Atmos. Chem. Phys. Discuss., 8, S4011–S4021, 2008
www.atmos-chem-phys-discuss.net/8/S4011/2008/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "A method for evaluating spatially-resolved NO_x emissions using Kalman filter inversion, direct sensitivities, and space-based NO_2 observations" *by* S. L. Napelenok et al.

S. L. Napelenok et al.

Received and published: 23 June 2008

** The difficulty in data assimilation is often in attributing realistic errors to the model and the observations. The authors discuss the effects of a range of combined measurement and a priori emission errors (Fig. 5) for the pseudodata case, but avoid stating what exact numbers they used in their case study and why. I think these numbers and some justification thereof should be given.*

We appreciate the reviewer noticing that we had left this information out of the manuscript. Our estimate of uncertainty in the emissions for the application to the southeastern United States case study was 2.0 and the uncertainty in the observations was 0.3. A discussion on the selection of these two parameters in the inverse now appears in section 4 as follows:

> For each case, the normalized uncertainty in the a priori emissions, U_E, was set 2.0, which allows for large adjustments and follows the estimates of Hanna et al. (2001) for possible errors in NO_x emissions. The uncertainty in observation, U_{obs}, was set to 0.3 according to the estimates of Martin et al. (2006) of mean monthly uncertainty for SCIAMACHY observations of polluted regions.

** The pseudodata analysis suggests that border regions and boundary conditions have minimal influences on the inversely modelled NO_x emissions within the region of interest. This is an important result, and I'm wondering why the authors do not include it in the abstract. Furthermore the lack of discussion of this finding is puzzling. The result seems to be specific for summertime southeastern US (short chemical lifetime, stagnant weather), and it needs to be discussed in that context.*

The modeling domain was designed such that the border region and the boundary conditions did not significantly influence the concentrations in the defined regions. DDM-3D has the flexibility to quantify the influence from the boundaries directly and was used for this purpose in the pseudodata analysis to verify that the influence was minimal. If a significant amount of NO_x in the domain was originating outside the source regions, the inverse results would be less trustworthy. The reasons for the minimal influence include mainly the factors pointed out by the reviewer - chemical and meteorological conditions during the episode. While we do not believe that this discussion is appropriate for the abstract, we have included the following text in section 3 in order to explain this result:

> Overall, the border region provided reasonable separation to neglect any impacts from the boundaries, mainly due to stagnant meteorological condi-

ACPD

8, S4011–S4021, 2008

tions common in southeastern summers and the consequently short chemical lifetime of NO_x relative to transport processes.

** The authors do not discuss the impact of the small number of SCIAMACHY observations (3-10 over the whole period) on their results whereas they could easily have done so. For instance, the pseudodata analysis was done for 1 August 2004 with base-case model simulations as pseudo observations. Such a test is useful, but not at all representative for the inverse method with real SCIAMACHY data. A test with pseudo observations sampled as SCIAMACHY observations would specify to what extent the observations contribute to finding the solutions in your real data. Similarly, the diagonal elements are set to 0.5 10^{15} molec. cm^{-2} (P6478, L26-27) to test the pseudodata analysis, and I'm wondering why the impact of more realistic observational errors has not been tested here.*

The number of available SCIAMACHY observations was indeed small (5 per each model grid cell on average). However, after these were mapped onto model domain space, the observations were averaged for the entire summer to achieve complete coverage in the domain. We agree that it is not clear in the manuscript that the inverse was performed using averaged modeling results and observations over the summer. The following text was added to section 4:

> To assure full coverage of the modeling domain by the infrequent and spatially varying observations, modeling results and SCIAMACHY observations were averaged over the three summer months in the inverse.

** Why have the authors chosen such large (200x300 km^2) source regions? This is hardly taking advantage of the high resolution SCIAMACHY data.*

As is mentioned in the text, inverse modeling frequently suffers from mathematically ill-posed scenarios. Initially, in an overly ambitious attempt, we assigned each modeling grid-cell as its own source region. However, it became quickly apparent that our

ACPD

8, S4011–S4021, 2008

Interactive Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

Kalman Filter implementation could not distinguish the difference in the signature of many of the source regions and thus could not converge on a unique solution to the system. The source regions presented in the manuscript were chosen for based on spatial and temporal emissions patterns in order to limit to assure convergence in the inverse. The high resolution of the SCIAMACHY allows for more accurate quantification of the impact from each source region on the domain, but is not always enough to assign adjustments to the emissions from smaller source regions in rural areas where concentrations were observed to be fairly homogeneous.

