Geosci. Model Dev. Discuss.,
doi:10.5194/gmd-2016-153-RC1, 2016

Geoscientific Model Development Discussions

Interactive comment on "An unusual way to validate regional chemistry-transport models" *by* Lauren Menut et al.

Anonymous Referee #1

Received and published: 3 August 2016

In this paper, the authors present an extension of the evaluation of (atmospheric chemistry) models by using measurements from other years than the year which was simulated by the model. New scores are introduced to quantify the ability of the model to capture the day to day variability as opposed to persistent patterns.

General comments:

While reading the paper I asked myself the question if the approach presented by the authors has a real added value as compared to a more traditional model evaluation based on bias, RMS and (one type of) correlation, and may be adopted by other groups. In the end I decided that it probably does, for the following reasons:

- The approach proposed quantifies the importance of day-to-day, weather dominated variability versus systematic patterns which are repeated from year to year.

- The approach naturally leads to an overview of the performance for multiple species in one graph (e.g. Fig.5), which is especially also useful (maybe even more useful) for comparisons between different models. This include both trace species as well as meteorological variables. This is a bit similar to the use of Taylor diagrams.

- The approach explicitly exploits both spatial and temporal correlations, which bring complementary information.

- The approach provides new insight into the performance of the WRF-CHIMERE model.

Because of this I am in favour of publication. However, to my opinion there are several major and minor points to be addressed before the paper can be considered by GMD. These are listed below:

- Is this approach really new? The authors provide a few interesting references in the paper, but I would like to see a more systematic overview of the model evaluation approaches and techniques/scores adopted in the past (e.g. including several European/American CTM intercomparison exercises) to better understand the added value of the approach proposed.

- The formulation is incomplete, and mathematical formulas are not well defined. In particular, the authors should provide the equations for R_s and R_t, and the mathematical formula for the MYV needs more discussion, see my comments below. Also, the authors should motivate why the R_s,t scores are chosen.

- The MYV is not really a model score to my opinion, but rather an indicator of how much the score is influenced by day-to-day variability. In particular one can argue that R=1 and D=0 is a good result. Also, I wonder if a formula for MYV is really needed. Showing D and R is maybe enough (see e.g. Fig. 5)? This should be more carefully presented/discussed.

Detailed comments:

p4, l13: “they are used as daily averaged in the present study”: why this choice to focus on daily averages instead of hourly values? Please motivate.

p4, table: Provide also the full names of the variables, e.g. “Temperature at 2m above ground” etc.

p4, last line: replace “same day for another is” by “same day for another year is”

p5, l4: “The correlation is the more appropriate statistical metric for such analysis.” Please explain and motivate this statement in detail. This is important for the rest of the paper!

p5, l8: “The spatial correlation, noted Rs, is calculated from the temporal mean averaged values of observations and model for each location where observations are available.” Please provide a detailed mathematical formula/recipe to be clear. Are observations and model first collocated for individual observations, or are means computed and then compared. Are these means of daily means or means of hourly values? It is important to define precisely how the correlations are computed: the devil is in the details.

p5, 13: Also for the temporal correlation: be more precise. Is it based on daily means, hourly values or something else.

p5, l14: “The longer the atmospheric lifetime of the species, the lower the relevance of temporal correlation” I would dispute this. For long-lived tracers the transport (wind direction) and location/strength of the sources becomes crucial, directly influencing temporal correlations. I suggest to remove this remark.

p5, eq.1: Why is there an absolute value introduced. Instead of absolute(s_i-s_N) I would suggest (S_N minus s_i) assuming higher values of “s” (or “s” close to 1) indicate better performance, which is the case for correlations.

p6, eq.2: Remove the “X” (multiplication) from the formula. This is not needed (in eq.1 there is also no X). Please introduce a one character symbol for the “Multi Year

Variability” instead of writing “MYV” in eq 2, which, in mathematical formula’s means M times Y times V. “D_s” has not been introduced: is it the same as “D” ?

p6, eq.2: Why this complicated exponential form?? It seems that you ideally would have the MYV to be =1 for (s=1 and D=1), and =0 for (s=0 or D=0). A much simpler form s_{MYV} = s_N D would do the trick. In fact, eq.2 is not =1 for s_N=D=1. Where does this formula come from? Is there a reference to a paper introducing this form? Also, it would be good if the formula has clear limits, e.g. 0 (very bad) and 1 (very good). This is not the case when D=1.

p7, l2: Where does the number 0.3 come from? It will depend a lot on how the score “s” is defined. The number seems arbitrarily chosen.

p7, l14: “ . . . is challenging because several uncertainties . . . “

Table 2: It would be helpful to remind the reader that these are Summer periods (1-5 to 1-9) and that the scores are based on daily mean values. Please also highlight the special situation for 2013 (I would suggest to start with 2013, add a thick line, and continue with 2008 2009 . . . Perhaps it can be stressed once more in the caption that observations for 2008-2012 (and 2013) are compared with 2013 model results.

Figure 4: Caption is incomplete.

Table 3: “. . .Values of MYV above 0.3 are shown in bold. . .”

p11, l18: . . .with differences above 0.5. . . .

Interactive comment on Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-153, 2016.