Analysis of atmospheric CO2 measurements in Mexico City

Yang Xu

université PARIS-SACLAY

Analysis of atmospheric CO_2 measurements in Mexico City

Interprétation des mesures atmosphériques de CO_2 à Mexico City

Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 129, sciences de l'environnement d'Île-de-France (SEIF)
Spécialité de doctorat : Géosciences
Graduate School : Géosciences, climat, environnement et planètes. Référent : Université de Versailles Saint-Quentin-en-Yvelines

Thèse préparée au **LSCE** (Université Paris Saclay, CNRS, CEA, UVSQ),
sous la direction de **Michel RAMONET** , chargé de recherche,
la co-direction de **François-Marie BREON** , directeur de recherche et
le co-encadrement de **Thomas LAUVAUX** , professeur

Thèse soutenue à Gif-sur-Yvette, le 13 avril 2023, par

Yang XU

Composition du Jury

Membres du jury avec voix délibérative

Philippe BOUSQUET Professeur, LSCE	Président
Gara Villalba MENDEZ Professeure, ITCA-UAB	Rapportrice & Examinatrice
Yao TÉ, Maître de conférences, LERMA	Rapporteur & Examinateur
Leena JÄRVI, Professeure, INAR, University of Helsinki	Examinatrice
Valérie THOURET Professeure, LAERO	Examinatrice

THESE DE DOCTORAT
NNT : 2023UPASJ013

université PARIS-SACLAY
ÉCOLE DOCTORALE
Sciences de l'Environnement d'Île-de-France (SEIF)

Titre : Interprétation des mesures atmosphériques de CO_2 à Mexico City

Mots clés : CO_2, Mexico, inversion atmosphérique

Résumé :

Les villes sont responsables de plus de 70% des émissions mondiales de CO_2. De nombreuses municipalités se sont engagées à réduire les émissions de CO_2 urbaines. Afin d'évaluer l'impact des Plans Climat, des inventaires d'émissions de gaz à effet de serre sont établis à l'échelle de la ville. La modélisation par inversion atmosphérique offre une solution complémentaire capable de réduire les incertitudes pour ces inventaires d'émissions. Elle combine la modélisation du transport atmosphérique et les mesures de concentrations en gaz à effet de serre pour affiner les estimations des émissions issues des inventaires.

La zone métropolitaine de Mexico (MCMA) est l'une des plus grandes mégalopoles du monde. Les émissions annuelles de CO_2 dans la MCMA sont passées de 42,1 millions de tonnes à 66,0 millions de tonnes entre 2012 et 2018. Le gouvernement mexicain a prévu de réduire les émissions de CO_2 de 65,2 millions de tonnes au cours de la période 2021-2030. Afin d'évaluer quantitativement les stratégies de réduction des émissions de CO_2, un projet franco-mexicain intitulé "Impacts du carbone dans la région de Mexico" (MERCI-CO_2) a déployé un réseau d'instruments d'observation du CO_2 in situ et en colonne dans la région de Mexico. Les gradients de concentration de CO_2 sont assimilés dans un système d'inversion basé sur le modèle de transport atmosphérique WRF-Chem pour améliorer l'estimation préalable des émissions de CO_2 à l'intérieur et à l'extérieur de la MCMA.

J'ai évalué les performances du modèle atmosphérique sous diverses configurations. Afin de quantifier les erreurs les plus impactantes pour la simulation des concentrations en CO_2 atmosphérique, le travail de thèse s'est focalisé sur les erreurs de simulation des températures de l'air, des vitesses et directions du vent et les hauteurs de mélange, en utilisant des données collectées aux stations météorologiques de la région mais également issues d'un instrument LiDar ainsi que de radiosondes. Ces tests de sensibilité ont permis de définir la configuration optimale du système de modélisation.

Des cartes de concentration de CO_2 au-dessus de la MCMA sur trois périodes caractéristiques des conditions météorologiques de la région (janvier, mai et juillet) ont été simulée par le modèle WRF-Chem à la résolution de 5 km, en utilisant deux inventaires d'émissions: inventaire local préparé par les chercheurs de l'UNAM et un inventaire d'émissions globales, ODIAC, ajusté par des facteurs d'échelle temporels. Parallèlement à l'évaluation sur la base de mesures in-situ de CO_2 en surface et de mesures de colonne (XCO_2) par FTIR, j'ai également analysé les distributions temporelle

et spatiale des signaux de CO_2, ainsi que la zone impactée par les flux anthropiques et la variation des flux biogéniques. Sur la base de cette analyse, j'ai évalué le potentiel de quantification du réseau actuel, y compris l'emplacement de la station de fond, tout en étudiant les emplacements potentiels de nouvelles stations.

Une inversion sur une année complète a été réalisée du 30 mars 2018 au 30 mars 2019. En se basant sur l'assimilation des gradients entre la station urbaine UNA et la station d'altitude ALZ, l'inversion a ajusté les émissions anthropiques issues des deux inventaires ainsi que les flux biogéniques du modèle CASA et les concentrations de fond du modèle CarbonTracker 2019B. Un ensemble de plusieurs inversions a été réalisé pour mieux quantifier les incertitudes en faisant varier les covariances d'erreur temporelles, en variant la fenêtre d'assimilation, en séparant les secteurs d'activités (trafic routier), en appliquant un masque sur la MCMA, et en filtrant les données de gradients en CO_2. Cet ensemble permet d'améliorer les performances de l'inversion et de spécifier l'impact du signal urbain et des différentes composantes du système. Une dernière expérience m'a permis d'étudier l'impact de l'assimilation des concentrations en CO, dont les observations débutent en décembre 2018.

Title : Analysis of atmospheric CO_2 measurements in Mexico City

Keywords : CO_2, Mexico City, atmospheric inversion

Abstract :

Cities are responsible for more than 70% of the global CO_2 emissions and thus play an important role in mitigating climate change. Mayors and local governments have been taking measures to reduce urban CO_2 emissions and to reach carbon neutrality. In order to evaluate their efforts, a series of high-resolution city-scale emission inventories were established. Top-down inversion modeling is a widely-used complementary solution to reduce the uncertainties in traditional bottom-up emission inventories. It combines atmospheric modeling and measurements to optimize the greenhouse gas estimates using Bayesian inference methods.

The Mexico City Metropolitan Area (MCMA) is one of the largest megacities in the world. Its annual CO_2 emissions have grown from 42.1 Mt to 66.0 Mt from 2012 to 2018. The Mexico government has planned to reduce 65.2 Mt CO_2 emission during the period 2021-2030. To assess local CO_2 emission reduction strategies, a French-Mexican project Mexico City regional Carbon impacts (MERCI-CO_2) deployed a network of in-situ and column CO_2 observation instruments in MCMA. The CO_2 concentration gradients are assimilated in our inversion system based on the WRF-Chem model to improve the inventory estimates of CO_2 emissions in and outside MCMA.

Various options on meteorological drivers, domain sizes, physics, dynamics schemes and spectral nudging of the WRF modeling system over MCMA were quantitatively evaluated for model performance. A series of meteorological parameters were taken into account for the comparison between simulations and in-situ observations, LiDar analysis as well as WMO radiosonde observations. For the purpose of CO_2 simulation, the most studied variables are those related to the dispersion of the ambient air, including air temperatures, wind speeds, wind directions and mixing heights. These sensitivity tests helped to define the optimal model configuration.

The CO_2 concentration maps over MCMA during 3 typical months (January, May and July) were reconstructed by the double-nesting 5-km resolution WRF-Chem model, coupled with the local emission inventories from UNAM and the global emission inventories ODIAC scaled by temporal scaling factors. The evaluation of CO_2 simulations were based on CO_2 in-situ measurements by PICARRO and column measurements (XCO_2) by FTIR at an urban site UNA and at the background station ALZ. Along with the evaluation, we also analyzed the temporal and spatial distribution of CO_2 signals, as well as the area impacted by anthropogenic fluxes and by biogenic fluxes. Based on our analysis, we assessed the potential of our network to constrain the urban emissions, defined the potential locations for future stations, and defined a "background index" to

represent the suitability to build a background station.

After the ground validation of the modeling system, we performed a 1-year inversion over the MCMA from 30 March 2018 to 30 March 2019. According to the assimilation of concentration gradients between the urban station and the rural station, the inversion adjusted the prior anthropogenic emission from UNAM and ODIAC estimates, in parallel with prior biogenic fluxes from the CASA model and background concentrations by CarbonTracker 2019B global inversion system. An ensemble of inversion configurations was constructed. The reference configuration optimizes three components: fossil fuel sources, biogenic fluxes and background concentrations to generate separate scaling factors for each block of 5 days. The sensitivity tests include several temporal error correlation length scales between continuous days, varying time windows over each day, a separation of the activity sectors (traffic), a filter over the MCMA, varying data screening and block sizes, to evaluate the performances of the inversion, and to specify the impact of our various configurations. The same system was also used to assimilate carbon monoxide concentrations, collected at the two stations since December 2018.

ACKNOWLEDGEMENT / REMERCIEMENT

First of all, many thanks to you, who is reading it. Thanks for your patience and tolerance.

I would like to express my deepest gratitude to my thesis advisors and supervisors, Dr. Michel Ramonet, Dr. François-Marie Bréon and Prof. Thomas Lauvaux, for their invaluable guidance, support, and encouragement throughout my doctoral studies. Their expertise and insights have been instrumental in shaping my research and developing my academic skills. Michel has been consistently supportive and Thomas makes the studies in Reims very effective and PhD-style.

The review of this dissertation was very convoluted, leaving very limited time for the reporters. I must thank the two reporters, Prof. Gara Villalba Mendez and Prof. Yao Té, for their precious time. I would like to extend my heartfelt thanks to them for their rigorous and thoughtful evaluation of my work. Their comments and suggestions have been invaluable in shaping the direction of my research and improving the overall quality of my dissertation. I would like to express my special thanks to all my jury members, to Prof. Leena Järvi, Prof. Phillipe Bousquet and Prof. Valérie Thouret, for their acceptance to be part of my PhD jury.

I am also indebted to the members of my thesis committee, Prof. Michel Grutter, Dr. Felix Vogel and Dr. Philippe Ciais, for their valuable feedback and suggestions on my research. Their constructive criticism and encouragement have been crucial in improving the quality of my work. When we visited Mexico City, Professor Grutter and his group's warm reception impressed me.

I would like to thank my collegues and friends. Julia gave me an enthusiastic welcome to ICOS, Jinghui is always answering my newbie questions and Xin gave me suggestions on my acdemic life. Yitong and Yunsong have been with me for most of these four years. Their sharing has saved me from many detours and their accompany helped me to survive the tough times during Covid-19. Thanks to the paragliding group, swmming group, traveling group and museum group, and the wonderful life we spent together. Thanks to my office mates, Rodrigo, Sara, Emeric, Ivonne and Josselin, for sharing our cultures and helping me navigate French-speaking situations.

My family deserves special thanks. Their love is my greatest source of courage. Additional thanks to my "idol" Ling Han and my buddies from the star-chasing journey. Han is my female role model, providing an example of her trust to the world and her power in career. With Han, Zhenzhen and Chuyi, I faced the death of our mutual friend, an important life lesson. Chuyi, Ivy and Siteng reminds me how fantastic it is to have friends with telepathy.

I'm also grateful to Paris. I never thought I would learn French, attend so many cultural activities, communicate with artists and designers, and see new landscapes I had only seen on textbooks.

Lastly, thanks to my boyfriend or girlfriend who never showed up and whose absence kept me focused on my studies☺

CONTENT

FIGURES

TABLES

1 CHAPTER 1 INTRODUCTION

1.1 CLIMATE CHANGE AND URBAN CO_2 CONTROL

With the increasing frequency of climate disasters, climate change becomes one of the greatest challenges to humanity. Since the earliest established temperature records at the end of the 19th century, the global surface temperature has increased by 1.09 (0.95 - 1.20) °C (Arias et al., 2021), and it is still increasing by 0.2°C per decade (Hansen et al., 2006). Global warming leads to a series of environmental calamities, including sea level rise, abnormal extreme precipitation, droughts, heat waves and tropical cyclones (Arias et al., 2021). Greenhouse gases (GHG) are the main driver of global warming and climate change. Carbon dioxide (CO_2), known as the most important greenhouse gas, is responsible for 66% of the radiative forcing by long-lived greenhouse gases and 82% of the increase in radiative forcing in the past decade. The atmospheric concentration of CO_2 has been rising sharply since the Industrial Revolution. By 2021, the globally averaged atmospheric CO_2 mole fraction has reached 414.7 ppm (NOAA, 2022), 149% above the pre-industrial level (Wigley, 1983). Although the quarantine during Covid-19 in 2020 reduced CO_2 emissions by 5 to 10% (Liu et al., 2020; Le Quéré et al., 2020; Friedlingstein et al., 2021), the current annual growth rate (2.5 ppm / year) was still higher than the average of the last 10 years (WMO, 2022).

The main anthropogenic sources of CO_2 are fossil fuel combustion (e.g., transportation, power plant and residential use), industrial processes (e.g. cement production and oil refining) and biomass burning (Metz, 2005). More than 70% of anthropogenic CO_2 emissions are from urban areas (Seto et al., 2014) and megacities contribute almost 50% of them (Wei et al., 2021). Meanwhile, the population in cities keeps growing. The urban population worldwide has increased from 1.35 billion to 4.38 billion in the last 50 years and will increase by 1 - 2% annually until 2050 (UN, 2018). Consequently, cities will play a pivotal role in the reduction of GHG emissions and bear the most pressure of the climate change related disasters.

The international community has taken measures to address carbon

emission reduction. 196 countries signed the Paris Agreement at the 21st Conference of the Parties (COP 21) in 2015, committing to limit global warming to less than 2 ℃, preferably to 1.5 ℃ (Paris Agreement, 2015). In order to mitigate CO_2 emissions and to achieve the goal of carbon neutrality by 2050, a series of protocols, agreements and projects between countries, sub-national governments and international organizations are concluded. The World Bank provides more than 200 governments, communities and individuals with financial support on reducing GHG emissions through their Emission Reductions Payment Agreements (ERPA) (The World Bank, 2021). C40 Cities, a network of mayors, promotes a collaboration between 97 large cities for approaching their commitment to the Paris Agreement (https://www.c40.org/about-c40/). The Global Covenant of Mayors have attracted more than 12,000 cities from 142 countries to commit to reduce their GHG emissions by 1.9 Gt CO_2 equal annually in 2030 (https://www.globalcovenantofmayors.org/who-we-are/). The URBACT Zero Carbon Cities project (https://energy-cities.eu/project/zero-carbon-cities/) established a network of 7 European cities of carbon reduction targets, policies and plans for the EU's strategic vision for carbon neutrality by 2050. To evaluate all their efforts and actions, accurate measurements of carbon emissions are necessary.

1.2 EMISSIONS INVENTORIES

The greenhouse gas emissions over a certain time and region are commonly estimated using a greenhouse gas emission inventory protocol. Few emission inventories are directly measured from sources - they are usually measured and calculated under specific guidelines. Most of the current national CO_2 emission inventories, including the official emission inventories reported to the United Nations Framework Convention on Climate Change (UNFCCC), Global Carbon Budget (Friedlingstein et al., 2022), Open-source Data Inventory for Anthropogenic CO_2 (ODIAC, Oda et al., 2011) and Emissions Database for Global Atmospheric Research (EDGAR, Olivier et al., 2015), are compiled following the Intergovernmental Panel on Climate Change (IPCC) Guidelines for National Greenhouse Gas Inventories (Eggleston et al., 2006). In addition, the World Resources Institute's (WRI) GHG

Protocol standards (GHG protocol, 2015; Fong et al., 2015) are used for emission inventory development by organizations such as the United States Environmental Protection Agency (U.S. EPA) and by projects such as C40 Cities.

Most emission inventories mostly come from self-reported databases for power production, transportation, and manufacturing (Dasgupta et al., 2022). This emission estimation model describes the relationship between the CO_2 emissions from a certain activity and the activity strength. The assumed relationship is usually linear with a fixed emission factor. The uncertainties of such a model come from various steps. Many emission sources are variable in time and space. For example, the transportation emissions are estimated with the traffic flow. Not every city and highway has its own traffic count dataset. Quite often, the traffic flow (volume and speed) is estimated on average, or thanks to another traffic-related parameter, like a congestion index. This will cause large uncertainties in both temporal and spatial dimensions, and in the estimated total amounts. Emission factors contribute to the emissions uncertainties. The real emissions are not linearly related to activity data, that is, emission factors are also variable. Engine type, engine capacity, speed and driving habits all affect the amount of CO_2 emissions per kilometer per vehicle. Advances in technology are changing emission factors as well. It remains nearly impossible to adjust these emission factors based on local situations. Additionally, this estimation model requires a knowledge of local emission patterns. If there is some unknown or ignored source, the emission will definitely be under-estimated (Rypdal et al., 2001).

Besides, the development of bottom-up emission inventories is time-consuming and costly, especially for those at high spatial and temporal resolution. When it comes to city scales, people focus primarily on the major cities. Most of the urban-scale emission inventories are in developed countries (Dasgupta et al., 2022). The Hestia program is dedicated to providing the high-resolution hourly emission inventories for 4 cities in the U.S.: Baltimore, Salt Lake City, Indianapolis and Los Angeles. It calculated the residual and commercial non-point sources, industrial point sources and transportation sources in the 4 cities at the scale of buildings, streets and factories in 2011 (Gurney et al., 2012).

The Anthropogenic Carbon Emissions System (ACES) is another system focusing on urban scale CO_2 emissions in the U.S. It provides the 1×1 km CO_2 emission inventories from fossil fuel combustion for 13 states across the Northeastern United States (covering New York), annually in 2011 and hourly in 2013/2014 (Gately et al., 2018). Cong et al. developed an annual direct fossil-fuel CO_2 emission estimate in the Tokyo metropolitan area for the year 2014 (Cong et al., 2021). Some of the megacity emission inventories keep being updated regularly. London Atmospheric Emissions Inventory (LAEI) is a yearly 1×1 km emission dataset including CO_2 of domestic, industrial/commercial, transport and other miscellaneous sources in the Greater London area (LAEI, 2021). It is updated about every three years. The AirParif emission inventory provides the 1×1 km CO_2 emission map of direct emissions in Ile-de-France (IdF, or Greater Paris) as well as those indirect emissions related to consumption in the Ile-de-France regions of electricity and urban heating in five typical months every two years (AirParif, 2022). These city-scale emission inventories are more localized and accurate than widely-used global emission inventories. There could be a difference as high as 20% in regional emission total amount, and the spatial patterns are also so different that in some of the sectors and at some urban core, the global emission inventories are 75% higher or lower than the local emission inventories (Gately et al., 2017). However, only about 80 cities have their own emission inventories (Dasgupta et al., 2022), they are not all regularly updated with the latest local spatial and temporal emission patterns, and the uncertainties could be 4-10%, up to 20% (Gately et al., 2017). For those cities which do not have a local emission inventory, the high-resolution global CO_2 emission inventory, ODIAC, with the energy consumption data compiled by the energy company BP p.l.c. and the spatial distribution of night-time light spatial data as well as individual power plant profiles, provides a widely-used solution.

1.3 BAYESIAN INVERSION MODELING

Top-down methodologies, or atmospheric inversions, offer a complementary solution to reduce the uncertainty of traditional bottom-up emission inventories. It takes advantage of a dense and local atmospheric CO_2 concentration observation network to improve the prior emissions estimate from an inventory. This inverse approach

can compensate for known gaps and errors in traditional emission inventories - large uncertainties in emission factors, lack of data, and errors in the spatial and temporal distributions. Bayesian inversions take advantage of accurate measurements to track the long-term trends and the major discrepancies in urban CO_2 emission estimation, supporting the implementation of mitigation policy at the city scale.

Inversion modeling is based on the Bayesian theory, minimizing the mismatches between modeled and observed concentrations according to their respective uncertainties. More specifically, we utilize an atmospheric transport model coupled to anthropogenic emissions, biogenic fluxes and boundary conditions, to simulate the local gradients in CO_2 concentrations. In our approach, we compare the simulated gradients to the observed gradients over short time-periods (a few days), while assigning uncertainties to the observation and the prior emissions. By applying the Bayesian inference algorithm described hereafter (**Figure 1.1**), we finally calculate the posterior emissions and their associated uncertainties.

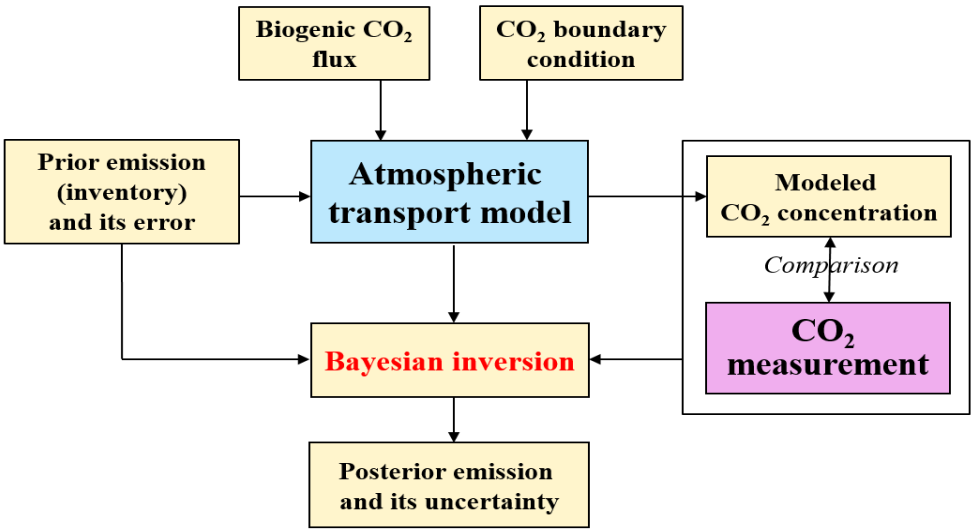


Figure 1.1 *Flowchart of the top-down (inversion) approach used to quantify the emissions from Mexico City, based on the Bayesian inversion system developed over Paris (Lian, 2020)*

The Bayesian inference estimates the conditional probability of a variable x (emissions) knowing the concentrations (y) by the probability of the concentrations (y) knowing the emissions (x), the uncertainty in ***x*** (emissions), and the uncertainty in ***y*** (concentrations):

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)}$$

Here we assumed a hidden Markov model:

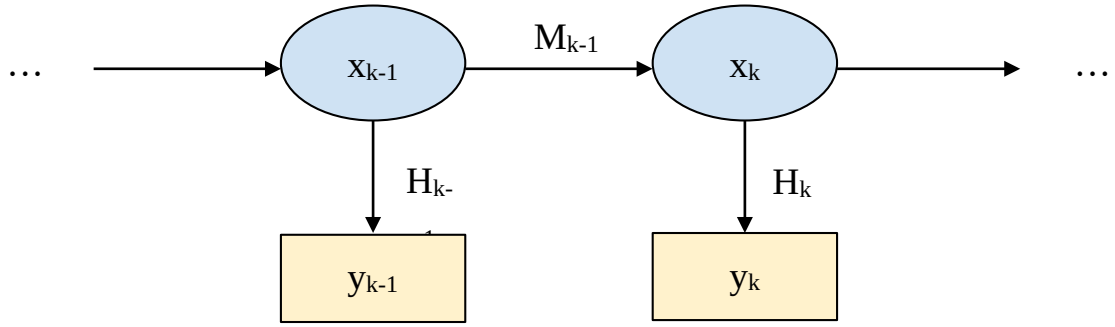


Figure 1.2 *Structure of the hidden Markov Model*

In this model, our target variable ***x*** (estimation) is not directly measurable. The measurable variable ***y*** (observation) is linked to ***x*** by an operator **H**. Our aim is to make a Bayesian inference of emission ***x***, based on the observation ***y*** and the prior estimate of ***x*** with its uncertainties. The forecast function and the observation function are:

$$x_k = M_{k-1}(x_{k-1}, \lambda) + b_k$$
$$y_k = H_k(x_k) + r_k$$

where k is the time step; **M** is the forecast operator and $\boldsymbol{\lambda}$ is its parameter independent of ***x***; **H** is the linear observation operator that generates the estimation of ***y*** (observation) by ***x*** (estimation); **b** and **r** are the errors of emissions and observations, which are assumed to be .

