

# Sources of carbon monoxide and formaldehyde in North America determined from high-resolution atmospheric data

S. M. Miller<sup>1</sup>, D. M. Matross<sup>2</sup>, A. E. Andrews<sup>3</sup>, D. B. Millet<sup>4</sup>, M. Longo<sup>1</sup>, E. W. Gottlieb<sup>1</sup>, A. I. Hirsch<sup>3</sup>, C. Gerbig<sup>5</sup>, J. C. Lin<sup>6</sup>, B. C. Daube<sup>1</sup>, R. C. Hudman<sup>1</sup>, P. L. S. Dias<sup>7</sup>, V. Y. Chow<sup>1</sup>, and S. C. Wofsy<sup>1</sup>

<sup>1</sup>Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, USA

<sup>2</sup>Department of Environmental Science, Policy and Management, University of California Berkeley, Berkeley, CA, USA

<sup>3</sup>NOAA Earth Systems Research Laboratory, Global Monitoring Division, Boulder, CO, USA

<sup>4</sup>Department of Soil, Water, and Climate, University of Minnesota, St. Paul, MN, USA

<sup>5</sup>Max-Planck Institute for Biogeochemistry, Jena, Germany

<sup>6</sup>Department of Earth and Environmental Sciences, University of Waterloo, Waterloo, ON, Canada

<sup>7</sup>National Laboratory of Scientific Computing, Petropolis, Brazil

Received: 24 April 2008 – Published in Atmos. Chem. Phys. Discuss.: 11 June 2008

Revised: 14 November 2008 – Accepted: 28 November 2008 – Published: 18 December 2008

**Abstract.** We analyze the North American budget for carbon monoxide using data for CO and formaldehyde concentrations from tall towers and aircraft in a model-data assimilation framework. The Stochastic Time-Inverted Lagrangian Transport model for CO (STILT-CO) determines local to regional-scale CO contributions associated with production from fossil fuel combustion, biomass burning, and oxidation of volatile organic compounds (VOCs) using an ensemble of Lagrangian particles driven by high resolution assimilated meteorology. In many cases, the model demonstrates high fidelity simulations of hourly surface data from tall towers and point measurements from aircraft, with somewhat less satisfactory performance in coastal regions and when CO from large biomass fires in Alaska and the Yukon Territory influence the continental US.

Inversions of STILT-CO simulations for CO and formaldehyde show that current inventories of CO emissions from fossil fuel combustion are significantly too high, by almost a factor of three in summer and a factor two in early spring, consistent with recent analyses of data from the INTEX-A aircraft program. Formaldehyde data help to show that sources of CO from oxidation of CH<sub>4</sub> and other VOCs represent the dominant sources of CO over North America in summer.

## 1 Introduction

Carbon monoxide is a key species for both atmospheric chemistry and public health. In the United States, it is one of the original six criteria air pollutants in the Clean Air Act of 1970, and many urban areas remain either in violation of ambient CO air quality standards or at risk of violation (US EPA, 2007b). Effective emissions control strategies require accurate emissions inventories and models that can forecast concentrations across North America. Carbon monoxide also plays important roles in ozone production, in regulating concentrations of OH radicals, and indirectly in climate forcing (Thompson, 1992; Daniel and Solomon, 1998; Warneke et al., 2006).

Primary emissions of CO arise from incomplete combustion. Motor vehicle exhaust accounts for 85–95% of fossil fuel sources (US EPA, 2007a). Other major sources include biomass burning and secondary production from oxidation of methane and other volatile organic compounds (VOCs) emitted from anthropogenic sources, wetlands, and vegetation (Granier et al., 2000; Goldstein and Galbally, 2007). The principal sink for CO is oxidation by the OH radical, giving a mean atmospheric lifetime of two months (Logan et al., 1981).

The present paper develops a model-data fusion framework to provide accurate CO source magnitudes on regional/continental scales and to attribute source strengths to specific processes. Despite a long history of emissions estimates, substantial uncertainty remains in knowledge of



Correspondence to: S. M. Miller  
(scot.m.miller@post.harvard.edu)

carbon monoxide sources. IPCC (2007) indicates that remote sensing efforts have helped constrain CO emissions, but several recent studies suggest that EPA's 1999 National Emissions Inventory (NEI-1999) may overestimate anthropogenic CO emissions by 50–300% (Parrish, 2006; Turnbull et al., 2006; Warneke et al., 2006; Hudman et al., 2008). Attempts to estimate another major CO source - secondary production from biogenic VOC emissions - stretch back as far as the 1970s (Zimmerman, 1979; Guenther et al., 1995; Stewart et al., 2003; Chang et al., 2005; Simon et al., 2006; Guenther et al., 2006). VOCs are emitted by anthropogenic and biogenic sources, but biogenic VOC emissions, particularly isoprene and monoterpenes from plants, constitute 80% of the total global source (Olivier et al., 2001).

Recent studies have attempted to improve knowledge of VOC sources by using remote sensing measurements of formaldehyde (e.g., Palmer et al., 2003, 2006). Nevertheless, the magnitude and distribution of VOC sources remain very controversial. For example, the commonly-used GEIA biogenic VOC inventory differs from prior estimates by as much as a factor of five (Guenther et al., 1995).

The combination of remote sensing and in situ data for CO and formaldehyde can help distinguish production of CO from different sources. When methane and VOCs decay to CO, both decay to a common intermediate species: formaldehyde (HCHO) (Duncan et al., 2007). The atmospheric lifetime of formaldehyde is only a few hours or less, and the CO yield is near unity (Palmer et al., 2003; Duncan et al., 2007). Formaldehyde data have been used to validate emissions estimates of VOCs and CO in a number of recent studies (Abott et al., 2003; Palmer et al., 2003; Martin et al., 2004; Shim et al., 2005; Millet et al., 2006; Palmer et al., 2006; Millet et al., 2008).

Sources of CO from biomass fires are also poorly constrained. Biomass burning contributes 15–30% of all global CO emissions (IPCC, 2001; Petron et al., 2004; Müller and Stavrou, 2005; Arellano and Hess, 2006; Duncan et al., 2007). Individual fires can be vast and persist for significant periods of time. For example, during one episode, Canadian fires enhanced carbon monoxide concentrations over Ireland by almost 60% (Forster et al., 2001). A variety of methods have been used to estimate biomass burning sources of CO: historical written fire records (e.g., Liu, 2004), inverse models (e.g., Wotowa and Trainer, 2000), and satellite data (e.g., Pfister et al., 2005; Wiedinmyer et al., 2006). Remote sensing instruments have been used to quantify monthly or even daily variations in biomass burning emissions (Duncan et al., 2003; Ito and Penner, 2004; Wiedinmyer et al., 2006). However, even after careful processing and assessment, satellite estimates still differ markedly from historical fire records; one of the most recent satellite estimates (Wiedinmyer et al., 2006), disagrees with its predecessors by as much as a factor of two. Uncertainties in emissions estimates arise from uncertainties in fuel loadings (estimates of biomass per area), in emissions factors (volume of emissions per mass burned),

in combustion efficiency (fraction of biomass burned), and from the inability of satellites to detect fires through cloud cover (Wiedinmyer et al., 2006).

Lagrangian models like STILT-CO are particularly well-suited to determine the magnitude and distribution of CO sources. If a measurement site is located in a rural area, the carbon monoxide record will show distinct peak event periods separated by discrete non-peak periods. The peaks reflect transport from intense localized sources (urban areas, fires). The background arises because CO has an atmospheric lifetime of about two months - enough time to transport the pollutant over long distances but not enough time for the pollutant to build up to very high levels in the absence of intense localized sources (Pfister et al., 2004). If the model overestimates or underestimates peak pollution events, the results suggest well-defined adjustments to the original emissions inventories.

Time-inverted Lagrangian models have been used in a number of studies to characterize regional pollution sources for a variety of trace gases. Moody et al. (1998) defined patterns of backward particle trajectories and matched these transport patterns with fluctuations in pollution measurements taken at Harvard Forest in Massachusetts. Vermeulen et al. (2006) used the Lagrangian model COMET on small spatial scales (5×5 km to 10×20 km) to explain the observed variance in methane concentrations at measurement sites downwind of urban sources in Europe. Warneke et al. (2006) applied Lagrangian modeling to carbon monoxide using the FLEXPART model to estimate CO concentrations at measurement sites in New England. Warneke et al. (2006) modeled CO only from anthropogenic and biomass burning sources with no photochemical loss; FLEXPART obtained a model-measurement fit of  $r^2=0.30-0.45$  and inferred that EPA's NEI-99 inventory may be too high by 50% in Boston/New York urban outflow.

The present paper describes the Stochastic Time-Inverted Lagrangian Transport Model for CO (STILT-CO), incorporating anthropogenic emissions, biogenic VOCs, biomass burning emissions, and associated atmospheric chemical processes into an hourly model of CO and formaldehyde concentrations over North America at a high spatial resolution (45 km). STILT-CO allows us to create very detailed representations of carbon monoxide and formaldehyde concentrations in time and space, which can be compared to a wide variety of observations from tall towers and aircraft. We then use a Bayesian optimization technique to refine current estimates of anthropogenic CO emissions and CO production from VOC emissions. We also examine more generally some of the challenges in source-receptor Lagrangian modeling that arise, for example, in coastal areas.

## 2 Methodology

### 2.1 The STILT-CO model

The Stochastic Time Inverted Lagrangian Transport Model (STILT) of Lin et al. (2003) and Gerbig et al. (2003) is a Lagrangian Particle Dispersion model (LPDM) that forms the foundation for STILT-CO. STILT calculates concentrations of a trace gas at a single point, known as a receptor point, defined as a location in space and time that corresponds to a measurement (e.g. at a tall tower or on an aircraft flight). A series of ambient air measurements taken every hour at a tower for one day would count as 24 different receptor points, all with the same location but at different times. A time-inverted LPDM releases an ensemble of imaginary air parcels or particles from the receptor point that travel upwind (backward in time), and the trace gas sources that these particles encounter while traveling upwind are then used to calculate concentrations at the receptor point. The very detailed rendition of concentration fluctuations provided by the LPDM can be validated against individual measurements taken at the receptor, providing a powerful framework for assessing upwind surface or volume sources. The following sections describe in detail the STILT transport model and its subsequent application to carbon monoxide and formaldehyde concentrations.

#### 2.1.1 The modeled advected boundary condition

The STILT lateral tracer boundary condition, developed by Gerbig et al. (2003), uses CO and CH<sub>4</sub> levels observed at Pacific stations of the NOAA monitoring network to derive a boundary condition for all altitudes at the 145° W meridian. Most particles (~65%) released in our domain cross the 145° W boundary after six days or less, while some of the remainder may stay a long time in the domain (Gerbig et al., 2003). When a particle reaches the terminal time step set in the model (typically 10 days), or when it reaches 145° W, the boundary condition is taken from its latitude and altitude projected, if needed, onto the 145° W meridian (Gerbig et al., 2003). The boundary condition has daily temporal resolution and 2.5° latitude by 0.5 km altitude spatial resolution (Matross et al., 2006). Because the lifetime of formaldehyde is a few hours or less, we set formaldehyde to zero at the boundary.

To form the lateral tracer boundary condition, Gerbig et al. (2003) used CO and CH<sub>4</sub> measurements from three different monitoring stations on the NOAA GMD network: Cape Kumakahi, Hawaii; Cold Bay, Alaska; and Barrow, Alaska. Matross et al. (2006) supplemented these measurements with aircraft data over Carr, Colorado; Poker Flats, Alaska; and Park Falls, Wisconsin. These data were interpolated over all latitudes and times on the western domain boundary. Gerbig et al. (2003) then used a Green's function fit to available aircraft data over the Pacific in order to derive a boundary

condition to all altitudes. Transport of CO and CH<sub>4</sub> from the boundary in STILT allows for chemical loss due to oxidation in transit to the receptor point, as outlined in Sects. 2.1.5 and 2.1.6.

We do not include any VOC boundary conditions. Measurements from Millet et al. (2004) provide an estimate of VOC concentrations in air at Trinidad Head, CA, advected from the Pacific Ocean. They found an average acetone concentration of 0.6 ppb from air transported over the Pacific. We estimate that the lack of an acetone boundary condition within the model contributes 0.07–0.08 ppb uncertainty in formaldehyde model results.

#### 2.1.2 The STILT meteorological transport model

The STILT model calculates the change in the concentration of a trace gas at a receptor point with location  $\mathbf{x}_r$  and time  $t_r$ ,  $\Delta C(\mathbf{x}_r, t_r)$ , by multiplying the spatially and temporally resolved source  $S(\mathbf{x}, t)$  (units:  $\mu\text{mol m}^{-2} \text{s}^{-1}$ ), by the influence function  $I(\mathbf{x}_r, t_r|\mathbf{x}, t)$  (units:  $\text{ppm}/(\mu\text{mol m}^{-2} \text{s}^{-1})$ ) of the source location on the receptor point and integrating over the domain  $V$  (Eq. 1, term 1). The second term in Eq. (1) provides the contribution from the advection of the initial tracer field, taken from the boundary condition (Lin et al., 2003).

$$\Delta C(\mathbf{x}_r, t_r) = \int_{t_0}^{t_r} dt \int_V d^3x I(\mathbf{x}_r, t_r|\mathbf{x}, t) S(\mathbf{x}, t) + \int_V d^3x I(\mathbf{x}_r, t_r|\mathbf{x}, t_0) C(\mathbf{x}, t_0) \quad (1)$$

$$I(\mathbf{x}_r, t_r|\mathbf{x}, t) = \frac{\rho(\mathbf{x}_r, t_r|\mathbf{x}, t)}{N_{\text{tot}}} \quad (2)$$

To calculate the influence  $I$  of a particular location  $(\mathbf{x}, t)$  in space and time, the model divides the density of particles computed by the LPDM at  $(\mathbf{x}, t)$ ,  $\rho(\mathbf{x}_r, t_r|\mathbf{x}, t)$ , by the total number of particles,  $N_{\text{tot}}$ , released backward in time from the receptor (see Eq. 2) (Lin et al., 2003).

The model computes the source function  $S(\mathbf{x}, t)$  associated with a surface flux  $F(\mathbf{x}, t)$  by distributing mass emitted at the surface through the atmosphere to a mixing height  $h$ , set as a fraction of the planetary boundary layer (PBL) height. Gerbig et al. (2003) found that varying  $h$  between 10% and 100% of the planetary boundary layer (PBL) did not significantly affect model results. We set this initial mixing height for surface sources equal to half the PBL height in the current paper.

