Are the TRACE-P measurements representative of the western Pacific during March 2001?

Juno Hsu,1 Michael J. Prather,1 Oliver Wild,2 Jostein K. Sundet,3 Ivar S. A. Isaksen,3 Edward V. Browell,4 Melody A. Avery,4 and Glen W. Sachse4

Received 23 July 2003; revised 29 October 2003; accepted 25 November 2003; published 31 January 2004.

[1] Observations of CO and O3 from the Transport and Chemical Evolution over the Pacific (TRACE-P) campaign are compared with modeled distributions from the FRSGC/UCI CTM driven by the Oslo T63L40 ECMWF forecast meteorology. The model-measurement comparison is made within the context of how well the TRACE-P observations represent the springtime chemistry and ozone distributions over eastern Asia and the western Pacific in March 2001 and uses the four-dimensional (4-D) extended domain from the model to provide unbiased statistics. A key question is whether the limited sampling density or mission strategy led to a statistically biased sample. To address this question, we examine a diverse range of statistical analyses of the observations of CO and O3. The middle percentiles of the cumulative probability functions for CO in the free troposphere are representative (and reproduced by the CTM), but those in the boundary layer are not. The frequency of low-CO, stratospheric influence is well matched along flight tracks but is atypical of the extended domain. The percentiles of the latitude-by-height distribution of lidar O3 show how the CTM reproduces the nonrepresentative clumpy nature of the observations but has too low a tropopause about the jet region (30–35N). Adaptive kernel estimation of the 2-D probability density of O3-CO correlations shows a very good simulation of two different chemical regimes (stratospheric and polluted) that is quite different from the extended domain but also highlights the failure to predict CO > 400 ppb. Empirical orthogonal function analysis of the O3 vertical profiles shows how six EOFs can effectively describe the 4-D structures of O3 over this entire domain. The latitude-by-longitude maps of the principal components provide an excellent test of the CTM simulation along flight tracks and clearly show the unique sampling of O3 events by the TRACE-P flights. In many cases the ability of the model to simulate the nonrepresentative observations implies a clear skill in matching the unique meteorological and chemical features of the region.

INDEX TERMS: 0322 Atmospheric Composition and Structure: Constituent sources and sinks; 0365 Atmospheric Composition and Structure: Troposphere—composition and chemistry; 0368 Atmospheric Composition and Structure: Troposphere—constituent transport and chemistry; KEYWORDS: TRACE-P measurements, CTM stimulations, representativeness, sampling bias, ozone and CO2 comparisons, EOF, PDFs


1. Introduction

[2] In trying to understand atmospheric composition on a global scale, instantaneous measurements of trace gases made at specific sites are often assumed to be representative of a larger region over a longer period. For example, the NOAA CMDL flask network [Dlugokencky et al., 1994; Conway et al., 1994] has provided the global mean latitudinal gradient as well as trends on the total atmospheric burdens of CH4 and CO2, based solely on biweekly sampling at a number of remote sites. The assumption that these limited, surface measurements can be used to integrate the total atmospheric burden is necessary but far-reaching. In other cases, many single-location, surface, or sonde measurement programs have argued that their data are representative of the entire boundary layer [e.g., Haszpra, 1999; Inoue and Masueda, 2001] or the free troposphere for a greater region [e.g., Navasques and Rus, 1991; Gallardo et al., 2000]. Compared with these examples, recent airborne regional campaigns have provided an extremely dense,
latitudinal, longitudinal, and vertical sampling over several weeks (e.g., TRACE-P, PEM, NARE, CEPEX, MINOS). These data sets provide extensive statistics on the atmospheric chemistry of the region and thus are often assumed to be fully representative of the location and period. Nevertheless, even such high-density campaign data are greatly undersampled compared with atmospheric variability, and, further, the design of campaigns to study specific processes may result in statistically biased sampling of the region. For example, the NASA Transport and Chemical Evolution over the Pacific (TRACE-P) measurement campaign [Jacob et al., 2003] had the primary goal of studying the export of pollution from eastern Asia. Understanding the representativeness of these campaign data would greatly strengthen their use in global studies.

In addition, understanding of the representativeness of a given set of observations can help evaluate the accuracy of model simulations of those observations. Matching the observed statistics when the observations are representative of the sampling region is one measure of skill; however, if the observations are a statistically biased sampling, then accurate model simulation implies a clear skill in matching the unique meteorological and chemical features of the region. In the case when there is representative sampling from a campaign, the discrepancy between the flight-track observations and the model simulation cannot be dismissed as meteorological error and is more likely due to a fundamental or systematic error in the model such as emission levels, chemistry, or large-scale spatial gradients. In the case of statistically biased sampling, the mismatch between observations and model can perhaps be attributed to errors in the modeled meteorological fields, such as the height of convective outflows or the timing of frontal passing. Here we evaluate the accuracy of our global chemistry-transport model (CTM) simulations of the TRACE-P observations of CO and O_3 within the context of how these measurements are representative of the western Pacific during March 2001.

