Regional NOx emission inversion through a four-dimensional variational approach using SCIAMACHY tropospheric NO2 column observations

Tianfeng Chai, Gregory R. Carmichael, Youhua Tang, Adrian Sandu, Andreas Heckel, Andreas Richter, John P. Burrows

A R T I C L E   I N F O

Article history: 
Received 11 November 2008 
Received in revised form 5 April 2009 
Accepted 26 June 2009

Keywords: 
NO2 column 
NOx emission 
Adjoint 
Emission inversion 
Assimilation

A B S T R A C T

In this paper, the NOx emission scaling factors applied over the 2001 National Emissions Inventory (NEI) are estimated through a four-dimensional variational (4D-Var) approach using SCIAMACHY (Scanning Imaging Absorption spectroMeter for Atmospheric CHartographY) tropospheric NO2 columns measured during summer 2004. In the “top-down” approach, two-month average NO2 columns are assimilated into a regional chemical transport model (CTM), STEM, using different assimilation setups. In a basic setup, NOx emissions are adjusted by assimilating the NO2 columns. A more general setup of emission inversion allows the initial O3 concentrations be adjusted along with the NOx emissions. A final case is set up to assimilate both the NO2 columns and O3 measurement from various platforms while allowing adjustments of both the NOx emissions and the initial O3 concentrations. It is found that the addition of O3 measurements did not improve the NOx emission inversion. With the NOx emission at surface and upper levels being adjusted separately, results from four cases show that the elevated NOx emission reduction ranges from 8.9% to 11.4%, and the surface NOx emission reduction is up to 6.6%. All the cases show NOx emission reduction in Ohio valley and Washington, District of Columbia areas.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

As the most important ozone precursor and a direct contributor to the local air pollution itself, nitrogen dioxide (NO2) is one of the key species in atmospheric chemistry of earth’s troposphere. Measurements by satellite instruments Global Ozone Monitoring Experiment (GOME, from August 1995 to June 2003) spectrometer (Burrows et al., 1999) and Scanning Imaging Absorption spectroMeter for Atmospheric CHartographY (SCIAMACHY, since August 2002) (Bovensmann et al., 1999) provide continuous global coverage of NO2 columns. Richter and Burrows (2002) presented a technique using the Differential Optical Absorption Spectroscopy (DOAS) to retrieve the tropospheric NO2 columns from the GOME satellite measurements. By applying the technique to GOME and SCIAMACHY observations from 1995 to 2004, Richter et al. (2005) found significant reductions of tropospheric NO2 over parts of Europe and over the Ohio valley region in the USA. An upward trend of tropospheric NO2 over the years was observed over parts of China and in the northeast of the USA.

To a first approximation, the changes of tropospheric NO2 columns reflect the NOx emission changes. However, transport and photochemical reactions that affect NO2/NOx partitioning have to be taken into account in order to attribute the changes in NO2 levels to changes in emissions. The nonlinear relationship between NOx emissions and NO2 columns was demonstrated by Stavrakou et al. (2008). Using the IMAGES global CTM and its adjoint along with the GOME/SCIAMACHY observations, they showed that the inferred emission growth rate in Beijing region from 1997 to 2006 as 9% year−1 in both summer and winter although there is dramatic differences in the growth rate of the observed NO2 columns between the two seasons (5.3% year−1 in summertime and 11.8% year−1 in wintertime). Therefore, to infer the emission of NOx via the “top-down” approach, it is important to fully take advantage of the CTMs which have our best understanding of the physical and chemical processes thoroughly implemented.

Emission inversion problems have drawn a lot of attention in recent years. For instance, Pétron et al. (2002) constrained the global emissions of carbon monoxide (CO) by using a three-dimensional inverse modeling scheme with the IMAGES model. Palmer et al. (2003) and Wang et al. (2004) used aircraft and surface
station observations of CO and NO$_x$ during the Transport and Chemical Evolution Over the Pacific (TRACE-P) mission in combination of an optimal estimation inverse model to constrain the regional CO and NO$_x$ emissions of different sources. Space-based observations of NO$_x$ columns have been utilized to constrain NO$_x$ emissions over both global and regional scales using different methods (Martin et al., 2003; Jaegle et al., 2005; Boersma et al., 2008; Napelenok et al., 2008). Recently, some CTMs and their adjoints have been utilized in the “top-down” emission estimations. The IMAGES global CTM and its adjoint were used to invert CO and NO$_x$ emissions with both surface and satellite observations (Müller and Stavrakou, 2005; Stavrakou and Müller, 2006; Stavrakou et al., 2008). The STEM regional CTM and its adjoint model were developed for emission inversion problems and were applied to analyze the black carbon (Hakami et al., 2005) and mercury (Pan et al., 2007) emission inventories using the observations during the Asian Pacific Regional Aerosol Characterization Experiment (ACE-Asia). Henze et al. (2008) developed an inverse modeling scheme for PM$_{2.5}$ precursor emissions using the adjoint of GEOS-Chem.

