Response to Reviewer 1's comments on manuscript egusphere-2022-1200 entitled 'Comparing detrital age spectra, and other geological distributions, using the Wasserstein distance.'

A. Lipp, P. Vermeesch

January 30, 2023

Sincerely, I enjoyed reading the manuscript. It is a concise and neatly written technical paper that presents reasoning and implementation of a numerical metric to map age/grains-size distributions in R and Python. The graphical quality of the figures and tables is good. Still, the axes labelling could be improved for readers unfamiliar with multidimensional scaling.

We thank the reviewer for the encouraging comments.

Minor point: The title seems to promise more than the manuscript delivers. "Comparing . . . ". The manuscript does not compare something. The presented "comparison" is a performance test of the Wasserstein-2 distance and the Kolmogorov-Smirnov distance. I think the title should reflect better what the manuscript tries to achieve: a presentation of an alternative metric in the realm of multidimensional scaling.

We believe the reviewer has slightly misunderstood the meaning of our title. Which is meant to describe how the Wasserstein distance can be used to compare distributions as a dissimilarity metric. Not, as the reviewer suggests, that we are comparing the Wasserstein distance itself. We propose to modify the title to make this meaning more clear and prevent confusion. The proposed new title is: *'The Wasserstein distance as a dissimilarity metric for comparing detrital age spectra, and other geological distributions'.*

The general idea of the manuscript fits within the scope of GChron. However, I am in a little bit of doubt about whether it justifies requesting a peer-review procedure and a peer-review publication. The numerical metric presented here is not new, and the manuscript does not (yet?) show significant scientific progress. The implementation in R appears limited to a few code lines. Perhaps under different circumstances, the implementation in R would have remained a single line in a news files, along with a few lines in the package manual or an entry in a science blog.

This is a fair comment that could be addressed in two opposite ways:

1. By revising the manuscript as a 'Short Communication' which, according to the *Geochronology* website '*report[s] new developments in or novel aspects of methods, techniques, or tools that are relevant for scientific investigations within the journal scope.*' but are not '*fully self-contained research articles*' could be well suited for this manuscript.

2. By significantly expanding the paper. This option may be unavoidable if we are to address all of Reviewer 2's more substantial comments. These additions will probably extend the paper beyond the length of a 'Short Communication'.

We argue that some form of manuscript, is necessary to communicate the use of this metric to the community. We further argue that the simplicity of what we propose, and its pre-existing implementations are a significant advantage, not a drawback. Both of these mean that its benefits can be immediately and rapidly utilised by the community. The limiting factor is instead a low awareness of the Wasserstein distance in the community. A short communication in a relevant journal is we argue the most appropriate way to overcome this.

As a non-expert in multidimensional scaling, I feel the manuscript would benefit from more context. The formal description is sufficient and easy to follow, but the likely impact of this manuscript seems low except for having announced a 'new' feature.

We would be happy to add further context to the Introduction about the importance of dissimilarity measures, in multi-dimensional scaling and elsewise. We strongly contest that the potential impact of this contribution would be 'low'. Since the introduction of multidimensional scaling to detrital geochronology by Vermeesch (2013), the method has been used literally hundreds of times and spawned a small ecosystem of software and derived methods.
How to sensibly evaluate the 'difference' between two observations is of fundamental importance to essentially all statistical methods. Additionally, sensible distance measures are required for developing misfit functions used in inversion schemes (e.g., De Doncker et al. 2020). Whilst MDS is *currently* the most common geochronology technique which uses distance metrics, the potential usage is more significant. The importance of this is reflected in the significant and on-going debate of how best to visualise and compare distributional data (e.g., Sircombe et al. 2004; Saylor et al. 2012; Vermeesch 2012; Satkoski et al. 2013; Vermeesch 2013, 2018; Sundell et al. 2021). We would be happy to incorporate a further discussion along those lines into a revised manuscript.

In other words: How does this new measure perform for real samples and their (new) interpretation? Section 4 reads interesting, but was a new conclusion reached? Did it lead to better (e.g., more accurate, more precise) results, or did the geoscientific interpretation essentially remains the same? If the latter is the case, perhaps you can present a real case underlining the point you want to make better.

