TOWARDS A ROBUST ESTIMATE OF THE GLOBAL LIGHTNING NITROGEN OXIDES SOURCE RATE AND ITS ERROR BOUND


(1) Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Physik der Atmosphäre, Oberpfaffenhofen, 82234 Wessling, Germany
(2) National Center for Atmospheric Research (NCAR), Boulder, Colorado, USA
(3) Max-Planck Institut (MPI) für Chemie, Mainz, Germany
(4) Royal Netherlands Meteorological Institute (KNMI), De Bilt, The Netherlands
(5) Central Aerological Observatory (CAO), Dogoprudny, Russia
(6) Instituto Nazionale di Ottica Applicata (INOA), Firenze, Italy

ABSTRACT/RESUME

A weighted least square fit method is introduced which provides best estimates for the annually mean global source strength $S$ of lightning induced nitrogen oxides (LNOx) together with random and systematic error bounds. The method is demonstrated using data measured in-situ with airborne instruments near deep convection over a tropical continent at altitudes up to 20 km in Southern Brazil in February/March 2004 and February 2005 during the EU-project “Tropical Convection, Cirrus, and Nitrogen Oxides Experiment” (TROCCINOX). Model results for air composition along the flight paths for at least two values of $S$ were provided by several global chemical transport models. LNOx estimates are obtained not only from NO mixing ratio data measured close to thunderstorms with fresh LNOx sources, but also from O3 and CO data because of their LNOx sensitivity due to tropospheric photochemistry during the chemical life-time of these species. The best estimate of the annual and global mean LNOx source value computed for the given data is $S = 4.8 \pm 2.5 \text{Tg a}^{-1}$ (in mass units of nitrogen per year). Ongoing model improvements and inclusion of data from further in-situ and remote sensing measurements (including SCIAMACHY NO2 columns) and data from other periods and latitudes may still change this result.

1 INTRODUCTION

The global LNOx source is the largest source of NOx (NOx = NO + NO2) in the upper troposphere, in particular in the tropics [1], and it is considered to be the least known source of the total NOx budget [2]. Present best estimates of the LNOx source rate, derived from theoretical, laboratory, and field studies, and literature reviews [3-6] range within at least 1 to 10 Tg a$^{-1}$, indicating an uncertainty range of more than $\pm 5 \text{Tg a}^{-1}$. None of the studies provides rigorous error bounds so far.

Lightning occurs most frequently in the tropics and mainly during the afternoon over continents [7]. A recent field experiment (TROCCINOX) for the first time provided data with strong LNOx signals over a continent in the tropics. Moreover, satellite data are now providing NO2 columns. In spite of the strong daily cycle of lighting activity with maximum activity during the afternoon, and large cloud effects affecting the measured radiances, these data, which are obtained at one fixed local time (10:00 for GOME [8] and 10:30 for SCIAMACHY [9]), can be used to provide LNOx estimates [10-12].

This paper describes a method to derive both a best estimate $S$ and an error bound $\Delta S$ of the source strength of LNOx from observed and modelled values of atmospheric properties (volumetric species concentrations and possibly other properties, including species column concentrations and lightning flash frequencies). The method used is basically known as a weighted least-square-fit or maximum likelihood solution, and is a special case of the so-called Bayesian least square method which provides the most probable solutions for the free parameters for unbiased data with Gaussian error statistics [13, 14]. Here, we generalize the approach for data which may contain biases.

2 METHOD

A best estimate for the source strength $S$ is determined minimizing the quadratic expression of the discrepancy variance

\[\sigma^2(S) = \sum \sum [c_i - m_i(S)] W_{ij} [c_j - m_j(S)].\]  

In practice, we use a normalised diagonal weight matrix \(W_{ij} = \frac{1}{N} \text{diag}(w^2_i)\), with \(N = \sum w^2_i\), so that

\[\sigma^2(S) = \sum w^2_i [c_i - m_i(S)]^2.\]

(2)

Here, \(c_i\) are the measured values; \(m_i\) are the model results of the same parameter sampled along the flight route; and \(w_i \geq 0\) are suitable weights, see below; the sum \(\sum\) is taken over all data \(i\); \(S\) is the open model parameter, i.e., the global and annually averaged LNOx source value in mass units of nitrogen per year (Tg a\(^{-1}\)). Minimization requires to find \(S\) such that \(\partial \sigma^2(S)/\partial S = 0\). After linearization of \(m(S)\) relative to a suitable first guess \(S_0\),

\[m_i(S) = m_i(S_0) + b_i(S_0) (S - S_0), \quad b_i(S_0) = \frac{\partial m_i(S_0)}{\partial S},\]

