Global carbon monoxide products from combined AIRS, TES and MLS measurements on A-train satellites

J. X. Warner¹, R. Yang¹, Z. Wei¹, F. Carminati¹,5,6, A. Tangborn², Z. Sun³, W. Lahoz⁴,5, J.-L. Attié⁵,6, L. El Amraoui⁵, and B. Duncan⁷

¹Department of Atmospheric & Oceanic Science, University of Maryland, 3433 Computer & Space Sciences Bldg, College Park, MD, USA
²Joint Center for Earth Systems Technology, University of Maryland Baltimore County, Baltimore, MD, USA
³Universities Space Research Association, 10211 Wincopin Circle, Columbia, MD 21044, USA
⁴Norwegian Institute of Air Research, Instituttveien 18, Kjeller 2027, Norway
⁵CNRM-GAME, Météo-France and CNRS, UMR3589, Toulouse, France
⁶Laboratoire d’Aérologie, Université de Toulouse, CNRS, UMR5560, Toulouse, France
⁷Atmospheric Chemistry and Dynamics Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA

Correspondence to: J. X. Warner (juying@atmos.umd.edu)

Received: 6 April 2013 – Published in Atmos. Chem. Phys. Discuss.: 11 June 2013
Revised: 15 October 2013 – Accepted: 19 November 2013 – Published: 3 January 2014

Abstract. This study tests a novel methodology to add value to satellite data sets. This methodology, data fusion, is similar to data assimilation, except that the background model-based field is replaced by a satellite data set, in this case AIRS (Atmospheric Infrared Sounder) carbon monoxide (CO) measurements. The observational information comes from CO measurements with lower spatial coverage than AIRS, namely, from TES (Tropospheric Emission Spectrometer) and MLS (Microwave Limb Sounder). We show that combining these data sets with data fusion uses the higher spectral resolution of TES to extend AIRS CO observational sensitivity to the lower troposphere, a region especially important for air quality studies. We also show that combined CO measurements from AIRS and MLS provide enhanced information in the UTLS (upper troposphere/lower stratosphere) region compared to each product individually. The combined AIRS–TES and AIRS–MLS CO products are validated against DACOM (differential absorption mid-IR diode laser spectrometer) in situ CO measurements from the INTEX-B (Intercontinental Chemical Transport Experiment: MILAGRO and Pacific phases) field campaign and in situ data from HIPPO (HIAPER Pole-to-Pole Observations) flights. The data fusion results show improved sensitivities in the lower and upper troposphere (20–30 % and above 20 %, respectively) as compared with AIRS-only version 5 CO retrievals, and improved daily coverage compared with TES and MLS CO data.

1 Introduction

Atmospheric carbon monoxide (CO) is simultaneously measured by three EOS (Earth Observing System) “A-train” satellite sensors: AIRS (Atmospheric Infrared Sounder) (Aumann et al., 2003) on Aqua, and TES (Tropospheric Emission Spectrometer) (Beer, 2006), and MLS (Microwave Limb Sounder) (Waters et al., 2006) on Aura. Based on the heritage of the A-train system with individual sensors taking measurements within short time intervals between each other (15–30 min), the combined data sets from these sensors are expected to provide three-dimensional composition information that incorporates the uniqueness of AIRS’s large spatial coverage, TES’s high spectral resolution, and MLS’s measurements in the upper troposphere and stratosphere. AIRS’s trace gas retrieval products provide high-density twice-daily global coverage with spatial resolution at 45 km × 45 km at 01:30 and 13:30 local time (LT) overpasses. However, as a nadir sounder, it measures relatively thick integrated layers with most of the sensitivities in the troposphere – and therefore lacks details in the relatively thin UTLS (upper...
troposphere and lower stratosphere) layer. The Aura limb sounder, MLS, on the other hand, possesses the advantage of high vertical resolution in the UTLS region but lacks horizontal coverage. We seek to combine these two types of CO data sets to provide both higher vertical structure and global coverage. This fused data set will benefit studies of stratosphere-troposphere exchange and the chemical and dynamical structures of the UTLS and TTL (tropical tropopause layer) regions as CO is a tracer of transport and has markedly different concentrations in the troposphere and the stratosphere.

A number of radiatively significant gases have strong vertical gradients in composition across the tropopause, including ozone and water vapor. The variability in their tropopause vertical gradients is due to the transport across the tropopause and their sources and sinks in this region. These interactions have a significant impact on chemistry–climate coupling (Holton et al., 1995). The CO–ozone correlation has been used extensively in the UTLS region as a diagnostic tool for studying extratropical tropopause processes, and stratosphere–troposphere exchange and mixing (e.g., Zahn et al., 2000; Hoor et al., 2002, 2004; Pan et al., 2004, 2007a, b). UTLS dynamics and transport studies will greatly benefit from trace gas data sets that show detailed structures of CO plumes, have high-density daily global coverage, and provide information for an extended time period (~10 yr).

