



# Errorchrons and anchored isochrons in IsoplotR

Pieter Vermeesch

Department of Earth Sciences, University College London, Gower Street, London WC1E 6BT, United Kingdom

**Correspondence:** Pieter Vermeesch (p.vermeesch@ucl.ac.uk)

Received: 31 January 2024 – Discussion started: 14 February 2024

Revised: 13 May 2024 – Accepted: 31 May 2024 – Published: 18 July 2024

**Abstract.** Isochrons are usually fitted by “York regression”, which uses a weighted least-squares approach that accounts for correlated uncertainties in both variables. Despite its tremendous popularity in modern geochronology, the York algorithm has two important limitations that reduce its utility in several applications. First, it does not provide a satisfactory mechanism to deal with so-called “errorchrons”, i.e. datasets that are overdispersed with respect to the analytical uncertainties. Second, York regression is not readily amenable to anchoring, in which either the slope or the intercept of the isochron is fixed based on some external information. Anchored isochrons can be very useful in cases where the data are insufficiently spread out to constrain both the radiogenic and non-radiogenic isotopic composition.

This paper addresses both of these issues by extending a maximum likelihood algorithm that was first proposed by Titterton and Halliday (1979). The new algorithm offers the ability to attribute any excess dispersion to either the inherited component (“model 3a”) or diachronous closure of the isotopic system (“model 3b”). It provides an opportunity to anchor isochrons to either a fixed non-radiogenic composition or a fixed age. Last but not least, it allows the user to attach meaningful analytical uncertainty to the anchor. The new method has been implemented in IsoplotR for immediate use in Ar/Ar, Pb/Pb, U/Pb, Th/Pb, Rb/Sr, Sm/Nd, Lu/Hf, Re/Os, K/Ca, and U–Th–He geochronology.

## 1 Introduction

Isochrons are mixing lines between radiogenic and inherited isotopic endmembers. They are an essential component of radiometric geochronology and exist in several forms. Sections 1–7 of this paper will deal with the simple case of “ $P/D$  isochrons”, which apply to geochronometers that are

based on the decay of a single radioactive parent nuclide ( $P$ ; e.g.  $^{87}\text{Rb}$ ,  $^{40}\text{K}$ ,  $^{147}\text{Sm}$ ) to a particular daughter nuclide ( $D$ ; e.g.  $^{87}\text{Sr}$ ,  $^{40}\text{Ar}$ ,  $^{143}\text{Nd}$ ). Sections 8 and 9 will discuss Pb/Pb and U/Pb isochrons, respectively. These are based on the paired decay of two parents ( $^{238}\text{U}$  and  $^{235}\text{U}$ ) to two daughters ( $^{206}\text{Pb}$  and  $^{207}\text{Pb}$ , respectively) and require linear regression in two or three dimensions.

$P/D$  isochrons come as two types. “Conventional” isochrons are straight-line relationships of the following kind:

$$\left[\frac{D}{d}\right]_i = \left[\frac{D}{d}\right]_0 + \left[\frac{P}{d}\right]_i (e^{\lambda t} - 1), \quad (1)$$

where  $i$  is the aliquot number ( $1 \leq i \leq n$ ),  $d$  is a non-radiogenic sister isotope of  $D$ ,  $[D/d]_0$  is the inherited component,  $\lambda$  is the decay constant, and  $t$  is the time elapsed since isotopic closure. A full list of definitions is provided in Appendix A. “Inverse” isochrons are obtained by permuting  $P$ ,  $D$ , and  $d$ :

$$\left[\frac{d}{D}\right]_i = \left[\frac{d}{D}\right]_0 \left\{ 1 - \left[\frac{P}{D}\right]_i (e^{\lambda t} - 1) \right\}. \quad (2)$$

The choice between conventional and inverse isochrons depends on the relative precision of the mass spectrometer measurements. Inverse isochrons are preferred if  $d \ll D$ . Otherwise, conventional isochrons are fine (Li and Vermeesch, 2021).

Equations (1) and (2) can be recast into a generic linear form:

$$y_i = a + bx_i. \quad (3)$$

Table A1 maps the parameters of this generic equation to the parameters of Eqs. (1) and (2) as well as subsequent variants thereof. Equation (3) is usually solved by York et al.

(2004) regression (hereafter simply referred to as “York regression”). York regression uses a least-squares algorithm to estimate the intercept ( $a$ ) and slope ( $b$ ) of the isochron line from an  $n \times 5$  table of paired isotopic ratio measurements, along with their standard errors and their error correlations.

Although this paper will use the York parameters  $a$  and  $b$  to fit the isochron, all results will be presented in terms of the geologically more meaningful inherited endmember ratio  $[D/d]_0$  and radiogenic endmember ratio  $[D/P]^* \equiv e^{\lambda t} - 1$ , where  $[D/d]_0 = a$  and  $[D/P]^* = b$  for conventional isochrons and  $[D/d]_0 = 1/a$  and  $[D/P]^* = -b/a$  for inverse isochrons (Li and Vermeesch, 2021).

The most accurate and precise results are obtained from samples that are evenly spread along the isochron line, spanning the entire range of values from the inherited to the radiogenic endmember. Unfortunately, this condition is not always fulfilled. It is not uncommon for most or all aliquots in a sample to cluster together at one point along the isochron, making it difficult to accurately estimate the endmember compositions. For example, no precise isochron ages can be obtained from samples whose radiogenic daughter component is dwarfed by the inherited daughter component. Conversely, the composition of the inherited component cannot be precisely estimated in extremely radiogenic samples. Finally, when the data cluster together in the middle, then neither endmember component can be reliably determined (Fig. 1).

Sometimes, such poorly constrained isochrons can be fixed using external information. For example, if the composition of the inherited component is known through some independent means (e.g. by analysing a cogenetic mineral that is naturally poor in  $P$  and rich in  $D$  and  $d$ ), then this information can be used to anchor the isochron. Anchoring reduces the number of unknown parameters by 1, benefitting numerical stability and precision. Conversely, if the age of the sample is known, then the inherited component can be estimated by anchoring to the radiogenic endmember.

There is currently no formally documented way to anchor York regression. A commonly used “hack” is to add an extra data point with infinite precision representing either the inherited or radiogenic endmember component. However, this hack does not provide a satisfactory mechanism to assign uncertainty to the anchor. This paper solves that problem. Section 2 introduces a maximum likelihood formulation of York regression, which is amenable to anchoring with uncertainty. Section 3 shows how this formulation can be used to model “errorchrons” that are overdispersed with respect to the analytical uncertainties.

Section 4 provides further details about the implementation of maximum likelihood regression, which can attribute overdispersion to either the intercept (“model 3a”) or slope (“model 3b”) of the linear fit. It turns out that model 3b is computationally more challenging than model 3a. Section 5 shows how this issue can be avoided by inverting and flipping the isochron axes.

Section 6 shows how anchored isochron regression represents a trivial special case of the maximum likelihood algorithm. Section 7 attaches two different geological interpretations to the uncertainty of isochron anchors. These two interpretations lead to two different fit models (“model 1” and “model 3”, which are so named for historical reasons). Sections 8 and 9 apply the same logic to Pb/Pb and U/Pb isochrons, respectively. Finally, Sect. 10 shows how the new algorithms can be used within the IsoplotR software package.