As for the optimal steps, we computed the Green's response functions. To avoid the complicated computation of operator **H**, a scaling factor was involved. This estimation vector ***x*** is used as the scaling factor to optimize all emission variables, including anthropogenic emissions in various sectors, areas, time steps or time windows, as well as biogenic fluxes and boundary conditions. We take $\boldsymbol{x}_A$ as the posterior estimation of the emission inventory scaling factor, and $\boldsymbol{x}_B$ as the prior (Bréon et al., 2015). The optimal solution is given by the following equation

(Enting, 2002; Tarantola, 2005):

$$\boldsymbol{x_A = x_B + (B^{-1} + H^T R^{-1} H)^{-1} H^T R^{-1} (y - H x_B)}$$

where **B** is the error covariance matrix of the prior emission variable vector, and **R** is the error matrix of the observation vector.

The posterior error matrix, defined as **A**, is

$$A = (B^{-1} + H^T R^{-1} H)^{-1}$$

With a reliable atmospheric transportation model and reasonable observations with errors, the inversion system was set up.

This inversion model has been widely used in urban CO_2 emission estimation during recent years. **Table 1.1** shows some recent studies on major cities across the world with various models, seasons, measurements and emission inventories. Lagrangian Backward Trajectory models are used most frequently as the atmospheric transport model in the previous studies using backward footprint simulations. The Weather Research and Forecasting model (WRF) coupled with various modules is also mentioned many times. Sometimes WRF and Lagrangian models are combined together. The spatial resolution of the models is usually 1-2 km. Considering the impact of growing vegetation, some studies restrict their target period in winter. Some of them use local emission inventories; some have to use global emission inventories, EDGAR, ODIAC, FFDAS (Fossil Fuel Data Assimilation System; Rayner et al.,, 2010) or a mix of them. Because of the ease of operation and the relatively lower cost, surface measurements are the most common, while there are some measurements from towers, aircrafts and satellites. With the development of computing science, there existed a new inversion approach named inversion ensemble, which combines a large number of emission inventories from different sources and bases, various measurements, several transportation models and configurations to make sensitivity tests (Lopez-Coto, 2020).

Reference	City	Emission Inventory	Transportation model	Target period	Measurement
Pitt, 2022	New York	Vulcan, ACES, EDGAR	HYSPLIT	2018-2020 non-growing season	aircraft
Kaminski, 2022	Berlin	self-prepared local	WRF-CMAQ	2008.2.3, 2008.7.3	satellite
Nalini, 2022	Paris	Origins Earth	WRF-LPDM	2019.3-2019.5, 2020.3-2020.5	tower
Lopez-Coto, 2020	Washington DC - Baltimore	Vulcan, ODIAC, FFDAS, ACES, DARTE	HYSPLIT, HRRR, WRF	2016.2	aircraft
Pisso, 2019	Tokyo	ODIAC, EDGAR	FLEXPART, flexpart-wrf, TRACZILLA	2005-2009 winter	surface, tower, aircraft
Kunik, 2019	Salt Lake City	Hestia, ODIAC	WRF-STILT	2015.9	surface
Nickless, 2018	Cape Town	self-prepared local	CCAM	2012.3-2013.6	surface
Boon, 2016	London	UK National	CHIMERE	2012 summer	surface
Feng, 2016	Los Angeles	Hestia, Vulcan	WRF-Chem	2010.5-2010.6	surface, tower
Staufer, 2016	Paris	AirParif, EDGAR	CHIMERE	2010.8-2011.7	surface
Lauvaux, 2016	Indianapolis	Hestia, ODIAC	WRF-LPDM	2012–2013 dormant season	tower

Table 1.1 *Summary of several typical studies of urban CO_2 inverse modeling.*

The studies in **Table 1.1** have found significant adjustments by

inversion modeling on city carbon budgets, either for the emission inventories or for the uncertainties. Nalini et al. (2022) found that the CO_2 emissions at Paris in 2019 and 2020 were 10% - 38% under-estimated in different time periods and sectors, and the impact of the confinement during Covid-19 over the Paris metropolitan area is 12 percentage points under-estimated. In Salt Lake City (Kunik, 2019), Cape Town (Nickless, 2018) and Indianapolis (Lauvaux, 2016), the inversion system reduced about 39%, 30% and 28% of the uncertainties in prior emissions, separately.

1.4 MEXICO CITY AND MERCI-CO_2 PROJECT

Mexico City (Ciudad de México, CDMX), the capital city of the United States of Mexico, is one of the largest cities in the world. According to the definition of Mexico and state government, the Mexico City Metropolitan Area (MCMA, Greater Mexico City, or Metropolitan Area of the Valley of Mexico, or in Spanish *Zone métropolitaine de la vallée de Mexico*, ZMVM), including all 16 boroughs of Mexico City, 59 municipalities in Mexico State and 1 municipality in Hidalgo State (Tizayuca), has a population of 21.6 million, ranking fifth in the world (UN, 2018). Almost half of the population are concentrated in CDMX.

The annual CO_2 emissions in MCMA have grown from 42.1 million tons to 66.0 million tons in 6 years, from 2012 to 2018. Mobile sources are the main source of CO_2 in MCMA contributing about 57% - 65% to total CO_2 emissions (SEDEMA, 2012; SEDEMA, 2021). By the end of 2020, there are 14.7 million motor vehicles registered for circulation (including automobiles, passenger trucks and vans, cargo trucks and motorcycles, as well as the type of service official, public and private) on this MCMA land of 7585 square kilometers, 102% more than 10 years ago, accounting for 29% of the national motor vehicle population (INEGI, 2021).

Despite the rapid urban development, there is another challenge for the environmental problems in MCMA - the topography. The MCMA valley is surrounded by several mountain systems, as shown in **Figure 1.3**. The highest mountain in elevation is La Sierra Nevada in the east, more than 5000 m a.s.l. The next highest is the south, about 3600-3930 m a.s.l. Mountains in the north are the lowest - about 2500-3000 m a.s.l., while the average altitude of Mexico City is about 2240 m (CDMX

Government et al., 2021). The surrounding mountainous topography is a natural physical barrier to wind circulation, which limits the dispersion of pollutants emitted within the MCMA. Due to the warm climate and the dense urban area, the convective boundary layers are deep most of the year. In general, the maximum heights of the mixing layer are observed between 2:00 p.m. and 4:00 p.m. in the months of April and May, while the minimum values are recorded at night and during the winter season (CDMX Government et al., 2021). There is a particular phenomena named thermal inversion, which prevents the natural convective movement of the atmosphere. It is caused by a heated air layer above a cold one, most observed in the winter months.

The climate in the MCMA is predominantly dry (sub-humid) subtropical with summer rains. Its average annual temperature ranges between 12 and 18 ºC, with year-round warmth. The year in Mexico City can be divided into a rainy season (May to October) and a dry season (November to April) (GMAO, 2015; CDMX Government et al., 2021). The most frequent rains are in July and August. There are often rainy episodes observed in the afternoon from June to August. In the northern part of MCMA it's slightly dryer and in the mountainous areas it's much colder.

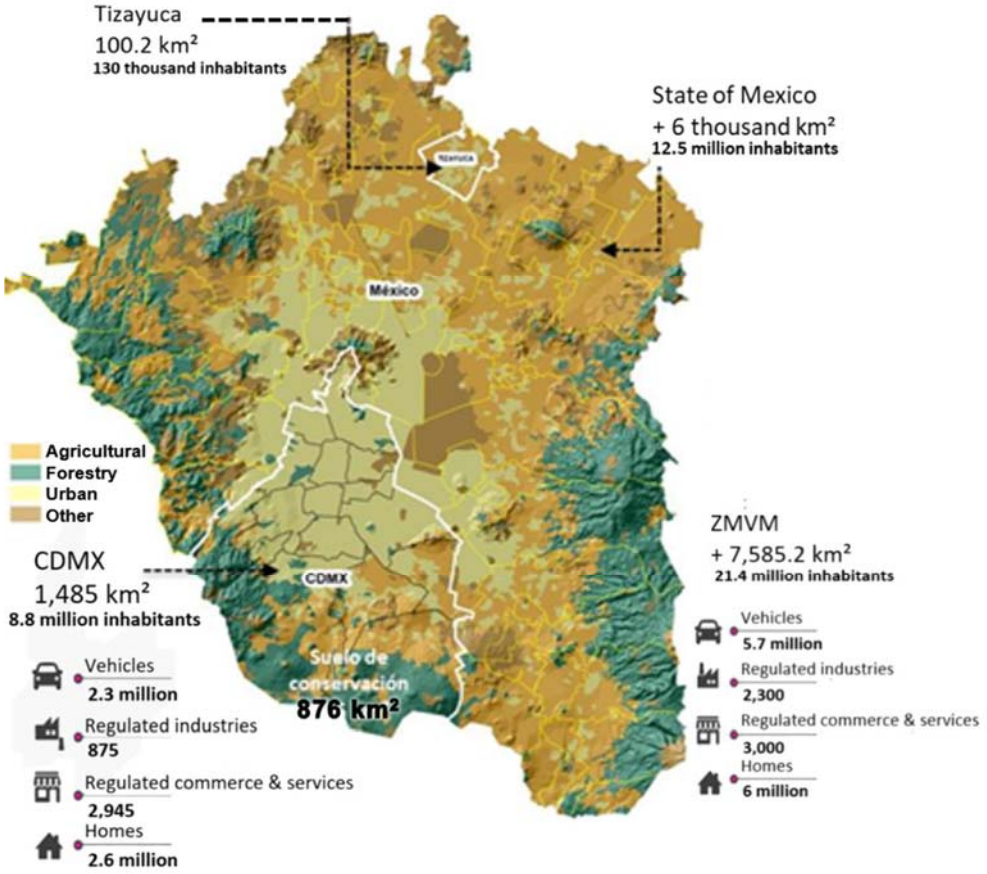


Figure 1.3 *Land use, topography, administrative division and social-economical statistics of MCMA in 2016*

The three local governments (CDMX, Mexico State and Hidalgo State) have engaged into aggressive carbon emission reduction policies. The standard NOM-163-SEMARNAT-ENER-SCFI-2013 was updated to restrict the emission of CO_2 by small vehicles. For large public and service vehicles, low-carbon technologies are introduced, including electric, hybrid or emission control systems (EURO VI or EPA 10 standard) (CDMX Government et al., 2021). However, strategies for the department of industry and power production are mainly focusing on contaminants instead of GHGs. The plan for CO_2 reduction in the period 2021-2030 is to reduce by 65.2 million tons, of which 44.0 million tons are from individual private transportation and 18.7 million tons are from public transportation of passengers (SEDEMA, 2021). There is a detailed time schedule for the strategies, and the barriers for the strategies are analyzed (CDMX Government et al., 2021), but the

effects and implementation of the strategies on CO_2 emission reduction have not been quantitatively evaluated.

Thanks to the French-Mexican project MExico City Regional Carbon Impacts (MERCI-CO_2), a dense observation network of atmospheric CO_2 concentrations was developed in MCMA (shown in **Figure 1.4**). Three different types of measurements were employed: high precise surface measurement using cavity ring-down spectrometers G2401 commercialized by PICARRO, low-cost medium-precise sensor or LCMP based on non-dispersive infrared sensors commercialized by Senseair (Arzoumanian et al., 2019), and column measurements using solar absorption spectrometers commercialized by Bruker (Frey et al., 2019). PICARRO, which is a cavity ring-down spectrometer (CRDS) (Crosson et al., 2008; Rella et al., 2013), is widely used in background and urban GHG monitoring (Yver Kwok et al., 2015; McKain et al., 2015; Heltfer et al., 2016; Wei et al., 2020), providing continuous measurements of CO, CO_2 and CH4. In the previous studies mentioned in **Table 1.1**, the city inversion project of New York (Pitt et al., 2022), Paris (Staufer et al., 2016; Nalini et al., 2022), Los Angeles (Feng et al., 2016), Cape Town (Nickless et al.,, 2018) and Indianapolis (Lauvaux et al.,, 2016) all benefit from the same high precision analyzers.

There are two stations equipped with PICARRO (model G2401) high precise analyzers in MCMA, National Autonomous University of Mexico (UNA) and Altzomoni (ALZ). Both are calibrated to the WMO/GAW requirements (WMO, 2020) since the end of 2018, providing measurements of atmospheric CO_2, CH4 and CO every few seconds. After calibration and quality control processes, the measurements were processed into hourly records with standard deviation. The uncertainties of the measures in the two stations are estimated to be 0.2 ppm and 0.05 ppm (González del Castillo et al., 2022). UNA is an urban station located on the campus of National Autonomous University of Mexico (*Universidad Nacional Autónoma de México*, UNAM), with an altitude of 2304 m. The ALZ background station is located on mountain tops (4000 m a.s.l.), about 60 km from UNA, avoiding most of the impacts of anthropogenic fluxes in the urban area.

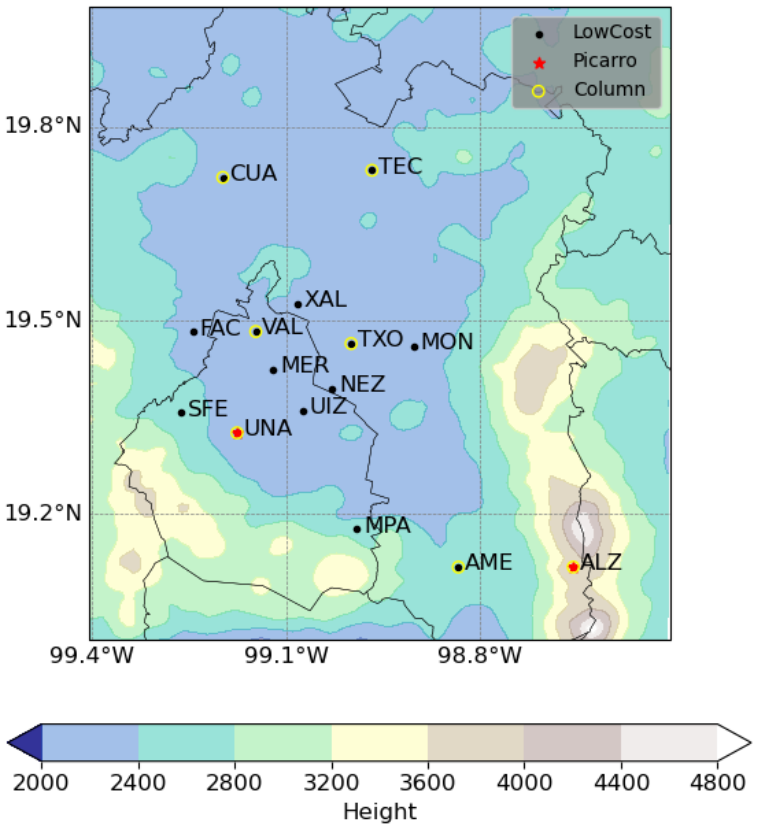


Figure 1.4 *Measurement network envisioned for MERCI-CO_2, with topography and administrative divisions in MCMA.*

Column sensors were also installed at these two sites, a mobile low resolution (0.5 cm^{-1}) EM27/SUN in UNA in 2016 and a high resolution spectrometer HR125 in ALZ in 2012. EM27/SUN is a solar-tracking Fourier transform spectrometer (Gisi et al. 2012) and HR125 is the Fourier Transform Infrared (FTIR) system, known as Bruker IFS 125HR, used as reference instrument for the TCCON network (Poberovskii, 2010). A dense network of seven EM27/SUN in MCMA started running in MCMA in September 2020 (Ramonet, 2020). Additionally, there is a network of LCMP sensors installed and planned all over MCMA, consisting of 13 LCMP sensors, based on the existing meteorological stations operated by SEDEMA (Secretaría del Medio Ambiente de la Ciudad de México, Secretary of the Environment of Mexico City). This later deployment of the LCMP sensors has been delayed due to COVID-19 restrictions starting in Spring 2020. The co-location of the in-situ and column sensors provides a unique opportunity to analyze

the urban CO_2 signals both at the surface level, and integrated through the vertical columns.

1.5 ATMOSPHERIC CO_2 MODELING

In order to quantify the transportation of contaminants or GHGs, atmospheric transport models (with or without chemistry depending on the tracers being considered) are widely used in atmospheric sciences. The various models could be divided into two types according to the form of the continuous function : Eulerian models (box models, as shown in **Figure 1.5**) and Lagrangian models (particle dispersion models) (Jacob, 1999). Eulerian models use fixed coordinates which look like a group of boxes, while Lagrangian models use moving coordinates, which look like a group of puffs.

An atmospheric chemistry model considers the behavior of chemicals in four types of processes: Emission, Transport, Chemistry and Deposition (Jacob, 1999). Because of the stability of CO_2, the chemistry part could be omitted. The main deposition of CO_2, over continents, is the uptake by photosynthesis.

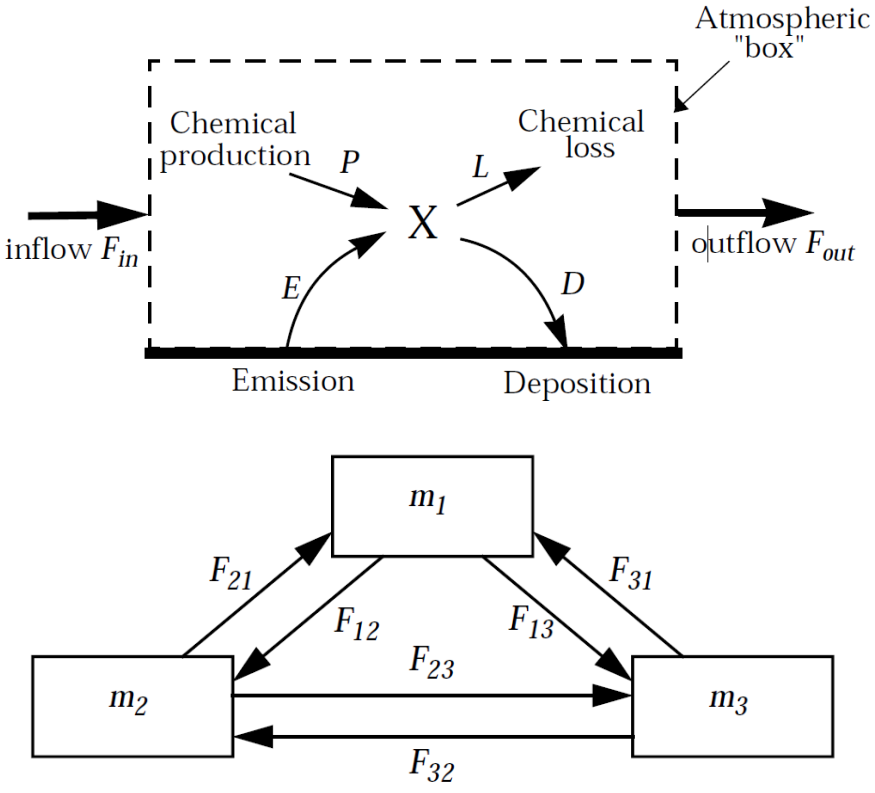


Figure 1.5 *The main mechanism "in the box" and "between boxes" in a typical Eularian model (Jacob, 1999).*

One of the most famous Eulerian models is the Weather Research and Forecasting model coupled with Chemistry (WRF-Chem) developed by the Mesoscale and Microscale Meteorology (MMM) Laboratory of NCAR (Skamarock et al., 2008). As the most widely-used mesoscale atmospheric model, WRF offers robust and reliable simulation of meteorological parameters. It is not only capable of being coupled with its own chemistry module, but also provides input fields for a series of chemistry transport models, such as another Eulerian model, e.g. CHIMERE (Menut et al., 2013; Briant et al., 2017) by LMD (Laboratoire de Météorologie Dynamique) and CMAQ (Community Multiscale Air Quality modeling system) (Byun and Schere, 2006) by U.S. EPA, or providing the meteorological input to Lagrangian models (Angevine et al., 2014).

Lagrangian models, including trajectory model and dispersion model (Lagrangian particle dispersion models, LPDMs), make it possible to track the activities of air mass, and make the footprint of CO_2. The air parcel transportation trajectories could be tracked either forward or backward from sources to measurements. The most popular Lagrangian models includes the Stochastic Time-Inverted Lagrangian Transport (STILT) model (Lin et al., 2003) used in Salt Lake City (Kunik et al., 2019), FLEXPART (FLEXible PARTicle dispersion model) (Pisso et al., 2019b; Brioude et al., 2013) used in Tokyo (Pisso et al., 2019a), the Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) (Stein et al., 2015) used in New York (Pitt et al., 2022).

Both types of models have a long history and wide range of application in urban atmosphere simulation, but there are some challenges. Mexico City lies in a valley with high altitude. The sharp topography changes make the model very unstable with fine vertical layers. In the city, the dense buildings lead to bias of wind speed and directions, heat and energy in the urban areas. Urban Canopy Model (UCM) is a common solution (Tewari et al., 2007). Previous sensitivity tests show that although the single-layer UCM tends to overestimate wind speeds and temperatures, it is enough for our purpose (Salamanca et al., 2011). Basin topography can cause the accumulation of signals under continuous low wind speeds, which may lead to errors in simulations

and inversions (Borsdorff, et al., 2020).

WRF-Chem, one of the most popular Eulerian models in the world, is an online-coupled mesoscale model that has successfully worked on the CO_2 transportation in cities (Feng et al., 2016; Zheng et al., 2019; Zhou W. et al., 2020; Lian et al., 2021). A modified version of WRF-Chem was settled in this study to trace CO_2 transport. There are also some successful cases combining WRF with Lagarangian models in urban CO_2 inversion with footprints, e.g. FLEXPART-WRF (Pisso, 2019), WRF-STILT (Kunik, 2019) and WRF-LPDM (Lauvaux et al., 2016; Nalini et al., 2022). Considering the data availability in our simulating period, we are not using any explicit tower footprints with a Lagrangian model in this study.

