SATELLITE-BASED INVERSION OF NO_X EMISSIONS USING THE ADJOINT VERSION OF CMAQ

Amir Hakami* and John H. Seinfeld
California Institute of Technology

Qinbin Li
Jet Propulsion Laboratory, NASA

Adrian Sandu and Kumaresh Singh
Virginia Polytechnic Institute and State University

Daewon Byun, Peter P. Percell, and Violeta Coarfa
University of Houston

1. INTRODUCTION

An important application of sensitivity analysis methods is in inverse modeling and data assimilation. Among various sensitivity methods, adjoint analysis is particularly attractive for such applications due to its receptor-based nature. In the adjoint method, sensitivities of a cost function with respect to a large number of input parameters are calculated efficiently. Therefore, adjoint-based inversions are capable of adjusting very large number of parameters, and providing high spatial and/or temporal resolution for such adjustments. Here we use a newly developed adjoint version of CMAQ for assimilating NO_2 column observations of SCIAMACHY over a regional domain.

In data assimilation, sensitivity information is used to minimize a cost function, i.e., a measure of mismatch between simulations and observations. In our application we define the cost function as:

$$J = \frac{1}{2}\left[\frac{1}{\mu}\sum_k B_k^{-1}\left(\alpha_k^b - \alpha_k\right)^2\right] + \frac{1}{2}\left[\sum_i O_i^{-1}\left(C_{sim} - C_{obs}\right)^2\right] \qquad (1)$$

where index i represents the time and location of observations, and index k represents the particular parameter to be estimated. *B* and *O* are error covariance matrices for backgrounds and observations, respectively. The goal is to reduce the observation-based portion of the cost function (first term) by adjusting the input variables away from the background values (first guess), with the least possible penalty for such adjustments (second term). The regularization parameter μ is introduced to balance the magnitudes of the observation and background portions of the cost function, or alternatively to put more emphasis on one or the other. Ideally, it is desirable to have both portions of the cost function at about the same magnitude. Since both parts of the cost function change as the assimilation progresses through iterations (observation part reduces, while background part increases), a-priori estimation of the optimal regularization parameter is not possible. However, our experience in adjoint data assimilation has shown that the optimization process and outcome is only modestly sensitive to the regularization parameter [Hakami *et al.*, 2005].

We assign scaling factors to the emissions of NO_x ($NO+NO_2$) for each location, assuming these factors are constant in time. The gradients of the cost function with respect to NO and NO_2 emissions are integrated during the adjoint calculations and for each iteration. These gradients for locations with no emissions of NO_x will be zero, and therefore, the scaling factors for those locations will remain unchanged during the optimization process. Note that some of the errors in predictions of CMAQ are likely to be introduced from inaccuracy in inputs other than emissions. However, in this work we do not account for the gradients of the cost function with respect to other inputs.

**Corresponding author:* Amir Hakami, M/C 210-41 California Institute of Technology, Pasadena, CA 91125; e-mail: amir@cheme.caltech.edu

2. SATELLITE RETRIEVALS

SCIAMACHY is a UV spectrometer on board the EAS ENVISAT satellite launched in March 2002. The sensor alternatively scans in the nadir and limb viewing directions. It has a local overpass time of approximately 10:00 UTC. Individual nadir pixels cover a surface area of 30×60 km^2, and SCIAMACHY achieves global coverage every 6 days.

We assimilate the SCIAMACHY tropospheric NO_2 column densities. The SCIAMACHY tropospheric NO_2 column data used in this study is obtained from the Tropospheric Emission Monitoring Internet Service (TEMIS) project (http://www.temis.nl). There are several steps involved in the retrieval of SCIAMACHY tropospheric NO_2. The first step is spectral fitting of SCIAMACHY Level 1 radiance to retrieve total NO_2 slant column densities. The spectral fitting technique used is Differential Optical Absorption Spectroscopy (DOAS) [*Platt,* 1994]. The second step involves the correction for stratospheric NO_2 and the air-mass factor (AMF) correction [*Boersma et al.,* 2004]. The SCIAMACHY NO_2 slant column data are then assimilated into a global 3D chemistry and transport model (CTM). The stratospheric NO_2 slant column is then deduced from the assimilation results and then subtracted from the retrieved SCIAMACHY total slant NO_2 column to get a tropospheric slant NO_2 column. Then the tropospheric NO_2 vertical column is retrieved using the global 3D CTM simulated tropospheric NO_2 profiles (co-located with each SCIAMACHY pixel individually) and combined with cloud information (i.e., cloud fraction and cloud top height). The retrieval includes surface albedo values constructed by using a combination (on a monthly basis) of TOMS and GOME spectral surface reflectivities. No aerosol correction is applied as the cloud retrieval will be influenced by aerosol and further supported by error analysis [*Boersma et al.,* 2004]. The individual retrievals have a typical precision of 35-60%. This uncertainty is dominated by the uncertainty in the estimate of the tropospheric air mass factor. The most important uncertainties associated with the computation of the tropospheric air mass factor are the cloud fraction, surface albedo and NO_2 profile shape. Detailed information about the retrieval method and error analysis is presented by *Boersma et al.* [2004].