Then I have concerns about a number of methodological issues: The authors assume that atmospheric NO_2 *concentrations at 10:00 reflect the* NO_x *emitted in the previous 16 hours only (P6478, L4-7). This suggests that* NO_x *emitted in the afternoon rushhour does not contribute in any way to* NO_2 *observed the following morning at 10:00 am. But some afternoon rush-hour* NO_2 *will live through the night and may contribute to* NO_2 *observed the next day at 10:00 am, especially downwind of strong sources. I understand that summertime* NO_2 *has a short (daytime) chemical lifetime of 2-4 hours, but at night the chemical lifetime is longer. I think the authors should justify the implicit assumption that 10:00 am* NO_2 *concentrations are unrelated to afternoon emissions. Related to this issue, I think it is important to include some discussion on the timing of the* NO_x *emissions in CMAQ.*

We agree with the reviewer that some of the previous day's NO_x emissions will contribute to the concentrations during the satellite overpass the next morning (16:00 UTC). However, the amount of the "evening"; NO_x relatively small compared to "morning"; NO_x particularly in the summer. We have analyzed this aspect in some detail while developing the method, since DDM-3D is able to track contributions from specific time intervals. We found that emissions had a significant impact in the emitting cell for less than five hours, while contributions to a downwind cells (over 50km) peaked at a five hour delay during this episode. Furthermore, the transported emissions did have a longer impact on downwind cells, but it was substantially lower than local emissions.

To clarify this point in the manuscript, the following text was added to this section:

> During the summer in the southeast, NO_x has a relatively short lifetime, therefore, this time interval captures the majority of emissions that would impact concentrations during the time of the satellite measurement.

Furthermore I suggest the authors discuss the representativity of the surface NO_2 measurements in more detail. How representative were the SEARCH sites for the average concentrations simulated for the 36 x 36 km^2 CMAQ grid cells?

Expert judgment was used to select the location of the SEARCH monitors in order to contrast the difference between urban and regional pollution. It is impossible to know how representative a given site is without many more measurement locations. However, the intention and design of the SEARCH network was to provide observations that are representative of either urban or regional conditions.

We've updated the text on line 353 to reflect this comment:

> The SEARCH network is designed to provide observations that are representative of either urban or regional conditions.

Why were the surface measurements averaged over the daytime concentrations rather than sampled at 16:00 UTC as was done in the CMAQ-SCIA analysis?

While it is important to accurately simulate the column density of NO_2, we are most interested in using this technique for improving our simulation of the atmosphere near the surface. Even though the adjustment is made due to the difference at a particular time, the surface NO_2 concentration throughout the day is important for air quality. Despite only providing a column NO_2 density observed at a specific time, does this inverse-technique improve the day-time surface NO_2 concentration?

ACPD

8, S4011–S4021, 2008

Interactive Comment

The nighttime NO_2 concentration is most impacted by chemistry and vertical mixing, and less influenced by emissions. We exclude the nighttime concentrations from our comparison because any improvement in the nighttime concentration would likely be due to coincidence.

We've updated the text on lines 339-346 to reflect that the comparison with the SEARCH network is both an independent test and used to determine if the results of the NO_2 column density derived emission estimates are general enough to also improve the simulation of surface NO_2 concentrations:

> While it is important to accurately simulate the NO_2 column density, we are most interested in using this technique for improving our simulation of surface air quality. However, the discrepancy in model and observed concentrations can be due to processes other than errors in emissions. Despite only providing a column NO_2 density observed at a specific time, can the SCIAMACHY data and this inverse-technique improve the day-time surface NO_2 concentration? This comparison is useful as an independent check on the results and to determine the extent to which the results can be generalized to phenomenon relevant to air quality.

In the Abstract, reasons for the underestimation of NO_2 columns are mentioned. The authors convincingly point out that the lightning NO_2 is likely too low in the CMAQ model, but also mention "a short modelled lifetime of NO_x aloft" as a likely reason but without substantiating this in the paper.

A possible explanation of the under prediction of NO_2 is an under prediction of the NO_x lifetime in the free troposphere. However, we do not have substantive evidence of this at this time, so have removed this sentence from the abstract.

Minor comments P6471, Ll28: please spell Müller not Muller.

Done.

ACPD

8, S4011–S4021, 2008

Interactive Comment

P6472, L1: please spell Quélo, not Quelo. Although Quélo's is an interesting paper, it doesn't use SCIAMACHY or any other space-based observations as the sentence now suggests. Please update.

The spelling was changed. In this part of the introduction, we wanted to acknowledge recent work in regional scale inverse modeling that is not limited to any specific source of observations. The sentence was updated as follows to clarify:

> These data, as well as ground-based and other observations, have been used previously in inverse modeling of "top-down"; inventories, but typically on the global scale (Martin et al., 2003; Müller and Stavrakou, 2005), and less frequently on the regional scale (Blond et al., 2007; Kim et al., 2006; Quélo et al., 2005; Konovalov et al., 2006; Konovalov et al., 2008; Wang et al., 2007).

P6472, L1: it would be appropriate to also cite the work by Blond et al. (JGR, 2007), Y. Wang et al. (GRL, 2007) and Konovalov et al. (ACPD, 2008) here. These papers use high-resolution CTM simulations, SCIAMACHY/OMI and surface NO_2 to better understand air pollution on the regional/urban scale.

Done.

P6472, L6: please provide a citation to DDM-3D here.

Done.

P6475, L10: I don't get this. At eq. (2), it is stated that N includes not only observation errors but also model uncertainties. This sentence suggests we're only dealing with observation errors here. Please clarify.

