Intercomparison methods for satellite measurements of atmospheric composition: application to tropospheric ozone from TES and OMI

L. Zhang1, D. J. Jacob1,2, X. Liu3,4,5, J. A. Logan2, K. Chance4, A. Eldering6, and B. R. Bojkov7

1Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA 02138, USA
2School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA
3Goddard Earth Sciences and Technology Center, University of Maryland, Baltimore County, Baltimore, Maryland, USA
4Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA
5NASA Goddard Space Flight Center, Greenbelt, Maryland, USA
6Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA
7European Space Agency (ESA/ESRIN), Science, Applications and Future Technologies Department, Via Galileo Galilei, Casella Postale 64, 00044 Frascati (RM), Italy

Received: 27 November 2009 – Published in Atmos. Chem. Phys. Discuss.: 20 January 2010
Revised: 13 May 2010 – Accepted: 17 May 2010 – Published: 26 May 2010

Abstract. We analyze the theoretical basis of three different methods to validate and intercompare satellite measurements of atmospheric composition, and apply them to tropospheric ozone retrievals from the Tropospheric Emission Spectrometer (TES) and the Ozone Monitoring Instrument (OMI). The first method (in situ method) uses in situ vertical profiles for absolute instrument validation; it is limited by the sparseness of in situ data. The second method (CTM method) uses a chemical transport model (CTM) as an intercomparison platform; it provides a globally complete intercomparison with relatively small noise from model error. The third method (averaging kernel smoothing method) involves smoothing the retrieved profile from one instrument with the averaging kernel matrix of the other; it also provides a global intercomparison but dampens the actual difference between instruments and adds noise from the a priori. We apply the three methods to a full year (2006) of TES and OMI data. Comparison with in situ data from ozonesondes shows mean positive biases of 5.3 parts per billion volume (ppbv) (10%) for TES and 2.8 ppbv (5%) for OMI at 500 hPa. We show that the CTM method (using the GEOS-Chem CTM) closely approximates results from the in situ method while providing global coverage. It reveals that differences between TES and OMI are generally less than 10 ppbv (18%), except at northern mid-latitudes in summer and over tropical continents. The CTM method further allows for CTM evaluation using both satellite observations. We thus find that GEOS-Chem underestimates tropospheric ozone in the tropics due to possible underestimates of biomass burning, soil, and lightning emissions. It overestimates ozone in the northern subtropics and southern mid-latitudes, likely because of excessive stratospheric influx of ozone.

1 Introduction

Tropospheric ozone is of environmental importance as a surface pollutant, a precursor of the hydroxyl radical (OH) oxidant, and an effective greenhouse gas. It is produced by photochemical oxidation of CO and volatile organic compounds (VOCs) in the presence of nitrogen oxides (NOx ≡ NO+NO2). This photochemical production dominates over stratospheric ozone influx on a global scale (Prather and Ehhalt, 2001). Ozone concentrations can vary from less than 10 parts per billion volume (ppbv) in clean surface air to over 100 ppbv in the upper troposphere and in polluted regions (Logan, 1999). Satellite observations of tropospheric ozone and its precursors are providing a growing
resource to better understand the processes that control ozone levels and the effect of human influence (NRC, 2008).

One difficulty in measuring tropospheric ozone from space is separating it from stratospheric ozone, which accounts on average for 90% of the total ozone column. The first approach to derive global tropospheric ozone from satellite measurements involved subtracting the stratospheric ozone column measured in the limb from the total ozone column measured independently in the nadir (Fishman and Larsen, 1987; Fishman et al., 1990; Ziemke et al., 1998, 2005). This approach has been refined and extended in recent years (Ziemke et al., 2006; Schoeberl et al., 2007).

Direct retrieval of global tropospheric ozone distributions from solar backscattered UV spectra was reported by Liu et al. (2005, 2006) for the Global Ozone and Monitoring Experiment (GOME) and more recently by Liu et al. (2010a) for the Ozone Monitoring Instrument (OMI) on the EOS Aura satellite launched in July 2004. Tropospheric ozone has also been directly retrieved from nadir measurements of thermal infrared (IR) emission in and around the 9.6 µm absorption band at high spectral resolution. This was first done for the Interferometric Monitor Greenhouse gases (IMG) instrument on the ADEOS platform, which operated for 10 months in 1996 (Turquety et al., 2002). A multi-year record (2004) of thermal IR observations of tropospheric ozone is available from the Tropospheric Emission Spectrometer (TES) also onboard Aura (Beer, 2006). Tropospheric ozone retrievals in the thermal IR have also been made from the Atmospheric Infrared Sounder (AIRS) on Aqua launched in 2002 (Aumann et al., 2003) and the Infrared Atmospheric Sounding Interferometer (IASI) on MetOp-A launched in October 2006 (Boynard et al., 2009).

The direct retrievals of tropospheric ozone from TES and OMI have generated great interest for better understanding the processes controlling ozone concentrations and testing chemical transport models (CTMs) (Liu et al., 2006; Zhang et al., 2006; Parrington et al., 2008). The reliability of the data is an issue. Both TES and OMI have been validated with ozonesonde and aircraft measurements (Nassar et al., 2008; Richards et al., 2008; Boxe et al., 2009; Liu et al., 2010b), but these in-situ measurements are very sparse. The quality of the satellite ozone retrieval depends on viewing angle, surface type, vertical structure of ozone and temperature, cloud and aerosol interferences, and other factors, requiring a greater validation space than can be achieved from in situ data alone.

