Interactive comment on “Numerical modeling of Asian dust emission and transport with adjoint inversion using LIDAR network observations” by K. Yumimoto et al.

Anonymous Referee #1

Received and published: 28 December 2007

Review of 'Numerical modeling of Asian dust emission and transport with adjoint inversion using LIDAR network observations' by K. Yumimoto et al.

General comments

The authors present an inversion of dust emissions using LIDAR observations of dust extinction coefficients during a heavy dust event over eastern Asia. The inversion is performed using a mesoscale model including dust emissions and transport in combination with a four-dimensional variational data assimilation framework.

The study is a nice example of the use of variational data assimilation for inverse mod-
eling of emissions based on atmospheric observations. For the case study presented, observations from a LIDAR network in Japan are effectively assimilated to modify estimates of central Asian dust emissions, leading to overall improved agreement with other (independent) observations. Despite these convincing results, I have three general comments:

- My main concern is the specification of the cost function (Eq. (5)). Apart from the usual background and observation terms, this contains a so-called smoothing term. It is not clear to me what makes this smoothing term fundamentally different from the background term. My impression is that it would not be needed if the $B$ matrix were modeled more carefully. Presently, $B$ is just a diagonal matrix (thus no spatial and temporal correlations) with very large variances, so that it probably provides only a minor constraint in the optimization. The authors should work on a more realistic representation of background error covariances, instead of introducing an additional ad-hoc regularization term.

- As a rule, in data assimilation those model parameters that are most uncertain should be optimized. Although my feeling is that emissions represent indeed the largest uncertainty in the dust model, it would be useful if the authors could briefly discuss other model uncertainties, in particular transport, and defend why those have been neglected in their assimilation system.

- The paper is syntactically poorly written. Below I have listed many technical comments, but this list is far from complete, and I advice the authors to have the paper checked by a native speaker.

**Specific comments**

P15955, Title: The title is not to the point. Suggest revising to, e.g., 'Inverse modeling of Asian dust emissions using LIDAR observations'. It is important to note that emissions, NOT transport, are inversely modeled.
P15956, L10-11: Remove list of validation data here, since this list is again given later in the abstract.

P15956, L17-18: This is selective reporting: these are the three best stations; at other stations the reduction is less or negligible.

P15957, L14: To which spatial and temporal scales is this 'factor of ten' related?

P15957, L16: What are 'atmospheric models within the models'?

P15957, L21: Provide basic references here (e.g., Talagrand et al. 1987).

P15959, L17: I assume that $t_i$ is the time and $dt$ is the time step.

P15959, L22: Describe not (only) what can be included in the control vector, but (also) what is included in the control vector in this study.

P15959, L22: What does 'etc.' include?

P15960, L8-9: This description of the adjoint makes no sense. For example: where are the emissions? How does the term $dJ/dH_i C$ enter in this equation? If the authors want to spell out the adjoint model, they should do it carefully, and add a separate equation for the adjoint emissions. The adjoint forcing $dJ/dH_i C$ only shows up in the equation of the gradient of the cost function.

P15960, general: Is preconditioning applied to the control vector?

P15960, L24-25: How many iterations are typically needed to reach this reduction? Has it been verified that the solution does not significantly change when more iterations are added?

P15961, Eq. (4): According to this equation, there is a negative dust flux if the friction velocity is smaller than the threshold. Is this true?

P15961, L4-5: If $C$ is a constant, how can it depend on the grid cell? What is the dimension of $C$? How does it typically depend on different types of surface information?
P15961, Eq. (5): What is $\gamma$? What is $\Delta$?

P15961, L14: Why are negatives replaced with zeroes? In this way, the solution is no longer optimal.

P15961, L17: There is no 'assimilation of dust transport' performed, but 'assimilation of dust observations'.

P15962, L3-11: Are the LIDAR observations sensitive to other aerosols than dust? If so, has this been corrected for?

P15962, L9: What is $S_1$?

P15962, L27-28: Why is the inversion performed over a period of two weeks, while only observations in the last week are assimilated? Regarding the transport of dust, a 'spin-up' time of 3 to 4 days appears to be sufficient.

