Evaluation of model-simulated source contributions to tropospheric ozone with aircraft observations in the factor-projected space

C. Shim1,*, Y. Wang1, and Y. Yoshida2,**

1Department of Earth and Atmospheric Sciences, Georgia Institute of Technology, 311 Ferst Drive, Atlanta, GA 30332, USA
2Goddard Earth Science & Technology Center, University of Maryland, Baltimore County, Baltimore, MD 21228, USA
* now at: Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive Pasadena, CA 91109, USA
** now at: Atmospheric Chemistry and Dynamics Branch, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

Received: 4 October 2007 – Published in Atmos. Chem. Phys. Discuss.: 2 November 2007
Revised: 7 February 2008 – Accepted: 29 February 2008 – Published: 26 March 2008

Abstract. Trace gas measurements of TOPSE and TRACE-P experiments and corresponding global GEOS-Chem model simulations are analyzed with the Positive Matrix Factorization (PMF) method for model evaluation purposes. Specially, we evaluate the model simulated contributions to O3 variability from stratospheric transport, intercontinental transport, and production from urban/industry and biomass burning/biogenic sources. We select a suite of relatively long-lived tracers, including 7 chemicals (O3, NOy, PAN, CO, C3H8, CH3Cl, and 7Be) and 1 dynamic tracer (potential temperature). The largest discrepancy is found in the stratospheric contribution to 7Be. The model underestimates this contribution by a factor of 2–3, corresponding well to a reduction of 7Be source by the same magnitude in the default setup of the standard GEOS-Chem model. In contrast, we find that the simulated O3 contributions from stratospheric transport are in reasonable agreement with those derived from the measurements. However, the springtime increasing trend over North America derived from the measurements are largely underestimated in the model, indicating that the magnitude of simulated stratospheric O3 source is reasonable but the temporal distribution needs improvement. The simulated O3 contributions from long-range transport and production from urban/industry and biomass burning/biogenic emissions are also in reasonable agreement with those derived from the measurements, although significant discrepancies are found for some regions.

1 Introduction

Tropospheric O3 has important environmental consequences. Photolysis of O3 and the subsequent reaction of O(1D) with water vapor (H2O) in troposphere produces the hydroxyl radical (OH), which is the most important oxidant in troposphere. This tropospheric oxidation by OH determines the lifetime of major greenhouse gases such as methane (CH4). The sources of tropospheric O3 include photochemical production within the troposphere and transport from the stratosphere. Many studies have investigated the main sources to tropospheric O3. Springtime O3 increase is attributed to photochemical production (e.g., Penkett and Brice, 1986 and Liu et al., 1987). A number of studies using 3-D chemical transport models have focused on the effect of intercontinental transport on tropospheric O3 concentrations from Asia to North America (e.g., Berntsen et al., 1999; Jaffe et al., 1999; Jacob et al., 1999; Jacob et al., 1999; Bey et al., 2001). The effect of trans-Pacific transport is particularly noticeable in the spring (e.g., Jacob et al., 1999; Mauzerall et al., 2000; Wild and Akimoto, 2001; Tanimoto et al., 2002; Wang et al., 1998, 2006). On the other hand, the studies based on the observed correlations between O3 and 7Be attributed this trend to transport of stratospheric O3 (e.g., Oltmans and Levy, 1992; Dibb et al., 1994).

The observed relationships between tropospheric O3 and CO provide additional diagnosis of O3 sources (e.g., Fishman and Seiler, 1983; Chameides et al., 1987; Parrish et al., 1993). Furthermore, those relationships between simulated CO and O3 offer a reasonable way to evaluate model simulations of O3 (e.g., Chin et al., 1994). However, direct attribution of O3 sources based on atmospheric measurements is
difficult since tropospheric O$_3$ is a secondary product from primary emissions of trace gases. One method we used previously to diagnose sources of tropospheric trace gases is the positive matrix factorization (PMF) method, an advanced multi-variant factor analysis (Wang et al., 2003b; Shim et al., 2007). PMF analysis of the measurements obtained during the Tropospheric Ozone Production about the Spring Equinox (TOPSE) experiment found that the increasing seasonal trend of springtime O$_3$ at northern mid and high latitudes is attributed more to tropospheric O$_3$ production and transport, even though O$_3$ transport from the stratosphere is the largest contributor to O$_3$ variability (Wang et al., 2003b).

The factor attributions based on atmospheric measurements provide additional constraints on model simulated tropospheric O$_3$ sources beyond those provided by direct comparisons of simulated and observed trace gas concentrations. In this work, we apply the PMF method to the simulation results of a global 3-D chemical transport model (GEOS-Chem). As such, we can compare the PMF results to model simulations and evaluate the performance of GEOS-Chem on the basis of aircraft measurements. A main issue is how model simulated factor contributions to O$_3$ variability compare with those based on the measurements. Unlike direct comparisons between observed and simulated trace gases, measurements and corresponding model results are first projected with PMF onto the factor space before model evaluation. In the factor space, a suite of chemicals can be evaluated simultaneously in a consistent and objective manner, which is difficult to achieve using direct comparisons between the measurements and model results. Aircraft measurements from two aircraft field campaigns, TOPSE and TRAnsport of Chemical Evolution over the Pacific (TRACE-P, March–April 2001) experiments are used. We describe data selections from TOPSE and TRACE-P and GEOS-Chem simulations in Sect. 2.1. The PMF method is explained in Sect. 2.2. Evaluation of model results in the projected factor space is discussed in Sect. 3. Conclusions are given in Sect. 4.

2 Methodology

2.1 Measurements and GEOS-Chem simulations

Figure 1 shows the measurement regions during TOPSE and TRACE-P. The TOPSE experiment (February–May 2000) was conducted to investigate the photochemical transition during spring at northern mid and high latitudes (Atlas et al., 2003). The TRACE-P experiment (March–April 2001) was conducted to investigate the Asian outflow to the Pacific (Jacob et al., 2003). Both experiments took place during spring when significant transport of O$_3$ from the stratosphere is expected (e.g., Wang et al., 1998b and references therein).