A key societal issue is air quality (AQ), which affects human health and has a significant impact on economical and policy issues (Lahoz et al., 2012). Air quality concerns span from the local to the global scales since pollution can be transported across continents and affect AQ at remote locations (Lin et al., 2012). Satellite measurements are among the most effective ways to monitor air pollution and the related transport due to their relatively high spatial and temporal coverage. CO is a good tracer of pollution emission and transport in the troposphere owing to its lifetime of 1–3 months (Crutzen and Zimmermann, 1991; Sze, 1977). In this study, we focus on CO measurements. We show that fusing of CO data from multiple sensors in space can benefit both air quality and climate studies.

In the following sections, we present the data fusion approach that we use to combine AIRS, TES, and MLS CO data sets, with TES and MLS CO as observations and AIRS CO retrievals as the background field. We also show that the combined data set provides a 3-D CO field with daily global coverage that incorporates the sensitivity from AIRS primarily in the mid-troposphere, TES in the lower troposphere, and MLS in the UTLS. Section 2 introduces the data sets that we use in this study, and Sect. 3 describes the data fusion methodology. Section 4 presents results and their validation, while Sect. 5 provides discussion and a summary of this study.

2 Observational data sets

The AIRS instrument on Aqua was launched in 2002 with the primary goal of determining the vertical profiles of temperature and water vapor in the Earth’s atmosphere (Aumann et al., 2003). CO retrievals are obtained from the 2160–2200 cm\(^{-1}\) portion of the spectrum on the edge of the 1-0 vibration–rotation band of CO (McMillan et al., 2005). The AIRS main CO sensitivity is broad and centered in the mid-troposphere between approximately 300 and 600 hPa (Warner et al., 2007). AIRS science team retrieval codes are used to determine the physical parameters measured by AIRS such as temperature, water vapor, and ozone (Susskind et al., 2003).

The TES instrument is an imaging infrared Fourier transform spectrometer with both nadir- and limb-viewing capability covering the spectral range 650–3050 cm\(^{-1}\) at either 0.08 cm\(^{-1}\) or 0.02 cm\(^{-1}\) spectral resolution. TES provides higher spatial and spectral resolutions compared to AIRS; however, the global coverage of TES is significantly lower (Luo et al., 2002, 2007). This is partly due to the fact that, in addition to the sensors’ respective swath widths, AIRS uses a cloud-clearing scheme that recovers, on average, more than half of the cloudy pixels, which increases the global coverage of the retrievals significantly (Susskind et al., 2003; Warner et al., 2013). Validations and intercomparison between AIRS and TES CO indicate that localized biases exist (Warner et al., 2010), which globally and annually generally cancel each other to a large extent.

MLS on Aura satellite was launched on 15 July 2004 and placed into a near-polar Earth orbit at ~705 km with an inclination of 98° and an ascending node at 13:45 LT. It makes about 14 orbits per day. The MLS instrument onboard Aura uses the microwave limb sounding technique to measure chemical constituents and dynamical tracers between the upper troposphere and the lower mesosphere. It provides dense spatial coverage with 3500 profiles daily between 82° N and 82° S. The standard retrievals for CO are given between 215 and 0.46 hPa with a vertical resolution of ~3 km in the upper troposphere and the stratosphere. The along-track resolution of the CO measurements is ~200 km between 215 and 10 hPa. The estimated single-profile precision in the extratropical UTLS is of the order of 0.04 ppmv (parts per million volume) from 215 to 100 hPa and between 0.05 and 0.2 ppmv from 46 to 10 hPa (Livesey et al., 2007; Read et al., 2007; Jiang et al., 2007).

3 Methodology

The data fusion methodology in this study takes a similar approach to that used in data assimilation (Kalnay, 2003) except we use AIRS measurements as the background field and other measurements with lower spatial coverage (in this case, TES and MLS) as the observations. We assume that the AIRS
data represent a snapshot of a three-dimensional geophysical field analogous to a model output. We do not attempt to evaluate any time variability in the CO field, but to simply combine CO of collocated pixels from AIRS, TES, and/or MLS that are part of the A-train satellite constellation and are measured within a short time period (15–30 min). A major difference between this technique and traditional data assimilation is that we do not use a model, which avoids uncertainties arising from model dynamics and transport issues. We follow a formulation identical to the Kalman filter method, commonly used in data assimilation (e.g. Lahoz et al., 2010), except we exclude the transport by model dynamics.