## 2 Maximum likelihood formulation of York regression

York et al. (2004) use the method of least squares to fit the general problem of weighted regression with correlated uncertainties in both variables. Titterton and Halliday (1979) obtained identical results using the method of maximum likelihood. The latter approach offers greater flexibility than least squares and will be used in the remainder of this paper.

The method of maximum likelihood is a standard statistical technique to estimate the parameters of a probability distribution from a set of measurements. In the case of two-dimensional isochron regression, the parameters are the intercept  $a$ , the slope  $b$ , and the true (but unknown)  $x_i$  values. The data are the  $X_i$  and  $Y_i$  measurements along with their assumed uncertainties and error correlations (see Appendix A for a full list of definitions). For a dataset of  $n$  aliquots, there are  $2n$  measurements ( $n$   $X_i$  values and  $n$   $Y_i$  values) and  $n + 2$  parameters ( $a$ ,  $b$ , and  $n$   $x_i$  values). That leaves  $2n - n - 2 = n - 2$  “degrees of freedom” to estimate the parameters.

This overconstrained problem can be solved by minimizing the paired differences (“residuals”) between the true  $x_i$  and  $y_i$  values (where  $y_i = a + bx_i$ ) and between the measured  $X_i$  and  $Y_i$  values.

$$\begin{cases} X_i = x_i + \epsilon_{x,i} \\ Y_i = a + bx_i + \epsilon_{y,i} \end{cases} \quad (4)$$

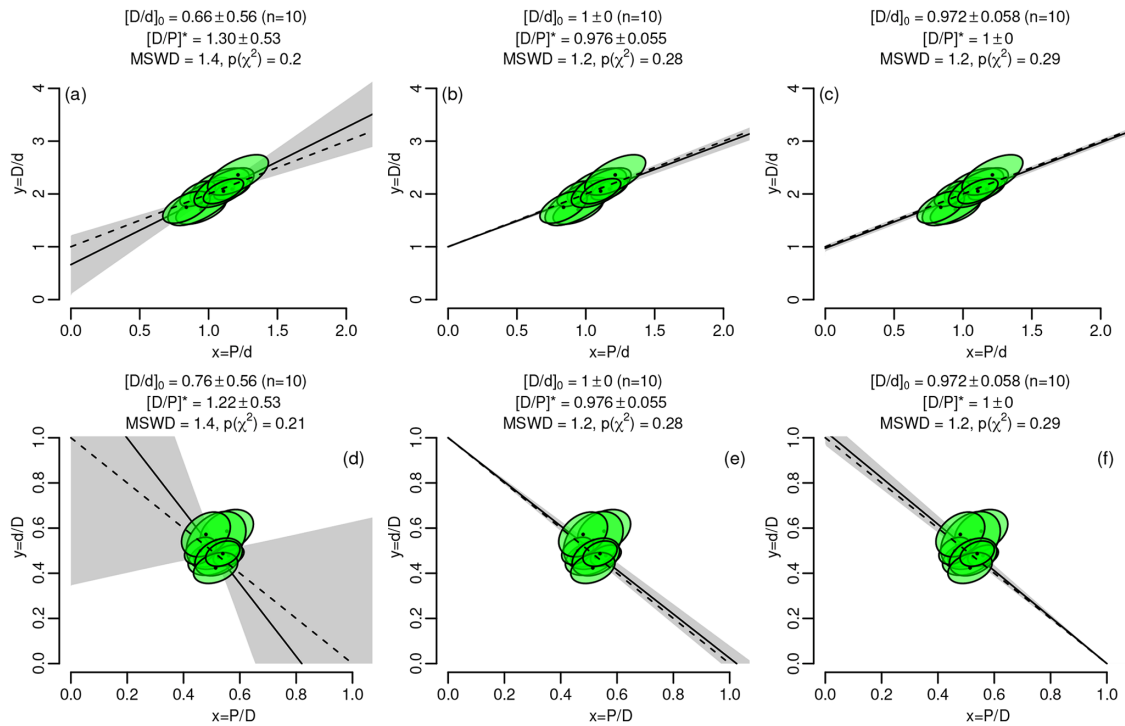
To make the problem tractable, the vector of residuals  $\Delta_i$  is assumed to be drawn from a bivariate normal distribution with zero mean and covariance matrix  $\Sigma_i$ :

$$\Delta_i \equiv \begin{bmatrix} X_i - x_i \\ Y_i - y_i \end{bmatrix} = \begin{bmatrix} \epsilon_{x,i} \\ \epsilon_{y,i} \end{bmatrix}$$

and

$$\Sigma_i \equiv \begin{bmatrix} s[X_i]^2 & s[X_i, Y_i] \\ s[X_i, Y_i] & s[Y_i]^2 \end{bmatrix}. \quad (5)$$

The product of the normal probabilities for all the aliquots (from  $1 \leq i \leq n$ ) is called the “likelihood” of the parameters given the data. Maximizing this number produces the most accurate possible parameter estimates  $\hat{a}$ ,  $\hat{b}$ , and  $\hat{x}_i$ . Alternatively, and equivalently, the same estimates can be obtained



**Figure 1.**  $P/D$  isochron diagrams for a synthetic dataset generated using the procedures of Appendix B. The true slope and intercept are  $a = 1$  and  $b = 1$  for the conventional isochrons (a–c) and  $a = 1$  and  $b = -1$  for the inverse isochrons (d–f). These parameter values correspond to  $[D/d]_0 = [D/P]^* = 1$ . The true trends are shown as dashed lines and the estimated trends as solid lines, with 95 % confidence envelopes shown in grey. The estimated parameter values are shown above the panels. Unconstrained isochron regression (a, d) yields imprecise and inaccurate results. Fixing the inherited component (b, e) or the radiogenic component (c, f) greatly improves both aspects of the fit.

by maximizing the logarithm of the likelihood. This is the preferred approach as it improves numerical stability and facilitates the estimation of the parameter uncertainties. The log-likelihood function can be formally defined as

$$-\mathcal{L}\mathcal{L}(\text{parameters}|\text{data}) = n \ln(2\pi) + \frac{1}{2} \sum_{i=1}^n \left( \ln |\Sigma_i| + \Delta_i^T \Sigma_i^{-1} \Delta_i \right). \quad (6)$$

This formulation allows for correlated uncertainties between the variables ( $s[X_i, Y_i] \neq 0$  in Eq. 5) but not between aliquots ( $s[X_i, Y_j] = 0$  if  $i \neq j$ ). Daëron and Vermeesch (2023) discuss the generalized case of “omnivariant” regression, where this requirement has been relaxed and Eq. (6) is replaced with a single matrix expression.

### 3 Dealing with overdispersion

The degree to which the residuals  $\Delta_i$  are consistent with the assumed analytical uncertainties  $\Sigma_i$  can be assessed using a chi-square test and the MSWD parameter. Readers who are not familiar with these concepts are referred to Appendix C for details. Isochrons that exhibit significant overdispersion with respect to the analytical uncertainties are colloquially referred to as errorchrons. Whether it is right to use such a

pejorative term for this common phenomenon is debatable (Schaen et al., 2021). Four approaches may be used to deal with errorchrons.

