Geosci. Model Dev. Discuss., 6, 4665–4704, 2013
www.geosci-model-dev-discuss.net/6/4665/2013/
doi:10.5194/gmdd-6-4665-2013


Geoscientific Model Development Discussions
Open Access



# 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 Environmental Protection Agency, Research Triangle Park, North Carolina, USA

Received: 25 July 2013 – Accepted: 15 August 2013 – Published: 9 September 2013

Correspondence to: B. H. Henderson (barronh@ufl.edu)

Published by Copernicus Publications on behalf of the European Geosciences Union.







## 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 2000–2010 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 vertical profiles. The results show performance is largely within uncertainty estimates for the Tropospheric Emission Spectrometer (TES) with some exceptions. The major difference shows a high bias in the upper troposphere along the southern boundary in January. This publication documents the global simulation database, the tool for conversion to LBC, and the fidelity 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 tightened (40 CFR 50.10); (2) influence of international activities has increased average





GMDD
6, 4665–4704, 2013



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 (AQM) 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 ($O_3$ = 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 LBC 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





GMDD
6, 4665–4704, 2013



to provide estimates of air pollution concentrations with time resolution ranges 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 output. The third section details the methods and results of evaluating GLBC using satellite observations. The conclusions review the usability of the tool and the





GMDD
6, 4665–4704, 2013



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 and conversion to regional model ready boundary files.

GEOS-Chem is a research-grade atmospheric model 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 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. To reduce computational burden, GEOS-Chem combines many 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.





GMDD
6, 4665–4704, 2013



All simulations are based upon GEOS-Chem's $NO_x$-$O_x$-hydrocarbon-aerosol application with the optional Secondary Organic Carbon Aerosol module enabled. An update in the chemistry mechanism between versions v8-02-01 and v8-02-04 contained a bug fix which led to a decrease in simulated ozone chemical loss. 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. Improvements to halogen and heterogeneous aerosol chemistry have shown promise in reducing this high-bias (Mao et al., 2013), but are not included in production simulations. As these updates are still underway, the work here does not include them. Our v8-02-03 GEOS-Chem simulations follow the recommended settings. A 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 (NEI; US EPA, 2013) (see note 3), Europe (UNECE/EMEP; Vestreng and Klein, 2002), Mexico (BRAVO; Kuhns et al., 2003), Canada (CAC, Environment Canada, 2013), and Asia (Streets et al., 2003, 2006). In addition, the emissions included additional source: lightning $NO_x$ (Ott et al., 2010; Pickering et al., 1998; Price and Rind, 1992), soil $NO_x$ (Wang et al., 1998; Yienger and Levy, 1995), oceanic Dimethyl Sulfide, volcanic $SO_2$, sea salt, wind-blown mineral dust, wildfires 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 simulations were conducted based upon the availability of meteorological data for the desired modeling period (2000–2010). For simulations prior to 2004, GEOS-Chem requires use of the MERRA (Rienecker et al., 2011) meteorological dataset and





GMDD
6, 4665–4704, 2013



version 9 of the GEOS-Chem model. For simulations after 2004, GEOS-Chem version 8 using the GEOS-5 assimilated meteorology dataset (Molod et al., 2012) is recommended. See the GEOS-Chem documentation (http://acmg.seas.harvard.edu/geos/) for a description of changes between model releases. Overall using a mature version 8 of GEOS-Chem and the GEOS-5 meteorology dataset is recommended, as some of the changes in the latest released version 9 have not been as fully evaluated. Additional details of the model setup for each of the available simulations are listed in Table 1.

## 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 preprocessor 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-





GMDD
6, 4665–4704, 2013



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 in green 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





GMDD
6, 4665–4704, 2013



9. ASO4K, 0.0776 * SALC + 0.02655 * (DST2 + DST3 + DST4) + 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.

Lines 4–8 demonstrate that additional lines are additive.

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 3–7 demonstrate how GEOS-Chem aerosols such as SALC and DST2 are mapped based on CMAQ emission profiles for assignment to coarse mode sulfate. Because the lines are additive, these lines could be re-written as line 8 but lines 3–7 and 8 should not both be included. The mapping expressions can include all standard python operators (+, –, *, /, **, %, etc), but math functions (e.g., sin) are not currently available.

There are 5 types of factors that might 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.





GMDD
6, 4665–4704, 2013



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 preprocessor 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 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 the 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. Using the centroid locations from METBDY3D, the software uses a nearest-neighbor approach to identify a corresponding GEOS-Chem row and column. Figure 1 shows the intersection of an example boundary and the GEOS-Chem





GMDD
6, 4665–4704, 2013



outputs. Using the sigma-P coordinates (native and derived), the method then linearly interpolates concentrations from GEOS-Chem layer centers to CMAQ layer centers.

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 then, 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), previously 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 satellites 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 the Tropospheric Emission Spectrometer (TES) satellite retrievals. The TES instrument uses infrared





GMDD
6, 4665–4704, 2013



Fourier transform spectroscopy to retrieve ozone vertical profiles (Bowman et al., 2011) from the Aura satellite using nadir scanning. 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.

To evaluate the model, we pair TES observations with GEOS-Chem grid cells from two years, 2006 and 2008. January results are selected to represent winter and August results are selected to represent the traditional ozone season. The GEOS-Chem grid cells are filtered for just those that would be used in creating CONUS boundary conditions (see Fig. 1). Grid cells are paired with TES when the swath centroid is contained within cell. After pairs have been identified, the GEOS-Chem prediction is processed using Eq. (1), which is adapted from Bowman et al. (2011, Eqs. 5–8).

