# Regional data assimilation of multi-spectral MOPITT observations of CO over North America

Zhe Jiang[1,2], Dylan B. A. Jones[1,3], John Worden[2], Helen M. Worden[4], Daven K. Henze[5], Yuxuan Wang[6,7]

[1]Department of Physics, University of Toronto, Toronto, ON, Canada, [2]Jet Propulsion Laboratory, California Institute of Technology, Pasadena CA, USA, [3]JIFRESSE, University of California, Los Angeles, Los Angeles, CA, USA, [4]National Center for Atmospheric Research, Boulder, CO, USA, [5]University of Colorado Boulder, CO, USA, [6]Department of Marine Sciences, Texas A&M University at Galveston, Galveston, TX, USA, [7]Ministry of Education Key Laboratory for Earth System Modeling, Center for Earth System Science, Institute for Global Change Studies, Tsinghua University, Beijing, China

## Abstract

Chemical transport models (CTMs) driven with high-resolution meteorological fields can better resolve small-scale processes, such as frontal lifting or deep convection, and thus improve the simulation and emission estimates of tropospheric trace gases. In this work, we explore the use of the GEOS-Chem four-dimensional variational (4D-Var) data assimilation system with the nested high-resolution version of the model (0.5°x0.67°) to quantify North American CO emissions during the period of June 2004 – May 2005. With optimized lateral boundary conditions, regional inversion analyses can reduce the sensitivity of the CO source estimates to errors in long-range transport and in the distributions of the hydroxyl radical (OH), the main sink for CO. To further limit the potential impact of discrepancies in chemical aging of air in the free troposphere, associated with errors in OH, we use surface level multispectral MOPITT CO retrievals, which have greater sensitivity to CO near the surface and reduced sensitivity in the free troposphere, compared to previous versions of the retrievals. We estimate that the annual

total anthropogenic CO emission from the contiguous US 48 states was 97 Tg CO, a 14% increase from the 85 Tg CO in the a priori. This increase is mainly due to enhanced emissions around the Great Lakes region and along the west coast, relative to the a priori. Sensitivity analyses using different OH fields and lateral boundary conditions suggest a possible error, associated with local North America OH distribution, in these emission estimates of 20% during summer 2004, when the CO lifetime is short. This 20% OH-related error is 50% smaller than the OH-related error previously estimated for North American CO emissions using a global inversion analysis. We believe that reducing this OH-related error further will require integrating additional observations to provide a strong constraint on the CO distribution across the domain. Despite these limitations, our results show the potential advantages of combining high-resolution regional inversion analyses with global analyses to better quantify regional CO source estimates.

## 1. Introduction

Inverse modeling is a powerful tool to improve our understanding of emissions of greenhouse gases and pollutant tracers, by combining observations of atmospheric composition with models. Despite more than a decade of inverse modeling work to better quantify emissions of atmospheric CO (e.g., Palmer et al., 2003; Pétron et al., 2004; Heald et al., 2004; Arellano et al., 2006; Jones et al., 2009; Kopacz et al., 2010; Gonzi et al., 2011; Fortems-Cheiney et al., 2012), there is significant uncertainty in regional CO source estimates, reflecting varying source estimates from the inverse modeling analyses. As noted in previous studies, the discrepancies between the estimated CO emissions from different inversion analyses are due, in part, to errors in the atmospheric models used in the inversions. Model errors in long-range transport, vertical convective transport, diffusion, and chemistry (e.g. Arellano et al. 2006; Fortems-Cheiney et al.,

2011; Locatelli et al., 2013; Worden et al., 2013; Jiang et al., 2011, 2013, 2015) all adversely impact the inverse modeling of CO and other trace constituents (such as methane), and mitigating these errors in global models is challenging.

One way to reduce the effects of some model errors is to carry out the model simulations at high spatial resolution, which allows an improved description of small-scale processes, particularly those associated with vertical convection and diffusion. There have been several studies using high-resolution mesoscale models for inversion analyses (e.g. Stroud et al. 2011; Valin et al. 2011; Klich and Fuelberg 2014; Stock et al. 2014) with the lateral boundary conditions provided from global, coarse resolution models (e. g. Curci et al., 2010; Peylin et al., 2011). However, the consistency of boundary conditions becomes a critical issue in these regional analyses (e.g. Gockede et al., 2010). The boundary conditions have also been imposed based on independent data, such as aircraft in-situ measurements (e.g. Brioude et al., 2012; Lauvaux et al. 2012; Wecht et al. 2014).

Regional inverse modeling of CO emissions with adequate boundary condition optimization will also reduce the impact of discrepancies in long-range transport and in the chemical sink of CO. Reducing the sensitivity to the chemical sink of CO also requires that transport across the regional domain is fast compared to the lifetime of CO. Jiang et al. (2015) compared CO source estimates inferred from inversion analyses of surface level and profile retrievals of CO from the MOPITT (Measurement of Pollution in The Troposphere) satellite instrument and found that they were generally consistent (to within 10%), except for source estimates for North America, Europe, and East Asia. In an earlier study, Jiang et al. (2013) noted that when comparing source estimates inferred from in situ surface data and from satellite observations, “in the absence of transport bias, the surface and satellite data should provide

74 consistent constraints on the sources, if the data coverage is representative of the spatiotemporal
75 variability in CO". They were the first to show a large discrepancy in the Asian source estimates
76 obtained from the MOPITT surface level and profile retrievals, and they argued that it was due to
77 errors in convective transport over Asia associated with the Asian summer monsoon. In addition
78 to transport biases, discrepancies in the chemical sink of CO will also impact the constraints on
79 the surface sources provided by the surface level and free tropospheric data. Jiang et al. (2015)
80 suggested that the differences in the North American and European sources that they estimated
81 from the MOPITT surface level and profile retrievals could be due to the fact that air in the free
82 troposphere over North America and Europe is more chemically aged, thus, the surface level and
83 profile data are sampling air with different CO characteristics, with the profile data being more
84 susceptible to biases in the chemical sink.

85 The work presented here is based on the global analysis of Jiang et al. (2015), but
86 employs the high-resolution, regional version of GEOS-Chem (e.g. Wang et al., 2004; Chen et
87 al., 2009) and the MOPITT surface level retrievals to better quantify North American emissions
88 of CO. We focus on the period June 2004 to May 2005 for consistency with Jiang et al. (2015)
89 and Kopacz et al. (2010). As mentioned above, regional inversion analyses are sensitive to the
90 lateral boundary conditions, but use of global models to provide these boundary conditions is
91 problematic if there are biases in transport and the chemistry in the models. Use of in situ
92 observations to provide boundary conditions is also problematic because observational coverage
93 is often limited in space and time. A better approach for imposing the boundary conditions is to
94 assimilate satellite observations that can provide a strong constraint on the distribution of CO
95 throughout the free troposphere. Here we explore the use of the MOPITT data to constrain the
96 lateral boundary conditions as well as the surface CO emissions. We also examine the potential

impact of discrepancies in the abundance of the hydroxyl radical (OH), the main CO sink, on the estimates CO sources in a regional inverse modeling context.

This paper is organized as follows: in Section 2 we describe the MOPITT instruments and the GEOS-Chem model. In Section 3 we outline the inversion framework used in this work. In Section 4, we describe our approach for initial and boundary condition optimization, and present the estimated monthly mean North American emissions. The sensitivity of the source estimates to the chemical sink is examined by comparing the inversion results obtained with two different OH fields. Our conclusions are then provided in Section 5.

## 2. Observations and Model

### 2.1. MOPITT

The MOPITT instrument was launched on December 18, 1999, on NASA's Terra spacecraft. We employ the multispectral version 5 (V5J) retrievals, in which the thermal infrared (TIR) radiances at 4.7μm are combined the near infrared (NIR) radiances at 2.3μm to provide greater sensitivity to lower tropospheric CO over land (Worden et al., 2010; Deeter et al., 2011). The retrievals are conducted with respect to the logarithm of the volume mixing ratio (VMR), and are reported on a 10-level pressure grid (surface, 900, 800, 700, 600, 500, 400, 300, 200, and 100 hPa). Although we use only the surface level MOPITT retrievals in our analysis, it is necessary to transform the modeled CO profile to account for the vertical resolution of the MOPITT retrieval. This transformation is carried out using the following observation operator

$$F(\mathbf{x}) = \mathbf{y}_a + \mathbf{A}(H(\mathbf{x}) - \mathbf{y}_a) \quad (1)$$

where $\mathbf{A}$ is the MOPITT averaging kernel, $H(\mathbf{x})$ is the GEOS-Chem profile of CO (interpolated onto the MOPITT retrieval grid), and $\mathbf{y}_a$ is the MOPITT a priori profile. After transforming the

119 modeled profile, the modeled CO at the surface is compared to the surface level MOPITT CO, as
120 described in Eq (2) in Section 3. Deeter et al. (2012, 2013) evaluated the multispectral MOPITT
121 data and reported a small positive bias of 2.7% at the surface and a larger positive bias of 14% at
122 200 hPa for the V5J data. The large bias in the upper troposphere is not an issue here since we
123 focus on the surface level data. Further details for the MOPITT instrument and the multispectral
124 retrievals are given in Jiang et al. (2015).

125 **2.2. GEOS-Chem**

126 The GEOS-Chem global chemical transport model (CTM) (http://www.geos-chem.org) is
127 driven by assimilated meteorological fields from the NASA Goddard Earth Observing System
128 (GEOS-5) at the Global Modeling and data Assimilation Office. The standard GEOS-Chem
129 chemical mechanism includes 43 tracers, and simulates a detailed description of tropospheric $O_3$-
130 $NO_x$-hydrocarbon chemistry, including the radiative and heterogeneous effects of aerosols. The
131 native horizontal resolution of GEOS-5 is 0.5°x0.667°, but the meteorological fields are usually
132 degraded to 4°x5° or 2°x2.5° for global-scale simulations.