**Model 1.** Inflate the analytical uncertainties by a factor  $\sqrt{MSWD}$ .

**Model 2.** Ignore the analytical uncertainties and replace York regression with orthogonal regression or a similar technique. This approach will not be discussed further in this paper.

**Model 3.** Quantify the dispersion as an additional free parameter. There are two options for doing so.

**Model 3a.** Attribute the excess dispersion to variability of the inherited component and, hence, the isochron intercept. For conventional isochrons, this means that the true isotopic ratios of the cogenetic aliquots do not belong to a single isochron line but to a family of parallel isochron lines whose intercepts are normally distributed (Titterton and Halliday, 1979). The standard deviation of this normal distribution ( $\sigma_a$ ) can be used to quantify the dispersion of the data around the “central” isochron line.

**Model 3b.** Attribute the excess dispersion to diachronous closure of the isotopic system. This mechanism produces families of conventional isochrons that share a common intercept but differ in their slopes. Again, the distribution of the slopes can be assumed to follow a normal distribution, with mean  $b$  and standard deviation  $\sigma_b$ . The latter value can be used to quantify the dispersion of the data, which in this case captures the degree of diachroneity.

Model 3a isochron regression can be formalized by modifying Eq. (4) as follows:

$$\begin{cases} X_i = x_i + \epsilon_{x,i} \\ Y_i = a + bx_i + \epsilon_a + \epsilon_{y,i}. \end{cases} \quad (7)$$

The dispersion parameter  $\sigma_a$  can be estimated (as  $\hat{\sigma}_a$ ) by incorporating it into the covariance matrices:

$$\Sigma_{a,i} \equiv \begin{bmatrix} s[X_i]^2 & s[X_i, Y_i] \\ s[X_i, Y_i] & s[Y_i]^2 + \sigma_a^2 \end{bmatrix}. \quad (8)$$

The equivalent expressions for model 3b regression are

$$\begin{cases} X_i = x_i + \epsilon_{x,i} \\ Y_i = a + (b + \epsilon_b)x_i + \epsilon_{y,i} \end{cases} \quad (9)$$

and

$$\Sigma_{b,i} \equiv \begin{bmatrix} s[X_i]^2 & s[X_i, Y_i] \\ s[X_i, Y_i] & s[Y_i]^2 + (x_i\sigma_b)^2 \end{bmatrix}, \quad (10)$$

respectively.

#### 4 Specific cases

Three specific cases of Eq. (6) are of interest in the present discussion.

**Model 1.**

$$\mathcal{LL}_1(a, b, \mathbf{x}|\mathbf{X}, \mathbf{Y}, \Sigma) = \text{constant} - \frac{1}{2} \sum_{i=1}^n \Delta_i^T \Sigma_i^{-1} \Delta_i \quad (11)$$

**Model 3a.**

$$\begin{aligned} \mathcal{LL}_{3a}(a, b, \sigma_a, \mathbf{x}|\mathbf{X}, \mathbf{Y}, \Sigma) = \\ \text{constant} - \frac{1}{2} \sum_{i=1}^n \left( \ln|\Sigma_{a,i}| + \Delta_i^T \Sigma_{a,i}^{-1} \Delta_i \right) \end{aligned} \quad (12)$$

**Model 3b.**

$$\begin{aligned} \mathcal{LL}_{3b}(a, b, \sigma_b, \mathbf{x}|\mathbf{X}, \mathbf{Y}, \Sigma) = \\ \text{constant} - \frac{1}{2} \sum_{i=1}^n \left( \ln|\Sigma_{b,i}| + \Delta_i^T \Sigma_{b,i}^{-1} \Delta_i \right) \end{aligned} \quad (13)$$

Note that the (logged) determinant of the covariance matrices is not included in the constant for model 3 fits because  $\Sigma_{a,i}$  and  $\Sigma_{b,i}$  are both functions of  $\mathbf{x}$ , whereas  $\Sigma_i$  is not.  $\mathcal{LL}$  can be numerically maximized using an iterative two-step procedure. For model 1 regression, the first step maximizes  $\mathcal{LL}_1(\mathbf{x}|a, b, \mathbf{X}, \mathbf{Y}, \Sigma)$  to find  $\mathbf{x}$  for any pair of  $a$  and  $b$  values. The second step repeats the first step for different values of  $a$  and  $b$  until the overall optimum is reached.

This procedure can easily be adapted to model 3 regression by adding  $\sigma_a$  or  $\sigma_b$  to the unknown parameters. It works well for model 3a, where the first step has a direct solution. However, it is slow for model 3b regression, in which finding  $\mathbf{x}$  requires an additional level of iteration. See Appendix D for further details. Fortunately, in practice model 3b is rarely or never needed for geochronology, as explained next.

#### 5 Flipped isochron regression

For inverse isochrons, model 3b regression does not actually provide any useful information. Unlike conventional isochrons, whose chronological information is contained in the slope, the chronological information of inverse isochrons is contained in their horizontal intercept. This information can be unlocked by flipping the axes of the isochron diagram around, inverting the isochron, and treating the  $D/P$  ratio as the dependent variable. Applying model 3a regression to the flipped data yields a dispersion estimate  $\sigma_a$  that can be used to quantify the age dispersion.

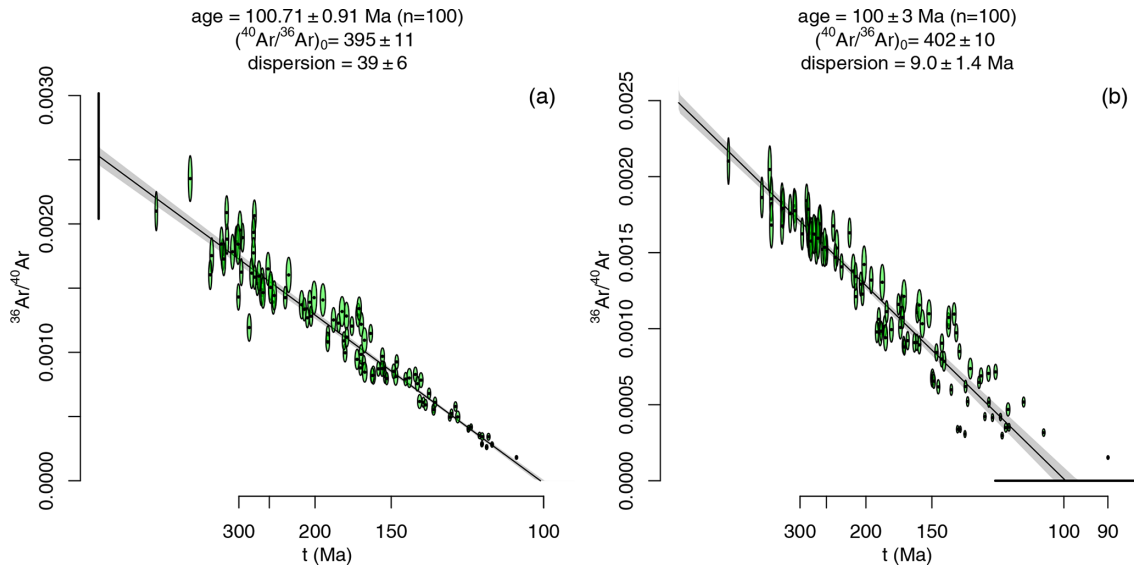
The same trick can be used to estimate the slope uncertainty of a conventional isochron. A pragmatic way to avoid the slow convergence rate of model 3b regression is to invert the isochron, flip the dependent and independent variables around, invert the isochron a second time, and carry out a model 3a regression on the transformed data. The resulting  $\sigma_a$  estimate can be converted to a  $\sigma_b$  estimate:

$$\hat{\sigma}_b(\text{conventional isochron}) \approx \frac{\hat{\sigma}_a}{\hat{a}^2}(\text{flipped inverse isochron}). \quad (14)$$

