Investigating Sources of Variability and Error in Simulations of Carbon Dioxide in an Urban Region

Cory R. Martin[a], N. Zeng[a,b], A. Karion[c], K. Mueller[c], S. Ghosh[d], I. Lopez-Coto[d], K. R. Gurney[e], T. Oda[f,g], K. Prasad[d], Y. Liu[a*], R. R. Dickerson[a,b], and J. Whetstone[c]

[a]Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD, USA
[b]Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA
[c]Special Programs Office, National Institute of Standards and Technology, Gaithersburg, MD, USA
[d]Fire Research Division, National Institute of Standards and Technology, Gaithersburg, MD, USA
[e]School of Life Sciences, Arizona State University, Tempe, AZ, USA
[f]Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, MD, USA
[g]Goddard Earth Sciences Research and Technology, Universities Space Research Association, Columbia, MD, USA
*Now at: College of Geosciences, Texas A&M University, College Station, TX, USA

Corresponding authors: Cory Martin (cmart90@umd.edu); Ning Zeng (zeng@umd.edu)

Ning Zeng

2417 Atlantic Building

University of Maryland, College Park

College Park, MD 20742

301-405-5377

Fax: 301-314-9482

Declarations of interest: none

Highlights:

- Evaluation of modeled urban carbon dioxide using multiple emissions inventories
- Modeled carbon dioxide mole fractions agree with observations on average within 1%
- Spread in emissions inventories secondary to error resulting from model meteorology
- Synoptic meteorology as important as time of day for simulating observations

Abstract

As cities embark upon greenhouse gas (GHG) mitigation efforts, there is an increasing need for accurate quantification of urban emissions. In urban areas, transport and dispersion is particularly difficult to simulate using current mesoscale meteorological models due, in part, to added complexity from surface heterogeneity and fine spatial/temporal scales. It is generally assumed that the errors in GHG estimation methods in urban areas are dominated by errors in transport and dispersion. Other significant errors include, but are not limited to, those from assumed emissions magnitude and spatial distribution. To assess the predictability of simulated trace gas mole fractions in urban observing systems using a numerical weather prediction model, we employ an Eulerian model that combines traditional meteorological variables with multiple passive tracers of atmospheric carbon dioxide (CO_2) from anthropogenic inventories and a biospheric model. The predictability of the Eulerian model is assessed by comparing simulated atmospheric CO_2 mole fractions to observations from four in situ tower sites (three urban and one rural) in the Washington DC/Baltimore, MD area for February 2016. Four different gridded fossil fuel emissions inventories along with a biospheric flux model are used to create an ensemble of simulated atmospheric CO_2 observations within the model. These ensembles help to evaluate whether the modeled observations are impacted more by the underlying emissions or transport. The spread of modeled observations using the four emission fields indicates the model's ability to distinguish between the different inventories under various meteorological conditions. Overall, the Eulerian model performs well; simulated and observed average CO_2 mole fractions agree within 1% when averaged at the three urban sites across the month. However, there can be differences greater than 10% at any given hour, which are attributed to complex meteorological conditions rather than differences in the inventories themselves. On average, the mean absolute error of the simulated compared to actual observations is generally twice as large as the standard deviation of the modeled mole fractions across the four emission inventories. This result supports the assumption, in urban domains, that the predicted mole fraction error relative to observations is dominated by errors in model meteorology rather than errors in the underlying fluxes in winter months. As such, minimizing errors associated with atmospheric transport and dispersion may help improve the performance of GHG estimation models more so than improving flux priors in the winter months. We also find that the errors associated with atmospheric transport in urban domains are not restricted to certain times of day. This suggests that atmospheric inversions should use CO_2 observations that have been filtered using meteorological observations rather than assuming that meteorological modeling is most accurate at certain times of day (such as using only mid-afternoon observations).

Key Words

Urban, climate, flux, emissions, CO_2

1 Introduction

Anthropogenic activities such as the combustion of fossil fuels contribute to the largest net flux of carbon into the atmosphere (Intergovernmental Panel on Climate Change, 2013). These emissions along with fluxes associated with deforestation have caused globally averaged CO_2 concentrations to rise from approximately 280 μmol mol^{-1} of dry air (parts per million, or ppm)

at the beginning of the preindustrial era (Etheridge et al., 1996), to a level today of over 400 ppm (NOAA, 2018), a value that continues to rise year after year. Urban areas contribute roughly 70% of the global total fossil fuel emissions (United Nations, 2011). This large contribution of CO_2 emissions from urban areas is due to the generation of electrical and industrial energy generation as well as vehicular transportation, among other things (United Nations, 2011). For regional or global scales, the uncertainty of fossil fuel CO_2 emissions is assumed to be small relative to that of the natural fluxes, but this assumption likely does not hold for smaller scales as biases in both spatial patterns and magnitudes may occur (e.g., Hutchins et al. 2016; Fischer et al., 2017; Gately and Hutyra 2017; Oda et al. 2018) As cities embark upon GHG mitigation efforts, accurate quantification of urban emissions may be able to significantly inform their effective management, in addition to providing quantitative substantiation of progress toward emission reduction goals, many of which extend well into the 21st century (Clarke et al., 2014).

As such, in recent years, several urban GHG measurement campaigns have been implemented to improve measurement, quantification, and attribution of fossil fuel CO_2 emission fluxes in metropolitan areas along with their associated uncertainties. These cities include, but are not limited to, Salt Lake City (McKain et al., 2012), Boston (Briber et al., 2013; Sargent et al., 2018), Indianapolis (Turnbull et al., 2015; Lauvaux et al., 2016; Miles et al., 2017), Paris (Breón et al., 2015), Davos (Lauvaux et al., 2013), and Los Angeles (Kort et al., 2013; Feng et al., 2016). Generally, these experiments feature a network of towers equipped with state-of-the-art greenhouse gas analyzers with mole fraction standards, however other networks with differing observation systems exist (Wunch et al., 2010; Strong et al., 2011; Shusterman et al., 2016; Vogel et al., under review).

The observations from towers upwind, downwind, and inside the metropolitan areas are used in conjunction with atmospheric transport and Lagrangian dispersion models and with prior emission distributions to estimate CO_2 fluxes (both spatially and temporally resolved) and associated uncertainties using a Bayesian approach. These statistical methods are known as inverse models and are commonly considered a “top down” approach since the estimated flux is calculated using atmospheric observations. Urban emissions can also be derived using emission factors applied to emission proxies such as buildings, traffic counts, and power plants (aka “bottom-up” method; Ciais et al., 2010; Leip et al., 2018) that yield anthropogenic flux inventories (e.g., Gurney et al., 2009; Oda and Maksyutov, 2011; Mcdonald et al., 2014; Gateley and Hutyra, 2017). Theoretically, these two approaches could be completely independent from one another. However, given limited atmospheric constraint along with the diffuse nature of the atmosphere, inventories are generally used as a priori information of GHG emissions to help constrain spatial patterns of emission estimates. Thus, transport and dispersion models and the representation of the underlying flux distribution play a large role in accurately estimating GHG emissions at the necessary spatial and temporal scales for mitigation purposes. Although both types of errors (in modeled fluxes and transport) in urban domains have been explored (e.g., Deng et al., 2017), relative magnitudes and relationships continue to be uncertain especially at observational time scales.

One of the most recently established urban GHG observation networks is the National Institute of Standards and Technology (NIST) Northeast Corridor (NEC-B/W), currently encompassing the Baltimore, Maryland and Washington, DC metropolitan areas (Lopez-Coto et al., 2017; Mueller et al., 2018). This network was implemented to demonstrate and improve measurement capabilities for quantifying anthropogenic GHG emissions from urban areas that cannot easily be

disentangled from one another. It is expected that meteorological conditions are not spatially uniform across the region and are temporally impacted by distinct synoptic events. Although this campaign will ultimately consist of a sixteen-tower network, in 2016 only three towers were operational. The observations from these three towers provide an opportunity to (1) investigate the ability to predict mole fractions using an atmospheric transport and dispersion model, along with prior flux distributions and (2) assess the relative impact of transport and prior errors on the simulated observations.

To achieve these objectives, we employ an Eulerian transport model that includes passive chemical tracers that use emission inventories as the surface flux along with initial and boundary conditions to generate 4D fields of atmospheric CO_2. In this way, we can vary the tracers to examine the sensitivity of the predicted mole fractions compared to the assumed meteorology. Eulerian models advect and disperse GHGs forward in time (e.g. Feng et al., 2016) compared to Lagrangian approaches that use particle dispersion models operating backward from an observational 4D location (e.g. McKain et al., 2012). These are analogous approaches, but we employ a Eulerian model so that we can examine simulated meteorology for the entire domain to help us interpret model performance (at added computational cost compared to a Lagrangian model). In addition, we focus on CO_2 given the availability of data, specifically inventory data so that multiple inventories can be used to estimate errors resulting from emissions inventories. Simulated CO_2 mole fractions are compared to CO_2 observations from four in situ towers sites (three urban and one rural) in the NEC-B/W for the month of February 2016.

The paper is outlined as follows: Section 2 describes the methods used for this analysis, including the model domain and configuration, the observation datasets used in the evaluation, and the emissions inventories as input to the transport model, Section 3 presents the observed and modeled CO_2 time series at specific locations, and compares the model meteorology and predicted mole fractions to observed values. A discussion of the results is featured in Section 4 and a summary of results and conclusions are presented in Section 5.

2 Methods

The main component of the modeling framework described in this study is the Weather Research and Forecasting model coupled with chemistry (WRF-Chem), a non-hydrostatic, compressible model that provides passive tracer transport online with mesoscale meteorology forecasting capabilities (Grell et al., 2005; Skamarock et al., 2008; Beck et al., 2011). WRF-Chem has been modified to allow for separate passive CO_2 tracers for four anthropogenic emissions inventories. To evaluate the modeled CO_2's sensitivity to the tracer input, we employ inventories that are commonly used as prior anthropogenic fluxes in inverse modeling studies. In this paper, we refer to a tracer as a 4D mole fraction field of CO_2 whereas the emissions inventory refers to the 3D (or 2D if it does not have temporal variability) flux field. Additionally, a tracer for the biogenic component of the CO_2 concentrations is also included in this modified version of WRF-Chem since the mole fractions observed at tower locations are the integrated signal of both biospheric and anthropogenic fluxes on top of the global atmospheric concentration. For the subsequent analysis presented in this paper, WRF-Chem was run for the month of February 2016. The month of February is used because it is assumed that anthropogenic emissions dominate the integrated atmospheric signals as observed from these tower locations during winter months, as wintertime fluxes from the biosphere are assumed to be smaller relative to anthropogenic emissions, which is not the case during summer months. Thus, the impact of differences between

various prior anthropogenic flux estimates can be better ascertained in the simulated observations. Future work will include evaluating the biospheric model by simulating a summertime period.

2.1 CO_2 Observations

The NEC-B/W will ultimately feature a network of 16 observation sites (12 urban/suburban sites and 4 rural sites) to measure CO_2 and methane (CH_4) continuously. Operated by Earth Networks (EN), each site will feature a high-precision CRDS greenhouse gas analyzer and a calibration and data processing system similar to the in situ sites in the Los Angeles Megacities project (Verhulst et al., 2017). Additionally, a software-controlled valve system switches between multiple inlets, where available, to pull ambient air to sample from different heights above ground level. Data are quality controlled and averaged to hourly mole fractions reported on the WMO X2007 scale (CO_2; Zhao and Tans, 2006) and X2004A scale (CH_4; Dlugokencky et al., 2005) for each inlet height.

In February 2016, three GHG observation tower sites had been established and were collecting continuous in situ CO_2 mole fraction measurements. The three sites are: HAL in Halethorpe, MD southwest of Baltimore (39.2552N, 76.6753W), NDC in the Tenleytown neighborhood of northwest Washington, DC (38.9499N, 77.0796W), and ARL in Arlington, VA (38.8917N, 77.1317W). Additionally, the NOAA / University of Virginia CO_2 observation site in Shenandoah National Park (SNP, 38.6170N, 78.3500W; Lee et al., 2012; Andrews et al., 2014; CarbonTracker Team, 2017) was used as a rural comparison site, as it is located at 1008 m above sea level putting it frequently above the surrounding local planetary boundary layer (PBL) at night (Poulida et al., 1991). This site is also over 10 km from the nearest town, over 25 km from the nearest major highway, and far from most local anthropogenic influences. This analysis uses observations from all inlet heights when available, but for plotting purposes only the lowest inlet

is shown for time series. These four site locations are shown in Fig. 1, with additional information in Table 1.