In this study, we use the SCIAMACHY tropospheric NO$_x$ columns during the International Consortium for Atmospheric Research on Transport and Transformation (ICARTT) (Singh et al., 2006) operations in the summer of 2004 and the STEM regional CTM to deduce time-independent scaling factors applied to grid-based NO$_x$ emissions generated using the 2001 U.S. EPA National Emissions Inventory (NEI). In the emission inversion, two separate sets of emission scaling factors are applied to surface and elevated NO$_x$ emissions, respectively. A general setup of emission inversion that also allows the adjustment of initial concentrations of chemical species is also tested. In addition, the effect of assimilating ozone observations from various platforms is discussed.

The paper is organized as follows. Section 2 describes the SCIAMACHY data. A brief description of the STEM model and the emission inversion method using the variational approach is given in Section 3. Section 4 presents the emission inversion results. A summary is given in Section 5.

2. SCIAMACHY tropospheric nitrogen dioxide observations

The tropospheric NO$_x$ columns from SCIAMACHY used in this study were prepared at the Institute of Environmental Physics, University Bremen, Germany for the ICARTT project. As the sensitivity of nadir measurements decreases strongly towards the surface, the NO$_x$ retrieval requires the knowledge of the vertical profile shape (Richter and Burrows, 2002). For the SCIAMACHY NO$_x$ column retrievals, daily MOZART (Horowitz et al., 2003) model profiles were used for the airmass factors (AMF) calculation. Our current STEM model runs used the same MOZART profiles to provide boundaries conditions. Thus, integrating the NO$_x$ mass over the model layers without averaging kernels generates comparable NO$_x$ columns.

Although SCIAMACHY can provide global coverage of NO$_x$ columns, the region of our interest is only partially covered each day. Fig. 1(a) shows the observed regions on July 20, 2004. The observing time is around 1030 local time. To eliminate the effect of clouds on the retrievals, a simple criterion, the normalized intensity < 0.15, was used to select cloud-free regions. Fig. 1(b) shows the tropospheric NO$_x$ columns on July 20, 2004 after removing the cloud regions. The number of measurements is significantly reduced after the cloud removal. Hereafter, only data associated with the normalized intensity less than 0.15 are used.

For the NO$_x$ columns shown in Fig. 1, not all the NO$_x$ emissions will be constrained if only such spatially incomplete samplings are incorporated in the emission inversion. To have a better constraint from the measurements in the “top-down” emission inversion, it is desirable to have a near complete sampling throughout the domain. Here we chose to merge the available NO$_x$ columns from July 1 to August 31, 2004 together by neglecting the daily variations.

Fig. 2 shows the cloud-free measurements during the weekends and weekdays from July 1 to August 31, 2004, respectively. Even with the data of 18 different weekends stacked together, there are still significant blank regions. Merging the 44 weekdays generates a much better coverage of the domain, although there are still areas left without observations, e.g., middle Pennsylvania. However, these areas are well represented by the data during the weekends. Fig. 2 also shows little contradictions between measurements from different days. Both weekday and weekend data sets show strong and similar spatial patterns revealing the various levels of anthropogenic activities at different locations. There are slight differences between the data sets of weekdays and weekends, but the differences are relative small compared with the spatial variations. In order to achieve a complete spatial coverage in the domain, we stacked both weekday and weekend data together. Such merging also helps to eliminate the effect of daily variation in NO$_x$ emission, which is not of our interest here.

After stacking all the data together, we calculated the mean and standard deviation of the observations inside each grid cell. Whether a measurement is inside a grid cell is decided by the

![Fig. 1. SCIAMACHY NO$_x$ columns on July 20, 2004. Unit: molecules cm$^{-2}$. Original data shown on the left (a); data with the normalized intensity less than 0.15 shown on the right (b).](image-url)
center of the pixel. The standard deviation was then normalized by the local mean to indicate the variation of the observation in the two-month period. Fig. 3 shows the distributions of the mean observations and normalized standard deviation in the computational domain. For 90.4% of the grid cells, the normalized standard deviation is smaller than 70%. Except for the New York City area, the locations with large variations are often associated with low NO2 columns. Note that the deviations inside each grid cell not only come from the daily variations, but also come from the spatial variations not resolved by the model. The footprint of a SCIAMACHY pixel is approximately 60 km x 30 km, i.e., about half of the 60 km x 60 km model grid cell. Thus, the standard deviations calculated inside grid cells include the representative errors. As Fig. 2 shows little variation between days and the daily variation is not of our interest here, we generate a pseudo-observation data set that has the two-month mean NO2 column at each grid cell as the measurement value. The normalized standard deviation at each grid cell reflects the uncertainty of the pseudo-observation and is assigned as the uncertainty of the pseudo-observations. Fig. 3 shows the normalized standard deviation is close to 50% for most of the domain (81% of grid cells fall between 30% and 70%). In addition, the average measurement time during the day inside each grid cell was given to the pseudo-observation and it is used by the model to reconstruct the NO2 columns. Note that it is difficult to give a good estimate of the NO2 column retrieval uncertainty. If the SCIAMACHY retrieval uncertainty is not significantly greater than 50%, as most pseudo-observation uncertainties are close to, including the retrieval uncertainty is not expected to add more uncertainties to the pseudo-observation uncertainties.