One approach to extend the satellite validation to a global scale is by comparison to a CTM that has been independently evaluated with accurate in situ measurements. Validation is measured by comparison of the CTM bias against the satellite data and against the in situ data. Assimilation of satellite tropospheric ozone observations into CTMs also provides an indirect validation through comparison with independent measurements (Geer et al., 2006; Parrington et al., 2008). The CTM can further serve as a common platform to intercompare measurements from different satellite instruments with different viewing scenes and vertical sensitivities (often measured by averaging kernel matrices). Aside from validation, using the CTM as a common intercomparison platform tests the consistency of the multi-instrument datasets for CTM evaluation and thus enables better diagnostics of CTM biases. Kopacz et al. (2010) used the GEOS-Chem CTM to test the consistency of multiple satellite CO data sets. Here we investigate the theoretical basis of the CTM intercomparison method with satellite retrievals of tropospheric ozone.

Global intercomparisons of satellite tropospheric ozone profiles have not been reported in the literature so far. Rodgers and Connor (2003) presented a general method to compare measurements from two satellite instruments with different averaging kernels, by smoothing the retrievals from the instrument with higher vertical resolution using the averaging kernels of the instrument with lower vertical resolution. The method was applied by Luo et al. (2007) and Ho et al. (2009) to compare retrieved CO profiles from TES and the Measurement Of Pollution In The Troposphere (MOPITT) instrument. As we will see, it cannot be used to diagnose biases between two instruments when the vertical sensitivities of both instruments are weak. We will show that using a CTM as an intercomparison platform is a far more accurate method.

We analyze here the theoretical basis for the above three different methods to validate and intercompare satellite retrievals on a global scale, focusing on tropospheric ozone measurements from TES and OMI both on the Aura satellite: (1) the in situ method (here using ozonesondes), (2) the CTM method and (3) the averaging kernel smoothing method of Rodgers and Connor (2003). We apply and compare the three methods for a full year (2006) of TES and OMI data, and also use 2005–2007 ozonesonde data to better constrain the in situ validation. We show how the different methods provide different results, and discuss the value of the CTM method as a versatile and accurate tool for instrument validation and intercomparison as well as CTM evaluation.

2 TES and OMI ozone profile retrievals

The Tropospheric Emission Spectrometer (TES) and the Ozone Monitoring Instrument (OMI) are both on the EOS Aura satellite launched in July 2004 into a polar, sun-synchronous orbit with an ascending equator crossing time of ~13:45. TES is a Fourier transform IR spectrometer with high spectral resolution (0.1 cm⁻¹ apodized in nadir) and a wide spectral range (650–3050 cm⁻¹) (Beer, 2006). The standard products of TES (“global surveys”) consist of 16 daily orbits of nadir-viewing measurements with a spatial resolution of 5x8 km² spaced 1.6° along the orbit track every other day. Global coverage is achieved in 16 days. We use TES V003 data. Nassar et al. (2008) and Richards
et al. (2008) presented the validation of TES V002 ozone retrievals with ozonesondes globally and with aircraft data over the Pacific, and they found a high bias of 3–10 ppbv. The TES V003 data have a similar positive bias (Boxe et al., 2009; also see Sect. 4.1).

OMI is a nadir-scanning instrument that measures backscattered solar radiation over the 270–500 nm wavelength range with a spectral resolution of 0.42–0.63 nm (Levelt et al., 2006). OMI has a large field of view of 114° with a swath of 2600 km. It has a spatial resolution of 13×24 km² at nadir and daily global coverage. We use the direct retrieval of ozone profiles developed by Liu et al. (2010a). Validation with ozonesondes showed that OMI has a positive bias of 5–10% in the troposphere (Liu et al., 2010b).

We use a full year of TES and OMI data for 2006, including 169 TES global surveys. We exclude data poleward of 60° where satellite sensitivities are weak due to low brightness temperature for TES and high solar zenith angle for OMI. OMI has daily global coverage; we select OMI observations along the TES sampling locations for comparison. We remove cloudy observations and other erroneous data following the TES ozone data quality flag as defined in Osterman et al. (2009). TES V003 ozone data include some unphysical retrievals with anomalously high ozone near the surface and anomalously low ozone in the middle troposphere (“C-curve” shape). We have developed an additional flag to filter these profiles (also described in Osterman et al., 2009). For OMI, we remove cloudy observations (effective cloud fraction >30%).

Ozone retrievals from TES and OMI are both based on the optimal estimation method of Rodgers (2000), as described by Bowman et al. (2006) and Liu et al. (2010a), respectively. The true vertical profile of concentrations is represented as a vector \( \hat{x} \) whose elements are the constituent amounts, either as mixing ratios or partial columns at different vertical levels. The retrieval algorithm minimizes the differences between observed and simulated radiances subject to constraints from a priori profiles \( x_a \). The retrieved vertical profile \( \hat{x} \) can be expressed as a linear combination of \( x \) and \( x_a \),

\[
\hat{x} = A x + (I - A)x_a + \varepsilon + b
\]  

(1)

where \( A \) is the averaging kernel matrix that describes the vertical sensitivity of the retrieval to the true profiles. It is determined by the instrument characteristics and by the a priori error covariance matrix. The term \( \varepsilon+b \) is the retrieval error including errors in the measurement and in the radiative transfer model. We separate it into a random noise (\( \varepsilon \)) and a systematic bias (\( b \)). TES retrieves natural logarithms of ozone volume mixing ratios (VMR) in ppbv at 67 pressure levels up to 0.1 hPa, while OMI retrieves partial ozone columns in Dobson Units (DU) for 24 layers with thicknesses of approximately 2.5 km.