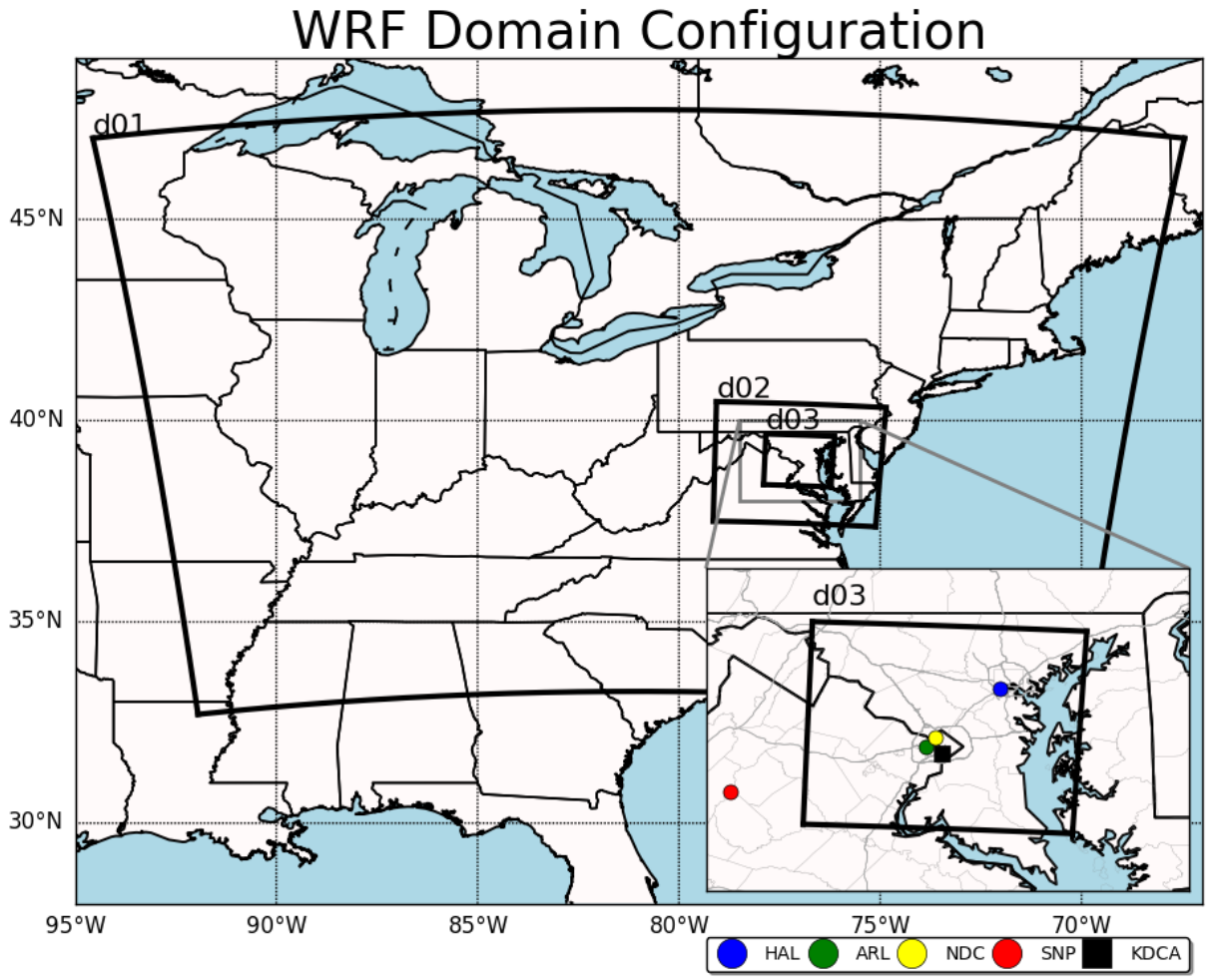

Figure 1. Map showing the WRF-Chem domain configuration used in this analysis. Domain d01 is modeled with 9km horizontal resolution, d02 with 3km, and d03 with a 1km horizontal resolution. The lower right inset shows the immediate area around d03 and the locations of the observing sites used: Shenandoah National Park (SNP; red circle), Arlington, VA (ARL; green circle), Northwest Washington, DC (NDC; yellow circle), Halethorpe, MD (HAL; blue circle), and Washington National Airport for wind direction (KDCA; black square). Major highways are plotted as dark gray lines on the inset map along with the county boundaries in light gray.

Site	SNP	ARL	NDC	HAL
Site Name	Shenandoah National Park	Arlington, VA	Northwest Washington, DC	Halethorpe, MD
Latitude (ºN)	38.6170	38.8917	38.9499	39.2552
Longitude (ºW)	78.3500	77.1317	77.0796	76.6753
Inlet Heights (meters above ground level)	17 m	50 m, 92 m	45 m, 91 m	29 m, 58 m
Site Elevation (meters above sea level)	1008 m	111 m	128 m	70 m
Data provider	NOAA/UVA	NIST/EN	NIST/EN	NIST/EN

Table 1. Summary of the four observation sites used in this study.

2.2 WRF-Chem

A triply nested grid was defined for our WRF-Chem model configuration (Fig. 1). The outermost domain (d01) covers roughly the northeastern quadrant of the United States at a horizontal resolution of 9 km. The d01 extent was chosen because the predominant wind direction for the NEC-B/W is from the North and Northwest (Whelpdale et al., 1984) in February, and this extent generally captures the incoming CO_2 from areas as far away as Chicago, IL. Within this parent domain is an intermediate two-way nested domain (d02) with a resolution of 3 km. An additional fine-scale domain (d03) is nested within d02; it features a horizontal resolution of 1 km that covers the metropolitan areas of the NEC-B/W. A description of the parameterizations and options used for each WRF-Chem domain is provided in Table S1 in the supplemental information (Chou et al., 2001; Hong et al., 2004,2006; Kain 2004; Mlawer et al., 1997; Tewari et al., 2004).

Meteorological initial and boundary conditions are provided by the National Oceanographic and Atmospheric Administration (NOAA) National Centers for Environmental Prediction (NCEP) North American Regional Reanalysis (NARR), a product with a horizontal resolution of 32 km, 30 vertical layers, and three-hourly output (Mesinger et al., 2006). Because the simulation runs for the entire month, sea surface temperatures are also included as boundary conditions in the WRF-Chem model. The high-resolution version of the NOAA NCEP real-time, global, sea surface temperature analysis (RTG_SST_HR) with a horizontal resolution of 1/12 degree and daily output is used (Thiébaux et al., 2003; Gemmil et al., 2007). No analysis nudging or data assimilation is performed as part of the modeling work, as there is concern about how these techniques may affect the simulated CO_2 fields when optimizing meteorology.

Initial and boundary conditions for the background CO_2 concentrations are provided by NOAA Earth System Research Laboratory's (ESRL) CarbonTracker Near Real-Time gridded product

(Peters et al., 2007; https://www.esrl.noaa.gov/gmd/ccgg/carbontracker/CT-NRT/index.php). This is a 3D mole fraction product with three-hourly output and a horizontal resolution of 1º over North America. This background value is available as a separate tracer at all hours of the simulation, and is added to the other tracers for the total predicted CO_2. All tracers resulting from the anthropogenic emissions inventories have initial and boundary conditions of zero ppm. The model-simulated CO_2 mole fraction determined at a point in time and space is the sum of the tracer associated with the specified anthropogenic emissions inventory, the biospheric flux tracer, and the advected background CarbonTracker mole fraction.

For atmospheric transport and trace gas dispersion, three of the most important meteorological variables are the PBL height, near-surface wind speed, and near-surface wind direction. Averaged over 10 meteorological surface sites, the WRF run over-predicts wind speed by 1.2 m/s and has a wind direction bias of 2.8º, however each hour can significantly vary. These results are consistent with previous comparisons of WRF to observations in similar work (e.g. Nehrkorn et al., 2012; Feng et al., 2016). Conversely, WRF's YSU PBL scheme tends to underpredict PBL heights, with an average value of -70 m from observations at three airports over the entire month. Observations of PBL height are computed from potential temperature profiles from both commercial aircraft and radiosondes; both are limited in temporal and spatial resolution resulting in gaps, particularly during the overnight hours. More details on the evaluation of WRF's performance relative to meteorological observations are available in the supplemental information, with specific scenarios presented in Sect. 3.2.

2.3 Emissions Inventories

To evaluate whether the modeled observations are impacted more by the underlying emissions or transport, an ensemble of tracers of atmospheric CO_2 resulting from different emissions inventories are used within the same transport model simulation using WRF-Chem. For this study, four different anthropogenic CO_2 emissions inventories are used: EDGAR, FFDAS, ODIAC, and Vulcan. Since the inventories were generated for a year differing from the modeled year, ratios are used to scale each emission product using national totals from the U.S. Energy Information Administration's (EIA) Monthly Energy Review (https://www.eia.gov/totalenergy/data/monthly/) for each day of February as shown in Equation 1. Emission products will be referred to as tracers from henceforward to correspond to WRF-Chem nomenclature. Fig. 2 shows a map of the hourly mean flux of CO_2 from each inventory interpolated to all three WRF-Chem domains. The following paragraphs and Table 2 provide details of each inventory.

$$Inventory_{2016} = Inventory_{Feb.Year} \times \frac{EIA\ Total_{Feb.2016}}{EIA\ Total_{Feb.Year}} \quad (1)$$

The first inventory employed is EDGAR, i.e. the Emissions Dataset for Global Atmospheric Research version 4.2 (Olivier et al., 2005; http://edgar.jrc.ec.europa.eu). EDGAR is a global emissions product with a horizontal resolution of 0.1º, and provides average fluxes for the year 2010 based on the International Energy Agency's (IEA) energy budget statistics (IEA, 2012). The emissions are then distributed on the 0.1º x 0.1º grid by incorporating population density, road networks, and the locations of point sources and industrial processes.

The Fossil Fuel Data Assimilation System (FFDAS; Rayner et al., 2010; Asefi-Najafabady et al., 2014) is also used. As with EDGAR, FFDAS is a global product with a horizontal grid of 0.1º x 0.1º but unlike EDGAR, it features hourly varying anthropogenic fluxes for the entirety of 2015. FFDAS utilizes the Kaya Identity, a method to estimate emissions based off of economic factors, as well as information on national fossil fuel CO_2 emissions, satellite-derived nightlights, population density, and power plant information to estimate flux at each grid point.

The Open-source Data Inventory for Anthropogenic CO_2 (ODIAC; Oda and Maksyutov, 2011; Oda and Maksyutov, 2015; Oda et al., 2018) is the third inventory used in WRF-Chem. It is the only dataset of the four chosen with a finer horizontal resolution of approximately 1 km, or ~0.01 º. Using the total emissions estimated by the Carbon Dioxide Information and Analysis Center (CDIAC) at the US Department of Energy's Oak Ridge National Laboratory, the locations of point sources and satellite-derived nightlights are then used to distribute the emissions onto the 1 km grid. Monthly total fluxes are provided by ODIAC for each month projected using statistical data from the energy company BP with the most recent version for the year 2015.

The fourth inventory is Vulcan (Gurney et al., 2009) which is a 10 km x 10 km fossil fuel emissions dataset for the United States for the year 2002. Unlike the others, the emissions of individual buildings, power plants, roadways, and other sectors are each characterized, and then aggregated to this 10 km x10 km grid. This provides a higher level of detail both spatially and temporally, but with the limitations of being much older than the other inventories, and only covers the coterminous United States, so parts of domain d01 have no emissions in this tracer.

Inventory	EDGAR	FFDAS	ODIAC	Vulcan
Version	4.2	2.2	2016	2.2
Horizontal Resolution	0.1 degree	0.1 degree	0.01 degree	10 km
Created for Year	2010	2015	2015	2002
Temporal Resolution Provided	Yearly	Hourly	Monthly	Hourly
Domain d01 Average Sum (Tonnes C / hr)	88416	104709	97732 106231 * (no TIMES scaling)	97469
Domain d03 Average Sum (Tonnes C / hr)	2112	3622	2332 2534 * (no TIMES scaling)	2882
Citation	Olivier et al., 2005	Rayner et al., 2010; Asefi-Najafabady et al., 2014;	Oda and Maksyutov, 2011; Oda et al. 2018	Gurney et al., 2009
Available at:	http://edgar.jrc.ec.europa.eu	http://www.gurneylab.org/portfolio-item/ffdas/	http://db.cger.nies.go.jp/dataset/ODIAC/	http://vulcan.project.asu.edu

Table 2. Summary of the four anthropogenic CO_2 emissions inventories used within this study. For ODIAC, both the domain average sum is provided with and without temporal scaling added to the dataset.