The diagram in Fig. 1 illustrates the data fusion system, where we start with AIRS operational V5 L2 retrievals, shown on the upper left corner, as the background field (H). Other products, such as TES or MLS, are added to the Kalman filter process as observations. We follow a formulation identical to the Kalman filter method, commonly used in data assimilation (e.g. Lahoz et al., 2010), except we exclude the transport by model dynamics.

The background error covariance $P^b$ is the error covariance for the AIRS satellite retrievals and is treated as the standard background error covariance in data assimilation. Satellite observations are often treated as spatially uncorrelated in data assimilation systems, but this is not strictly accurate (Talagrand, 1997). In the application of the data fusion technique, we assume a simple spatially decaying correlation for $P^b$ that can be easily tuned by changing the correlation length scale, and by comparing the results with in situ data.

The background error covariance $P^b$ is constructed from the error correlation matrix between any two locations in the background state vector and the error variances of AIRS retrievals. It is defined as

$$P^b = (P^{1/2})^T CD^{1/2},$$  \hspace{1cm} (3)

where $C$ is the error correlation function and $D$ is a diagonal matrix containing the error variances of the analyzed species (El Amraoui et al., 2004). Following a general approach in data assimilation, we separate the vertical and horizontal correlations (Daley and Barker, 2000; Errera and Ménard, 2012). The spatial correlation matrix is equal to the tensor product of the horizontal correlation matrix $\rho_h$ and the vertical correlation matrix $\rho_v$.

$$X^a = X^b + K(X^o - HX^b),$$ \hspace{1cm} (1)

where the Kalman gain, $K$, is

$$K = P^b H^T (HP^b H^T + R)^{-1},$$ \hspace{1cm} (2)

where $H$ is the linear observation operator which maps the background state into the observational space, and $P^b$ is the background error covariance. The analysis state vector $X^a$ is the new estimate of the state vector, which has the same dimension as the background field $(X^o)$, i.e., the current size of the AIRS vector. The size of the observation vector $X^o$ is determined by the number of assimilated observations (TES or MLS). The matrix $R$ is the observational error covariance. To prescribe $R$, we assumed that there is no global bias in the TES CO measurements, nor a global bias between the AIRS and TES CO measurements. Results from a bias correction experiment where a crude approximation to the global bias between AIRS and TES CO is made (not shown) indicate that the assumption of no bias does not affect the data fusion results adversely. Since the number of variables in the control vector is not prohibitively large and $R$ is assumed diagonal, we compute the inverse matrix in the gain matrix (i.e., $K$) directly. In data assimilation, including that applied to trace gases, it is often a standard assumption to use a diagonal $R$ matrix, which assumes that there is no vertical or horizontal correlation between different observations (see Stajner et al., 2001; Tangborn et al., 2009). We realize that including observation error correlations would have the effect of reducing the relative impact of each individual observation slightly (depending on the correlation), and its omission here does slightly degrade the accuracy of the final data set. However, the purpose of this work is to demonstrate the viability of the data fusion approach to combining satellite data sets, and future enhancements to this work will include a more refined observation error covariance.

When we apply the data fusion algorithm, $P^b$ is the error covariance for the AIRS satellite retrievals and is treated as the standard background error covariance in data assimilation. Satellite observations are often treated as spatially uncorrelated in data assimilation systems, but this is not strictly accurate (Talagrand, 1997). In the application of the data fusion technique, we assume a simple spatially decaying correlation for $P^b$ that can be easily tuned by changing the correlation length scale, and by comparing the results with in situ data.

The background error covariance $P^b$ is constructed from the error correlation matrix between any two locations in the background state vector and the error variances of AIRS retrievals. It is defined as

$$P^b = (P^{1/2})^T CD^{1/2},$$ \hspace{1cm} (3)

where $C$ is the error correlation function and $D$ is a diagonal matrix containing the error variances of the analyzed species (El Amraoui et al., 2004). Following a general approach in data assimilation, we separate the vertical and horizontal correlations (Daley and Barker, 2000; Errera and Ménard, 2012). The spatial correlation matrix is equal to the tensor product of the horizontal correlation matrix $\rho_h$ and the vertical correlation matrix $\rho_v$.

