Global estimates of CO sources with high resolution by adjoint inversion of multiple satellite datasets (MOPITT, AIRS, SCIAMACHY, TES)

M Kopacz  
Harvard University

D J. Jacob  
Harvard University

J A. Fisher  
University of Wollongong, jennyf@uow.edu.au

J A. Logan  
Harvard University

L Zhang  
Harvard University

See next page for additional authors

Follow this and additional works at: https://ro.uow.edu.au/scipapers

Part of the Life Sciences Commons, Physical Sciences and Mathematics Commons, and the Social and Behavioral Sciences Commons

Recommended Citation
Kopacz, M; Jacob, D J.; Fisher, J A.; Logan, J A.; Zhang, L; Megretskaia, I A.; Yantosca, R M.; Singh, K; Henze, D K.; Burrows, J P.; Buchwitz, M; Khlystova, I; McMillan, W. W; Gille, J C.; Edwards, D P.; Eldering, A; Thouret, V; and Nedelec, P: Global estimates of CO sources with high resolution by adjoint inversion of multiple satellite datasets (MOPITT, AIRS, SCIAMACHY, TES) 2010, 855-876.  
https://ro.uow.edu.au/scipapers/4693

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au
Global estimates of CO sources with high resolution by adjoint inversion of multiple satellite datasets (MOPITT, AIRS, SCIAMACHY, TES)

Abstract
We combine CO column measurements from the MOPITT, AIRS, SCIAMACHY, and TES satellite instruments in a full-year (May 2004–April 2005) global inversion of CO sources at 4° × 5° spatial resolution and monthly temporal resolution. The inversion uses the GEOS-Chem chemical transport model (CTM) and its adjoint applied to MOPITT, AIRS, and SCIAMACHY. Observations from TES, surface sites (NOAA/GMD), and aircraft (MOZAIC) are used for evaluation of the a posteriori solution. Using GEOSChem as a common intercomparison platform shows global consistency between the different satellite datasets and with the in situ data. Differences can be largely explained by different averaging kernels and a priori information. The global CO emission from combustion as constrained in the inversion is 1350 Tg a⁻¹. This is much higher than current bottom-up emission inventories. A large fraction of the correction results from a seasonal underestimate of CO sources at northern mid-latitudes in winter and suggests a larger-than-expected CO source from vehicle cold starts and residential heating. Implementing this seasonal variation of emissions solves the long-standing problem of models underestimating CO in the northern extratropics in winter-spring. A posteriori emissions also indicate a general underestimation of biomass burning in the GFED2 inventory. However, the tropical biomass burning constraints are not quantitatively consistent across the different datasets.

Keywords
airs, mopitt, datasets, satellite, multiple, inversion, adjoint, tes, resolution, sciamachy, high, sources, co, estimates, global

Disciplines
Life Sciences | Physical Sciences and Mathematics | Social and Behavioral Sciences

Publication Details

Authors

This journal article is available at Research Online: https://ro.uow.edu.au/scipapers/4693
Global estimates of CO sources with high resolution by adjoint inversion of multiple satellite datasets (MOPITT, AIRS, SCIAMACHY, TES)

M. Kopacz1,*, D. J. Jacob1, J. A. Fisher1, J. A. Logan1, L. Zhang1, I. A. Megretskaia1, R. M. Yantosca1, K. Singh2, D. K. Henze3, J. P. Burrows3, M. Buchwitz4, I. Khlystova4, W. W. McMillan5, J. C. Gille6, D. P. Edwards6, A. Eldering7, V. Thouret8,9, and P. Nedelec8,9

1School of Engineering and Applied Science, Harvard University, Cambridge, MA, USA
2Department of Computer Science, Virginia Polytechnic Institute, Blacksburg, VA, USA
3Department of Mechanical Engineering, University of Colorado at Boulder, CO, USA
4Institute of Environmental Physics (IUP), University of Bremen, Bremen, Germany
5Department of Physics, University of Maryland Baltimore County, Baltimore, MD, USA
6National Center for Atmospheric Research, Boulder, Colorado, USA
7Jet Propulsion Laboratory, Pasadena, CA, USA
8Univ. de Toulouse, UPS, LA (Laboratoire d’Arologie), 14 avenue Edouard Belin, 31400, Toulouse, France
9CNRS, LA (Laboratoire d’Aérologie), 31400 Toulouse, France

*now at: Woodrow Wilson School of International and Public Affairs, Princeton University, Princeton, NJ, USA

Received: 18 August 2009 – Published in Atmos. Chem. Phys. Discuss.: 24 September 2009
Revised: 12 January 2010 – Accepted: 12 January 2010 – Published: 1 February 2010

Abstract. We combine CO column measurements from the MOPITT, AIRS, SCIAMACHY, and TES satellite instruments in a full-year (May 2004–April 2005) global inversion of CO sources at 4°×5° spatial resolution and monthly temporal resolution. The inversion uses the GEOS-Chem chemical transport model (CTM) and its adjoint applied to MOPITT, AIRS, and SCIAMACHY. Observations from TES, surface sites (NOAA/GMD), and aircraft (MOZAIC) are used for evaluation of the a posteriori solution. Using GEOS-Chem as a common intercomparison platform shows global consistency between the different satellite datasets and with the in situ data. Differences can be largely explained by different averaging kernels and a priori information. The global CO emission from combustion as constrained in the inversion is 1350 Tg a⁻¹. This is much higher than current bottom-up emission inventories. A large fraction of the correction results from a seasonal underestimate of CO sources at northern mid-latitudes in winter and suggests a larger-than-expected CO source from vehicle cold starts and residential heating. Implementing this seasonal variation of emissions solves the long-standing problem of models underestimating CO in the northern extratropics in winter-spring. A posteriori emissions also indicate a general underestimation of biomass burning in the GFED2 inventory. However, the tropical biomass burning constraints are not quantitatively consistent across the different datasets.

1 Introduction
Carbon monoxide (CO) is a product of incomplete combustion and atmospheric oxidation of volatile organic compounds (VOCs). It has an atmospheric lifetime of about two months against oxidation by the OH radical. It is of interest as a sink for OH, the main tropospheric oxidant (Logan et al., 1981), as an indirect greenhouse gas (Forster et al., 2007), as a tracer of long-range transport of pollution (Logan et al., 1981), and as a correlative constraint for inverse analyses of CO₂ surface fluxes (Palmer et al., 2006). Understanding CO sources also places constraints on emissions of other pollutants released during combustion and whose emissions are often referenced to CO (Andreae and Merlet, 2001).

CO has strong absorption lines in the thermal infrared (4.7 µm) and in the solar shortwave infrared (2.3 µm), enabling its observation from space. A number of satellite instruments have been measuring tropospheric CO globally over the past decade including MOPITT (2000-) (Edwards...
et al., 2006b; Emmons et al., 2007, 2009), SCIAMACHY (2002-) (Bovensmann et al., 1999; Buchwitz et al., 2007; Burrows et al., 1995; de Laat et al., 2007), AIRS (2002- ) (McMillan et al., 2005, 2008, 2010; Warner et al., 2007; Yurganov et al., 2008), ACE-FTS (2003-) (Clerbaux et al., 2005, 2008), TES (2004-) (Lopez et al., 2008; Luo et al., 2007a; Rinsland et al., 2006), and IASI (2007-) (Fortems-Chiney et al., 2009; Turquety et al., 2009). These satellite data expand the perspective offered by in situ observations, such as from the NOAA/GMD surface monitoring network (Novelli et al., 2003) and from commercial aircraft (Nedelec et al., 2003).