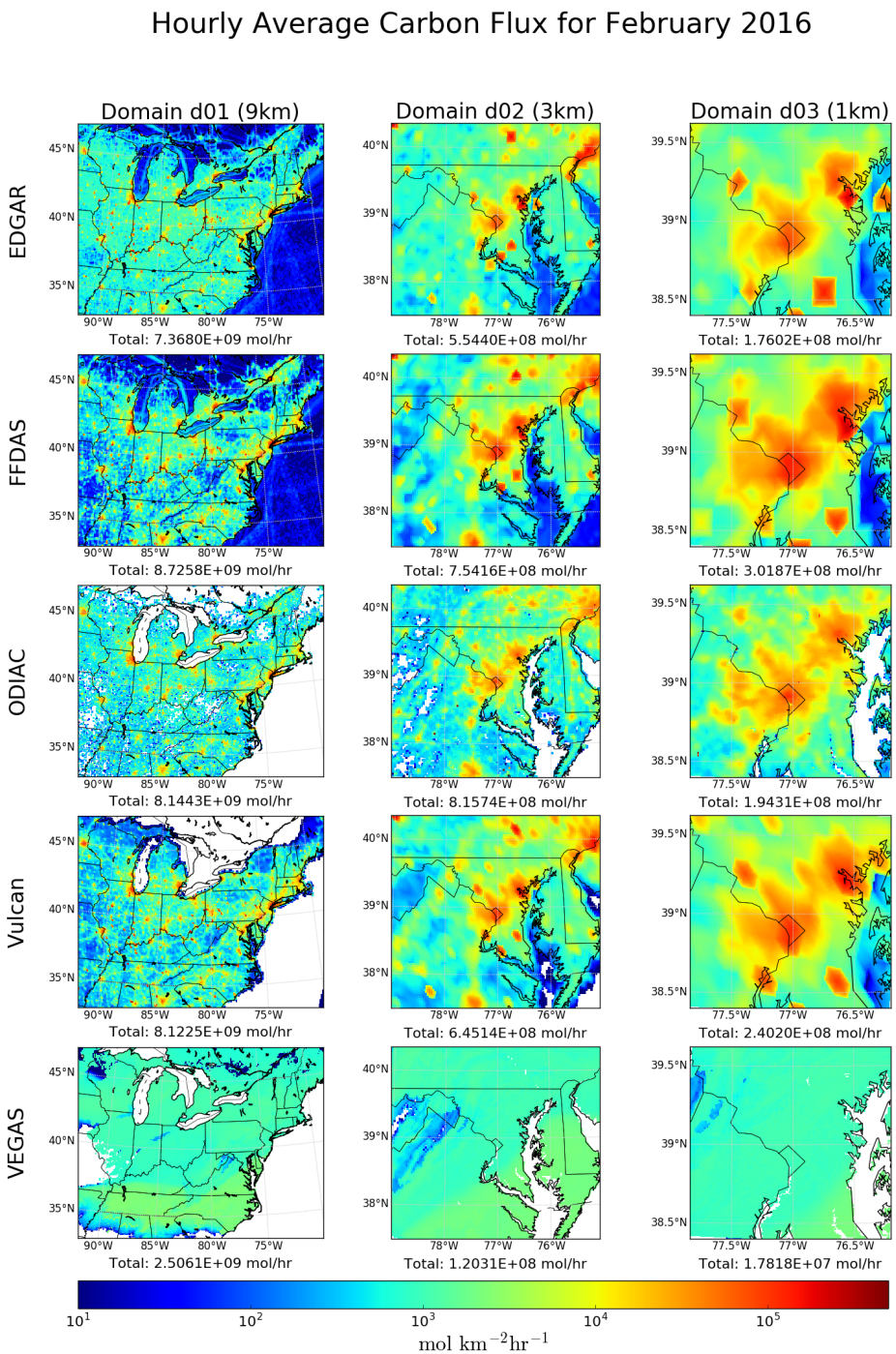

Figure 2. Average CO_2 hourly fluxes for the four emissions inventories and the VEGAS biospheric model for all three WRF-Chem domains in February 2016. Non-positive values (zero

and negative) are shown as white. The hourly average flux for the month for each dataset summed over the entire domain, is shown below each map.

The inventories are interpolated in time and space to ensure consistency. Each one is linearly interpolated from its native grid to the three WRF-Chem domains. Mass is mostly conserved however slight differences from the original datasets may occur due to the domain not lining up exactly with the lower resolution pixels. Additionally, the Temporal Improvements for Modeling Emissions by Scaling (TIMES; Nassar et al., 2013) scale factors are applied to ODIAC and EDGAR to provide weekly and diurnal variations to these two inventories. However, we also use the native monthly ODIAC product as input to WRF-Chem (aka ODIACFIX) as it allows us to investigate the impact of diurnal and weekly varying fluxes on simulated observations. Note that TIMES scaling results in an approximate decrease of emissions of 8.5% when averaged over the entire month because while the daily average remains the same for weekdays, the scaling factor causes a reduction for weekend hours and the number of each day of the week is not the same in any given month. The impact of the TIMES scaling on the simulated observations will be discussed in Section 3.1. We further ensure consistency between the inventories by shifting the inventories with time information so that the calendar days and hours are the same across all emission products. For example, the fluxes for February 2, 2015 of FFDAS are used for February 1, 2016 as they are both Mondays.

2.4 Vegetation Model

As mentioned earlier, a dynamical vegetation model has also been coupled to WRF-Chem to provide the contribution of biogenic fluxes to the simulated observations. The VEgetation-Global-Atmosphere-Soil (VEGAS) model (Zeng et al., 2005) is coupled offline with WRF-Chem to provide hourly biospheric CO_2 flux. Because VEGAS features carbon pools and dynamic vegetation growth, the model must first be spun up on the domain to achieve a climatology. For this analysis, VEGAS is first initialized by forcing it with the Climate Forecast System Reanalysis version 2 (CFSRv2; Saha et al., 2014) calibration climatologies for the years 1981 to 2010. The CFSR climatology dataset is first regridded to the WRF-Chem domains, and then the model is run for 100 years using this calibration climatology repeatedly to reach equilibrium. To generate the land to atmosphere carbon flux, VEGAS uses the WRF-Chem meteorological output variables (2 m temperature, 2 m specific humidity, hourly precipitation, 10 m winds, skin temperature, and total net radiation) as well as the WRF domain topography, emissivity, and albedo. Figure 2 shows the hourly average biospheric flux from VEGAS on all three WRF-Chem domains in February 2016; including areas of net uptake (the white region in the south part of domain d01). The biospheric VEGAS tracer has been added to all the simulated CO_2 values shown although its average contribution to the simulated CO_2 mole fraction across all sites for the month is approximately 1 ppm, with a standard deviation of 0.97 ppm due to the diurnal cycle and meteorological fluctuations. Additionally, no evaluation has been performed on the biospheric fluxes from VEGAS at these spatial scales further highlighting why we focused on winter months for this analysis.

3 Results

In this section we assess WRF-Chem's ability to simulate the atmospheric CO_2 in the NEC-B/W by comparing modeled CO_2 mole fractions at four locations to high-accuracy in situ observations from the three urban and one rural tower sites. First, these datasets are compared over the entire month-long simulation to determine the overall performance of the model. We then select

specific time periods of the month to diagnose possible causes of both high and low performing scenarios. These two analyses help us to evaluate the performance of WRF-Chem in modeling transport and dispersion of urban CO_2 and whether there is sufficient skill in the model for use with various GHG flux estimation methods.

3.1 Overall Model Performance

3.1.1 Tower Observations

Not surprisingly, the magnitudes and variability of the observations from the towers are different for the rural site compared to those from the urban towers (Fig. 3). Over the four locations, the lowest observed CO_2 of the four sites was typically at SNP (with an average of 412.2 ppm vs. the mean of the urban sites at just over 421 ppm) due to its rural location and high altitude, frequently above the PBL in the free troposphere. The amplitude of the observational diurnal cycle at SNP is also smaller than the urban towers but can vary from day to day depending on the synoptic weather situation. The variability in the diurnal cycle at the urban sites is much greater (frequently as high as 50 ppm, but occasionally under 5 ppm) which indicates that synoptic events have a large impact on urban CO_2 observations given the magnitude and variation of the underlying flux distribution in such areas.

3.1.2 Simulated Observations

In general, WRF-Chem generated mole fractions with similar magnitudes and variabilities to the observed mole fractions from the four tower sites. Consistent with the observations, the relative magnitudes of the simulated mole fractions at urban towers are larger than those from the rural site and they exhibit more diurnal variability (the right panels of Fig. 3). In addition, when looking at the model predicted mole fractions in Fig. 3 across all sites, one can note the variations in synoptic and diurnal cycles are similar to the observed time series. These results provide evidence that the model is able to reasonably recreate the time series of CO_2 mole fractions when looking at the complete time series, but an in-depth analysis is required to determine its performance for a particular day or period.

During certain unfavorable meteorological conditions, the spread of the individual emission tracers increases even though their overall variability remains proportionally the same. In terms of overall magnitudes, the differences in the minimum and maximum daily values (i.e. differences between the observed and modeled mole fractions) can be at times quite significant. These large differences correspond to synoptic scale weather patterns (3 d to 5 d) that also create the variability in the observations as discussed in Section 3.1.1. Depending on the day, the differences between modeled and observed CO_2 mole fractions throughout the day can vary by an order of magnitude from less than 5 ppm to over 50 ppm. These synoptic weather conditions will be discussed further in Section 3.2.

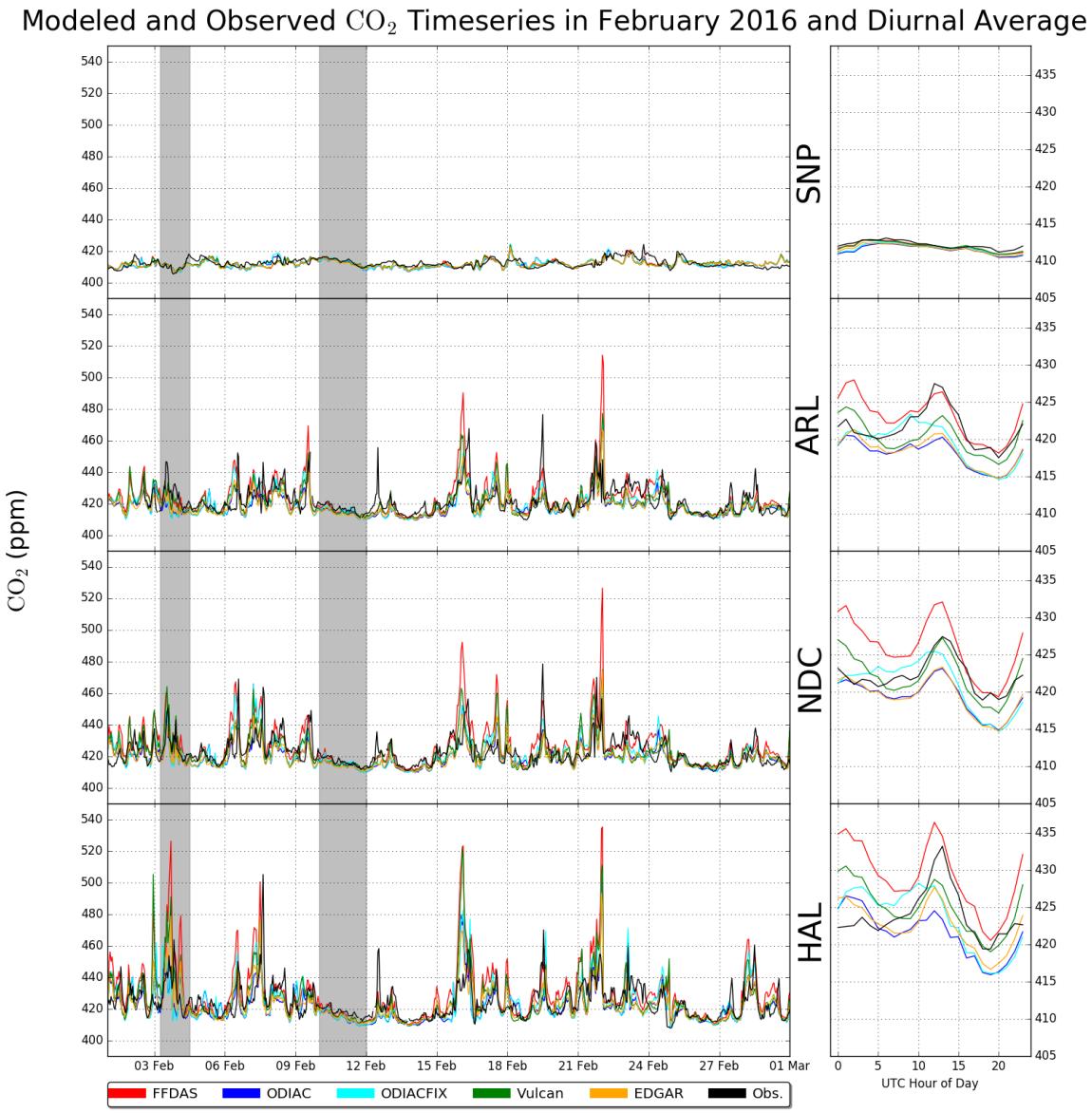

Figure 3. Time series of hourly averaged modeled versus observed CO_2 mole fractions at four observing sites for all hours of the day on the left and diurnal hourly averages on the right. The black lines are the observed values, and each color represents the model-simulated CO_2 interpolated to that location and inlet height (only the lowest inlet levels are plotted at the 3 urban sites). The model-simulated mole fraction at a point in time and space is the sum of an anthropogenic tracer generated from a specific inventory plus the VEGAS biospheric flux tracer plus the background CarbonTracker advected value. From top to bottom: Shenandoah National Park (SNP), Arlington, VA (ARL), Northwest Washington, DC (NDC), and Halethorpe, MD (HAL). FFDAS predicted values are in red, ODIAC in blue, ODIAC without temporal scaling in light blue, Vulcan in green, and EDGAR in orange. Gray shaded areas are scenarios described in detail in Section 3.2.

