**Geoscientific
Model Development
Discussions**

Interactive comment on "Impact of a time-dependent background error covariance matrix on air quality analysis" *by* E. Jaumouillé et al.

Anonymous Referee #2

Received and published: 15 June 2012

The subject of the present paper is the data assimilation of surface ozone measurements to correct simulated surface ozone fields in the context of air quality monitoring and forecasting. More specifically the paper presents how the background error covariance matrix (hereafter called BECM) can be specified in such assimilation systems using aposteriori diagnostics and/or ensemble approaches. It shows that the BECM is time and space dependent. Various experiments have been done on a 10 days period to show how the formulation of this BECM can impact the analysis but also the ozone forecast. The topic of the paper is of great interest for the air quality/assimilation community. The paper is in general well written and very educational on some points (i.e a good equilibrium is found to present theoretical aspects). The approach of the study

is sound. For this reasons, it deserve to be published in Geoscientific Model Development. Nevertheless, i think that some aspects of the paper could be improved especially authors should clarify more carefully the position of their work/results compared to the current knowledge of the community. Moreover, considering that the results do not show clear improvements of the system with the various specification of the BECM (which is a result in itself), we do wait more analysis of these results. That is why i recommend some minor revisions in the following that could, i hope, increase the interest of this paper for the community.

Minor revisions

p875 – lines 17-26: maybe some references are missing concerning air quality and especially the way that they are prescribing their BECM (Elbern et al, Blond et al, Wu et al . . .).

p883: concerning the hypothesis made in equation 13, is there a way (a priori or a posteriori) to verify that this hypothesis is sound?

p883: authors explain that the estimate of the background error correlation is out of the scope of the paper. It is a bit disappointing for the reader but can you at least discuss the way it could impact the results of the assimilation? Have we some insight on the respective weight of variances and correlation formulation on the assimilation results? Is it a way that could explain that results remains close what ever your BECM choice is? If the answer is yes maybe it should be mention in the conclusion.

p887 –section 4.1: Here authors are describing how their system (ensemble and choices for OECM and BECM) is built. Some choices should maybe be explained such like 50% for emissions perturbations, 45km for length-scales, why only emissions are perturbed? is it enough? Probably that reference to the existing literature would help (Hanea, Boynard, Garaud, Mallet). Moreover, the results of previous analysis tend to show that OECM > BECM, is it meaning that the system tends to have more confidence in the model than in the observations? Could authors discuss this? Lastly

for this section, authors explain that ocean emissions are missing? Considering the amount of NOx emissions by ships is it not a problem (especially along coastal areas)?

p887-888-889– section 4.2: Authors consider a 10 days period for their study. Such a period do not sample a statistically representative number of synoptic situations. Now, we do suspect that synoptic situations (stream direction, cloud cover ...) could modify the length-scales. I suggest that authors discuss this in this section. Maybe, it would be interesting to know the meteorological conditions of this period. It could help to interpret more deeply the results.

p892âĂŤline1-11: It will improve the section 5.2 if authors could recall (briefly) why they use this Joly and Peuch classification instead of the metadata classification proposed classically by airbase. I think it is obvious for the authors but it could be an interesting point for less skilled readers.

p893âĂŤline 25: Why authors do choose this particular date? Is it representative of the whole period? Is is a particular situation?

p895âĂŤline 8-10: I think that it would be a real plus for the paper if authors could discuss the fact that it is difficult to distinguish the best configuration of the BECM. Do authors have any leads to explain that? You do not consider different formulation of the correlation, could it be important? Other things ??? Whatever author can propose maybe they can emphasize more on the fact that there are few differences in their results as a function of the BECM choice. Indeed, it is valuable information for other group working on such topic.

Section 5.5: I'm surprised that authors choose so few validation stations. How the choice of the stations has been done? I can understand this choice for operational set-up but for this study it appears not so well appropriate. Authors should discuss this point. We can imagine that if the station network is dense enough the choice you are making to prescribe properly the BECM especially concerning length-scales is useless. This will be impossible to verify if validation stations are few and located just

in the neighbourhood of assimilation stations. It would be interesting to consider this aspect.

Section 5.6: This section as well as the corresponding conclusions missed comparisons with previous studies of this type. Considering papers of Blond, Elbern will help authors to strengthen their analysis and conclusions on these forecasting aspects.

Remarks and technical corrections

p875 - line 7 : "...the best estimate of the physical state given the input ..."

p877 – line 22 : observation instead of observations

p881 – line 22 : flow-dependent instead of flow- independent ?

p884 – line 5 : maybe a web site or a reference on the MACC project is needed.

p884 –line12 : I understand that authors do not spend too much time on the description of tropospheric ozone but maybe that an "English" reference could also be added there such like the books of Seinfeld & Pandis or Finlayson-Pitts or Warneck.

p885 –line 9-19: Can you specified the reference there where the assimilation system your are using is described in details?

p885 –line 16-17 : "...mesurements are available at hourly frequency from the MACC project, ..."

p885 –line 21: If you do specify the web site and more specifically the link toward the informations about ground stations, this sentence is almost useless.

P885 –line23: I do not think that you need to show the station's location for both summer and winter. One of these two figure is enough. If you want to underline differences between winter and summer (is it significative, I don't think though?) just tell it or show a map of differences.

p885: You should probably mention what kind of ground stations are used and from

which database.

p894âĂŤline 10: a word is missing after quite ?

C300