3. Method

3.1. Chemical transport model

In this study, the STEM-2K3 (Tang et al., 2004) regional chemical transport model is employed. It is a flexible regional-scale chemical transport model using SAPRC99 chemical mechanism (Carter, 2000) with on-line photolysis solver (Tang et al., 2003). Meteorological inputs to the model came from the fifth-generation Mesoscale Model (MM5) using NCEP (National Centers for Environmental...
Prediction) FNL (Final Global Data Assimilation System) analyzed
data during post-analysis. A grid with a 60 km horizontal resolution
(25 cells in longitude, and 22 cells in latitude) is used over the
northeast US domain, as shown in Fig. 3. Vertically the model had 21
layers, extending from the surface to 100 hPa using 0.999, 0.9965,
0.9925, 0.985, 0.97, 0.945, 0.91, 0.87, 0.825, 0.77, 0.71, 0.65, 0.59,
0.53, 0.47, 0.41, 0.35, 0.285, 0.21, 0.125, and 0.04 in sigma coordi-
nates. The emissions inventory was based on the 2001 NEI, with
updated large point source emissions (Frost et al., 2006). Upper
troposphere lightning NOX emissions were added to the model
based on the National Lightning Detection Network (NLDN),
modulated by signal strength and multiplicity of flashes. Further
information about the lightning emissions can be found in Tang
et al. (2007). Biogenic emissions were estimated using Biogenic
Emissions Inventory System 2 (BEIS2) which generates time-varied
isoprene and monoterpene emissions driven by meteorological
variables from MM5. Forest fires that occurred during the ICARTT
period were largely outside the model domain (in Alaska and
Northwestern Canada), therefore their influence was incorporated
through lateral boundary conditions from MOZART global chemical
model predictions. The boundary conditions are provided by STEM-
2K3 run over a bigger domain that covers the continental United
States (see Tang et al., 2007 for detail).

The evolution of the chemical constituent concentration vector $c$
in time ($t$) can be described as

$$
\frac{\partial c}{\partial t} = -u \cdot \nabla c + \frac{1}{\rho} \nabla \cdot (\rho K \nabla c) + \frac{1}{\rho} f + E
$$

Here we denote by $u$ the wind field vector, $\rho$ the air density, $K$
the turbulent diffusivity tensor, $f$ the chemical transformation rate, and $E$
the emission rate.

### 3.2. Emission inversion via 4D-Var

In this study, the emission inversion problem is solved via the
STEM 4D-Var system (Sandu et al., 2005; Chai et al., 2006, 2007).
The discrepancy between the available observations and model
counterparts is built into a cost functional. Optimal solutions of
model parameters such as the NOx emissions are obtained by
finding new parameters which minimize the cost functional. To
solve the minimization problems efficiently, the sensitivity of the
cost functional with respect to the control variables (parameters to
be adjusted) need to be calculated. Adjoint models provide the
most efficient way to calculate the gradients of a scalar cost func-
tional with respect to a large number of control variables ( Talag-
rand and Courtier, 1987).

The cost functional $J$ is defined as

$$
J = \frac{1}{2} \left[ (r - 1)^T E^{-1} (r - 1) + \beta (c_0 - c_b)^T B^{-1} (c_0 - c_b) \right] + \frac{1}{2} \left[ (y - h(c))^T O^{-1} (y - h(c)) \right]
$$

$E$, $B$, and $O$ are error covariance matrices for emission scaling
factors $r$, a priori (background) initial states $c_0$, and observations in
discrete spaces, respectively. $h$ is a projection operator, calculating
the observation vector $y$ from the model space $c$. In the current
study the control variables can include both initial states $c_0$
and emission rates. The subscript “0” is used to denote variables at the
instant $t = 0$. Assuming that the operator $h$ is linear, $h(c)$ can be
written as $h(c) = H \cdot c$. In our application, $H$ reflects vertical inte-
gration and linear interpolation in time when constructing model
counterparts of the NO2 columns at the same measurement time.