To characterize the performance of the simulated mole fractions using WRF-Chem relative to observations while accounting for the differences in emissions inventories, a linear regression was performed for all five modeled time series at each site (and both inlets where applicable) against their respective observed time series. The results from these regressions are available in the supplemental information Table S2. For this regression, outliers, defined as differences larger than three standard deviations between the mean of both the observations and the WRF predicted values are removed. The resulting statistics help discern whether the simulated mole fractions, on average, deviate strongly from the observations, are linearly related, and reflect the true variability.

Overall, for all of the urban locations, FFDAS has the strongest linear relationship to the observations with a slope closest to one. However, FFDAS has the largest R^2 for only four out of the seven timeseries (those from two inlet heights at each of the three urban towers and from SNP) for the February 2016 model simulation. However, the lowest R^2 value is associated with the modeled mole fractions using FFDAS at both the HAL inlet levels. This could potentially be due to the location of HAL near large FFDAS point sources in Baltimore that are redistributed onto the native WRF-Chem grid. For the other inventories, the slopes and R^2 varies across all towers.

We also performed a two-day running correlation on each modeled CO_2 timeseries against the observations and a plot for each of the four observing sites is available in the supplemental information in Fig. S2. Similar to what is shown in the residuals of the linear regression, there are times when the modeled CO_2 mole fraction is highly correlated with the observations and periods where correlation is low, and even occasionally negatively correlated. All four observing sites have similar ranges, with two-day correlations ranging from as low as approximately -0.5 to over 0.98. The curves across the four observing sites all look very similar, suggesting that the correlations are consistent across the domain and not limited to one location.

The slopes associated with simulated observations from the other four anthropogenic tracers have a small range of spread between inventories, with the spread being between 0.03 and 0.16 depending on observing site. The slopes closest to zero tend to be either ODIAC or EDGAR depending on the observing point location. The fact that these two inventories have similar slopes is not totally unexpected, as their derived emissions may be distributed spatially in a similar manner, albeit at different resolutions, and the TIMES scaling factors are applied to both of them.

The other statistics indicate that the performance of the modeled time-series is dependent on local conditions (i.e. meteorology or flux distribution) around each tower site since no single tracer consistently out-performs the others. For example, the mean absolute error (MAE) associated with the five anthropogenic tracers varies across tower inlets for a single inventory (such as FFDAS from 3.97 ppm to 7.56 ppm or Vulcan from 4.40 ppm to 6.73 ppm). FFDAS generally has the highest MAE with EDGAR or Vulcan usually having the lowest, but again, there is no consensus on any best or worst performer at each observing site. This suggests that model performance should not be assessed on average across the entire domain but rather locally around tower sites.

After calculating these linear regressions, the fitted datasets can be subtracted from the original modeled time series to see where the linear fit is not valid. Figure 4 shows the residuals of each linear fit from the observed CO_2 at each site where the five colors represent the different anthropogenic tracers in WRF-Chem. As with the slope, FFDAS (red) has a consistently larger absolute residual value (5.1 ppm) than the other four datasets (ODIAC (blue): 3.6 ppm; ODIACFIX (cyan): 4.3 ppm; Vulcan (green): 4.3 ppm; EDGAR (orange) 3.5 ppm) for February 2016 across the observing sites, likely due to the periodic high values skewing the linear fit as noted earlier. The residual plot also shows clearly periods where the simulated CO_2 deviates greatly from the observations for all tracers. This suggests that at times 1) the synoptic scale background CO_2 provided by CarbonTracker may not be resolved correctly, 2) there are sufficient errors in the meteorological transport, or 3) VEGAS is under-predicting respiration during this period. Two of these cases (the dark gray shaded regions in Fig. 4) will be described in detail in Section 3.2.

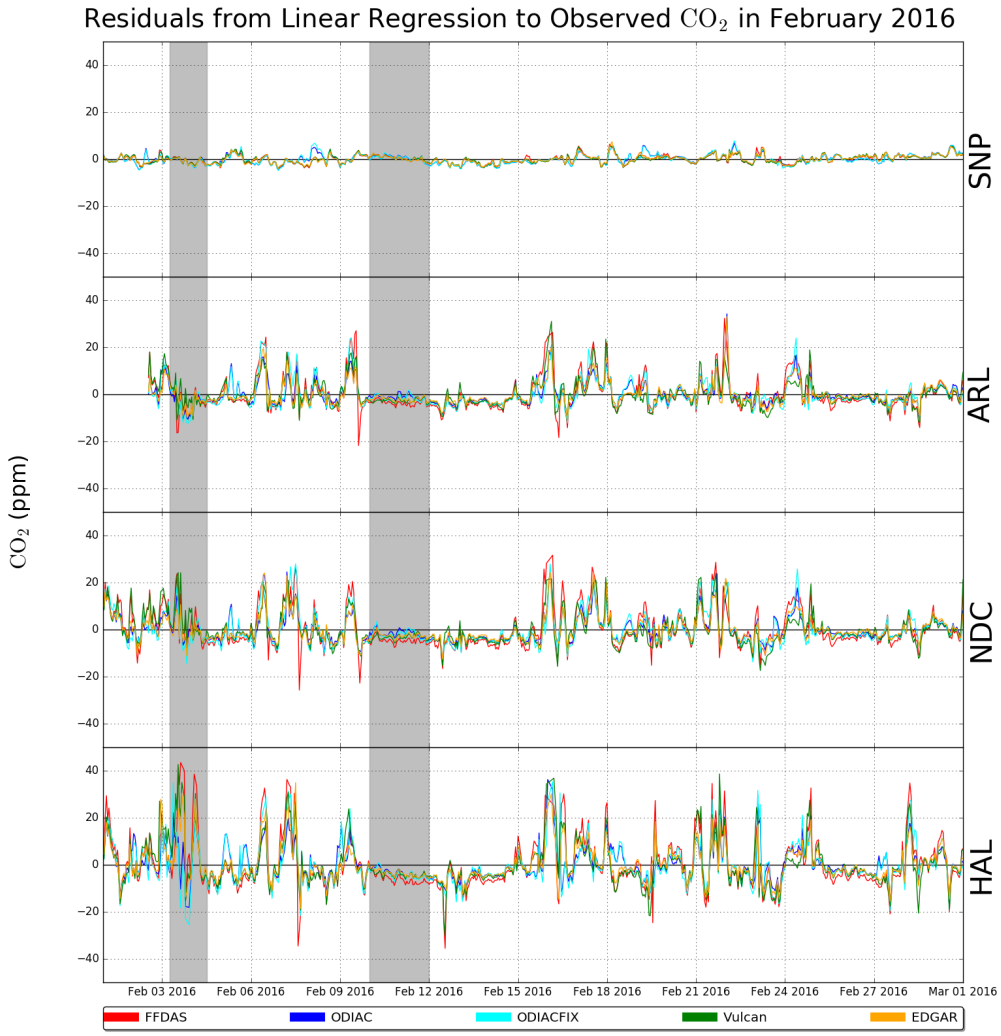

Figure 4. Residuals of a linear regression between the observed CO_2 and each tracer at all four observing sites (lowest inlet only at the 3 urban sites). See Section SI2 of the supplemental information for the regression equation used for this analysis. The different colors represent the

five different tracers from the multiple emissions inputs. The dark gray shaded areas are scenarios described in detail in Section 3.2 which were also shown on Figure 3.

To investigate the impact of bias on the modeled vertical mixing and its representation of the PBL, the simulated mole fractions are analyzed using (1) all hours of the time series, and (2) afternoon hours only (12 pm to 4 pm local time; 17 UTC to 21 UTC), both without removing any outliers. Afternoon observations are typically used in inversions since it is assumed that the meteorological models can better represent well mixed conditions (i.e. the lower relative error of the higher PBL heights) typically found in the middle of the day (e.g., McKain et al., 2012; Kort et al., 2013; Breón et al., 2015; McKain et al., 2015; Lauvaux et al., 2016; Sargent et al., 2018). Figure 5 shows the monthly mean bias of simulated CO_2 mole fractions for all hours and afternoon hours only for all five tracers at the observing sites and different inlet heights.

The spread of the monthly bias from all tracers ranges from -4.5 ppm to 5.2 ppm for the five simulated observational timeseries encompassing all hours of the day. FFDAS has a slight positive mean bias at all three urban sites (the largest being at HAL and at NDC for the lowest inlet). This is consistent with the domain mean hourly averaged flux in domain d03 being the largest for FFDAS as shown in Fig. 2 and Table 2. Conversely, as expected from the flux summaries in Section 2.3, the most negative bias tends to be from either ODIAC or EDGAR as they are diurnally scaled using the TIMES dataset and have the lowest domain mean hourly averaged flux in the urban domain. The impact of the TIMES scaling is clearly demonstrated by the fact that ODIACFIX has a smaller bias than that of ODIAC.

When considering only afternoon hours, the spread in the monthly bias is smaller, ranging from (-3.9 ppm to 2.1 ppm). The FFDAS tracer yields the highest simulated CO_2 mole fraction (similar to all hours), and is the only inventory that has a clear positive bias. Although mostly negative, the Vulcan tracer has a near zero bias at both the HAL inlets. When looking at afternoon hours only, the ODIAC and ODIACFIX tracers are virtually the same as the TIMES scaling factors are based on a value of one for the mid-afternoon (the emissions are scaled down overnight). The mean bias from EDGAR during these periods is similar to that from ODIAC, with a negative mean bias of approximately 3 ppm to 4 ppm from the observations. This range of biases (FFDAS positive to ODIAC and EDGAR being the lowest) is consistent with their respective rankings in the areal total anthropogenic flux for February 2016. Using afternoon hours only appears to reduce the spread of the modeled time series overall compared to the complete datasets but this is dominated by the large reduction in spread at HAL and the lower inlet at NDC.

Generally, WRF-Chem using these emissions inventories tends to underpredict near-surface CO_2 as shown in Fig. 5 (red and blue lines represent the average associated with all hours and afternoon hours respectively). On average, the mean of the five modeled time series is below the observed values, except for the all-hours datasets at HAL, likely largely the result of an underestimation of emissions. The smaller bias at SNP implies that a bias in the CarbonTracker background or the biospheric flux may also contribute to the low bias across the domain. Virtually no spread (< 1 ppm) is found at SNP for both all hours and afternoon only, with very little changes between the two periods, consistent with its rural location and lack of enhanced CO_2 values from emissions sources. This general underprediction trend is exacerbated when only

including afternoon hours (the blue line in Fig. 5), where the mean bias becomes more negative for all datasets except for the higher inlet level at ARL. However, when including all hours (the red line in Fig. 5), the absolute value of the average mean bias across all inventories is smaller than when only looking at afternoon hours. This is likely due to the overprediction that commonly occurs at night balancing out the underprediction occurring in the afternoon hours. Thus, there may be some advantage to using additional observations from hours outside of the afternoon period, as generally speaking the more data used, the better constrained the inversion estimate would be.

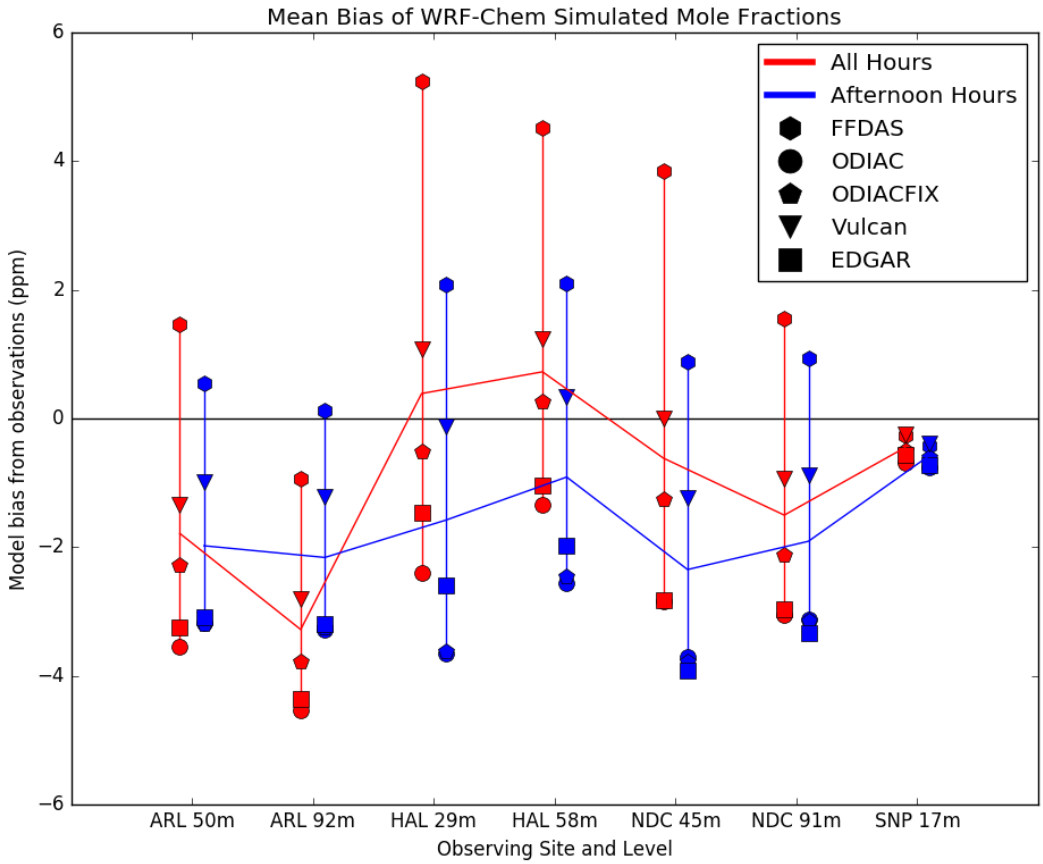

Figure 5. Mean bias of WRF-Chem simulated CO_2 mole fractions (ppm) compared to observations at each observing site and for both inlets where applicable for all five tracers (FFDAS: square, ODIAC: circle, ODIACFIX: pentagon, Vulcan: triangle, and EDGAR: square) during all hours (red) and during afternoon (12 pm to 4 pm local time; blue) hours only. Means of bias at each inlet are connected with additional lines: red for all hours and blue for afternoon.

