A database and tool for boundary conditions for regional air quality modeling: description and evaluation

B. H. Henderson[1], F. Akhtar[2], H. O. T. Pye[3], S. L. Napelenok[3], and W. T. Hutzell[3]

[1]Environmental Engineering Sciences, University of Florida, Gainesville, FL, USA
[2]CSC, Alexandria, VA, USA
[3]Atmospheric Modeling and Analysis Division, US EPA, RTP, NC, USA

January 16, 2014

Abstract. Transported air pollutants receive increasing attention as regulations tighten and global concentrations increase. The need to represent international transport in regional air quality assessments requires improved representation of boundary concentrations. Currently available observations are too sparse vertically to provide boundary information, particularly for ozone precursors, but global simulations can be used to generate spatially and temporally varying Lateral Boundary Conditions (LBC). This study presents a public database of global simulations designed and evaluated for use as LBC for air quality models (AQMs). The database covers the contiguous United States (CONUS) for the years 2001–2010 and contains hourly varying concentrations of ozone, aerosols, and their precursors. The database is complimented by a tool for configuring the global results as inputs to regional scale models (e.g., Community Multiscale Air Quality or Comprehensive Air quality Model with extensions). This study also presents an example application based on the CONUS domain, which is evaluated against satellite retrieved ozone and carbon monoxide vertical profiles. The results show performance is largely within uncertainty estimates for ozone from the Ozone Monitoring Instrument and carbon monoxide from the Measurements Of Pollution In The Troposphere (MOPITT), but there were some notable biases compared to Tropospheric Emission Spectrometer (TES) ozone. Compared to TES, our ozone predictions are high-biased in the upper troposphere, particularly in the southern during January. This publication documents the global simulation database, the tool for conversion to LBC, and the evaluation of concentrations on the boundaries. This documentation is intended to support applications that require representation of long-range transport of air pollutants.

1 Introduction

The role of hemispheric transport of air pollutants is increasingly a focus of regional pollution studies (Lin et al., 2000, 2012; Reidmiller et al., 2009). The growing emphasis reflects three factors: (1) the National Ambient Air Quality Standards have been tightened (40 CFR 50.10); (2) influence of international activities has increased average hemispherically transported pollutants (Cooper et al., 2010; Fiore et al., 2009; Oltmans et al., 2006, 2010) and (3) long-range transport can have episodic strong influence (Fiore et al., 2002). Thus, model attainment demonstrations must achieve lower pollutant concentrations fields with a higher uncontrollable fraction. Under these conditions, it is imperative for the model to include long-range transported air pollution concentrations and accurately represent their variability in time and space. The long-range transported air pollutants are primarily communicated to air quality models (AQMs) through the lateral boundary conditions (LBC). This paper documents the development and availability of a resource that provides LBC for the air quality modeling community.

The surface level ozone concentrations have a 10–15 ppb sensitivity to LBC values even in locations relatively far from the boundary (Napelenok et al., 2011). Much of the model sensitivity can be attributed to high mixing ratios ($\mathrm{O_3}$=100–1000 ppb) in the upper troposphere/lower stratosphere (Krueger and Minzner, 1976; Lacis et al., 1990; Warneck and Williams, 2012). The high concentrations aloft are influenced by local emissions, international transport (Dentener et al., 2010; Lin et al., 2012), and stratosphere-troposphere-exchanges (Bourqui et al., 2012; Cui et al., 2009; Lefohn et al., 2011). The LBC, particularly at high altitude, is a mechanism of communicating each of these sources to the contiguous domains often used in regional air quality simulations.

Previously, LBC have come from a variety of sources and have been evaluated indirectly. The Community Multiscale Air Quality (CMAQ; Foley et al., 2010) model originally used "clean air" estimates or observations averaged over space and time, but preserving the vertical dimension where possible (e.g., ozone based on Logan et al., 1999). These vertical profile lateral boundary conditions (PLBC) have obvious limitations. The observations used to construct PLBC are sparse in space and time and, therefore, interpolation and extrapolation are unavoidable. As a result, variability in space and time is lost. Although utilizing "clean air" estimates is still common (Gégo et al., 2008; Godowitch et al., 2008; Smyth et al., 2009; Zhang et al., 2004), increasingly publications recognized these limitations and the growing availability of global simulations to provide estimates of air pollution concentrations with time resolution ranging from hourly to seasonal mean (Appel and Gilliland, 2006; Barna and Knipping, 2006; Fu et al., 2009; Hogrefe et al., 2008; Jiménez et al., 2007; Lam and Fu, 2009; Nghiem and Oanh, 2008; Schichtel et al., 2005; Valari et al., 2011). By themselves, these global simulations are too coarse for regional/urban air quality standard attainment demonstrations, but they offer a potential source of LBC for regional/urban AQM (Appel and Gilliland, 2006; Lam and Fu, 2009; Song et al., 2008).

The importance of evaluating LBC is evident in sensitivity analysis (Barna and Knipping, 2006; Jiménez et al., 2007; Napelenok et al., 2008), but most LBC evaluations are indirect. When modeling the contiguous United States (CONUS), most of the LBC are over water. As mentioned above, these locations have a paucity of observational data. As a result, the accuracy of the LBC inputs are evaluated based on alternate locations. For example, Lam and Fu (2009) first evaluated model predictions based on three ozonesondes sites over the CONUS (Trinidad Head, CA; Boulder, CO; Huntsville, AL). They further indirectly evaluated the LBC fitness based on model performance at surface locations. Although air quality models have many degrees of freedom to isolate LBC, this type of indirect evaluation has been necessary. Even these indirect evaluations concluded that GEOS-Chem LBC (GLBC) outperformed clean air profiles and climatological averages (Appel and Gilliland, 2006; Lam and Fu, 2009; Song et al., 2008). This conclusion gives some credence to the GLBC values, but in this report, we will further evaluate the GLBC using space/time coincident measurements available from satellite retrievals.

This document is structured according to the process of creating and evaluating LBC. The first section describes the details of the GEOS-Chem simulations used to create a database of global concentration fields for LBC. The second section documents the design, components, and functionality of the tool designed to create GLBC from GEOS-Chem outputs. The third section details the methods and results of evaluating GLBC using satellite observations. The conclusions review the usability of the tool and the fitness of database results. Finally, we discuss the availability of the LBC tool and global simulation database for the community.

2 GLBC simulation database

While LBC may be improved by global atmospheric modeling, the development and testing of global models is beyond the resources and scope of many air quality modeling studies. In order to provide users of regional AQM with global model information for boundary conditions in regional domains, a series of GEOS-Chem simulations have been conducted and are available for download with tools to produce to regional model ready boundary files.

GEOS-Chem is active engaged in research projects with scientific groups across the world continuously improving the model code, chemistry formulation, and input information (Details of the ongoing work on GEOS-Chem can be found at the model wiki page: http://wiki.seas.harvard.edu/geos-chem/). Continual improvements to the model and a variety of chemistry, meteorology, and emission options within GEOS-Chem poses a challenge for regional air quality modelers in choosing the optimal model setup for generating LBC.

To address this, we have conducted a series of preliminary GEOS-Chem simulations at $2^\circ \times 2.5^\circ$ horizontal resolution spanning multiple model release versions and input options. Hourly concentrations for North America from all of these simulations are archived and available for download. Due to data storage considerations, only the hourly values for gridcells containing and surrounding the contiguous United States are archived (Fig. 1). Plans are underway to expand availability to global coverage. For each day, we archive to composition files. GEOS-Chem requires two output files because some explicit species are not typicall saved. To reduce computational burden, GEOS-Chem combines several chemical species into "tracer" groups at time of advection. These tracer groups are then converted back into chemical species ("cspec") during the chemical calculations. Since some chemical species are important when mapped to regional models (Pye and Napelenok, 2013), both the GEOS-Chem tracer and cspec arrays are included in the LBC archive.

Details of the model setup for each of the available simulations are listed in Table 1. Table 1 details combinations of GEOS-Chem model versions, chemistry version, meteorology datasets, shipping emissions, and time period covered. All simulations used GEOS-Chem's $\mathrm{NO_x}$-$\mathrm{O_x}$-hydrocarbon-aerosol configuration with the optional Secondary Organic Carbon Aerosol module enabled. Versions of the chemical mechanism will be discussed further below. Whenever possible, the simulations follow GEOS-Chem manual recommended settings. The Sparse Matrix Vectorized Gear-based solver (Jacobson and Turco, 1994) is employed to solve the system of partial differential equations representing emissions and chemistry. Convection was solved using non-local planetary boundary layer and solving cloud convection.

Emissions for these simulations closely follow the default configuration of GEOS-Chem. For emissions, the Emission Database for Global Atmospheric Research (EDGAR) provided global anthropogenic emissions (Berdowski et al., 2001) with regions being overwritten where available. Regional anthropogenic emissions were provided by specific databases for the United States (NEI2005; US EPA, 2013), Europe (UNECE/EMEP; Vestreng and Klein, 2002), Mexico (BRAVO; Kuhns et al., 2003), Canada (CAC, Environment Canada, 2013), and Asia (INTEX-B, Streets et al., 2003, 2006). In addition, the emissions included additional source: lightning $\mathrm{NO_x}$ (Ott et al., 2010; Pickering et al., 1998; Price and Rind, 1992), soil $\mathrm{NO_x}$ (Wang et al., 1998; Yienger and Levy, 1995), oceanic Dimethyl Sulfide, volcanic $\mathrm{SO_2}$, sea salt, wind-blown mineral dust, wild-fires from the Global Fire Emissions Database (Werf et al., 2006) and biogenic volatile organic compound emissions from Model of Emissions of Gases and Aerosols from Nature (MEGAN) version 2.1 (Guenther et al., 2012).

Two versions of the chemistry were initially evaluated because of recent updates to GEOS-Chem's the chemical mechanism. An update in the chemistry mechanism between Between GEOS-Chem versions v8-02-01 and v8-02-04, the isoprene nitrate yield was decreased. Decreasing the yield of isoprene nitrate enhances radical cycling, which will increase in simulated ozone concentrations. Because modeled ozone concentrations already have high positive biases in North America (Mao et al., 2013), this bug fix may lead to increased ozone biases in regional models by inflating the amount of ozone entering the regional domain from the boundaries. Improvements to halogen and heterogeneous aerosol chemistry have shown promise in reducing this high-bias (Mao et al., 2013), but are not included, as these updates are still the subject of continuing research. Since the goal of including boundary conditions is to reduce overall bias within the regional simulation, we recommend using the chemistry mechanism from GEOS-Chem version v8-02-01 without updated isoprene nitrate when results are used for boundaries in applied regional simulations.

Performing simulations from 2001 to 2010 required using two meteorological datasets. The GEOS-5 dataset (Molod et al., 2012) was used to drive GEOS-Chem simulations from 2004 to 2012, but was not available before. The MERRA dataset (Rienecker et al., 2011) was available from 2001 to 2010. Using the MERRA dataset, however, is only supported by GEOS-Chem version 9. The version 9 also includes other updates (full documentation available at http://acmg.seas.harvard.edu/geos/).

Several simulations using different combinations of model code and meteorological datasets were conducted (Table 1). Preliminary evaluations showed best performance from the GEOS-Chem version 8 simulations with GEOS-5 meteorology, version 8-02-01 chemistry, and ICOADS shipping emissions. Model version 9 with MERRA meteorology is also made available for the years 2001 through 2003. When 2001-2003 boundaries are necessary, additional evaluation should be performed for that application. Based on preliminary analysis, only the results from MERRA (model v9) and GEOS-5 (model v8) with v8-02-1 chemistry have been archived. The evaluation section of this paper will focus on the GEOS-Chem version 8-03-02 model with version 8-02-01 chemistry, and ICOADS shipping from 2006 to 2010.

