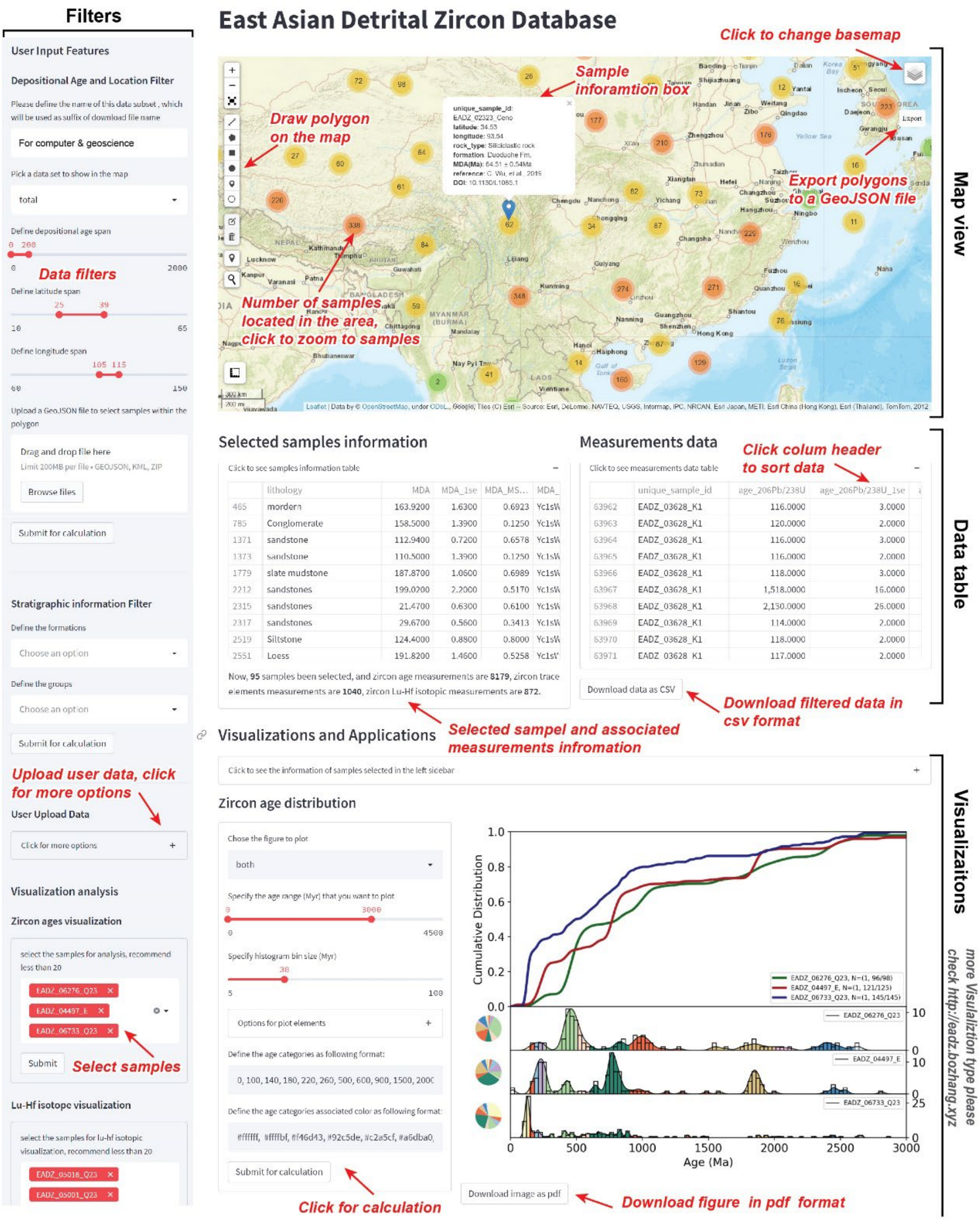
EaDz

A web-based, relational database for detrital zircons from East Asia



Bo Zhang (bzhang@cugb.edu.cn)

Shaofeng Liu (shaofeng@cugb.edu.cn)

China University of Geosciences (Beijing)

Contents

The EaDz website consists of a sidebar on the left and a content area on the right. The content area is further divided into a map view section, a data table section, and a visualization section according to its functionality. The left sidebar serves as global parameters holder and sample selection filters; the right content area serves as an interactive data visualization panel, where you can further fine-tune how the data/figures look with parameters/options specific to each module or analysis tool. All these parameters/options are quite self-explanatory.

1 Data query, download, and upload

With online Map, EaDz allows users to query and select data by different criteria with a user-friendly interface. The selected data can be easily downloaded in CSV format (Fig. 1). The left sidebar provides a portal for users to upload their own data (Fig. 2).

