Abstract. Comparisons of aircraft measurements of carbon monoxide (CO) to the retrievals of CO using observations from the Measurements of Pollution in The Troposphere (MOPITT) instrument onboard the Terra satellite are presented. Observations made as part of the NASA INTEX-B and NSF MIRAGE field campaigns during March–May 2006 are used to validate the MOPITT CO retrievals, along with routine samples from 2001 through 2006 from NOAA and the MOZAIC measurements from commercial aircraft. A significant positive bias, around 20% for total column CO, in MOPITT CO was found in the comparison to in situ measurements during 2006. Comparisons to the long-term records of measurements from NOAA and MOZAIC revealed an increasing bias in the V3 MOPITT CO retrievals over time. The impact of an instrumental drift is illustrated through retrieval simulations.

1 Introduction

Retrievals of carbon monoxide (CO) from the Measurements of Pollution in The Troposphere (MOPITT) instrument onboard the Terra satellite have been available continuously with a constant instrument configuration for more than six years (since September 2001). While MOPITT started making measurements in March 2000, one of the two coolers failed in May 2001, requiring a change in the channels used in the retrievals (Deeter et al., 2004). No major instrument issues have affected MOPITT operations since the cooler failure in 2001.

Validation of the MOPITT Version 3 (V3) retrievals against in situ measurements from aircraft has been performed on a regular basis since the start of the mission (Emmons et al., 2004, 2007). This has included comparison of MOPITT CO retrievals to aircraft in situ measurements as part of routine sampling performed by NOAA at several sites, intensive field campaigns, and sampling from commercial aircraft.

This paper follows the precedent of publishing scientific papers on the validation of satellite observations (e.g. Clerbaux et al., 2008; Livesey et al., 2008; Warner et al., 2007; Heue et al., 2005; Sussmann and Buchwitz, 2005; Sussmann et al., 2005). Validation papers such as these do not necessarily contain new scientific results. The goal is to compare measurements in the most scientifically rigorous manner possible, often across spatial and temporal scales, and account for representativeness in each dataset. In addition, taking into account differences in vertical sensitivity is non-trivial. This paper covers a wide variety of validation exercises covering diverse geographical and seasonal cases including both monitoring and intensive field campaigns. The MOPITT observations are the longest global record of tropospheric CO and are used widely by the scientific community, therefore communicating this information to the community is essential.

This paper extends previous work by presenting the results of MOPITT validation using aircraft measurements made during the NASA INTEX-B and NSF MIRAGE field campaigns in 2006, as well as the long-term record from NOAA observations and the MOZAIC experiment. Details of the
in situ measurements used in this validation study are presented in the next section. Section 3 gives the details of the validation comparisons and presents the results. Section 4 discusses the possible causes of bias and illustrates the impact of a drift in the MOPITT instrument on the bias with time.

2 Aircraft measurements

The aircraft measurements used for the MOPITT validation presented here are made by several different techniques. The details of these instruments and the experiments in which they were used are described below. The location of the measured profiles used for validation are shown in Fig. 1.

2.1 INTEX-B DC-8

As part of the NASA Intercontinental Chemical Transport Experiment, part B (INTEX-B) the NASA DC-8 made flights in 2006 from Texas 1–22 March and from Hawaii and Alaska 17 April–15 May. While no spirals were planned to coincide with a Terra overpass, 11 sampled profiles were used in these comparisons, generally reaching to an altitude of about 10 km. The in situ CO measurements made from the DC-8 during INTEX-B were made by the fast response tunable diode laser (TDL) instrument, called DACOM (Differential Absorption CO Measurement) (Sachse et al., 1987). The time response of the measurements is 1 s and their precision is 1% or 1 ppbv, whichever is greater. Measurement accuracy is tied to that of the reference gases provided by NOAA that are based on the NOAA/WMO 2000 reference scale (Novelli et al., 2003).

2.2 INTEX-B C-130

The NCAR/NSF C-130 made flights from Veracruz, Mexico, 1–31 March 2006, as part of the NSF Megacities Impact on Regional and Global Environment (MIRAGE) experiment and from Seattle, Washington, 17 April–15 May as part of INTEX-B. The maximum altitude of the C-130 is about 6 km. A total of 14 profiles over both phases were coincident with MOPITT observations, with two of those being spirals that were timed to be within half an hour of the
Table 1. NOAA ESRL CO profile sampling sites with dates and number of profiles used in this study.