In this study, we analyze relatively long-lived chemical tracers including O$_3$, total reactive nitrogen (NO$_x$), peroxycetyl nitrate (PAN), CO, C$_3$H$_8$, CH$_3$Cl, and Beryllium-7 ($^7$Be) and one dynamic tracer (potential temperature). Those tracers other than O$_3$ generally have specific primary source characteristics. NO$_x$ is a good tracer for air masses influenced by tropospheric NO$_x$ emissions (mostly from fossil fuel combustion, biomass burning, and soils) or transport from the stratosphere. PAN is produced only in the troposphere during oxidation of $>$C$_2$ hydrocarbons and its lifetime increases rapidly with increasing altitude. Therefore, it is a good tracer for photochemically aged air masses in the free troposphere. CO is for combustion influence primarily from fossil fuel, biofuel, and biomass and C$_3$H$_8$ is a good liquefied petroleum gas (LPG) tracer. CH$_3$Cl has its major sources from terrestrial biosphere and biomass burning (Yoshida et al., 2004, 2006). $^7$Be is produced mainly by cosmic rays in the stratosphere and upper troposphere and is generally used as a tracer for stratospheric air mass (Dibb et al., 2003). Potential temperature is a useful dynamic tracer since it is conserved in adiabatic processes. The analytical approach for the observed species is similar to the work by Wang et al. (2003b), but the number of chemicals used is smaller because only measured species that are also simulated by GEOS-Chem are selected. The resulting discrepancies with
the previous work by Wang et al. (2003b) will be discussed in Sect. 3.

Photochemical and dynamical environments vary dramatically with latitude. We separate the analysis regions to low, mid, and high latitudes. The TOPSE measurement data set is over mid latitudes (30–45° E) and high latitudes (60–85° N, 61–94° W). We consider only coincident measurements, which are mostly limited by availability of 7Be measurements (144 coincident data points of all selected tracers for mid latitudes and 200 data points for high latitudes). Concentrations of O3 in the selected subset have a similar probability distribution as the whole dataset, and the derived spring O3 trends (TOPSE) of the subset are similar to the whole dataset. We exclude missing data because including large amounts of missing data (by assigning a large uncertainty to these data) would lead to a large underweight of the 7Be measurements and a loss of Be and O3 correlation signal (Wang et al., 2003b). The 7Be and O3 correlation is critical for analyzing the effect of stratospheric transport. The selected data have a bias towards high altitudes of 5–8 km (~70% of the data); therefore the evaluation results are more relevant for the middle and upper troposphere. The TRACE-P measurements data set is over mid latitudes (30–45° N, 125–240° E, 65 data points) and low latitudes (15–30° N, 120–205° E, 78 data points). The selected data also have a bias towards 7–12 km (40–50% of the data) due to the availability of 7Be measurements.

GEOS-Chem is a global 3-D chemical transport model driven by assimilated meteorological data from the Global Modeling Assimilation Office (GMAO) (Schubert et al., 1993). The 3-D meteorological fields are updated every six hours, and the surface fields and mixing depths are updated every three hours. We use version 7.24 with a horizontal resolution of 2° x 2.5° and 30 vertical layers (GEOS-3 meteorological fields were used). GEOS-Chem includes a comprehensive tropospheric O3-NOx-VOC chemical mechanism (Bey et al., 2001), which includes the oxidation mechanisms of 6 VOCs (ethane, propane, lumped >C3 alkanes, lumped >C2 alkenes, isoprene, and terpenes). Climatological monthly mean biomass burning emissions are from Duncan et al. (2003). The fossil fuel emissions are from the Global Emission Inventory Activity (GEIA) for other chemical compounds (Benkovitz et al., 1996; Olivier et al., 2001). The cross-tropopause O3 transport from the stratosphere is simulated in the model using a passive ozone-like tracer (Synoz) (Mclinden et al., 2000). The annual net flux is ~475 Tg of O3. For standard simulations, the model was first spun up for one year. The GEOS-Chem simulations for the selected five tracers and one dynamic tracer (O3, NOx, PAN, CO, C3H8, and potential temperature) are sampled at the same time and locations as the aircraft measurements.

Simulated total reactive nitrogen (NOx) is estimated by the sum of simulated NOX, HNO3 (nitric acid), HNO4 (pernitric acid), PAN, and N2O5 (dinitrogen pentoxide).

We follow Liu et al. (2001 and 2004) in 7Be simulations. The 7Be source in GEOS-Chem is taken from the study by Lal and Peters (1967) as a function of altitude and latitude and ~70% of 7Be is emitted in the stratosphere. The seasonal and longitudinal dependence of 7Be productions is very small and not considered. The major sink of atmospheric 7Be is by wet deposition; the model considers scavenging in convective updrafts as well as first-order rainout and washout from both convective and large-scale precipitation (Liu et al., 2001). Liu et al. (2004) reduced the stratospheric 7Be source by a factor of 3. This simulation was first spun up for one year as well.

For model simulated CH3Cl, we used the GEOS-Chem results by Yoshida et al. (2004). Contributions from the six sources (pseudo-biogenic, oceanic, biomass burning, incineration/industrial, salt marsh and wetland) are considered. The model results are evaluated extensively with surface and aircraft measurements; the model simulations are usually in good agreement with measurements in the northern hemisphere.

In order to investigate the stratospheric O3 contributions in the model, we conducted tagged O3 simulations to track the fractions of O3 transported from the stratosphere (Liu et al., 2002). Photochemistry is considered in the simulations by taking archived O3 production and loss rates from the GEOS-Chem standard simulations on a daily basis. In this manner, when projecting simulated O3 variability in the factor space using PMF, we can examine the fractional contribution from the stratosphere as compared to tropospheric production (Sect. 3) in each factor.

