Top-Down Constraints on Anthropogenic CO₂ Emissions Within an Agricultural-Urban Landscape

Cheng Hu1, Timothy J. Griffis1, Xuhui Lee2,3, Dylan B. Millet1, Zichong Chen1, John M. Baker1,4, and Ke Xiao1

1Department of Soil, Water, and Climate, University of Minnesota, Twin Cities, St. Paul, MN, USA, 2School of Forestry and Environmental Studies, Yale University, New Haven, CT, USA, 3Yale-NUIST Center on Atmospheric Environment, Nanjing University of Information Science and Technology, Nanjing, China, 4U.S. Department of Agriculture-Agricultural Research Service, St. Paul, MN, USA

Abstract Anthropogenic carbon dioxide (CO₂) emissions dominate the atmospheric greenhouse gas radiative forcing budget. However, these emissions are poorly constrained at the regional (10⁶–10⁸ km²) and seasonal scales. Here we use a combination of tall tower CO₂ mixing ratio and carbon isotope ratio observations and inverse modeling techniques to constrain anthropogenic CO₂ emissions within a highly heterogeneous agricultural landscape near Saint Paul, Minnesota, in the Upper Midwestern United States. The analyses indicate that anthropogenic emissions contributed 6.6, 6.8, and 7.4 μmol/mol annual CO₂ enhancements (i.e., departures from the background values) in 2008, 2009, and 2010, respectively. Oil refinery, the energy industry (power and heat generation), and residential emissions (home heating and cooking) contributed 2.9 (42.5%), 1.4 (19.8%), and 1.1 μmol/mol (15.8%) of the total anthropogenic enhancement over the 3-year period according to a priori inventories. The total anthropogenic signal was further partitioned into CO₂ emissions derived from fuel oil, natural gas, coal, gasoline, and diesel consumption using inverse modeling and carbon isotope ratio analyses. The results indicate that fuel oil and natural gas consumption accounted for 52.5% of the anthropogenic CO₂ sources in winter. Here the a posteriori CO₂ emission from natural gas was 79.0 ± 4.1% (a priori 20.0%) and accounted for 63% of the total anthropogenic signal. The energy industry (power and heat generation), and residential emissions (home heating and cooking) contributed 2.9 (42.5%), 1.4 (19.8%), and 1.1 μmol/mol (15.8%) of the total anthropogenic enhancement over the 3-year period according to a priori inventories. The total anthropogenic signal was further partitioned into CO₂ emissions derived from fuel oil, natural gas, coal, gasoline, and diesel consumption using inverse modeling and carbon isotope ratio analyses. The results indicate that fuel oil and natural gas consumption accounted for 52.5% of the anthropogenic CO₂ sources in winter. Here the a posteriori CO₂ emission from natural gas was 79.0 ± 4.1% (a priori 20.0%) and accounted for 63% of the total anthropogenic enhancement. Further analyses indicate that fuel oil and natural gas consumption accounted for 52.5% of the anthropogenic CO₂ sources in winter. Here the a posteriori CO₂ emission from natural gas was 79.0 ± 4.1% (a priori 20.0%) and accounted for 63% of the total anthropogenic enhancement. Further analyses indicate that fuel oil and natural gas consumption accounted for 52.5% of the anthropogenic CO₂ sources in winter. Here the a posteriori CO₂ emission from natural gas was 79.0 ± 4.1% (a priori 20.0%) and accounted for 63% of the total anthropogenic enhancement.

1. Introduction

Over the past 25 years significant advances have been made in measuring carbon dioxide (CO₂) exchange above natural and managed ecosystems across the globe (Ballantyne et al., 2012; Berthelot et al., 2002; Friend et al., 2007; Peters et al., 2007). These observations have been critical in assessing feedbacks among photosynthesis, respiration, and climate, and for determining changes in the terrestrial carbon sink and source strength (Ballantyne et al., 2012; Berthelot et al., 2002; Gray et al., 2014). The atmospheric CO₂ balance is primarily determined by the very small difference between large gross fluxes of photosynthesis (120 Pg C/a) and respiration (110 Pg C/a; Kesselmeier et al., 2002) with a disequilibrium of about 10 Pg C/a, that is, more-or-less in balance at longer time scales. Current estimates indicate a residual terrestrial sink of about 2–3 Pg C/a, which is comparable to the oceanic CO₂ sink strength (2.3 ± 0.7 Pg C/a; Intergovernmental Panel on Climate Change (IPCC), 2013; Quéré et al., 2016; Rotach et al., 2013). Anthropogenic CO₂ emissions (9.8 Pg C/a for 2015), therefore, represent the main contribution to the increasing atmospheric CO₂ burden and the increase in greenhouse gas radiative forcing (Boden et al., 2013; Peters et al., 2007). Anthropogenic CO₂ emissions dominate the atmospheric greenhouse gas radiative forcing budget, but these emissions are poorly constrained at regional to subregional (10⁶–10⁸ km²), and seasonal scales (Lauvaux et al., 2016; Newman et al., 2013; Turner et al., 2016).

Previous research has applied “bottom-up” and “top-down” methods to quantify anthropogenic CO₂ emissions (Gurney et al., 2017; Brioude et al., 2013; Peters et al., 2007; Turner et al., 2016). The IPCC uses a bottom-up methodology to estimate fossil fuel CO₂ emissions at different scales using activity data and emission factors derived from the synthesis of field observations and modeling studies. The emission...
factors have been shown to vary spatially even for the same source types, and the use of one single emission factor has been shown to cause large biases in emission inventories for some regions (i.e., emission factors are much higher than IPCC default values based on measurements from the 100 largest coal mining regions in China; Liu et al., 2015) and contain large uncertainties (Rypdal & Winiwarter, 2001; Zhao et al., 2011, 2012). Further, it is challenging to quantify the activity data for all source types at finer spatial scales (i.e., urban or city scales). The uncertainties in the activity data and emission factors have been shown to be 10–40% at the country scale and over 150% at finer spatial scales (Peylin et al., 2013; Wang et al., 2013). It is well established that the uncertainty increases at finer spatial and temporal resolutions (Marland, 2008; Shiga et al., 2014). Given these relatively large uncertainties, there is a need for atmospheric-based measurements to help constrain the problem. Unfortunately, CO₂ concentration and flux measurements within urban environments remain relatively rare (Bréon et al., 2015; Lauvaux et al., 2016; Menzer & Mcfadden, 2017).

Eddy covariance flux measurements have also been used as a “bottom-up” approach to assess the net ecosystem CO₂ exchange of urban landscapes (Crawford et al., 2011; Diem et al., 2006; Menzer & Mcfadden, 2017; Song & Wang, 2012; Velasco et al., 2005; Ward et al., 2015). This approach represents the net exchange of natural and anthropogenic fluxes (Liu et al., 2012; Nemitz et al., 2002). Such measurements have been used to help constrain anthropogenic emissions, but relatively large uncertainties exist when scaling them up, because of their relatively small source footprint (10–100 times the measurement height) and the inherent heterogeneity of urban landscapes.

Top-down methodologies use atmospheric measurements of CO₂ mixing ratios (and isotope ratios) observed from rooftops, tall/short towers, aircraft, or satellite observations to help constrain the surface emissions (Berezin et al., 2013; Lauvaux et al., 2016; Suntharalingam et al., 2004; Wang et al., 2013; Worden et al., 2012). The scaling ratio (or enhancement factor) methodology has been widely used to constrain CO₂ emissions at regional scales (Vardag et al., 2015; Wang et al., 2013). This approach uses a source tracer to isolate the anthropogenic CO₂ contribution and assumes the following: (1) that the tracer is transported similarly and well mixed with anthropogenic CO₂ during transport and (2) the flux ratio of these two tracers is constant and equal to their concentration enhancement ratio. Berezin et al. (2013) derived multiannual CO₂ emissions based on satellite observed NO₂ columns and compared it with “bottom-up” Emission Database for Global Atmospheric Research, version 4.2 (EDGAR4.2) inventories. The results showed similar trends for both products. Wang et al. 2010 compared dCO₂:dCO (i.e., the concentration enhancement above background values) with emission ratios at a downwind site near Beijing, China, and found its value was 25% higher than the emission ratio. They attributed this enhancement to urban CO₂ emissions. Recently, Vardag et al. (2015) compared modeled CO₂ concentrations with retrieved results by using three different tracers (i.e., CO, δ¹³C-CO₂, and δ¹⁴C(CO₂)) and concluded that the method performed well at an urban site where the anthropogenic CO₂ signal was relatively strong but performed poorly for rural locations.