P15963, L14: Which 'surface boundary data' are meant here?

P15963, L15: On which spatial and temporal scales are these difference present?

P15963, L16-19: The authors should improve the specification of the background error covariance matrix. (see general comments)

P15964, L12: Is this the same low-pressure system?

P15964, L25-26: How do we know that the dust has been captured between those potential temperature levels?

P15964, L22-23: On the next page a blacked-out area is attributed to the presence of a too dense dust layer below, through which the LIDAR signal could not penetrate. Apparently, black does not only refer to rain or clouds.

P15965, L14-16: The upper dense dust appears not to be reproduced very well, but only to a limited extent.

P15965, L16-19: I guess the authors mean 'undefined' instead of 'not undefined'? It
may be clearer to rephrase this sentence: e.g., 'The LIDAR could not measure the dust above 2000 m between 31 March and early 1 April at Seoul because the signal was not able to penetrate through the very dense dust below that altitude. This indicates'.

P15965, L21-24: There are significant increases around 26 April in the posterior vs. the prior simulation, both at Seoul and at Matsue. How can this be explained, given that only observations after 28 April are assimilated?

P15966, L24-25: I am guessing that for the calculation of observed AOT, the 'black regions’ are ignored (effectively set to 0). How is the modeled AOT calculated for cases in which only part of the column could be observed by the LIDAR?

P15966, L25 - P15967, L5: Most of this description is also given in the figure caption, which is the right place. Duplication should be avoided. Thus, most of this paragraph must be removed.

P15968, L1-2: Table 1 lists the RMS of the difference between observed and modeled AOT. (The term error refers to a difference between an observed/modeled value and the truth. This is not what is meant here.)

P15968, L7-9: Is the period 25 to 29 March taken into account for the calculation of RMS and mean values in Table 1? Then the result is not surprising, because only observations after 29 March have been assimilated. Suggest to calculate statistics for the period in which observations have been assimilated.

P15969, L13: It is surprising that the assimilation degrades the RMS for all stations (and, since the mean difference between observations and model generally decreases, the standard deviation is degraded even more). Can you comment a bit further on this result? From the time series in Figure 5 it seems that the assimilation generally leads to improved agreement. Why is this not reflected in the RMS?

P15970, L16-21: Again, this is a redundant duplication of the figure caption. Should be removed.
P15973, L12: $u_{*,th}$ is constant.

P15973, L24-25: I don’t understand this sentence.

P15974, L24 - P15975, L2: This is a selective (too positive) summary of the results.

P15975, L6-8: Again, too positive summary: RMS became worse after assimilation.

P15976, L9-13: Suggest to mention that probably biases between different types of observations will be present, and that proper correction for these mutual biases will be a prerequisite for joint assimilation of different observation types.

P15980, caption Table 1: 'error' should be replaced by 'difference between observed and modeled' as mentioned in comment P15968, L1-2.

P15980, caption Table 1: What is the time window over which the statistics have been calculated?

P15980, Table 1: According to the note, the table contains RMS errors and mean of AOT, which is unitless. However, according to the table 'dust ext., 1/km' are reported.

P15980, Table 1: For the two stations added in experiment B, the mean modeled AOT differs more from the observed mean in experiment B than in experiment A. How can this result be explained?

P15981, Fig. 1: Dust emissions are not 'assimilated' but 'analysed' or 'optimized' or 'inverted'. (Observations are assimilated.)

P15982, Fig. 2: In the right column plots, only SYNOP sites reporting dust are indicated. It would be useful to add also the SYNOP sites reporting NO dust.

P15983, Fig. 3: Explain what black is in upper row (is in text but should be in caption). Explain what white is in lower three rows.

P15984, Fig. 4: Explain more clearly the difference between dashed and solid lines. After some thinking, I assume that they represent the full and partial column integral,
respectively, 'partial' meaning that the integral is only taken over the height range where valid LIDAR observations were present.

P15987, caption Fig. 7: What is 'surface speed'?

**Technical comments**

P15956, L9: observation → observational.

P15956, L26: the heavy dust uplift flux → strong dust uplift fluxes.

P15956, L27: Remove 'the'.

