprinted from Analytical Chemistry, v. 31] in Division, A. C., ed., Interlaboratory Testing Techniques: Milwaukee, WI, Am. Soc. Quality Control, 1978, p. 26–34.

Schmitz, B.; Aberg, G.; Werdelin, L.; Forey, P.; and Bendix-Almgreen, S. E., 1991, $^{87}Sr/^{86}Sr$, Na, F, Sr and La in skeletal fish debris as a measure of the paleosalinity of fossil-fish habitats: Geol. Soc. America Bull., v. 103, p. 786–794.

Shackleton, N. G.; Crowhurst, S.; Hagelburg, T.; Pisias, N.; and Schneider, D., 1994, A new late Neogene time scale: Application to Leg 138 Sites, *in* Proc. Ocean Drilling Program, Scientific Results, College Station, Texas, p. 73–101.

Smalley, P. C.; Higgins, A. C.; Howarth, R. J.; Nicholson, H.; Jones, C. E.; Swinburne N. H. M.; Bessa, J., 1994, Marine Strontium Isotopes: A Phanerozoic seawater curve for practical sediment dating and correlation: Geology, v. 22, p. 431–434.

Sugarman, P. J.; Miller, K. G.; Bukry, D.; and Feigenson, M. D., 1995, Uppermost Campanian-Maastrichtian strontium isotopic, biostratigraphic, and sequence stratigraphic framework of the new Jersey Coastal Plain: Geol. Soc. America Bull., v. 107, p. 19–37.

Thirlwall, M. F., 1991, Long-term reproducibility of multicollector Sr and Nd isotope ratio analysis: Chem. Geol. (Isotope Geosci. Sect.), v. 94, p. 85–104.

Thisted, R. A., 1988, Elements of Statistical Computing: New York, Chapman and Hall, 427 p.

Thompson, M., and Howarth, R. J., 1973, The Rapid Estimation and Control of Precision by Duplicate Determinations: Analyst, v. 98, p. 153–160.

———, and ———, 1976, Duplicate Analysis in Geochemical Practice. Part I. Theoretical Approach and Estimation of Analytical Reproducibility: Analyst, v. 101, p. 690–698.

Veizer, J., 1989, Strontium isotopes in seawater through time: Ann. Rev. Earth Planet. Sci., v. 17, p. 141–167.

Wahba, G., 1975, Smoothing Noisy Data with Spline Functions: Numer. Math., v. 24, p. 383–393.

Wei, W., 1994, Age conversion table for different time scales: Jour. Nannoplankton Res., v. 16, p. 71–73.