Bulletin of the Geological Survey of Japan, vol. 68 (3), p. 131–140, 2017

Notes and Comments

# A new tool for calculation and visualization of U–Pb age data: UPbplot.py

Atsushi Noda[1,*]

Atsushi Noda (2017) A new tool for calculation and visualization of U–Pb age data: UPbplot.py. *Bull. Geol. Surv. Japan*, vol. 68 (3), p. 131–140, 4 figs, 1 table.

**Abstract:** This paper presents usages, examples, and mathematical backgrounds of a Python script, UPbplot.py, which was newly developed for calculation and visualization of U–Pb age data. The script is a collection of various functions to deal with the one- and two-dimensional weighted means, concordia ages, and concordia-intercept ages on the conventional (Wetherill) and Tera–Wasserburg concordia diagrams for U–Pb age data. This script can calculate those ages and output images including concordia diagrams, bar plots, and histograms.

**Keywords:** geochronology, open-source software, Python, U–Pb age

## 1. Introduction

For visualization of geochronological data, such as uranium–lead (U–Pb) ages, Isoplot (Ludwig, 2012) has been successfully used for a long time. Unfortunately, because it was written in Microsoft® Visual Basic®, there are some dependent problems for specific versions of Excel® or operating systems. Although a new open-source project (e.g., Topsoil; Bowring and PI CIRDLES. org Open Source Development Team, 2016) is now developing to replace Isoplot, it can only deal with the conventional diagrams at present. In order to overcome such problems, a new script, UPbplot.py (Noda, 2016), was written in Python which is a commonly used language in the scientific community. It enables us to

- plot scattered data with error ellipses on conventional ($^{207}Pb^*/^{235}U$–$^{206}Pb^*/^{238}U$; Wetherill, 1956) and Tera–Wasserburg ($^{207}Pb^*/^{206}Pb^*$–$^{238}U/^{206}Pb^*$; Tera and Wasserburg, 1972) concordia diagrams,
- calculate the one- or two-dimensional weighted mean, concordia, and concordia-intercept ages with errors on both concordia diagrams, and
- work on any operating systems which can run Python scripts.

The purposes of this short article are to introduce the usage of this script, show examples of output images, and explain details of the calculations used in the script.

## 2. Usage

### 2.1 Preparations

The script was written in Python version 2.7 series. If the script is run by Python version 3 series, some modifications are required by reference to the comments in the script.

Mandatory libraries of matplotlib (Hunter, 2007), pandas (McKinney, 2010), and SciPy (Jones *et al.*, 2001–) should be installed (Table 1). When the script is executed in the GUI mode, further libraries of PySide, wxPython, and quickgui will be required (Table 1).

Table 1 List of libraries used in the script of UPbplot.py.

| Library | Version | License | URL |
|---|---|---|---|
| matplotlib[1] | 2.0.0 | PSF BSD[2] | http://matplotlib.org |
| Numpy[1] | 1.12.0 | BSD3[3] | http://www.numpy.org |
| pandas[1] | 0.19.2 | BSD3 | http://pandas.pydata.org |
| SciPy[1] | 0.18.1 | BSD3 | https://www.scipy.org |
| PySide | 1.2.2 | LGPLv2.1[4] | https://wiki.qt.io/PySide |
| quickgui | 1.5.6 | MIT[5] | https://pypi.python.org/pypi/quickgui |
| wxPython | 3.0.2.0 | wxWindows-3.1[6] | http://www.wxpython.org |

1 Mandatory libraries
2 http://matplotlib.org/users/license.html
3 https://opensource.org/licenses/BSD-3-Clause
4 https://opensource.org/licenses/LGPL-2.1
5 https://opensource.org/licenses/mit-license.php
6 https://www.wxwidgets.org/about/licence/

[1] AIST, Geological Survey of Japan, Research Institute of Geology and Geoinformation
[*] Corresponding author: A. Noda,Central 7,1-1-1 Higashi, Tsukuba, Ibaraki 305-8567, Japan. Email: a.noda@aist.go.jp

## 2. 2 Data and configuration files

Data files (extension of file name must be .csv) are comma- or tab-separated data sheets that must have at least six columns for isotopic ratios of $^{207}Pb^*/^{235}U$, $^{206}Pb^*/^{238}U$, $^{207}Pb^*/^{206}Pb^*$, and their errors (1σ or 2σ). Optional columns of Th/U ratios and their errors are also acceptable.

Configuration files (.cfg as extension) are needed to run the script, which define many variables used in the script. At first, it is recommended to copy and modify example files in Noda (2016).

## 2. 3 Command-line mode

### 2. 3. 1 Options

The script accepts some options as arguments.

```
Options:
 -h, --help              Show this help message and exit
 -i FILE, --in=FILE      Name of input data file(*.csv)
 -c FILE, --cfg=FILE     Name of configuration file
                         (*.cfg)
 -o FILE, --out=FILE     Name of output file (*.pdf,
                         when DRIVER is pdf)
 -g, --gui               Use GUI for file selection
 -n, --no-gui            Do not use GUI (default)
 -d DRIVER,              Choose driver [pdf (default),
    --driver=DRIVER      qt4agg]
 -f, --force-overwrite   Force overwrite the pre-
                         existing pdf
```

### 2. 3. 2 Running the script

If the script (UPbplot.py), data (data1.csv), and configuration (data1.cfg) files are stored in the current working directory, type like below in a terminal window.

```
python UPbplot.py -i data1.csv
```

Name of the configuration file is assumed to be the same with that of the data file as default, but a different name can be set by using -c option. The following command line means that the script do not use GUI mode (-n), input data file name is data1.csv (-i), configuration file name is all.cfg (-c), pdf is the driver, and the pre-existing pdf file will be overwritten without any notice (-f).

```
python UPbplot.py -n -i data1.csv -c all.cfg -d pdf -f
```

### 2. 3. 3 Output

After the script successfully works in the command-line mode, two types of output (standard output and a pdf file) will be given by this script. The standard output includes the file names, first few lines of the input data, and results of calculation related to generate diagrams (Figure 1). A pdf file has four diagrams (A, B, C and D) in one file. The diagrams A and B are designated for plots of the measured isotopic ratios with error ellipses on the conventional (A) and Tera–Wasserburg (B) concordia diagrams, respectively. Decay constants used in this script are listed in Appendix A.1. Error ellipses for confidence regions of the measurements are illustrated by using covariances between the coordinates in each diagram (Appendix A.2). The script optionally plots two-dimensional (2D) weighted means (Appendix A.5), concordia ages (Appendix A.6), and regression lines (Appendix A.7), and concordia-intercept ages (Appendix A.7) with the confidence regions. Configuration files can set ranges of axes, styles of symbols and lines, labels, and confidence levels (e.g., 68%, 95%, or 99%) for error ellipses, weighted means, concordia ages, and concordia-intercept ages. Several examples in Noda (2016) may be helpful to customize diagrams.