Stable carbon isotope techniques were originally used to determine regional CO₂ sources using the Keeling plot (mixing line) method. Early work was based on air samples collected in flasks and analyzed using isotope ratio mass spectrometers (Bowing et al., 2008; Keeling, 1958, 1961). Keeling’s early work showed that the increase of CO₂ concentration and relative depletion of the δ¹³C-CO₂ were mainly caused by anthropogenic CO₂ emissions (Keeling, 1960). With the development of optical isotope ratio infrared spectroscopy, in situ and high-frequency carbon isotope measurements have allowed wide application in tracing and partitioning CO₂ fluxes (Grift et al., 2013; Xu et al., 2017). Stable carbon isotope techniques have been used to constrain the relative contribution of different CO₂ sources within urban airsheds (Newman et al., 2013, 2015; Pang et al., 2016; Pataki et al., 2006; Xu et al., 2017). A recent study conducted in Nanjing, China, found that CO₂ emissions from cement production were significantly underestimated compared with the a priori estimate in bottom-up inventories (Xu et al., 2017). However, as of 2017, only 11 urban sites have established long-term (>30 days) high-frequency stable carbon isotope measurements (Newman et al., 2015; Pang et al., 2016) to help constrain anthropogenic CO₂ emissions.

Atmospheric inversion techniques combine atmospheric transport models and CO₂ mixing ratio observations to constrain surface emissions (Brioude et al., 2013; Pillai et al., 2016; Turnbull et al., 2015). These techniques have been applied to large cities such as Paris, France (Bréon et al., 2015), Berlin, Germany (Pillai et al., 2016), the California Bay Area and Los Angeles, United States (Brioude et al., 2013; Turner et al., 2016), and
Indianapolis, United States (Lauvaux et al., 2016; Turnbull et al., 2015). These studies have demonstrated significant progress in constraining the high spatial and temporal variation of CO₂ emissions and evaluating related CO₂ mitigation strategies. These methods have also been used to constrain net ecosystem CO₂ exchange in highly heterogeneous terrain at the continent scale, with anthropogenic emissions prescribed as “true” values (Gurney et al., 2002; Ogle et al., 2015; Peters et al., 2007).

To the best of our knowledge, there have been very few attempts to quantify the anthropogenic emissions within the Upper Midwest, United States, using atmospheric observations. There has been a significant effort to constrain anthropogenic emissions for the dense urban center of Indianapolis, Indiana, as part of the Indianapolis Flux Experiment (Gurney et al., 2017; Lauvaux et al., 2016). In these studies state-of-the-art emission inventories such as the Hestia-Indianapolis Version 3.0 were combined with Bayesian inversion methods (Gurney et al., 2017). Gurney et al. (2017) demonstrated very good agreement (within the statistical uncertainty) between this bottom-up inventory and the top-down inversion but only after accounting for the importance of ecosystem respiration during the nongrowing season for the period September 2012 to April 2013. Their inventory emissions indicate that vehicle and electricity production dominated (70%) the anthropogenic CO₂ budget. Menzer and McFadden (2017) used the eddy covariance approach to measure net CO₂ fluxes within a residential and park neighborhood in the Saint-Paul-Minneapolis Metropolitan area in Minnesota. They used statistical modeling to partition the net CO₂ exchange into gross fluxes including anthropogenic emissions from fossil fuel burning. They estimated that anthropogenic emissions were dominated (857 to 952 g C/(m²·year)) by natural gas consumption associated with space heating over the period 2007 to 2008. CO₂ emissions associated with vehicle traffic were also relatively large, ranging from 496 to 509 g C/(m²·year). It is not clear to what extent these observational studies are representative of Minnesota, Indiana, or the Upper Midwestern, United States.

Here we apply a combination of the above approaches at a very tall tower located near the Minneapolis-Saint Paul metropolitan area in the Upper Midwestern United States. We combine high-resolution concentration footprints with high spatial and temporal anthropogenic CO₂ flux estimates to simulate year-round CO₂ mixing ratios and its source components. Further, stable carbon isotope ratios ($\delta^{13}$C-CO₂), measured at high frequency (seconds to minutes) using tunable diode laser spectroscopy, were used to further inform the atmospheric inversion mainly in the Saint Paul metropolitan area for the year 2008. The objectives of this research were to (1) quantify the extent to which urban anthropogenic emissions contribute to the tall tower CO₂ mixing ratio enhancements, (2) investigate the potential uncertainty in CO₂ concentration modeling for this heterogeneous (agricultural to urban) landscape, and (3) constrain anthropogenic CO₂ emissions for different source categories by including high-density $\delta^{13}$C-CO₂ information.

2. Materials and Methodology

2.1. Research Site

The tall tower CO₂ observations reported in this study were measured at the University of Minnesota Rosemount Research and Outreach Center located approximately 25 km south of Minneapolis-Saint Paul, MN, United States (Figure 1). This area is in the northern domain of the U.S. Midwest/Corn Belt, where approximately 46% of the land use is agriculture (Griffis et al., 2013; Zhang et al., 2014). The tower is located on the urban/rural interface, with the Minneapolis-Saint Paul metropolitan area immediately to the north and northwest, and agricultural land extending in all other directions. While there are broad trends in wind direction on a seasonal basis, winds can occur from any direction throughout the year, offering the opportunity to observe seasonal patterns in CO₂ mixing ratios that contain both anthropogenic and terrestrial biosphere CO₂ signals.

2.2. CO₂ Concentration, Stable Isotope, and Eddy Covariance Measurements

CO₂ mixing ratios have been measured at the tall tower from April 2007 to February 2018. Air is pulled continuously from four sample inlets (Synflex tubes, ID: 6.25 mm and with a flow rate of 16 L/min) located at 32, 56, 100, and 185 m above the ground using a large diaphragm pump (1023-101Q-SG608X, GAST Manufacturing Inc., Benton Harbor, Michigan, United States). From 2007 to 2009 the sample air was dried and subsampled at 3 Standard Liters Per Minute (SLPM) prior to analysis using a tunable diode laser spectrometer (TGA100A, Campbell Scientific Inc., Logan, Utah, United States), which measured the mixing ratios of three isotopologues including $^{12}$CO₂, $^{13}$CO₂, and $^{18}$O$^{16}$O (Griffis et al., 2010). The air samples of the four
inlets were measured sequentially every 15 s. An omit time of 5 s was used in our data processing to eliminate any residual air from the previously selected sample line. The CO2 measurements were calibrated every 12 min with standards traceable to the National Oceanic and Atmospheric Administration, Earth System Research Laboratory. Half-hourly precisions of $^{12}$CO$_2$, $^{13}$CO$_2$, and $\delta^{13}$C were ±0.03 μmol/mol, 0.0002 μmol/mol, and 0.07‰, respectively (Griffis et al., 2007, 2010). Eddy covariance systems were installed at the 100- and 185-m levels using sonic anemometer-thermometers (CSAT3, Campbell Scientific, Inc.) with air subsampled to an infrared gas analyzer (LI-7000, Licor Inc., Lincoln, Nebraska, United States). Net ecosystem CO2 exchange (NEE) was calculated as the sum of the eddy flux and change in CO2 storage. The net CO2 flux was obtained using 60-min block averaging, and the changes in storage were calculated using the CO2 mixing ratio profiles following Griffis et al. (2010) and Zhang et al. (2014).

2.3. Weather Research and Forecasting Model (WRF)-STILT Model Setup

The Stochastic Time-Inverted Lagrangian Transport model (hereafter the STILT model) is based on the hysplit model, and it helps to define how an observation is influenced by the upwind scalar flux (i.e., the source footprint influence) of the measurement site (Lin et al., 2003). Here the footprint function is calculated by the integration of released particle numbers and their residence time in the planetary boundary layer.