WRF provides rich options on the physical parameters. In order to cooperate with the options of UNAM study, the Planet Boundary Layer (PBL) scheme was MYJ (Mellor–Yamada–Janjić; Janjić, 1994; Janjić, 2001) at the first test. For the future application of turbulent kinetic energy (TKE) and the adaptation to the modified version of WRF-Chem, it was changed to MYNN (Mellor–Yamada–Nakanishi–Niino; Nakanishi and Niino, 2004; Nakanishi and Niino, 2009). The two are very similar. It is reported that nudging would improve the model performance in wind conditions (Deng et al., 2017), so an attempt of nudging was tested as well.

1.6 OBJECTIVES AND OUTLINE OF THE THESIS

The main objective of this thesis is to analyze the signals induced by the anthropogenic and biogenic CO_2 fluxes in the atmosphere of Mexico-City. The analysis of the temporal and spatial characteristics of the CO_2 concentrations, at the surface and through the vertical columns, should inform us about the intensity of the emissions in MCMA. To carry out this task I used the atmospheric transport model WRF in an optimal configuration thoroughly evaluated from meteorological measurements in and around MCMA. To simulate atmospheric CO_2 concentrations with this model, I also had to prepare a set of inventories of anthropogenic and biospheric emissions. In the end, the thesis aims to use the atmospheric concentrations by an inversion system assimilating the measurements to the modeled concentrations, to reduce the uncertainties on the inventory of

anthropogenic emissions of MCMA.

More specifically, the research questions contained in this thesis are:

(1) How does the WRF model with different configurations work on the simulation of meteorological parameters in MCMA?
(2) Could the WRF-Chem modeling system reproduce the surface CO_2 and total column CO_2 (XCO_2) atmospheric concentrations in MCMA in various weather conditions? What are the characteristics of CO_2 concentration variations in MCMA?
(3) Is ALZ representing the background of CO_2 in UNA? What makes a good background station?
(4) Can we successfully reduce the uncertainties from the prior emission inventories with the available measurements? How do the CO_2 emissions in MCMA vary with time and space?
(5) What could we interpret from the data assimilation between the CO_2 measurements and simulations?

Following this introductory chapter, the thesis consists of 3 main chapters. Chapter 2 and 3 form a publishable paper together, with separate summaries, while Chapter 4 is another publishable paper with a summary. There are also unpublished tests in Chapter 2 for a series of attempts.

Chapter 2 describes the various options on drivers, domains, physics schemes and spectral nudging of the WRF modeling system over MCMA and makes a quantitative evaluation of the model performance. The modeled meteorological parameters are compared with in-situ surface observations, LiDar analysis as well as WMO radiosonde observations. For the purpose of CO_2 simulation in the next chapters, the most focused variables are the ones related to dispersion of the air, containing air temperatures, wind speeds, wind directions and PBL heights. The results of these sensitivity tests are applied in the final modeling system used in the next chapters.

Chapter 3 aims at the reconstruction of the CO_2 and XCO_2 concentration maps over MCMA during 3 typical months (January, May and July) in the dry and rainy seasons in 2018, the characteristics of CO_2 temporal variation, and the network design of CO_2 observation

stations. The CO_2 concentration maps are reconstructed by the 3-km resolution WRF-Chem coupled with the urban canopy model (UCM) through the local emission inventories from UNAM (García, 2020) and global emission inventories ODIAC scaled by temporal scaling factors (Nassar et al., 2013). The evaluation of CO_2 simulations are made on the basis of CO_2 in-situ measurements and column measurements (XCO_2) by FTIR in the urban station UNA and the background station ALZ. Along with the evaluation, we also analyze the temporal and spatial distribution of CO_2 signals, including diurnal cycle, weekly variation and seasonal variation, as well as the area impacted by anthropogenic fluxes and the variation of biogenic fluxes. Based on the analysis, we assessed the reasonability of the current location of the background station (ALZ), and defined a "background index" to represent the suitability to build a background station on the map.

Chapter 4, the key part of the thesis, describes the Bayesian inversion system and its results. The inversion adjusts the prior anthropogenic emission from UNAM and ODIAC, and prior biogenic fluxes from the Carnegie Ames Stanford Approach (CASA) model previously optimized using eddy-flux tower data (Zhou et al., 2020), based on the data assimilation between the gradients from the urban station to the rural station. while the boundary conditions from CarbonTracker 2019B (Peters et al., 2007) remain the same. A series of inversion plans were applied to a 1-year simulation covering MCMA from March 30 2018 to March 30 2019. The plans basically contain 3 variables, standing for fossil fuel sources, biofluxes and background concentrations to generate separate scaling factors for each block of 5 days. Covariances between continuous days, time windows in a day, sectors from mobile/area/point, a MCMA mask and changed block sizes are added to different plans, to test the sensitivities of these options, to improve the performance of the inversion, and to specify the impact of urban signal and of various departments. The same system was also applied on carbon monoxide (CO), which observations start from December 2018. The interest of CO is that it is a tracer co-emitted with CO_2 by anthropogenic sources, but without major emissions from the biosphere, excluding biomass fires.

Chapter 5 summarizes the main conclusions from this thesis and presents the perspectives of this project, as well as the potential future

research directions.

REFERENCES

AirParif: 2019 emissions and consumption in Île-de-France (estimates made in 2022), available at: https://www.airparif.asso.fr/surveiller-la-pollution/les-emissions (last access: 29 December 2022), 2022

Arias, P.A., N. Bellouin, E. Coppola, R.G. Jones, G. Krinner, J. Marotzke, V. Naik, M.D. Palmer, G.-K. Plattner, J. Rogelj, M. Rojas, J. Sillmann, T. Storelvmo, P.W. Thorne, B. Trewin, K. Achuta Rao, B. Adhikary, R.P. Allan, K. Armour, G. Bala, R. Barimalala, S. Berger, J.G. Canadell, C. Cassou, A. Cherchi, W. Collins, W.D. Collins, S.L. Connors, S. Corti, F. Cruz, F.J. Dentener, C. Dereczynski, A. Di Luca, A. Diongue Niang, F.J. Doblas-Reyes, A. Dosio, H. Douville, F. Engelbrecht, V. Eyring, E. Fischer, P. Forster, B. Fox-Kemper, J.S. Fuglestvedt, J.C. Fyfe, N.P. Gillett, L. Goldfarb, I. Gorodetskaya, J.M. Gutierrez, R. Hamdi, E. Hawkins, H.T. Hewitt, P. Hope, A.S. Islam, C. Jones, D.S. Kaufman, R.E. Kopp, Y. Kosaka, J. Kossin, S. Krakovska, J.-Y. Lee, J. Li, T. Mauritsen, T.K. Maycock, M. Meinshausen, S.-K. Min, P.M.S. Monteiro, T. Ngo-Duc, F. Otto, I. Pinto, A. Pirani, K. Raghavan, R. Ranasinghe, A.C. Ruane, L. Ruiz, J.-B. Sallée, B.H. Samset, S. Sathyendranath, S.I. Seneviratne, A.A. Sörensson, S. Szopa, I. Takayabu, A.-M. Tréguier, B. van den Hurk, R. Vautard, K. von Schuckmann, S. Zaehle, X. Zhang, and K. Zickfeld: Technical Summary. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2021, pp. 33–144. doi:10.1017/9781009157896.002.

Boon, A., Broquet, G., Clifford, D. J., Chevallier, F., Butterfield, D. M., Pison, I., Ramonet, M., Paris, J.-D., and Ciais, P.: Analysis of the potential of near-ground measurements of CO_2 and CH4 in London, UK, for the monitoring of city-scale emissions using an atmospheric transport model, Atmos. Chem. Phys., 16, 6735–6756, https://doi.org/10.5194/acp-16-6735-2016, 2016.

Borsdorff, T., García Reynoso, A., Maldonado, G., Mar-Morales, B., Stremme, W., Grutter, M., and Landgraf, J.: Monitoring CO emissions of the metropolis Mexico City using TROPOMI CO observations, Atmos. Chem. Phys., 20, 15761–15774, https://doi.org/10.5194/acp-20-15761-2020, 2020.

Bréon, F. M., Broquet, G., Puygrenier, V., Chevallier, F., Xueref-Remy, I., Ramonet, M., Dieudonné, E., Lopez, M., Schmidt, M., Perrussel, O., and Ciais, P.: An attempt at estimating Paris area CO_2 emissions from atmospheric concentration measurements, Atmos. Chem. Phys., 15, 1707–1724, https://doi.org/10.5194/acp-15-1707-2015, 2015.

Briant, R., Tuccella, P., Deroubaix, A., Khvorostyanov, D., Menut, L., Mailler, S., and Turquety, S.: Aerosol-radiation interaction modelling using online coupling between the WRF 3.7.1 meteorological model and the CHIMERE 2016 chemistry-transport model, through the OASIS3-MCT coupler, Geosci. Model Dev., 10, 927-944, doi:10.5194/gmd-10-927-2017, 2017

Brioude, J., D. Arnold, A. Stohl et al. (2013): The Lagrangian particle dispersion model FLEXPART-WRF version 3.1, Geosci. Model Dev., 6, 1889-1904, https://doi.org/10.5194/gmd-6-1889-2013.

Byun, D. & Schere, K.L. (2006). Review of the governing equations, computational algorithms, and other components of the Model-3 Community Multiscale Air Quality (CMAQ) Modeling System. Appl. Mech. Rev. 59(2): 51-77. doi: 10.1115/1.2128636

CDMX Government, Mexico State Government, Hidalgo State Goverment: Management program to improve AIR QUALITY in the Mexico City Metropolitan Area, available at: http://www.aire.cdmx.gob.mx/descargas/publicaciones/flippingbook/proaire2021-2030/ (last access: 29 December 2022), 2021.

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.

Deng, A., Lauvaux, T., Davis, K. J., Gaudet, B. J., Miles, N., Richardson, S. J., ... & Gurney, K. R. (2017). Toward reduced transport errors in a high resolution urban CO_2 inversion system. Elementa: Science of the Anthropocene, 5.

Eggleston, H. S., Buendia, L., Miwa, K., Ngara, T., and Tanabe, K.: Volume 4: Agriculture, forestry and land use, in: 2006 IPCC guidelines for national greenhouse gas inventories, https://www.ipccnggip.iges.or.jp/public/2006gl/vol4.html (last access: 25 September 2022), 2006.

Enting, I. G. (2002). Inverse problems in atmospheric constituent transport. Cambridge University Press.

Feng, S., Lauvaux, T., Newman, S., Rao, P., Ahmadov, R., Deng, A., Díaz-Isaac, L. I., Duren, R. M., Fischer, M. L., Gerbig, C., Gurney, K. R., Huang, J., Jeong, S., Li, Z., Miller, C. E., O'Keeffe, D., Patarasuk, R., Sander, S. P., Song, Y., Wong, K. W., and Yung, Y. L.: Los Angeles megacity: a high-resolution land–atmosphere modelling system for urban CO_2 emissions, Atmos. Chem. Phys., 16, 9019–9045, https://doi.org/10.5194/acp-16-9019-2016, 2016.

Fong, W. K., Sotos, M., Schultz, S., Deng-Beck, C., Marques, A., & Doust, M. (2015). Global protocol for community-scale greenhouse gas emission inventories. desLibris. 2015.

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Bakker, D. C., Hauck, J., ... & Zeng, J. (2022). Global carbon budget 2021. Earth System Science Data, 14(4), 1917-2005.

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Gregor, L., Hauck, J., ... & Zheng, B. (2022). Global carbon budget 2022. Earth System Science Data, 14(11), 4811-4900.

García Reynoso J.A.: National Emissions Inventory 2016, accessible at: https://github.com/JoseAgustin/emis_2016 (last access: 29 December 2022), 2019.

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

Gately, C., and L.R. Hutyra. 2018. CMS: CO_2 Emissions from Fossil Fuels Combustion, ACES Inventory for Northeastern USA. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1501

GHG Protocol: GHG Protocol Corporate Accounting and Reporting Standard (GHG Protocol Corporate Standard), revised edition, 2015.

Gisi, M., Hase, F., Dohe, S., Blumenstock, T., Simon, A., and Keens, A.: XCO_2-measurements with a tabletop FTS using solar absorption spectroscopy, Atmos. Meas. Tech., 5, 2969–2980, https://doi.org/10.5194/amt-5-2969-2012, 2012.

GMAO (Global Modeling and Assimilation Office) (2015), inst3_3d_asm_Cp: MERRA-2 3D IAU State, Meteorology Instantaneous 3-hourly (p-coord, 0.625x0.5L42), version 5.12.4, Greenbelt, MD, USA: Goddard Space Flight Center Distributed Active Archive Center (GSFC DAAC), Accessed Enter User Data Access Date at doi: 10.5067/VJAFPLI1CSIV, re-analyzed by weatherspark: https://es.weatherspark.com/y/5674/Clima-promedio-en-Ciudad-de-M%C3%A9xico-M%C3%A9xico-durante-todo-el-a%C3%B1o

González del Castillo, E., Taquet, N., Bezanilla, A., Stremme, W., Ramonet, M., Laurent, O., ... & Grutter, M. (2022). CO_2 variability in the Mexico City region from in situ measurements at an urban and a background site. Atmosphere , 35 (2), 377-393.

Gurney, K. R., Razlivanov, I., Song, Y., Zhou, Y., Benes, B., & Abdul-Massih, M. (2012). Quantification of fossil fuel CO_2 emissions on the building/street scale for a large US city. Environmental science & technology, 46(21), 12194-12202.

Hansen, J., Sato, M., Ruedy, R., Lo, K., Lea, D. W., & Medina-Elizade, M: Global temperature change. Proceedings of the National Academy of Sciences, 103(39), 14288-14293, 2006.

Helfter, C., Tremper, A. H., Halios, C. H., Kotthaus, S., Bjorkegren, A., Grimmond, C. S. B., ... & Nemitz, E. (2016). Spatial and temporal variability of urban fluxes of methane, carbon monoxide and carbon dioxide above London, UK. Atmospheric Chemistry and Physics, 16(16), 10543-10557.

I. Pisso, E. Sollum, H. Grythe, N. I. Kristiansen, M. Cassiani, S. Eckhardt, D. Arnold, D. Morton, R. L. Thompson, C. D. Groot Zwaaftink, N. Evangeliou, H. Sodemann, L. Haimberger, S. Henne, D. Brunner, J. F. Burkhart, A. Fouilloux, J. Brioude, A. Philipp, P. Seibert, and A. Stohl. (2019), The Lagrangian particle dispersion model FLEXPART version 10.4. Geosci. Model Dev., 12, 4955–4997, URL: https://doi.org/10.5194/gmd-12-4955-2019.

INEGI (Instituto Nacional de Estadística y Geografía, Mexico National Institute of Statistics and Geography): Registered Motor Vehicles in Circulation, available at: https://www.inegi.org.mx/programas/vehiculosmotor/ (last access: 29 December 2022), 2021.

Jacob, D. J. (1999). Introduction to atmospheric chemistry. Princeton University Press.

Janjić, Z. I., 1994: The step-mountain Eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. Mon. Wea. Rev., 122, 927–945.

Janjić, Z. I., 2001: Nonsingular implementation of the Mellor-Yamada level 2.5 scheme in the NCEP Meso Model. NCEP Office Note 437, 61 pp.

Kaminski T, Scholze M, Rayner P, Houweling S, Voßbeck M, Silver J, Lama S, Buchwitz M, Reuter M, Knorr W, Chen HW, Kuhlmann G, Brunner D, Dellaert S, Denier van der Gon H, Super I, Löscher A and Meijer Y (2022) Assessing the Impact of Atmospheric CO_2 and NO2 Measurements From Space on Estimating City-Scale Fossil Fuel CO_2 Emissions in a Data Assimilation System. Front. Remote Sens. 3:887456. doi: 10.3389/frsen.2022.887456

Kunik, L., Mallia, D. V., Gurney, K. R., Mendoza, D. L., Oda, T., & Lin, J. C. (2019). Bayesian inverse estimation of urban CO_2 emissions: Results from a synthetic data simulation over Salt Lake City, UT. Elementa: Science of the Anthropocene, 7.

LAEI (London Atmospheric Emissions Inventory): London Datastore, available at: https://data.london.gov.uk/dataset/london-atmospheric-emissions-inventory--laei--2019 (last access: 29 December 2022), 2021

Lauvaux, T., Miles, N. L., Deng, A., Richardson, S. J., Cambaliza, M. O., Davis, K. J., ... & Wu, K. (2016). High‐resolution atmospheric inversion of urban CO_2 emissions during the dormant season of the Indianapolis Flux Experiment (INFLUX). Journal of Geophysical Research: Atmospheres, 121(10), 5213-5236.

Le Quéré, C., Jackson, R.B., Jones, M.W. et al. Temporary reduction in daily global CO_2 emissions during the COVID-19 forced

confinement. Nat. Clim. Chang. 10, 647–653 (2020). https://doi.org/10.1038/s41558-020-0797-x

Lian, J., Bréon, F. M., Broquet, G., Lauvaux, T., Zheng, B., Ramonet, M., ... & Ciais, P. (2021). Sensitivity to the sources of uncertainties in the modeling of atmospheric CO 2 concentration within and in the vicinity of Paris. Atmospheric Chemistry and Physics, 21(13), 10707-10726.

Lian, J., Bréon, F. M., Broquet, G., Zheng, B., Ramonet, M., & Ciais, P. (2020). Quantitative evaluation of the uncertainty sources for the modeling of atmospheric CO_2 concentration within and in the vicinity of Paris city. Atmospheric Chemistry and Physics Discussions, 2020, 1-22.

Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C., Davis, K. J., and Grainger, C. A.: A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, Journal of Geophysical Research-Atmospheres, 108, 4493, doi:10.1029/2002JD003161, 2003.

Liu, Z., Ciais, P., Deng, Z. et al. Near-real-time monitoring of global CO_2 emissions reveals the effects of the COVID-19 pandemic. Nat Commun 11, 5172 (2020). https://doi.org/10.1038/s41467-020-18922-7

Lopez-Coto, I., Ren, X., Salmon, O. E., Karion, A., Shepson, P. B., Dickerson, R. R., ... & Whetstone, J. R. (2020). Wintertime CO_2, CH4, and CO emissions estimation for the Washington, DC–Baltimore metropolitan area using an inverse modeling technique. Environmental science & technology, 54(5), 2606-2614.

McKain, K., Down, A., Raciti, S. M., Budney, J., Hutyra, L. R., Floerchinger, C., ... & Wofsy, S. C. (2015). Methane emissions from natural gas infrastructure and use in the urban region of Boston, Massachusetts. Proceedings of the National Academy of Sciences, 112(7), 1941-1946.

Menut L, B.Bessagnet, D.Khvorostyanov, M.Beekmann, N.Blond, A.Colette, I.Coll, G.Curci, G.Foret, A.Hodzic, S.Mailler, F.Meleux, J.L.Monge, I.Pison, G.Siour, S.Turquety, M.Valari, R.Vautard and

M.G.Vivanco, 2013, CHIMERE 2013: a model for regional atmospheric composition modelling, Geoscientific Model Development, 6, 981-1028, doi:10.5194/gmd-6-981-2013,

Metz, B., Davidson, O., De Coninck, H. C., Loos, M., & Meyer, L.: IPCC special report on carbon dioxide capture and storage. Cambridge: Cambridge University Press, 2005.

Nakanishi, M., and Niino H. , 2004: An improved Mellor–Yamada level-3 model with condensation physics: Its design and verification. Bound.-Layer Meteor., 112, 1–31.

Nakanishi, M., and Niino H. , 2009: Development of an improved turbulence closure model for the atmospheric boundary layer. J. Meteor. Soc. Japan, 87, 895–912.

Nalini, K., Lauvaux, T., Abdallah, C., Lian, J., Ciais, P., Utard, H., ... & Ramonet, M. (2022). High - Resolution Lagrangian Inverse Modeling of CO_2 Emissions Over the Paris Region During the First 2020 Lockdown Period. Journal of Geophysical Research: Atmospheres, 127(14), e2021JD036032.

Nickless, A., Rayner, P. J., Engelbrecht, F., Brunke, E.-G., Erni, B., and Scholes, R. J.: Estimates of CO_2 fluxes over the city of Cape Town, South Africa, through Bayesian inverse modelling, Atmos. Chem. Phys., 18, 4765–4801, https://doi.org/10.5194/acp-18-4765-2018, 2018.

NOAA, https://gml.noaa.gov/ccgg/trends/global.html (last access: 2022/12/29)

Oda, T. and Maksyutov, S.: A very high-resolution (1 km×1 km) global fossil fuel CO_2 emission inventory derived using a point source database and satellite observations of nighttime lights, Atmos. Chem. Phys., 11, 543–556, https://doi.org/10.5194/acp-11-543-2011, 2011.

Olivier, J.G.J., Janssens-Maenhout, G., Muntean, M. and Peters, J.A.H.W. (2015) Trends in global CO_2 emissions: 2015 Report. PBL Netherlands Environmental Assessment Agency, The Hague; European Commission, Joint Research Centre (JRC), Institute for Environment and Sustainability (IES). Internet: http://edgar.jrc.ec.europa.eu/news_docs/jrc-2015-trends-in-

global-CO_2-emissions-2015-report-98184.pdf. JRC report 98184/ PBL report 1803, 2015.

Paris Agreement: Paris agreement. In Report of the Conference of the Parties to the United Nations Framework Convention on Climate Change (21st Session, 2015: Paris). Retrived December (Vol. 4, p. 2017). HeinOnline.

Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., ... & Tans, P. P. (2007). An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker. Proceedings of the National Academy of Sciences, 104(48), 18925-18930.

Pisso, I., Patra, P., Takigawa, M. et al. Assessing Lagrangian inverse modelling of urban anthropogenic CO_2 fluxes using in situ aircraft and ground-based measurements in the Tokyo area. Carbon Balance Manage 14, 6 (2019). https://doi.org/10.1186/s13021-019-0118-8

Pitt, JR, Lopez-Coto, I, Hajny, KD, Tomlin, J, Kaeser, R, Jayarathne, T, Stirm, BH, Floerchinger, CR, Loughner, CP, Gately, CK, Hutyra, LR, Gurney, KR, Roest, GS, Liang, J, Gourdji, S, Karion, A, Whetstone, JR, Shepson, PB. 2022. New York City greenhouse gas emissions estimated with inverse modeling of aircraft measurements. Elementa: Science of the Anthropocene 10(1). DOI: https://doi.org/10.1525/elementa.2021.00082

Poberovskii, A.V. High-resolution ground measurements of the IR spectra of solar radiation. Atmos Ocean Opt 23, 161–163 (2010). https://doi.org/10.1134/S1024856010020132

R. Cong, M. Saito, T. Oda, T. Fukui, R. Hirata, and A. Ito, A high-resolution, spatially explicit estimate of fossil-fuel CO_2 emissions from the Tokyo Metropolis, Japan, 2020, PREPRINT (Version 4) available at Research Square (doi:10.21203/rs.3.rs-17565/v1).