Our objective here is to combine information from four different satellite sensors (MOPITT, SCIAMACHY, AIRS, TES) to provide global high-resolution constraints on CO sources using an adjoint inverse modeling method. The four instruments all observe in the nadir from sun synchronous polar orbits. MOPITT, AIRS, and TES observe thermal emission in the 4.7 µm absorption band and thus are most sensitive to the mid-troposphere. SCIAMACHY observes backscattered solar radiation in the 2.3 µm absorption band and is thus sensitive to the full depth of the atmosphere. AIRS and TES are on the same orbit (A-train) with the Equator crossing time within 8 min of 01:30 local solar time (LST). MOPITT and SCIAMACHY are on different orbits with the Equator crossing times of 10:30 and 10:00.

A central component of our work is to assess the consistency and complementarity of the data from the different satellite instruments. This is challenging because of the differences across instruments in sensitivity, retrieval techniques and observing schedule. Some limited intercomparisons between satellite pairs have been reported in the literature (Buchwitz et al., 2007; George et al., 2009; Luo et al., 2007b; Turquety et al., 2008; Warner et al., 2007; Yurganov et al., 2008). Aircraft vertical profiles can provide accurate validation but are sparse. A more general approach that we exploit here is to use a chemical transport model (CTM) as an intercomparison platform. The CTM provides a global, continuous, and consistent 3-D representation of CO concentrations, albeit with some error. Comparison of the observed and modeled CO concentrations sampled for the different orbits, overpass times, and retrievals of the individual instruments are used to examine the consistency of the observations relative to the model. This is particularly useful in an inverse modeling framework where, as here, the CTM serves as the forward model for the inversion.

Despite long-standing interest in atmospheric CO and the abundance of data, our understanding of the CO budget remains inadequate, as illustrated by a recent CTM comparison exercise showing large disagreements between models and observations (Shindell et al., 2006). Simulation of the spatial, seasonal, and interannual variability of CO involves a complex interplay of sources, transport, and chemistry (Duncan et al., 2008). Errors in sources can exceed a factor of two on continental scales (Bian et al., 2007; Hudman et al., 2008).

A number of inverse modeling studies have used MOPITT satellite data as constraints on CO sources (Arellano et al., 2004, 2006; Heald et al., 2004; Pétrot et al., 2004; Pfister et al., 2005), including several by the adjoint method (Chevallier et al., 2009; Kopacz et al., 2009; Stavrakou and Müller, 2006; Yumimoto and Uno, 2006). The study by Fortems-Chiney et al. (2009) combined MOPITT and IASI data. Results of these and other inverse studies using surface CO measurements as constraints (e.g. Kasibhatla et al., 2002; Pétrot et al., 2002) are often not quantitatively consistent, which could reflect insufficient constraints from observations, errors from model transport, and unrecognized errors in the inverse modeling approach. The adjoint method is particularly efficient at extracting the information content from observations by retrieving sources at the resolution of the underlying CTM, thus overcoming large-region aggregation errors in the more standard analytical method (Kopacz et al., 2009).

Exploitation of multi-sensor satellite data in a global inversion by the adjoint method holds the potential for significant advance over previous studies and we follow that approach here.

We use a full year (May 2004–May 2005) of satellite data from MOPITT, SCIAMACHY, AIRS, and TES. This time period corresponds to the best overlap of data from these instruments. The Short Wave Infrared channels of SCIAMACHY were experimental and the first of their kind to fly in space. The 2.3 µm channel suffered most from the growth of the ice layer in 2003 and later also from an increasing number of bad and dead detector pixels (2005-) arising from radiation damage. Hence, 2004 was the best year for SCIAMACHY data (Buchwitz et al., 2007). The TES record begins in September 2004. We use the GEOS-Chem CTM as the forward model for the inversion and apply its adjoint (Henze et al., 2007; Kopacz et al., 2009) to optimize the CO sources on a 4°×5° horizontal grid with monthly temporal resolution. We begin by describing the satellite datasets (Sect. 2) and the GEOS-Chem CTM (Sect. 3). In Sect. 4, we intercompare the data from the different satellite instruments using GEOS-Chem as the intercomparison platform. The inverse analysis is described in Sect. 5 and results are presented in Sect. 6. Testing of the optimized sources with independent datasets including in situ data from the surface (NOAA/GMD network) and aircraft (MOZAIC) is presented in Sect. 7.

2 Satellite data

2.1 MOPITT

The Measurements Of Pollution In The Troposphere (MOPITT) instrument was launched aboard EOS Terra in December 1999. The Equator crossing time is 10:30/22:30 with global coverage every 3 days. MOPITT measures thermal emission in the 4.7 µm absorption band, which results in highest vertical sensitivity in the mid-troposphere but also provides some boundary layer information (Deeter et
al., 2003, 2007; Kar et al., 2008). The retrieval uses the Rodgers (2000) optimal estimation technique to retrieve CO concentrations. The sensitivity of the retrieval (to the true profile) is defined by its averaging kernel matrix A:

\[ \hat{z} = z_a + A(z - z_a) \]  

(1)

where \( \hat{z} \) is the retrieved vertical profile vector consisting of mixing ratios on a fixed pressure grid (Deeter et al., 2003), \( z \) is the true profile on the same grid, and \( z_a \) is a globally uniform a priori profile derived from an ensemble of aircraft observations (Deeter et al., 2003). Only cloud-free scenes are retrieved. The Degrees Of Freedom (DOF) for signal, representing the number of pieces of information in the vertical profile and estimated as the trace of the averaging kernel matrix, are typically about 1.5 (Deeter et al., 2004). Therefore we only use the altitude-weighted CO column \( y \) obtained by summing the vertical profile \( \hat{z} \) with the corresponding pressure weights. MOPITT version 3 data for \( \hat{y} \) and A are collected from ftp://airspar1u.ecs.nasa.gov/MOPITT/MOP02.003/. MOPITT daytime observations have been validated against aircraft data from several campaigns (mostly in the Northern Hemisphere), indicating a positive bias of about 5±11% on the column (Emmons et al., 2004, 2007, 2009; Jacob et al., 2003). Nighttime observations have not been validated and appear subject to larger bias (Heald et al., 2004). We use the daytime data only.

2.2 AIRS

The Atmospheric Infrared Sounder (AIRS) instrument was launched aboard EOS Aqua in May 2002. The Equator crossing time is 01:30/13:30 with daily global coverage due to a 1650-km cross-track scanning swath. AIRS measures thermal emission in the 4.7 µm absorption band, as does MOPITT (McMillan et al., 2005, 2010; Warner et al., 2007). However, unlike MOPITT and other instruments in this comparison, AIRS possesses a cloud clearing capability (Susskind et al., 2003, 2010) that enables it to retrieve partly cloudy scenes and thus achieve 70% effective daily coverage. Version 5 CO retrieval is described by and McMillan et al. (2010). Retrieval of partial columns \( \hat{z} \) follows the equation (Olsen, 2007):

\[ \ln \hat{z} = \ln z_a + F A F' (\ln z - \ln z_a) \]  

(2)

where \( z \) is a vertical profile of partial columns on the 100 levels of the radiative transfer model, \( F \) is a matrix that defines the nine trapezoidal layers on which AIRS CO is retrieved, \( F' \) is its pseudo inverse, \( A \) is a 9×9 averaging kernel matrix in the trapezoidal space, and \( z_a \) is an a priori profile of partial columns, which is the same as for MOPITT for the common levels and AFGL standard atmosphere above that. AIRS retrievals have DOF for signal on average about 0.8, with higher values over land than ocean and typically higher in daytime than at night. We use the columns \( \hat{y} \) obtained by summing the vertical profiles \( \hat{z} \) of partial columns. AIRS version 5 data for \( \hat{y} \), \( F \) and \( A \) are collected from ftp://airspar1u.ecs.nasa.gov/MOPITT/MOP02.003/. MOPITT daytime observations have been validated against aircraft data from several northern hemisphere campaigns finding a positive bias of approximately 9%±12% between 300 and 900 hPa (McMillan et al., 2010). However, AIRS v5 CO retrievals are known to exhibit a larger high bias in the Southern Hemisphere (Yurganov et al., 2008, 2010). For consistency, we use daytime CO column data only as for MOPITT. For best quality, we subsample for retrievals with surface temperature greater than 250 K.