Figure 1. Maps of (a) three domains setup in North America centered in KCMP tall tower and (b) land use categories around tall tower and observation site.
(PBL). Since in the near field, the upstream influence is stronger with decreasing distance from the receptor (Gerbig et al., 2003), accurate transport simulation is especially important within the inner domain of the WRF setup. In this study, three nested domains and a two-way feedback option were applied, with spatial resolution of 27, 9, and 3 km, respectively (Figure 1a). The PBL and microphysical options used in WRF are the same as described in Chen et al. 2016 and are summarized briefly in Table 1. The previous simulation of PBL height showed good agreement with observations at the tall tower site (Chen et al., 2016; Fu et al., 2017). Fu et al. (2017) reported the root mean square error (RMSE) of the PBL height simulation to be 199, 270, and 274 m for the same domain in summer, fall, and winter, respectively. Lin et al. (2003) and Gerbig et al. (2003) were the first to apply this innovative receptor-oriented STILT model for CO₂ observations. It has since been used for many other trace gases including CH₄ (Fournaise & Chaurand, 2015; Jeong et al., 2013, 2016; Zhao et al., 2009), N₂O (Chen et al., 2016; Griffis et al., 2017; Jeong et al., 2012; Xiang et al., 2013), and CO (Bagley et al., 2017; Hooghiemstra et al., 2011; Kim et al., 2013). Here we aim to constrain the regional CO₂ budget with our independent observations and also evaluate the footprint representation by modeling the tall tower hourly CO₂ mixing ratios for the year 2008. Next we simulate the concentration footprint for 3 years (2008 to 2010) and analyze the diurnal, seasonal, and interannual variations.

2.4. Modeled CO₂ Mixing Ratios and Prior Fluxes

Our study domain mainly contains the U.S. Corn Belt in the Upper Midwestern United States and includes the greater Minneapolis-Saint Paul metropolitan area to the north. The land surface characteristics of the domain are highly heterogeneous and include natural and managed vegetation and anthropogenic CO₂ sources (Griffies et al., 2010). A number of prior CO₂ flux products have been evaluated (Gurney et al., 2009; Peters et al., 2007; Peylin et al., 2013), ranging from IPCC inventories, biophysical models, and inverse products. We used prior flux information from Carbon Tracker (Peters et al., 2007), EDGAR42 (European Commission, 2011), and VULCAN (Gurney et al., 2009). For the anthropogenic CO₂ flux, each product has its strengths and weaknesses (as discussed below). We used high spatial (0.1° × 0.1°) and temporal (hourly) anthropogenic emissions that were derived based on these three products. The inverse strategy consisted of four main steps: (1) CO₂ enhancements were modeled using a prior fluxes that included source categories from the above inventory products; (2) stable carbon isotope (δ¹³C-CO₂) end members were specified for different fuel categories (i.e., natural gas, coal, and diesel) and were used to derive a modeled source value (δ¹³Cₘₛ); (3) the carbon isotope ratio of wintertime NEE (δ¹³Cₛ) was obtained using the tall tower observations and the Miller-Tans mixing model approach (Miller & Tans, 2003); and (4) based on the difference between the modeled (δ¹³Cₘₛ) and observed (δ¹³Cₛ) values, a Monte Carlo simulation was used to repartition the relative proportion of each fuel category.

2.4.1. Simulation of CO₂ Mixing Ratios

The CO₂ mixing ratio (CO₂ₘ) can be modeled as the sum of the initial background (CO₂ₜₙₙ) and the enhancement from upstream sources and sinks (S(x,t)). Here we estimated the background CO₂ mixing ratio by tracing the airflow using back trajectory analyses. Each hour, we released 500 particles from the receptor (tall tower 100-m level) and tracked these particles backward in time for 7 days and identified their locations in 3D following the methods of Karion et al. (2015) and Chen et al. (2016). Intercomparison of two Global 3-D background concentration data sets and comparisons with the global background CO₂ network observations demonstrated that these products have relatively high accuracy (Peters et al., 2007; Pillai et al., 2012). Both products (Carbon Tracker and Jana inversion) are generated from the TMS transport model with optimized CO₂ fluxes (Peters et al., 2007; Pillai et al., 2012). Considering the relatively high spatiotemporal resolutions of Carbon Tracker (1° by 1°, and 3-hourly for North America), we used the Carbon Tracker 3-D CO₂ background data product in our study. The modeled CO₂ mixing ratio (CO₂ₘ) was calculated as follows:

\[ CO₂ₘ = CO₂ₜₙₙ + ΔCO₂ₖ + ΔCO₂ₚ \] (1)

\[ ΔCO₂ₖ = ΔCO₂ₕ + ΔCO₂ₚ \] (2)

Table 1

<table>
<thead>
<tr>
<th>Category</th>
<th>Model Configuration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic equations</td>
<td>Nonhydro model</td>
</tr>
<tr>
<td>Microphysics</td>
<td>WSM 3-Class</td>
</tr>
<tr>
<td>Longwave radiation</td>
<td>Rapid Radiative Transfer Model</td>
</tr>
<tr>
<td>Shortwave radiation</td>
<td>Dudhia Scheme</td>
</tr>
<tr>
<td>Surface layer</td>
<td>Monin-Obukhov Similarity Scheme</td>
</tr>
<tr>
<td>Land surface</td>
<td>Noah Land Surface Model</td>
</tr>
<tr>
<td>Boundary layer</td>
<td>YSU Scheme</td>
</tr>
<tr>
<td>Cumulus</td>
<td>Kain-Fritsch (New Eta) Scheme (Domain1 and Domain2)</td>
</tr>
</tbody>
</table>
where \( \Delta CO_2 \) with the subscripts NEE and comb indicate \( CO_2 \) enhancement from net ecosystem exchange and combustion. Here combustion contains fossil fuel burning (ff) and biomass burning (bb).

To avoid “aggregation” errors (i.e., the errors caused by aggregating heterogeneous fluxes in a grid cell into a single average flux) in the modeled \( CO_2 \) mixing ratio (Kaminski et al., 2001; Turner & Jacob, 2015; Zhao et al., 2009), high spatial (0.1° by 0.1°) and temporal (hourly) resolution footprint modeling was performed using the WRF-STILT model. The source footprint functions were then multiplied by different spatial and temporal prior flux to obtain the \( CO_2 \) enhancement contributions from different sources/sinks (i.e., fossil fuel, natural gas, and croplands). The details of the prior flux estimates are described in section 3.1. Here the \( CO_2 \) enhancement was calculated as (Gerbig et al., 2003; Newman et al., 2013; Pillai et al., 2012):

\[
\Delta CO_2,m = \sum_{i=1}^{n} [\text{foot}_{i} \times S(x, t_i)]
\]

where \( \Delta CO_2,m \) is the modeled hourly \( CO_2 \) enhancement from the respective sources (anthropogenic and biogenic fluxes). Hourly \( \Delta CO_2,m \) is the accumulated \( CO_2 \) enhancement occurring over the 7-day back trajectory (i.e., following the released particles backward in time \( [n = 168 \text{ hr}], i \) is the respective hour over the 7-day trajectory from 1 to 168; foot represents the footprint function, and \( S(x, t) \) is the corresponding flux. For example, the concentration at the receptor is given by the source footprint function (units: ppm m\(^2\) · s/\( \mu \text{mol} \)) multiplied by the surface fluxes (\( \mu \text{mol}/(\text{m}^2 \cdot \text{s}) \)) and then adding the background concentration.

### 2.4.2. Anthropogenic \( CO_2 \) Fluxes

High spatial (0.1° by 0.1°) and temporal (hourly) anthropogenic \( CO_2 \) emissions were derived from three data sets (EDGAR42, Carbon Tracker, and VULCAN) and are summarized in Table 2. EDGAR42 provides 13 different categories of anthropogenic combustion and has high spatial resolution (Figure 3a) and good distribution (0.1° by 0.1°) from 1970 to 2010 (Saito et al., 2012). However, it does not account for the diurnal or seasonal variation of fossil fuel burning emissions. The VULCAN data set has high spatial resolution (0.1° by 0.1°) and temporal resolution (hourly) but is only available for the year 2002 (Gurney et al., 2009). Previous studies conducted for Salt Lake City, Utah, United States found that the derived \( CO_2 \) mixing ratio was biased low, which was attributed to the underestimation of the VULCAN anthropogenic \( CO_2 \) flux or an overestimation of mixing in the model transport (McKain et al., 2012; Nehrkorn et al., 2013). Since VULCAN is limited to the year 2002, previous studies have interpolated to the required target year (Mathias et al., 2010; Nassar et al., 2013), which ignores the potential change in spatial distributions and can lead to large uncertainties at the local scale. The anthropogenic emissions specified in Carbon Tracker represent products from both “Miller” (Boden et al., 2013) and “ODIACT” (Oda & Maksyutov, 2011), with a spatial resolution of 1° × 1° and with emissions specified on a monthly basis. Therefore, to utilize the above three data sets, we derived time of day scaling factors in VULCAN and monthly scaling factors in Carbon Tracker (Figure 2). The seasonal variation of these scaling factors accounted for how fuel combustion temporal patterns were influenced by human activities. Then we applied these to EDGAR to get both high spatial (0.1° by 0.1°) and temporal (hourly) anthropogenic emissions following Mallia et al. (2015) and Nassar et al. (2013).

