# NIST Special Publication 1237

# Carbon Dioxide Biogenic vs Anthropogenic Sectoral Contribution for the Indianapolis Flux Experiment (INFLUX)

Israel Lopez-Coto
Kuldeep Prasad
James R. Whetstone

**NIST**
**National Institute of Standards and Technology**
U.S. Department of Commerce

# NIST Special Publication 1237

# Carbon Dioxide Biogenic vs Anthropogenic Sectoral Contribution for the Indianapolis Flux Experiment (INFLUX)

Israel Lopez-Coto
Kuldeep Prasad
*NIST Engineering Laboratory*

James R. Whetstone
*NIST Office of Special Programs*

January 2017

**Abstract**

The Indianapolis Flux Experiment (INFLUX) aims to use a top-down inversion methodology to quantify sources of Greenhouse Gas (GHG) emissions over an urban domain with high spatial and temporal resolution. Analyzing the contribution of different source types or sectors is a fundamental step in order to achieve an accuracy level desired for such applications. To this end, the Weather Research and Forecasting model (WRF-ARW) was used along with a modified version of the GHG chemistry module for simulating the carbon dioxide ($CO_2$) mole fraction transport during September and October 2013. Sectoral anthropogenic $CO_2$ emissions were obtained from Hestia 2013 and from Vulcan 2002. Biogenic $CO_2$ emissions were simulated by using an augmented version of the "Vegetation Photosynthesis and Respiration Model" (VPRM) included in WRF-CHEM. The atmospheric model performed well in capturing day to day variability of the concentration. The errors were largely reduced during the well-mixed conditions indicating that a large share of the errors was due to the misrepresentation of nighttime stable conditions in the model, causing an overestimation of the nighttime peaks. In general, the averaged daily cycle was well represented by the model. In addition, 4 towers presented a slightly delayed collapse of the nocturnal stable boundary layer. Last, the biogenic contribution was important, being larger than the anthropogenic contribution for 7 out of 12 towers during this period. In general, the anthropogenic contribution was dominated by the Electricity Production, 6.5 % to 39 %, Mobility, 11 % to 56 %, and Industrial sectors, 4.8 % to 10 %.

**Key words**

Atmospheric Transport, Biogenic vs Anthropogenic, Carbon Dioxide, Emissions, INFLUX.

# Table of Contents

# List of Tables

# List of Figures

## 1. Introduction

Carbon dioxide ($CO_2$) is the major long-lived, anthropogenic greenhouse gas (GHG) that has substantially increased in the atmosphere since the industrial revolution due to human activities, raising serious climate and sustainability issues [1]. Development of methods for determination of GHG flows to and from the atmosphere independent of those used to develop GHG inventory data and reports will enhance that data scientific basis thereby increasing confidence in them.

Cities play an important role in emissions mitigation and sustainability efforts because they intensify energy utilization and greenhouse gas emissions in geographically small regions. Urban areas are estimated to be responsible for over 70% of global energy-related carbon emissions [2]. This percentage is anticipated to grow as urbanization trends continue; cities will likely contain 85% – 90% of the U.S. population by century's end. Urban carbon studies have increased in recent years with diverse motivations ranging from urban ecology research to testing methods for independently verifying GHG emission inventory reports and estimates. Examples of these are Salt Lake City [3], Houston [4], Paris [5], Los Angeles [6] and Indianapolis (INFLUX; [7, 8]).

The Indianapolis Flux Experiment (INFLUX) aims to use a top-down inversion methodology to quantify sources of Greenhouse gas (GHG) emissions over an urban domain with high spatial and temporal resolution. This project is an experimental test bed which is intended to establish reliable methods for quantifying and validating GHG emissions independently of the inventory methods. Analyzing the contribution of different source types or sectors is a fundamental step in order to achieve an accuracy level desired for such applications. This is especially challenging when attempting to determine anthropogenic emissions during the growing season since biological GHG fluxes reach a maximum at this time.

Main goals of this work were to test the ability of a high resolution model (WRF) to reproduce the atmospheric $CO_2$ variability within and around an urban domain (Indianapolis) and to identify the contribution to the atmospheric $CO_2$ coming from vegetation during the end of the growing season as well as the contribution of the different anthropogenic sectors.

The rest of the paper is structured as follow. Section 2 describes the model configurations, the datasets used in analysis nudging, the biogenic $CO_2$ emissions model (VPRM) and the optimization method for its parameters, the sectoral anthropogenic emissions and the towers measurements. Section 3 presents the results obtained. First, the meteorological verification is discussed and one model configuration is selected. Then, the VPRM optimization results are discussed. Afterwards, the simulated $CO_2$ is compared to observations at 10 towers for two months and last, the averaged sectoral contribution is described. Section 4 highlight the main conclusions obtained in this work.

## 2.  Methodology

### 2.1   Weather Research and Forecasting Model (WRF)

We simulated the $CO_2$ atmospheric transport by using version 3.5.1 of the Weather Research and Forecasting model with coupled chemistry (WRF-Chem) [9], which uses the Advanced Research WRF dynamical core [10]. Specifically, the module GHG was used to simulate the transport of $CO_2$ tracers. WRF-GHG allows the "online" passive tracer transport of $CO_2$, CO and $CH_4$. The tracers undergo the advection, boundary layer and convective mixing as the chemical species but with no chemical reactions [11, 12]. The module includes a model (VPRM) to simulate the $CO_2$ biogenic emissions (described in the next section). WRF-GHG includes tracers to track the biogenic emissions as well as the anthropogenic emissions. However, the anthropogenic emissions are tracked in a single tracer while our focus is to track the emissions by sectors. Thus, several additional tracers (8) were implemented in the module to allow us to track independently the emissions coming from Aircraft (AIR), Cement (CEM), Commercial (COM), Industrial (IND), Nonroad (NON), Onroad (MOB), Residential (RES) and Electricity Production (UTL) sectors.

A 'two-way nesting' strategy (with feedback) was selected for downscaling the three telescoping domains that have 9, 3 and 1 km horizontal resolutions. These domains are centered on Indianapolis, Indiana, (N39.79999, W86.19999) and they have 101 x 101, 100 x 100 and 88 x 88 horizontal grid cells, respectively. This domain configuration was chosen in order to limit the influence of the North American Regional Reanalysis (NARR, 32 km resolution) provided boundary conditions on the area of interest. A configuration of 60 vertical levels with higher resolution between the surface and 3 km was selected to better reproduce the boundary layer dynamics. To ensure model stability, the time-step size was defined dynamically using a Courant-Friedrichs-Lewy (CFL) criterion of 1. Initial (0000 UTC) and boundary conditions (each 3 hours) were taken from NARR data provided by NCEP. Simulations were run continuously for 2 months, September and October 2013, with the first 24 hours being used for spinning up the model.