2.3 TES

The Tropospheric Emission Spectrometer (TES) instrument was launched aboard EOS Aura in July 2004 (observations available starting September 2004). The overpass time lags 8 min behind AIRS. TES measures thermal emission at 4.7 µm, as do MOPITT and AIRS. It obtains global coverage every 16 days and has no cross-track scanning capability, yielding a much sparser dataset than MOPITT or AIRS (Rinsland et al., 2006). The optimal estimation retrieval (Rodgers, 2000) provides vertical profiles \( \hat{z} \) of logarithms of mixing ratios:

\[ \ln \hat{z} = \ln z_a + F A F' (\ln z - \ln z_a) \]  

(3)

Unlike AIRS and MOPITT, the TES a priori profiles \( z_a \) vary by region and season (Osterman et al., 2007). As for MOPITT and AIRS, we only use daytime column data \( \hat{y} \) computed from the vertical profile \( \hat{z} \). TES V002 data for \( \hat{y} \) and A were collected from http://eosweb.larc.nasa.gov/PRODOCS/tes/table_tes.html. Limited validation of these data with aircraft show no consistent bias (Lopez et al., 2008; Luo et al., 2007a). The quality of the TES CO data improved greatly (four-fold increase in signal-to-noise ratio) following a warm-up of the optical bench in early December 2005 (Rinsland et al., 2006). Therefore we consider here not only the period September 2004–April 2005 overlapping with the other satellite datasets, but also the period May 2005–April 2006 (with available data starting in July 2005), which includes data after the December 2005 bench warm-up.

2.4 SCIAMACHY

The Scanning Imaging Absorption SpectroMeter for Atmospheric CHartographY (SCIAMACHY) instrument was launched aboard ENVISAT in March 2002 with the Equator crossing time of 10:00 (Bovensmann et al., 1999; Burrows et al., 1995). SCIAMACHY measures solar backscattered radiation at 2.3 µm, which allows for nearly uniform
sensitivity through the tropospheric column though with no vertical resolution (Buchwitz et al., 2004, 2005; de Laat et al., 2006). Global coverage is obtained by SCIAMACHY for its nadir measurements in 6 days, but actual coverage is much reduced because of the low ocean reflectivity and the presence of clouds in the relatively large pixels (30 km × 120 km) (Buchwitz et al., 2007; Gloudemans et al., 2005, 2008). We use the version 0.6 retrieval from the University of Bremen (Buchwitz et al., 2004, 2005, 2007). The retrieval provides CO columns, 〈y〉, with vector averaging kernels a, related to the true vertical profile (z) by the following equation:

\[ \hat{y} = z_a + a(z - z_a) \]  

where \( z_a \) is a fixed a priori profile. Buchwitz et al. (2007) found that the Bremen retrieval was on average 10% higher than MOPITT CO columns with 20% standard deviation.

SCIAMACHY data have considerable noise, typically 10–100% of the total column. Here we use daily averaged data weighted by the reported instrument error and use available quality flags for data screening, which filter any negative columns that are produced during the retrieval process. We select data with the quality flag, which is part of the data product, and which to some extent corrects for cloud effects using simultaneously retrieved methane. We further sample using the cloud-free flag (as also done by Tangborn et al., 2009), which corresponds to a cloud fraction of no more than 0.1. Buchwitz et al. (2007) used data with maximum cloud fraction of 0.3 in their comparison with MOPITT. The cloud-free screening significantly reduces the number of measurements, especially over the oceans (Khlystova et al., 2009).

3 CO simulation in the GEOS-Chem CTM

GEOS-Chem is a global 3-D chemical transport model (CTM) driven by GEOS assimilated meteorological data from the NASA Global Modeling and Assimilation Office (GMAO) (http://acmg.seas.harvard.edu/geos/). The GEOS-Chem CO simulation was originally described by Bey et al. (2001) and more recently by Duncan et al. (2007). Here we use version 7-04-11 for the period spanning from 1 May 2004 through 30 April 2005. We use GEOS-4 meteorological data with 1° × 1.25° horizontal resolution and degrade the resolution in GEOS-Chem to 2° × 2.5° for the satellite data intercomparison and to 4° × 5° for the inverse model analysis. Combustion sources of CO include fossil fuel, biofuel, and biomass burning emissions, augmented following Duncan et al. (2007) by 19%, 19%, and 11% respectively to account for co-emitted nonmethane VOCs (NMVOCs). Additional CO sources include oxidation of methane, which produces CO in the atmosphere with an instantaneous yield of unity, and NMVOCs, which produce CO at the point of emission with a yield of 0.09–1.00 (Duncan et al., 2007). We compute CO loss and production from methane by using monthly mean 3-D OH concentration fields archived from a detailed oxidant-aerosol GEOS-Chem simulation (version 5-07-08) (Park et al., 2004). Our global mean tropospheric OH concentration is 10.8 × 10^6 molec/cm^3, which compares well with the multimodel mean of 11.1 ± 1.7 × 10^6 molec/cm^3 reported by Shindell et al. (2006). Our corresponding tropospheric lifetime of methyl chloroform against oxidation by OH is 5.3 years, somewhat shorter than those reported by Prinn et al. (2005) and Spivakovsky et al. (2000), 6.0 (+0.5, −0.4) years and 5.7 years, respectively. We initialize our simulation with a year-long GEOS-Chem spin-up simulation and subsequent rescaling to MOPITT CO columns (corrected for the 5% high bias), as done previously in Kopacz et al. (2009).

Previous versions of the GEOS-Chem CO simulation have been evaluated against observations from surface sites (Duncan et al., 2007; Goldstein et al., 2004; Liang et al., 2004; Weiss-Penzias et al., 2004), aircraft (Heald et al., 2003a; Hudman et al., 2008; Zhang et al., 2008), and satellites, including MOPITT and TES (Arellano et al., 2006; Heald et al., 2004; Kopacz et al., 2009; Zhang et al., 2006). The comprehensive model evaluation by Duncan et al. (2007) showed biases relative to the NOAA/GMD (Novelli et al., 2003) surface network data in the range ±10% in the Northern Hemisphere and up to −19% in the southern tropics.

Duncan et al. (2007) estimated a direct global emission of CO (excluding co-emitted NMVOCs) of 956–1086 Tg a\(^{-1}\) for 1988–1997, a period of downward emission trends in Europe and the US, but upward trend in Asia. They assigned an error of less than 25% on this global estimate. Their mean tropospheric OH concentration simulated for that period was 8.7–9.3 × 10^6 molec/cm^3, in agreement with CH\(_2\)CCl\(_3\) lifetimes estimates (Prather, 2001). The models in the Shindell et al. (2006) comparison included higher OH concentrations, and found a consistent underestimate of CO concentrations across models (including GEOS-Chem) of up to 40–60 ppb in spring at northern midlatitudes and in excess of 60 ppb over south-central Africa during the biomass burning season.