<table>
<thead>
<tr>
<th>Site Description</th>
<th>Code</th>
<th>Location</th>
<th>Dates</th>
<th>No. Profiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harvard Forest, Mass.</td>
<td>HFM</td>
<td>42.5 N, 72.2 W</td>
<td>8/9/2001–30/10/2006</td>
<td>47</td>
</tr>
<tr>
<td>Carr, Colorado</td>
<td>CAR</td>
<td>40.9 N, 104.8 W</td>
<td>29/8/2001–9/11/2006</td>
<td>48</td>
</tr>
<tr>
<td>Virginia</td>
<td>VAA</td>
<td>32.9 N, 79.4 W</td>
<td>28/7/2004–3/10/2006</td>
<td>9</td>
</tr>
<tr>
<td>Rarotonga, Cook Islands</td>
<td>RTA</td>
<td>21.2 S, 159.8 W</td>
<td>29/8/2001–16/10/2006</td>
<td>61</td>
</tr>
</tbody>
</table>

The measurements made on the C-130 used a vacuum UV resonance fluorescence instrument similar to that of Gerbig et al. (1999). These data have a 3 ppbv precision, 1 s resolution, and a typical accuracy better than ±10% for a 100 ppbv ambient mixing ratio. Calibrations are tied to the NOAA/WMO 2000 reference scale (Novelli et al., 2003).

2.3 NOAA

As described in Emmons et al. (2004), canister samples have been collected from small aircraft at several NOAA/ESRL/GMD monitoring sites since MOPITT began operating in 2000 at times coincident with Terra overpasses. Table 1 lists the locations of these sites, along with the dates and number of profiles used in these comparisons. Analysis is performed in the NOAA laboratory in Boulder, Colorado, by gas chromatography, followed by HgO reduction detection, using the NOAA/WMO 2000 reference scale (Novelli et al., 1998, 2003). The long-term stability of the reference scale is good to ±4% below 100 ppb and 1–2% through 225 ppb.

2.4 MOZAIC

The MOZAIC (Measurement of OZone, water vapour, carbon monoxide and nitrogen oxides by Airbus In-service airCraft) program includes measurements of CO on several commercial aircraft. The majority of these flights are from Europe to North America or eastern Asia, and back, but also include flights to South America, the Mideast and West Africa. CO measurements were made with an improved infrared correlation instrument with a time resolution of 30 s, and a precision of ±5 ppbv ±5%, with calibrations performed with NIST standards (Nedelec et al., 2003). Profiles of CO on descent to or ascent from airports were used in comparisons to MOPITT retrievals when within 200 km and 12 h of MOPITT measurements. A total of 4312 profiles between 1 January 2002 and 1 September 2006 were used for the validation presented here. The same procedure as described in Emmons et al. (2007) was used for determining coincidences, i.e., the mean latitude and longitude of the MOZAIC measurements between 800 and 500 hPa are used for matching with MOPITT overpasses. Because the ascents and descents are sampled over 150–400 km in distance, they are not precisely vertical profiles. However, the MOPITT measurements are relatively less sensitive to the CO concentrations at the surface or in the upper troposphere, so the location of the MOZAIC measurements in the lower to middle troposphere is most relevant for validation.

3 Validation results

Carbon monoxide is retrieved from the MOPITT measured radiances using a maximum a posteriori (MAP) algorithm (Deeter et al., 2003). Insufficient information is contained in
Fig. 2. Validation results for January–October 2006. MOPITT CO retrievals at 700 hPa, 250 hPa and for the total column are plotted against in situ aircraft measurements that have been transformed by the MOPITT averaging kernels and a priori. Black symbols are for the C-130 (+) and DC-8 (×) measurements during INTEX-B and MIRAGE; colored symbols are for the various NOAA sampling sites (see Table 1).

The results of the validation with the measurements that were part of the INTEX-B and MIRAGE field experiments, along with the NOAA GMD samples during 2006, are shown in Fig. 2. Each point in Fig. 2 represents the median value of the MOPITT retrievals within the time and spatial constraints described in Sect. 2, plotted against the aircraft profile transformed by the averaging kernels and a priori. Error bars represent the inner quartile range of all the MOPITT pixels. Aircraft measurements from the NOAA sites and the field campaigns have consistent correlation and bias with respect to MOPITT. A distinct positive bias is apparent at all retrieval levels with a mean bias of 19% for the total column. The bias is somewhat stronger at lower altitudes, being 25% at 750 hPa, while 9% at 250 hPa. These bias values are significantly larger than those published in our earlier validation studies, indicating a drift in the MOPITT retrievals. The column bias was 5±11% determined for measurements in 2000 (Emmons et al., 2004) and approximately 7±8% in 2004 (Emmons et al., 2007).