A larger-scale bound-constrained limited-memory quasi-
Newton code, L-BFGS-B (Zhu et al., 1997) is used for the mini-
mization. The maximum number of iterations is set to be 25. Tests that
add 25 extra iterations show little improvements in cost functiona-
reduction. The initial O3 background error covariance $B$ were esti-
minated using both NMC and observational methods. Truncated
singular value decomposition (SVD) regularization is used for the
inversion of $B$ matrix (see Chai et al., 2007 for detail). Assuming
uncorrelated emission scaling factor errors in space, the uncer-
tainty of emission scaling factor $r$ was chosen as 0.5 uniformly.
The upper and lower bounds of $r$ during the minimization with L-BFGS-
B subroutine were assigned as 10 and 0.1, respectively. The obser-
vation errors of NO2 pseudo-columns were given by the standard
deviations at individual grid cells during the two-month period.
Following Chai et al. (2007), the observation errors of O3 were set to
be 8 ppbv everywhere. Errors of NO2 pseudo-columns and O3
observations were assumed uncorrelated among themselves and
between each other.

The assimilation time window is chosen to be 24 h for the
following emission inversion cases. We first started our emission
inversion tests by only adjusting the NOx emission rates. Later we
included the initial ozone concentrations as control variables as
well. It is designed to eliminate the effect of the notable errors in

Fig. 4. Comparison between SCIAMACHY NO2 columns and “model counterparts”, before (left) and after (right) averaging in each grid cell. “Model counterparts” were generated by assuming all the observations were measured on July 20, 2004.
the initial ozone states. The most recent work by Elbern et al. (2007) studied the effect of combining both emission rate and initial state as control variables. Also note that we chose emission scaling factors \( \epsilon \) instead of the emission rates as control parameters (see Hakami et al. (2005); Pan et al. (2007) for detail). However, the linear penalty applied to the emission scaling factors is prone to low-biased results. To avoid the problem, a logarithmic function of the scaling factors could be introduced into the cost functional.

Appendix A illustrates the association between emission sensitivity and adjoint variables through the “continuous adjoint” approach. In our application, the “discrete adjoint” approach is implemented to assure consistency. This is briefly described in Appendix B. More discussion on “continuous adjoint” and “discrete adjoint” can be found in Sandu et al. (2005).

4. Results

4.1. Results without assimilation

First we included all the measurements and assumed the two-month observations were measured on July 20, 2004. Fig. 4 shows the comparison between the SCIAMACHY NO\(_2\) columns and the “model counterparts” on July 20, 2004. It is seen that the model mostly overestimates the low NO\(_2\) column values (\(< 10^5\) molecules cm\(^{-2}\)). Then we considered the pseudo-columns which were generated by averaging multiple observations during the two-month period at each grid cell. Fig. 4 shows a good agreement between such pseudo-observations and their “model counterparts”. Note that the low NO\(_2\) column values were effectively removed after the averaging for the pseudo-observations.

Fig. 5 shows model predictions of NO\(_2\) columns on July 17 (Saturday), 18 (Sunday), 19 (Monday), and 20 (Tuesday).

<table>
<thead>
<tr>
<th>Date</th>
<th>Bias</th>
<th>Mean error</th>
<th>RMS error</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>July 17, 2004</td>
<td>13.6</td>
<td>25.2</td>
<td>41.4</td>
<td>0.336</td>
</tr>
<tr>
<td>July 18, 2004</td>
<td>8.7</td>
<td>18.1</td>
<td>29.6</td>
<td>0.491</td>
</tr>
<tr>
<td>July 19, 2004</td>
<td>11.7</td>
<td>17.7</td>
<td>28.9</td>
<td>0.587</td>
</tr>
<tr>
<td>July 20, 2004</td>
<td>6.7</td>
<td>14.6</td>
<td>33.4</td>
<td>0.673</td>
</tr>
</tbody>
</table>
respectively. The NO$_2$ columns were constructed at the same time of the day when the measurement was made. The current emission scheme differentiates Saturdays and Sundays from weekdays, i.e., it has 3 different temporal daily profiles. Since the NO$_2$ columns are dominated by the lower level NO$_2$ concentrations, the different NOx daily emission profiles generated significantly different NO$_2$ columns. It is shown in Fig. 5 that NO$_2$ predictions on July 17 and 18 are dramatically different from those predicted during the weekdays. The differences between July 19 and 20 predictions are relatively small, albeit apparent. The distribution of NO$_2$ column predictions on the weekdays resemble that of the SCIAMACHY observations, showing a small bias of $6.7 \times 10^{14}$ molecules cm$^{-2}$, and correlation coefficient as 0.673.

4.2. Assimilation results

The emission inversion tests are listed in Table 2. In all four tests, the NOx emission scaling factors are adjusted. Considering the different nature of the two major NOx sources, transportation and power plants, we adjust the surface and elevated (above the first level) NOx emissions with two different sets of emission scaling factors. No temporal variation is assumed for the scaling factors. That is, the diurnal variations of the original emissions are preserved. In the emission inversion, the valid range of the emission scaling factors is set to be between 0.1 and 10.0.