Figure 2 applies model 3 regression to two synthetic Ar–Ar datasets with a non-radiogenic  $^{40}\text{Ar}/^{36}\text{Ar}$  ratio of 400 and an age of 100 Ma. The dataset of Fig. 2a exhibits overdispersion of the  $y$  intercept ( $[^{40}\text{Ar}/^{36}\text{Ar}]_0 = 400, \sigma_a = 40$ ), whereas the dataset of Fig. 2b is overdispersed in the  $x$  intercept ( $t = 100 \text{ Ma}, \sigma_t = 10 \text{ Ma}$ ). In both cases, the maximum likelihood algorithm has retrieved the correct solution from the noisy data.

#### 6 Anchored isochron regression

Anchored isochron regression requires just a trivial modification of the maximum likelihood algorithm. It suffices to treat the anchored parameter as data. For example, the slope and intercept of a model 1 isochron can be anchored by maximizing  $\mathcal{LL}_1(a, \mathbf{x}|b, \mathbf{X}, \mathbf{Y}, \Sigma)$  and  $\mathcal{LL}_1(b, \mathbf{x}|a, \mathbf{X}, \mathbf{Y}, \Sigma)$ , respectively. Note that these calculations assume that the anchor



**Figure 2.** Model 3 regression of two synthetic Ar–Ar datasets. **(a)** Model 3a regression through 100 randomly drawn aliquots from an isochron with a true non-radiogenic  $^{40}\text{Ar}/^{39}\text{Ar}$  ratio of  $400 \pm 40$  ( $1\sigma$ ) and a true age of 100 Ma; **(b)** model 3b regression through 100 aliquots from an isochron with a true non-radiogenic  $^{40}\text{Ar}/^{39}\text{Ar}$  ratio of 400 and a true age of  $100 \pm 10$  Ma ( $1\sigma$ ). Error bars on the y and x axis show the overdispersion at 95 % confidence.

is known exactly. This is reflected in the zero uncertainties of the non-radiogenic component in Fig. 1b and e and of the radiogenic components in Fig. 1c and f. Such absolute certainty is unrealistic in the noisy world of geology. Section 7 introduces two ways to formally account for uncertainty in anchored isochron regression.

### 7 Accounting for anchor uncertainty

When assigning uncertainty to statistical parameters, it is important to clearly define the meaning of this uncertainty. In the case of anchored isochron regression, the uncertainty of the intercept or slope can carry two meanings. For example, when the intercept of an isochron is anchored at a value  $\bar{a} \pm \sigma_{\bar{a}}$ , this can mean either of the following.

**Model 1.** The data are underlain by a single isochron whose intercept is only approximately known, with a most likely value of  $\bar{a}$  and a precision (standard error) of  $\sigma_{\bar{a}}$ .

**Model 3.** The data were drawn from a family of isochron lines whose intercepts follow a normal distribution with mean  $\bar{a}$  and standard deviation  $\sigma_{\bar{a}}$ .

These two different approaches can be implemented by replacing the log-likelihood functions of Eq. (13) with the fol-

lowing.

**Model 1.**

$$\mathcal{L}\mathcal{L}_1(a, b, \mathbf{x}|\bar{a}, \sigma_{\bar{a}}, \mathbf{X}, \mathbf{Y}, \Sigma) = \text{constant} + \mathcal{L}\mathcal{L}_1(a, b, \mathbf{x}|\mathbf{X}, \mathbf{Y}, \Sigma) - \frac{1}{2} \left( \frac{a - \bar{a}}{\sigma_{\bar{a}}} \right)^2 \quad (15)$$

**Model 3.**

$$\mathcal{L}\mathcal{L}_{3a}(b, \mathbf{x}|\bar{a}, \sigma_{\bar{a}}, \mathbf{X}, \mathbf{Y}, \Sigma)$$

The equivalent expressions for anchored slopes with uncertainty are as follows.

**Model 1.**

$$\mathcal{L}\mathcal{L}_1(a, b, \mathbf{x}|\bar{b}, \sigma_{\bar{b}}, \mathbf{X}, \mathbf{Y}, \Sigma) = \text{constant} + \mathcal{L}\mathcal{L}_1(a, b, \mathbf{x}|\mathbf{X}, \mathbf{Y}, \Sigma) - \frac{1}{2} \left( \frac{b - \bar{b}}{\sigma_{\bar{b}}} \right)^2 \quad (16)$$

**Model 3.**

$$\mathcal{L}\mathcal{L}_{3b}(a, \mathbf{x}|\bar{b}, \sigma_{\bar{b}}, \mathbf{X}, \mathbf{Y}, \Sigma)$$

Note that model 1 treats the slope and intercept as unknowns, so their maximum likelihood estimates generally do *not* equal the anchored values  $\bar{a}$  and  $\bar{b}$ . However, the smaller  $\sigma_{\bar{a}}$  and  $\sigma_{\bar{b}}$  are relative to  $\bar{a}$  and  $\bar{b}$ , the closer that  $\hat{a}$  and  $\hat{b}$  are to  $\bar{a}$  and  $\bar{b}$ . In contrast, anchored model 3 regression fixes the dispersion of the anchored parameters so that  $\hat{\sigma}_a = \sigma_{\bar{a}}$  for anchored intercepts, whereas  $\hat{\sigma}_b = \sigma_{\bar{b}}$  for anchored slopes.

Figure 3 applies the two approaches to the synthetic dataset of Fig. 1 using anchors of  $[D/d]_0 = 1 \pm 0.05$  and  $[D/P]^* = 1 \pm 0.05$  ( $1\sigma$ ). As expected, models 1 and 3 produce slightly different outcomes, with model 1 resulting in  $[D/d]_0$  and  $[D/P]^*$  estimates that are slightly different than the anchored values, whereas model 3 reproduces the anchors exactly. The model 3 fits produce more precise estimates for



the unanchored quantities. This reflects the fact that some of the uncertainty in the isochron is captured by the dispersion parameter, which is absent from the model 1 fit.