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

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

The evaluation has been performed on groups of individual retrievals based upon similar bias features. Based on TES swath centroid locations, there are a total of 2139 GEOS-Chem and TES pairs during the comparison time period. For these comparisons, biases were initially reviewed for 24 categories (3 yr × 2 months × 4 perimeter cardinal edges). From those categories, four emerged as distinct cases. The difference between years was nominal and is not highlighted here, but is included in the Appendix.





GMDD
6, 4665–4704, 2013



Instead results are grouped only by month across years (individual years are shown in the Figs. A3–A8). The difference between perimeter edges was most interesting between north and south. As a result, the west and east perimeter have been bisected equally and allocated to either north or south. The west and east boundaries, with bisected north and south, are shown in the Appendix. As discussed further below, the north/south divide dominates the bias in a somewhat offsetting manner. As a result, the west and east overall performance (Figs. A1 and A2) is nominally better than either the south or north.

Figure 3 shows ozone (ppb) for the remaining four categories (northern-January, southern-January, northern-August, southern-August) from GEOS-Chem, GEOS-Chem retrievals, and TES retrievals. To aid in interpretation, GEOS-Chem biases have been highlighted using triangles on the *y* scale (red = high; blue = low) when greater than the TES observation uncertainty. The mean and range of profiles show good correspondence most of the time. In January, mixing ratios and variability increases with altitude. In August, mixing ratios increase with altitude, while variability decreases. In the north, the model and TES have lower mixing ratios than the prior above 700 hPa, which demonstrates the sensitivity of this retrieval to the model/TES signal.

In Fig. 3, the two main areas of bias can be seen in more detail. The first is evident along the northern boundary in both January and August. In the north, GEOS-Chem shows a low bias relative to TES in the upper troposphere (350–125 hPa). The southern boundary in January is high-biased in the upper troposphere (> 350 hPa). This high-bias is only found in winter.

To further explore these aggregate biases, Fig. 4 shows the distribution of individual retrieval biases. The biases in Fig. 4 are shown as the ratio of retrieval mixing ratios (i.e., ppb). To reiterate, this type of comparison is not dependent upon the prior – only the sensitivity of the instrument. A comparison of the range of biases is shown in Fig. 4, rather than range of concentration. As in Fig. 3, the southern boundary in August has the best performance. Southern-August results show that 53 % of the results are within ±10 % (80 % within ±20 %). The northern-August results show slightly





GMDD
6, 4665–4704, 2013



worse performance with 44 % within ±10 % (73 % within ±20 %). In January when performance was worse, the north has 40 % of results within ±10 % (67 % within ±20 %) and the south has 40 % within ±10 % (68 % within ±20 %). Given the potential bias in TES results (Bowman et al., 2011; Nassar et al., 2008), this level of performance is acceptable.

## 5 Conclusions

We describe and evaluate a tool for using global simulations to produce LBC for regional air quality models. In general, the LBC had biases that are acceptable given the uncertainties of the TES retrievals. Our evaluation showed better performance for August compared to January. A higher bias was found in the upper troposphere on the southern boundary in January that has longer ozone life times. The altitude and time of the bias suggests an over-estimation of long-distance transport. Whether the error comes from simulated transport or emissions is not known, but Southeast Asia emissions are suspect based on their high uncertainty.

We propose that the presented tool provides a resource to better represent global transport through boundary conditions in regional air quality studies. The tool's evaluation demonstrates the fitness of produced LBC.

## Appendix A

The Appendix contains species mapping for common gas-phase and aerosol mechanisms and more detailed evaluation of ozone 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 boundaries as split between north and south and without year segregation. The Appendix provides information on east and west boundaries and on annual segregation of north/south boundaries.





GMDD
6, 4665–4704, 2013



## A1 Additional temporal and spatial details

See Figs. A1–A8.

## 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 ($SO_4$s) and nitrate formed on seasalt ($NO_3$s) (Alexander, 2005) are mapped to the CMAQ coarse mode. 99.9 % of primary carbonaceous





GMDD
6, 4665–4704, 2013



aerosols from GEOS-Chem are attributed to the accumulation mode while 0.1 % are assigned to the Aitken mode consistent with CMAQ emissions processing. 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.





GMDD
6, 4665–4704, 2013



## 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.

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.






GMDD
6, 4665–4704, 2013



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.
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.





GMDD
6, 4665–4704, 2013



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.





GMDD
6, 4665–4704, 2013



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.





GMDD
6, 4665–4704, 2013



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.

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





GMDD
6, 4665–4704, 2013



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.





GMDD
6, 4665–4704, 2013



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.





GMDD
6, 4665–4704, 2013



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.

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.





GMDD
6, 4665–4704, 2013



**Table 1.** GEOS-Chem Annual Simulations for CMAQ boundaries (recommended in bold).

| GEOS-Chem version | Chemistry version | Meteorology | Shipping emissions[a] | Simulation years[b] |
|---|---|---|---|---|
| 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] Years shown are inclusive. First year is spinup.





