Geosci. Model Dev. Discuss., 6, C1076–C1085, 2013
www.geosci-model-dev-discuss.net/6/C1076/2013/

Geoscientific Model Development Discussions

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 #1

Received and published: 22 July 2013

Review of

Regional scale ozone data assimilation using an Ensemble Kalman Filter and the CHIMERE Chemical-Transport Model

by

B. Gaubert, A. Coman, G. Foret, F. Meleux, A. Ung, L. Rouil, A. Ionescu,Y. Candau, and M. Beekmann

General comments:

The paper presents an application of an Ensemble Kalman Filter using the CHIMERE

regional Chemical Transport Model for the assimilation of hourly surface ozone observation from 350 stations in Europe (Airbase). The period for the assimilation was 10 days in August 2009. A reference run (without assimilation) was also carried out for June, July and August 2009.

The authors take a scientifically sound approach in the treatment of the ozone observation since they stratify the data according an existing air-quality-regime classification (FLEM05) in mountain, rural and urban regimes. The stratification is useful for the evaluation and is used (to some extent) for the specification of the observation errors statistics, in particular the representativeness error. The validation applies a cross-validation approach by randomly splitting (albeit one realisation only) the observations in one ensemble for the assimilation and one for the evaluation. The evaluation focuses on the diurnal cycle, which is an interesting aspect. Sensitivity studies about the specifics of the background error and observation error statistics are carried out, but overall there seemed to be only little impact. An interesting comparison of different formulations of the observation error variance is presented based on Reduced Centred Random Variable Diagnostics.

The authors test basically only the ability of the Kalman Filter to interpolate dense surface observations in the spatial domain. Use of the analysis as initial conditions for forecast is not considered in the paper. One could argue that mapping algorithms such as Kriging or Optimum Interpolation would have done a similar job to interpolate these observations. The presented study somehow fails to convince why a 4D data assimilation method was applied to the problem, i.e. the interpolation of a rather dense network of ozone observations to 0.5x0.5 grid. The study period (10 days) is too short for air-quality assessment according to EU guidelines.

The authors compare the ozone surface analysis with ozone surface observation, which were not assimilated, but these observations are still very close to the assimilated observations and probably highly correlated, in other words, they are not really independent. This might has consequences for the assumption that observation er-

rors are not correlated. See for example Liu and Rabier (2001) for a discussion of the problem. I also think that the large number of assimilated observations is actually not helpful for a key point of the study i.e. to test the impact of different formulations of the observation and background errors. The presented ozone analysis is constrained by the large number of observations and not sensitive to changes in the observations errors. It would have been helpful to thin the assimilated observations further and further to see at what point the analysis quality starts to deteriorate and if at this point changes in the background observation error formulation make a difference. The question would be can a Kalman-Filter make better analyses with a thinned network than an interpolation method ?

The model and the assimilation show greater errors during the night. It could be related to the vertical resolution of the model. This aspect is hardly discussed in the paper. The strong vertical gradients of ozone during the night might not be resolved. This has consequences for the vertical representativeness of the observation during the night.

Please differentiate more clearly between the “model error” (no assimilation) and the “background error” (forecast started from the previous analysis).

The text would have greatly benefited from further proofreading both for the English language, consistency of wording and structure of the document. For example, headings 4.2 and 5.1 have the same title “Evaluation of reference simulation”.

Specific comments

P 3034 L1 The abstract should avoid general statements but should clearly state the purpose of the paper and summarise methods and results. I would omit the first sentence but say what the purpose of the study is. L7 “quadratic error” or RMSE ? – the latter is more common L9 please make clear that you use an existing method for the classification L26, “which increases . . .” omit – it is obvious and the relation is not linear over all possible ozone concentrations.

