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.

Anonymous Referee #2

Received and published: 23 July 2013

1   General comments

The authors propose a study on ozone at the regional scale. They built their study on surface ozone measurements from the European database Airbase and on the CHIMERE regional chemical transport model. To better constrain their model, they adapted a local Ensemble Kalman filter (EnKF) square root scheme previously developed to assimilate tropospheric ozone columns, so that they can assimilate surface ozone measurements.

Airbase provides information on the type of each measurement. But the authors de-

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

cided to produce their own classification of the measurements by further developing the idea of Flemming et al. (2005). They therefore derived four station types: background/mountain stations, rural stations, suburban stations and urban stations. Their approach seems easy to implement and seems consistent with the Airbase classification. So it should be useful for everybody looking for comparing surface ozone data with model outputs.

Among the Airbase ozone data, they selected randomly about a half for the analysis and the other half for validating their simulation (without assimilation) and their analyses with various configurations. The different configurations they proposed aimed at providing insight of the sensitivity of their EnKF to the background and observation error covariance matrices. The evaluation of their analyses is provided for a period of 10 days during summer 2009 when high surface ozone concentrations were observed. The discussion is well conduced by separating the results of the day from the results of the night or by looking at the station types. Nevertheless they assumed that the weak sensibility of their system to the background and observation error statistics proves the robustness of their system. But they did not discuss why they have such a robustness. Is it due to the large amount of assimilated observation? Or is it due to some property of the EnKF?

As they conclude, they demonstrated that their system is robust enough and suitable for implementation in operational systems supported by the European FP7 MACC project. My main concern is to know if a 10 days period is enough to demonstrate the ability of the system? The chosen period furthermore corresponds to an episode of high ozone concentration when the model already perform well without assimilation. It would have be useful to have a winter period as well in the study even if this period of the year has less media exposure (the ozone concentrations being generally lower). My other concern about this 10 days period is to know if the restriction to 10 days is due or not to computational costs? If not, why the model is evaluated with a period of 3 months and not the analysis? If yes, it is an issue to operationally run the system in MACC?

I finally strongly suggest the author to pay more attention at the structure of the paragraphs (even of the document itself sometime) and at the English language. See few examples in specific comments hereafter. Concerning the structure of the paper, section 2 could be divided into two parts, one on the EnKF and one for the model description (currently section 4.1) as the model is a part of the assimilation system. The current section 4.2 could be include in section 5. Section 5 should also contain a subsection with a better description of the period chosen for the study as well as the repartition of the stations (by type and by purpose - assimilation or validation) for these periods. Note that this repartition is already described in the first paragraph of section 5.1.

2 Specific comments

- page 3034 lines 2 to 4: "The Ensemble Kalman Filter is ... analysis step." This kind of sentences are useless in an abstract.

- page 3034 lines 6 to 9: "The analyzed ozone field is evaluated ...". There are three ideas in this sentence: 1. the comparison with consistent set of observations (do you mean "ozone ground measurements" by observations?); 2. the reduction of the quadratic error and the improvement of the hourly correlation coefficient ; 3. an improvement despite a low ensemble size. Please split the sentence to highlight these 3 ideas.

- page 3034, introduction: you write that ozone impacts human health and vegetation growth. You have then one sentence to explain the impact on "human health" but nothing for the impact on the "vegetation growth". Please think of adding a sentence to explain this impact.

- page 3035 line 1: GMES is not focused only on "ozone and other pollutants" as suggested by your sentence. GMES services are indeed Land Monitoring, Marine

Interactive
Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

Monitoring, Atmosphere Monitoring, Emergency Management, Security and Climate Change. Please correct.

• page 3035 lines 9 and 10: What is the difference between "direct model simulation" and "forecast"? What do you mean by performing a forecast in synergy with observations?

• page 3035 line 16: Please precise here what you mean by "analysis". Is it a result of a data assimilation system? In your document, does "analysis" always refer to as the result of a data assimilation system?

• page 3035 line 21: What is the "analysis product"?

• page 3035 line 23: What are these "reporting activities"? On which analysis are they based? CHIMERE one?

• page 3035 line 26: What in the link between the first sentence of this new paragraph and the following sentence?

• page 3036 line 9: Are Courtier et al. (1998) and Houtekamer and Mitchell (1998) really the right references for pioneering work in data assimilation for numerical weather predictions? Without going to Kalman (1960), pioneering works were realised by Talagrand and Courtier (1987) and Courtier and Talagrand (1987). Moreover, it looks like Elbern et al. (1997) was a study done before the pioneering work for numerical weather predictions you assumed to be done in 1998!

• page 3036 line 27: In Geer et al. (2006) for example you can see that there is more than only two major strategies that are employed to obtain an accurate 4D-analysis of ozone concentrations. Please rephrase. Moreover, in your sentence, you forgot to mention that the 4D-Var analysis allows the correction of the initial condition.

• page 3040 line 16: What do you mean by "these" in the sentence "These are derived from ..."? Do you mean that in your study the ensemble of perturbed ozone concentration fields are derived from a two-dimensional Gaussian distribution? In this case, why

do you add the reference to Evensen et al.(2003)? And you also should specify what you are using as a variance and spatial correlation.

• page 3040 lines 21 to 22: Diagnostics from Desroziers et al. (2005) allow to derive the forecast or the background error variance (in the observation space) not the model error variance. Please correct here and all over the document if needed to better differentiate background error and model error.