Model 3 partitions the analytical and geological uncertainty between the standard errors of  $\hat{a}$  and  $\hat{b}$  on the one hand and the dispersion parameter on the other hand. It is not easy to visualize this partitioned uncertainty. Figure 3 uses a grey confidence envelope to show the analytical uncertainty and a black error bar to visualize the geological dispersion. Together these two items represent the total uncertainty budget.

It is, of course, also possible to treat the dispersion parameter as an unknown by maximizing  $\mathcal{L}\mathcal{L}_{3a}(b, \sigma_a, \mathbf{x}|\hat{a}, \mathbf{X}, \mathbf{Y}, \Sigma)$  or  $\mathcal{L}\mathcal{L}_{3b}(a, \sigma_b, \mathbf{x}|\hat{b}, \mathbf{X}, \mathbf{Y}, \Sigma)$ . Section 9 will give an example of this in the context of U–Pb geochronology.

## 8 Pb/Pb isochrons

The  $^{207}\text{Pb}/^{206}\text{Pb}$  method is based on the paired decay of two radioactive parents  $P_1$  and  $P_2$  ( $^{238}\text{U}$  and  $^{235}\text{U}$ ) to two daughters  $D_1$  and  $D_2$  ( $^{206}\text{Pb}$  and  $^{207}\text{Pb}$ , respectively) in the presence of a non-radiogenic sister isotope  $d$  ( $^{204}\text{Pb}$ ). This gives rise to the following implicit age equation:

$$\frac{[D_2/d]_i - [D_2/d]_0}{[D_1/d]_i - [D_1/d]_0} = \left[ \frac{P_2}{P_1} \right] \frac{e^{\lambda_2 t} - 1}{e^{\lambda_1 t} - 1}, \quad (17)$$

where  $\lambda_1$  and  $\lambda_2$  are the decay constants of  $P_1$  and  $P_2$ , respectively, and  $[P_2/P_1]$  is constant. Equation (17) can be recast into the generic form of Eq. (3) to form a “ $D/D$  isochron” using the mapping of Table A1. This gives rise to conventional ( $^{207}\text{Pb}/^{204}\text{Pb}$  vs.  $^{206}\text{Pb}/^{204}\text{Pb}$ ) and inverse ( $^{207}\text{Pb}/^{206}\text{Pb}$  vs.  $^{204}\text{Pb}/^{206}\text{Pb}$ ) isochrons.

Unlike inverse  $P/D$  isochrons, which are characterized by negative slopes, inverse  $D/D$  isochrons have positive slopes. There is perfect symmetry between conventional and inverse  $D/D$  isochrons in the sense that the intercept of a conventional isochron equals the slope of an inverse isochron and vice versa. This slightly changes and, in fact, simplifies the procedure for anchored isochron regression.

There is no need for flipped isochron regression of Pb/Pb data. Instead, all possible scenarios can be handled by straightforward model 1 and model 3a regression. Isochrons can be anchored to the non-radiogenic component in conventional isochron space and to the radiogenic component in inverse isochron space.

## 9 U/Pb isochrons

The U/Pb method, like the Pb/Pb method, is based on the paired decay of  $^{238}\text{U}$  to  $^{206}\text{Pb}$  and of  $^{235}\text{U}$  to  $^{207}\text{Pb}$ . The two chronometers can be treated separately and plotted as  $P/D$  isochrons. These two-dimensional isochrons can be anchored to the radiogenic and non-radiogenic composition using the

methods described in Sects. 2–7.

$$\begin{cases} \left[ \frac{^{206}\text{Pb}}{^{204}\text{Pb}} \right]_i = \left[ \frac{^{206}\text{Pb}}{^{204}\text{Pb}} \right]_0 + \left[ \frac{^{238}\text{U}}{^{204}\text{Pb}} \right]_i (e^{\lambda_{238} t} - 1) \\ \left[ \frac{^{207}\text{Pb}}{^{204}\text{Pb}} \right]_i = \left[ \frac{^{207}\text{Pb}}{^{204}\text{Pb}} \right]_0 + \left[ \frac{^{235}\text{U}}{^{204}\text{Pb}} \right]_i (e^{\lambda_{235} t} - 1) \end{cases} \quad (18)$$

Alternatively, the two ingrowth equations can also be coupled, forming a three-dimensional “total Pb/U isochron”. This problem can be solved using the method of maximum likelihood (Ludwig, 1998). There is just one important difference between this solution and the York-like algorithm of Sect. 2. York regression is formulated in terms of a generic intercept ( $a$ ) and slope ( $b$ ). In contrast, the “Ludwig” algorithm is formulated in terms of the actual non-radiogenic  $^{206}\text{Pb}/^{204}\text{Pb}$  ( $=\alpha$ ) and  $^{207}\text{Pb}/^{204}\text{Pb}$  ( $=\beta$ ) ratio, as well as the age ( $t$ ) of the system.

It is relatively straightforward to generalize the concept of model 3 regression to total Pb/U isochrons when the dispersion is attributed to the isochron age. However, it is not immediately clear how to do the same for the non-radiogenic composition because this would require partitioning the dispersion between the  $\alpha$  and  $\beta$  parameters. This paper will not attempt to answer this question. Instead, let us conclude the technical discussion by shifting to a far more common type of U/Pb data.

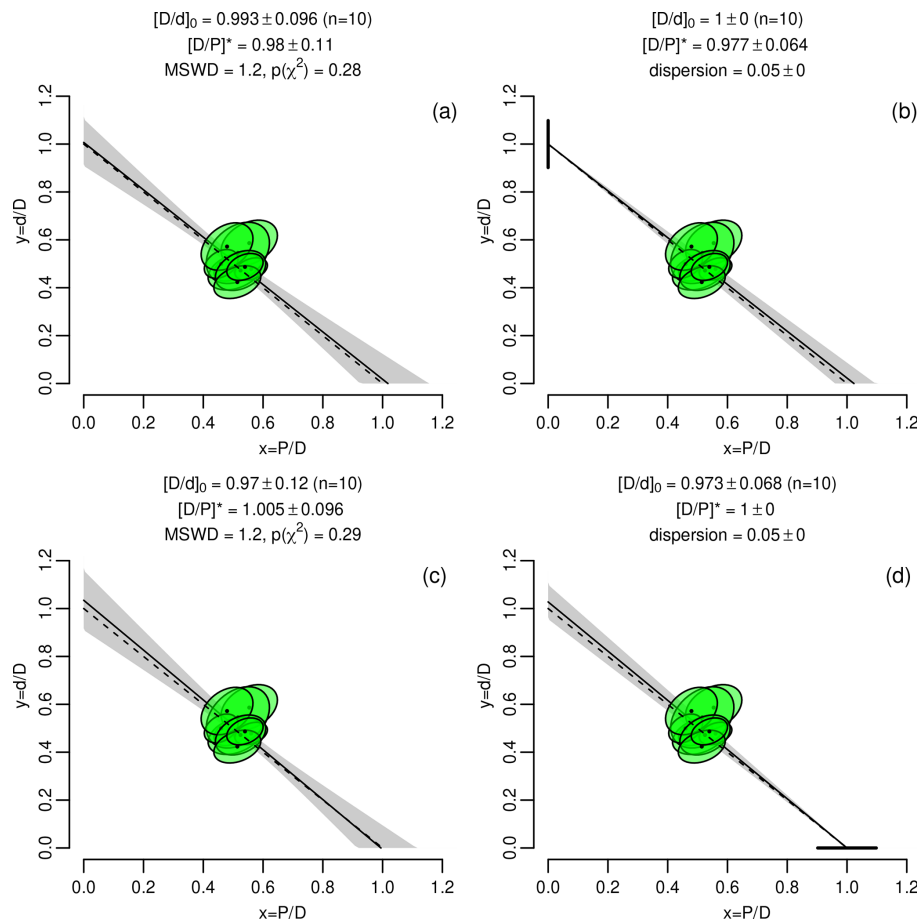
Most published U/Pb studies do not report  $^{204}\text{Pb}$  because this nuclide is rare and difficult to measure. Instead, “semitotal Pb/U” regression is done in Wetherill or Tera–Wasserburg concordia space, which are akin to conventional and inverse isochron space, respectively. In the absence of a non-radiogenic sister isotope of Pb, the isochron is then redefined as a mixing line between an inherited  $^{207}\text{Pb}/^{206}\text{Pb}$  ratio and the radiogenic  $^{206}\text{Pb}/^{238}\text{U}$  and  $^{207}\text{Pb}/^{235}\text{U}$  ratios.