Ramonet, M., Taquet, N., & Grutter, M. (2020, May). Intensive CO_2 and CH4 measurement campaign at Mexico-City. In EGU General Assembly Conference Abstracts (p. 20364).

Rayner, P. J., M. R. Raupach, M. Paget, P. Peylin, and E. Koffi (2010), A new global gridded data set of CO_2 emissions from fossil fuel

combustion: Methodology and evaluation, J. Geophys. Res., 115, D19306, doi:10.1029/2009JD013439

Rella, C. W., Chen, H., Andrews, A. E., Filges, A., Gerbig, C., Hatakka, J., ... & Zellweger, C. (2013). High accuracy measurements of dry mole fractions of carbon dioxide and methane in humid air. Atmospheric Measurement Techniques, 6(3), 837-860.

Rypdal, K., & Winiwarter, W.: Uncertainties in greenhouse gas emission inventories—evaluation, comparability and implications. Environmental Science & Policy, 4(2-3), 107-116. 2001.

Salamanca, F., Martilli, A., Tewari, M., & Chen, F. (2011). A study of the urban boundary layer using different urban parameterizations and high-resolution urban canopy parameters with WRF. Journal of Applied Meteorology and Climatology, 50(5), 1107-1128.

SEDEMA (Secretaría del Medio Ambiente de la Ciudad de México, Secretary of the Environment of Mexico City): Contaminantes y de efecto invernadero 2012. Secretaría del Medio Ambiente de la Ciudad de México, available at: http://www.aire.cdmx.gob.mx/descargas/publicaciones/flippingbook/inventario-emisioneszmvm2012/inventario-emisioneszmvm2012.pdf

SEDEMA (Secretaría del Medio Ambiente de la Ciudad de México, Secretary of the Environment of Mexico City): Inventario de emisiones de la Zona Metropolitana del Valle de México 2018. Contaminantes criterio, tóxicos y compuestos de efecto invernadero. Secretaría del Medio Ambiente de la Ciudad de México (datos preliminares), 2021.

Seto, K. C., Dhakal, S., Bigio, A., Blanco, H., Delgado, G. C., Dewar, D., Huang, L., Inaba, A., Kansal, A., Lwasa, S., McMahon, J., Müller, D. B., Murakami, J., Nagendra, H., and Ramaswami, A.: Climate Change 2014: Mitigation of Climate Change. IPCC Working Group III Contribution to AR5. Chap. 12 - Human settlements, infrastructure and spatial planning. Cambridge University Press, Cambridge, UK and New York, NY, USA, 2014.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, M. G Duda, X.-Y. Huang, W. Wang, and J. G. Powers, 2008: A Description

of the Advanced Research WRF Version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp. doi:10.5065/D68S4MVH

Staufer, J., Broquet, G., Bréon, F. M., Puygrenier, V., Chevallier, F., Xueref-Rémy, I., ... & Ciais, P. (2016). The first 1-year-long estimate of the Paris region fossil fuel CO 2 emissions based on atmospheric inversion. Atmospheric Chemistry and Physics, 16(22), 14703-14726.

Stein, A. F., R. R. Draxler, G. D. Rolph, B. J. B. Stunder, M. D. Cohen, and F. Ngan. " NOAA's HYSPLIT Atmospheric Transport and Dispersion Modeling System", Bulletin of the American Meteorological Society 96, 12 (2015): 2059-2077, accessed Oct 7, 2021, https://doi.org/10.1175/BAMS-D-14-00110.1

SUSMITA DASGUPTA, SOMIK LALL, DAVID WHEELER: Cutting global carbon emissions: where do cities stand?. Sustainable Cities, JANUARY 05, 2022.

Tarantola, A.: Inverse problem theory and methods for model parameter estimation, Society for Industrial and Applied Mathematics, Philadelphia, PA, xii, 342 pp., 2005.

Tewari, M., Chen, F., Kusaka, H., & Miao, S. (2007). Coupled WRF/Unified Noah/urban-canopy modeling system. In Ncar WRF Documentation, NCAR, Boulder (Vol. 122, pp. 1-22).

The World Bank: What You Need to Know About Emission Reductions Payment Agreements (ERPAs), https://www.worldbank.org/en/news/feature/2021/05/19/what-you-need-to-know-about-emission-reductions-payment-agreements, 2021(Accessed 2022-12-29).

Tomohiro Oda, Shamil Maksyutov (2015), ODIAC Fossil Fuel CO_2 Emissions Dataset (Version name : ODIAC2019), Center for Global Environmental Research, National Institute for Environmental Studies, doi:10.17595/20170411.001. (Reference date : 2022/05/22)

UN (United Nations), Department of Economic and Social Affairs, Population Division (2018). The World's Cities in 2018—Data Booklet (ST/ESA/ SER.A/417).

UN: World Urbanization Prospects: The 2018 Revision, File 03 - Urban_Population, 2019.

Wayne M. Angevine, Jérôme Brioude, Stuart Mckeen, S. Holloway John. Uncertainty in Lagrangian pollutant transport simulations due to meteorological uncertainty from a mesoscale WRF ensemble. Geoscientific Model Development, 2014, 7 (6), pp.2817 - 2829. ff10.5194/gmd-7-2817-2014. hal01348995

Wei, C., Wang, M., Fu, Q., Dai, C., Huang, R., & Bao, Q. (2020). Temporal characteristics of greenhouse gases (CO_2 and CH4) in the megacity Shanghai, China: Association with air pollutants and meteorological conditions. Atmospheric Research, 235, 104759.

Wei, T., Wu, J., & Chen, S.: Keeping track of greenhouse gas emission reduction progress and targets in 167 cities worldwide. Frontiers in Sustainable Cities, 3, 2021.

Wigley, T. (1983). The pre-industrial carbon dioxide level. Climatic change, 5(4), 315-320.

WMO (2020), 20th WMO/IAEA Meeting on Carbon Dioxide, Other Greenhouse Gases and Related Tracers Measurement Techniques (GGMT-2019), Jeju Island, South Korea, GAW Rep. No. 255, Geneva, Switzerland.

World Meteorological Organization: WMO GREENHOUSE GAS BULLETIN No. 17: The State of Greenhouse Gases in the Atmosphere Based on Global Observations through 2020, 2021.

Yver Kwok, C., Laurent, O., Guemri, A., Philippon, C., Wastine, B., Rella, C. W., ... & Ramonet, M. (2015). Comprehensive laboratory and field testing of cavity ring-down spectroscopy analyzers measuring H 2 O, CO 2, CH 4 and CO. Atmospheric Measurement Techniques, 8(9), 3867-3892.

Zheng, T., Nassar, R., & Baxter, M. (2019). Estimating power plant CO_2 emission using OCO-2 XCO_2 and high resolution WRF-Chem simulations. Environmental Research Letters, 14(8), 085001.

Zhou, W., Niu, Z., Wu, S., Xiong, X., Hou, Y., Wang, P., ... & Zhu, Y. (2020). Fossil fuel CO_2 traced by radiocarbon in fifteen Chinese cities. Science of The Total Environment, 729, 138639.

Zhou, Y., C. A. Williams, T. Lauvaux, K. J. Davis, S. Feng, I. Baker, S. Denning, & Y. Wei. 2020. A multiyear gridded data ensemble of surface biogenic carbon fluxes for North America: Evaluation and

analysis of results. Journal of Geophysical Research: Biogeosciences, 125(2), e2019JG005314. https://doi.org/10.1029/2019JG005314

2 CHAPTER II : EVALUATION OF CONFIGURATIONS AND OPTIONS IN WRF MODELING SYSTEM OVER MEXICO CITY METROPOLITAN AREA

SUMMARY

The quality of atmospheric inversion results depends heavily on the performance of the atmospheric transport model. The Weather Research and Forecasting model (WRF) is a widely used Eulerian atmospheric transport model, which has been applied in various urban-scale inversion studies, but the best selection of physical parameters for urban-scale simulation varies from city to city. The main objective of this chapter is to investigate the performance of various configurations of the model and to determine the configuration in the future simulations. A period of 2 weeks in January 2018 is selected for our experiments.

A set of triple-nesting domains was initialized in WRF over the Mexico City Metropolitan Area (MCMA). Mexico City is a basin city at high altitude, and such complex terrains are prone to cause mismatch in simulations. In order to evaluate how the WRF modeling system reproduces the transport conditions of CO_2, validation with observations was made on the meteorological variables, e.g, wind speed, wind direction, temperature and Planetary Boundary Layer (PBL) height. The first three variables are measured by the SEDEMA (Secretaría del Medio Ambiente de la Ciudad de México, Secretary of the Environment of Mexico City) network. PBL heights and the vertical profile of horizontal wind speed are diagnosed from a Lidar installed at the National Autonomous University of Mexico (UNAM).

For the sake of the optimization of model inputs, several key configurations and parameters in the WRF modeling system are tested, including PBL scheme, nudging strategy, urban canopy model, meteorological drivers, and the CO_2 concentrations for boundary conditions. The reanalysis product ERA-5 repaired the temperature bias in the old version ERA-interim. The PBL scheme MYNN reproduces the PBL heights better than MYJ and YSU. Observation nudging failed in our domains and grid nudging worsened the fit between simulations

and observations.

A relatively optimal configuration was then applied in WRF-Chem and simulated the CO_2 transportations. I then evaluated the simulations of CO_2 mixing ratio with two high-precise observation stations from MERCI-CO_2 observation network, located at the south of the city center (UNA) and in the background area (ALZ). The urban canopy model BEP with default parameters is better than UCM in most meteorological variables, and succeeded in capturing several peaks in CO_2 concentrations, but it didn't improve CO_2 mean absolute errors (MAE). The moderate resolution 5-km domain tends to overestimate the CO_2 concentrations at night and in the morning than the high resolution 1-km domain, but it captured the observation trends in the afternoon. There is no significant difference in the MAEs during the afternoon. This makes it possible to apply lower resolution models to increase the computational speed.

With the 5-km domain WRF coupled with UCM, I analyzed the model performance and the characteristics of the meteorological conditions in 3 time periods: January for the dry season, July for wet season, and May for the transition between the two. PBL heights, surface temperatures, surface wind speed and wind directions, and vertical profile of horizontal wind speeds in the three seasons are all satisfying for inversion.

The main part of this chapter, as well as Chapter 3 *Atmospheric CO_2 sources, spatial distribution and seasonal variability over Mexico City Metropolitan Area*, forms a manuscript *Evaluation of atmospheric CO_2 simulation over the Mexico City metropolitan area by WRF-chem* submitted to *Sustainable Cities and Society*.

Evaluation of atmospheric CO_2 simulation over the Mexico City metropolitan area by WRF-chem

Yang Xu[1], T. Lauvaux[1,2], M. Grutter[3], N. Taquet[3], JA. García-Reynoso[3], O.Laurent[1], M. Lopez[1], J. Lian[1], X. Lin[1], W. Stremme[3], M. Ramonet[1]

[1] Laboratoire des Sciences du Climat et de l'Environnement (LSCE), IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France

[2] H (GSMA), Université de Reims Champagne Ardenne, UMR CNRS 7331, Reims, France

[3] Centro de Ciencias de la Atmósfera, Universidad Nacional Autónoma de México, México

ABSTRACT

Atmospheric inversions have been utilized across multiple cities to improve self-reported emission inventories, but the atmospheric dynamics over large metropolitan areas located in complex topography remain challenging for atmospheric transport models. Here, we evaluate the performance of an atmospheric CO_2 modeling system and assess the optimal deployment of measurement sites to capture the whole-city emissions. As part of the French-Mexican project "Mexico City's Regional Carbon Impacts (MERCI-CO_2)", two observation stations of surface and two column CO_2 sensors were deployed in the Mexico City metropolitan area. We show that our modeling system is able to capture the local atmospheric dynamics with WRF at high resolution. The mean error in wind speed is 0.69 m/s and the relative error in PBL is 6.4%. The model also simulates CO_2 gradients observed by in situ and column sensors during dry and rainy seasons. The relative error of CO_2 surface mixing ratio in the afternoon is -4.3%. The current network is deemed sufficient to capture city emissions (as a whole - not with intra-urban information) and to define the background conditions. Mountains are ideal to capture the background but the biogenic signals are responsible for large CO_2 spatial gradients and temporal variations. Therefore, we conclude that biogenic fluxes must be optimized within future inversion systems. However, we demonstrate how stations in the valley and mountain are

better simulated compared to those on the basin slopes, a deterrent to the deployment of sites in the suburbs of Mexico.

Keywords: urban emissions, fossil fuel, greenhouse gas measurements, atmospheric modeling, Mexico City, WRF-chem, network design

2.1 INTRODUCTION

Atmospheric greenhouse gas (GHG) concentrations have significantly increased at global scale since the industrial revolution. Carbon dioxide (CO_2), the main contributor to climate change, has seen its concentration reaching 414.7 ppm in 2021 (NOAA, 2022) compared to its pre-industrial level of 260-270 ppm (Wigley, 1983). Among the major sources of CO_2, emissions from cities represent more than 70% of global fossil fuel CO_2 emissions (IPCC, WG2, 2014). Countries and sub-national governments around the world are implementing policies to mitigate CO_2 emissions (IPCC, WG3, 2014). The European Climate Law sets a target to reduce the net GHG emissions of the European Union by at least 55% by 2030 compared to 1990 levels[1]. At the sub-national level, cities have also pledged to reduce their emissions, organized into large international consortiums. About 100 large metropolitan areas have joined the C40 City consortium, a network of mayors advocating for climate policies in line with the 1.5°C ambition of the Paris Agreement[2]. In parallel, more than 12,000 cities from 142 countries joined the Global Covenant of Mayors, committed to reduce their annual GHG emissions by 1.9 Gt CO_2 in 2030, compared to a business-as-usual (BAU) trajectory[3].

Accounting of GHG emissions is performed following varied established protocols resulting in uncertainties and inconsistencies across Self-Reported Inventories (SRI) of CO_2 emissions (Chen et al., 2019; Mueller et al., 2021; Gurney et al., 2021). Other non-standard approaches have also been developed to estimate city-scale emissions using spatial and temporal disaggregation methods of national-scale emissions, based on satellite night light data, a proxy of human activities, in combination with activity data from power plants and

[1] https://ec.europa.eu/clima/eu-action/european-green-deal/european-climate-law_en
[2] https://www.c40.org/about-c40/
[3] https://www.globalcovenantofmayors.org/who-we-are/

large industries (Oda et al., 2015, Olivier et al., 2015). Both methods (inventories and spatial disaggregation) are prone to large uncertainties, impairing the ability to verify the effectiveness of emission reduction policies over time. Common protocols, recent and detailed local activity data, and independent verification remain essential to improve current fossil fuel CO_2 estimates (Mueller et al., 2021). More recently, atmospheric methods (inverse modeling) have been developed and applied to monitor CO_2 emissions from megacities, e.g. Paris (Bréon et al., 2015; Staufer et al., 2016), Indianapolis (Lauvaux et al., 2016; 2020), Cape Town (Nickless et al., 2018), Washington D.C. and Los Angeles (Feng et al., 2016; Yadav et al., 2021). However, atmospheric approaches require modeling systems affected by transport errors, especially over complex terrain and near coastal areas, where the local dynamics vary rapidly at sub-kilometric scales (Ye et al., 2020). Atmospheric inversion systems also rely on the quality of the emissions inventories, often delayed by several years, increasing the uncertainties in fast-growing metropolitan areas. The assessment of atmospheric models remains key to producing reliable and accurate estimates, taking advantage of existing meteorological measurement networks at the ground and collected during aircraft campaigns (Feng et al., 2016; Deng et al., 2017; Lian et al., 2018; Lopez-Coto et al., 2020a).

The Mexico City Metropolitan Area (MCMA), or Greater Mexico City (incl. La Ciudad de México, CDMX) is one of the largest megacities in the world. As the capital city of Mexico, its population of about 22 million inhabitants (UN, 2018) ranks fifth in the world with a population growth rate of about 0.9% per year. Motor vehicle ownership has grown by 30% from 2014 to 2020 (INEGI Mexico, 2021), leading to a 25% increase in CO_2 emissions. The Secretaría del Medio Ambiente de la Ciudad de México (SEDEMA, https://www.sedema.cdmx.gob.mx/) updates an inventory of pollutant emissions, including CO_2, every 2 years, with a delay of about 2 years. The latest reported inventory of GHG emissions of Mexico City is available for the year 2018 (SEDEMA, 2021). The total CO_2 emissions for 2018 amount to 66 Mt CO_2 in the MCMA, with about 85% from the transport and industry sectors. This represents a 50% increase in estimated CO_2 emissions compared to the inventory for 2014 (SEDEMA, 2016), mainly due to a 60% increase

in the mobile sources sector. However, methodological changes and additional activity data explain the strong increase in only four years. In 2015, Mexico committed to reduce by 22% its greenhouse gas emissions by 2030 and to achieve carbon neutrality in Mexico City by 2050 as stated in its Energy Transition Law.

The French-Mexican project "Mexico City's Regional Carbon Impacts (MERCI-CO_2)", funded by the French *Agence Nationale de la Recherche* (ANR) and the Mexican *Consejo Nacional de Ciencia y Tecnología* (CONACYT), enabled the strengthening of measuring atmospheric CO_2 concentrations over the MCMA by the deployment of additional instrumentation. The project was successful to incorporate three types of CO_2 measurements in and around Mexico City. First, two high-precision spectrometers (Picarro) (Crosson et al., 2008; Rella et al., 2013), collect hourly CO_2, CH_4, and CO concentration measurements (Gonzalez del Castillo et al., 2022) calibrated on international measurement scales. Second, about eight CO_2 sensors of intermediate cost and precision (Mid-cost medium precision sensors, MCMP) are to be deployed around the area. Those mid–cost instruments are based on the development done by Arzoumanian et al. (2019). The periods of confinement due to COVID-19 in Mexico City have delayed the deployment of these instruments, initially planned for 2020. And finally, two FTIR measurements of total columns of CO_2, CH_4, and CO are being collected, one (model IFS-120/5HR, Bruker) from the mountain tops (Altzomoni station, ALZ; 19.12°N, 98.66°W) and a second (EM27/SUN) downtown (UNAM university campus, UNA; 19.33°N, 99.18°W). The objective of these surface and column measurements is to characterize the atmospheric CO_2 enhancements caused by anthropogenic emissions from the MCMA. To quantify the physical relationship between surface emissions from urban sources and atmospheric concentrations measured by the various instruments, an atmospheric transport model is required. This atmospheric approach comes up against several obstacles and requires, on the one hand, sufficient measurement density to characterize the evolution of concentrations as a function of exposure to emissions; and on the other hand an atmospheric circulation model capable of representing horizontal and vertical mixing in an urban environment (Feng et al., 2016). The application of atmospheric top-down inversion methods to

CO_2 on a regional scale must also face specific difficulties, such as the strong overlap of atmospheric signals related to anthropogenic and biospheric emissions; as well as atmospheric signals related to local/regional emissions from those on a larger spatial scale (Schuh et al., 2021).

As part of MERCI-CO_2, an atmospheric CO_2 modeling system was configured over the MCMA and its surroundings. The location of Mexico-City spreading through a wide valley surrounded by mountains can prove to be advantageous in favoring the accumulation of CO_2 in the planetary boundary layer (PBL) and therefore generating relatively high atmospheric gradients, more easily detectable by the analyzers. On the other hand, the complex topography around the basin represents a major difficulty for the simulation of atmospheric transport. In this study, we present the meteorological evaluation of atmospheric CO_2 simulations using the Weather Research and Forecasting model (WRF, Skamarock et al., 2008) adapted to the regional scale, and the results of a simulation of the atmospheric concentrations of CO_2 (WRF-Chem, Grell et al., 2005). The reliability of the model was tested based on wind speed, wind direction, PBL heights and temperature monitored at several locations disseminated in and around the city. We then compared the simulated surface and column CO_2 concentrations to the available observations in Mexico-City.

2.2 METHOD

2.2.1 Atmospheric modeling system

The atmospheric simulations were firstly conducted with WRF version 4.0 (Skamarock et al., 2008, Grell et al., 2005). Sensitivity tests were made on various urban canopy models, meteorological drivers, PBL schemes, and nudging parameters before we coupled the chemistry module to WRF.

Table 2.1 shows all the tests made on WRF configurations. The atmospheric turbulence in the PBL for WRF was simulated using the eta (η) operational Mellor-Yamada-Janjic (MYJ, Xu et al., 2015) scheme with local vertical mixing, while the UNAM runs the same model using

Mellor-Yamada Nakanishi and Niino (MYNN, Nakanishi and Niino, 2009) and Yonsei University scheme (YSU, Hu et al., 2010; Fekih et al., 2017). The Run 3 series was coupled with the single-layer urban canopy model (UCM, Tewari et al., 2007; Chen et al., 2011).

The first two simulations (Run 1 & 2) were performed for the period March 16-20, 2016 and the other three were January 4-18 2018, the same as the two rounds of UNAM simulations. Run 3 xn (no nudging) and Run 3 xs (grid analysis nudging) are compared to figure out how the nudging options work. The nesting domain's lateral boundary conditions are initially processed by WRF preprocessing system (WPS) and then obtained from the parent domain. Grid nudging processes the other times and assimilates them with the conditions obtained from parent domains.

No.	Dates	Run time	Test object
Run 1	20160316 - 20160317	5h	PBL scheme MYJ compared to YSU
Run 2	20160316 - 20160320	26h	Id. with eta levels
Run 3 xn	20180104 - 20180118	6d	reference run for Jan, 2018
Run 3 xs	20180104 - 20180118	6d	grid nudging
Test UCM	20180104 - 20180118	6d	without urban canopy model (UCM)
Test ERA-5	20180104 - 20180118	6d	meteorological drivers

Table 2.1 *Sensitivity tests on WRF*

WRF was configured with two-way triple-nested domains for the Valley of Mexico City, with the spatial resolutions of 15, 5 and 1 km (**Figure 2.1**). The largest domain (domain 1, d01) covers almost the entire

territory of Mexico and several neighboring countries; the middle domain (domain 2, d02) is one-third of the length and width of domain 1 and domain 3 is one-third the length and width of domain 2. The smallest domain (domain 3, d03) covers the MCMA as well as several surrounding states, in order to include the whole valley and to avoid the instability when domain boundaries cut the mountains. Domain 3 is the same size and location as the UNAM run, for the sake of comparison to the UNAM simulations with the MYNN and YSU boundary layer scheme.

There are 49 vertical layers in each domain, the first two of which are around 16 and 45 meters above ground level. Due to the complex terrain of Mexico City's valley, the model would be very unstable with finer vertical layers near the ground.