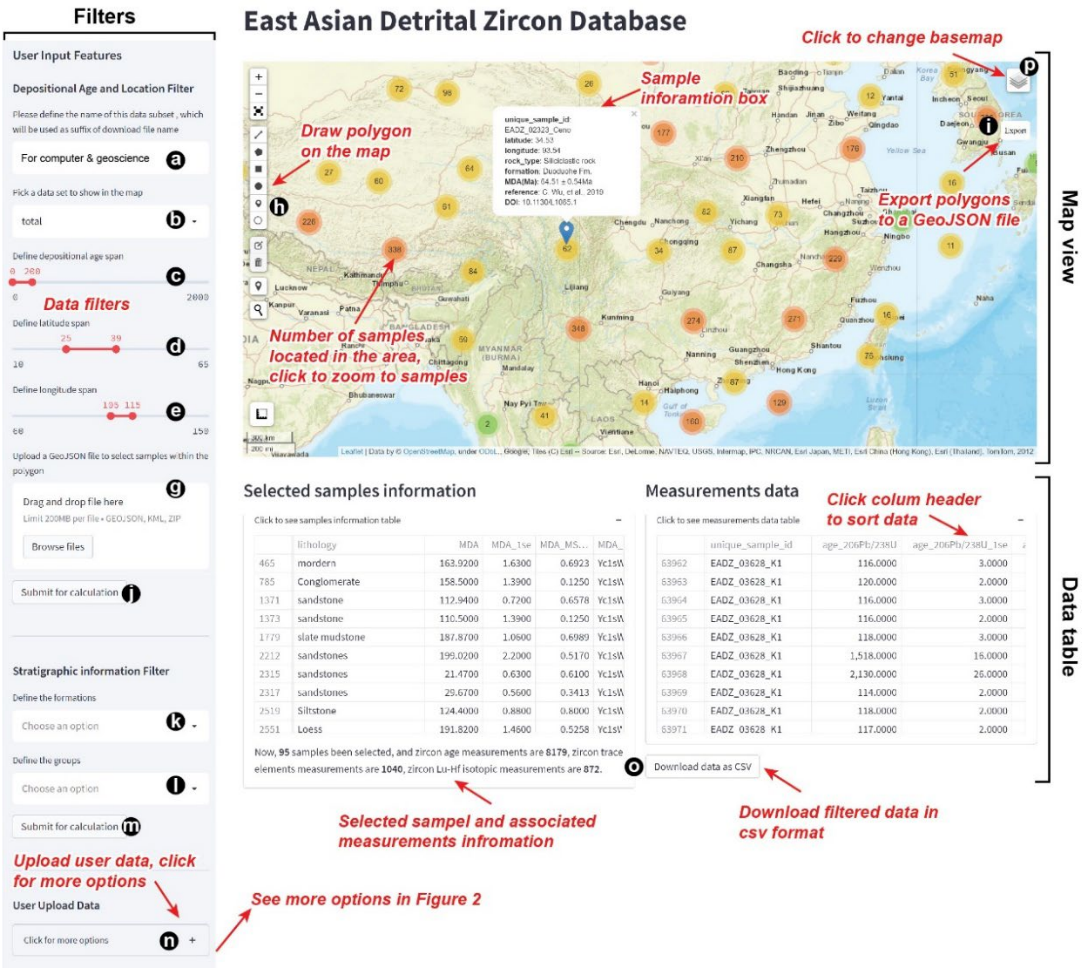


Figure 1. Screenshot of the EaDz website, showing the data query, download, and upload functions

1.1 Data query, download

a. Set name of this data subset, e.g., project name, which will be used as suffix of the downloaded file name. This option can be left empty (Fig. 1a).

b. Set how many samples are to be displayed in the interactive map. There are two options, ***"total"*** will plot all samples from the EaDz database, whereas "selected" will only plot the samples selected by the user. The default value is ***"total"*** (Fig. 1b).

c-e. Sample filters with selection slider, **c** for depositional age, **d** for latitude, and **e** for longitude (Fig. 1c-e).

g-i. Sample selection with user-defined polygons. Step 1, draw a polygon on the map using the drawing tools provided on the left side of the map view section (Fig. 1h); Step 2, Export this polygon as a GeoJSON file by clicking the "Export" button in the map view (Fig. 1i); Step 3, Upload the GeoJSON file though the file uploading portal. The GeoJSON file generated from other applications can also be uploaded here (Fig. 1g).

j. Click to search for samples based on the specified criteria. The results will display in the "Selected sample information" and "Measurements data" tables. If you set it to plot only the selected samples on the map, the map view will also be updated.

k-m. Options to further filter data based on Stratigraphic information, **k** is the host formations of samples, and **l** is the host groups of samples. Click on "Submit for calculation" to run the search task.

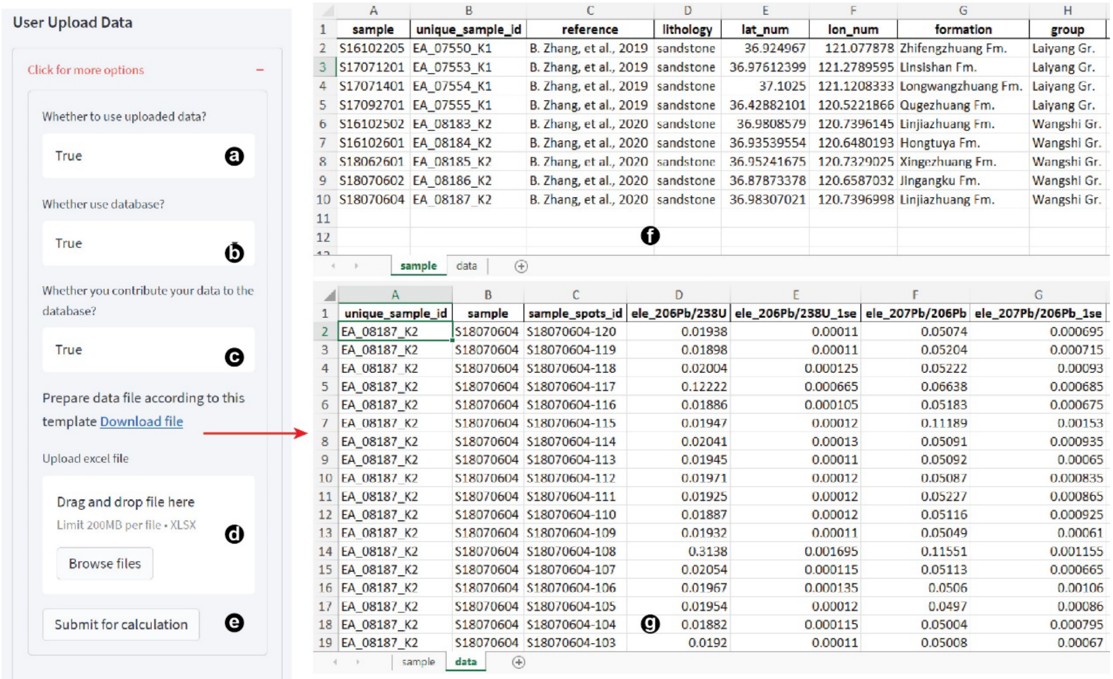
n. Options for users to upload data. See later for more detail.

p. Click on this button to change the basemap.

o. Download the queried data in CSV format for further usages apart from those provided by EaDz.

1.2 Options for uploading data

As an ongoing project, the database will be continuously expanded by incorporating newly published data and data that have not yet been considered. We encourage users to collaborate with us on data compilation by submitting data through the website or contacting the corresponding authors directly after their data are published. It is important to note that the EaDz application will not keep the data uploaded by users if they do not want to submit their data to the database. Our website



sample	unique_sample_id	reference	lithology	lat_num	lon_num	formation	group
S16102205	EA_07550_K1	B. Zhang, et al., 2019	sandstone	36.924967	121.077878	Zhifengzhuang Fm.	Laiyang Gr.
S17071201	EA_07553_K1	B. Zhang, et al., 2019	sandstone	36.97612399	121.2789595	Linsishan Fm.	Laiyang Gr.
S17071401	EA_07554_K1	B. Zhang, et al., 2019	sandstone	37.1025	121.1208333	Longwangzhuang Fm.	Laiyang Gr.
S17092701	EA_07555_K1	B. Zhang, et al., 2019	sandstone	36.42882101	120.5221866	Qugezhuang Fm.	Laiyang Gr.
S16102502	EA_08183_K2	B. Zhang, et al., 2020	sandstone	36.9808579	120.7396145	Linjiazhuang Fm.	Wangshi Gr.
S16102601	EA_08184_K2	B. Zhang, et al., 2020	sandstone	36.93539554	120.6480193	Hongtuya Fm.	Wangshi Gr.
S18062601	EA_08185_K2	B. Zhang, et al., 2020	sandstone	36.95241675	120.7329025	Xingezhuang Fm.	Wangshi Gr.
S18070602	EA_08186_K2	B. Zhang, et al., 2020	sandstone	36.87873378	120.6587032	Jingangku Fm.	Wangshi Gr.
S18070604	EA_08187_K2	B. Zhang, et al., 2020	sandstone	36.98307021	120.7396998	Linjiazhuang Fm.	Wangshi Gr.

