

Interactive comment on “A community diagnostic tool for Chemistry Climate Model Validation” by A. Gettelman et al.

A. Gettelman et al.

andrew@ucar.edu

Received and published: 16 July 2012

REVIEWS and REPLIES:

We thank the two reviewers for their comments. We have made the minor changes suggested by the reviewers. The major changes in response to the reviewers' comments are:

1. Revised figure 2 with common vertical scale and revised caption. 2. While quantitative performance metrics can be calculated by the tool, they are not a major part of the tool. Accordingly, we have de-emphasized quantitative performance metrics in the revised text by: (a) removal of mention from the abstract, (b) reorganization and reduction of discussion in section 3.4, and (c) shortened discussion with further caveats in

C426

section 4.3

We thank the reviewers for their careful consideration of this manuscript, and think that the changes have improved it.

Anonymous Referee #2 Received and published: 6 July 2012

The paper describes a modular analysis and plotting tool mainly aimed at climate and chemistry-climate models. It explains well the workflow and the requirements for its application. Examples are provided to illustrate the capabilities of the tool. General:

I am sure the tool will prove useful in tracing and characterising model changes (as are some other tools), but I would have preferred a slightly more neutral way of describing this fact. Not every model change will result in a model improvement (a word slightly overused in this paper); I would encourage the use of the more neutral word change instead of improvement.

» We have replaced the word 'improvement' with 'change' in the abstract and in the introduction to section 4 and section 4.2. The word 'improve' or 'improvement' appears 2 more times in section 4.2, but here we are specifically discussing the fact that models have improved.

Given the undoubted abilities of the tool to summarise and compare physical quantities I found the focus on performance metrics slightly odd.

» Quantitative metrics are a minor part of the text, consisting of one figure and a 1 paragraph sub-section (4.3) with a description in section 3.4. They are not a focus of the paper. They are not mentioned in the conclusions. To further de-emphasize them, we have removed the sentence from the abstract that discusses quantitative metrics we have shortened the paragraph in Section 4.3, and added caveats and further explanation to 4.3 and section 3.4. Section 3.4 has been reorganized to highlight the quantitative trends calculations. We think these changes should satisfy the reviewer.

Quantifying model changes and illustrating model improvements are non-trivial tasks,

C427

and even though those are helped by tools such as the one described here, the overall model assessment will still rely on the choice of diagnostics used.

» We agree with this statement, and have highlighted this in the text in section 3.4 and 4.3

This choice will hopefully be influenced/guided by the “fit for purpose” principle. Therefore, instead of focusing on performance metrics I would have preferred more details about the practical implementation of derived physical diagnostics. For example another example like the tropopause trend calculation would be nice, maybe deriving streamlines from wind components, or similar.

» Instead of adding more, we have further explained the derived physical diagnostics in another sentence or two, and some other possibilities for comparison.

Even though I do not doubt the statement that “The code can produce performance metrics and is designed to enable comparison of models to observations.” I find emphasising “quantitative grades” (relative to observations) slightly irritating in the context of describing this tool. Obviously, the results of such calculation do not depend on the tool discussed here, but on the observational data used. Some observational data seems to be part of the package, but I wasn’t quite clear about which and how error characterisations are considered, please clarify.

» An extra sentence has been added to explain broadly what data go into the figure, mostly with a reference to the other work that produced it. As noted above, the point is not the quantitative metrics specifically, but that the tool can be used to produce them.

Consequently I would like to ask the authors to deemphasise the quantitative grades and to highlight the ability to characterise changes between model versions. Maybe a comparison of two very similar runs could be shown, where the significance of the zonal-mean zonal wind changes (or another quantity) is assessed.

» We have de-emphasized the quantitative metrics in the paper as noted, including

C428

removing the mention from the abstract, and reorganizing the small sections discussing metrics, and adding appropriate caveats. The current standard data sets do not have individual versions of models, except those submitted to CCMVal-1 and CCMVal-2, and some of these differences are assessed and discussed in section 4.2.

Minor comments:

When talking about the input data for the tool it would be helpful to spell out in a little more detail the link between the CF-netCDF and CCMVal-2 format conventions.

» Done, in response to this comment and the other review: CCMVal-2 format uses CF conventions and names, and adds some meta-data.

Small technical question regarding Figure 2: I assume pressure labels could be defined as non-overlapping? Presumably the data plotted was on a pressure grid and the height scale was converted on the fly using a constant scale height?

» Yes. The height values are a logarithmic interpolation from Standard Atmosphere of the pressure array. This is now noted in the caption for Figure 2.

Please also note the supplement to this comment:

<http://www.geosci-model-dev-discuss.net/5/C426/2012/gmdd-5-C426-2012-supplement.pdf>

Interactive comment on Geosci. Model Dev. Discuss., 5, 1229, 2012.

C429