Direct comparison of TES and OMI ozone profile retrievals is not appropriate because they have different averaging kernels (\( A \)) and use different a priori profiles (\( x_a \)). TES a priori profiles are monthly mean MOZART CTM values (Brasseur et al., 1998) averaged over a 10° latitude × 60° longitude grid (Bowman et al., 2006). OMI a priori profiles are based on a latitude- and month-dependent ozone profile climatology (McPeters et al., 2007) derived from 15 years of ozonesonde and Stratospheric Aerosol and Gas Experiment (SAGE) data (Liu et al., 2010a). To remove the discrepancy from the use of different a priori profiles, we reprocess the retrievals to a common fixed a priori (\( x_c \)) (Rodgers and Connor, 2003):

\[
\hat{x}' = \hat{x} + (A - I)(x_a - x_c)
\]  

(2)

We choose \( x_c \) by averaging the original OMI a priori profiles within the latitude band of 30° S–30° N and applying it to all TES and OMI retrievals. Results from Eq. (2) are very similar to retrievals with the fixed a priori for both TES (Kulawik et al., 2008) and OMI because the retrieval is only weakly non-linear. We remove the prime henceforth for simplicity of notation and refer to \( \hat{x} \) as the retrieved vertical profile reprocessed as per Eq. (2).

Intercomparison between satellite instruments is much easier to analyze mathematically if the retrieved vertical profiles \( \hat{x} \) have the same dimensions and units. To achieve this we interpolate TES retrievals on the OMI pressure grid, and convert the TES log(VMR)-based averaging kernels to partial columns on the OMI pressure grid, as described in Appendix A. In this manner, the reprocessed TES (\( \hat{x}_{\text{TES}} \)) and OMI (\( \hat{x}_{\text{OMI}} \)) retrievals are both partial ozone column profiles on the OMI pressure grid:

\[
\hat{x}_{\text{TES}} = A_{\text{TES}} x + (I - A_{\text{TES}})x_c + \varepsilon_{\text{TES}} + b_{\text{TES}}
\]  

(3)

\[
\hat{x}_{\text{OMI}} = A_{\text{OMI}} x + (I - A_{\text{OMI}})x_c + \varepsilon_{\text{OMI}} + b_{\text{OMI}}
\]  

(4)

with all terms on the right hand side of Eqs. (3) and (4) also computed on the OMI pressure grid.

Figure 1 shows sample averaging kernel matrices for TES and OMI. The trace of the matrix gives the number of independent columns on the OMI pressure grid, and does not significantly modify the structure, as shown in the central panel of Fig. 1. The converted TES averaging kernel matrix \( A_{\text{TES}} \) can be compared directly with the OMI
Fig. 1. Sample averaging kernel matrices for TES (left) and OMI (right) ozone retrievals below 100 hPa (cloud-free ocean scene at 28° N, 58° W on 6 August 2006). Each line is a row of the averaging kernel matrix; row elements are dimensionless. The central panel shows TES averaging kernels recomputed as sensitivities to partial ozone columns on the OMI pressure level grid as described in Appendix A. The colored numbers are centers of the OMI pressure grid (hPa). Red colors represent relatively high pressure levels and blue colors represent relatively low pressure levels. The degrees of freedom for signal (DOFS) in the troposphere are shown inset. The horizontal dotted line shows the tropopause pressure (120 hPa) used for this retrieval.

Fig. 2. Mean tropospheric DOFS for TES (left) and OMI (right) retrievals of tropospheric ozone in January (top) and July 2006 (bottom). The data are averaged on a 4° × 5° grid. Red colors indicate relatively high DOFS, and blue colors indicate relatively low DOFS. White areas indicate lack of data meeting the retrieval criteria described in the text.

averaging kernel matrix $A_{\text{OMI}}$. The OMI averaging kernel matrix shows weaker sensitivity than TES with DOFS=1.0 in the troposphere, although this is partly due to a weaker assumed a priori error constraint in TES (Kulawik et al., 2006; Liu et al., 2010a). The different sensitivities result in large differences between TES and OMI observations as will be discussed below. The OMI sensitivity peaks at 700–500 hPa, overlapping with that of TES.

Figure 2 shows the global distributions of TES and OMI tropospheric DOFS for January and July 2006. There are typically 1–2 DOFS for TES and 0.5–1 for OMI, with lower values at high latitudes (>45°). Both TES and OMI show higher DOFS at northern mid-latitudes in summer than in winter, reflecting higher surface temperatures (TES) and lower solar zenith angles (OMI). TES has higher DOFS than OMI everywhere. The TES V003 ozone has ~0.5 higher DOFS than the V002 data due to improvement of the nadir temperature retrievals (Osterman et al., 2009).

3 Tropospheric ozone distributions from TES and OMI

Figure 3 shows the seasonally averaged TES and OMI retrievals of tropospheric ozone at 500 hPa. OMI data have been sampled at the TES observation locations. Both TES and OMI data have been reprocessed with a single fixed a priori following Eq. (2), and thus the variability is driven solely by the satellite information. The general geographic features and seasonal variability observed by TES and OMI are very similar. They both observe the zonal wave-one pattern in the
tropics, with higher ozone over the Atlantic than the Pacific. They both show enhanced ozone pollution in the summertime (JJA: June-July-August) of northern mid-latitudes due to increasing photochemical production, and high ozone over southern Africa in SON (September-October-November) due to lightning (Sauvage et al., 2007b) and biomass burning (Duncan et al., 2003).

