ARTICLE IN PRESS

Geoscience Frontiers xxx (2017) 1–16

HOSTED BY

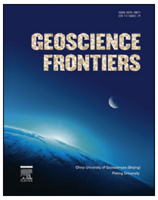
ELSEVIER

Contents lists available at ScienceDirect

China University of Geosciences (Beijing)

Geoscience Frontiers

journal homepage: www.elsevier.com/locate/gsf



Research Paper

A relational database of global U–Pb ages

Stephen J. Puetz

Progressive Science Institute, Honolulu, HI 96815, USA

A R T I C L E   I N F O

Article history:
Received 14 November 2017
Received in revised form
11 December 2017
Accepted 12 December 2017
Available online xxx
Handling Editor: M. Santosh

Keywords:
U–Pb age-distribution
Histogram
Global database
Sampling
Filtering
Zircon

A B S T R A C T

Enhanced understanding of how sampling techniques affect estimates of the global U–Pb age-distribution have, in turn, constrained U–Pb database design. Recent studies indicate that each continent has a unique age-distribution, as determined by zircon ages dated by the U–Pb isotope method. Likewise, broad regions within a continent also exhibit diverse age-distributions. To achieve a reliable estimate of the global distribution, the heterogenous composition of the continental crust requires sampling as many regions as feasibly possible. To attain this goal, and to provide a method for calculating age histograms, the records from a recent global U–Pb compilation are supplemented with 281,631 new records. These additions increase the database size to 700,598 records. In addition, the data are now restructured and made available as a relational database. After filtering the records by the six age-models included with the database, the results reveal two problems that might generally be unrecognized. First, an abrupt switch in the best-age at any given point (such as 1000 Ma) from $^{206}Pb/^{238}U$ ages to $^{207}Pb/^{206}Pb$ ages artificially depresses the age-distribution at the cutoff point. Second, rejecting analyses based on either absolute discordance or the magnitude of 2σ precision errors artificially depresses the age-distribution between 900 Ma and 2000 Ma. The results indicate that, when estimating the global U–Pb age-distribution, the methods for determining best-age and for rejecting records both require some attention. Possible solutions include using either an Accuracy Model or a Precision Model for estimating best-age, and then including all U–Pb records in the estimate, rather than rejecting any of them.

1. Introduction

As awareness increases of how sampling techniques affect estimates of the global U–Pb age-distribution, the enhanced understanding constrains the design of new U–Pb databases. In particular, recent studies show that each continent has a unique age-distribution (Condie et al., 2017; Puetz et al., 2017). Even regions within a continent generally have diverse zircon age-distributions. To achieve an accurate estimate of the global age-distribution, the heterogenous nature of crustal ages requires sampling as many continental regions as possible. The estimation process is complicated by the non-random selection of sample-sites and then using multiple zircon grains from a single sample. These methods are often referred to as convenience-sampling and cluster-sampling (Lo and Watson, 1998; Stehman and Selkowitz, 2010; Etikan et al., 2016). Most sampling sites for zircon U–Pb analyses are selected to address local geologic problems. For example, resource exploration is sometimes a key consideration. Regardless of the reasons, an analysis of sample-sites in a recent global U–Pb database (Puetz et al., 2018) shows that some regions are sampled extensively (over-sampled) while other regions are barely sampled at all (under-sampled).

Weighting the records partially improves estimates from uneven sampling (Lo and Watson, 1998; Puetz et al., 2017, 2018), but not entirely. The primary method, grid-area sampling (Puetz et al., 2017, 2018), involves assigning every sample to one of 10,678 global grids, and then weighing the records based on the number of analyses within each grid. For instance, samples from a grid with 400 records are weighted twice that of samples from a grid with 800 records. Indeed, weighting reduces the influence of over-sampled regions; however, weighting can never adjust for regions that are not sampled at all. Thus, to further improve the global U–Pb estimate, the current effort focuses on obtaining more data for under-sampled regions. From this effort, 281,631 new records are now available from 1397 research articles. These additions increase the total size of the database to 700,598 U–Pb records, from 2656 research articles.

E-mail address: puetz.steve@gmail.com.

Peer-review under responsibility of China University of Geosciences (Beijing).

https://doi.org/10.1016/j.gsf.2017.12.004

Please cite this article in press as: Puetz, S.J., A relational database of global U–Pb ages, Geoscience Frontiers (2017), https://doi.org/10.1016/j.gsf.2017.12.004

In addition to the new data, the increased size adds complexities that limit the platforms available for holding the data. To minimize calculation times and to increase filtering capabilities, the U–Pb records are now transported to a Microsoft® Access™ relational database, hereafter referred to as the Relational U–Pb Database. The new database uses referential integrity, which limits valid entries to the unique elements in its base-tables. This is fundamental to the design because filters that query the data always return all appropriate records when linked with referential integrity. For example, in a free-format system, "North America" could be abbreviated as "N America", "North Amer", or "N Amer", and a viewer would generally recognize all three as representing North America. However, in a database query, a filter using "N America" as the criterion would only return some of the North American records. Accordingly, the Relational U–Pb Database utilizes referential integrity to ensure that all elements within its classification tables are defined in only one way, and thus, the referential integrity eliminates the possibility of partial-filtering.

2. Data

The ages of the U–Pb records in the database span most of Earth's existence – ranging from the present into the Hadean. The data are from samples of all types of rocks, primarily from published research articles in major and minor geological journals, but the database also includes records from dissertations and theses. The database itself cannot address biases that might be related to age, rock type, zircon recycling (the same zircon might be transferred through multiple sedimentary cycles), zircon fertility (zircon is often more abundant in some rock types and less abundant in others), and preservation potential (younger rocks are typically more abundant than older rocks). However, the database includes details such as country, GPS coordinates, rock type, and depositional age that might provide a means for studying these types of biases in detail.