Carbon Tracker also provides monthly fossil fuel burning \( CO_2 \) flux at the 1° by 1° spatial resolution. We compared the differences between EDGAR and Carbon Tracker anthropogenic \( CO_2 \) emissions and found that within the target area with radius increasing from 4° to 20°, centered on the tall tower, the difference of its average flux decreased from 5.9% to 1.8% for the annual average. However, its application to our study domain is not ideal given the strong spatial heterogeneity of anthropogenic emissions. For example, the EDGAR fossil fuel prior emissions map shows that emissions can vary up to 1,000 times in neighboring grid cells such that the average of fine spatiotemporal scales (0.1°) to coarse scales (1°) can result in large errors (i.e., the errors caused by aggregating heterogeneous fluxes in a grid cell into a single average flux) in the modeled \( CO_2 \) mixing ratio (Kaminski et al., 2001; Turner & Jacob, 2015; Zhao et al., 2009), high spatial (0.1° by 0.1°) and temporal (hourly) resolution footprint modeling was performed using the WRF-STILT model. The source footprint functions were then multiplied by different spatial and temporal prior flux to obtain the \( CO_2 \) enhancement contributions from different sources/sinks (i.e., fossil fuel, natural gas, and croplands). The details of the prior flux estimates are described in section 3.1. Here the \( CO_2 \) enhancement was calculated as (Gerbig et al., 2003; Newman et al., 2013; Pillai et al., 2012):

\[
\Delta CO_2,m = \sum_{i=1}^{n} [\text{foot}_{i} \times S(x, t_i)]
\]

where \( \Delta CO_2,m \) is the modeled hourly \( CO_2 \) enhancement from the respective sources (anthropogenic and biogenic fluxes). Hourly \( \Delta CO_2,m \) is the accumulated \( CO_2 \) enhancement occurring over the 7-day back trajectory (i.e., following the released particles backward in time \( [n = 168 \text{ hr}], i \) is the respective hour over the 7-day trajectory from 1 to 168; foot represents the footprint function, and \( S(x, t) \) is the corresponding flux. For example, the concentration at the receptor is given by the source footprint function (units: ppm m\(^2\) · s/\( \mu \text{mol} \)) multiplied by the surface fluxes (\( \mu \text{mol}/(\text{m}^2 \cdot \text{s}) \)) and then adding the background concentration.

### 2.4.2. Anthropogenic \( CO_2 \) Fluxes

High spatial (0.1° by 0.1°) and temporal (hourly) anthropogenic \( CO_2 \) emissions were derived from three data sets (EDGAR42, Carbon Tracker, and VULCAN) and are summarized in Table 2. EDGAR42 provides 13 different categories of anthropogenic combustion and has high spatial resolution (Figure 3a) and good distribution (0.1° by 0.1°) from 1970 to 2010 (Saito et al., 2012). However, it does not account for the diurnal or seasonal variation of fossil fuel burning emissions. The VULCAN data set has high spatial resolution (0.1° by 0.1°) and temporal resolution (hourly) but is only available for the year 2002 (Gurney et al., 2009). Previous studies conducted for Salt Lake City, Utah, United States found that the derived \( CO_2 \) mixing ratio was biased low, which was attributed to the underestimation of the VULCAN anthropogenic \( CO_2 \) flux or an overestimation of mixing in the model transport (McKain et al., 2012; Nehrkorn et al., 2013). Since VULCAN is limited to the year 2002, previous studies have interpolated to the required target year (Mathias et al., 2010; Nassar et al., 2013), which ignores the potential change in spatial distributions and can lead to large uncertainties at the local scale. The anthropogenic emissions specified in Carbon Tracker represent products from both “Miller” (Boden et al., 2013) and “ODIACT” (Oda & Maksyutov, 2011), with a spatial resolution of 1° × 1° and with emissions specified on a monthly basis. Therefore, to utilize the above three data sets, we derived time of day scaling factors in VULCAN and monthly scaling factors in Carbon Tracker (Figure 2). The seasonal variation of these scaling factors accounted for how fuel combustion temporal patterns were influenced by human activities. Then we applied these to EDGAR to get both high spatial (0.1° by 0.1°) and temporal (hourly) anthropogenic emissions following Mallia et al. (2015) and Nassar et al. (2013).

Carbon Tracker also provides monthly fossil fuel burning \( CO_2 \) flux at the 1° by 1° spatial resolution. We compared the differences between EDGAR and Carbon Tracker anthropogenic \( CO_2 \) emissions and found that within the target area with radius increasing from 4° to 20°, centered on the tall tower, the difference of its average flux decreased from 5.9% to 1.8% for the annual average. However, its application to our study domain is not ideal given the strong spatial heterogeneity of anthropogenic emissions. For example, the EDGAR fossil fuel prior emissions map shows that emissions can vary up to 1,000 times in neighboring grid cells such that the average of fine spatiotemporal scales (0.1°) to coarse scales (1°) can result in large

### Table 2

Description of Three Different Anthropogenic \( CO_2 \) Emissions

<table>
<thead>
<tr>
<th>Data sets</th>
<th>Spatial resolution</th>
<th>Temporal resolution</th>
<th>Representative years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carbon Tracker (Peters et al., 2007)</td>
<td>1° × 1°</td>
<td>Monthly</td>
<td>1970–2010</td>
</tr>
<tr>
<td>Emission Database for Global Atmospheric Research, version 4.2</td>
<td>0.1° × 0.1°</td>
<td>Yearly</td>
<td>2000–2010</td>
</tr>
<tr>
<td>VULCAN (Gurney et al., 2009)</td>
<td>0.1° × 0.1°</td>
<td>Hourly</td>
<td>~2002</td>
</tr>
</tbody>
</table>
spatial and temporal aggregation errors (Kaminski et al., 2001; Turner & Jacob, 2015). To quantify the extent of aggregation errors, we aggregated the EDGAR CO2 emissions to 1° by 1°, and the modeled CO2 enhancement was only 0.4 times that of using a 0.1° spatial resolution. The effects of these aggregation errors on the inverse analyses are described in greater detail in section 3.1.

2.4.3. Biogenic Flux

The terrestrial biosphere prior flux was provided by Carbon Tracker assimilation systems (Peters et al., 2007). The biospheric CO2 flux from Carbon Tracker is an optimized product (1° × 1°, Figure 3b). It is derived from the a priori information of modeled NEE using the Carnegie-Ames-Stanford approach and satellite-observed Normalized Difference Vegetation Index (NDVI) as a first guess and then was optimized by atmospheric CO2 measurements. However, it should be noted that the current optimization process assumed the fossil fuel emission uncertainty to be zero. Therefore, its true uncertainty may be distributed into the biospheric flux (Gurney et al., 2002; Nassar et al., 2013; Peters et al., 2007). Finally, the prior CO2 fluxes from biomass burning and oceans (Figures 3c and 3d) were also from Carbon Tracker (Peters et al., 2007). We note that although oceans contribute to overall tall tower source footprint (Figure 1a), the contribution to the calculated enhancement is typically less than 0.02 ppm and can, therefore, be ignored.

2.5. Carbon Isotope Ratio Partitioning Method

2.5.1. Partitioning Process

Here we derived the observed carbon isotope ratio of the source (δs) and then compared it with the modeled value (δ13C_ms) as determined from the a priori emissions. Based on the carbon isotope ratio mass conservation equation, we examined the posteriori values for each source (i.e., respiration, natural gas, diesel, fuel oil, gasoline, and coal) as

\[ \sum_{i=1}^{n} \delta_i e_i = \delta_s \]  

(4)

where δ_i is the carbon isotope ratio for each source category and e_i is the posteriori proportion of the CO2 mixing ratio enhancement for each source type. By giving random uncertainty in δ13C end members and solving the mass conservation equation in equation (4), we used a Monte Carlo simulation to investigate how the uncertainty of the δ13C end members for each fuel category could impact our partitioning results (Brown et al., 2012; Shen et al., 2014). Here we applied a uniform probability distribution, using 10,000 iterations, to assign uncertainty of δ13C in all categories. Note that 2.5% and 97.5% probability values are considered the lower and upper bounds for the retrieved posteriori proportion of each source category.