The diagram C shows bar plots of a selected age from among $^{206}Pb^*/^{238}U$, $^{207}Pb^*/^{235}U$, and $^{207}Pb^*/^{206}Pb^*$ ages with the one-dimensional weighted mean (Appendix A.3). The mean square of the weighted deviation (MSWD) can also be calculated (Appendix A.4).

The diagram D is a histogram of the age used in the diagram C with or without kernel density estimations (KDE), which is a way to estimate the probability density function (PDF) of a random variable in a non-parametric way. In this script, the function of stats.gaussian_kde in SciPy is used to obtain the KDE (SciPy.org, 2016). If the input data file has Th/U ratios, they can be plotted in this diagram.

Figure 2 is an example of plots for the Cretaceous granitoids in the Abukuma Highland (Ishihara and Orihashi, 2015). In the diagrams A and B, solid red, dashed blue, and dotted black ellipses represent 95% confidence regions of accepted, discordant (>10%), and manually excluded measurements, respectively. Calculation of the discordance is explained in Appendix A.8. Concordia ages are indicated by solid circles on the concordia curves, which are calculated from the accepted data (red ellipses in this case). Numbers of all ($N$) and accepted ($n$) data points, calculated ages, errors, and MSWD are listed in the upper left sides of the diagrams.

Figure 2C shows the weighted mean (blue line) of $^{206}Pb^*/^{238}U$ ages with the 95% confidence region (shaded band). Red square, gray and open circles with error bars are accepted, discordant, and excluded data, respectively. Figure 2D contains a histogram (left side vertical axis) for the same age with the diagram C, which is stacked by blue (accepted), gray (discordant), and open (excluded) boxes. The Th/U ratios (right side vertical axis) are plotted by the same symbols with those in the diagram C. In addition, kernel-density estimations of all (dashed red) and accepted (solid red) data are also shown in the diagram D.

Figure 3 is another example of the output image for the Cretaceous granitic rocks in the Setouchi area (Iida *et al*., 2015). In addition to Figure 2, the diagrams A and B include confidence regions of 2D weighted means (green ellipses) of the accepted measurements and the regression lines with the errors (thick blue lines with purple zones). Solid black and dashed gray ellipses represent confidence regions of accepted and discordant data, respectively.

```
Terminal — -zsh — 80×44
% UPbplot.py -i Ishihara2015ia_TableS1_68A-64.csv -f
Output file Ishihara2015ia_TableS1_68A-64.pdf already exists.
------------------------------------------------------------
Data filename is Ishihara2015ia_TableS1_68A-64.csv
Configuration filename is Ishihara2015ia_TableS1_68A-64.cfg
Output filename is Ishihara2015ia_TableS1_68A-64.pdf
# Input data (first 5 lines)
------------------------------------------------------------
207Pb/235U  2s         206Pb/238U  2s         207Pb/206Pb  2s
column[9]   column[11] column[6]   column[8] column[3]    column[5]
------------------------------------------------------------
207Pb/235U  1s         206Pb/238U 1s          207Pb/206Pb 1s
------------------------------------------------------------
0.11190     0.00280    0.01664     0.00034    0.04880     0.00070
0.11770     0.00295    0.01725     0.00035    0.04950     0.00070
0.11380     0.00265    0.01681     0.00034    0.04910     0.00060
0.11740     0.00395    0.01753     0.00034    0.04860     0.00135
0.18110     0.00410    0.01871     0.00036    0.07020     0.00085
------------------------------------------------------------
Discordant data (> 10.00%) are excluded from analysis.
Discordance is calculated by 100*(1-([206Pb/238U age]/[207Pb/235U age])
Discordant data points are
4 29.29%
6 54.01%
12 11.12%
20 54.07%
22 26.21%
Manually excluded data points are [13, 19]
Accepted data points are [ 0  1  2  3  5  7  8  9 10 11 14 15 16 17 18 21 23]
--------------------------------------------------
Plotting A: Conventional concordia diagram
Error ellipses are 95% for data points
Concordia age = 109.70 ± 1.04 Ma [95% conf.]
    (MSWD of concordance=17.78, p(chi^2)=0.00)
Plotting B: Tera-Wasserburg concordia diagram
Concordia age = 109.70 ± 1.04 Ma [95% conf.]
    (MSWD of concordance=17.23, p(chi^2)=0.00)
Plotting C: One-dimensional bar plot
1D weighted mean age = 109.72 ± 1.05 Ma [95% conf.] (MSDW=2.34)
Plotting D: Histogram
All done.
Saving Ishihara2015ia_TableS1_68A-64.pdf
%
```

Figure 1 An example of standard output in a terminal window. This output includes file names to be processed, a part of the input data, and results of calculation. In this case, the script was applied to the data of sample 68A-64 in Ishihara and Orihashi (2015).

Other symbols and lines are the same with those in Figure 2.

### 2. 4 GUI mode

To run the script as the GUI mode, add -g option in the command line. A new window will be opened to select an input data file. Please notice that the GUI mode cannot accept any command-line options related with file names (-i, -c, and -o) at the present version of the script. After the selection of input data and/or configuration files, a message window will pop up to show the calculation process.

If PySide is installed and the user's matplotlib has a backend of qt4agg, the script accepts an option of "-d qt4agg" as the driver.

```
python UPbplot.py -g -d qt4agg
```

This driver enables us to modify ranges of axes, labels, and styles of lines, interactively. To output images in the GUI mode, press "Save" bottom in the toolbar. Several image formats, such as png, jpg, and tiff, can be chosen in the save dialog, in addition to pdf. Figure 4 shows main, preference, and message windows of the GUI mode for the same data of Figure 2.

## 3. Summary

In this paper, I introduced usages and examples of the script, UPbplot.py, which had developed in order to offer a new tool for calculation and visualization of U-Pb age data. It is an operating system-independent software and can deal with the conventional and Tera–Wasserburg diagrams with error ellipses of arbitrary confidence levels. It can calculate and plot the one- and two-dimensional weighted means, MSWDs, concordia ages, and concordia-intercept ages. This script provides an alternative tool to a well-known Visual Basic® add-in program for Microsoft® Excel® "Isoplot".