For TRACE-P the community developed several, independent, high-resolution four-dimensional (4-D) CTM simulations for that region and period that do a commendable job on matching many of the measurements [e.g., Kiley et al., 2003; Wild et al., 2003; Carmichael et al., 2003; Pierce et al., 2003; Liu et al., 2003; C. Mari et al., The effect of clean warm conveyor belts on the export of pollution from East Asia, submitted to Journal of Geophysical Research, 2004, hereinafter referred to as Mari et al., submitted manuscript, 2004]. Here we use the FRSGC/UCI/Oslo chemistry-transport model (T63L40 resolution with EC forecast fields) to generate a densely sampled 4-D data set for each chemical species for the TRACE-P region. The accuracy of our CTM simulations is determined by comparing a wide range of statistical features of O_3 (in situ and lidar) and CO (in situ) from (1) the observations and (2) the CTM simulations along flight tracks. The representativeness of the TRACE-P sampling is determined by parallel comparisons between (2) the CTM flight-track data and (3) the CTM simulations over the extended 4-D domain (i.e., the eastern Asia-western Pacific region for the month of frequent flight measurements).

A first approach in comparing such TRACE-P species measurements with model simulations is to plot the two overlapping time series for each flight. From parallel measurement-model plots of several species one can visually identify the temporal and spatial scales of variability and also the correlation of different species. Another, more quantitative method plots the modeled versus measured abundances as a scatter plot, yielding a measure of the accuracy of the simulation through a linear regression (for CO from TRACE-P, see Figure 1 in the work of Kiley et al. [2003]). Here we resort to a new range of statistical methods that allow us to compare not only the measurements with the model at the specific measurement locations but also the model sampled along flight tracks versus an extended 4-D domain.

Section 2 describes the FRSGC/UCI version of the CTM and our simulations of the extended TRACE-P domain. In addition, two statistical techniques used in this study are described: the adaptive kernel estimation for construction of 2-D Probability Density Functions (PDFs) to characterize the O_3-CO correlations and Empirical Orthogonal Functions (EOFs) for analysis of vertical structures in the lidar O_3 profiles. Cumulative probability distributions for in situ CO, in situ O_3, and lidar O_3 are examined in section 3. Latitude-height sections of the O_3 abundance from the lidar sampling are shown for the 10th, 50th, and 90th percentiles of both observation and model in section 4. The same CTM statistics are also presented for the extended TRACE-P domain rather than just the flight tracks. In section 5 the O_3-CO correlations observed along the flight tracks are compared with the model for individual flights. These data are combined into a single, two-dimensional probability distribution for all flights to compare with CTM simulation of both the flight tracks and the extended TRACE-P domain. In section 6, EOFs are used to describe the vertical features of the O_3 distribution and to show where these features are prevalent off the coast of Asia. Conclusions regarding the accuracy of the CTM simulation of TRACE-P observations in relation to the representativeness of the observing strategy are given in section 7.

2. Methodology

2.1. Chemical Transport Model

The chemical fields used in this study are generated by the Frontier Research System for Global Change (FRSGC) version of the University of California, Irvine (UCI) global chemical transport model (CTM), described by Wild and Prather [2000]. The model is run at T63 resolution (1.9° × 1.9°), with 37 eta-levels in the vertical and is driven by 3-hour averaged meteorological fields for Spring 2001 generated with the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecast System (IFS). The configuration of the model and the meteorological data used for the TRACE-P period are described by Wild et al. [2003]. The model simulates the tropospheric production of ozone during the TRACE-P period reasonably well, with notable exceptions in the polluted boundary layer, where production is overestimated, and in the upper free troposphere, where it may be underestimated [Wild et al., 2003]. In addition, a linearized chemistry is used for stratospheric ozone [McLinden et al., 2000], and in combination with the
IFS meteorological fields this is shown to reproduce the stratospheric intrusions observed over the western Pacific during TRACE-P [Wild et al., 2003]. The simulation of CO shows a small negative model bias [Kiley et al., 2003], but the variation of CO with latitude, altitude, and time are captured reasonably well.

We define an extended TRACE-P 4-D domain covering most of the measurement flights during the campaign over the east Asia-western Pacific region: Each CTM grid cell from 14N to 46N and 100E to 150E up to 18 km height from 3 March to 3 April 2001. Photochemical results from the CTM are saved hourly over this domain to make a 4-D data set. Figure 1 shows DC-8 and P-3B flight tracks in both plan view and vertical cross section. Only sampling points that fall in the extended 4-D domain (14N–46N, 100–150E) are shown. CTM simulations of flight-track observations of CO and O$_3$ are interpolated from the 4-D data to the exact location of the DC-8 and P-3B aircraft at 1-min intervals along their flight tracks. Similarly, CTM simulations of lidar data are constructed to compare with the DIAL O$_3$ lidar measurements made on the DC-8 aircraft [Browell et al., 2003]. The reported 1-min O$_3$ lidar observations have data missing occasionally due to cloud obscuration or instrumental downtime. We derive simulated-lidar profiles from the 4-D CTM data that miss the same data points. We also calculate a second set simulated O$_3$ vertical profiles along each DC-8 flight track assuming perfect profiling from the surface to 18 km.

2.2. Adaptive Kernel Estimation Method

Adaptive kernel estimation method has been used widely in climate research to construct PDFs and locate the maxima identifying climate regimes [Kimoto and Ghil, 1993; Corti et al., 1999; Hsu and Zwiers, 2001]. The method is able to analyze unevenly and sparsely distributed data by considering a varying window width of the kernel, and thus it produces PDFs where data are sparse without spurious local maxima [Silverman, 1986]. Such an approach is needed to derive continuous 2-D PDFs from the O$_3$-CO correlations observed in TRACE-P.