The semitotal Pb/U problem can be cast into a maximum likelihood format by replacing Eq. (4) with

$$\begin{cases} X_i = x_i + e^{\lambda_{235} t} - 1 + \epsilon_{x,i} \\ Y_i = a + bx_i + \epsilon_{y,i}, \end{cases} \quad (19)$$

where  $X_i$  and  $Y_i$  are the measured  $^{207}\text{Pb}/^{235}\text{U}$  and  $^{206}\text{Pb}/^{238}\text{U}$  ratios of the  $i$ th aliquot, respectively, and  $x_i$  is the true  $^{207}\text{Pb}_0/^{235}\text{U}$  ratio of the  $i$ th aliquot, whilst  $a = e^{\lambda_{238} t} - 1$  and  $b = [^{206}\text{Pb}/^{207}\text{Pb}]_0 [^{235}\text{U}/^{238}\text{U}]$ . Equation (19) can be solved using the same recipe as the York reformulation of Sects. 2–7, with the caveat that the isochron terminates on the concordia line at the U/Pb composition corresponding to  $t$ .

Figure 4 shows a model 3a anchored semitotal Pb/U isochron in Tera–Wasserburg space. The true age for this synthetic dataset is 1500 Ma. The unconstrained model 1 isochron produces an imprecise isochron age of  $1491 \pm 100$  Ma, which is improved to  $1504 \pm 38$  Ma by anchoring to  $[^{207}\text{Pb}/^{206}\text{Pb}]_0 = 1.1$ . Just like the  $P/D$  isochrons of Fig. 3, the uncertainty budget for the U/Pb isochrons of Fig. 4 is partitioned between the analytical and geological dispersion,



**Figure 3.** Anchored isochrons for the data of Fig. 1, shown on inverse isochrons. Panels (a) and (b) anchor the non-radiogenic endmember composition to  $[D/d]_0 = 1 \pm 0.05$  ( $1\sigma$ ). Panels (c) and (d) anchor the radiogenic endmember composition to  $[D/P]^* = 1 \pm 0.05$  ( $1\sigma$ ). Panels (a) and (c) use model 1 isochron regression, whereas panels (b) and (d) use model 3. Note how the estimated values for  $[D/d]_0$  and  $[D/P]^*$  are exactly the same as the anchored values for model 3 but not for model 1. The vertical error bar in panel (b) and the horizontal error bar in panel (d) represent the fixed dispersion of the vertical and horizontal intercept, respectively.

where the latter is treated as a free parameter with a maximum likelihood estimate of  $\hat{\sigma}_a = 0.053 \pm 0.026$ .

## 10 Implementation in IsoplotR

All the algorithms presented in this paper have been implemented in the free and open geochronological toolbox IsoplotR (Vermeesch, 2018). The ability to fit isochrons with models 1, 2, and 3 dates back to IsoplotR's first public release. Versions 2.1 and 5.2 added model 1 anchors to U/Pb isochrons without and with uncertainty, respectively. Anchored York regression and model 3 anchors with uncertainty were added to IsoplotR 6.0. These functions can be accessed via the GUI or from the command line.

In the GUI, anchored regression is available from the “isochron” menu (so not from the “concordia” menu for U/Pb data). The relevant settings can be selected from the “options” menu and should be self-explanatory. Moving on to the R console, use a built-in  $^{40}\text{Ar}/^{39}\text{Ar}$  dataset to illustrate

the command-line functionality and start with an unanchored inverse isochron fit using York regression.

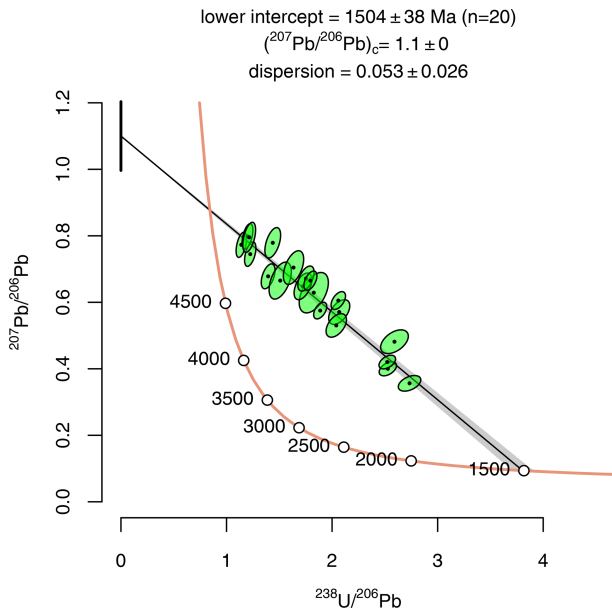
```
library(IsoplotR)
attach(examples)
isochron(ArAr)
```

Carry out model 3a isochron regression anchored to a variable initial  $^{40}\text{Ar}/^{36}\text{Ar}$  ratio of  $300 \pm 5$  ( $1\sigma$ ).

```
settings('iratio', 'Ar40Ar36', 300, 5)
isochron(ArAr, anchor=1, model=3, taxis=TRUE)
```

Here, `settings(...)` is used to change the non-radiogenic argon composition, `anchor=1` fixes the isochron to the inherited component, `model=3` tells IsoplotR to assign the dispersion to the reported uncertainty of the atmospheric  $^{40}\text{Ar}/^{36}\text{Ar}$  ratio, and `taxis=TRUE` changes the  $x$  axis to a timescale (as in Fig. 2).

Finally, carry out model 3b isochron regression anchored to a radiogenic composition that was set at  $62 \pm 1$  Ma ( $1\sigma$ ).



