

Revision notes

ESSD-2021-280

Manuscript title: Ordovician to Silurian graptolite specimen images for global correlation and shale gas exploration

Corresponding author: Hong-He Xu (hhxu@nigpas.ac.cn)

Revised title: A multi-dimensional dataset of Ordovician to Silurian graptolite specimens images for virtual examination, global correlation and shale gas exploration

Revisions were made based on reviewer's comment and our discussions, which greatly update my understanding of the dataset. They include following aspects.

REASON TO PUBLISH

As in my first review, I welcome this approach to **sharing the contents of paleontological collections in museums and research institutes. It could save considerable travel costs for research based in part upon remote collections.** It may allow the verification or updating of identifications published in the papers cited in the XLSX file. I imagine that the quality and quantity of the new images **far exceeds that in some of the original publications.** Naturally, they cannot exceed the quality of the specimens themselves.

Author reply:

the potential value of our dataset is shown, the idea of virtual examination to specimens (VES) is proposed and emphasized to the revised manuscript. This is the major revision of the manuscript – a new understanding of the dataset. It is not only a collection of images, but a multi-dimensional data with scientific value.

AN IMPROVEMENT and A SERIOUS NEW PROBLEM

The revision notes indicate that the authors have now added the essential scale bars that were promised but missing on images in Version 1 (August 16 2021) of the Xenodo data repository. The new version is presumably enlarged by adding a second set of images that include scales. Although this would be a very welcome correction, I am still struggling to verify it by opening the Version 2 (Jan 31, 2022) ZIP files. The downloaded ZIP files are not readable or extractable by my Windows computer. Neither the file explorer nor the customary unzipping utility can read them. I have experimented with different utilities and tried renaming the files, but to no avail. I can still read and extract the Version 1 files. So, my computer is presumably not the problem. Until this failing is addressed, there are effectively no scaled images in the database. It should not be published in this condition. I would have

Author reply:

The dataset is updated and the newly uploaded to share. According to the editor, we still use the same DOI, for it includes updating. We re-organized all images and assured that every specimen image shows with the scale in its own photo or separated photo. Now the updated dataset includes 2951 images and the whole size is 10.4 G. every scaled photo file is named with a postfix S. The image name consists of the number of specimen and the species name. One only trouble is that the volume of the whole dataset is quite large. It takes a while to upload and download all files.

Now it is ensured that every specimen is shown with additional scale bar and that any measurement to the fossil is available. The visualization tool will be improving for better fulfil the VES.

Although the data are surely intended for use by experts, the manuscript includes some very elementary facts about graptolites and Paleozoic stratigraphy. **Not all are strictly correct.** The local first appearances of graptolite taxa, for example, are a means to correlate with locations away from GSSPs for many Ordovician and Silurian stage boundaries. **The graptolite taxa are indicative, not definitive. First appearances may be diachronous and earlier occurrences may even be found at stratotype sections. The “spikes” are definitive, even if less practical.**

Author reply:

This is a good question. Bio-stratigraphy, or correlation using fossil data, so far is still a common and normal method in stratigraphy and GSSP, the work of which is still ongoing globally and approved by international commission on stratigraphy. Of course fossil record might be diachronous in different sediment settings, but such discussion is beyond the scope of this study. We here just show the dataset-concerned fossils have significance in current study.

The various sections of the paper and its **figure captions tend to be repetitive and still contain many common errors of English.** The repetitiveness is irritating and serves no purpose. The grammatical flaws do not obscure meaning and are not surprising for non-native English-speaking authors. **The usual mismatches of singular subjects with plural verbs occur in the first two words of section 1 – “Graptolite was . . .”** The database managers want readers to trust taxonomic revisions made by un-named experts of their choosing. It would be unfortunate to undermine this by failing to find an English language proof-reader. At one time, trained journal staff would undertake such corrections, but that is less common in large publishing houses today. Fortunately, universities are more cosmopolitan.

REASON TO REQUIRE FUNDAMENTAL REVISION

Given the concerns about quality-control, which range from trivial to nullifying, it is possible that readers will lose trust in the authors' attention to other aspects of their data. Certainly, it begins to shake my confidence in the project. Am I making an elementary unzipping mistake or are the database managers careless? Because the credibility of the data is at risk, it is important to address even minor concerns. I recommend easy remedies for the authors. **They should enlist a set of volunteers with different computer expertise and hardware to test the accessibility of their data files.** It should also be a simple matter to find a native English-speaking paleontologist who can quickly correct the English grammar and syntax, given an editable text file. I can correct English grammar, but I am worried about database managers and editors whose revised datafiles are unreadable on a very standard Windows computer that can still read the previous version.

Author reply:

Duplicate part of data description, of the whole manuscript, was deleted. Contents were re-organized. And the English wording was checked by software (Grammar) and authors. Graptolite experts who help curating specimens are mentioned in acknowledgments part. Opinion data, or comment from different authors are emphasized and was merged into the dataset.