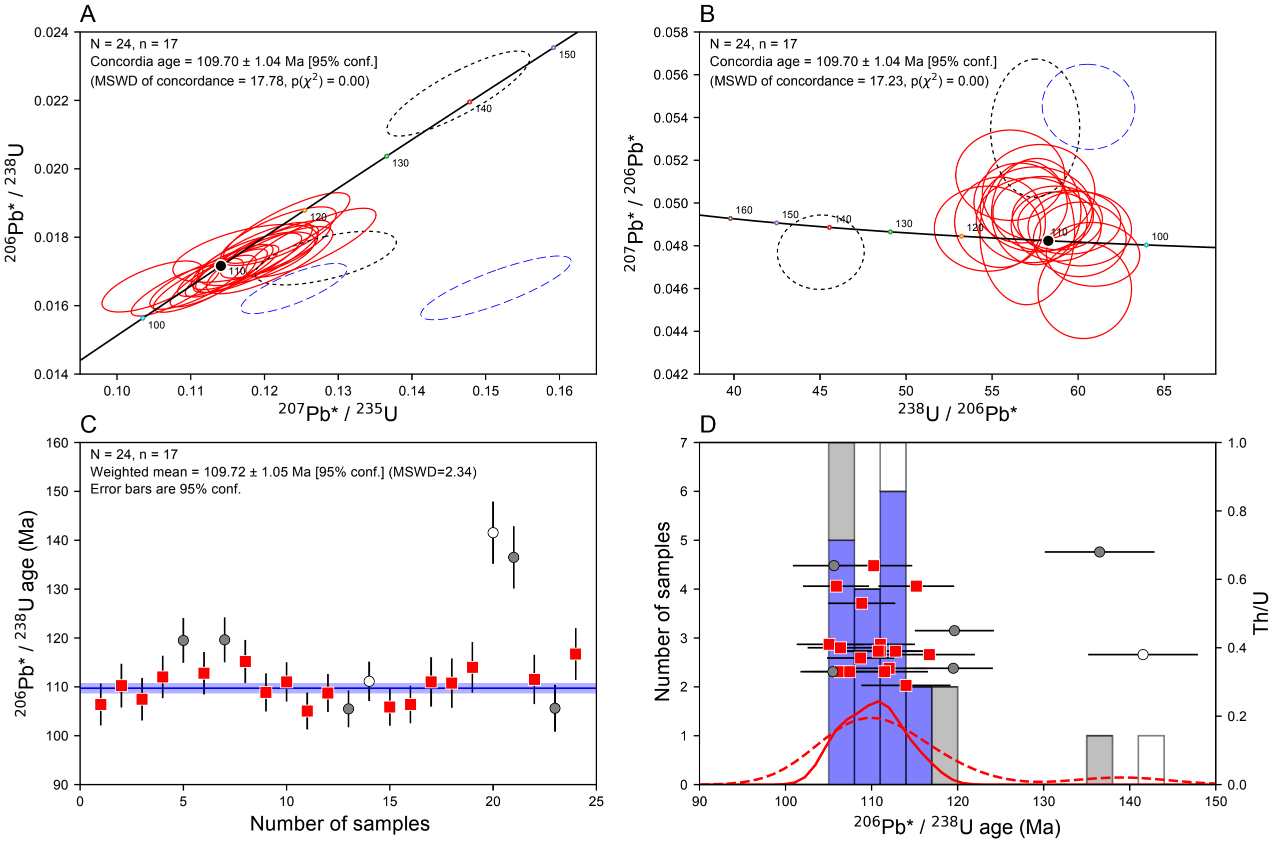


Figure 2 An example of output image derived from the same data with those of Figure 1. A: Conventional concordia diagram with error ellipses for 95% confidence regions of the measurements. Solid red, dashed blue, and dotted black ellipses are accepted, discordant, and manually excluded measurements from the calculation, respectively. A solid black circle on the concordia curve indicates the concordia age. *N* and *n* indicate numbers of total and accepted measurements, respectively. B: Tera–Wasserburg concordia diagram for the same legend with A. C: Bar plots of $^{206}Pb^*/^{238}U$ ages and their weighted mean (blue line) with its 95% confidence region (shaded band). Red squares, gray circles, and open circles with error bars are accepted, discordant, and excluded measurements, respectively. D: Histogram of $^{206}Pb^*/^{238}U$ age (left side of the vertical axis). Blue, gray, and, open boxes are accepted, discordant, and excluded measurements, respectively. Kernel density estimates of all (dashed red line) and accepted (solid red line) measurements are also illustrated. Th/U ratios (right side of the vertical axis) are indicated by the same symbol with C.

## Appendix

### A. 1 Decay constants

The following decay constants are used in the script, based on Jaffey *et al*. (1971), Steiger and Jäger (1977) and Hiess *et al*. (2012) (see Schoene, 2014).

$$\lambda_{^{238}U} = 1.55125 \times 10^{-10} \quad [\text{year}^{-1}] \qquad \pm 0.107\% \quad (2\sigma)$$

$$\lambda_{^{235}U} = 9.8485 \times 10^{-10} \quad [\text{year}^{-1}] \qquad \pm 0.137\% \quad (2\sigma)$$

$$\lambda_{^{232}Th} = 4.9475 \times 10^{-11} \quad [\text{year}^{-1}] \qquad \pm \sim 1\% \quad (2\sigma)$$

$$\frac{^{238}U}{^{235}U} = 137.818 \qquad \pm 0.045 \quad (2\sigma)$$

### A. 2 Concordia diagrams

Coordinates of $X$ and $Y$ for the conventional concordia diagram (Wetherill, 1956) are written as functions of time $t$,

$$X = \frac{^{207}Pb^*}{^{235}U} = \exp(\lambda_{^{235}U} t) - 1 \qquad (A1)$$

$$Y = \frac{^{206}Pb^*}{^{238}U} = \exp(\lambda_{^{238}U} t) - 1 \qquad (A2)$$

where asterisks denote radiogenic components. Coordinates of $x$ and $y$ for the Tera–Wasserburg concordia diagram (Tera and Wasserburg, 1972) can be expressed by using eqs. (A1)–(A2) and the constants in Appendix A.1.