In case EM01, NO$_2$ columns are assimilated on July 20, 2004 with a 24-h time window. The distributions of the NOx emission scaling factors at the surface and upper levels are shown in Fig. 6. In most of the domain there is little NOx emission adjustment, indicated by the regions with the emission scaling factors close to one. Close to the area where the states of Ohio, West Virginia, and Pennsylvania join each other hereinafter referred as the OWP area, both surface and elevated emissions show downward adjustment to alleviate the model overestimation (see Figs. 3 and 5). Similar emission reductions are found at Detroit and Washington DC areas for both surface and elevated emissions as well. At the surface, there are several other locations with NOx emissions adjusted up or down. However, these locations are often not associated with large NOx emission sources. To show the overall emission adjustment in the domain, we multiply the emission scaling factors with their base emission rates. The results are shown in Table 3. It shows that the elevated emissions are reduced by 12.9% and the surface emissions are reduced by only 6.6%. Combined together, the total NOx emissions are reduced by 8.0% after assimilating the SCIAMACHY data. After the emission adjustment, the model results of NO$_2$ columns improve significantly. As shown in Tables 1 and 3, the bias is reduced from $6.7 \times 10^{14}$ molecules cm$^{-2}$ to $0.6 \times 10^{14}$ molecules cm$^{-2}$, and root-mean-square (RMS) error is reduced from 33.4 to 12.2 $\times 10^{14}$ molecules cm$^{-2}$. The correlation between the model and observation also improves, with correlation coefficient increased from 0.673 to 0.819.

The setup of case EM02 is same as case EM01, except that the assimilation time window is shifted one day earlier. Note that there are not only changes in the meteorological fields, but also differences in the atmospheric chemistry states. The distributions of the NOx emission scaling factors at the surface and upper levels are shown in Fig. 6. In most of the domain there is little NOx emission adjustment, indicated by the regions with the emission scaling factors close to one. Close to the area where the states of Ohio, West Virginia, and Pennsylvania join each other hereinafter referred as the OWP area, both surface and elevated emissions show downward adjustment to alleviate the model overestimation (see Figs. 3 and 5). Similar emission reductions are found at Detroit and Washington DC areas for both surface and elevated emissions as well. At the surface, there are several other locations with NOx emissions adjusted up or down. However, these locations are often not associated with large NOx emission sources. To show the overall emission adjustment in the domain, we multiply the emission scaling factors with their base emission rates. The results are shown in Table 3. It shows that the elevated emissions are reduced by 12.9% and the surface emissions are reduced by only 6.6%. Combined together, the total NOx emissions are reduced by 8.0% after assimilating the SCIAMACHY data. After the emission adjustment, the model results of NO$_2$ columns improve significantly. As shown in Tables 1 and 3, the bias is reduced from $6.7 \times 10^{14}$ molecules cm$^{-2}$ to $0.6 \times 10^{14}$ molecules cm$^{-2}$, and root-mean-square (RMS) error is reduced from 33.4 to 12.2 $\times 10^{14}$ molecules cm$^{-2}$. The correlation between the model and observation also improves, with correlation coefficient increased from 0.673 to 0.819.

The setup of case EM02 is same as case EM01, except that the assimilation time window is shifted one day earlier. Note that there are not only changes in the meteorological fields, but also differences in the atmospheric chemistry states. The distributions of the NOx emission scaling factors at the surface and upper levels are shown in Fig. 6. In most of the domain there is little NOx emission adjustment, indicated by the regions with the emission scaling factors close to one. Close to the area where the states of Ohio, West Virginia, and Pennsylvania join each other hereinafter referred as the OWP area, both surface and elevated emissions show downward adjustment to alleviate the model overestimation (see Figs. 3 and 5). Similar emission reductions are found at Detroit and Washington DC areas for both surface and elevated emissions as well. At the surface, there are several other locations with NOx emissions adjusted up or down. However, these locations are often not associated with large NOx emission sources. To show the overall emission adjustment in the domain, we multiply the emission scaling factors with their base emission rates. The results are shown in Table 3. It shows that the elevated emissions are reduced by 12.9% and the surface emissions are reduced by only 6.6%. Combined together, the total NOx emissions are reduced by 8.0% after assimilating the SCIAMACHY data. After the emission adjustment, the model results of NO$_2$ columns improve significantly. As shown in Tables 1 and 3, the bias is reduced from $6.7 \times 10^{14}$ molecules cm$^{-2}$ to $0.6 \times 10^{14}$ molecules cm$^{-2}$, and root-mean-square (RMS) error is reduced from 33.4 to 12.2 $\times 10^{14}$ molecules cm$^{-2}$. The correlation between the model and observation also improves, with correlation coefficient increased from 0.673 to 0.819.
Emission reductions are seen around the OWP area and Washington DC in both surface and upper levels, similar as case EM01 shown in Fig. 6. However, the emission reduction in the OWP area for case EM02 is greater than that in case EM01. For the surface NOx emissions, there are differences in the direction of the adjustments at several locations, such as the northeast and northwest corners of the domain. For the elevated emissions, results of the two cases resemble each other except for slight differences in northern Virginia and western Erie lake areas. The emission changes over the whole domain after the adjustment for case EM02 are very close to case EM01, especially for the emissions at the upper levels, with both cases showing a 12.1% reduction over the domain.