![Fig. 1. Design of the satellite data fusion system, where AIRS operational V5 L2 retrievals are used as the background field, and TES or MLS measurements are added to the Kalman filter process as observations.](image-url)
In terms of matrix elements, the background error covariance matrix $P^b$ is defined by

$$P^b(i, j) = \sigma(i)\sigma(j)\rho_h(i, j)\rho_v(i, j),$$  

where $\sigma_i$ and $\sigma_j$ are the standard deviations of background error at location $i$ and $j$, respectively; $\rho_h$ is the horizontal correlation model; and $\rho_v$ is the vertical correlation model. We use the power law for the $\rho_h$ model as shown in Fig. 2, left panel (Dee and da Silva, 1999):

$$\rho_h(i, j) = \left(1 + \frac{1}{2} \left(\frac{r_{i,j}}{L_0}\right)^2\right)^{-1},$$  

where $r_{i,j}$ denote the horizontal distance between the two locations $(i, j)$, and $L_0$ is the prescribed horizontal error correlation length.

The vertical resolution of the analysis for the combined AIRS and TES CO is based on the AIRS pressure grid in the troposphere, whereas the vertical resolution of the analysis for the combined AIRS and MLS CO is based on the AIRS pressure grid in the stratosphere and the UTLS regions. Table 1 lists all the pressure levels used in the AIRS and TES combined CO experiments (left columns) and the AIRS and MLS combined CO experiments (right columns). These analyses’ pressure levels are fixed except for near the surface. The vertical correlation model for AIRS and TES is shown in Fig. 2 right panel and has the form

$$\rho_v(l, n) = \exp(-\Delta z_{l,n}^2),$$  

where

$$\Delta z_{l,n} = \frac{\ln p_n - \ln p_l}{L_z},$$  

$p_n$ and $p_l$ are the pressure values at levels $n$ and $l$, and $L_z$ is the prescribed vertical error correlation length (Daley and Barker, 2000; Errera and Ménard, 2012).

In the examples given in Fig. 2, the horizontal correlation length $L_0 = 750$ km and the vertical correlation length $L_z = 0.17$ were chosen to match the measurement characteristics and sensitivities of AIRS and TES. Specifically, we use a horizontal correlation length of 750 km so that the entire AIRS swaths are influenced by TES measurements from AIRS nadir points considering AIRS swaths are approximately 1600 km wide. The vertical correlation length of 0.17 is chosen to match the number of vertical levels correlated as shown in Fig. 2 right panel. Figure 3 shows the prescribed error $1\sigma$ standard deviations (SDs) for AIRS (solid blue line), as the background error, and for TES (dashed red line), as the observational error. These errors originated from the retrieval error covariance matrices from the respective teams and were averaged and/or smoothed to form one global set of error profiles (Dee and da Silva, 1999). Three-dimensional retrieval errors are difficult to quantify due to the limited amount of in situ measurements in both temporal and spatial dimensions. The retrieval errors distributed by the individual teams (AIRS and TES in this case) are defined and computed differently, which means they cannot be used together. We used the statistics of the validation summaries from various studies to derive one global set of error profiles (Luo et al., 2007; McMillan et al., 2011; Warner et al., 2007, 2010; Yurganov et al., 2008) to reduce prescribed errors for AIRS in the mid-troposphere and for TES in the lower troposphere. Additionally, we considered the statistics of the averaging kernels (AKs) from AIRS and TES to weight the relative importance of the data in the vertical dimension. We have
Table 1. The pressure levels used in combined AIRS (background and analysis) and TES (observations) CO (left columns), and combined AIRS (background and analysis) and MLS (observations) CO (right columns).