Equation (1) can be made more directly applicable to surface fluxes by integrating over the grid elements and time step of the model ( $\tau$ ), to obtain Eq. (3). Here  $\Delta C_{m,i,j}(\mathbf{x}_r, t_r)$  is the contribution to the concentration from the volume element at the surface due to gases emitted between time  $t_m$  and  $t_m + \tau$ .  $f(x_i, y_j, t_m)$  is the footprint function defined by the expression in brackets.  $m_{\text{air}}$  is the molar mass of air. The total concentration change due to all surface sources in the domain,  $\Delta C_{\text{CO}}(\mathbf{x}_r, t_r)$ , is obtained by summing over all  $m, i,$

and  $j$  (Eq. 4). Equation (4) accounts for direct CO emissions at the surface with flux  $F_{\text{CO}}$  (1st term), CO produced from the chemical degradation of  $\text{CH}_4$  and VOCs emitted from the surface with flux  $F_{[\text{VOCs,CH}_4]}$  (2nd term), and CO loss by reaction with OH to  $\text{CO}_2$  (3rd term). The second term includes a chemistry function  $R(x_i, y_j, t_m | \mathbf{x}_r, t_r)$  that accounts for creation of CO due to chemistry on precursor gases (emission fluxes  $F_{[\text{VOCs,CH}_4]}$ ) during particle transit to the receptor, and the third term describes CO loss due to chemistry en route to the receptor point (OH oxidation with rate constant  $k_{\text{OH}}$ ). Summing over all footprint elements for different CO and VOC sources yields the concentration due to surface sources that is seen at the receptor.

$$\begin{aligned} \Delta C_{m,i,j}(\mathbf{x}_r, t_r) &= \left[ \frac{m_{\text{air}}}{(h\rho(x_i, y_j, t_m))} \int_{t_m}^{t_m+\tau} dt \int_{x_i}^{(x_i+\Delta x)} dx \int_{y_j}^{(y_j+\Delta y)} dy \int_0^h dz I(\mathbf{x}_r, t_r | \mathbf{x}, t) \right] \\ &\cdot F(x_i, y_j, t_m) = f(\mathbf{x}_r, t_r | x_i, y_j, t_m) F(x_i, y_j, t_m) \end{aligned} \quad (3)$$

$$\begin{aligned} \Delta C_{\text{CO}}(\mathbf{x}_r, t_r) &= \sum \text{CO from direct anthropogenic emissions} + \text{CO} \\ &\text{from VOCs and CH}_4 - \text{loss of CO (from direct fluxes,} \\ &\text{VOCs, and CH}_4) \text{ en route to the receptor} \\ &= \sum_{i,j,m} \left\{ f(x_i, y_j, t_m) F_{\text{CO}}(x_i, y_j, t_m) \right. \\ &+ f(x_i, y_j, t_m) F_{[\text{VOCs,CH}_4]}(x_i, y_j, t_m) \int_{t_m}^{t_r} R(x_i, y_j, t_m | \mathbf{x}_r, t) dt \\ &\left. - f(x_i, y_j, t_m) F_{[\text{CO,VOCs,CH}_4]}(x_i, y_j, t_m) \int_{t_m}^{t_r} k_{\text{OH}}[\text{OH}] dt \right\} \end{aligned} \quad (4)$$

Equation 5 describes the similar approach taken for formaldehyde. The HCHO signal at the tower from surface sources ( $\Delta C_{\text{HCHO}}(\mathbf{x}_r, t_r)$ ) equals the influence of formaldehyde surface sources (1st term), the influence of VOC and  $\text{CH}_4$  surface fluxes that decay to formaldehyde (2nd term) and the decay of formaldehyde given by the decay rate  $j_{\text{HCHO}}$  (described in more detail in Sect. 2.1.5) (3rd term). The emissions from each back trajectory location and time ( $x_i, y_i, t_m$ ) are summed to find the influence of advected continental sources ( $\Delta C_{\text{HCHO}}(\mathbf{x}_r, t_r)$ ).

$$\begin{aligned} C_{\text{HCHO}}(\mathbf{x}_r, t_r) &= \sum_{i,j,m} \left\{ f(x_i, y_j, t_m) F_{\text{HCHO}}(x_i, y_j, t_m) \right. \\ &+ f(x_i, y_j, t_m) F_{[\text{VOCs,CH}_4]}(x_i, y_j, t_m) \int_{t_m}^{t_r} R(x_i, y_j, t_m | \mathbf{x}_r, t) dt \\ &\left. - f(x_i, y_j, t_m) F_{[\text{HCHO,VOCs,CH}_4]}(x_i, y_j, t_m) \int_{t_m}^{t_r} j_{\text{HCHO}} dt \right\} \end{aligned} \quad (5)$$

The domain for the STILT model over North America extends from  $11^\circ \text{N}$  latitude to  $70^\circ \text{N}$  and from  $-145^\circ$  longitude to  $-51^\circ$ . The transport grid size is 45 km and the surface emissions fluxes are gridded with maximum resolution

of  $1/6^\circ$  latitude by  $1/4^\circ$  longitude. All particle trajectories are run ten days backward in time, or until the particles leave the domain, whichever is shorter.

STILT utilizes a dynamic re-gridding scheme when calculating the influence footprint. As particles track far from the receptor, the footprint covers larger areas, and the statistical probability of finding a particle in a particular grid square becomes small. The STILT-CO model produces results with less statistical noise by dynamically aggregating the grid of surface fluxes as the particle ensemble disperses (Gerbig et al., 2003).

We initially used three assimilated meteorological drivers to run the particle ensembles back in time: the final data assimilation of the National Centers for Environmental Prediction model (FNL) (Stunder, 1997), the Eta Data Assimilation System 40 km (EDAS-40) (NOAA ARL, 2004), and the Brazilian adaptation of the Regional Atmospheric Modeling System (BRAMS) (Pielke et al., 1992; Cotton et al., 2004; Sanchez-Ccoylo et al., 2006). The FNL and EDAS-40 fields produced substantial mass violation, and therefore, BRAMS is the primary meteorological driver used for all model runs.

Our BRAMS core model (v. 3.2) is strongly based on RAMS solver, with several optimizations for faster solution, developments for enhanced portability, and new parameterizations for convection (shallow and deep) and turbulence. We modified the diagnostic outputs from BRAMS to ensure mass conservation to very high accuracy and applied a specific mass conservation fix from Medvigy et al. (2005). The domain consisted of a single, 45-km horizontal resolution grid, covering most of North America. The simulated period was from 1 February 2004 to 1 March 2005. The vertical coordinate was terrain-following with a resolution ranging from 150 m at the bottom of the domain to 850 m at the top of the domain (20 600 m maximum altitude).

Interactions between the atmosphere, biosphere, and soil were solved using LEAF-3 surface sub-model (Walko et al., 2000). Sub-grid convective clouds were parameterized using the Grell and Devenyi (2002) scheme, from which we retrieved mass fluxes due to convection, entrainment and detrainment. We also computed the average vertical Lagrangian time scale, based on Hanna (1982), and retrieved the boundary layer height, following Vogelesang et al. (1996). The model timestep was 60 s. The variables needed for STILT were output every 10 min to ensure consistency between RAMS and STILT transport and to enhance mass conservation.

### 2.1.3 Overview of CO and HCHO sources

STILT-CO incorporates primary CO sources at the surface from two distinct processes: fossil fuel combustion and biomass burning, plus volume sources of CO produced from the oxidation of biogenic VOCs emitted at the surface. The CO model also accounts for CO production from the oxidation of methane (from surface sources and

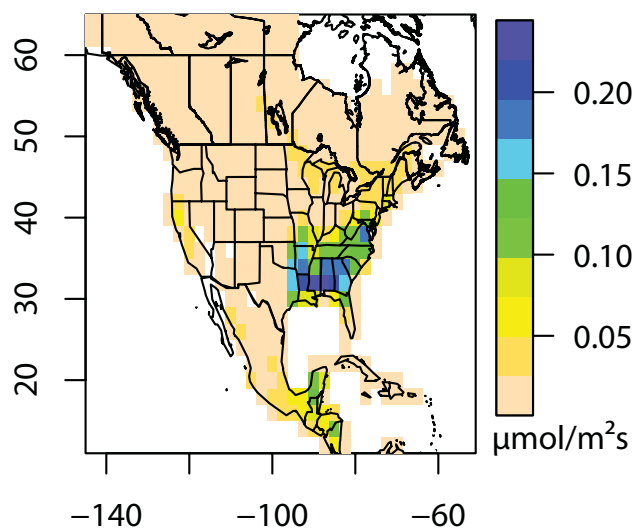
the boundary condition) and CO loss due to oxidation to CO<sub>2</sub>. The formaldehyde model incorporates HCHO from anthropogenic formaldehyde sources, from the decay of biogenic VOCs, and from methane decay (both from continental sources and from the methane boundary condition). The formaldehyde model also accounts for HCHO losses to CO via oxidation and photolysis. We assume negligible HCHO loss due to deposition. The following sections describe the surface flux emissions inventories and the chemistry mechanisms within the model.

#### 2.1.4 Surface fluxes

The STILT-CO model utilizes a variety of different emissions inventories for the purpose of comparing different source estimates. This paper primarily relies upon the US EPA's 1999 National Emissions Inventory (NEI-1999) for anthropogenic CO and formaldehyde emissions over the US, Canada, and Mexico (US EPA, 2004; Frost and McKeen, 2007). Newer less detailed EPA inventories through 2006 are relatively similar in magnitude: EPA estimates that 2006 emissions are a modest 14% lower than in 1999, due mostly to reductions in highway vehicle emissions (US EPA, 2007c). Two other inventories for anthropogenic CO provide comparison: an EPA 1993 northeastern regional inventory interpolated across the United States using the correlation between CO and NO<sub>x</sub> emissions (Benkovitz et al., 1996; Gerbig et al., 2003) and the Emissions Database for Global Atmospheric Research 2000 inventory (EDGAR-2000) (Olivier et al., 2005). The EDGAR-2000 inventory has a 1° latitude × 1° longitude resolution whereas both EPA inventories have been re-gridded from counties to a 1/6° latitude by 1/4° longitude resolution. The EPA-1993 and EDGAR-2000 inventories average emissions over the year; the STILT-CO model then applies hourly and weekday/weekend scaling factors (Ebel et al., 1997). The NEI-1999 gives average hourly emissions rates over summer months and weekday/weekend scaling factors are applied.

For biogenic VOC emissions, STILT-CO uses the MEGAN (Model of Emissions from Gases and Aerosols from Nature) inventory (Guenther et al., 2006). The MEGAN framework calculates ecosystem-specific emissions scaled to leaf area, light, and temperature. Here we use GEOS-Chem (<http://www.as.harvard.edu/chemistry/trop/geos/index.html>), a global Eulerian atmospheric chemistry model driven by GEOS-4 meteorological fields, to calculate MEGAN fluxes for STILT-CO simulations. We utilize biogenic isoprene, monoterpenes, acetone, and alkenes emissions over the North American continent with a 2-hourly, 2° × 2.5° resolution. Figure 1 displays a map of mean midday biogenic VOC fluxes (at 01:00 p.m. CST/11:00 a.m. PST) over North America from 1 June to 15 August from the MEGAN inventory (Millet et al., 2006; Hudman et al., 2008). STILT-CO does not include anthropogenic VOC sources: we

### Mean July Mid-Day VOC Emissions



**Fig. 1.** Mean midday vegetation VOC fluxes over North America from the MEGAN inventory (Guenther et al., 2006). The highest VOC emissions occur over the American Deep South.

estimate this likely contributes ~2–4 ppb to model uncertainty for CO.

The EDGAR 1995 inventory provides anthropogenic emissions estimates of methane (Olivier et al., 2001). Kort et al. (2008) applied the STILT model to methane and found that natural wetland sources over North America contributed only 3% of the total methane model enhancement (the model prediction minus the advected boundary condition). Most of the methane contribution in the STILT-CO model comes from the decay of methane at the boundary condition (97.5%). Given the relatively small contribution of continental methane sources to the CO budget, we include only anthropogenic methane sources in the model. We estimate a possible model uncertainty from natural methane fluxes of ~3–5% in the boundary layer near wetland source regions.

The biomass burning component of the STILT-CO model uses daily satellite estimates of biomass burning from Wiedinmyer et al. (2006), who used the MODIS Aqua and Terra satellites to identify fires of size 100 m<sup>2</sup> or larger in North America. The Wiedinmyer et al. (2006) study produces a daily 1 km by 1 km grid estimate of biomass burning emissions across North America (10–71° N and –175 to –55° W). The emissions were subsequently regrided to a 1/6 × 1/4 degree latitude/longitude resolution for the STILT-CO model.

#### 2.1.5 VOC chemistry

In order to reduce the computational expense of the Lagrangian model, VOC chemistry in the STILT-CO model

is simplified from the VOC reactions that occur in nature. We follow as tracers isoprene, monoterpenes, acetone, and higher order alkenes, and represent their decay to HCHO, CO, and finally CO<sub>2</sub>. Reactions (R1)–(R3) show the simplified model chemistry. Because not all carbon atoms in VOCs are converted to HCHO or CO, the model applies a yield factor ( $\alpha$ ) to R1. The yield of CO from HCHO is subsequently assumed to be one. The model utilizes yield factors of 0.28 (Palmer et al., 2006), 0.15 (Granier et al., 2000), 0.25 (Somnitz et al., 2005), and 0.24 (Duncan et al., 2007) for isoprene, monoterpenes, acetone, and alkenes respectively. The yield factor for isoprene is based upon low NO<sub>x</sub> concentrations, the condition most likely to prevail over the domain sampled by the WLEF tower (see Fig. 9). We take mean VOC decay lifetimes to formaldehyde from empirical satellite formaldehyde column observations: seven hours for isoprene and five hours for monoterpenes (Palmer et al., 2006). We also use a decay lifetime to formaldehyde for acetone of 15 days (Singh et al., 2004; de Reus et al., 2003). These lifetimes scale inversely with diurnal fluctuations in OH from Martinez et al. (2003).



Equations (6)–(9) show how VOCs decay to HCHO and CO in the STILT-CO model, where  $k_1$  is the decay constant for VOCs,  $j_2$  is the decay constant of HCHO,  $k_3$  is the oxidation rate constant for CO from NASA's Jet Propulsion Laboratory (2006), and  $\alpha$  represents the yield factor. HCHO loss rates are taken from the GEOS-Chem model on a 2-hourly 2×2.5 degree latitude-longitude resolution, and CO and CH<sub>4</sub> loss rates are calculated from chemical rate equations. Equation (6) describes the decay of VOCs to HCHO (1st term) and the decay of HCHO to CO (2nd term). Solving Eq. (6) for HCHO gives an expression (Eq. 7) for the concentration of HCHO from VOCs at the receptor after the gases have been transported for time  $t$ . Similarly, Eq. (8) expresses the creation of CO from decaying formaldehyde (1st term) and the loss of CO to oxidation (2nd term). Solving this equation produces Eq. (9): the concentration of CO produced from VOCs at the receptor after gases have been transported for time  $t$  from the source ( $x_i, y_j, t_m$ ) to the receptor ( $x_r, t_r$ ).

$$d \frac{[\text{HCHO}]}{dt} = \alpha k_1 [\text{VOC}] - j_2 [\text{HCHO}] \quad (6)$$

$$[\text{HCHO}]_{\text{VOCs}}(t) = \frac{k_1 \alpha [\text{VOC}_{t=0}]}{(k_1 - j_2)} (e^{-j_2 t} - e^{-k_1 t}) \quad (7)$$

$$d \frac{[\text{CO}]}{dt} = j_2 [\text{HCHO}] - k_3 [\text{OH}][\text{CO}] \quad (8)$$

$$[\text{CO}]_{\text{VOCs}}(t) = \frac{\alpha [\text{VOC}_{t=0}] (k_1 * j_2)}{(j_2 - k_1)(j_2 - k_3 [\text{OH}])} (e^{-j_2 t} - e^{-k_3 [\text{OH}] t}) - \frac{\alpha [\text{VOC}_{t=0}] (k_1 * j_2)}{(j_2 - k_1)(k_3 [\text{OH}] - k_1)} (e^{-k_1 t} - e^{-k_3 [\text{OH}] t}) \quad (9)$$

### 2.1.6 Additional model chemistry

In addition to VOC chemistry, the model incorporates chemistry from CH<sub>4</sub> loss to HCHO, HCHO loss to CO, and CO loss to CO<sub>2</sub>. These reactions are included in calculating the CO advected boundary condition, CO and HCHO from the CH<sub>4</sub> boundary condition and surface fluxes, and chemical losses of HCHO surface fluxes. The model calculates CH<sub>4</sub> losses using the reaction constant from NASA JPL (2006) and 2-hourly OH concentrations from the GEOS-Chem model.

## 2.2 Model optimization framework

Inverse modeling provides a powerful tool for using hourly model results to improve emissions estimates and reduce the uncertainty in these inventories. Many existing studies use inverse models to characterize CO sources (e.g., Kasibhatla et al., 2002; Heald et al., 2003; Petron et al., 2004; Pfister et al., 2005). None of these previous inversions use regional-scale Lagrangian models where source-receptor relationships are highly resolved and transparent.

The Bayesian inversion framework used here closely follows the framework of Gerbig et al. (2003) and Matross et al. (2006) in their studies of CO<sub>2</sub> fluxes from vegetation. Here we optimize for overall scaling factors for the anthropogenic CO, biomass burning, and biogenic VOC inventories respectively, incorporating estimates of prior uncertainties in the model and emissions inventories in order to produce a posteriori emissions scaling factors. This optimization cannot correct for problems in the spatial distribution of emissions or errors in the transport field, as discussed below.