Also shown in Fig. 3 are the GEOS-Chem CTM profiles ($x_{CTM}$) sampled along the TES orbit track at the observation time, and then smoothed with the averaging kernels from TES and OMI over the OMI pressure grid:

$$\hat{x}_{CTM, TES} = A_{TES} x_{CTM} + (I - A_{TES}) x_c$$

$$\hat{x}_{CTM, OMI} = A_{OMI} x_{CTM} + (I - A_{OMI}) x_c$$

Details on the GEOS-Chem CTM are given in Appendix B. The model reproduces the large-scale spatial variability observed by TES and OMI (correlation coefficient $r > 0.8$ for all seasons), although it tends to have lower values as will be discussed in Sect. 6. Figure 3 shows the effect of smoothing by the TES vs. OMI averaging kernels when applied to the same model fields. By comparing in Fig. 4 the differences between model fields smoothed by TES vs. OMI averaging kernels ($\hat{x}_{CTM, TES} - \hat{x}_{CTM, OMI}$) to the observed differences between TES and OMI ($\hat{x}_{TES} - \hat{x}_{OMI}$), we find that most of the observed differences are explainable by instrument sensitivity (different averaging kernel matrices in Eqs. 5 and 6). The residuals represent the actual bias between the instruments as computed by the CTM method and will be discussed in Sect. 5.2.

4 Validation and intercomparison methods

Intercomparison of TES and OMI ozone profiles needs to account for their different vertical resolutions. The difference from directly comparing TES and OMI is given by:

$$\Delta = \hat{x}_{TES} - \hat{x}_{OMI} = b_{TES} - b_{OMI} + (A_{TES} - A_{OMI})(x - x_c)$$

Here and in what follows we consider the average over a sufficiently large number of retrievals (e.g. 5–25 retrieved profiles over each 4° × 5° grid in Fig. 3) so that the random error terms $\sigma_{TES}$ and $\sigma_{OMI}$ average out to zero. The observed difference $\Delta$ thus computes the true difference between TES and OMI $b_{TES} - b_{OMI}$, but with an additional noise term reflecting biases in the a priori profile weighted by the difference between the averaging kernel matrices of the two instruments. Directly intercomparing the two ozone profiles is not appropriate because the noise term can dominate over the true difference as shown in Fig. 4.

We describe here three different methods for validating and intercomparing TES and OMI measurements of tropospheric ozone: (1) independent validation of each with in
situ sonde measurements; (2) use of the GEOS-Chem CTM as a comparison platform; and (3) comparison of OMI ozone profiles with TES profiles smoothed by OMI averaging kernels. The first method provides absolute validation, assuming that ozonesondes measure the true profiles, but is limited by the sparsity of ozonesonde observations (see Sect. 5.1). The second method provides a global intercomparison and also an indirect validation through independent evaluation of the CTM with ozonesonde data. The third method also provides a global intercomparison (with dampening of the difference between instruments, as we will see) but no validation. Results from each method will be presented in Sect. 5.

4.1 In situ method

Validations against ozonesonde data have been presented by Worden et al. (2007) and Nassar et al. (2008) for older versions of TES data (V001 and V002), and by Boxe et al. (2009) for TES V003 and V004 data in the Arctic (>60° N). Validation of OMI ozone data is given by Liu et al. (2010b). Here ozonesonde vertical profiles for a satellite viewing scene are interpolated to the OMI retrieval pressure grid and smoothed with the instrument averaging kernel matrix:

\[
\hat{x}_{\text{sonde}_{\text{TES}}} = A_{\text{TES}} \hat{x} + (I - A_{\text{TES}})\hat{x}_c
\]

(8)

\[
\hat{x}_{\text{sonde}_{\text{OMI}}} = A_{\text{OMI}} \hat{x} + (I - A_{\text{OMI}})\hat{x}_c
\]

(9)

where \(\hat{x}_{\text{sonde}_{\text{TES}}}\) and \(\hat{x}_{\text{sonde}_{\text{OMI}}}\) are the smoothed ozonesonde profiles. Comparisons with satellite retrievals for that scene (\(\hat{x}_{\text{OMI}}\) and \(\hat{x}_{\text{TES}}\)) measure the retrieval error (\(\epsilon + b\)). Averaging over a number of comparison scenes provides an estimate for the systematic error or bias (\(b_{\text{TES}}\) and \(b_{\text{OMI}}\)), and the residuals provide statistics for the random error (\(\epsilon_{\text{TES}}\) and \(\epsilon_{\text{OMI}}\)). The average difference \(\Delta\) between TES and OMI derived from common ozonesonde validation profiles measures the true difference in bias between the two instruments, i.e., their internal consistency:

\[
\Delta = \left(\hat{x}_{\text{TES}} - \hat{x}_{\text{sonde}_{\text{TES}}}\right) - \left(\hat{x}_{\text{OMI}} - \hat{x}_{\text{sonde}_{\text{OMI}}}\right)
\]

(10)

\[b_{\text{TES}} - b_{\text{OMI}}\]

4.2 CTM method

We propose an intercomparison method that uses a CTM as a transfer platform. We apply the instrument averaging kernels to the CTM simulated vertical profiles as described in Sect. 3. From Eqs. (5) and (6), the differences between model and observation are:

\[
\hat{x}_{\text{CTM}_{\text{TES}}} - \hat{x}_{\text{TES}} = A_{\text{TES}}(x_{\text{CTM}} - x) - b_{\text{TES}}
\]

(11)
5 Results from the different methods

We examine here how the CTM method (Sect. 4.2) and the averaging kernel smoothing method (Sect. 4.3) compare to the standard set by the in situ method (Sect. 4.1). The in situ method gives the true measure of consistency between the two instruments $A_{\text{CTM}} = b_{\text{TES}} - b_{\text{OMI}}$. We will assess the importance of this shortcoming below.