**Figure 4.** Model 3a semitotal Pb/U isochron for a synthetic dataset with a true age  $t = 1500$  Ma, anchored to  $^{207}\text{Pb}/^{206}\text{Pb} = 1.10$  with the dispersion parameter treated as a free parameter.

```
isochron(ArAr, anchor=c(2, 62, 1),
        model=3, t axis=TRUE)
```

The dispersion can be changed to a free parameter by setting the third element of the `anchor` argument to a non-positive value. Documentation for additional options and settings can be accessed from the console using the `help(isochron)` or `?isochron` commands.

## 11 Conclusions

This paper builds on previous work by McIntyre et al. (1966), Titterton and Halliday (1979), and Ludwig (2012). McIntyre et al. (1966) introduced the concept of model 3 isochron regression, attributing the excess dispersion of errorchrons to variability of the inherited component. This approach is implemented in Isoplot (Ludwig, 2012) and is referred to as model 3a regression in this paper. Note that the McIntyre et al. (1966) definition of model 2 regression differs from that of Ludwig (2012). It is more akin (but not identical) to the  $\sqrt{\text{MSWD}}$  approach of Sect. 3.

Titterton and Halliday (1979) recast the model 3 algorithm of McIntyre et al. (1966) in a maximum likelihood framework. This paper extends the maximum likelihood approach to a second type of geological scenario, in which excess dispersion of the data around the isochron is not attributed to the inherited component but to the radiogenic component. This approach is referred to as model 3b regression.

The maximum likelihood formulation of isochron regression can also be used to anchor isochrons to either the inherited or the radiogenic component and to assign geologically meaningful uncertainty to the anchor. In reality it is possible that some samples are affected by both mechanisms so that different aliquots of the same sample differ in their initial ratios as well as the timing of their isotopic closure. Unfortunately, it is not possible to simultaneously capture both types of dispersion using the algorithms of this paper.

The difference between model 1 and model 3 regression represents two contrasting views of geological reality. The model 1 approach assumes that the isotopic composition of minerals represents discrete components recorded at distinct events. In contrast, model 3 isochrons represent a “fuzzier” reality, in which the initial composition or timing of isotopic closure is allowed to vary within a rock. Under the latter model, the concept of an errorchron no longer makes sense.



**Appendix A: Definitions**

$P$	a radioactive parent (e.g. $^{87}\text{Rb}$ )
$D$	the radiogenic daughter of $P$ (e.g. $^{87}\text{Sr}$ )
$d$	a non-radiogenic sister isotope of $D$ (e.g. $^{86}\text{Sr}$ )
$a, b, x_i, y_i$	the true (but unknown) values of the intercept, slope, and isotopic ratio of the $i$ th aliquot (out of $n$ aliquots) from a sample
$\hat{a}, \hat{b}, \hat{x}_i, \hat{y}_i$	the estimated values of $a, b, x_i,$ and $y_i,$ respectively
$X_i, Y_i$	the measured values of $x_i$ and $y_i,$ respectively
$s[X_i], s[Y_i], s[X_i, Y_i]$	the standard errors and covariance of $X_i$ and $Y_i$
$\epsilon_{x,i} = X_i - x_i, \epsilon_{y,i} = Y_i - y_i$	the residuals of the data around the best-fit line
$\Sigma_i$	the $2 \times 2$ covariance matrix of $X_i$ and $Y_i$
$\mathbf{x}, \mathbf{y}, \mathbf{X}, \mathbf{Y}, \Sigma$	arrays and sets containing all $n$ $x_i, y_i, X_i, Y_i,$ and $\Sigma_i$ values

All uncertainties in this paper are reported as 95 % confidence intervals except if noted otherwise.

**Table A1.** Assignment of the isotopic ratios for  $P/D$  and  $D/D$  isochrons to the generic parameters of Eq. (3).

Parameter	$P/D$		$D/D$	
	Conventional	Inverse	Conventional	Inverse
$x_i$	$[P/d]_i$	$[P/D]_i$	$[D_1/d]_i$	$[d/D_1]_i$
$y_i$	$[D/d]_i$	$[d/D]_i$	$[D_2/d]_i$	$[D_2/D_1]_i$
$a$	$[D/d]_0$	$[d/D]_0$	$[D_2/d]_0$	$[D_2/D_1]^*$
$b$	$[D/P]^*$	$-[d/D]_0/[P/D]^*$	$[D_2/D_1]^*$	$[D_2/d]_0$

**Appendix B: Data**

The synthetic data in Figs. 1 and 3 were generated as follows.

1. Generate  $\mathbf{x}$  by drawing 10 random numbers from a logit-normal distribution where the mean and standard deviation of the logits are 0 and 1/5, respectively.
2. Generate  $\mathbf{y}$  by plugging  $\mathbf{x}$  into Eq. (3) with  $a = b = 1$ .
3. Turn the resulting 10 pairs of  $\{x_i, y_i\}$  values into 10 trios of  $\{P_i, D_i, d_i\}$  values where  $d_i$  is a random number between 100 and 500.
4. Use each of the  $\{P_i, D_i, d_i\}$  values obtained in the previous step as parameters for a Poisson experiment.
5. Use the Poisson values obtained in the previous step to form 10 pairs of  $\mathbf{X}$  and  $\mathbf{Y}$  measurements and their (correlated) uncertainties.
6. Add some excess dispersion by shrinking the uncertainties by 20 %.

Note that this procedure fits the definition of a model 1 isochron, even though panels (b) and (d) of Fig. 3 use the resulting data to illustrate model 3 regression. The synthetic data in Figs. 2 and 4 are generated using a slightly different algorithm that simulates model 3a and 3b overdispersion. This modified procedure uses steps 1 and 3–4 of the procedure for Fig. 1 whilst replacing step 2 with a random effects model, in which  $a$  and  $b$  are drawn from random normal distributions with standard deviations corresponding to the dispersion parameters  $\sigma_{\hat{a}}$  (10 % for Fig. 2a and 5 % for Fig. 4) and  $\sigma_{\hat{b}}$  (10 % for Fig. 2b).

**Appendix C: Testing overdispersion**

Whether a dataset is overdispersed with respect to the analytical uncertainties can be formally assessed using the chi-square statistic and test. To this end, calculate the  $\chi^2$  statistic using the sum of squares term of Eq. (6).

$$\chi^2 = \sum_{i=1}^n \Delta_i^T \Sigma_i^{-1} \Delta_i \tag{C1}$$

If analytical uncertainty is the only source of dispersion in the dataset, then  $\chi^2$  is expected to follow a chi-square distribution with  $n - 2$  degrees of freedom (DF =  $n - 2$ , or, for anchored regression, DF =  $n - 1$ ). This is called the “null distri-

bution". The probability of observing a value greater than  $\chi^2$  under this distribution is called the  $p$  value. A  $p$  value cutoff of 0.05 is generally used to distinguish isochrons ( $p \geq 0.05$ ) from errorchrons ( $p < 0.05$ ). Dividing Eq. (C1) by the number of degrees of freedom produces a "reduced chi-square statistic", which is known to geologists as an MSWD (McIntyre et al., 1966).

$$\text{MSWD} = \frac{\chi^2}{\text{DF}} \quad (\text{C2})$$

In the absence of overdispersion and for sufficiently large samples ( $n > 20$ ), the expected null distribution of the MSWD is approximately normal with a mean of 1 and a standard deviation  $\sqrt{2/\text{DF}}$  (Wendt and Carl, 1991). This information can be used as an alternative way to test the null hypothesis.