Following the general inverse methods outlined by Rogers (2000), CO measurements at a tall tower can be related to CO surface sources through the following equation:

$$\mathbf{y} = \mathbf{K}\mathbf{\Gamma} + \boldsymbol{\epsilon} \quad (10)$$

where  $\mathbf{y}$  is the hourly measured concentration at the tall tower,  $\mathbf{K}$  is the Jacobian matrix relating the vector of measured values to the state vector,  $\mathbf{\Gamma}$  is the vector of a posteriori scaling factors, and  $\boldsymbol{\epsilon}$  is a vector of errors in the measurements and hourly model results. In the inversion framework, the state space refers to the elements being optimized by the inversion, in this case anthropogenic CO emissions, biogenic VOC emissions, and biomass burning estimates. The non-state space refers to elements other than those being optimized. More specifically, Eq. (11) calculates the a posteriori scaling factors and Eq. (12) calculates the a posteriori uncertainty in trace gas sources.

$$\mathbf{\Gamma}_{\text{post}} = (\mathbf{K}^T \mathbf{S}_{\varepsilon}^{-1} \mathbf{K} + \mathbf{S}_{\text{prior}}^{-1})^{-1} (\mathbf{K}^T \mathbf{S}_{\varepsilon}^{-1} \mathbf{y} + \mathbf{S}_{\text{prior}}^{-1} \mathbf{\Gamma}_{\text{prior}}) \quad (11)$$

$$\mathbf{S}_{\text{post}} = (\mathbf{K}^T \mathbf{S}_{\varepsilon}^{-1} \mathbf{K} + \mathbf{S}_{\text{prior}}^{-1})^{-1} \quad (12)$$

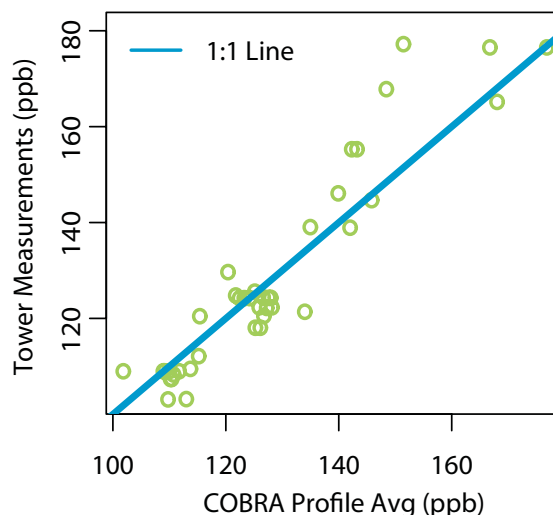
In Eq. (11),  $\mathbf{y}$  is the measured CO source signal at the tower: the hourly tall tower CO measurements minus the modeled lateral tracer boundary condition. For a given model run with  $m$  hourly data points,  $\mathbf{y}$  is a vector of length  $m$ . The Jacobian matrix  $\mathbf{K}$  of dimensions  $m \times n$  relates the measurements to the state vector (where  $n$  is the number of factors being optimized). In this case, the first column of the matrix contains the modeled CO fossil fuel signal (for all  $m$  receptor points). The second and third columns list the modeled CO signal from VOCs and from biomass burning respectively.

$\mathbf{\Gamma}_{\text{prior}}$ , a vector of length  $n$ , represents the a priori scaling factors in the state space. Because none of the sources are scaled prior to the inversion, the  $\mathbf{\Gamma}_{\text{prior}}$  vector is set to one.  $\mathbf{\Gamma}_{\text{post}}$ , calculated in Eq. (11), gives the a posteriori scaling factors that optimize the CO sources in the state space. The Bayesian framework presented here produces three a posteriori scaling factors to scale the CO source from anthropogenic emissions, VOCs, and biomass burning, respectively.

The  $\mathbf{S}_{\text{prior}}$  matrix with dimensions  $n \times n$  is the prior error covariance matrix of the elements in the state space. The diagonal elements of the matrix represent the uncertainty in each of the three elements. EPA does not provide error estimates for the NEI-1999, so the uncertainty is estimated at 60% in accordance with the inventory error as estimated by Hudman et al. (2008). Wiedinmyer et al. (2006) and Pfister et al. (2005) estimate uncertainty in the biomass burning inventory at a factor of two. Therefore, we use 100% as the a priori uncertainty in the biomass burning estimates. For the CO contribution from VOCs, we use an uncertainty of 30%. Palmer et al. (2006) used the GOME satellite to validate the MEGAN biogenic VOC inventory. They find that during the summer of 2001, MEGAN falls within 30% of isoprene measurements inferred from satellite-measured HCHO columns. We do not include uncertainty from CO/HCHO yields or chemistry in  $\mathbf{S}_{\text{prior}}$  because this error is not random – it is a single number affecting only the scaling of the prior, not its error covariance. To obtain the variance for the  $\mathbf{S}_{\text{prior}}$  matrix, we multiply these relative uncertainties by the respective CO signal and then square the result to obtain the weighted variance. The errors in the different source emissions are uncorrelated, so we set the non-diagonal elements of the covariance matrix at zero. The  $\mathbf{S}_{\text{post}}$  matrix (dimensions  $n \times n$ ) given in Eq. (12) lists the a posteriori uncertainty of the elements in the state space.

$\mathbf{S}_{\varepsilon}$  is the covariance matrix for all non state space elements (dimensions  $m \times m$ ). Non state space errors include uncertainties in the lateral tracer boundary condition, tall tower CO measurements, and the number of particles used in the STILT-CO model. The variance, or diagonal elements of  $\mathbf{S}_{\varepsilon}$ , can be represented by the following equation:

### COBRA Profile Avg vs. Argyle Measurements

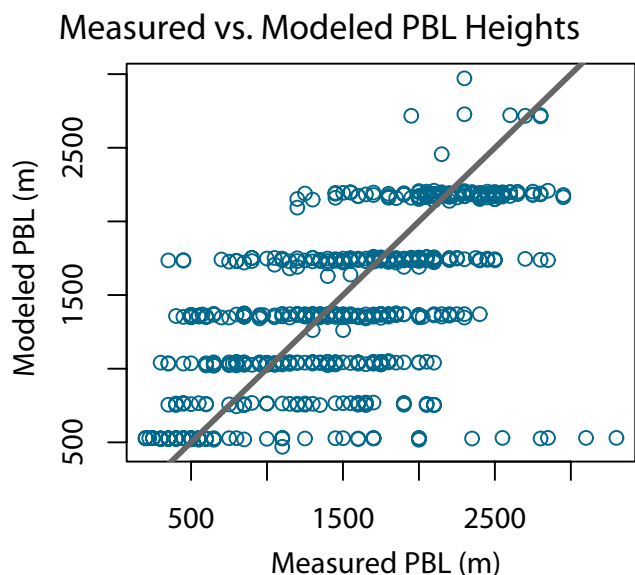


**Fig. 2.** Argyle tower CO measurements plotted against mean CO values in the PBL taken from COBRA-2004 aircraft measurements within 50 km of the tower. The mean bias between COBRA and Argyle measurements is  $-0.423$  ppb and the standard deviation is  $7.8$  ppb.

$$\mathbf{S}_{\varepsilon} = \mathbf{S}_{\text{obs}} + \mathbf{S}_{\text{background}} + \mathbf{S}_{\text{part}} + \mathbf{S}_{\text{eddy}} + \mathbf{S}_{\text{transp}} + \mathbf{S}_{\text{aggr}} \quad (13)$$

We neglect non-diagonal covariance elements in the error covariance matrix.  $\mathbf{S}_{\text{obs}}$  represents the instrumentation error in observed CO concentrations at the WLEF tower. We estimate the uncertainty in measured CO values at 5 ppb based on the high and low calibration values measured at the tower.  $\mathbf{S}_{\text{obs}}$  is therefore  $(5 \text{ ppb})^2$ .  $\mathbf{S}_{\text{background}}$  represents error in the modeled background. Gerbig et al. (2003) estimate the uncertainty in modeled background at 22 ppb.  $\mathbf{S}_{\text{part}}$  quantifies the error introduced by using a finite number of particles, in this case 100 particles. Gerbig et al. (2003) estimate  $\mathbf{S}_{\text{part}}$  at 13% of the modeled surface flux CO signal, making  $\mathbf{S}_{\text{part}} = [0.13 * (\text{modeled signal})]^2$ .

$\mathbf{S}_{\text{eddy}}$  represents the variance in the data caused by unresolved turbulent eddies within the planetary boundary layer. Entrainment of surface sources into the boundary layer and uneven vertical mixing cause significant variance of CO concentrations within the boundary layer; this unresolved variability is estimated by  $\mathbf{S}_{\text{eddy}}$ . We quantified  $\mathbf{S}_{\text{eddy}}$  by sampling all COBRA-2004 CO aircraft vertical measurement profiles for the PBL within 50 km of the Argyle tower in Bangor, ME. We subtracted measured CO values at Argyle tower from CO aircraft measurements averaged over the entire height of the PBL (see Fig. 2). The square of the standard deviation in the mean difference represents the variance of  $\mathbf{S}_{\text{eddy}}$ ,  $59.1 \text{ ppb}^2$ .

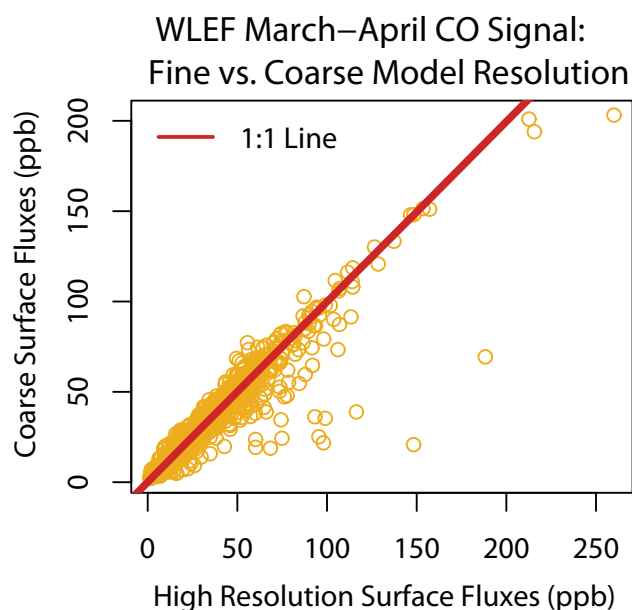


**Fig. 3.** Height of the PBL as measured in COBRA aircraft profiles is plotted against the PBL height set by the BRAMS domain. The grey line is a 1:1 line superimposed over the data.

$S_{\text{transp}}$  represents the effect of errors in the modeled height of the planetary boundary layer on modeled CO (see Fig. 3). Matross et al. (2006) calculated observed PBL heights for over 900 COBRA-2004 vertical profiles by examining potential temperature profiles. To approximate  $S_{\text{transp}}$ , we run the STILT transport model a very small step backward in time and record the PBL height as set by the BRAMS meteorological driver. The BRAMS driver sets the PBL at the midpoint between two vertical layers in the meteorological driver, resulting in discrete modeled PBL heights. We calculate percentage bias in modeled PBL height as outlined in Eq. (14), where  $z_{\text{measured}}$  is the measured PBL height and  $z_{\text{modeled}}$  is PBL height as modeled by BRAMS. The correlation ( $r$ ) between modeled and observed PBL height was 0.64. We multiply the variance in the percentage error by the hourly modeled CO fossil fuel and VOC signals at the tall tower. The method presented here follows that of Gerbig et al. (2003) and Matross et al. (2006). The modeled PBL height shows a mean bias of  $-96$  m, relatively small compared to the typical height of the PBL (1000–2000 m).

$$\text{Percentage error} = \frac{(z_{\text{measured}} - z_{\text{modeled}})}{z_{\text{modeled}}} \quad (14)$$

$S_{\text{aggr}}$  represents the aggregation error, the uncertainty from formulating the state space as a single scalar for the three source components. To make a conservative estimate of  $S_{\text{aggr}}$ , we model hourly CO levels at WLEF tower using both the highest possible inventory resolution and the coarsest resolution for surface fluxes (16 times larger grid cells than the original inventory) (see Fig. 4). The results show an average



**Fig. 4.** A plot showing the hourly modeled CO signal with surface fluxes at the highest and lowest resolutions. This plot shows the modeled CO signal, meaning that the results do not include CO from the boundary condition. The red line is a 1:1 line superimposed over the data.

bias of only 0.67 ppb, but the variance of  $89.17 \text{ ppb}^2$  is comparable to other error variances and is therefore included in the model.

The Bayesian inversion minimizes the cost function adapted from Rogers (2000) given by Eq. (15).

$$J(\Gamma) = (y - K\Gamma)^T S_{\epsilon}^{-1} (y - K\Gamma) + (\Gamma - \Gamma_{\text{prior}})^T S_{\text{prior}}^{-1} (\Gamma - \Gamma_{\text{prior}}) \quad (15)$$

### 2.3 Study site descriptions

Two instrumented towers of the NOAA tall tower network and several aircraft missions provide the data for testing the STILT-CO model and for deriving CO emission rates over the Midwest and Northeast of North America. We focus on data from 1 March–15 August 2004, when comprehensive atmospheric observations are available. The 107 m cell phone tower at Argyle, Maine, just north of Bangor ( $45.03^{\circ}$  N,  $68.68^{\circ}$  W), was the anchor ground station of the CO<sub>2</sub> Boundary-layer Regional Airborne Experiment in Maine (COBRA-2004), an extensive measurement program using the University of Wyoming King Air platform (NOAA ESRL-GMD, 2007). Modeling carbon monoxide at the Argyle tower allows for direct comparison with a substantial body of previous work on CO<sub>2</sub> modeling at that site. WLEF, a 450 m tall TV tower near Park Falls in northern Wisconsin ( $45.93^{\circ}$  N,  $90.27^{\circ}$  W) (Bakwin et al., 1998), provides a second important site for assessing the STILT-CO model and the



a priori emission inventories. Because WLEF lies in the middle of the continent, it sees very different synoptic transport patterns and emission sources than the coastal region near Argyle. WLEF also received significant CO emissions from biomass burning in both northern Canada and in the southeastern US during the study period, summer 2004 (Turquety et al., 2007).

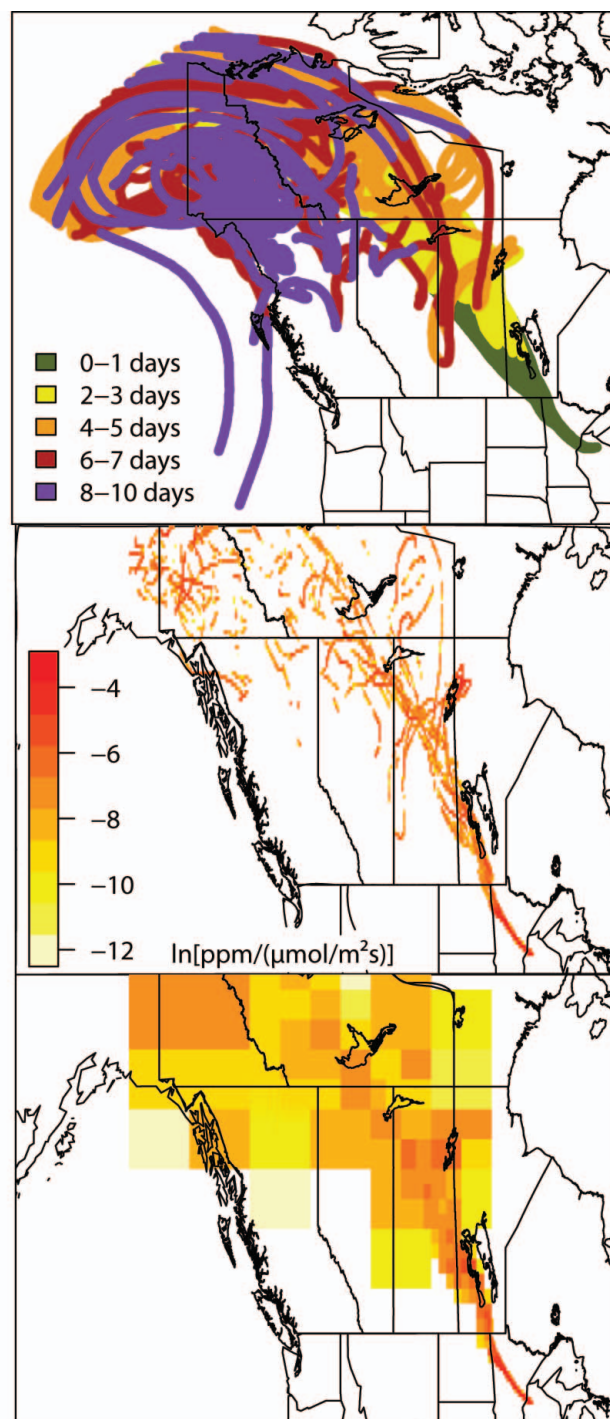
COBRA-2004 aircraft missions, originating at Bangor, ME, complement the tower data. There were 59 flights during the summer of 2004 with over 900 vertical measurement profiles recording CO concentrations at 1 Hz (Lin et al., 2006; Matross et al., 2006). We also use aircraft data on temperature and water vapor to aid in assessing the model boundary condition through comparisons with free tropospheric measurements. Aircraft flights in the Intercontinental Chemical Transport Experiment (INTEX-A) also measured carbon monoxide at 1 Hz, along with formaldehyde, and other trace gas concentrations in the troposphere over the continental United States from 1 July to 15 August 2004 (Singh et al., 2006). We examine here INTEX-A vertical measurement profiles within 1000 km of WLEF tower in order to evaluate the model's ability to simulate the transport of CO from surface sources to altitude. Two different instruments aboard the DC8 aircraft reported HCHO data during INTEX-A: from the National Center for Atmospheric Research (NCAR) and from University of Rhode Island (URI). They disagree by 30%, apparently reflecting a difference in calibration (Heikes et al., 2001; Wert et al., 2003; Roller et al., 2006). This difference creates challenges in validating model results, as discussed in Sect. 3.4.