Accurately reproducing the planetary boundary layer (PBL) structure is a key point in atmospheric transport models since the species mixing within the boundary layer are primarily driven by the turbulent structures found there. Therefore, the Mellor-Yamada-Nakanishi-Niino 2.5-level (MYNN2, [13, 14]) was selected as PBL parametrization, because this scheme is a local PBL scheme that diagnoses potential temperature variance, water vapor mixing ratio variance, and their covariances to solve a prognostic equation for the turbulent kinetic energy (TKE). It is an improved version of the former Mellor-Yamada-Janjic scheme, (MYJ, [15, 16]) where the stability functions and mixing length formulations are based on large eddy simulation (LES) results instead of observational datasets. MYNN2 has been shown to be nearly unbiased in PBL depth, moisture, and potential temperature in convection-allowing configurations of WRF-ARW. This alleviates the typical cool, moist bias of the MYJ scheme in convective boundary layers upstream from convection [17]. The RRTMG scheme [18] for short and long wave radiation was selected as the radiative heat transfer scheme since good performances has been reported

2

[19]. To model the microphysics, we selected the Thompson scheme [20, 21] because of its improved treatment of the water/ice/snow effective radius coupled to the radiative scheme (RRTMG). For the cumulus cloud scheme, the widely used Kain-Fritsch scheme [22] was selected only for the outermost domain (9 km). The Noah model was selected as Land Surface Model (LSM) since it has been showed the Noah-LSM simulated skin temperature and energy fluxes are reasonably comparable with observations [23] and, thus, an acceptable feedback to the PBL scheme is expected.

We updated the land use by using a more recent version of the IGBP land surface cover types provided by the Moderate-resolution Imaging Spectroradiometer (MODIS) dataset MCD12Q1 [24, 25] for the year 2012.

We tested two dataset to be used in analysis nudging in WRF. First, the NCEP ADP Global Upper Air Observational Weather Data (http://rda.ucar.edu/datasets/ds351.0/) and NCEP ADP Global Surface Observational Weather Data (http://rda.ucar.edu/datasets/ds461.0/) provided by the Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory (RDA-CISL) and second, the Meteorological Assimilation Data Ingest System (MADIS, https://madis.ncep.noaa.gov/) a National Oceanic and Atmospheric Administration (NOAA) program that collects, integrates, quality-controls, and distributes observations from NOAA and other organizations. MADIS observations were converted to "little_r" format by using the MADIS2LITTLER converter provided by National Center for Atmospheric Research (NCAR).

First, an objective analysis was applied in order to blend the boundary conditions towards the observations. We used here the OBSGRID tool provided by WRF. Buddy check and maximum difference check quality control methods were applied and the observations with not satisfactory quality control flag were discarded. Cressman scheme [26] was selected as objective analysis method with radius of influence being automatically selected by the code on the base of the available observations at the specific time. Then, grid nudging was performed in the 3 domains. For profile observations (3d nudging), just the wind was nudged within the PBL. A nudging coefficient of 0.0003 s$^{-1}$, a 1 h of time ramp and an updating frequency of 3 h were used (as the boundary conditions). Wind, temperature and water vapor mixing ratio for surface observations (surface nudging) were nudged with the same nudging coefficient and time ramp than before but with an updating cycle of 1h. Radius of influence was set to 25 km.

In total, we ran 6 WRF configurations: A control run with no nudging and the original land use dataset; RDA, using RDA dataset for analysis nudging and the original land use dataset; MADIS, using MADIS for analysis nudging and the original land use; MODIS, using the updated land use but no nudging; MODIS + RDA, using the updated land use and nudging with RDA dataset and MODIS + MADIS, using updated land use and nudging with MADIS dataset.

Fifteen surface stations (METAR) were used to compare temperature, wind speed and direction with simulations, (Figure 1).

## 2.2 Biogenic $CO_2$ emissions

The vegetation photosynthesis and respiration model (VPRM) of Xiao et al., [27] and Mahadevan et al., [28] is a simple diagnostic terrestrial flux model that is included in WRF-Chem GHG module [11, 12]. In spite of its simplicity, VPRM captures daily and annual cycles in $CO_2$ fluxes reasonably well. VPRM structure and skills are described in great detail by Mahadevan et al., [28] and Hilton et al., [29]. VPRM calculates Net Ecosystem Exchange (NEE) as a sum of Gross Ecosystem Exchange (GEE) and ecosystem Respiration (R). In WRF-GHG implementation [30, 31], VPRM calculates GEE for 8 vegetation categories using (1) shortwave radiative flux (SWDOWN) and temperature at 2 meters (T2) provided by the WRF simulation; (2) enhanced vegetation index (EVI), which represents the fraction of shortwave radiation absorbed by leaves; and (3) the land surface water index (LSWI), which reflects changes in both leaf water content and soil moisture and accounts for the effects of leaf age on photosynthesis at the canopy level [27]. Respiration fluxes are calculated as a linear function of T2 from WRF [28].

LSWI is derived from the Moderate-resolution Imaging Spectroradiometer (MODIS) reflectances from the dataset MCD43A4 [32]. EVI is obtained from the dataset M*D13A2 [33] and IGBP land surface cover types are from dataset MCD12Q1 [24, 25]. The "*" in M*D is either "O", representing the data from the Terra satellite, or "Y" representing the data from the Aqua satellite. The "C" represents the composite product including data from both satellites.

IGBP land use groups all the croplands into the same category and thus does VPRM as implemented in WRF. However, the abundance of soybean and corn fields surrounding Indianapolis and the very different photosynthesis and respiration of these two crops imposes the need of adding an explicit distinction into the model. Thus, we added an extra vegetation category into the WRF-VPRM implementation to account for the presence of corn fields. We mapped the corn vegetation fraction from the USDA National Agricultural Statistics Service Cropland Data Layer for the year 2013.

VPRM has four user-estimated parameters for each vegetation category ($\lambda$, $PAR_0$, $\alpha$, $\beta$). The first two are related to the GEE and the last two are related to the respiration. We optimized the parameters of VPRM for the 3 dominant vegetation categories in our computational domain (Deciduous broadleaf forest, Soybean and Corn) which account for more than 95 % of the total area. We optimized the four parameters at the same time by minimizing the cost function in Eq. 1, using an unconstrained nonlinear optimization method [34] from the Stats package for the R language and platform for statistical computing [35].

$$score = \mid BIAS \mid + RMSD + 10 \cdot LEPS \tag{1}$$

$$BIAS = \frac{1}{N} \sum_{i=1}^{N} \left( x_i - x_i^{observed} \right) \tag{2}$$

4

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left( x_i - x_i^{observed} \right)^2} \tag{3}$$

$$LEPS = \frac{1}{N} \sum_{i=1}^{N} \left( ECDF(x) - ECDF(x^{observed}) \right)_i \tag{4}$$

Where x is the simulated NEE, $x^{observed}$ is the measured NEE and ECDF is the empirical cumulative distribution function.

The parameters were estimated by using the observed NEE data obtained from eddy covariance flux towers from AmeriFlux network, Table 1. The two locations selected are the closest location to the area of interest. The optimization period was September and October for the 3 vegetation categories. However, only Monroe State Forest had data for 2013. The site Fermi Agricultural had a Corn/Soybean rotation and therefore we used the corresponding year to optimize for each category.