unique_sample_id	sample	sample_spots_id	ele_206Pb/238U	ele_206Pb/238U_1se	ele_207Pb/206Pb	ele_207Pb/206Pb_1se
EA_08187_K2	S18070604	S18070604-120	0.01938	0.00011	0.05074	0.000695
EA_08187_K2	S18070604	S18070604-119	0.01898	0.00011	0.05204	0.000715
EA_08187_K2	S18070604	S18070604-118	0.02004	0.000125	0.05222	0.00093
EA_08187_K2	S18070604	S18070604-117	0.12222	0.000665	0.06638	0.000685
EA_08187_K2	S18070604	S18070604-116	0.01886	0.000105	0.05183	0.000675
EA_08187_K2	S18070604	S18070604-115	0.01947	0.00012	0.11189	0.00153
EA_08187_K2	S18070604	S18070604-114	0.02041	0.00013	0.05091	0.000935
EA_08187_K2	S18070604	S18070604-113	0.01945	0.00011	0.05092	0.00065
EA_08187_K2	S18070604	S18070604-112	0.01971	0.00012	0.05087	0.000835
EA_08187_K2	S18070604	S18070604-111	0.01925	0.00012	0.05227	0.000865
EA_08187_K2	S18070604	S18070604-110	0.01887	0.00012	0.05116	0.000925
EA_08187_K2	S18070604	S18070604-109	0.01932	0.00011	0.05049	0.00061
EA_08187_K2	S18070604	S18070604-108	0.3138	0.001695	0.11551	0.001155
EA_08187_K2	S18070604	S18070604-107	0.02054	0.000115	0.05113	0.000665
EA_08187_K2	S18070604	S18070604-106	0.01967	0.000135	0.0506	0.00106
EA_08187_K2	S18070604	S18070604-105	0.01954	0.00012	0.0497	0.00086
EA_08187_K2	S18070604	S18070604-104	0.01882	0.000115	0.05004	0.000795
EA_08187_K2	S18070604	S18070604-103	0.0192	0.00011	0.05008	0.00067

Figure 2. Options for uploading data

also allows users only to use their own data, in which case the site serves as an interactive platform for analyzing the detrital zircon data. User can upload data through data uploading portal (Fig. 2d). Refer to the Template provided on the EaDz website for data preparation (Fig. 2f-g). Below are some more options you can specify during data uploading.

a-c. These options control how EaDz handles data uploaded by users. **a.** Whether to use uploaded data? If True, the user's uploaded data will be merged with the data selected by users in the EaDz database, **b.** Whether use database, if False, users only use their own data, in which case the site only serves as an interactive platform for analyzing the detrital zircon data; **c.** Whether you contribute your data to the database? If False, the EaDz application will not keep the data uploaded by the user.

d. User data uploading portal. EaDz requires that uploading data be prepared using a Microsoft Excel file containing "***sample***" and "***data***" worksheets. These two worksheets are linked through a unique sample identifier ("***unique_sample_id***" column) that must be unique and assigned to each sample.

The "***sample***" worksheet contains information related to individual samples. Rows must contain column names and at least a "***unique_sample_ID***" column. Each unique sample in the worksheet must have its row and be labeled with a unique sample identifier. ***"lat_num"*** and ***"lon_num"*** columns are the latitude and longitude of samples, respectively must be entered in decimal degrees format using the WGS84 datum. These two columns are required for plotting the sample on the map and reconstructing the palaeolocation of the sample. Other columns, such as '***reference***', '***lithology***', '***formation***', etc., are not mandatory. However, we do recommend the users to provide related information, when available, if they are willing to generously share their data with us.

The "data" worksheet contains information related to each detrital analysis. At least two columns are required: "***unique_sample_id***" and "***sample_spots_id***." The ***"unique_sample_id"*** column must contain the same unique sample identifier that is used to identify the sample within the "sample" worksheet. "***sample_spots_id***" is the unique identifier for each analysis. The column name of analysis items must be the same as those shown in the Template file.

e. Click to upload the file and execute the associated task.

2 Interactively visualization and analyzing detrital zircon data

The visualization section provides several tools to explore, analyze and visualize the selected data from the EaDz database and, optionally, user-uploaded data. Each analysis tool has its parameter sidebar or option input box for further customization. The visualization tools can be either used for sample or sample groups. EaDz provides a convenient way to incorporate samples into groups. Here, we focus on the visualization of samples and how to incorporate samples into groups. The visualization of user-customized sample groups is the same as the visualization of samples.

2.1 Visualizing and analyzing age data

2.1.1 Visualization of Zircon age distribution

The zircon age distribution can be analyzed using the most popular visualization methods, including cumulative distribution plots (CDF), pie charts, histograms, and relative probability distribution plots. A cumulative distribution refers to the sum of all probabilities of occurrence less than or equal to a certain age. Probability density plots (PDPs) and kernel density estimate plots (KDEs) are provided for the relative probability distribution. PDPs are constructed by summing and normalizing a Gaussian distribution for each analysis, where the mean and standard deviation of each Gaussian distribution is equal to the age and standard error of the analysis (Andersen *et al.*, 2018). KDEs are constructed similarly to PDPs, except that each Gaussian distribution uses bandwidth for the standard deviation rather than the standard error of analytical uncertainty (Vermeesch, 2012, 2018).

