

Characterization of transport errors in chemical forecasts from a global tropospheric chemical transport model

I. Bey

Swiss Federal Institute of Technology Lausanne (EPFL), Switzerland

D. J. Jacob, H. Liu¹, R. M. Yantosca

Division of Engineering and Applied Science and Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, USA

G. W. Sachse

NASA Langley Research Center, Hampton, VA, USA

¹Also at National Institute of Aerospace,
Hampton, VA, USA

Abstract. We propose a new methodology to characterize errors in the representation of transport processes in chemical transport models. We constrain the evaluation of a global three-dimensional chemical transport model (GEOS-CHEM) with an extended dataset of carbon monoxide (CO) concentrations obtained during the Transport and Chemical Evolution over the Pacific (TRACE-P) aircraft campaign. The TRACE-P mission took place over the western Pacific, a region frequently impacted by continental outflow associated with different synoptic-scale weather systems (such as cold fronts) and deep convection, and thus provides a valuable dataset for our analysis. Model simulations using both forecast and assimilated meteorology are examined. Background CO concentrations are computed as a function of latitude and altitude and subsequently subtracted from both the observed and the model datasets to focus on the ability of the model to simulate variability on a synoptic scale. Different sampling strategies (i.e., spatial displacement and smoothing) are applied along the flight tracks to search for systematic model biases. Statistical quantities such as correlation coefficient and centered root-mean-square difference are computed between the simulated and the observed fields and are further inter-compared using Taylor diagrams. We find no systematic bias in the model for the TRACE-P region when we consider the entire dataset (i.e., from the surface to 12 km). This result indicates that the transport error in our model is globally unbiased, which has important implications for using the model to conduct inverse modeling studies. Using the First-Look assimilated meteorology only provides little improve-

ment of the correlation, in comparison with the forecast meteorology. These general statements can be refined when the entire dataset is divided into different vertical domains, i.e., the lower troposphere (<2 km), the middle troposphere (2–6 km), and the upper troposphere (>6 km). The best agreement between the observations and the model is found in the lower and middle troposphere. Downward displacements in the lower troposphere provide a better fit with the observed value, which could indicate a problem in the representation of boundary layer height in the model. Significant improvement is also found for downward and southward displacements in the upper troposphere. There are several potential sources of errors in our simulation of the continental outflow in the upper troposphere which could lead to such biases, including the location and/or the strength of deep convective cells as well as that of wildfires in Southeast Asia.

1. Introduction

Quantifying model errors due to transport processes in chemical transport models (CTMs) is a current concern for a number of reasons. Significant uncertainties still remain on regional estimates of emission inventories that are derived from bottom-up approaches. An increasing number of inverse modelling studies have been carried out in an attempt to reduce uncertainties in regional sources; for carbon monoxide (CO) sources, these include for example *Bergamaschi et al.* [2002], *Kasibhatla et al.* [2002], *Arellano et al.* [2003], *Palmer et al.* [2003], and *Heald et al.* [2004]. The inverse methods seek to relate atmospheric observations, interpreted with a CTM (the forward model), with regional emission inventories through an optimal estimation methodology. However, characterizing the errors in the forward model used to carry out the inversion is a critical issue. Systematic errors in the model transport and/or chemistry will lead to biased estimates of the sources. In particular, *Palmer et al.* [2003] showed that transport errors in the forward model can be the largest source of uncertainties in an inversion analysis of atmospheric CO.

In addition, intercontinental transport of pollution is currently raising a lot of interest [*Parrish and Law*, 2004]. The major pathways for the transfer of pollution from the planetary boundary layer (PBL) to the free troposphere (FT) over the source regions include lifting by warm conveyor belt (WCB) associated with mid-latitude cyclones [e.g., *Stohl*, 2001; *Bey et al.*, 2001a; *Purvis et al.*, 2003; *Liu et al.*, 2003] and lifting by convection [e.g., *Jacob et al.*, 1993; *Horowitz et al.*, 1998; *Liu et al.*, 2003; *Duncan and Bey*, 2004]. Global CTMs are the unique tools that can provide a quantification of the impact of intercontinental transport of pollution on the global distribution of key chemical compounds.

Current CTMs offer however a horizontal resolution of about 1.5° to 4° (150 to 400 km) with a vertical spacing of about 50 to 100 m within the boundary layer, 500 to 1000 m in the middle/upper troposphere and 1200–1500 m at the tropopause levels [*Bey et al.*, 2001b; *Horowitz et al.*, 2003; *Wild et al.*, 2003]. The processes involved in the vertical lifting of trace compounds from the PBL are often not well represented with such resolutions [*Donnell et al.*, 2001]. There is thus a need to better quantify the model errors due to transport, especially those associated with vertical lifting.

In this paper, we describe a new methodology which quantitatively characterizes the model transport errors in CTMs. We used in our analysis the atmospheric observations provided by the NASA aircraft mission Transport and Chemical Evolution over the Pacific (TRACE-P) to evaluate the CTM GEOS-CHEM [*Bey et al.*, 2001b]. The TRACE-P mission took place in February–April 2001 over the western Pacific. The primary goals of the mission were to conduct measurements of various chemical tracers to provide a thorough quantification and characterization of the continental outflow of trace gases and aerosols out of Asia [*Jacob et al.*, 2003]. The TRACE-P mission thus provides an extended dataset of chemical tracers sampled over a large range of synoptic-scale weather patterns that is particularly valuable to evaluate the model’s capabilities to reproduce the processes associated with transport. The flight planning of the two planes engaged during the mission (DC-8 and P-3B) was strongly guided by several chemical and weather forecasts to ensure a comprehensive sampling of the weather patterns that lead to episodes of pollution export from the Asian continent [*Jacob et al.*, 2003]. In particular, five CTMs were used to provide chemical forecasts (including the GEOS-CHEM model). Providing chemical weather forecast is becoming a standard approach for planning aircraft campaigns [*Flatoy*

et al., 2000; *Lawrence et al.*, 2002; *Chin et al.*, 2003; *Carmichael et al.*, 2003], but there is a need to have a quantitative estimate of the extent to which chemical weather forecast products can be trusted for flight planning of field measurement campaigns. In this analysis, we evaluate the transport errors in CO simulations from the GEOS–CHEM model driven by both forecast and assimilated meteorology.

After a brief description of the forecast products and of the CO observations used in that study (Section 2), we describe the statistical tools as well as the Taylor diagrams used to evaluate the model (Section 3). Transport errors in the model are discussed in Section 4 and Section 5 summarizes our findings.

2. GEOS–CHEM Chemical Forecast Products

The GEOS–CHEM model is driven by assimilated meteorological data provided by the NASA Global Modeling and Assimilation (GMAO) office. Typical meteorological products from the GEOS Data Assimilation System (DAS) version 3 (GEOS–3) include 3–D fields archived every 6–hour (winds, temperature, cloud mass fluxes, cloud fractions, etc.) and 2–D fields archived every 3–hour (surface pressure, boundary layer depth, etc.) with a horizontal resolution of $1^\circ \times 1^\circ$ and a vertical resolution of 48 layers from the surface to 0.01 mbar. We degraded the horizontal resolution of the fields to $2^\circ \times 2.5^\circ$ for speeding up the generation of the chemical forecasts.

The GEOS–3 system is run by GMAO in two assimilation modes: the First–Look assimilation (available about 8–15 hours behind real time and which assimilates meteorological data from conventional and satellite observations available at the time), and the Late–Look assimilation (available with a 2-week delay but including a comprehensive set of observations) [*Chin et al.*, 2003]. In addition, the First–Look system provides a 5–day

meteorological forecast twice a day (0 GMT and 12 GMT). During the TRACE-P mission, we applied the GEOS-CHEM model in a forecast mode using a set of 1-day First-Look GEOS-3 assimilated fields and a set of GEOS-3 forecast meteorological fields (0 GMT or 12 GMT) subsequently for each forecast run, to generate a 5-day chemical forecasts twice a day. To speed up the generation of the chemical forecasts, only CO concentrations were calculated during the mission. Previously archived monthly mean hydroxyl radical (OH) fields were used to compute the chemical loss due to the reaction of CO with OH as well as the chemical production associated with hydrocarbon oxidation following the method described in *Bey et al.* [2001a]. Comprehensive emissions inventories for anthropogenic, biomass burning and biogenic sources are included in the GEOS-CHEM model as described in *Bey et al.* [2001]. We used these global climatological inventories during the mission, except for the anthropogenic fossil fuel and biofuel sources over East Asia for which we used the 2000 emission inventory provided by *Streets et al.* [2003] in support of the campaign. Generation of the chemical forecast was started 3 weeks before the aircraft campaign begins to initialize model CO concentrations. During the mission, restart files were archived at each run and used to initialize the following chemical forecast. For the present analysis, we used the 24-hour chemical forecast products.