133 Our analysis is based on the CO-only simulation in GEOS-Chem v8-02-01, with relevant
134 updates through v9-01-01, using archived monthly OH fields from the full chemistry run. The
135 standard OH field used in this work is from GEOS-Chem version v5-07-08 (Evans et al. 2005).
136 In order to study the influence of the OH distribution on the inversion analyses, we also archive
137 the OH fields from a v8-02-01 GEOS-Chem full chemistry simulation. Additional details about
138 OH fields and emission inventories can be found in Jiang et al. (2015). Briefly, the annual North
139 America sources of CO are 134 Tg CO from fossil fuel and biofuel combustion and biomass
140 burning, 61 Tg CO from the oxidation of biogenic volatile organic compounds (VOCs), and 71
141 Tg CO from the oxidation of $CH_4$. Figure 1 shows the distribution of the annual mean CO

emissions for June 2004 to May 2005.

The inversion analyses here are carried out using the GEOS-Chem four-dimensional variational (4D-Var) data assimilation system, which was first described by Henze et al. (2007) and has been widely used in the chemical assimilation of CO and other tracer gases (e.g. Kopacz et al., 2009, 2010; Singh et al., 2011, Wells et al., 2014, Deng et al. 2014). Previous GEOS-Chem CO inversion analyses were conducted with the global version of the model. Here we extend the 4D-Var system to enable regional inverse modeling of CO using the nested version of GEOS-Chem.

The nested capability was first implemented in GEOS-Chem by Wang et al. (2004) for the GEOS-3 version of the meteorological fields. The model was then updated by Chen et al. (2009) to support the GEOS-5 meteorological fields, with 0.5°x0.667° resolution, which are used here. In the nested simulation the boundary conditions are based on fields archived from a global simulation (at 4°x5° or 2°x2.5°) with a 3-hour temporal resolution, which are used to rewrite the tracer concentrations in a buffer zone around the nested domain before every transport step. Along the boundary of the nested domain, the direction of the wind field is used to identify whether the flow is directed into or out of the domain, and the mixing ratios of the tracers in the buffer zone are used to provide the necessary upstream information. A key benefit of using the nested model was shown by Wang et al. (2004), who found that the CO mixing ratios in the high-resolution nested simulation were lower than in the coarse resolution global model, which they attributed to the failure of the coarse global model to capture subgrid vertical motions. Figure 2 shows the simulated CO mixing ratio on May 1, 2006, obtained with the 4°x5° global simulation and with the nested North America simulation. The yellow box in Figure 2 indicates the buffer zone in which the boundary conditions are applied. As shown in the figure, the high-

165 resolution CO distribution better reflects the influence of the mid-latitude cyclone present over
166 central North America, and the urban emission centers can be more clearly identified in the high-
167 resolution simulation.

## 168 3. Inversion Approach

169 The inverse method seeks an optimal estimate of the CO sources that is consistent with
170 both the observed atmospheric concentrations and the a priori constraints on the sources by
171 minimizing the cost function $J(\mathbf{x})$,

172 $$J(\mathbf{x}) = \sum_{i=1}^{N} (F_i(\mathbf{x}) - \mathbf{y}_i)^T \mathbf{S}_\Sigma^{-1} (F_i(\mathbf{x}) - \mathbf{y}_i) + (\mathbf{x} - \mathbf{x}_a)^T \mathbf{S}_a^{-1} (\mathbf{x} - \mathbf{x}_a) \quad (2)$$

173 where $\mathbf{x}$ is the state vector of emissions, $N$ is the total number of observations assimilated over
174 the assimilation window (which is one month), $\mathbf{y}_i$ is the $i$th vector of observed concentrations
175 (the MOPITT surface level retrievals), and $F(\mathbf{x})$ is the forward model, which accounts for the
176 vertical smoothing of the MOPITT retrieval and is described in Eq. 1. Here $\mathbf{x_a}$ is the a priori
177 estimate and $\mathbf{S}_\Sigma$ and $\mathbf{S_a}$ are the observational and a priori error covariance matrices, respectively.
178 The first term on the right in Eq 2 represents the mismatch between the simulated and observed
179 concentrations weighted by the observation error covariance. The second term represents the
180 departure of the estimate from the a priori. The cost function is iteratively minimized using the
181 L-BFGS algorithm (Liu and Nocedal, 1989). The inversion approach is exactly the same as
182 described in Jiang et al. (2015). We, therefore, refer the reader to Jiang et al. (2015) for details of
183 the optimizing scheme, the MOPITT data selection criteria, and the specification of the error
184 covariance matrices. We employed an Observing System Simulation Experiments (OSSE) to
185 evaluate our high-resolution 4D-Var system in the Appendix, which suggested the nested
186 inversion has similar reliability as the global scale assimilation system.

# 4. Results and Discussion

## 4.1. Optimization on the initial and boundary conditions

We produce initial conditions following the approach of Jiang et al. (2015), by assimilating MOPITT V5J tropospheric profile data using the sequential sub-optimal Kalman filter (Parrington et al. 2008) from 1 January 2004 to 1 May 2005. Because of the bias in the V5J data at 200 hPa, we assimilate the profile data only below 200 hPa. The optimized CO distribution from the Kalman filter is archived at the beginning of each month, providing the initial conditions at the beginning of each month for the 4D-Var inversion analyses.

As mentioned above, the lateral boundary conditions for the nested simulation could be specified from the global model. However, a better approach would be to constrain the global-scale and regional-scale emissions within the same inversion framework, so that the optimized emissions on the global-scale will provide less biased boundary conditions for the regional inversion. Such an approach has been used to constrain $CH_4$ and $N_2O$ emissions over South America and Europe (Meirink et al., 2008; Bergamaschi et al., 2010; Corazza et al., 2011) with the nested TM5 model. An issue with this approach is that the adjustment in the emissions on the global scale will have to be projected through long-range transport to the nested domain. If there are any biases in the model transport, those biases will also be projected onto the nested inversion.

Because the GEOS-Chem nesting is one-way, we cannot implement the same approach that is used in TM5. Instead, we conduct a global-scale inversion analysis and use the a posteriori CO fields as boundary conditions for the regional inversion. The a posteriori simulation from the global scale inversion should provide less biased boundary conditions for the regional scale inversion than the free running model (without assimilation). However, as mentioned above, the

boundary conditions could potentially be problematic if there are biases in the model transport. Alternatively, one could optimize the model CO distribution (using the sequential sub-optimal Kalman filter, for example), over writing any potential discrepancy in the tracer distribution associated with errors in the model transport (or chemistry). Here we compare the CO distribution obtained from optimizing the CO distribution using the Kalman filter and from optimizing the CO sources using the 4D-Var system.

The relative differences in the CO distribution obtained from these two assimilation approaches are shown in Figure 3. The assimilation of the MOPITT tropospheric profiles with the sequential sub-optimal Kalman filter from January 1 2004 to June 1 2005 is referred as CO_KF. The a posteriori CO distribution obtained from optimizing the monthly mean CO emissions using the 4D-Var scheme is referred as CO_EMS. Shown in Figure 3 are the relative differences of the lower tropospheric (surface – 500 hPa) partial columns, calculated as (CO_EMS – CO_KF) / CO_KF. Since both approaches used the same initial conditions, archived from the Kalman Filter assimilation, at the beginning of each month, the relative differences shown in Figure 3 can be considered as the residual bias in the a posteriori simulation that cannot be effectively removed within the one-month assimilation period by adjusting only the surface emissions. The most significant feature is the positive residual bias along the Intertropical Convergence Zone (ITCZ), suggesting errors in convective transport in the model (e.g., Jiang et al., 2013; Worden et al., 2013).

Since the objective of this work is to constrain the North American CO emissions, using CO_EMS for the boundary conditions may lead to biases in Mexico, the southern US, and along the North American west coast. Consequently, we decided to use the a posteriori fields from the Kalman filter as our optimized boundary conditions. The impact of the initial and boundary

conditions from the Kalman filter assimilation is shown in Figure 4. Driven with original initial and boundary conditions (without assimilation), the modeled CO columns (Figure 4b) are obviously lower than that of the MOPITT observations over the North America continent. The difference over the continent is much smaller when the initial and boundary conditions are optimized (Figure 4c) with the assimilation of the MOPITT profiles. It is clear that a significant bias will be introduced in the a posteriori regional emission estimates if the original initial and boundary conditions were used in the inversion analyses. As described in Jiang et al. (2015), we do not assimilate MOPITT data at high latitudes due to a potential positive bias in the CO retrievals at high latitudes; we only assimilate MOPITT data equator-ward of 40° over oceans and 52° over land, as shown in Figure 4.

The distribution of the relative differences between the modeled and observed CO fields is shown in Figure 5. With both the free running model and the optimized initial and boundary conditions, the distribution of the differences with respect to the MOPITT data are approximately Gaussian. The free running model has a low bias of -13.3%. Assimilating the MOPITT profile data to optimize the initial and boundary conditions reduced the mean bias to 3.5%, which should produce a better constraint on local North American emissions.

Although the mean a posteriori bias in the initial and boundary conditions is small, the largest residual bias in the boundary conditions is found on the southern boundary (near 10°N), where the a posteriori bias can be as large as 20% (between 80°W – 100°W). The vertical distribution of the relative bias along the southern boundary is shown in Figure 6. In lower troposphere, the original model simulation has a large positive bias (Figure 6a), approaching 50%, which the Kalman Filter assimilation significantly reduces (Figure 6b). The inability of the assimilation to more strongly reduce the bias is likely due to limited MOPITT observational

256 coverage over South America, associated with cloud cover (Keller et al., 2015).