The kernel-type probability density function estimator has the form

$$\hat{f}(x) = \frac{1}{N} \sum_{i=1}^{N} H^{-\alpha} \eta^n \kappa \left( \frac{x - X_i}{H^n} \right),$$

where $x$ is an arbitrary point in the n-dimensional space, $X_i$ represents a data point in this space, and N is the sample size. For our specific problem, we compute the 2-D PDF over the CO-O$_3$ domain (i.e., $n = 2$). $\kappa(\cdot)$ assigns weights that depend upon the relative proximity of $x$ and the data points. This function is called the kernel function and is
chosen to be the Gaussian kernel, i.e., $K(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)$, with $t$ denoting transpose. The width of the kernel function is determined by a smoothing parameter, $H$, weighted by the local data density $\eta_i$, which is inversely proportional to the square-root of the probability density at sample point $i$. Thus a pilot estimation of $f(X_i)$ is required to compute $\eta_i$ for which $f(x)$ is calculated with a fixed window width $h_{\text{pilot}} = 0.96N^{-1/6}$. The optimal value of the smoothing parameter, $H$, is determined by minimizing the score function $M(H)$ which is the integrated square error between the estimator $f(x)$ and the true density $f(x)$. The score function $M(H)$, can be approximated as:

$$M(H) \equiv \int (\hat{f} - f)^2 \, dx \approx \int \hat{f}^2 \, dx - 2/N \sum_{i=1}^{N} \hat{f}_i^2(X_i)$$ (2)

where $\hat{f}_i(X_i)$ is the probability density estimated at sample $i$ with the estimator that is set up to find the minimum of the score function and extract the optimal smoothing parameter and compute the PDF as in equation (1).

[11] We apply this method to the TRACE-P in situ measurements, combining all DC-8 and P-3B flights. The same method is used for the CTM simulations of these in situ measurements. For the large and evenly distributed 4-D CTM data set, however, this method is computationally infeasible and unnecessary, and we use the conventional 2-D histogram with a Laplacian filter to smooth the edges of the distributions.

2.3. Empirical Orthogonal Analysis of Ozone Lidar Profiles

[12] EOF analysis has become one of the more commonly used multivariate statistical tools in atmospheric sciences since Lorenz [1956] introduced this method (see also Wilks [1995]). In climate studies, for example, recent EOF analysis [Thompson and Wallace, 1998, 2000] has sparked tremendous interest in detecting the climate signal of downward propagation from stratosphere to troposphere [e.g., Baldwin and Dunkerton, 1999], and in atmospheric chemistry it has been used to relate the spatial distribution of lightning (and the odd-nitrogen it generates) with the expected anomalies in tropospheric O$_3$ [Martin et al., 2000]. EOF analysis is particularly effective when the EOF patterns can be associated with geophysically meaningful patterns, e.g., in the climate studies above the first EOF is characterized with a more or less annular structure and is named the Arctic Oscillation or Northern Annual Mode. On the other hand, such studies can ignite disputes over whether the EOFs obtained from statistical analysis are real physical modes [e.g., Dommenget and Latif, 2002].

[13] In our study we use EOFs to identify the typical vertical structures of the O$_3$ profiles as measured by the lidar and as simulated in the CTM. Assume that there are $N$ O$_3$ profiles, spanning a range of horizontal space and time, and that each profile $Y_k$ is resolved by $k = 1$ to $K$ vertical levels. The O$_3$ anomaly matrix $Y_{K \times N}$ is constructed from these $N$ vectors of rank $K$ with the mean profile $\bar{Y}$ subtracted (i.e., centered data). Because the vertical levels are uneven, we choose to weight each row of $Y'$ by the square root of that level thickness $(dZ_k)^{1/2}$. The covariance matrix $S_{K \times K}$ is basically the product of the anomaly matrix $Y'$ and its transpose $Y'^T$.

$$S = \frac{1}{N-1} Y' Y'^T.$$ (3)

Points in an O$_3$ profile where data are missing are skipped and do not contribute to the covariance matrix. The degrees of freedom are no longer $N - 1$ as in the denominator in equation (3), but this factor is taken inside the matrix and varies with each element of $S$ depending on the number of nonzero elements in the cross-multiplication between different vertical levels of O$_3$ anomalies.

[14] The eigenvectors $e_i$ of the covariance matrix are the $K$ solutions to equation, $S e_i = \lambda_i e_i$, where $\lambda_i$ are the eigenvalues. The eigenvectors are of rank $K$, and each element corresponds to a vertical level. The EOF vectors $e_i$ are then the eigenvectors rescaled at each level by the inverse of the weights used in calculating the anomaly matrix, i.e., dividing each element of the vector $e_i$ by the square root of the level thickness. We use normalized EOFs such that the inner product of each EOF vector using the level thickness as the weighting is unity, $\Sigma_{k=1}^{K} e_i^k e_i^k dZ_k = 1$. The proportion of the total variance represented by the $i$th EOF vector is $\lambda_i / \Sigma_{k=1}^{K} \lambda_k$.

[15] Any O$_3$ profile $y$ can be expressed in terms of the mean profile $\bar{y}$, the EOFs, and their Principal Components $a_i$ (PCs),

$$y = \bar{y} + y' = \bar{y} + \sum_{i=1}^{K} e_i a_i.$$ (4)

Because of orthogonality of the EOFs, the PCs are readily calculated by projecting the centered data onto the normalized EOFs and have units of ppb (parts per billion as mole fraction) of O$_3$,

$$a_i = e_i^T \cdot y'.$$ (5)

The number of EOFs can be very large and depends on the vertical resolution in this case, but as with most studies, we focus only on those EOFs with the largest eigenvalues since they describe most of the variance. What is different in this study, however, is that each O$_3$ profile is collected over a range of horizontal space and time. Hence the PCs are not just time series as in most published analyses but vary with geographic location as well.