GMDD
6, 4665–4704, 2013



**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 |
| N2O5, N2O5 | 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, HNO4 | TOL, TOLU | ALDX, RCHO |
| H2O2, H2O2 | XYL, XYLE | ETH, ETH |
| NTR, R4N2 | ISPD, MACR + MVK | HO2, HO2 |
| FORM, $CH_2O$ | SO2, SO2 | HONO, HONO |
| ALD2, 1./2 * ALD2 | ETHA, C2H6 | MGLY, MGLY |
| CO, CO | BENZENE, BENZ | NO, NO |
| MEPX, MP | ISOP, ISOP | NO2, NO2 |
| PAN, PAN | PAR, 1.5 * C3H8 | $NO_3$, $NO_3$ |
| TERP,ALPH + LIMO + ALCO | PAR, 4. * ALK4 | |





GMDD
6, 4665–4704, 2013



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

| | | |
|---|---|---|
| ACETONE, ACET | MACR, MACR | $RNO_3$, R4N2 |
| ALK1, C2H6 | MAPAN, PMN | ROOH, ETP |
| ALK2, C3H8 | 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 |
| BUTADIENE13, | N2O5, N2O5 | ROOH, PP |
| CCHO, ALD2/3 | NH3, NH3 | ROOH, PRPN |
| CCOOH, ACTA | NO, NO | ROOH, R4P |
| CCOOOH, MAP | NO2, NO2 | ROOH, RA3P |
| CO, CO | $NO_3$, $NO_3$ | ROOH, RB3P |
| COOH, MP | $O_3$, $O_x$ - $NO_x$ | ROOH, RIP |
| HCHO, $CH_2O$ | OXYL, XYLE/3 | ROOH, RP |
| $HNO_3$, $HNO_3$ | PAN, PAN | ROOH, VRP |
| HNO4, HNO4 | PAN2, PPN | SO2, SO2 |
| HO2H, H2O2 | PRD2, MEK * 2/3 | TERP,ALPH + LIMO + ALCO |
| HOCCHO, GLYC | PROPENE, PRPE | |
| HONO, HNO2 | PXYL, XYLE/3 | TOLUENE, TOLU |
| ISOPRENE, ISOP | RCHO, RCHO | |





GMDD
6, 4665–4704, 2013



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

| | | |
|---|---|---|
| AALJ, 0.05695 * DST1 | $ANO_3J$, 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, | AXYL1J, 0.03 * SOA5 |
| ACLK, 0.01190 * DST2 | 0.4 * 0.001 * OCPI | AXYL2J, 0.01 * SOA5 |
| ACLK, 0.01190 * DST3 | APNCOMI, | AXYL3J, 0.11 * SOA5 |
| ACLK, 0.01190 * DST4 | 0.4 * 0.001 * OCPO | NH3, NH3 |
| ACLK, 0.5538 * SALC | APNCOMJ, | NUMACC, NUMACC |
| ACORS, ACORS | 0.4 * 0.999 * OCPI | NUMATKN, NUMATKN |
| AECI, 0.001 * BCPI | APNCOMJ, | NUMCOR, NUMCOR |
| AECI, 0.001 * BCPO | 0.4 * 0.999 * OCPO | SRFACC, SRFACC |
| AECJ, 0.999 * BCPI | APNCOMJ, 0.0043 * DST1 | SRFATKN, SRFATKN |
| AECJ, 0.999 * BCPO | APOCI, 0.001 * OCPI | SRFCOR, SRFCOR |
| AFEJ, 0.03355 * DST1 | APOCI, 0.001 * OCPO | SULF, SULF |
| AISO1J, 0.75 * SOA4 | APOCJ, 0.999 * OCPI | SV_ALK, SV_ALK |
| AISO2J, 0.25 * SOA4 | APOCJ, 0.999 * OCPO | SV_BNZ1, 0.06 * SOG5 |
| $AISO_3J$, $AISO_3J$ | APOCJ, 0.01075 * DST1 | SV_BNZ2, 0.23 * SOG5 |
| AKJ, 0.0114 * SALA | ASEACAT, 0.3685 * SALC | SV_ISO1, 0.75 * SOG4 |
| AKJ, 0.03770 * DST1 | ASIJ, 0.19435 * DST1 | SV_ISO2, 0.25 * SOG4 |
| AMGJ, 0.0368 * SALA | ASO4I, 0.01 * SO4 | SV_SQT, SOG3 |
| AMNJ, 0.00115 * DST1 | ASO4J, 0.99 * SO4 | SV_TOL1, 0.23 * SOG5 |
| ANAJ, 0.3086 * SALA | ASO4J, 0.0225 * DST1 | SV_TOL2, 0.23 * SOG5 |
| ANAJ, 0.03935 * DST1 | ASO4J, 0.0776 * SALA | SV_TRP1, 0.33 * SOG1 |
| ANH4I, 0.01 * NH4 | ASO4K, 0.0776 * SALC | SV_TRP1, 0.33 * SOG2 |
| ANH4J, 0.00005 * DST1 | ASO4K, 0.02655 * DST2 | SV_TRP2, 0.67 * SOG1 |
| ANH4J, 0.99 * NH4 | ASO4K, 0.02655 * DST3 | SV_TRP2, 0.67 * SOG2 |
| $ANO_3I$, 0.01 * NIT | ASO4K, 0.02655 * DST4 | SV_XYL1, 0.19 * SOG5 |
| $ANO_3J$, 0.00020 * DST1 | ASO4K, SO4s | SV_XYL2, 0.06 * SOG5 |
| | ASOIL, 0.95995 * DST2 | |
| | ASOIL, 0.95995 * DST3 | |
| | ASOIL, 0.95995 * DST4 | |
| | ASQTJ, SOA3 | |