### 5.1 In situ method

We use the ensemble of worldwide ozonesonde measurements for 2005–2007 available from the Aura Validation Data Center for tropospheric studies (Schoeberl et al., 2007; Liu et al., 2010b; http://avdc.gsfc.nasa.gov/), giving us more comparison profiles than for 2006 alone. We find the validation results for individual years are similar. For comparison to the satellite data we require spatial coincidence within 5° longitude and 2° latitude, and temporal coincidence within 10 h, similar to the coincidence criteria applied by Nassar et al. (2008). This results in 528 TES/sonde coincidences and 2568 OMI/sonde coincidences within 60° S–60° N for 2005–2007. OMI has more coincidences because of its daily global coverage. About 80% of the comparison profiles are at northern mid-latitudes (20°–60° N).

Table 1 summarizes the mean differences and standard deviations for TES and OMI relative to the ozonesonde data for three latitudinal bands at 860, 500, and 300 hPa. Our validation results for TES V003 data are very similar to those obtained for V002 data by Nassar et al. (2008). At northern mid-latitudes (20° N–60° N), TES has a positive bias of 5.7 ± 3.5 ppbv (mean ± standard deviation) at 500 hPa.

### Table 1. TES and OMI ozone biases relative to ozonesondes$^a$.

<table>
<thead>
<tr>
<th>Latitude band</th>
<th>N (TES)</th>
<th>TES bias ± σ (ppbv)</th>
<th>N (OMI)</th>
<th>OMI bias ± σ (ppbv)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>860 hPa</td>
<td>500 hPa</td>
<td>300 hPa</td>
<td>860 hPa</td>
</tr>
<tr>
<td>20° N–60° N</td>
<td>398</td>
<td>4.6±12.1</td>
<td>5.7±13.5</td>
<td>17.9±36.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.1±4.4)$^b$</td>
<td>(3.3±6.0)</td>
<td>(11.1±29.2)</td>
</tr>
<tr>
<td>20° S–20° N</td>
<td>102</td>
<td>7.5±12.6</td>
<td>3.2±7.5</td>
<td>−1.9±11.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(−0.2±3.6)</td>
<td>(0.3±7.8)</td>
<td>(1.0±7.3)</td>
</tr>
<tr>
<td>60° S–20° S</td>
<td>28</td>
<td>3.3±4.6</td>
<td>6.5±5.8</td>
<td>4.6±20.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.4±4.5)</td>
<td>(0.9±6.0)</td>
<td>(4.4±9.5)</td>
</tr>
</tbody>
</table>

$a$ Mean difference and standard deviation as determined by difference with ozonesonde data for 2005–2007 (see text). $N$ is the number of comparison profiles.

$^b$ Values in parentheses are the OMI validation statistics if subsampling the OMI/sonde comparisons at the TES observation locations.
Fig. 5. TES and OMI ozone retrieval biases relative to ozonesondes in unit of ppbv at 500 hPa averaged over the four seasons of 2005–2007. Each point represents the mean bias over a sonde launch site. Values are shown as differences between retrieved ozone concentrations and the ozonesonde data with averaging kernels applied. Red colors indicate a positive bias in the satellite data.

17.9±36.5 ppbv at 300 hPa. In the tropics and southern mid-latitudes, TES has a positive bias of 3.2–7.5 ppbv in the lower troposphere and less in the upper troposphere. The OMI biases are also positive but smaller than TES and less variable. The OMI validation statistics are similar if we subsample the OMI/sonde comparisons at the TES observation locations (see Table 1).

Figure 5 shows the distribution of satellite minus ozonesonde differences at 500 hPa averaged over each sonde site for the four seasons. The global mean bias relative to the sondes at 500 hPa is 5.3±12.3 ppbv for TES ($n=528$) and 2.8±6.6 ppbv for OMI ($n=2568$). The seasonal variation of TES biases is small. At northern mid-latitudes OMI has a positive bias of 1.2 ppbv in summer and 4 ppbv in other seasons.

5.2 CTM method

Figure 6 (left) compares the TES minus OMI differences $\Delta_2$ obtained by the CTM method (Sect. 4.2) to those obtained by the in situ validation method at 500 and 860 hPa. Results are shown for the 180 TES/OMI/sonde coincidences in the year 2006 where we can measure the true instrument differences $\Delta_1=b_{\text{TES}}-b_{\text{OMI}}$ through the calibration provided by the sonde profiles (x-axis in Fig. 6). The CTM method provides a close approximation to the true results from the in situ method. Correlation with results of the in situ method is high at both 500 and 800 hPa ($r=0.89–0.91$) and slopes are near unity (0.96–0.99).

The close agreement between the CTM and in situ methods lends confidence in using the CTM method for deriving
Fig. 7. Seasonally averaged TES minus OMI ozone differences (ppbv) at 500 hPa computed by the in situ method (left) and the CTM method (right). The in situ method is applied to 2005–2007 data while the CTM method is applied to 2006 data. Red colors indicate that TES ozone is higher than OMI. Gray areas indicate lack of data meeting the retrieval quality criteria.

global patterns of differences between TES and OMI. Figure 7 shows the global distribution of TES minus OMI differences at 500 hPa obtained by the CTM method for the four seasons of 2006, and compares these differences with those from the in situ method. We previously showed in Fig. 4 how differences in instrument sensitivity contributed most of the TES vs. OMI differences at 500 hPa; Fig. 7 shows the residuals not attributable to instrument sensitivity. These are generally less than 10 ppbv. The largest inconsistencies between TES and OMI occur in the summertime northern mid-latitudes where TES is higher than OMI (TES − OMI differences > 10 ppbv), and over South America, Africa and Indonesia where TES is lower than OMI (TES − OMI differences < −10 ppbv). At northern mid-latitudes TES is generally higher than OMI, but in winter there is a broad region within 20° N–40° N where TES is lower than OMI. The CTM method shows that TES is generally lower than OMI in the tropics, which is not apparent from the in situ method because of sparse spatial sampling of ozonesonde measurements.