P15958, L24-29: Sentence is too long and hard to read.

P15958, L28: Explain what 'AOT' means.

P15958, L29: Start new paragraph at 'This paper is structured '.

Section 2: This section jumps back and forth between description of model and 4D-Var. Separate these, first describing the model (including parameterization of emissions) and then the 4D-Var.

P15959, L11: advections and diffusions → advection and diffusion.

P15959, Eq. (1) and further: Matrices and vectors should be bold; scalars normal font.

P15959, L14: $M_{\text{diffs}}$ → $M_{\text{diff}}$.

P15959, L21: parameter → vector (also at further locations in the text).

P15959, L23: Also in that equation → Further, in Eq. (2).

P15960, L12: '; it' → '.

P15960, L21: Remove 'these'.

P15960, L25: reduced by 1/1000 → reduced by a factor 1000.
P15961, L2: What is $F_k$? What is $f_k$?
P15961, L4: second 'the' → 'a'.
P15961, L5: Remove 'of'.
P15961, L11: the → a.
P15961, L12: Remove second 'the'.
P15962, L28: Add 'the' between 'from' and 'surface'.
P15962, L29: Remove 'e.g.' (all observations used for validation are listed).
P15963, L20: Remove 'as'.
P15963, Eq. (6): $B \rightarrow R$ (or better use indices, $R_{ii}$ and $y_i$, since otherwise left-hand side is a matrix, while right-hand side is a vector). Also, use of letter $E$ for errors is confusing, since $E$ without subscript was used for emissions in Eq. (1).
P15964, L26: '$\theta = 285-295$ K' → 'between potential temperature ($\theta$) levels of 285 K and 295 K.'
P15965, L2: '<' → '>'.
P15965, L11: Remove 'For that reason'.
P15965, L12: might not → does not have to.
P15967, L28: Remove 'a'.
P15968, L10: lesser → smaller.
P15968, L18: are → were.
P15968, L20: for → from.
P15969, L15: presents → yields.
P15969, L16: lesser → smaller.
P15969, L22: Add 'a' between 'make' and 'comparison'.
P15969, L27: average → averaging.
P15970, L1: Add 'a' between 'makes' and 'classification'.
P15971, L5: falls → fall.
P15971, L6: Add 'of' between 'type' and 'aerosol'.
P15971, L7: Remove 'small'.
P15971, L9: Remove 'the' before 'misclassifications'.
P15972, L2: 'modeled AOT and observations shows' → 'modeled and observed AOT show'.
P15972, L5: 'on' → 'in the'.
P15972, L14-15: 'gives a smaller AOT level' → 'is lower'.
P15972, L17-18: I don’t understand this sentence.
P15972, L20-21: Figure → Figures.
P15973, L20: grids → grid cells.
P15973, L24: Put 'Eq. (3)' between brackets.
P15973, L26: grid → grid cell.
P15973, L26-27: What are 'observations that can enable measurement'?
P15974, L20: 'through the assimilations' → 'after assimilation'.
P15985, L11: see comment P15964, L26.
P15975, L16: Add 'The' before 'CALIPSO'.

P15976, L1: to → in.

P15976, L4: 'could derive the' → 'yielded'.

P15976, L8: Remove 'the'.

P15980, caption Table 1: Title of the table should start with 'RMS ', not with 'Observation sites'.

P15980, Table 1: Remove the last column of the table containing the units of PM$_{10}$, as these are already given in the first column.

P15980, Table 1: Suggest to separate the NIES LIDAR AOT and JADS PM observations by a horizontal line for clarity.

P15981, Fig. 1: What is the blue square? What are the black circles?

P15981, caption Fig. 1: 'Dust emission increment for w/o assimilation and assimilated' → 'Dust emission analysis increment'.

P15982, Fig. 2: The OMI-AI contours are very hard to read. Suggest to add a separate contour plot for OMI-AI.

P15982, caption Fig. 2: 'assimilated AOT' → 'modeled dust AOT from assimilation experiment A (or B?)'.

P15986, Fig. 6: The plots in the lowermost row are very small and yet very busy. This makes them very difficult to read. Suggest to increase size and put them in a separate figure.

References