GMDD
6, 4665–4704, 2013



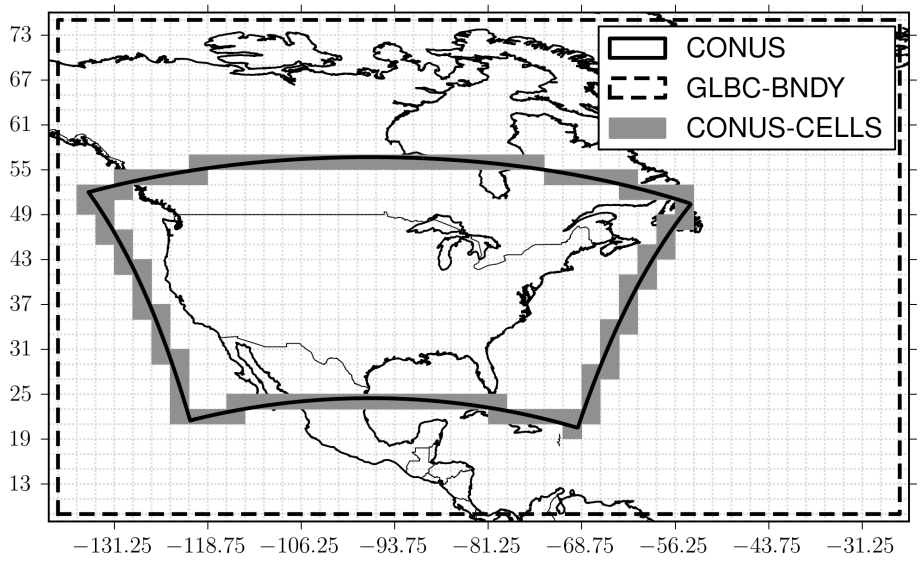


**Fig. 1.** GEOS-Chem lateral boundary condition output domain (GLBC; black dashed line) with the CONUS domain (black line) and grid cells that intersect the CONUS domain boundary.





GMDD
6, 4665–4704, 2013



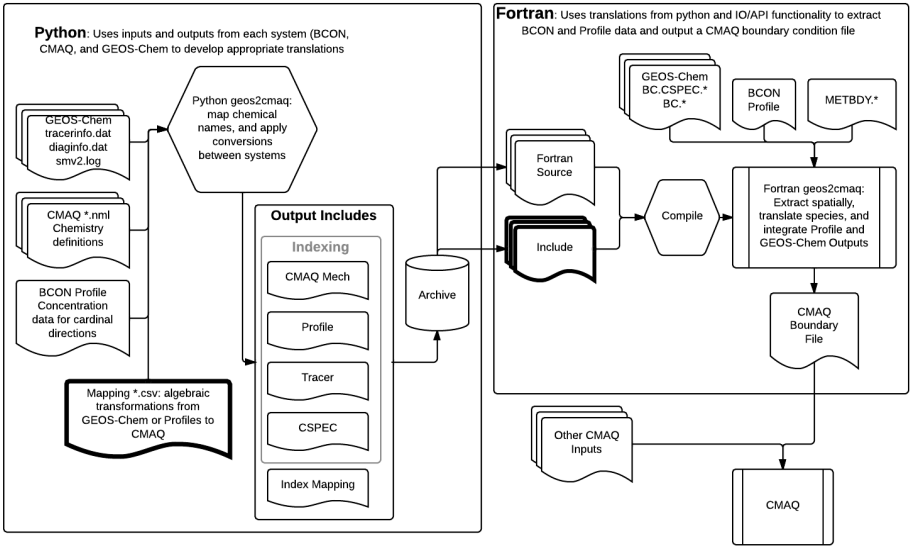


**Fig. 2.** Program description and flow UML diagram. The BCON and BC.CSPEC.* files are not required. Heavy lined inputs represent geos2cmaq specific inputs or outputs (i.e., not also necessary for standard run).