### 3 Results and discussion

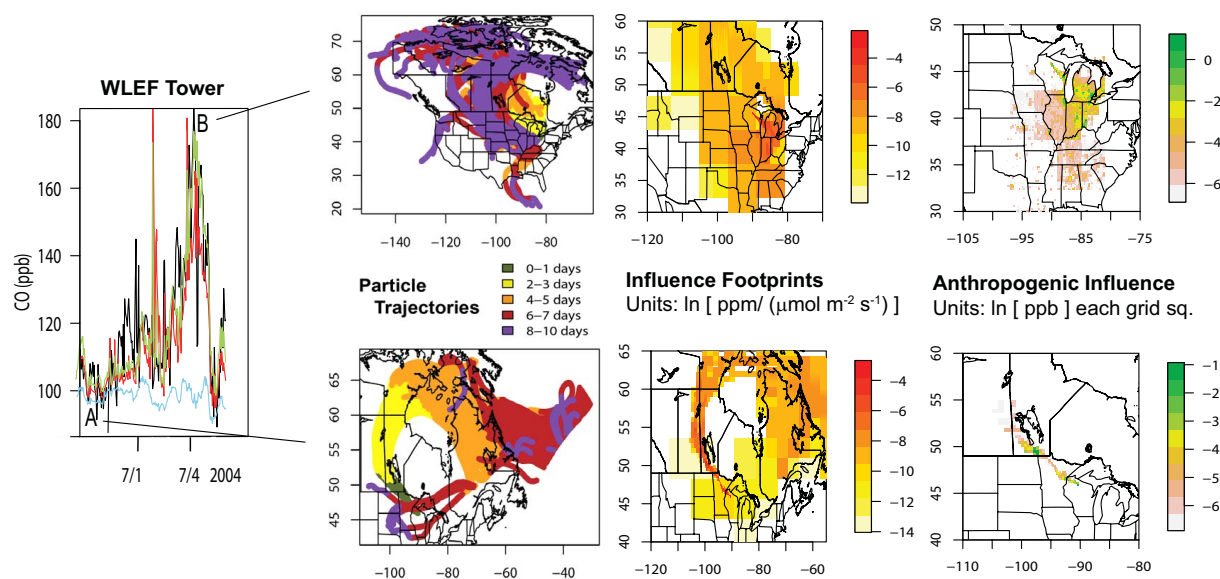
#### 3.1 STILT-CO model characteristics

Particles traveling ten days backward in time from WLEF may reach as far as eastern Russia. Figure 5 shows sample particle trajectories for midday on 18 August 2004. The top panel of the figure displays the model influence function color coded by the time since the particles left the tower. The middle panel shows the footprint influencing the WLEF tower at this time, using the full resolution within the model the entire way back in time, whereas the bottom third of the figure shows the influence footprint after dynamically aggregating surface sources and particle locations far from the receptor. With 100 particles sent out from the receptor point, the influence function calculated from an individual particle is disjoint, but a smooth pattern emerges once STILT aggregates surface fluxes and footprints.

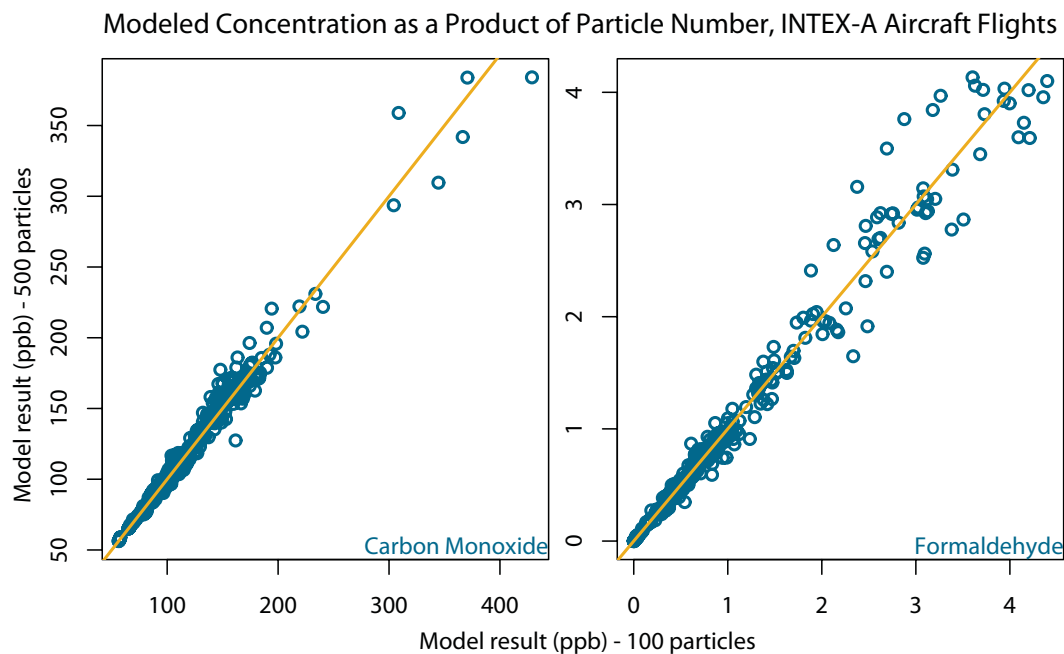
The model multiplies the surface source influence by the surface source inventories, and the sum of the influence-weighted surface fluxes is incrementally added to the advected boundary value to obtain the model concentration at the receptor point, as illustrated in Fig. 6.



**Fig. 5.** An example of STILT-CO particle trajectories from 18 August 2004. The top panel of the figure shows the particles traveling backward in time away from the WLEF tower. The particles are color-coded by time away from the tower. The middle panel shows the logarithmic influence footprint in units  $\ln(\text{ppm}/(\mu\text{mole m}^{-2} \text{s}^{-1}))$  with the maximum resolution while the bottom panel displays the logarithmic influence footprint with dynamic gridding.



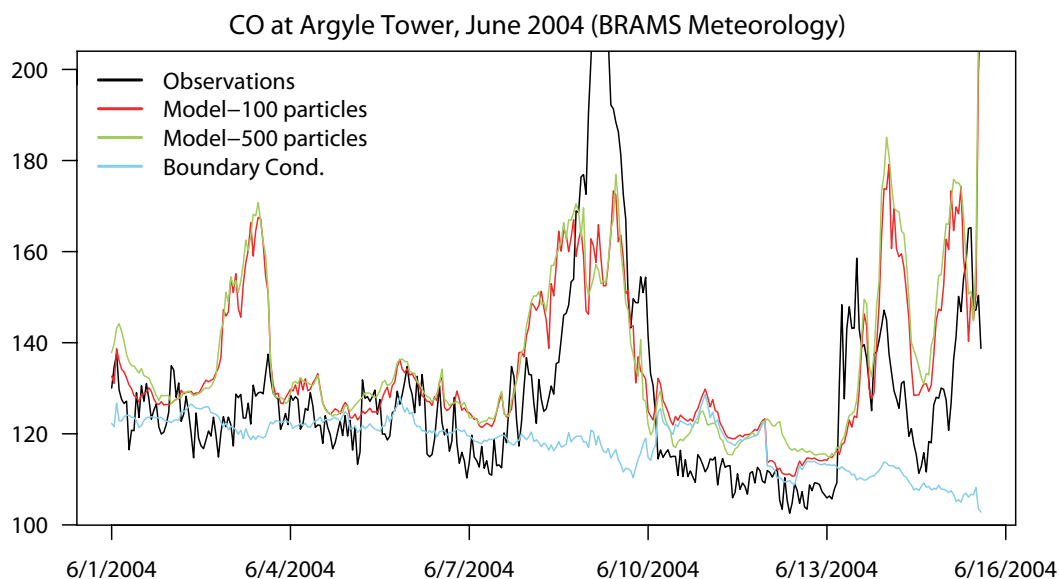
**Fig. 6.** An illustration of how the particle trajectories and influence footprint work in tandem to produce modeled hourly concentrations. Particle trajectories are color coded by time away from the tower. The influence footprint is multiplied by the emissions inventories to produce the enhancement maps. The enhancement contributions are then summed over the entire modeled domain to produce modeled hourly concentrations. Period A illustrates a time of low concentrations whereas period B is highly influenced by urban areas such as Detroit and Chicago.



**Fig. 7.** Model results for the INTEX-A aircraft missions as a product of the number of particles used in model simulations.

Gerbig et al. (2003) found a typical standard deviation of 13% in the  $\text{CO}_2$  surface source signal, due to statistical fluctuation associated with the use of a finite ensemble of 100 particles. Since sources tend to be more spatially concentrated for CO than for  $\text{CO}_2$ , we tested model simulations

using both 100 and 500 particles to determine the number required for accurate simulations. Figure 7 shows a scatter plot of simulated CO and formaldehyde concentrations for INTEX-A aircraft flights using both 100 and 500 particles. At low trace gas concentrations, both plots show relatively



**Fig. 8.** The STILT-CO model result at Argyle with both 100 and 500 particles.

little scatter. These points represent model results at high altitudes with small surface flux influence, and particle number makes little difference. Model results for higher CO concentrations represent aircraft receptor points within the planetary boundary layer that experience significant influence from surface fluxes. These results show significantly higher variance; model particle number is associated with incomplete sampling of the surface emissions. The effect on modeled CO of increasing particle number from 100 to 500 was relatively small, however, with a mean difference of only 0.32 ppb for the INTEX-A data and standard deviation of 7.1 ppb. For HCHO, the mean difference and standard deviation were 0.03 ppb and 0.15 ppb, respectively. In addition, we ran 500 particle simulations at Argyle tower but found no improvement in model-measurement fit over 100 particle simulations (see Fig. 8). Since the associated statistical variance is much smaller than other sources of error, the marginal improvement in model performance did not justify the ( $5\times$ ) computational cost.

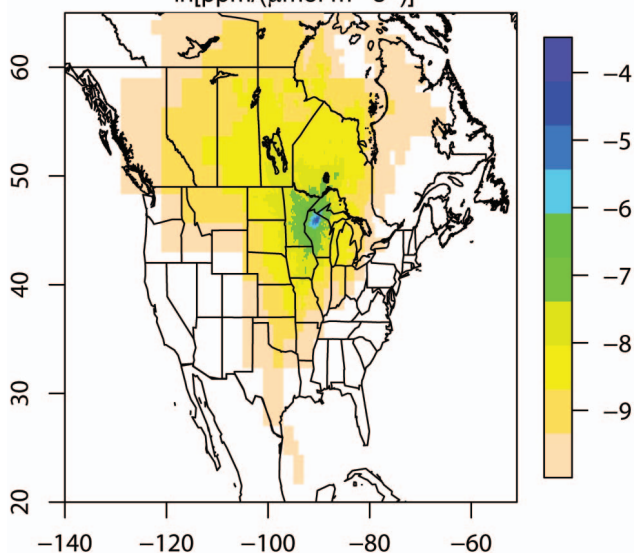
### 3.2 Regional CO sources derived from comparing model and data at a tall tower

The WLEF tower saw substantial influence from northern Canada during summer months (Fig. 9). Areas of Nunavut and Northwest Territories in Canada exerted as much influence on WLEF tower data as air from Indianapolis and Detroit, even though the tower is over a thousand kilometers closer to these American cities.

The a priori model systematically overestimates CO concentrations at WLEF compared to measurements (Fig. 10). The model time series for both EPA NAPAP 1993 and EPA NEI-1999 show pollution-related peaks well-correlated with

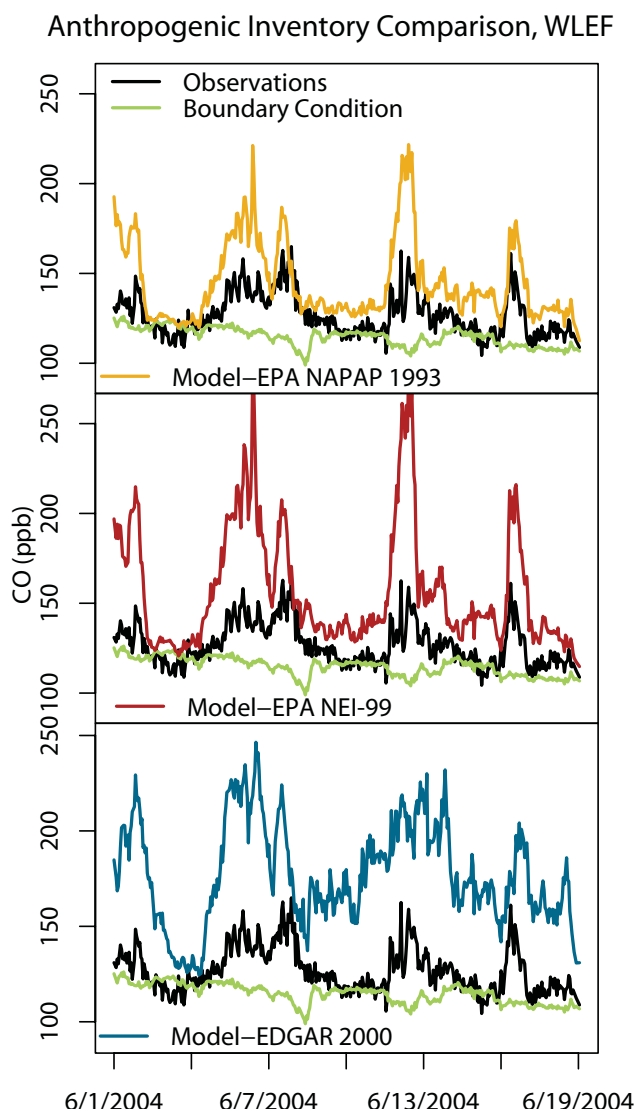
### Mean Logarithmic Footprint, WLEF, Summer 2004

$\ln[\text{ppm}/(\mu\text{mol m}^{-2} \text{s}^{-1})]$



**Fig. 9.** The mean influence footprint in units  $\ln(\text{ppm}/(\mu\text{mol m}^{-2} \text{s}^{-1}))$  for the WLEF tower averaged over the summer of 2004.

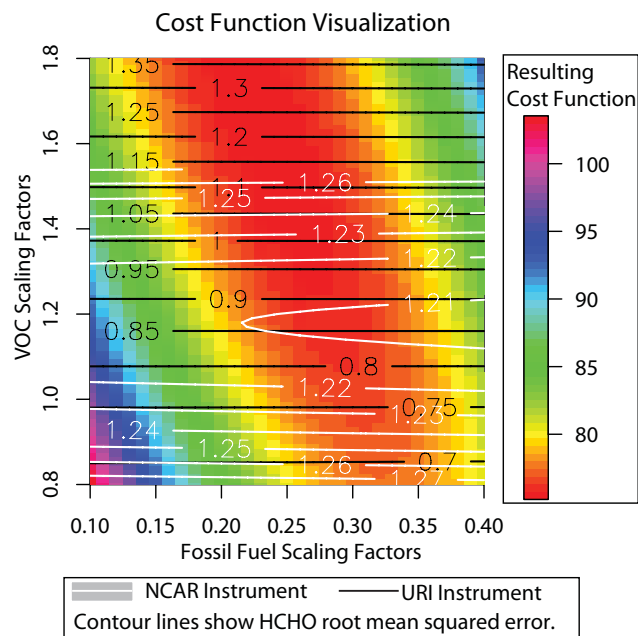
observations, indicating good spatial accuracy for the assumed emissions, but the magnitudes of pollution peaks are far too large. The bias appears therefore to be directly attributable to errors in the magnitude of the fossil fuel emissions from the inventories. Model results with EDGAR-2000 reveal large inaccuracies in the inventory both spatially and in terms of total emissions, with peaks and troughs appearing where none exist in the tower measurements (see Fig. 10).



**Fig. 10.** A comparison of CO concentrations computed by several a priori anthropogenic CO emissions inventories. All inventories overestimate anthropogenic CO concentrations.

For the first 20 days of June, we fit each fossil fuel inventory along with MEGAN VOC sources using a simple least squares model. The EPA NEI-1999 and EPA NAPAP 1993 showed the best fit ( $r=0.73$  and  $r=0.72$ , respectively), while EDGAR-2000 showed a substantially lower fit ( $r=0.62$ ). Evidently EDGAR-2000 does not represent a good prior for analysis of CO emissions in this region, and we therefore use the most recent EPA-1999 inventory as our prior (results are similar using EPA-1993).