Figures 1, 2, and Table 1 show seasonal and annual emissions in our current GEOS-Chem simulation for May 2004–April 2005, taken as a priori for our source inversion. CO emissions from combustion amount to 858 Tg a\(^{-1}\), with an additional 140 Tg a\(^{-1}\) from oxidation of co-emitted NMVOCs. They are drawn from EDGAR 3.2FT2000 inventory (Olivier et al., 1999; Olivier and Berdowski, 2001) for the year 2000, implemented in GEOS-Chem by van Donkelaar et al. (2008). These are overwritten with the following regional inventories: the US Environmental Protection Agency National Emission Inventory for 1999 (EPA-NEI99) for the US with a 60% downward correction following Hudman et al. (2008) (NEI99_Hudman), the Big Bend Regional Aerosol and Visibility Observational (BRAVO) Study Emissions Inventory for Mexico (Kuhns et al., 2003), the Cooperative Programme for Monitoring and Evaluation of the Long-range Transmission of Air Pollutants in Europe (EMEP) inventory for Europe in 2000 (Vestreng and Klein, 2002), as well as Streets et al. (2006, 2003) anthropogenic
Figure 1. Average seasonal a priori CO emissions from fossil fuel, biofuel, and biomass burning for May 2004 - April 2005. See text for details.

Fig. 1. Average seasonal a priori CO sources from fossil fuel, biofuel and biomass burning for 1 May 2004–30 April 2005. See text for details.

Figure 2. Seasonal variation of total CO combustion sources from the contiguous US (NEI99 region), Europe (EMEP region) and E. Asia (20–50°N, 70–150°E). A priori values for fossil fuel are from the NEI99 inventory for the US (with Hudman et al. [2008] 60% downward correction), EMEP inventory for Europe and Streets et al. [2006] inventory for E. Asia. A posteriori values are from the inversion. Both a priori and a posteriori reflect total emission source, including direct emissions and rapid oxidation from co-emitted NMVOCs.

Fig. 2. Seasonal variation of total CO combustion sources from the contiguous US (NEI99 region), Europe (EMEP region) and E. Asia (20–50°N, 70–150°E). A priori values for fossil fuel are from the NEI99 inventory for the US (with Hudman et al. (2008) 60% downward correction), EMEP inventory for Europe and Streets et al. (2006) inventory for E. Asia. A posteriori values are from the inversion. Both a priori and a posteriori reflect total emission source, including direct emissions and rapid oxidation from co-emitted NMVOCs.
Table 1. Annual CO emissions.

<table>
<thead>
<tr>
<th>Region</th>
<th>Fossil fuel</th>
<th>Biofuel</th>
<th>Biomass burning</th>
<th>Total</th>
<th>Inverse model results</th>
</tr>
</thead>
<tbody>
<tr>
<td>US</td>
<td>35.2</td>
<td>2.5</td>
<td>2.6</td>
<td>40.2</td>
<td>49.5</td>
</tr>
<tr>
<td>Alaska and Canada</td>
<td>1.4</td>
<td>0.4</td>
<td>15.4</td>
<td>17.2</td>
<td>21.4</td>
</tr>
<tr>
<td>Europe</td>
<td>60.4</td>
<td>15.2</td>
<td>2.5</td>
<td>78.1</td>
<td>94.7</td>
</tr>
<tr>
<td>E Asia</td>
<td>136</td>
<td>67.1</td>
<td>12.8</td>
<td>216</td>
<td>354</td>
</tr>
<tr>
<td>SE Asia</td>
<td>43.6</td>
<td>45.7</td>
<td>83.4</td>
<td>173</td>
<td>306</td>
</tr>
<tr>
<td>S. America</td>
<td>15.8</td>
<td>16.6</td>
<td>86.6</td>
<td>119</td>
<td>183</td>
</tr>
<tr>
<td>Africa (NH)</td>
<td>27.4</td>
<td>21.4</td>
<td>74.9</td>
<td>124</td>
<td>175</td>
</tr>
<tr>
<td>Africa (SH)</td>
<td>6.48</td>
<td>10.1</td>
<td>74.0</td>
<td>90.3</td>
<td>168</td>
</tr>
<tr>
<td>Australia</td>
<td>4.1</td>
<td>1.3</td>
<td>17.2</td>
<td>22.6</td>
<td>40.5</td>
</tr>
<tr>
<td>Global</td>
<td>319</td>
<td>160</td>
<td>379</td>
<td>858</td>
<td>1350</td>
</tr>
</tbody>
</table>

Values are in units of Tg a⁻¹ for May 2004–April 2005. Oxidation of co-emitted NMVOCs from combustion contributes an additional 140 Tg a⁻¹ (a priori) and 217 Tg a⁻¹ (a posteriori). Oxidation of methane and biogenic NMVOCs contributes an additional 853 Tg a⁻¹ and 426 Tg a⁻¹ (a priori) and total of 1290 Tg a⁻¹ (a posteriori).

From the bottom-up emission inventories described in Sect. 3 and used as a priori for the inversion.

Inversion using MOPITT, AIRS, and SCIAMACHY (Bremen) data for May 2004–April 2005.

Contiguous 48 states. The a priori fossil fuel source is from the EPA NEI99 inventory, reduced by 60% on the basis of constraints from ICARTT aircraft observations in summer 2004 (Hudman et al., 2008).

The summer of 2004 saw unusually large boreal forest fire activity in Alaska and Canada (Pfister et al., 2005; Turquety et al., 2007)

European region (including European Russia) as defined by the EMEP emission inventory.

Includes China, Korea, and Japan, same as in Fig. 2.


Africa is separated into Northern and Southern Hemispheres as given by Chevallier et al. (2009), and includes the Arabian peninsula.

emissions for Asia in 2000 and China in 2001. Biomass burning emissions are from the interannual GFED2 inventory with monthly resolution (van der Werf et al., 2006). The combustion emissions differ from those used by Duncan et al. (2007). In the high summer fire season in the North American boreal region, we do not assume any emission injection above the boundary layer. This could cause an underestimate of vertical transport and thus an overestimate of surface emissions in the. A recent analysis of the heights of plumes from these fires shows that at least 10% of plumes were injected above the boundary layer at local time of 11:00–13:00 (Kahn et al., 2008; Val Martin et al., 2009). Additional CO sources come from oxidation of methane (853 Tg) and biogenic NMVOCs (426 Tg), which include isoprene, monoterpenes, methanol and acetone as described in previous studies (Arellano et al., 2006; Duncan et al., 2007; Heald et al., 2004; Kopacz et al., 2009).

Figure 3 compares our CO simulation with a priori sources (in red) with monthly mean CO concentrations (climatological 1988–2001 in black, 2004–2005 in blue) from the same NOAA/GMD surface sites as in Duncan et al. (2007). A posteriori model results shown in green will be discussed in Sect. 7. The model is too low in winter-spring in the extratropical Northern Hemisphere, consistent with the previous model studies reported by Shindell et al. (2006) and Duncan et al. (2007). One notable difference with Duncan et al. (2007) is our underestimate of Bermuda in winter-spring (2004–2005 not shown due to scarcity of data), reflecting our decrease of US CO emissions following Hudman et al. (2008). This will be discussed further in the context of the inverse model results. Our simulation in the southern tropics (Samoa) improves on Duncan et al. (2007), who found a larger underestimate for the seasonal maximum; this could reflect our use of GFED2 biomass burning emissions or more recent meteorological fields, and again will be discussed further in the context of the inverse model results.