2.2 PMF applications

The PMF method (Paatero and Tapper, 1994) explores factor categorization through the covariant structures of observed or simulated chemical and dynamical parameters (e.g., Paatero, 1997; Wang et al., 2003b; Liu et al., 2005). PMF generates only positive factor contributions, which enables a better physical interpretation of the results. In contrast, conventional the principal component analysis method lumps positively and negatively correlated tracers together. The data matrix X of m measurements by n tracers are decomposed in PMF analysis for p factors as

\[
X = GF + E
\]

Or

\[
x_{ij} = \sum_{k=1}^{p} g_{ik} f_{kj} + e_{ij}
\]

where the m by p matrix G is the mass contributions of kth factor to ith sample (factor score), the p by n matrix F is the gravimetric average contributions of kth factor to jth chemical species (factor loadings), and the m by n matrix E is the
error. In the PMF model, the solution is a weighted least squares fit, where the data uncertainties are used for determining the weights of the residuals in the error matrix. We also use the explained variation (EV),

$$EV_{kj} = \frac{\sum_{i=1}^{m} |g_{ik}f_{kj}|}{\sum_{i=1}^{m} \left( \sum_{k=1}^{K} |g_{ik}f_{kj}| + |e_{ij}| \right)}$$

(3)

to define the relative contributions of each factor to chemical species since the mixing ratios of different compounds are not directly comparable.

During PMF analysis, it is important to choose the number of factors that provide physically meaningful results. In this analysis, the order factor is determined by sorting the center-of-mass locations of the G or F matrix in ascending order. By evaluating the error matrix E, we define the range of mathematically acceptable number of factors (Paatero et al., 2002). We then inspect the factor profiles to choose the number of factors that gives the best physically meaningful results. In general, we pick as small a number of factors as possible to reduce the potential of overinterpreting of the dataset. Rotation is further used to improve factor separation (Paatero et al., 2002). However, the results presented in this work are insensitive to rotation.

As in the work by Wang et al. (2003b), the values of tracers are linearly scaled to a nondimensional range of 0–1 and assigned uniformly small uncertainty for the dataset. The scaling is applied because the chemical and dynamical tracers have very different scales that affect the least square fitting in PMF. By scaling and assigning a uniform uncertainty, we assure that all tracers are weighted equally in the PMF analysis. In the analysis, we selected only coincident measurements of the selected tracers and corresponding model results. Missing measurement data are not used in order to reduce the uncertainty in the analysis. Following the procedure described above, PMF resolved 5 factors for TOPSE and 4 factors for TRACE-P in both observed and simulated datasets.

PMF was often used for source apportionments of surface aerosols (e.g., Lee et al., 1999). For that purpose, it is often necessary to assume that the composition of the air mass from a specific source does not change during transport. That assumption is unnecessary in this analysis since we evaluate how the simulated contributions to tropospheric ozone from different processes compare to the contributions derived from observations. Obviously the chemical characteristics of air masses are affected by transport. Our previous analyses (Wang et al., 2003b; Shim et al., 2007) indicate that although some collocated sources are mixed during transport, clear air mass separation based on the covariance of chemical and dynamical tracers can be obtained. We did not find evidence that transport and mixing “create” chemically distinct air mass.

3 Results and discussion

3.1 TOPSE

As mentioned in Sect. 2.1, the TOPSE results are biased toward the middle and upper troposphere. In order to capture the correlation between the stratospheric O3 and 7Be using PMF, we have included the data points that have O3 concentrations >100 ppbv (5% of the data set), which are generally associated with the lower stratospheric air. When analyzing the PMF results, however, we only use data points with O3<100 ppbv to minimize the effect of these lower stratospheric data (Wang et al., 2003b). The simulated O3 mixing ratios do not exceed 100 ppbv. We analyze the datasets for mid (40–60°) and high (60–85°) latitudes separately.

3.1.1 TOPSE at mid latitudes

PMF derived EV profiles from the observed and simulated datasets for TOPSE mid latitudes are shown in Fig. 2. Each factor is named after the tracers that show the largest variability (7Be, θ, CH3Cl, NOy/PAN, and hydrocarbons). The figure shows reasonably consistent factor profiles between observed and simulated datasets (black and yellow/red bars). Direct comparison of the EV profiles between the observed and simulated datasets can be misleading when the simulated variability differ significantly from the observations. For the simulated datasets, we therefore also show the scaled EV

Fig. 2. Explained variation (%), defined in Sect. 2.2) profiles of the observed (shown in black) and simulated datasets (shown in yellow and red) at mid latitudes (40–60° N) during TOPSE. Also shown is the scaled EV profiles (in blue and green) for the simulated datasets for direct comparison with the measurements (Eq. 4, see text for details). The two-color bars for O3 show the model simulated stratospheric (upper bar, red or green) and tropospheric (lower bar, yellow or blue) fractions. The results are for measurement data with O3 concentrations <100 ppbv (and the corresponding model dataset).
Table 1. The factor scores correlations ($r$) with latitude, altitude, and C$_2$H$_6$/C$_3$H$_8$ for TOPSE mid latitudes.

<table>
<thead>
<tr>
<th>Factors</th>
<th>Latitude</th>
<th>Altitude</th>
<th>C$_2$H$_6$/C$_3$H$_8$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Obs</td>
<td>Mod</td>
<td>Obs</td>
</tr>
<tr>
<td>$^7$Be</td>
<td>-0.26</td>
<td>-0.37</td>
<td>0.42</td>
</tr>
<tr>
<td>$\theta$</td>
<td>-0.59</td>
<td>-0.51</td>
<td>0.65</td>
</tr>
<tr>
<td>CH$_3$Cl</td>
<td>-0.34</td>
<td>-0.2</td>
<td>0.38</td>
</tr>
<tr>
<td>NO$_y$/PAN</td>
<td>0.32</td>
<td>0.32</td>
<td>-0.12</td>
</tr>
<tr>
<td>HC$^1$</td>
<td>0.52</td>
<td>0.31</td>
<td>-0.65</td>
</tr>
</tbody>
</table>

Extreme factor scores (outside 2σ range) and the measurements that have O$_3$ greater than 100 ppbv are excluded. HC denotes the hydrocarbon factor.