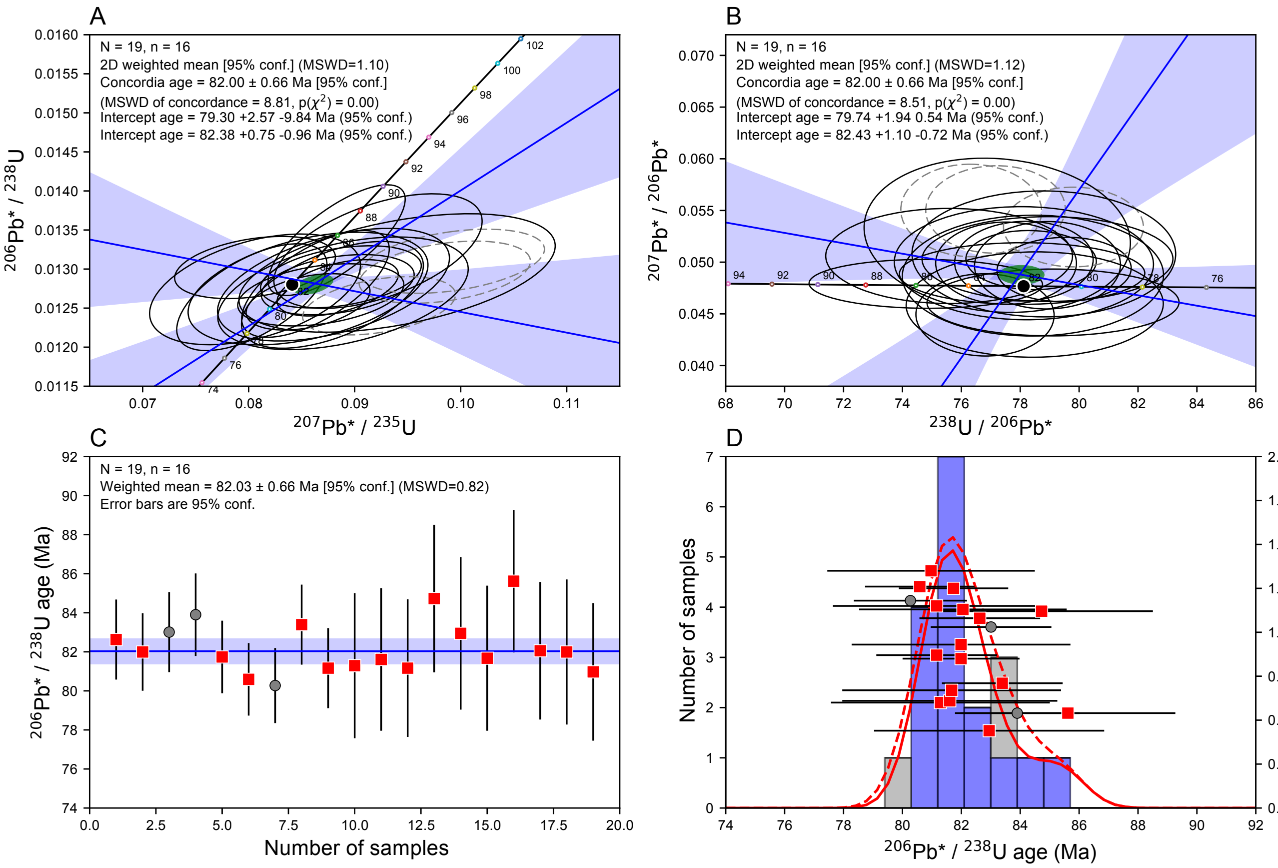


Figure 3 Another example of output image. A: Conventional concordia diagram with 95% confidence regions of the measurements. Solid and gray dashed ellipses are accepted and discordant data, respectively. Two-dimensional weighted mean (solid green ellipse), concordia age (solid black circle), two regression lines (thick blue lines) with the errors (shaded areas) are also plotted. B: Tera–Wasserburg concordia diagram for the same legend with A. C: Bar plots of $^{206}Pb^*/^{238}U$ ages. Symbols are same with Figure 2C. D: Histogram of $^{206}Pb^*/^{238}U$ age and Th/U ratios. Symbols are same with Figure 2D. The data come from sample 080122I01 in Iida *et al*. (2015).

$$x = \frac{^{238}\mathrm{U}}{^{206}\mathrm{Pb}^*} = \frac{1}{\exp(\lambda_{238\mathrm{U}}t) - 1} = \frac{1}{Y} \tag{A3}$$

$$y = \frac{^{207}\mathrm{Pb}^*}{^{206}\mathrm{Pb}^*} = \frac{^{235}\mathrm{U}}{^{238}\mathrm{U}} \frac{\exp(\lambda_{235\mathrm{U}}t) - 1}{\exp(\lambda_{238\mathrm{U}}t) - 1} = \frac{1}{137.818} \frac{X}{Y}. \tag{A4}$$

Because both coordinates of $X$ and $Y$ ($x$ and $y$) are not independent each other, the covariance of $X$ and $Y$ ($x$ and $y$) is needed to obtain confidence regions of the measurements, which is defined as

$$\mathrm{cov}(X, Y) \equiv \sigma_X \sigma_Y \rho_{XY} \tag{A5}$$

$$\mathrm{cov}(x, y) \equiv \sigma_x \sigma_y \rho_{xy} \tag{A6}$$

where $\sigma_X$ and $\rho_{XY}$ mean the standard deviation of $X$ and the error correlation between $X$ and $Y$, respectively.