Figure 4 compares our a priori model results at 710, 480, and 305 hPa with 2002–2007 monthly mean MOZAIC aircraft observations over selected locations. The winter-spring model underestimate in the Northern Hemisphere is apparent at all altitudes, consistent with the data from surface sites, although it dampens with altitude. Here also, a posteriori model results are shown in green and will be discussed in Sect. 7.

4 Intercomparison of satellite datasets

Figure 5 shows annual mean (May 2004–April 2005) CO columns from MOPITT, AIRS, SCIAMACHY Bremen. For TES, the mean is computed for July 2005–April 2006. There are obvious differences, which could reflect differences in instrument/retrieval properties (as described by the averaging
Figure 3. Seasonal variation of CO concentrations at remote surface sites. Climatological observations from NOAA/GMD (1988–2001) (Novelli et al., 2003) are shown in black, 2004–2005 observations are in blue. GEOS-Chem model values are shown in red (a priori sources) and in green (a posteriori sources). Note the differences in scale between panels.

Figure 6 shows scatterplots of satellite versus model CO columns for May 2004–April 2005 (except for TES, where we show July 2005–April 2006 observations and model). Individual points represent daily observations averaged over the 2° × 2.5° grid of the model. We report the resulting correlation coefficient (r) and slope of the reduced-major axis (RMA) regression line, which allows for error in both datasets, as well as the mean model-observed percentage difference. Also included in Fig. 6 is the model correlation with in situ measurements from the GMD and MOZAIC datasets (from Figs. 3 and 4), which provides an absolute reference. It shows r=0.84 with a slope of 0.75. The relative difference of annual mean model versus annual mean data is ~12%, indicating a mean model underestimate as discussed previously.

Differences between the model and satellite observations in Fig. 6 reflect model, retrieval, and instrument errors. The smoothing error described by the averaging kernel is applied to both the observations and the model and thus is not a cause of the differences. In fact, variability of this smoothing error from scene to scene could lead to the appearance of strong correlation in cases where the DOF are low (Luo et al., 2007b; Rodgers, 2000).

Figure 6 shows strong consistency between MOPITT and AIRS, based on their correlations with the model. The correlations reflect actual information from the instruments, as opposed to variability in the a priori, since the a priori is globally uniform and identical for both retrievals. There is no indication from Fig. 6 that variability in DOF contributes to the correlation of the observations with the model. DOF shown in Fig. 6 for MOPITT and AIRS are about 0.5 in the polar regions but much larger than 0.5 otherwise, indicating that most of the information comes from the measurement as opposed to the a priori; average values...
Figure 4. Seasonal variation of CO concentrations throughout the troposphere. Climatological aircraft observations from MOZAIC (2002–2007) (Nedelec et al., 2003) are shown in black, 2004–2005 observations are in blue. Vertical lines show interannual variability of monthly mean concentrations. GEOS-Chem model values are shown in red (a priori sources) and in green (a posteriori sources).

are 1.1 for MOPITT (1.4 in extra-polar regions), and 0.78 for AIRS (0.81 in extra-polar regions). Cloud screening is a likely reason for the higher MOPITT DOF. McMillan et al. (2010) demonstrate the anticorrelation of AIRS DOF with retrieved cloud fraction, indicating lower information content for cloudier AIRS retrievals. Further examination of MOPITT-AIRS comparisons with GEOS-Chem for individual hemispheres, land versus ocean, and individual seasons indicate statistics similar to the global values in Fig. 6. MOPITT shows a stronger winter-spring maximum than AIRS as well as a larger interhemispheric difference (Fig. 5). We elaborate further on regional differences between MOPITT and AIRS in the context of the inversion results in Sect. 6.

TES shows stronger correlation and less difference with GEOS-Chem compared to MOPITT or AIRS (Fig. 6). For the 2005–2006 data shown in Fig. 6 (mean DOF of 0.99), the correlation coefficient is 0.91 and the regression slope is 0.88. We find similar statistics for the 2004–2005 data, with mean DOF of 0.74. The high correlation reflects the variable a priori used by TES. To test the effect of the smoothing, we reprocessed TES retrieved columns and the corresponding GEOS-Chem columns using the same globally uniform a priori profile as used by MOPITT and AIRS. We find that the TES versus GEOS-Chem correlation coefficient drops to 0.81 and the slope drops to 0.74, yielding statistics similar to MOPITT and AIRS vs. GEOS-Chem. Although the model-TES correlation is very close to that of MOPITT and AIRS when TES and corresponding model are reprocessed with MOPITT a priori, the absolute values of the reprocessed TES CO columns and the corresponding model columns are much lower. Also the global annual mean model-data difference is −5% for TES vs. −16% for MOPITT and −13% for AIRS.
**Fig. 5.** Annual daytime average CO columns observed by the MOPITT, AIRS, TES and SCIAMACHY satellite instruments over the period 1 May 2004–30 April 2005 (TES data for July 2005–April 2006). White space indicates lack of data. SCIAMACHY data include “cloud-free” data only, AIRS data are restricted to retrievals with surface temperature > 250 K.

**Fig. 6.** Scatterplots of CO observational datasets vs. the GEOS-Chem model using a priori sources. Points represent daily observations averaged over the 2° × 2.5° grid of the model for the period May 2004–April 2005, with the exception of TES (July 2005–April 2006) and the GMD/MOZAIC data (monthly climatological averages as described in Figs. 3 and 4). The green dashed line is the 1:1 relationship. The red solid line is a reduced-major-axis (RMA) fit. Correlation coefficients and slopes are given inset. Symbols on the top three panels are colored by their degrees of freedom (DOF) for signal. Units are 10^{18} molecules cm^{-2} for the satellite panels and 10^2 ppb for the GMD/MOZAIC panel.
SCIAMACHY daily CO data have considerable noise and most assessments of these data have been on a monthly average basis to reduce the noise error (Buchwitz et al., 2007; de Laat et al., 2007). Figure 6 shows a correlation coefficient of 0.44 relative to GEOS-Chem and a regression slope of 0.56; the mean model-observations difference is −20%. This is at least qualitatively consistent with the thermal IR instruments.

5 The inverse model

Our inverse problem consists of optimizing the sources of CO by minimizing the mismatch between simulated (GEOS-Chem) and observed CO columns, accounting for constraints from a priori knowledge. Let the vector \( y_o \) represent the ensemble of CO column observations used in the inversion (as described in Sect. 2). \( y_m \) the corresponding model values, \( x \) the ensemble of CO sources to be optimized (state vector), and \( x_a \) the a priori estimate (described in Sect. 3 and shown in Fig. 1). Bayesian optimization assuming Gaussian errors involves minimization of the least-squares scalar cost function \( J(x) \):

\[
J(x) = (y_m - y_o)^\mathrm{T} \Sigma_1^{-1} (y_m - y_o) + (x - x_a)^\mathrm{T} \Sigma_a^{-1} (x - x_a)
\]

(5)

where \( \Sigma_2 \) and \( \Sigma_a \) are the observational and a priori error covariance matrices described below.

We use the GEOS-Chem model adjoint to solve the minimization problem \( \nabla_x J = 0 \) numerically, as described previously by Kopacz et al. (2009) in an inverse analysis of CO sources in East Asia in spring 2001 using MOPITT data. The GEOS-Chem adjoint was originally developed by Henze et al. (2007). We extend it here to include the adjoint of GEOS-4 convective transport, derived using the Tangent Linear and Adjoint Model Compiler (TAMC) software, and advective transport, using negative winds. We also include the satellite observation operators and their adjoints. The observation operators apply retrieval Eqs. (1)–(4) and compute corresponding columns for each observation scene.

