Author Response to Referee #1

A global zircon U–Th–Pb geochronology database

Yujing Wu et al.

Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2023-20

RC: Referee Comment, AR: Author Response

Dear referee,

Thank you very much for your positive response, and for your precious time and effort spent reviewing the manuscript and the dataset. Your suggestions are of great help. Please find a point-by-point reply below.

Kind regards,
Yujing Wu (on behalf of the author team)

Comments and responses:

RC: General Comments:

The U-Pb database discussed in this manuscript is certainly publishable, and it could provide the research community with valuable. However, at this point, I'm still unsure about its actual value for multiple reasons. Despite these concerns, I recommend publishing the database, but only after major revisions to the manuscript. The areas of concern are threefold: (a) using outdated and inaccurate methods for determining the best U-Pb age, which is more directly related to the degree of concordance than it is to the age-uncertainty, (b) concerns about the percentage of ages that have null values, these should be stated in the revised manuscript, and (c) significant grammar errors and/or poor word choices that will likely require a reliable editing service to correct. Further details related to these items follow.

AR: General Responses:

Thank you very much for your comments. We will revise the manuscript as you suggested. The following is our reply to your concerns.

(a) We will reselect the best U-Pb ages using the new non-iterative probability method you suggested and discuss its advantages. But, we would like to keep the original age series in the main text or supplementary materials for readers to compare.

(b) We will add the $^{206}Pb/^{238}U$, $^{207}Pb/^{235}U$, and $^{207}Pb/^{206}Pb$ ratios and uncertainties to the Zenodo repository and give detailed statements of the null values in the revised manuscript.

(c) We feel so sorry for our poor English expression. Thank you very much for your kind suggestions on the grammar and word choices. We will find a more reliable editing service for the revised manuscript.

RC: Specific Comments:

RC: Table at line 118: The database lacks key details, such as the $^{206}Pb/^{238}U$, $^{207}Pb/^{235}U$, and $^{207}Pb/^{206}Pb$ ratios and uncertainties, the depositional/stratigraphic ages, and many records have ages and GPS coordinates that are missing. Despite these deficiencies, the database still has considerable potential for solving outstanding geological problems, but less so than if all data items were completed. The authors should mention the percentage of records that have null values.

AR: We appreciate very much your affirmation of our work. First, we did collect the $^{206}Pb/^{238}U$, $^{207}Pb/^{235}U$, and $^{207}Pb/^{206}Pb$ ratios and uncertainties. We will add these data to the Zenodo repository. Second, for the missing ages and GPS coordinates, we will state the percent of the null values in detail in the revised manuscript. We wish we could fill these values but the original papers publishing the zircon records don't include them. We have to leave these items empty for the sake of authenticity.

RC: Lines 156-157: This sentence currently states: "Although TIMS is more precise, other methods are more efficient and widely used (Gehrels, 2014)." This is not the exact reason. Perhaps rephrase this as: "Although TIMS is more precise, methods such as LA-ICP-MS are more cost effective and thus are more widely used (Gehrels, 2014)."

AR: Will be implemented. Thanks for the suggestion.

RC: Line 161-164: Using an arbitrary cutoff-age at 1000 Ma to select the best U-Pb age, as proposed in this manuscript, is flawed. Instead, Puetz et al. (2021) and Puetz & Spencer (2023) published a non-iterative probability method that (a) eliminates the artificial depression in the U-Pb age distribution at 1000 Ma caused by the arbitrary cutoff method, and (b) produces consistent age-distribution based on the degree of discordance without producing the artificial depression at 1000 Ma. It is suggested that the authors review these papers and mention these advantages, as discussed in detail in the references below:

Puetz, SJ; Spencer, CJ; Ganade, CE (2021). Analyses from a validated global U-Pb detrital zircon database: Enhanced methods for filtering discordant U-Pb zircon analyses and optimizing crystallization age estimates. Earth-Science Reviews 220, 103745. https://doi.org/10.1016/j.earscirev.2021.103745

Puetz, SJ; Spencer, CJ (2023). Evaluating U-Pb accuracy and precision by comparing zircon ages from 12 standards using TIMS and LA-ICP-MS methods. Geosystems and Geoenvironment 2, 100177. http://dx.doi.org/10.1016/j.geogeo.2022.100177

AR: We appreciate you very much for your references. We have carefully read these two papers and will discuss the non-iterative probablity method in the revised manuscript. Although it doesn't matter anymore, we used cutoff ages as shown in Table 4 rather than 1000 Ma.