We conducted separate model inversions for the WLEF tower in spring and summer and produced posteriori scaling factors ( $\lambda$ ) simultaneously for three factors:  $\lambda_{ff}$  (anthropogenic CO emissions),  $\lambda_{bb}$  (biomass burning emissions, prior from Wiedinmyer et al. (2006)), and  $\lambda_{VOC}$  (prior from

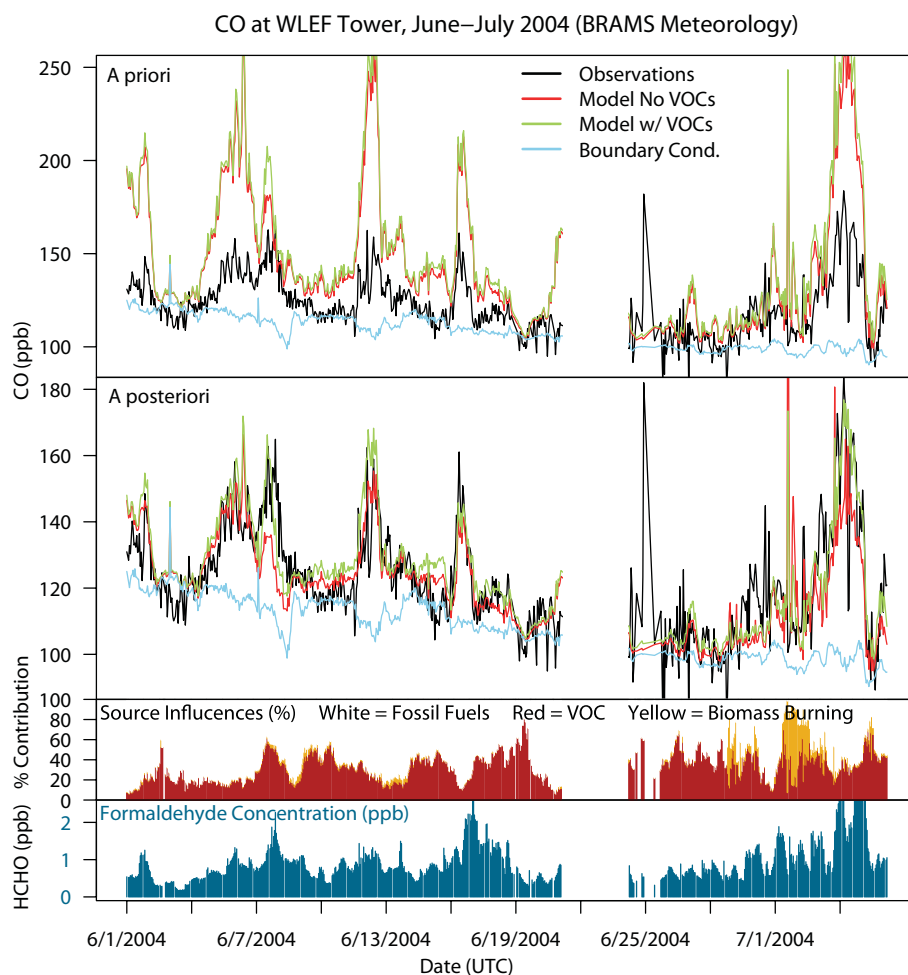


**Fig. 11.** A visualization of the Bayesian inversion cost function at WLEF for the summer months. The fossil fuel and VOC scaling factors are set to discrete values. The biomass burning scaling factor is allowed to float with the inversion. The surface plot shows that there is no clear single optimum in the Bayesian Inversion. The contour lines show the RMSE of model results for INTEX-A formaldehyde aircraft profiles near the WLEF tower. Based on this plot, we choose a final fossil fuel scaling factor of 0.3 and a VOC factor of 1.2.

the MEGAN inventory). For early summer (1 June–23 July), prior to the arrival of large signals from boreal biomass fires, the optimal values for  $\lambda_{ff}$ ,  $\lambda_{bb}$ , and  $\lambda_{VOC}$  were  $0.24 \pm 0.07$ ,  $0.50 \pm 0.30$ ,  $1.57 \pm 0.52$ , respectively. (Note, that value of  $\lambda_{VOC}$  is strictly a constraint on  $(\text{MEGAN fluxes}) \times (\text{CO yield})$ , not on VOC fluxes themselves.)

The scaling factor for VOC emissions in summer (1 June–23 July) was highly correlated with fossil fuel emissions ( $r=0.81$ ), as illustrated in Fig. 11. For this figure, we set fossil fuel scaling factors (x-axis) and VOC scaling factors (y-axis) at prescribed values, and optimized only for biomass burning scaling factors. The cost function has a narrow valley: the minimum fell along a line given by  $\lambda_{ff} = -0.11(\lambda_{VOC}) + 0.41$ .

Formaldehyde data, available for INTEX-A flights, provide an additional constraint on the inversion to help determine optimal scaling factors for both VOCs and fossil fuel CO emissions. The contour lines overlaying the surface plot show the root mean squared error of the formaldehyde model result for INTEX-A vertical aircraft profiles with the prescribed set of CO and VOC scaling factors (see discussion in Sect. 3.4, below). Since HCHO emissions from fossil sources are small, model results are almost independent of



**Fig. 12.** Hourly model results from WLEF during June 2004. The a priori results are shown on top followed by the a posteriori results below. The bottom two panels respectively display the relative importance of different CO sources and the corresponding formaldehyde model results during the period.

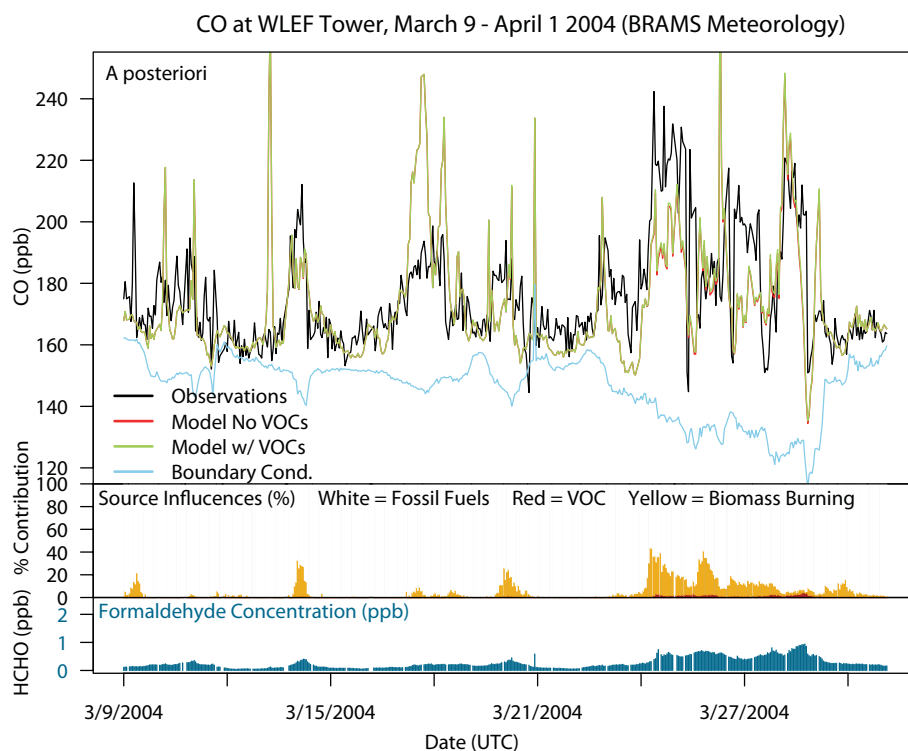
the scaling factor for fossil fuels. We can therefore find optimum scaling factors on the cost function minimum that also minimize RMSE for the HCHO model for INTEX-A flights near the WLEF tower.

Due to the difference in calibrations for URI and NCAR data, we obtain two different optimal scaling factors (0.65 or 1.2, respectively) for the effective CO sources from VOCs as indicated by fidelity with INTEX-A HCHO data. The corresponding optimal fossil fuel scaling factor changes relatively little (0.3 and 0.34, respectively) for the NCAR or URI calibration. If we use the NCAR results, the minimum cost function lies very close to the global minimum for the three-factor optimization on CO data from WLEF, whereas the lower VOC sources implied by the URI give results that are less consistent. By this measure the URI calibration appears to be too low. Additionally, GEOS-Chem simulations from Millet et al. (2006) capture 70% of the variability in NCAR formaldehyde measurements but capture only 42%

of the variability in URI measurements (along with a 34% bias). We therefore use NCAR measurements for the model optimization.

Table 1 lists the resulting optimization factors, and Table 2 lists indicators for the resulting model-measurement fit. The inversion reduces the cost function at WLEF (summer) from 353.1 to 76.1 (a 78.5% reduction). The upper two panels of Fig. 12 display the a priori and a posteriori time series at WLEF for June and early July, before the advent of high levels of CO from boreal fires (third panel). The bottom panel of Fig. 12 shows the model-generated concentrations of HCHO at WLEF, showing that measurements of formaldehyde at this site would be very effective in distinguishing CO from fossil fuels versus VOCs.

Our estimated VOC scaling factor (1.2) can be compared to the results of Palmer et al. (2006), who found that during early summer months, MEGAN estimates of VOC emissions were 10% lower than GOME satellite-derived estimates (i.e.,



**Fig. 13.** Hourly WLEF model results during the spring of 2004.

**Table 1.** STILT-CO posterior scaling factors (including 95% confidence intervals). We use the same scaling factors at WLEF (summer) for the posterior model at Argyle. The inversion does not produce a reliable scaling factor for biomass burning at WLEF during summer months, so we have omitted the value below.

Tower	$\lambda_{ff}$	$\lambda_{voc}$	$\lambda_{bb}$
Wlef summer	$0.3 \pm 0.05$	$1.2 \pm 0.4$	NA
Wlef winter	$0.55 \pm 0.05$	$1.01 \pm 0.24$	$0.47 \pm 0.28$

scaling factor of 1.1). Our estimate of  $\lambda_{ff}$  also corresponds roughly to the scaling factor  $\lambda_{ff}=0.4$  derived independently by Hudman et al. (2008) using INTEX-A data for CO, which we did not use in our optimization.

During the spring, CO emissions from fossil fuel combustion are expected to be higher than in summer, while sources from VOCs and biomass fires are lower. Table 1 summarizes the posterior scaling factors (see Fig. 13 for the a posteriori time series). The results suggest notably stronger seasonal variations than adopted in the inventories (see below). The inversion reduces the cost function from 246.7 to 97.4 (60.5% reduction), with a correlation coefficient ( $r$ ) of 0.57 (see Table 2). Since fossil fuel emissions are the dominant source during this period, constraints are strongest for the a posteriori fossil fuel scaling factor ( $\lambda_{ff}$ ). There were some biomass fires to the south of the site, only a few hundred

**Table 2.** Summary of STILT posterior model performance for CO.

Tower	Prior RMSE	Posterior RMSE	$r$
Wlef summer	34.4	10.4	0.81
Wlef winter	35.2	16.2	0.57
Argyle summer	61.3	22.9	0.40

km away, and these sources are moderately well constrained. Figure 14 provides a contour plot of the cost function for the spring inversion. The plot for the spring months shows a clearer minimum than for summer, albeit with a fairly large range for  $\lambda_{voc}$ .

The NEI-99 inventory data is only available for typical mid-week summer days and typical mid-week winter days. We use the summer inventory as the a priori for all simulations. We note that our scaling factor for the NEI-1999 inventory for simulations in the summer (1 June–15 August) makes a much larger reduction than the scaling factor for simulations in the spring months (1 March–30 April). During colder months in the upper Midwest, CO emissions could be higher because of less efficient combustion from mobile sources, plus sources from home heating using wood fuel. In contrast to our model results, the NEI-99 winter inventory predicts that total national CO emissions will be slightly lower during winter months than during the summer. On-road sources are predicted to be 5% higher during winter

months, and area sources such as home fuel burning are about three times higher during the winter. But non-road sources such as tractors and construction equipment are estimated to be 97% higher during summer months (Frost and McKeen, 2007). Our results suggest that anthropogenic CO emissions are higher in spring months than summer months. We therefore cast doubt on the seasonal adjustments used in NEI-1999. The EPA-1999 inventory might overestimate the relative increase in non-road sources from winter to summer (i.e. tractors, diesel from construction, etc.) and/or underestimate the relative increase in area sources (home heating, fire places, etc.) or increases in CO emissions from power generation from summer to winter.

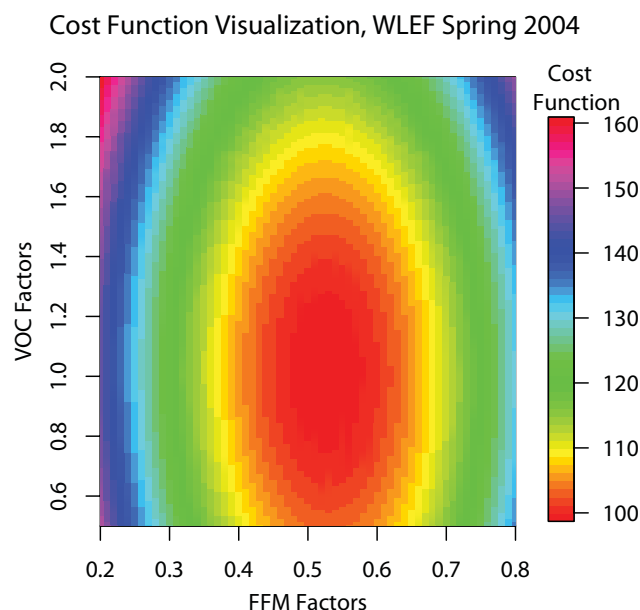
The results from WLEF during the spring months could reflect regional differences in the seasonal variability of emissions - because the upper Midwest has particularly cold winter and spring months, and more use of wood fuel than other regions (Fernandes et al., 2007). A study by Meszaros et al. (2004) found seasonal adjustments in European CO emissions that predicted 10% higher emissions during spring than during summer. Europe is not as cold as the region around WLEF, so the seasonal trend might be expected to be even larger in Wisconsin.

Figure 15 shows scatter plots of hourly data (model vs observed) for WLEF in summer and spring and for Argyle, ME, in summer (using the a posteriori scaling from WLEF). As discussed below, model results at Argyle tower are significantly affected by problems in modeling transport near the coast, and INTEX-A model results appear to be significantly influenced by errors associated with the boundary condition. Below we examine model results for these data sets to understand the factors that limit model performance in order to guide future model development and to help design strategies for future observing programs.

### 3.3 Biomass burning and STILT-CO

In general, STILT-CO appeared to do well at capturing the influence of CO from biomass burning emissions in the near field, but it was inconsistent in capturing emissions influence from very large fires that were far away. The time series from WLEF tower in spring 2004 (Fig. 13) shows that even during the spring months, biomass burning can substantially influence pollution levels at the tower site. The influence of biomass burning events in Missouri and Arkansas were accurately characterized by STILT-CO during this time period.

Figure 16, an example from WLEF tower during August 2004, displays an example where STILT-CO did provide a very detailed, high resolution prediction of distant CO source regions. The particle trajectories left the WLEF tower, traveled backward in time toward northern Canada, and intersected large forest fires near Great Slave Lake, in the Yukon Territories, and in eastern Alaska. The pollution influence was modestly overestimated. Figure 17 displays a time series from the WLEF tower during the latter half of summer 2004



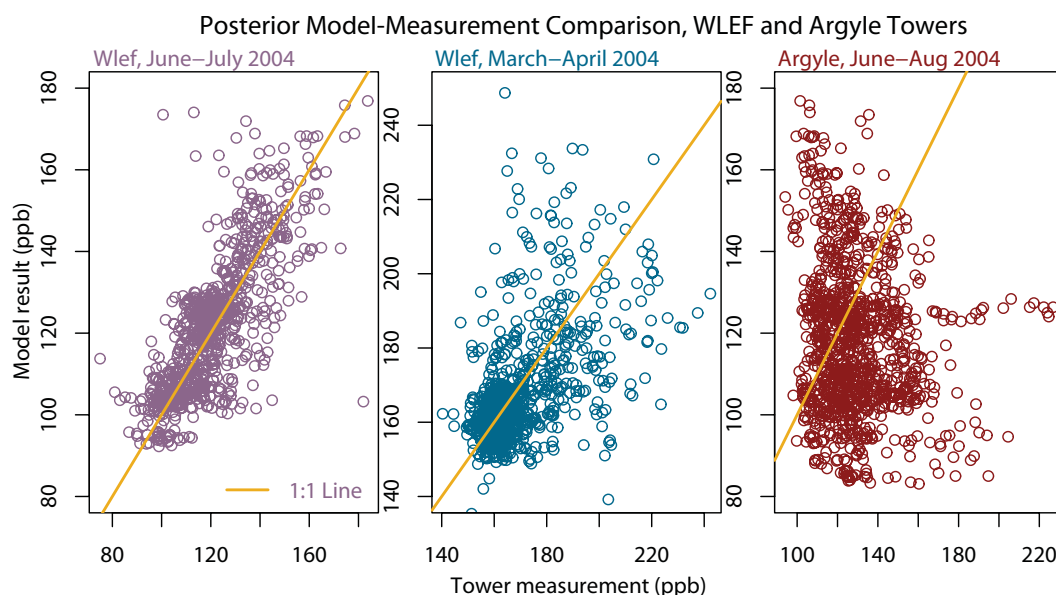
**Fig. 14.** A visualization of the Bayesian inversion cost function for WLEF during March–April. The plot is constructed in the same way as Fig. 11. Unlike the summer model results, which show no clear minimum in the cost function, the spring 2004 model results show a much clearer optimum.

when WLEF experienced significant pollution from forest fires in Alaska and northern Canada. Most of the time, the model did find influence from these forest fire influences but incorrectly computed the magnitude of this influence.