**Table 1.** Ameriflux towers used for optimizing the VPRM parameters

| Name | Latitude | Longitude | Vegetation category | year |
|------|----------|-----------|---------------------|------|
| Monroe State Forest(US-MMS) | 39.3231 | -86.4131 | Deciduous broadleaf forest | 2013 |
| Fermi Agricultural (US-IB1) | 41.8593 | -88.2227 | Corn | 2008 |
| Fermi Agricultural (US-IB1) | 41.8593 | -88.2227 | Soybean | 2009 |

## 2.3 Anthropogenic CO$_2$ Emissions

For anthropogenic sectoral CO$_2$ emissions, we used the inventories developed by Arizona State University, Vulcan and Hestia, which provide CO$_2$ emissions separated by sectors: Aircraft (AIR), Cement (CEM), Commercial (COM), Industrial (IND), Nonroad (NON), Onroad (MOB), Residential (RES) and Electricity Production (UTL). The Vulcan inventory [36] provides the United States sectoral fossil fuel CO$_2$ emissions for the year 2002 at a spatial scale that is less than 10 km and hourly. This inventory was then re-sampled to the two WRF outermost domains (9 and 3 km) in order to represent the emissions outside of the city. For the inner most domain, the Hestia inventory was used. The Hestia inventory [37] provides the sectoral fossil fuel CO$_2$ emissions for the year 2013 at street and building level and hourly for the city of Indianapolis. The inventory was then re-sampled to 1 km spatial resolution to match the WRF domain.

## 2.4 Tower CO$_2$ Measurement Network

As part of the INFLUX experiment, continuous measurements are collected at 12 tower locations within and around the urban area of Indianapolis, IN, USA (Table 2 and Figure 1). INFLUX utilizes existing cell phone and communications towers, and the sampling altitudes are 39 – 136 m above ground level (agl). Seven tower locations ring the outskirts of the urban area, four more towers are closer to the city center, and one tower is further

downwind in the prevailing wind direction. Continuous measurements of $CO_2$ are made at all 12 towers, and continuous methane ($CH_4$) and CO measurements are made at a subset of the towers, all utilizing cavity ring-down spectrometers (CRDS, Picarro Inc.) [38, 39]. During the period of interest for this study, September and October 2013, towers 8 and 9 were not collecting data.

Richardson et al. [39] document the quality assessment of the instruments during the Mid-Continent Intensive (MCI) campaign; 8 months of testing against a NOAA-ESRL nondispersive infrared (NDIR) system in West Branch, Iowa, yielded median daytime-only differences of -0.13 ± 0.63 ppm

**Table 2.** INFLUX tower locations.

| TID | Name | Latitude | Longitude | magl | Instrumentation |
|---|---|---|---|---|---|
| 1 | Mooresville | 39.5805 | -86.4207 | 121 | $CO_2/CH_4/CO$ |
| 2 | E 21 St | 39.7978 | -86.0183 | 136 | $CO_2/CH_4/CO$ |
| 3 | Downtown | 39.7833 | -86.1652 | 54 | $CO_2/CH_4/CO$ |
| 4 | Greenwood | 39.5926 | -86.1099 | 60 | $CO_2$ |
| 5 | W 79 St | 39.8949 | -86.2028 | 125 | $CO_2/CO$ |
| 6 | Lambert | 39.9201 | -86.028 | 39 | $CO_2$ |
| 7 | Wayne Twp Comm | 39.7739 | -86.2724 | 58 | $CO_2$ |
| 8 | Noblesville | 40.0411 | -85.9734 | 41 | $CO_2$ |
| 9 | Greenfield | 39.8627 | -85.7448 | 130 | $CO_2/CO$ |
| 10 | Garfield Park | 39.7181 | -86.1436 | 40 | $CO_2/CH_4$ |
| 11 | Butler | 39.8403 | -86.1763 | 130 | $CO_2/CH_4$ |
| 13 | Pleasant View | 39.7173 | -85.9417 | 87 | $CO_2/CH_4$ |

## 2.5 $CO_2$ Boundary Conditions & Low Frequency Trend Correction

The initial and boundary conditions for $CO_2$ were set to a constant value (380 ppm) for all the tracers in order to compute the enhancements properly for every sector without the influence of changing boundary conditions. This fact indeed makes it difficult to compare the simulations to the observations. Therefore, we corrected the simulations low frequency trend as follow: First, we calculated the observations low frequency trend as a linear function. Then, we computed the simulations low frequency trend and removed it from the time series. At the end, we added the observations low frequency trend to the detrended simulations. We anticipate that this correction will remove most of the bias but still will allow us to compare simulations to observations variability at the towers.
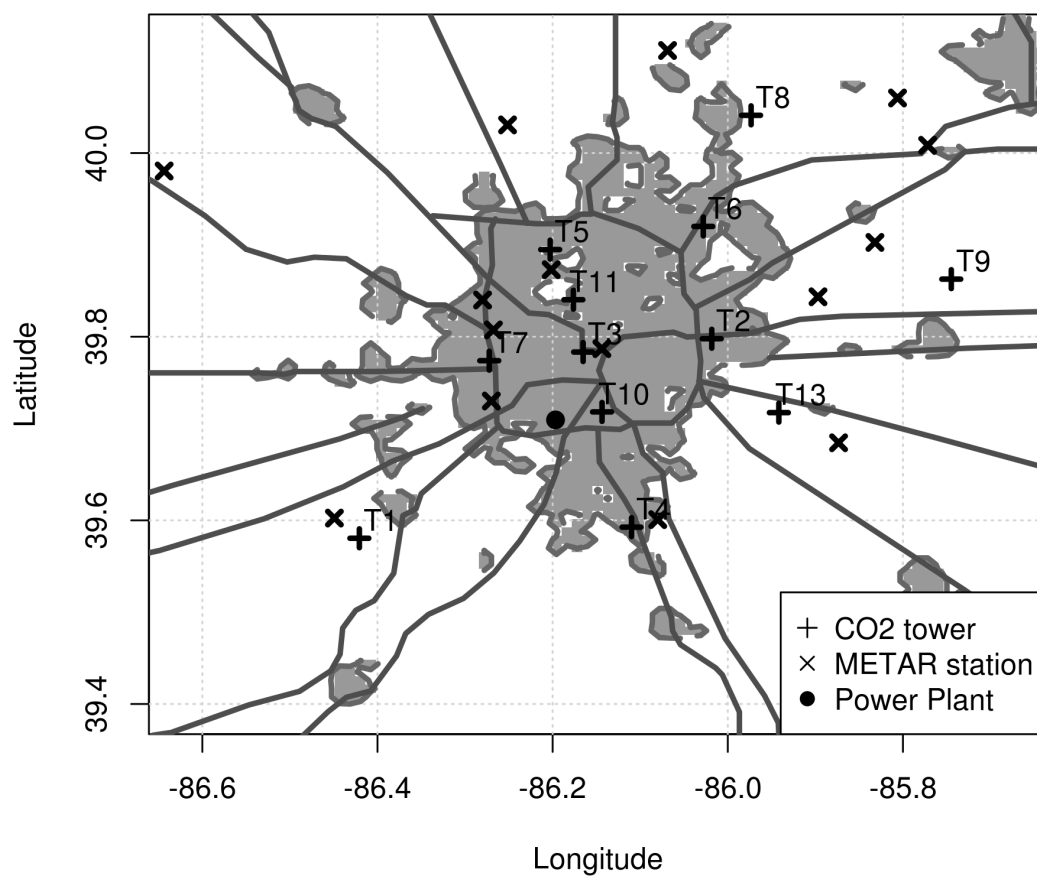
**Fig. 1.** Innermost computational domain showing the urban land use of Indianapolis, the main roads, the $CO_2$ towers, the METAR stations and the power plant.