The profiles (blue/green bars in Fig. 2–5) by the following equation,

$$
EV_{kj}^{\text{scaled}} = \frac{\sum_{j=1}^{m} |g_{kj}|_{\text{model}}}{\sum_{j=1}^{m} \sum_{i=1}^{p} |g_{ij}|_{\text{model}} + |e_{ij}|_{\text{measurements}}}
$$

The $^7$Be factor in Fig. 2 shows the largest variability of $^7$Be for both observations and model (60% and 237 femtocurie per square centimeter (fCi/SCM) for the observations, 85% and 188 fCi/SCM for the model), indicating the stratospheric origin of the air masses. The stratospheric O$_3$ fraction from tagged O$_3$ simulation in this factor show ~75% of the stratospheric origin, and a small fraction of the tropospheric $^7$Be origin (~25%), due likely to upper tropospheric $^7$Be production, which is also evident in the small fractions of PAN in this factor. The $^7$Be factor is associated with the largest O$_3$ variability at mid latitudes (40% and 20 ppbv for the observations, 58% and 14 ppbv for the model). A notable underestimation in the simulated $^7$Be mean concentration is found (435 and 234 fCi/SCM for the observations and model, respectively). Including measured O$_3$ data >100 ppbv in the $^7$Be factor profile do not alter the results except a small decrease of the scaled EV of O$_3$. The same result is found for other regions, we therefore only show the profiles with measured O$_3$$\leq$100 ppbv.

We examine the factor correlations with latitudes, altitude, and C$_2$H$_6$/C$_3$H$_8$ ratio in order to further investigate the factor characteristics (Table 1). The higher C$_2$H$_6$/C$_3$H$_8$ ratio reflects photochemically aged air masses (Wang and Zeng, 2004). The positive correlations of the $^7$Be factor with altitude ($r=0.42$ and 0.62 for the observed and simulated datasets, respectively) are expected for a factor dominated by transport from the stratosphere. The weak negative correlations with latitude indicate that stratosphere-troposphere exchange is likely more active at lower latitudes in 40–60° N region (Table 1).

The potential temperature ($\theta$) factor has large variability of $\theta$ (Fig. 2). It explains 14% of observed O$_3$ variability (3.6 ppbv) and 12.8% of simulated O$_3$ variability (3.1 ppbv). The negative factor correlations with latitude and CO ($r=-0.64$ and -0.52, respectively, not shown in the table), and positive correlations with altitudes and C$_2$H$_6$/C$_3$H$_8$ ratio (Table 1) imply that this factor is likely associated with intercontinental long-range transport of O$_3$ from lower latitudes, which is consistent with the result by Wang et al. (2003b).

The CH$_3$Cl factor is characterized by large signals of CH$_3$Cl and no O$_3$ variability is explained by this factor (Fig. 2). This factor contains significant CO variability, which can imply the biomass burning influence. However, the very small factor correlations with C$_2$H$_6$/C$_3$H$_8$ ratio and negative correlation with latitude (Table 1) may support the large biogenic CH$_3$Cl emissions from the tropics (e.g., Yoshida et al., 2004, 2006) rather than biomass burning.

The NO$_y$/PAN factor has large signals of NO$_y$ and PAN (Fig. 2) and this factor is the second important factor for tropospheric O$_3$ variability at mid latitudes (25.6% and 6.6 ppbv for the observations, 18.4% and 4.4 ppbv for the model). The positive correlation with latitude and much weaker correlations with altitude and C$_2$H$_6$/C$_3$H$_8$ ratio (Table 1) imply that the factor represents long-range transport of reactive nitrogen.

The hydrocarbon factor is characterized by large variability of CO and C$_3$H$_8$ (Fig. 2). There is no contribution to tropospheric O$_3$ variability. The positive correlation with latitude and clear negative correlations with altitudes and C$_2$H$_6$/C$_3$H$_8$ ratio reflect the influence of relatively fresh emissions from the surface (Table 1).

3.1.2 TOPSE at high latitudes

Five factors are identified for high latitudes ($^7$Be, $\theta$, CH$_3$Cl, NO$_y$/PAN, and hydrocarbons; Fig. 3). As mid latitudes, there is also significant difference in $^7$Be variability...
between observations and simulation (415 fCi/SCM and 206 fCi/SCM, respectively) in the $^{7}$Be factor, reflecting serious underestimation of $^{7}$Be by GEOS-Chem. Liu et al. (2001) artificially scaled down the stratospheric $^{7}$Be source by a factor of $\sim$3 in order to adjust for some surface measurements of $^{7}$Be. However, the simulated $^{7}$Be mean concentrations and the variability accounted for in the $^{7}$Be factor show consistent underestimations by about a factor of 2 in TOPSE and TRACE-P (to be shown) datasets. It implies that the factor of 3 reduction in the stratospheric $^{7}$Be source in the standard GEOS-Chem model is too large.

The $^{7}$Be factor shows comparable $O_3$ variabilities between observations and simulation (34% and 13.7 ppbv for the observations, 45% and 12.4 ppbv for the model). The stratospheric $O_3$ fraction from the tagged $O_3$ simulation suggests that $\sim$70% is of the stratospheric origin (Fig. 3). The tropospheric fraction for $O_3$ is $\sim$30% in this factor. The positive correlations with altitude and negative correlations with CO support its stratospheric origin (Table 2).

The potential temperature ($\theta$) factor shows large variabilities of $O_3$ levels as much as that of the $^{7}$Be factor (39% and 15.4 ppbv for the observations, 41% and 10.7 ppbv for the model), which is different from mid latitudes. The positive correlation with altitude and $C_2H_6/C_3H_8$ ratio (Table 2) indicates that this factor is likely associated with transport of reactive-nitrogen poor air masses from lower latitudes. Tagged $O_3$ simulation shows that $O_3$ variability accounted for in this factor is produced in the troposphere (Fig. 3).