We agree with the reviewer and will add several new examples. See the beginning of our reply to Reviewer 1 for further details.

The manuscript comes without a proper discussion. Section 4 is an application example that includes elements of a discussion. However, for a scientific manuscript, I would expect to see more. In particular I would like to see a discussion about the question: Does it likely change the outcome of studies working with this 'new' metric.

We agree. The revised manuscript will be expanded with an in-depth discussion of the pros and cons of the KS and W_2 distances. This will be illustrated with specific examples, as detailed in our response to Reviewer 1. This will give the reader clearer guidance as to when to use which metric.

The synthetic data outlines the general problem you want to address. I suggest leading with an example based on a case study where the Kolmogorov-Smirnov distance did not perform as expected for the reasons you have mentioned.

We agree. See examples 1-4 in our response to Reviewer 1.

Line-by-line comments

L111: I've played a bit with the proposed synthetic data and found that it depends to some extent on the standard deviation. A more narrow standard deviation for the same fixed mean values leads to more complex KS-distance patterns. The higher the degree of overlap (higher standard deviation), the more conclusive the KS distance becomes. Perhaps you can add a few lines about it in the text.

This is an great observation and something we would explore in a revised manuscript.

L150-L175: I think this paragraph can be improved in order to provide a better experience to readers.

We are happy to revise this section in line with the constructive feedback provided.

The R code snippet produces a plot output. However, if I reduce the dataset, it fails. This appears to be a bug in the package 'IsoplotR' because it returns an uncontrolled error:

```
DZ <- IsoplotR::read.data("scandinavia_short.csv", method = "detritals")
```

We assume that the reviewer has created `scandinavia_short.csv` themself? MDS analysis of two samples is pointless so it is not surprising that this throws an error. It is a bit harsh to call this a bug. A three sample data set does work, however.

L152: If I look into the R code (file mds.R), I read in line 199 of the code: modified after the wasserstein1d function of the transport package. It is normal to look up open-source code of others, however, if it helped for the own implementation and since the code line in question seems identical, credit should be given in the manuscript to authors of the package 'transport' (Schuhmacher et al. 2022)

This is a pertinent comment. Proper credit has been given to Schumacher et al. (2022) in the `@author` field of the documentation.

L164: Please consider adding the example data to the manuscript or the R package

We would be happy to add the example data as a supplement to the manuscript, in addition to its current location in the GitHub repository.

L167+ (footnote): The repository pvermees/IsoplotRbeta does not exist, but I guess the branch beta was meant and it should read: remotes::install_github('vermeesch/IsoplotR@beta')

The reviewer is correct and we would fix this in a revised manuscript.

Figure 1: How did you modify the data to "aid illustration"? It appears that you have shifted the 'Byskealven' dataset by 1 Ma. This should be mentioned. If so, how does it affect the W1 distance? Your pink area becomes considerably smaller if the non-shifted dataset is used. Perhaps you have a better dataset at hand; one that does not need such manipulation.

This is correct, the Byskealven sample is simply translated by 1000 Ma, and we will add this to manuscript. This translation adds (exactly) 1000 to W_1 between the two samples. Whilst we could seek a 'real' dataset to illustrate this point, we don't see how this would provide any greater insight as by definition, we would be seeking 'real' data that produces the same insights as presently demonstrated. Additionally, introducing too many datasets we believe could be confusing for the reader.

Figure 2: The upper plot would benefit from y-axis labelling

We will improve labelling of the MDS figures throughout.

Figure 3: Something is at odds with the lower figures (j and k), if I try to reproduce them with IsoplotR::mds() and the example data from GitHub. Figure 3j does not look as in the manuscript, because 'Ljusnan' and 'Byskealven' seem to be no different. Figure 3k is somewhat mirrored. Below my code, I first used IsoplotR():mds(), here more manually to show the calculation steps. The mirrored figure is not a big deal because the interpretation should not change, but it should be presented as the users would see it running the code.