<table>
<thead>
<tr>
<th>Combined AIRS and TES CO</th>
<th>Combined AIRS and MLS CO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Background</td>
<td>Observations</td>
</tr>
<tr>
<td>1085.39343</td>
<td>1000.00</td>
</tr>
<tr>
<td>1056.51999</td>
<td>980.00</td>
</tr>
<tr>
<td>1028.04346</td>
<td>970.00</td>
</tr>
<tr>
<td>999.94598</td>
<td>941.20</td>
</tr>
<tr>
<td>972.26398</td>
<td>908.51</td>
</tr>
<tr>
<td>944.99750</td>
<td>825.40</td>
</tr>
<tr>
<td>918.12610</td>
<td>749.89</td>
</tr>
<tr>
<td>891.68030</td>
<td>681.29</td>
</tr>
<tr>
<td>865.63971</td>
<td>618.96</td>
</tr>
<tr>
<td>840.01459</td>
<td>562.34</td>
</tr>
<tr>
<td>814.80499</td>
<td>510.89</td>
</tr>
<tr>
<td>790.02087</td>
<td>464.16</td>
</tr>
<tr>
<td>765.64209</td>
<td>421.69</td>
</tr>
<tr>
<td>741.68878</td>
<td>383.11</td>
</tr>
<tr>
<td>718.16119</td>
<td>348.06</td>
</tr>
<tr>
<td>695.04901</td>
<td>316.22</td>
</tr>
<tr>
<td>672.36230</td>
<td>287.29</td>
</tr>
<tr>
<td>650.10120</td>
<td>261.01</td>
</tr>
<tr>
<td>628.26569</td>
<td>237.13</td>
</tr>
<tr>
<td>606.84552</td>
<td>215.44</td>
</tr>
<tr>
<td>585.85107</td>
<td>195.73</td>
</tr>
<tr>
<td>565.28198</td>
<td>177.82</td>
</tr>
<tr>
<td>545.13861</td>
<td>161.56</td>
</tr>
<tr>
<td>525.42078</td>
<td>146.77</td>
</tr>
<tr>
<td>506.11841</td>
<td>133.35</td>
</tr>
<tr>
<td>487.23141</td>
<td>121.15</td>
</tr>
<tr>
<td>468.78009</td>
<td>110.06</td>
</tr>
<tr>
<td>450.73410</td>
<td>100.00</td>
</tr>
<tr>
<td>433.11371</td>
<td>91.7866</td>
</tr>
<tr>
<td>415.91891</td>
<td>82.9380</td>
</tr>
<tr>
<td>399.12930</td>
<td>74.9984</td>
</tr>
<tr>
<td>382.75519</td>
<td>67.2314</td>
</tr>
<tr>
<td>366.78641</td>
<td>59.8751</td>
</tr>
<tr>
<td>351.23300</td>
<td>52.4078</td>
</tr>
<tr>
<td>336.09500</td>
<td>45.1129</td>
</tr>
<tr>
<td>321.35220</td>
<td>38.2539</td>
</tr>
<tr>
<td>307.01471</td>
<td>31.5320</td>
</tr>
<tr>
<td>293.07239</td>
<td>25.0000</td>
</tr>
<tr>
<td>279.53540</td>
<td>20.0000</td>
</tr>
<tr>
<td>266.39359</td>
<td>15.0000</td>
</tr>
<tr>
<td>253.63670</td>
<td>10.0000</td>
</tr>
<tr>
<td>241.26500</td>
<td>5.0000</td>
</tr>
<tr>
<td>229.28830</td>
<td>0.0000</td>
</tr>
<tr>
<td>217.68660</td>
<td>0.0000</td>
</tr>
<tr>
<td>206.45979</td>
<td>0.0000</td>
</tr>
<tr>
<td>195.60789</td>
<td>0.0000</td>
</tr>
<tr>
<td>185.12080</td>
<td>0.0000</td>
</tr>
<tr>
<td>174.99840</td>
<td>0.0000</td>
</tr>
<tr>
<td>165.24080</td>
<td>0.0000</td>
</tr>
<tr>
<td>155.83780</td>
<td>0.0000</td>
</tr>
<tr>
<td>146.77940</td>
<td>0.0000</td>
</tr>
<tr>
<td>138.07561</td>
<td>0.0000</td>
</tr>
<tr>
<td>129.70610</td>
<td>0.0000</td>
</tr>
<tr>
<td>121.67110</td>
<td>0.0000</td>
</tr>
<tr>
<td>113.97040</td>
<td>0.0000</td>
</tr>
<tr>
<td>106.58380</td>
<td>0.0000</td>
</tr>
<tr>
<td>99.52550</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

![Fig. 3](https://www.atmos-chem-phys.net/14/103/2014/)

Fig. 3. Prescribed 1σ standard deviation of the CO error for AIRS (solid blue line) and TES (dashed red line). found that constructing the global error profiles from our experience of validation and the retrieval AKs provides a better estimate than the errors distributed by the retrieval teams. Although we used 1-D background error SDs as an approximation for this study, the extension to 2-D error fields and the inclusion of seasonal variations is possible although not trivial, and will be the subject of a future study.

4 Results

4.1 Combined AIRS and TES CO

We show in this section results from applying the data fusion methodology to AIRS and TES CO data. AIRS V5 L2 and TES V3 CO volume mixing ratios (VMRs) at 500 hPa for 4 March 2006 are shown in Fig. 4 to demonstrate the data coverage pattern, where the squares represent AIRS footprints and the filled circles represent TES pixels. TES step and stare (SS) special observations (SO) were taken to coincide with the INTEX-B experiment (Singh et al., 2009) (http://www.espo.nasa.gov/intex-b). The pixel sizes in Fig. 4 do not represent the correct proportions of the sizes for the AIRS field of regards (FORs) at approximately 45 km × 45 km and the TES field of views (FOVs) at 5 km × 8 km.