GMDD
6, 4665–4704, 2013



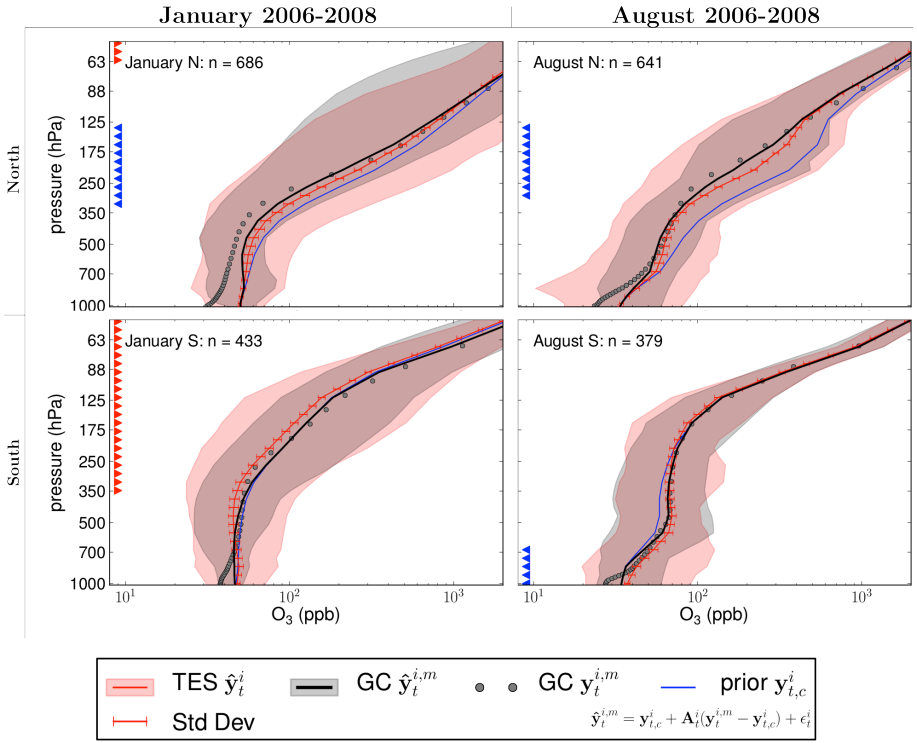


**Fig. 3.** Ozone retrievals as observed by TES (TES $\hat{y}_t^i$, red) and GEOS-Chem (GC $\hat{y}_t^{i,m}$, black). GEOS-Chem retrievals are calculated by applying the TES averaging kernel to the GEOS-Chem prediction (GC $\hat{y}_t^{i,m}$, grey dots), which relies on the prior (prior $\hat{y}_{t,c}^i$, 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.





GMDD
6, 4665–4704, 2013



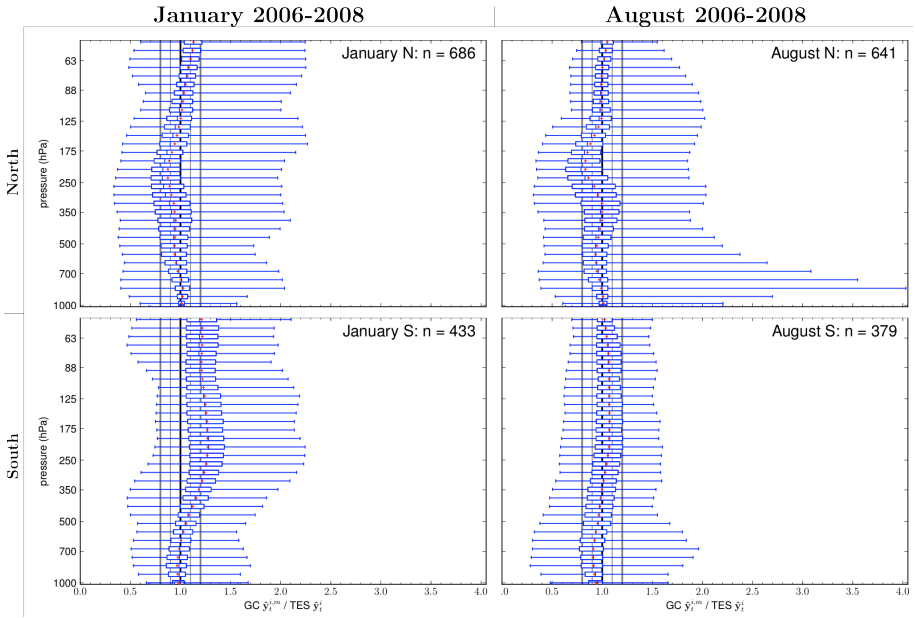


**Fig. 4.** Individual retrieval bias shown as boxplots for each altitude bin in the TES product. Whiskers indicate min/max, the box represents the interquartile range, the blue line in the box is the median and the red cross is the mean. Vertical gray lines delineate the ±10 % (fine) and ±20 % (heavy) bias ranges.





GMDD
6, 4665–4704, 2013



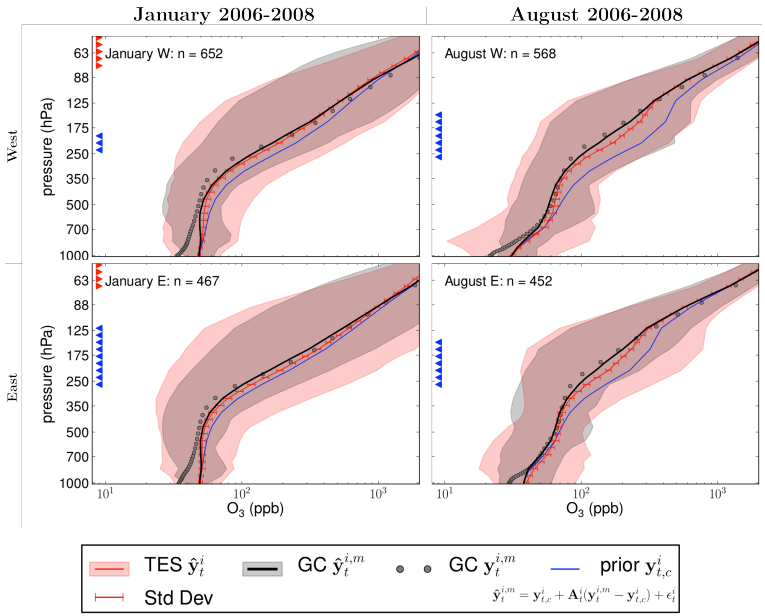


**Fig. A1.** Ozone retrievals for 2006, 2007, and 2008 for east and west boundaries as observed by TES (TES $\hat{y}_t^i$, red) and GEOS-Chem (GC $\hat{y}_t^{i,m}$, black). GEOS-Chem retrievals are calculated by applying the TES averaging kernel to the GEOS-Chem prediction (GC $\hat{y}_t^{i,m}$, grey dots), which relies on the prior (prior $\hat{y}_{t,c}^i$, 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.