257 A similar bias over the eastern Pacific (0°-12°N, 77.5°W – 122.5°W) was reported by Liu
258 et al. (2010) in their comparison of CO data from the Tropospheric Emission Spectrometer (TES)
259 and GEOS-Chem driven by GEOS-4 and GEOS-5 meteorological fields, for 2005 and 2006 (see
260 their Figures 3 and 4). Relative to TES, the GEOS-Chem was biased high at 681 hPa over the
261 region, which Liu et al. (2010) attributed to excessive export of South American biomass burning
262 emissions, which typically peak in August-September. The magnitude of the bias was larger with
263 GEOS-4 fields than with GEOS-5 (which are used here), and varied from year to year, reflecting
264 the variability in the biomass burning emissions; the model bias was larger in 2005 than in 2006.
265 In Figure 6 we see that relative to MOPITT, the model is biased low in the upper troposphere.
266 This reflects, in part, the high bias in the MOPITT V5J retrievals in the upper troposphere.
267 However, Liu et al. (2010) found that relative to CO data from the Microwave Limb Sounder
268 (MLS), the model (driven by GEOS-5) was also biased low over the eastern Pacific at 215 hPa in
269 August-September 2006, whereas it was biased high in 2005 (see their Figures 5 and 6). The
270 residual bias on the southern boundary will clearly impact the source estimates obtained here,
271 with the high bias in the lower troposphere resulting in an over adjustment (i.e., underestimation)
272 of the CO emission estimates in June – October in southern North America (Mexico and the
273 southern US). Since the TES retrievals are carried out in the presence of clouds, the TES data
274 may provide additional information on CO in the outflow region. Thus, assimilating TES
275 together MOPITT with the Kalman filter may help further reduce the bias the southern boundary
276 conditions. Another promising approach is the weak-constrained 4D-Var technique, recently
277 implemented in GEOS-Chem by Keller et al. (2015), in which the cost function Eq. (2) is
278 augmented with an addition term to mitigate the model transport errors.

279 **4.2. CO Source Estimates for June 2004 – May 2005**

280 Figure 7 shows the monthly scaling factors, which are the ratio of the a posteriori to a
281 priori emissions, for June 2004 – May 2005. The figure shows enhancement of anthropogenic
282 CO emissions in the Great Lakes region and along the US west coast. The annual total
283 anthropogenic emission for the contiguous US 48 states is increased by 14%, from 85 Tg to 97
284 Tg. This estimate is consistent with the results of our global inversion analysis presented in Jiang
285 et al. (2015). The annual total North America CO emission from the oxidation of biogenic VOCs
286 is reduced by 17%, from 61 Tg to 51 Tg with the largest reduction around the Gulf of Mexico in
287 July- September 2004. A possible reason for this reduction is the overestimation of isoprene
288 emission in the MEGANv2.0 inventory used in this work. As discussed in Jiang et al. (2015)
289 several previous studies have suggested that the MEGANv2.0 isoprene emissions are biased high
290 over North America. Hu et al. (2015) found that using the MEGANv2.1, together with an
291 improved land cover distribution, in GEOS-Chem successfully reproduced isoprene observations
292 in North America. However, the MEGANv2.1 inventory is not yet available in the GEOS-Chem
293 adjoint model. It is also possible that the reduction in the biogenic emissions, which are strongest
294 in the southern US, is due, in part, to the high bias in the southern boundary conditions.

295 The time series for the a priori and a posteriori estimates for different emission categories
296 are shown in Figure 8 for all of North America (15°N-65°N) and for the contiguous US 48 states.
297 For the whole continent, generally, the "bottom-up" inventory shows high CO emissions in
298 summer and lower values in winter. This seasonal variation is driven by the oxidation of
299 biogenic VOCs, which is significant in May-September and peaks in July, and by biomass
300 burning, which is at a maximum in April in Mexico and in August in boreal Canada. The
301 uncertainty of the anthropogenic emissions is assumed to be small. The analysis, however, does

suggest greater anthropogenic emissions in January-May 2005, which accounts for the larger total CO emissions for this period, as shown in Figure 8.

The United States is the largest CO source in North America, contributing 63% to the total North America source. The monthly a priori source in July is 16 Tg, twice as large as the source in winter. The distinct seasonal variation is driven by the strong biogenic VOCs source in summer. During January – April 2005, the total a posteriori CO source in this region is 59 Tg, 36% higher than the a priori value. On the contrary, during June – August 2004, the total a posteriori CO source in this region is 31 Tg, 29% lower than the a priori value. This significant discrepancy between summer and winter was also observed by Kopacz et al. (2010). The estimated winter emissions of Kopacz et al. (2010) are about 50% larger than the summer emissions. Kopacz et al. (2010) and Stein et al. (2014) attributed the higher wintertime emissions in the Northern Hemisphere to vehicular emissions, which are not account for in the a priori emission inventory.

The monthly total CO emissions for the contiguous US 48 states agrees well with the results from the global 4°x5° resolution inversion of Jiang et al. (2015). The largest difference is observed in Dec 2004, when the a posteriori emission estimate from the coarse-resolution inversion is 17% higher. The smallest difference is observed in Oct 2004, when the a posteriori emission estimate of the coarse-resolution inversion is 2% higher. As discussed in Jiang et al. (2015), the seasonal variations of the a posteriori source estimates obtained here are consistent with those of Kopcz et al. (2010), but the magnitude of the source estimates differ significantly, reflecting differences in the configuration of the inversions analayses. We refer the reader to Jiang et al. (2015) for a more detailed discussion of the differences between the source estimates obatined here and those from Kopcz et al. (2010).

Although there is good agreement at continental scales between our high-resolution

inverse analysis and the course-resolution inversion of Jiang et al. (2015), we do observe significant differences at regional scales. In June 2004, the high-resolution inversion shows large changes, along the west coast, whereas the spatial variation is much smaller in the coarse-resolution inversion, due to averaging on the model grid. More pronounced discrepancies are observed in Dec 2004. The coarse-resolution inversion shows the a priori CO emission estimates over Eastern US should be increased by 20-50%. However, the high-resolution inversion shows much more variation on regional or urban scales. For example, there is a significant CO emission decrease in the Toronto area, whereas there are large increases to the east of Lake Ontario. There is also significant CO emission reduction in West Virginia, and large increases in North and South Carolina. The greater spatial structure in the regional emission estimates is somewhat expected because of the higher spatial resolution, but it is unclear as to how reliable these features are, given the information content of the MOPITT data.

The Intercontinental Transport Experiment – North America, Phase A (INTEX-A) campaign was conducted during July 1 - August 15, 2004, over North America (Singh et al., 2006). A DC-8 aircraft was used to measure gas and aerosol abundances, including CO, $NO_2$, formaldehyde (HCHO), and $H_2O$, over an altitude range from 0.2 to 12.5 km. In this work, the aircraft measurements from the INTEX-A DC-8 aircraft in July 2004 are used to evaluate the inversion results obtained from MOPITT data. Figure 9 shows the difference between the GEOS-Chem simulation and the INTEX-A DC-8 aircraft observation in free troposphere. The inverse model significantly reduces the positive bias in the model bias relative to the aircraft measurements, from 7.2 ppb to 0.5 ppb, suggesting that the a posteriori CO does indeed provide a better regional fit to the independent aircraft data. The reduction in the bias relative to the aircraft data also suggests that vertical transport within North America is unbaised, since such a

transport bias in the inversion would degrade the agreement with aircraft data that was obtained with the optimized initial and boundary conditions (the a priori).

### 4.3. Sensitivity of Regional Source Estimates to OH

Following Jiang et al. (2015) we assess the sensitivity of the source estimates to the OH fields by repeating the inversion for June – August 2004 with the OH fields from version v8-02-01 of GEOS-Chem. We focus on just the summer months for comparison with Jiang et al. (2015). Also, the OH impact is expected to be greater in summer, when the CO lifetime is short. The initial and boundary conditions for this inversion were obtained by assimilating the MOPITT V5J profiles into the model with the Kalman filter and the v8-02-01 OH fields from 1 January 2004 to 1 September 2004. The inversion based on these initial and boundary conditions is referred to as v8OH_BCv8. Our standard inversion with initial and boundary conditions based on the v5-07-08 OH is referred to as v5OH_BCv5. As discussed in Jiang et al. (2015), the OH concentrations in v8-02-01 are significantly higher than the v5-07-08 version in the Northern Hemisphere, and consequently, the CO lifetime is about 30% shorter.

Figures 10a-c show the scaling factors for the v8OH_BCv8 inversion. The differences relative to the standard inversion (v5OH_BCv5) are shown in Figs 10d-f. The relative difference in the a posteriori CO emission estimates in the contiguous US 48 states inferred from the two OH fields is 32%, suggesting that the OH fields still have a significant impact on the a posteriori estimates, even with the optimized boundary conditions. The relative difference of 32% is 20% smaller than the relative difference (40%) obtained by Jiang et al. (2015) in their 4°x5° resolution global-scale inversion. Although the OH–related error is smaller than in the global, coarse-resolution inversion, it is still suprisingly large. Figures 10j-l show the relative differences between the boundary conditions obtained from the v5-07-08 and the v8-02-01 OH fields, which

were used in the v5OH_BCv5 and v8OH_BCv8 inversions, respectively. Under ideal conditions, the difference between two boundary conditions should be small because both were optimized with the same MOPITT data using the same approach. However, we do see large differences of more than 15% along the northern and northeastern boundaries, suggesting our optimization of the boundary conditions is inadequate. We believe that the main reason for this is that we neglected MOPITT data at high latitudes (see Figure 4) to avoid a potential positive bias in the data. Assimilating data from multi-instruments, such as was done by Kopacz et al (2010), could provide better data coverage at high-latitutdes, and thus a better constraint on the northern boundary conditions. Moving the northern boundary to lower latitude would be also helpful.

To quantify the contribution of the differences in the boundary conditions to the differences in the source estimates, we repeated the inversion using the v8-02-01 OH fields, but with the initial and boundary conditions obtained with v5-07-08 OH. This inversion is referred to as v8OH_BCv5. Since the initial and boundary condtions in the v8OH_BCv5 and v5OH_BCv5 inversions are identical, the differences in the source estimates obtained from these will reflect only the influence of the OH differences over North America. Figures 10g-i show the differences between v8OH_BCv5 and v5OH_BCv5 (our standard inversion). The relative difference in the a posteriori CO emission estimates for the contiguous US 48 states is only 20%, which is 50% smaller than the relative difference obtained by Jiang et al. (2015) in their 4°x5° global inversion. The large reduction in the impact of OH on the source estimates compared to the global-scale inversion is encouraging, and demonstrates the potential advantages of high-resolution regional inversion analyses.