The text was not clear on this point and was modified to include the passage below. Our model evaluation revealed large underestimations in NO_2 aloft due to possibly missing

ACPD

8, S4011–S4021, 2008

Interactive Comment

sources and possibly incorrect chemistry. The goal of our implementation the Kalman filter inverse is to correct systematic biases in the emissions. We decided to neglect random model errors from the noise matrix, because these are difficult to estimate and appear to be less important than the biases we are aiming at correcting. At the same time, we realize that correcting ground-level emissions should not be done to compensate for missing processes aloft. To address this issue, we added a simple correction factor to account for missing NO_2 concentrations aloft as discussed in section 4.

> In this application, the noise matrix did not include an estimate of model uncertainties. An evaluation of model results revealed a clear systematic bias in NO_2 column predictions overwhelming any Gaussian type errors that would be included in the noise matrix. Possible sources of this bias and the approach of addressing it are presented later in section 4.

P6475, Eq. (6): I think the right-hand side of the expression should be squared since it is a (co)variance matrix. Furthermore, there is now inconsistency with Eqs. (4), (5) and P6478, L27.

The inconsistency was addressed as the reviewer suggests.

P6475: I feel the paper would be strengthened if the authors give a range of numbers for U_E and U_{obs}, and some justification for these estimates.

Ranges for uncertainties estimates are discussed in some detail in section 3 of the manuscript and are illustrated in Figure 5. We realize that we did not provide sufficient justification to our choice of specific values. To address this and a previous reviewer comment, we added the following text in section 4:

> For each case, the normalized uncertainty in the a priori emissions, U_E, was set 2.0, which allows for large adjustments and follows the estimates

ACPD

8, S4011–S4021, 2008

Interactive Comment

of Hanna et al. (2001) for possible errors in NO_x emissions. The uncertainty in observation, U_{obs}, was set to 0.3 according to the estimates of Martin et al. (2006) of mean monthly uncertainty for SCIAMACHY observations.

P6477, L14-16 ("The inverse was ... data as X^{mod}"): I don't understand what this sentence should tell us. Could you please clarify?

X^{mod} and X^{obs} are both vectors that contain the number of components equal to the number of paired model and observation points. We are trying to convey the fact that all grid cells contained by the source regions were used in the inverse. This sentence was poorly worded and was rewritten as follows:

> Each model grid cell contained by a defined source region was paired with the spatially matched averaged satellite observation to develop X^{mod} and X^{obs} vectors in the inverse.

P6478, L20: I think the reference should be to Eq. (6) here, not (5).

Corrected.

P6479, L14: I think the authors also want to refer to Eq. (6) here.

Corrected.

P6479, L16: there is no Eq. 5a to refer to.

Corrected to Eq. 6.

P6479, L24-25: I think it should be stressed here that Fig. 5 relates to the Atlanta case only.

The text was changed as follows to make this point:

> As expected, larger uncertainties in emissions and lower uncertainties in

ACPD
8, S4011–S4021, 2008

observations allow for larger adjustments to the emission fields in the case of the Atlanta source region (Fig. 5) and elsewhere.

P6481, L5-11: this raises the question if and how lightning NO_x is simulated in CMAQ. Please describe the (lack of) of LNO_x simulation. P6481, L13: Konovalov et al. have not found a systematic bias between satellite observation their model simulation over Europe. They just report that NO_x emissions from lightning are not included in CHIMERE, and state that this likely leads to an underestimation of the NO_2 column of at most 0.08 10^{15} (not 0.8 10^{15} as the authors suggest here). Their deficiency is thus more than an order of magnitude smaller than the 1.07 10^{15} reported here. Please clarify.

While our group and others have made some attempts to address this problem, lightning produced NO_x is currently not simulated in CMAQ. Anonymous Referee 1 also raised the issue of attempting to spatially resolve this source. Once again, we are actively pursing the improvement of emissions and their chemical fate aloft, but decided to focus this manuscript on the inverse methodology.

Konovolov et al. (2006) suggest that the lightning produced average NO_2 column over Europe during their simulation period is likely less than 8 x 10^{13} (molecules/cm^2) (Pg. 1749). However, they go on to say that the difference between measured and simulated NO_2 column values show a systematic difference of "about 8 x 10^{14} cm^{-2}"; (Pg. 1752). They do not offer an explanation besides suggesting possible sources that include the omission of the upper troposphere in CHIMERE, and other errors in the model, emissions, or satellite data.

P6482, L15-16: I suggest the authors provide some more information on the surface NO_2 measurement technique, the Hansen et al. paper was not readily available to this reviewer.

The following text was added to the manuscript in section 4 to describe the measure-

ACPD

8, S4011–S4021, 2008

Interactive Comment

ment technique and provide a reference:

> These sites measure NO_2 by photolytic conversion to NO followed by chemiluminescence (Ryerson et al., 2000).

Interactive comment on Atmos. Chem. Phys. Discuss., 8, 6469, 2008.

ACPD
8, S4011–S4021, 2008
Interactive Comment