In cases EM01 and EM02, only the NOx emissions are adjusted to fit the model predictions of the SCIAMACHY NO2 columns. This approach assumes that the only source of error is the emission, thus it attempts to minimize the model prediction errors by adjusting emissions only. With significant uncertainties in many other parameters, such as initial and boundary conditions, reaction rates, and meteorological fields, the emission adjustments may yield faulty results due to the errors in the other model parameters. In cases IE01 and IE02 which are listed in Table 2, we extend the emission inversion to include the simultaneous adjustment of additional parameters. Adjoint sensitivity shows that NO2 columns are more sensitive to the initial concentrations of O3 than those of the other species, including NO2. It is probably because NO2 plays a key role in the troposphere ozone chemistry and NO2 has a short lifetime. We chose to add initial ozone concentrations as control variables.

Fig. 8 shows the distributions of the NOx emission scaling factors from case IE01. Great similarity is found between Figs. 6 and 8, especially for the elevated emissions. At the surface, the magnitude
of the emission adjustment tends to be smaller than for the case EM01 in most regions. As listed in Table 3, the total emission reduction at surface over the computational domain (2.7%) is less than half of case EM01 (6.6%). The emission reduction for IE01 at the upper levels (12.4%) is close to the result of case EM01 (12.1%). By allowing additional parameters to be adjusted, the model predictions of NO2 are slightly better than case EM01, as indicated by model bias, RMS error, and correlation coefficient listed in Table 3.

In case IE01, it is found that allowing the NO2 column observations to impact both the emission scaling factors and the initial O3 improves the NO2 predictions. We also explored whether assimilating additional observations would improve the results. In case IE02, ozone observations from various platforms on July 20, 2004 (see Chai et al., 2007 for details) are added to the observation set assimilated. Same as case IE01, both the emission scaling factors and the initial O3 concentrations are treated as control variables. Fig. 9 shows the distributions of NOX emission scaling factors yielded from case IE02, which are quite different from the previous cases. However, the emission reduction of 8.9% over the domain is close to the previous cases. Compared with cases IE01 and EM01 (see Table 3), IE02 generates NO2 predictions with worse bias ($3.6 \times 10^{14}$ molecules cm$^{-2}$), larger RMS error ($15.1 \times 10^{14}$ molecules cm$^{-2}$), and smaller correlation coefficient (0.794). This suggests that the additional O3 observations do not help constraining the NOX emissions. It is probably due to the high uncertainty in other parameters involved in the ozone chemistry, such as the emissions of volatile organic compounds (VOCs).

5. Summary

In a “top-down” approach, SCIAMACHY NO2 column data in the summer of 2004 are assimilated into a chemical transport model (CTM), STEM. We demonstrate that the four-dimensional variational (4D-Var) approach allows a more general framework in emission inversion. In this general setup, more uncertain model parameters can be adjusted in addition to the emission fields. Observations of different species in various formats can also be used to constrain the inverse problem.

The test cases show that the emission inversion results are sensitive to the problem setup. Emission scaling factors from the last case where both ozone observations and NO2 column data are assimilated while allowing initial ozone and NOx emissions to be adjusted are very different from the other cases. It is probably not beneficial to add ozone observations to constrain the NOx emissions because of the great uncertainties associated with VOCs.

When only NO2 column data are assimilated, adding initial ozone concentrations as control variables to be adjusted generates less adjustment in NOx emissions, especially for the surface emissions. All the cases show consistent results over the Ohio valley region and Washington, District of Columbia area, revealing the NOx emission reduction. With emissions at surface and upper levels adjusted separately, we found the results at upper levels are quite robust. The elevated NOx emission reduction results from four test cases range from 8.9% to 11.4%, indicating the power plant NOX emission reduction from 2001 to 2004. Stavrakou et al. (2008) reported their inferred posterior NOx emissions decreased by 35% between July 1997 and 2006 in the Ohio River Valley, with annual change rate of $-4.3\%$ year$^{-1}$ in summertime. This agrees well with our results.