## 5. Summary

High-resolution CTM simulations have obvious accuracy advantages over coarse

resolution simulations, particularly for small-scale processes (e.g., Stroud et al. 2011; Klich and Fuelberg 2014; Stock et al. 2014). In this work, we used the adjoint of the nested GEOS-Chem model, at a resolution of 0.5° x 0.67°, to constrain North American CO emissions during the period of June 2004 – May 2005. To reduce the potential impact of discrepancies in the modeled OH field on the source estimates, we used the V5J surface level MOPITT retrievals. Our results show that the annual total anthropogenic CO emissions for the contiguous US 48 states should be increased by 14%, from 85 Tg to 97 Tg for 2004 – 2005. The adjustment was mainly caused by an increase in emissions near the Great Lakes and along the west coast. The inversion analysis also suggested that the total CO emissions should be increased by 36% during January – April 2005, and decreased by 29% during June – August 2004. This seasonal variation was also observed by Kopacz et al. (2010), and could be associated with an underestimation of road traffic emission in winter (Stein et al. 2014). The inversion results were evaluated with in-situ measurements from the DC-8 aircraft during the INTEX-A campaign in July 2004. The mean bias between the model and the aircraft data in the free troposphere was reduced from 7.2 ppb with the a priori emissions to 0.5 ppb with the a posteriori emissions.

Reliable initial and boundary conditions are critical for regional inversion analyses. We used a sequential sub-optimal Kalman filter (Parrington et al. 2008) to assimilate MOPITT CO profiles to optimize the distribution of CO (rather than the emissions) to produce improved initial and lateral boundary conditions for the regional inversion analyses. Because of the restricted domain of the regional analyses, the optimized boundary conditions should significantly reduce the sensitivity of the estimated CO sources to errors in long-range transport and in the OH distribution. We found that the Kalman filter assimilation significantly improved the initial and lateral boundary conditions, reducing the bias from -13.3% to 3.5%. However, there was a large

417 residual bias in the southern boundary (near 10°N, between 80°W – 100°W), in the outflow
418 region for biomass burning emissions from South America, which could result in an over
419 adjustment of the CO emissions in Mexico and the southern US. Comparison of the inversion
420 results driven with two different OH fields, from version v5-07-08 and v8-02-01 of GEOS-Chem,
421 produced relative differences in the North American source estimates of 32% for June – August
422 2004. This OH-related difference in the source estimates is about 20% smaller than the
423 differences obtained by Jiang et al. (2015) for the same period in their global-scale inversion
424 analysis.

425 Examination of the differences in the boundary conditions based on the two OH fields
426 showed large relative differences (greater than 15%) in the northern and northeastern boundaries,
427 suggesting that our optimization of the boundary conditions was inadequate. In our assimilation
428 of the MOPITT data we neglected data pole-ward of 52° and 40° over land and oceans,
429 respectively, to avoid the influence of a potential high-latitude bias in the data, and we believe
430 that this accounted for the weaker constraint on the northern boundary conditions in the analysis.
431 To assess the influence of the boundary conditions we repeated the inversion with the v8-02-01
432 OH fields using the same initial and boundary conditions from the standard inversion using the
433 v5-07-08 OH. In this case, we estimated a relative difference between the source estimates based
434 on the v8-02-01 and v5-07-08 OH fields of 20%, which is 50% smaller than that reported by
435 Jiang et al. (2015). Thus, our best estimate for North American CO emissions for 2004 – 2005 is
436 97 Tg, with a potential error of 20%, associated with discrepancies in local North America OH.

437 Our results demonstrate that high-resolution, regional inversion analyses can reduce the
438 sensitivity of the inferred CO source estimates to errors in long-range transport and in the OH
439 distributions. However, the 20% OH-related discrepancy that we estimated is still large, and

could indicate that more stringent constraints on the regional boundary conditions are needed. This may be achieved by integrating data from multiple sources. The OH-related discrepancies could also reflect that fact that in summer, air in the middle and upper troposphere over North America is trapped by a semipermanent anticyclone, which allows greater chemical aging than direct lateral export from the continent (Li et al., 2005; Cooper et al., 2007). Although the MOPITT surface level retrievals have peak sensitivity to CO near the boundary layer, their sensitivity extend up to the middle troposphere (see Figure 1a of Jiang et al. 2013). Consequently, the inversion analyses could be sensitivite to chemical aging of air in the North American anticyclone. Work is needed to determine the residence time for air in the anticyclone compared to the spatio-temporal variability of the constraints on the North American source estimates provided by the MOPITT surface level retrievals. Improving the source estimates will likely require assimilating sufficient information to obtain a strong constraint on the CO distribution on a timescales shorter than the timescale for chemical aging in the domain. Despite these limitations, we believe that our results show the potential advantages of combining high-resolution regional inversion analyses with global analyses to better quantify regional CO source estimates.

## Appendix: Observing System Simulation Experiments (OSSE)

The reliability of the nested inversion is examined with an OSSE for the period June 1-15, 2004. In the OSSE, we firstly create pseudo-observations, by archiving model output with CO emission unchanged. In the pseudo-inversion, we reduced the CO emission by 50% and the objective of the OSSE is to observe whether the scaling factors can return to true state (1.0). Figure A1a shows the result of the reference global scale inversion with 4°x5° resolution. The a posteriori estimation converges to the true state in all major emission regions. In the nested

463 inversion (Figure A1c), the model converges to the true state in Eastern US, whereas the result is
464 noisy in Western US and Canada, which is consistent with the global scale inversion, as shown
465 in Figure A1b. The OSSE demonstrates the nested inversion has similar reliability as the global
466 scale assimilation system.

467 ## Acknowledgments.

468 This work was supported by funding from the Natural Science and Engineering Research
469 Council of Canada, the Canadian Space Agency, and NASA grants NNX10AT42G and
470 NNX11AI54G.

471 ## References

472 Arellano, A. F. Jr., Kasibhatla, P. S., Giglio, L., van der Werf, G. R., Randerson, J. T., and
473 Collatz, G. J.: Time dependent inversion estimates of global biomass-burning CO emissions
474 using Measurement of Pollution in the Troposphere (MOPITT) measurements, J. Geophys.
475 Res., 111, D09303, doi:10.1029/2005JD006613, 2006.
476 Bergamaschi, P., Krol, M., Meirink, J. F., Dentener, F., Segers, A., van Aardenne, J., Monni, S.,
477 Vermeulen, A. T., Schmidt, M., Ramonet, M., C., Yver, Meinhardt, F., Nisbet, E. G., R. E.,
478 Fisher, O'Doherty, S. and Dlugokencky, E. J.: Inverse modeling of European $CH_4$ emissions
479 2001–2006, J. Geophys. Res., 115, D22309, doi:10.1029/2010JD014180, 2010.
480 Brioude, J., Petron, G., Frost, G. J., Ahmadov, R., Angevine, W. M., Hsie, E.-Y., Kim, S.-W.,
481 Lee, S.-H., McKeen, S. A., Trainer, M., Fehsenfeld, F. C., Holloway, J. S., Peischl, J., Ryerson,
482 T. B. and Gurney, K. R.: A new inversion method to calculate emission inventories without a
483 prior at mesoscale: Application to the anthropogenic $CO_2$ emission from Houston, Texas, J.
484 Geophys. Res., 117, D05312, doi:10.1029/2011JD016918, 2012.
485 Chen, D., Wang, Y., McElroy, M. B., He, K., Yantosca, R. M., and Le Sager, P.: Regional CO

pollution and export in China simulated by the high-resolution nested-grid GEOS-Chem model, Atmos. Chem. Phys., 9, 3825-3839, doi:10.5194/acp-9-3825-2009, 2009.

Cooper, O. R., et al., Evidence for a recurring eastern North America upper tropospheric ozone maximum during summer, J. Geophys. Res., 112, D23304, doi:10.1029/2007JD008710, 2007.

Corazza, M., Bergamaschi, P., Vermeulen, A. T., Aalto, T., Haszpra, L., Meinhardt, F., O'Doherty, S., Thompson, R., Moncrieff, J., Popa, E., Steinbacher, M., Jordan, A., Dlugokencky, E., Brühl, C., Krol, M., and Dentener, F.: Inverse modelling of European $N_2O$ emissions: assimilating observations from different networks, Atmos. Chem. Phys., 11, 2381-2398, doi:10.5194/acp-11-2381-2011, 2011..

Curci, G., Palmer, P. I., Kurosu, T. P., Chance, K., and Visconti, G.: Estimating European volatile organic compound emissions using satellite observations of formaldehyde from the Ozone Monitoring Instrument, Atmos. Chem. Phys., 10, 11501-11517, doi:10.5194/acp-10-11501-2010, 2010.

Deeter, M. N., Worden, H. M., Gille, J. C., Edwards, D. P., Mao, D., and Drummond, J. R.: MOPITT multispectral CO retrievals: Origins and effects of geophysical radiance errors, J. Geophys. Res., 116, D15303, doi:10.1029/2011JD015703, 2011.

Deeter, M. N., Worden, H. M., Edwards, D. P., Gille, J. C., and Andrews, A. E.: Evaluation of MOPITT retrievals of lower-tropospheric carbon monoxide over the United States, J. Geophys. Res., 117, D13306, doi:10.1029/2012JD017553, 2012.

Deeter, M. N., Martínez-Alonso, S., Edwards, D. P., Emmons, L. K., Gille, J. C., Worden, H. M., Pittman, J. V., Daube, B. C., and Wofsy, S. C.: Validation of MOPITT Version 5 thermal-infrared, near-infrared, and multispectral carbon monoxide profile retrievals for 2000–2011, J. Geophys. Res. Atmos., 118, 6710–6725, doi:10.1002/jgrd.50272, 2013.