3 GLBC tool description

Model compound translation (GEOS-Chem to regional model compounds) and spatial mapping of the global output to LBC are served by two distinct components in the GLBC tool. Model compound translation is performed by a Python (python.org) pre-processor, and a Fortran program handles spatial mapping. A flowchart of the overall program is shown in Fig. 2 and each component is described below.

3.1 Python pre-processing

The Python pre-processor interprets model configurations and user inputs to apply appropriate scaling. Both GEOS-Chem and CMAQ have several chemistry/aerosol configurations that continue to evolve. The pre-processor interprets configurations files and provides failsafe measures to prevent mapping of incorrect model versions and highlight potential errors. In addition, the pre-processor is able to apply appropriate unit conversions when appropriate.

To perform these tasks, the pre-processor must first interpret the model gas-phase and aerosol-phase configurations. From CMAQ, the pre-processor requires the namelists (*.nml) or include files (*.EXT) that describe the gas-phase (GC_*), aerosol (AE_*), non-reactive (NR_*), and tracer (TR_*) species. From GEOS-Chem, the pre-processor requires the tracer_info.dat. The final input is a user configuration file that will be described further below.

Mapping between GEOS-Chem and CMAQ species requires human interpretation. Each model has its own definition of gas-phase and aerosol-phase speciation. Even common elements are named inconsistently (e.g., formaldehyde = FORM = HCHO = CH_2O). The default compound-mapping file shown as a csv file with a bold outline in Fig. 2 is described in detail below to facilitate user creation of new mapping files. For the most common configurations of GEOS-Chem and CMAQ, species mapping are already provided for several chemical mechanisms (e.g., Carbon Bond '05, SAPRC07T – provided in supplemental Tables A1 and A2). For these mechanisms, the species mapping has already been done and no manual interpretation is necessary. Ideally, any new mapping configuration files will be submitted back to the software package for subsequent distribution to other users. The mapping file contains one or more lines for each output boundary species. The individual lines represent algebraic transformations excluding unit conversion, which is mostly automatic. The numbered lines below are example lines from the species-mapping file with the regional model (e.g., CMAQ) species listed first followed by the global model (GEOS-Chem) formula.

1. O_3, O_x–NO_x
2. ALD2, 1./2 * ALD2
3. PAR, 4. * ALK4
4. ASO4K, 0.0776 * SALC
5. ASO4K, 0.02655 * DST4
6. ASO4K, 0.02655 * DST3
7. ASO4K, 0.02655 * DST2
8. ASO4K, SO4s

Mapping assumes that the formula is based on GEOS-Chem tracers. If the name indicated is not found in the tracer file, the species (cspec) file will be searched. Line 1 is currently configured for the GEOS-Chem tracer file. The GEOS-Chem version 8 tracer file does not include ozone explicitly, but rather O_x or odd oxygen. The "cspec" file includes ozone explicitly as "O_3", so if line 1 is replaced with "O_3, O_3" and the mapping tool would first try to find O_3 in the tracer file, not find it, and then search and find "O_3" in the "cspec" file.

Caution is advised when using values contained in the "cspec" file. For example, in the stratosphere, the "cspec" file does not contain meaningful values. These values are generally not updated or accessed by the GEOS-Chem simulation, and should not be used for LBC if information is available in the tracer file.

Line 2 illustrates a difference between the quantities stored in CMAQ LBC files and GEOS-Chem tracer files. ALD2, or acetaldehyde, is stored as parts per billion of carbon (ppbC) in GEOS-Chem and ppb in CMAQ. Since acetaldehyde has two carbons, the GEOS-Chem value must be halved for use by CMAQ.

Aerosol species in GEOS-Chem, such as wind-blown mineral dust and sea-salt, are speciated into individual aerosol constituents (Appel et al., 2013), and lines 4–7 demonstrate how GEOS-Chem aerosols such as SALC and DST2 are mapped based on CMAQ emission profiles for assignment to coarse mode sulfate.

Lines 4–8 above demonstrate that additional lines are additive. Because the lines are additive, these lines could have been re-written as a single line, "ASO4K, 0.0776 * SALC + 0.02655 * (DST2 + DST3 + DST4) + SO4s"

The mapping expressions can include all standard python operators (+, –, *, /, **, %, etc). Thus, any combination of GEOS-Chem simulated species may be mapped to modeled species using basic algebra. With more complex mathematical representations, a user could develop algorithms for mapping tracers to models with, for example, modal size distributions. In addition, empirical regression relationships could be used to develop boundary concentrations for species that are not simulated by GEOS-Chem. In a beta version of the code, arbitrary math functions (e.g., $\sin$) are available, but require advanced knowledge of Python and NumPy (http://www.numpy.org/). The flexible syntax allows for creative applications to other studies.

There are 5 types of factors that are routinely be applied:

1. Speciation of lumped GEOS-Chem things (like seasalt, dust, PRPE, etc.) to individual CMAQ species when the CMAQ representation is more detailed/speciated.
2. Conversion of real species to CB05/SAPRC mechanism species (like multiplying ACETONE by 3 for PAR).
3. Conversion of tracers in ppbC to ppb (like dividing benzene by 6).
4. Conversion of tracers to functional groups (e.g., ALK4 = 4 * PAR).
5. Conversion to regional model units.

Type 1 and 2 require algebraic expressions in the mapping file. Type 3 does not require expressions because the python pre-processor will automatically convert ppbC to ppb. Type 4 is a special case of type 3 where the regional model's conversion to ppb must be overridden in the file. Type 5 are treated automatically, converting ppb to $\mu g\,m^3$ for aerosols and ppb to ppm for gas-phase species.

3.2 Fortran spatial mapping

The Fortran-based spatial mapping program uses 3 required inputs and 2 optional inputs. The software first requires the output from the species mapping Python pre-processor described above. The species mapping is simply applied in concert with the spatial mapping.

The software also requires a GEOS-Chem tracer output file and, optionally, a chemical species ("cspec") output file. The GEOS-Chem files have sufficient meta-data to identify the files spatial location and extent based on the well-documented GEOS-Chem domains (Yantosca et al., 2012). The vertical coordinate is specified in the GEOS_DOMAIN.INC file, which re-writes the GEOS-Chem hybrid-eta coordinates as a sigma-P coordinate.

Finally, the software requires a meteorological input file, METBDY3D produced by a CMAQ utility (Otte and Pleim, 2010), which contains sufficient information to describe the centroid locations of each boundary cell, the vertical location on a sigma-P coordinate, and air density. The Fortran program selects a GEOS-Chem column/row using the "nearest neighbor" algorithm based on the regional model and GEOS-Chem centroids. Figure 1 shows the intersection of an example boundary and the GEOS-Chem outputs. The GEOS-Chem concentrations are then interpolated from the GEOS-Chem hybrid-eta levels to the regional model coordinate. This is done by first calculating each layer-center pressure for GEOS-Chem and the regional model, and then linearly interpolating. The archive only has GEOS-Chem results up to either the 34th layer (100 hPa) or the 38th layer (40 hPa). In the case where the regional model coordinate is outside the range of GEOS-Chem, the concentrations are extrapolated by default. This extrapolation can be disabled in the code.

High ozone in the upper troposphere is known to have undue influence on models with coarse vertical resolution. Simulations using coarse vertical resolution may need to reduce the influence of aloft ozone LBC. For example, previous work has shown that coarse vertical resolution can cause bias due to high ozone near the tropopause (Lam and Fu, 2009). We include tools for excluding stratospheric air from LBC, but do not recommend its use unless specifically desired.

Exclusion of stratospheric air has been suggested on the basis that AQM do not explicitly treat the stratosphere (Lam and Fu, 2009). Since that publication, there has been more work identifying the importance of stratospheric air in air quality (e.g., Lefohn et al., 2011). Air quality models have increased their vertical extent and now often include stratospheric influence, if not stratospheric air (e.g., Carlton et al., 2010). To account for the stratosphere, efforts have been made to scale the upper layer concentrations based on stratospheric indicators (Lin et al., 2008). As such, LBCs that specifically exclude stratospheric air are not consistent with the need to include stratospheric influence in air quality models. Further, reports show that vertically coarse models, like that used in Lam and Fu (2009), transported too much aloft air to the surface. This suggests that, while stratospheric air is an important contributor to variability, previous models would have optimal solutions that minimized aloft LBC values. The use of indirect evaluation, like interior domain surface concentrations, is inherently subject to canceling errors (e.g., Oreskes et al., 1994).

4 GLBC evaluation

This section describes the evaluation of GLBC using satellite retrievals. While ozonesondes are often considered the gold standard for evaluating satellite products (Nassar et al., 2008; Worden et al., 2007), they are not available at the boundary locations. In this analysis, we evaluate the LBC ozone values using two satellite products for ozone and one for carbon monoxide. Aerosol species are provided in the database to provide consistent boundary conditions, but have not been evaluated here. To evaluate the model, we pair satellite retrievals with GEOS-Chem grid cells from five years, 2006 to 2010 for two months. January results are selected to represent winter and August results are selected to represent the traditional ozone season. Details of the satellite products and model processing for comparison with retrievals are discussed below, followed by satellite and model processing details.

Ozone retrievals are taken from the Tropospheric Emission Spectrometer (TES) and the Ozone Monitoring Instrument (OMI). The TES instrument uses infrared Fourier transform spectroscopy to retrieve ozone vertical profiles (Bowman et al., 2011) from the Aura satellite and are limited to nadir scanning in this paper. We are using version 4 (V004) that has improved performance compared to V001 evaluated by Worden et al. (2007), but has a 5–15 % high-bias consistent with Nassar et al. (2008). Although the evaluation below will be performed in an absolute sense, the interpretation of these results must account for TES's unresolved high bias. Although the evaluation is nominally for 2006 to 2010, TES profiles were not available for January of 2010. Data for all other months was downloaded from http://avdc.gsfc.nasa.gov/download_2.php?site=634280718&id=60.

The OMI instrument measures ozone from the Aura satellite. We use the Level 2 OMI ozone profile (OMO3PR) version 3. Files were downloaded from NASA's Mirador website, and filtered using the recommended bitwise and calculation of the ProcessingQualityFlags (i.e., ProcessingQualityFlags & 43679). The OMI ozone data was available for the all years and months.

The MOPITT instrument is aboard the Terra satellite and measures carbon monoxide. MOPITT retrieves carbon monoxide by differential absorption of light in infrared absorption bands. The carbon monoxide measurement is translated into a vertical profile using a retrieval algorithm described by Deeter et al. (2003). We are using the MOPITT carbon monoxide Level 2 product version 6, which uses only thermal infrared radiances (MOP02T). Data files were downloaded from NASA's Reverb website http://reverb.echo.nasa.gov/reverb with no additional filtering. The MOPITT carbon monoxide data was not available for August of 2009, so that month will not be evaluated.