We include in the inversion the observations for May 2004–April 2005 from MOPITT, AIRS, and SCIAMACHY Bremen, averaged over the \( 4^\circ \times 5^\circ \) resolution of GEOS-Chem used for the inversion. We exclude MOPITT and SCIAMACHY Bremen data in polar regions (>60° latitude), where they are of lower quality. We thus have 305,484 observations from MOPITT, 923,234 observations from AIRS, and 25,773 observations from SCIAMACHY Bremen on the \( 4^\circ \times 5^\circ \) grid. Aircraft validation data (Sect. 2) show a 5% MOPITT positive bias and we correct for it for the purposes of the source inversion. AIRS validation indicates a positive bias of 7–10% between 300 and 900 hPa in the Northern Hemisphere (McMillan et al., 2010), but we do not correct for it here. There is no clear indication that bias correction for SCIAMACHY data is needed. TES data is used as an independent set of observations to evaluate our a posteriori results.

We optimize the CO combustion sources at the \( 4^\circ \times 5^\circ \) grid resolution of the GEOS-Chem model and monthly temporal resolution, over the whole year from 1 May 2004 to 30 April 2005. Optimization is only for grid squares with non-zero combustion sources in the a priori (Fig. 1). We also optimize the global CO source from oxidation of methane and biogenic NMVOCs as a single variable with monthly temporal resolution. Our state vector \( x \) thus has 18,420 elements.

The observational error covariance matrix \( \Sigma_2 \) includes contributions from the measurement error, GEOS-Chem model error, and representation error. We estimate the total observational error with the Relative Residual Error (RRE) method (Heald et al., 2004; Kopacz et al., 2009; Palmer et al., 2003). This method attributes the mean of model-observation differences for a given grid square and season (month in the case of AIRS) to an error in CO sources, and the residual to observational error. We thus find that the highest observational errors are for SCIAMACHY (up to 70–100% in high northern latitudes). MOPITT observational errors are in the 10–30% range, highest over pollution outflow regions. AIRS errors are similar to MOPITT but lower (as low as 5% in remote ocean regions), reflecting the lower DOF. Error correlations between observations can be neglected at the \( 4^\circ \times 5^\circ \) resolution used for the inversion (Heald et al., 2004), so that \( \Sigma_2 \) is diagonal.

The a priori error covariance matrix \( \Sigma_a \) includes a uniform error of 50% for combustion sources and 25% for the global oxidation source, the latter as used in previous studies (Heald et al., 2004; Kopacz et al., 2009). The monthly errors are assumed uncorrelated so that \( \Sigma_a \) is diagonal. The a priori terms in Eq. (5) do not contribute substantially to minimization of the cost function. This implies that a posteriori sources differ relatively little from the a priori values, and not necessarily that they are independent of the choice of a priori sources.

6 Optimized monthly CO sources

6.1 General results

Figure 7 shows the annual mean correction factors to the a priori emission estimates and Table 1 gives the annual total emissions for the largest source regions. The emission correction factors are ratios of a posteriori to a priori emissions. Emissions increase almost everywhere relative to the a priori, but the global CO source from oxidation of methane and biogenic NMVOCs increases by less than 1% (Table 1). Our posteriori annual global estimate for direct CO emissions is 1350 Tg a\(^{-1}\) (+217 Tg a\(^{-1}\) from oxidation of co-emitted VOCs), a 60% increase from a priori. This is within 25% of results from previous (global and annual) inversions of satellite (MOPITT) measurements: 1091 Tg a\(^{-1}\) using the MOZART model (Pétron et al., 2004), 1342–1502 Tg a\(^{-1}\) using GEOS-Chem (Arellano et al., 2004, 2006) and 1695 Tg a\(^{-1}\) using the IMAGES model (Stavrakou et al., 2003). This method attributes the mean of model-observation differences for a given grid square and season (month in the case of AIRS) to an error in CO sources, and the residual to observational error. We thus find that the highest observational errors are for SCIAMACHY (up to 70–100% in high northern latitudes). MOPITT observational errors are in the 10–30% range, highest over pollution outflow regions. AIRS errors are similar to MOPITT but lower (as low as 5% in remote ocean regions), reflecting the lower DOF. Error correlations between observations can be neglected at the \( 4^\circ \times 5^\circ \) resolution used for the inversion (Heald et al., 2004), so that \( \Sigma_2 \) is diagonal.

The a priori error covariance matrix \( \Sigma_a \) includes a uniform error of 50% for combustion sources and 25% for the global oxidation source, the latter as used in previous studies (Heald et al., 2004; Kopacz et al., 2009). The monthly errors are assumed uncorrelated so that \( \Sigma_a \) is diagonal. The a priori terms in Eq. (5) do not contribute substantially to minimization of the cost function. This implies that a posteriori sources differ relatively little from the a priori values, and not necessarily that they are independent of the choice of a priori sources.
and Müller, 2006). We also compare our results to regional and seasonal studies as we describe details of our results in Sect. 6.2.

Fortems-Chiney et al. (2009) applied MOPITT and IASI 700-hPa CO concentrations individually to constrain global CO sources during July to November 2008 (outside the biomass burning season), using the adjoint of the LMDZ-INCA model. They obtain a 643 Tg global total (for 5 months) using IASI data and 649 Tg using MOPITT data. Our a posteriori estimate for the same months is comparable, 628 Tg of direct emissions, and 95 Tg of co-emitted VOCs.

6.2 Seasonal and regional results

A striking result of the inversion is the seasonal variation of the source correction at northern mid-latitudes. Figure 8a shows this seasonal variation for North America. We find no need for correction over the US in summer, supporting the previous 60% downward correction to the NEI99 emission inventory, which was derived by Hudman et al. (2008), and which we included in our a priori. This correction was based on ICARTT summer aircraft measurements. In an independent analysis using aircraft and tower CO data, Miller et al. (2008) found the NEI99 emissions to be too high by a factor of three in summer and two in spring. They suggest that spring emissions are higher because of domestic wood burning and less efficient combustion for mobile sources. We find indeed that emissions are higher in seasons other than summer, in a way that is not represented by the seasonal variation in NEI99 inventory (Fig. 2). A posteriori US emissions in winter (DJF) are on average 50% higher than in summer, while spring (MAM) and fall (SON) are 25% higher with the largest effects (exceeding a factor of two) in the Northeast and Midwest. The spring estimate in Miller et al. (2008) may be larger than ours because they focused their analysis on the Midwest and Northeast. Parrish (2006) in his evaluation of NEI99 emission estimates against fuel-based inventory and surface measurements, suggests that while US on-road emissions are overestimated, as corrected by Hudman et al. (2008), many uncertainties remain including vehicle cold starts during the cold months. The cause of the dramatic emission overestimate in summer also (60%) remains unclear.

The spring underestimate in the Yucatan Peninsula occurs during a period of large biomass burning in the region. Inverse model results for the boreal forest fire regions of Alaska and western Canada in summer 2004 indicate a 30% underestimate in the GFED2 biomass burning inventory, corresponding to an a posteriori emission estimate of 24 Tg. A previous inversion for that region and season by Pfister et al. (2005) using MOPITT data indicated an a posteriori estimate of 30 Tg. A detailed bottom-up fire emission inventory for the region also found a total of 30 Tg (Turquety et al., 2007).

Figure 8b shows a qualitatively similar picture for Europe. Summer a posteriori emissions are close to the a priori (the EMEP inventory), but we see large underestimates in other seasons ranging up to 30–70% in urban regions of western Europe. The seasonal correction corresponds to urban areas and cold months and is weaker than in the US (Fig. 2), which might reflect a smaller effect from vehicle cold starts. The upward corrections also correspond to areas of large emissions and as with any least-squares inversion, there is a possibility that the optimization might disproportionately focus on largest sources. Our previous study (Kopacz et al. 2009) analyzed this possibility, however and found no evidence of consistent bias.