GMDD
6, 4665–4704, 2013



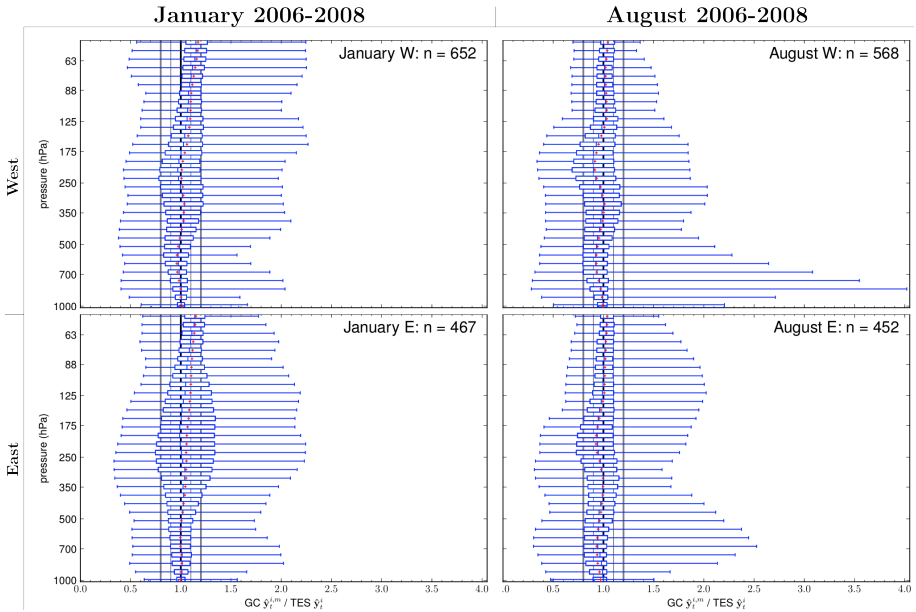


**Fig. A2.** Individual retrieval bias for 2006–2008 for east and west boundaries shown as boxplots for each altitude bin in the TES product. Whiskers indicate min/max, the box represents the interquartile range, the blue line in the box is the median and the red cross is the mean. Vertical gray lines delineate the ±10 % (fine) and ±20 % (heavy) bias ranges.