The GEOS-Chem grid cells are filtered for just those that would be used in creating CONUS boundary conditions (see Fig. 1). GEOS-Chem grid cells are then paired with satellite pixel centroids when the pixel is contained within the grid cell. After pairs have been identified, the satellite retrieval algorithms are applied to GEOS-Chem using Eq. (1) for TES and OMI. Equation (1) follows Bowman et al. (2011, Eqs. 5–8) methodology and has the effect of smoothing model results vertically. Smoothing is required because the satellites estimates at each pressure level sensitive to concentrations at other pressure levels.

$$\widehat{\boldsymbol{y}}_t^{i,m} = \boldsymbol{y}_{t,c}^i + \mathbf{A}_t^i \left(\boldsymbol{y}_t^{i,m} - \boldsymbol{y}_{t,c}^i \right) + \varepsilon_t^i \tag{1}$$

where all $\boldsymbol{y}$ values are the natural log of the mixing ratio for ozone or carbon monoxide, $\boldsymbol{y}_t^{i,m}$ is the original model prediction, $\boldsymbol{y}_{t,c}^i$ is the a prior estimate, $\mathbf{A}_t^i$ is the averaging kernel, and ε_t^i is an unknown error component. $\widehat{\boldsymbol{y}}_t^{i,m}$ is the model retrieval that can be directly compared to the satellite retrieval. In the evaluation shown here, the results have all been converted to mixing ratios (ppbV). Although the absolute value of $\widehat{\boldsymbol{y}}_t^{i,m}$ depends on the a prior ($\boldsymbol{y}_{t,c}^i$), a comparison between $\widehat{\boldsymbol{y}}_t^{i,m}$ and the retrieval ($\widehat{\boldsymbol{y}}_t^i$) does not (Bowman et al., 2011). This independence is mathematically shown in the TES User Guide.

The evaluation has been performed by grouping grid cells by boundary face (West, North, East, South) on the CONUS domain. Based on pixel centroid locations during the 5 years, there are a total of 274316 pairs with MOPITT carbon monoxide (Jan: 165246, Aug: 109070), 128186 pairs with OMI ozone (Jan: 64216, Aug: 63970), and 1753 pairs with TES ozone (Jan: 841, Aug: 912). The larger number of pixel pairs for MOPITT and OMI is expected because there are more pixels in their arrays.

For each satellite, biases were initially reviewed for 40 categories (5 yr $\times$ 2 months $\times$ 4 perimeter cardinal edges). The difference between years was nominal and is not highlighted here, but is included in the Appendix (Figs. A1–A6). Instead this paper will focus on results aggregated by month and boundary face (West, North, East, South). As previously noted, the GEOS-Chem data base only has 34 or 38 layers. To minimize the influence of extrapolation on this analysis, above layer 34 extrapolated results are replaced with satellite estimates. Because extrapolation is optional, this seemed most appropriate.

Figure 3 and 4 shows ozone and carbon monoxide (ppb) for each boundary face for January (Fig. 3) and August (Fig. 4). Each panel shows raw GEOS-Chem results, GEOS-Chem retrievals (Eq. 1), and satellite retrievals. To aid in interpretation, GEOS-Chem biases have been highlighted using triangles on the y axis (red = high; blue = low) when the bias is greater than

the twice observation uncertainty. To prevent spurious differences, we require that a student's t-test reject the null hypothesis that the distribution of model retrievals is the same as the satellite ($p < 0.001$). The mean and range of profiles show good correspondence most of the time. The MOPITT carbon monoxide and OMI ozone are in good agreement with GEOS-Chem. For TES, however, the evaluation shows some discrepancies.

The TES bias also exhibits time and space dependence. Figures 3 and 4 show distinct performance regimes above and below 350 hPa. Below 350 hPa, there is a transient low-bias that is most pronounced in August. Above 350 hPa, there is a more persistent high-bias. The high bias above 350 hPa is higher in the West, South, and East faces compared to the North. The analysis thus far is based on the vertical profile of means and basic distribution statistics.

To further explore these aggregate biases, Fig. 5 and 6 show the distribution of individual retrieval biases for January (Fig. 5) and August (Fig. 6). The biases in Fig. 5 and 6 are shown as the ratio of retrieved mixing ratios (i.e., ppb). To reiterate, this type of comparison is not dependent upon the a prior – only the sensitivity of the instrument. Table 2 shows the percentage of pixels for each boundary face and for each month where the model and observed value are within $\pm 10\,\%$ and $\pm 20\,\%$. For most categories, 70-81% of MOPITT and OMI results are within $\pm$20% and 45-56% are within $\pm$10% of satellite retrievals. Only the OMI south faces has less than 70% (January 61% and August 66%) within $\pm 20\,\%$, and less than 45% within $\pm 10\,\%$. TES shows more variable performance. Except for the North face, 56% or less of TES comparisons are within $\pm$20%, and 32% or less are within $\pm$10%.

The TES high bias above 350 hPa is more pronounced in January than in August, and this significant bias correlates with an enhanced bias in the MOPITT carbon monoxide. For MOPITT, the biases are not outside of precision, but the correlation is intriguing. The biases in Figs. 3 and 4 and Tab. 2 correlate with latitude, with a stronger relationship aloft. For TES retrievals, the ratio model to satellite retrieval was regressed against latitude and longitude. The regression was performed for each layer for all January and August months. The January slope is strongest, as shown in Figure 6, and ranges from -1.05% at 316 hPa to -2.08% at 162 hPa. Although this explains only 15% of bias variability, the slope is statistically significant for latitude. For longitude, the slope is negligible and never significant.

5 Conclusions

We describe and evaluate a tool for using global simulations to produce LBC for regional air quality models. In general, the LBC performed well in evaluation for ozone and carbon monoxide. There was a bias seen when comparing to TES retrievals. A persistent high bias was found in the upper troposphere (above 350 hPa). This bias is counter balanced by good performance compared to OMI ozone evaluation and may be a limitation of our dataset.

The model bias compared to TES may be the result of limited outputs at high altitudes. Our archived GEOS-Chem only includes levels below approximately 100 hPa. In comparing to TES, the averaging kernel effectively includes data from above 100 hPa at several layers below 100 hPa. As such, the model bias compared to TES may be the result of assumption made above our archived model data.

The altitude and timing of the bias compared to TES suggests an over-estimation of long-distance transport. Our evaluation showed that the model performed better in August than in January. In January and at high altitudes, temperatures are low and ozone lifetimes are long. These conditions are ideal to highlight ozone from continental outflow. More research is needed to understand the source of this bias, which could be transport or emissions. The emissions are implicated by the correlation between biases of carbon monoxide and ozone. This suggests up wind emissions, possibly from Asia, are over-estimated. Asian emissions have grown rapidly and future projections shown continued growth (Ohara et al., 2007). This leads to high uncertainty in simulated Asian emissions that could contribute to the observed bias.

The presented tool provides a resource to better represent global transport through boundary conditions in regional air quality studies. Evaluations showed good mean performance, but the maximum bias was over a factor of five. This bias could either be the result of satellite uncertainties or model uncertainties. The role of uncertainty in boundary conditions can have strong impacts on regional model results. This will be particularly true for longer-lived compounds with direct impacts, like ozone. When specific episodes are critical to the model application, further application specific evaluation will be necessary. The database's overall evaluation demonstrates the fitness for producing LBC.

Both the tool and the database are freely available. The database can be downloaded from the University of Florida's FTP server at ftp://data.as.essie.ufl.edu/pub/geos2cmaq and the tool can be downloaded from http://github.com/barronh/geos2cmaq. At the tool website, an example dataset can be found with step-by-step instructions. The availability and usability of this tool serves the community need for lateral boundary conditions for regional modeling.

Appendix A

The Appendix contains species mapping for common gas-phase and aerosol mechanisms and more detailed evaluation of ozone and carbon monoxide lateral boundary conditions. Tables A1 and A2 provide mapping details for Carbon Bond '05 and SAPRC07. These tables are followed by detailed discussion of aerosol mapping for CMAQ's aerosol mechanism. Finally, the body of the paper discusses aggregated years 2006 to 2010. The Appendix provides information on individual years.

A1 Individual Year Evaluation

See Figs. A1–A6.

A2 Species mapping for gas-phase

See Tables A1 and A2.

A3 Species mapping for CMAQ aerosols

The CMAQ AERO6 aerosol module generally contains more detailed information regarding aerosol speciation and size than standard GEOS-Chem output. As a result, factors are applied to GEOS-Chem aerosols to appropriately convert them to CMAQ-ready boundary conditions. The conversions we recommend are shown in Table A3 and discussed below.

Both seasalt and dust in GEOS-Chem contain size information. Accumulation (SALA) and coarse (SALC) mode seasalt from GEOS-Chem are matched with the accumulation (J) and coarse (K) mode in CMAQ. Based on the particle size of the four GEOS-Chem dust size bins, the smallest dust (DST1) is mapped to the accumulation mode while all other bins (DST2-4) are mapped to the coarse mode. Speciation of seasalt into trace metals and other aerosol constituents is based on the same speciation profile that CMAQ uses for seasalt emissions diagnosed within the model. The speciation of wind-blown mineral dust also follows a speciation profile in CMAQ and is based on a composite of four desert dust profiles (Appel et al., 2013).

Sulfate, nitrate, and ammonium aerosol in GEOS-Chem (Park et al., 2004; Pye et al., 2009) do not explicitly contain size information, but are generally assumed to be representative of the accumulation mode. As a result 99 % of sulfate, nitrate, and ammonium are assigned to the accumulation (J) mode while 1 % is attributed to the Aitken (I) mode. Sulfate formed on seasalt (SO_4s) and nitrate formed on seasalt (NO_3s) (Alexander, 2005) are mapped to the CMAQ coarse mode. 99.9 % of primary carbonaceous aerosols from GEOS-Chem are attributed to the accumulation mode while 0.1 % are assigned to the Aitken mode consistent with CMAQ emissions processing (Binkowski and Roselle, 2003, paragraph 12). Both hydrophobic (BCPO) and hydrophilic (BCPI) forms of black carbon in GEOS-Chem are summed together and mapped to elemental carbon (EC). Similarly, hydrophobic and hydrophilic organic carbon is mapped to primary organic carbon. The non-carbon organic matter (NCOM) associated with primary organic aerosols is not calculated by GEOS-Chem, so a OM/OC ratio of 1.4 is assumed for boundary condition purposes (Park, 2003).

Although CMAQ and GEOS-Chem both treat secondary organic aerosol from the same set of parent hydrocarbons, the species lumping schemes differ. In CMAQ, lumping is based on precursor hydrocarbon identity as well as volatility while the GEOS-Chem SOA lumping scheme (Chung, 2002; Henze et al., 2008; Liao et al., 2007) generally does not separate based on volatility. The mapping of SOA as well as gas-phase semivolatiles is based on identifying the equivalent parent hydrocarbon in each model. Speciation to the different volatility species within CMAQ is based on the expected relative amounts of each species in outflow of the Eastern US as predicted by a typical CMAQ simulation.

The particle number and surface area for the boundary conditions are calculated in the Fortran code based on the mass mapped into each mode.

The following CMAQ aerosol species boundary conditions are not mapped since there is not an analogous GEOS-Chem model species: AOLGBJ, AOLGAJ, AALKJ, SV_ALK, ACORS. Aerosol water is also not mapped as it is readily computed within CMAQ and does not need to be transported.

Acknowledgements. Barron H. Henderson was supported in part by the Research Participation Program at the Environmental Protection Agency administered by the Oak Ridge Institute for Science and Education, and in part by startup funds from the University of Florida.

The United States Environmental Protection Agency (EPA) through its Office of Research and Development collaborated in the research described here. This paper has been subjected to the Agency's administrative review and approved for publication.

References

Alexander, B.: Sulfate formation in sea-salt aerosols: constraints from oxygen isotopes, J. Geophys. Res., 110, D10307, doi:10.1029/2004JD005659, 2005.

Appel, K. W. and Gilliland, A. B.: Effects of vertical-layer structure and boundary conditions on CMAQ – v4.5 and v4.6 models, Chapel Hill, NC, available at: http://www.cmascenter.org/conference/2006/abstracts/appel_session4.pdf, last access: 23 August 2013, 2006.