509 Deng, F., et al., Inferring regional sources and sinks of atmospheric $CO_2$ from GOSAT $XCO_2$
510 data, Atmos. Chem. Phys., 14, 3703-3727, doi:10.5194/acp-14-3703-2014, 2014.
511 Evans, M. J., and Jacob, D. J.: Impact of new laboratory studies of $N_2O_5$ hydrolysis on global
512 model budgets of tropospheric nitrogen oxides, ozone, and OH, Geophys. Res. Lett., 32,
513 L09813, doi:10.1029/2005GL022469, 2005.
514 Fortems-Cheiney, A., Chevallier, F., Pison, I., Bousquet, P., Szopa, S., Deeter, M. N., and
515 Clerbaux, C.: Ten years of CO emissions as seen from Measurements of Pollution in the
516 Troposphere (MOPITT), J. Geophys. Res., 116, D05304, doi:10.1029/2010JD014416, 2011.
517 Fortems-Cheiney, A., Chevallier, F., Pison, I., Bousquet, P., Saunois, M., Szopa, S., Cressot, C.,
518 Kurosu, T. P., Chance, K., and Fried, A.: The formaldehyde budget as seen by a global-scale
519 multi-constraint and multi-species inversion system, Atmos. Chem. Phys., 12, 6699-6721,
520 doi:10.5194/acp-12-6699-2012, 2012.
521 Göckede, M., Turner, D. P., Michalak, A. M., Vickers, D., and Law, B. E.: Sensitivity of a
522 subregional scale atmospheric inverse CO2 modeling framework to boundary conditions, J.
523 Geophys. Res., 115, D24112, doi:10.1029/2010JD014443, 2010.
524 Gonzi, S., Feng, L., and Palmer, P. I.: Seasonal cycle of emissions of CO inferred from MOPITT
525 profiles of CO: Sensitivity to pyroconvection and profile retrieval assumptions, Geophys. Res.
526 Lett., 38, L08813, doi:10.1029/2011GL046789, 2011.
527 Heald, C. L., Jacob, D. J., Jones, D. B. A., Palmer, P. I., Logan ,J. A., Streets, D. G., Sachse, G.
528 W., Gille, J. C., Hoffman, R. N., and Nehrkorn, T.: Comparative inverse analysis of satellite
529 (MOPITT) and aircraft (TRACE-P) observations to estimate Asian sources of carbon monoxide,
530 J. Geophys. Res. 109, D23306, doi:10.1029/2004JD005185, 2004.
531 Henze, D. K., Hakami, A., and Seinfeld, J. H.: Development of the adjoint of GEOS- Chem,

Atmos. Chem. Phys.,7, 2413-2433, doi:10.5194/acp-7-2413-2007, 2007.

Jiang, Z., D. Jones, B. A., Kopacz, M., Liu, J., Henze, D. K., and Heald, C., Quantifying the impact of model errors on top-down estimates of carbon monoxide emissions using satellite observations, J. Geophys. Res., 116, D15306, doi:10.1029/2010JD015282, 2011.

Jiang, Z., Jones, D. B. A., Worden, H. M., Deeter, M. N., Henze, D. K., Worden, J., Bowman, K. W., Brenninkmeijer, C. A. M., and Schuck, T. J.: Impact of model errors in convective transport on CO source estimates inferred from MOPITT CO retrievals, J. Geophys. Res. Atmos., 118, 2073–2083, 2013.

Jiang, Z., Jones, D. B. A., Worden, H. M., and Henze, D. K.: Sensitivity of top-down CO source estimates to the modeled vertical structure in atmospheric CO, Atmos. Chem. Phys., 15, 1521-1537, doi:10.5194/acp-15-1521-2015, 2015.

Jones, D. B. A., Bowman, K. W., Logan, J. A., Heald, C. L., Liu, J., Luo, M., Worden, J., and Drummond, J.: The zonal structure of tropical $O_3$ and CO as observed by the Tropospheric Emission Spectrometer in November 2004 – Part 1: Inverse modeling of CO emissions, Atmos. Chem. Phys., 9, 3547 – 3562, 2009.

Keller, M., Jones, D. B. A., Jiang, Z., Henze, D. K., Worden, H. M.: Quantifying Model Biases in CO Emission Estimation Using Weak Constraint 4D-Var, submitted to J. Geophys. Res., 2015.

Klich, C. A. and Fuelberg, H. E.: The role of horizontal model resolution in assessing the transport of CO in a middle latitude cyclone using WRF-Chem, Atmos. Chem. Phys., 14, 609-627, doi:10.5194/acp-14-609-2014, 2014.

Kopacz, M., Jacob, D., Henze, D., Heald, C., Streets, D. and Zhang, Q.: Comparison of adjoint and analytical Bayesian inversion methods for constraining Asian sources of carbon monoxide

using satellite (MOPITT) measurements of CO columns, J. Geophys. Res., 114(D4), doi:10.1029/2007JD009264, 2009.

Kopacz, M., Jacob, D. J., Fisher, J. A., Logan, J. A., Zhang, L., Megretskaia, I. A., Yantosca, R. M., Singh, K., Henze, D. K., Burrows, J. P., Buchwitz, M., Khlystova, I., McMillan, W. W., Gille, J. C., Edwards, D. P., Eldering, A., Thouret, V. and Nedelec, P.: Global estimates of CO sources with high resolution by adjoint inversion of multiple satellite datasets (MOPITT, AIRS, SCIAMACHY, TES), Atmos. Chem. Phys.,10, 855-876, doi:10.5194/acp-10-855-2010, 2010.

Lauvaux, T., Schuh, A. E., Uliasz, M., Richardson, S., Miles, N., Andrews, A. E., Sweeney, C., Diaz, L. I., Martins, D., Shepson, P. B., and Davis, K. J.: Constraining the $CO_2$ budget of the corn belt: exploring uncertainties from the assumptions in a mesoscale inverse system, Atmos. Chem. Phys., 12, 337-354, doi:10.5194/acp-12-337-2012, 2012.

Liu, D. C. and J. Nocedal, On the Limited Memory Method for Large Scale Optimization, Mathematical Programming, 45, 503-528, 1989.

Liu, J.-H., Logan, J. A., Jones, D. B. A., Livesey, N. J., Megretskaia, I., Carouge, C. and Nedelec, P.: Analysis of CO in the tropical troposphere using Aura satellite data and the GEOS-Chem model: insights into transport characteristics of the GEOS meteorological products, Atmos. Chem. Phys., 10, 12207-12232, doi:10.5194/acp-10-12207-2010, 2010.

Locatelli, R., Bousquet, P., Chevallier, F., Fortems-Cheney, A., Szopa, S., Saunois, M., Agusti-Panareda, A., Bergmann, D., Bian, H., Cameron-Smith, P., Chipperfield, M. P., Gloor, E., Houweling, S., Kawa, S. R., Krol, M., Patra, P. K., Prinn, R. G., Rigby, M., Saito, R., and Wilson, C.: Impact of transport model errors on the global and regional methane emissions estimated by inverse modelling, Atmos. Chem. Phys., 13, 9917-9937, doi:10.5194/acp-13-9917-2013, 2013.

578 Meirink, J. F., Bergamaschi, P., and Krol, M. C.: Four-dimensional variational data assimilation
579 for inverse modelling of atmospheric methane emissions: method and comparison with
580 synthesis inversion, Atmos. Chem. Phys., 8, 6341-6353, doi:10.5194/acp-8-6341-2008, 2008.
581 Palmer, P., Jacob, D., Jones, D., Heald, C., Yantosca, R., Logan, J., Sachse, G. and Streets, D.:
582 Inverting for emissions of carbon monoxide from Asia using aircraft observations over the
583 western Pacific, J. Geophys. Res., 108(D21), doi:10.1029/2003JD003397, 2003.
584 Parrington, M., Jones, D. B. A., Bowman, K. W., Horowitz, L. W., Thompson, A. M., Tarasick,
585 D. W., and Witte, J. C.: Estimating the summertime tropospheric ozone distribution over North
586 America through assimilation of observations from the Tropospheric Emission Spectrometer, J.
587 Geophys. Res., 113, D18307, doi:10.1029/2007JD009341, 2008.
588 Parrington, M., Palmer, P. I., Henze, D. K., Tarasick, D. W., Hyer, E. J., Owen, R. C., Helmig,
589 D., Clerbaux, C., Bowman, K. W., Deeter, M. N., Barratt, E. M., Coheur, P.-F., Hurtmans, D.,
590 Jiang, Z., George, M. and Worden, J. R.: The influence of boreal biomass burning emissions on
591 the distribution of tropospheric ozone over North America and the North Atlantic during
592 2010, Atmos. Chem. Phys., 12, 2077-2098, doi:10.5194/acp-12-2077-2012, 2012.
593 Pétron, G., Granier, C., Khattatov, B., Yudin, V., Lamarque, J.-F., Emmons, L., Gille, J., and
594 Edwards, D. P.: Monthly CO surface sources inventory based on the 2000– 2001 MOPITT
595 satellite data, Geophys. Res. Lett., 31, L21107, doi:10.1029/2004GL020560, 2004.
596 Peylin, P., Houweling, S., Krol, M. C., Karstens, U., Rödenbeck, C., Geels, C., Vermeulen, A.,
597 Badawy, B., Aulagnier, C., Pregger, T., Delage, F., Pieterse, G., Ciais, P., and Heimann, M.:
598 Importance of fossil fuel emission uncertainties over Europe for $CO_2$ modeling: model
599 intercomparison, Atmos. Chem. Phys., 11, 6607-6622, doi:10.5194/acp-11-6607-2011, 2011.
600 Singh, H. B., Brune, W. H., Crawford, J. H., Jacob, D. J., and Russell, P. B.: Overview of the

summer 2004 Intercontinental Chemical Transport Experiment– North America (INTEX-A), J. Geophys. Res., 111, D24S01, doi:10.1029/2006JD007905, 2006.