(3)

an improved solution \(S\) is found from

\[S = S_0 + \left(\sum (b_i(S_0) w_i)^2\right)^{-1} \sum b_i(S_0) w^2_i [c_i - m_i(S_0)].\]

(4)

The sensitivity vector \(b_i\), to which the “Jacobian” matrix degenerates in this univariate case with only one fit parameter, is computed from a discrete approximation

\[b_i(S_0) = \frac{m_i(S_i+1) - m_i(S_i)}{S_i+1 - S_i},\]

(5)

from at least two sets of model results \(m_i(S_j), j=1,2,...,M, (M \geq 2)\), for different \(S_j\) values in the neighbourhood of the expected result \(S\). In the linear case, with \(b_i(S_0)\) independent of \(S_0\), Eq. (4) provides the solution directly without iterations. Linearity is also implied by using only 2 model solutions to approximate \(b_i(S_0)\). In the non-linear case, with \(b_i(S_0)\) depending on \(S_0\), the procedure has to be applied iteratively replacing \(S_0\) by \(S\) for the next iteration.

The solution \(S\) has uncertainties \(\Delta S\) because of random and systematic errors (or uncertainties) in the model and observation data, which are estimated from \(\varepsilon_i = c_i - m_i(S_i)\) (with local estimates \(S_i\), see below). For random errors, the value \(S\) is also random. According to the error propagation law [15], the expectation \(E\) of the error is estimated from

\[\Delta S_{\text{ran}}^2 = E(\Delta S^2) \cong \sum g_i \varepsilon_i^2, \quad \text{with \ "gains" } g_i = \frac{\partial S}{\partial \varepsilon_i} = \frac{(\sum b_i w_i)^2}{\sum (b_i w_i)^2} b_i w^2_i,\]

(6)

From this we may compute a “random” error, applying for independent random errors:

\[\Delta S_{\text{ran}}^2 = \sum (g_i \varepsilon_i)^2\]

(7)

Because of Schwarz inequality, we have a systematic error bound which is satisfied always:

\[\Delta S_{\text{sys}} \leq \sum |g_i| \varepsilon_i|

(8)

The value of \(\Delta S_{\text{sys}}\) estimates the upper error bounds including random and systematic errors.

The method provides solutions \(S\) and uncertainty bounds \(\Delta S\) for arbitrarily selected weights \(w_i\), e.g. for constant \(w_i\). However, the optimal solution, i.e. the solution with smallest random and systematic errors \(\Delta S\), is obtained if the weights \(w_i\) are optimally selected. The optimal weights are those which minimize the errors.

For random correlated errors, \(\Delta S_{\text{ran}}\) is minimized by a nondiagonal, positive definite and symmetric weight matrix \(W_{ij}\) representing the inverse of the correlation matrix \(X_{ij}\) of errors, i.e. the expectation of correlations between errors \(\varepsilon_i\) and \(\varepsilon_j\) at individual data points. For random independent errors \(\varepsilon_i\), the correlation matrix is a diagonal matrix with diagonals equal to the variances \(\sigma^2\varepsilon_i\) of the uncertainties \(\varepsilon_i\). For the diagonal case the proof is obtained simply by looking for the minimum of \(\Delta S_{\text{ran}}^2\) with respect to variations of the \(w_i\). For the full matrix \(W_{ij}\) the proof is less simple ([16], p. 61; [17], appendix E). It is noteworthy that the optimal weights for random errors are independent of the sensitivity \(b_i\). The random error decreases with the square root of the number \(n\) of independent data ([15], p. 102). This can be seen easily in the special case of constant \(b_i w_i = b_i w\), for which \(g_i = \text{const}\) and \(\sum (b_i w_i)^2 = n (b w)^2\) with \(\sum\) summing over \(n\) data. In general, the number of statistically independent data is smaller than the number of data measured, e.g. because of
temporal correlation. One may reasonably assume that errors in data from different flights are statistically independent. Since we have data from about 90 flight-parameter combinations (32 flights, with up to 4 sets of values for various species), the random errors of the mean of $S$ obtained from all flights is about one order of magnitude smaller than the random errors of the individual results. Therefore, we take each parameter/flight combination as one independent data and approximate the weight matrix $W_0$ by a diagonal matrix $\text{diag}(w_j^2)$. Of course, the use of a diagonal matrix also simplifies the practical computations in this method.

