**Atmospheric
Chemistry
and Physics
Discussions**

Interactive
Comment

Interactive comment on "PM$_{10}$ data assimilation over Europe with the optimal interpolation method" *by* M. Tombette et al.

Anonymous Referee #2

Received and published: 21 July 2008

General remarks:

1. I think that the conclusions drawn from the study is based on too a small data basis.

January 2001 is taken for the statistical part of validation. The authors should recognize that biogenic organic aerosols are missing for larger parts of Europe and France, rendering some discussion in the paper concerning aerosol splits questionable. More problematic is the fact that the temporal and horizontal impacts for different DA parameter configurations described in section 6 includes only a single validation forecast starting from January 6th, 2001, a Saturday, with the preceding five days for spin up with hourly data assimilation. This is virtually 1 winter case with a specific meteorological situation and emission conditions for a weekend.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

2. Moreover, the authors do not include any a posteriori validation of the DA procedure in terms of the now state of the art Observation-minus-Forecast (O-F) and Observation-minus-Analysis (O-A) and chi-square validation as for example described in Talagrand (Proceedings of the Workshop on diagnosis of data assimilation systems, ECMWF, 1998), to assure at least a rough consistency between forecast and observation error covariance matrices.

Any statistical interpretation should be performed on the improved basis of consistent error covariances.

In general, it is difficult to draw conclusions from this paper beyond the very specific setup and selected January 2001 case. My recommendation is to extend at least those portions of the paper, seeking for improving DA parameters like decorrelation lengths Lh/v and introduce appropriate means for a posteriori validation.

The paper is not publishable in its present form and should be subjects to major revisions.

Specific remarks

Introduction, line 106,. It is claimed that aerosol CTMs do not reproduce observed highest PM10 peaks due to missing processes of exceptional events (Saharan dust, . . .).

Then, in section 3, line 231, this statement is generalized to usual conditions, with the exception of nitrate winter conditions. Literature references are indispensable for these claims. Please provide.

Section 4, line 304: Table 1 does not contain information, which is suitable for presentation in a table. Does "model" mean background field for assimilation? It is suggested to remove Table 1.

Subsection3.2 To what extend can EMEP and AirBase data be used for validation of BDQA data? Do EMED and AirBase data repositories not include BDQA data? Or is

care taken, that there is no coincidence for validation?

Section 4, line 307: On which information has the 2 grid cell selection of the scale parameter L_h been made, preferred to be used for the full month? Given previous remarks that some stations have been dropped from the assimilation procedure due to their poor representativity, a uniform 2 grid cell scale parameter (decorrelation lenghts) appears to be questionable.

Same paragraph: What does the statement mean: "The error variance for observations is lower than the instrumental uncertainty."? In terms of data assimilation relevance, the observational error variance an the error variance for representativity is required.

Section 4, line 315: As mentioned above, care should be taken to keep error covariances consistent. Formally, there is no freedom to "assume" observations to be "highly accurate" to assess the potential benefits of data assimilation. On the contrary, OI (as variational and Kalman filter approaches) is based on statistical assumptions of error characteristics to be followed as much as possible. In the specific case here, if observation errors are larger than assumed in the DA set-up, later comparisons with observations will likely to be outside the to tightly selected error margins and hence engender worse skill scores. Unexplained impacts like phenomena described around line 350 may be a consequence.

Section 4, line 323: Figure 2 lacks a scale. There appears to be no more information than in Table 2. It is suggested to remove Figure 2.

Section 4, lines around 385: Does this mean that EMEP stations are not appropriate for validation? Is a 0.5 degree mesh size grid not more appropriate for the EMEP site deployment policy?

Section 4 last paragraph, around line 333: Which lessons can be taken from the given description? Are today's aerosol modules oversophisticated for data assimilation, if only "lump" information in terms of PM10 is given?

Section 5, 1st and 2nd paragraph: The description is confusing. First sentence: Is Near Real Time data of PM10 not available to the authors? What does mean "assimilation during the first 3 days, then model forecasts to the next two days". Does it mean hour by hour data assimilation during the first 3 days, and the free forecasts after that? Please clarify.

Tables 3, 4, and 6 should also include a column with biases, in addition to other statistical quantities.

As a rule, figures should include SI units of presented quantities.

Fig 6: For completeness SI units and time should be included in the caption.

Interactive comment on Atmos. Chem. Phys. Discuss., 8, 9607, 2008.