3.2 Typical Scenarios

The model-data comparisons presented in Section 3.1 cover the entire month of February 2016. While the mean biases of the modeled CO_2 mole fractions are relatively small, the variation between days, and even between individual hours of the day, can be significant, as shown in the residual plots in Fig. 4 (as well as in the running correlation plots in Fig. S2). Two sample cases

(the dark gray shaded regions in Fig. 4) are presented to show a range of scenarios. One of these cases is when the model error is large (greater than 10 % of the observed total value) and the proportion of variability (the ratio of the standard deviation of the differences over the period of interest and the standard deviation over the entire month) is greater than 100 %. The second scenario occurs when the model-observation difference is small (~<1 % of the total CO_2) and the proportion of variability relative to the whole month is under 40 %. For trace gases in the atmosphere, winds (both speed and direction) as well as the height of the PBL are the most important meteorological factors in estimating near-surface CO_2. In the subsequent subsections, two scenarios are presented that show examples of how different synoptic weather situations can affect the ability of WRF to predict these variables, and in turn, can impact the quality of the predicted CO_2 in a forward transport model.

3.2.1 Frontal Passages

A typical mid-latitude location such as the NEC-B/W experiences frontal passages every 3 to 7 days. When a front passes over an area, such as this urban domain, it causes sharp changes in wind speed and direction, temperature and moisture content, as well as other defining features of an air mass including CO_2 mole fractions (Parazoo et al., 2008). Figure 6 shows an example of the impact on CO_2 mole fractions when two fronts pass over the NEC-B/W as a mid-latitude cyclone moves east. The gradual slope of the warm front (~12 UTC February 3) can create a shallow PBL, allowing CO_2 to accumulate near the surface, whereas a cold front (~8 UTC February 4) is much steeper in its vertical structure. For the latter, the observed wind shifts are much more abrupt both in speed and direction. Both frontal passages can also create surface convergence ahead of the front, allowing CO_2 to build up in the PBL immediately before the wind shift occurs. Figure 7 shows surface CO_2 concentrations predicted using the FFDAS emissions inventory and 10 m wind vectors during the model simulated frontal passage (5 UTC February 4), illustrating the spatial gradient of CO_2 during this period.

For both the warm and cold fronts, WRF simulates the frontal passage and associated wind shift 4 to 5 hours before the observed passage (Fig. 6). Generally, correlations are observed between the wind direction, the PBL height, and the simulated CO_2. The simulated warm front caps the PBL, allowing for predicted mole fractions to increase across the domain. The diurnal range is well resolved by WRF-Chem at NDC, but it is greatly overestimated at HAL and underestimated at ARL, likely related to the timing and position of the passing front. The spread between the various inventories at HAL is also much larger than at the other two sites. During the first frontal passage (12 UTC February 3), it is clear that the predicted CO_2 values are influenced by the shallow PBL depth, which is also observed at a number of periods during the month where the model and observations diverge. However, during the simulated cold front (~8 UTC February 4), the PBL heights do not change significantly, but the predicted CO_2 peaks and then drops rapidly as the wind shifts from southerly to northerly. This feature is seen in the simulated time series at all three sites with WRF-Chem results underestimating CO_2 mole fractions at the observing locations when the front actually passes through. This example illustrates that the meteorological error, in both the timing of the front and the PBL depth, dominates the error in the simulated CO_2 values rather than the underlying emissions, as often the spread in the emissions inventories (shaded red area) is generally smaller than the difference between the model mean (red line) and the observations (black).

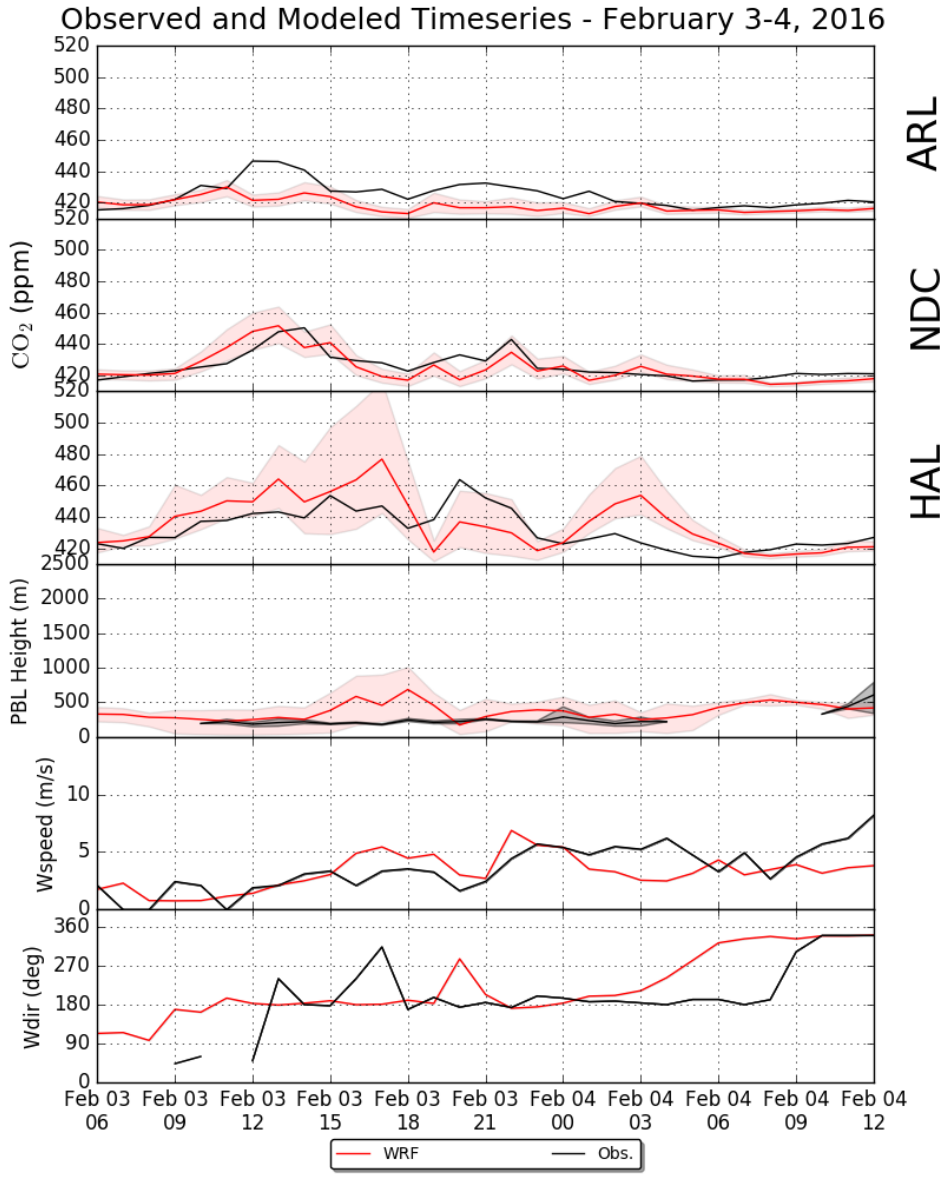

Figure 6. Observed (black) and modeled (red line average; shaded red spread of the five emissions inventories) hourly averaged CO_2 at all three urban sites for a typical frontal passage period (February 3-4, 2016). The fourth panel shows the mean (black line) and spread (shaded gray) of the inferred PBL heights from aircraft profiles and radiosonde observations, with the YSU PBL scheme predicted PBL height from WRF at each observing site mean (red line) and spread (shaded red). Bottom two panels show 10 m observed (black) and modeled (red) wind speed and direction at KDCA. See the supplemental information for details on the meteorological observations used.

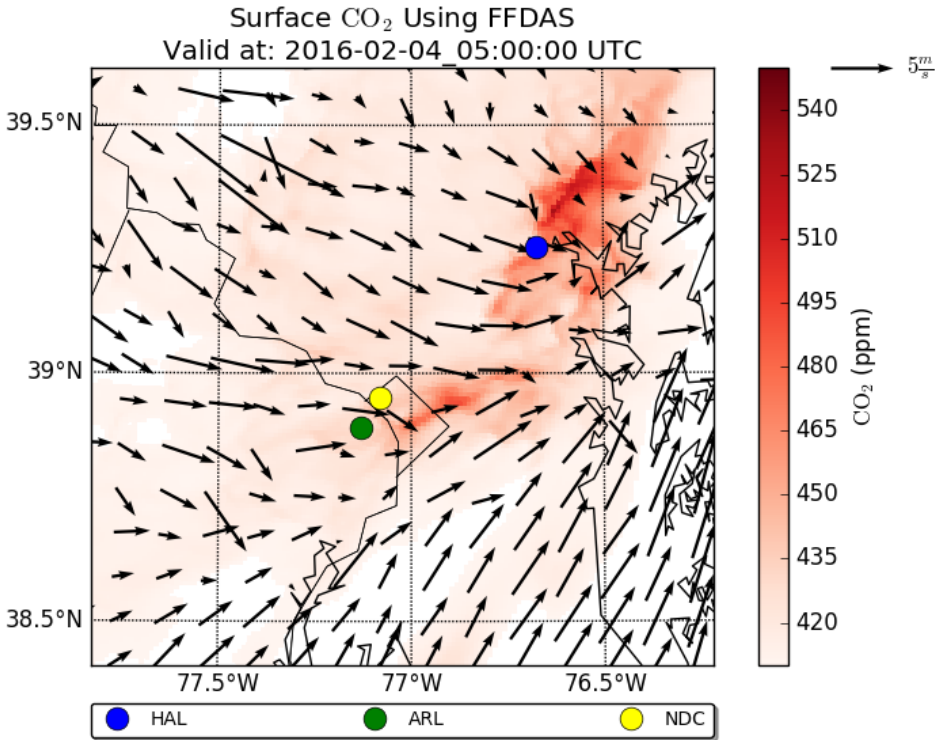

Figure 7. Simulated surface CO_2 concentrations using FFDAS emissions and 10 m wind vectors during a cold front passage at 5 UTC February 4, 2016 in domain d03. Locations of the three urban observing sites are shown (HAL in blue; ARL in green; NDC in yellow).

3.2.2 Persistent Winds

While the first scenario presents an example period where WRF-Chem significantly under or overpredicts CO_2 compared to the observations, there are times when the model simulated CO_2 is within 1 % of the observed value. One such example is from February 10 to February 11, 2016, when winds are steady and from a direction where the upwind CO_2 mole fractions are more representative of the global average. Figure 8 shows the modeled and observed CO_2 as well as wind direction and PBL height for this period. During these two days, the wind is persistently from the west or northwest, bringing in a steady stream of air into the urban area where the mole fraction observations are strongly influenced by the incoming atmosphere, or regional background, values.

The average modeled CO_2 mole fractions at all three sites for this 48 hour subset are always within 5 ppm of the observed value. At ARL and HAL most hours are underpredicted, possibly due to the overprediction of the PBL height during this period as the modeled wind speed and direction agree well with the observations, but at NDC the average modeled value differs by 1 ppm to 2 ppm from the observations from 0 UTC to 18 UTC on February 11. Additionally, the predicted CO_2 mole fractions from all five inventories do not vary significantly from one another during this case, with a spread of only 2 ppm to 3 ppm on either side of the mean throughout the period. Despite this agreement, there are still some subtleties that can be observed in the data. All three sites have a local maximum in the observations at around 20 UTC on February 10, but the modeled time series have a local minimum there. At the same time, WRF overpredicts the PBL

height compared to the observed height, likely causing dilution in the predicted CO_2 mole fractions. These features are present in all the inventory tracers, and often the observations are not within the spread of the five models, even during this period of relatively good model performance. In addition to the meteorological errors described above, biases in the CarbonTracker background or the biospheric tracer (both common to all five tracers) could also contribute to the overall error, but are likely limited to the magnitude of the bias observed at SNP.