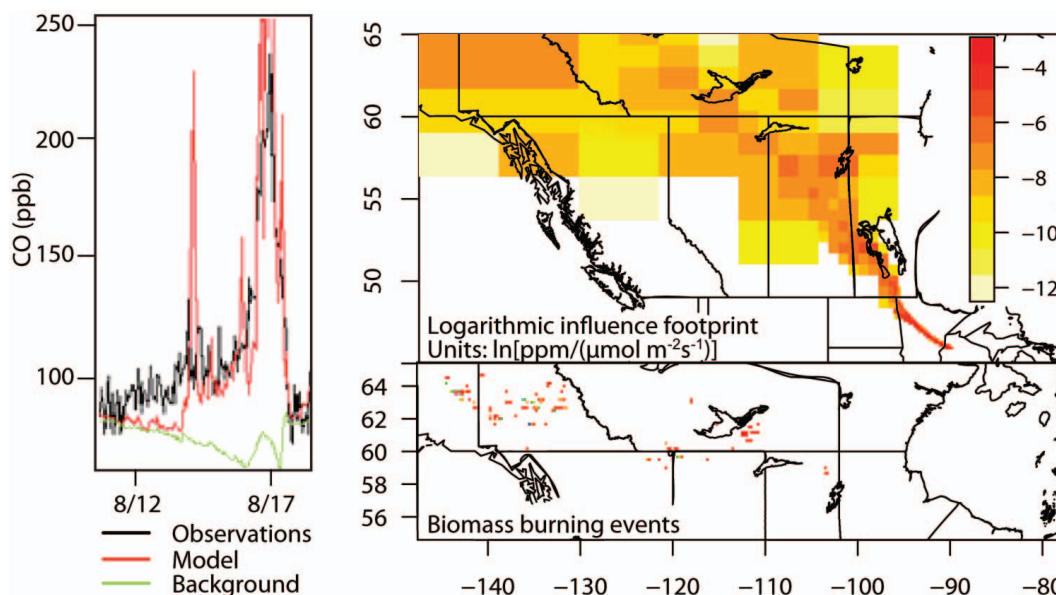
Issues affecting very large biomass sources very far away represent a major challenge to any modeling framework, and their resolution lies beyond the scope of this paper. The lack of pyro-convective injection in the model might, in part, account for why the model performed very well on relatively small fires in the near field but showed mixed performance in capturing the influence of very large fires at a long distance. Time periods affected by long distance biomass burning emissions are not included in the assessment of source inventories.

### 3.4 CO concentration trends with altitude: insights from the INTEX-A aircraft campaign

We noted above that formaldehyde data provide a potentially powerful way for independently constraining the influence of summertime emissions of VOCs on CO. The STILT-CO model using the MEGAN inventory and the HCHO yields from Palmer et al. (2006) captures measured HCHO for many INTEX-A vertical aircraft profiles in the US continental interior during summer 2004 (see a posteriori results in Fig. 18), although the results have a fairly large variance (Fig. 19). The vertical profiles of HCHO could validate model chemistry and provide a confirmation of inverse



**Fig. 15.** A comparison of model and measurements at Argyle and WLEF towers. The WLEF summer and Argyle summer plots use a posteriori scaling factors from the inversion at WLEF in June and July. WLEF from spring months is inverted separately.



**Fig. 16.** An example of the influence footprint (top) and biomass burning inventory (bottom) on 17 August 2004 - a period when biomass burning significantly influences pollution levels at the WLEF tower.

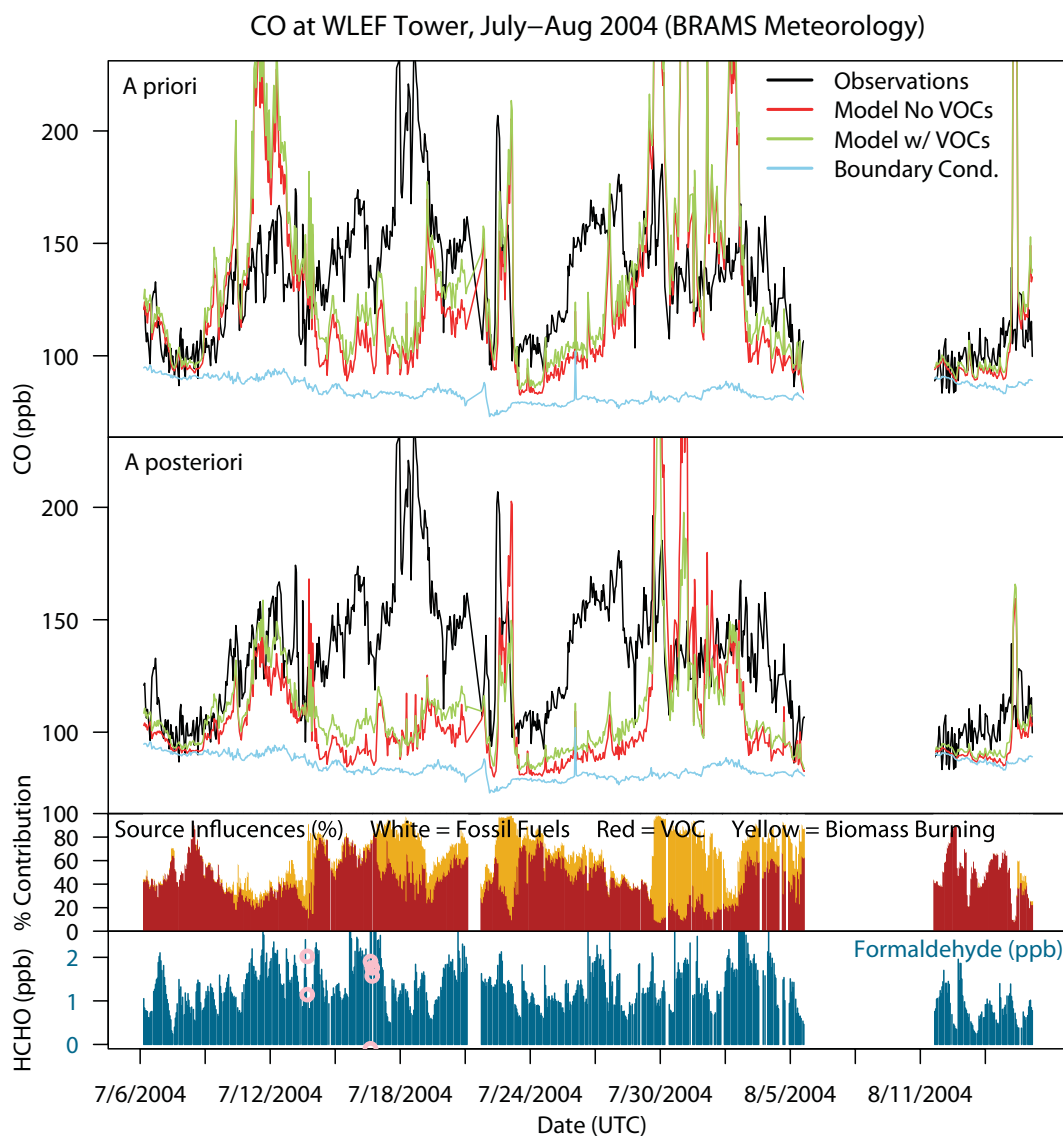
model results independent of fossil fuel and biomass burning CO influence, if the calibration difference could be resolved. INTEX-A vertical profiles of CO confirm that the STILT-CO model replicates CO measurements at aircraft receptor points as well as at tall tower sites, using our optimized values  $\lambda_{\text{VOC}}=1.2$  and  $\lambda_{\text{ff}}=0.3$ , although background values in the model appear to be about 20 ppb too low (Fig. 19).

### 3.5 Model limitations: case studies from Argyle tower, Maine

#### 3.5.1 Coastal meteorology

The correlation between model and measurement is generally much lower at Argyle than at WLEF (RMSE=22.9 ppb,  $r=0.40$ ; Fig. 15). Figure 20 displays a time series of STILT-CO results from Argyle during the summer of 2004. The





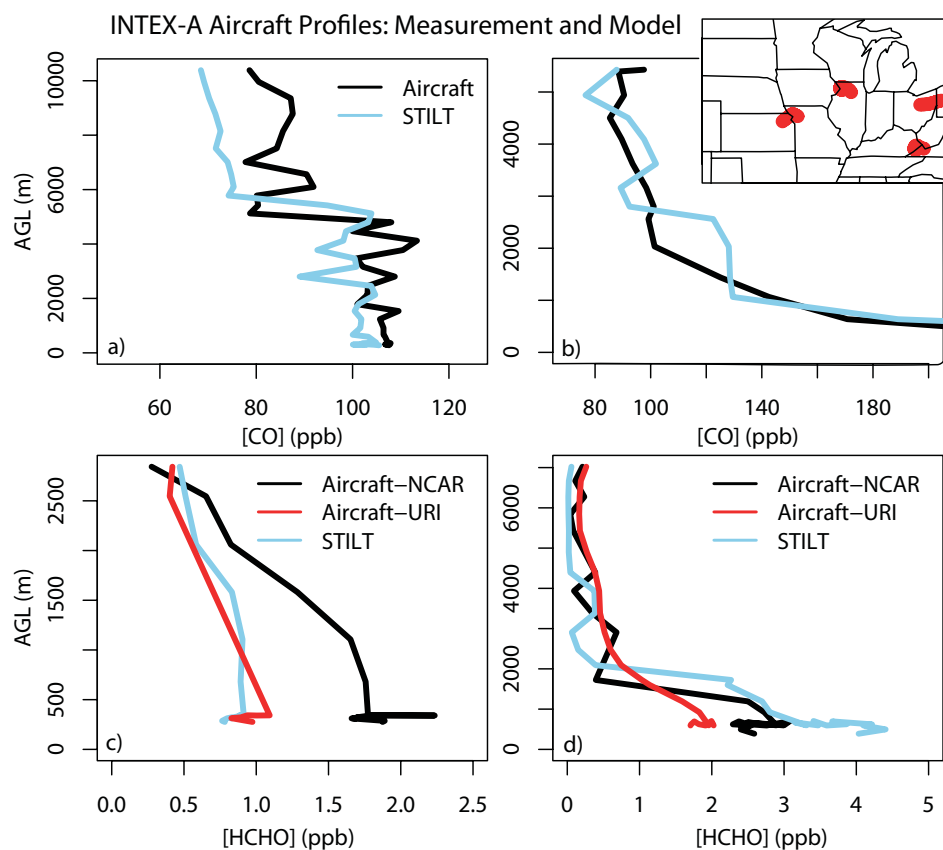
**Fig. 17.** Hourly modeled result at WLEF tower during July and August 2004, a period that saw significant pollution influence from biomass burning events in Alaska and northern Canada. INTEX-A aircraft measurements of HCHO taken near the WLEF tower are shown as pink dots.

model missed pollution peaks much more frequently than at WLEF. For example, on 14 July modeled air parcels became caught in low pressure front and just missed urban coastal sources, whereas the observations indicate strong influence from those sources (see Fig. 21). The model also created pollution peaks that do not exist, such as on 3 June when parcels traveled along an occluded front and pushed too close to coastal urban sources, sources that apparently did not influence Argyle at that time.

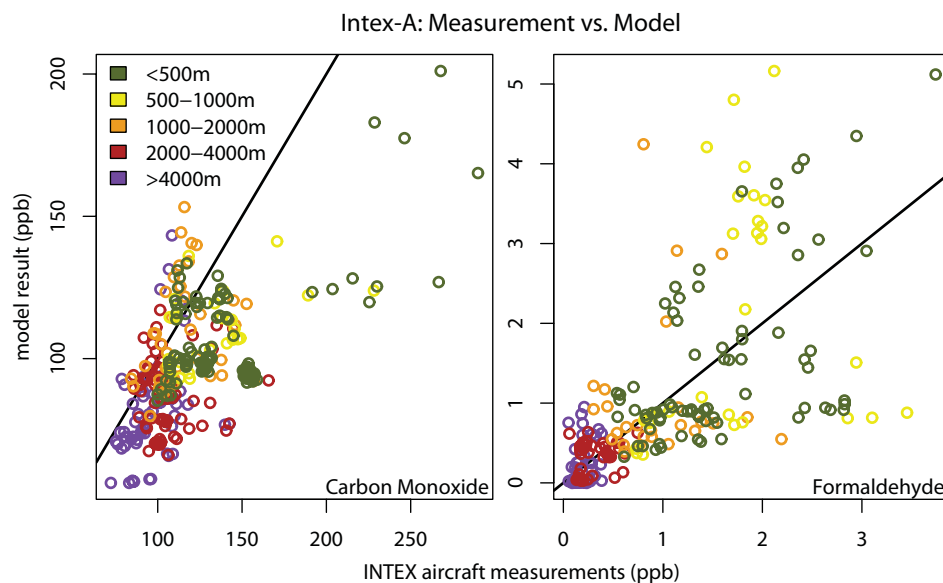
The BRAMS assimilated meteorological driver has a 45-km resolution. Many large sources affecting Argyle lie right on the coast, and Argyle itself lies within one grid square of the coast. We infer that our 45-km meteorological grid is not able to reliably resolve the influence of strong, very compact

sources that lie on the land-ocean boundary. BRAMS might also inaccurately simulate the PBL near coastal and ocean areas. Over the summer of 2004, an average 54% of particles in each ensemble traveled into the coastal domain for at least a portion of the particle trajectory, a feature that makes the receptor at Argyle more difficult to model than WLEF.

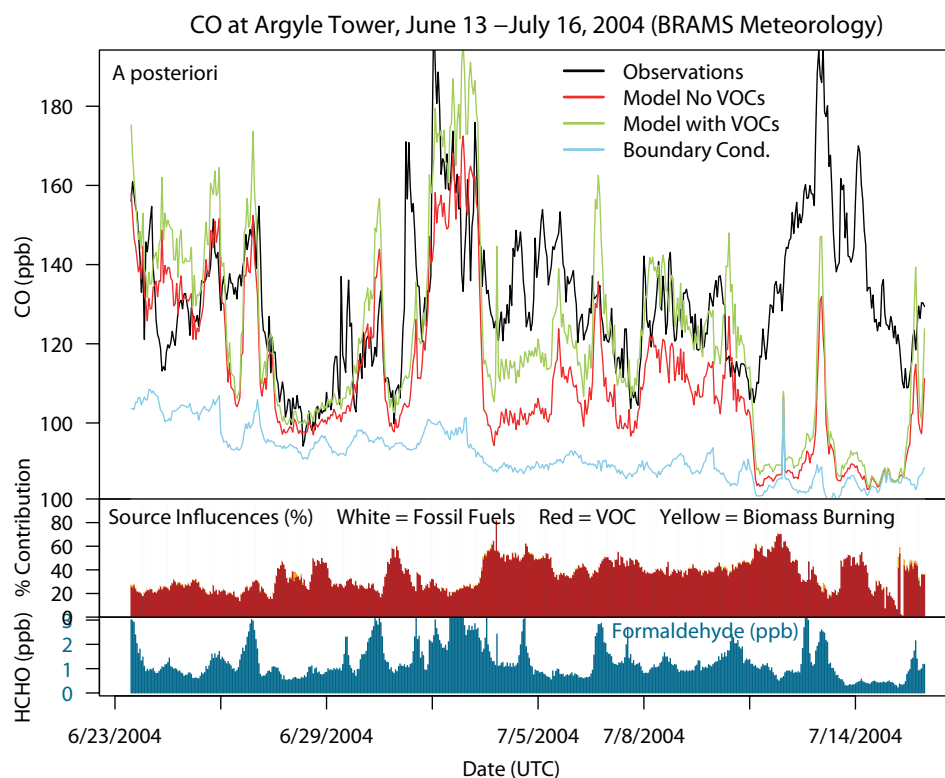
The WLEF tower lies close to large water bodies such as Lake Superior and Lake Michigan. These lakes are large enough to generate land/water mesoscale circulations, but they are much smaller than synoptic scales and exert much less influence on synoptic meteorology than the ocean. Lake Superior also lacks large anthropogenic CO sources on the coastline. Model results for COBRA-2004 aircraft flights near the New England coast (not shown) similarly show



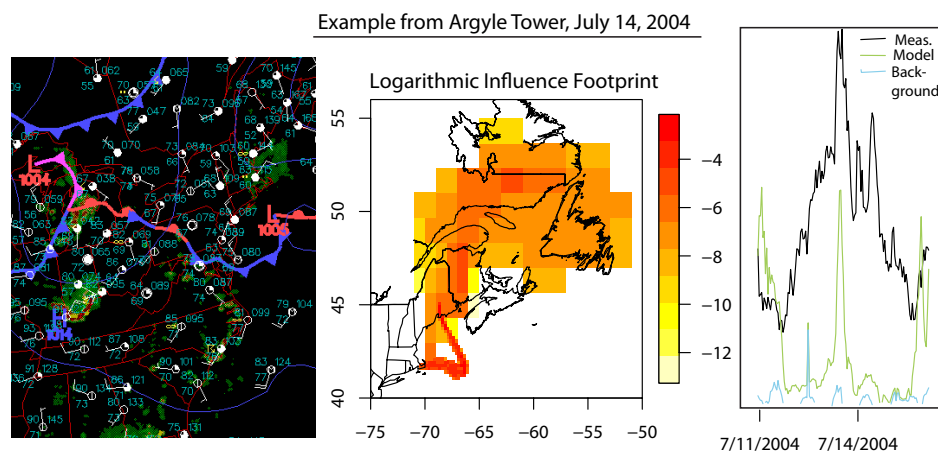
**Fig. 18.** Model results for INTEX-A vertical measurement profiles taken over the Midwestern US during the summer of 2004. The map displays the locations of the different profiles. The top two profiles show CO model results from (a) 8 July and (b) 10 July. Profiles (c) and (d) show formaldehyde model results from 6 July and 11 August, respectively.



