

## ***Interactive comment on “Distributed visualization of gridded geophysical data: a web API for carbon flux” by K. A. Endsley and M. G. Billmire***

**K. A. Endsley and M. G. Billmire**

endsley@umich.edu

Received and published: 7 November 2015

Thanks for your feedback! We'll address your comments one at a time below.

“The tool is called ‘Carbon Data Explorer’, though the authors indicate that it is not limited to carbon data sources alone. Why not name [it] more generically at this point?”

Other reviewers have also remarked on the name. However, as the software we developed with the aim of visualizing carbon science datasets, and as its features are likely best suited to that purpose, we feel the name is appropriate.

C2817

“In the introduction, one of the key innovations in CDE is that it has a new API for text based representations of data cubes. Does this mean, the data representations are converted to text-based representations internally (as most earth science datasets have large volumes and are not in text formats)? How does this scale for large data? What are the limitations of this approach? Presumably the initial cost of registering the data to the database must be high.”

The data are indeed converted to a text-based representation. This has proven to be adequate global gridded datasets at 1-degree resolution. The CDE was designed for regional and global carbon science datasets at this similar scales. Performance will suffer at significantly higher resolutions. The initial cost of registering the data to the database is indeed high but needs to be done only once by the data manager. Also, the current high cost of registration is largely a reflection of a MongoDB API implementation on our end that is syntactically simpler but not as efficient as other currently available methods. This part of the overall CDE architecture, the “Models” and “Mediators” of the Python API, is amenable to revision by end-users with experience in Python and does not currently represent the potential performance of data registration. MongoDB was also rapidly evolving during our development of the CDE and has since incorporated important performance improvements.

“It is also mentioned that the text based representations has the added benefits of compressing data and enabling rapid filtering and aggregation. Generally text data don't lend themselves to compression formats and methods. How does the JSON document style compare to other typically compression methods (HDF5, grib)? (The article mentions that the data is only 'slightly' compressed).”

The use of the word “compress” in the manuscript was an error and has regrettably

C2818

caused confusion. We will clarify this in revising the manuscript. The text-based representation is not compressed relative to the original data, rather, we have eliminated redundancies that would come from a naive implementation of the dataset as text. Specifically, the spatial structure of the data have been separated from the measurement values so they can be transmitted to the client separately and without redundancy.

“The background makes no mention of similar systems that have been developed and are being widely used. NASA itself has a whole host of similar tools (Giovanni, Mirador, etc.). It will be good to describe CDE in the context of such tools and by describing how different CDE is from these tools.”

We will place CDE in the context of the tools you mentioned in our revisions. Thanks for your suggestions!

“The authors claim that the tool supports scalable analysis which is very important when working with large datasets. Can you include some computational estimates that demonstrate this fact?”

We will include performance metrics in our revisions.

“The data example shown in the paper is very coarse (1 deg x 1 deg) and is not representative of modern day satellite products that get down to resolutions of meters (SRTM, MODIS, etc.). To really claim that this technology is viable for such large data, examples should be presented using such datasets (and with associated computational estimates).”

We did not mean to imply that SRTM or MODIS data would be viewed in the CDE. Rather, “geophysical variables...derived from Earth observation satellites or models”

C2819

are the expected data. OCO-2, rather, is an example of the satellite platform we had in mind. Regional and global gridded products derived from satellite measurements, such as bias-corrected carbon concentrations, are the kinds of datasets for which the CDE was developed. We will strike the phrase “Earth observation” from the quoted sentence to help mitigate any confusion.

---

Interactive comment on Geosci. Model Dev. Discuss., 8, 5741, 2015.