GMDD
6, 4665–4704, 2013



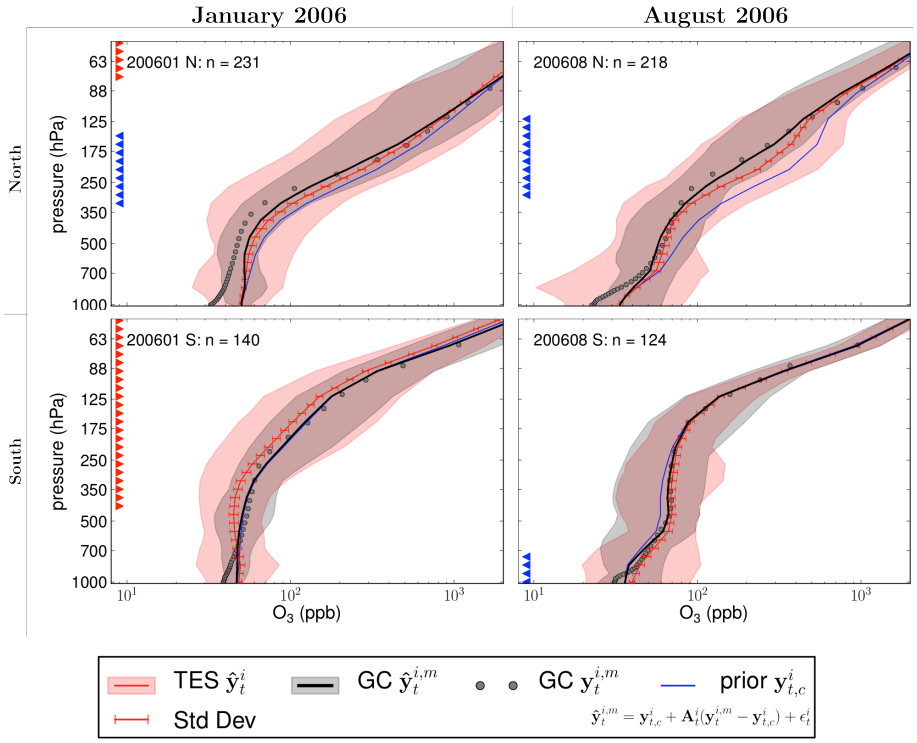


**Fig. A3.** Same as A1 for 2006 northern and southern boundaries.





GMDD
6, 4665–4704, 2013

LBC: description and evaluation

B. H. Henderson et al.



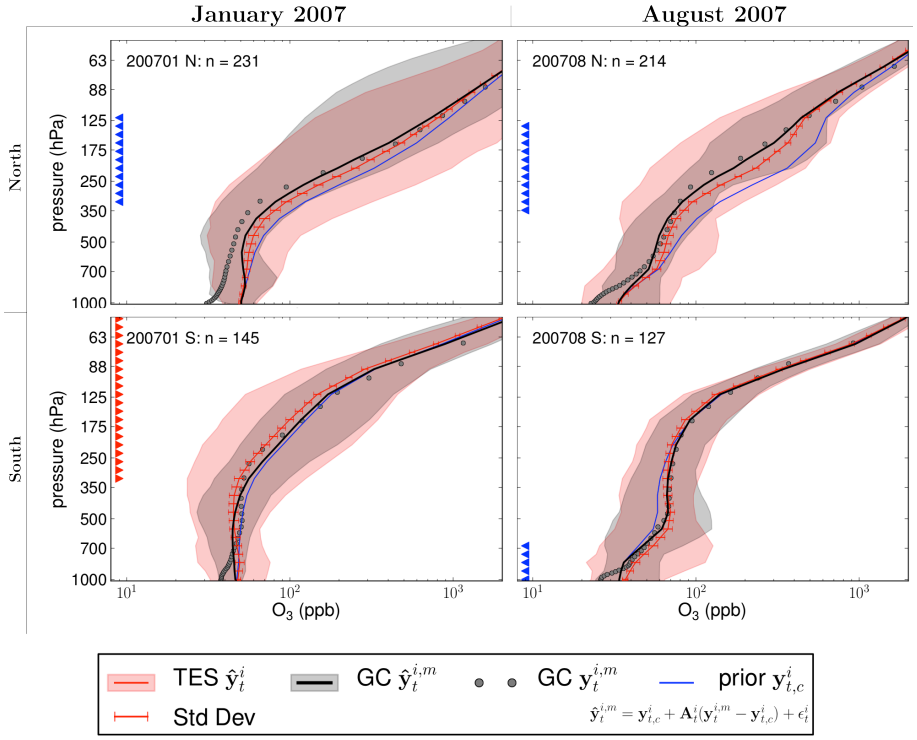


**Fig. A4.** Same as A1 for 2007 northern and southern boundaries.





GMDD
6, 4665–4704, 2013



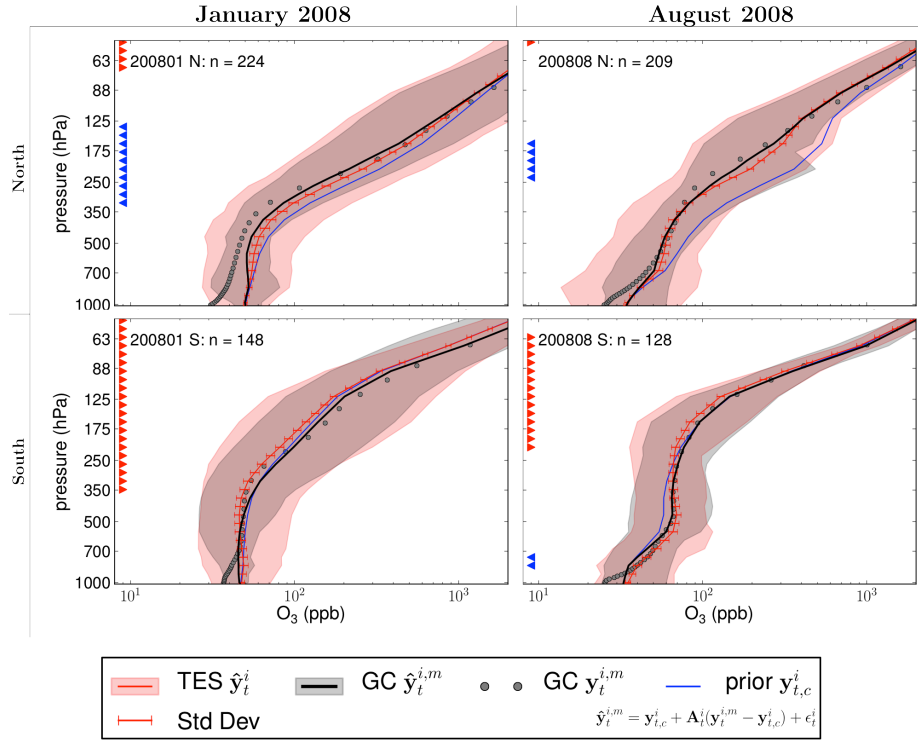


**Fig. A5.** Same as A1 for 2008 northern and southern boundaries.





GMDD
6, 4665–4704, 2013



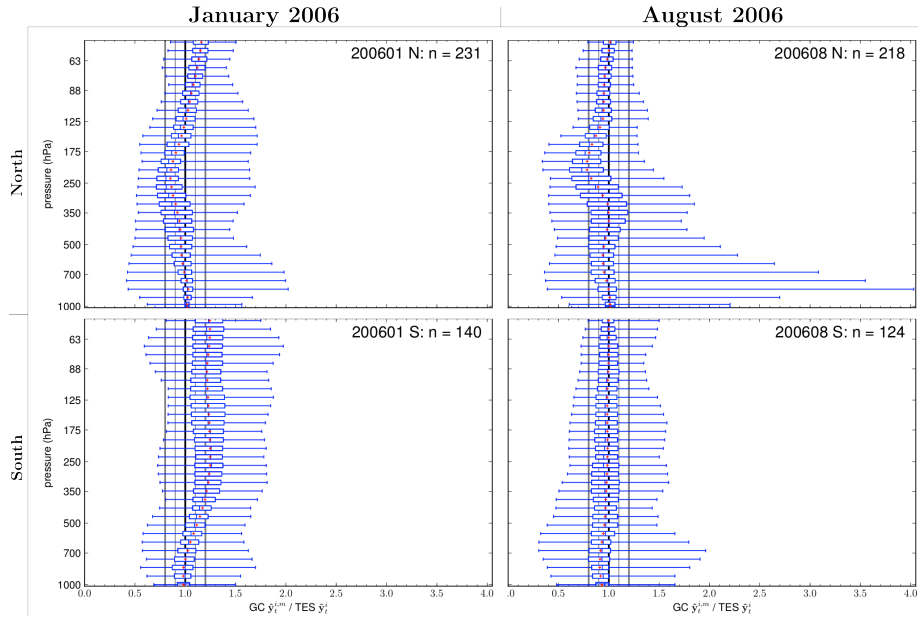