[16] The extended 4-D CTM data set for the TRACE-P domain gives $N = 373248$ O$_3$ profiles. For the observed lidar data, $N = 4871$ profiles. For the CTM data interpolated onto DC-8 Flight tracks to match the 1-min lidar profiles, two data sets are used: Exact lidar simulations with missing data and complete profiles along the flight tracks. To prevent the overwhelmingly large variance of stratospheric ozone from dominating the analysis, we confine analysis to heights below 8.3 km. For CTM data this corresponds to $K = 24$ vertical levels, and for the lidar observations this is interpolated onto 140 levels.

3. Cumulative Probability Distributions of CO and O$_3$

[17] The cumulative probability distributions of the observed CO and O$_3$ abundances show a population that
in situ measurements along the combined in situ data

OBS

lidar data

Talbot et al.

Percentile Levels (in ppb)

measurements from HSU ET AL.: ARE THE TRACE-P MEASUREMENTS REPRESENTATIVE?

TRACE-P In Situ CO and O₃

Simpson et al.

Number of Data Points (N) for CO and O₃

From DC-8

Lidar Data Along DC-8 Flight Tracks, and CTM Data for the Extended 4-D Domain

25 – 46N 1.4 M 5.2 M 0.5 M

14 – 25N 1.0 M 2.8 M 0.3 M

14 – 25N 35814 0.29 M 0.13 M

25 – 46N 33000 0.27 M 75431

N for extended 4-D domain

0 – 1 km 1 – 10 km 10+ km

Sample size (N) for CO in situ data

14 – 25N 793 3135 434

25 – 46N 1624 4337 241

N for O₃ in situ data

14 – 25N 821 3305 486

25 – 46N 1736 4338 271

N for O₃ lidar data

14 – 25N 35814 0.29 M 0.13 M

25 – 46N 33000 0.27 M 75431

N for extended 4-D domain

14 – 25N 1.0 M 2.8 M 0.3 M

25 – 46N 1.4 M 5.2 M 0.5 M

The sample size of the extended 4-D data for 10+ km listed below is from 10–12 km. To compare with the lidar data, the 4-D domain is extended to 18 km and the sample size is 4 times larger than the numbers shown above.

can be separated into background levels (typically the central 50% or more of the population) and into pollution events or stratospheric intrusions (evident in the extreme abundance ranges). Our CTM simulated probability distributions of this flight-track data are a direct test of our ability to represent these populations for the TRACE-P sampling of the domain. Moreover, the same probability distributions from the extended 4-D domain can identify whether the TRACE-P sampling is representative of the larger domain.

We recognize that different parts of the domain will have different extreme populations and background levels, and thus we have split the domain into tropical (14N–25N) and extratropical (25N–46N) as well as boundary layer (0–1 km), free troposphere (1–10 km), and region of stratospheric influence (10–12 km) where air of stratospheric origin are more likely to be sampled. The boundaries are somewhat arbitrary but we found that these choices sufficiently highlighted the different probability distributions in the observations. Table 1 gives the number of data points, N, for CO and O₃ in situ measurements along the combined DC-8 and P-3B flight tracks, O₃ measurements from DC-8 lidar sampling, and the sampling from CTM in the extended 4-D domain. N increases by about 2 orders of magnitude from in situ measurement to lidar sampling, and by about 1 order of magnitude from lidar sampling to 4-D model data sampling. Table 2 summarizes the 25th percentile (first quartile or Q) and the 50th percentile (median or M) for the observations, the simulated observations along flight tracks, and the simulated distributions of the extended 4-D domain. The cumulative probability distributions for CO are plotted in Figure 2a as a function of sigma, σ, the standard deviation of the normal distribution. Vertical dashed lines mark the 25% (Q, σ = –0.675) and 50% (M, σ = 0) probabilities. In each of the six regions, the observed distribution (solid line) is compared with the simulated observations (dashed line). The overlap of these distributions is a measure of the accuracy of the CTM in simulating the TRACE-P observations.

Below the 50th percentile, the CO distribution is extremely well simulated (typically within 5 ppb) by the model for all regions except the tropical boundary layer where the CTM uniformly underestimates the observed CO by about 12 ppb. In CTM sensitivity tests with a range of CO-like tracers (not shown here), we find that much of the observed variance (e.g., as measured by M - Q), including fine-scale features, is driven by large and synoptic-scale systems acting on the global-scale latitudinal gradients in CO, rather than by the nearby east Asian emissions. Thus we take this agreement to mean that the large-scale CO gradients and meteorological systems are well simulated. Above the 75th percentile, however, the simulations are uniformly much smaller than observed. One cause might be the failure of the CTM to resolve urban plumes, for example, the intense, small-scale pollution events such as the Shanghai plume [Russo et al., 2003; Simpson et al., 2003; Talbot et al., 2003]. However, for the distributions shown in the figure (CO < 300 ppb), the observed probability distributions are unaffected by spatial filtering at the CTM resolution, and hence these probability distributions should be resolved by the model. Thus the uniform underprediction of the CO probabilities at the upper end of the distribution as shown are likely due to an underestimate of CO emissions from east Asian sources [Palmer et al., 2003] or possibly to chemical influences rather than lack of model