For the calculation of the error correlations in eqs. (A5) and (A6), error propagation (Taylor, 1997) should be considered. As for $x = 1/Y$ from eq. (A3) and $y = uX/Y$ ($u$ is a constant) from eq. (A4),

$$\sigma_x^2 = \left(\frac{\partial y}{\partial Y}\right)^2 \sigma_Y^2 = x^2 S_Y^2 \tag{A7}$$

$$\sigma_y^2 = \left(\frac{\partial y}{\partial X}\right)^2 \sigma_X^2 + \left(\frac{\partial y}{\partial Y}\right)^2 \sigma_Y^2 + 2\frac{\partial y}{\partial X}\frac{\partial y}{\partial Y}\mathrm{cov}(X, Y)$$
$$= y^2\left(S_X^2 + S_Y^2 - 2S_X S_Y \rho_{XY}\right) \tag{A8}$$

where $S_X = \sigma_X/X$ and $S_Y = \sigma_Y/Y$ are relative errors in the measures. Then, eqs. (A7) and (A8) yield relative errors of

$$S_x^2 = S_Y^2 \tag{A9}$$

$$S_y^2 = S_X^2 + S_Y^2 - 2S_X S_Y \rho_{XY} \tag{A10}$$

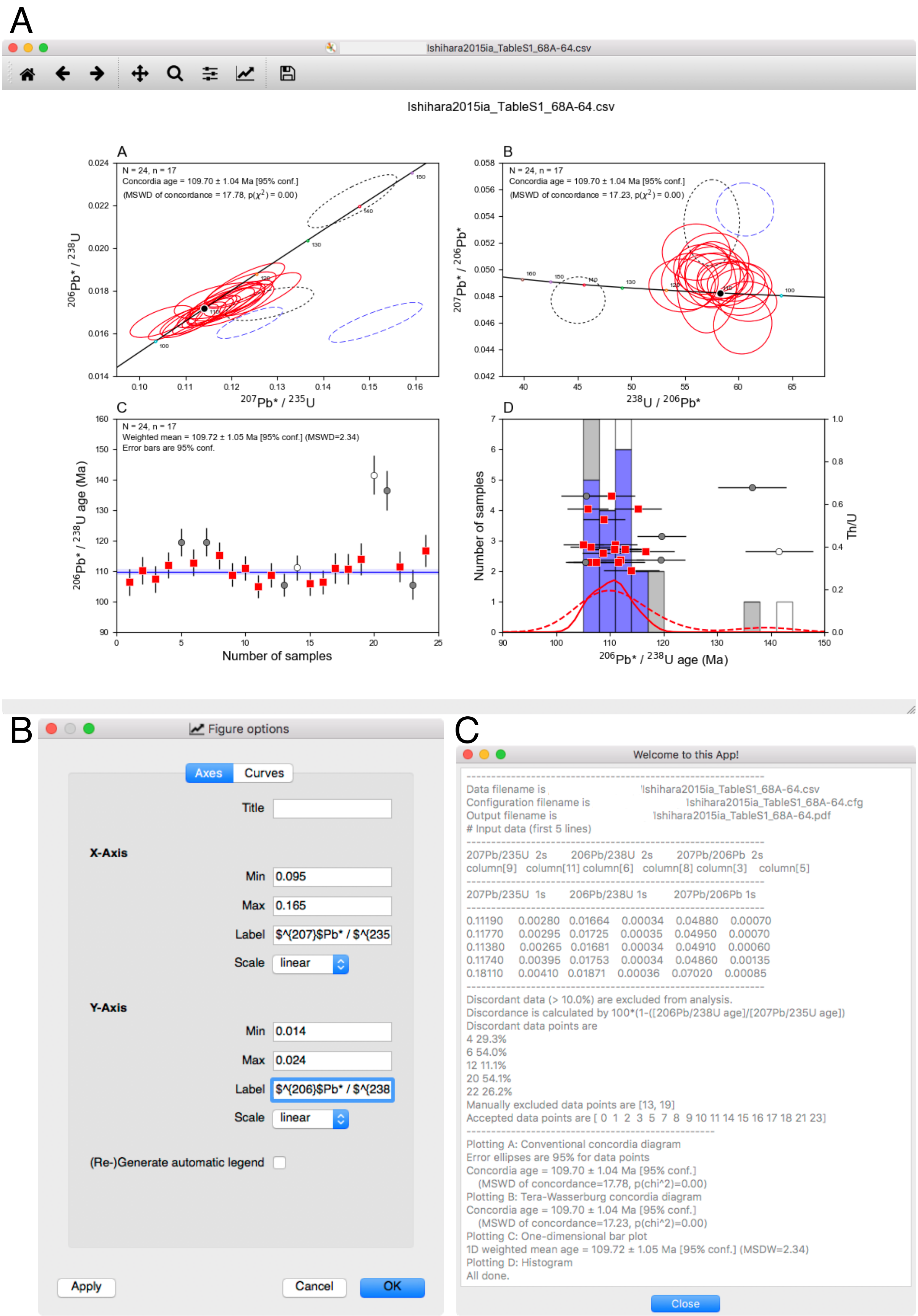


Figure 4 A collection of screenshots of the GUI mode. A: Main window of the diagrams. B: Preference panel to control ranges of axes and labels in the diagrams. C: Message window showing result of calculation. The U-Pb data are same with Figure 1.

where $S_x = \sigma_x/x$ and $S_y = \sigma_y/y$. For the inverse case of eq. (A8), $X = y/(ux)$, we obtain

$$S_X^2 = S_x^2 + S_y^2 - 2S_x S_y \rho_{xy}. \tag{A11}$$

Therefore, the error correlations can be calculated from eqs. (A9)–(A11),

$$\rho_{XY} = \frac{S_X^2 + S_Y^2 - S_y^2}{2S_X S_Y} \tag{A12}$$

$$\rho_{xy} = \frac{S_Y^2 - S_X S_Y \rho_{XY}}{S_x S_y}. \tag{A13}$$

Please notice that eq. (A13) is different from the equation of $\rho_{xy}$ in p. 27 of Ludwig (2012).

By using the covariances of eqs. (A5)–(A6) with error correlations of eqs. (A12)–(A13), we can draw error ellipses showing the two-dimensional "normally distributed" areas with certain confidences, especially 68% (≈1σ) and 95% (≈2σ).

### A. 3 One-dimensional weighted mean

Weighted mean (weighted average) is a kind of mean with consideration of errors. For the one-dimensional case, the weighted mean is written as

$$\overline{X} = \sum_{i=1}^{N} \omega(X_i) X_i$$

where $X_i$ is the measurement of the $i$th component of a total of $N$ measurements, and $\omega(X_i)$ is weight of each measurement. The weight is an inverse square of the standard deviation $\sigma_{Xi}$ divided by sum of them,

$$\omega(X_i) = (\sigma_{X_i})^{-2} \Big/ \sum_{i=1}^{N} (\sigma_{X_i})^{-2}.$$

Variance of the weighted mean is

$$\sigma_{\overline{X}}^2 = 1 \Big/ \sum_{i=1}^{N} (\sigma_{X_i})^{-2}$$

when each of the measurements are independent (McLean *et al.*, 2011).

If we use 95% confidence region of the weighted mean $\overline{X}$, it can be simply calculated by using $\sigma_{\overline{X}}$ times 1.96 (Student's $t$ for an infinite sample size).

