Atmos. Chem. Phys. Discuss., 14, C12563–C12570, 2015
www.atmos-chem-phys-discuss.net/14/C12563/2015/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Data assimilation in atmospheric chemistry models: current status and future prospects for coupled chemistry meteorology models" *by* M. Bocquet et al.

M. Bocquet et al.

seigneur@cerea.enpc.fr

Received and published: 12 March 2015

Summary This manuscript presents a review of data assimilation in atmospheric chemistry models and contains a wealth of information. I appreciate that the authors addressed some of my comments from my "short review" before this manuscript was published in ACPD. Nonetheless, my overall opinion is nearly unchangedâĂŤI still think the manuscript is too long and unfocused and that the writing and presentation are the main shortcomings of this manuscript. However, I have little concern regarding the scientific content, as I believe the authors appropriately encapsulated most of the work to date on data assimilation in atmospheric chemistry models. I have identified

several places where I think the authors can shorten their paper. However, ultimately, I will defer to the authors' choices. If the authors do not wish to make any substantial omissions, that is fine, but I expect that many readers will be turned-off from this article because of its size and often unfocused writing.

Reply: Since this is a review paper, we feel that it is appropriate to provide fairly comprehensive descriptions of methods, data sets, past applications, and selected case studies. Nevertheless, we eliminated some material where we felt that it was appropriate to do so and we also followed some recommendations concerning the organization of Section 3.

Bigger comments and suggestions 1. I feel you should strongly consider removing section 5 and all the figures because they add little to the paper. Section 5.2 is essentially just Pagowski and Grell (2012) restated, and section 5.3 is already-published work from P. Saide. I found section 5.4 to be the most interesting of the case studies, but even that can be safely removed, in my opinion. While it's nice to have figures in an article, I feel that in this case, they don't contribute to further understanding of the topics already described in the text. I feel that section 2.4 can be omitted. A few lines about nonlinearity and non-Gaussianity can easily be slipped into other earlier material in section 2. Is section 2.5 really necessary? The point of this paper is data assimilation, not verification approaches. If you're going to keep section 2.5, then, within it, I suggest removing the "leave-one-out-approach" because, as you mention, this approach is very expensive, and quite frankly, I believe a bit silly and unpractical. Can section 3.3 be omitted? I felt it added little to the text. The first paragraph of section 4.2 can be safely omitted. Further, I feel that the text in section 4.2 beginning "Most retrieval products" through the end of the section can be removed. I feel that section 4.3 can be safely omitted tooâĂŤof course observations are used in chemical data assimilation. Most of this content has been said somehow earlier.

Reply: We feel that it is important to show some examples of data assimilation in atmospheric chemistry models, as those illustrate some of the associated advantages and

ACPD
14, C12563–C12570, 2015

limitations. We debated whether the case studies could be incorporated into Section 3. However, we decided to keep them as a separate section because they not only provide illustrations of the data assimilation methods, but also exemplify the use of observational data sets (ground-level and satellite data), which have been described in Section 4.

We agree that section 2.4 is rather short. Nevertheless, we believe it deserves a subsection on its own because this issue is likely to become a major mathematical and technical hardship of CCMM, when coupling heterogeneous variables, some of them physically bounded. These assumptions often contradict mathematical axioms of standard data assimilation methods such as Gaussianity of the errors. Coupled climate models (with sea ice for instance) and coupled ocean-biogeochemical models also face the same class of issues and addressing this non-Gaussianity issue is already considered a major challenge.

We agree that the leave-one-out approach is not numerically feasible and we have modified Section 2.5 accordingly.

Section 3.3 is useful as a link between the data assimilation methods, which are described in Section 3 and the observational data sets, which are described in Section 4.

The first paragraph of Section 4.2 introduced the major agencies operating satellites. This paragraph has been removed. Acronyms have been defined in other parts of the texts where needed.

The end of Section 4.2 starting with “Most retrieval products. . .” is useful as a reminder of the necessary components of the retrieval products. In particular, DOAS is a popular retrieval approach, but providing kernels with the DOAS approach has become common practice only very recently.

Section 4.3 is important as it exemplifies the methods to use observations for data as-

ACPD
14, C12563–C12570, 2015

similation in an optimal manner. Therefore, it is complementary, rather than redundant, of the earlier section and it provides a bridge with the case studies section.

2. Section 3.1 should be broken into subsections to make it easier to read. Perhaps one subsection could contain studies looking at inverse modeling and another those that examined modifying initial conditions. Similarly, section 3.2 should also be broken into subsections. I'd suggest one subsection for gaseous chemistry data assimilation and another for aerosol data assimilation.

Reply; We have reorganized Section 3.1 along the suggested lines. However, it was not possible to break it down into only two sub-sections and it has been organized into four sub-sections.

It was not possible to break down Section 3.2 into sub-sections along the same lines as Section 3.1 since inverse modeling has not been performed with CCMM yet. To break it down into assimilation of gaseous and aerosol data was not feasible either, because some applications have assimilated both gaseous and aerosol data. Furthermore, it appears that data assimilation into CCMM tends to differ at the moment by their data assimilation techniques (4D-Var, 3D-Var, Kalman filter) as mentioned in the introductory paragraph. Therefore, we kept the current organization. Since Section 3.2 is shorter than Section 3.1, it seems appropriate not to break it down into sub-sections.

3. In general, I strongly urge you to remove all unnecessary text, primarily in section 3. The details of the various studies do not have to be mentioned here. For example, in the paragraph about Schutgens et al. (2010), beginning on page 32253, the sentences starting with "To obtain" and "In addition" can probably be safely removed without detracting from the main point of this study. If readers want more information, they can consult the reference.