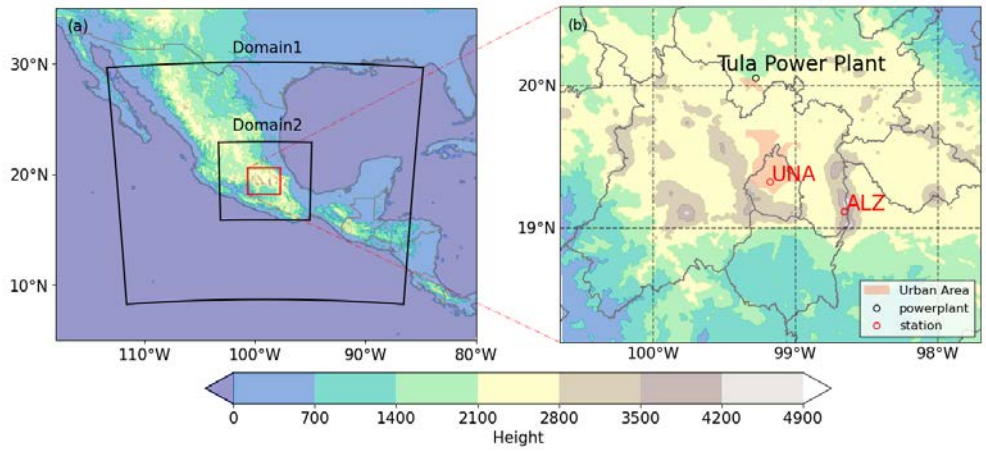


Figure 2.1 *Simulation domains (a) of the three nested domainsat 15-km (Domain 1, d01), 5-km (Domain 2, d02) and 1-km (Domain 3, d03, red square) resolutions with their corresponding terrain heights (in meters) and (b) Domain 3 showing Mexico City and its surroundings including the CO_2 observation stations for surface and total columns measurements (red circles), a large power plant nearby (black circle), the urban area of MCMA (in pink), the administrative boundaries of Mexico City (in gray) and the model topography in meters.*

The meteorological drivers used as boundary and initial conditions are the 6-hourly ECMWF Reanalysis interim (ERA-interim, Dee et al., 2011) and version 5 (ERA-5, Hersbach et al., 2017), at 0.25° × 0.25° resolution.

The North American Mesoscale Model (NAM), which is used in the simulations run by UNAM, was tested as well, but was excluded because our Domain 1 exceeds the boundary of NAM.

Due to the lack of available high-quality meteorological observations in and around Mexico City, the meteorological simulations were performed without data assimilation. However, several surface stations, not referenced at the World Meteorological Organization (WMO), were used to evaluate the model performances. A summary of the configurations and parameters are shown in **Table 2.2**, left column.

	WRF	WRF-Chem
model version	4.0	3.9
simulation dates	2018/1/4 00:00:00 - 2018/1/18 00:00:00	
spatial resolution	15km, 5km, 1km	15km, 5km(, 1km)
meteorological drivers	ERA-interim (ERA-5)	ERA-interim (ERA-5)
PBL scheme	MYJ (YSU, MYNN by UNAM)	MYNN (YSU)
urban canopy	UCM	UCM (BEP)
anthropogenic emission	/	ODIAC, ODIAC (UNAM), UNAM
biogenic fluxes	/	CASA
boundary conditions	/	300 ppm (CarbonTracker 2019B)
nudging	/ (grid nudging)	/ (grid nudging)

Table 2.2 *Summary of the configurations and parameters for the reference run and sensitivity test (shown in brackets) for WRF and WRF-Chem*

After the meteorological evaluation of WRF version 4.0, we coupled

the Chemistry module to WRF (**Table 2.2**) and continued the tests on CO_2 simulations (**Table 2.3**). The CO_2 simulations were conducted with WRF-Chem version 3.9, with 2 tracers for anthropogenic CO_2 and CO_2 biogenic fluxes, separately. The anthropogenic emissions of domain 1 and domain 2 are interpolated from the global CO_2 emission product, the Open-source Data Inventory for Anthropogenic CO_2 (ODIAC) (Oda et al., 2018) and those of domain 3 are interpolated from the local emission inventory provided by the National Autonomous University of Mexico (UNAM). The biofluxes are interpolated from the Carnegie Ames Stanford Approach (CASA) model (Zhou et al., 2020). Interpolations were made by a mass-conserving python package (xESMF - conservative method, https://doi.org/10.5281/zenodo.1134365). In the test C-bdy, the global model CarbonTracker 2019 (Peters et al., 2007; Jacobson et al., 2020) was applied for optimized boundary conditions. The details of the emission inventories and boundary products, as well as interpolation details, are described in the next Chapter (section 3.1.2).

No.	Dates	Test object
C0	20180104 - 20180118	reference run of WRF-Chem
C1-BEP	20180104 - 20180118	multi-layer urban canopy BEP
C2-UNAMd2	20180104 - 20180118	change d02 emission to UNAM
C3-nudging	20180104 - 20180118	grid nudging in d01
C4-ERA5	20180104 - 20180118	meteorological drivers ERA-5
C-bdy	20180104 - 20180118	boundary condition from CarbonTracker
C25-d02	20180104 - 20180118	spatial resolution

C-YSU	20180104 - 20180118	meteorological drivers
C345-feedback	20180104 - 20180118	one-way nesting

Table 2.3 *Sensitivity tests on WRF-Chem*

In the WRF-Chem modeling system, we switched to the MYJ PBL scheme for the potential future application of Turbulent Kinetic Energy (TKE), and tested YSU for the optimization on the wind speed and direction in mountainous areas. The drivers (C4-ERA5) and grid nudging (C3-nudging) are re-tested under MYJ. The two-way nesting doesn't show significant difference without the chemistry module, but it has become imperative in WRF-Chem. C345-feedback compared the two nesting ways. It took around 6 days to run a 2-week simulation on the cluster. For higher computational efficiency on long-term simulations, we tested the model performance on the middle 5-km domain, domain 2, to determine whether it is sufficient for inversion (C25-d02). The emission inventory source of domain 2 is then switched to UNAM 3-km national emission inventory.

Another urban canopy model coupled with WRF, the multi-layer Building Environment Parameterization (BEP) (Martilli et al., 2002) was tested in WRF-Chem (test C1-BEP). Lian et al. (2019) found that BEP could significantly reduce the Root Mean Square Error (RMSE) and Mean Error (ME) of the CO_2 concentrations from a one-year simulation compared to 2 GreenLITE^{TM} transceivers and 4 in-situ stations in Paris. BEP uses a series of parameters to describe the characteristics of the urban area, including street directions, road widths, and building width, height and percentage. Due to lack of study on local building, these parameters remain default.

2.2.2 Meteorological observations

2.2.2.1 Meteorological surface network

We evaluated the WRF meteorological performances using measurements from the air quality network installed and operated by SEDEMA. Horizontal wind speed and direction, temperature, pressure,

and relative humidity were measured across 26 surface stations located in and around Mexico City, publicly available from SEDEMA[4]. The 26 stations are listed in **Table 2.4**, including the station's names, their geographic coordinates, and the data availability for the year 2018. In previous reports, we found specific comments on possible disturbances affecting wind measurements at three sites. At Merced (MER), located near downtown Mexico City, a three-story school located to the south of the station blocks the air flow from that direction, in addition to an elevated metro railway to the west. At Xalostoc (XAL), the inlet location is partially blocked by a building on one side of the shelter. Finally, at Cuajimalpa (CUA), the site is located on a two-story building with clear air flow in three out of four directions. There is a line of trees on one side of the shelter that may slightly impact wind direction and wind speed measurements, but the meteorological station is at the same altitude or slightly higher than the treeline. In conclusion, most of the SEDEMA meteorological observations match international standards thanks to rigorous data quality control.

ID	Station	Entity	Latitude	Longitude	Altitude	Sampling height/m	Since	% data in 2018
ACO	Acolman	Mexico State	19.6355	-98.9120	2198	13.3	2011	84.35%
AJM	Ajusco Medio	CDMX	19.2722	-99.2077	2548	15	2015	96.21%
AJU	Ajusco	CDMX	19.1543	-99.1626	2942	4	2015	68.29%
BJU	Benito Juarez	CDMX	19.3717	-99.1591	2257	5.5	2015	82.19%
CHO	Chalco	Mexico State	19.2670	-98.8861	2253	7.3	2011	87.03%
CUA	Cuajimalpa	CDMX	19.3653	-99.2917	2704	10.6	2000	98.05%
CUT	Cuautitlán	Mexico State	19.7222	-99.1986	2263	4.5	2012	64.93%
FAC	FES Acatlán	Mexico State	19.4825	-99.2435	2299	10.8	1986	61.37%
GAM	Gustavo A. Madero	CDMX	19.4828	-99.0947	2242	6	2015	98.86%
HGM	Hospital General de México	CDMX	19.4116	-99.1522	2234	15	2012	86.94%
INN	Investigaciones Nucleares	Mexico State	19.2920	-99.3805	3080	4	2015	3.84%

[4] http://www.aire.cdmx.gob.mx/default.php?opc=%27aKBhnmI=%27&opcion=Zw==

LAA	Lab. de Análisis Ambiental	CDMX	19.4838	-99.1473	2255	5.1	2016	97.59%
MER	Merced	CDMX	19.4246	-99.1196	2245	17.2	1986	98.85%
MGH	Miguel Hidalgo	CDMX	19.4041	-99.2027	2327	11.5	2015	99.77%
MON	Montecillo	Mexico State	19.4604	-98.9029	2252	3.5	2000	97.24%
MPA	Milpa Alta	CDMX	19.1769	-98.9902	2592	8	2016	72.74%
NEZ	Nezahualcóyotl	Mexico State	19.3937	-99.0282	2235	12.1	2011	99.47%
PED	Pedregal	CDMX	19.3252	-99.2041	2326	11.5	1986	96.05%
SAG	San Agustín	Mexico State	19.5330	-99.0303	2241	8.6	1986	78.31%
SFE	Santa Fe	CDMX	19.3574	-99.2629	2599	5.5	2012	84.87%
TAH	Tláhuac	CDMX	19.2465	-99.0106	2297	3.6	2000	75.79%
TLA	Tlalnepantla	Mexico State	19.5291	-99.2046	2311	7	1986	92.36%
UAX	UAM Xochimilco	CDMX	19.3044	-99.1036	2246	18.7	2015	99.71%
UIZ	UAM Iztapalapa	CDMX	19.3608	-99.0739	2221	19.6	2014	64.29%
VIF	Villa de las Flores	Mexico State	19.6582	-99.0966	2242	4.5	2000	86.16%
XAL	Xalostoc	Mexico State	19.5260	-99.0824	2160	30	1986	96.72%

Table 2.4 *Coordinates, heights, first year of data acquisition, and percentage of valid data in 2018 of the SEDEMA meteorological stations in and around Mexico City.*

2.2.2.2 Lidar profiling

The Doppler Lidar located at UNA (Wind Cube 100, Leosphere) was used for wind profile and mixing- and residual-layer height observations. The instruments measure with 50 m vertical resolution backscattered light from a pulsed laser in the infrared that is transmitted to the sky at four different directions and Doppler-shifted, 15º from the zenith, from which a 3-dimensional wind field is reconstructed. Mixing and residual-layer heights calculated from the aerosol backscattering signal (carrier-to-noise ratio) and the variance of the vertical velocity are reported by the manufacturer. Since January 2017, data is continuously collected by this instrument and stored in the RUOA repository (Red Universitaria de Observatorios Atmosféricos, http://www.ruoa.unam.mx) and the results have been evaluated against other data sources (Burgos-Cuevas et al 2022).

As supplementary information, PBL heights measured by Radiosonde in international airports are also included in the comparison. The PBL heights from Radiosonde were obtained in the Mexico City international airport (19.43°N, 99.08°W), WMO station 76679 (http://weather.uwyo.edu/upperair/sounding/html), and were computed from potential temperature (Θ) gradient: For each scan, PBL

height was determined by identifying the point in the Theta inversion where the temperature is 2 K warmer than at the base of the inversion and the Theta gradient is greater than 0.005 K / m. Most of the Radiosonde data were collected at 12:00 UTC, which corresponds to 6:00 local time in winter and 7:00 local time in summer. The PBL heights inferred from the radiosonde data were reliable in dry seasons but not in rainy seasons, as clouds significantly affected the temperature vertical profile and led to unreasonable inferred PBL heights that could reach thousands of meters. Therefore, I have excluded the inferred PBL heights more than 2000 meters in the present analysis.

2.3 RESULT: METEOROLOGICAL PERFORMANCES

2.3.1 Sensitivity tests

2.3.1.1 Meteorological drivers

I have made systematic comparisons of the observations with the meteorological dataset used as the drivers of the WRF simulations (met_em files generated by WPS, the pre-processing system of WRF). Three datasets have been investigated: ERA-interim, ERA-5 and NAM. ACO, an observation station from SEDEMA measurement network, is taken as an example of the comparisons. As **Figure 2.2** shows, NAM seems to fit the wind speed observations the most, with an R square of 0.503 (that of ERA-interim and ERA-5 are 0.15 and 0.27 respectively). During January 5-6, ERA-interim significantly underestimated the wind speed at ACO station in the city center. ERA-interim also failed to capture the peak on January 13 in the evening (UTC time January 14 00:00). However, the area which NAM covers is slightly smaller than the domain 1 defined for MCMA, and for this reason I could not use NAM in the simulations.

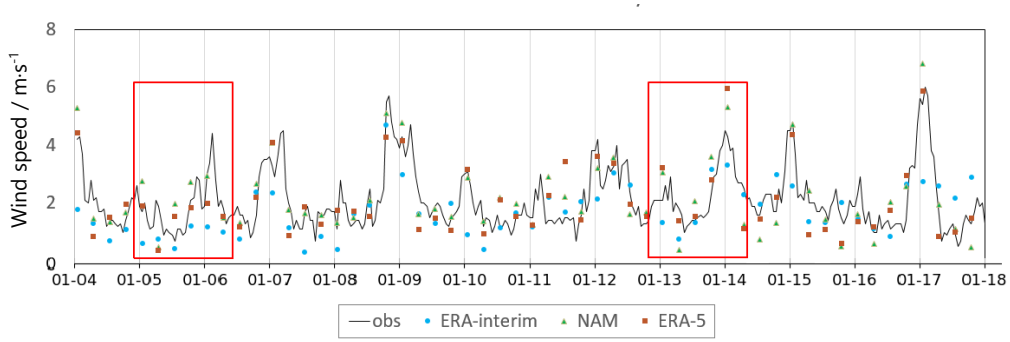


Figure 2.2 *Wind speed time series (in m/s) from hourly observation and the interpolated 6-hour WRF input file from three meteorological products: ERA-interim, ERA-5 and NAM, at the observation station ACO in the city center of MCMA. (Two periods with abnormal values of ERA-interim are marked with red square)*

The tests on meteorological drivers ERA-interim and ERA-5 are implemented on both WRF version 4.0 and WRF-Chem version 3.9, separately. As a more mature product, ERA-5 significantly improved the performance of the model, especially on the 2-meter a.g.l. temperature. On average over the 26 meteorological stations, the mean error of 2-week hourly temperature has been reduced from -4.25 ℃ to -0.94℃ (**Figure 2.3**).

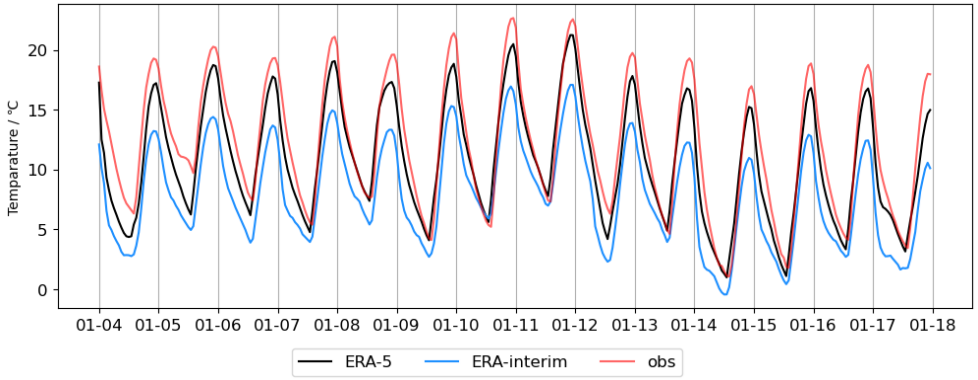


Figure 2.3 *Temperature time series of 26-station average from hourly observation by SEDEMA and the simulation driven by three meteorological products: ERA-interim, ERA-5 and NAM.*

2.3.1.2 PBL schemes

I have also compared the vertical wind measurements collected by the Lidar deployed at UNA (South city-side) to the simulations. The scanning elevation of the wind reconstruction covers a 15-degree angle from the zenith (75° to 90°; every scan contains 4 directions). The average wind speed from the 4 directions is used for the model evaluation of the wind speed. Two atmospheric layer heights are diagnosed from raw Lidar data: a mixing layer height and a residual layer height. As shown in **Figure 2.4**, mixing layer heights (red dots) diagnosed by the Lidar correspond to the top of the mixed layers but fail to capture the maximum elevations deduced from the residual heights (red line) during the afternoons. Conversely, the residual layer

heights correspond to the top of the mixed layer during the afternoon, while measuring the actual residual layer at night. To avoid a negative bias in Lidar measurements during daytime, we combined the two products to derive the most representative height of the mixed layer during daytime. The mixing layer diagnostic is used in the morning (local time 8:00-13:00) while residual layer heights are used to evaluate the mixed layer height in the afternoon (local time 14:00-18:00).

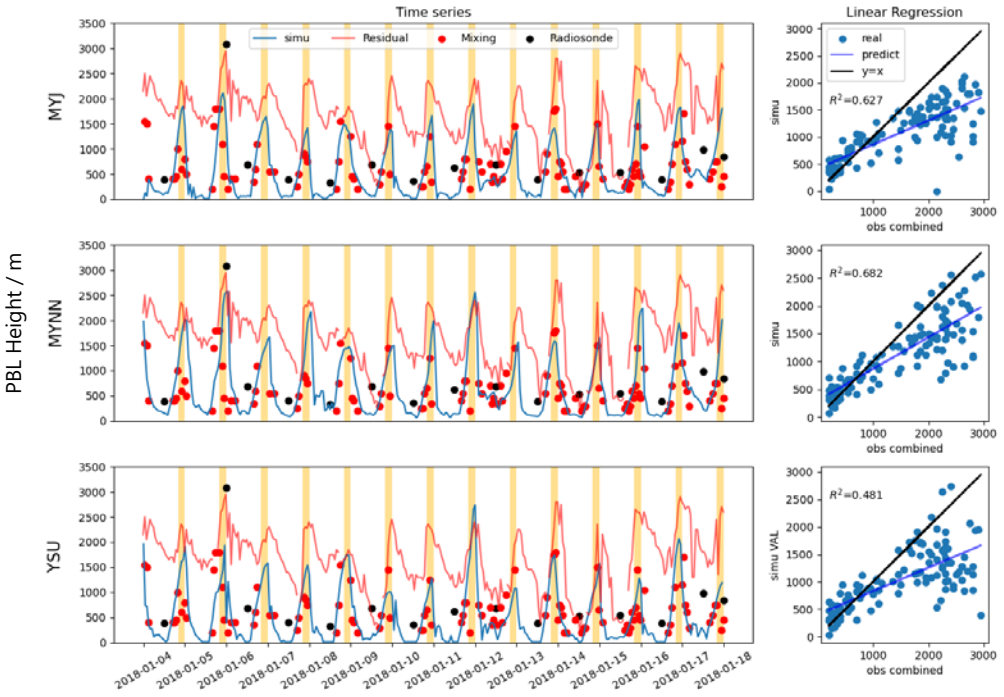


Figure 2.4 *Time series and linear regression of PBL height (in meters) simulation (in blue) and observation ("residual layer" heights in red lines and mixing layer heights in red dots) in UNA (time in UTC, golden background stands for local time UTC-6 afternoon: 14:00-18:00), for the three PBL schemes MYJ, MYNN and YSU.*

As for the model, the approximate height above the surface was calculated using the geopotential height:

$$Height = (PH + PHB)/g - HGT$$

in which *PH*, *PHB*, *HGT* and g stand for perturbation geopotential height, base-state geopotential height, terrain height and gravitational acceleration, respectively.

Three PBL schemes were tested in our modeling system: MYJ (with

WRF version 4.0), MYNN and YSU (with WRF-Chem version 3.9). **Figure 2.4** shows the time series and linear regressions for the simulations and observations. MYJ and MYNN are more similar, but MYJ didn't capture some of the high developments of the PBL. For example, MYNN catched the sharp decrease in the evening of January 3 (UTC time 2018-01-04 00:00:00 - 03:00:00), as well as the rapid ascent in the afternoon of January 5 and 7 (UTC time 2018-01-05 19:00:00 - 2018-01-06 00:00:00 and 2018-01-07 19:00:00 - 2018-01-08 00:00:00). The better performance of MYNN is shown in the regression plots. Linear fit of MYNN is closer to the line y=x, with a higher R square (0.68 vs 0.63), and MYJ often underestimates the PBL heights in the morning. As for YSU, it tends to underestimate the nighttime PBL heights even more. In 6 of the 14 days, the simulated PBL heights, using YSU scheme, remain close to zero for several hours. All of the three schemes are severely underestimating the afternoon PBL heights. The daytime mean errors of PBL heights are -425 m for MYJ, -465m for YSU, and -354 m for MYNN.

Regarding the wind speeds, the comparison of the three PBL schemes clearly shows that YSU reduced the overestimation of wind speeds in the mountains. **Table 2.5** shows the errors in different PBL schemes. The mean error is reduced by about 75%. On the other hand, the MAE wasn't reduced. The "topo_wind" option of YSU usually decreases the wind speed enabling a better comparison to the observations, but in some cases it may drive to extremely high wind speed in complicated topography (see for example the INN station on **Figure 2.5**).