RC: Line 183-185: Regarding the sentence: "Therefore, the amount of zircon

production can be used to understand the past intensity of geological activity (Hawkesworth et al., 2010).” … Hawkesworth et al. (2010) is a poor reference to support this statement. Instead, the following reference is suggested:

Arndt, N; Davaille, A (2013). Episodic Earth evolution. Tectonophysics 609, 661-674. https://doi.org/10.1016/j.tecto.2013.07.002

AR: Thanks for the reference. We will revise it as you suggested.

RC: Lines 185-194: These lines discuss the results in Figures 8 and 9. Importantly, the age distributions in these figures are raw age counts. The usage of raw age counts is less than optimal because it favors age peaks in heavy sampled regions while failing to show significant age peaks in sparsely sampled regions. For instance, age distributions from a database heavily populated with samples from China, as the database here has, will show strong age-peaks at 800 Ma and 2500 Ma. However, another database with minimal samples from China will tend to show a weak peak at 800 Ma, and a peak at 2700 Ma that is far stronger than the 2500 Ma peak. One way around this problem of disproportionate sampling is to weight the records inversely proportionally to the sampling densities. Then, the resulting age distributions will be remarkably consistent, despite the divergent sampling densities for each database. This suggestion is easy to test simply by first weighting the records inversely proportional to sampling densities, and then summing the age-counts by using the weights. For details about this method, refer to Puetz et al. (2017), *Quantifying the evolution of the continental and oceanic crust,* which is already in the reference list.

AR: Thank you very much for your suggestion. We will calculate new zircon production series by adding weights in the revised manuscript. However, we want to keep the original series using raw age counts in the manuscript. In this way, readers can have more a direct understanding of our database and compare the differences brought about by the weights.

RC: Line 212: Regarding the sentence: “At different geological times, the places where zircons grew in large quantities are also different.” This is already well known and is commonly referred to as the globally heterogeneous distribution of magmatic ages (Hawkesworth et al., 2010; Puetz et al., 2017; and many others). Suggest replacing this sentence by stating that the database here supports the globally heterogeneous distribution of magmatic U-Pb ages.

AR: Will be revised as you suggested.

RC: Line: 220: The authors propose a very subjective approach, with no details on how to accomplish adjustments for the different regional sampling densities. As already explained in the comments related to lines 185-194, the simple and standard approach to solving this problem is to weight the records inversely proportional to sampling densities.

AR: We will weigh the records inversely proportional to sampling densities in the revised manuscript.

RC: Lines 252-254: The method that the authors propose here is seriously flawed, based on tests in Puetz et al. (2021) and Puetz & Spencer (2023) – which compared highly accurate and precise TIMS ages with LA-ICP-MS ages. Using $^{206}Pb/^{238}U$ ages for 0-1163 Ma; $^{207}Pb/^{235}U$ ages for 1163-2390 Ma, and $^{207}Pb/^{206}Pb$ ages when > 2390 Ma is a flawed system. Specifically, the magnitude of the uncertainty (the imprecision) is not directly related to the accuracy of the age. Read Puetz et al. (2021) and Puetz & Spencer (2023) for details about this method. Studies in those papers show that the best U-Pb age gradually transitions from the $^{206}Pb/^{238}U$ age at ~400 Ma to the $^{207}Pb/^{206}Pb$ age at ~1600 Ma. Between those points, the best age gradually transitions from ~400 Ma to ~1600 Ma based on a non-iterative probability model.

AR: Thank you for this new method. We will use this non-iterative probability model to recalculate our data. Again, we want to keep the original series either in the main text or in the supplementary materials for readers to compare differences. Another reason is that we are not sure how much our series are influenced by the cutoff ages since we applied various bin sizes and Monte Carlo simulation to minimize the influence of age uncertainty, which were not used in the papers you suggested. Presenting the results calculated by two methods is a good opportunity to compare and test.

RC: Line 290: Once again, the statement that "The zircon production peaks of the global continental crust are…" is biased by using raw age counts rather than weighting the records inversely proportional to sampling densities.

AR: We will weigh the records inversely proportional to sampling densities in the revised manuscript.

RC: Lines 303-309: This regionally based approach is good, and in this instance, does not necessarily require weighting the records inversely proportional to sampling densities.