To quantify this change in the bias, the validation data sets have been examined for each year. Figure 3 shows the MOPITT column retrievals plotted against the transformed aircraft measurements. The validation results from the INTEX-A and COBRA-2004 field campaigns, as presented in Emmons et al. (2007), are shown in the 2004 plot. The comparisons have been sorted by the individual NOAA sites,
field campaign data for 2004 and 2006, and for geographical regions of the MOZAIC data. The correlation between the MOPITT retrievals and aircraft measurements is generally good, as illustrated in Fig. 3. Several of the very high MOZAIC observations are much larger than the MOPITT CO retrievals. These are likely strong pollution plumes with limited spatial extent that are not captured by the coarse horizontal resolution of the 22 km MOPITT observations.

Figures 4–6 show the biases in the column, 700 hPa and 250 hPa retrievals for the site and regional groupings for each year. Each point in these figures represents the mean and standard deviation of the biases (MOPITT minus aircraft) of the corresponding points shown in Fig. 3. The mean biases for each year for the long-term NOAA sites (leaving out Virginia), the field campaign data, and all of the MOZAIC data are given in Table 2. While there is large uncertainty and scatter in the biases, there is clearly an increase in the bias from 2002 to 2006 at almost all sites and regions. In addition, the NOAA sites generally show larger biases than the MOZAIC data.

4 Possible causes of bias and changes in bias

Retrieval bias has a variety of potential sources. Biases might be identified with either instrument modeling, radiative transfer modeling, biases in ancillary datasets (e.g., meteorological data) or the retrieval algorithm. Two particular sources of bias in the V3 product are now fairly well characterized. One type of bias is associated with the assumption of Gaussian VMR variability rather than log-normal VMR variability. As shown by Deeter et al. (2007), the assumption of Gaussian VMR variability in the retrieval algorithm is inconsistent with in-situ datasets and leads to positive retrieval bias in especially clean conditions (e.g., VMR values less than 60 ppbv). V4 retrievals will use a state vector based on log(VMR) and will therefore not be subject to this effect.

A second source of potential retrieval bias in the V3 product occurs only in particularly polluted conditions. Retrieved profiles with CO total column values larger than the maximum value in the forward model training set (approximately $4 \times 10^{18}$ mol cm$^{-2}$) are rejected by the retrieval algorithm because of the inability of the forward model used in V3 to handle such profiles. This effect may be a significant source of negative retrieval bias for the MOZAIC retrievals, which are primarily located near major urban centers.

The cause of the increasing bias in MOPITT CO retrievals is not entirely understood, but an instrumental drift appears to be a contributing cause. Throughout the operational phase of the MOPITT mission, both the static pressure in LMC3 (the sole remaining length-modulation cell) and “working pressure” (or mean pressure) in PMC2 (the sole remaining pressure modulation cell) have been slowly but steadily decreasing. Analysis of temperature and pressure engineering data indicates that in both cases, the pressure loss is at least partially a consequence of a slow leak, or possibly adsorption of

![Fig. 3. MOPITT CO column retrievals versus aircraft measurements, including NOAA, Field Campaigns and MOZAIC for each year. Symbols are colored by site or region of measurement.](image-url)
the gas to the cell walls. In the MOPITT V3 product, the Forward Model and retrieval processing do not account for time-dependent instrument cell parameters. Hence, the changing state of the instrument could lead to a time-dependent retrieval bias, or drift. Qualitatively, the effect of decreasing absorber amount in the LMC and PMC can be understood in terms of the Difference (D) signal instrument response functions (Drummond and Mand, 1996; Edwards et al., 1999). Spectrally, these D signal response functions are small in the gaps between the CO spectral lines and larger immediately around the CO spectral lines. As the amount of absorber and pressure in either the LMC or PMC decreases, the peaks in

Fig. 4. Bias between MOPITT and aircraft in situ measurements for the column retrieval for each year, sorted by NOAA site, Field Campaign or MOZAIC geographical region. Each symbol and error bar indicates the mean and standard deviation of the biases for each site or region.

Fig. 5. As Fig. 4, for the 700 hPa retrieval.
the D signal response functions will generally become more narrow and decrease in magnitude. This leads to a negative bias in the Level 1 radiances for the CO-sensitive D signals. Because increasing atmospheric CO concentrations also tend to decrease the CO-sensitive D signals, this instrumental bias will be transformed by the retrieval algorithm into a positive bias in retrieved CO concentrations.

The MOPITT retrieval algorithm incorporates a fast radiative transfer model based on prescribed instrument parameters, including the pressures and temperatures of the length- and pressure-modulation cells. For the MOPITT Version 3 algorithm, the assumed instrument parameters were based on pre-launch values measured in a laboratory. Over the duration of the MOPITT mission, drifts have been observed in the pressures and temperatures of the modulation cells. Because these drifts lead to discrepancies between the actual instrument state and the assumed state used for developing the fast radiative transfer model, these long-term instrument drifts can potentially lead to drifts in both the Level 1 radiances and resulting retrievals.