## 3. Results

### 3.1 Temperature, Wind Speed and Direction

Simulated temperature by the control run presents a cold bias of -0.74 K and RMSD of 2.67 K. The correlation coefficient is 0.95 indicating the temperature is well capture by the model. Including meteorological forcing with the RDA dataset changes the sign of the bias and slightly increases the RMSD. The correlation coefficient is slightly reduced. However, by including the dataset MADIS as forcing benefits the simulations, substantially reducing the bias, RMSD and increasing the correlation coefficient up to 0.99 On the other hand, the updated land use configuration (MODIS) shows a degradation of the performances, with a larger cold bias and larger RMSD than the control run. The inclusion of the forcing, both RDA and MADIS, benefits this configuration, alleviating the cold bias, reducing the RMSD and improving the correlation coefficient, especially the MODIS+MADIS configuration, showing the best performances, (Table 3).

**Table 3.** Bias, RMSD (K) and correlation coefficient (R) for the predicted temperature.

| Temperature | Bias | RMSD | R |
|---|---|---|---|
| CONTROL | -0.74 | 2.67 | 0.95 |
| RDA | 0.42 | 2.82 | 0.93 |
| MADIS | 0.16 | 1.13 | 0.99 |
| MODIS | -1.50 | 3.36 | 0.94 |
| MODIS+RDA | -0.29 | 1.50 | 0.96 |
| MODIS+MADIS | 0.15 | 0.98 | 0.99 |

The control run overestimates the wind speed with a bias of 1 m s$^{-1}$, RMSD of 1.81 m s$^{-1}$ and correlation coefficient of 0.45. The configuration using RDA dataset as meteorological forcing shows lower bias and RMSD and slightly better correlation coefficient. By using MADIS dataset, the bias is reduced down to -0.18 m s$^{-1}$ and the RMSD to 1.38 m s$^{-1}$, however, the correlation coefficient is also reduced. The configuration using the updated land use (MODIS) presents lower bias and RMSD but it also shows a lower correlation coefficient than the control run. The configuration using analysis nudging with the RDA dataset improve the RMSD but worsen the correlation coefficient. By using MADIS instead, the bias and RMSD are the smallest of all configuration while the correlation coefficient is better than the MODIS configuration and similar to the control run, (Table 4).

8

**Table 4.** Bias, RMSD (m s$^{-1}$) and correlation coefficient for the predicted wind speed.

| wind speed | Bias | RMSD | R |
|---|---|---|---|
| CONTROL | 1.01 | 1.81 | 0.45 |
| RDA | 0.84 | 1.60 | 0.46 |
| MADIS | -0.18 | 1.38 | 0.35 |
| MODIS | 0.78 | 1.75 | 0.39 |
| MODIS+RDA | 0.78 | 1.59 | 0.33 |
| MODIS+MADIS | -0.18 | 1.30 | 0.43 |

The wind direction presents a bias of 4.24° for the control run with a RMSD of 52.09° which is the smallest for all the configurations. The correlation coefficient, 0.73, is also the best for all configurations. The configuration using the RDA dataset for meteorological forcing shows a larger bias and RMSD, with a decreased correlation coefficient of 0.61 By using MADIS dataset, the bias changes sign, being better than the control run but with larger RMSD and smaller correlation coefficient. By updating the land use (MODIS) the bias is reduced down to -0.54° but the RMSD is slightly increased. The correlation coefficient is similar to the control run. By using the RDA dataset, the bias increases, the RMSD slightly decreases and the correlation coefficient stays the same. The forcing with MADIS dataset slightly worsen the bias, RMSD and the correlation coefficient, (Table 5).

**Table 5.** Bias, RMSD (°) and correlation coefficient for the predicted wind direction.

| wind direction | Bias | RMSD | R |
|---|---|---|---|
| CONTROL | 4.24 | 52.09 | 0.73 |
| RDA | 8.99 | 61.72 | 0.61 |
| MADIS | -3.66 | 59.67 | 0.65 |
| MODIS | -0.54 | 53.99 | 0.71 |
| MODIS+RDA | 1.71 | 53.16 | 0.71 |
| MODIS+MADIS | -2.97 | 56.37 | 0.70 |

No big improvements were found by using updated land use or analysis nudging. However, the configuration with updated land use and analysis nudging with MADIS (MODIS+MADIS) presented the smallest bias and RMSD for temperature and wind speed with acceptable values for wind direction and, therefore, we selected this configuration for the rest of the analysis in this paper. More research optimizing the analysis nudging parameters and perhaps including observation nudging might lead to better results.

## 3.2 VPRM Parameters Optimization

Figure 2 shows the observed and predicted NEE for the three vegetation categories during the two months used for optimizing the parameters. Each vegetation category show different photosynthesis and respiration. Deciduous broadleaf forest shows the lowest amplitude in the cycle, with minimum values of -25 $\mu$mol m$^{-2}$ s$^{-1}$ and respiration peaks around
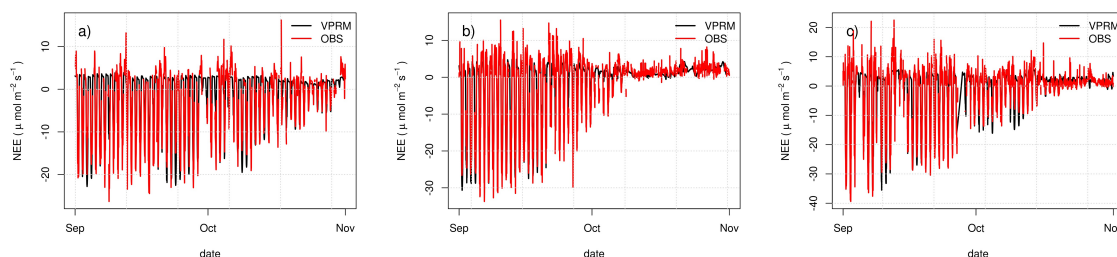
**Fig. 2.** Observed and simulated NEE evolution for the optimization period for a) Deciduous broadleaf forest, b) Soybean and c) Corn.