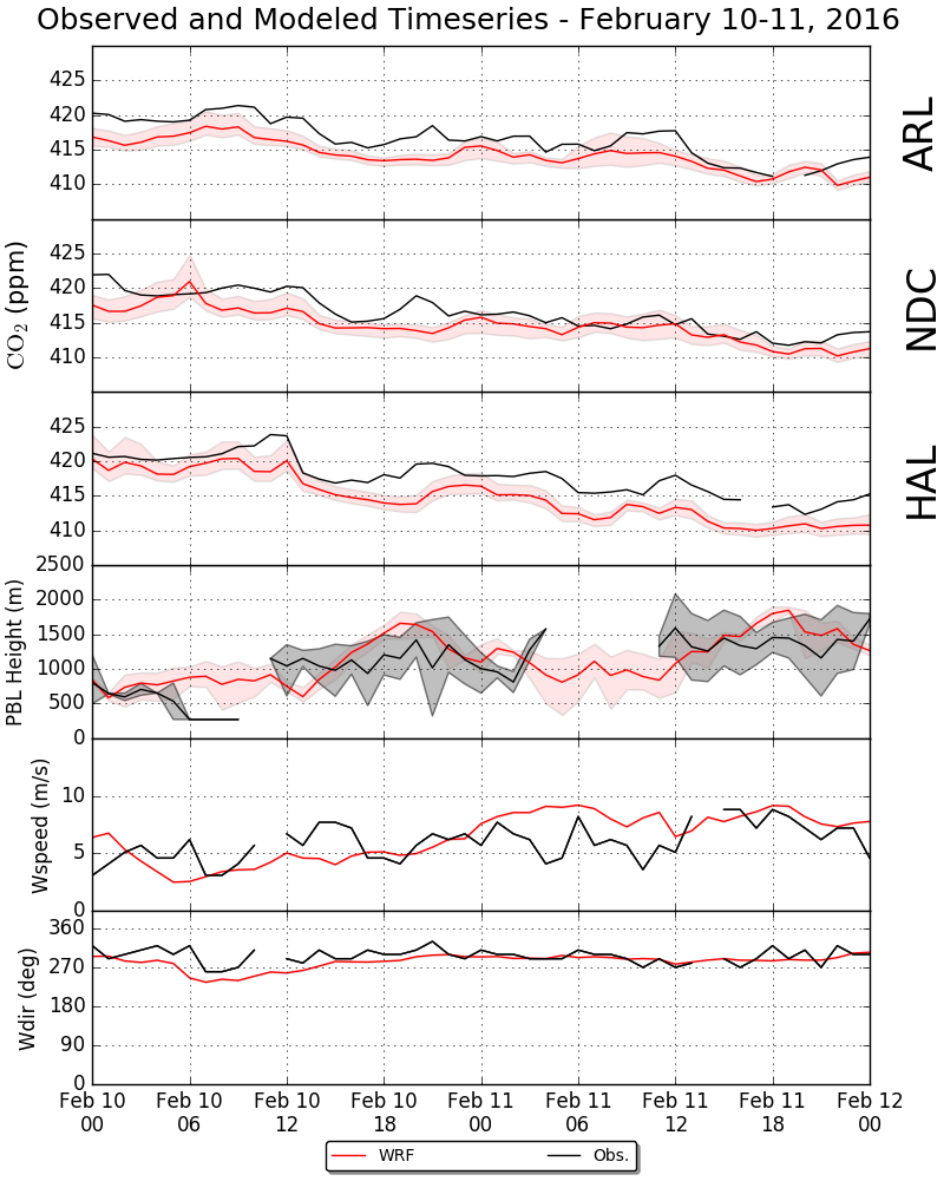

Figure 8. Observed (black) and modeled (red line average; shaded red spread of the five emissions inventories) hourly averaged CO_2 mole fractions at all three urban sites for a typical

period with persistent winds from a rural area (February 10-11, 2016). The fourth panel shows the mean (black line) and spread (shaded gray) of the inferred PBL heights from aircraft profiles and radiosonde observations, with the YSU PBL scheme predicted PBL height from WRF at each observing site mean (red line) and spread (shaded red). Bottom two panels shows 10 m observed (black) and modeled (red) wind speed and direction at KDCA. These figures indicate that WRF-Chem is able to resolve both the wind direction and height of the PBL with reasonable skill, although deviations do occur. See the supplemental information for details on the meteorological observations used.

4 Discussion

As discussed in Section 3.1, WRF-Chem tends to underestimate hourly-averaged values of near-surface CO_2 mole fractions when compared to observations averaged over the entire month, however this is not necessarily the case when using FFDAS as shown in Fig. 5. The daily maxima in simulated mole fractions from certain emissions inventories are often high relative to observations, particularly when they occur during the overnight hours when the modeled PBL depth is under predicted. However, the larger range of predicted values during these periods as well as the overall underprediction during the afternoon hours result in lower averaged values from the model than from observed values when including all sites, hours, and inventories as shown in Fig. 5. Previous work comparing simulated CO_2 to observed time series in an urban region also found an overall low bias, with predicted levels over certain hours/days exceeding observed levels (Feng et al., 2016). There are some synoptic situations, e.g., February 10 and February 11 (Section 3.2.2), where persistent winds allow for minimal errors in predicted CO_2 across the domain over an entire diurnal cycle, not just in the afternoon hours. During the afternoon of February 10, in fact, WRF overpredicts the PBL height, and thus underpredicts the near-surface CO_2. On this day, the modeled CO_2 may be more representative of reality during the overnight hours than it is in the afternoon. Conversely, other days, such as February 3 (Section 3.2.1), with a passing mid-latitude cyclone and its associated fronts, yield much different results. During this case, because of the predicted wind shift timing and the magnitude of the PBL height varying from observations, WRF tends to either overestimate or underestimate near-surface CO_2 depending on the hour.

In addition to the synoptic meteorology, variations in emissions inventories are also reflected in the predicted CO_2 mole fractions as shown in the mean biases described in Fig. 5. For the outermost domain (d01), the areal sum of the hourly averaged emissions is similar in magnitude (all within 10 % of the mean of the 4 inventories) (Fig. 2; Table 2). However, in the innermost domain (d03), the areal sum of FFDAS is over 36 % higher than the mean and the lowest inventory EDGAR is 20% below the mean of the inventories. Even though on the national scale each inventory is similar, there can be substantial differences between them due to the emission disaggregation methods (e.g., Hutchins et al. 2016; Oda et al. 2018) when considering mesoscale modeling of CO_2, a problem being studied further in other works (e.g., Fischer et al., 2017). The differences in totals and local sources could also be attributed to differing methodologies and datasets included in each emissions inventory, including the exact location of point sources and grid cell locations, among other things. These differences are generally reflected in the simulated CO_2 levels, with FFDAS being the highest averaged over the entire month, and EDGAR the lowest.

Traditionally, atmospheric inversions utilize data and meteorological model output from afternoon hours (12 pm to 4 pm local time) only (Kort et al., 2013; Breón et al., 2015; Lauvaux et al., 2016; McKain et al., 2015; Sargent et al., 2018). This is because the PBL is generally considered to be well-mixed during this time with stronger winds and higher PBL heights, allowing the relative error of these modeled quantities to be smaller during the day than during the overnight hours. By including prior emissions inventories, these inversions are used to estimate the total flux of carbon from an area of interest. However, in the forward modeling framework presented here with WRF-Chem, using afternoon hours only may not be the best metric to determine whether the model meteorology is sufficient to accurately predict CO_2 mole fractions. Figure 9 shows the mean absolute error (the dark bars) as well as the mean standard deviation (the lightly colored bars) of the five predicted CO_2 mole fractions for each time series, both for all hours (red) and afternoon hours only (blue). When considering only the three urban sites, the mean absolute error for all hours is between 3.6 ppm and 7.0 ppm across the different observing sites and inlets whereas the mean standard deviations are between 1.9 ppm and 4.1 ppm. When including afternoon hours only, the mean absolute error of the model does decrease by an average of 1.42 ppm, and the standard deviations decrease by an average of 0.58 ppm. Despite the improvement in MAE when only including afternoon hours, the mean error of the model is still approximately twice the variation in the predicted CO_2 values from each emissions inventory (as shown in the difference between MAE and standard deviation in Fig. 9). This result indicates that although limiting inversion analysis to afternoon hours may reduce overall meteorological model error it also limits analysis to time periods when local and regional emissions influence the observations the least and the differences between modeled mole fractions from various emissions inventories are smallest (due to deeper PBLs and stronger mixing).

Additionally, the mean absolute error is roughly a factor of two larger than the mean standard deviation at the urban sites, with the ratio of the two ranging from 1.64 to 2.58 for all hours, and 1.57 to 1.89 for afternoon hours only, depending on the site and inlet level. This result suggests that on average, factors common to all five tracers (meteorological error, background error, or error in the biosphere tracer) contribute more to the overall model performance than the choice of anthropogenic emissions inventory. Given the low bias (~2 ppm) at SNP, the extent of our largest WRF domain, and small relative contribution of the VEGAS tracer to the monthly mean values, we expect that the errors shown in Fig. 9 are dominated by meteorological conditions during winter. This conclusion is further supported by the results of prior studies (Kretschmer et al., 2014; Locatelli et al., 2015) as well as the two contrasting examples illustrated in Section 3.2. However, to fully validate this conclusion, an experiment would need to be performed in contrast to what is presented here, where the emissions inventory chosen is held constant and meteorological transport is varied in an ensemble of simulations. It is also important to note that these meteorological conditions or errors can exacerbate the differences in the emissions inventories as well, like shown in Fig. 6.

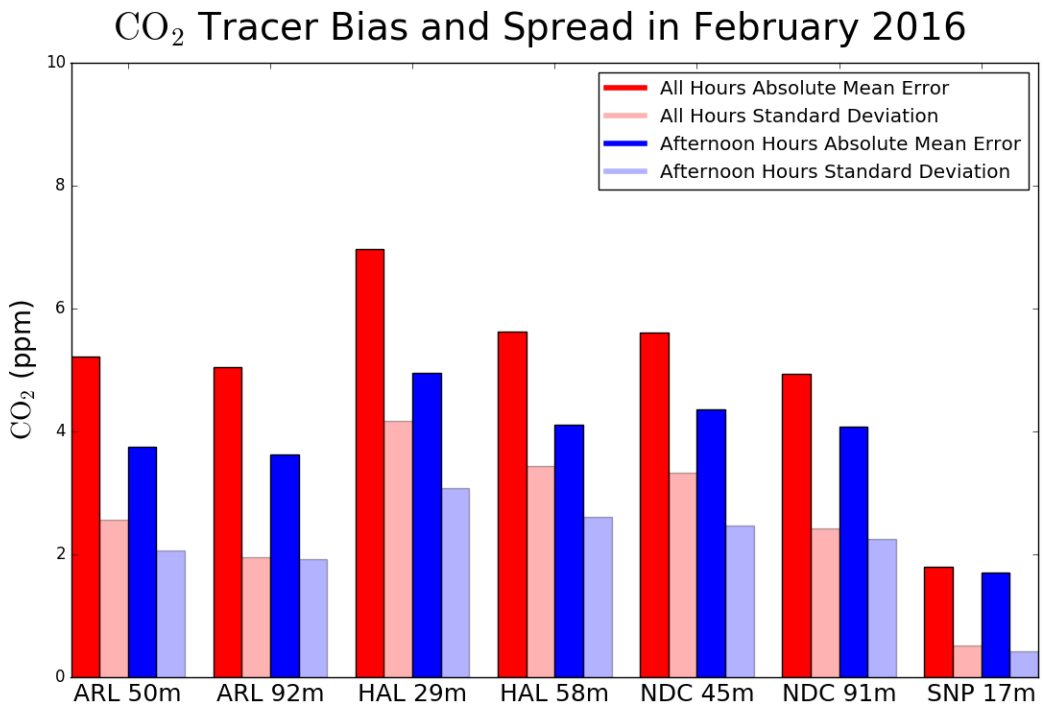

Figure 9. The mean absolute error (the dark bars) and the mean standard deviation (the lightly colored bars) of the five predicted CO_2 mole fractions for each observing site and inlet height (where applicable), both for all hours (red) and afternoon hours only (blue).

Rather than looking at the mean absolute error for the entire month and the mean standard deviation, it is also useful to consider the mean absolute error of the five tracers compared to the observed mole fraction at each hour and the associated standard deviation to get an idea of how the error and spread of the modeled CO_2 are related. These two values are reasonably related with R^2 values of approximately 0.3 across each of the three urban sites over the month. At times, generally when the modeled CO_2 is at its highest, approximately a factor of two difference between the highest and lowest modeled enhancements can be found. This is also when transport model errors tend to be largest such as during the overnight hours or frontal passages. But for many other cases, and on average as described above, the differences between the various emissions inventory tracers are smaller than the absolute error relative to observations. During periods of low modeled CO_2 error (i.e. small differences from the observations), the variation among the different emissions models is small but still discernable (as shown in Fig. 8 for example). Regardless, the error in modeled CO_2 compared to observations for any given hour or day appears to be influenced more by common factors such as meteorological error than the differences among the various emissions inventories. This suggests that atmospheric inversions need to attempt to quantify errors in modeled transport and dispersion as well as the uncertainty in the surface fluxes. It also suggests the need for methods that identify time periods when meteorological conditions are best represented by simulations (such as the case presented in Sect. 3.2.2 compared to that in Sect. 3.2.1), rather than only including afternoon hours, and when it may be anticipated that inversions will perform with higher fidelity to actual emissions conditions. The meteorological skill could potentially be improved by using analysis nudging or

data assimilation techniques, but for a forward transport modeling study such as this, there are some potential difficulties that could arise from assimilating both meteorological variables and chemical constituents (Bocquet et al., 2015). This is particularly important for urban applications where the impact of synoptic variability may strongly and adversely impact inversion analyses.