In Asia (Fig. 8c), the inversion finds again that the Streets et al. (2006) inventory for China is largely correct for summer but too low by up to a factor of 2 in winter. The a priori anthropogenic sources other than Streets et al. (2003) in China do not include seasonal variability. A recent update of the Asian inventory (not used here) does include seasonality (Zhang et al., 2009). Our a posteriori annual estimate for Chinese combustion emissions is 267 Tg. We see from Fig. 2 that the seasonal amplitude of the correction is much larger than for Europe or the US. The Streets et al. (2006) inventory does not include seasonal variation from residential heating, and the seasonal peak in the a priori in March–April in Fig. 2 is from biomass burning. Vehicle cold starts might again be an explanation. Analyzing hourly evolution of Beijing surface CO concentrations, Han et al. (2009) find that residential heating emissions in Beijing might in fact be overestimated in current inventories, while emissions from non-domestic sources such as transport (and including cold starts) might be underestimated. Cold start emissions depend on many non-temperature factors, such as time since last operation of the vehicle as well as its operation after starting (Wenzel et al., 2000). Meanwhile, emission factors for domestic burning of coal and wood can differ as well. Therefore, more detailed analysis of our findings requires further regional perspective.

Figure 8c also indicates a consistent underestimate of Indian emissions with little seasonal variation except for northern India in spring. The underestimate could be due to biomass burning, which in India is largely absent in the GFED2 inventory and would be expected to peak in northeastern India in March–April (Heald et al., 2003b). The underestimate over western Siberia in spring could also reflect a large seasonal biomass burning source missing from current inventories (Warneke et al., 2009).

The large biomass burning areas in southeastern Asia, in particular Indonesia and Malaysia appear to be consistently underestimated by more than 100% with respect to the GFED2 inventory. Our previous work (Kopacz et al., 2009) focused on spring 2001 using MOPITT data derived an annual source of 113 Tg in the SE Asia-Indonesia-Philippines region (Kopacz et al., 2009). Our current estimate for the region is 256 Tg. One reason for the difference is the ENSO cycle: 2001 was a La Niña year, while 2004 was a weak
Fig. 7. Annual mean correction factors to the a priori combustion sources of CO from Fig. 1 as derived from the adjoint inversion of MOPITT, AIRS, and SCIAMACHY CO columns for May 2004–April 2005.

Fig. 8. Ratio of a posteriori to a priori CO emission estimates for different seasons in (a) North America (b) Europe and Middle East, (c) Asia, (d) Africa and S. America.
El Niño year with above-average biomass burning (Edwards et al., 2006a). Also, the MOPITT v3 (unlike v4) retrieval algorithm is not applied to high signal values and thus high CO concentrations are not obtained, introducing a low bias over high emission regions like this one. Here both model-MOPITT and model-AIRS a priori differences are negative, indicating low model bias, but a posteriori model biases are positive (smaller with respect to AIRS), indicating overcompensation for the original bias. This overcompensation can be expected, given a large least-squares correction to a large source, and should be kept in mind when comparing a posteriori results to independent observations.

In biomass burning dominated emission regions of Africa and S. America our a posteriori annual estimates are 343 Tg and 183 Tg (Table 1), much larger than the GFED2 inventory as shown by the seasonal correction factors in Fig. 8d. The large underestimates in the GFED2 inventory are found especially during the biomass burning season (August–October in S. America and July–October in southern Africa), with some underestimate seen as early as August and as late as March.

Figure 9 shows that while the inversion improves model bias over eastern Brazil, it worsens it over the interior. The inversion points to the same difficulty in southern Africa, where model bias is reduced more over the source region than over the outflow. In the tropical biomass burning regions we find inconsistencies among the datasets as visible in the a posteriori model bias in Fig. 9.

Chevallier et al. (2009) previously applied 2000–2006 MOPITT 700-hPa CO concentrations to an adjoint inversion of CO sources in Africa using the LMDZ-INCA model and the GFED2 biomass burning inventory as a priori. The aim of their study was to constrain African biomass burning emissions (bounded by 40S–40N, 25W–60E) over the years and seasons. They also considered the GMD station Ascension, where their prior model bias of –5% was reduced to 0, as well as other stations with comparable a posteriori improvement. This contrasts with our a posteriori disagreement at that station. Their results indicate the need for both increases and decreases of African emissions depending on season and location. Their a posteriori continental estimates for 2004
Fig. 8. Continued.
Fig. 9. Relative a priori and a posteriori model biases against MOPITT, AIRS and SCIAMACHY during September–November 2004. The middle row shows results from the inversion using observations from all three instruments, while the bottom row shows results from inversions using individual instruments.

(255 Tg) and 2005 (283 Tg) are lower, than our value of 343 Tg. This reflects some inconsistency between MOPITT and AIRS, with AIRS implying higher emissions for that region than MOPITT as discussed below.

In addition to data inconsistencies, our estimates are subject to uncertainty in OH concentrations, particularly in northern extratropics, where OH concentrations are not well constrained by methyl chloroform lifetime, as well as unknown errors in meteorological data and errors in VOC concentrations. Lower OH concentrations would lower our emission estimates, while errors in meteorology and VOC concentrations could have a varied effect, the latter especially affecting CO seasonal cycle in some regions (Arellano and Hess, 2006).

6.3 Individual versus combined datasets

Figure 9 shows the relative model biases against the MOPITT, AIRS, and SCIAMACHY observations, for the a priori and a posteriori simulations. The largest bias reduction is for AIRS, reflecting the much larger number of AIRS observations (923,234) compared to MOPITT (305,484) and SCIAMACHY (25,773). There is no objective way to weigh each dataset differently, other than through proper characterization of observational error. To better understand the contributions of each dataset to the inversion we performed source inversions using single-instrument data for the 3-month period of September–November 2004. This also tests the consistency among the inverse results and, by extension, the consistency of the datasets themselves. Figure 9 shows model a priori bias, model a posteriori bias from the three satellite inversion and model a posteriori bias derived from source inversion using individual datasets, all with respect to AIRS, SCIAMACHY and MOPITT data.

Globally, the model bias changes from a priori to a posteriori in the joint inversion are from −6% to −0.3% for AIRS, from −4% to +10% for MOPITT, and from −10% to +2% for SCIAMACHY. In contrast, individual dataset inversions yield a larger a posteriori bias for AIRS (−0.7%), a smaller a posteriori bias for MOPITT (+1%), and no significant improvement for SCIAMACHY (a posteriori bias still −10%). The lack of success in the inversion using SCIAMACHY only reflects the small amount of data along with the high observational error. Our inversion corrected for the 5% MOPITT high bias determined from aircraft validation data, as described in Sect. 5, but made no such correction for AIRS due to lack of validation information.

In summary, it is overall beneficial to combine the data to improve the model bias, but based on the model bias amounts, MOPITT column concentrations (with the correction for the 5% high bias) appear lower than AIRS or SCIAMACHY, especially in the Southern Hemisphere. The largest contribution to the cost function (78%) and largest
difference comes from the model-AIRS discrepancy, which is much lower in the individual inversion. Figure 9 shows that the three satellite inversion best improves the model-AIRS disagreement, further suggesting that AIRS data tends to dominate the overall source estimates. Unless AIRS observational errors were much larger than those of MOPITT and SCIAMACHY, we expect AIRS CO to dominate the a posteriori source corrections, given the relatively large number of AIRS data and potential inconsistencies. If the datasets were perfectly consistent, improvement in model-AIRS agreement would perfectly map onto model-MOPITT and model-SCIAMACHY agreement (as seen over NH Pacific and Middle East). In fact, AIRS observational errors are lower than those of MOPITT and SCIAMACHY, but that reflects lower AIRS DOF and should not affect the information balance. Since the difference between model and observation for low DOF is also small, it prevents large contributions to the inversion from low signal data.