5.3 Averaging kernel smoothing method

Figure 6 (right) compares the TES minus OMI differences Δ3 obtained by the Rodgers and Connor (2003) averaging kernel smoothing method (Sect. 4.3) to those obtained by the in situ method. We see that the averaging kernel smoothing method generally underestimates Δ3, as would be expected from the application of the OMI averaging kernel matrix AOMI to the TES systematic bias bTES, which tends to be positive (Fig. 5). The slopes of the reduced-major-axis (RMA) regression lines are less than 1, especially at 860 hPa where the OMI sensitivity is low. The Δ values obtained by the two methods are only moderately correlated (r = 0.70 for 500 hPa and 0.52 for 860 hPa). We find that the inability of the averaging kernel smoothing method to reproduce the true intercomparison from the in situ method is mostly due to the bias smoothing term AOMIbTES in Eq. (15). The additional noise term AOMI(ΔTES−I)x is small in comparison, although this reflects our use of the OMI geographically varying profile as common a priori x for both retrievals. It would be the dominant term had we used the single fixed a priori profile.

The averaging kernel smoothing method has the advantage over the in situ method of extending the intercomparison to a global scale, although one has to be wary of results in view of...
Fig. 8. Differences of simulated GEOS-Chem (GC) ozone concentrations (ppbv) at 500 hPa with ozonesondes (left), TES (central), and OMI (right), averaged for the four seasons of 2006. Comparisons with TES and OMI use the GEOS-Chem profiles smoothed by the respective instrument averaging kernels as shown in Fig. 3. Satellite measurements have been corrected for the mean positive biases of 5.3 ppbv for TES and 2.8 ppbv for OMI. Red colors indicate that the model simulation is biased high and blue colors indicate that the model simulation is biased low. Black areas in the central and right columns indicate where differences between TES and OMI computed by the CTM method as shown in Fig. 7 are larger than 10 ppbv or smaller than −10 ppbv, indicating that the satellite measurements are not consistent.

the deficiencies shown in Fig. 6. We find that the spatial patterns of TES minus OMI differences at 500 hPa obtained by the averaging kernel smoothing method are similar to those obtained from the CTM method \((r>0.8\) for all seasons).

6 Application to CTM evaluation

The consistency between TES and OMI data in most regions lends confidence to using these data to evaluate the GEOS-Chem CTM simulation. Figure 8 shows the differences of the GEOS-Chem ozone simulation with TES and OMI measurements at 500 hPa for the four seasons of 2006. Comparison with the sonde measurements is also shown. We have subtracted from the TES and OMI data the global mean positive biases of 5.3 and 2.8 ppbv respectively, as revealed by the ozonesonde comparisons. For purpose of model evaluation, we consider TES and OMI to be consistent if their differences do not exceed 10 ppbv. Regions where the differences exceed 10 ppbv are shown in black in Fig. 8. These include some tropical continental regions as well as large areas at northern mid-latitudes.

Both TES and OMI measurements at 500 hPa show that GEOS-Chem underestimates ozone by more than 10 ppbv over Africa and South America; weaker negative biases extend over most of the tropics. Available sonde observations also show the model underestimate, but with very limited spatial and seasonal resolution. The underestimate could reflect a number of factors. Lightning is the dominant contributor to tropical tropospheric ozone, but its magnitude is highly uncertain and may contribute to the model bias, particularly over the South Atlantic (Sauvage et al., 2007a, b). Jaeglé et al. (2004, 2005) found that soil NOx emissions in GEOS-Chem are a factor of 2 too low over north tropical Africa in spring and summer due to rain-induced microbial pulsing, and this could cause 5–7 ppbv seasonal underestimates of ozone over Africa (Sauvage et al., 2007b). Some underestimates occur over the seasonal biomass burning regions, such as southern Africa and South America in September-October-November (SON), possibly reflecting a negative bias of the GFEDv2 biomass burning emission inventory (van der Werf et al., 2006) used in the model. Nassar et al. (2009) previously conducted a detailed analysis of the GEOS-Chem model simulation of tropical tropospheric ozone in the context of the El Niño event in October–December 2006, and found that overly strong convection in the model can also contribute to the ozone underestimate.

Both satellite instruments reveal a year-round model overestimate in the northern subtropics, which had been noticed in previous GEOS-Chem model simulations (Liu et al., 2006), and also in the extra-tropical southern hemisphere, where ozonesonde measurements are very sparse. Downward stratospheric ozone fluxes are expected to be largest over these regions due to the subtropical jet streams (Hsu et al., 2005). The GEOS-Chem model simulates transport of ozone from stratosphere using the “Synoz” flux boundary
condition of McLinden et al. (2000), with a global stratosphere- tropopause ozone flux of 495 Tg ozone per annum (a\(^{-1}\)), but this may not adequately represent the stratosphere-troposphere exchange over these regions. Further investigation of these model errors is warranted but is beyond the scope of the paper.