The CH$_3$Cl factor is characterized by large signals of CH$_3$Cl (78% and 38.6 pptv for the observations, 98% and 20.3 pptv for the model), but its contribution to $O_3$ variability is insignificant. This factor is positively correlated with altitude, consistent with long-range transport of high CH$_3$Cl air masses from lower latitudes since there are no significant sources of CH$_3$Cl at high latitudes.

The NO$_y$/PAN factor has large signals of NO$_x$ and PAN (Fig. 3). This factor also has clear chemical signals of CO but not $^{7}$Be, implying that the air masses are influenced by industrial/fossil fuel emissions at high latitudes (Table 2). This factor, however, contributes to less than 5% of $O_3$ variability, reflecting the largely inactive photochemical environment at high latitudes in spring (Wang et al., 2003a).

### Table 2. Same as Table 1, but for TOPSE high latitudes.

<table>
<thead>
<tr>
<th>Factors</th>
<th>Latitude</th>
<th>Altitude</th>
<th>$C_2H_6/C_3H_8$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^{7}$Be</td>
<td>-0.08</td>
<td>-0.08</td>
<td>0.48</td>
</tr>
<tr>
<td>$\theta$</td>
<td>-0.1</td>
<td>-0.2</td>
<td>0.51</td>
</tr>
<tr>
<td>CH$_3$Cl</td>
<td>-0.07</td>
<td>-0.32</td>
<td>0.38</td>
</tr>
<tr>
<td>NO$_y$/PAN</td>
<td>0.16</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td>HC$^1$</td>
<td>-0.08</td>
<td>0.04</td>
<td>-0.43</td>
</tr>
</tbody>
</table>

### Table 3. Factor contributions to $O_3$ seasonal increase (ppbv/month) for TOPSE.

<table>
<thead>
<tr>
<th>Factors</th>
<th>Mid latitudes</th>
<th>High latitudes</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^{7}$Be</td>
<td>2.7</td>
<td>1.2</td>
</tr>
<tr>
<td>$\theta$</td>
<td>0.3</td>
<td>0.4</td>
</tr>
<tr>
<td>CH$_3$Cl</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NO$_y$/PAN</td>
<td>3.5</td>
<td>1.3</td>
</tr>
<tr>
<td>HC$^1$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>6.5</td>
<td>3.0</td>
</tr>
</tbody>
</table>

Only the measurements $O_3<100$ ppbv are analyzed.

The hydrocarbon factor is characterized by a large variability of CO and C$_3$H$_8$ (Fig. 3). It does not contribute to tropospheric $O_3$ variability. Just as mid latitudes, the negative correlations with altitude and $C_2H_6/C_3H_8$ ratio reflect air masses affected by relatively fresh emissions (Table 2).

#### 3.1.3 Springtime $O_3$ trends at northern mid and high latitudes

Understanding the contributions to the seasonal $O_3$ trend is another important purpose of this study. As stated in Sect. 2.1, this study analyzed only eight tracers due to the limited availability of simulated tracers, while the previous study (Wang et al., 2003b, hereafter referred to as the previous study) included fourteen tracers with seven factors. At mid latitudes, the seasonal increase of all factors of measurements is 6.48 ppbv/month (Table 3), consistent with the previous study (6.3 ppbv/month). The largest contributor to the $O_3$ seasonal trend is the NO$_x$/PAN factor (3.55 ppbv/month, Table 3) followed by the $^{7}$Be factor (2.66 ppbv/month). That is also consistent with the previous study (3.5 ppbv/month, and 2.5 ppbv/month, respectively). In contrast, the simulated overall seasonal increase is only 3.01 ppbv/month, indicating a large underestimation. The increase from the NO$_x$/PAN factor is underestimated (1.32 ppbv/month in the model), and the $^{7}$Be factor increase is also much smaller than that of observation (1.29 ppbv/month in the model).

At high latitudes, the overall springtime increase from the measurements is 4.29 ppbv/month (Table 3), comparable with the previous study (4.6 ppbv/month). In comparison, the simulated increase is only 1.3 ppbv/month, indicating a significant underestimation. The most contributions to the seasonal increases at high latitudes are from $^{7}$Be, $\theta$, and NO$_x$/PAN factors (1.78, 1.16, and 1.10 ppbv/month, respectively) in the measurement dataset. In comparison, the corresponding trends in the model are much lower (0.76, 0.77, and 0.11 ppbv/month). The underestimation is particularly large for the NO$_x$/PAN factor, implying that simulated $O_3$ production in reactive-nitrogen rich air masses does not increase as

much as in the observations. The negative O\textsubscript{3} trend in the hydrocarbon factor in the simulation but not in the measurements is a likely reflection of the problematic simulations of its major components (C\textsubscript{2}H\textsubscript{6} and C\textsubscript{3}H\textsubscript{8}) in May by GEOS-Chem (Wang and Zeng, 2004).

The NO\textsubscript{y}/hydrocarbon factor trend is consistent with the previous study. However, the contributions of \textsuperscript{7}Be and \theta factors are different from those of the previous study (0.8 ppbv/month and 0.6 ppbv/month, respectively). The previous study had additional tracers resulting in the CH\textsubscript{4}-halocarbon factor. It accounts for transport from lower latitudes, which contributes to the largest increase of O\textsubscript{3} at 1.7 ppbv/month at high latitudes. In this study, that large increase trend is apportioned into the \textsuperscript{7}Be and \theta factors since we do not have CH\textsubscript{4} and halocarbon (other than CH\textsubscript{3}Cl) simulations in GEOS-Chem. Because the PMF factor projections are for the same number of tracers, model results can still be evaluated in this analysis. The CH\textsubscript{4}-halocarbon factor contribution to O\textsubscript{3} variability is, however, <10\% (3 ppbv) at high latitudes in the previous study; thus the effect of the missing factor on factor apportioned O\textsubscript{3} variability is fairly insignificant in this study.

