Geosci. Model Dev. Discuss., 6, C1778–C1783, 2013
www.geosci-model-dev-discuss.net/6/C1778/2013/
© Author(s) 2013. This work is distributed under
the Creative Commons Attribute 3.0 License.

Geoscientific Model Development Discussions
Open Access

Interactive comment on "Regional scale ozone data assimilation using an Ensemble Kalman Filter and the CHIMERE Chemical-Transport Model" *by* B. Gaubert et al.

B. Gaubert et al.

gaubert@lisa.u-pec.fr

Received and published: 28 October 2013

1 General comments

A particular strong point is the emphasis on the diurnal cycle of concentrations and diagnostics. The majority of the figures show diurnal cycles, and within the system, hourly profiles for assimilation parameters are set based on diagnostic profiles obtained for the day before. This focus might get some more attention in the abstract and conclusions, since to my opinion this is one of the novelties of this application.

Answer:We thank the referee to appreciate the analysis of the diurnal cycle of assimilation results and diagnostics. We reinforced this focus in the abstract and conclusions.

The experiments described have been performed for a time period of 10 days only. Although this does not hamper the evaluation and the conclusions, some outlook on the performance over a longer time period would be useful. Is the system in this case in particular tuned for a high-ozone episode for example ? Is the performance in area with low ozone (west of the domain) expected to be typical for 'normal' conditions ?

Answer: Following the referee's suggestion, in the revised version of the paper an evaluation of the assimilation system was added for a whole three months summer period in one specific configuration. This paragraph is iterated with the answer to referee #2 (2nd general comment). It shows that performances are even better for the summer period because the analysis errors increase during the ozone episode. Then, it is true that performance over the west of the domain (i.e. with typical background ozone of 30/40 ppb) is typical of the normal conditions.

2 Specific comments Eq. 5 and 6. It is not always clear over which 'p' observations the statistics are computed. From the text it seems that within the assimilation these numbers are computed over all assimilated observations available at a single hour, but for some evaluations of the overall performance it seems also to include the 10 days of the experiment.

P 3040: Diagnostics were always calculated for all ('p') assimilated observations. Diagnosed values were then averaged over the ten days of the assimilation period in figures 8 and 11.

p 3047, line 14. Only ozone fields are included in the ensemble. With 24 ensemble members, the cost of propagation of the ensemble is probably less than the cost of propagation of the full model, is this correct? If this is correct, it makes the system a very cheap assimilation tool. Please add some lines on the computational costs of the system. P 3047:Each ensemble member evolves with the model, so the cost is a bit less than 20 times the propagation of the full model. p. 3048, line 16: "..., we first use for all types of stations an observation error standard devaiton of 5 ppb, ...".

GMDD
6, C1778–C1783, 2013

I guess the "second" choice is the "R_EXPLICIT" experiment in section 5.4, but this is not imediately clear. A table showing the essential differences between the various experiments would be useful.

P 3048:Indeed, in other experiments, this error is diagnosed (section 5.4). However, the R_EXPLICIT experiment has been removed following your suggestion below. This is made clear in the table 2 which shows the different set-ups for the different experiments.

p 3049, lines 16-18. I guess that what is refered to as the "correction" means the innovation from forecast to analysis field. But the spatial extend of the innovation is limitted by the local analysis range of 250 km. Thus, the observed change over the North Sea is the result of transport (this is also what is mentioned in the conclusion at p. 3057 lines 25-27). The same result could be obtained with an optimal interpolation method. I would say that extension of innovation with the flow is only possible if the ensemble size is large enough to avoid spurious correlations and no form of localization is used. Please clarify this statement.

P 3049:Most of the correction in the North Sea results from the transport of innovation or correction. The new sentence has been corrected as following: “However, the spatial shape of the corrections, for instance over the North Sea illustrates the ability of the sequential assimilationto extend innovations along with the ozone flow (in the north-west direction) during the forecast step.”

p 3051, line 5. The temporal profile of the noise in the reference run just a switch between 10% and 20% for day and night. How would the results for the "MOD_DESR" look like ?

P 3051: The average noise for each hour in the MOD_DESR experiment is added on the figure 3 below. The applied perturbations are generally lower, except at 7-8 and 18-19 hour UTC.

GMDD

6, C1778–C1783, 2013

Figure (see below): Prescribed noise in the REF_ASSIM experiment (brown) and average profile (10 days) of the MOD_DESR noise.

p 3052, line 1. It seems obvious that in a region with only a single observation station the spatial impact of this site is large: there are no other observations available that could counter act. I think the main issue here is the presence of a model bias that is persistent over a large area: if this is present, then it is indeed sufficient to have a single station, but only if the spatial scales in the BECM are large too.