Appel, K. W., Pouliot, G. A., Simon, H., Sarwar, G., Pye, H. O. T., Napelenok, S. L., Akhtar, F., and Roselle, S. J.: Evaluation of dust and trace metal estimates from the Community Multiscale Air Quality (CMAQ) model version 5.0, Geosci. Model Dev. Discuss., 6, 1859–1899, doi:10.5194/gmdd-6-1859-2013, 2013.

Barna, M. G. and Knipping, E. M.: Insights from the BRAVO study on nesting global models to specify boundary conditions in regional air quality modeling simulations, Atmos. Environ., 40, Supplement 2, 574–582, doi:10.1016/j.atmosenv.2006.01.065, 2006.

Berdowski, J., Guicherit, R., Heij, B., and Dutch National Research Programme on Global Air Pollution and Climate Change: The Climate System, A. A. Balkema Publishers, Lisse, Exton, PA, 2001.

Binkowski, F. S. and Roselle, S. J.: Models-3 Community Multiscale Air Quality (CMAQ) model aerosol component 1. Model description, Journal of Geophysical Research: Atmospheres, 108(D6), doi:10.1029/2001JD001409, 2003.

Bourqui, M. S., Yamamoto, A., Tarasick, D., Moran, M. D., Beaudoin, L.-P., Beres, I., Davies, J., Elford, A., Hocking, W., Osman, M., and Wilkinson, R.: A new global real-time Lagrangian diagnostic system for stratosphere-troposphere exchange: evaluation during a balloon sonde campaign in eastern Canada, Atmos. Chem. Phys., 12, 2661–2679, doi:10.5194/acp-12-2661-2012, 2012.

Bowman, K., Eldering, A., Fisher, B., Jacob, D., Jourdain, L., Kulawik, S. S., Luo, M., Monarrez, R., Osterman, G., Paradise, S., Payne, V., Poosti, S., Rischards, N., Rider, D., Shepard, D., Shephard, M., Vilnrotter, F., Worden, H., Worden, J., Yun, H., and Zhang, L.: Earth Observing System (EOS) Tropospheric Emission Spectrometer (TES) Level 2 (L2) Data User's Guide (Up to and including Version 5 data), 5.0 edn., edited by: Herman, R. and Kulawik, S., available at: http://tes.jpl.nasa.gov/uploadedfiles/TES_L2_Data_Users_Guide-3.pdf, last access: 23 August 2013, 2011.

Carlton, A. G., Bhave, P. V., Napelenok, S. L., Edney, E. O., Sarwar, G., Pinder, R. W., Pouliot, G. A., and Houyoux, M.: Model representation of secondary organic aerosol in CMAQv4.7, Environ. Sci. Technol., 44, 8553–8560, doi:10.1021/es100636q, 2010.

Chung, S. H.: Global distribution and climate forcing of carbonaceous aerosols, J. Geophys. Res., 107, 4407, doi:10.1029/2001JD001397, 2002.

Cooper, O. R., Parrish, D. D., Stohl, A., Trainer, M., Nedelec, P., Thouret, V., Cammas, J. P., Oltmans, S. J., Johnson, B. J., Tarasick, D., Leblanc, T., McDermid, I. S., Jaffe, D., Gao, R., Stith, J., Ryerson, T., Aikin, K., Campos, T., Weinheimer, A., and Avery, M. A.: Increasing springtime ozone mixing ratios in the free troposphere over western North America, Nature, 463, 344–348, doi:10.1038/nature08708, 2010.

Cui, J., Sprenger, M., Staehelin, J., Siegrist, A., Kunz, M., Henne, S., and Steinbacher, M.: Impact of stratospheric intrusions and intercontinental transport on ozone at Jungfraujoch in 2005: comparison and validation of two Lagrangian approaches, Atmos. Chem. Phys., 9, 3371–3383, doi:10.5194/acp-9-3371-2009, 2009.

Deeter, M. N., Emmons, L. K., Francis, G. L., Edwards, D. P., Gille, J. C., Warner, J. X., Khattatov, B., Ziskin, D., Lamarque, J.-F., Ho, S.-P., Yudin, V., Attié, J.-L., Packman, D., Chen, J., Mao, D. and Drummond, J. R.: Operational carbon monoxide retrieval algorithm and selected results for the MOPITT instrument, Journal of Geophysical Research: Atmospheres, 108(D14), doi:10.1029/2002JD003186, 2003.

Dentener, F., Keating, T. J., and Akimoto, H.: Hemispheric transport of air pollution. Part A: Ozone and Particulate Matter, Economic Commission For Europe United Nations, Geneva, 2010.

Environment Canada: National Pollutant Release Inventory, Environ. Can., available at: http://www.ec.gc.ca/inrp-npri/, last access: 23 August 2013, 2013.

Fiore, A. M., Jacob, D. J., Bey, I., Yantosca, R. M., Field, B. D., Fusco, A. C., and Wilkinson, J. G.: Background ozone over the United States in summer: origin, trend, and contribution to pollution episodes, J. Geophys. Res.-Atmos., 107, ACH 11-1–ACH 11-25, doi:10.1029/2001JD000982, 2002.

Fiore, A. M., Dentener, F. J., Wild, O., Cuvelier, C., Schultz, M. G., Hess, P., Textor, C., Schulz, M., Doherty, R. M., Horowitz, L. W., MacKenzie, I. A., Sanderson, M. G., Shindell, D. T., Stevenson, D. S., Szopa, S., Dingenen, R. V., Zeng, G., Atherton, C., Bergmann, D., Bey, I., Carmichael, G., Collins, W. J., Duncan, B. N., Faluvegi, G., Folberth, G., Gauss, M., Gong, S., Hauglustaine, D., Holloway, T., Isaksen, I. S. A., Jacob, D. J., Jonson, J. E., Kaminski, J. W., Keating, T. J., Lupu, A., Marmer, E., Montanaro, V., Park, R. J., Pitari, G., Pringle, K. J., Pyle, J. A., Schroeder, S., Vivanco, M. G., Wind, P., Wojcik, G., Wu, S., and Zuber, A.: Multimodel estimates of intercontinental source–receptor relationships for ozone pollution, J. Geophys. Res., 114, D04301, doi:10.1029/2008JD010816, 2009.

Foley, K. M., Roselle, S. J., Appel, K. W., Bhave, P. V., Pleim, J. E., Otte, T. L., Mathur, R., Sarwar, G., Young, J. O., Gilliam, R. C., Nolte, C. G., Kelly, J. T., Gilliland, A. B., and Bash, J. O.: Incremental testing of the Community Multiscale Air Quality (CMAQ) modeling system version 4.7, Geosci. Model Dev., 3, 205–226, doi:10.5194/gmd-3-205-2010, 2010.

Fu, J. S., Streets, D. G., Jang, C. J., Hao, J., He, K., Wang, L., and Zhang, Q.: Modeling Regional/Urban Ozone and Particulate Matter in Beijing, China, J. Air Waste Manage., 59, 37–44, doi:10.3155/1047-3289.59.1.37, 2009.

Gégo, E., Gilliland, A., Godowitch, J., Rao, S. T., Porter, P. S., and Hogrefe, C.: Modeling analyses of the effects of changes in nitrogen oxides emissions from the electric power sector on ozone levels in the Eastern US, J. Air Waste Manage., 58, 580–588, doi:10.3155/1047-3289.58.4.580, 2008.

Godowitch, J. M., Gilliland, A. B., Draxler, R. R., and Rao, S. T.: Modeling assessment of point source NO_x emission reductions on ozone air quality in the eastern United States, Atmos. Environ., 42, 87–100, doi:10.1016/j.atmosenv.2007.09.032, 2008.

Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., and Wang, X.: The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions, Geosci. Model Dev., 5, 1471–1492, doi:10.5194/gmd-5-1471-2012, 2012.

Henze, D. K., Seinfeld, J. H., Ng, N. L., Kroll, J. H., Fu, T.-M., Jacob, D. J., and Heald, C. L.: Global modeling of secondary organic aerosol formation from aromatic hydrocarbons: high- vs. low-yield pathways, Atmos. Chem. Phys., 8, 2405–2420, doi:10.5194/acp-8-2405-2008, 2008.

Hogrefe, C., Civerolo, K. L., Hao, W., Ku, J.-Y., Zalewsky, E. E., and Sistla, G.: Rethinking the assessment of photochemical modeling systems in air quality planning applications, J. Air Waste Manage., 58, 1086–1099, doi:10.3155/1047-3289.58.8.1086, 2008.

Jacobson, Z. M. and Turco, R. P.: SMVGEAR: a sparse-matrix, vectorized gear code for atmospheric models, Atmos. Environ., 28, 273–284, doi:10.1016/1352-2310(94)90102-3, 1994.

Jiménez, P., Parra, R., and Baldasano, J. M.: Influence of initial and boundary conditions for ozone modeling in very complex terrains: a case study in the northeastern Iberian Peninsula, Environ. Modell. Softw., 22, 1294–1306, doi:10.1016/j.envsoft.2006.08.004, 2007.

Krueger, A. J. and Minzner, R. A.: A mid-latitude ozone model for the 1976 US Standard Atmosphere, J. Geophys. Res., 81, 4477, doi:10.1029/JC081i024p04477, 1976.

Kuhns, H., Green, M., and Etyemezian, V.: Big Bend Regional Aerosol and Visibility Observational (BRAVO) Study Emissions Inventory, Desert Res. Inst., available at: http://acmg.seas.harvard.edu/geos/word_pdf_docs/BRAVOEI_Report_d2.pdf, last access: 23 August 2013, 2003.

Lacis, A. A., Wuebbles, D. J., and Logan, J. A.: Radiative forcing of climate by changes in the vertical distribution of ozone, J. Geophys. Res., 95, 9971–9981, doi:10.1029/JD095iD07p09971, 1990.

Lam, Y. F. and Fu, J. S.: A novel downscaling technique for the linkage of global and regional air quality modeling, Atmos. Chem. Phys., 9, 9169–9185, doi:10.5194/acp-9-9169-2009, 2009.

Lefohn, A. S., Wernli, H., Shadwick, D., Limbach, S., Oltmans, S. J., and Shapiro, M.: The importance of stratospheric–tropospheric transport in affecting surface ozone concentrations in the western and northern tier of the United States, Atmos. Environ., 45, 4845–4857, doi:10.1016/j.atmosenv.2011.06.014, 2011.

Liao, H., Henze, D. K., Seinfeld, J. H., Wu, S., and Mickley, L. J.: Biogenic secondary organic aerosol over the United States: comparison of climatological simulations with observations, J. Geophys. Res., 112, D06201, doi:10.1029/2006JD007813, 2007.

Lin, C., Jacob, D., Munger, J., and Fiore, A.: Increasing background ozone in surface air over the United States, Geophys. Res. Lett., 27, 3465–3468, doi:10.1029/2000GL011762, 2000.

Lin, H., Mathur, R., McKeen, S. A., and McQueen, J.: Application of Potential Vorticity in a comprehensive air quality forecast model for Ozone, available at: https://ams.confex.com/ams/88Annual/techprogram/paper_132967.htm, last access: 23 August 2013, 2008.

Lin, M., Fiore, A. M., Horowitz, L. W., Cooper, O. R., Naik, V., Holloway, J., Johnson, B. J., Middlebrook, A. M., Oltmans, S. J., Pollack, I. B., Ryerson, T. B., Warner, J. X., Wiedinmyer, C., Wilson, J., and Wyman, B.: Transport of Asian ozone pollution into surface air over the Western United States in spring, J. Geophys. Res., 117, D00V07, doi:10.1029/2011JD016961, 2012.