2.5.2. Miller-Tans Plot Approach

We applied a simple carbon isotope ratio mixing model to characterize the source contribution of the anthropogenic CO2 emissions. The δs was obtained by applying the Miller-Tans method (equation (5)) to the tall tower CO2 mixing ratio and carbon isotope ratio data. Here the Miller-Tans method is used because it has proved to be more robust when the background atmospheric CO2 values are varying with time (Miller & Tans, 2003),

\[ \delta_a C_a = \delta_s (C_a - C_b) + \delta_b C_b \]  

(5)

where δ with the subscripts of a, b, and s indicate the 13C isotope signals of total, background, and additional CO2 contributions by all the sources and sinks. The C_a is the observed CO2 concentration, and C_b is the background CO2 concentration. Here δ_s of the regional sources was derived from the linear regression slope between δ_sC_a and (C_a - C_b). The Miller-Tans method was performed for daytime (10:00–16:00 LT) and nighttime (22:00–06:00 LT) to assess if atmospheric boundary layer dynamics influenced the carbon isotope ratio of the enhancements at the receptor site. Although there is still some debate regarding the best regression type and fitting approach (Zobitz et al., 2006; Kayler et al., 2010), we applied the geometric mean regression for daytime and the weighted least squares for nighttime.
to fit equation (5) to the tall tower data as recommended by Pataki et al. (2003), Kayler et al. (2010), and Pang et al. (2016). The slope of the Miller-Tans plot was then compared with the modeled value ($\delta^{13}C_{\text{Ms}}$) that was derived from the modeled CO$_2$ enhancements from the different source categories with the best available knowledge of their corresponding $\delta^{13}C$ end members (Pataki et al., 2003).

The inherent isotope ratios in different sources and sinks have been measured and applied worldwide. Based on a literature review, we used the following $\delta^{13}C$-CO$_2$ values: $-39.5 \pm 1.1$‰ for natural gas, $-28.9 \pm 0.5$‰ for gasoline (Pang et al., 2016; Takahashi et al., 2002; Widory & Javoy, 2003; Widory et al., 2006), $-25.5 \pm 0.4$‰, $-29.8 \pm 0.3$‰, $-29.3 \pm 0.2$‰ for coal, diesel, and fuel oil, respectively (Widory et al., 2006; Xu et al., 2017). Considering the requirement for 10% ethanol content for gasoline, we applied an end member value of $-25.5 \pm 0.5$‰ for gasoline following recent experiments conducted in Los Angeles (Newman et al., 2015). Further, to account for the influence of ecosystem respiration, we used an end member value of $-28\pm 0.15$‰ based on nongrowing season observations from within the region (Billmark & Griffis, 2009; Griffis et al., 2004, 2007).

In this study we used the following generalized classifications for $\delta^{13}C$ assignment: oil production (refineries) was classified as fuel oil and residential emissions as natural gas. While CO$_2$ emissions in the road transportation and energy industry are complex, based on statistics from the U.S. Energy Information Administration, the proportion combusted by gasoline and diesel is given as 72% and 28%, respectively, for road transportation in Minnesota (International Energy Agency (IEA) 2015). For the energy industry sector, approximately

---

**Figure 3.** Annual average CO$_2$ flux for 2008, (a) anthropogenic fossil fuel emissions, units: $10^{-7}$ mol/(m$^2$ s)), (b) net ecosystem exchange and units: 10$^{-7}$ mol/(m$^2$ s) and (c) ocean flux, units: 10$^{-7}$ mol/(m$^2$ s)).
90% is from power and heat generation (IEA, 2015) and the remainder 10% is from nuclear energy, which is not considered a CO2 emission category. The proportion of natural gas used in electricity power and heat industry was 7.5% in 2005 and changed to 13.3% in 2010 for Minnesota. We assumed a linear trend for its variation over the 5-year period. Further, the emission factor for natural gas (2.7 kgCO2/kg) is very similar to coal (2.5~2.6; IPCC, 2013), suggesting that 11% of CO2 emissions from the energy industry is from natural gas combustion and 89% from coal combustion. For the manufacturing industry, according to statistics for the United States, 60% is from natural gas, 23% from coal, and 17% from oil combustion (IEA, 2015).

3. Results and Discussion

3.1. Aggregation Error Analysis

To quantify the aggregation error when using different spatial resolutions, we conducted 16 paired tests by aggregating the 0.1° prior CO2 emissions to multiple spatial resolutions and then simulated the CO2 mixing ratio enhancement ($\Delta$CO2) for tall tower receptor over the whole year. Here we define the “aggregation ratio” as the simulated CO2 enhancement ratio between different spatial resolutions versus the standard of 0.1°. The greater the bias from 1 means larger aggregation error. For the biogenic CO2 flux map (Carbon Tracker), the domain is more homogeneous, and for this reason we did not have 0.1° biospheric flux; rather, here we treated soil CO2 flux (from EDGAR42) aggregation error as a surrogate for the biospheric flux. The results show that the aggregation ratios are almost the same for soil flux (Figure 4), with a relative difference less than 10%. Here the difference between the 0.1° and 1° grid results is within 2%, so that the aggregation error when using Carbon Tracker 1° by 1° can be ignored for the biospheric flux. Mallia et al. (2015) also used this data set in Salt Lake City, where the land type is less homogeneous than our domain. However, the influence of aggregation error with respect to the anthropogenic emissions is very important. Here the modeled CO2 mixing ratio enhancement was only 0.4 times that of using 0.1° flux, indicating the necessity of applying a high spatial resolution emission map. The aggregation error for biomass burning was not considered here because the monthly average CO2 enhancement from biomass burning was less than 0.06 ppm and was negligible when compared with anthropogenic CO2 emissions and NEE.

Figure 4. Aggregation error analyses for biogenic and anthropogenic CO2 flux, the error bar indicates for root mean square error for 12 months.

3.2. Variations of Concentration Footprint and Cumulative CO2 Mixing Ratio Enhancement

The cumulative CO2 enhancement from anthropogenic emissions for different seasons was assessed for December–February (winter), March–May (spring), June–August (summer), and September–November (autumn) in 2008. As defined by equation (3), the cumulative CO2 mixing ratio enhancement, and its proportional change as a function of time, was analyzed for each period. Figure 5 shows that the CO2 mixing ratio enhancement was characterized by a logarithmic increase with the variation during the last 12-hr back trajectory accounting for 84.5%, 84.2%, 77.9%, and 80.6% of the total CO2 enhancement relative to the past 7 days for spring, summer, autumn, and winter, respectively.
Given a measured annual average wind speed of 6.7 m/s at the tall tower (100-m level), the main source area (i.e., the area representing the previous 12 hr) that contributed to the CO2 mixing ratio enhancement was within about 289 km of the tower. Given the seasonal variation of wind speed for spring (7.1 m/s), summer (5.9 m/s), fall (7.0 m/s), and winter (6.9 m/s), the main source areas were within 307, 255, 302, and 298 km of the tall tower, respectively. This is consistent with the footprint functions shown in Figure 6, where the heavy black color illustrates the most intense or sensitive zone. The smallest source area was associated with summer because of the relatively low wind speed and great atmospheric instability. This intense footprint zone contains strong anthropogenic CO2 emissions from the Saint Paul-Minneapolis area. Overall, the anthropogenic CO2 mixing ratio enhancement for each period was greater than 5 μmol/mol, and the enhancement in autumn (8.5 μmol/mol) was larger than in winter (6.8 μmol/mol), summer (6.6 μmol/mol), and spring (5.3 μmol/mol). These enhancement differences appear to be partly explained by the significantly lower PBL height and the more stable boundary layer that typically occur in autumn and winter (Ahmadov et al., 2009; Guha & Ghosh, 2010). Direct PBL height observations were only available for 06:00 and 18:00 (local time). Based on these observations, we found that the CO2 enhancement in autumn was larger than in winter because the PBL height at 06:00 was lower (130 m) compared to winter (200 m) for these observations. Similar PBL height differences were observed by Kim et al. (2013) for the tall tower site. The distinct difference in daytime and nighttime PBL height can cause much larger CO2 enhancement in nighttime than in daytime. Thus, given similar anthropogenic emissions in each season, the boundary layer meteorology characteristics (including PBL dynamics, wind speed, wind direction, and source area) can contribute up to 1.6-fold CO2 mixing ratio enhancement for different seasons.