While SCIAMACHY satellite observations provide more than adequate data for global model emission inversion, their temporal and spatial resolutions are lacking for a regional model application to resolve detailed grid-based emission inversion. Ignoring the daily variability, we stacked together the NO2 columns in two months to generate a pseudo-observation set that has the two-month mean at each grid cell. Such pseudo-observations do not reflect the daily variations in chemistry and physics of the troposphere. Also note that the satellite observations are usually least sensitive to the boundary layer where most of the emissions occur. In addition, the 24-h assimilation time window chosen for the NOx emission inversion tests has its drawback as well. Although we showed that the NOx emission inversion tests using two different days generated very similar results, a full 2-month inversion wherein NO2 columns from each day are assimilated separately is still preferable. While our main focus here is on the formulating of the emission inversion problem, we plan to apply such pseudo-observations as well as the original data over a larger domain for an
extended time period in the future. Such emission inversion results will be able to provide more valuable information to evaluate and improve the “bottom-up” emission inventories. In the end, we want to emphasize that the current emission inversion formulation can be easily applied to other types of measurements, although here we focus on the emission inversion using the SCIAMACHY NO2 columns.

Acknowledgments

The authors gratefully thank NASA, NOAA, and the National Science Foundation for their support grants.

Appendix A: Emission sensitivity via continuous adjoint

Here we first consider a simple one dimensional problem involving only one species. By assuming constant air density, and ignoring the advection and reaction terms, the transport equation is further simplified and becomes

\[
\frac{\partial c}{\partial t} + \frac{\partial c}{\partial z} + E(z,t) = 0
\]  

(3)

where \( c = c(z,t) \) and \( z \in (0,L), t \in (0,T) \). As we only aim to illustrate how the emission sensitivity associates with the adjoint variables, instead of a cost functional in the least square form as in equation (2), we define a general response functional as

\[
J = \frac{1}{LT} \int_{0}^{L} \int_{0}^{T} g(z,t) dt dz
\]  

(4)

To derive the adjoint equation, we introduce the Lagrange multiplier \( \lambda_c(z,t) \). Multiply equation (3) with it and integrate over computational domain (a factor of 1/L is added). Subtract the result from equation (4), we get

\[
f' = J - \frac{1}{L} \int_{0}^{L} \int_{0}^{T} \lambda_c \left[ \frac{\partial c}{\partial t} - \frac{\partial c}{\partial z} \left( \frac{\partial c}{\partial z} \right) E(z,t) \right] dt dz
\]  

(5)

The variation of equation (5) yields

\[
f' = \delta f - \frac{1}{L} \int_{0}^{L} \int_{0}^{T} \delta \lambda_c \left[ \frac{\partial c}{\partial t} - \frac{\partial c}{\partial z} \left( \frac{\partial c}{\partial z} \right) E(z,t) \right] dt dz
\]  

(6)

Integrating by parts gives

\[
f' = \frac{1}{L} \int_{0}^{L} \int_{0}^{T} \delta g(z,t) dt dz - \frac{1}{L} \int_{0}^{L} \int_{0}^{T} \delta \lambda_c \left[ \frac{\partial c}{\partial t} - \frac{\partial c}{\partial z} \left( \frac{\partial c}{\partial z} \right) \right] dt dz
\]

\[
- E(z,t) \right] dt dz \frac{1}{L} \int_{0}^{T} \left[ - \lambda_c \frac{\partial c}{\partial z} + \frac{\partial c}{\partial z} \left( \frac{\partial c}{\partial z} \right) \right] dt dz
\]

\[
+ \frac{1}{L} \int_{0}^{L} \left[ \lambda_c \delta c \right]_{t=0}^{t=T} dz + \frac{1}{L} \int_{0}^{T} \lambda_c \delta E(z,t) dt dz
\]  

(7)

The second term on the right-hand side vanishes as the state variable satisfies equation (3). Combine the first and the third terms and force them to vanish. This gives the adjoint equation

\[
\frac{\partial \lambda_c}{\partial t} = \frac{\partial}{\partial z} \left( \frac{K \lambda_c}{\partial z} \right) - \frac{1}{T} \frac{\partial}{\partial t} E(z,t)
\]  

(8)

The fourth term would vanish with proper boundary conditions (e.g., \( \lambda_c = 0 \) at Dirichlet boundaries). Forcing the fifth term to be zero gives the initial condition for the adjoint variable \( \lambda_c \) at \( t = T \) as

\[
\lambda_c(z,T) = 0
\]

Then equation (7) becomes

\[
\delta f' = \frac{1}{L} \int_{0}^{L} \int_{0}^{T} \left[ \delta \lambda_c \right]_{t=0}^{t=T} dz + \frac{1}{L} \int_{0}^{T} \lambda_c \delta E(z,t) dt dz
\]  

(9)

This provides the variational sensitivity information for both the initial and emission functions.

\[
\frac{\delta f'}{\delta \lambda_c(z,t)} = \lambda_c(z,t) + \frac{\delta E(z,t)}{E(z,t)}
\]  

(10)

If we introduce emission scaling factors \( \varepsilon \) as

\[
\varepsilon(z) = \frac{E(z,t)}{E_0(z,t)}
\]