Table 1. Number of Data Points (N) for CO and O₃ From DC-8 and P-3B In Situ Flight Measurements, O₃ Lidar Data Along DC-8 Flight Tracks, and CTM Data for the Extended 4-D Domain

<table>
<thead>
<tr>
<th>Region</th>
<th>CO N</th>
<th>O₃ N</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–1 km</td>
<td>14–25N</td>
<td>1.0 M</td>
</tr>
<tr>
<td></td>
<td>25–46N</td>
<td>1.4 M</td>
</tr>
</tbody>
</table>

Table 2. TRACE-P In Situ CO and O₃ Percentile Levels (in ppb) Compared With the CTM Simulation Along the Flight Tracks and for the Extended Domain

<table>
<thead>
<tr>
<th>Region</th>
<th>CO Q</th>
<th>CO M</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–1 km</td>
<td>14–25N</td>
<td>1.0 M</td>
</tr>
<tr>
<td></td>
<td>25–46N</td>
<td>1.4 M</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Region</th>
<th>O₃ Q</th>
<th>O₃ M</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–1 km</td>
<td>14–25N</td>
<td>1.0 M</td>
</tr>
<tr>
<td></td>
<td>25–46N</td>
<td>1.4 M</td>
</tr>
</tbody>
</table>

*The sample size of the extended 4-D data for 10+ km listed below is from 10–12 km. To compare with the lidar data, the 4-D domain is extended to 18 km and the sample size is 4 times larger than the numbers shown above.
Figure 2.
resolution (more supporting evidence from the O₃-CO correlations is presented in section 4).

[20] The difference between the flight-track simulations (dashed line) and the extended-domain probabilities (dotted line) is also shown in Figure 2. These latter distributions have three orders of magnitude more points than the in situ sampling and hence smoother curves. In the free troposphere the Q and M values show no obvious statistical bias, but in the boundary-layer there is a preference for sampling higher CO, a possible indication of chasing pollution outflow from the continent. For CO abundances greater than 200 ppb at all heights, the extended-domain sampling includes values over the continent and thus shows greater probabilities for these high-CO events than the simulated flight tracks. The extremely low CO abundances in the extratropics (1–10 km and 10–12 km) indicate air of stratospheric origin, and their frequency ranges from a few percent below 10 km height to as much as 50% in the 10–12 km region. The flight track sampling greatly underestimates their frequency both above and below 10 km height; this reflects the strong latitude and height gradients over this domain and sampling that is preferentially toward the southern and the lower part of the range.

[21] Ozone comparisons show both successful simulations and some obvious model errors. In the free troposphere the observed extratropics probability distribution for the in situ data (Figure 2b) is well matched by the model for the central 50% of the distribution. In the tropics, however, the model accurately matches only the lowest 25% of the distribution and consistently underestimates ozone in the remaining 75 percent of the distribution by 10 ppb or more. For the boundary layer, the model is biased high for both tropics and extratropics. The offset between boundary layer and free troposphere is large and consistently in error for all latitudes: Observations have a shift of about +5 ppb (boundary layer being less than free troposphere for both Q and M); the model predicts an opposite shift of about −10 ppb. This model error can best be explained if the continental boundary-layer sources of ozone from Asia are exaggerated [Wild et al., 2003]. A separate error, the underestimate of tropical ozone (and also the upper 50% of CO) could be due to a missing source, most likely from biomass burning [Wild et al., 2003].

[22] For stratospherically influenced regions the comparison with in situ data is erratic due to the small number of points and the large variability induced mostly by stratospheric intrusions. If we expand the comparison to the DC-8 lidar data (Figure 2c), the number of these high-altitude points increases from a few hundred to a hundred thousand and the probability distributions become well defined. For this sampling the model successfully matches the observations for the lowest 50% of the distribution and predicts about the right frequency of the stratospheric influence (O₃ > 100 ppb). Including the lidar data does not change the previous conclusions and only reemphasizes the systematic error in the boundary layer found with the in situ data (see also later discussion on ozone EOFs).

4. Latitude-Height Distributions of O₃ From the Lidar Sampling

[23] The latitude-height distribution of tropospheric O₃ can help identify the tropopause, stratospheric intrusions, pollution events, and the large-scale gradient between tropics and extratropics. The O₃ lidar data from TRACE-P provide the extensive sampling needed to define this latitude-height section [Browell et al., 2003]. Here, we examine the probability distributions of these latitude-height sections, showing how the sequence from 10th, to 50th, to 90th percentiles (Figure 3) can be used to identify the statistical distribution of tropopause heights, those regions impacted by stratospheric intrusions, and the representativeness of the TRACE-P sampling. Each percentile figure shows the observations (top panel), the CTM simulation along flight tracks (middle panel) and the CTM simulation of the extended 4-D domain (bottom panel). The CTM simulation along the flight tracks follows the lidar sampling with data points missing; while the extended domain statistics assume that O₃ is measured from 0 to 18 km, even in the presence of clouds. The white reference line in the extended-domain plots marks the upper boundary of the flight-track data used in the analysis.