P 3035 L1 provide reference for GEMS, L3 “ and in-situ ..” L6 delete “the” before “monitoring” L10 What is a “modelling platform”, please clarify , Do you mean a model ? L16 please clarify that by “analysis” you mean the result of a data assimilation method in the remainder of the test L23, why are the analyses “ a fundamental result” – Consider reformulating the whole sentence

P 3036 L16 please clarify the difference between model error statistics and background error statistics L28 There are variational methods (4DVAR) and sequential Extended Kalman Filter methods. The Ensemble Kalman filter is just one example of the possible Kalman Filter approaches. Both methods have been used to correct emission rates. Brunner et al. 2012 and Miyazaki et al. 2012 are important examples which should be mention in the paper.

P 3037 L18 It would be good to mention the purpose of the study much earlier, i.e. at the beginning of the introduction.

P 3038 L2 Please clarify if you mean all Extended Kalman Filters or just the specific Ensemble Kalman Filter implementation. L14 Clarify what is x in your study, only the surface ozone field ? Eq 2 for Pf is not a notation that leads to the covariance matrix. It is only the variance. One has to use the transpose and/or a second index. P 3039 Eq 2 to 4 are the equations for the optimum interpolation. The forecast of the state vector and P with the model M are important steps of any Kalman Filter algorithm. If the forecast of P with the model M is not part of your EnKF, please make this clear. The forecast of P usually contains a noise term (Model error), which makes sure that P does not deflate. L22. In all data assimilation approaches the observations have a random error – please clarify what you mean.

P 3040 L2 unclear sentence , also please correct “measurement’s perturbations” L4 again, this a property of all Extended Kalman Filters. Do you use Ensemble Kalman Filter as synonym for Extended Kalman Filter?

P 3041 L5 Please explain what the Desroziers Diagnostics tell us, i.e. what they mean.

L10 What “assimilation exercises”? Please provide reference. It is not clear in the whole section if you make statements based on references in the literature, or if you want to test these assumption later in the paper. L16 How do you know it is reliable for a dense network, please provide reference.

P 3042 L14 It is not clear how the classification (either based on observation time series or meta data) is used to quantify the representativeness, which is needed for data assimilation. A rural station might have a larger area of representativeness but the actual values (100 , 50 or 10 km?) are not provided by the classification. L26, Why are you using a new label for the air-quality regimes in FLEM05. Later in the text you only use the term “urban”, “rural” etc. Please be precise and consistent.

P 3043 L2 Do you calculate P50DA and P50DV from the annual values or only for the JJA season (Table 1 seems to indicate that). L6 all urban stations according to FLEM05 or to Airbase ? How do you justify this assumption that they are not representative for the 0.5 x 0.5 grid. The urban emissions should be part of the emissions in any grid box containing urban stations. L10 It would be good to give some indication of the observation density, i.e. something like the average distance to the nearest neighbour station, both for the total observation set and the split one. L18 what “variability” - daily variability ??

P 3044 L15 not sure what GEMS means here. Please provide vertical extent of the eight layers in m. In particular the height of the surface layer is importance for the interpretation of the study L18 please rephrase “mandatory” L20 “Analyses” use plural L24 Is it MOZART 3.5 ? please double check, also the reference for MOZART 3.5.

P 3045 L24 Is it a temporal correlation of the hourly values - please specify. L27 Please discuss this also w.r.t to the vertical resolution of the model, i.e. the depth of the surface layer.

P 3046 L5 Why did you not choose an appropriate model level for the mountain stations? Would have a different level with a better match during the night led to worse

results during the day? L8 Please consider also the NO gradients in the nocturnal PBL. L19 -24. I believe this short introduction of the chapter is not helpful and can be deleted. The titles of the 2nd level headings should be clear enough. Please keep explanation of Table 2.