In addition to random errors we have systematic measurement and model errors (such as calibration errors in the measurements; numerical approximation errors, errors in emissions data, or uncertainties in the physical approximations). In principle, the systematic errors get minimized by $w_j^2 = \delta_n$, where $j$ is the index of the most accurate data, with smallest systematic error. In other words, the result $S$ with smallest systematic error is obtained by using only one model and one set of flight data, namely those with smallest systematic errors.

Even for perfect measurements and perfect performance of the models, representativeness errors remain [18]. For example, the data along the flight path differ from the mean values within the model grid cell used in the models. The representativeness errors are considered to be included in the random error estimates. Moreover, we have unknown sampling errors because of the finite extent of the set of observations and model results. The data obtained from a limited time period and a limited region, with limited set of weather conditions, may not be representative for the whole globe and long term annual mean. For this reason, one seeks to include data obtained under a wide set of ambient conditions and not just form the best case with smallest systematic error.

The magnitude of the random and systematic errors $\Delta S$ depend on the individual errors $e_i$ in the data and sensitivities $g_i$ of the data relative to $S$. The “gain” value $g_i$ is the larger the larger $b_i$, i.e. the larger the contribution $S b_i = m_i(S) - m_i(0)$ of the LNOx source to changes in the model result $m(S)$. This is best seen from the relative error of $S$. From Eqs. (4, 6, and 8), and assuming linear model response and $S_0 = 0$, we obtain for the relative error in $S$,

$$\Delta S/S = \sum b_i w_i^2 \frac{c_i - m_i(S)}{c_i - m_i(0)} \approx \frac{\sum b_i w_i^2 (c_i - m_i(S))}{\sum b_i w_i^2 (c_i - m_i(S) + m_i(S) - m_i(0))} \approx \frac{\sum c_i - m_i(S)}{\sum c_i - m_i(S) + m_i(S) - m_i(0)}$$

The sensitivity $b_i$ enters the first equation directly in the product $b_i w_i^2$ and via the difference $m_i(S) - m_i(0) = b_i(S - 0)$. The products $b_i w_i^2$ occur in both the numerator and denominator of this expression and are unimportant for the relative error if of similar magnitude for all data points; therefore they are omitted in the last expression. More important for the relative error is the discrepancy between measured and modelled values for optimal $S$. $\sigma \sim |c_i - m_i(S)|$, which should be small for small errors, and the differences $|m_i(S) - m_i(0)|$, which should be large. In other words, we obtain best estimates with smallest uncertainties from the data with smallest discrepancy $\sigma$ between measured and modelled results for optimal $S$ and largest contribution $|b_i S|$ from the LNOx source to the model result $m_i(S)$.

When using several models, we use the “best” model which provides consistent results with smallest systematic error $\Delta S_{sys}$. The model is “consistent” if the individual $S_i$ values for all cases $i$ are positive, and if the frequency distribution of the $S_i$ values close to normal (i.e., unimodal, close to symmetric, with few “outliers” far from mean value).

In order to be able to check for consistency, and to minimise the systematic error by selecting the most accurate model and most sensitive measurements for the present purpose, we apply the method to subsets of the measured and modelled results. In principle, one could compute individual estimates