Singh, K., Jardak, M., Sandu, A., Bowman, K., Lee, M., and Jones, D.: Construction of non-diagonal background error covariance matrices for global chemical data assimilation, Geosci. Model Dev., 4, 299-316, doi:10.5194/gmd-4-299-2011, 2011.

Stein, O., Schultz, M. G., Bouarar, I., Clark, H., Huijnen, V., Gaudel, A., George, M., and Clerbaux, C.: On the wintertime low bias of Northern Hemisphere carbon monoxide found in global model simulations, Atmos. Chem. Phys., 14, 9295-9316, doi:10.5194/acp-14-9295-2014, 2014.

Stock, Z. S., Russo, M. R., and Pyle, J. A.: Representing ozone extremes in European megacities: the importance of resolution in a global chemistry climate model, Atmos. Chem. Phys., 14, 3899-3912, doi:10.5194/acp-14-3899-2014, 2014.

Stroud, C. A., Makar, P. A., Moran, M. D., Gong, W., Gong, S., Zhang, J., Hayden, K., Mihele, C., Brook, J. R., Abbatt, J. P. D., and Slowik, J. G.: Impact of model grid spacing on regional- and urban- scale air quality predictions of organic aerosol, Atmos. Chem. Phys., 11, 3107-3118, doi:10.5194/acp-11-3107-2011, 2011.

Valin, L. C., Russell, A. R., Hudman, R. C., and Cohen, R. C.: Effects of model resolution on the interpretation of satellite $NO_2$ observations, Atmos. Chem. Phys., 11, 11647-11655, doi:10.5194/acp-11-11647-2011, 2011.

Wang, Y. X., McElroy, M. B., Jacob, D. J., and Yantosca, R. M.: A nested grid formulation for chemical transport over Asia: Applications to CO, J. Geophys. Res., 109, D22307, doi:10.1029/2004JD005237, 2004.

Wecht, K. J., Jacob, D. J., Sulprizio, M. P., Santoni, G. W., Wofsy, S. C., Parker, R., Bösch, H.,

and Worden, J.: Spatially resolving methane emissions in California: constraints from the CalNex aircraft campaign and from present (GOSAT, TES) and future (TROPOMI, geostationary) satellite observations, Atmos. Chem. Phys., 14, 8173-8184, doi:10.5194/acp-14-8173-2014, 2014.

Wells, K.C., et al., Quantifying global terrestrial methanol emissions using observations from the TES satellite sensor, Atmos. Chem. Phys., 14, 2555–2570, doi:10.5194/acp-14-2555-2014, 2014.

Worden, H. M., Deeter, M. N., Edwards, D. P., Gille, J. C., Drummond, J. R., and Nédélec, P.: Observations of near-surface carbon monoxide from space using MOPITT multispectral retrievals, J. Geophys. Res., 115, D18314, doi:10.1029/2010JD014242, 2010.

Worden, J., et al., El Niño, the 2006 Indonesian peat fires, and the distribution of atmospheric methane,Geophys. Res. Lett., 40, 4938–4943, doi:10.1002/grl.50937, 2013.

## Tables and Figures

**Figure 1**. Annual mean CO emissions from fossil fuel (FF), biofuel (BF), biomass burning (BB) and the oxidation of biogenic NMVOC and $CH_4$, averaged for June 2004 to May 2005. The unit is $10^{12}$ molec/cm$^2$/sec.

**Figure 2**. CO mixing ratio in GEOS-Chem on level 10 (about 850 hPa) on May 1, 2006. The influence of a mid-latitude cyclone is clearly shown in the high-resolution (0.5°x0.667°) simulation (right), whereas it is not obvious in that with coarse resolution (4°x5°) simulation (left). The light yellow line on the 0.5°x0.667° plot demarcates the buffer zone in which the coarse resolution boundary conditions are imposed.

**Figure 3**. Relative difference between the optimized lower tropospheric partial columns (surface – 500 hPa) between Kalman Filter assimilation, referred as CO_KF and a posteriori simulation of global scale inversion, referred as CO_EMS. The value is calculated as (CO_EMS – CO_KF) / CO_KF.

**Figure 4.** Mean tropospheric CO columns ($10^{18}$ molec/cm$^2$) in the GEOS-Chem North America nested domain in June 2004 from (a) MOPITT version 5; (b) GEOS-Chem model with the original initial and boundary conditions; (c) with the optimized initial and boundary conditions.

Note that MOPITT data poleward of 40° and 52° over oceans and land, respectively, are not used in this work to reduce the influence of potential positive bias in MOPITT CO retrievals, as described in Jiang et al. (2015).

**Figure 5**. Distribution of the relative bias between the model and MOPITT with the oiriginal initial and lateral boundary conditions (red) and the the optimized initial and lateral boundary conditions (blue), after assimilation of the MOPITT data using the Kalman filter. The numbers are the mean relative difference.

**Figure 6**. Vertical distribution of relative difference in June 2004 along the southern boundary, between model and MOPITT, calculated as (Mod - MOP) / MOP. (a) original model simulation. (b) optimized model simulation by assimilating MOPITT data using Kalman Filter.

**Figure 7**. Monthly scaling factor for total CO emissions (from combustion sources and the oxidation of biogenic NMVOC) during June 2004 – May 2005.

**Figure 8**. Monthly CO emissions during June 2004 – May 2005 for different emission categories: total emissions (black), anthropogenic emissions (blue), biomass burning (red), and the oxidation of biogenic NMVOCs (green). The a priori estimates are shown with the solid line and a posteriori values are indicated with the dashed line. The unit is Tg/month.

**Figure 9**. Difference between the GEOS-Chem simulation with INTEX-A DC-8 aircraft observation in free troposphere in July 2004. (a) A priori model simulation (based on the optimized initial and boundary conditions). (b) A posteriori model simulation. The model is sampled at the aircraft measurements time, location and alitude.

**Figure 10.** (a–c): Scaling factors of v8OH_BCv8 inversion, based on the v8-02-01 OH; (d–f): Difference between scaling factors of the v5OH_BCv5 and v8OH_BCv8 inversions; (g–i): Similar to Panel (d-f), but for the v8OH_BCv5 inversion, with the initial and boundary conditions from the standard inversion with the v5-07-08 OH fields; (j-l) Relative difference between the lower tropospheric CO partial columns (surface – 500 hPa) of the boundary conditions for the v5OH_BCv5 (v5) and the v8OH_BCv8 (v8) inversions, calculated as: 2 * (v8 - v5) / (v8 + v5).

**Figure A1**. OSSE scaling factors for June 1-15, 2004. (a) Global reference inversion with 4°x5° resolution. (b) Zoomed in North America region of the global inversion. (c) Nested inversion results. The scaling factor for the first guess is 0.5 and for the true state is 1.0.

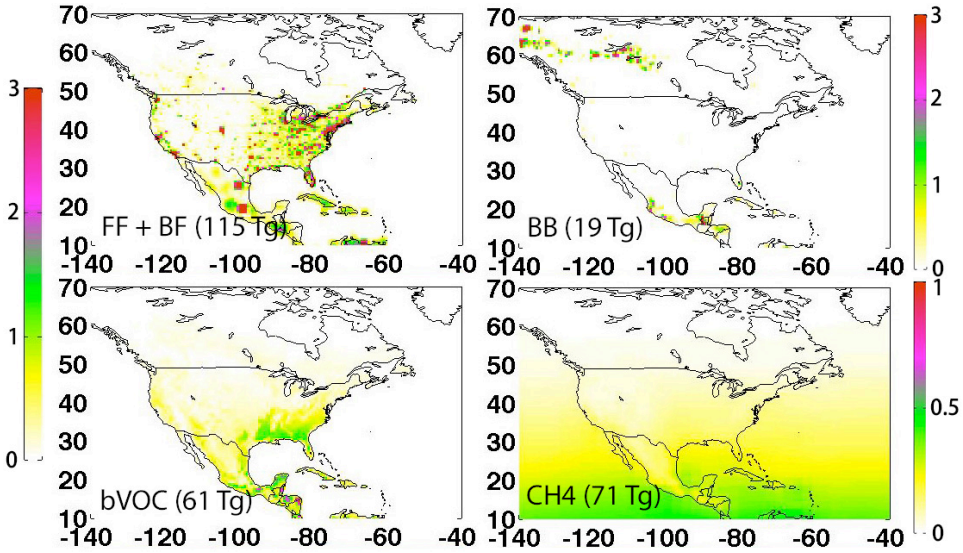


**Figure 1**. Annual mean CO emissions from fossil fuel (FF), biofuel (BF), biomass burning (BB) and the oxidation of biogenic NMVOC and $CH_4$, averaged for June 2004 to May 2005. The unit is $10^{12}$ molec/cm$^2$/sec.

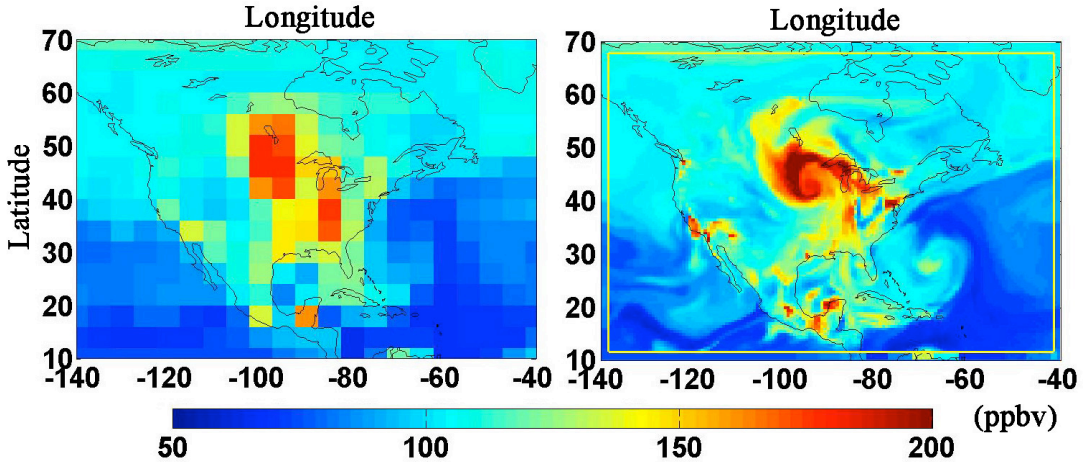


**Figure 2**. CO mixing ratio in GEOS-Chem on level 10 (about 850 hPa) on May 1, 2006. The influence of a mid-latitude cyclone is clearly shown in the high-resolution (0.5°x0.667°) simulation (right), whereas it is not obvious in that with coarse resolution (4°x5°) simulation (left). The light yellow line on the 0.5°x0.667° plot demarcates the buffer zone in which the coarse resolution boundary conditions are imposed.