3. Methodology for Quantifying the Model Errors

3.1. Data Collected During TRACE-P

In the present study, we used observed CO, a continental chemical tracer of fossil fuel and biomass burning origin. Diode laser spectroscopic measurements of CO [*Sachse et al.*, 1987] were taken during TRACE-P on board of both the DC-8 and the P-3B planes. Figure 1 gives a general view of the CO dataset collected during TRACE-P in the lower

(<2 km), middle (2–6 km) and upper (>6 km) troposphere. Several continental outflow events were sampled during the mission as shown by the elevated CO concentrations measured over the TRACE–P region, revealing the strong impact of the Asian continent on the atmospheric composition over the western Pacific [*Jacob et al.*, 2003]. Both anthropogenic (fossil fuel and biofuel) and biomass burning emissions contributed to the Asian continental outflow [*Liu et al.*, 2003]. Lifting ahead of cold fronts associated with conveyor belts (WCBs), low level flows in the boundary layer behind cold fronts, and deep convection have been found to be the major meteorological processes that drive the continental pollution outflow from Asia [*Liu et al.*, 2003]. These pathways contribute to impact different regions of the troposphere, as further discussed in the following.

For our analysis, we consider CO data for all the flights from both planes (excluding the P–3B flight 19 for which no forecast data is available due to a computer failure). We used the 1–minute merge files and further averaged them over the 2°x2.5° model grid along the flight tracks. We only consider CO data for the geographical area with longitudes from the Pacific Rim to the date line, since we are mostly interested in evaluating the model’s performance in the representation of chemical features associated with continental outflow events. CO data obtained further east during the transit flights were mostly dedicated to the sampling of trace gas concentrations representative for background air [*Jacob et al.*, 2003]. We excluded stratospheric air masses by removing CO observations associated with ozone concentrations higher than 100 ppbv.

3.2. Statistical Quantities Used to Discuss Model Performances

Quantitative evaluations of the model CO concentrations were the subject of a number of previous studies, [e.g., *Heald et al.*, 2003a, 2003b; *Palmer et al.*, 2003]. According to

an inverse model analysis by *Palmer et al.* [2003], the *Streets et al.* [2003] inventory used to produce the chemical forecasts is too low by 30%. This is however not an issue for the present analysis since we are not interested in the model's capabilities to reproduce the actual CO observed concentrations. Instead, we are interested in the model's capabilities to reproduce the transport processes associated with synoptic-scale weather patterns, which can be seen as the distributions of ΔCO concentrations, where Δ is defined as the increase in a given concentration over the background concentration. Background CO concentrations were estimated for the lower and the upper troposphere in each 2° bin of latitude over the TRACE-P region by computing the 10th percentiles for the observed values over the entire period (Figure 2). Background CO concentrations present strong latitudinal and vertical variations reflecting the sources at the northern mid-latitudes, the structure of the continental outflow to the western Pacific (i.e., strongest in the lower troposphere and 30–40°N) [*Bey et al.*, 2001a; *Liu et al.*, 2003]), and the latitudinal and vertical variation in CO photochemical sink. In the following, we systematically subtracted the background concentrations estimated as a function of latitude and altitude from each data point of the observed and simulated fields, a standard approach used in evaluation of weather forecast skills [see for example, *Hollingsworth et al.*, 1980; *Murphy and Epstein*, 1989; *Kalnay et al.*, 1990; *Storch and Zwiers*, 1999].

The usual way to produce a simulated dataset to be compared with an aircraft-observed dataset is to sample the model results in such a way that the location and timing of the two 4-D (3 dimensions plus time) fields correspond. We refer to that sampling as the reference sampling. We then searched for systematic biases in the model by sampling the CO model fields using different strategies, as indicated in Table 1. Two main sampling

strategies, displacement and smoothing, are tested both in the horizontal and vertical directions. This allows us to give a measure of how the forecast errors can be accounted for by moving or smoothing the simulated fields to best fit the observations [*Hoffman et al.*, 1995]. Displaced simulated fields are obtained by applying a systematic displacement from the reference model grid point while the smoothed simulated fields are obtained by averaging a number of model grid points located around the reference model grid point. The sampling tests were carried out using instantaneous forecast CO concentrations with a 2-hour temporal resolution as well as monthly mean forecast CO concentrations to indicate in a quantitative manner the usefulness of chemical weather forecast during a field campaign. Model results from simulations driven by assimilated meteorology were also examined to assess the skill of the assimilated meteorology versus that of the forecast meteorology (Table 1).

The reference, displaced and smoothed simulated fields obtained by the various sampling strategies were then evaluated against the observed field. Figure 3 shows for example a scatter-plot of observed field (2687 observations) against the simulated field obtained from a reference sampling of the instantaneous forecast CO concentrations. The two fields present a relatively low correlation coefficient of 0.44. However it should be mentioned that the correlation coefficient is significantly higher (0.60) when the background concentrations are not removed from the observed and the simulated fields, reflecting the fact that model and observed backgrounds are correlated. *Kiley et al.* [2003] intercompared four state-of-the-art global models (including the GEOS-CHEM model) using the DC-8 observations and found that correlations between model and observed values typically range from 0.56 to 0.75, in agreement with our value. Figure 3 also reveals the difficulties

of the model to reproduce the highest CO plumes, a problem also observed in the CTMs examined in *Kiley et al.* [2003].

The different statistical quantities (such as correlation coefficient, root-mean-square error, and standard deviation) indicated on Figure 3 are subsequently used to quantitatively compare the skills of the simulated fields obtained from the different samplings following *Taylor* [2001]'s approach. The so-called Taylor diagrams are built to provide a concise summary of several statistical quantities indicating the degree of correspondence between two datasets (i.e., an observed and a simulated field). Indeed, on the Taylor diagrams, the correlation coefficient (R), the centered root-mean-square difference (RMS_c) between a tested field (model) and a reference field (observations), as well as the ratio of the standard deviation of the two fields are indicated by a single point on a two-dimensional plot, which provide a way to compare the performances of different simulations (or simulated fields) toward the same observed field.

Let us note f the simulated field and r the observed field defined at N discrete points, and \bar{f} and \bar{r} are the mean values of f and r , respectively. For each simulated field, we computed the standard deviation σ_f , the correlation coefficient R between the simulated and the observed field and the centered root-mean-square difference RMS_c as follows:

$$\sigma_f = \left[\frac{1}{N} \sum_{n=1}^N (f_n - \bar{f})^2 \right]^{\frac{1}{2}} \quad (1)$$

$$R = \frac{\frac{1}{N} \sum_{n=1}^N (f_n - \bar{f})(r_n - \bar{r})}{\sigma_r \sigma_f} \quad (2)$$

$$RMS_c = \left[\frac{1}{N} \sum_{n=1}^N ((f_n - \bar{f}) - (r_n - \bar{r}))^2 \right]^{\frac{1}{2}} \quad (3)$$

where σ_r is the standard deviation of the observed field.

R , RMS_c , and σ_f provide complementary information about the similarities and differences between the observed and the simulated fields, and thus, the model performances. The geometrical relationship between these quantities and on which Taylor diagram is based is shown on Figure 4. The radial distance from the origin of each given point is proportional to σ_f ; the R between the observed and the simulated dataset is given by the azimuthal position; the RMS_c between the simulated and the observed dataset is proportional to their distance apart [Taylor, 2001]. We used here a normalized version of Taylor diagrams, where the variables σ_f , σ_r , and RMS_c are normalized by the standard deviation σ_r of the observed field.

4. Evaluating Transport Errors in the Model

4.1. Displacement and Smoothing Errors in the Chemical Forecast Products

Figure 5 shows the results of the statistical analysis when data points within the whole troposphere (2687 observations) are considered, while Figures 6 to 8 focus on specific regions, <2 km (1221 observations), 2–6 km (979 observations), and >6 km (487 observations), respectively. The reference point corresponding to the observed dataset (black dot on Figure 5–8) is normalized by itself and thus is plotted at unit distance from the origin. The red dot on Figures 5–8 corresponds to the reference sampling of the instantaneous model results obtained with the forecast meteorology (with a correlation coefficient of 0.44, also shown in Figure 3) while blue and green symbols display the statistical quantities for the displaced and smoothed simulated forecast fields, respectively.