P 3047 L1 Please check the title “Evaluation of the reference . . .” – it should be “setup of assimilation experiment “ L5 - It would be good to give some indication of the observation density, i.e. something like the average distance to the nearest neighbour station, both for the total observation set and the split one. Even after splitting in the observations in evaluation stations and assimilating stations one can not assume (as it is sometimes implicitly done) that the evaluation stations are truly independent of the assimilated stations. L13. The PBL heights (I suggest to use this term rather BLH) are a crucial point in your method and needs more explanation. How is PBL height diagnosed. How is it defined during night time conditions, when there is no mixed layer? What happens if the PBL height is smaller than the height of the surface level? L14 please correct “tri-dimensional” L17 please give value (in ppb) of the magnitude of the added noise, How does it compare to the observation error standard deviation.

P 3048 L9 Above you say that the classification of the station regime is used to define different representativeness error – now all stations get the same value (5pp). Is contradicts a bit the statement before. Do you distinguish between observation error (instrument error) and representativeness error. What is the assumed resolution for the estimate of the representativeness error in Flemming (2004). In your case you also need to discuss the representativeness in the vertical direction, in particular during the night. The vertical ozone gradient can be quite pronounced. The vertical representativeness also raises the issue if the error is only “random”. I would argue it is more bias, which needs to be corrected before the assimilation. L27. The caption of the Figure say 3.00 UTC – here you say 15 UTC!

P 3049 L1 What about the low values over the North Sea in the analysis. Are they realistic or a consequence of the extrapolation of urban stations in the UK? Night time

values over the sea are in general higher than over land because of the reduced dry deposition. L8 How far away is the nearest assimilated station to Odense, what type is it ? L18 I am not convinced that OI (i.e. a Kalman Filter run without update of P) which includes a model would not be able to transport the information from the observations to other area. Is there evidence that the change over the North Sea comes from the analysis step in the KF and not from the model forecast started from an analysis. (The latter would also work with OI). Finally, as said above, it is not clear if the changes over the North Sea are actually an improvement.

P 3050 L7 “error profile” is perhaps misleading, “diurnal cycle of error” is a better way of calling what is shown in Fig. 5 L7 “globally” ? use “overall” or else L15 It is not the “model error” but the something like the “background error covariance description” L20 again it is not the “model error” but the “background error”, i.e. the error of the model started from the analysis one hour earlier L24 a bit more detail of what the RCRV is required at this point (or later)

P 3051 L1 it would be interesting to see graph of the original BECM diurnal cycle (perhaps the diagonal elements average) and the modified one of MOD_DESR. One would like to know if the BECM changes a lot or only a little. L10 The following discussion seems interesting but is difficult to follow. Perhaps you should provide formulas for “diagnosed errors”, “ensemble standard deviation “ – is it before or after the analysis step. L12 “model error” ? Do you mean “background error” ? L13 “shape of the diagnostics” more realistic than the “model error” – I don’t understand this L16 Please use a consistent Figure order in the text. L19 Do you want to say that the BECV of MOD_DESR are better (because the analyses have a smaller error?) during evening and morning than the standard case. If yes, are the analysis errors smaller during this time?

P 3052 L8 Please mention in the text what parameters you perturb. Reference to the supplement is no sufficient. Please include the table in the main paper. L9 Could not find reference for Hanna et al. 2001 L26 – please remove brackets. It would be better to introduce RCRV this point and not already on page 3050 without explanation

P 3053 L11 – please check sentence. L12 Do you mean that biases are not taken into account ? L19 “fading” ?, now you use the term PBL without explanation , before it was BLH.

P 3054 L9 “error budget” perhaps better “behaviour of error statistics” L21 I don’t understand this. You said the spread of the ensemble needs to be increased before, now it is “preferable to reduce the perturbations”

P 3055 L1 If you change the observation error for the assimilated stations, this also needs to be reflected for the evaluation. Hence, a larger observation error (that’s is what we assume to be the correct value) should also mean that a certain RMSE or bias of the analysis is less problematic. One can not conclude that a smaller analysis error is a confirmation that the chosen observation error is more correct. L6 please discuss also the vertical representativeness L28, RMSE, bias, Correlation are no “skill scores” - they are accuracy measures. Skill scores compare accuracy measures against a reference (see the textbook by Wilks “Statistical Methods in the Atmospheric Sciences”)

P 3056 L1 I believe the little impact can be explained with the large number of assimilated observations. They dominate the analysis and the error description is of minor importance. L2 The evaluation stations are not independent of the assimilated stations because they come from the same network. A smaller analysis error (if we assume the observation error is larger) is not an indication that the analysis is better.