	MYJ	MYNN	YSU
Mean error	0.41	0.42	0.09
Mean absolute error	0.87	0.98	0.99

Table 2.5 *Mean Errors and Mean Absolute Errors on wind speed of the three PBL schemes: MYJ, MYNN and YSU.*

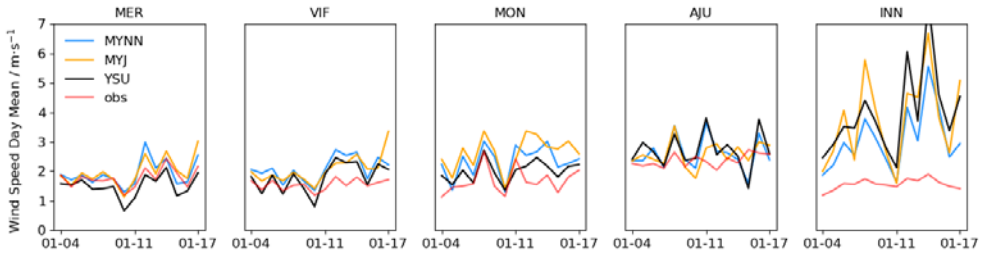


Figure 2.5 *Wind speed daily mean observations and simulations under three PBL schemes in January 2018 at 5 of the 26 SEDEMA stations: MER in the city center, VIF in the north, MON in the east, AJU in the south (mountains), and INN in the west (mountains), in UTC time (local time=UTC-6).*

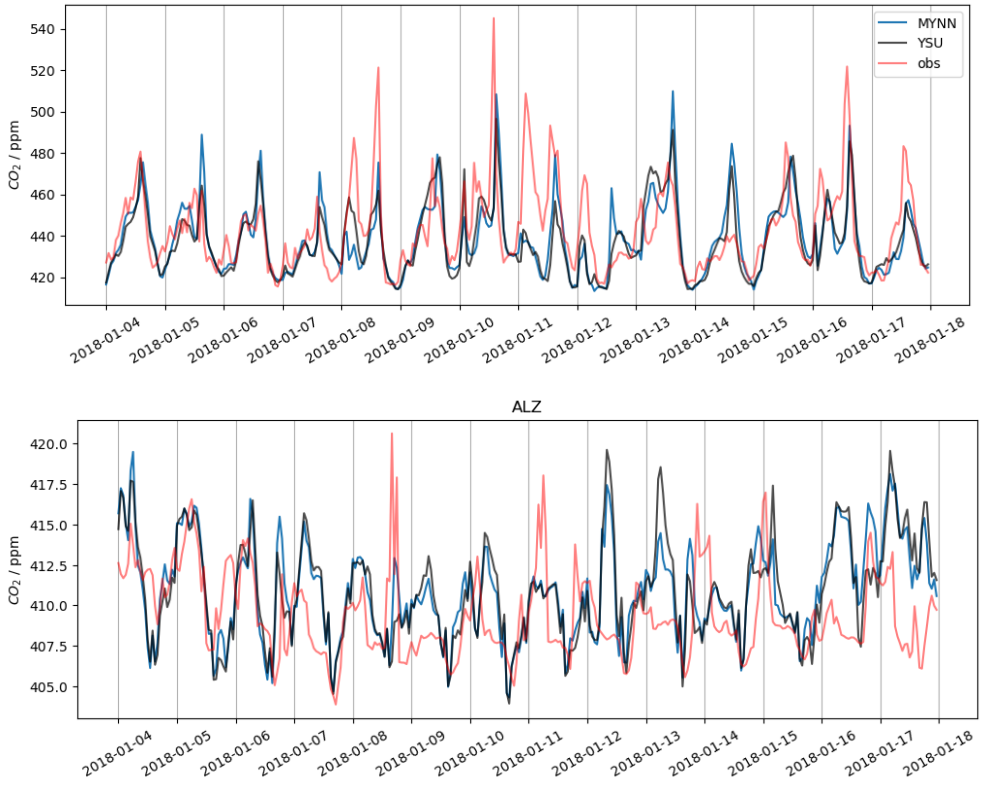


Figure 2.6 *Time series of CO_2 mixing ratio observed (red lines) at UNA (above) and ALZ (below) stations in the 2-week simulation in January, 2018, with PBL scheme MYNN (blue lines) and YSU (black lines).*

I have also also compared the CO_2 mixing ratios to the WRF-Chem simulations performed with two different PBL schemes (MYJ hasn't been tested in WRF-Chem). **Figure 2.6** shows that the overestimation of peaks (in January 5, 7, 12, 13) simulated with MYNN scheme could be fixed with YSU, probably thanks to its better performance in the

simulation of the wind. On the contrary, there are some peaks captured by MYNN but missed by YSU, on January 8 and 11. The CO_2 variations at ALZ are much smaller than that of UNA and no significant difference was found between the two simulations under the two schemes. Overall, despite the fact that YSU improved the wind speed in most cases it doesn't show obvious improvement in the CO_2 simulation.

2.3.1.3 Urban canopy models

Two urban canopy models coupled with WRF were tested: the single-layer UCM and the multi-layer BEP. With default parameters, the BEP module was not as stable as UCM and the model coupled with BEP sometimes crashes. Despite that, BEP improves the performance of the model in the simulation of the PBL heights, temperatures and CO_2 concentrations. The mean error of PBL heights at UNA in the afternoon simulated with BEP has decreased by 39%, from -588 m to -359 m.

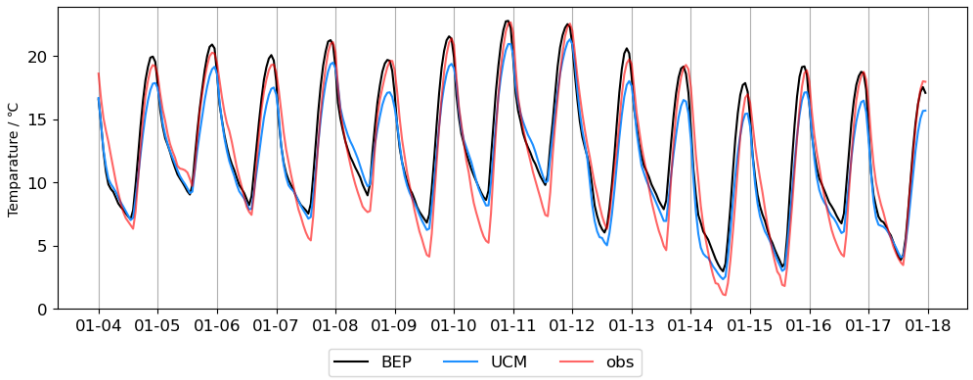


Figure 2.7 *Time series of the temperature averaged for the 26 SEDEMA stations. Observations are shown in red, and the simulations coupled with single-layer urban canopy model UCM and multi-layer urban canopy model BEP are shown in blue and black respectively.*

BEP has optimized the temperature of the 26 stations in MCMA. In particular, the underestimation in the afternoon is largely reduced with BEP compared to UCM. On the other hand it still overestimates the minimum temperature observed in the morning (shown in **Figure 2.7**). The mean error increased from -0.34℃ (UCM) to 0.65℃ (BEP). The mean error and mean absolute error (MAE) maps in **Figure 2.8** show that the mean temperatures are increased for almost all SEDEMA stations. For the stations in the city, they are underestimated with UCM and slightly overestimated with BEP. For several stations in the rural

area, e.g. AJU and CUT in the mountains, we observe a higher overestimation of the temperature. In total, the fit between observations and simulations didn't change a lot, with an R square of 0.83.

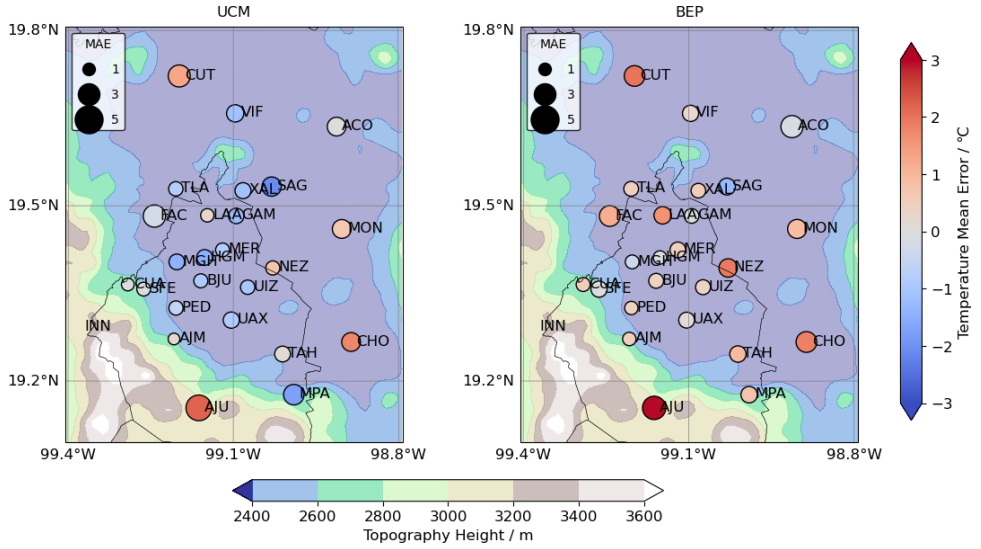


Figure 2.8 *The mean error and mean absolute error (MAE) map of the 26 SEDEMA stations, in the 2-week simulation in January, coupled with urban canopy models UCM (left) and BEP (right).*

As for the simulation of the wind, using BEP with the default parameters, we do not observe a systematic improvement of the wind speed at all SEDEMA stations. As **Figure 2.9** shows, in the city station MER and rural station AJU in the southern mountains, BEP even degrades the performance of the model. However, it reduces the large bias when the wind speed is high in the mountain, e.g. at INN. For the average of all the 26 stations, the mean error decreased from 0.42 m/s to -0.10 m/s with BEP, and the mean absolute errors decreased from 0.98 m/s to 0.87 m/s.

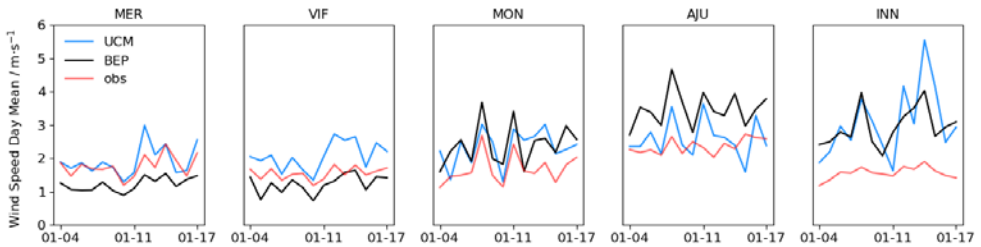


Figure 2.9 *Wind speed daily mean in January 2018 at 5 of the 26*

SEDEMA stations: MER in the city center, VIF in the north, MON in the east, AJU in the south (mountains), and INN in the west (mountains), in UTC time (local time=UTC-6), coupled with urban canopy model UCM (in blue) and BEP (in black).

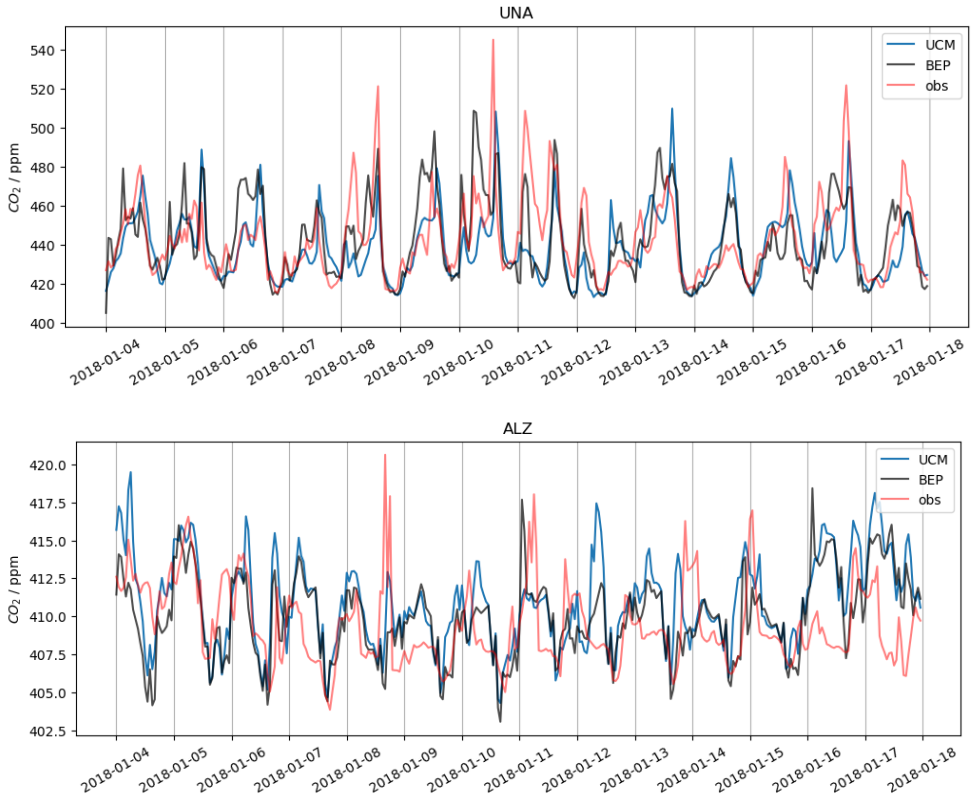


Figure 2.10 *Time series of CO_2 mixing ratio at UNA and ALZ station in the 2-week simulation in January, 2018, coupled with urban canopy models UCM (in blue) and BEP (in black).*

In CO_2 simulations, BEP (with default parameters) is prone to generate CO_2 concentration peaks in the urban area. It captured relatively well several peaks observed at UNA in the city center, that UCM missed on January 8, 11 and 12, but also created some peaks that didn't exist, e.g. on January 6, 13 and 17, shown in **Figure 2.10** (UNA, upper panel). Overall the mean error has decreased from -5.6 ppm to -1.1 ppm with BEP, but the mean absolute error keeps 12.8 ppm, and the mean absolute error in the afternoon even increased by 0.5 ppm.

2.3.1.4 Nudging

The fdda option nudges the WRF towards a gridded analysis. It is

supposed to nudge wind, temperature and water vapor. Grid nudging in the first domain didn't make a significant difference in the model. The mean errors of PBL heights without and with nudging are -87 m and -99 m in the morning (local time 7:00-12:00), and -971 m and -986 m in the afternoon (local time 13:00-18:00). In departure from earlier studies (Lian et al., 2018), grid nudging sometimes worsen the model performance. As shown in **Figure 2.11**, the model captured the peaks of PBL height in the afternoons of January 4, 15 and 17 without nudging, but missed them with nudging.

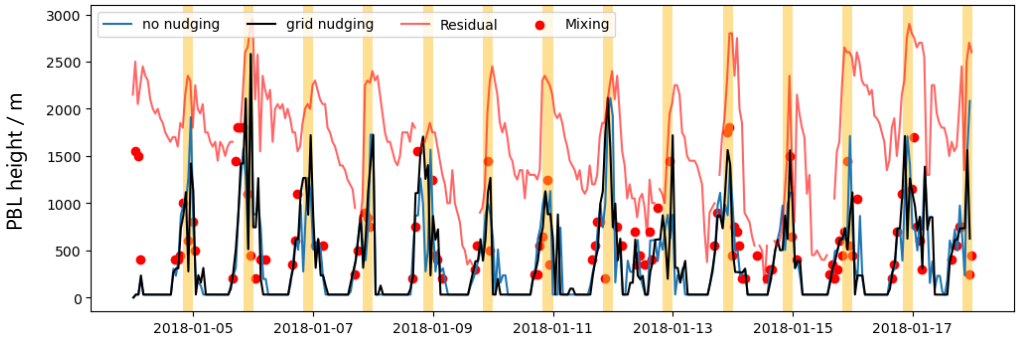


Figure 2.11 *Time series of PBL height (in meters) simulation with and without nudging as well as Lidar observation ("residual layer" heights in red lines and mixing layer heights in red dots) at UNA (time in UTC, golden background stands for local time UTC-6 afternoon: 14:00-18:00)*

The deterioration due to grid nudging is more obvious in wind speed. The R2 of the linear regression between simulations and observations has dropped from 0.40 to 0.23, as shown in **Figure 2.12**. The mean errors in the morning and the afternoon have increased by 0.5 - 0.6 m/s and the mean absolute errors have increased by 0.4 - 0.5 m/s. Grid nudging is not going to be applied in the future modeling.

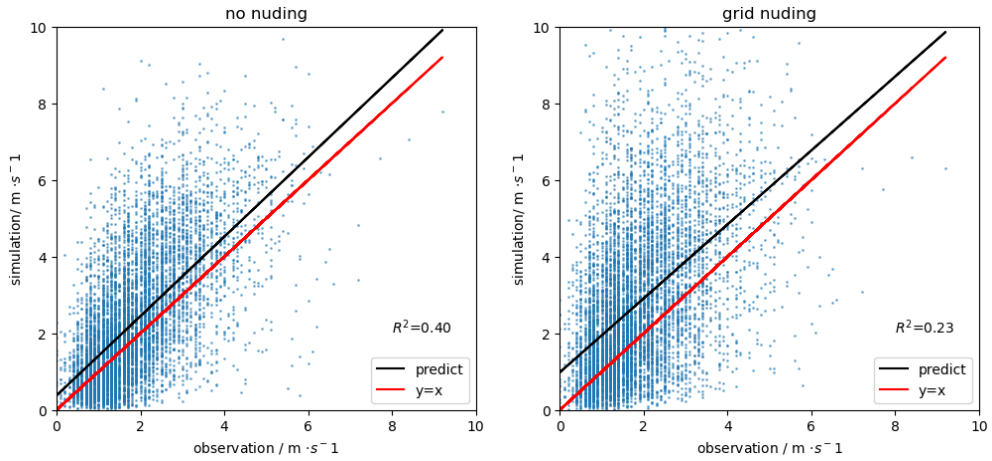


Figure 2.12 *Linear regression of hourly wind speed observation and simulation in the 2-week simulation at the 26 stations of SEDEMA network, in January 2018, without and with grid nudging in domain 1.*

2.3.1.5 Spatial resolution

Feng et al. (2016) found that the high resolution configurations (1.3km) resolved PBL heights and vertical profile of the horizontal winds better than moderate-resolution (4 km) in Los Angeles (LA) basin, while it did not cause a significant difference in the simulation of the surface meteorological parameters. From the comparison I have done over a two-week period, unlike the LA case, the PBL heights at UNA from the 5-km domain d02 and 1-km domain d03 have no significant difference. **Figure 2.13** shows the time series of PBL heights derived from the two domains. Indeed, there are some differences, but all of them are at nighttime. The mean error of PBL heights in the afternoon 14:00-18:00 (compared with the so-called "residual layer heights") of 5-km simulation is -684 m and that of 1-km simulation is -686 m, almost the same with 5-km simulation. The mean absolute error in the morning (7:00-12:00) is 179 m and 178 m.

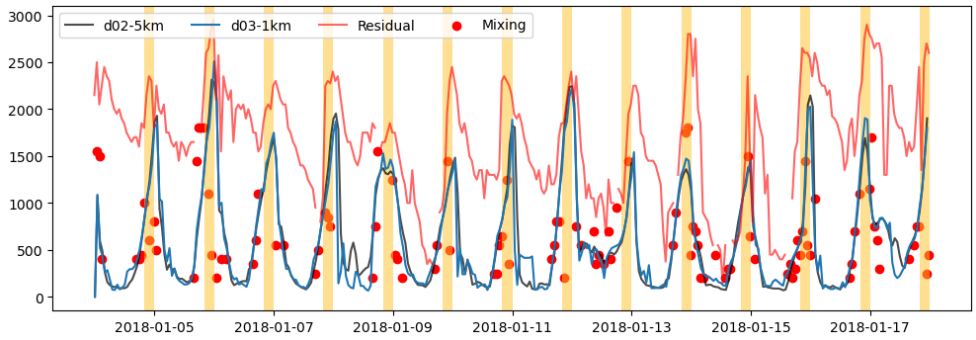


Figure 2.13 *Time series of PBL height (in meters) of 5 km (black line) and 1 km (blue line) simulations as well as Lidar observation ("residual layer" heights in red lines and mixing layer heights in red dots) at UNA (time in UTC, golden background stands for local time UTC-6 afternoon: 14:00-18:00).*

I have also compared the vertical profile of horizontal wind speed to the Lidar Dopller observations at the UNAM site (**Figure 2.14**). The mean error and mean absolute error of the two domains are rather close: mean errors are -0.75 m/s for 1 km and -0.60 m/s for 5 km; MAEs are 1.59 m/s for 1 km and 1.58 m/s for 5 km. The errors for the high resolution are even slightly larger, especially in the lower atmosphere (**Figure 2.14**).

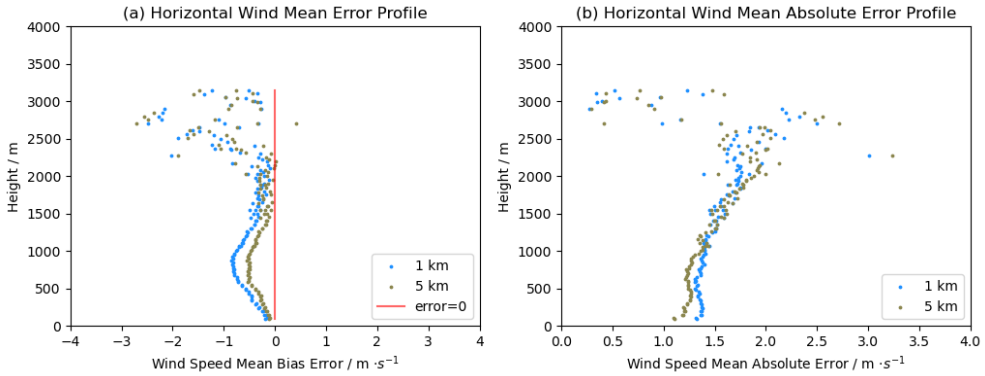


Figure 2.14 *Vertical profiles of 2-week mean horizontal wind speed errors at UNA for two spatial resolutions (1 km in blue and 5 km in green) in the afternoon: (a) mean error profile (b) mean absolute error profile.*

Nevertheless, the moderate resolution (5-km) simulation appears to be not as good as the high resolution (1-km) domain to simulate the CO_2 concentrations. I have compared the CO_2 concentrations simulated using the two resolutions, with the same fossil fuel emissions, biogenic

fluxes and boundary conditions. The 5-km simulation overestimated the gradients between the two stations (UNA-ALZ) by about 9.7%, while the 1-km domain overestimated them by 5.6%. As shown in **Figure 2.15**, the overestimations of the gradients between the two stations are mainly occurring at night time (UTC time 19:00 - 11:00) and in the morning (UTC time 12:00 - 17:00) when the CO_2 concentrations at UNA are higher with the 5-km resolution. For the simulation in the afternoon, when the PBL is stable, the mean absolute errors of the gradients for the two resolutions are similar: 7.6 ppm for the 5-km domain 2 and 7.7 ppm for the 1-km domain 3. The similar MAEs during the afternoon built a sufficient basis for the potential use of a moderate resolution domain in exchange for computational efficiency in future inversions.