$$S_i = S_0 + (S_i - S_0) \frac{c_i - m_i(S_0)}{m_i(S_i) - m_i(S_0)}$$

for each individual pair of measured and modelled results. However, the individual data are strongly affected by random errors. Therefore, we apply the method to larger subsets of measured values $c_i = c_{f,p,t}$ and model results $m_i(S) = m_{f,p,t}(S)$. Here, the index $i$ corresponds to the triple index $f,p,t$ denoting flight $f$, parameter (or species) $p$, and time $t$, for which the data apply. The parameters $p=1,2,...,N_p$ cover the mixing ratios of the species measured (4 species in this application, see below). The flights have numbers $f=1,2,...,N_f$ ($N_f = 32$ in this application). The times $t$ cover flight times of order $10000$ s with temporal resolution $\Delta t = 1$ s. In order to reduce representativeness errors and to provide data with similar spatial resolution and hence similar variance in both the model and the measured data and to reduce temporal correlations, the data are smoothed temporally with a running mean over an averaging time interval ($\Delta t = 1200$ s)}
similar to the time step size of the model data. The weights $w_i = w_{f,p}$ are computed in two steps as follows: In step 1 we first use constant weights $w_i = \text{const}$. With these weights we apply the above procedure individually for each parameter set $i = f,p$ to compute the best fit of $S_i$ with related error measures $\Delta S_{\text{ran},i}$ and $\Delta S_{\text{sys},i}$, and compute a suitable error variance estimate, e.g. the variances of the data, $\sigma_{c,i}^2$. (For inhomogeneous relative measurement errors, the measurement error variances should be used instead of $\sigma_{c,i}^2$) In step 2, the results of step 1 are used to specify the final weights $w_i$:

$$w_i = 1/\left[\frac{\sigma_{c,i}^2 \Delta S_{\text{sys},i}}{\text{max}(0, S_i)}\right].$$  \hspace{1cm} (11)

The weighting with the error variances $\sigma_{c,i}^2$ minimizes the random error $\Delta S_{\text{ran}}$. The weighting with the maximum function gives zero weights to cases $i$ with negative $S_i$ values, and gives highest weight to those cases $i$ which exhibit the smallest systematic errors and hence reduces the final systematic error $\Delta S_{\text{sys}}$.

3 MEASUREMENT AND MODEL DATA

The measurement used were performed with the research aircraft Falcon of the Deutsches Zentrum für Luft- und Raumfahrt (DLR) and the M55 Geophysica of the company Myasishchev Design Bureau (MDB) during the EU-project TROCCINOX over a region including the State of São Paulo in Brazil (between 10° - 28°S, 38° - 55°W), at altitudes up to 12.5 km (Falcon) or 20 km (Geophysica), in the time periods February 14 - March 10, 2004 (Falcon only), and February 1 – 19, 2005 (both aircraft). Mixing ratio data are available among others for CO, O₃, NO, and NOₓ from the Falcon (provided by H. Schlager et al., DLR) and from the Geophysica (NO and NOₓ from H. Schlager et al., DLR; O₃ from A. Ulanovsky, CAO, and F. Ravegnani, CNR; and CO from S. Viciani and P. Mazzinghi et al., INOA), for 12 Falcon flights in 2004, 12 Falcon flights in 2005, and 8 Geophysica flights in 2005, with 1 second nominal time resolution. The flights include local flights in the vicinity of thunderstorm events (< 500 km radius from the local airport), partially in the thundercloud anvils, or survey flights along more or less straight paths over larger distances (< 1500 km radius). Most measurements were taken in the afternoon. The measurements cover tropical and subtropical air masses, which we classify mainly on the base of equivalent potential temperature. Some of the flights are impacted by fresh outflow from deep convective events; others represent non-convective conditions.

The model results provided by various teams in support of TROCCINOX include results from the model ECHAM5/MESSY (provided by Max-Planck-Institute for Chemistry in a preliminary (E5M0) and a first version (E5M1); simulations performed by C. Kurz, DLR), MOZART 4 (MZ4, L. Emmons, NCAR), TM4 (E. Meijer, KNMI), and MATCH-MPIC (L. Labrador and M. Lawrence, MPI for Chemistry, Mainz). The MATCH results and details of all models will be given in a future paper. The models are global three-dimensional atmospheric chemistry transport models (CTM) or climate-chemistry models (CCM) simulating atmospheric photochemistry driven offline by numerical weather analysis data. ECHAM5/MESSY computes its own meteorology following ECMWF divergence, vorticity, temperature and surface pressure analysis data at time scales of 48, 6, 24, and 24 h, respectively [19]. The model fields are initiated at least three months before the observation period starts using available analysis for similar conditions, so that the integrations proceed longer than the relevant chemical life-times (a few days for NO, days to weeks for O₃, and weeks to about a month for tropical CO), and so that the results are insensitive to the chemical initial conditions. Each of the models uses its own parameterisation to compute the amount of LNOₓ released by lightning. The horizontal distribution is computed as a function of either cloud-top-height (CTH [20] for MATCH and MZ4), mean updraft velocity (UPD [21] in E5M0 and E5M1), or convective precipitation (CPR [22] in TM4). Vertically the LNOₓ emissions are released between surface and local cloud top using C-shaped profiles [19]. The total globally and annually averaged amount of LNOₓ emission is a free parameter in these models and set to a prescribed value (mostly 2 and 5 Tg a⁻¹). Its actual value may differ by an order 10 % from the given nominal value because of interannual fluctuations, and has to be reevaluated a posteriori from the simulation results for at least one year. From the model output mixing ratio data of NO, O₃, CO, and NOₓ, are derived as a function of flight time (with 1 or 10 s temporal resolution) by interpolating in time and space to the flight paths of the aircraft.