### A. 4 Mean square of the weighted deviation (MSWD)

Mean square of the weighted deviation (MSWD) was originally developed for statistical evaluation of a regression line (e.g.,Wendt and Carl, 1991), which indicates how well the line describes the data $X_i$ . When a value of MSWD is less than 1, the observed deviations from the regression line are considered within analytical error. Otherwise, if it is more than 3, interpretation of such line may be doubtful. It is written as

$$\text{MSWD} = S / df \tag{A14}$$

where $S$ is sum of "residual" that is sum of distances from each data point to the regression line (to the weighted mean for the one-dimensional case),

$$S = \sum_{i=1}^{N} \frac{(X_i - \overline{X})^2}{\sigma_i^2} \tag{A15}$$

and $df$ is degree of freedom, $N-1$.

### A. 5 Two-dimensional weighted mean

The two-dimensional weighted mean ($\overline{X}$, $\overline{Y}$) can be obtained by minimizing the sum of the squares of the $N$ error weighted residuals (Ludwig, 1998) as,

$$S = \sum_{i=1}^{N} v_i^T \Omega_i v_i = \sum_{i=1}^{N} \left( R_i^2 \Omega_i^{11} + r_i^2 \Omega_i^{22} + 2R_i r_i \Omega_i^{12} \right) \tag{A16}$$

where $v_i$ is a vector of residuals written as

$$v_i = \begin{pmatrix} X_i - \overline{X} \\ Y_i - \overline{Y} \end{pmatrix} = \begin{pmatrix} R_i \\ r_i \end{pmatrix} \tag{A17}$$

and $\Omega_i$ is

$$\Omega_i = \begin{pmatrix} \Omega_i^{11} & \Omega_i^{12} \\ \Omega_i^{12} & \Omega_i^{22} \end{pmatrix} = \begin{pmatrix} \sigma_{X_i}^2 & \text{cov}(X_i, Y_i) \\ \text{cov}(X_i, Y_i) & \sigma_{Y_i}^2 \end{pmatrix}^{-1}. \tag{A18}$$

Solving eqs. (A16)–(A18) gives ($\overline{X}$, $\overline{Y}$) that minimize $S$ as

$$\overline{X} = \frac{\sum_{i=1}^{N} \Omega_i^{22} \sum_{i=1}^{N} (X_i \Omega_i^{11} + Y_i \Omega_i^{12}) - \sum_{i=1}^{N} \Omega_i^{12} \sum_{i=1}^{N} (Y_i \Omega_i^{22} + X_i \Omega_i^{12})}{\sum_{i=1}^{N} \Omega_i^{11} \sum_{i=1}^{N} \Omega_i^{22} - \left( \sum_{i=1}^{N} \Omega_i^{12} \right)^2} \tag{A19}$$

$$\overline{Y} = \frac{\sum_{i=1}^{N} \Omega_i^{11} \sum_{i=1}^{N} (Y_i \Omega_i^{22} + X_i \Omega_i^{12}) - \sum_{i=1}^{N} \Omega_i^{12} \sum_{i=1}^{N} (X_i \Omega_i^{11} + Y_i \Omega_i^{12})}{\sum_{i=1}^{N} \Omega_i^{11} \sum_{i=1}^{N} \Omega_i^{22} - \left( \sum_{i=1}^{N} \Omega_i^{12} \right)^2} \tag{A20}$$

Standard deviations of the means ($\sigma_{\overline{X}}$ and $\sigma_{\overline{Y}}$) are derived from eq. (9) in Ludwig (1998),

$$\sigma_{\overline{X}} = \sqrt{\frac{\sum_{i=1}^{N}\Omega_i^{22}}{\sum_{i=1}^{N}\Omega_i^{11}\sum_{i=1}^{N}\Omega_i^{22} - \left(\sum_{i=1}^{N}\Omega_i^{12}\right)^2}} \quad \text{(A21)}$$

$$\sigma_{\overline{Y}} = \sqrt{\frac{\sum_{i=1}^{N}\Omega_i^{11}}{\sum_{i=1}^{N}\Omega_i^{11}\sum_{i=1}^{N}\Omega_i^{22} - \left(\sum_{i=1}^{N}\Omega_i^{12}\right)^2}}. \quad \text{(A22)}$$

The MSWD can be calculated by eq. (A14) with $S$ of eq. (A16) and the degree of freedom, $df = 2N - 2$.

## A. 6 Concordia ages

### A. 6. 1 Conventional concordia curve

A point on the conventional concordia curve ($X_c$, $Y_c$) is expressed as a function of time $t$ using eqs. (A1) and (A2). The vector of residuals $v_i$ represents the difference between ($X_c$, $Y_c$) and each measurement ($X_i$, $Y_i$),

$$v_i = \begin{pmatrix} X_i - X_c \\ Y_i - Y_c \end{pmatrix} = \begin{pmatrix} R_i \\ r_i \end{pmatrix}. \quad \text{(A23)}$$

The best $t$ can be obtained, when the least sum of the weighted squared residuals of eq. (A16) with eqs. (A18) and (A23) is minimum. As mentioned in Ludwig (1998), the two-dimensional weighted mean of eqs. (A19)–(A20) and the standard deviation of eqs. (A21)–(A22) are practically used to calculate the minimum $S$ and $t$ for ($X_c$, $Y_c$). The least square method (the function of optimize. leastsq in SciPy) is applied in this script to calculate $t$ to minimize $S$.