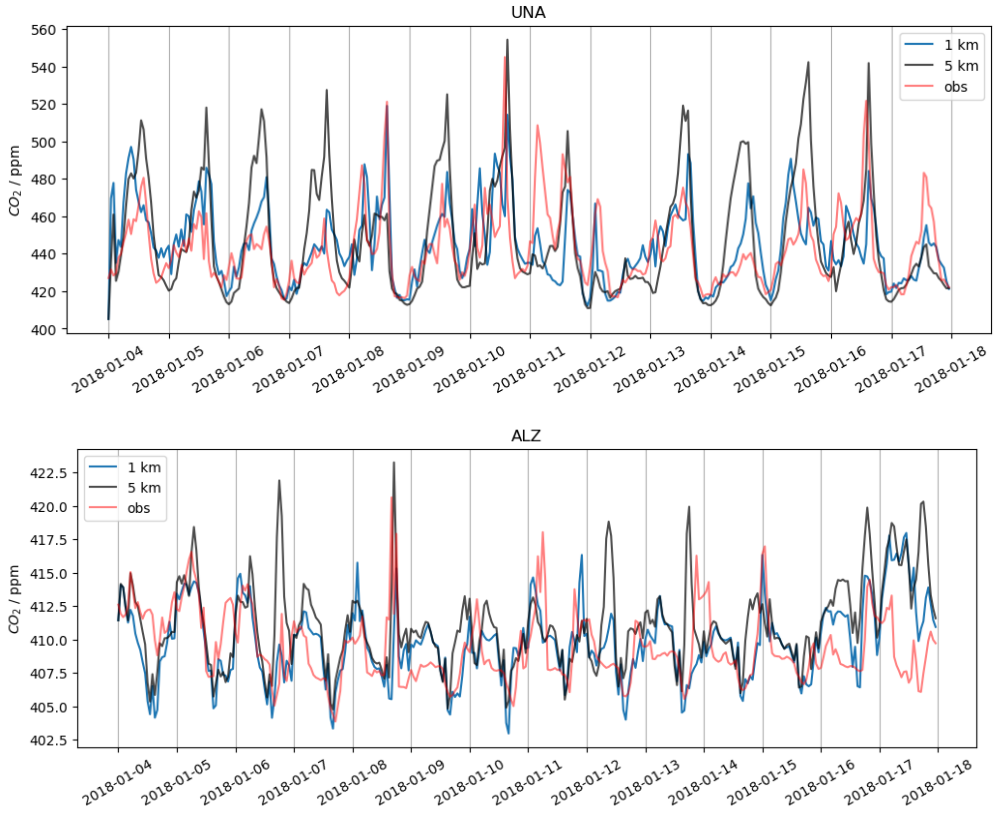


Figure 2.15 *Time series of simulated and observed CO_2 mixing ratio (red) at UNA and ALZ station in the 2-week simulation in January, 2018, in 1-km (blue) and 5-km (black) simulation, in UTC time.*

2.3.1.6 *Boundary conditions and its spinning up*

With the two-way nesting approach, there is an interaction between the two model domains at each hour, and the input concentration map of domain 2 is affecting the smaller domain 1. As **Figure 2.16** (right column) shows, the two-way nesting introduced a bias of the low-value area in the center of domain 1, due to the default input concentration of domain 2 (300 ppm), which is much lower than the background concentrations from CarbonTracker (410 – 420 ppm).

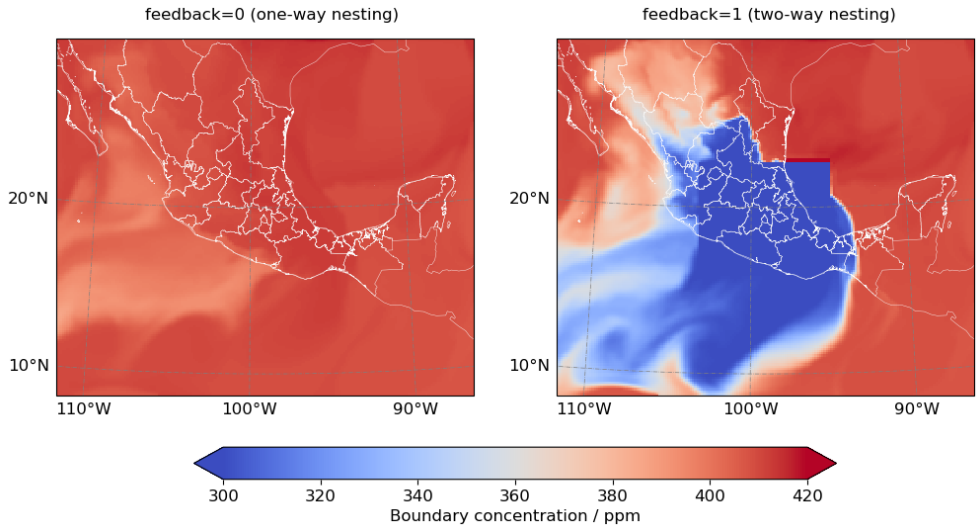


Figure 2.16 *Background CO_2 concentrations from WRF boundary conditions for Domain 1, in one-way and two-way nesting after the 2-week simulation, at 23:00:00 UTC time, 2018-01-08.*

To solve the problem, we have two choices: either we keep the two-way nesting and edit the WRF input file, that is, to interpolate the input concentration map from CarbonTracker every time we run real.exe; or we switch to one-way nesting and reserve a period of time for spinning up. **Figure 2.17** shows the time required for spinning up. The time depends on the large-scale diffusion conditions, e.g. wind speed in the high altitude. For our case in January, it takes about 5 days for CarbonTracker to cover half of the domain and about 10 days to cover the other half. Until the end of the 2 weeks, the domain has totally adapted the CarbonTracker background concentrations. 2 weeks would be enough for the concentration spinning up from the domain boundary.

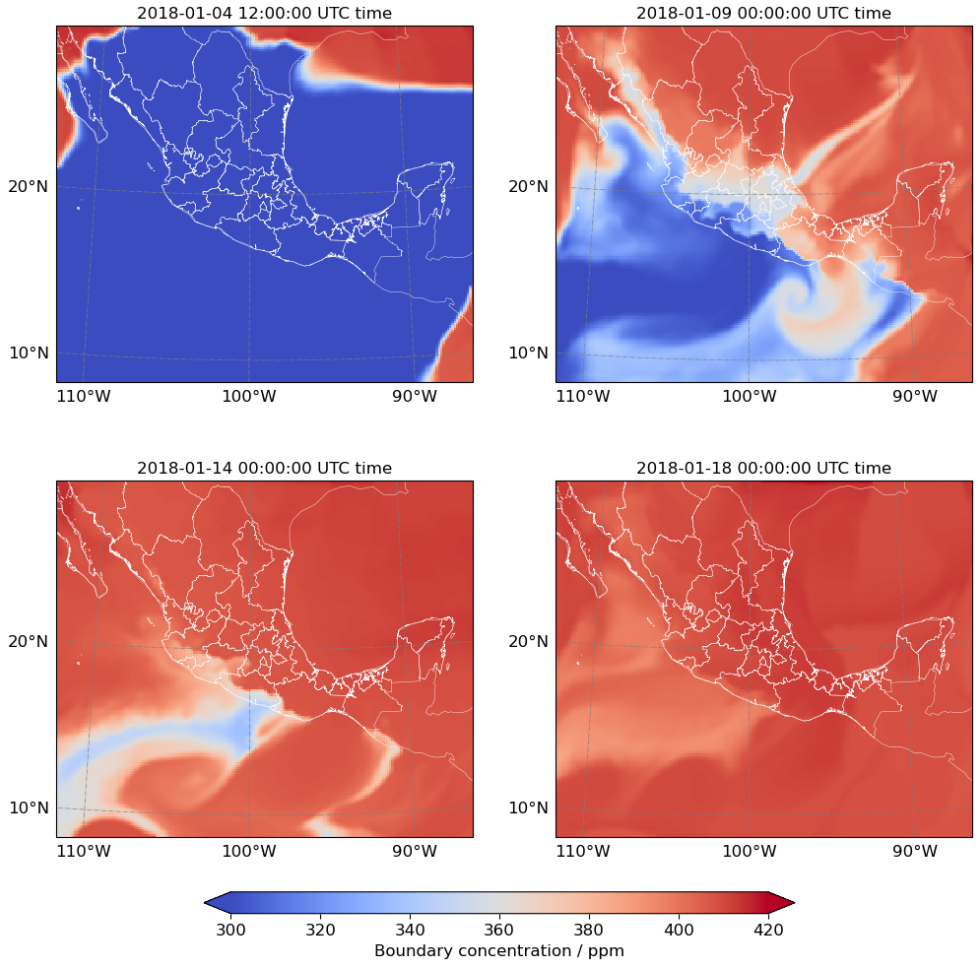


Figure 2.17 *Background CO_2 concentrations from WRF boundary conditions for Domain 1, in one-way nesting, at four different times.*

2.3.2 Surface meteorological evaluation

The sensitivity tests that I carried out made it possible to evaluate the sensitivity of the simulations to several parameters, with the aim of defining the configuration providing the best compromise between results and calculation time. According to the results of the sensitivity tests, I selected the configuration shown in **Table 2.6**, to perform longer simulations at different seasons. According to the tests on boundary conditions, it takes around 10 days for the background concentrations spinning up from domain boundaries. Consequently, I reserved 2 weeks for this spinning up in each round of running. In this section I present the evaluation of the simulations made with this configuration, by comparisons with meteorological observations.

	WRF-Chem
model version	3.9
simulation dates	Jan: 2017/12/21 00:00:00 - 2018/1/18 00:00:00 May: 2018/5/1 00:00:00 - 2018/5/29 00:00:00 July: 2018/7/1 00:00:00 - 2018/7/28:00:00:00
spatial resolution	15km, 5km
meteorological drivers	ERA-5
PBL scheme	MYNN
urban canopy	UCM
anthropogenic emission	ODIAC, UNAM
biogenic fluxes	CASA
boundary conditions	CarbonTracker 2019B
nudging	/

Table 2.6 *Summary of the configurations for WRF-Chem in meteorological evaluations*

Three simulations were performed over January, May, and July 2018, representing three contrasted seasons in the region of Mexico-City (Jáuregui, 2002). January is the coldest month of the year with temperatures ranging between 6°C and 22°C, with low precipitation, during the dry season. During the 2 weeks of January 2018, the average temperature of Mexico City, calculated from the 26 weather stations, varies from 2.3° to 21.3°C with a mean of 11.7 ± 4.5 °C (**Figure 2.18**). Conversely, the month of May is the hottest month of the year in Mexico City (13 to 26°C), with a mean temperature of 20 ± 3.4 °C recorded in 2018 (**Figure 2.18**). The month of July is characterized by lower temperatures compared to May (18.4 ± 3 °C in July 2018), but with higher precipitation rates in the city (52 mm) corresponding to the rainy season.

The model succeeds in simulating most of the seasonal, synoptic and

diurnal variability of the surface temperature deduced from the network of 26 SEDEMA stations (**Figure 2.18**). In May and July, the mean daily biases are respectively equal to 0.6°C and 0.5°C, but the peak-to-peak amplitude is underestimated by 3.0 and 2.3°C. For most of the days, night minimums are overestimated by about 2°C and daily maximums are underestimated by about 0.4-0.8°C. On the contrary, the average temperatures of the 26 stations over the month of January are underestimated by about 0.3℃. This positive anomaly may be related to residential and commercial heating in winter. Correlation coefficients for the near-surface temperature (3.5 to 30 m a.g.l.) between observation and simulation are very high (0.95, 0.97 and 0.92 for January, May and July separately). The fit of our regression model (R-square) is larger than 0.85. The simulated diurnal cycle in May is highly correlated with observations (R=0.93), including the sharp drop after sunset on clear nights and the smooth variations on cloudy nights.

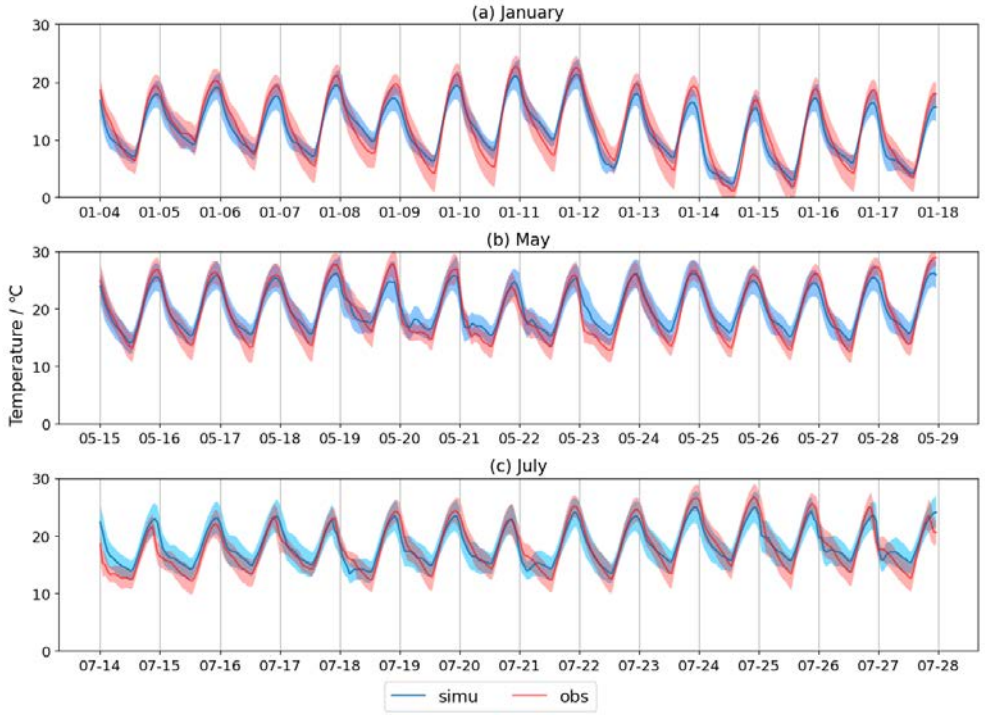


Figure 2.18 *Time series of the mean hourly temperature over the 26 meteorological stations in and around Mexico City, in UTC time (local time=UTC-6), for the months of (a) January, (b) May, and (c) July of 2018.*

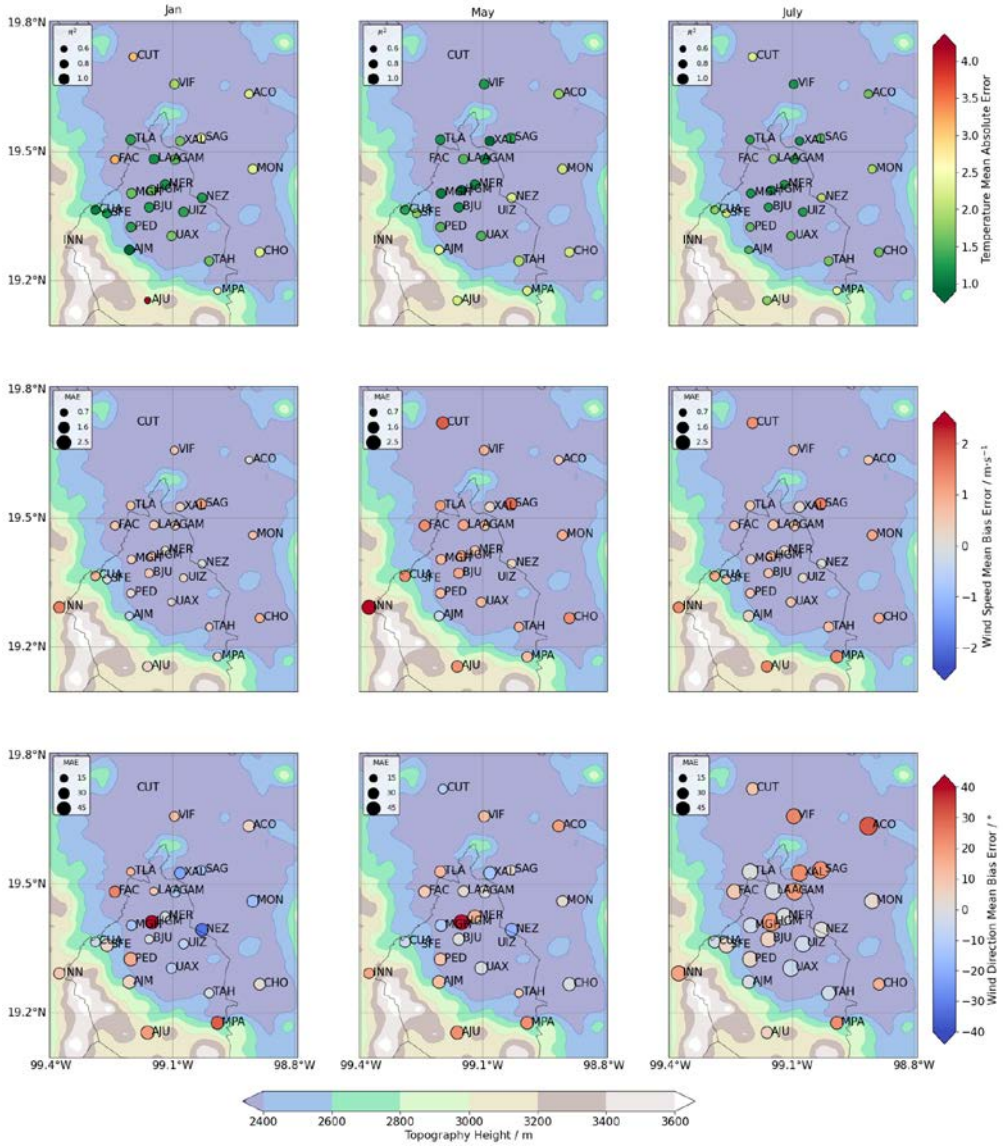


Figure 2.19 *Mean Absolute Error (MAE) in first-level temperature (0-16m; upper row) and Mean Error in wind speed (middle row) and wind direction (bottom row) for the months of January (left column), May (middle column) and July (right column) of the year 2018.*

Observed and simulated surface wind speeds were compared at the 26 meteorological stations. For January, May and July, the network-wide Mean Absolute Error (MAE) of the 26 stations is 0.98, 1.40 and 1.29 m/s, respectively. The average s MAE for wind speed vary from 0.72 to 1.67 m/s across the station network. As shown in **Figure 2.19** and **Figure 2.20**, mountain stations (e.g. INN, AJU and MPA) show

degraded performances (1.5-2.5 m/s overestimated). Urban stations in the city center correspond to a mean bias of less than 1 m/s and a regression fit (R-square) larger than 0.5. Surface wind directions vary across stations with MAE's between 50° to 70° (cf. **Figure 2.21**). Unlike wind speed, the simulated wind directions vary across the basin independently of the topography, possibly due to the misrepresentation of building heights and the use of default parameters in the urban canopy model. Overall, meteorological performances are similar to previous urban studies without data assimilation (Feng et al., 2016; Lian et al., 2018).

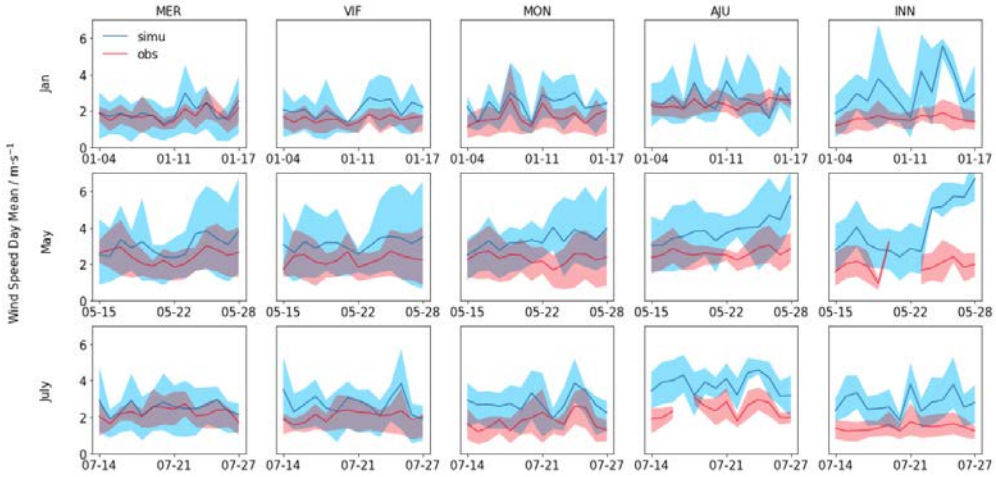


Figure 2.20 *Wind speed daymean with interquartile range in January, May and July 2018 at 5 of the 26 stations: MER in the city center, VIF in the north, MON in the east, AJU in the south (mountains), and INN in the west (mountains), in UTC time (local time=UTC-6).*

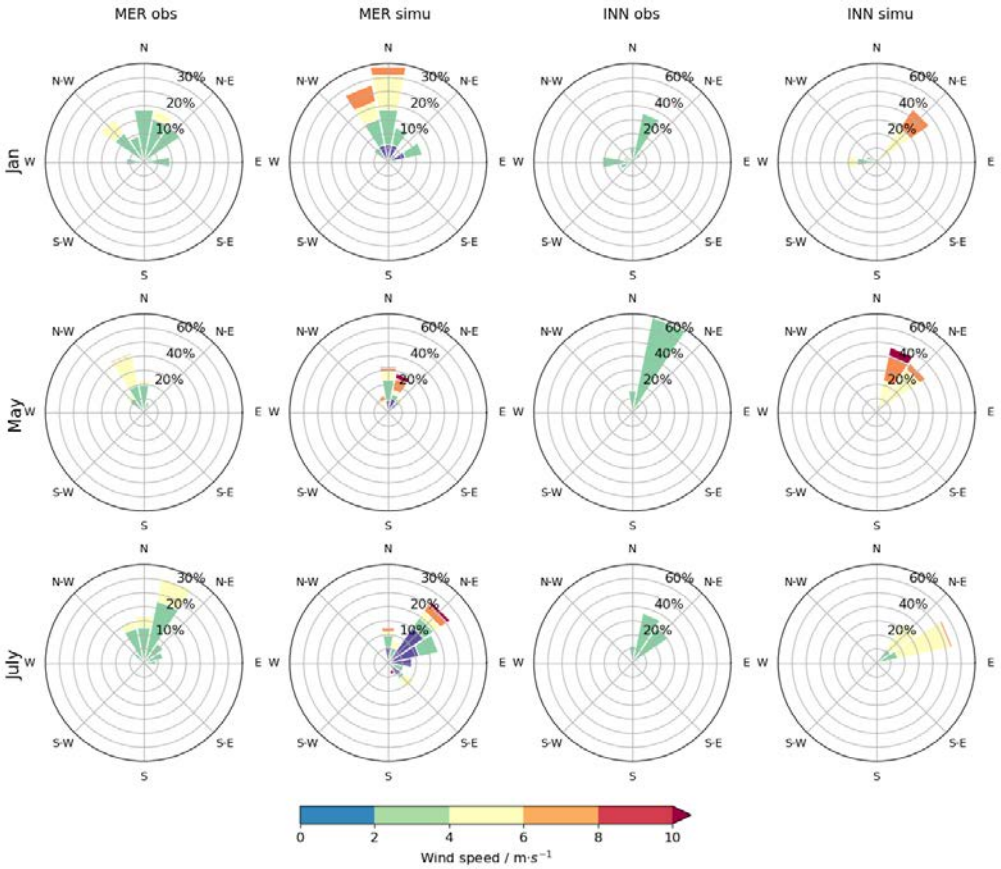


Figure 2.21 *Simulated and observed wind roses during daytime hours (07:00-17:00 local time, UTC-6) at two different meteorological stations: MER (in the city center) and INN (in surrounding mountains) over the 3 comparison period in January, May and July 2018, when observed wind speed > 2m/s.*

2.3.3 Vertical mixing and horizontal wind evaluation

Two vertical wind error profiles from the Lidar and the model are shown in **Figure 2.22** to illustrate our model-data performances (left panel), in addition to the MAEs from 0 to 4,000 m above ground level (right panel). No observations were available for the month of July. Since the sample heights of the model and the Lidar are different, MAE was calculated based on the observation records and the simulation wind speed (49 layers) interpolated to observation layer heights (99 heights). MAEs are less than 2 m/s in both months. In January, the mean error and mean absolute errors in the horizontal mean wind are

-0.4 m/s and 1.3 m/s respectively, with standard deviations of 0.9 m/s and 0.6 m/s. In May, they are equal to -0.7 m/s and 2.0 m/s, with standard deviations of 1.7 m/s and 1.1 m/s. Both errors increase for layers higher than 2000 m, which corresponds to the maximum observing elevation of the Lidar instrument.