The footprint function was calculated using high spatial (0.1° × 0.1°) and temporal (hourly) resolution for 2008. The ensemble average spring, summer, autumn, and winter footprint functions are shown in Figure 6. These results illustrate that the entire North American continent contributes to the KCMP tall tower CO2 mixing ratio enhancement, with 80% of the contribution coming from a source area with radius of 289 km for CO2. In all four seasons, the prevailing wind direction was from the northwest to southeast, resulting in a footprint function that shows relatively strong similarity for each season.

However, there are some key differences among the seasons. For summer and autumn, the source footprint was influenced by the prevailing wind direction that was from 270° to 360°. During spring, the source footprint was influenced by a dominant wind direction from 315° to 45°. The total footprint area was greatest in winter followed by spring, autumn, and summer, which was mainly caused by differences in atmospheric stability and turbulence. Here we use color contours to show the variation in sensitivity or weighting function. Previous work used a threshold of 1e^-4 ppm m^2 · s/μmol, to define the most sensitive zone (shown as red) for the investigation of CO, N2O, and CH4 (Kim et al., 2013; Chen et al., 2016, 2018). At our same site, Chen et al. 2018 concluded that about 30% of the CH4 enhancement was contributed by an area defined by a footprint function threshold of 1e^-4 ppm m^2 · s/μmol, and most of the U.S. Corn Belt is contained within the most
sensitive and intense footprint zone and, thereby, included the dominant influence of agricultural crops (mainly corn and soybean). For the case of CO₂, the most sensitive zone is quite different because of the strongly heterogeneous distribution of anthropogenic CO₂ emissions. We further analyzed the averaged anthropogenic CO₂ flux within the source area with footprint >1e⁻³ and >1e⁻⁴ ppm m² · s/μmol. The results indicate an emission of 0.048 and 0.018 μmol/(m² · s), respectively. This highlights the large CO₂ emissions from the greater Minneapolis-Saint Paul metropolitan area and its strong influence on the tall tower observations. As shown in section 3.1 and Figure 5, nearly 80% of the CO₂ mixing ratio enhancement was attributed by the more intense footprint zone within a radius of 289 km (i.e., in close agreement with the black contour of Figure 6). The source area contained by the red contours accounted for about 10% of the CO₂ enhancement, despite having an area that was about 24 times greater than the source area contained within the black contours. Therefore, we define the intense source area with footprint >1e⁻³ ppm m² · s/μmol for CO₂ in this study. Within this zone the land use consisted mainly of cropland (~40%), forest (~25%), and grass/pasture (~21%), according to the U.S. Department of Agriculture land use layer data in 2008. Though urban areas represent a small fraction (<5%) of the intense footprint, they exert a disproportionately strong anthropogenic effect on the observed tall tower CO₂ mixing ratios and will be examined in greater detail below.

Figure 6. Seasonal average footprints for spring, summer, autumn, and winter (units: ppm m² · s/μmol, shown in the form of logarithm to base 10).
3.3. Comparing Modeled Versus Observed CO₂ Mixing Ratios

The tall tower CO₂ observations show strong diurnal, seasonal, and annual variations influenced by regional CO₂ sinks and sources (i.e., ecosystem respiration and photosynthesis, anthropogenic emissions) and boundary layer stratification. Observed CO₂ mixing ratios at different heights show strong annual average diurnal variations (peak to peak) of 13.8 μmol/mol (32 m) > 11.5 μmol/mol (56 m) > 8.8 μmol/mol (100 m) > 5.2 μmol/mol (185 m). The differences among these four sample heights reached a maximum at 8:00–9:00 (local time, modeled PBL height at ~250 m for summer), while the differences were less than 0.1 μmol/mol from 13:00–18:00 (modeled PBL height ~1,250 m for summer) due to strong turbulent mixing through a relatively deep convective boundary layer (Ahmadov et al., 2009; Ballav et al., 2016; Griffis et al., 2010; Guha & Ghosh, 2010). Because the 100-m sample level includes near continuous CO₂ mixing ratio variation of the CO₂ mixing ratios during the growing season. We performed a sensitivity test using anthropogenic or biological CO₂ emissions with and without diurnal variations through the whole year, to assess its influence on the diurnal CO₂ mixing ratio amplitude. The relatively large regression intercept (7.26 ± 0.03) when using the daytime data indicates bias in one or more of the emission categories including anthropogenic emissions with and without diurnal variations through the whole year, to assess its influence on the diurnal CO₂ mixing ratio amplitude. The geometric regression results indicate a slope of 0.83 ± 0.06 and an intercept of 2.44 ± 0.35 μmol/mol (100 m) – 8:00 (modeled PBL height at ~250 m for summer), while the differences were less than 0.1 μmol/mol from 13:00–18:00 (modeled PBL height ~1,250 m for summer) due to strong turbulent mixing through a relatively deep convective boundary layer (Ahmadov et al., 2009; Ballav et al., 2016; Griffis et al., 2010; Guha & Ghosh, 2010). Because the 100-m sample level includes near continuous CO₂ mixing ratio and carbon isotope (δ¹³C-CO₂) observations, we examine these data in further detail to provide better constraints on the anthropogenic CO₂ emissions.

Overall, the modeled CO₂ mixing ratios (a priori) were in good agreement with the observed seasonal and diurnal variations (Figure 7a) for 2008, with a RMSE of 10.6 ppm and Pearson correlation coefficient of 0.44 (n = 7784, p < 0.001). When restricting the analysis to the afternoon data (13:00–18:00), the RMSE dropped to 7.48 ppm and the correlation coefficient increased to 0.8 (n = 2184, p < 0.001). This improvement appears to be related to well-mixed boundary layer condition, which is better represented by the WRF model compared to the nighttime stable conditions. From January to April 2008, both observed and modeled CO₂ mixing ratios were above the background values because CO₂ emissions from anthropogenic sources and ecosystem respiration dominated the boundary layer budget. In May, the CO₂ mixing ratios dropped below the background values as photosynthesis from forests, grassland, and eventually corn and soybean (Chen et al., 2015; Zhang et al., 2014) dominated the boundary layer budget. The average NEE measured at the tall tower 100-m level was −0.15 μmol/(m² · s), supporting the conclusion that the domain switched from a net carbon source to a net sink in May. During the growing season (June–September) photosynthesis exceeded ecosystem respiration and the observed average NEE was −1.73, −4.21, −2.93, and −0.73 μmol/(m² · s) for June, July, August, and September, respectively. After harvest in September/October, the region transitioned to being a source, leading to enhanced CO₂ mixing ratios that exceeded the background values. The average NEE was +0.65 μmol/(m² · s) for this period (October 2008).

Large diurnal variation of the biogenic CO₂ flux (Figure 7b) combined with a relatively large amplitude of the diurnal PBL height variations (i.e., ~220 m (6:00–7:00) to 1,280 m (14:00–13:00) resulted in pronounced diurnal variation of the CO₂ mixing ratios during the growing season. We performed a sensitivity test using anthropogenic CO₂ emissions with and without diurnal variations through the whole year, to assess its influence on the diurnal CO₂ mixing ratio amplitude. Our sensitivity results showed that the annual average diurnal amplitude changed from 4.3 to 3.4 μmol/mol, indicating that anthropogenic emissions played a minor role compared to the biogenic CO₂ fluxes and PBL height variations.