10 $\mu$mol m$^{-2}$ s$^{-1}$. The photosynthetic activity is almost constant during September and is reduced slowly during October. Soybean shows lower values of NEE during September, reaching more than -30 $\mu$mol m$^{-2}$ s$^{-1}$. In addition, the respiration peaks are more frequent during this month. Unlike before, the activity is reduced during September reaching the minimum after the first week of October. Corn presents the more negative values of NEE, reaching -40 $\mu$mol m$^{-2}$ s$^{-1}$. It also reaches the maximum values of respiration being well over 10 $\mu$mol m$^{-2}$ s$^{-1}$. The activity is reduced during September, reaching the minimum after the second week of October. The model captures well the daily cycle and the photosynthetic activity, showing the reduction of activity in a timely manner. However, the model does not capture the respiration peaks.

Table 6 shows the VPRM optimized parameters, which show the differences between each vegetation category. Table 7 shows the statistical evaluation of the modeled NEE during the optimization period. The optimization method achieves a nearly unbiased NEE for the 3 vegetation categories, with RMSD values between 3.29 $\mu$mol m$^{-2}$ s$^{-1}$ and 4.09 $\mu$mol m$^{-2}$ s$^{-1}$. The correlation coefficients are in the range 0.84 – 0.93, showing a very good agreement between observed and predicted NEE, especially for Soybean and Corn fields.

**Table 6.** Optimized VPRM parameters.

|  | RAD$_0$ | $\lambda$ | $\alpha$ | $\beta$ |
|---|---|---|---|---|
| Deciduous | 1266 | 0.0983 | 0.0975 | 1.044 |
| Soybean | 1667 | 0.0770 | 0.2507 | -0.40 |
| Cornfields | 2203 | 0.0916 | 0.2059 | 1.176 |

**Table 7.** Bias, RMSD ($\mu$mol m$^{-2}$ s$^{-1}$) and correlation coefficient for the predicted NEE after optimization of the VPRM parameters for the optimization period.

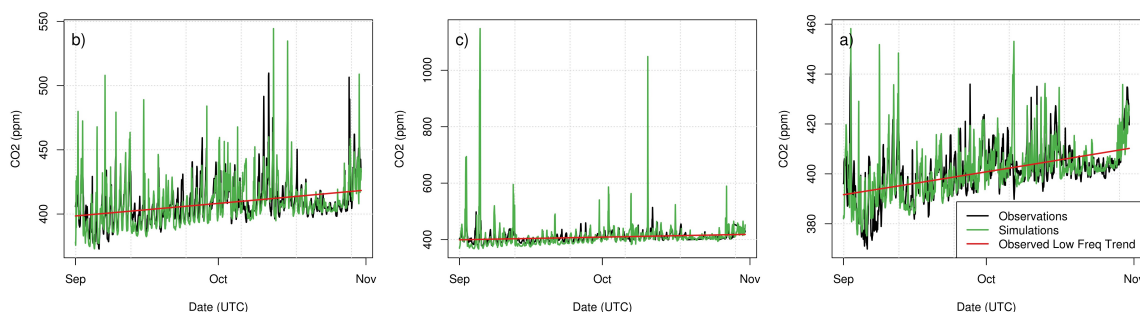|  | Bias | RMSD | R |
|---|---|---|---|
| Deciduous | -0.02 | 4.09 | 0.84 |
| Soybean | 0 | 3.29 | 0.93 |
| Cornfields | 0 | 3.7 | 0.92 |

**Fig. 3.** Simulated vs observed total $CO_2$ at towers a) 1, b) 7 and c) 10.

### 3.3 Total $CO_2$ Mixing Ratio

Figure 3 shows the hourly $CO_2$ concentration measurements at towers 1, 7 and 10 during the two months together with simulated $CO_2$. The low frequency trend correction applied to the model appears to be suitable for the purpose of comparing simulations to observations. The figure shows how the model performs in capturing day to day variability of the concentration. Also the amplitude of the signal is captured reasonably well for both daytime minimum and nighttime maximum by the model in many cases. It is worth noting the reduction in the amplitude of the signal during the end of the third week and the beginning of the fourth week of October that occurs in all the towers is well captured by the model. However, the model tends to overestimate the nighttime maximums in occasions. This is specially important for tower 10, Figure 3.c, where the spikes reach unreasonably high values. There are 9 events with unreasonable high values and 2 with extremely large values. We argue that this is due to the direct influence of the power plant plume during nighttime plus the fact that the stack height is misrepresented in the model, causing extremely large accumulations at the tower.

Table 8 shows the statistical evaluation of the simulated total $CO_2$ mixing ratio at the different tower locations. The low frequency trend correction applied removed most of the bias, ranging from 0.01 ppm at tower 1 to 1.81 ppm at tower 10. The RMSD ranged from 8.8 ppm at tower 2 to 49.4 ppm at tower 10. Towers 3, 6 and 10 presented the largest RMSD, coinciding with the largest bias as well. The correlation coefficient ranged from 0.31 to 0.67. However, only tower 10 presented such a low correlation coefficient while the rest of towers presented values larger than 0.54. Towers 3 and 10 are under the direct influence of the plume of the power plant during the prevailing wind conditions (SW) and the misrepresentation of the emissions height may cause part of the error in the model.

11

**Table 8.** Bias, RMSD (ppm) and correlation coefficient for the predicted $CO_2$ mixing ratio at the different tower locations.

| Tower ID | Bias | RMSD | R |
|----------|------|------|------|
| 1 | 0.01 | 8.9 | 0.65 |
| 2 | 0.03 | 8.8 | 0.65 |
| 3 | 0.30 | 22.6 | 0.54 |
| 4 | 0.03 | 10.6 | 0.64 |
| 5 | 0.07 | 10.0 | 0.65 |
| 6 | 0.76 | 23.1 | 0.67 |
| 7 | 0.14 | 15.0 | 0.64 |
| 10 | 1.81 | 49.4 | 0.31 |
| 11 | 0.10 | 11.3 | 0.64 |
| 13 | 0.08 | 11.7 | 0.59 |

Table 9 shows the statistical evaluation of the simulated total $CO_2$ mixing ratio at the different tower locations during the afternoon hours (17 - 21 UTC). It is clear that the errors are largely reduced during the well-mixed conditions of the afternoon hours. In this case, the RMSD ranged from 6.3 ppm at tower 13 to 11.8 ppm at tower 10. The correlation coefficients were also improved, ranging from 0.59 to 0.78. This fact indicates that a large share of the errors is due to the misrepresentation of nighttime stable conditions in the model, causing an overestimation of the nighttime peaks.