5 Conclusions

An evaluation of WRF-Chem simulated CO_2 mole fractions using multiple anthropogenic CO_2 emissions inventories at four CO_2 observing sites in the Baltimore, MD and Washington, DC metropolitan areas was presented above. For all emissions inventories the modeled CO_2 is within 5 ppm of observations when averaged over all observing sites for the month of February in 2016. However, for any given hour, at any particular site, the differences between the ensemble of simulated CO_2 values and the observed CO_2 can vary from near zero to as high as 100 ppm (the left panels of Fig. 3). The differences between the simulated time series for the different emissions inventories vary significantly in time, but tend to be proportional to the magnitude of the enhancement over the background CO_2 value. When averaged over the entire month all simulated CO_2 mole fractions are within 8 ppm of each other (Fig. 5) representing a range of approximately 2% of the total mole fraction.

This analysis suggests that the predicted mole fraction error relative to observations is dominated by model meteorology and not the underlying emissions inventory in winter months when looking at individual observing sites. Not only do certain synoptic setups allow for minimum absolute errors in the predicted values, but the timing and location of frontal passages can significantly impact the model performance at predicting CO_2 mole fractions. We also find that the errors associated with atmospheric transport are not restricted to certain times of day. This suggests that filtering data based on model performance rather than time of day (such as using only mid-afternoon observations) for atmospheric inversions might yield better overall results. Thus, further methods, such as machine learning algorithms, are needed to better identify time periods where the simulated transport performs well. To improve the simulated CO_2 mole fractions error relative to observations, the prediction of key meteorological variables such as wind speed and direction and the height of the PBL must be improved, either through more advanced physics schemes or through data assimilation techniques. As such, minimizing errors associated with atmospheric transport and dispersion generally will improve the performance of estimated fossil fuel CO_2 emissions more than improving emission priors.

Acknowledgments

We acknowledge support for this project from the FLAGG-MD grant from NIST's Greenhouse Gas Measurements program (Cooperative Agreement #70NANB14H333). We would like to thank A. Andrews (NOAA) and S. de Wekker (UVA) for use of the in situ observations of CO_2 at SNP. We also would like to thank the team at Earth Networks for their help and support. Certain commercial equipment, instruments, or materials are identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose. The observational data for SNP are available from NOAA, and the observations from the other sites, as well as the modeled time series, are available from NIST at http://data.nist.gov (doi TBD).

References

Andrews, A. E., Kofler, J. D., Trudeau, M. E., Williams, J. C., Neff, D. H., Masarie, K. A., Chao, D. Y., Kitzis, D. R., Novelli, P. C., Zhao, C. L., Dlugokencky, E. J., Lang, P. M., Crotwell, M. J., Fischer, M. L., Parker, M. J., Lee, J. T., Baumann, D. D., Desai, A. R., Stanier, C. O., De Wekker, S. F. J., Wolfe, D. E., Munger, J. W., and Tans, P. P.: CO_2, CO, and CH_4 measurements from tall towers in the NOAA Earth System Research Laboratory's Global Greenhouse Gas Reference Network: instrumentation, uncertainty analysis, and recommendations for future high-accuracy greenhouse gas monitoring efforts, Atmos. Meas. Tech., 7, 647-687, 10.5194/amt-7-647-2014, 2014.

Asefi-Najafabady, S., Rayner, P. J., Gurney, K. R., McRobert, A., Song, Y., Coltin, K., Huang, J., Elvidge, C., and Baugh, K.: A multiyear, global gridded fossil fuel CO2 emission data product: Evaluation and analysis of results, Journal of Geophysical Research: Atmospheres, 119, 10,213-210,231, 10.1002/2013JD021296, 2014.

Beck, V., T. Koch, R. Kretschmer, J. Marshall, R. Ahmadov, C. Gerbig, D. Pillai, and M. Heimann, 2011: The WRF Greenhouse Gas Model (WRF-GHG) Technical Report No. 25, Max Planck Institute for Biogeochemistry, Jena, Germany, available online at http://www.bgc-jena.mpg.de/bgc-systems/index.shtml

Bocquet, M., Elbern, H., Eskes, H., Hirtl, M., Žabkar, R., Carmichael, G. R., Flemming, J., Inness, A., Pagowski, M., Pérez Camaño, J. L., Saide, P. E., San Jose, R., Sofiev, M., Vira, J., Baklanov, A., Carnevale, C., Grell, G., and Seigneur, C.: Data assimilation in atmospheric chemistry models: current status and future prospects for coupled chemistry meteorology models, Atmos. Chem. Phys., 15, 5325-5358, https://doi.org/10.5194/acp-15-5325-2015, 2015.

Breon, F. M., Broquet, G., Puygrenier, V., Chevallier, F., Xueref-Remy, I., Ramonet, M., Dieudonne, E., Lopez, M., Schmidt, M., Perrussel, O., and Ciais, P.: An attempt at estimating Paris area CO2 emissions from atmospheric concentration measurements, Atmospheric Chemistry and Physics, 15, 1707-1724, 10.5194/acp-15-1707-2015, 2015.

Briber, B., Hutyra, L., Dunn, A., Raciti, S., and Munger, J.: Variations in Atmospheric CO2 Mixing Ratios across a Boston, MA Urban to Rural Gradient, Land, 2, 304, 2013.

Carbontracker Team: Compilation of near real time atmospheric carbon dioxide data provided by NOAA and EC; obspack_co2_1_NRT_v3.3_2017-04-19; NOAA Earth System Research Laboratory, Global Monitoring Division. http://doi.org/10.15138/G3G01J, 2017.

Chou, M.-D., Suarez, M. J., Liang, X.-Z., Yan, M. M.-H., and Cote, C.: A thermal infrared radiation parameterization for atmospheric studies, 2001.

Ciais P, Rayner P, Chevallier F, Bousquet P, Logan M, Peylin P, Ramonet M (2010) Atmospheric inversions for estimating CO_2 fluxes: methods and perspectives. Clim Change 103(1-2):69-92. https://doi.org/10.1007/s10584-010-9909-3

Clarke L., K. Jiang, K. Akimoto, M. Babiker, G. Blanford, K. Fisher-Vanden, J.-C. Hourcade, V. Krey, E. Kriegler, A. Löschel, D. McCollum, S. Paltsev, S. Rose, P.R. Shukla, M. Tavoni, B.C.C. van der Zwaan, and D.P. van Vuuren: Assessing Transformation Pathways. In: Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Edenhofer, O., R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P.

837 Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel and J.C.
838 Minx (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY,
839 USA, 2014.

840 Deng, A., T. Lauvaux, K. J. Davis, B. J. Gaudet, N. Miles, S. J. Richardson, K. Wu, D. P.
841 Sarmiento, R. M. Hardesty, and T. A. Bonin (2017), Toward reduced transport errors in a high
842 resolution urban CO2 inversion system, Elem Sci Anth, 5.

843 Dlugokencky, E. J., Myers, R. C., Lang, P. M., Masarie, K. A., Crotwell, A. M., Thoning, K. W.,
844 Hall, B. D., Elkins, J. W., and Steele, L. P.: Conversion of NOAA atmospheric dry air CH_4 mole
845 fractions to a gravimetrically prepared standard scale, Journal of Geophysical Research:
846 Atmospheres, 110, doi:10.1029/2005JD006035, 2005.

847 Etheridge, D. M., L. P. Steele, R. L. Langenfelds, R. J. Francey, J.-M. Barnola, and V. I. Morgan
848 (1996), Natural and anthropogenic changes in atmospheric CO_2 over the last 1000 years from air
849 in Antarctic ice and firn, J. Geophys. Res., 101(D2), 4115–4128, doi: 10.1029/95JD03410.

850 Feng, S., Lauvaux, T., Newman, S., Rao, P., Ahmadov, R., Deng, A. J., Diaz-Isaac, L. I., Duren,
851 R. M., Fischer, M. L., Gerbig, C., Gurney, K. R., Huang, J. H., Jeong, S., Li, Z. J., Miller, C. E.,
852 O'Keeffe, D., Patarasuk, R., Sander, S. P., Song, Y., Wong, K. W., and Yung, Y. L.: Los Angeles
853 megacity: a high-resolution land-atmosphere modelling system for urban CO2 emissions,
854 Atmospheric Chemistry and Physics, 16, 9019-9045, 10.5194/acp-16-9019-2016, 2016.

855 Fischer, M. L., Parazoo, N., Brophy, K., Cui, X., Jeong, S., Liu, J., Keeling, R., Taylor, T. E.,
856 Gurney, K., Oda, T., and Graven, H.: Simulating estimation of California fossil fuel and
857 biosphere carbon dioxide exchanges combining in situ tower and satellite column observations,
858 Journal of Geophysical Research: Atmospheres, 122, 3653-3671, 10.1002/2016JD025617, 2017.

859 Gately, C. K., & Hutyra, L. R. (2017). Large uncertainties in urban-scale carbon emissions.
860 Journal of Geophysical Research: Atmospheres, 122, 11,242–11,260.
861 https://doi.org/10.1002/2017JD027359

862 Gemmill, W., Katz, B., and Li, X.: Daily real-time global sea surface temperature-high
863 resolution analysis at NOAA/NCEP, NCEP Off. Note, 260, 39, 2007.

864 Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and
865 Eder, B.: Fully coupled "online" chemistry within the WRF model, Atmospheric Environment,
866 39, 6957-6975, 10.1016/j.atmosenv.2005.04.027, 2005.

867 Gurney, K. R., Mendoza, D. L., Zhou, Y., Fischer, M. L., Miller, C. C., Geethakumar, S., and de
868 la Rue du Can, S.: High Resolution Fossil Fuel Combustion CO2 Emission Fluxes for the United
869 States, Environmental Science & Technology, 43, 5535-5541, 10.1021/es900806c, 2009.

870 Hong, S.-Y., Dudhia, J., and Chen, S.-H.: A Revised Approach to Ice Microphysical Processes
871 for the Bulk Parameterization of Clouds and Precipitation, Monthly Weather Review, 132, 103-
872 120, 10.1175/1520-0493(2004)132<0103:aratim>2.0.co;2, 2004.

873 Hong, S.-Y., Noh, Y., and Dudhia, J.: A New Vertical Diffusion Package with an Explicit
874 Treatment of Entrainment Processes, Monthly Weather Review, 134, 2318-2341,
875 10.1175/mwr3199.1, 2006.

876 Hutchins, M. G., Colby, J. D., Marland, G., and Marland, E.: A comparison of five high-
877 resolution spatially-explicit, fossil-fuel, carbon dioxide emission inventories for the United

States, Mitigation and Adaptation Strategies for Global Change, 1-26, 10.1007/s11027-016-9709-9, 2016.

IEA, 2012. CO_2 emissions from fuel combustion 1971–2010, 2012 Edition, Paris: International Energy Agency (IEA), 2012.

Intergovernmental Panel on Climate Change (2013). Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, In T. F. Stocker, et al. (Eds.), (1535 pp.). Cambridge, U. K., and New York: Cambridge University Press. https://doi.org/10.1017/CBO9781107415324

Kain, J. S.: The Kain–Fritsch Convective Parameterization: An Update, Journal of Applied Meteorology, 43, 170-181, 10.1175/1520-0450(2004)043<0170:tkcpau>2.0.co;2, 2004.

Kort, E. A., Angevine, W. M., Duren, R., and Miller, C. E.: Surface observations for monitoring urban fossil fuel CO2 emissions: Minimum site location requirements for the Los Angeles megacity, Journal of Geophysical Research-Atmospheres, 118, 1-8, 10.1002/jgrd.50135, 2013.

Kretschmer, R., C. Gerbig, U. Karstens, G. Biavati, A. Vermeulen, F. Vogel, S. Hammer, and K. U. Totsche: Impact of Optimized Mixing Heights on Simulated Regional Atmospheric Transport of CO_2, Atmos. Chem. Phys. 14, no. 14 (July 16, 2014): 7149-72. https://doi.org/10.5194/acp-14-7149-2014, 2014.

Lauvaux, T., N.L. Miles, S.J. Richardson, A. Deng, D.R. Stauffer, K.J. Davis, G. Jacobson, C. Rella, G. Calonder, and P.L. DeCola, 2013: Urban Emissions of CO_2 from Davos, Switzerland: The First Real-Time Monitoring System Using an Atmospheric Inversion Technique. J. Appl. Meteor. Climatol., 52, 2654–2668, https://doi.org/10.1175/JAMC-D-13-038.1

Lauvaux, T., Miles, N. L., Deng, A., Richardson, S. J., Cambaliza, M. O., Davis, K. J., Gaudet, B., Gurney, K. R., Huang, J., and O'Keefe, D.: High-resolution atmospheric inversion of urban CO2 emissions during the dormant season of the Indianapolis Flux Experiment (INFLUX), Journal of Geophysical Research: Atmospheres, 2016.