During TOPSE, the major contributions to the seasonal O\textsubscript{3} increase in springtime is from intercontinental transport of polluted air masses, while the major contributions to O\textsubscript{3} variability is from the stratospheric influences and long-range transport of O\textsubscript{3} from lower latitudes. While the model generally captures the factor contributions to O\textsubscript{3}, factor contributions to the springtime increasing trend of O\textsubscript{3} in the measurements are severely underestimated. These model underestimations are also consistent with the results by Wang et al. (2006). Improvements in the seasonal transitions of stratosphere and intercontinental transport are needed in the model.

### 3.2 TRACE-P

The TRACE-P experiment was conducted to investigate the effects of Asian outflow to the Pacific during spring (Jacob et al., 2003). As mentioned in Sect. 2.1, the TRACE-P results are biased toward the middle and upper troposphere (more than 40\% of the data is above 7 km). Compared to TOPSE analysis, there are fewer coincident measurements limited mostly by the availability of \textsuperscript{7}Be measurements (65 and 79 for mid and low latitudes, respectively). We analyze the datasets for low (15–30\°) and mid (30–45\°) latitudes separately.

#### 3.2.1 TRACE-P at mid latitudes

Four factors are identified for mid latitudes (\textsuperscript{7}Be, \theta, CH\textsubscript{3}Cl, and NO\textsubscript{y}/hydrocarbons, Fig. 4). The \textsuperscript{7}Be factor shows larger O\textsubscript{3} variability in the observations than model results (68\% and 20.8 ppbv for the observations, 48.5\% and 13.8 ppbv for the model). There is also a large underestimation in simulated \textsuperscript{7}Be variability (428 fCi/SCM and 211 fCi/SCM, respectively) for the reason discussed in Sect. 3.1.2. The tagged O\textsubscript{3} simulation shows that \sim80\% of O\textsubscript{3} variability in this factor is of the stratospheric origin (Fig. 4). While this factor in the simulated dataset showed a positive correlation with altitude (r=0.63), it has a much weaker correlation (r=0.12) in the measurements (Table 4). One possible reason for the large difference is that transport from the stratosphere occurs too close to TRACE-P regions in the model, resulting in stronger correlations. If the stratosphere-troposphere exchange occurs in regions farther away, further downward transport or mixing with low-altitude polluted air would reduce the gradients in altitude.

The potential temperature (\theta) factor shows large signals of \theta (Fig. 4). Positive correlations with altitudes and C\textsubscript{2}H\textsubscript{6}/C\textsubscript{3}H\textsubscript{8} ratio in Table 4 characterize this factor as long-range transport of air masses from the tropics. While this factor accounts for 25.6\% of O\textsubscript{3} variability in the simulated datasets, it has no contribution in the measurement dataset.

The CH\textsubscript{3}Cl factor is characterized by the large signals of CH\textsubscript{3}Cl and the contributions of this factor to O\textsubscript{3} variability are small in measured and simulated datasets (Fig. 4). The significant contributions to CO (Fig. 4) as well as negative correlations with altitude and C\textsubscript{2}H\textsubscript{6}/C\textsubscript{3}H\textsubscript{8} ratio (Table 4), and the reactive nitrogen signals suggest a strong influence from biomass burning. This factor contributes to NO\textsubscript{y} variability (168 pptv) only in the simulated dataset, and PAN variability (164 pptv) only in the observed dataset. Since PAN is an

---

**Table 4.** Same as Table 1, but for TRACE-P mid latitudes.

<table>
<thead>
<tr>
<th>Factors</th>
<th>Latitude</th>
<th>Altitude</th>
<th>C\textsubscript{2}H\textsubscript{6}/C\textsubscript{3}H\textsubscript{8}</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textsuperscript{7}Be</td>
<td>-0.37</td>
<td>0.05</td>
<td>0.12</td>
</tr>
<tr>
<td>\theta</td>
<td>-0.33</td>
<td>-0.67</td>
<td>0.79</td>
</tr>
<tr>
<td>CH\textsubscript{3}Cl</td>
<td>-0.02</td>
<td>0.11</td>
<td>-0.25</td>
</tr>
<tr>
<td>NO\textsubscript{y}/HC</td>
<td>0.18</td>
<td>-0.42</td>
<td>-0.77</td>
</tr>
</tbody>
</table>

---

www.atmos-chem-phys.net/8/1751/2008/
important component of NO\textsubscript{y}, the signal in PAN will propagate to become a signal of NO\textsubscript{y}. However, another large component of NO\textsubscript{y} is HNO\textsubscript{3}, which can be removed rapidly by wet deposition in the atmosphere. It appears to suggest that the scavenging of HNO\textsubscript{3} (a major component of NO\textsubscript{y}) during transport and the production of PAN from biomass burning NO\textsubscript{y} are underestimated by the model. The stronger negative correlation of the factor with altitude (Table 4) suggests that the altitude of biomass burning transport is lower in the model. This model bias also leads to a higher negative correlation with \textit{C}_2\text{H}_6/\text{C}_3\text{H}_8 as is found here because mixing with locally emitted \textit{C}_2\text{H}_6 and \textit{C}_3\text{H}_8 tends to destroy the negative correlation (Table 4).

In TRACE-P analysis, the NO\textsubscript{y}/PAN and hydrocarbon factors in TOPSE are combined (now NO\textsubscript{y}/hydrocarbon factor) because the separation of those factors leads to incomparable factor profiles between the measurements and model results. The NO\textsubscript{y}/hydrocarbon factor is characterized by a large variability of NO\textsubscript{y}, PAN, CO, and \textit{C}_3\text{H}_8 (Fig. 4). This factor shows a contribution to tropospheric O\textsubscript{3} variability only in the simulation (14.5\%). This factor has negative correlations with altitudes and \textit{C}_2\text{H}_6/\text{C}_3\text{H}_8 ratio, likely reflecting relatively the influence of fresh industrial/fossil fuel emissions over Asia (Table 4). The stronger negative correlations with \textit{C}_2\text{H}_6/\text{C}_3\text{H}_8 ratio and altitude in the measurements than the simulations imply that mixing is too fast at low altitudes in the model.