The analysis of CO₂ mixing ratio enhancement, rather than the absolute mixing ratio, can provide greater insights regarding the influence of source footprint and individual CO₂ source contribution. Here we calculated the observed enhancement by removing the background mixing ratio estimate for the domain (Figure 7c). The results showed that the modeled enhancements were in good agreement with the observed enhancements (R = 0.52, p < 0.001), indicating that the source footprint modeling does a reasonably good job of capturing the spatial and temporal variability. The geometric regression results indicate a slope of 1.08 ± 0.03 and an intercept of 7.26 ± 0.03 μmol/mol. Restricting the analysis to well-mixed daytime conditions yielded a regression with a slope of 0.83 ± 0.06 and an intercept of 2.44 ± 0.35 μmol/mol (100 m) – 8:00. The relatively large regression intercept (7.26 μmol/mol) when using all the data or given the bias in the slope (0.83) when using the daytime data indicates bias in one or more of the emission categories including anthropogenic or biological CO₂ fluxes. Here we note that the biogenic CO₂ flux has been previously optimized based on atmospheric inversion studies that include our study domain. For example, Peters et al. (2007) reported that the NEE for North America, based on inverse analyses, varied between −0.03 and −0.12 Pg C/a (average of −0.11 Pg C/a) and showed that the uncertainty and bias at the regional scale were highly dependent on the temporal and spatial distribution of CO₂ observations. Schuh et al. (2013) used CO₂ observations from five tall towers as part of the Midwest Continental Intensive and three different transport models.
(CT: Carbon Tracker; CSU: Sib-RAMS-LPDM; and PSU: WRF-LPDM) to optimize the CO2 sink/source in the midwestern United States. Their results were compared to bottom-up statistical inventories mainly including crop yield, forest biomass, and woody production. The model results reproduced the spatial distribution of the CO2 flux with a slight overestimation of the CO2 sink strength by 8–20%, 10–20%, and 21% for PSU, CSU and CT, respectively. Ogle et al. (2015) compared the biogenic CO2 inventory flux with that estimated from the atmospheric inversion results in 2007 for the Mid-Continent region and found that they were not

Figure 7. Comparison between (a) hourly simulated and observed CO2 concentration for the year of 2008, (b) observed net ecosystem CO2 exchange at the tall tower, (c) simulated and observed CO2 enhancement for growing seasons (June–September), and (d) comparison of observed and modeled used net ecosystem CO2 exchange around tall tower (June–September).
To further evaluate the influence of the biogenic flux on the tall tower CO$_2$ mixing ratio enhancement, we performed linear regression analyses using our 100-m eddy covariance observations and Carbon Tracker NEE for the grid cell location that includes the tall tower (Figure 7d). Half-hourly NEE values were averaged to 3 hr to eliminate random bias, with data between 10:00 a.m. and 15:00 p.m. chosen to represent daytime and 1:00 a.m. to 6:00 a.m. for nighttime. The regression slopes were 2.26 ($R = 0.69, p < 0.001$) and 1.56 ($R = 0.41, p < 0.001$) for daytime and nighttime, respectively. Although there is a mismatch in the source footprint function for the tall tower NEE versus the Carbon Tracker NEE (i.e., the eddy covariance footprint varies from 10 to 100 times the observation height depending on different PBL conditions (Davis et al., 2003; Horst & Weil, 1992; Zhang et al., 2014)), the land use characteristics are remarkably similar as distance increases away from the tall tower. Considering the similarity of the underlying surface (Griffis et al., 2010; Zhang et al., 2014), the results suggest an overestimation of NEE by Carbon Tracker. The uncertainties in anthropogenic emissions for these sources are calculated based on the IPCC inventory method with values of 5–20% at the national scale and substantially larger (~60%) at regional scales (Ciais et al., 2010; Mallia et al., 2015; Nassar et al., 2013). In the next section, we add $\delta^{13}$C-CO$_2$ to the inverse modeling approach to help better constrain the anthropogenic emissions from different sources.

### 3.4. CO$_2$ Sources and Their Influence on Observed and Modeled Enhancements

Figure 8 shows the monthly growing season CO$_2$ enhancements contributed by biogenic and anthropogenic CO$_2$ fluxes. The modeled average diurnal amplitude in the growing season was 24.3 μmol/mol, in close agreement with the observations (20.2 μmol/mol). The biogenic modification caused a maximum reduction of 25 μmol/mol by daytime photosynthesis, while nighttime ecosystem respiration caused a maximum nighttime enhancement of more than 40 μmol/mol. Biomass burning had a very small impact on the CO$_2$ enhancement. The monthly average enhancement was 0.06 μmol/mol, with hourly contributions rarely reaching 1 μmol/mol. The maximum monthly average occurred in June (0.12 μmol/mol), and the minimum was observed in September (0.001 μmol/mol). These contributions were largely driven by forest fires in central Canada. The anthropogenic emissions caused significant enhancement that is most clearly evident in the relatively shallow and stable nocturnal boundary layer. Throughout 2008, emissions from oil production (refineries), energy industry, road transportation, residential emissions, and the manufacturing industry contributed annual mean enhancement of 2.55, 1.43, 1.11, 0.67, and 0.49 μmol/mol, respectively.

The designations A1–A6 in Figure 8 represent different anthropogenic CO$_2$ categories including oil production (refineries), energy industry, residential emission, road transportation, manufacturing industry combustion, and the left anthropogenic categories (i.e., mineral process and solid waste disposal), respectively. Oil production was the most important anthropogenic CO$_2$ source and accounted for more than 40% of the total anthropogenic CO$_2$ enhancement in all 4 months of the growing season and varied between 42% and 57%. This largely reflects the fact that the largest oil refinery in the Upper Midwest United States, processing an average of 339,000 barrels/day, is located only 8 km NNW of the tall tower. Further, a second large refinery (98,000 barrels/day) is located 18 km away in the same direction (shown in Figure 1b). The energy industry ranked second, contributing between 15% and 22%. Surprisingly, residential emissions contributed only 13% to 17% of the total CO$_2$ enhancement, despite the close proximity of the Twin Cities Metropolitan Area, with a population of 3.6 million. The contribution from residential emissions was much higher during the winter as a consequence of home heating and will be discussed further in the next section.

To provide an additional constraint on the CO$_2$ source partitioning, we compared the observed tall tower $\delta^{13}$C$_{\text{fl}}$ with $\delta^{13}$C$_{\text{CMs}}$. The tall tower Miller-Tans analyses (Figure 9) are shown for the winter months (December, January, and February) when ecosystem photosynthesis was negligible compared to ecosystems respiration. The winter season was also selected to minimize the influence of biogenic CO$_2$ fluxes (Pang et al., 2016). The winter a priori CO$_2$ fluxes contribute an anthropogenic and biogenic CO$_2$ enhancement of 64% and 36% of the total, respectively. The wintertime isotope mixing line analyses indicate values that ranged between $-34.81\%$ and $-35.47\%$o for daytime (10:00–16:00 LT) and nighttime (22:00–06:00 LT), respectively. These results are very similar to the findings reported by Pataki et al. (2003) for Salt Lake City, Utah, United States, which ranged from $-37.2\%$o to $-30.0\%$o in winter. Tanaka et al. (2013) reported $\delta^{13}$C$_{\text{T}}$ values that were statistically different, with $-408 \pm 136$ Tg CO$_2$ for inventory and $-478 \pm 136$ Tg CO$_2$ for the atmospheric inversion.
Figure 8. Observed and simulated hourly CO₂ concentration in June, July, August, and September, the mean components of concentration enhancement for each month are also shown to the right. Designations A1–A6 represent different anthropogenic CO₂ categories including oil production (refineries), energy industry, residential emission, road transportation, manufacturing industry combustion, and the other anthropogenic categories.
significantly more depleted (i.e., –40.1‰) values for Tokyo, Japan. These carbon isotope ratio values are consistent with CO₂ emitted from fossil fuel burning.

Based on the a priori emissions, we used the following weighting factors: oil production (refineries; 31%), residential emissions (12%), energy industry (23%), road transportation (19%), and manufacturing industry (9%) estimated for the winter. Following the classifications for the fuel categories, as described in section 2.5.2, the a priori fuel oil, natural gas, coal, diesel, and gasoline accounted for 32.5%, 20.0%, 22.5%, 5.3%, and 13.7%, respectively, of the total anthropogenic CO₂ enhancement. All other anthropogenic sources accounted for only 6% of the total anthropogenic CO₂ enhancement. We attributed these other sources with an average carbon isotope ratio end member value of –29.8 ± 0.3‰ (equivalent to diesel). Using these weighting factors and the literature values for carbon isotope ratios typical for these sources (see section 2.5), the modeled carbon isotope ratio was –29.3 ± 0.4‰. The modeled carbon isotope ratio was relatively enriched compared to the observations. As is noted in section 2.5, δ¹³C-CO₂ values of CO₂ emitted from natural gas are substantially more depleted than –29.3‰. This supports that CO₂ emissions from natural gas burning were likely underestimated. We now repartition the proportion of each fuel source based on the strategy described below.