Lee, T. R., De Wekker, S. F. J., Andrews, A. E., Kofler, J., and Williams, J.: Carbon dioxide variability during cold front passages and fair weather days at a forested mountaintop site, Atmospheric Environment, 46, 405-416, https://doi.org/10.1016/j.atmosenv.2011.09.068, 2012.

Leip, A., Skiba, U., Vermeulen, A., and Thompson, R. L.: A complete rethink is needed on how greenhouse gas emissions are quantified for national reporting, Atmospheric Environment, 174, 237-240, https://doi.org/10.1016/j.atmosenv.2017.12.006, 2018.

Locatelli, R., P. Bousquet, M. Saunois, F. Chevallier, and C. Cressot: Sensitivity of the Recent Methane Budget to LMDz Sub-Grid-Scale Physical Parameterizations, Atmospheric Chemistry and Physics 15, no. 17 (September 1, 2015): 9765-80. https://doi.org/10.5194/acp-15-9765-2015, 2015.

Lopez-Coto, I., Ghosh, S., Prasad, K., and Whetstone, J.: Tower-based greenhouse gas measurement network design—The National Institute of Standards and Technology North East Corridor Testbed, Advances in Atmospheric Sciences, 34, 1095-1105, 10.1007/s00376-017-6094-6, 2017.

McDonald, B. C., Z. C. McBride, E. W. Martin, and R. A. Harley (2014), High-resolution mapping of motor vehicle carbon dioxide emissions, J. Geophys. Res. Atmos., 119, 5283–5298, doi: 10.1002/2013JD021219.

McKain, K., Wofsy, S. C., Nehrkorn, T., Eluszkiewicz, J., Ehleringer, J. R., and Stephens, B. B.: Assessment of ground-based atmospheric observations for verification of greenhouse gas emissions from an urban region, Proceedings of the National Academy of Sciences of the United States of America, 109, 8423-8428, 10.1073/pnas.1116645109, 2012.

McKain, K., Down, A., Raciti, S. M., Budney, J., Hutyra, L. R., Floerchinger, C., Herndon, S. C., Nehrkorn, T., Zahniser, M. S., Jackson, R. B., Phillips, N., and Wofsy, S. C.: Methane emissions from natural gas infrastructure and use in the urban region of Boston, Massachusetts, Proceedings of the National Academy of Sciences, 112, 1941-1946, 10.1073/pnas.1416261112, 2015.

McRae, J. E., and Graedel, T. E.: Carbon Dioxide in the Urban Atmosphere: Dependencies and Trends, J. Geophys. Res.-Oceans, 84, 5011-5017, 10.1029/JC084iC08p05011, 1979.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., Jović, D., Woollen, J., Rogers, E., Berbery, E. H., Ek, M. B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D., and Shi, W.: North American Regional Reanalysis, Bulletin of the American Meteorological Society, 87, 343-360, 10.1175/bams-87-3-343, 2006.

Miles, N. L., Richardson, S. J., Lauvaux, T., Davis, K. J., Balashov, N. V., Deng, A., Turnbull, J. C., Sweeney, C., Gurney, K. R., and Patarasuk, R.: Quantification of urban atmospheric boundary layer greenhouse gas dry mole fraction enhancements in the dormant season: Results from the Indianapolis Flux Experiment (INFLUX), Elem Sci Anth, 5, 2017.

Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., and Clough, S. A.: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave, Journal of Geophysical Research: Atmospheres, 102, 16663-16682, 10.1029/97JD00237, 1997.

Mueller, K., Yadav, V., Lopez-Coto, I., Karion, A., Gourdji, S., Martin, C., & Whetstone, J. (2018). Siting background towers to characterize incoming air for urban greenhouse gas estimation: A case study in the Washington, DC/Baltimore area. Journal of Geophysical Research: Atmospheres, 123. https://doi.org/10.1002/2017JD027364

Nassar, R., Napier-Linton, L., Gurney, K. R., Andres, R. J., Oda, T., Vogel, F. R., and Deng, F.: Improving the temporal and spatial distribution of CO2 emissions from global fossil fuel emission data sets, Journal of Geophysical Research: Atmospheres, 118, 917-933, 10.1029/2012JD018196, 2013.

Nehrkorn, T., Henderson, J., Leidner, M., Mountain, M., Eluszkiewicz, J., McKain, K., and Wofsy, S.: WRF simulations of the urban circulation in the Salt Lake City area for CO2 modeling, Journal of Applied Meteorology and Climatology, 52, 323-340, 2013.

National Oceanic and Atmospheric Administration (NOAA) Earth System Research Laboratory, Global Monitoring Division, Trends in Atmospheric Carbon Dioxide, https://www.esrl.noaa.gov/gmd/ccgg/trends/full.html, Last accessed: January 5, 2018.

Oda, T., and Maksyutov, S.: A very high-resolution (1 km×1 km) global fossil fuel CO2 emission inventory derived using a point source database and satellite observations of nighttime lights, Atmos. Chem. Phys., 11, 543-556, 10.5194/acp-11-543-2011, 2011.

961 Oda, T., and Maksyutov, S. (2015), ODIAC Fossil Fuel CO2 Emissions Dataset (Version name:
962 ODIAC2016), Center for Global Environmental Research, National Institute for Environmental
963 Studies, doi:10.17595/20170411.001. (Reference date : 2017/02/02)

964 Oda, T., Maksyutov, S., and Andres, R. J.: The Open-source Data Inventory for Anthropogenic
965 Carbon dioxide (CO2), version 2016 (ODIAC2016): A global monthly fossil fuel CO2 gridded
966 emissions data product for tracer transport simulations and surface flux inversions, Earth Syst.
967 Sci. Data, 10, 87-107, https://doi.org/10.5194/essd-10-87-2018, 2018

968 Olivier, J. G. J., Van Aardenne, J. A., Dentener, F. J., Pagliari, V., Ganzeveld, L. N., and Peters,
969 J. A. H. W.: Recent trends in global greenhouse gas emissions:regional trends 1970–2000 and
970 spatial distribution of key sources in 2000, Environmental Sciences, 2, 81-99,
971 10.1080/15693430500400345, 2005.

972 Parazoo, N. C., Denning, A. S., Kawa, S. R., Corbin, K. D., Lokupitiya, R. S., and Baker, I. T.:
973 Mechanisms for synoptic variations of atmospheric CO_2 in North America, South America and
974 Europe, Atmos. Chem. Phys., 8, 7239-7254, https://doi.org/10.5194/acp-8-7239-2008, 2008.

975 Poulida, O., R. R. Dickerson, B. G. Doddridge, J. Z. Holland, R. G. Wardell, and J. G. Watkins:
976 Trace Gas Concentrations and Meteorology in Rural Virginia: 1. Ozone and Carbon-Monoxide,
977 Journal of Geophysical Research-Atmospheres, 96(D12), 22461-22475, 1991.

978 Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J.
979 B., Bruhwiler, L. M., Petron, G., and Hirsch, A. I.: An atmospheric perspective on North
980 American carbon dioxide exchange: CarbonTracker, Proceedings of the National Academy of
981 Sciences, 104, 18925-18930, 2007.

982 Rayner, P. J., Raupach, M. R., Paget, M., Peylin, P., and Koffi, E.: A new global gridded data set
983 of CO2 emissions from fossil fuel combustion: Methodology and evaluation, Journal of
984 Geophysical Research: Atmospheres, 115, n/a-n/a, 10.1029/2009JD013439, 2010.

985 Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.-T.,
986 Chuang, H.-y., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M. P., Dool, H. v. d., Zhang, Q.,
987 Wang, W., Chen, M., and Becker, E.: The NCEP Climate Forecast System Version 2, Journal of
988 Climate, 27, 2185-2208, 10.1175/jcli-d-12-00823.1, 2014.

989 Sargent, M., Barrera, Y., Nehrkorn, T., Hutyra, L. R., Gately, C. K., Jones, T., McKain, K.,
990 Sweeney, C., Hegarty, J., Hardiman, B., and Wofsy, S. C.: Anthropogenic and biogenic CO2
991 fluxes in the Boston urban region, Proceedings of the National Academy of Sciences,
992 10.1073/pnas.1803715115, 2018.

993 Shusterman, A. A., Teige, V., Turner, A. J., Newman, C., Kim, J., and Cohen, R. C.: The
994 BErkeley Atmospheric CO_2 Observation Network: initial evaluation, Atmos. Chem. Phys.
995 Discuss., 2016, 1-23, 10.5194/acp-2016-530, 2016.

996 Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-
997 Y., Wang, W., and Powers, J. G.: A Description of the Advanced Research WRF Version 3,
998 2008.

999 Strong, C., C. Stwertka, D. R. Bowling, B. B. Stephens, and J. R. Ehleringer, Urban carbon
1000 dioxide cycles within the Salt Lake Valley: A multiple-box model validated by observations, J.
1001 Geophys. Res., 116, D15307, doi: 10.1029/2011JD015693, 2011.

Tewari, M., Chen, F., Wang, W., Dudhia, J., LeMone, M., Mitchell, K., Ek, M., Gayno, G., Wegiel, J., and Cuenca, R.: Implementation and verification of the unified NOAH land surface model in the WRF model, 20th conference on weather analysis and forecasting/16th conference on numerical weather prediction, 2004.

Thiébaux, J., Rogers, E., Wang, W., and Katz, B.: A new high-resolution blended real-time global sea surface temperature analysis, Bulletin of the American Meteorological Society, 84, 645-656, 2003.

Turnbull, J. C., Sweeney, C., Karion, A., Newberger, T., Lehman, S. J., Tans, P. P., Davis, K. J., Lauvaux, T., Miles, N. L., Richardson, S. J., Cambaliza, M. O., Shepson, P. B., Gurney, K., Patarasuk, R., and Razlivanov, I.: Toward quantification and source sector identification of fossil fuel CO2 emissions from an urban area: Results from the INFLUX experiment, Journal of Geophysical Research-Atmospheres, 120, 292-312, 10.1002/2014jd022555, 2015.

United Nations: Cities and Climate Change: Global Report on Human Settlements 2011. 2011. Available online: https://unhabitat.org/?mbt_book=cities-and-climate-change-global-report-on-human-settlements-2011 (accessed on 15 September 2018).

Verhulst, K. R., Karion, A., Kim, J., Salameh, P. K., Keeling, R. F., Newman, S., Miller, J., Sloop, C., Pongetti, T., Rao, P., Wong, C., Hopkins, F. M., Yadav, V., Weiss, R. F., Duren, R. M., and Miller, C. E.: Carbon dioxide and methane measurements from the Los Angeles Megacity Carbon Project – Part 1: calibration, urban enhancements, and uncertainty estimates, Atmos. Chem. Phys., 17, 8313-8341, 10.5194/acp-17-8313-2017, 2017.

Vogel, F. R., Frey, M., Staufer, J., Hase, F., Broquet, G., Xueref-Remy, I., Chevallier, F., Ciais, P., Sha, M. K., Chelin, P., Jeseck, P., Janssen, C., Te, Y.-V., Groß, J., Blumenstock, T., Tu, Q., and Orphal, J.: XCO_2 in an emission hot-spot region: the COCCON Paris campaign 2015, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-595, in review, 2018.

Whelpdale, D. M., Low, T. B., and Kolomeychuk, R. J.: Advection climatology for the east coast of North America, Atmospheric Environment (1967), 18, 1311-1327, 10.1016/0004-6981(84)90040-4, 1984.

Wunch, D., G. C. Toon, P. O. Wennberg, S. C. Wofsy, B. Stephens, M. L. Fisher, O. Uchino, J. B. Abshire, P. F. Bernath, S. C. Biraud, J.-F. L. Blavier, C. D. Boone, K. P. Bowman, E. V. Browell, T. Campos, B. J. Connor, B. C. Daube, N. M. Deutscher, M. Diao, J. W. Elkins, C. Gerbig, E. Gottlieb, D. W. T. Griffith, D. F. Hurst, R. Jiménez, G. Keppel-Aleks, E. A. Kort, R. Macatangay, T. Machida, H. Matsueda, F. L. Moore, I. Morino, S. Park, J. Robinson, C. M. Roehl, Y. Sawa, V. Sherlock, C. Sweeney, T. Tanaka, and M. A. Zondlo (2010), Calibration of the Total Carbon Column Observing Network using aircraft profile data, Atmospheric Measurement Techniques, 3(5), 1351-1362, doi:10.5194/amt-3-1351-2010.

Zeng, N., Mariotti, A., and Wetzel, P.: Terrestrial mechanisms of interannual CO2 variability, Global Biogeochemical Cycles, 19, 2005.

Zhao, C. L., and Tans, P. P.: Estimating uncertainty of the WMO mole fraction scale for carbon dioxide in air, Journal of Geophysical Research: Atmospheres, 111, doi:10.1029/2005JD006003, 2006.