Reply: We feel that some summary description of the cited studies is needed in order to provide sufficient information regarding those applications. Therefore, only minimal text removal was performed.

ACPD

14, C12563–C12570, 2015

Smaller comments and suggestions 1. P 32236, L 24: Clarify how this paper differs from Zhang et al. (2012b)

Reply: We added the following text: "..., however, only data assimilation in CTM was addressed".

2. I feel the paragraph beginning on line 17 on page 32237 can be shortened.

Reply: This paragraph was slightly reduced.

3. Suggest rewriting the first sentence of section 2.1

Reply: This sentence was rewritten as follows: "Data assimilation in geosciences has been initially applied to meteorology where methods...".

4. P 32238, L 14: 90's should be "1990s"

Reply: This has been corrected.

5. P 32238, L 18-20: What errors? Please be precise.

Reply: We meant all errors (background, observation, posterior). This has been rewritten as: "...on all errors...".

6. P 32239, L 27: "of" not "in", specify it's the background error covariances

Reply: "in" is correct; "of" is appropriate only when several elements are listed after "consist of...", meaning "composed of...".

The definition of inflation is valid for any type of errors. In practice, inflation could be (and often is) applied to any type of error covariance: background, posterior but also observation.

7. P 32240, L 20: This sentence can probably be omitted.

Reply: We feel that this sentence is a crucial remark backed up by recent numerical experiments: It tells that 4D-Var has an advantage over EnKF. Because of the pop-

ACPD

14, C12563–C12570, 2015

ularity of EnKF, it is often forgotten that 4D-Var should outperform EnKF in strongly nonlinear conditions if it were not for the flow dependence. Therefore, this remark is quite relevant for CTM and perhaps also for CCMM.

8. P 32241, L 5-10: How are the “hybrid ensemble/variational” and “ensemble variational schemes” different? I believe you’re referring to the same thing.

Reply: Hybrid methods consist in coupling two different data assimilation schemes such as an ensemble scheme (EnKF), and a variational scheme (3D-Var and 4D-Var). Because of the use of 3D-Var and 4D-Var, it usually entails using climatological information. Ensemble variational schemes are not always the result of the coupling of two data assimilation schemes, and/or do not necessarily use climatological information (for instance, the iterative ensemble Kalman smoother). There is a very smart account on the issue by Andrew Lorenc (however, it is meteorology-oriented): http://www.wcrp-climate.org/WGNE/BlueBook/2013/individual-articles/01_Lorenc_Andrew_EnVar_nomenclature.pdf. We changed "hybrid ensemble/variational" into "hybrid" to avoid any confusion.

9. In section 2.3, it might be appropriate to mention the NMC method as a way of obtaining background errors.

Reply: Yes, we agree.

"Algorithms relying on consistency check, cross validation and statistical likelihood have been used in meteorology (Hollingsworth and Lönnberg,1986; Desroziers and Ivanov, 2001; Chapnik et al., 2004; Desroziers et al., 2005) to better assess those pivotal statistics." was modified as follows: "Algorithms relying on consistency check, cross validation, statistical likelihood (Hollingsworth and Lönnberg, 1986; Desroziers and Ivanov, 2001; Chapnik et al., 2004; Desroziers et al., 2005) or the empirical but efficient National Meteorological Center (NMC) technique (Parrish and Derber, 1992) have been used in meteorology to better assess those pivotal statistics."

ACPD
14, C12563–C12570, 2015

10. P 32250: Suggest omitting the paragraph beginning in line 14.

Reply: The first sentence has been deleted.

11. P 32252, L 12: “led” not “lead”

Reply: This has been corrected.

12. P 32255: Please rewrite the sentence beginning in line 11. I suggest omitting lines 13-17.

Reply: This sentence has been rewritten as follows: “The authors showed that data assimilation of a combination of different observations (including multiple species) is a very effective way to remove systematic model errors.”

We preferred to keep the end of that paragraph. Although it sounds intuitive, it is nevertheless relevant to future prospects of data assimilation in CCMM as data from different sources are more and more likely to be used.

13. I suggest omitting the text beginning in line 18 on page 32255 through the end of the section. Seems out of place to me.

Reply: This paragraph and the following one have been deleted, along with the associated figures.

14. I believe lines 4-15 on page 32265 could be removed, since IMPROVE and STN network observations are not suitable for data assimilation purposes.

Reply: Such data, which are not available in near real-time, are not suitable for air quality forecasting; however, they can be used for re-analyses of air pollutant concentrations.

15. Suggest omitting the paragraph beginning “MPLNET is a global lidar” on page 32266.

Reply: Assimilation of lidar data has recently been shown to improve air quality fore-

ACPD

14, C12563–C12570, 2015

casts; therefore, it seems appropriate to keep this paragraph on lidar networks unchanged.

16. P 32271, L 18, “past” not “passed”

Reply: This has been corrected.

17. P 32284, L 18: Please rewrite this sentence.

Reply: This sentence has been rewritten as follows: “Assimilating distinct data sets that influence the same model variable could lead to some contradictory information concerning that model variable when the error statistics are misspecified (e.g., unknown bias in semi-volatile PM components); therefore, it will be essential to properly specify those measurement error statistics.” 18. P 32287, Lines 1-9: This material was just said nearly verbatim in section 6. Please consider removing.

Reply: It is not uncommon for the main conclusions of an article to appear in the main text, the conclusion, and the abstract. Some journals accordingly do not accept conclusion sections. However, since Atmos. Chem. Phys. articles typically include a conclusion section, we prefer to keep this part of the conclusion unchanged.

Interactive comment on Atmos. Chem. Phys. Discuss., 14, 32233, 2014.