Logan, J. A., Megretskaia, I. A., Miller, A. J., Tiao, G. C., Choi, D., Zhang, L., Stolarski, R. S., Labow, G. J., Hollandsworth, S. M., Bodeker, G. E., Claude, H., Muer, D. D., Kerr, J. B., Tarasick, D. W., Oltmans, S. J., Johnson, B., Schmidlin, F., Staehelin, J., Viatte, P., and Uchino, O.: Trends in the vertical distribution of ozone: a comparison of two analyses of ozonesonde data, J. Geophys. Res., 104, 26373–26399, doi:10.1029/1999JD900300, 1999.

Mao, J., Paulot, F., Jacob, D., Cohen, R., Crounse, J., Wennberg, P., Keller, C., Hudman, R., Barkley, M., and Horowitz, L.: Ozone and Organic Nitrates over the Eastern US: Sensitivity to Isoprene Chemistry, 2013.

Molod, A., Takacs, L., Suarez, M., Bacmeister, J., Song, I.-S., and Eichmann, A.: The GEOS-5 Atmospheric General Circulation Model: Mean Climate and Development from MERRA to Fortuna, 2012.

Napelenok, S. L., Cohan, D. S., Odman, M. T., and Tonse, S.: Extension and evaluation of sensitivity analysis capabilities in a photochemical model, Environ. Modell. Softw., 23, 994–999, doi:10.1016/j.envsoft.2007.11.004, 2008.

Napelenok, S. L., Foley, K. M., Kang, D., Mathur, R., Pierce, T., and Rao, S. T.: Dynamic evaluation of regional air quality model's response to emission reductions in the presence of uncertain emission inventories, Atmos. Environ., 45, 4091–4098, doi:10.1016/j.atmosenv.2011.03.030, 2011.

Nassar, R., Logan, J. A., Worden, H. M., Megretskaia, I. A., Bowman, K. W., Osterman, G. B., Thompson, A. M., Tarasick, D. W., Austin, S., Claude, H., Dubey, M. K., Hocking, W. K., Johnson, B. J., Joseph, E., Merrill, J., Morris, G. A., Newchurch, M., Oltmans, S. J., Posny, F., Schmidlin, F. J., Vömel, H., Whiteman, D. N., and Witte, J. C.: Validation of Tropospheric Emission Spectrometer (TES) nadir ozone profiles using ozonesonde measurements, J. Geophys. Res., 113, 17, doi:10.1029/2007JD008819, 2008.

Nghiem, L. H. and Oanh, N. T. K.: Evaluation of the Mesoscale Meteorological Model (MM5)-Community Multi-Scale Air Quality Model (CMAQ) performance in hindcast and forecast of ground-level ozone, J. Air Waste Manage., 58, 1341–1350, doi:10.3155/1047-3289.58.10.1341, 2008.

Ohara, T., Akimoto, H., Kurokawa, J., Horii, N., Yamaji, K., Yan, X. and Hayasaka, T.: An Asian emission inventory of anthropogenic emission sources for the period 1980–2020, Atmos. Chem. Phys., 7(16), 4419–4444, doi:10.5194/acp-7-4419-2007, 2007.

Oltmans, S. J., Lefohn, A. S., Harris, J. M., Galbally, I., Scheel, H. E., Bodeker, G., Brunke, E., Claude, H., Tarasick, D., Johnson, B. J., Simmonds, P., Shadwick, D., Anlauf, K., Hayden, K., Schmidlin, F., Fujimoto, T., Akagi, K., Meyer, C., Nichol, S., Davies, J., Redondas, A., and Cuevas, E.: Long-term changes in tropospheric ozone, Atmos. Environ., 40, 3156–3173, doi:10.1016/j.atmosenv.2006.01.029, 2006.

Oltmans, S. J., Lefohn, A. S., Harris, J. M., Tarasick, D. W., Thompson, A. M., Wernli, H., Johnson, B. J., Novelli, P. C., Montzka, S. A., Ray, J. D., Patrick, L. C., Sweeney, C., Jefferson, A., Dann, T., Davies, J., Shapiro, M., and Holben, B. N.: Enhanced ozone over western North America from biomass burning in Eurasia during April 2008 as seen in surface and profile observations, Atmos. Environ., 44, 4497–4509, doi:10.1016/j.atmosenv.2010.07.004, 2010.

Oreskes, N., Shrader-Frechette, K., and Belitz, K.: Verification, validation, and confirmation of numerical models in the Earth sciences, Science, 263, 641–646, doi:10.1126/science.263.5147.641, 1994.

Ott, L. E., Pickering, K. E., Stenchikov, G. L., Allen, D. J., Decaria, A. J., Ridley, B., Lin, R.-F., Lang, S., and Tao, W.-K.: Production of lightning NO_x and its vertical distribution calculated from three-dimensional cloud-scale chemical transport model simulations, J. Geophys. Res., 115, D04301, doi:10.1029/2009JD011880, 2010.

Otte, T. L. and Pleim, J. E.: The Meteorology-Chemistry Interface Processor (MCIP) for the CMAQ modeling system: updates through MCIPv3.4.1, Geosci. Model Dev., 3, 243–256, doi:10.5194/gmd-3-243-2010, 2010.

Park, R. J.: Sources of carbonaceous aerosols over the United States and implications for natural visibility, J. Geophys. Res., 108, 4355, doi:10.1029/2002JD003190, 2003.

Park, R. J., Jacob, D. J., Field, B. D., Yantosca, R. M., and Chin, M.: Natural and transboundary pollution influences on sulfate-nitrate-ammonium aerosols in the United States: Implications for policy, J. Geophys. Res., 109, D15204, doi:10.1029/2003JD004473, 2004.

Pickering, K. E., Wang, Y., Tao, W.-K., Price, C., and Müller, J.-F.: Vertical distributions of lightning NO_x for use in regional and global chemical transport models, J. Geophys. Res., 103, 31203–31216, doi:10.1029/98JD02651, 1998.

Price, C. and Rind, D.: A simple lightning parameterization for calculating global lightning distributions, J. Geophys. Res.-Atmos., 97, 9919–9933, doi:10.1029/92JD00719, 1992.

Pye, H. O. T. and Napelenok, S. L.: Chemical mapping of GEOS-Chem to CMAQv5.0, available at: http://wiki.seas.harvard.edu/geos-chem/index.php/GEOS-Chem_to_CMAQv5.0, last access 24 July 2013.

Pye, H. O. T., Liao, H., Wu, S., Mickley, L. J., Jacob, D. J., Henze, D. K., and Seinfeld, J. H.: Effect of changes in climate and emissions on future sulfate-nitrate-ammonium aerosol levels in the United States, J. Geophys. Res.-Atmos., 114, D01205, doi:10.1029/2008JD010701, 2009.

Reidmiller, D. R., Fiore, A. M., Jaffe, D. A., Bergmann, D., Cuvelier, C., Dentener, F. J., Duncan, B. N., Folberth, G., Gauss, M., Gong, S., Hess, P., Jonson, J. E., Keating, T., Lupu, A., Marmer, E., Park, R., Schultz, M. G., Shindell, D. T., Szopa, S., Vivanco, M. G., Wild, O., and Zuber, A.: The influence of foreign vs. North American emissions on surface ozone in the US, Atmos. Chem. Phys., 9, 5027–5042, doi:10.5194/acp-9-5027-2009, 2009.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., and Woollen, J.: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, J. Climate, 24, 3624–3648, doi:10.1175/JCLI-D-11-00015.1, 2011.

Schichtel, B. A., Gebhart, K. A., Malm, W. C., Barna, M. G., Pitchford, M. L., Knipping, E. M., and Tombach, I. H.: Reconciliation and interpretation of Big Bend National Park particulate sulfur source apportionment: results from the Big Bend regional aerosol and visibility observational study – Part I, J. Air Waste Manage., 55, 1709–1725, doi:10.1080/10473289.2005.10464769, 2005.

Smyth, S. C., Jiang, W., Roth, H., Moran, M. D., Makar, P. A., Yang, F., Bouchet, V. S., and Landry, H.: A comparative performance evaluation of the AURAMS and CMAQ air-quality modelling systems, Atmos. Environ., 43, 1059–1070, doi:10.1016/j.atmosenv.2008.11.027, 2009.

Song, C.-K., Byun, D. W., Pierce, R. B., Alsaadi, J. A., Schaack, T. K., and Vukovich, F.: Downscale linkage of global model output for regional chemical transport modeling: method and general performance, J. Geophys. Res., 113, D08308, doi:10.1029/2007JD008951, 2008.

Streets, D. G., Bond, T. C., Carmichael, G. R., Fernandes, S. D., Fu, Q., He, D., Kilmont, Z., Nelson, S. M., Tsai, N. Y., Wang, M. Q., Woo, J.-H., and Yarber, K. F.: An inventory of gaseous and primary aerosol emissions in Asia in the year 2000, J. Geophys. Res., 108, 8809, doi:10.1029/2002JD003093, 2003.

Streets, D. G., Zhang, Q., Wang, L., He, K., Hao, J., Wu, Y., Tang, Y., and Carmichael, G. R.: Revisiting China's CO emissions after the Transport and Chemical Evolution over the Pacific (TRACE-P) mission: synthesis of inventories, atmospheric modeling, and observations, J. Geophys. Res., 111, D14306, doi:10.1029/2006JD007118, 2006.

US EPA: National Emissions Inventory (NEI) Air Pollutant Emissions Trends Data, available at: http://www.epa.gov/ttnchie1/trends/, last access 24 July 2013.

Valari, M., Menut, L., and Chatignoux, E.: Using a chemistry transport model to account for the spatial variability of exposure concentrations in epidemiologic air pollution studies, J. Air Waste Manage., 61, 164–179, doi:10.3155/1047-3289.61.2.164, 2011.

Vestreng, V. and Klein, H.: Emission data reported to UNECE/EMEP: quality assurance and trend analysis & presentation of WebDab, MSC-W Status Rep 2002, available at: http://emep.int/publ/reports/2002/mscw_note_1_2002.pdf, last access: 23 August 2013, 2002.

Wang, Y., Logan, J. A., and Jacob, D. J.: Global simulation of tropospheric O_3–NO_x – hydrocarbon chemistry: 2. Model evaluation and global ozone budget, J. Geophys. Res., 103, 10727, doi:10.1029/98JD00157, 1998.

Warneck, P. and Williams, J.: The Atmospheric Chemist's Companion: Numerical Data for Use in the Atmospheric Sciences, 2012th Edn., Springer, 2012.

van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Kasibhatla, P. S., and Arellano Jr., A. F.: Interannual variability in global biomass burning emissions from 1997 to 2004, Atmos. Chem. Phys., 6, 3423–3441, doi:10.5194/acp-6-3423-2006, 2006.

Table 11. GEOS-Chem Annual Simulations for CMAQ boundaries (recommended in bold).

GEOS-Chem version	Chemistry version	Meteorology	Shipping emissions[a,b]	Simulation years[c]
v9-01-01	v8-02-04	GEOS-5	EDGAR	2004–2006
v9-01-02	v8-02-04	MERRA	EDGAR	2001–2008
v8-03-02	v8-02-04	GEOS-5	EDGAR	2004–2007
v8-03-02	**v8-02-01**	**GEOS-5**	**ICOADS**	**2004–2012**
v9-01-02	**v8-02-01**	**MERRA**	**ICOADS**	**2001–2010**

[a] ICOADS is the default (recommended) ship emission inventory (http://wiki.seas.harvard.edu/geos-chem/index.php/EDGAR_anthropogenic_emissions#Ship_emissions). [b] In GEOS-Chem simulations below v9-01-01, U.S. biofuel emissions were erroneously excluded when using the NEI2005 inventory. In versions v9-01-01 and later, NEI1999 biofuel emissions are used. [c] Years shown are inclusive. First year is spinup.