In principle the comparison between SCIAMACHY NO_2 column observations and CMAQ simulated NO_2 fields consists of three steps. First, the CMAQ 3D NO_2 filed is interpolated in time and space to produce a NO_2 vertical profile $\boldsymbol{X}(k, t)$ (in partial subcolumns) at each of the SCIAMACHY measurement locations k and times t (i.e., each pixel) on the pressure grid of the averaging kernel. Second, the averaging kernel vector $\boldsymbol{A}$ of the retrieval is applied to CMAQ simulated NO_2 vertical profile to produce a model-prediction of the "retrieved" column $\boldsymbol{Z}(k, t)$, to be compared with the SCIAMACHY retrieved NO_2 vertical column

$$\boldsymbol{Z}(k, t) = \boldsymbol{A}(k, t) \cdot \boldsymbol{X}(k, t) \qquad (2)$$

Third, we bin the SCIAMACHY and CMAQ (after steps 1 and 2) data onto the CMAQ grid. This comparison approach is independent of *a priori* assumptions on the vertical profile of NO_2 in the troposphere as needed in the retrieval.

For our purpose, we are interested in the tropospheric NO_2 columns only. Therefore equation (1) becomes

$$\boldsymbol{Z}_{trop}(k, t) = \boldsymbol{A}_{trop}(k, t) \cdot \boldsymbol{X}_{trop}(k, t) \qquad (3)$$

where $\boldsymbol{X}_{trop}$ is CMAQ simulated tropospheric NO_2 profile and $\boldsymbol{A}_{trop}$ is the averaging kernel for tropospheric retrievals defined as:

$$\boldsymbol{A}_{trop} = \boldsymbol{A} \cdot \boldsymbol{AMF} / \boldsymbol{AMF}_{trop} \qquad (4)$$

The results can then be directly compared with SCIAMACHY retrieved tropospheric NO_2 vertical columns.

An alternative approach that is easier to implement is to interpolate the SCIAMACHY averaging kernel into the CMAQ grid. This averaging kernel remains the same during the iterative process, and is used in the backward simulation to inject the forcing terms for the adjoint calculations. At each observation time step, the forcing term for each observed pixel/grid is comprised of the derivative of the integrated column density with respect to the concentrations at each CMAQ layer. For practical purposes, the forcing terms are injected at the synchronization time step closest to the actual scan time. CMAQ

column densities are integrated up to the tropopause (among SCIAMACHY's data). Since the retrieved pixels are larger than resolution used in this study, each retrieval may fall over few grid cells. In that case, the value of the retrieval is assigned to all the cells, however, the variance of the retrieval in each cell is considered to be inversely proportional to the area of the cell that is covered by the SCIAMACHY pixel. In other words, if a grid cell is completely covered by a pixel, the retrieval's error is assigned to the observation uncertainty in that cell. But if only a quarter of the area of a cell is covered by the pixel, the uncertainty in that cell will be considered to be twice the size of the retrieval error. Cells with less than 10% coverage are excluded.

3. RESULTS

To evaluate the effectiveness of the adjoint method for emission inversion, an ideal test case was developed, where emissions of NO_x were increased by factors of 1.5 and 2.5 at either side of 25th column in the domain to create a set of observations. These observations were then used to adjust the emissions. After 20 iterations the cost function was reduced by 95% and reasonable estimate of the emissions scaling factors were estimated (Figure 1), where the distinction between the scaling factors at the 25th column is evident. Note that areas with zero or negligible emissions produce no gradient and therefore remain unchanged (scaling factor of unity) in the optimization.

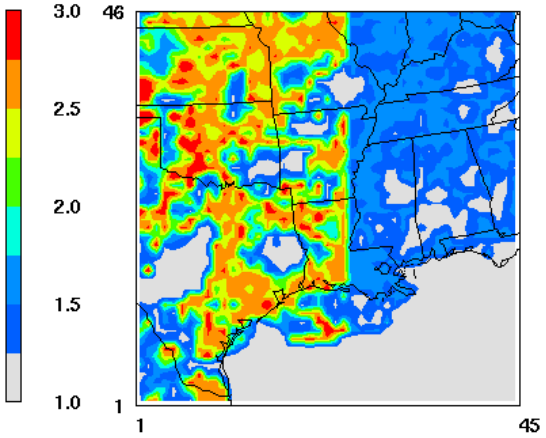

Fig. 1. Scaling factors for NO_x emissions after 20 iterations. The "true" answer is 1.5 to the west of column 25, and 2.5 to its east.