The vast majority of the AIRS pixels are not at nadir nor collocated with TES due to the coverage differences between the two sensors, so the data fusion scheme described in Sect. 3 is needed to distribute TES narrow swaths onto
the global coverage of AIRS to produce daily maps. Results show that in vertical regions where AIRS has lower measurement sensitivity, i.e., the lower and upper troposphere, the retrievals can still provide sufficient spatial variability (or patterns), even though they may not be able to produce the correct magnitude (see Fig. 4). Consequently, the TES CO profiles are horizontally distributed based on the AIRS variances.

We first demonstrate the overall performance of data fusion applied to global AIRS and TES CO data using the methodology described in Sect. 3. The monthly grid-averaged zonal means for March 2006 are shown in Fig. 5 for AIRS CO (left top panel), the analyzed CO (left middle panel), the analysis increment (i.e., the analysis minus background) (left bottom panel), and AIRS CO interpolated to TES locations (right top panel), TES CO (right middle panel), and TES minus AIRS (TMA) (right bottom panel). CO data are averaged in $1^\circ \times 1^\circ$ latitude and longitude grid boxes. The CO data range of the zonal mean analysis has increased significantly, especially in the lower troposphere, compared with the background AIRS CO values shown by the left column top two panels, as mirrored by the differences between TES and AIRS on the right. This is an improvement over AIRS V5 CO products based on our previous experience.

The horizontal CO distribution at approximately 944 hPa from a monthly mean grid-averaged data set for March 2006 is shown in Fig. 6 with TES CO (top panel), AIRS CO (middle panel), and the analyzed CO (bottom panel). The analysis from the data fusion scheme presents much more realistic CO values based on our previous knowledge and validation results compared with the background CO values from AIRS. In particular, the CO analyses show distinct plumes in the NH associated with emissions and the SH shows the expected relatively low values of CO. The analyzed CO agrees better with TES CO than with AIRS CO, with higher CO values in the NH and lower values in the SH compared to AIRS, and it provides much better continuity in data coverage compared with TES CO values. Similarly, Fig. 7 shows the monthly mean grid-averaged CO distribution at approximately 506 hPa. Unlike at 944 hPa, the analyzed CO distribution is very similar to the AIRS background CO. Note that these examples use monthly mean results to demonstrate the stability of the technique over time, and the main advantage
of this technique is the availability of daily maps that matches AIRS coverage, whereas there are only a few narrow tracks per day from TES routine survey products.

For validation purposes, we have collocated the TES and AIRS footprints with the independent INTEX-B in situ measurements by selecting the retrievals from the nearest pixels. Figure 8 shows the curtain plots for CO VMRs along TES footprints for the DACOM CO measurements (top panel), TES V3 CO retrievals (second panel), AIRS V5 L2 CO retrievals (third panel), and the combined CO VMRs produced by data fusion (bottom panel). The geographical region measured on the DC-8 plane on that particular day was affected by agricultural fires in the southeast US, which resulted in high CO values in the lower troposphere as shown by DACOM CO. AIRS V5 retrievals, using a single profile as the first guess, show good sensitivity in the mid-troposphere (300–600 hPa), but the magnitudes of the retrieved results are lower than DACOM CO by as much as ~40 ppbv below the 900 hPa level, in what is typically the boundary layer, due to reduced measurement sensitivity. TES senses higher CO signals in the lower troposphere especially over land (north of 30° N); however, TES shows lower CO values between 200 and 500 hPa over parts of the ocean by ~10–30 ppbv. Large portions of TES measurements show similar features as observed by DACOM in the lower troposphere where the AIRS observational signals are relatively weak. As shown in the bottom panel, the mid-troposphere CO concentrations resemble AIRS retrievals in the mid-troposphere while following the lower-tropospheric CO values from TES. The combined products show higher CO VMRs in the lower troposphere, which improves the agreement with the in situ measurements to within 10–20 ppbv (see red color pixels in Fig. 8 bottom panel vs. top panel).

A profile validation against the DACOM in situ CO measurements is presented in Fig. 9. The DACOM in situ CO spiral profile used in this comparison was obtained during INTEX-B on 4 March 2006, near the center of the Gulf of Mexico, as denoted by a black square in Fig. 4 (Sachse et al., 1987). The CO VMR profiles are shown on the left panel in Fig. 9 as the in situ data (green), AIRS data (cyan), TES data (blue), and the combined AIRS and TES data (red). The combined profile agrees with the in situ profile much better than AIRS and TES profiles individually in the troposphere, especially over the vertical regions below 400 hPa. The combined profile is enhanced by AIRS information in the mid-troposphere by approximately 10–15 ppbv compared to TES, and by TES information below 750 hPa by nearly 20 ppbv. This shows an advantage of the data fusion method, in that it
combines the best information characteristics from each sensor. The fact that the combined CO profile agrees with the in situ data better than both AIRS and TES is largely due to the effect of the influence from neighboring pixels and from different height regions. The center panel shows the AIRS AKs, which indicate that the AIRS measurement information is primarily in the mid-troposphere, peaking at approximately 500 hPa. The right panel shows TES AKs at three selected retrieval pressure levels (825, 510, and 215 hPa), indicating that, for this example, the maximum information responsible for the lower-troposphere CO retrievals peaks at levels just above 900 hPa where the top of the boundary layer is usually located (see blue line in right hand panel of Fig. 9). Although the final form of the AKs for the combined CO profiles, a nonlinear summation of the AKs from AIRS and TES, is not presented here, the combined information content should increase in the lower atmosphere based on the contribution from TES.