4 RESULTS

From the measured and modelled data, we compute best estimates $S_i$ and (error) variances $\sigma_{c,i}$, and from these the weights $w_i$ as explained in section 2, and then compute the best fit $S$ from all data for various period/species combinations together, separately for each model. From comparisons between measured and modelled data from individual flights (not shown in this paper because of limitation to 6-pages), we find that the classical CTH-based LNOₓ parameterisation gives too little variability and does not fit the high NO concentrations measured close to tropical
thunderstorms. As a consequence, results computed with CTH exhibit large relative errors $\Delta S_{sys}/S$. The UPD and CPR variants fit better. The largest sensitivity $[m(S) - m(0)]/m(S)$ of the model results $m(S)$ to LNOX sources is found for NO, followed by NO$_y$, O$_3$, and CO (positive for O$_3$: LNOx causes enhanced O$_3$ production; negative for CO: LNOx increases OH and hence CO destruction, in particular in the warm lower tropical troposphere). The sensitivity is a maximum at mid tropospheric altitudes between about 500 and 200 hPa in the tropics. Hence, we obtain more accurate results, with smaller $\Delta S_{sys}/S$, from the Falcon data than from either the Geophysica measurements in the tropopause region or from any measurements in the boundary layer. The most accurate results are obtained from NO data along long-distance flights in air masses with LNOX influence from the last hours or day, outside individual thunderstorms anvils. Here, the models resolve well the spatial variability of the mixing ratios. We do not average over the various model results, which would be appropriate if all include random errors only, but select the result form the “best” model which provides consistent results with smallest systematic errors. The individual $S_i$ results are mostly positive and close to normally distributed for E5M0, but E5M1 often computes negative $S_i$ values, for yet to be understood reasons. Also the results from TM4 and partially MZ4 deviate strongly from a normal distribution, in particular when including NO$_y$ and CO.

The following Table presents the results for the best-fit $S$ with the systematic error bounds $\Delta S$ (the random error bounds are smaller) separately for various flights, i.e. for the Falcon flights in 2004 (F04), Falcon in 2005 (F05), Geophysica in 2005 (G05), and from all flights during TROCCINOX together (F04+F05+G05), and from either only one species (NO) or from several species together (NO and NO$_y$, or NO and O$_3$, or NO and CO, or “all” species).

<table>
<thead>
<tr>
<th>Flights</th>
<th>Species</th>
<th>TM4</th>
<th>E5M1</th>
<th>MZ4</th>
<th>E5M0</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$S$ Tg a$^{-1}$</td>
<td>$\Delta S$</td>
<td>$S$</td>
<td>$\Delta S$</td>
<td>$S$</td>
</tr>
<tr>
<td>F05</td>
<td>NO</td>
<td>3.6</td>
<td>2.0</td>
<td>4.0</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td>NO+NO$_y$</td>
<td>2.3</td>
<td>1.8</td>
<td>4.3</td>
<td>2.7</td>
</tr>
<tr>
<td></td>
<td>NO+O$_3$</td>
<td>3.2</td>
<td>1.2</td>
<td>6.3</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td>NO+CO</td>
<td>11.4</td>
<td>2.5</td>
<td>6.2</td>
<td>1.0</td>
</tr>
<tr>
<td>F05+G05</td>
<td>NO</td>
<td>3.7</td>
<td>2.1</td>
<td>3.5</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td>NO+NO$_y$</td>
<td>3.9</td>
<td>3.6</td>
<td>3.3</td>
<td>2.9</td>
</tr>
<tr>
<td></td>
<td>NO+O$_3$</td>
<td>3.2</td>
<td>1.2</td>
<td>6.5</td>
<td>4.3</td>
</tr>
<tr>
<td></td>
<td>NO+CO</td>
<td>10.1</td>
<td>3.9</td>
<td>6.2</td>
<td>1.0</td>
</tr>
<tr>
<td>F04+F05+G05</td>
<td>NO</td>
<td>4.1</td>
<td>2.7</td>
<td>3.5</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td>NO+NO$_y$</td>
<td>4.3</td>
<td>3.7</td>
<td>3.3</td>
<td>2.9</td>
</tr>
<tr>
<td></td>
<td>NO+O$_3$</td>
<td>4.8</td>
<td>2.5</td>
<td>6.5</td>
<td>4.3</td>
</tr>
<tr>
<td></td>
<td>NO+CO</td>
<td>10.3</td>
<td>4.2</td>
<td>6.2</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>all-CO</td>
<td>4.8</td>
<td>2.5</td>
<td>6.5</td>
<td>4.3</td>
</tr>
</tbody>
</table>