The variance in $t$ is

$$\sigma_t^2 = \left(Q_{235}^2\Omega^{11} + Q_{238}^2\Omega^{22} + 2Q_{235}Q_{238}\Omega^{12}\right)^{-1}$$

where

$$Q_{235} = \lambda_{235\mathrm{U}}\exp(\lambda_{235\mathrm{U}}t)$$

$$Q_{238} = \lambda_{238\mathrm{U}}\exp(\lambda_{238\mathrm{U}}t)$$

$$\Omega^{11} = \sigma_{\overline{Y}}^2 \Big/ \left(\sigma_{\overline{X}}^2\sigma_{\overline{Y}}^2 - \mathrm{cov}(\overline{X},\overline{Y})^2\right)$$

$$\Omega^{22} = \sigma_{\overline{X}}^2 \Big/ \left(\sigma_{\overline{X}}^2\sigma_{\overline{Y}}^2 - \mathrm{cov}(\overline{X},\overline{Y})^2\right)$$

$$\Omega^{12} = -\mathrm{cov}(\overline{X},\overline{Y}) \Big/ \left(\sigma_{\overline{X}}^2\sigma_{\overline{Y}}^2 - \mathrm{cov}(\overline{X},\overline{Y})^2\right)$$

and

$$\begin{pmatrix} \sigma_{\overline{X}}^2 & \mathrm{cov}(\overline{X},\overline{Y}) \\ \mathrm{cov}(\overline{X},\overline{Y}) & \sigma_{\overline{Y}}^2 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{N}\Omega_i^{11} & \sum_{i=1}^{N}\Omega_i^{12} \\ \sum_{i=1}^{N}\Omega_i^{12} & \sum_{i=1}^{N}\Omega_i^{22} \end{pmatrix}^{-1}.$$

Three MSWDs are considered for the concordia ages, which include the MSWD for X–Y equivalence that is sum of $S$ ($X_i$, $Y_i$) divided by $2N - 1$, the MSWD for concordance that is $S$ ($\overline{X}$, $\overline{Y}$), and the MSWD for combined equivalence and concordance (p. 667 in Ludwig, 1998). The script outputs one of the MSWDs according to the settings in the configuration file.

### A. 6. 2 Tera–Wasserburg concordia curve

Transformed ratios of ($x_i$, $y_i$), errors ($\sigma_{x_i}$, $\sigma_{y_i}$) and error correlation ($\rho_{xy}$) from the Tera–Wasserburg coordinate to the conventional coordinate ($X_i$, $Y_i$) are used to solve the concordia age on the Tera–Wasserburg curve (Ludwig, 1998). The best $t$ and its standard deviation $\sigma_t$ can be obtained by substituting $X_i = 137.818 y_i / x_i$ and $Y_i = 1/x_i$ from eqs. (A3)–(A4) with eqs. (A9), (A11), and (A13) into the procedures of the conventional concordia age (Appendix A.6. 1).

## A. 7 Regression line by Maximum Likelihood Estimate (MLE)

The best fit line ($X'$, $Y'$) with the uncertainty can be expressed as Ludwig (1980),

$$Y' = a + bX' \pm \sqrt{\sigma_a^2 + \sigma_b^2 X'(X' - 2\overline{X})} \quad \text{(A24)}$$

where $a$ and $b$ are intercept and slope, respectively, $\overline{X}$ is the centroid of measured data $X$, which can be calculated by the method of maximum likelihood estimate (York, 1969; Titterington and Halliday, 1979) of

$$\overline{X} = \sum_{i=1}^{N} Z_i X_i \Big/ \sum_{i=1}^{N} Z_i$$

$$\overline{Y} = \sum_{i=1}^{N} Z_i Y_i \Big/ \sum_{i=1}^{N} Z_i$$

where $Z_i$ is

$$Z_i = \frac{\omega(X_i)\omega(Y_i)}{b^2\omega(Y_i) + \omega(X_i) - 2b\rho_{XY}\sqrt{\omega(X_i)\omega(Y_i)}}$$

with

$$\omega(X_i) = \sigma_{X_i}^{-2},\ \omega(Y_i) = \sigma_{Y_i}^{-2},$$

and

$$\rho_{XY} = \mathrm{cov}(X_i, Y_i)/\sigma_{X_i}\sigma_{Y_i}.$$

The best-fit line can be obtained, when

$$S = \sum_{i=1}^{N} Z_i (Y_i - bX_i - a)^2$$

is minimized. Because this equation cannot be solved explicitly, the MLE method is used to calculate the best $b$,

$$b = \frac{-B \pm \sqrt{B^2 + 4AC}}{2A}$$

where

$$A = \sum_{i=1}^{N} Z_i^2 \left( \frac{(X_i - \overline{X})(Y_i - \overline{Y})}{\omega(X_i)} - \frac{\rho_{XY}(X_i - \overline{X})^2}{\sqrt{\omega(X_i)\omega(Y_i)}} \right)$$
$$B = \sum_{i=1}^{N} Z_i^2 \left( \frac{(X_i - \overline{X})^2}{\omega(Y_i)} - \frac{(Y_i - \overline{Y})^2}{\omega(X_i)} \right)$$
$$C = \sum_{i=1}^{N} Z_i^2 \left( \frac{(X_i - \overline{X})(Y_i - \overline{Y})}{\omega(Y_i)} - \frac{\rho_{XY}(Y_i - \overline{Y})^2}{\sqrt{\omega(X_i)\omega(Y_i)}} \right).$$

Therefore, $b$ has two solutions, meaning two lines are obtained.

The variances of $a$ and $b$ are given by

$$\sigma_a^2 = \sum_{i=1}^{N} X_i^2 Z_i \Big/ \sum_{i=1}^{N} Z_i \sum_{i=1}^{N} (X_i - \overline{X})^2 Z_i$$
$$\sigma_b^2 = 1 \Big/ \sum_{i=1}^{N} (X_i - \overline{X})^2 Z_i.$$

The concordia-intercept age is the age ($t$) at which the regression line of eq. (A24) intersects the concordia curve, meaning $|Y_c - Y'|$ is zero.

### A. 8 Data rejection

Discordant data can be excluded from calculation of the weighted mean, concordia, and concordia-intercept ages. The script can choose one from among the following three methods to calculate the discordance (%).

$$\text{Discordance} = \left(1 - \frac{^{206}\mathrm{Pb}^*/^{238}\mathrm{U\ age}}{^{207}\mathrm{Pb}^*/^{206}\mathrm{Pb}^*\ \mathrm{age}}\right) \times 100,$$
$$= \left(1 - \frac{^{207}\mathrm{Pb}^*/^{235}\mathrm{U\ age}}{^{207}\mathrm{Pb}^*/^{206}\mathrm{Pb}^*\ \mathrm{age}}\right) \times 100, \text{or}$$
$$= \left(1 - \frac{^{206}\mathrm{Pb}^*/^{238}\mathrm{U\ age}}{^{207}\mathrm{Pb}^*/^{235}\mathrm{U\ age}}\right) \times 100$$

where

$$\frac{^{207}\mathrm{Pb}^*}{^{235}\mathrm{U}}\ \mathrm{age} = \frac{1}{\lambda_{^{235}\mathrm{U}}} \log\left(\frac{^{207}\mathrm{Pb}^*}{^{235}\mathrm{U}} + 1\right)$$
$$\frac{^{206}\mathrm{Pb}^*}{^{238}\mathrm{U}}\ \mathrm{age} = \frac{1}{\lambda_{^{238}\mathrm{U}}} \log\left(\frac{^{206}\mathrm{Pb}^*}{^{238}\mathrm{U}} + 1\right).$$