Displacement in the horizontal direction of one grid point corresponds to 2° (230 km) in the north–south direction and to 2.5° (about 150–250 km depending on the latitude) in the east–west direction. Most of the displaced fields (blue symbols noted as W1, W2, E1,

E2, S1, S2, N1, N2) lead to lower R and higher RMS_c than that of the reference simulated field. This indicates that most of the simulated and observed fields show less similarities in comparison with the reference sampling in terms of both the location and timing of the patterns (e.g., the continental outflow signatures induced by different synoptic-scale weather patterns) and the amplitudes of these patterns. One exception to this general result is displacement to the south by one box, which provides a slightly higher correlation coefficient of 0.47 (to be compared with 0.44). For a southerly displacement greater than one box, the correlation starts to decrease again and is even much smaller (0.33) than that of the reference sampling. In most of the cases, a higher displacement lead to lower R , as expected. Note for example the evolution from the point N1 (northerly displaced by one box) to the point N2 (northerly displaced by two boxes).

Horizontal smoothing of the model results (downward green triangle H1, H2, and H3 on Figure 5) does not provide a better fit with the observed field. The smoothed simulated fields lead to correlation close to the reference one, but to smaller standard deviations, indicating a decrease in the representation of the variability and amplitude of the observed continental outflow signatures. In general, an increase of the size of the smoothing-region (i.e., the region over which the average is computed) leads to a decrease of the standard deviation, as expected. For example, from the horizontal averaging over 9 boxes (H1 symbol) to the horizontal averaging over 49 boxes (H3 symbol), the normalized standard deviation decreases from 0.56 to 0.41.

Figure 5 also shows statistical quantities calculated applying vertical displacements (upward blue triangles on Figure 5) of 0.5, 1, and 1.5 km above (Va1, Va2, Va3) and below (Vb1, Vb2, Vb3) the reference level using vertical interpolations between the data

points of a model column. Similarly to what we find for the horizontal direction, none of the vertically smoothed and displaced fields show significantly higher similarities to the observed fields than the reference simulated field. Note however that this finding results in fact from compensating effects between the different vertical domains of the troposphere. As further discussed in the following section, downward displacements lead to better fits in the region <2 km (Figure 6) and >6 km (Figure 8), in contrast to the result found for the intermediary 2–6 km domain (Figure 7).

4.2. Displacement and smoothing Errors in the Lower, Middle and Upper Troposphere

The correlation coefficient found for the <2 km (0.48) and 2–6 km (0.56) domains (Figures 6 and 7, respectively) are higher than that of the whole troposphere (0.44, Figure 5), while the correlation for the upper troposphere is much lower (0.33, Figure 8), in agreement with the findings of *Kiley et al.* [2003] who also noted better model performances for the lower regions of the troposphere.

For the lower troposphere, similarly to what we found for the whole troposphere (Figure 5), only a limited improvement is obtained from a sampling strategy different from the reference one, indicating that there is no major transport bias in that region of the troposphere. A better fit with the observed field is however found when model results are sampled below the reference level (Vb1, Vb2, and Vb3). Displacement by 500 m below the reference level increases the R from 0.48 to 0.56 and decrease the RMS_c from 77 ppbv to 72 ppbv. The distribution of chemical tracers over the western Pacific below 2 km is strongly impacted by events of low-level continental outflow taking place behind cold fronts and are usually capped below 2 km by strong subsidence [*Liu et al.*, 2003;

Carmichael et al., 2003]. The better results observed for a displacement below the reference level in the < 2 km domain over the TRACE-P region could reflect a problem in the representation of the boundary layer heights in the model, either over the continental region from which the pollution is exported or over the marine areas over which the pollution is transported in the western Pacific.

Figure 7 shows the statistical quantities for the domain 2–6 km. In that region, the signature of continental outflow in the middle troposphere is mainly induced by the passage of cold fronts. Pollution is lifted out of the continental boundary layer by WCBs ahead of the fronts and further transported to the Western Pacific between 2 and 6 km [*Liu et al.*, 2003, *Carmichael et al.*, 2003]. A much larger dispersion of the points over the Taylor diagram is seen for the domain 2–6 km (Figure 7) in comparison with Figures 5 and 6. All the different sampling strategies lead to a significant degradation of at least one important statistical quantity (the correlation or/and the standard deviation). This indicates that the synoptic-scale patterns contributing to the transport in that region are well reproduced in the model, in term of location, timing and amplitude. The only exception to that general result is the displacement by one grid point (230 km) toward the south which leads to a small increase (decrease) of the R (RMSc) from 0.56 to 0.57 (from 58 ppbv to 53 ppbv). A small improvement of the R to 0.59 is also found for a westward displacement but it also leads to an increase of the RMS_c ; therefore overall, this displacement does not provide a better fit with the observations.

The poorer correlation ($R=0.33$) between the reference simulated and the observed fields is found for the domain >6 km (Figure 8). For this region, some of the smoothing and displacement sampling strategies actually improve the statistical quantities (correlation

and/or RMS_c), indicating that there are some systematic biases for that region in the model representation of the features associated with continental outflow. In particular, a better fit to the observed values is found with a downward displacement below the reference level ($R=0.42$) as well as a southward displacement ($R=0.43$). The continental outflow observed in the upper troposphere is mainly due to deep convection over Southeast Asia, especially at low latitudes [*Liu et al.*, 2003] showing a clear biomass burning signature. The DC-8 flights 9, 10, and 14, for example, flew between 20°N and 30°N and encountered convective outflow in the upper troposphere which were attributed to biomass burning combustion over Southeast Asia [*Browell et al.*, 2003; *Jacob et al.*, 2003; *Liu et al.*, 2003].

To further investigate these systematic biases, we examined Taylor diagrams for the 13 individual flights (not shown) that flew at higher altitudes and we found that each of them gives very different fits to the observed values. Six out of the 13 flights were more specifically dedicated to the sampling of convective outflow (i.e., flights 6, 9, 10, 12, 14, 15, see Table 4a in *Jacob et al.* [2003]). Figure 9 shows the results of vertical and horizontal displacements when data points from only the convective-impacted flights are considered. We extended the displacement lengths to quantify the distance that provides the best fit with the observations. We find that on average, downward displacements lead to an improvement of the statistical quantities, and that a downward distance of 500 m provides the best fit with the observations (with the highest R and the lowest RMS_c). This indicates that convection in our model delivers the continental pollution at too high altitudes. This result is consistent with the findings of *Allen et al.* [1997] who reported that an earlier version of the GEOS DAS archive (GEOS-1) overestimated the extent of

convection in subtropical regions such as Southeast Asia. *Kiley et al.* [2003] also found that the GEOS–CHEM model exhibits too strong convection over Southeast Asia and eastern India, as least for one DC–8 flight during the campaign but they did not report a systematic problem in the simulation of convective events with the GEOS–CHEM model.

Southward displacements also provide some improvements, and a better fit with the observations is found for a distance of about 500 km. This finding could point out to systematic errors in the location of the deep convective cells, but also could indicate a problem in the location of wild fires in the model. Note that a slightly better fit is also found for a southward displacement in the middle and lower troposphere (Figures 6 and 7), which are also impacted, to a lesser extent, by biomass burning effluents. For the model simulations presented here, we used a climatological and monthly mean inventory for biomass burning emissions. *Duncan et al.* [2003] reported however that biomass burning exhibits a large interannual variability and *Heald et al.* [2003a] found that using a 2001 monthly versus a climatological monthly emissions in their model provides a significant improvement of their TRACE–P simulation. A displacement in the biomass burning fires would also lead to systematic errors in the location of the outflow. More recent work on biomass burning emission inventory (Rynda Hudman, personal communication, 2004) reveals that the geographical distribution of wildfires used in our study for the wildfires is likely incorrect. In particular, emissions over north of India and in the Bangladesh region are possibly too high. Because of the systematic errors found for the dataset in the upper troposphere, it is questionable whether these data should be considered in inverse modelling studies.

4.3. Forecast Meteorology Versus Late–Look Assimilated Meteorology

Simulated fields were also generated with the reference sampling strategy using instantaneous model results obtained with a simulation driven by the Late–Look assimilated meteorology instead of the forecast meteorology (yellow dots on Figures 5–8). Over the whole domain, using the Late–Look assimilated meteorology improves R from 0.44 to 0.50, and does not change significantly the RMS_c , indicating that using the Late–Look meteorology improves the simulation (Figure 5). This general result reflects in fact different tendencies depending on the region of the troposphere. The improvement induced by the use of assimilated meteorology instead of the forecast fields is seen in the middle troposphere where the R improves from 0.56 to 0.63 and the RMS_c decreases from 58 ppbv to 55 ppbv (Figure 7). Only a slight improvement is seen in the lower troposphere (Figure 6) while, surprisingly, the assimilated First–Look meteorology does not lead to any improvement of the simulation in the upper troposphere.