First, the biogenic flux was estimated from the wintertime tall tower eddy covariance observations. The mean observed NEE (ecosystem respiration) was 0.28 μmol/(m² · s), while the grid cell a priori NEE (Carbon Tracker) value was 0.42 μmol/(m² · s). Second, a scaling ratio of 0.67(0.28/0.42) was applied to the a priori NEE values. Therefore, the posteriori biogenic and anthropogenic CO₂ enhancement accounted for 21.4% and 78.6% of the total CO₂ enhancement, respectively. Third, we applied the observed δ¹³C-CO₂ value of –35.1 ± 1.7‰ (average of daytime and nighttime data) as the true value for the net CO₂ source (described in section 2.5.1) to perform the repartitioning. To solve the partitioning equations (equation (5)), only one unknown (x) was allowed and we set it as the posteriori proportion of natural gas. Here we assume that the posteriori values for other sources (i.e., coal, diesel, and gasoline) are in the same proportion of their a priori values and recalculate the partitioning based on the carbon isotope mass conservation equation:

$$\begin{align*}
0.21 \times (-28.0‰) + 0.79 & \quad \times \left\{ (-39.5‰) x + (1 - x) \left[ \frac{0.32}{0.8} (-29.3‰) + \frac{0.23}{0.8} (-25.5‰) + \frac{0.14}{0.8} (-25.5‰) + \frac{0.11}{0.8} (-29.8‰) \right] \right\} \\
& = -35.1‰
\end{align*}$$

Further, we applied the Monte Carlo method to investigate how the uncertainty of δ¹³C in each of the sources can impact our partitioning conclusions as applied to equation (5) above. The resolved posteriori values for natural gas was x = 79.0 ± 4.1%; it accounted for 62.8% of total CO₂ enhancement including both biological respiration and anthropogenic CO₂ emissions.
These analyses provide strong support that natural gas was substantially underestimated and that the posteriori value is constrained to be approximately 63%. We conclude that the sum of CO2 emitted from natural gas should increase from the a priori value of 20.0% to 79.0 ± 4.1% and that the emission of CO2 from fuel oil be reduced from 32.5% to 8.4 ± 3.8% of total anthropogenic CO2 enhancement in winter. This partitioning is supported by the fact that more than 66% of homes in Minnesota are heated by natural gas, 17.9% higher than the whole country of Unites States (U.S. Energy Information Administration). The predominance of fuel sources have been shown to vary substantial worldwide depending on different local energy structure. Lopez et al. (2013) conducted a pilot study in Paris, French, and found that natural gas accounted for 70% of the total fossil fuel use in the winter. Further, more than 50% of natural gas combustion attributed to atmospheric CO2 (including biological contributions) in wintertime was reported in Salt Lake City (Pataki et al., 2007); Wada et al. (2011) reported that natural gas combustion and biogenic respiration contributed 60% and 11% of the total enhancement in the winter for Nagoya, Japan, and is in excellent agreement with our results. In Beijing, China, coal accounts for >80% of the CO2 enhancement in winter (Pang et al., 2016). Djuricin et al. (2010) found that the use of natural gas accounted for about 30% to 50% of total fossil in October through December, in Los Angeles, United States.

3.5. Sensitivity of CO2 Enhancement to the Source Footprint Climatology

Here we examine the extent to which the above findings are sensitive to the source footprint climatology and extend our analyses from 2009 to 2010. In these analyses the anthropogenic emissions map was static for the 3 years in order to assess the influence of varying source footprint. The results show that the annual average CO2 enhancement was 6.62, 6.83, and 7.43 μmol/mol for 2008 to 2010, respectively, indicating relatively small variations among these years.

Previous studies have addressed this sensitivity question by examining the concentration source footprint for nighttime versus daytime, with source area size of about 100 km (a city scale) and more than 1,000 km (a regional scale), respectively (Gloor et al., 2001; Shen et al., 2014; Wunch et al., 2009). Here we averaged the WRF-STILT source footprint according to midday (10:00–16:00 LST) and midnight (23:00–5:00) for comparison. As displayed in Table 3, the source area in daytime and nighttime showed no obvious change among the 3 years. Using a much stronger footprint sensitivity criterion of −3 (i.e., where 80% of the CO2 concentration enhancement originates), the areas for the 3 years varied between 294 to 310 × 10^2 km^2 in the daytime and 483.3 to 528.3 × 10^2 km^2 in the nighttime, and when choosing criterion of −4 (where around 90% of the CO2 concentration enhancement originates), little variations were observed, indicating only small footprint source area variations for these 3 years.

4. Conclusions

We combined the WRF-STILT model with high spatial and temporal CO2 flux information to simulate 3 years of tall tower CO2 mixing ratios within a transition zone of agricultural to urban land use. Tall tower CO2 mixing ratios and carbon isotope ratios were used to help constrain the anthropogenic CO2 emissions. Our findings indicate that

1. The anthropogenic CO2 emissions contribute to a tall tower CO2 enhancement of 6.6, 6.8, and 7.4 μmol/mol for 2008 to 2010, respectively. These findings show relatively low sensitivity to the overall source footprint climatology at the annual time scale. However, at finer time scales, careful

<table>
<thead>
<tr>
<th>Year</th>
<th>Period</th>
<th>lg(F) &gt; −1</th>
<th>&gt; −1.5</th>
<th>&gt; −2</th>
<th>&gt; −2.5</th>
<th>&gt; −3</th>
<th>&gt; −3.5</th>
<th>&gt; −4</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>Midday</td>
<td>0</td>
<td>0.9</td>
<td>6.3</td>
<td>46.8</td>
<td>294.3</td>
<td>1,989</td>
<td>10,584</td>
</tr>
<tr>
<td></td>
<td>Midnight</td>
<td>0</td>
<td>3.6</td>
<td>22.5</td>
<td>114.3</td>
<td>483.3</td>
<td>2,075.4</td>
<td>9,650</td>
</tr>
<tr>
<td>2009</td>
<td>Midday</td>
<td>0</td>
<td>0.9</td>
<td>7.2</td>
<td>49.5</td>
<td>300.6</td>
<td>2,187</td>
<td>11,488</td>
</tr>
<tr>
<td></td>
<td>Midnight</td>
<td>0.9</td>
<td>3.6</td>
<td>23.4</td>
<td>117</td>
<td>493.2</td>
<td>2,142</td>
<td>10,048</td>
</tr>
<tr>
<td>2010</td>
<td>Midday</td>
<td>0</td>
<td>0.9</td>
<td>7.2</td>
<td>51.3</td>
<td>310.5</td>
<td>2,082.6</td>
<td>10,071</td>
</tr>
<tr>
<td></td>
<td>Midnight</td>
<td>0</td>
<td>3.6</td>
<td>19.8</td>
<td>118</td>
<td>528.3</td>
<td>2,239.2</td>
<td>11,135</td>
</tr>
</tbody>
</table>
attention must be paid to the potential influence of aggregation errors on the anthropogenic emissions.

2. The combination of carbon isotope observations and inverse modeling indicates that CO₂ emissions from natural gas are underestimated for this region during the winter. The posteriori CO₂ emission from natural gas was 79.0 ± 4.1% (a priori 20.0%), while the posteriori CO₂ emission from fuel oil was 8.4 ± 3.8% (a priori 32.5%)—suggesting a more important role of residential heating than previously estimated.

3. The uncertainty in the anthropogenic source may be greater than biogenic because of more heterogeneous distribution and emission uncertainty at the regional scale.

References


**Erratum**

In the originally published version of this article, several instances of text were incorrectly typeset. The following have since been corrected, and this version may be considered the authoritative version of record.

The ‘2’ in ‘m²’ was changed to a superscript to represent square meters throughout the body, and the captions for Figures 3 and 6. Y. Zhao et al., 2011, 2012 was changed to Zhao et al., 2011, 2012. R. Wang et al., 2013 was changed to Wang et al., 2013. Y. Wang et al. (2010) was changed to Wang et al., 2010. Z. Chen et al. (2016) was changed to Chen et al. 2016. C. Zhao et al., 2009 was changed to Zhao et al., 2009. Z. Chen et al. (2018) was changed to Chen et al. 2018. M. Chen et al., 2015 was changed to Chen et al., 2015.