#### Appendix D: Further details about the maximum likelihood solution

Section 4 describes how the log-likelihood functions of Eq. (13) can be maximized in a two-step process: (1) find the  $x_i$  values that maximize  $\mathcal{LL}$  for any pair of  $a$  and  $b$  values, and (2) search the space of  $a$  and  $b$  values until the overall maximum is found. For model 1 regression, the first step has a direct solution. Some shorthand notation for the inverse of the covariance matrix  $\Sigma_i$  is introduced:

$$\Sigma_i^{-1} \equiv \begin{bmatrix} \Omega_{1,1}^i & \Omega_{1,2}^i \\ \Omega_{2,1}^i & \Omega_{2,2}^i \end{bmatrix}, \quad (\text{D1})$$

where  $\Omega_{1,2} = \Omega_{2,1}$ . The fitted points  $x_i$  can be found by solving

$$\left. \frac{\partial \mathcal{LL}_1}{\partial x_i} \right|_{\hat{x}_i} = 0, \quad (\text{D2})$$

which yields

$$\hat{x}_i = \frac{X_i(\Omega_{1,1}^i + b\Omega_{1,2}^i) + (Y_i - a)(\Omega_{2,1}^i + b\Omega_{2,2}^i)}{(\Omega_{1,1}^i + b\Omega_{1,2}^i) + b(\Omega_{2,1}^i + b\Omega_{2,2}^i)}. \quad (\text{D3})$$

The same approach can be used for model 3a regression (with  $\partial \mathcal{LL}_{3a}/\partial x_i = 0$ ). However, it does not work for model 3b isochrons because  $\partial^2 \mathcal{LL}_{3b}/\partial x_i \partial x_j \neq 0$  when  $i \neq j$ . Therefore, the  $x_i$  values are interdependent and cannot be solved separately. Although it is possible to jointly optimize the entire  $\mathbf{x}$  vector, this takes considerably longer than model 1 and model 3a regression. Fortunately, as explained in Sect. 5, model 3b regression can be reformulated in terms of the model 3a algorithm.

**Code availability.** IsoplotR is available from CRAN (<https://CRAN.Rproject.org/package=IsoplotR>, last access: 15 July 2024) and Zenodo (<https://doi.org/10.5281/zenodo.11114310>; Vermeesch and Lloyd, 2024).

**Data availability.** This paper only uses synthetic data, which were generated using the procedures outlined in Appendix B.

**Competing interests.** The author is a member of the editorial board of *Geochronology*. The peer-review process was guided by an independent editor, and the author also has no other competing interests to declare.

**Disclaimer.** Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

**Acknowledgements.** The author would like to thank reviewers Donald Davis and John Rudge for suggesting to expand the discussion of model 3 regression and for checking the equations. This work was completed during a sabbatical stay at the Institute of Geology and Geophysics, Chinese Academy of Sciences, which was supported by the CAS President's International Fellowship Initiative (PIFI). The author would like to thank Xian-Hua Li of IGG-CAS and Yang Li of Peking University for hosting him at IGG-CAS. He would also like to thank IsoplotR users Guilhem Hoareau (CNRS, France), Shitou Wu (IGG-CAS, China), and Stijn Glorie (Adelaide, Australia) for discussing the utility of anchored isochrons and moving this feature up the priority list for IsoplotR.

**Financial support.** This research has been supported by the Natural Environment Research Council (grant no. NE/T001518/1).

**Review statement.** This paper was edited by Noah M McLean and reviewed by Donald Davis and John Rudge.

#### References

- Daëron, M. and Vermeesch, P.: Omnivariant generalized least squares regression: Theory, geochronological applications, and making the case for reconciled  $\Delta_{47}$  calibrations, *Chem. Geol.*, 647, 121881, <https://doi.org/10.1016/j.chemgeo.2023.121881>, 2023.
- Li, Y. and Vermeesch, P.: Short communication: Inverse isochron regression for Re–Os, K–Ca and other chronometers, *Geochronology*, 3, 415–420, <https://doi.org/10.5194/gchron-3-415-2021>, 2021.
- Ludwig, K. R.: On the treatment of concordant uranium–lead ages, *Geochim. Cosmochim. Ac.*, 62, 665–676, [https://doi.org/10.1016/S0016-7037\(98\)00059-3](https://doi.org/10.1016/S0016-7037(98)00059-3), 1998.
- Ludwig, K. R.: User's manual for Isoplot 3.75: a geochronological toolkit for Microsoft Excel, Berkeley Geochronology Center, 5, <https://doi.org/10.5281/zenodo.12744084>, 2012.

- McIntyre, G. A., Brooks, C., Compston, W., and Turek, A.: The Statistical Assessment of Rb-Sr Isochrons, *J. Geophys. Res.*, 71, 5459–5468, 1966.
- Schaen, A. J., Jicha, B. R., Hodges, K. V., Vermeesch, P., Stelten, M. E., Mercer, C. M., Phillips, D., Rivera, T. A., Jourdan, F., Matchan, E. L., Hemming, S. R., Morgan, L. E., Kelley, S. P., Cassata, W. S., Heizler, M. T., Vasconcelos, P. M., Benowitz, J. A., Koppers, A. A. P., Mark, D. F., Niespolo, E. M., Sprain, C. J., Hames, W. E., Kuiper, K. F., Turrin, B. D., Renne, P. R., Ross, J., Nomade, S., Guillou, H., Webb, L. E., Cohen, B. A., Calvert, A. T., Joyce, N., Ganerød, M., Wijbrans, J., Ishizuka, O., He, H., Ramirez, A., Pfänder, J. A., Lopez-Martínez, M., Qiu, H., and Singer, B. S.: Interpreting and reporting  $^{40}\text{Ar}/^{39}\text{Ar}$  geochronologic data, *Bulletin*, 133, 461–487, <https://doi.org/10.1130/B35560.1>, 2021.
- Titterton, D. M. and Halliday, A. N.: On the fitting of parallel isochrons and the method of maximum likelihood, *Chem. Geol.*, 26, 183–195, 1979.
- Vermeesch, P.: *IsoplotR*: a free and open toolbox for geochronology, *Geosci. Front.*, 9, 1479–1493, 2018.
- Vermeesch, P. and Lloyd, J. C.: *pvermeesch/IsoplotR*: 6.2 (6.2), Zenodo [code], <https://doi.org/10.5281/zenodo.11114310>, 2024.
- Wendt, I. and Carl, C.: The statistical distribution of the mean squared weighted deviation, *Chem. Geol.: Isotope Geoscience Section*, 86, 275–285, 1991.
- York, D., Evensen, N. M., Martínez, M. L., and De Basabe Delgado, J.: Unified equations for the slope, intercept, and standard errors of the best straight line, *Am. J. Phys.*, 72, 367–375, 2004.