[24] In the tropics, there is clear evidence of a high-ozone region at 5–10 km height near 17°N in the TRACE-P sampling. It is seen in both the observations and the CTM. This region is clearly seen at all percentiles from 10th to 90th, and moreover the ozone abundance increases slowly from about 45 ppb at the 10th percentile to about 75 ppb at the 90th percentile indicating an extensive region of low variability. Even at the 90th percentile, the ozone abundance remains well below stratospheric intrusion levels. This region was sampled on DC-8 Flight 6, and the high O₃ levels have been attributed to biomass burning [Browell et al., 2003]. The TRACE-P sampling clearly singles out this event (i.e., it is not seen on the extended domain) and shows

Figure 2. (a) CO cumulative probability distributions (in ppb of CO) as a function of standard deviation of the normal distribution (σ). The 25 percentile (σ = −0.675) and 50 percentile (median, σ = 0) are shown as vertical dashed lines. The distribution from the combined DC-8 and P3-B in situ observations (solid line) is compared with that from the CTM simulation of these in situ measurements (dashed line). Also shown is the cumulative probability distribution from the 4-D CTM data for the extended TRACE-P domain below 12 km height (dotted line). The data set has been broken into six panels by latitude (tropical = 14°N–25°N (top) versus extratropical = 25°N–46°N (bottom)) and height range (0–1 km (left) versus 1–10 km (center) versus 10–12 km (right)). See Table 1 for the number of data points defining each distribution and Table 2 for the 1st Quartile and Median values. (b) Same as Figure 2a, but for O₃ cumulative probability distributions (in ppb of O₃) based on in situ sampling along flight tracks. (c) Same as Figure 2b, but for O₃ lidar observations and simulations along DC-8 flight tracks. The CTM simulated lidar profiles have the same missing data as the lidar observations. The 4-D data extend to 18 km height.
Figure 3. The (a) 10th, (b) 50th, and (c) 90th percentile of the \(O_3\) latitude-height distribution (ppb) in the extended domain. The data are binned by 1° latitude by 1 km height bins centered at integer latitude and height. (Top row) the TRACE-P lidar observations are compared with (middle row) the CTM simulations of the lidar and with (bottom row) the extended 4-D CTM data set (100E–150E, 3 March to 3 April 2001).
the success of the CTM simulation in reproducing it at all statistical levels. Overall in the free troposphere, the CTM underestimates ozone abundance by about 10 ppb as seen also in the probability distributions in Figures 2b and 2c.

[25] In the midlatitudes, the region of high ozone abundance at 6–12 km near 28 N can be clearly seen as the remnant of a stratospheric intrusion; at the 10th percentile it has similar enhancements to the 17 N region, but the ozone abundance jumps to more than 75 ppb at the 50th percentile and become merged into the stratosphere (>100 ppb) at the 90th percentile. Several flights intercepted stratospheric intrusions in this region behind midlatitude cyclones, indicated in situ measurements by low CO and high wet-bulb potential temperatures (Mari et al., submitted manuscript, 2004). However, the greatest contributions to the O3 feature at 28N come from DC-8 Flight 16, which made a transect at this latitude on 30 March and intercepted a deep intrusion. The greater sampling along this transect explains why the feature is clearly seen in the median as well as the high end of the distribution. Another stratospheric intrusion, visible only at the 90th percentile, is observed in the tropics at 22N between 7–14 km; while several flights contributed to the statistics over this region, a stratospheric intrusion was sampled on only one flight (DC-8 Flight 14). The CTM simulation captures this intrusion, but it occurs at a slightly lower altitude and a little further north. In summary, the CTM captures the basic statistical features of these intrusions along the flight tracks. It simulates the distribution of elevated-ozone (55–95 ppb) as it mixes into the troposphere through the range of percentiles. This success, plus the overall excellent simulation of TRACE-P ozone, ozone sondes, TOMS ozone columns, and the global mean stratosphere-to-troposphere ozone flux (Wild et al., 2003) indicates a good simulation of the dispersion and mixing of stratospheric intrusions.

[26] The statistics for the extended domain show a layer of enhanced ozone, apparently of stratospheric origin, extending from the subtropical break in the tropopause down into the tropics at about 6 km height. At all percentiles, however, they show none of the individual features picked up by the TRACE-P flight tracks. Thus tropical ozone appears to be dominated by stratospheric ozone mixing down from the subtropical jet. In comparing the CTM flight tracks with the extended domain, it is clear that the TRACE-P mission favored sampling of pollution sources with enhanced boundary-layer ozone, but this only emphasizes the CTM exaggeration of boundary-layer ozone when compared with observations, as discussed above. All three altitude-height sections show the descent of stratospheric air (as measured by the 100 ppb contour) by 2–3 km in height as one progresses from 10th to 90th percentile.

5. Probability Density Function in CO-O3 Domain

[27] The patterns of correlation between CO and O3 can identify mixing between different chemical regimes in the atmosphere and further provide information on the photo-chemical evolution of O3 in polluted plumes (Parrish et al., 1993). A first approach to O3-CO correlations is to examine the scatter plots for all in situ measurements. As an example, the one-minute in situ measurements from DC-8 flights 13 and 15 compared with the CTM simulation in Figure 4. In this analysis, the 1-min observations are taken as is, with no spatial smoothing to match the CTM grid. In the figure, the dashed lines (the same in all panels of Figures 4 and 5) are a least-squares fit to all observations for low-CO and high-CO regions (see Figure 5 caption or text below). DC-8 flights 13 and 15 show occurrences of low-CO stratospheric air that are well simulated, including both magnitude and slope.

[28] Model simulations of individual flights are generally excellent, agreeing with the observation of stratospheric air (low CO, very high O3) and pollution plumes (high CO, moderate O3). A combined scatter plot with all the data points would not be easy to interpret, and thus we apply adaptive kernel estimation (section 2.2) to derive two-dimensional probability density functions (PDF) for both measurements and model. The adaptive kernel method generates smooth PDFs without spurious maxima from the more than 10,000 individual points as shown in Figure 5.