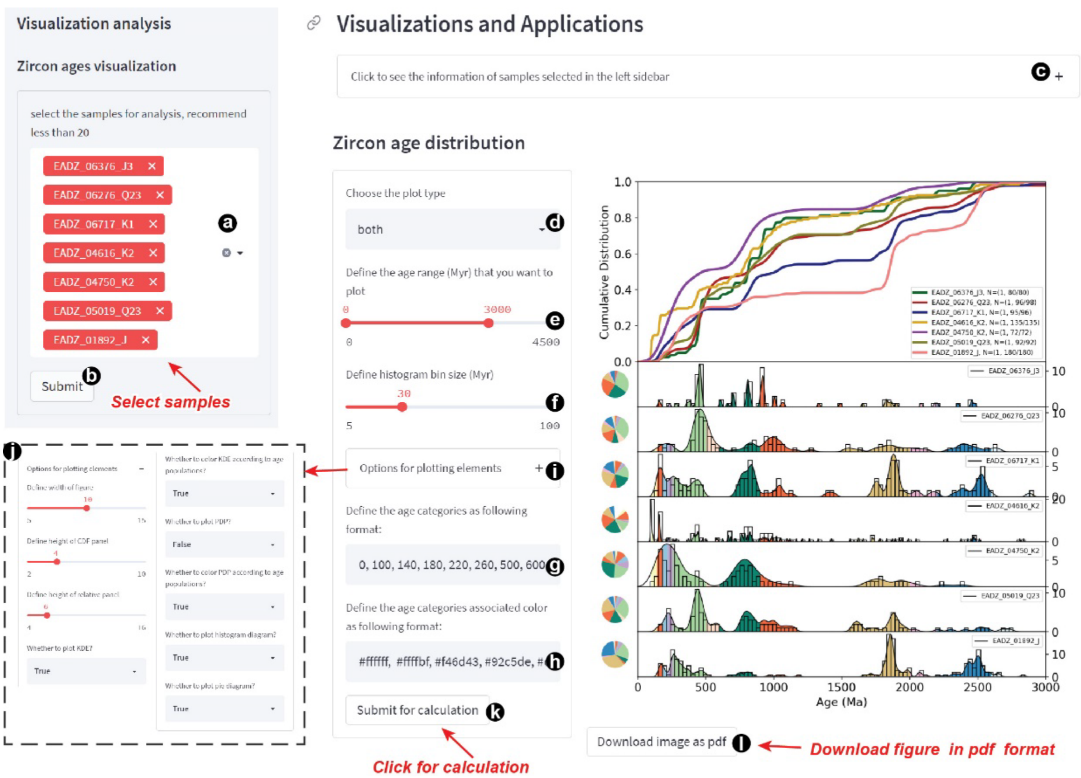


Figure 3. Options for visualization of Zircon age Distribution

a-c. Select samples for visualization and analysis. The selected samples can also be downloaded in CSV format through the function provided in Fig. 3c. After selecting the samples and clicking on the "Submit" button (Fig. 3b), the EaDz will perform all detrital zircon age analysis function within the EaDz with default parameters, including "calculation of maximum depositional age" and "multidimensional scaling analysis." Then we can fine-tune the results by customizing parameters and the "submit for calculation" button inside each function, as shown in Fig. 3d-k.

d. Choose plot types. The plotting function is divided into a cumulative distribution plot in the upper panel and relative distribution plots in the lower panel. The option here can be set to ***'both'***, ***'cumulative',*** or ***'relative'.*** If it is set to ***"both"***, the cumulative distribution is shown on top, and the relative distribution for each sample is shown below. The default value is ***"both".***

e. Specify age range for plot in Myr by slide bar. The default value is from 0 to 3000.

f. Specify width of the histogram bin in Myr by slide bar. The default value is thirty.

g. An array contains comma-separated age range boundaries to define age categories. The default value is ***"0, 100, 140, 180, 220, 260, 500, 600, 900, 1500, 2000, 2200, 2600, 2900, 4500".***

h. A comma-separated array of Hexadecimal color values corresponding to the specified age bin boundaries. The default value is ***"#ffffff, #ffffbf, #f46d43, #92c5de, #c2a5cf, #a6dba0, #fddbc7, #018571, #f46d43, #dfc27d, #f1b6da, #3288bd, #e0f3f8, #fee08b".***

i-j. Other options for specifying the drawing style, such as height and width of the drawing.

k. Click to update drawing based on the above custom parameters.

l. Download figure in PDF format.

2.1.2 Calculation of Maximum Depositional Age

Under the assumption that sediments cannot be older than the youngest crystallization age of their source rocks, using the youngest U–Pb ages of detrital zircons as constraints on the maximum depositional age (MDA) of strata has grown rapidly over the past two decades (Coutts et al., 2019; Dickinson and Gehrels, 2009; Johnstone et al., 2019; Sharman and Malkowski, 2020). We used the methods following Dickinson & Gehrels (2009) to calculate the MDA for the sample in the database. For each sample, three ages will be calculated: the youngest single grain (YSG), the youngest cluster of 2 or more ages with overlapping 1σ uncertainties (YC1σ(2+)), and the youngest cluster of 3 or more ages with overlapping 2σ uncertainties (YC2σ(3+)) (see Sharman et al., 2018, for more details).

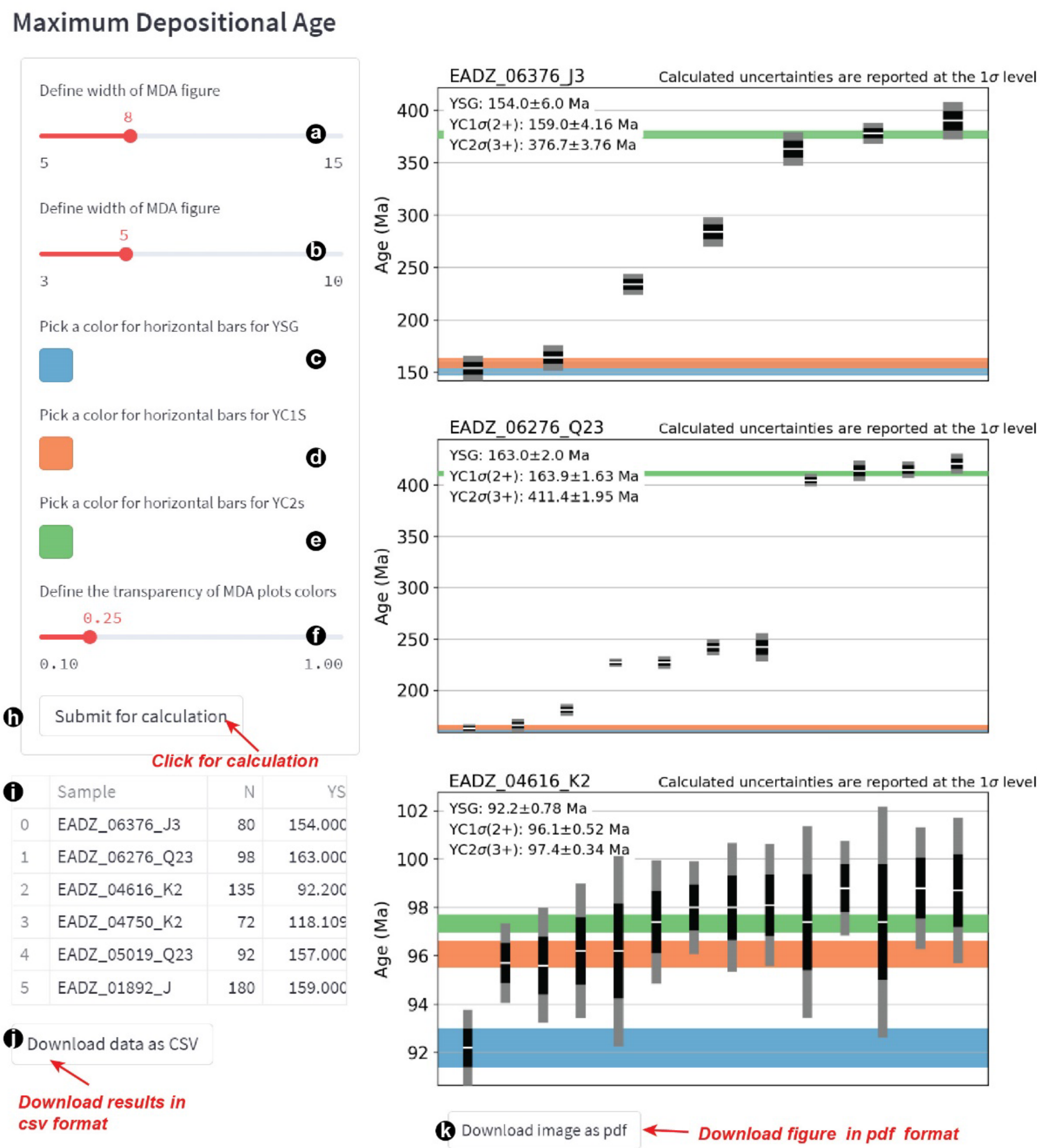


Figure 4. Options for calculating the maximum depositional age.

a-b. Specify width and height of the panel for each sample by the slide bar.

c-d. Specify color of horizontal bars for YSG, YC1σ(2+), and YC2σ(3+), respectively.

f. Specify transparency of the colors.

h. Click to update the drawing based on the customized parameters.

i. Table showing calculation results.

j. Download calculation results in CSV format.

k. Download figure in pdf format.