P 3052: In fact, when peaks are underestimated, the larger background error variance in the REF_ASSIM experiment gives better results. When only one observation is assimilated the shape is mainly controlled by the inflation factor. We don't believe that these corrections are realistic, particularly for suburban stations (i.e. with less representativeness). For instance the ozone observations in Madrid or Athens are not representative for the surrounding rural areas.

p 3056, eq8 . The formulation of a spatialldependend observation error in Eq. 8 feels a bit as violating the idea of a Kalman filter. The spatial relations between grid cells are supposed to be described in the P matrix in eq. 3, while the relation between observations is in the R matrix (which are usually set to zero). The impact of an observation site on a grid cell further away is small if the gain matrix is small, and this is most controlled by the P matrix; eventually localization is applied to explicitly limit this spatial correlation at large distance. The observation error matrix R is not changed by a localization. But in the choosen formulation, R has become a mixture of observation properties and spatial (physical) properties. It almost feels as if the authors have constructed a fixed gain, that is very similar to what you get from an optimal interpolation. What benefit from the ensemble is left here ? Please clarify.

P 3056: It is true that this formulation is a bit violating the formulation of the algorithm. We decided to remove this experiment in the new version of the paper.

3 Technical corrections

GMDD

6, C1778–C1783, 2013

CC BY

p. 3035, line 24: are the '\dots' intended ?

P 3035: These dots represent air quality indicators such as evaluation of human health or vegetation impact indicators.

Interactive comment on Geosci. Model Dev. Discuss., 6, 3033, 2013.

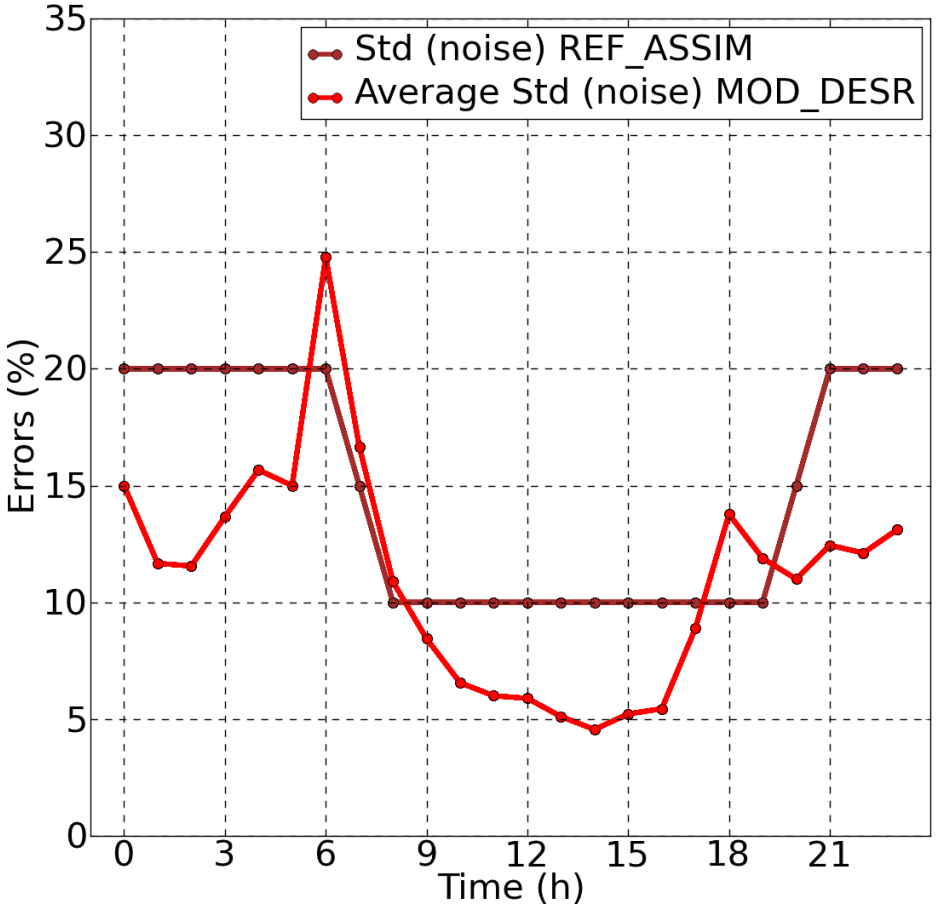

Fig. 1.

GMDD
6, C1778–C1783, 2013

CC BY