Table 12. Percentage of retrieval values below 50 hPa that are within 10 and 20 percent of TES, OMI, and MOPITT for each boundary face (West, North, East, and South) from 2006 to 2010.

	TES		OMI		MOPITT	
Boundary	±10 %	±20 %	±10 %	±20 %	±10 %	±20 %
January						
West	27 %	48 %	50 %	77 %	49 %	75 %
North	49 %	77 %	59 %	86 %	45 %	70 %
East	31 %	52 %	48 %	75 %	48 %	76 %
South	19 %	34 %	34 %	61 %	45 %	73 %
August						
West	32 %	56 %	49 %	81 %	56 %	81 %
North	47 %	77 %	57 %	90 %	51 %	75 %
East	30 %	56 %	46 %	76 %	48 %	75 %
South	27 %	50 %	36 %	66 %	48 %	75 %

Worden, H. M., Logan, J. A., Worden, J. R., Beer, R., Bowman, K., Clough, S. A., Eldering, A., Fisher, B. M., Gunson, M. R., Herman, R. L., Kulawik, S. S., Lampel, M. C., Luo, M., Megretskaia, I. A., Osterman, G. B., and Shephard, M. W.: Comparisons of Tropospheric Emission Spectrometer (TES) ozone profiles to ozonesondes: methods and initial results, J. Geophys. Res., 112, D03309, doi:10.1029/2006JD007258, 2007.

Yantosca, R. M., Long, M. S., Payer, M., and Cooper, M.: GEOS-Chem v9-01-03 Online User's Guide, available at: http://acmg.seas.harvard.edu/geos/doc/man/, last access: 22 December 2012, 2012.

Yienger, J. J. and Levy, H.: Empirical model of global soil-biogenic NO_x emissions, J. Geophys. Res., 100, 11447, doi:10.1029/95JD00370, 1995.

Zhang, Y., Pun, B., Wu, S.-Y., Vijayaraghavan, K., and Seigneur, C.: Application and evaluation of two air quality models for particulate matter for a Southeastern US episode, J. Air Waste Manage., 54, 1478–1493, doi:10.1080/10473289.2004.10471012, 2004.

Table A1. Carbon Bond '05 (CB05) species mapping in the form CB05 Species, GEOS-Chem expression.

O_3, O_x–NO_x	PANX, PPN + PMN	PAR, 3. * ACET
N_2O_5, N_2O_5	OLE, 0.5 * 1./2. * 3. * PRPE	PAR, 4. * MEK
HNO_3, HNO_3	IOLE, 0.5 * 1./4. * 3. * PRPE	PAR, 1. * BENZ
PNA, HNO_4	TOL, TOLU	ALDX, RCHO
H_2O_2, H_2O_2	XYL, XYLE	ETH, ETH
NTR, R4N2	ISPD, MACR + MVK	HO_2, HO_2
FORM, CH_2O	SO_2, SO_2	HONO, HONO
ALD2, 1./2 * ALD2	ETHA, C_2H_6	MGLY, MGLY
CO, CO	BENZENE, BENZ	NO, NO
MEPX, MP	ISOP, ISOP	NO_2, NO_2
PAN, PAN	PAR, 1.5 * C_3H_8	NO_3, NO_3
TERP,ALPH + LIMO + ALCO	PAR, 4. * ALK4	

Table A2. SAPRC07 species mapping in the form SAPRC07 Species, GEOS-Chem expression.

ACETONE, ACET	MACR, MACR	RNO_3, R4N2
ALK1, C_2H_6	MAPAN, PMN	ROOH, ETP
ALK2, C_3H_8	MEK, MEK/3	ROOH, IAP
ALK3, ALK4/2	MEOH, MOH	ROOH, INPN
ALK4, ALK4/4	MGLY, MGLY	ROOH, ISNP
ALK5, ALK4/4	MVK, MVK	ROOH, MAOP
BENZENE, BENZ	MXYL, XYLE/3	ROOH, MRP
CCHO, ALD2/3	N_2O_5, N_2O_5	ROOH, PP
CCOOH, ACTA	NH_3, NH_3	ROOH, PRPN
CCOOOH, MAP	NO, NO	ROOH, R4P
CO, CO	NO_2, NO_2	ROOH, RA3P
COOH, MP	NO_3, NO_3	ROOH, RB3P
HCHO, CH_2O	O_3, O_x - NO_x	ROOH, RIP
HNO_3, HNO_3	OXYL, XYLE/3	ROOH, RP
HNO_4, HNO_4	PAN, PAN	ROOH, VRP
HO_2H, H_2O_2	PAN2, PPN	SO_2, SO_2
HOCCHO, GLYC	PRD2, MEK * 2/3	TERP,ALPH + LIMO + ALCO
HONO, HNO_2	PROPENE, PRPE	TOLUENE, TOLU
ISOPRENE, ISOP	PXYL, XYLE/3	
	RCHO, RCHO	

Table A3. CMAQ Aerosols version 6 (AE6) in the form AE6 Species, GEOS-Chem expression.

AALJ, 0.05695 * DST1	ANO3J, 0.99 * NIT	ATIJ, 0.0028 * DST1
AALKJ, AALKJ	ANO3K, 0.0016 * DST2	ATOL1J, 0.04 * SOA5
ABNZ1J, 0.12 * SOA5	ANO3K, 0.0016 * DST3	ATOL2J, 0.04 * SOA5
ABNZ2J, 0.04 * SOA5	ANO3K, 0.0016 * DST4	ATOL3J, 0.29 * SOA5
ABNZ3J, 0.32 * SOA5	ANO3K, NITs	ATRP1J, 0.33 * SOA1
ACAJ, 0.0118 * SALA	AOLGAJ, AOLGAJ	ATRP1J, 0.33 * SOA2
ACAJ, 0.07940 * DST1	AOLGBJ, AOLGBJ	ATRP2J, 0.67 * SOA1
ACLJ, 0.00945 * DST1	AOTHRJ, 0.50219 * DST1	ATRP2J, 0.67 * SOA2
ACLJ, 0.5538 * SALA	APNCOMI, 0.4 * 0.001 * OCPI	AXYL1J, 0.03 * SOA5
ACLK, 0.01190 * DST2	APNCOMI, 0.4 * 0.001 * OCPO	AXYL2J, 0.01 * SOA5
ACLK, 0.01190 * DST3	APNCOMJ, 0.4 * 0.999 * OCPI	AXYL3J, 0.11 * SOA5
ACLK, 0.01190 * DST4	APNCOMJ, 0.4 * 0.999 * OCPO	NH3, NH3
ACLK, 0.5538 * SALC	APNCOMJ, 0.0043 * DST1	NUMACC, NUMACC
ACORS, ACORS	APOCI, 0.001 * OCPI	NUMATKN, NUMATKN
AECI, 0.001 * BCPI	APOCI, 0.001 * OCPO	NUMCOR, NUMCOR
AECI, 0.001 * BCPO	APOCJ, 0.999 * OCPI	SRFACC, SRFACC
AECJ, 0.999 * BCPI	APOCJ, 0.999 * OCPO	SRFATKN, SRFATKN
AECJ, 0.999 * BCPO	APOCJ, 0.01075 * DST1	SRFCOR, SRFCOR
AFEJ, 0.03355 * DST1	ASEACAT, 0.3685 * SALC	SULF, SULF
AISO1J, 0.75 * SOA4	ASIJ, 0.19435 * DST1	SV_ALK, SV_ALK
AISO2J, 0.25 * SOA4	ASO4I, 0.01 * SO4	SV_BNZ1, 0.06 * SOG5
AISO3J, AISO3J	ASO4J, 0.99 * SO4	SV_BNZ2, 0.23 * SOG5
AKJ, 0.0114 * SALA	ASO4J, 0.0225 * DST1	SV_ISO1, 0.75 * SOG4
AKJ, 0.03770 * DST1	ASO4J, 0.0776 * SALA	SV_ISO2, 0.25 * SOG4
AMGJ, 0.0368 * SALA	ASO4K, 0.0776 * SALC	SV_SQT, SOG3
AMNJ, 0.00115 * DST1	ASO4K, 0.02655 * DST2	SV_TOL1, 0.23 * SOG5
ANAJ, 0.3086 * SALA	ASO4K, 0.02655 * DST3	SV_TOL2, 0.23 * SOG5
ANAJ, 0.03935 * DST1	ASO4K, 0.02655 * DST4	SV_TRP1, 0.33 * SOG1
ANH4I, 0.01 * NH4	ASO4K, SO4s	SV_TRP1, 0.33 * SOG2
ANH4J, 0.00005 * DST1	ASOIL, 0.95995 * DST2	SV_TRP2, 0.67 * SOG1
ANH4J, 0.99 * NH4	ASOIL, 0.95995 * DST3	SV_TRP2, 0.67 * SOG2
ANO3I, 0.01 * NIT	ASOIL, 0.95995 * DST4	SV_XYL1, 0.19 * SOG5
ANO3J, 0.00020 * DST1	ASQTJ, SOA3	SV_XYL2, 0.06 * SOG5

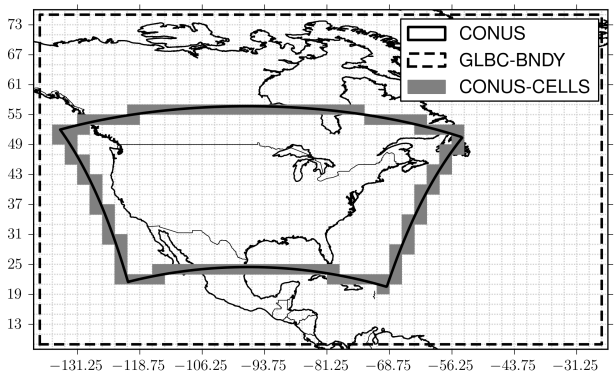

Fig. A1. GEOS-Chem lateral boundary condition output domain (GLBC; black dashed line) with the CONUS domain (black line) and grid cells that intersect the CONUS domain boundary.

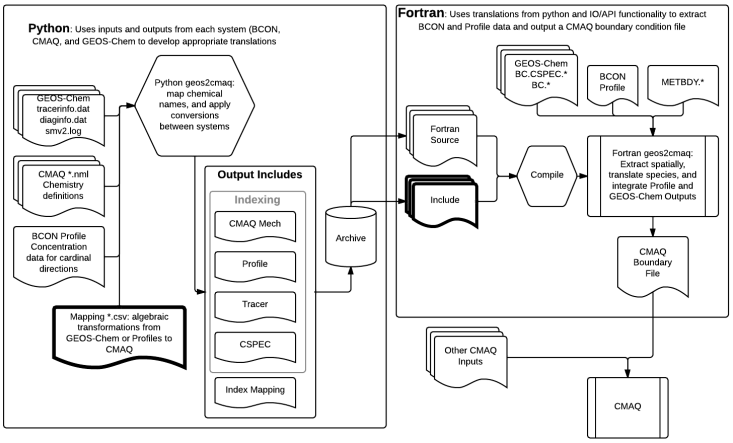

Fig. A2. Program description and flow UML diagram. The BCON and BC.CSPEC.* files are not required. Heavy lined inputs represent geos2cmaq specific inputs or outputs (i.e., not also necessary for standard run).

