Atmos. Chem. Phys. Discuss., 11, C2563–C2566, 2011
www.atmos-chem-phys-discuss.net/11/C2563/2011/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Improvement of ozone forecast over Beijing based on ensemble Kalman filter with simultaneous adjustment of initial conditions and emissions" *by* X. Tang et al.

Anonymous Referee #1

Received and published: 2 May 2011

Tang et al. describe an application of data assimilation to improve Ozone forecasts in the Beijing area. From the methodology point of view the work is not on the cutting edge. The Ensemble Kalman filter algorithm used, the model error term introduced and the localization method used have all been employed before for Ozone prediction problems (see also the references mentioned by the authors). The interesting part might be the application to the Beijing area. But then it should become more clear what we learn from this specific application.

The presentation of the methodology is a weak point. It isn't sufficiently related to other work and the justification is iffy. In my view the paper is not yet ready. I suggest a major

revision.

Major comments

The notation used for the state space formulation of the model is not clear. To achieve a better understanding of the paper an established notation should be used. For applications of data assimilation in the field of weather prediction, oceanography and air quality there is a standard notation that is often used: See "Unified Notation for Data Assimilation: Operational, Sequential and Variational.", K. Ide et al, 1997.

An important issue in data assimilation is the representation of the model error. If one starts with reading the equations (1) and (2) and then read the equation (7) for the model error it is not clear how the model error is related to the state vector and how the model error is actually implemented. The use of an established notation could help to solve this problem.

The choice of the Ensemble Kalman filter algorithm used is not motivated. There are many options here. The scheme used is generally not the best one. A number of very attractive schemes have been introduced recently, see e.g. "Implications of the form of the ensemble transformation in the ensemble square root filters", Sakov P, Oke PR, MONTHLY WEATHER REVIEW, 2008. At least a discussion should be added to convince the reader that the chosen algorithm is a good option for this application.

Regarding the colored noise introduced in equation (7) the authors should motivate the specific choice of the equation: qt is now a stationary process with variance sigma/(1-alphaˆ2). Also they should motivate the choice of the decorrelation length. Has a sensitivity study been carried out to come up with the value used in this study?

As one can see in figure 1 the model covers 3 different domains: The largest domain D1 and the nested domains D2 and D3. It is not clear from the text how the boundary conditions have been taken into account in case of the nested models. It is well known that there might be difficulties in the treatment of the boundary conditions in these kind

of problems.

It is important to validate the results presented, e.g. by considering validation data that is not used in the assimilation experiments. Validation issues are hardly discussed in the paper.

Minor remarks

page 7813, line 18, there is a reference to "30 ppbv" without defining what the abbreviation "ppbv" means.

line 25, where the authors define data assimilation: "... into three-dimensional model...". DA is not restricted to three-dimensional models.

page 7314, equation (2): It seems to me that the noise gamma should also depend on i.

page 7815. Concepts as 4D variational methods and the adjoint are presented here without introduction and explanation, and without references.

line 17, The following sentence needs rewriting, because it is not clear what the authors would like to say here: "The system is employed for constraining input uncertainties of ozone modeling including ozone initial conditions, precursors (NOx and VOC) initial conditions and emissions".

page 7816. After equation (2) the choice of 50 ensemble is made. References to others are added, but how similar are the referred applications with the application studied in the submitted paper? Why would 50 also be good for the application presented here?

page 7817, line 11. The operator M cannot be the same as the operator M in equation (1) since the dimension of V is larger then the dimension of x. M should be redefined here.

line 16: It seems to me that tau1, tau2, ... should depend on t+1 too.

page 7818, line 19. The authors discuss the " ... underestimation of analysis error covariance". A reference to this result should be added or an additional explanation is needed.

line 23. The authors discuss the localization used. They are referring to an "optimal" one. The localization option is optimal with respect with what? What optimality criterion is used here?

Interactive comment on Atmos. Chem. Phys. Discuss., 11, 7811, 2011.