7 Summary

We have analyzed the theoretical basis of three different methods to validate and intercompare satellite retrievals of atmospheric composition. The first method (in situ method) uses in situ vertical profiles for absolute instrument validation. The second method (CTM method) uses a chemical transport model (CTM) as an intercomparison platform. The third method (averaging kernel smoothing method) involves smoothing the retrieved profile from one instrument with the averaging kernel matrix of the other. The three methods were illustrated and compared using a full year (2006) of tropospheric ozone data from the Tropospheric Emission Spectrometer (TES) and the Ozone Monitoring Instrument (OMI), both aboard the EOS Aura satellite.

An important preliminary step to the intercomparison is to convert the retrievals for the different instruments to use the same units, grid, and a priori information. We showed how to perform the unit conversion between TES (original retrieval in logarithm of mixing ratio units) and OMI (original retrieval in partial pressure units), by operating on the TES averaging kernel matrices without significantly degrading the information content. Both TES and OMI averaging kernels show peak sensitivity to ozone concentration in the middle troposphere (700–400 hPa). There are typically 1–2 degrees of freedom for signal (DOFS) in the troposphere for TES and 0.5–1 for OMI. Comparison of collocated TES and OMI data for the full year of 2006 shows similar geographic features and seasonal variability. Application of the averaging kernels of TES and OMI to ozone profiles from the GEOS-Chem CTM indicates that much of the difference in the data between the two instruments simply reflects differences in instrument sensitivity.

Direct intercomparison of measurements from two satellite instruments introduces a noise term due to error in the a priori profile (the difference between the a priori profile and the true profile) that can dominate over the true difference. We examined how the three different validation/intercomparison methods can improve on this situation. The in situ method provides absolute validation of the satellite instruments and true intercomparison, but it is limited by the sparseness of the in situ data. The CTM method by contrast provides a globally complete intercomparison. It introduces noise from error in the CTM profile but that is generally smaller than the error from the generic a priori profile. The averaging kernel smoothing method also provides a global intercomparison but it dampens the actual difference between the two instruments (particularly when sensitivity is low). The CTM method has three major advantages: (1) it allows intercomparison of satellite instruments over the full range of operating conditions, (2) it enables indirect validation against in situ data using the CTM as a transfer function, and (3) it exploits the satellite data for CTM evaluation.

We applied each method to the analysis of differences between TES and OMI tropospheric ozone retrievals. For the in situ method, using the global ozonesonde database, we find 528 TES/sonde coincidences and 2568 OMI/sonde coincidences within 60° S–60° N for 2005–2007 (80% of these are between 20° N and 60° N). Both instruments show an overall positive bias. The global mean bias at 500 hPa is 5.3±12.3 ppbv for TES and 2.8±6.6 ppbv for OMI. Application of the CTM method to the 180 TES/OMI/sonde coincidences for 2006 shows that it closely reproduces the results of the in situ method while providing a globally complete intercomparison perspective. The averaging kernel smoothing method does not perform as well.

We find from the CTM method that differences between TES and OMI are generally within ±10 ppbv (18%). The largest differences are at summertime northern mid-latitudes where TES is higher than OMI (TES – OMI differences >10 ppbv), and over tropical continents where TES is lower than OMI (TES – OMI differences <−10 ppbv).

We used the CTM method to diagnose GEOS-Chem model biases for regions where the two satellite instruments are consistent, exploiting the much better global coverage afforded by the satellite data relative to ozonesondes. We first removed the mean instrument biases revealed by the ozonesonde validation (+5.3 ppbv for TES, +2.8 ppbv for OMI). Both satellite (TES and OMI) and in situ measurements show that GEOS-Chem underestimates ozone at 500 hPa in the tropics; this could reflect model errors in convective transport and in NO\(_x\) emissions from lightning, soil, and biomass burning. Both TES and OMI reveal model overestimates in the northern subtropics and southern extratropics that could reflect excessive stratospheric ozone influx in these latitudinal bands. The consistency of measurements from TES and OMI allows integration of the two into models and combining their advantages (better vertical structure from TES, daily global coverage from OMI) to improve our understanding of tropospheric ozone.

Appendix A

Converting TES averaging kernels to the OMI pressure grid

TES averaging kernels are based on retrieval of logarithms of ozone volume mixing ratio (VMR) on a 67-layer pressure grid at fixed pressure levels. We convert them to partial columns on the OMI 24-layer pressure grid to enable direct comparison of TES and OMI vertical sensitivities and to
facilitate the mathematical interpretation of the intercomparison. The conversion involves the following steps:

1. Convert the TES log(VMR)-based averaging kernel ($A_1$) to VMR-based ($A_1$). This is done by rewriting Eq. (1) for the TES retrieval in the text so that $\hat{x}$, $x_a$, and $x$ are profiles of ozone VMR:

$$\ln \hat{x} = \ln x_a + A_1 (\ln x - \ln x_a) + \epsilon$$  \hspace{1cm} (A1)

By definition of $A_1$, we must have:

$$\hat{x} = x_a + A_1 (x - x_a) + \epsilon'$$  \hspace{1cm} (A2)

Let $x_i$ denote the VMR for layer $i$, which is the $i$-th element of $x$, and assume that the difference between $x_i$ and $x_{a,i}$ is relatively small so that

$$\ln x_i - \ln x_{a,i} \approx \frac{x_i - x_{a,i}}{x_{a,i}}$$  \hspace{1cm} (A3)

with the same relationship holding between $\hat{x}_i$ and $x_{a,i}$. The elements $a_{1,ij}$ of the converted averaging kernel $A_1$ are then related to the elements $a_{ij}$ of the averaging kernel $A$ by:

$$a_{1,ij} = \frac{x_{a,i}}{x_{a,j}} a_{ij}$$  \hspace{1cm} (A4)