2.1.3 Multidimensional scaling analysis

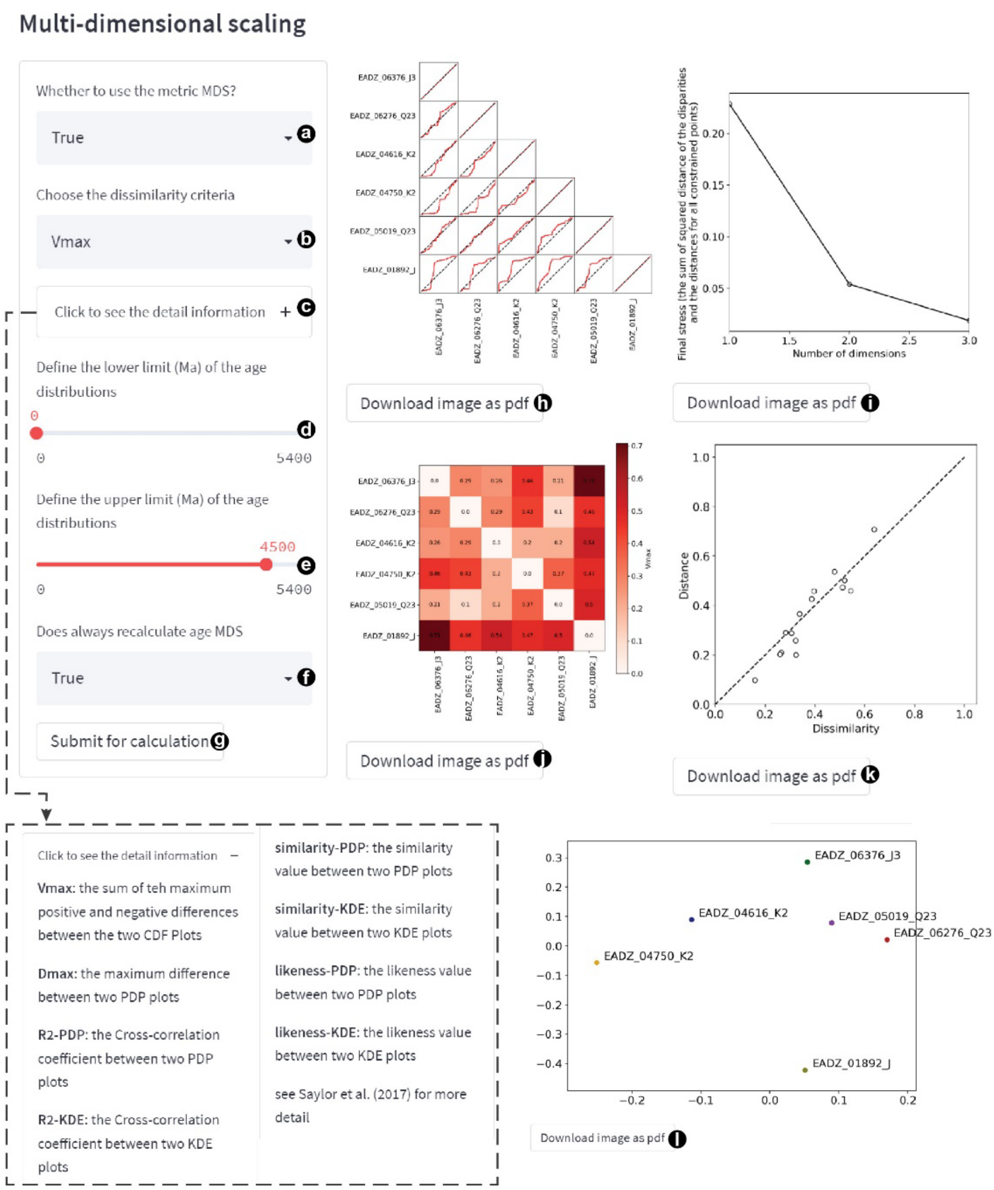


Figure 5. Options for Multidimensional scaling analysis

Converting age distributions to points on a scatter plot is an efficient and effective data visualization method when comparing multiple samples. Multidimensional scaling (MDS) analysis achieves this goal by converting dissimilarities between samples into distances in N-dimensional space (Saylor and Sundell, 2016; Vermeesch, 2013). MDS plots can be created in EaDz using the detritalPy library implemented by the sklearn module, with the option of metric or non-metric MDS, as well as multiple metrics to measure dissimilarities between samples following Sharman et al. (2018). Additional information on MDS implemented by the sklearn module can be found on this site (https://scikit-learn.org/stable/modules/manifold.html#multidimensional-scaling). A heatmap can visualize the dissimilarity metric between sample pairs in MDS space.

a. Specify whether use the metric MDS or not. True for metric MDS, whereas False for non-metric MDS.

b-c. Choose the dissimilarity criteria, which can be set to any of the followings:

- ***"Vmax"***: the sum of the maximum positive and negative differences between the two cumulative distribution plots (CDFs). The default value used in EaDz.
- ***"Dmax"***: the maximum difference between the CDFs
- ***"R2-PDP"***: the Cross-correlation coefficient between two PDP plots
- ***"R2-KDE"***: the Cross-correlation coefficient between two KDE plots
- ***"similarity-PDP"***: the similarity value between two PDP plots
- ***"similarity-KDE"***: the similarity value between two KDE plots
- ***"likeness-PDP"***: the likeness value between two PDP plots
- ***"likeness-KDE"***: the likeness value between two PDP plots

d-e. Specify the lower and upper limits of the age distribution in Myr.

f. Since MDS calculations are time-consuming, you can prevent automatic recalculation after each sample reselection by setting this option to False. If set to False, you need to manually click the "submit for calculation" button to activate the calculation.

g. Click to update drawing based on the customized parameters.

h-l. Download figures in pdf format.

2.2 Visualizing and analyzing Lu-Hf isotopic data

Lu-Hf isotope data can be visualized as violin plots and bivariate kernel density estimates (2D-KDE) combined with paired U–Pb age data.