P 3057 L10 Again the numbers of the assimilated stations are high. One can perhaps not expect too much. L13 “new” I understand that the Kalman Filter was already developed and used in Coman et al. 2012. So please be specific about what is new. L13 “chain” is unclear L22 Please discuss possible causes for the night-time problem. L25 Please mention how many observations sites you assimilate and what the average density of the assimilated stations is.

P 3058 L9 Please discuss the differences between your approach and Hanae et al. 2004 in more detail. Both papers assimilates surface ozone in Europe using a Kalman

Filter. This should also be mentioned in section 1 (Introduction) L13 The sentence "Considering .." does not make sense to me. L15 please describe the RCRV briefly with half a sentence or so. L24 How were the observation errors estimated? Give value of 5 ppb for the standard deviation. L24 Please mention that the RCRV diagnostics have indicated higher errors for the rural stations. This could be an important finding of the paper L27 Again they are no "skill scores" but accuracy measures. See textbook by Wilks.

P 3059 L7 Not just the robustness of the system but the very high number of stations. If you reduce the stations numbers you will see a much larger effect of the error statistics on the results. L13 Please mention references of papers on this topic. (Brunner et al, 2012, Miyazaki et al. 2012) L18 Is it also used for NRT forecasting? It would be interesting to discuss this in the paper.

Table1 Please spell out MOU, RUR, SUB. It is confusing that you introduce new abbreviations for the FLEM05 classification, and sometimes only the adjectives "rural" etc. Please be consistent throughout the paper. Are the mean and P50DV valid for the whole year or only JJA?

Table 2 Spell out OECM and PECM What means fixed (= constant)? Why profile ?Provide indicative value in ppb for all variances.

Table 3 Not skill scores – it is better called accuracy measures

Tables 3 and 5 and also 4 should be merged into on table.

Table 4 – why no discrimination of different regimes?

Table 5 "Accuracy measures" - Explain MOU, RUR SUB etc.

Fig1 I found it difficult to see the cyan square for Odense.

Fig2 Explain MOU, RUR, SUB

Fig3 Is the top CHIMERE identical to Fig 2 ? Perhaps Fig 2 can be omitted. Why is

SUB obs in red and the rest not ? Mention length of assimilation period.

Fig 4. What is "prescribed noise". Green and blue colours are difficult to discriminate. Consider changing the colours.

Fig 5. It seems that some of it is already shown in Fig. 3.

Fig 6. The shapes are difficult to distinguish. The text says the time is 15 UTC.

Fig 9. Please spell out RCRV.

References: Brunner, D., Henne, S., Keller, C. A., Reimann, S., Vollmer, M. K., O'Doherty, S., and Maione, M.: An extended Kalman-filter for regional scale inverse emission estimation, Atmos. Chem. Phys., 12, 3455-3478, doi:10.5194/acp-12-3455-2012, 2012.

K. Miyazaki, H. J. Eskes, K. Sudo, M. Takigawa, M. van Weele, and K. F. Boersma: Simultaneous assimilation of satellite NO2, O3, CO, and HNO3 data for the analysis of tropospheric chemical composition and emissions, Atmos. Chem. Phys., 12, 9545-9579, doi:10.5194/acp-12-9545-2012, 2012.

Liu, Z.-Q. and Rabier, F. (2002), The interaction between model resolution, observation resolution and observation density in data assimilation: A one-dimensional study. Q.J.R. Meteorol. Soc., 128: 1367–1386. doi: 10.1256/00359000232

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