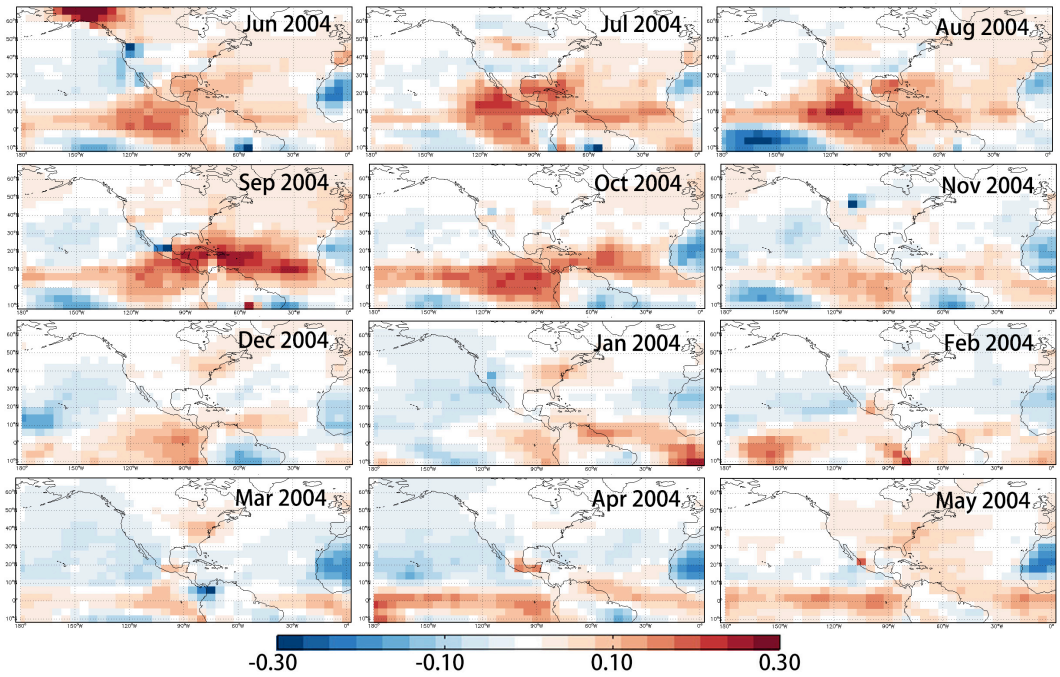


**Figure 3**. Relative difference between the optimized lower tropospheric partial columns (surface – 500 hPa) between Kalman Filter assimilation, referred as CO_KF and a posteriori simulation of global scale inversion, referred as CO_EMS. The value is calculated as (CO_EMS – CO_KF) / CO_KF.

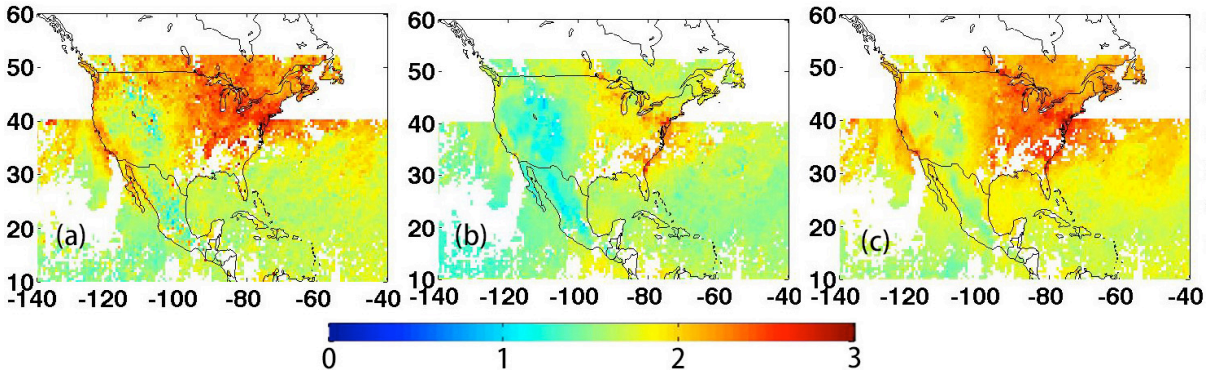


**Figure 4.** Mean tropospheric CO columns ($10^{18}$ molec/cm$^2$) in the GEOS-Chem North America nested domain in June 2004 from (a) MOPITT version 5; (b) GEOS-Chem model with the original initial and boundary conditions; (c) with the optimized initial and boundary conditions. Note that MOPITT data poleward of 40° and 52° over oceans and land, respectively, are not used in this work to reduce the influence of potential positive bias in MOPITT CO retrievals, as described in Jiang et al. (2015).

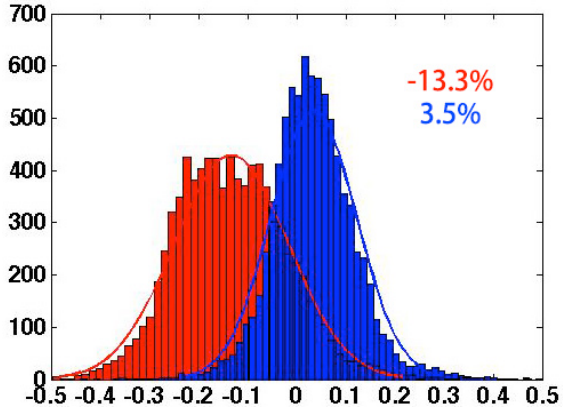


**Figure 5**. Distribution of the relative bias between the model and MOPITT with the oiriginal initial and lateral boundary conditions (red) and the the optimized initial and lateral boundary conditions (blue), after assimilation of the MOPITT data using the Kalman filter. The numbers are the mean relative difference.

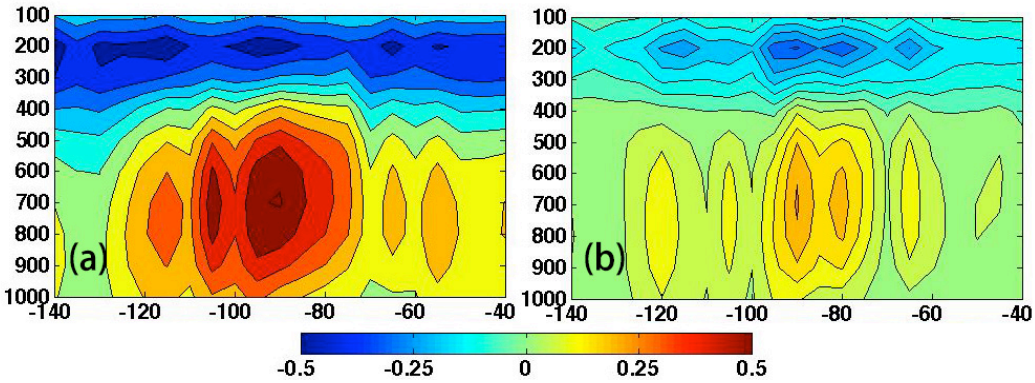


**Figure 6**. Vertical distribution of relative difference in June 2004 along the southern boundary, between model and MOPITT, calculated as (Mod - MOP) / MOP. (a) original model simulation. (b) optimized model simulation by assimilating MOPITT data using Kalman Filter.