To extend our validation efforts, we compared the data fusion output with recently available in situ measurements from the HIAPER Pole-to-Pole Observations (HIPPO) experiment, which took place from January of 2009 to September 2011 (Wofsy et al., 2012). The merged 10 s best CO measurements are used (http://hippo.ornl.gov/dataaccess) (HIPPO_Merged_Data_Users_Guide_20121130). TES CO data are available for HIPPO-1 (8–30 January 2009) and HIPPO-2 (31 October–22 November 2009) flights, but not for HIPPO-3 (24 March–16 April 2010), HIPPO-4 (14 June–11 July 2011), and HIPPO-5 (9 August–9 September 2011) periods. We initially carried out this data fusion study when only TES V3 CO products were available, which apply to all results presented in earlier sections. Although newer TES CO products are available, the principle and advantages of the data fusion methodology still apply using TES V3 CO. We attempt, however, to demonstrate the full potential of the technique and use the latest TES V5 CO in the HIPPO validation.

Figure 10 shows a summary of the differences between the combined CO from AIRS and TES and CO from both HIPPO-1 and HIPPO-2 flights. The averaged differences (HIPPO − retrievals) for all available cases are plotted at 41 vertical levels throughout the troposphere for AIRS (blue), TES (cyan), and for the combined product (red), respectively. Related error bars are computed from the 1σ standard deviations of the differences between the observations and the retrievals. The CO field errors below 700 hPa are reduced by more than half, from 19.9 ppbv to 8.7 ppbv at 766 hPa and from 33.4 ppbv to 11.9 ppbv at 918 hPa. In the mid-troposphere (350–600 hPa), TES average errors (10–14.3 ppbv) are twice as high as for AIRS (−1.1 to −5.4 ppbv) in magnitude with similar SDs. Note that the errors are defined by HIPPO CO minus the retrievals, so the negative values indicate an overestimate by the retrievals. The errors for both AIRS and the combined AIRS and TES are below 5 ppbv, with AIRS-retrieved CO lower than the HIPPO measurements and the combined CO slightly higher than HIPPO.
The differences between AIRS (blue), TES (cyan), and the combined CO (red) and CO from both HIPPO-1 and HIPPO-2 flights at 41 vertical levels. Related error bars are computed from the 1σ standard deviations of the differences between the observations and the retrievals. Positive differences indicate that the data fusion approach works well such that the lower-troposphere analyzed CO compared to AIRS CO and mid-troposphere analyzed CO compared to TES CO have errors that have been reduced by more than half in both cases.

CO measurements. The 1σ SDs are approximately the same for AIRS CO and the combined AIRS and TES CO at ∼15–20 ppbv in the mid-troposphere, and reduced in the lower troposphere from ∼20–30 ppbv to ∼10–20 ppbv, with the combined product having the lower error.

4.2 Combined AIRS and MLS CO retrievals

MLS CO retrievals are combined with AIRS CO retrievals using the same method as for AIRS and TES described in Sect. 3. Livesey et al. (2007) describe the error characteristics of these retrievals, and recommend that data not be used below the 215 hPa vertical level. Although we are using a newer version (V3.3) of MLS data, based on the errors provided by MLS team (http://mls.jpl.nasa.gov/data/v3-3_data_quality_document.pdf), we choose to only use data above 215 hPa. We modified the data fusion methodology by using gridded data for AIRS and MLS as opposed to using L2 individual pixels in the case of AIRS and TES. AIRS CO retrievals are averaged into 3° × 3° grids and MLS CO retrievals are averaged into 6° × 6° grids, in both cases separated into daytime and nighttime measurements. This is done to reduce noise in the MLS CO data and to take into account that the MLS pixel represents a relatively long horizontal path due to limb viewing.