2.2.1 Violins plots of Lu-Hf isotopic data

The violin plot with the kernel density estimates of the underlying data shows the distribution of Lu-Hf isotopes.

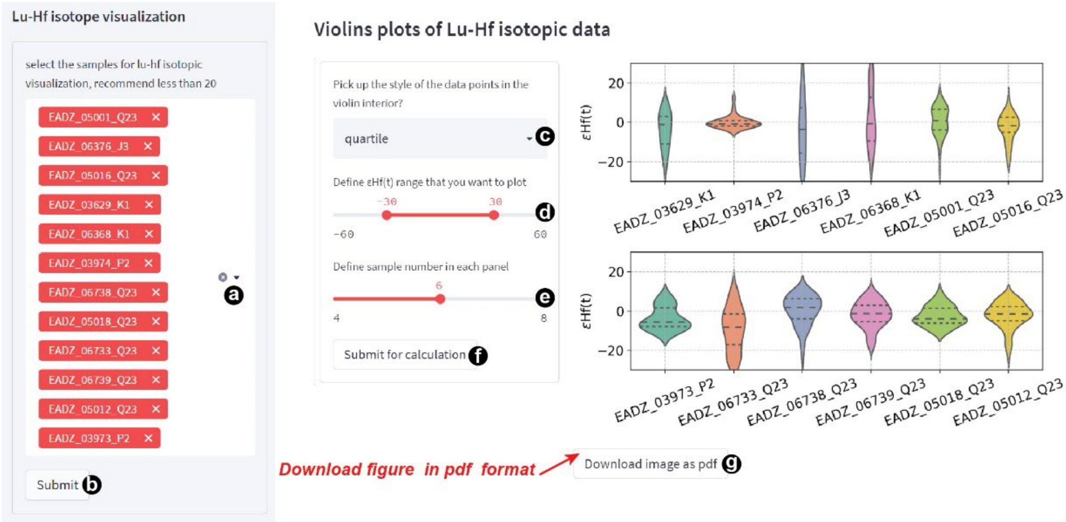


Figure 6. Options for Violins plots of Lu-Hf isotopic data

a-b. Select samples with Lu-Hf isotopic data for analysis. The EaDz automatically scans samples containing Lu-Hf isotope data for user selection. After selecting the samples and clicking on the ***"Submit"*** button (Fig. 6b), the EaDz will perform both the "Violins plots" and "2D-KDE plots" analysis with default parameters. Then we can adjust the results by customizing parameters and clicking the "submit for calculation" button inside each function, as shown in Fig. 6c-g.

c. Specify the representation of the data points in the violin interior. If ***"box"***, it will draw a miniature boxplot. If ***"quartiles"***, it will draw the quartiles of the distribution. If ***"point"*** or ***"stick"***, it will show each underlying data point. Using None will draw unadorned violins.

d. Specify the $\varepsilon Hf_{(t)}$ range of the plot by a slide bar. The default value is from -30 to 30.

e. Specify sample number in each panel. The default value is six.

f. Click to update drawing based on the customized parameters.

g. Download figure in pdf format.

2.2.2 2D-KDE plots of Lu-Hf isotopic data

To construct the bivariate kernel density estimation plot, each data point on the 2D scatter plot is converted to a 2D Gaussian kernel density estimate (Fig. 7). The respective kernel bandwidths for the X-axis (U–Pb age) and Y-axis (εHf(t)) values are determined by optimization. More information can be found here (https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.gaussian_kde.html#scipy.stats.gaussian_kde).

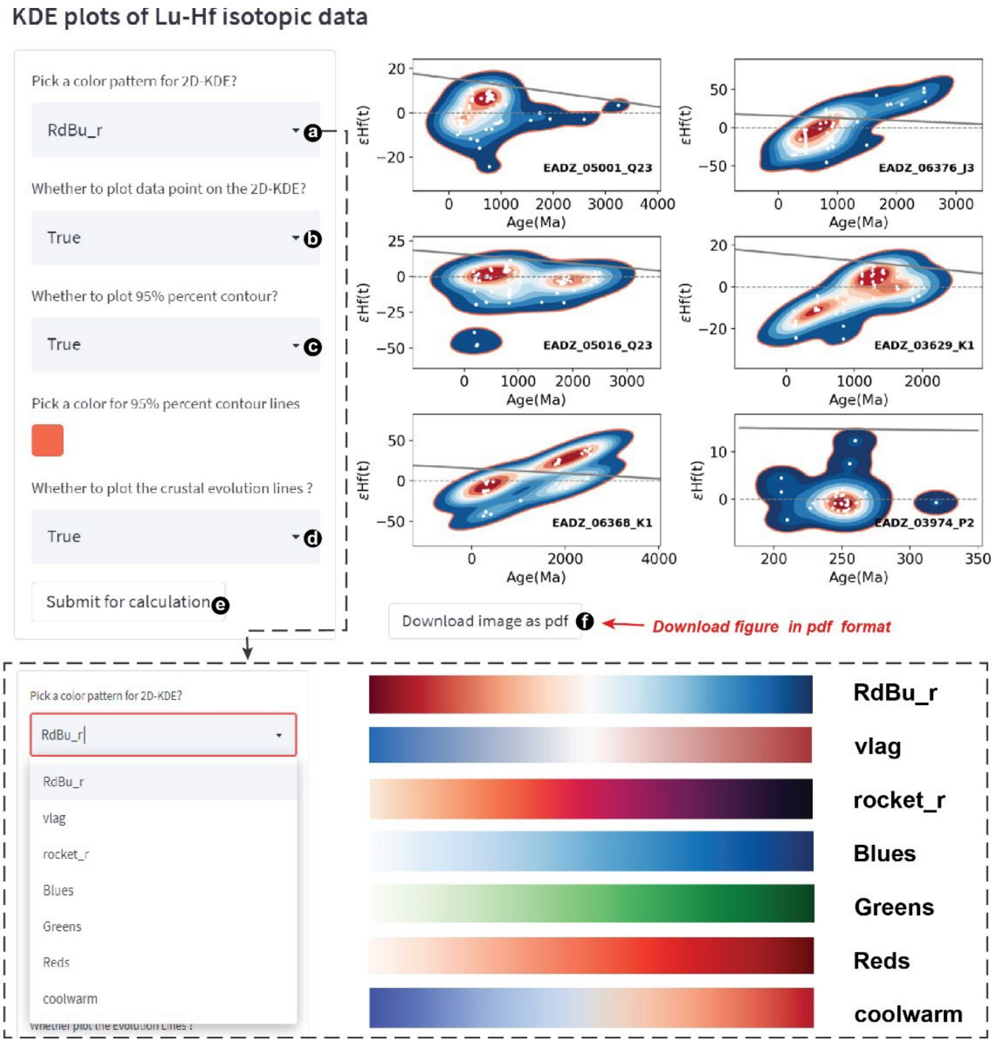


Figure 7. Options for bivariate kernel density estimates of Lu-Hf isotopic data.

a. Specify the colormap, which can be set to ***"RdBu_r"***, ***"vlag"***, ***"rocket_r"***, ***"Blues"***, etc. The default value is ***"RdBu_r".***

b. Specify whether to plot the data point on the 2D-KDE. If ***True***, the data will be plotted as white dots on the 2D-KDE plot.

c. Specify whether to plot the contours at 95% of peak density. If True, the color of the contour lines can also be specified.

d. Specify whether to plot the crustal evolution lines. If True, the solid gray line for the Hf-isotope evolution depleted mantle (DM) reservoir and the gray dash line for the Hf-isotope evolution Chondritic (CHUR) reservoir.

e. Click to update drawing based on the customized parameters.

f. Download figure in pdf format.