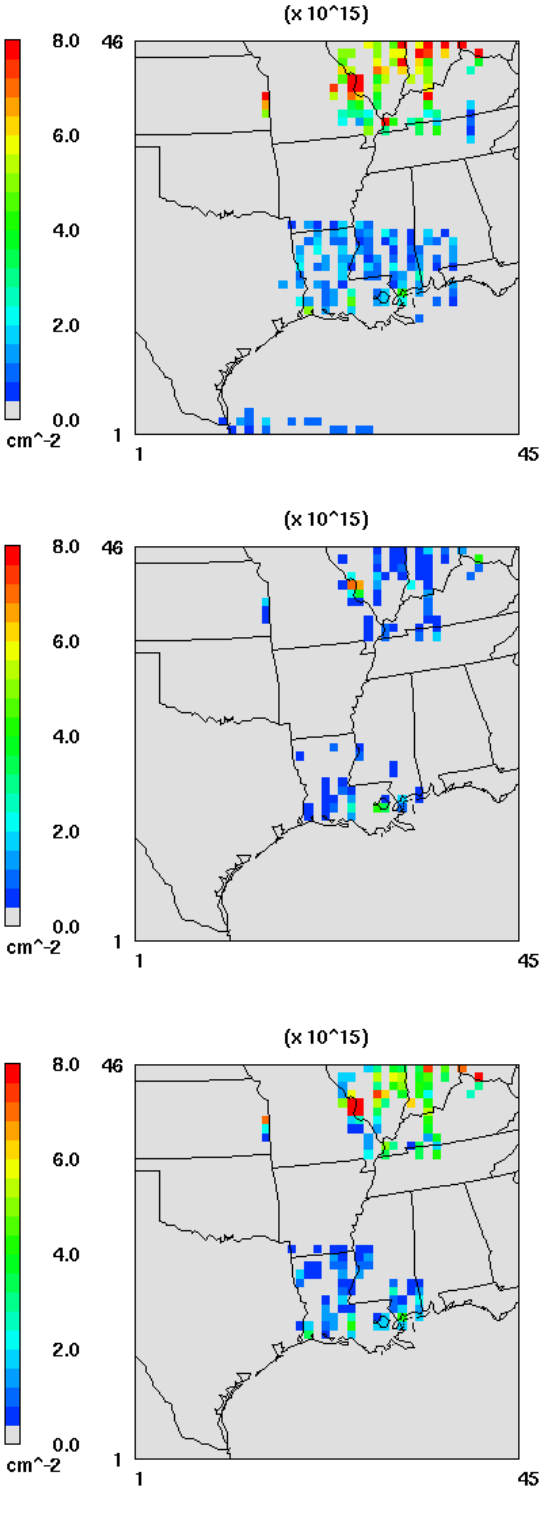

Fig. 2. Observed (top), before assimilation (middle), and after assimilation (bottom) NO_2 column densities.

Assimilation of NO_2 column densities are carried out for the domain shown in Figure 1, and for the retrievals on June 21, 2005. The domain has a horizontal resolution of 36 km, and consists of 23 vertical layers, with the tropopause typically occurring between the layers 20 and 21. We used a regularization parameter of 100.0 [equation (1)]. The results shown are for test cases with 17-hour simulations, but we will extend the assimilation window to 4 days in future. A long assimilation window is essential for emission inversion, as it allows for the regional transport to fully manifest itself. Extended assimilation windows may be less important for NO_2 which has a fairly short lifetime; however, longer assimilation periods will simply integrate more information into the model. Figure 2 shows the effect that assimilation has on the predicted NO_2 column densities. After about 30 iterations the cost function is reduced to half. The resulting scaling factors are shown in Figure 3. Note that our assimilation only adjusts the emissions, and as a result, errors and inaccuracies in other parameters/processes (e.g., boundary conditions) are incorrectly compensated by the adjustments in emissions. This may be true for the area of significant underestimations (large scaling factors) north of the domain, which may be more appropriately explained by underestimations in NO_x boundary conditions. In the continuation of our work, we will include other parameters such as boundary and initial conditions and dry deposition velocities in the assimilation process.

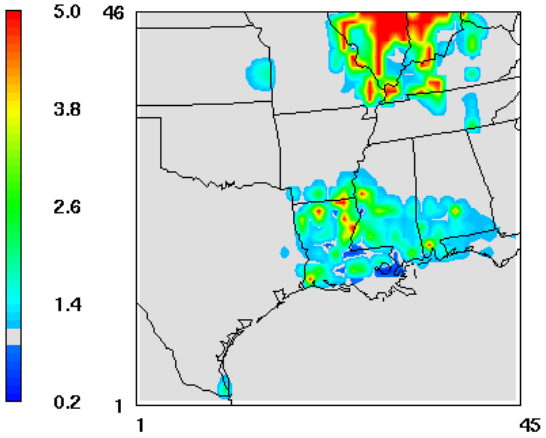

Fig. 3. NO_x emission scaling factors after 30 iterations

4. REFERENCES

Boersma, K.F. et al., 2004. Error analysis for tropospheric NO2 retrieval from space, *J. Geophys. Res.,* 109, doi:10.1029/2003JD003962.

Hakami, A. et al., 2005. Adjoint inverse modeling of black carbon during the Asian Pacific Regional Aerosol Characterization Experiment, *J. Geophys. Res.,* 110, D14301.

Platt, U., 1994. Differential optical absorption spectroscopy (DOAS), *Chem. Anal. Series,* 127, 27-83.