• page 3040 line 24: The whole paragraph (and the section as well) deals with the BECM and suddenly you give an equation for the observation error. You could introduce this equation later within a paragraph on the observation error for example.

• page 3043 line10 : Is the "daily average ozone profile" also averaged over the whole domain? If yes, please specify.

• page 3044 line 20: What are the horizontal and vertical resolutions of meteorological variables from ECMWF? What are you using as spatial interpolation?

• page 3045 line 7: The selection of the period you propose to study is an important information. I found it strange to find the information for the period for which assimilation is performed in parenthesis. And I found it strange to find this information in the section of the evaluation of the reference simulation. I suggest you to create a new section (or subsection) in which you could describe the period you studied and the reasons why you have chosen that one.

• page 3045 line 10: Fig. 6a displays the ozone fields for 15 August but you introduced it in a sentence where you discussed the period between 19 and 21 August.

• page 3045 line 21: Table 4 does not support the fact that you have a minimum of the daily maximum during the episode.

• page 3045 line 24: I guess that Tab. 3 contains statistics only for the assimilation period and not for the whole summer. Please make it clear that you are not discussing anymore about statistics over the whole summer (as suggested by the first sentence of

your paragraph).

• page 3048 line 11: What do you mean by "observation error of 8ppb at 60km horizontal resolution"?

• page 3048 line 25: I suggest you add the reference to your analysis, i. e. "the analysis from REF_ASSIM experiment". As you gave names to your analyses experiments, please always refer to these names in your text.

• page 3049 line 17: I found it difficult to see "the ability of the ensemble to extend innovations along with the ozone flow" with the figures you proposed. For instance, the differences in Spain for example between Fig. 6a and Fig. 6b can come from the transport from France as the ozone concentrations should have been modified in France the previous assimilation cycles due to the numerous stations there. For me, you did not demonstrate the "clear advantage of the Kalman filter".

• page 3049 line 23: Does the "8-h mean maxima" is the same as the "daily maximum of the running eight hour mean" of line 2 page 3050? If they are the same, please explain its role ("indicator for the impact of ozone exposure to human health") before in the text. This could enter in a section "evaluation method".

• page 3051 line 28: It is difficult to see differences between Figs. 6b and c.

• page 3052 line 26: You present the RCRV you have already introduced in the same section but earlier (page 3050 line 27). You could introduce it once and provide the information in a section "evaluation method".

• page 3056 line 6: What sd_{meas} stands for?

• page 3056 Eq. 8: Please explain the coefficient 1.2 in the equation.

• page 3070 Tab 3: Why are you changing the the name of your REF ASSIM experiment with ANA just for this table? Please keep REF ASSIM and do not introduce a new reference.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

• page 3071 Tab. 4: Please specify in the caption the period you used to compute the statistics. Be careful you did it in each caption of each table or figure.

• page 3075 Fig. 3: Please rephrase the caption. Moreover, please add the label (a) and (b) on the figure.

• page 3079 Fig. 7: Please specify the analysis (REF ASSIM?).

3 Minor revisions/comments

• page 3034 line 13: I would suggest "a posteriori diagnostics" instead of "a posteriori Desroziers diagnostics" as "Desroziers diagnostics" is not a generic term. But further in the text, I would suggest to keep "Desroziers diagnostics" if you precise clearly that by this you mean the diagnostics from Desroziers et al. (2005).

• page 3034 line 18: "parameters" instead of "parameter".

• page 3038 Eq. (2): Please add a transpose sign on the last right parenthesis.

• page 3040 line 21: "diagnostics" instead of "diagnostic".

• page 3041 lines 2 to 3: I suggest to rewrite the sentence with "These quantities have to be computed a posteriori because they require the analysis y^a (in the observation space), the observations y^o and the ensemble mean of the forecast y^f (in the observation space)". Moreover, link this sentence with the following or start a new paragraph.

• page 3041 line 10: Please provide the number of the section where this "has been demonstrated".

• page 3042 line 18: Please rephrase. I suggest: "In this paper, we choose this approach because it has the advantage of being directly targeted on the pollutant species of interest (ozone) and it does not require any other metadata."

• page 3044 line 17: As the aerosols are not included in our simulation, I suggest to remove the reference Bessagnet et al. (2004).

• page 3071 Tab 4: "averaged" instead of "average"

• page 3073 Fig 1: Could you please remove the background grid.

• page 3083 Fig. 11: The legends on the right of the figure and on the top left are difficult to read.

4 Bibliography

• Courtier P. and O. Talagrand: Variational assimilation of meteorological observations with the adjoint vorticity equation. II: Numerical results, Quarterly Journal of the Royal Meteorological Society, 1329-1347, 1987.

• Geer A.J. and W.A. Lahoz and S. Bekki and N. Bormann and Q. Errera and H.J. Eskes and D. Fonteyn and D.R. Jackson and M.N. Juckes and S. Massart and V.-H. Peuch and S. Rharmili and A. Segers: The ASSET intercomparison of ozone analyses: method and first results, Atmos. Chem. Phys., 6, 5445-5474, 2006.

• Kalman R.E: A new approach to linear filtering and prediction problems, Journal of Basic Engineering, 82, p 35-45, 1960.

• Talagrand O. and P. Courtier: Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory, Quarterly Journal of the Royal Meteorological Society, 1311-1328, 1987.

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