The mirroring is a result of the random state of the MDS algorithm. As this varies from computer to computer, there is little that we can do to prevent this. We will however add text clarifying that reflections are possible but, as the reviewer says, have no impact on any interpretations.

Additional Comments

The following comments refer to something I discovered along the way. It was not considered for my recommendation to the editor.

We really appreciate the detailed comments and expert suggestions regarding the `IsoplotR` package.

I have not used 'IsoplotR' before, but it appears to be an interesting and mature package. However, something I noticed missing while searching for the new feature implementation was a NEWS file, as it is custom for R packages. The authors may want to consider adding such a file because it helps users understand package changes without inspecting code changes. This is something I find very useful for scientific software packages.

Updates were previously listed at `https://isoplotr.es.ucl.ac.uk/home` but have now been added to `NEWS.md` as well. Thank you for the suggestion.

The R code example returns a default plot (IsoplotR::mds()). Perhaps it is obvious to users familiar with the package, but I found it confusing having numbers on the x and y axis but no axes labels. I guess what is shown is some scaled distance. In Section 4, you have explained it better, but personally, I prefer to understand figures without reading the text.

The default axis labels have been changed to 'Dim 1' and 'Dim 2'.

The default plot IsoplotR::mds() returns only the names, but the location in the plot remains a little bit obscure because the default plot setting is pch = NA. I find it awkward because it reduces the value of the plot when what is shown remains too vague to be positioned correctly. In particular, because the text() position in R can be very different from what you would expect depending on the setting in pos.

The default value of `pch=NA` reduces clutter. However, the default data set in the online GUI uses short labels, which are enclosed in circles. At least 20 times more people use `IsoplotR` via the GUI than do via the command line. Those people who do use the command line tend to know how to change `pch`.

In the function IsoplotR::plot.MDS(), the standard plot output has the argument asp = 1 hard coded. This is at least unexpected and leads to inconclusive plot behaviour when ylim and xlim are modified

`asp=1` protects the user from over-interpreting the variations in the second dimension if most of the variability is contained in the first one. One example of this is shown in Figure 3 of the response to Reviewer 2. However, Reviewer 1 is correct about the inconclusive behaviour when both `xlim` and `ylim` are specified. To avoid

this issue, we have added the default argument `asp=1` to the main `mds()` function, where it can be changed by the user.

References

De Doncker, F., F. Herman, and M. Fox (2020). "Inversion of provenance data and sediment load into spatially varying erosion rates". *Earth Surface Processes and Landforms* 45.15, pp. 3879–3901.

Satkoski, A. M., B. H. Wilkinson, J. Hietpas, and S. D. Samson (2013). "Likeness among detrital zircon populations—An approach to the comparison of age frequency data in time and space". *GSA Bulletin* 125.11-12, pp. 1783–1799.

Saylor, J., D. Stockli, B. Horton, J. Nie, and A. Mora (2012). "Discriminating rapid exhumation from syndepositional volcanism using detrital zircon double dating: Implications for the tectonic history of the Eastern Cordillera, Colombia". *Bulletin of the Geological Society of America* 124.5-6, pp. 762–779.

Sircombe, K. N. and M. L. Hazelton (2004). "Comparison of detrital zircon age distributions by kernel functional estimation". *Sedimentary Geology*. Quantitative Provenance Analysis of Sediments 171.1, pp. 91–111.

Sundell, K. E. and J. E. Saylor (2021). "Two-Dimensional Quantitative Comparison of Density Distributions in Detrital Geochronology and Geochemistry". *Geochemistry, Geophysics, Geosystems* 22.4, e2020GC009559.

Vermeesch, P. (2012). "On the visualisation of detrital age distributions". *Chemical Geology* 312-313, pp. 190–194.

– (2013). "Multi-sample comparison of detrital age distributions". *Chemical Geology* 341, pp. 140–146.

– (2018). "Dissimilarity measures in detrital geochronology". *Earth-Science Reviews* 178, pp. 310–321.