

Interactive comment on “Assimilation of ground versus lidar observations for PM₁₀ forecasting” by Y. Wang et al.

Anonymous Referee #1

Received and published: 9 October 2012

General Comments:

This paper tests the usefulness of assimilating Lidar observations of PM₁₀ concentrations in order to improve air quality forecasts, relative to the ground-based stations from the AirBase network. This is an interesting question and has been thoroughly examined in this paper.

However, the paper is at times quite difficult to read and the main result of the paper – whether or not the Lidar stations are truly a meaningful improvement – gets lost in the details. Several statements are made redundantly, and details that aren't really that relevant to the analysis (e.g. a long-winded description of the choice of background errors) could be greatly shortened. At the moment it's rather easy to get lost in the details of this paper – I would like to see the authors make the experimental details

C7914

simpler and more clear, so that the main point of the paper (i.e. that the assimilation of the LIDAR observations is more effective than that of the AirBase stations) is clear.

Very much space is dedicated to details of the assimilation experiments, while the space dedicated to actually discussing the results (section 7) is quite small. My overall recommendation is to simplify the discussion of experimental details (mostly by removing redundant statements, but also by carefully deciding which details are relevant to understanding the study and which are not), and devoting more time to explaining and contextualizing the results. By "contextualizing" I mean that the relevance of the results should be clear even to people who don't use this particular model or this particular assimilation approach. Presently, section 7 really only tells us that, given the current experiment settings, the Lidar observations happen to give lower forecast errors on day 2 and higher forecast errors on day 1. Is that a good reason to expand and eventually assimilate the lidar network? You tell me!

The writing could also be made more clear throughout by careful editing, and there are occasional grammatical and spelling errors throughout the document. I've made some technical suggestions below, but the manuscript should be carefully edited before put in final form.

Finally, I would like to remind the editor that my expertise is primarily in data assimilation, rather than air quality modeling, so that my review is heavily geared towards the DA-aspects of the paper.

Note: In the below, I've abbreviated the page numbers to the last three digits, for simplicity.

————— Specific comments:

P293,L22: Is "applications of DA to PM₁₀" equivalent to "applications of DA to air quality"? In that case, the beginning of this sentence is redundant. If there are other ways that DA has been done in the field of air quality, but with different aerosols, that

C7915

could be mentioned (or at least this sentence made more clear.)

P294,L2: "was needed for the DA system" is pretty vague. It would make more sense to say something like, "was needed to yield useful air quality forecasts", or whatever the case may be.

P295, L26: DA by definition always combines models and observations, not just in an OSSE.

P296, Para1: The focus of this paragraph should be to justify why OI was used in this study over more complex algorithms like the EnKF and 4D-Var. However, a lot of detail are given here that make it easy for the reader to get lost. I think this sentence could be greatly simplified by mentioning that it has been shown that (a) 4D-Var apparently has certain weaknesses that make it suboptimal for PM10 assimilation, and (b) that the EnKF has both been shown to perform better and worse than OI/SI. If the statistical interpolation method used in the Denby et al (2008) reference is effectively similar to the OI method used here (I assume that they are, since the two terms are often used to describe similar algorithms in the literature), it might be easier to just refer to both algorithms as "OI".

P296, L21-22: I would strongly suggest showing the OI analysis equations rather than referring the reader to the Tombette et al (2009) study – it seems to me that these are crucial to understanding what is done. Without seeing the basic DA update equation, a reader has only a weak handle on what the covariance matrix is there for.

Section 4.1: This section introduced four somewhat-complicated measures of the performance of the assimilation. – Are any of these redundant, do they all contain unique information? This subsection should make clear why we need all four measures, i.e., what unique information each measure gives us. (If there is a lot of redundancy, you could very much simplify this paper by getting rid of one or two measures).

P299, L12-14: At the end of this sentence, it would be good to append, "of the simu-

C7916

lated PM10 concentrations with respect to the truth" – this will give the reader an idea of why these measures are useful. Alternatively, put the sentence where the $\{o_i\}$ and $\{s_i\}$ are defined first in this paragraph.

P300, L4-5: There is no reason for the reader to be convinced here that the criterion specified by Boylan & Russel (2006) is applicable to this particular study. Is the Boylan & Russel study similar? Can you give a concise reason why we should accept these criteria?

P300, L11-13: The sentence beginning with "Even though for an OSSE..." nicely explains why we are interested in the MFB and MFE measures. This sentence should go to the beginning of this paragraph! Also, it looks like the MFB and MFE measures are used only to meet the model performance criteria, while the RMSE and correlation are going to be used to evaluate the assimilation (analysis) relative to the truth. This could be made more clear.

P300, L16: It is confusing to distinguish between "the truth" and "the true states (e.g. concentrations)" – for the latter, why not just say, "hypothetical PM10 observations"?

P300, L23: It should be explained what is meant by "representativeness errors". Where does the 35% come from? Why are different assumptions made about observation error in AirBase and in LIDAR?

P301: It is interesting that the observations are perturbed with a spatial covariance structure, as opposed to just adding random noise to each measurement. Can you explain why this more complex approach was chosen?

P302, L28: I don't entirely understand "allows us to increase the duration of this impact". If I understood correctly, both the initial values of all (?) chemical constituents and aerosols are perturbed in the assimilation run relative to the "true" run, in order to make the error between them larger. If this was not done, in which case both runs would in some sense forget the initial conditions, what would happen? Would the two

C7917

runs collapse back to the same state? If so, why are we interested in assimilating observations, if we can just recover the truth by letting the model run long enough?