3.2.2 TRACE-P at low latitudes

Four factors also are identified for low latitudes (\textsuperscript{7}Be, \(\theta\), CH\textsubscript{3}Cl, and NO\textsubscript{y}/hydrocarbons, Fig. 5). The \textsuperscript{7}Be factor shows smaller O\textsubscript{3} variability in the measurements than the model simulation (17.4\% and 7.1 ppbv for the observations, and 30.8 \% and 9.9 ppbv for the model). Large underestimation by a factor of 3 is found in simulated \textsuperscript{7}Be variability (Fig. 5). There are no data with O\textsubscript{3} above 100 ppbv in both observations and simulation results at low latitudes. The stratospheric O\textsubscript{3} fraction from the tagged O\textsubscript{3} simulation shows that \(\sim 50\%\) is due to transport from the stratosphere (Fig. 5), which is the smallest stratospheric influence among the datasets. The positive factor correlations with altitude reflect in part the contribution from the stratosphere (Table 5). The weaker correlations with altitude and \textit{C}_2\text{H}_6/\textit{C}_3\text{H}_8 ratio in the observed than simulated datasets likely reflect either a problem in the transport locations from the stratosphere or the mixing between stratospheric and tropospheric air masses in the model (Table 5).

The potential temperature (\(\theta\)) factor shows large signals of \(\theta\) (Fig. 5). While the \textsuperscript{7}Be factor is the largest contributor to simulated O\textsubscript{3} variability at low latitudes, the \(\theta\) factor is the largest contributor to observed O\textsubscript{3} variability (27.4\% and 11.1 ppbv for the observations, and 21\% and 6.7 ppbv for the model). The model estimates a small stratospheric fraction of 15\% in this factor. This factor contains small signals of simulated NO\textsubscript{y}, PAN, and CO, which are absent in the observed dataset, indicating again that mixing of different air masses in the model is overestimated. The correlation coefficients are more consistent between observed and simulated datasets for this factor. The positive correlations with altitude and \textit{C}_2\text{H}_6/\textit{C}_3\text{H}_8 ratio (Table 5) and negative correlations with CO \((r=-0.66\) and \(-0.49\), not shown in the table) suggest the dominance of photochemically aged upper tropospheric air in this factor.

The CH\textsubscript{3}Cl factor is characterized by large signals of CH\textsubscript{3}Cl and a significant contribution to O\textsubscript{3} variability is found in this factor (24.6\% and 9.9 ppbv for the observations, 15.7\% and 5 ppbv for the model, Fig. 5). This factor contributes more to O\textsubscript{3} in the observations than the model. The larger contribution in the observations is associated with CO, NO\textsubscript{y}, and PAN. In comparison, this factor in the simulated dataset has a smaller contribution from CO (22\% and 17.4 ppbv) and negligible contributions from NO\textsubscript{y} and PAN. The observed profile is consistent with the characteristics of biomass burning.

It appears that the contributions to PAN, NO\textsubscript{y}, and CO from biomass burning are attributed to the NO\textsubscript{y}/hydrocarbon factor in the model. Comparing the profiles between CH\textsubscript{3}Cl and NO\textsubscript{y}/hydrocarbon factors, a major separation factor between these factors is the correlation between \textit{C}_3\text{H}_8 and CH\textsubscript{3}Cl. In both datasets, almost all the \textit{C}_3\text{H}_8 signals are in the NO\textsubscript{y}/hydrocarbon factor. There is no correlation between

### Table 5. Same as Table 1, but for TRACE-P low latitudes.

<table>
<thead>
<tr>
<th>Factors</th>
<th>Latitude</th>
<th>Altitude</th>
<th>(\textit{C}_2\text{H}_6/\textit{C}_3\text{H}_8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textsuperscript{7}Be</td>
<td>0.33</td>
<td>-0.06</td>
<td>0.39</td>
</tr>
<tr>
<td>(\theta)</td>
<td>-0.05</td>
<td>0</td>
<td>0.98</td>
</tr>
<tr>
<td>CH\textsubscript{3}Cl</td>
<td>0.22</td>
<td>0.4</td>
<td>0.03</td>
</tr>
<tr>
<td>NO\textsubscript{y}/HC</td>
<td>0.25</td>
<td>0.35</td>
<td>-0.65</td>
</tr>
</tbody>
</table>
C$_3$H$_8$ and CH$_3$Cl in the observed dataset. Consequently there is no CH$_3$Cl signal in the observed NO$_x$/hydrocarbon factor. The opposite is true in the simulated dataset, leading to a significant contribution to the CH$_3$Cl variability (29% and 17 pptv). The inadequate separation of C$_3$H$_8$ and CH$_3$Cl in the model may result from two sources. The first is that noise is overestimated in the model, which results in excessive mixing of biomass burning and industrial/urban air masses. The second is that the locations of biomass burning or industrial/urban sources are misplaced in the model, which also leads to unrealistic mixing.

The NO$_x$/hydrocarbon factor is characterized by large variabilities of NO$_y$, PAN, CO, and C$_3$H$_8$ (Fig. 5). The factor contributions to these trace gases are lower in the observations than the model results because some of the enhancements in the model are due in part to biomass burning emissions. Interestingly, the factor contributions to tropospheric O$_3$ variability are comparable in the observations (16% and 6.5 ppbv) and model results (18.5% and 6 ppbv) even though the enhancements in NO$_y$, PAN, and CO are higher in the model results. The two datasets have comparable positive factor correlations with latitude, and large negative correlations with altitude and C$_2$H$_6$/C$_3$H$_8$ ratio indicating fresh pollution plumes from East Asia (Table 5).