2.3 Visualizing and analyzing based on user-customized sample groups

The visualization tools described above can be used in the same way for user-customized sample groups. EaDz provides a convenient way to incorporate samples into groups (Fig. 8).

Figure 8. Options for user-customized sample groups.

a. Specify the group number by a slide bar. The default value is six.

b-e. After defining the group number, the EaDz will create several groups based on the group number. Each group has a unique name that can be renamed via a text box (e.g., Fig. 8b, d). The default group names are group_1, group_2, group_3, etc. The samples within each group can be selected from the samples list (e.g., Fig. 8c, e).

f. Click to update the drawing based on the above custom parameters. The options for each function are the same as those used in the sample-based visualization. See sections 2.1 and 2.2 for more details.

3 Palaeogeography reconstruction with Gplates

GIS and age information enables the correlation of detrital zircon datasets to plate reconstruction models for inspection or performing paleogeographic reconstructions within full-plate reconstructions. By linking samples' depositional age and location information with GPlates reconstructions through pyGPlates, EaDz allows palaeoposition reconstruction of samples even back to 1000 Ma (Gurnis et al., 2018; Muller et al., 2018).

3.1 Plot the reconstructed Palaeolongitude and Palaeolatitude

The reconstructed palaeoposition of the samples can be plotted as line charts of palaeolatitude and palaeolongitude, showing their continuous position changes over time.

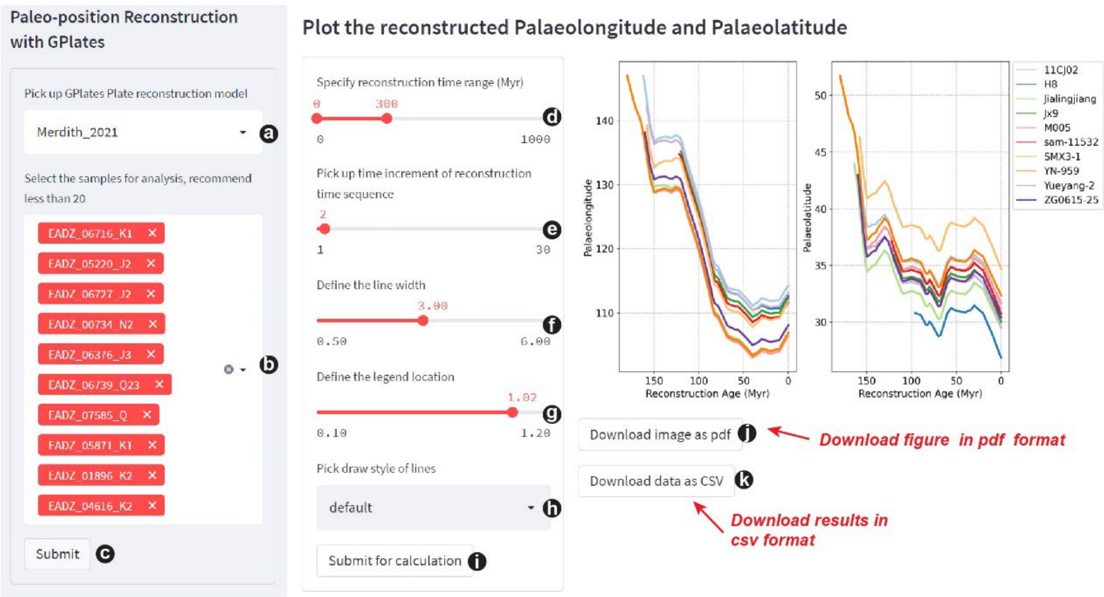


Figure 9. Options for palaeoposition reconstruction of samples.

a. Specify which global GPlates plate reconstruction model you want to use. "Medith_2021" represents the model published by Merdith *et al.* (2021) in the ***Earth-Science Reviews***, whereas "Young_2019" represents the model published by Young *et al.* (2019) in ***Geoscience Frontiers.***

b-c. Select samples from the results of the above data filters. After selecting the samples and clicking the "Submit" button (Fig. 9c), the EaDz will perform palaeoposition reconstruction and plot the results in both line chart and map view. Then we can then adjust the results by customizing parameters and the "submit for calculation" button inside each function, as shown in Fig. 9d-h.

d. Specify the reconstruction time range in Myr. The default value is from 0 to 300.

e. Specify the time increments of the reconstruction time sequence. The default value is two.

f. Specify the line width in pt. The default value is three.

g. Specify the legend location. When the value is less than 1.00, the legend will be plotted inside the box; When the value is greater than 1.00, the legend will be plotted outside the box. The default value is 1.02.

h. Specify how the points are connected. For ***"default"***, the points are connected with straight lines. The steps variants connect the points with step-like lines, i.e., horizontal lines with vertical steps. They differ in the location of the step: ***"steps-pre"***: The step is at the beginning of the line segment; ***"steps-mid"***: The step is halfway between the points; ***"steps-post"***: The step is at the end of the line segment.

i. Click to update drawing based on the customized parameters.

j. Download figure in PDF format.

k. Download reconstructed palaeoposition of samples in CSV format.

3.2 Palaeoposition of samples on the map

The reconstructed palaeoposition of the samples can also be plotted as maps with reconstructed plate boundaries showing the palaeogeographic setting of several deep time slices.