The error estimates obtained for NO data with the models TM4, E5M1 and E5M0 are less than 2.7 Tg a$^{-1}$ and smaller than error estimates of previous studies. Hence, this study provides progress compared to the state of the art. Inclusion of data from O$_3$ and partly CO in addition to NO reduces the error bounds at least for some sets of flight and model data. Moreover, by use of O$_3$ and CO, the representativity of the results is extended over essentially the whole tropics and subtropics because of the integrating effect of LNOx on O$_3$ and CO along the backward trajectory over the chemical life times of these species. The smallest error bounds are obtained from the whole set of data from E5M0 (1.8 Tg a$^{-1}$) followed by TM4 (2.5 Tg a$^{-1}$). The scatter of $S$ values is smallest for E5M0 and E5M1: 3.9 - 6.6 Tg a$^{-1}$ for E5M0, and 3.3 - 6.5 Tg a$^{-1}$ from E5M1. For the other models, the scatter is larger (2.3 - 11.4 Tg a$^{-1}$ for TM4, and 3.4 - 17.1 Tg a$^{-1}$ from MZ4). Hence, some models fit individual data sets well but do not provide the most consistent result, possibly due to systematic errors in the emission data for CO and other O$_3$ precursors. E5M0 is a preliminary model versions and E5M1 is still to be improved. Therefore, we consider the result $4.8 \pm 2.5$ Tg a$^{-1}$ from TM4 as the presently best available result. It overlaps with most of the more recent results (a review paper is under preparation).

5 CONCLUSIONS

The weighted least square fit method introduced in this paper provides objective estimates of the global LNOx source rate $S$ and a upper limit for systematic errors, $\Delta S_{sys}$. The present paper describes the method. The LNOx value obtained so far from the TROCCINOX data and the given set of model runs are preliminary. The analysis shows that $S$ can be determined not only from NO (and NO$_y$) but also from O$_3$ and CO data. In order to come to the best practically
achievable result, we plan to include many more of the measurements form various regions of the world since the mid 1980s (both from in situ and remote sensing), including NO$_2$ column data from SCIAMACHY, GOME and OMI.

The present method may also be applied to determine other or more than one parameters of Earth system models. For example, by comparing results from one model with two LNOx parameterisations with the data, one may decide which one gives more accurate LNOx estimates. Also, we may run models with two set of aviation emissions and compare with observational data at mid-latitudes. Such a study would reveal the consistency of state-of-the-art photochemical models in representing aviation effects. If the models simulate an aviation NO$_x$ impact on NO$_x$, O$_3$, and CO concentrations consistent with the data and among each other, then this provides some validation for the suitability of the models for aviation related assessments. If however, the consistency is weak, this would indicate that important processes are still to be included into the models. By careful analysis of the discrepancy and further parameter studies one should be able to detect "missing chemistry" such as heterogeneous chemical processes. Hence, this approach opens a whole new venue for research and assessments.

Acknowledgments
Funding from the EU, and contributions from all TROCCINOX partners, in particular the Brazilian partners, are gratefully acknowledged (http://www.pa.op.dlr.de/troccinox/). We thank the model teams for their work, and in particular Mark Lawrence for providing the flight interpolation algorithm and valuable suggestions. The model runs were driven with meteorological data from the European Center for Medium Range Weather Forecast (ECMWF), or the National Center for Environmental Prediction (NCEP), USA.

REFERENCES