4.4. Instantaneous Model Results Versus Monthly Mean Averages

Figures 5–8 also show the statistical quantities obtained for simulated fields sampled with the reference strategy but using monthly mean model outputs (instead of the instantaneous model outputs) obtained with both the forecast and the assimilated meteorology (red and yellow stars, respectively). These simulated fields lead to smaller correlation (0.2 for the whole domain) and slightly higher RMS_c and also clearly show a much lower standard deviation. Indeed, as expected, the variability of the simulated fields obtained from the monthly mean archived concentrations is much smaller than that obtained with the instantaneous archived model results. Note that the statistical quantities calculated with the archived monthly mean in the upper troposphere do not appear on the Tay-

lor diagram as the correlation coefficient is negative (Figure 8), which indicates that the model has some success in representing the timing of continental outflow events induced by deep convection. The general degradation of the statistical quantities obtained with the monthly mean model outputs in comparison with that obtained with the instantaneous values clearly show the usefulness of chemical weather forecasts in designing aircraft campaigns.

5. Conclusions

We used the extensive dataset provided by the TRACE-P aircraft mission to evaluate model transport biases in the trace gas distributions obtained from a global chemical transport model. CO concentrations collected by the two planes engaged during the mission offer a unique opportunity to evaluate model errors as they were sampled over a rather large region, and under a variety of weather patterns, namely, frontal systems and deep convection events. Background CO concentrations were computed from the observations as a function of altitude and latitude and were subsequently subtracted from both the simulated and the observed dataset to evaluate the performances of the model at reproducing the elevations of CO over the background. The model outputs were sampled along the plane flight tracks in different ways to produce a number of displaced and smoothed simulated fields which allows one to characterize systematic biases in the transport model. The different simulated datasets were evaluated against the observed dataset using Taylor diagrams to compare simultaneously the statistical quantities obtained from all the sampling strategies. We find no systematic bias in the model transport for the TRACE-P region when we consider all data points from the surface to 12 km, as none of the sampling strategies tested provided a significant improvement of R and RMS_c . This finding

has important implications for inverse modeling studies that use the GEOS-CHEM, as it indicates that the transport error in our model is globally unbiased.

We further examined model errors by dividing the entire dataset into different vertical domains, i.e., the lower troposphere (<2 km), the middle troposphere (2–6 km), and the upper troposphere (>6 km). In general, continental outflow in the different vertical regions of the troposphere are associated with different processes, i.e., low level flows behind cold fronts, lifting ahead of cold fronts, and deep convection for the regions <2 km, 2–6 km, and >6 km, respectively. The best agreement between observed and simulated reference fields are found in the lower ($R=0.48$) and middle troposphere ($R=0.56$). In these two domains, we find that the model biases are insignificant, especially in the middle troposphere, where most of the different sampling strategies induce a significant degradation of at least one of the important statistical quantities. For the lower troposphere, downward displacements provide a better fit with the observed value, which could indicate a problem in the representation of boundary layer height in the model. Our statistical analysis reveals that the upper troposphere is likely the region where transport biases are the more important, since the simulated fields generated from smoothing and displacement sampling strategies improve in a significant way the statistical quantities. There are several potential sources of model errors in our simulation of the continental outflow in the upper troposphere, including the location and/or the strength of deep convective cells and wildfires.

Acknowledgments. The GEOS-CHEM model is managed by the Atmospheric Chemistry Modeling Group at Harvard University with support from the NASA Atmospheric

Chemistry Modeling and Analysis Program. We are grateful to Colette L. Heald and Guergana Guerova for helpful comments on the text.

References

- Allen, D. J., K. E. Pickering, and A. Molod, An evaluation of deep convective mixing in the Goddard Chemical Transport Model using International Satellite Cloud Climatology Project cloud parameters, *J. Geophys. Res.*, 102, 25467-25476, 1997.
- Arellano, A. F., Jr., P. S. Kasibhatla, L. Giglio, G. R. van der Werf, and J. T. Randerson, Top-down estimates of global CO sources using MOPITT measurements, *Geophys. Res. Lett.*, 31, L01104, doi:10.1029/2003GL018609, 2004.
- Bergamaschi, P., R. Hein, M. Heimann, and P. Crutzen, Inverse modeling of the global CO cycle, 1. Inversion of CO mixing ratios, *J. Geophys. Res.*, 105, 1909-1928, 2000.
- Bey I., D. J. Jacob, J. A. Logan, and R. M. Yantosca, Asian chemical outflow to the Pacific: origins, pathways and budgets, *J. Geophys. Res.*, 106, 23,097-23,114, 2001a.
- Bey I., D. J. Jacob, R. M. Yantosca, J. A. Logan, B. Field, A. M. Fiore, Q. Li, H. Liu, L. J. Mickley, and M. Schultz, Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, *J. Geophys. Res.*, 106, 23,073-23,096, 2001b.
- Browell, E. V., et al., Large-scale ozone and aerosol distributions, air mass characteristics, and ozone fluxes over the western Pacific Ocean in late winter/early spring, *J. Geophys. Res.*, 108(D20), 8805, doi:10.1029/2002JD003290, 2003.
- Carmichael, G. R., et al., Regional-scale chemical transport modeling in support of the analysis of observations obtained during the TRACE-P experiment, *J. Geophys. Res.*,

108(D21), 8823, doi:10.1029/2002JD003117, 2003.

Chin M., P. Ginoux, R. Lucchesi, B. Huebert, R. Weber, T. Anderson, S. Masonis, B. Blomquist, A. Bandy, and D. Thornton, A global aerosol model forecast for the ACE-Asia field experiment, *J. Geophys. Res.*, 108 (D23), 8654, doi:10.1029/2003JD003642, 2003.

Donnell, E. A., D. J. Fish, E. M. Dicks, and A. J. Thorpe, Mechanisms for pollutant transport between the boundary layer and the free troposphere, *J. Geophys. Res.*, 106(D8), 7847-7856, 10.1029/2000JD900730, 2001.

Duncan, B. N., and I. Bey, A modeling study of export pathways of pollution from Europe: Seasonal and Interannual variations (1987-1997), *J. Geophys. Res.*, 109, D08301, doi:10.1029/2003JD004079, 2004.

Duncan, B. N., R. V. Martin, A. C. Staud, R. Yevich, and J. A. Logan, Interannual and seasonal variability of biomass burning emissions constrained by satellite observations, *J. Geophys. Res.*, 108, 4040, doi:10.1029/2002JD002378, 2003.

Flatøy, F., Ø. Hov, and H. Schlager, Chemical forecasts used for measurements flight planning during POLINAT 2, *Geophys. Res. Lett.*, 27, 951-954, 2000.

Heald, C.L., D. J. Jacob, P. I. Palmer, M. J. Evans, G. W. Sachse, H. B. Singh, and D. R. Blake, Biomass burning emission inventory with daily resolution: Application to aircraft observations of Asian outflow, *J. Geophys. Res.*, 108 (D21), 8811, doi:10.1029/2002JD003082, 2003a.

Heald C. L., et al., Asian outflow and trans-Pacific transport of carbon monoxide and ozone pollution: An integrated satellite, aircraft, and model perspective, *J. Geophys. Res.*, 108 (D24), 4804, doi:10.1029/2003JD003507, 2003b.

- Heald, C. L., D. J. Jacob, Dylan B. A. Jones, P. I. Palmer, J. A. Logan, D. G. Streets, G. W. Sachse, J. C. Gille, R. N. Hoffman, and T. Nehrkorn, Comparative inverse analysis of satellite (MOPITT) and aircraft (TRACE-P) observations to estimate Asian sources of carbon monoxide, *J. Geophys. Res.*, submitted, 2004.
- Hoffman, R. N., Z. Liu, J.-F. Luis, and C. Grassoti, Distorsion representations of forecast errors, *Month. Weather Rev.*, 123, 2758-2770, 1995.
- Hollingsworth, A., K. Arpe, M. Tiedtke, M. Capaldo, and H. Savijrvi, The Performance of a medium-range forecast model in winter, *Mon. Wea. Rev.*, 108, 11, 17361773, 1980.
- Horowitz, L. W., et al., A global simulation of tropospheric ozone and related tracers: Description and evaluation of MOZART, version 2, *J. Geophys. Res.*, 108(D24), 4784, doi:10.1029/2002JD002853, 2003.
- Horowitz, L. W., J. Liang, G. M. Gardner, and D. J. Jacob, Export of reactive nitrogen from North America during summertime, *J. Geophys. Res.*, 103, 13,451-13,476, 1998.
- Jacob D. J., J. H. Crawford, M. M. Kleb, V. S. Connors, R. J. Bendura, J. L. Raper, G. W. Sachse, J. C. Gille, L. Emmons, and C. L. Heald, Transport and Chemical Evolution over the Pacific (TRACE-P) aircraft mission: Design, execution, and first results, *J. Geophys. Res.*, 108 (D20), 9000, doi:10.1029/2002JD003276, 2003.
- Jacob, D. J., J. A. Logan, G. M. Gardner, R. M. Yevich, C. M. Spivakovsky, S. C. Wofsy, S. Sillman and M. J. Prather, Factors regulating ozone over the United States and its export to the global atmosphere, *J. Geophys. Res.*, 98:14817, 1993.
- Kalnay, E., M. Kanamitsu, and W. E. Baker, Global Numerical Weather Prediction at the National Meteorological Center, *Bull. Amer. Meteor. Soc.*, 71(10), 1410-1428, 1990.