The database is built from the 418,967 records in Global U–Pb Database 2017 (Version 1), and is expanded by 281,631 new records (Version 2), primarily from data in numerous country-specific journals. Even though the current database contains all records from Version 1, some of the Version 1 data are revised and others re-classified. Enhancements include (1) some minor revisions to the GPS coordinates, (2) the addition of estimated depositional ages and maximum depositional ages for all sedimentary rocks, (3) reorganization of the rock-type classifications, and (4) many in-press references now have final publication details.

Maps of the continents (Figs. 1–4) illustrate the expanded coverage, with blue dots (Version 1) and pink dots (Version 2) designating the sample-sites that are now available. Even though Version 1 already contains an extensive set of samples from Asia (Fig. 1), most of the under-sampled regions now have extensive coverage. For example, Version 2 includes samples from the Tarim Basin, Southeast Asian, Western Asia, and Russia – all being regions with very limited sampling in Version 1. Similarly, Version 2 samples from Europe (Fig. 2) fill many of the gaps left from Version 1. In Africa (Fig. 2), sampling coverage improves significantly in western regions and moderately in parts of central and eastern Africa. However, when compared to the other continents, Africa remains significantly under-sampled.

Coverage also improves for the regions in and around Australia (Fig. 3), although there are few samples from Papua-New Guinea. Likewise, the Version 2 sampling coverage in Antarctica (Fig. 3) barely improves beyond the regions already sampled in Version 1. Finding new U–Pb samples for many regions of Africa and Antarctica remains an ongoing challenge.

In the Western Hemisphere, new samples from Canada and the USA plains fill many of the North American gaps (Fig. 4), and Version 2 samples from the Amazon and eastern South America expand the coverage for under-sampled regions of South America (Fig. 4). Collectively, the expanded sampling helps to determine if the coverage is already sufficient for estimating the global U–Pb age-distribution, or if further sampling is required. At a minimum, the maps (Figs. 1–4) illustrate the regions that remain under-sampled, and thus, require future research.

Two versions of the database are available as Supplementary Files S1 (spreadsheet version) and S2 (relational database version), and Supplementary document S3 describes the installation and operational procedures for the Relational U–Pb Database, and includes illustrations of its usage.

3. Methods

The results (Section 4) provides examples of outputs for most of the filtering options. The best-age model is a key parameter for generating the outputs, and six options are available for making the calculations: Accuracy Model, Precision Model, Standard Model, only ^{206}Pb/^{238}U ages, only ^{207}Pb/^{235}U ages, and only ^{207}Pb/^{206}Pb ages.

The first model, the Accuracy Model, is so named because it uses the two ages normally presumed to be the most accurate – the ^{206}Pb/^{238}U and ^{207}Pb/^{206}Pb ages. However, the ^{207}Pb/^{235}U age is used to estimate the probability of the ^{207}Pb/^{206}Pb age being the most accurate age. Then, the probability determines the weights that are assigned to the ^{206}Pb/^{238}U and ^{207}Pb/^{206}Pb ages for estimating best-age. The probability, $P_{206\mathrm{Pb}}$, of the ^{207}Pb/^{206}Pb age being the most accurate is given by Eq. (1).

$$P_{206\mathrm{Pb}} = \frac{\arctan\left(\frac{A_{235\mathrm{U}}}{150} - 10\right) + 1.472}{2.993} \quad (1)$$

where arctan ($A_{235\mathrm{U}}$) is the arctangent of the ^{207}Pb/^{235}U age. An estimate of the probability, $P_{238\mathrm{U}}$, of the ^{206}Pb/^{238}U age being the most accurate is given by Eq. (2).

$$P_{238\mathrm{U}} = 1 - P_{206\mathrm{Pb}} \quad (2)$$

Then, the Accuracy Model uses the weights ($P_{238\mathrm{U}}$ and $P_{206\mathrm{Pb}}$) and the ages ($A_{238\mathrm{U}}$ and $A_{206\mathrm{Pb}}$) to estimate the best-age, A_{best}, as given in Eq. (3).

$$A_{\mathrm{best}} = P_{238\mathrm{U}}A_{238\mathrm{U}} + P_{206\mathrm{Pb}}A_{206\mathrm{Pb}} \quad (3)$$

Fig. 5 illustrates the probabilities of the ^{206}Pb/^{238}U ages (red curve) and ^{207}Pb/^{206}Pb ages (blue curve) being the most accurate, as a function of time. The Precision Model considers all three chronometers (^{206}Pb/^{238}U age, ^{207}Pb/^{235}U age, and ^{207}Pb/^{206}Pb age) as candidates, and uses the age with the smallest 2σ precision error as the best-age. The Standard Model is widely used in practice. For concordant analyses with ages less than 1000 Ma, the ^{206}Pb/^{238}U age is considered best. Conversely, for ages greater than or equal to 1000 Ma, the ^{207}Pb/^{206}Pb age is considered best.

The other three models only include a single chronometer: either ^{206}Pb/^{238}U ages, ^{207}Pb/^{235}U ages, or ^{207}Pb/^{206}Pb ages. Plotting the chronometers individually can provide insights beyond those possible when they are combined in a best-age model.

4. Results

In this section, the figures are from the Relational U–Pb Database outputs generated by selecting various combinations of the 20 available filters. The filtering combinations are immense and

Please cite this article in press as: Puetz, S.J., A relational database of global U–Pb ages, Geoscience Frontiers (2017), https://doi.org/10.1016/j.gsf.2017.12.004