4 Discussion and conclusions

Trace gas measurements of TOPSE and TRACE-P experiments and corresponding GEOS-Chem simulations are analyzed with the PMF method. The factor attributions based on the projections in the factor space allow a direct evaluation of model performance in simulating source contributions to tropospheric O$_3$ variability and its springtime increase (during TOPSE). We select a suite of relatively long-lived variables, which are available both in observations and GEOS-Chem model; seven chemicals (O$_3$, NO$_y$, PAN, CO, C$_3$H$_8$, CH$_3$Cl, and $^7$Be) and one dynamic tracer (potential temperature). The evaluation has a bias towards a high altitude of 5–8 km (~70% of the data) for TOPSE and 7–12 km (~50% of the data) for TRACE-P, due to the availability of $^7$Be measurements.

In general, the factor loadings between the observations and simulations are in better agreement during the TOPSE experiment than TRACE-P. The former experiment took place in remote regions. Therefore, the model results are not as sensitive to source locations as for the latter experiment. There are also slightly more data points (determined largely by the availability of $^7$Be measurements) in the former experiment. We summarize the factor contributions to O$_3$ variability in Figs. 6 and 7.

The $^7$Be factor is found in all regions. Among all the factors, the largest discrepancy is found in the variability of $^7$Be, which is controlled largely by its source in the stratosphere. The simulated results are a factor of 2–3 lower than those observed. The large underestimation is due to the default reduction of the stratospheric $^7$Be source by a factor of ~3. Inadvertently, the default reduction provides a test for the PMF analysis.

Tagged O$_3$ simulations in the model indicate that the O$_3$ signal in the $^7$Be factor is controlled largely (70–80%) by transport from the stratosphere at mid and high latitudes. Only over the lower latitude does the stratospheric contribution drop to ~50%. The $^7$Be factor explains 34–40% of O$_3$ variability in the measurement dataset during TOPSE, in agreement with the simulated dataset. During TRACE-P, this factor contributes 68% and 17% at mid and low latitudes, respectively in the measurement dataset. In comparison, the contributions in the simulated datasets are also higher at mid latitudes (49%) and lower at low latitudes (31%). In general, we find that the decrease of stratospheric O$_3$ contributions (and the increase of tropospheric O$_3$ contributions) from mid to low latitudes during TRACE-P are much larger in the measured than simulated datasets. One potential reason is that...
mixing is overestimated between mid and low latitudes in the model, reducing the gradients between the two latitude bands.

Another common factor is the $\theta$ factor. There are consistent positive correlations of this factor with altitude and $C_2H_6/C_3H_8$ ratio, indicating long-range transport in the upper troposphere. The contribution of this factor to reactive nitrogen is small, reflecting likely chemical processing during transport. The large contribution to $O_3$ variability at high latitudes during TOPSE ($\sim$40%) in the measurement dataset is in agreement with the simulated dataset. In comparison, its contributions to mid latitudes during TOPSE are much lower in both datasets. During TRACE-P, there is no contribution from this factor to $O_3$ variability in the measurement dataset at mid latitudes. However, 26% contribution is found in the simulated dataset. A similar situation is found for the $NO_3/\text{hydrocarbon}$ factor. Excessive mixing between mid and low latitudes could explain some of the discrepancy. Further, the unresolved portion of $O_3$ variability is $\sim$30% in this case, much higher than the range of 11–19% in the other cases. Some of the unresolved portion is due to $O_3$ production in the troposphere.

A third common factor found is the $\text{CH}_3\text{Cl}$ factor. The contributions of this factor to $O_3$ are usually small. The exception is at low latitudes during TRACE-P, when biomass burning contributes to both $\text{CH}_3\text{Cl}$ and $O_3$. Some of the biomass burning contribution in the simulated datasets is attributed to the $NO_3/\text{hydrocarbon}$ factor since simulated $C_3H_8$ is correlated with $\text{CH}_3\text{Cl}$. The latter correlation was not found in the measurement dataset. Thus, we combine the $\text{CH}_3\text{Cl}$ and $NO_3/\text{hydrocarbon}$ factor contributions to $O_3$ variability; it is somewhat higher in the measurement dataset (41%) than the simulated dataset (34%). As discussed previously, the difference can be reduced if mixing is reduced between mid and low latitudes in the model.

During TOPSE, the $NO_3/\text{PAN}$ factor is resolved separately from the hydrocarbon factor. The latter made no contribution to $O_3$ variability. The $NO_3/\text{PAN}$ factor contributions are much higher at mid latitudes (18–26%) than high latitudes (<5%) in measured and simulated datasets, reflecting more active photochemistry at mid latitudes in spring.

Since the TOPSE experiment lasted longer than TRACE-P, we compared the factor contributions to the seasonal trend of $O_3$ in the observed and simulated datasets. Despite reasonably good agreements in the averaged contributions, the trends of factor contributions are quite different. The observed springtime $O_3$ increase is higher than simulated by a factor 2 at mid latitudes (6.5 vs. 3 ppbv/month) and a factor of 3 at high latitudes (4.3 vs. 1.3 ppbv/month). The increasing trend from the stratospheric contribution (the $^7\text{Be}$ factor) is underestimated by a factor of 2. The increasing trend from the tropospheric contribution is simulated well for the $\theta$ factor. However, the increasing trend from $O_3$ production by reactive nitrogen (the $NO_3/\text{PAN}$ factor) is underestimated by a factor of $>3$ (3.5 ppbv/month vs. 1.3 ppbv/month at mid latitudes and 1 ppbv/month vs. 0.1 ppbv/month at high latitudes). These results suggest that more attention needs to be placed on improving the simulations of the temporal trends of trace gases in chemical transport models.

Acknowledgement. This work was supported by the National Science Foundation Atmospheric Chemistry Program. The GEOSCHEM model is managed at Harvard University with support from the NASA Atmospheric Chemistry Modeling and Analysis Program.

Edited by: R. von Glasow

References