Figure 11 shows the global CO retrieval comparisons at 147 hPa for 4 March 2006 daytime-only tracks, with AIRS V5 CO VMRs over-plotted by MLS CO VMRs (ppbv) in the upper panel and the combined AIRS and MLS CO VMRs in the lower panel. The combined AIRS and MLS CO fields show a large increase in their variability, with higher CO at low latitudes by approximately 10 to 20 ppbv (∼20%) or higher, and lower CO at mid- and high latitudes (by ∼10 ppbv), in both cases compared with the AIRS retrievals. Higher CO concentrations are identified as three large plumes over the tropics that cannot be measured by any nadir instrument; these features are expected due to the high tropopause and strong convection in the tropics. The decrease in CO values in the NH is also expected considering that the retrieved CO from a nadir-sensing instrument may be affected by the higher tropospheric CO signals. It is difficult to resolve vertically the location of the high CO concentrations with thermal nadir-viewing instruments, since the weighting functions for CO are generally broad and have very few (1–2) pieces of information in the troposphere.

Figure 12 shows the validation of the combined AIRS and MLS CO in the UTLS using in situ data from HIPPO flights. Measurements from all 5 HIPPO flights are summarized into V5 CO VMRs over-plotted by MLS CO VMRs in the upper panel and the combined AIRS and MLS CO VMRs in the lower panel.
the difference plot between HIPPO CO and the satellite products (HIPPO minus retrievals), shown as the blue curve for AIRS V5 CO, cyan curve for MLS CO, and the red curve for the combined AIRS and MLS CO in Fig. 12. The solid lines are the average errors showing 5 to 10 ppbv improvements for the height levels above 220 hPa over using AIRS data alone. The SDs have been improved by approximately 5 ppbv, from 20–25 ppbv to 15–20 ppbv above 220 hPa. The MLS average error is very low (∼2.1 ppbv) at 146.8 hPa, and it provides CO information to the combined product, which is an improvement over AIRS alone. Meanwhile, the SDs are reduced from 18.9 ppbv to 16.7 ppbv compared to the MLS-only CO. The MLS CO average error is relatively high (∼17.8 ppbv) (CO retrievals greater than HIPPO CO), with SDs at 36.0 ppbv at pressure level 215 hPa, which is given less weight in the data fusion of AIRS and MLS. The improvement of SDs using the combined CO at 217.7 hPa over MLS CO is significant from 36.0 ppbv to 19.8 ppbv.

5 Summary

Multiple sensors (AIRS, TES, and MLS) on the A-train system make correlative CO measurements within a short time period (15–30 min). Based on differences between their measurement techniques, each sensor provides unique characteristics. This study seeks to combine these data sets without using a chemistry transport model, an approach that differs from traditional data assimilation. This approach, called data fusion, avoids uncertainties arising from model dynam-ics and transport issues, and provides a new set of measurements only from satellite sensors in that the analysis is not dependent on model information. Results from this study have shown that the technique works well, generally providing improved CO products compared with those from each individual sensor. The combined AIRS and TES CO products showed improvements of 10 to 40 ppbv, and the combined AIRS and MLS CO products showed improvements of 10 to 20 ppbv above 175 hPa, based on validation using INTEX-B and HIPPO in situ measurements. The combined data sets retain the twice-daily nearly global coverage due to AIRS’s wide swaths and the implementation of a cloud-clearing algorithm that reconstructs cloudy pixels to recover more than 50 % of the daily measurements, compared to less than 10 % clear-sky coverage.

This study provides a new framework to combine different measurements without the use of a model. We have demonstrated the feasibility of the data fusion technique using A-train measurements from AIRS, TES, and MLS. However, this technique can be easily adapted to include direct and in situ measurements and/or other satellite products, and take advantage of the mature approaches developed in the field of data assimilation. Further studies will examine in detail error estimation and the constraints provided by satellite products. Data fusion for other atmospheric species such as ozone and water vapor using this technique is also possible and should benefit chemical and dynamical studies of the UTLS.

Acknowledgements. This study was partially supported by the NASA Atmospheric Composition Program (NNX07AM45G), Climate Data Record Uncertainty Analysis (NNX11AL22A), and Modeling, Analysis and Prediction Program. We have also been partially supported by the RTRA/STAE foundation from Toulouse, France. We acknowledge the AIRS, TES, and MLS science teams for the satellite products used in this study. We also thank the NASA INTEX-B and NSF and NOAA HIPPO science teams for providing high-quality in situ measurements.

Edited by: Q. Errera

References


J. X. Warner et al.: Global carbon monoxide products

Fig. 12. Validation of the combined AIRS and MLS CO VMRs (ppbv) (red lines) in the UTLS regions using HIPPO (HIPPO retrievals) with 1σ SDs, compared with AIRS CO VMRs (ppbv) (blue lines). Positive differences indicate that the data fusion approach works well in the UTLS regions with an error reduction of approximately half at most levels.
J. X. Warner et al.: Global carbon monoxide products