P303, L12-14: Same comment as for P302, L28

P304, L19-25: The end of section 4 states that the specification of the background error covariance / correlation field is critical to the success of the assimilation. Then the beginning of section 5 states that "the definition of background correlations are relatively trivial". Which is it? It's also confusing that section 4 discusses using the Balgovind approach but then section the NMC method is mentioned; one has to read both parts over a few times to figure out what used for which part. I would suggest integrating the discussion of background error statistics into section 5. Then make it clear that background error covariances are extremely important, but also explain what makes their specification easier in this (special) case where the model is a perfect representation of the truth. Some redundant statements can also be removed, e.g. "...are crucial for the success of the method" (P304, L7) and "...is crucial to the quality of the analysis" (P304, L17).

Figure 8: What do the vertical black lines denote? Are the RMSE and Correlation here defined between the assimilated analysis and truth? If so (or if otherwise), this should again be mentioned here. Also, the caption points top & bottom figures, but they are actually side by side. It would make the plots much more clear if they were clearly labeled "RMSE" and "Correlation" as headings. As in other figures, the axis labels could be much larger.

General comment on section 5: It isn't entirely clear to me why the NMC method of estimating background error covariances is explained in such detail, only to manually test different decorrelation lengths anyway. Why not just show Figure 8 and then talk about the effects of the different decorrelation lengths? Also, it is very clear from Figure 8 that assimilating the LIDAR observations yields lower errors and a higher correlation to the truth than does the assimilation of AirBase stations only. It seems odd not to

C7918

mention this, either at the end of Section 5 or at the beginning of Section 6.

P307, L1-3: "...because assimilation only influences initial conditions..." – this is only true if the assimilation is only applied at the initial time, but not if observations are assimilated at regular intervals.

P307, last para: Lines 18-19 state, "the AirBase DA leads to lower RMSE than column DA for most forecasts," but then lines 28-29 say, "the column DA leads to lower or similar RMSE as the AirBase DA for most forecasts." I had to read the paragraph several times to understand how this is not completely contradictory. Please clarify the text by (clearly!!) stating what the difference is between what is shown in Fig. 9 and in Fig. 10. Since the experimental set-up is fairly complicated, it really needs to be more clear where we are looking at continuously-assimilated runs and where we are looking at pure forecasts.

P309, L7-8: "The results shown in this paper suggest that the assimilation of lidar observation would improve PM10 forecast over Europe" – this statement should be made much more specific. Section 7 shows that the value of the lidar observations greatly increases on the second forecast day, relative to the first. Can the authors list other respects in which (according to this study) the lidar observations offer an advantage, and the extent of this advantage. Something along the lines of, "Assimilation of the lidar observations improves forecasts by x% relative to the standard AirBase measurements."

————— Technical corrections:

P293, L6-7: "Aerosols influence gaseous molecules photo-dissociation" → "Aerosols influence the photo-dissociation of gaseous molecules".

P294, L27: "analysis" → "analyses"

P295, L4: No need to write "OSSE system" since OSSE already contains the word "Experiment" – Just writing "An OSSE is constituted..." is enough. Same goes for

C7919

similar uses later on. This sentence also doesn't make it clear what a twin run actually is, since really any model run can be considered "an approximate atmosphere". The next sentence explains it, so the first sentence should be slightly restructured.

P296, L5: "the ensemble" → "an ensemble"

P298, L13: "stations types" → "station types"

P298, L23-24: Change the end of the sentence to "in order to better cover Western Europe".

P299, L3-4: This sentence is confusing and can be worded much more simply. I'd suggest something like, "Observation impact experiments for not-yet-existing observing systems require the simulation of an atmospheric state, from which the hypothetical observations can be generated." Then all that is needed in the rest of the paragraph is to explain that we call this state the "truth", and to describe the truth run used in this study.

Section 4.3: Since this run is afterwards only referred to as the "truth", perhaps it would be better to entitle this section "Truth simulation", or something like that.

P302, L8: Technically, the "twin run" wouldn't be called that if this wasn't a twin experiment. It's probably better here to refer to the non-control run as the "assimilation run" – then it's clear to the reader that this is the run where the observations are assimilated.

P302, L15: There is a lot of redundant information in this paragraph. It's better and more straightforward to state right away that this particular study performs identical-twin experiments, so move the sentence starting with "We follow..." to the beginning of the paragraph. The part about fraternal-twin experiments isn't really relevant to this study and should probably be parenthetical. The most important point made in this paragraph is that the impact of the observations is usually overestimated in identical twin experiments – if the other details are trimmed, this will be more clear and then you won't have to repeat it at the beginning of the next paragraph.

C7920

P303, L17-19: It is not clear why there should be a compromise between lower assimilation error and short assimilation time. – What is the benefit of keeping the assimilation time N low? Why would we not want to assimilate as many observations as possible?

Figure 5: I would suggest a colormap that is white in the center (i.e. for small to no differences between truth and assimilation run), to make it easier to spot where the main differences are, but I'll leave this up to the authors. Also, the figure labels could be much larger. Also, it would be good to remind the reader where in the assimilation time 00:00 on 15 July falls (i.e. that it's the initial time).

P303, L21-22: "The simulations use the same setup..." – this seems to have been made abundantly clear in the description of the identical twin experiments.

P306, L25: "of Fig. 8" seems misplaced here, since both tests are shown in Figure 8. (Or the sentence needs to be restructured.)

Interactive comment on Atmos. Chem. Phys. Discuss., 12, 23291, 2012.

C7921