AR: Thank you for your support.

RC: Lines 316-319: The age distributions (and thus the periodicities) for detrital, igneous, and metamorphic samples should be nearly identical. Again, if the authors recalculate the age distributions by weighting the records inversely proportional to sampling densities, then I suspect the periodicities will be essentially the same. Inaccurate age-distributions will produce incorrect periodicities. Another important requisite for testing periodicity is to de-trend the data. My question to the authors: Were the age-distributions de-trended prior to spectral analysis?

AR: First, we will weigh the records inversely proportional to sampling densities in the revised manuscript. Second, the age distributions were de-trended before spectral analysis.

RC: Lines 325-347: These are interesting studies that require more rigorous analyses to

determine their reliability.

AR: Great point! In this data description paper, we tend to introduce more potential research values of this database. These intriguing but controversial studies might be better verified in the future using this zircon database as one of the supporting materials.

RC: Lines 373-374: Regarding the sentences: "To solve hot data issues, Puetz et al. (2017) proposed the methods of grid-area and modern-sediment sampling using the surface area to weigh the zircon data. However, this approach is more suitable for studying the exposed crust than it is for studying the evolution of the crust." … This statement is false and it is suggested that it be removed. Weighting records inversely proportional to sampling densities is a STANDARD approach (refer to references in Puetz et al., 2017). However, If the authors actually believe this statement is true, then the authors should present the test that they used to demonstrate this. However, I suspect this is an unsupported statement. For instance, numerous studies over the past 50 years have shown that the age distributions are remarkably similar regardless of depth or height. Parman (2015) shows similar findings – the age distributions remain remarkably similar over time (each involving samples of different depths).

AR: Thanks. We will revise or remove this statement as you suggested.

Grammar related items:

RC: Line 11: Grammar error / typo. Delete the words "and theses"

AR: The words "and theses" will be replaced by "and dissertations". Sorry for the confusion.

RC: Line 14: Poor word choice. Suggest replacing the "weaken" with "minimize"

AR: Will be implemented. Thank you.

RC: Line 17: Instead of mining and energy, is the intent to state: "mining and energy exploration"?

AR: Yes, correct. Will be revised as you suggested.

RC: Line 45-46: As it is currently written, this sentence does not make sense, and in fact, is false: "However, in most cases, these zircon samples were used for independent regional studies and would probably not be used thereafter (Wu et al., 2019)." … In fact, over the past 20 years, numerous authors have re-used the data for these regional studies for further regional analyses and well as in global compilations for global analyses.

AR: This wrong sentence will be deleted.

RC: Lines 46-51: These lines should be deleted and rephrased in one sentence to state something like the following: Here, we expand upon previous global databases of U-Pb dated zircon, which could provide a means for enhanced academic and

commercial geological analyses.

AR: Will be implemented.

RC: Line 62: Suggest deleted the unnecessary words at the end of this sentence "in which the earth is located"

AR: Will be implemented.

RC: Line 63-64 currently state: "However, if the amount of data is not sufficient, the resolution of zircon age series will be lower, leading to possible analysis bias. In addition, limited sampling locations will also affect the objectivity of statistics." This sentence is too wordy and confusing. Thus, suggest making this clearer by simply stating something like the following: "Insufficient data with limited global coverage can affect results, which in turn can contribute to misleading interpretations."

AR: Will be implemented.

RC: Line 67-68: Suggest deleting this sentence: "Undoubtedly, this database provides a more comprehensive and objective chronology data source on both the time and space dimensions for future earth system science research." This interpretation is too strongly worded and even questionable. Only further independent studies (from research teams other than the current set of authors) will determine the usefulness of this global database.

AR: Will be implemented.

RC: Line 70: Suggest revising "other geological events and astronomical environments" to state "other geological and astronomical events"

AR: Will be implemented.

RC: From this point forward, I will no longer make suggestions related to grammar and interpretations. Even while the manuscript is generally understandable, it is riddled with grammar errors, poorly phrased sentences, and awkwardly phrased sentences. Therefore, it is suggested that the authors find a proficient proofreader or editing service to revise the entire manuscript to conform to standard English grammar and phrasing of words.

AR: Sorry for the inconvenience. We will find a more advanced English editing service for the revised manuscript.

RC: Line 268: Again, use the word "minimized" rather than "weakened"

AR: Will be implemented.