- Kasibhatla, P., A. Arellano, J. A. Logan, P. I. Palmer, and P. Novelli, Top-down estimate of a large source of atmospheric carbon monoxide associated with fuel combustion in Asia, *Geophys. Res. Lett.*, 29(19), 1900, doi:10.1029/2002GL015581, 2002.
- Kiley, C. M., et al., An intercomparison and evaluation of aircraft-derived and simulated CO from seven chemical transport models during the TRACE-P experiment, *J. Geophys. Res.*, 108 (D21), 8819, doi:10.1029/2002JD003089, 2003.
- Lawrence, M., P. Rasch, R. Kuhlmann, J. Williams, H. Fischer, M. Reus, M., J. Lelieveld, P. Crutzen, M. Schultz, P. Stier, H. Huntrieser, J. Heland, A. Stohl, C. Forster, H. Elbern, H. Jakobs, and R. Dickerson, Global chemical weather forecasts for field campaign planning: predictions and observations of large-scale features during MINOS, CONTRACE, and INDOEX, *Atmos. Chem. Phys.*, 3, 267-289, 2003.
- Liu, H., D. J. Jacob, I. Bey, R. M. Yantosca, B. N. Duncan, and G. W. Sachse, Transport pathways for Asian pollution outflow over the Pacific: Interannual and seasonal variations, *J. Geophys. Res.*, 108(D20), 8786, doi:10.1029/2002JD003102, 2003.
- Murphy, A. H., and E. S. Epstein, Skill scores and correlation coefficients in model verification., *Mon. Wea. Rev.*, 117, 3, 572-582, 1989.
- Palmer, P. I., D. J. Jacob, D. B. A. Jones, C.L. Heald, R.M. Yantosca, J.A. Logan, G. W. Sachse, and D. G. Streets, Inverting for emissions of carbon monoxide from Asia using aircraft observations over the western Pacific, *J. Geophys. Res.*, 108 (D21), 8828, doi:10.1029/2003JD003397, 2003.
- Parrish, D., and K. Law, Intercontinental transport and chemical transformation Lagrangian-2k4, *IGACTivities Newsletters*, No.29, 2003.

- Sachse, G. W., G. F. Hill, L. O. Wade, and M. G. Perry, Fast-reponse, high precision carbon monoxide sensor using a tunable diode laser absorption technique, *J. Geophys. Res.*, 92 (D2): 2071-2081, 1987
- Stohl, A., A 1-year Lagrangian climatology of airstreams in the Northern Hemisphere troposphere and lowermost stratosphere, *J. Geophys. Res.*, 106(D7), 7263-7280, 10.1029/2000JD900570, 2001.
- Storch, H. V., and W. Zwiers, Statistical analysis in climate research, Cambridge University Press, Cambridge, U.K., 1999.
- Streets, D. G., et al., An inventory of gaseous and primary aerosol emissions in Asia in the year 2000, *J. Geophys. Res.*, 108 (D21), 8809, doi:10.1029/2002JD003093, 2003.
- Taylor, K. E., Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.*, 106, 7183-7192, 2001.
- Wild, O., J. Sundet, M. J. Prather, I. Isaksen, H. Akimoto, E. V. Browell, and S. J. Oltmans, CTM ozone simulations for spring 2001 over the western Pacific: Comparisons with TRACE-P lidar, ozonesondes, and TOMS columns, *J. Geophys. Res.*, 108(D21), 8826, doi:10.1029/2002JD003283, 2003.

Figure 1. CO concentrations (1-min averaged) obtained during the TRACE-P mission for three vertical domains, <2 km, 2–6 km and >6 km.

Figure 2. Background CO values estimated from the TRACE-P observations as a function of latitude for the lower troposphere (from the surface to 2 km, black circle) and the free troposphere (from 2 km to 12 km, open circle). Original CO observations were averaged over the $2^\circ \times 2.5^\circ$ GEOS-CHEM grid. Stratospheric influence was removed by eliminating data point with O_3 observations higher than 100 ppbv. Background values are evaluated by calculating the 10th percentile for each latitude and altitude bins.

Figure 3. Scatter plot of the TRACE-P ΔCO data versus the corresponding GEOS-CHEM model ΔCO obtained from a simulation driven by the forecast meteorology. ΔCO corresponds to the enhancement above the background, (i.e., both the observed and the simulated fields are subtracted from the background CO concentrations). Data west of the date line from all the DC-8 and the P-3B flights are averaged over the $2^\circ \times 2.5^\circ$ model grid (except the P-3B flight 19), which lead to a total of 2687 observations. A number of statistical quantities are indicated, including the correlation coefficient (R), the mean bias, the root-mean-square difference (RMS), the centered RMS difference (RMS_c) and the standard deviations for the observed (σ_{obs}) and the simulated dataset (σ_{mod})

Figure 4. Geometrical relationship between the standard deviation of the observed field (σ_r), the standard deviation of the simulated field (σ_f), the correlation coefficient (R) and the centered RMS difference (RMS_c) in a normalized Taylor diagram [Taylor, 2001]. The radial distance from the origin of each given point is proportional to the standard deviation of the simulated field. The correlation coefficient between the observed and the simulated dataset is given by the azimuthal position. The centered RMS difference between the simulated and the observed datasets is proportional to the distance between the reference point and that of the simulated field. We used here a normalized version of the Taylor diagram, i.e., σ_r , σ_f and RMS_c are normalized by σ_r .

Figure 5. Taylor diagrams for different sampling strategies of the simulated dataset over the TRACE-P region (west of the date line). The black dot corresponds to the observed dataset and is used as the reference to which the simulated fields are compared. Each sampling strategy is represented by a given symbol as described in Table 1. See also Figure 4 for a description of the Taylor diagrams.

Figure 6. Same as Figure 5 but for the lower troposphere (< 2 km).

Figure 7. Same as Figure 5 but for the middle troposphere (2–6 km).

Figure 8. Same as Figure 5 but for the upper troposphere (> 6 km). Note that the points corresponding to the reference sampling obtained from monthly mean model outputs do not appear on this Figure as they have a negative correlation coefficient.

Figure 9. Taylor diagrams for flights impacted by convective outflow in the upper troposphere (> 6 km). A total of 215 data points are considered from the DC-8 flights 6, 9, 10, 12, 14, 15. Displacement lengths are extended in all directions to quantify the distance that provides the best fit with the observations.

Table 1. List of the sampling strategies used to produce simulated dataset

Symbols ^a	Meteorology ^b	Sampling description ^c
<i>Reference sampling</i>		
Red dot	Forecast	Reference sampling (correspondence in space and time with the observations)
<i>Horizontal displacement (blue downward triangle)^{a d f}</i>		
W1	Forecast	1 grid point to the west of the reference model grid point
W2	Forecast	2 grid points to the west of the reference model grid point
E1	Forecast	1 grid point to the east of the reference model grid point
E2	Forecast	2 grid points to the east of the reference model grid point
N1	Forecast	1 grid point to the north of the reference model grid point
N2	Forecast	2 grid points to the north of the reference model grid point
S1	Forecast	1 grid point to the south of the reference model grid point
S2	Forecast	2 grid points to the south of the reference model grid point
<i>Vertical displacement (blue upward triangle)^{a d f}</i>		
Va1	Forecast	0.5 km above the reference model grid point
Va2	Forecast	1 km above the reference model grid point
Va3	Forecast	1.5 km m above the reference model grid point
Vb1	Forecast	0.5 km below the reference model grid point
Vb2	Forecast	1 km below the reference model grid point
Vb3	Forecast	1.5 km below the reference model grid point
<i>Horizontal smoothing (green downward triangle)^{a g}</i>		
H1	Forecast	Mean value between 9 grid points, the reference grid point and the 8 grid points around it (1 grid points in both the east-west and the north-south directions) at the same model level
H2	Forecast	Mean value between 25 grid points, the reference grid point and the 24 grid points around it (2 grid points in both the east-west and the north-south directions) at the same model level
H3	Forecast	Mean value between 49 grid points, the reference grid point and the 48 grid points around it (3 grid points in both the east-west and the north-south directions) at the same model level
<i>Vertical smoothing (green upward triangle)^{a g}</i>		
V1	Forecast	Mean value of 3 grid points, <i>i.e.</i> , the reference grid point and the grid points above and below it (0.5 km)
V2	Forecast	Mean value of 5 boxes, <i>i.e.</i> , the reference grid point and the 2 grid points above and below it (0.5 and 1 km)
V3	Forecast	Mean value of 7 boxes, <i>i.e.</i> , the reference grid point and the 6 grid points above and below it (0.5, 1 and 1.5 km)
<i>Other samplings^{a c}</i>		
Red star	Forecast	Reference sampling (using the monthly mean archived concentrations)
Yellow dot	Assimilated	Reference sampling
Red star	Assimilated	Reference sampling (using the monthly mean archived concentrations)

^a Symbols and colors refer to those used on the Taylor diagrams (see Figures 5-8).