Note \( \varepsilon(z) \) does not to vary in time. The last term in equation (7) would become

\[
\frac{\delta f'}{\delta \varepsilon(z)} = \frac{1}{L} \int_{0}^{L} \int_{0}^{T} \lambda_c E_0(z,t) dt
\]

(11)

Thus, the sensitivity of the response function to the emission scaling factor has the following form

\[
\delta f' = \frac{T}{\lambda_c E_0(z,t) dt}
\]

(12)

Without detail, here we give the emission sensitivity for species \( i \) in a four-dimensional air quality model,

\[
\frac{\delta f'}{\delta q_i(x,y,z)} = \int_{0}^{T} \lambda_c(x,y,z) E_0(x,y,z,t) dt
\]  

(13)

In the current study, emission rates are adjusted separately using two sets of 2-D functions, \( \varepsilon'(x,y) \) for the surface and \( \varepsilon'(x,y) \) for higher levels.

\[
\frac{\delta f'}{\delta q_0(x,y)} = \int_{0}^{T} \lambda_c(x,y,z) q_0(x,y,t) \Delta z dt
\]  

(14)

\[
\frac{\delta f'}{\delta q_0(x,y,z)} = \int_{0}^{T} \lambda_c(x,y,z) E_0(x,y,z,t) \Delta z dt
\]  

(15)

where \( \Delta z_1 \) is the depth of the first layer. \( z_1 \) and \( z_{\text{top}} \) are the height of the first and top layer, respectively. Note that the surface area emission \( q_0(x,y,t) \) includes the first level volume emission rate, i.e.

\[
q_0(x,y,t) = q_0 \text{surface}(x,y,t) + \int_{0}^{z_1} E_0(x,y,z,t) dz
\]  

(16)
Appendix B: Emission sensitivity via discrete adjoint

As operator-splitting technique is implemented in solving equation (1), the emission sensitivity is only directly associated with vertical transport. Using Crank–Nicholson time stepping the concentrations and forward Euler for boundaries and the surface emissions, the forward discrete model evolving the concentration column vector \( \mathbf{C}_i \) from time step \( n \) to \( n + 1 \) for vertical transport reads

\[
\mathbf{C}^{n+1}_i = \left( I - \Delta t \nabla^2 \right)^{-1} \left( I + \Delta t \nabla^2 \right) \mathbf{C}^n_i + \Delta t \mathbf{B}(t^n) \mathbf{e}_N + \sum_{j=2}^{N} \mathbf{G}(t^n) \mathbf{e}_j + \frac{q_i(t^n)}{\Delta z_1} \mathbf{e}_1
\]

(19)

where matrix \( A \) depends on the wind field, the diffusion tensor, and the air density. \( B \) is a scalar that accounts for the top boundary, \( e_j \) is the \( j \)th column of the identity matrix. The adjoint sensitivity with respect to emission rates can be calculated as

\[
\frac{\partial f}{\partial q_i(t^n)} = \frac{\partial \mathbf{C}^{n+1}_i}{\partial q_i(t^n)}^T \left( I - \Delta t \nabla^2 \right)^{-1} \left( I + \Delta t \nabla^2 \right) \frac{\partial \mathbf{C}^n_i}{\partial q_i(t^n)} = \frac{\partial \mathbf{C}^{n+1}_i}{\partial q_i(t^n)}^T \frac{\partial \mathbf{C}^n_i}{\partial q_i(t^n)} + \frac{\Delta t}{\Delta z_1} \frac{\partial q_i(t^n)}{\partial q_i(t^n)} \frac{\partial q_i(t^n)}{\partial q_i(t^n)} \Delta t
\]

(20)

Since the vector \( (I - \Delta t \nabla^2) A^T \)\( (I + \Delta t \nabla^2) \) is already computed during the update of \( A \), there is little additional cost to calculate the emission sensitivity. If we choose the non-time-varying emission scaling factors \( e_i \) as control variables, their sensitivities can be written as

\[
\frac{\partial f}{\partial e_i} = \sum_{n=0}^{N-1} \frac{1}{\Delta z_1} \mathbf{G}(t^n) \left( I - \frac{\Delta t}{2} \nabla^2 \right)^{-1} \left( I + \frac{\Delta t}{2} \nabla^2 \right) \frac{\partial q_i(t^n)}{\partial q_i(t^n)} \Delta t
\]

(22)

\[
\frac{\partial f}{\partial e_i} = \sum_{n=0}^{N-1} \sum_{j=2}^{N} \frac{1}{\Delta z_1} \mathbf{G}(t^n) \left( I - \frac{\Delta t}{2} \nabla^2 \right)^{-1} \left( I + \frac{\Delta t}{2} \nabla^2 \right) \frac{\partial q_i(t^n)}{\partial q_i(t^n)} \Delta t
\]

(23)

where \( N \) and \( N_e \) are total number of time steps and vertical levels, respectively.

References