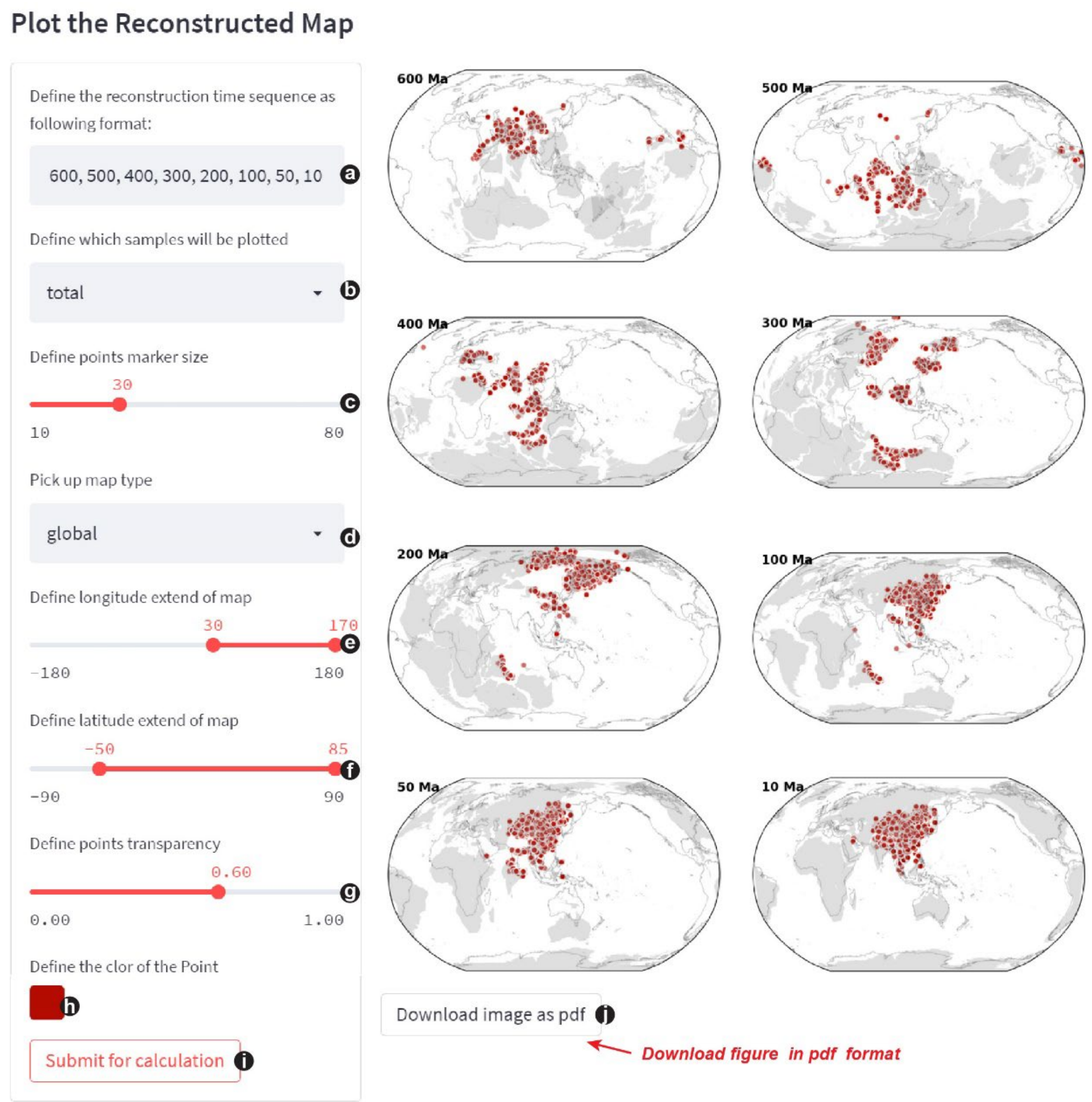


Figure 10. Options for plotting reconstructed palaeoposition of samples in the map.

a. An array contains comma-separated ages to define the reconstruction time slices. The default value is ***"600, 500, 400, 300, 200, 100, 50, 10"***.

b. Set the data set to be displayed on the map. There are two options, ***"total"*** will plot all samples in the EaDz database, and ***"selected"*** will plot only the samples selected by the user. The default value is ***"total"***.

c. Specify the size of the points maker in the pt. The default value is thirty.

d. Specify the map type. For ***"global"***, the data will be plotted on a world map in the Robinson projection; For ***"regional"***, the data will be plotted on a regional map in the Universal Transverse Mercator (UTM) projection, and the map extent can be specified later.

e-f. Specify the map extent via the sliders, which only works when the map type is set to ***"regional"***.

g-h. Specify the color and transparency of the points.

i. Click to update the drawing based on the customized parameters.

j. Download the figure in PDF format.

Reference

Andersen, T., Kristoffersen, M., Elburg, M.A., 2018. Visualizing, interpreting and comparing detrital zircon age and Hf isotope data in basin analysis - a graphical approach. Basin Research 30, 132-147.

Coutts, D.S., Matthews, W.A., Hubbard, S.M., 2019. Assessment of widely used methods to derive depositional ages from detrital zircon populations. Geoscience Frontiers 10, 1421-1435.

Dickinson, W.R., Gehrels, G.E., 2009. Use of U-Pb ages of detrital zircons to infer maximum depositional ages of strata: A test against a Colorado Plateau Mesozoic database. Earth and Planetary Science Letters 288, 115-125.

Gurnis, M., Yang, T., Cannon, J., Turner, M., Williams, S., Flament, N., Muller, R.D., 2018. Global tectonic reconstructions with continuously deforming and evolving rigid plates. Computers & Geosciences 116, 32-41.

Johnstone, S.A., Schwartz, T.M., Holm-Denoma, C.S., 2019. A Stratigraphic Approach to Inferring Depositional Ages From Detrital Geochronology Data. Frontiers in Earth Science 7.

Merdith, A.S., Williams, S.E., Collins, A.S., Tetley, M.G., Mulder, J.A., Blades, M.L., Young, A., Armistead, S.E., Cannon, J., Zahirovic, S., Müller, R.D., 2021. Extending full-plate tectonic models into deep time: Linking the Neoproterozoic and the Phanerozoic. Earth-Science Reviews 214.

Muller, R.D., Cannon, J., Qin, X.D., Watson, R.J., Gurnis, M., Williams, S., Pfaffelmoser, T., Seton, M., Russell, S.H.J., Zahirovic, S., 2018. GPlates: Building a Virtual Earth Through Deep Time. Geochemistry Geophysics Geosystems 19, 2243-2261.

Saylor, J.E., Sundell, K.E., 2016. Quantifying comparison of large detrital geochronology data sets. Geosphere 12, 203-220.

Sharman, G.R., Malkowski, M.A., 2020. Needles in a haystack: Detrital zircon U Pb ages and the maximum depositional age of modern global sediment. Earth-Science Reviews 203.

Sharman, G.R., Sharman, J.P., Sylvester, Z., 2018. detritalPy: A Python-based toolset for visualizing and analysing detrital geo-thermochronologic data. Depositional Record 4, 202-215.

Vermeesch, P., 2012. On the visualisation of detrital age distributions. Chemical Geology 312, 190-194.

Vermeesch, P., 2013. Multi-sample comparison of detrital age distributions. Chemical Geology 341, 140-146.

Vermeesch, P., 2018. IsoplotR: A free and open toolbox for geochronology.

Geoscience Frontiers 9, 1479-1493.
Young, A., Flament, N., Maloney, K., Williams, S., Matthews, K., Zahirovic, S., Müller, R.D., 2019. Global kinematics of tectonic plates and subduction zones since the late Paleozoic Era. Geoscience Frontiers 10, 989-1013.