^b "Assimilated" means that the First-Look assimilated meteorology was used to produce the simulated datasets.

^c All the sampling are conducted using the instantaneous concentrations fields archived every 2-hour unless otherwise indicated (see "Other samplings")

^d "Displacement" means that the simulated field is produced by sampling the model results with a systematic displacement from the reference model grid point.

^e Displacement in the horizontal direction of one grid point corresponds to 2° (230 km) in the north-south direction and to 2.5° (about 150-250 km depending on the latitude) in the east-west direction.

^f Displacement in the vertical direction are calculated for distances of 0.5, 1, and 1.5 km, by vertically interpolating model results from a model column.

^g "Smoothing" means that the simulated fields is produced by averaging a number of grid points, including the reference model grid point and grid points around it over a given zone.

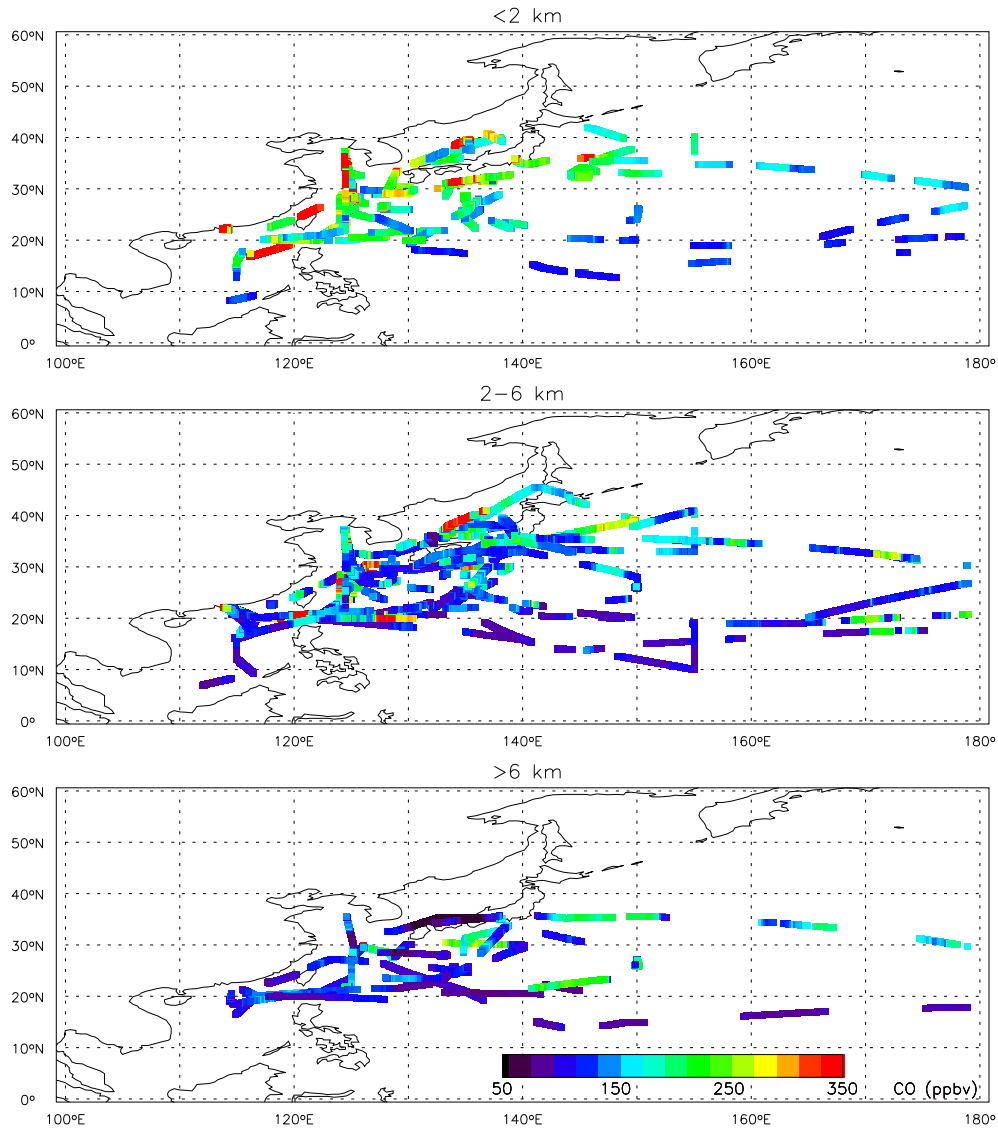


Figure 1. CO concentrations (1-min averaged) obtained during the TRACE-P mission for three vertical domains, <2 km, 2-6 km and >6 km.

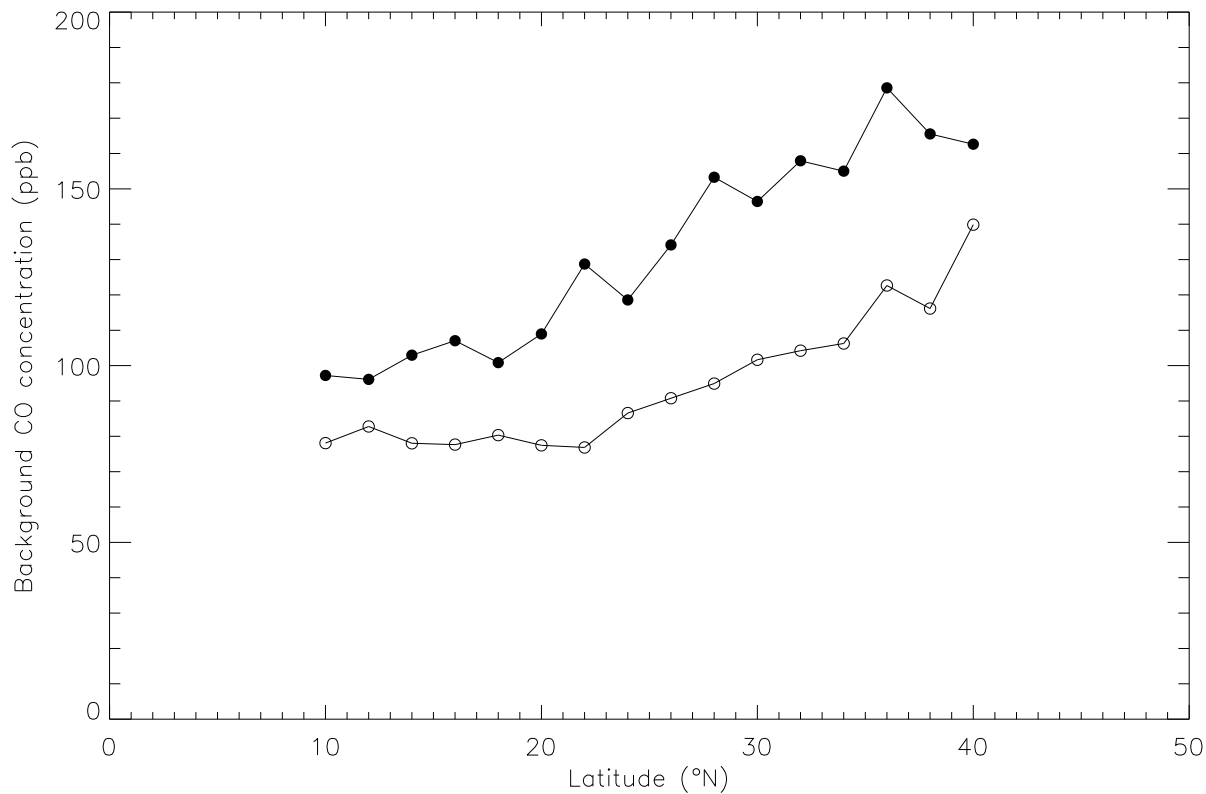


Figure 2. Background CO values estimated from the TRACE-P observations as a function of latitude for the lower troposphere (from the surface to 2 km, black circle) and the free troposphere (from 2 km to 12 km, open circle). Original CO observations were averaged over the $2^{\circ}\times 2.5^{\circ}$ GEOS-CHEM grid. Stratospheric influence was removed by eliminating data point with O_3 observations higher than 100 ppbv. Background values are evaluated by calculating the 10th percentile for each latitude and altitude bins.