$^{207}\mathrm{Pb}^*/^{206}\mathrm{Pb}^*$age is the time $t$, when

$$\left| \frac{^{238}\mathrm{U}}{^{235}\mathrm{U}} \frac{^{207}\mathrm{Pb}^*}{^{206}\mathrm{Pb}^*} - \frac{\exp(\lambda_{^{235}\mathrm{U}} t) - 1}{\exp(\lambda_{^{238}\mathrm{U}} t) - 1} \right| = 0.$$

## References


Bowring, J. and PI CIRDLES.org Open Source Development Team (2016) Topsoil: A community driven replacement for ISOPLOT. Apache License, Version 2.0, https://github.com/CIRDLES/topsoil, Topsoil v1.0.0-alpha.10 (accessed 2016-04-05).

Hiess, J., Condon, D. J., McLean, N. and Noble, S. R. (2012) $^{238}$U/$^{235}$U systematics in terrestrial Uranium-bearing minerals. *Science*, **335**, 1610–1614, doi: 10.1126/science.1215507.

Hunter, J. D. (2007) Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.*, **9**, 90–95, doi: 10.5281/zenodo.61948.

Iida, K., Iwamori, H., Orihashi, Y., Park, T., Jwa, Y.-J., Kwon, S.-T., Danhara, T. and Iwano, H. (2015) Tectonic reconstruction of batholith formation based on the spatiotemporal distribution of Cretaceous–Paleogene granitic rocks in southwestern Japan. *Island Arc*, **24**, 205–220, doi: 10.1111/iar.12103.

Ishihara, S. and Orihashi, Y. (2015) Cretaceous granitoids and their zircon U–Pb ages across the south-central part of the Abukuma Highland, Japan. *Island Arc*, **24**, 159–168, doi: 10.1111/iar.12105.

Jaffey, A. H., Flynn, K. F., Glendenin, L. E., Bentley, W. C. and Essling, A. M. (1971) Precision measurement of half-lives and specific activities of $^{235}$U and $^{238}$U. *Phys. Rev. C*, **4**, 1889–1906, doi: 10.1103/PhysRevC.4.1889.

Jones, E., Oliphant, T., Peterson, P. *et al*. (2001–) SciPy: Open source scientific tools for Python. http://www.scipy.org/ (accessed 2016-09-16).

Ludwig, K. R. (2012) *Isoplot 3.75: A geochronological toolkit for Microsoft Excel*, Spec. Publ., no. 5, Berkeley Geochronology Center, Berkeley, California, 75p.

Ludwig, K. R. (1980) Calculation of uncertainties of U–Pb isotope data. *Earth Planet. Sci. Lett.*, **46**, 212–220, doi: 10.1016/0012-821X(80)90007-2.

Ludwig, K. R. (1998) On the treatment of concordant Uranium-Lead ages. *Geochim. Cosmochim. Acta*, **62**, 665–676, doi: 10.1016/S0016-7037(98)00059-3.

McKinney, W. (2010) Data structures for statistical computing in Python. *In* van der Walt, S. and

Millman, J. eds. *Proceedings of the 9th Python in Science Conference*, 51–56.
McLean, N. M., Bowring, J. F. and Bowring, S. A. (2011) An algorithm for U–Pb isotope dilution data reduction and uncertainty propagation. *Geochem. Geophys. Geosyst*., **12**, Q0AA18, doi: 10.1029/2010GC003478.
Noda, A. (2016) *UPbplot.py: A python script for U–Pb age data analysis*, Open-File Report, no. 634, Geol. Surv. Japan, AIST.
Schoene, B. (2014) U–Th–Pb Geochronology. *In* Rudnick, R. L. ed. *The Crust*. Treatise on Geochemistry, 2nd ed., **4**, Ch. 4.10, 341–378, Elsevier, Oxford, doi: 10.1016/B978-0-08-095975-7.00310-7.
SciPy.org (2016) scipy.stats.gaussian_kde. https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.gaussian_kde.html (accessed 2016-11-14).
Steiger, R. H. and Jäger, E. (1977) Subcommission on geochronology: Convention on the use of decay constants in geo- and cosmochronology. *Earth Planet. Sci. Lett*., **36**, 359–362, doi: 10.1016/0012-821X(77)90060-7.
Taylor, J. R. (1997) *An Introduction to Error Analysis, 2nd ed*., Ch. 9, 209–226, University Science Books, Sausalito, California, United States.
Tera, F. and Wasserburg, G. J. (1972) U–Th–Pb systematics in three Apollo 14 basalts and the problem of initial Pb in lunar rocks. *Earth Planet. Sci. Lett*., **14**, 281–304, doi: 10.1016/0012-821X(72)90128-8.
Titterington, D. M. and Halliday, A. N. (1979) On the fitting of parallel isochrons and the method of maximum likelihood. *Chem. Geol*., **26**, 183–195, doi: 10.1016/0009-2541(79)90045-7.
Wendt, I. and Carl, C. (1991) The statistical distribution of the mean squared weighted deviation. *Chem. Geol.*, **86**, 275–285, doi: 10.1016/0168-9622(91)90010-T.
Wetherill, G. W. (1956) Discordant uranium-lead ages. *Trans. Amer. Geophys. Union*, **37**, 320–326.
York, D. (1969) Least squares fitting of a straight line with correlated errors. *Earth Planet. Sci. Lett*., **5**, 320–324, doi: 10.1016/S0012-821X(68)80059-7.

Recieved April 28, 2016
Accepted January 23, 2017

## U–Pb 年代データのための新しい計算・可視化ツールの開発：UPbplot.py

野田　篤

### 要　旨

本稿は，U–Pb年代データの計算・可視化のために新しく開発したスクリプト（UPbplot.py）の使用方法と使用例及び数学的背景についての解説である．このスクリプトは，U–Pb年代値の1次元または2次元の加重平均，標準（Wetherill）及び Tera–Wasserburg コンコーディア図におけるコンコーディア年代・コンコーディア曲線とのインターセプト年代を求めるための関数を含んでおり，それらの計算結果やコンコーディア図・棒グラフ・ヒストグラムなどのグラフを出力することができる．