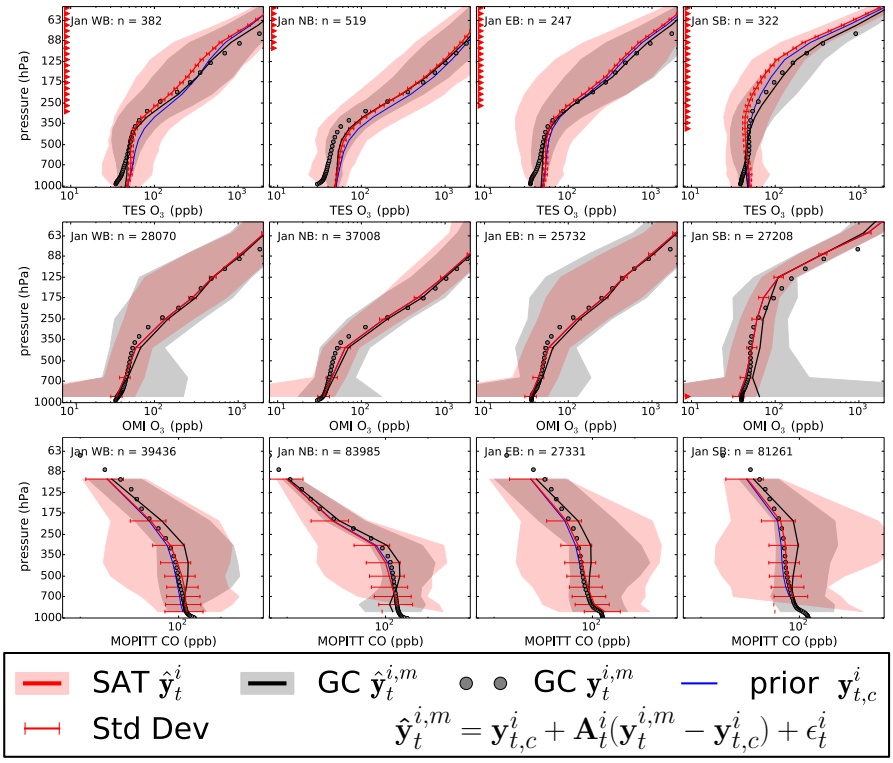

Fig. A3. Ozone and carbon monoxide mixing ratios (ppb) for January as observed by TES (O_3, row 1), OMI (O_3, row 2), and MOPITT (CO, row 3) (SAT $\widehat{y}_t^i$, red) and retrievals from GEOS-Chem (GC $\widehat{y}_t^{i,m}$, black). GEOS-Chem retrievals are calculated by applying the satellite averaging kernel to the GEOS-Chem prediction (GC $\widehat{y}_t^{i,m}$, grey dots), which relies on the a prior (a prior $\widehat{y}_{t,c}^i$, blue). Lines or dots represent median values, the shaded area represents the range of values, and satellite uncertainty is shown as error bars. Red and blue triangles show high (red) and low (blue) biases as defined by 2 times the satellite error for the median value.

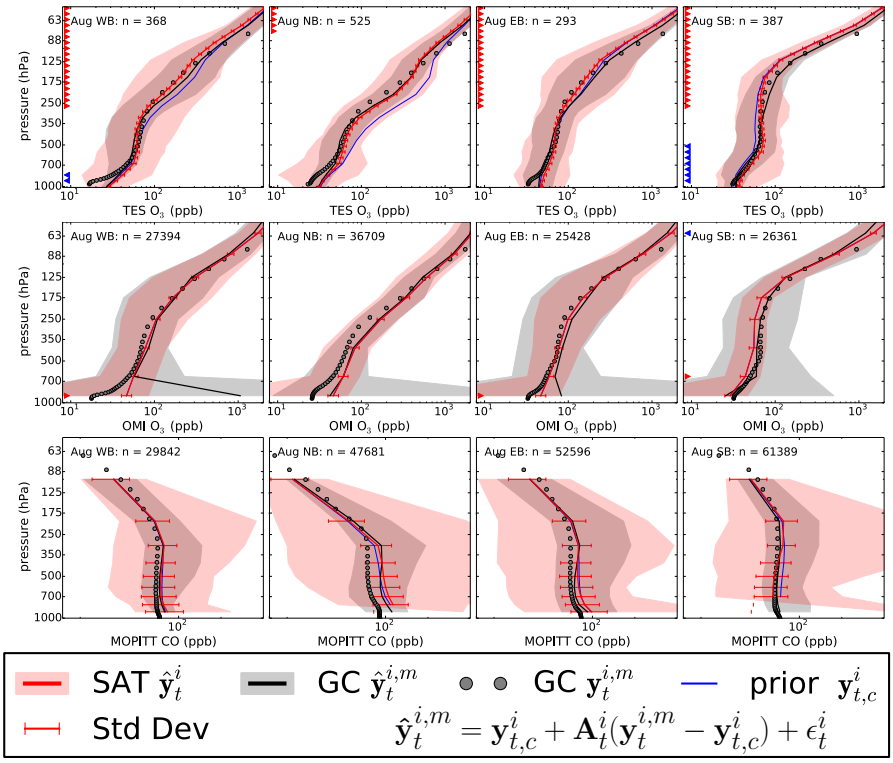

Fig. A4. same as Fig. A3 for August.

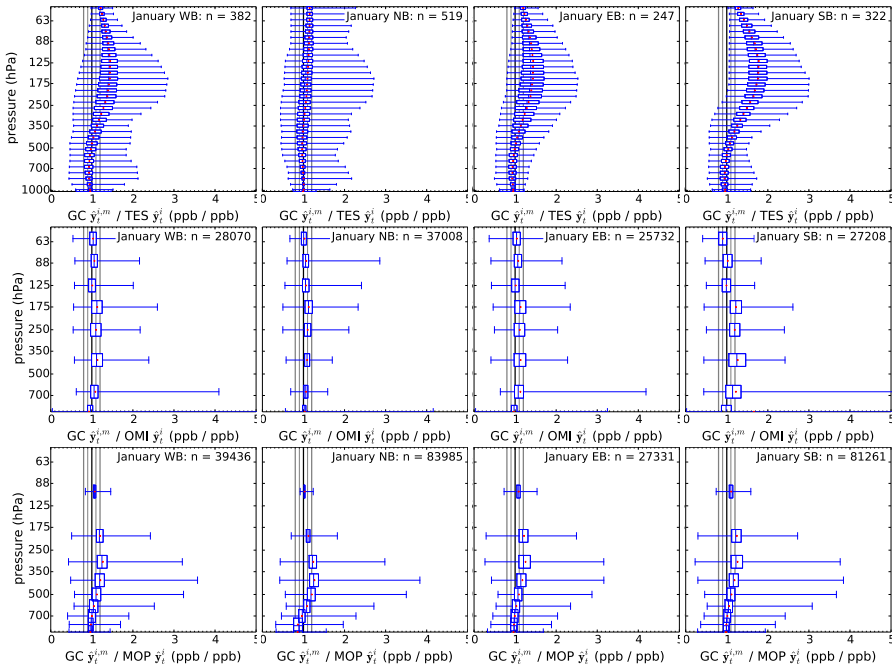

Fig. A5. Individual retrieval relative bias shown as boxplots for each altitude bin in each satellite product (TES, OMI, and MOP=MOPITT). Whiskers indicate min/max, the box represents the interquartile range, the blue line in the box is the median and the red cross is the mean. Vertical gray lines delineate the $\pm 10\,\%$ (fine) and $\pm 20\,\%$ (heavy) bias ranges.

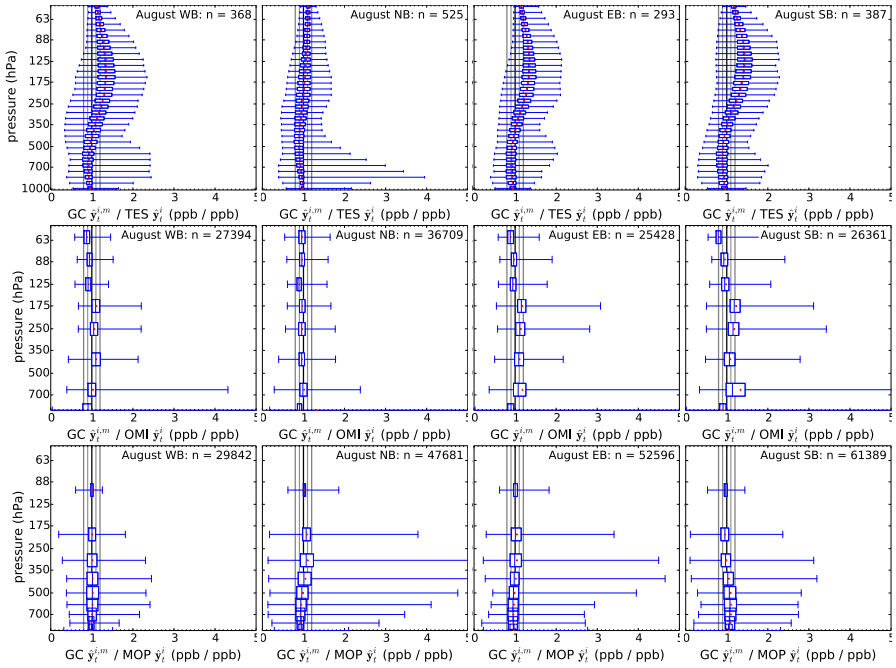

Fig. A6. same as Fig. A5 for August.

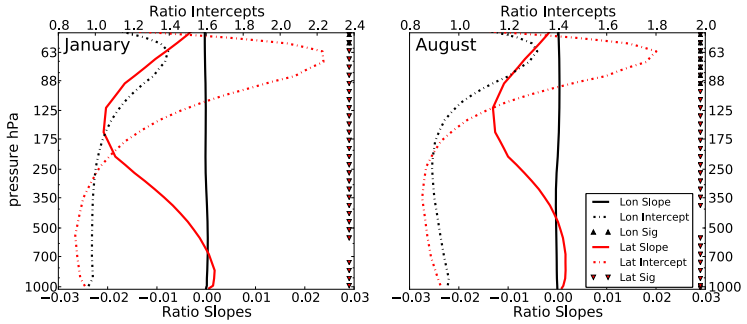

Fig. A7. Linear regression of slope (solid) and intercept (dash-dot) for the ratio of simulated retrieval to TES satellite retrieval as a function of longitude (black) and latitude (red) for January (left) and August (right).

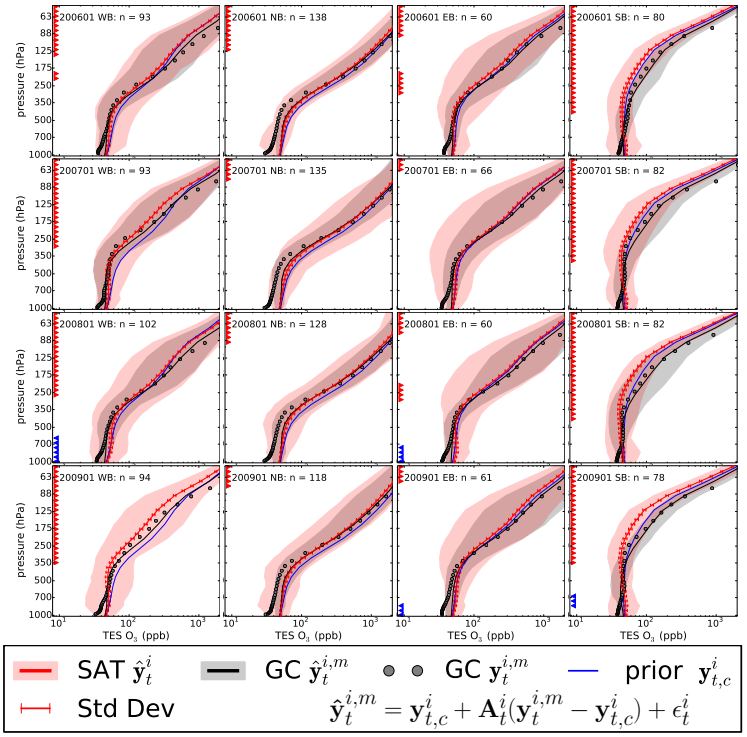

Fig. A1. Ozone retrievals from TES for January of 2006, 2007, 2008, and 2009 for each boundary face (West, North, East, South) observed by TES (TES $\widehat{y}_t^i$, red) and as retrieved from GEOS-Chem (GC $\widehat{y}_t^{i,m}$, black). GEOS-Chem retrievals are calculated by applying the TES averaging kernel to the GEOS-Chem prediction (GC $\widehat{y}_t^{i,m}$, grey dots), which relies on the a prior (a prior $\widehat{y}_{t,c}^i$, blue). Lines or dots represent median values, the shaded area represents the range of values, and TES uncertainty is shown as error bars. Red and blue triangles show high (red) and low (blue) biases as defined by 2 times the TES error for the median value.