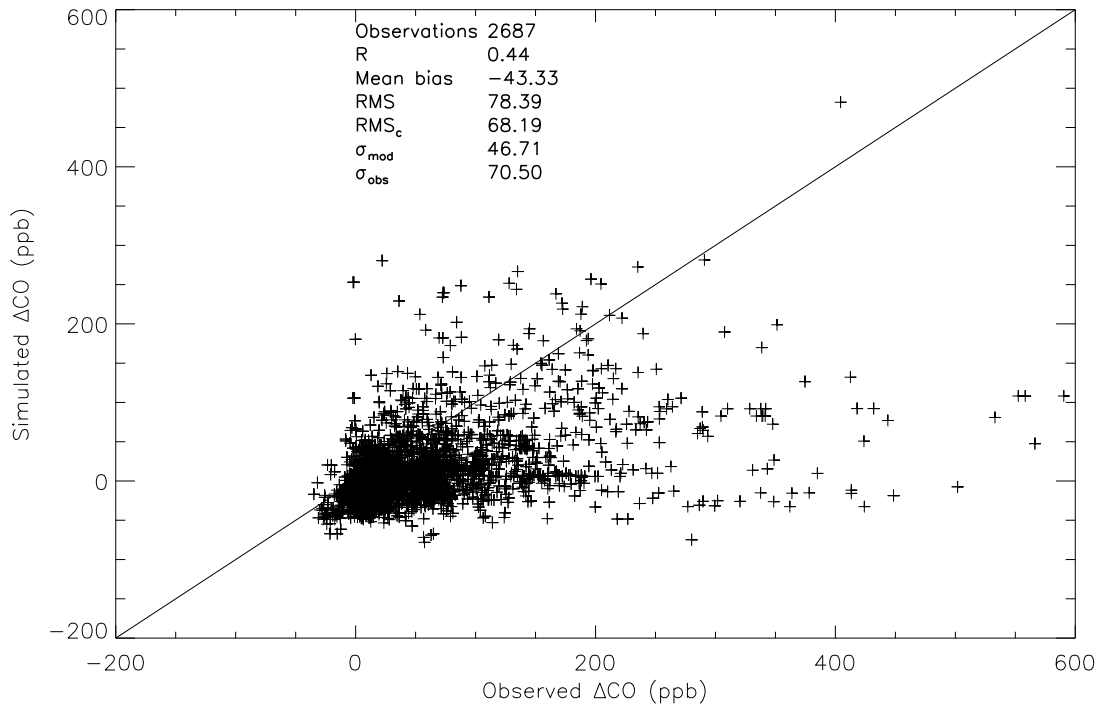


Figure 3. Scatter plot of the TRACE-P ΔCO data versus the corresponding GEOS-CHEM model ΔCO obtained from a simulation driven by the forecast meteorology. ΔCO corresponds to the enhancement above the background, (i.e., both the observed and the simulated fields are subtracted from the background CO concentrations). Data west of the date line from all the DC-8 and the P-3B flights are averaged over the $2\times 2.5^\circ$ model grid (except the P-3B flight 19), which lead to a total of 2687 observations. A number of statistical quantities are indicated, including the correlation coefficient (R), the mean bias, the root-mean-square difference (RMS), the centered RMS difference (RMS_c) and the standard deviations for the observed (σ_{obs}) and the simulated dataset (σ_{mod})

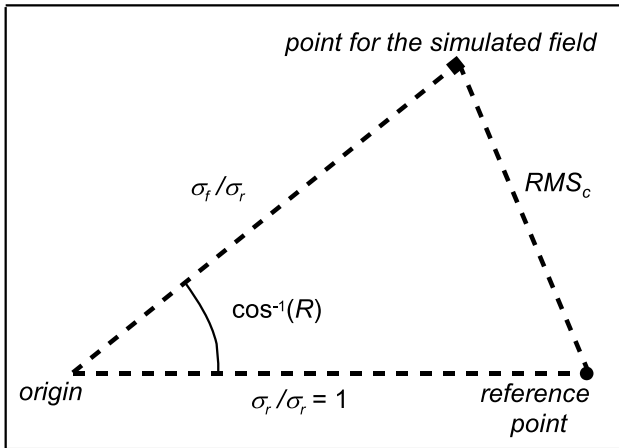


Figure 4. Geometrical relationship between the standard deviation of the observed field (σ_r), the standard deviation of the simulated field (σ_f), the correlation coefficient (R) and the centered RMS difference (RMS_c) in a normalized Taylor diagram [Taylor, 2001]. The radial distance from the origin of each given point is proportional to the standard deviation of the simulated field. The correlation coefficient between the observed and the simulated dataset is given by the azimuthal position. The centered RMS difference between the simulated and the observed datasets is proportional to the distance between the reference point and that of the simulated field. We used here a normalized version of the Taylor diagram, i.e., σ_r , σ_f and RMS_c are normalized by σ_r .

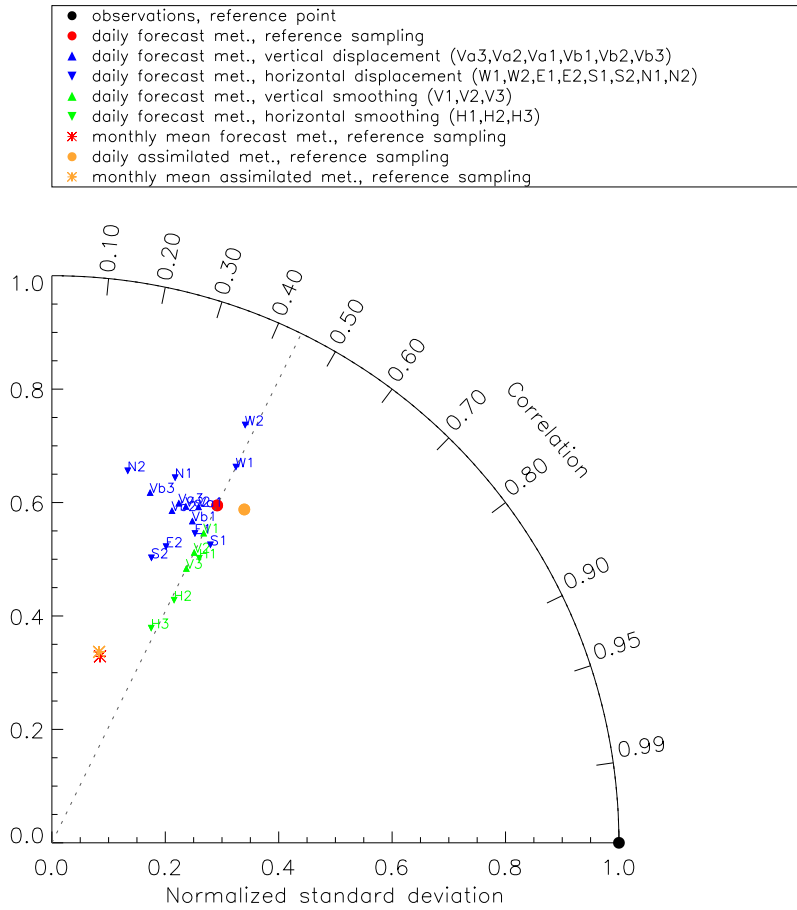


Figure 5. Taylor diagrams for different sampling strategies of the simulated dataset over the TRACE-P region (west of the date line). The black dot corresponds to the observed dataset and is used as the reference to which the simulated fields are compared. Each sampling strategy is represented by a given symbol as described in Table 1. See also Figure 4 for a description of the Taylor diagrams.