**Fig. A6.** Same as A2 for 2006 northern and southern boundaries.





GMDD
6, 4665–4704, 2013

LBC: description and evaluation

B. H. Henderson et al.



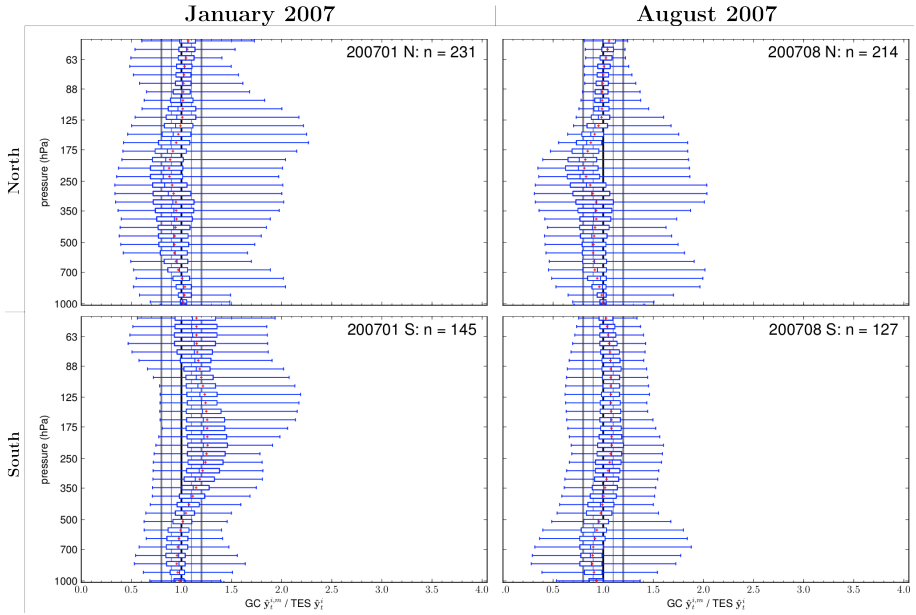


**Fig. A7.** Same as A2 for 2007 northern and southern boundaries.





GMDD
6, 4665–4704, 2013

LBC: description and evaluation

B. H. Henderson et al.



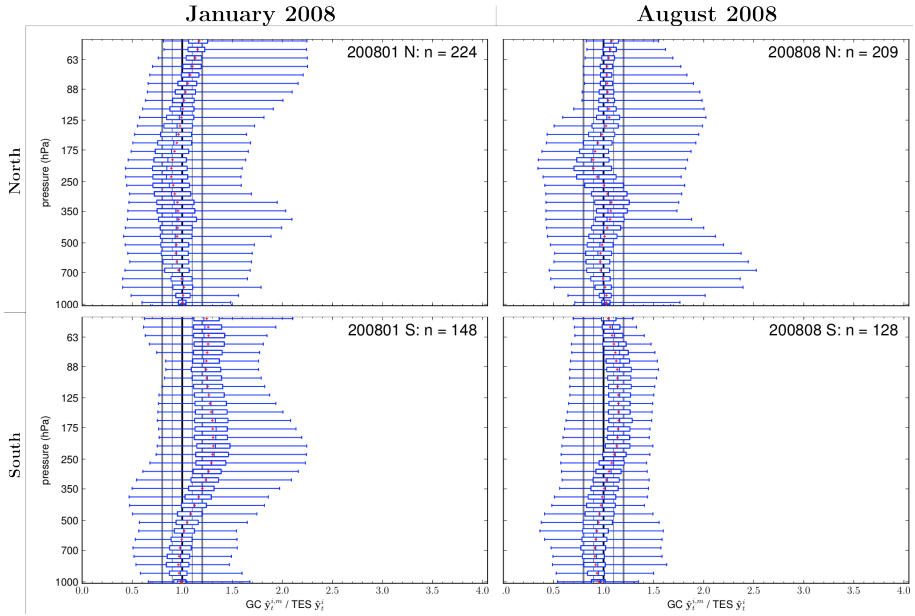


**Fig. A8.** Same as A2 for 2008 northern and southern boundaries.





GMDD
6, 4665–4704, 2013

LBC: description and evaluation

B. H. Henderson et al.