**Fig. 19.** A comparison of INTEX-A model results and aircraft measurements. Results are color coded by height above ground level (in meters). The black lines are 1:1 lines.



**Fig. 20.** Hourly model results from Argyle tower, Maine, during the summer of 2004.



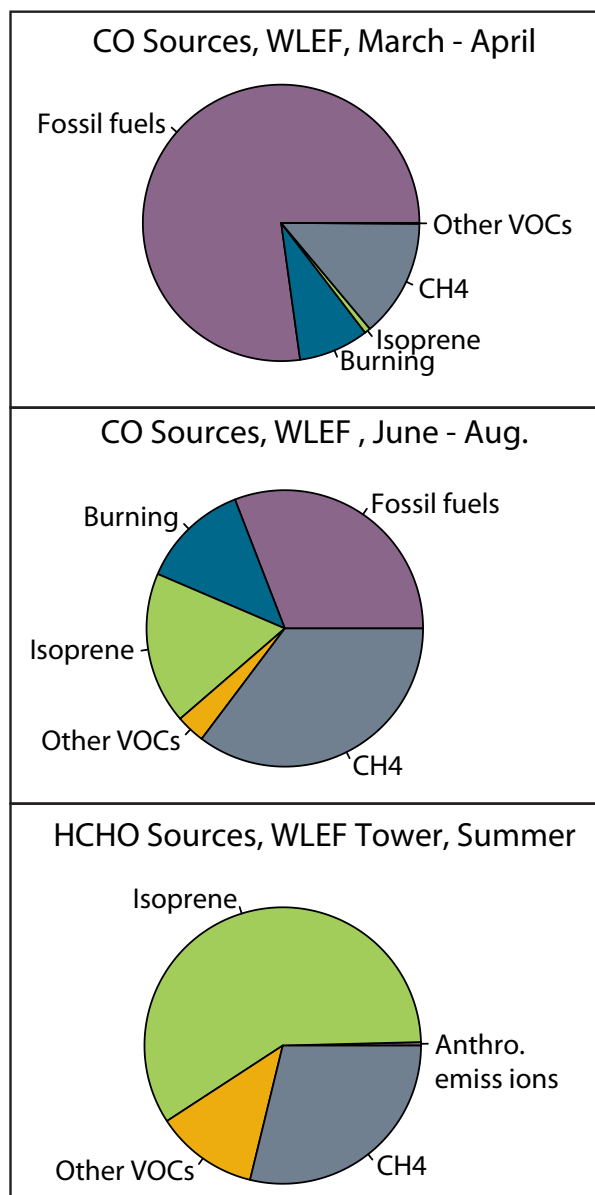
**Fig. 21.** An example of problematic meteorological transport from 14 July 2004. The modeled trajectories presumably miss major coastal sources, and the model underestimates concentrations at the receptor. (Meteorological map from Unisys, 2008).

lower model-measurement fit than INTEX-A aircraft profiles taken in the continental interior.

### 3.5.2 Advected boundary condition

Previous Lagrangian models of carbon monoxide have used a constant boundary condition (e.g., Warneke et al., 2006). However, at the WLEF tower, the advected boundary condi-

tion varies by as much as 40 ppb over 10 days (see Fig. 13), suggesting the variability in the boundary condition may contribute significantly to variability in observed CO. Washenfelder et al. (2006) drew a similar conclusion for CO<sub>2</sub>. Extensive high-altitude measurements (5000–7000 m) from COBRA-Maine airborne near the Argyle tower during the summer of 2004 provide an excellent opportunity to examine the model treatment of the lateral tracer boundary



**Fig. 22.** The relative importance of different CO and HCHO sources at WLEF tower during both spring and summer months.

condition in STILT-CO for receptors at the eastern edge of North America (Lin et al., 2006). During times with relatively little biomass burning influence, these altitudes provide free troposphere trace gas concentrations that often experience little influence from anthropogenic and vegetation fluxes of CO and CO<sub>2</sub>. Measurements therefore approximate the lateral tracer boundary condition, within a few ppb for CO and a few ppm for CO<sub>2</sub> (Matross et al., 2006).

Comparison between modeled and measured CO in the free troposphere near Argyle tower shows substantial scatter (not shown), but in general, the model appears to moderately underestimate the advected boundary concentrations of CO

– by  $19.5 \text{ ppb} \pm 28.1 \text{ ppb}$ . The notable variability in the free troposphere measurements indicates that a variable boundary condition is important. The present version of STILT-CO uses a model boundary condition applied at the western edge of the modeled domain because neither measurement stations nor global model results were available to define concentrations at other boundaries. This rough approximation affects Argyle more than WLEF: as the particles traveled backward in time from the Argyle tower during the summer of 2004, 10.5% of particles exited the domain to the North, 3% exited to the East, and 86% remained in the modeled domain. Hence, East Coast receptors need to be modeled using a different approach for the boundaries, as has been done for CO<sub>2</sub> by using Carbon Tracker concentrations (Peters et al., 2007; Matross, 2006). Global Eulerian models such as GEOS-Chem may provide another option for boundary conditions in Lagrangian modeling.

### 3.6 Relative importance of combustion CO and formaldehyde sources

According to our analysis, at WLEF during the summer of 2004, anthropogenic emissions accounted for only 31% of CO contributions to observation concentrations, biogenic VOCs contributed 21.2%, methane decomposition accounted for 35.3%, and biomass burning for 12.7% (see Fig. 22). In the formaldehyde model, primary emission sources accounted for only 0.4% of modeled atmospheric HCHO while VOC and methane decomposition accounted for 70.8% and 28.8% of advected HCHO respectively.

The regional influence of VOCs as determined by the STILT-CO model is generally consistent with recent literature. Granier et al. (2000) found that VOCs contributed 21% of the global CO burden over the course of the a year, and Holloway et al. (2000) found biogenic VOCs to be 27% of the global CO source. Hudman et al. (2008) estimated that biogenic VOCs contributed 56% of the CO source over the continental US during the summer. We found a lower relative VOC contribution likely because our study sites did not experience significant influence of VOCs from high biogenic source regions such as the American Southeast. Hudman et al. (2008) also estimated a higher VOC contribution because they did not include the methane CO source in calculating the relative significance of different sources. STILT-CO gives a regional perspective for CO sources in the central US, complementing, at very high resolution, earlier attempts to refine CO sources (e.g., Palmer et al., 2006; Parrish, 2006; Warneke et al., 2006; Hudman et al., 2008). The results show that VOC and CH<sub>4</sub> sources of CO significantly exceed anthropogenic CO emissions during summer months, even in areas of relatively lower biogenic VOC emissions such as in Wisconsin and the upper Midwest. The model reveals that estimates of anthropogenic CO sources in current inventories are too high by up to a factor of three in summer and a factor of two in spring.

## 4 Conclusions

STILT-CO produces hourly results for carbon monoxide concentrations that closely correlate with measurements at a tall tower site and from aircraft. The model performs well in the continental interior, resolving the influence of fossil fuel combustion, degradation of VOCs and CH<sub>4</sub>, and forest fires (when biomass burning events lie within a few hundred km). STILT-CO can give accurate hourly concentration levels throughout the lower troposphere at high spatial and temporal resolutions, providing tests of source inventories on scales spanning regional to continental.

Our results showed that the fossil fuel inventories were spatially accurate, at least within the fairly broad footprint of the WLEF tower, as inferred from the very good model measurement fit at WLEF during the first half of the summer, when biomass burning contributions are relatively small. When using models of trace gases with more diffuse sources, like CO<sub>2</sub> (e.g., Gerbig et al., 2003; Matross et al., 2006), it can be difficult to differentiate between model errors that are caused by trace gas sources and those caused by modeled meteorology. STILT-CO provides a diagnostic tool for models of trace gases like CO<sub>2</sub> that have more diffuse sources because it helps distinguish when model-data differences reflect errors in transport, versus errors in the underlying emission field.

Model results demonstrate that current fossil fuel emissions inventories systematically overestimate surface sources, by roughly a factor of three in summer for EPA's NEI-1999, and a factor of two in spring. The seasonal adjustment factors in NEI-1999 also appeared to be inaccurate according to our analysis, at least for the upper Midwest. VOC emissions estimates from Guenther et al. (2006) (multiplied by the HCHO yield) appeared to be reasonably accurate, as inferred from both CO and HCHO simulations.

Data from sites in the continental US showed limited ability to validate biomass burning emissions estimates because sampling errors and transport errors overshadowed errors in estimated emission rates for regions in Alaska and northern Canada, where very large sources at long distances represent major contributors.

Trace gas modeling of any kind is difficult in coastal regions like the Eastern Seaboard because most meteorological drivers do not perform well at the ocean-land interface. Especially where the sources of the modeled trace gas are diffuse, these transport problems may not be readily apparent. But for CO, with concentrated source regions on the coast, the problems are obvious. Any modeling study must approach coastal areas with caution, unless the study can assure accurate simulation of coastal meteorology.

Lateral tracer boundary conditions require close attention. A variable modeled boundary condition is crucial to modeling regardless of the spatial resolution of the model result. The boundary condition holds special importance when building models intended for examining trace gas sources on

regional scales. Without an accurate, time-varying boundary condition, a source-receptor model cannot accurately characterize emissions estimates because the modeled fossil fuel signal can be conflated with background errors.

Our study demonstrates that, from only one tall tower receptor site, the high-resolution STILT-CO model can produce very detailed information on emission sources and meteorology over a large geographic area. The model can also test and correct emissions inventories for a large portion of the continent. STILT-CO shows wide promise to model carbon monoxide and other trace gases to determine sources and sinks with accuracy and reliability, applicable to any trace gas that undergoes relatively simple chemistry during transport.

*Acknowledgements.* We thank NOAA scientists Stuart McKeen and Gregory Frost for providing the EPA NEI 1999 inventory data. Paul Novelli also provided crucial data from NOAA GMD measurement sites. We are grateful to Christine Wiedinmyer for her help with the biomass burning inventory. Without the help of Daniel Jacob, from Harvard, this project would not have been nearly as successful. We thank RCC Atlantic, Inc. and Unicel for access to the transmitter tower in Argyle, Maine and the State of Wisconsin Educational Communications Board for use of the WLEF-TV transmitter tower and facilities in Park Falls, Wisconsin. Dan Baumann and John Lee provided local technical support for the tower sites, and Peter Bakwin, Conglong Zhao and Bill Munger made substantial contributions to the tower measurement effort. Paul Novelli provided calibrated CO standards for the tower measurements.

This research would not have been possible without a Harvard University Center for the Environment Grant and the Harvard Dean's Summer Research Award. This work was supported by NSF Award ATM-0221850 (Biocomplexity in the Environment) and by grants DE-FG02-07ER64358 (US Dept. of Energy), NNX06AE83G (NASA) and NCC5-684 to Harvard University. The opinions expressed herein are not represented as expressing the views of any of the sponsoring agencies.

Edited by: L. Carpenter

## References

- Abott, D. S., Palmer, P. I., Martin, R. V., et al.: Seasonal and interannual variability of North American isoprene emissions as determined by formaldehyde column measurements from space, *Geophys. Res. Lett.*, 30(17), 1886–1889, 2003.
- Arellano, A. F. and Hess, P. G.: Sensitivity of top-down estimates of CO sources to GCTM transport, *Geophys. Res. Lett.*, 33(21), L21807, doi:10.1029/2006GL027371, 2006.
- Bakwin, P. S., Tans, P. P., Hurst, D. F., and Zhao, C. L.: Measurements of carbon dioxide on very tall towers: results from the NOAA/CMDL program, *Tellus B*, 50(5), 401–415, 1998.
- Benkovitz, C. M., Scholtz, M. T., Pacyna, J., et al.: Global gridded inventories of anthropogenic emissions of sulfur and nitrogen, *J. Geophys. Res.*, 101(D22), 29 239–29 253, 1996.
- Chang, K. H., Chen, T. F., and Huang, H. C.: Estimation of biogenic volatile organic compounds emissions in subtropical island - Taiwan, *Sci. Total Environ.*, 346, 184–199, doi:10.1016/j.scitotenv.2004.11.022, 2005.

- Crawford, J. H., Heald, C. L., Fuelberg, H. E., et al.: Relationship between measurements of pollution in the troposphere and in situ observations of CO based on a large-scale feature sampled during TRACE-P, *J. Geophys. Res.*, 109, D15S04, doi:10.1029/2003JD004308, 2004.
- Cotton, W. R., Pielke, R. A., Walko, R. L., et al.: RAMS 2001: Current status and future directions, *Meteorol. Atmos. Phys.* 82, 5–29, 2003.
- Daniel, J. S. and Solomon, S.: On the climate forcing of carbon monoxide, *J. Geophys. Res.*, 103(D11), 13 249–13 260, 1998.
- Duncan, B. N., Martin, R. V., Staudt, A. C., et al.: Interannual and seasonal variability of biomass burning emissions constrained by satellite observations, *J. Geophys. Res.*, 108(D2), 4100, doi:10.1029/2002JD002378, 2003.
- Duncan, B. N., Logan, J. A., Bey, I., et al.: The global budget of CO, 1988–1997: Source estimates and validation with a global model, *J. Geophys. Res.*, 112, D22301, doi:10.1029/2007JD008459, 2007.
- Ebel, A., Friedrich, R., and Rodhe, H. (Eds.): *Transport and Chemical Transformation of Pollutants in the Troposphere Vol 7: Tropospheric Modeling and Emission Estimation*, Springer-Verlag, New York, 1997.
- Edwards, D. P., Emmons, L. K., Hauglustaine, D. A., et al.: Observations of carbon monoxide and aerosols from the Terra satellite: Northern Hemisphere variability, *J. Geophys. Res.*, 109, D24202, doi:10.1029/2004JD004727, 2004.
- Fernandes, S. D., Trautmann, N. M., Streets, D. G., et al.: Global biofuel use, 1850–2000, *Global. Biogeochem. Cy.*, 21, GB2019, doi:10.1029/2006GB002836, 2007.
- Forster, C., Wandinger, U., Wotawa, G., et al.: Transport of boreal forest fire emissions from Canada to Europe, *J. Geophys. Res.*, 106(D19), 22 887–22 906, 2001.
- Frost, G. and McKeen, S. A.: Emissions inventory mapviewer, US Department of Commerce, National Oceanic and Atmospheric Administration, National Satellite, Data, and Information Service, National Geophysical Data Center, Boulder, CO, <http://map.ngdc.noaa.gov/website/al/emissions/viewer.htm>, 2007.
- Goldstein, A. and Galbally, I.: Known and unexplored organic constituents in the earth's atmosphere, *Environ. Sci. Technol.*, 41(5), 1514–1521, 2007.
- Gerbig, C., Lin, J. C., Wofsy, S. C., et al.: Toward constraining regional-scale fluxes of CO<sub>2</sub> with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor oriented framework, *J. Geophys. Res.*, 108(D23), 4757, doi:10.1029/2003JD003770, 2003.
- Granier, C., Gabrielle, P., Müller, J., Brasseur, G., et al.: The impact of natural and anthropogenic hydrocarbons on the tropospheric budget of carbon monoxide, *Atmos. Environ.*, 34, 5255–5270, 2000.
- Grell, G. A. and Devenyi, D.: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques, *Geophys. Res. Lett.*, 29(14), 1693, doi:10.1029/2002GL015311, 2002.
- Guenther, A., Hewitt, C. N., Erickson, D., et al.: A global model of natural volatile organic compound emissions, *J. Geophys. Res.*, 100(D5), 8873–8892, 1995.
- Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), *Atmos. Chem. Phys.*, 6, 3181–3210, 2006, <http://www.atmos-chem-phys.net/6/3181/2006/>.
- Hanna, S.: Applications in air pollution modeling, in: *Atmospheric Turbulence and Air Pollution Modeling*, edited by: Nieuwstadt, F. T. M. and van Dop, H., D. Reidel, Norwell, Mass., chapter 7, 275–310, 1982.
- Heald, C. L., Jacob, D. L., Palmer, P. I., et al.: Biomass burning emission inventory with daily resolution: Application to aircraft observations of Asian outflow, *J. Geophys. Res.*, 108(D21), 8811, doi:10.1029/2002JD003082, 2003.
- Heikes, B., Snow, J., Egli, P., et al.: Formaldehyde over the central Pacific during PEM-Tropics B, *J. Geophys. Res.*, 106(D23), 717–732, 2001.
- Holloway, T., Levy, H., and Kasibhatla, P.: Global distribution of carbon monoxide, *J. Geophys. Res.*, 105(D10), 12 123–12 147, 2000.
- Hudman, R. C., Murray, L. T., Jacob, D. J., et al.: Biogenic vs. anthropogenic sources of CO over the United States, *Geophys. Res. Lett.*, 35(4), L04801, doi:10.1029/2007GL032393, 2008.
- Intergovernmental Panel on Climate Change (IPCC): *Climate change 2001: The science of climate change, Third assessment report of the Intergovernmental Panel on Climate Change*, edited by: Houghton, J. T., Ding, Y., Griggs, D. J., et al., Cambridge University Press, New York, 2001.
- Intergovernmental Panel on Climate Change (IPCC): *Climate change 2007: The Physical Science Basis, Fourth assessment report of the Intergovernmental Panel on Climate Change*, edited by: Solomon, S., Qin, D., Manning, M., et al., Cambridge University Press, New York, 2007.
- Ito, A. and Penner, J. E.: Global estimates of biomass burning emissions based on satellite imagery for the year 2000, *J. Geophys. Res.*, 109, D14S05, doi:10.1029/2003JD004423, 2004.
- Kasibhatla, P., Arellano, A., and Logan, J.: Top-down estimate of a large source of atmospheric carbon monoxide associated with fuel combustion in Asia, *Geophys. Res. Lett.*, 29(19), 1900–1903, 2002.
- Kort, E. A., Eluszkiewicz, J., Stephens, B. B., et al.: Emissions of CH<sub>4</sub> and N<sub>2</sub>O over the United States and Canada based on a receptor-oriented modeling framework and COBRA-NA atmospheric observations, *Geophys. Res. Lett.*, 35(18), L18808, doi:10.1029/2008GL034031, 2008.
- Lin, J. C., Gerbig, C., Wofsy, S. C., et al.: A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, *J. Geophys. Res.*, 108(D16), 4493, doi:10.1029/2002JD003161, 2003.
- Lin, J. C., Gerbig, C., Wofsy, S. C., et al.: What have we learned from intensive atmospheric sampling field programmes of CO<sub>2</sub>?, *Tellus B*, 58(5), 331–343, 2006.
- Liu, Y.: Variability of wildland fire emissions across the contiguous United States, *Atmos. Environ.*, 38, 3489–3499, 2004.
- Logan, J. A., Prather, M. J., Wofsy, S. C., and McElroy, M. B.: Tropospheric chemistry: A global perspective, *J. Geophys. Res.*, 86, 7210–7255, 1981.
- Martin, R. V., Parrish, D. D., Ryerson, T. B., et al.: Evaluation of GOME satellite measurements of tropospheric NO<sub>2</sub> and HCHO using regional data from aircraft campaigns in the southeastern United States, *J. Geophys. Res.*, 109, D24307, doi:10.1029/2004JD004869, 2004.