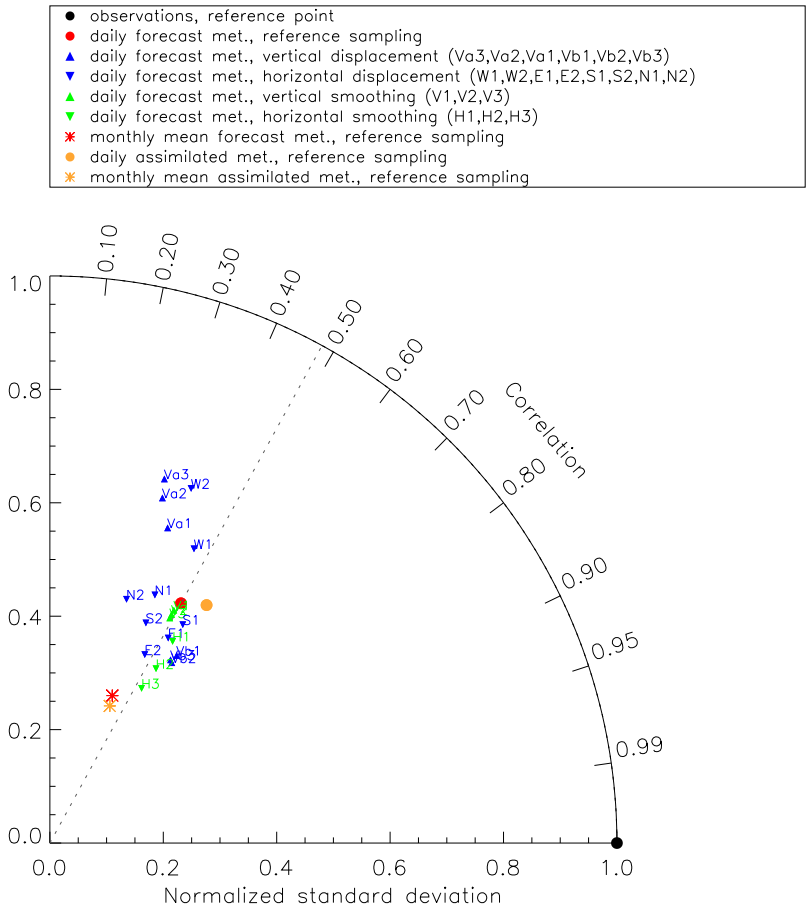


Figure 6. Same as Figure 5 but for the lower troposphere (< 2 km).

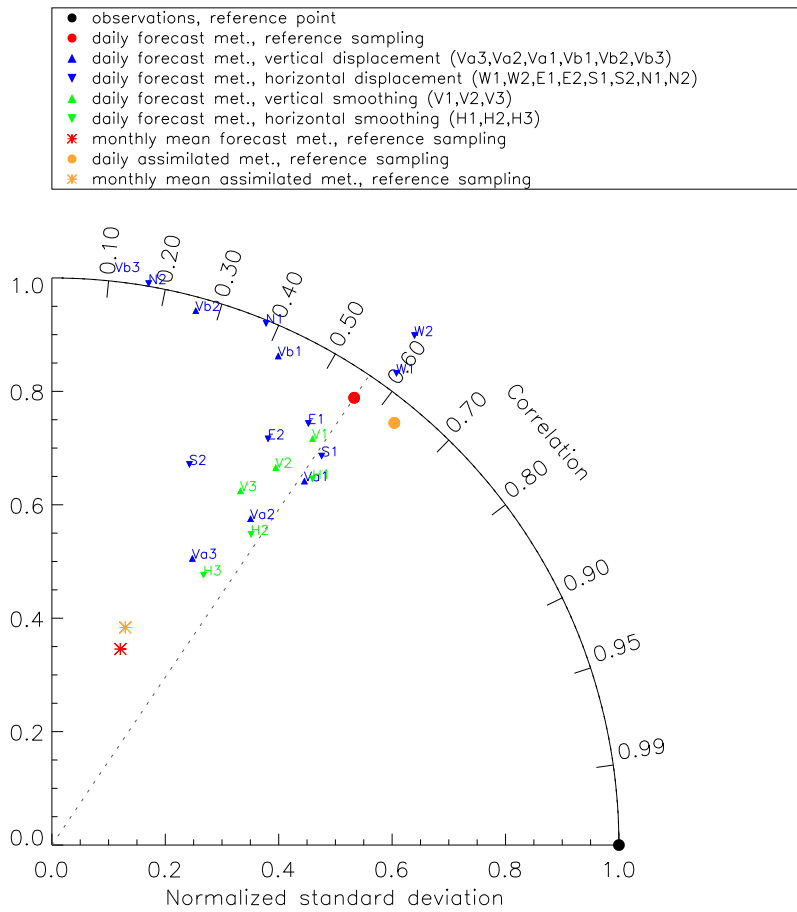


Figure 7. Same as Figure 5 but for the middle troposphere (2–6 km).

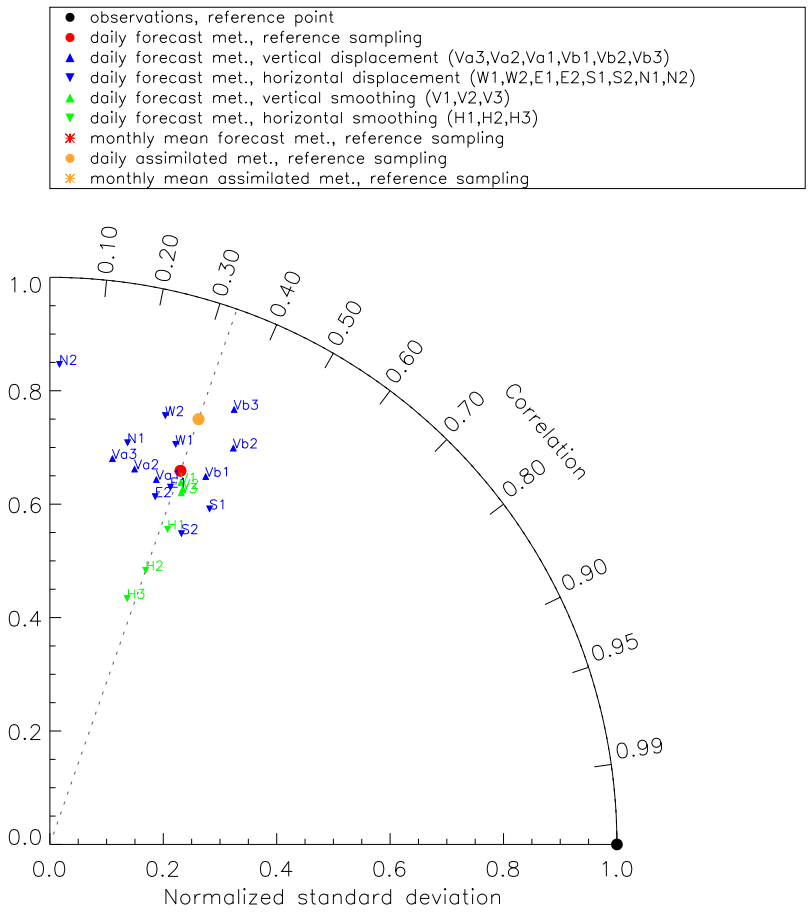


Figure 8. Same as Figure 5 but for the upper troposphere (> 6 km). Note that the points corresponding to the reference sampling obtained from monthly mean model outputs do not appear on this Figure as they have a negative correlation coefficient.

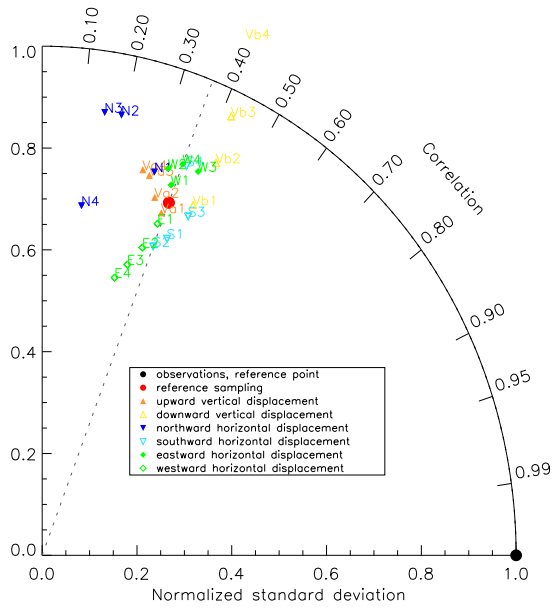


Figure 9. Taylor diagrams for flights impacted by convective outflow in the upper troposphere (> 6 km). A total of 215 data points are considered from the DC-8 flights 6, 9, 10, 12, 14, 15. Displacement lengths are extended in all directions to quantify the distance that provides the best fit with the observations.