2. Convert the averaging kernel for VMR ($A_1$) to an averaging kernel for partial ozone column ($A_2$). Let $x_1$ denote an ozone profile in unit of ppbv and $x_2$ denote the same profile in Dobson Units (1 DU=2.69×10$^{-16}$ molecules cm$^{-2}$). Their elements in layer $i$ ($x_{1,i}$ and $x_{2,i}$) are related through:

$$x_{2,i} = x_{1,i} \left( \frac{C A P_i}{m_a g} \right)$$  \hspace{1cm} (A5)

where $\Delta P_i$ is the layer thickness in unit of Pa, $m_a=4.808\times10^{-26}$ kg is the mean molecular mass of air, $g=9.81\times10^2$ cm s$^{-2}$ is the acceleration of gravity, and $C=3.72\times10^{-28}$ DU cm$^2$ is a unit conversion factor. Both averaging kernels $A_1$ and $A_2$ follow Eq. (A2), so that the elements $a_{2,ij}$ of $A_2$ are given by:

$$a_{2,ij} = \frac{\Delta P_i}{\Delta P_j} a_{1,ij}$$  \hspace{1cm} (A6)

3. Interpolate the resulting averaging kernel matrix ($A_2$) to the OMI grid ($A_3$). Transformation of averaging kernels to a different grid has been described by von Clarmann and Grabowski (2007). Let $M$ represent the mapping matrix that interpolates the retrieved profile from the coarse 24-layer OMI pressure grid to the fine 67-layer TES pressure grid. The regridded averaging kernel matrix ($A_3$) is then given by:

$$A_3 = M^T A_2 M$$  \hspace{1cm} (A7)

where $M^T=(M^T M)^{-1}M^T$ is the pseudo inverse of $M$ and $M^T$ is its transpose. The resulting averaging kernel matrix $A_3$ for TES ($A_{\text{TES}}$ in the text) applies to partial columns on the OMI pressure grid and can be directly compared with the OMI averaging kernel matrix $A_{\text{OMI}}$.

**Appendix B**

**The GEOS-Chem CTM**

GEOS-Chem is a global 3-D model of atmospheric composition (v8-01-04; http://acmg.seas.harvard.edu/geos/) driven by GEOS assimilated meteorological observations from the NASA Global Modeling and Assimilation Office (GMAO). General descriptions of GEOS-Chem are given by Bey et al. (2001) and Park et al. (2004). The model is applied here to a global simulation of tropospheric ozone-NO$_x$-VOC-aerosol chemistry for January–December 2006 using GEOS-4 data and with an 8-month spin-up. Meteorological fields in the GEOS-4 data have a temporal resolution of 6 h (3 h for surface variables and mixing depths) and a horizontal resolution of 1° latitude by 1.25° longitude, with 55 layers in the vertical from surface to 0.01 hPa. We degrade here the horizontal resolution to 4° latitude by 5° longitude for input to GEOS-Chem.

We use global anthropogenic emissions for 2000 from the Emission Database for Global Atmospheric Research (EDGAR) inventory (Oliver and Berdowski, 2001), replaced with the following regional inventories: the European Monitoring and Evaluation Programme (EMEP) for Europe in 2000 (Vestreng and Klein, 2002), Zhang et al. (2009) for Asia in 2006, and the NEI-99 for the United States with downward corrections of 50% in NO$_x$ emissions from point sources and 60% in CO emissions as described by Hudman et al. (2007, 2008). Biomass burning emissions are from the Global Fire Emission Database version 2 (GFEDv2) (van der Werf et al., 2006). Soil NO$_x$ emissions are computed using a modified version of the Yienger and Levy (1995) algorithm with canopy reduction factors described in Wang et al. (1998). Emissions of NO$_x$ from lightning are linked to deep convection following the parameterization of Price and Rind (1992) with vertical profiles taken from Pickering et al. (1998). We use a NO$_x$ yield per flash of 125 moles in the tropics and 500 moles at northern mid-latitudes (north of 30° N) (Hudman et al., 2007). The resulting lightning source is scaled to be 6 Tg nitrogen per annum ($\text{a}^{-1}$) globally. Transport of ozone from the stratosphere is simulated using the “SynoZ” flux boundary condition of McLinden et al. (2000), with a global cross-tropopause ozone flux of 495 Tg ozone a$^{-1}$.

The GEOS-Chem simulation of tropospheric ozone has been evaluated in many studies with measurements from surface sites (Fiore et al., 2002, 2003; Wang et al., 2009), ozonesondes (Sauvage et al., 2007a; Nassar et al., 2009), aircraft (Jaeglé et al., 2003; Hudman et al., 2004; Auvray et al., 2007; Hudman et al., 2007; Zhang et al., 2008), and satellites (Martin et al., 2002; Liu et al., 2006; Parrington et al., 2008).
Acknowledgements. This work was funded by the NASA Atmospheric Composition Modeling and Analysis Program and by NASA Headquarters under the Earth and Space Science Fellowship Program Grant NNX07AN65H to Lin Zhang, and by the New Investigator Program in Earth Science (NNX08AN98G) to Xiong Liu. Xiong Liu and Kelly Chance also acknowledge support from the NASA Atmospheric Composition Program (NNG06GH99G) and the Smithsonian Institution.

Edited by: W. Lahoz

References


Atmos. Chem. Phys., 10, 4725–4739, 2010

L. Zhang et al.: Intercomparison methods for satellite measurements of atmospheric composition