The magnitude of the retrieval biases resulting from long-term drifts in the MOPITT length- and pressure-modulation cells have been estimated through retrieval simulations. These simulations explicitly quantify the effect of exploiting a static radiative transfer model (based on fixed instrument parameters for one point in time) to process radiances produced by the instrument with perturbed instrument parameters. For these simulations, the retrieval algorithm incorporated a radiative transfer model based on MOPITT instrument parameters averaged over 2006. Radiances were simulated over a set of test atmospheres for two dates representing both an “early-mission” date (2 December 2002) and a much more recent “late-mission” date (1 February 2008). Comparisons of the simulated retrievals with the “true profiles” (processed appropriately with the averaging kernels as in Eq. 1) for the two dates yields an estimate of the effect of changing instrument parameters on long-term retrieval bias drift. Results of these simulations are shown in Fig. 7. These simulations were performed with the soon-to-be-released Version 4 product, which exploits a ten-level grid, with retrieval levels every 100 hPa. For the early-mission simulations, retrieval biases are typically negative and largest in the mid-troposphere. At 600 hPa, the mean retrieval bias is approximately −3 ppbv. For the late-mission date, retrieval biases are typically positive and largest in the upper troposphere. At 300 hPa, the mean retrieval bias is approximately 3 ppbv. For the current study, however, the most important statistic is the difference in retrieval biases for the two dates. Inspection of the figure shows that at all levels, the bias drift over the period of a little more than five years is positive (i.e., biases increase with time) with a maximum drift in the upper troposphere of approximately 5 ppbv. Thus, this simulation study indicates that long-term changes in the instrument cell parameters produce a retrieval bias drift on the order of 1 ppbv/yr in the upper troposphere, and somewhat weaker bias drift in the lower and middle troposphere.
The helpful comments provided by Helen Worden and Avelino Arellano are greatly appreciated. The authors acknowledge for their strong support the European Commission, Airbus and the Airlines (Lufthansa, Austrian, Air France) who carry free of charge the MOZAIC equipment and perform the maintenance since 1994. MOZAIC is supported by INSU-CNRS (Institut National des Sciences de l’Univers – Centre National de la Recherche Scientifique, France), Météo-France, and FZJ (Forschungszentrum Jülich, Germany). This material is based upon work supported by the National Aeronautics and Space Administration under Contract Nos. NNG04GO58G issued by the Earth Observing System Program and NNG06GB27G issued by the Tropospheric Chemistry Program. The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Edited by: H. Singh

References


Acknowledgements. The helpful comments provided by Helen Worden and Avelino Arellano are greatly appreciated. The authors acknowledge for their strong support the European Commission, Airbus and the Airlines (Lufthansa, Austrian, Air France) who carry free of charge the MOZAIC equipment and perform the maintenance since 1994. MOZAIC is supported by INSU-CNRS (Institut National des Sciences de l’Univers – Centre National de la Recherche Scientifique, France), Météo-France, and FZJ (Forschungszentrum Jülich, Germany). This material is based upon work supported by the National Aeronautics and Space Administration under Contract Nos. NNG04GO58G issued by the Earth Observing System Program and NNG06GB27G issued by the Tropospheric Chemistry Program. The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Edited by: H. Singh

Fig. 7. Simulated change in retrieval bias due to drifts in the MOPITT length- and pressure- modulation cells.

5 Conclusions

While the techniques of satellite validation are well understood, and the validation of MOPITT CO retrievals for short time periods has been published previously, the results presented here are new in that they show the MOPITT validation for the full time record. The current operational (V3) retrievals of MOPITT CO show a significant bias in 2006 (19% for column retrievals) determined by validation with in situ measurements from aircraft. Comparison of the MOPITT retrievals to the long-term records of the NOAA/GMD aircraft sampling and the MOZAIC program on commercial aircraft have revealed that the bias in MOPITT CO retrievals has been increasing in time. However, the uncertainty in the biases is too great to determine if there has been a constant rate of increase in the bias. The cause of these trends is not entirely understood, but attempts will be made to reduce the bias and bias trend in the next version (V4) of the MOPITT CO retrievals. Preliminary studies indicate that the V4 retrievals will have less of a drift in the bias, due in part to the use of log(VMR)-based algorithms. In addition, the Forward Model in that version will account for the fact that the modulation cell pressure has dropped over the mission.

While the MOPITT CO record is approaching a length that can be fully quantified, until the next data version has been produced and this effect is taken into consideration. It is therefore premature to use the MOPITT data as the basis for long-term trend analysis until the next data version has been produced and this effect is fully quantified.
L. K. Emmons et al.: MOPITT validation through 2006