[29] In calculating these PDFs, we have chosen to spatially smooth the observations to more closely match the CTM grid. The observations show distinct features on very small scales, such as 100-m thick laminae, which cannot be resolved by the CTM grid (about 500 m vertical by 180 km horizontal). Thus we define a triangular weighting function with a half-height, half-width of 500 m in the vertical and 180 km in the horizontal, and we process the 1-min in situ observations from each flight according to their vertical and horizontal separations. For each point all measurements that fall within a 180 km radius and within 500 m in height contribute to the value at that point.

[30] The top panel in Figure 5 shows the PDF for the observed CO-O3 data points during TRACE-P. The contours are logarithmic (base 10) and denote the probability per unit area in ppb^2. For example, the probability of observing CO between 200 and 201 ppb at the same time as O3 between 50 and 51 ppb is about 10^{-4}. The integral of the PDF over the entire range up to 1000 ppb in CO and O3 is nearly 1. The middle panel of Figure 5 shows the
Figure 5. Probability density (per unit ppb$^2$ area) in the CO-O$_3$ domain plotted in a log base-10 scale. The probability density over the entire CO-O$_3$ domain up to 1000 ppb (not plotted) integrates to 1. Top panel shows the in situ measurements from the combined DC-8 and P-3B flights, middle panel shows the CTM simulation of those in situ data, and the bottom panel shows the extended CTM 4-D data set from the TRACE-P domain sampled below 12 km height. The thick dashed lines, the same in all three panels, are simple linear regression lines from the observed data: The left line for O$_3$ > 100 ppb and CO < 200 ppb, and the right line from the remainder of the data.
Figure 6. The O₃ profiles mean value (ppb) and first 6 normalized EOFs (dimensionless) up to 8.3 km height. Different sets of O₃ data include the extended CTM 4-D domain (solid), the complete CTM profiles along DC-8 flight tracks (dashed), the observed lidar profiles (dotted), and the CTM simulation of the lidar which miss the same data points as the observations (dash-dot). See text.
section 3. The variability structures are quite reasonable up to EOF-5 or EOF-6, and hence this error is likely due to a systematic, time-independent model bias that is apparently related to the over-production of O$_3$ in the boundary layer [Wild et al., 2003].

As shown in the left panels in Figure 8, PC1-PC3 have more uniform zonal distributions while PC4 and PC5 show high values just off the Asian coast in the subtropics. This pattern is consistent with the decomposition of the single transect in Figure 7: The first three EOFs capture mostly the variance associated with the large-scale O$_3$ background, and the higher EOFs explain fine structures that tend to have more localized distributions. PC1 is mostly associated with the stratospheric influence below 8.3 km and exhibits a maximum around 135E and 45N, decreasing monotonically toward the tropics. PC2, whose EOF has a deep boundary-layer structure, has a maximum band around 32N and decreasing both northward and southward, with an equally large negative minimum in the tropics. PC3, whose EOF has a maximum around 5–6 km, has a positive maximum amplitude in the tropics near 14N and a smaller negative minimum near 22N in the subtropics. The fact that the maximum occurs near the tropics seems to suggest that this feature is related to biomass burning. PC4 has a positive maximum distribution along the southeast coast of Asia; PC5 has a maximum located west of Taiwan and North of Hong Kong. These features, in contrast to PC1-PC3, are probably associated with variability from local pollution plumes.

The CTM flight track data (center panel) capture more or less similar PC distributions to the extended 4-D results. Nevertheless, the values are higher than those from the 4-D data indicating a statistical bias, for example, in stratospheric intrusions north of Japan (PC1) and in pollution plumes near the coast (PC4 and PC5). Comparing the observed PCs (right panel) with those from the CTM flight tracks, the agreement is excellent for PC1 and PC2, quite good still for PC3 and PC4 (at least in terms of general pattern), but loses much of the coherence by PC5.

In summary, this EOF/PC analysis of the TRACE-P O$_3$ profiles has clearly quantified the statistical biases in TRACE-P sampling and identified them with specific profile structures and specific locations. In addition there is generally good agreement between model and observations for the geographic patterns of PC1 through PC4, however, some caution on this approach as a model-measurement validation tool is needed. The filling of lidar missing data with model data may have enhanced this
Figure 7. O$_3$ latitude-height transect (in ppb) from the CTM 4-D simulation at 126E at 0500Z, 21 March 2001. The top six panels show the reconstruction of this transect from the latitude-independent mean profile (solid line, first panel in Figure 6) with the cumulative addition of EOF-1 through EOF-6 (latitude independent) multiplied respectively by the values of PC1 through PC6 (latitude dependent). The CTM transect is shown in the next panel below with the same color scale. The residual (transect minus reconstr   ) is shown in the bottom panel with an expanded color scale.
agreement, and additional approaches to analyzing the lidar data for vertical structures are needed.

7. Conclusion

[41] We present a range of atypical statistical analyses of the TRACE-P observations of CO and O\textsubscript{3} to evaluate the representativeness of the TRACE-P observations and to provide possible new insights on the accuracy of chemistry-transport models. Representative is used here to mean that the data along the flight tracks has the same statistical properties as a uniform sampling of an extended region over eastern Asia and the western Pacific. This evaluation uses the modeled distributions from the FRSGC/UCI CTM driven by the Oslo T63L40 ECMWF forecast meteorology (1.9° × 1.9° × 500 m) to compare flight track data with those from an extended 4-D domain defined arbitrarily as 14N to 46N, 100E to 150E, up to 18 km in height, and from 3 March to 3 April 2001. We assume the extended domain as providing unbiased statistics.