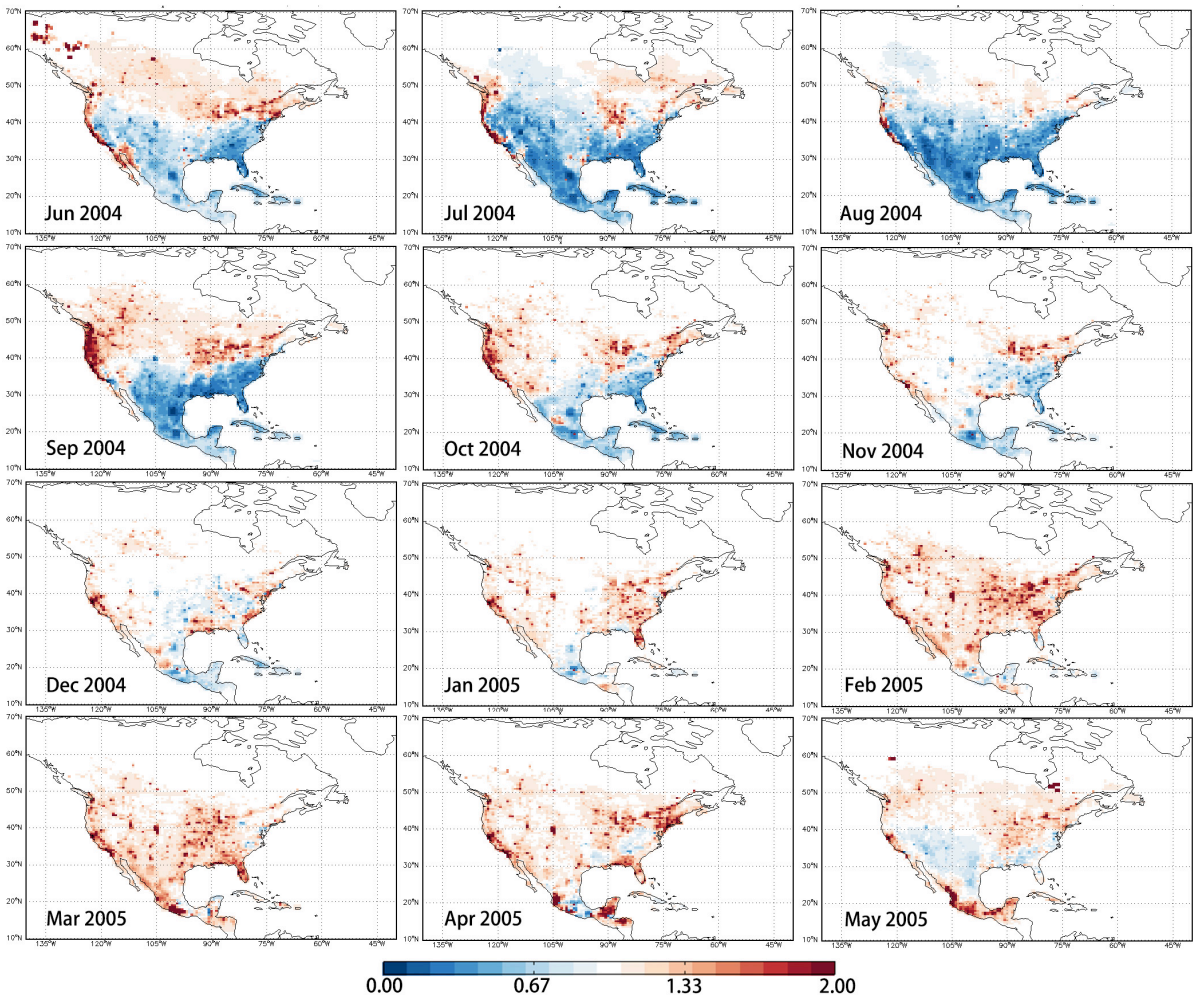


**Figure 7**. Monthly scaling factor for total CO emissions (from combustion sources and the oxidation of biogenic NMVOC) during June 2004 – May 2005.

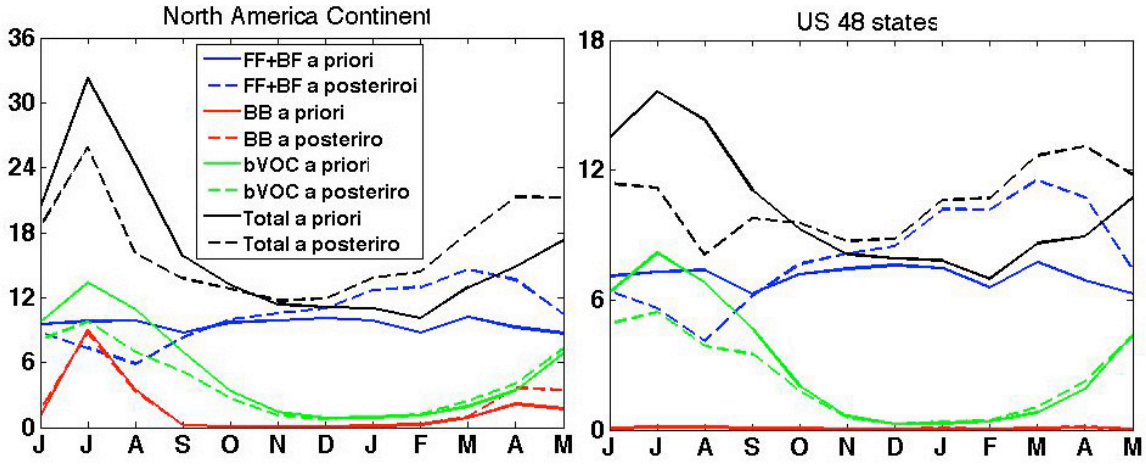


**Figure 8**. Monthly CO emissions during June 2004 – May 2005 for different emission categories: total emissions (black), anthropogenic emissions (blue), biomass burning (red), and the oxidation of biogenic NMVOCs (green). The a priori estimates are shown with the solid line and a posteriori values are indicated with the dashed line. The unit is Tg/month.

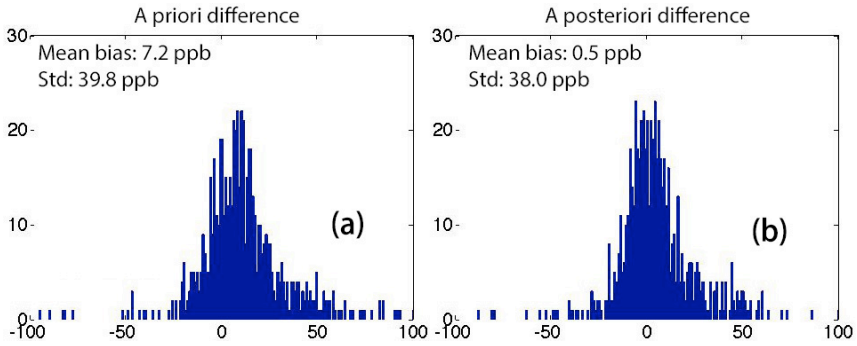


**Figure 9**. Difference between the GEOS-Chem simulation with INTEX-A DC-8 aircraft observation in free troposphere in July 2004. (a) A priori model simulation (based on the optimized initial and boundary conditions). (b) A posteriori model simulation. The model is sampled at the aircraft measurements time, location and alitude.

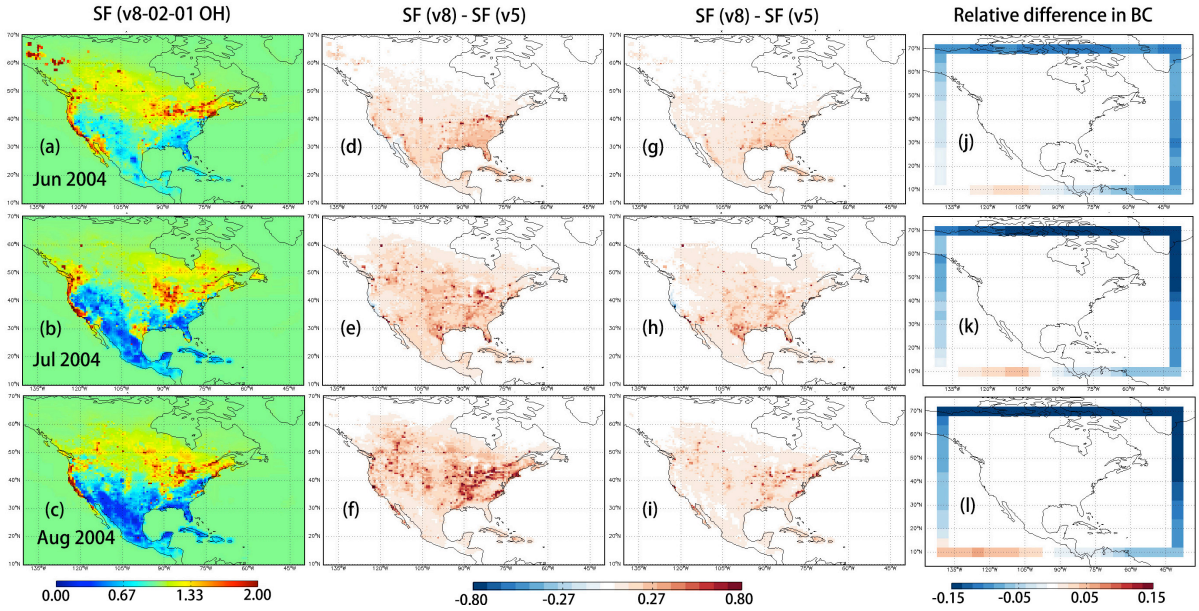


**Figure 10.** (a–c): Scaling factors of v8OH_BCv8 inversion, based on the v8-02-01 OH; (d–f): Difference between scaling factors of the v5OH_BCv5 and v8OH_BCv8 inversions; (g–i): Similar to Panel (d-f), but for the v8OH_BCv5 inversion, with the initial and boundary conditions from the standard inversion with the v5-07-08 OH fields; (j-l) Relative difference between the lower tropospheric CO partial columns (surface – 500 hPa) of the boundary conditions for the v5OH_BCv5 (v5) and the v8OH_BCv8 (v8) inversions, calculated as: 2 * (v8 - v5) / (v8 + v5).

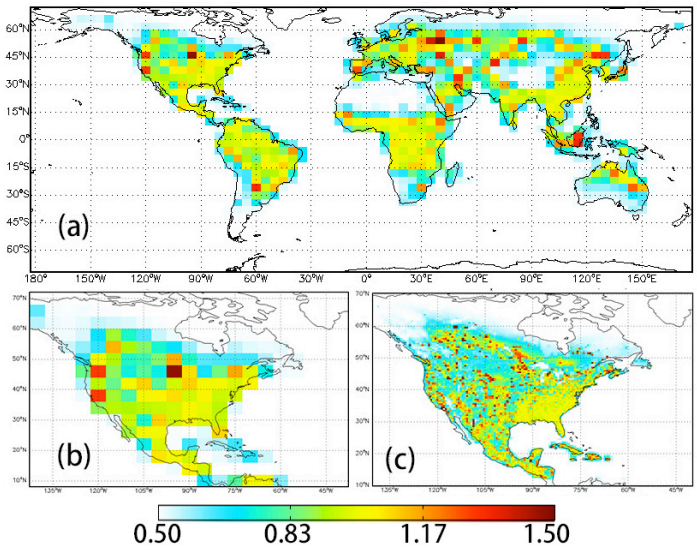


**Figure A1**. OSSE scaling factors for June 1-15, 2004. (a) Global reference inversion with 4°x5° resolution. (b) Zoomed in North America region of the global inversion. (c) Nested inversion results. The scaling factor for the first guess is 0.5 and for the true state is 1.0.