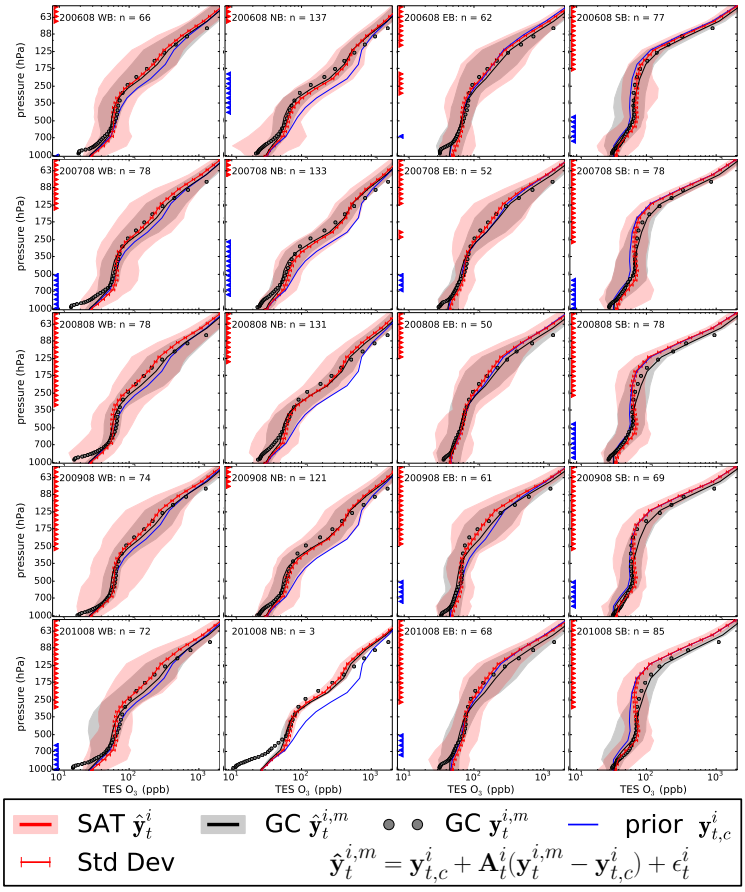

Fig. A2. Same as Fig. A1, but for August and includes year 2010.

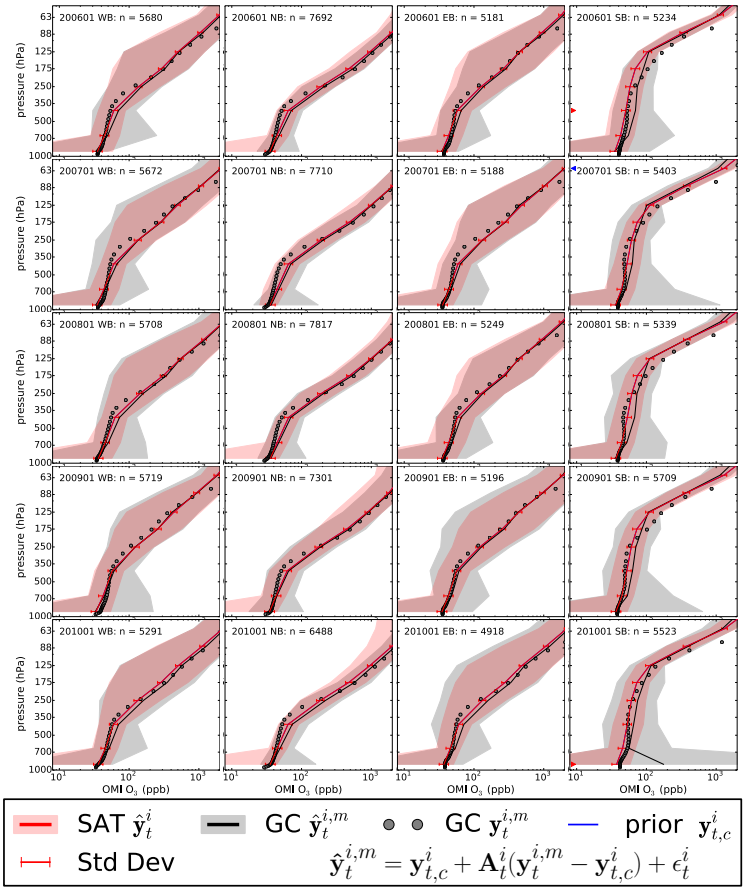

Fig. A3. Same as Fig. A1, but for OMI.

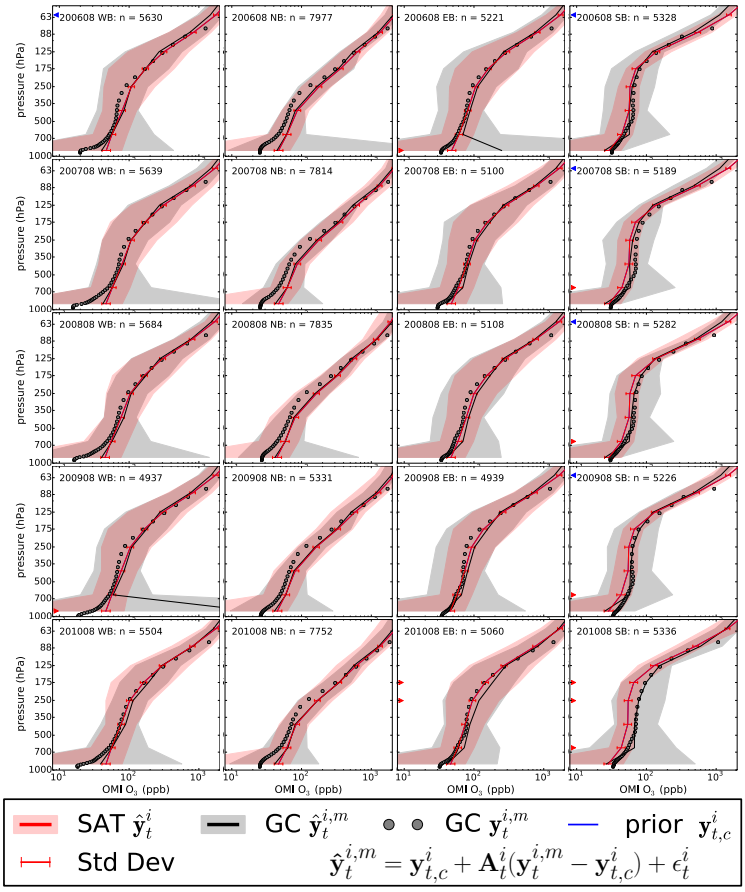

Fig. A4. Same as Fig. A2, but for OMI.

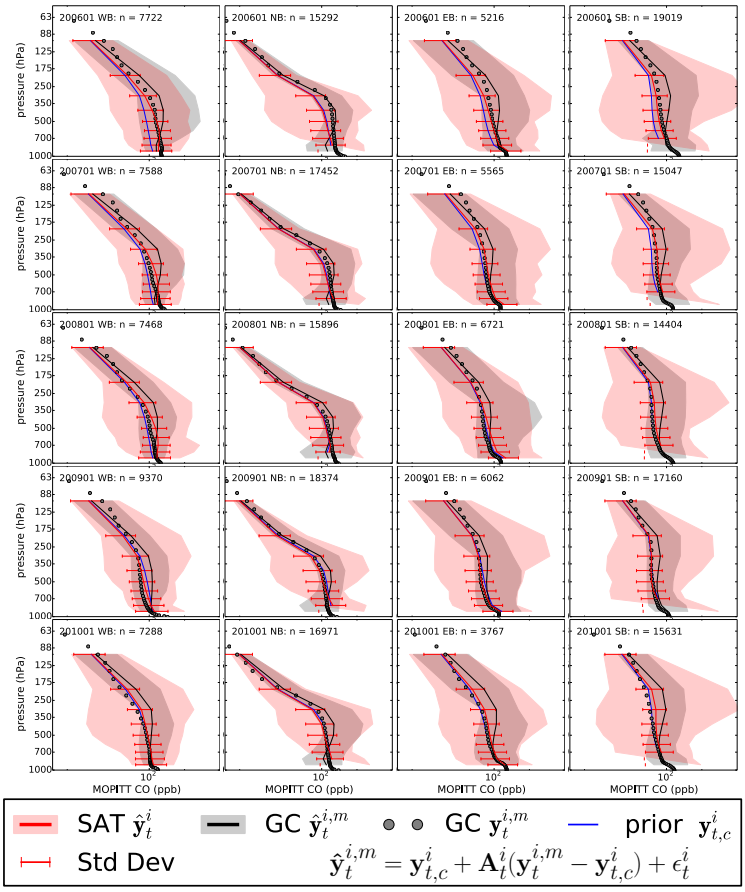

Fig. A5. Same as Fig. A1, but for MOPITT.

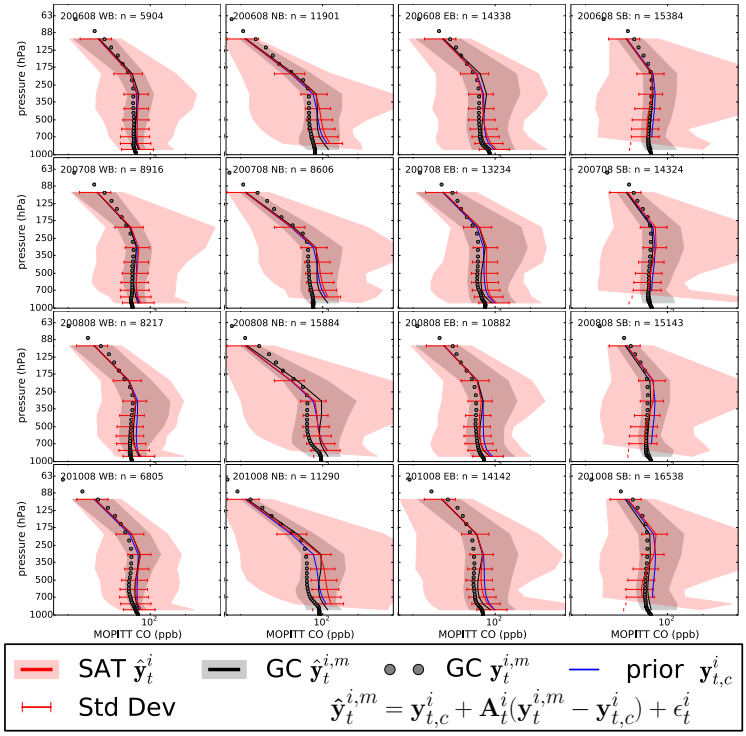

Fig. A6. Same as Fig. A2, but for MOPITT, does not have year 2009, and includes year 2010.