[42] We first focus on the central 50% of the distribution for CO and O\textsubscript{3}, since background air and generally avoid pollution plumes or stratospheric influences that appear at the extreme probabilities. For CO, outside of the boundary layer (0–1 km) and the region of dominant stratospheric influence (>10 km in midlatitudes), the 25th and 50th percentiles from the CTM along the flight tracks are basically the same as from the extended 4-D domain, and, moreover, these agree with the TRACE-P observations. In the boundary layer, the CTM flight tracks are systematically 10–20 ppb greater than the extended domain even though the domain includes continental emissions. This indicates that TRACE-P sampling is biased toward sampling pollution plumes. Furthermore, both CTM extended domain and flight tracks are less than the observations, particularly so for the tropical region. Thus the cumulative probability functions for CO support the generally excellent CTM simulation of the observations but for a systematic underestimate of nearby emissions. Even the frequency of low-CO stratospheric influence is well matched along flight tracks but is atypical of the extended domain. For O\textsubscript{3} these same probability functions clearly point out problems: excessive boundary-layer production in midlatitudes but a missing source in the tropics.

Figure 8. The time-averaged PCs (amplitudes of each EOF in ppb) obtained by projecting different centered data onto the EOF-1 to EOF-5 computed from the CTM 4-D data. The left panels show the PCs from the CTM 4-D data, the middle panels show the pseudo-PCs from the CTM simulation of the lidar, and the right panels show the pseudo-PCs from the lidar observations. For the lidar and simulated-lidar data here, the missing points are filled with the CTM simulated O\textsubscript{3} values.
The 10th, 50th, and 90th percentiles of the latitude-by-height distribution of lidar \( O_3 \) show how the CTM reproduces the nonrepresentative clumpy nature of the observations, which is dramatically different than the smooth patterns from the extended domain even at the extreme (10th and 90th) percentiles. This model-measurement comparison also shows good agreement for the statistical height of the stratosphere-troposphere transition (defined here as 100 ppb \( O_3 \)), except about the jet region (30°–35°N) where the model shows intrusion of the 100-ppb air to much lower heights.

Adaptive kernel estimation of the 2-D probability density of \( O_3-\text{CO} \) correlations shows a very good simulation of two different chemical regimes (stratospheric and polluted) that is quite different from the extended domain. It also clearly points out the model failure to predict \( \text{CO} \) > 400 ppb. For the EOF analysis of the vertical \( O_3 \) profiles, the lidar curtain sampling along the flight tracks has the EOF structures shifted downward about 1 km as compared with the extended domain. The latitude-by-longitude maps of the principal components show larger amplitudes for the CTM flight tracks as compared with the extended CTM domain indicating inadequate sampling or bias toward sampling anomalous events. In summary, for most tests we find that the TRACE-P data set shows some statistical biases in sampling and cannot be simply taken as representative of the chemistry and ozone distributions over eastern Asia and the western Pacific in March 2001.

In evaluating model error using these new statistical measures, we find that the FRSGC/UCI CTM simulation of the TRACE-P flight-track data is in most cases quite good and is even better when one takes into account the biased sampling of the extended domain by the specific flight tracks. For example, the CTM does an excellent job in simulating the stratospheric influence in the upper troposphere for the TRACE-P flights, and this influence is quite different from that averaged over the larger region. In previously noted cases in which the model failed to match high-CO events or produced too high \( O_3 \) abundances in the boundary layer, these new analyses point out that the errors are most likely due to source-region errors (e.g., \( \text{CO} \) emissions or near-field \( O_3 \) production) rather than meteorological errors. In most cases the modeled flight-track data look much more like the observations than the model averaged over the region, indicating that the specific spatial and meteorological characteristics of the observations are captured.

Overall, the TRACE-P sampling is not representative of the larger domain we selected. Similar results would likely apply for any useful domain size. We believe that the simplest explanation for this is a combination the limited number of observations plus TRACE-P strategy of sampling the chemical processes in pollution plumes leaving Asia and stratospheric intrusion events associated with cyclones. If one uses such campaign data to detect systematic long-term changes (e.g., between overlapping campaigns such as TRACE-P and PEM-West B [Davis et al., 2003]) or to provide long-term calibration for satellite observations, then the representativeness of the different data sets needs to be evaluated.

Acknowledgments. This work was supported by the NASA Tropospheric Chemistry Program part of the TRACE-P mission.

References


M. A. Avery, E. V. Browell, and G. W. Sachse, Atmospheric Sciences Competency, NASA Langley Research Center, Hampton, VA 23681-0001, USA. (m.aavery@larc.nasa.gov; edward.v.browell@nasa.gov; glen.w.sachse@nasa.gov)

J. Hsu and M. J. Prather, Department of Earth System Science, University of California, Irvine, 2101 Croul Hall, Irvine, CA 92697, USA. (juno@halo.ps.uci.edu; mprather@uci.edu)

I. S. A. Isaksen and J. K. Sundet, Department of Geophysics, University of Oslo, Box 1022, 0315 Oslo, Norway. (ivar.isaksen@geo.uio.no; j.k.sundet@geofysikk.uio.no)

O. Wild, Frontier Research System for Global Change, 3173-25 Showamachi, Yokohama, Kanagawa 236-0001, Japan. (oliver@jamstec.go.jp)