Figure 9 further allows inspection of the robustness of regional features in the source correction. We see that while the bulk calculations reveal overall consistency, regional discrepancies increase the model-data disagreement. The correction at northern mid-latitudes is consistent across all three instruments, but we also see that the joint inversion leads to positive biases with respect to MOPITT and SCIAMACHY through much of the Southern Hemisphere. This implies that AIRS is higher than MOPITT and SCIAMACHY (at least during September–November 2004), and that the difference cannot be fully explained by lower AIRS DOF. This conclusion is consistent with findings of earlier comparison studies (Yurganov et al., 2008, 2010). Areas where the joint inversion did not improve model-data agreement are also not well constrained by individual dataset inversions (e.g. in S. America). Figure 9 also shows that using individual datasets to constrain CO sources can yield different results than combining the data. However, as each dataset has been thoroughly evaluated by its retrieval team, they should be combined together for a balance of information.

It follows then that the emission correction factors from the individual dataset inversions corresponding to model bias shown in Fig. 9 are generally, but not entirely consistent. The three satellite inversion correction patterns are common in each of the individual dataset inversions. For September–November 2004, all datasets find a large (~100%) underestimate of southern African biomass burning and a similar pattern of underestimate and overestimate of biomass burning in the Amazon, but of different magnitudes. The only consistent difference is that MOPITT and SCIAMACHY corrections are more localized, while AIRS finds large areas to be underestimated. A few other regions also show opposite signs of corrections from different instruments, but generally, the differences are confined only to the magnitude of the correction.

7 Comparison with independent measurements

Figure 3 shows the a posteriori model CO compared against in situ observations from the GMD network. All stations in the Northern Hemisphere show considerable improvement relative to the a priori simulation. The winter-spring underestimate is largely corrected. The phase and amplitude of the seasonal cycle in the model match the observations, supporting the seasonally varying corrections to the northern mid-latitude emissions and implying consistency between the satellite and surface data. The inconsistency at Barrow in summer reflects the anomalous fire conditions in summer 2004, not reflected in the GMD data, which represent background conditions.

The a posteriori model comparison with MOZAIC aircraft observations in Fig. 4 also shows large improvement at all extratropical locations and complete correction of the winter-spring underestimate at the different altitudes. The seasonal phase and amplitude are well reproduced.

No such improvement in fitting the surface observations is found for the Southern Hemisphere sites in Fig. 3. The simulation with a posteriori sources fares generally worse than the a priori, although there is an improvement in the amplitude and phase of the seasonal cycle at all stations in the extratropics. This suggests an overestimate of the biomass burning source in the southern tropics constrained by the AIRS data, as suggested also in Fig. 9 by the results for MOPITT.

Since we did not use TES CO data in the source inversion, we use it as an additional independent set of measurements to verify our a posteriori results. Figure 10 shows a priori and a posteriori global TES correlations against GEOS-Chem for the 2004–2005 period. The a posteriori correlation coefficient is $r=0.91$, same as the a priori, but the slope of the regression line increases from 0.89 to 1.04, indicating a better fit.

8 Conclusions

We applied the adjoint of the GEOS-Chem CTM to a global inversion of CO sources as constrained by three satellite datasets (MOPITT, AIRS, SCIAMACHY). The inversion used a full year of data (May 2004–April 2005) and optimized CO combustion sources at a spatial resolution of $4^\circ \times 5^\circ$ and monthly temporal resolution. The optimization also included a monthly global source from oxidation of methane and biogenic NMVOCs. Results were evaluated with independent CO observations from surface sites (NOAA GMD network), aircraft (MOZAIC), and satellite (TES).

An important first step was to evaluate the consistency of the satellite datasets used in the inversion. Here we used GEOS-Chem with a priori sources as an intercomparison platform. We showed that MOPITT, AIRS, and TES (all observing in the 4.7 µm thermal IR band) are consistent overall,
and that apparent differences in the data are driven mainly by different averaging kernels and a priori information. SCIAMACHY (observing in the 2.3 µm solar IR band) is considerably noisier. All instruments as well as the GMD and MOZAIC in situ measurements show a consistent low bias in the GEOS-Chem simulation with a priori sources.

Our inversion yields an a posteriori best estimate of 1350 Tg for direct CO emissions from combustion, with 217 Tg from oxidation of co-emitted NMVOCs. This represents a 60% underestimate of bottom-up inventories, but is within 25% of recent top-down estimates (Arellano et al., 2004, 2006; Pétron et al., 2004; Stavrakou and Müller, 2006). The CO source from oxidation of methane and biogenic NMVOCs changed by <1% from our a priori of 1280 Tg.

Using the a posteriori sources gives a significantly improved fit of the GEOS-Chem simulation to the in situ and TES observations taken as independent datasets.

A striking feature of our results is the larger-than-expected seasonal variation of CO emissions at northern mid-latitudes. Emissions in winter are 50% higher than in summer in the US and Europe, and up to 100% higher in winter in E. Asia. Such large seasonal variation is not recognized in current bottom-up inventories. We hypothesize that it could be due to a combination of emissions from residential heating and vehicle cold starts. Our annual a posteriori estimate is 49.5 Tg for the US (48 states), 94.7 Tg for Europe, and 354 Tg for E. Asia (with 267 Tg for China alone). Implementing this seasonal variation of emissions in the GEOS-Chem simulation corrects the long-standing model biases in simulating CO in the northern extratropics in winter-spring (Shindell et al., 2006), and provides in particular a good fit to the ensemble of GMD and MOZAIC observations.

Our inverse model results indicate a large underestimate of tropical biomass burning in the GFED2 inventory (van der Werf et al., 2006). Annual a posteriori emission estimates are 343 Tg a⁻¹ for Africa and 183 Tg a⁻¹ for South America. However, the consistency among the satellite datasets is not as good in the Southern Hemisphere as in the north. In particular, AIRS implies larger biomass burning estimates than MOPITT or SCIAMACHY or the GMD surface sites, likely due to AIRS previously found high southern bias. Despite using numerous datasets and a high resolution optimization, our study is still only a first step towards a detail top-down understanding of CO emissions. Many regional details of source estimates and regional dataset discrepancies can be merely highlighted here.

Acknowledgements. This work was supported by the NASA Atmospheric Chemistry Modeling and Analysis Program and by NASA Headquarters under the Earth System Science Fellowship Grant NGT5 06-ESSF06-45 to Monika Kopacz. The authors acknowledge the strong support of 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 presently funded by INSU-CNRS (France), Meteo-France, and Forschungszentrum (FZJ, Julich, Germany). The MOZAIC database is supported by ETHER (CNES and INSU-CNRS). MK would also like to thank Christopher Holmes, Eric Leibensperger, Kevin Wecht, Jos de Laat, Annemieke Gloudemans and Ilse Aben for useful insight and discussions.

Edited by: B. N. Duncan
References


Edwards, D. P., Petron, G., Novelli, P. C., Emmons, L. K., Gille, J. C., and Drummond, J. R.: Southern Hemisphere carbon monoxide interannual variability observed by Terra/Measurement of


