

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

Anonymous Referee #2

Received and published: 6 October 2015

The article describes the Carbon Data Explorer (CDE), a web-based API for storing carbon (and earth science) data. It does seem like a useful tool, but the authors could do a better job of describing the uniqueness of this tool compared to similar systems.

* The tool is called 'Carbon Data Explorer', though the authors indicate that it is not limited to carbon data sources alone. Why not name is more generically at this point?

* 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.

C2419

* 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 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.

* 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?

* 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).

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

C2420