**Table 9.** Bias, RMSD (ppm) and correlation coefficient for the predicted $CO_2$ mixing ratio at the different tower locations during the afternoon hours (17 – 21 UTC).

| Tower ID | Bias | RMSD | R |
|----------|------|------|------|
| 1 | 0.07 | 8.1 | 0.69 |
| 2 | 0.05 | 6.9 | 0.75 |
| 3 | 0.09 | 9.5 | 0.70 |
| 4 | 0.04 | 7.5 | 0.68 |
| 5 | 0.09 | 7.8 | 0.71 |
| 6 | 0.10 | 9.6 | 0.70 |
| 7 | 0.07 | 9.0 | 0.71 |
| 10 | 0.01 | 11.8 | 0.59 |
| 11 | 0.12 | 9.1 | 0.74 |
| 13 | 0.04 | 6.3 | 0.78 |

Figure 4 shows the observed and simulated averaged $CO_2$ daily cycle for towers 7, 10 and 11. The daily cycle at tower 7 is well reproduced by the model, Figure 4.a. Towers 1, 3, 4 and 13 present similar behavior. Tower 10 presents a maximum at 3 EST (0800 UTC) that deviates largely from observations, Figure 4.b. This maximum is probably due to the misrepresentation of the power plant stack height, the stable conditions during the nighttime and the direct influence of the plume on this tower. Tower 6 shows a similar
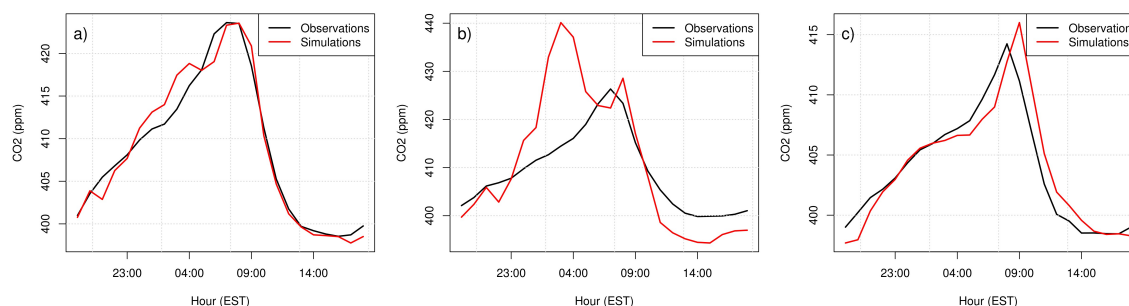
**Fig. 4.** Observed and simulated averaged $CO_2$ daily cycle for the towers a) 7, b) 10 and c) 11.

overestimation at the same time, showing a excessive nocturnal accumulation of the surrounding emissions. Towers 2, 5, 10 and 11 show a delayed collapse of the nocturnal stable boundary layer, similar to that shown in Figure 4.c for Tower 11, occurring, in average, at 9 EST, an hour later than in the observations, 8 EST (1300 UTC). The rest of the towers present a timely collapse of the nocturnal stable boundary layer.

## 3.4 Simulated $CO_2$ Sectoral Contribution

Figure 5 shows the averaged and relative simulated $CO_2$ contributions at the INFLUX towers for the different sectors. The biogenic contribution is important, being larger than the anthropogenic contribution for the towers, 1, 2, 4, 5, 8, 9 and 13 during this period. In absolute values, the largest biogenic contribution is seen at tower 8 (12.6 ppm), followed by towers 6, 13 and 4, (Figure 5, a) The minimum average biogenic contribution is seen at tower 1 (5 ppm). In relative terms, towers 8, 13, 4 and 9 present the largest contribution (Figure 5, b), being larger than 60 % while tower 6 presents the lowest, 26.7 %. On the other hand, towers 6, 10 and 3 present the largest anthropogenic contribution, 23, 18 and 15 ppm respectively, while tower 9 presents the lowest, 3.5 ppm. In general, the anthropogenic contribution is dominated by the Electricity Production, ranging from 6.5 % to 39 %, Mobility, ranging from 11 % to 56 %, and Industrial sectors, ranging from 4.8 % to 10 %. Towers 10 and 3 are clearly dominated by the electricity production sector due to the proximity to the power plant, with contributions of 10 ppm (39 %) and 6 ppm (24 %) respectively. Tower 6 is clearly dominated by traffic emissions with 17.5 ppm, that accounts for 56 % of the total contribution at this tower. In addition, towers 3, 7 and 10 present a large share of traffic contribution. Contributions from the Residential, Commercial and Non-road sectors are present in all the towers with similar values, ranging from 1.3 % to 3.6 %, from 1.1 % to 3.6 % and from 1.3 % to 2.5 % respectively. Lastly, Aircraft and Cement sectors contribute very little to the total $CO_2$.
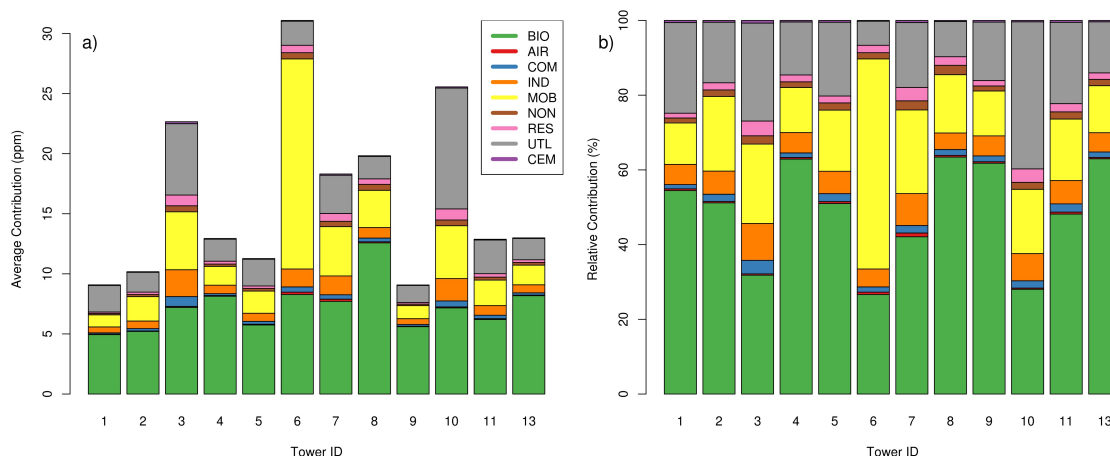
13

**Fig. 5.** a) Averaged and b) relative simulated sectoral $CO_2$ contributions for the INFLUX towers. (Biogenic (BIO), Aircraft (AIR), Cement (CEM), Commercial (COM), Industrial (IND), Nonroad (NON), Onroad (MOB), Residential (RES) and Electricity Production (UTL)).

## 4. Conclusions

We have used the high-resolution Weather Research and Forecasting Model (WRF-ARW) along with a modified version of the Greenhouse Gases chemistry module for simulating the $CO_2$ mole fraction transport during September and October 2013. Sectoral anthropogenic $CO_2$ emissions were obtained from Hestia 2013 and from Vulcan 2002 beyond the spatial coverage of Hestia. Biogenic $CO_2$ emissions were simulated by using an augmented version of the "Vegetation Photosynthesis and Respiration Model" (VPRM) included in WRF-CHEM. The results were compared with a network of stations deployed in Indianapolis as part of the Indianapolis Flux Experiment (INFLUX).

We ran 6 WRF configuration including different land use datasets and different source of meteorological data used in analysis nudging. No big improvements were found by using updated land use or analysis nudging. However, the configuration with updated land use and analysis nudging with MADIS (MODIS+MADIS) presented the smallest bias and RMSD for temperature and wind speed with acceptable values for wind direction.