- Martinez, M., Harder, H., Kovacs, T. A., et al.: OH and HO<sub>2</sub> concentrations, sources, and loss rates during the Southern Oxidants Study in Nashville, Tennessee, summer 1999, *J. Geophys. Res.*, 108(D19), 4617, doi:10.1029/2003JD003551, 2003.
- Matross, D.: Regional scale land-atmosphere carbon dioxide exchange: data design and inversion within a receptor, Ph.D. thesis, Harvard University, 2006.
- Matross, D. M., Andrews, A., Pathmathevan, M., et al.: Estimating regional carbon exchange in New England and Quebec by combining atmospheric, ground-based and satellite data, *Tellus B*, 58, 344–358, 2006.
- Medvigy, D., Moorcroft, P. R., Avissar, R., and Walko, R. L.: Mass conservation and atmospheric dynamics in the regional atmospheric modeling system (RAMS), *Environ. Fluid. Mech.* 5, 109–134, 2005.
- Meszaros, T., Haszpra, L., and Gelencser, A.: The assessment of the seasonal contribution of the anthropogenic sources to the carbon monoxide budget in Europe, *Atmos. Env.*, 38, 4147–4154, 2004.
- Meszaros, T., Haszpra, L., Gelencser, A.: Tracking changes in carbon monoxide budget over Europe between 1995 and 2000, *Atmos. Environ.*, 39, 2005, 7297–7306.
- Millet, D. B., Goldstein, A. H., Allan, J. D., et al.: Volatile organic compound measurements at Trinidad Head, California, during ITCT 2K2: Analysis of sources, atmospheric composition, and aerosol residence times, *J. Geophys. Res.*, 109, D23S16, doi:10.1029/2003JD004026, 2004.
- Millet, D. B., Jacob, D. J., Turquety, S., et al.: Formaldehyde distribution over North America: Implications for satellite retrievals of formaldehyde columns and isoprene emissions, *J. Geophys. Res.*, 111, D24S02, doi:10.1029/2005JD006853, 2006.
- Millet, D. B., Jacob, D. J., Boersma, K. F., et al.: Spatial distribution of isoprene emissions from North America derived from formaldehyde column measurements by the OMI satellite sensor, *J. Geophys. Res.*, 113(D2), D02307, doi:10.1029/2007JD008950, 2008.
- Moody, J. L., Munger, J. W., Goldstein, A. H., Jacob, D. J., and Wofsy, S. C.: Harvard Forest regional scale air mass composition by Patterns in Atmospheric Transport History (PATH), *J. Geophys. Res.*, 103(D11), 13 181–13 194, 1998.
- Müller, J.-F. and Stavrou, T.: Inversion of CO and NO<sub>x</sub> emissions using the adjoint of the IMAGES model, *Atmos. Chem. Phys.*, 5, 1157–1186, 2005, <http://www.atmos-chem-phys.net/5/1157/2005/>.
- Olivier, J. G. J. and Berdowski, J. J. M.: Global emissions sources and sinks, in: *The Climate System*, edited by: Berdowski, J., Guicherit, R., and Heijl, B.J., Balkema/Swets and Zeitlinger, Lisse, The Netherlands, 33–78, 2001.
- Olivier, J. G. J., Van Aardenne, J. A., Dentener, F., et al.: Recent trends in global greenhouse gas emissions: regional trends and spatial distribution of key sources, in: *Non-CO<sub>2</sub> Greenhouse Gases (NCGG-4)*, edited by: van Ambstel, A., Millpress, Rotterdam, 325–330, 2005.
- Palmer, P. I., Jacob, D. J., Fiore, A. M., and Martin, R. V.: Mapping isoprene emissions over North America using formaldehyde column observations from space, *J. Geophys. Res.*, 108(D6), 4180, doi:10.1029/2002JD002153, 2003.
- Palmer, P. I., Abott, D. S., Fu, T. M., Jacob, D. J., et al.: Quantifying the seasonal and interannual variability of North American isoprene emissions using satellite observations of the formaldehyde column, *J. Geophys. Res.*, 111, D12315, doi:10.1029/2005JD006689, 2006.
- Parrish, D.: Critical evaluation of US on-road vehicle emissions inventories, *Atmos. Environ.*, 40(13), 2288–2300, doi:10.1016/j.atmosenv.2005.11.033, 2006.
- Peng, L., Zhao, C., Lin, Y., et al.: Analysis of Carbon Monoxide Budget in North China, *Chemosphere*, 66(8), 1383–1389, 2007.
- Peters, W., Jacobson, A. R., Sweeny, C., et al.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, *P. Natl. Acad. Sci. USA*, 104(48), 18 925–18 930, 2007.
- Petron, G., Granier, C., Khattatov, B., et al.: Monthly CO surface sources inventory based on the 2000–2001 MOPITT satellite data, *Geophys. Res. Lett.*, 31, L21107, doi:10.1029/2004GL020560, 2004.
- Pfister, G., Petron, G., Emmons, L.K., et al.: Evaluation of CO Simulations and the Analysis of the CO Budget for Europe, *J. Geophys. Res.*, 109, D19304, doi:10.1029/2004JD004691, 2004.
- Pfister, G., Hess, P. G., Emmons, L. K., et al.: Quantifying CO emissions from the 2004 Alaskan wildfires using MOPITT CO data, *Geophys. Res. Lett.*, 32, L11809, doi:10.1029/2005GL022995, 2005.
- Pielke, R. A., Cotton, W. R., Walko, R. L., et al.: A comprehensive meteorological modeling system - RAMS, *Meteorol. Atmos. Phys.*, 49, 69–91, 1992.
- de Reus, M., Fischer, H., Arnold, F., de Gouw, J., et al.: On the relationship between acetone and carbon monoxide in different air masses, *Atmos. Chem. Phys.*, 3, 1709–1723, 2003, <http://www.atmos-chem-phys.net/3/1709/2003/>.
- Rogers, C. D.: Inverse methods for atmospheric sounding: Theory and practice, World Scientific, Singapore, 2000.
- Roller, C., Fried, A., Walega, J., et al.: Advances in hardware, system diagnostics software, and acquisition procedures for high performance airborne tunable diode laser measurements of formaldehyde, *Appl. Phys. B-Lasers O.*, 82(2), 247–264, 2006.
- Sanchez-Ccoyllo, O. R., Silva Dias, P. L., Andrade, M. D., et al.: Determination of O<sub>3</sub>, CO, and PM<sub>10</sub> transport in the metropolitan area of Sao Paula, Brazil through synoptic-scale analysis of back trajectories, *Meteorol. Atmos. Phys.*, 92(1–2), 83–93, 2006.
- Shim, C. Y., Wang, Y., Choi, Y., et al.: Constraining global isoprene emissions with Global Ozone Monitoring Experiment (GOME) formaldehyde column measurements, *J. Geophys. Res.*, 110, D24301, doi:10.1029/2004JD005629, 2005.
- Simon, V., Dumergues, L., Ponche, J. L., and Torres, L.: The biogenic volatile organic compounds emission inventory in France: Application to plant ecosystems in the Berre-Marseilles area (France), *Sci. Total Environ.*, 372, 164–182, doi:10.1016/j.scitotenv.2006.08.019, 2006.
- Singh, H. B., Salas, L. J., Chatfield, R. B., et al.: Analysis of the atmospheric distribution, sources, and sinks of oxygenated volatile organic chemicals based on measurements over the Pacific during TRACE-P, *J. Geophys. Res.*, 109, D15S07, doi:10.1029/2003JD003883, 2004.
- Singh, H. B., Brune, W. H., Crawford, J. H., et al.: Overview of the summer 2004 Intercontinental Chemical Transport Experiment North America (INTEX-A), *J. Geophys. Res.*, 111, D24S01, doi:10.1029/2006JD007905, 2006.

- Somnitz, H., Fida, M., Ufer, T., and Zellner, R.: Pressure dependence for the CO quantum yield in the photolysis of acetone at 248 nm: A combined experimental and theoretical study, *Phys. Chem. Chem. Phys.*, 7, 3324–3352, 2005.
- Stewart, H. E., Hewitt, C. N., Bunce, R. G. H., et al.: A highly spatially and temporally resolved inventory for biogenic isoprene and monoterpene emissions: Model description and application to Great Britain, *J. Geophys. Res.*, 108(D20), 4644, doi:10.1029/2002JD002694, 2003.
- Stunder, B. J. B.: NCEP model output - FNL archive data, US Department of Commerce, National Oceanic and Atmospheric Administration, National Climate Data Center, Air Resources Laboratory, Silver Springs, MD, <http://www.arl.noaa.gov/fnl.php>, 1997.
- Thompson, A. M.: The oxidizing capacity of the earth's atmosphere: probable past and future changes, *Science*, 256, 1157–1165, 1992.
- Turnbull, J. C., Miller, J. B., Lehman, S. J., et al.: Comparison of  $^{14}\text{CO}_2$ , CO, and SF<sub>6</sub> as tracers for recently added fossil fuel CO<sub>2</sub> in the atmosphere and implications for biological CO<sub>2</sub> exchange, *Geophys. Res. Lett.*, 33, L01817, doi:10.1029/2005GL024213, 2006.
- Turquety, S., Logan, J. A., Jacob, D. J., et al.: Inventory of boreal fire emissions for North America in 2004: Importance of peat burning and pyroconvective injection, *J. Geophys. Res.*, 113(D12), D12S03, doi:10.1029/2006JD007281, 2007.
- Unisys Weather: Image and Map Archive: <http://weather.unisys.com/archive/index.html>, access: 15 September 2008, 2006.
- US EPA: Emissions modeling clearinghouse related spatial allocation files: "New" surrogates, <http://www.epa.gov/ttn/chief/emch/spatial/newsurrogate.html>, 2004.
- US EPA: The Green Book Nonattainment Areas for Criteria Pollutants, <http://www.epa.gov/air/oaqps/greenbk/cindex.html>, 2007b.
- US EPA: Six common air pollutants: Carbon monoxide, <http://www.epa.gov/air/urbanair/co/effrt1.html>, 2007a.
- US EPA: 1970–2006 Average annual emissions, all criteria pollutants, <http://www.epa.gov/ttn/chief/trends/index.html>, 2007c.
- US NASA JPL: Chemical kinetics and photochemical data for use in atmospheric studies, edited by: Sander, S. P., Ravishankara, A. R., Golden, D. M., et al., US National Aeronautics and Space Administration, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, 2006.
- US NOAA ARL: Eta Data Assimilation System (EDAS40) archive information, US Department of Commerce, National Oceanic and Atmospheric Administration, Air Resources Laboratory, Silver Springs, MD, <http://www.ready.noaa.gov/ss/transport/edas40.html>, 2004.
- US NOAA ESRL: GMD data archive. US Department of Commerce, National Oceanic and Atmospheric Administration, Global Monitoring Division, Earth Systems Research Laboratory, <http://www.esrl.noaa.gov/gmd/dv/ftpdata.html>, 13 March 2007, 2007.
- US NOAA ESRL-GMD: NOAA ESRL GMD Tall Tower Network, US Department of Commerce, National Oceanic and Atmospheric Administration, Earth Systems Research Laboratory, Global Monitoring Division, <http://www.esrl.noaa.gov/gmd/ccgg/towers/>, 28 December 2007, 2007.
- Vermeulen, A. T., Pieterse, G., Hensen, A., et al.: COMET: a Lagrangian transport model for greenhouse gas emission estimation forward model technique and performance for methane, *Atmos. Chem. Phys. Discuss.*, 6, 8727–8779, 2006, <http://www.atmos-chem-phys-discuss.net/6/8727/2006/>.
- Vogelezang, D. H. P. and Holtzman, A. A. M.: Evaluation and model impacts of alternative boundary-layer height formulations, *Bound.-Lay. Meteorol.*, 81(3–4), 245–269, 1996.
- Walko, R. L., Cotton, W. R., Feingold, G., and Stevens, B.: Efficient computation of vapor and heat diffusion between hydrometeors in a numerical model, *Atmos. Res.*, 53, 171–183, 2000.
- Warneke, C., de Gouw, J. A., Stohl, A., et al.: Biomass burning and anthropogenic sources of CO over New England in the summer of 2004, *J. Geophys. Res.*, 111, D23S15, doi:10.1029/2005JD006878, 2006.
- Washenfelder, R. A., Toon, G. C., Blavier, J. F., et al.: Carbon dioxide column abundances at the Wisconsin Tall Tower site, *J. Geophys. Res.-Atmos.*, 111(D22), D22305, doi:10.1029/2006JD007154, 2006.
- Wert, B.P., Fried, A., Rauenbuehler, S., et al.: Design and performance of a tunable diode laser absorption spectrometer for airborne formaldehyde measurements, *J. Geophys. Res.-Atmos.*, 108(D12), 4350, doi:10.1029/2002JD002872, 2003.
- Wiedinmyer, C., Quayle, B., Geron, C., et al.: Estimating emissions from fires in North America for air quality modeling, *Atmos. Environ.*, 40, 3419–3432, 2006.
- Wotowa, G. and Trainer, M.: The influence of Canadian forest fires on pollutant concentrations in the United States, *Science*, 288, 324–328, 2000.
- Zimmerman, P.: Testing of hydrocarbon emissions from vegetation, leaf litter, and aquatic surfaces and development of a method for compiling biogenic emissions inventories, US Environmental Protection Agency, Research Triangle Park, NC, 1979.