Each vegetation category showed different photosynthesis and respiration with deciduous broadleaf forest having the lowest amplitude while corn had the largest amplitude. The biogenic emissions model (VPRM) captured well the NEE daily cycle and the photosynthetic activity, showing the reduction of activity in a timely manner. However, the model did not capture the respiration peaks.

The atmospheric model performed well in capturing day to day variability of the concentration. Also the amplitude of the signal was captured reasonably well for both daytime minimum and nighttime maximum by the model in many cases. However, the model tended to overestimate the nighttime maximums in occasions. This was specially important for tower 10 due to the direct influence of the power plant plume during nighttime plus the

14

fact that the stack height was misrepresented in the model. The errors were largely reduced during the well-mixed conditions of the afternoon hours indicating that a large share of the errors was due to the misrepresentation of nighttime stable conditions in the model, causing an overestimation of the nighttime peaks.

In general, the averaged daily cycle was well represented by the model with the exception of towers 10 and 6 were the presence of nearby strong sources caused an overestimation during the nighttime. In addition, 4 towers presented a slightly delayed collapse of the nocturnal stable boundary layer.

The biogenic contribution was important, being larger than the anthropogenic contribution for 7 out of 12 towers during this period. In general, the anthropogenic contribution was dominated by the Electricity Production, ranging from 6.5 % to 39 %, Mobility, ranging from 11 % to 56 %, and Industrial sectors, ranging from 4.8 % to 10 %.

## Acknowledgments

## References

[1] IPCC (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 (Intergovernmental Panel on Climate Change), 5AR.

[2] Rosenzweig C, Solecki W, Hammer SA, Mehrotra S (2010) Cities lead the way in climate–change action. *Nature* 467(7318):909–911. https://doi.org/10.1038/467909a. Available at https://doi.org/10.1038%2F467909a

[3] McKain K, Wofsy SC, Nehrkorn T, Eluszkiewicz J, Ehleringer JR, Stephens BB (2012) Assessment of ground-based atmospheric observations for verification of greenhouse gas emissions from an urban region. *Proceedings of the National Academy of Sciences* 109(22):8423–8428.

[4] Brioude J, Petron G, Frost GJ, Ahmadov R, Angevine WM, Hsie EY, Kim SW, Lee SH, McKeen SA, Trainer M, Fehsenfeld FC, Holloway JS, Peischl J, Ryerson TB, Gurney KR (2012) A new inversion method to calculate emission inventories without a prior at mesoscale: Application to the anthropogenic CO2emission from houston, texas. *Journal of Geophysical Research: Atmospheres* 117(D5):n/a–n/a. https://doi.org/10.1029/2011jd016918. Available at https://doi.org/10.1029%2F2011jd016918

[5] Bréon FM, Broquet G, Puygrenier V, Chevallier F, Xueref-Remy I, Ramonet M, Dieudonné E, Lopez M, Schmidt M, Perrussel O, Ciais P (2015) An attempt at estimating paris area $CO_2$ emissions from atmospheric concentration measure-

ments. *Atmospheric Chemistry and Physics* 15(4):1707–1724. https://doi.org/10.5194/acp-15-1707-2015. Available at https://doi.org/10.5194%2Facp-15-1707-2015

[6] Duren RM, Miller CE (2012) Measuring the carbon emissions of megacities. *Nature Climate Change* 2(8):560–562. https://doi.org/10.1038/nclimate1629. Available at https://doi.org/10.1038%2Fnclimate1629

[7] Cambaliza MOL, Shepson PB, Caulton DR, Stirm B, Samarov D, Gurney KR, Turnbull J, Davis KJ, Possolo A, Karion A, Sweeney C, Moser B, Hendricks A, Lauvaux T, Mays K, Whetstone J, Huang J, Razlivanov I, Miles NL, Richardson SJ (2014) Assessment of uncertainties of an aircraft-based mass balance approach for quantifying urban greenhouse gas emissions. *Atmospheric Chemistry and Physics* 14(17):9029–9050. https://doi.org/10.5194/acp-14-9029-2014. Available at https://doi.org/10.5194%2Facp-14-9029-2014

[8] Turnbull JC, Sweeney C, Karion A, Newberger T, Lehman SJ, Tans PP, Davis KJ, Lauvaux T, Miles NL, Richardson SJ, Cambaliza MO, Shepson PB, Gurney K, Patarasuk R, Razlivanov I (2015) Toward quantification and source sector identification of fossil fuel CO2emissions from an urban area: Results from the INFLUX experiment. *Journal of Geophysical Research: Atmospheres* 120(1):292–312. https://doi.org/10.1002/2014jd022555. Available at https://doi.org/10.1002%2F2014jd022555

[9] Grell GA, Peckham SE, Schmitz R, McKeen SA, Frost G, Skamarock WC, Eder B (2005) Fully coupled "online" chemistry within the wrf model. *Atmospheric Environment* 39(37):6957–6975.

[10] Skamarock WC, Klemp JB, Dudhia J, Gill DO, Barker D, Duda MG, yu Huang X, Wang W, Powers JG (2008) A description of the advanced research wrf version 3 (National Center for Atmospheric Research), NCAR/TN-475+STR.

[11] Beck V, Koch T, Kretschmer R, Marshall J, Ahmadov R, Gerbig C, Pillai D, Heimann M (2011) The wrf greenhouse gas model (wrf-ghg) (Max Planck Institute for Biogeochemistry, Jena, Germany), 25.

[12] Beck V, Gerbig C, Koch T, Bela MM, Longo KM, Freitas SR, Kaplan JO, Prigent C, Bergamaschi P, Heimann M (2013) WRF-chem simulations in the amazon region during wet and dry season transitions: evaluation of methane models and wetland inundation maps. *Atmospheric Chemistry and Physics* 13(16):7961–7982. https://doi.org/10.5194/acp-13-7961-2013. Available at https://doi.org/10.5194%2Facp-13-7961-2013

[13] Nakanishi M, Niino H (2004) An improved mellor–yamada level-3 model with condensation physics: Its design and verification. *Boundary-layer meteorology* 112(1):1–31.

[14] Nakanishi M, Niino H (2006) An improved mellor–yamada level-3 model: Its numerical stability and application to a regional prediction of advection fog. *Boundary-Layer Meteorology* 119(2):397–407. https://doi.org/10.1007/s10546-005-9030-8. Available at https://doi.org/10.1007%2Fs10546-005-9030-8

[15] Mellor GL, Yamada T (1982) Development of a turbulence closure model for geophysical fluid problems. *Reviews of Geophysics* 20(4):851. https://doi.org/10.1029/rg020i004p00851. Available at https://doi.org/10.1029%2Frg020i004p00851

[16] Janjić ZI (1994) The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Monthly Weather Review* 122(5):927–945. https://doi.org/10.1175/1520-0493(1994)122⟨0927:TSMECM⟩2.0.CO;2. Available at https://doi.org/10.1175/1520-0493(1994)122⟨0927:TSMECM⟩2.0.CO;2

[17] Coniglio MC, Correia J, Marsh PT, Kong F (2013) Verification of convection-allowing WRF model forecasts of the planetary boundary layer using sounding observations. *Weather and Forecasting* 28(3):842–862. https://doi.org/10.1175/waf-d-12-00103.1. Available at https://doi.org/10.1175%2Fwaf-d-12-00103.1

[18] Mlawer EJ, Taubman SJ, Brown PD, Iacono MJ, Clough SA (1997) Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *Journal of Geophysical Research: Atmospheres* 102(D14):16663–16682. https://doi.org/10.1029/97jd00237. Available at https://doi.org/10.1029%2F97jd00237

[19] Ruiz-Arias JA, Dudhia J, Santos-Alamillos FJ, Pozo-Vázquez D (2013) Surface clear-sky shortwave radiative closure intercomparisons in the weather research and forecasting model. *Journal of Geophysical Research: Atmospheres* 118(17):9901–9913. https://doi.org/10.1002/jgrd.50778. Available at https://doi.org/10.1002%2Fjgrd.50778

[20] Thompson G, Rasmussen RM, Manning K (2004) Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. part i: Description and sensitivity analysis. *Monthly Weather Review* 132(2):519–542.

[21] Thompson G, Field PR, Rasmussen RM, Hall WD (2008) Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. part ii: Implementation of a new snow parameterization. *Monthly Weather Review* 136(12):5095–5115.

[22] Kain JS (2004) The kain–fritsch convective parameterization: an update. *Journal of Applied Meteorology* 43(1):170–181.

[23] Patil M, Waghmare R, Halder S, Dharmaraj T (2011) Performance of noah land surface model over the tropical semi-arid conditions in western india. *Atmospheric Research* 99(1):85–96. https://doi.org/10.1016/j.atmosres.2010.09.006. Available at https://doi.org/10.1016%2Fj.atmosres.2010.09.006

[24] Loveland TR, Belward AS (1997) The IGBP-DIS global 1km land cover data set, DISCover: First results. *International Journal of Remote Sensing* 18(15):3289–3295. https://doi.org/10.1080/014311697217099. Available at https://doi.org/10.1080%2F014311697217099

[25] Strahler AH, Muchoney D, Borak J, Friedl M, Gopal S, Lambin E, Moody A (1999) Modis land cover product: algorithm theoretical basis document version 5.0, . Available at https://modis.gsfc.nasa.gov/data/atbd/atbd_mod12.pdf.

[26] Cressman GP (1959) An operational objective analysis system. *Monthly Weather Review* 87(10):367–374. https://doi.org/10.1175/1520-0493(1959)087⟨0367:AOOAS⟩2.0.CO;2. Available at https://doi.org/10.1175/1520-0493(1959)087⟨0367:AOOAS⟩2.0.CO;2

[27] Xiao X, Zhang Q, Braswell B, Urbanski S, Boles S, Wofsy S, Moore III B, Ojima D (2004) Modeling gross primary production of temperate deciduous broadleaf forest using satellite images and climate data. *Remote Sensing of Environment* 91(2):256–270.

[28] Mahadevan P, Wofsy SC, Matross DM, Xiao X, Dunn AL, Lin JC, Gerbig C, Munger JW, Chow VY, Gottlieb EW (2008) A satellite-based biosphere parameterization for net ecosystem CO2exchange: Vegetation photosynthesis and respiration model (VPRM). *Global Biogeochemical Cycles* 22(2):n/a–n/a. https://doi.org/10.1029/2006gb002735. Available at https://doi.org/10.1029%2F2006gb002735

[29] Hilton TW, Davis KJ, Keller K, Urban NM (2013) Improving north american terrestrial $CO_2$ flux diagnosis using spatial structure in land surface model residuals. *Biogeosciences* 10(7):4607–4625. https://doi.org/10.5194/bg-10-4607-2013. Available at https://doi.org/10.5194%2Fbg-10-4607-2013

[30] Ahmadov R, Gerbig C, Kretschmer R, Koerner S, Neininger B, Dolman AJ, Sarrat C (2007) Mesoscale covariance of transport and CO2fluxes: Evidence from observations and simulations using the WRF-VPRM coupled atmosphere-biosphere model. *Journal of Geophysical Research* 112(D22). https://doi.org/10.1029/2007jd008552. Available at https://doi.org/10.1029%2F2007jd008552

[31] Ahmadov R, Gerbig C, Kretschmer R, Körner S, Rödenbeck C, Bousquet P, Ramonet M (2009) Comparing high resolution WRF-VPRM simulations and two global $CO_2$ transport models with coastal tower measurements of $CO_2$. *Biogeosciences* 6(5):807–817. https://doi.org/10.5194/bg-6-807-2009. Available at https://doi.org/10.5194%2Fbg-6-807-2009

[32] Strahler AH, Muller J, Lucht W, Schaaf C, Tsang T, Gao F, Li X, Lewis P, Barnsley MJ (1999) Modis brdf/albedo product: algorithm theoretical basis document version 5.0, . Available at https://modis.gsfc.nasa.gov/data/atbd/atbd_mod09.pdf.

[33] Huete A, Didan K, Miura T, Rodriguez E, Gao X, Ferreira L (2002) Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment* 83(1-2):195–213. https://doi.org/10.1016/s0034-4257(02)00096-2. Available at https://doi.org/10.1016%2Fs0034-4257%2802%2900096-2

[34] Nelder JA, Mead R (1965) A simplex method for function minimization. *The computer journal* 7(4):308–313.

[35] R Core Team (2013) *R: A Language and Environment for Statistical Computing* R Foundation for Statistical Computing Vienna, Austria, . Available at http://www.R-project.org/.

[36] Gurney KR, Mendoza DL, Zhou Y, Fischer ML, Miller CC, Geethakumar S, de la Rue du Can S (2009) High resolution fossil fuel combustion CO2emission fluxes for the united states. *Environmental Science & Technology* 43(14):5535–5541. https://doi.org/10.1021/es900806c. Available at https://doi.org/10.1021%2Fes900806c

[37] Gurney KR, Razlivanov I, Song Y, Zhou Y, Benes B, Abdul-Massih M (2012) Quantification of fossil fuel CO2 emissions on the building/street scale for a large

18

u.s. city. *Environmental Science & Technology* 46(21):12194–12202. https://doi.org/10.1021/es3011282. Available at https://doi.org/10.1021%2Fes3011282

[38] Crosson E (2008) A cavity ring-down analyzer for measuring atmospheric levels of methane, carbon dioxide, and water vapor. *Applied Physics B* 92(3):403–408. https://doi.org/10.1007/s00340-008-3135-y. Available at https://doi.org/10.1007%2Fs00340-008-3135-y

[39] Richardson SJ, Miles NL, Davis KJ, Crosson ER, Rella CW, Andrews AE (2012) Field testing of cavity ring-down spectroscopy analyzers measuring carbon dioxide and water vapor. *Journal of Atmospheric and Oceanic Technology* 29(3):397–406. https://doi.org/10.1175/jtech-d-11-00063.1. Available at https://doi.org/10.1175%2Fjtech-d-11-00063.1