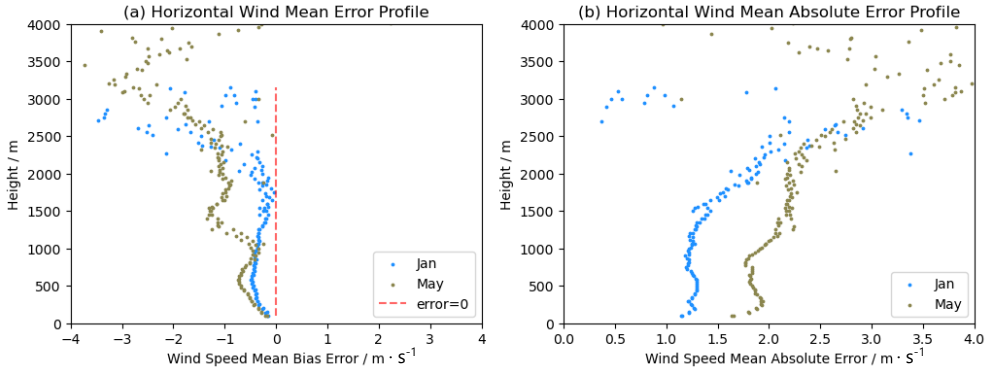


Figure 2.22 *Vertical profile of 2-week mean horizontal wind speed errors at UNA in the afternoon: (a) mean error profile (b) mean absolute error profile.*

To evaluate the temporal variations of the modeled PBL heights, I performed a linear regression for the months of January and May, with estimated R-square values of 0.68 and 0.64, respectively. The mean bias during morning hours is about 15 m in January and about 280 m in May. During the afternoon hours, the bias reaches -580 m (in January) and 110 m (in May). These values are similar to the study of Feng et al. (2016) over the basin of Los Angeles using aircraft PBL heights. During the afternoons, the modeled PBL heights are usually lower than the Lidar measurements during the month of January. We note here that no Lidar measurement is available for the month of July 2018.

In addition to the Lidar measurements, I have also tried to use the PBL heights derived from Radiosonde measurements. Again, due to the weather, there are many radiosonde scans in the rainy season that don't have a reasonable PBL height. Sometimes the inferred PBL height could reach unrealistic values of 10.000 meters. In the rainy season, it is cloudier and the temperature structure is complicated. This causes the huge uncertainty in locating the PBL boundaries. The remaining measures in July are all at UTC time 12:00, which is 7:00 AM local time. As shown in **Figure 2.23**, the Radiosonde measures match our

simulated PBL heights during July 25-27, but are much higher on the other days. The situation for May is similar: sometimes the Radiosonde matched the simulation or Lidar observation, sometimes not with a large bias. In January, the Radiosonde measures are much more reasonable: none of them is larger than 5000 m. The mean absolute errors of Radiosonde in the three months are 430 m, 2241 m and 1567 m, including the abnormal measurements under 5000 m.

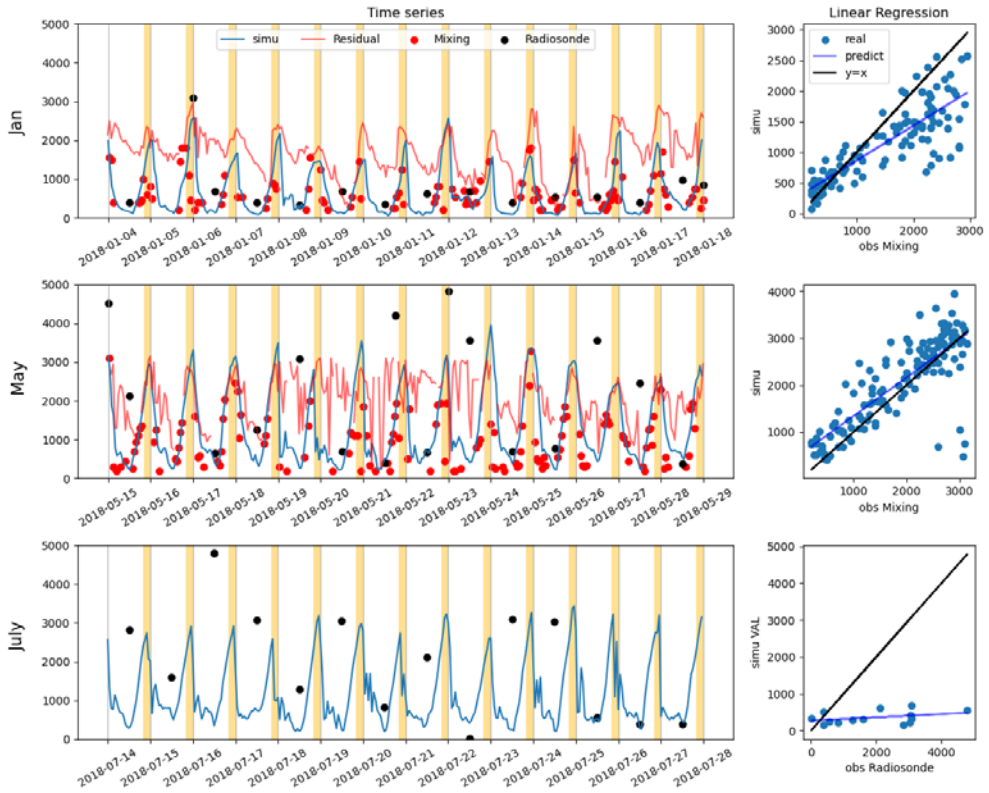


Figure 2.23 *Time series and linear regression of PBL height (in meters) simulation (in blue) and observation ("residual layer" heights in red lines and mixing layer heights in red dots) in UNA, as well as radiosonde inference in black dots (time in UTC, golden background stands for local time UTC-6 afternoon: 14:00-18:00)*

2.4 CONCLUSIONS AND DISCUSSIONS

In this chapter I evaluated the performances of the WRF and WRF-Chem model to simulate the CO_2 transportation conditions over MCMA. The meteorological conditions (surface temperature, wind speed and wind direction) were evaluated across the SEDEMA

observation network of 26 stations, and the development of the PBL was compared to the Lidar Doppler observations at UNAM campus.

Of all the configurations I tested, the updated meteorological drivers from ERA-5 significantly improved the model performance in all surface meteorological parameters, especially the temperatures, compared to ERA-Interim and NAM. The PBL scheme MYNN simulated the PBL heights better than MYJ and YSU. Both MYNN and MYJ overestimated the surface wind speed. YSU corrected the bias, but YSU did not reproduce the wind speed time series satisfactorily. My results don't show the necessity of using YSU in the future simulations. Just like YSU, neither grid nudging nor high-resolution domain seems to significantly improve the results. The multi-layer urban canopy model BEP is better than the single-layer UCM in meteorological parameters. It reduced the errors in PBL heights, repaired the bias of surface temperatures and corrected the overestimation of wind in a mountain station. However, it tends to create extra peaks in CO_2 concentrations, and the model coupled with BEP is not as stable as with UCM. For this reason, the UCM urban canopy model was finally selected. Domain 2's lower input will be brought to domain 1 through two-way nesting domains. To resolve the issue, a simple solution would be to switch to one-way nesting and allocate 2 weeks for the CO_2 to spin up from a distance.

Further evaluations have been conducted on the model with the optimized configuration in 3 months, representing various seasons. Within and around the city, the Mean Absolute Errors (MAE's) in wind speed are around 1 m/s, while the wind speed errors were 1.5-2.5 m/s over-estimated by the model in the mountainous areas outside the city. The regression model of measured and modeled PBL heights at UNA is significant, with R-square values exceeding 0.64 for both seasons. Additionally, our simulation accurately captures local dynamics across the basin, as indicated by mean errors in horizontal wind speed at UNA being less than 1 m/s within the PBL.

REFERENCES

Arzoumanian, E., Vogel, F. R., Bastos, A., Gaynullin, B., Laurent, O., Ramonet, M., & Ciais, P. (2019). Characterization of a commercial lower-cost medium-precision non-dispersive infrared sensor for

atmospheric CO_2 monitoring in urban areas. Atmospheric Measurement Techniques, 12(5), 2665-2677.

Bréon, F. M., Broquet, G., Puygrenier, V., Chevallier, F., Xueref-Remy, I., Ramonet, M., ... & Ciais, P. (2015). An attempt at estimating Paris area CO_2 emissions from atmospheric concentration measurements. Atmospheric Chemistry and Physics, 15(4), 1707-1724.

Burgos-Cuevas, A., Magaldi, A., Adams, D.K. et al. Boundary Layer Height Characteristics in Mexico City from Two Remote Sensing Techniques. Boundary-Layer Meteorol (2022). https://doi.org/10.1007/s10546-022-00759-w

Chen, F, Kusaka, H, Bornstein, R, Ching, J, Grimmond, CSB, Grossman-Clarke, S, Loridan, T, Manning, KW, Martilli, A, Miao, S, Sailor, D, Salamanca, FP, Taha, H, Tewari, M, Wang, X, Wyszogrodzki, AA and Zhang, C. (2011). The integrated WRF/urban modelling system: development, evaluation, and applications to urban environmental problems. International Journal of Climatology 31(2): 273– 288. DOI: https://doi.org/10.1002/joc.2158

Chen, G., Shan, Y., Hu, Y., Tong, K., Wiedmann, T., Ramaswami, A., ... & Wang, Y. (2019). Review on city-level carbon accounting. Environmental science & technology, 53(10), 5545-5558.

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.

Deng, A., Lauvaux, T., Davis, K. J., Gaudet, B. J., Miles, N., Richardson, S. J., ... & Gurney, K. R. (2017). Toward reduced transport errors in a high resolution urban CO_2 inversion system. Elementa: Science of the Anthropocene, 5.

Feng, S., Lauvaux, T., Newman, S., Rao, P., Ahmadov, R., Deng, A., ... & Yung, Y. L. (2016). Los Angeles megacity: a high-resolution land–atmosphere modeling system for urban CO 2 emissions. Atmospheric Chemistry and Physics, 16(14), 9019-9045.

González del Castillo, E., Taquet, N., Bezanilla, A., Stremme, W., Ramonet, M., Laurent, O., ... & Grutter, M. (2022). CO_2 variability in the Mexico City region from in situ measurements at an urban and a background site. Atmosphere , 35 (2), 377-393.

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., & Eder, B. (2005). Fully coupled "online" chemistry within the WRF model. Atmospheric environment, 39(37), 6957-6975.

Gurney, K. R., Liang, J., Roest, G., Song, Y., Mueller, K., & Lauvaux, T. (2021). Under-reporting of greenhouse gas emissions in US cities. Nature communications, 12(1), 553.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horάnyi, A., Muñoz‐Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J-N. (2017): Complete ERA5 from 1979: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service (C3S) Data Store (CDS). (Accessed on 21-Jan-2021)

INEGI (Instituto Nacional de Estadística y Geografía, Mexico National Institute of Statistics and Geography) Mexico (2021). Registered Motor Vehicles in Circulation, available at: https://www.inegi.org.mx/programas/vehiculosmotor/ (last access: 29 December 2022)

Lauvaux, T., Gurney, K. R., Miles, N. L., Davis, K. J., Richardson, S. J., Deng, A., ... & Turnbull, J. (2020). Policy-relevant assessment of urban CO_2 emissions. Environmental Science & Technology, 54(16), 10237-10245.

Lauvaux, T., Miles, N. L., Deng, A., Richardson, S. J., Cambaliza, M. O., Davis, K. J., ... & Wu, K. (2016). High‐resolution atmospheric inversion of urban CO_2 emissions during the dormant season of the Indianapolis Flux Experiment (INFLUX). Journal of Geophysical Research: Atmospheres, 121(10), 5213-5236.

Lian, J., Wu, L., Bréon, F. M., Broquet, G., Vautard, R., Zaccheo, T. S., ... & Ciais, P. (2018). Evaluation of the WRF-UCM mesoscale model and ECMWF global operational forecasts over the Paris region in the prospect of tracer atmospheric transport modeling. Elementa:

Science of the Anthropocene, 6.

Lopez-Coto, I., Hicks, M., Karion, A., Sakai, R. K., Demoz, B., Prasad, K., & Whetstone, J. (2020). Assessment of planetary boundary layer parameterizations and urban heat island comparison: Impacts and implications for tracer transport. Journal of applied meteorology and climatology, 59(10), 1637-1653. https://doi.org/10.1175/JAMC-D-19-0168.1

Lopez-Coto, I., Ren, X., Salmon, O. E., Karion, A., Shepson, P. B., Dickerson, R. R., ... & Whetstone, J. R. (2020). Wintertime CO_2, CH4, and CO emissions estimation for the Washington, DC–Baltimore metropolitan area using an inverse modeling technique. Environmental science & technology, 54(5), 2606-2614.

Mueller, K. L., Lauvaux, T., Gurney, K. R., Roest, G., Ghosh, S., Gourdji, S. M., ... & Whetstone, J. (2021). An emerging GHG estimation approach can help cities achieve their climate and sustainability goals. Environmental Research Letters, 16(8), 084003.

Nakanishi, M., & Niino, H. (2009). Development of an improved turbulence closure model for the atmospheric boundary layer. Journal of the Meteorological Society of Japan. Ser. II, 87(5), 895-912.

Nickless, A., Rayner, P. J., Engelbrecht, F., Brunke, E. G., Erni, B., & Scholes, R. J. (2018). Estimates of CO_2 fluxes over the City of Cape Town, South Africa, through Bayesian inverse modelling. Atmospheric Chemistry and Physics, 18(7), 4765-4801. https://doi.org/10.5194/acp-18-4765-2018

NOAA, https://gml.noaa.gov/ccgg/trends/global.html (last access: 2022/12/29)

Olivier, J.G.J., Janssens-Maenhout, G., Muntean, M. and Peters, J.A.H.W. (2015) Trends in global CO_2 emissions: 2015 Report. PBL Netherlands Environmental Assessment Agency, The Hague; European Commission, Joint Research Centre (JRC), Institute for Environment and Sustainability (IES). Internet: http://edgar.jrc.ec.europa.eu/news_docs/jrc-2015-trends-in-global-CO_2-emissions-2015-report-98184.pdf. JRC report 98184/ PBL report 1803, 2015.

Rella, C. W., Chen, H., Andrews, A. E., Filges, A., Gerbig, C., Hatakka, J., ... & Zellweger, C. (2013). High accuracy measurements of dry mole fractions of carbon dioxide and methane in humid air. Atmospheric Measurement Techniques, 6(3), 837-860.

Sánchez-León, E., Castro, T., Peralta, Ó., Álvarez-Ospina, H., Espinosa, M. D. L. L., & Martínez-Arroyo, A. (2016). Estimated carbon dioxide exchange for three native species in an ecological reserve of Mexico City. Atmósfera, 29(3), 189-196.

Schuh, A. E., Otte, M., Lauvaux, T., & Oda, T. (2021). Far-field biogenic and anthropogenic emissions as a dominant source of variability in local urban carbon budgets: A global high-resolution model study with implications for satellite remote sensing. Remote Sensing of Environment, 262, 112473.

SEDEMA (2016). Inventario de Emisiones de la CDMX 2014. Dirección General de Calidad del Aire, Dirección de Proyectos de Calidad del Aire. Ciudad de México (http://www.aire.cdmx.gob.mx/descargas/publicaciones/flippingbook/inventario-emisiones-cdmx2014-2/mobile/index.html)

SEDEMA (2021). Inventario de Emisiones de la Zona Metropolitana del Valle de México 2018. Dirección General de Calidad del Aire, Dirección de Proyectos de Calidad del Aire. Ciudad de México. Agosto (http://www.aire.cdmx.gob.mx/descargas/publicaciones/flippingbook/inventario-emisiones-cdmx-2018/Inventario-de-emisiones-cdmx-2018.pdf)

Seto, K., Bigio, A., Blanco, H., Delgado, G. C., Dewar, D., Huang, L., Inaba, A., Kansal, A., Lwasa, S., McMahon, J., Müller, D. B., Murakami, J., Nagendra, H., & Ramaswami, A. (2015). Mitigation of climate change. In C. B. Field et al. (Eds.), Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (pp. 361–409). United Kingdom and New York, NY: Cambridge University Press.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D., Duda, M. G., ... Powers, J. G. (2008). A Description of the Advanced Research WRF Version 3 (No. NCAR/TN-475+STR). University

Corporation for Atmospheric Research. doi:10.5065/D68S4MVH

Somanathan E., T. Sterner, T. Sugiyama, D. Chimanikire, N.K. Dubash, J. Essandoh-Yeddu, S. Fifita, L. Goulder, A. Jaffe, X. Labandeira, S. Managi, C. Mitchell, J. P. Montero, F. Teng, and T. Zylicz, 2014: National and Sub-national Policies and Institutions. 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. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel and J.C. Minx (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Staufer, J., Broquet, G., Bréon, F. M., Puygrenier, V., Chevallier, F., Xueref-Rémy, I., ... & Ciais, P. (2016). The first 1-year-long estimate of the Paris region fossil fuel CO 2 emissions based on atmospheric inversion. Atmospheric Chemistry and Physics, 16(22), 14703-14726.

Tewari, M, Chen, F, Kusaka, H and Miao, S. (2007). Coupled WRF/Unified Noah/urban-canopy modeling system. NCAR WRF Documentation 122: 1–22. NCAR, Boulder

Tomohiro Oda, Shamil Maksyutov (2015), ODIAC Fossil Fuel CO_2 Emissions Dataset (Version name : ODIAC2019), Center for Global Environmental Research, National Institute for Environmental Studies, doi:10.17595/20170411.001. (Reference date : 2022/05/22)

United Nations, Department of Economic and Social Affairs, Population Division (2018). The World's Cities in 2018—Data Booklet (ST/ESA/ SER.A/417).

Wigley, T. (1983). The pre-industrial carbon dioxide level. Climatic change, 5(4), 315-320.

Yadav, V., Ghosh, S., Mueller, K., Karion, A., Roest, G., Gourdji, S. M., ... & Whetstone, J. (2021). The impact of COVID‐19 on CO_2 emissions in the Los Angeles and Washington DC/Baltimore metropolitan areas. Geophysical research letters, 48(11), e2021GL092744.

Ye, X., Lauvaux, T., Kort, E. A., Oda, T., Feng, S., Lin, J. C., ... & Wu, D. (2020). Constraining fossil fuel CO_2 emissions from urban area using

OCO - 2 observations of total column CO_2. Journal of Geophysical Research: Atmospheres, 125(8), e2019JD030528.

3 CHAPTER III : ATMOSPHERIC CO_2 SIMULATION AND VARIABILITY OVER MEXICO CITY METROPOLITAN AREA

SUMMARY

In this chapter, the main objective is to investigate the potential of future inversion of surface CO_2 and XCO_2 based on the current simulation system and observation network. The seasonal variability and the component of CO_2 and XCO_2 at the urban station UNA and the background station ALZ and over the Mexico City Metropolitan Area (MCMA) is analyzed as well, including the contribution of biogenic fluxes. For the inversion in the next chapter, I examined the spatial distribution pattern of the background CO_2 and XCO_2 from long distance transportation, discussed the ideal location of background stations and fossil fuel signal stations, and assessed the quality of the background station ALZ. Additionally, the sensitivity test on emission inventories mentioned in the last chapter (ODIAC and UNAM) was made to compare the impact of emission inventory in the modeling system, to decide whether to include ODIAC in the future inversion.

Based on the sensitivity tests made on the WRF model for January 2018 in the last chapter, the optimal modeling configurations were decided. Further simulations in different seasons (the dry season, the rainy season, and the transition between the two) were made to evaluate the model performance on CO_2 simulation in various climate conditions. The meteorological drivers, ECMWF Reanalysis interim (ERA-interim), were updated to the latest version ECMWF Reanalysis version 5 (ERA-5). Both emission products (UNAM and ODIAC), as well as the three sectors (area / residential & commercial, mobile / transportation, point / industrial) provided by UNAM, are applied in separate tracers in WRF-Chem. Single-layer urban canopy model (UCM) is coupled with WRF-Chem instead of the multi-layer canopy scheme BEP. I chose the one-way double nesting domain plan for reasons of computational efficiency. The resolution of the inner domain remains 5 km. The simulations and observations in January (dry season), May (transition of the two seasons) and July (rainy season) of 2018 were compared to evaluate the performance of our

WRF-Chem modeling system on CO_2 simulation at the relatively lower resolution in various seasons and weathers. Each of the three simulations lasts for 4 weeks, 2 weeks for background spinning up from a long distance and 2 weeks for comparison.

For the performance of the model, the time series, diurnal cycle and weekly distribution of surface CO_2 are evaluated by the observation records from MERCI-CO_2 measurement network. Despite the lack of measurement due to rainy weather in rainy seasons, XCO_2 measurements by FTIR are decomposed by the model in the same way. Additionally, the CO_2 concentration maps of background, fossil fuel and biogenic flux contributions are shown to describe the spatial distribution of fossil fuel signals in different sectors, the characteristics of biogenic fluxes, and the network design.

The main conclusions from this chapter are: i) The model reproduced most of the observed variations in the afternoon (2PM-5PM UTC-6) in both seasons, though it tends to under-estimate the peaks in the morning and there remains a 1-hour time lag in CO_2 accumulation. Sometimes the morning and night trends are captured as well, but only the data in the afternoon when the PBL is stable are going to be assimilated in the future inversion. ii) The model indicates that the correlation between the background concentration of the two stations are high enough and the fossil fuel signals at our background station (ALZ) are low enough, leading to the high sufficiency to constrain the city emissions by the concentration gradients between the two stations (UNA-ALZ). ODIAC could be included in the future inversion. The biogenic fluxes should be optimized separately. iv) The mismatch of XCO_2 during the first 3 days in May is mainly driven by high altitude air mass. Boundary conditions should also be optimized separately in future inversions of XCO_2.

The main points of this chapter, as well as Chapter 2 Evaluation of Configurations and Options in WRF modeling system over Mexico City Metropolitan Area, forms a manuscript Evaluation of atmospheric CO_2 simulation over the Mexico City metropolitan area by WRF-chem submitted to Sustainable Cities and Society.

3.1 SURFACE CO_2 EMISSIONS AND BOUNDARY CONDITIONS

3.1.1 Fossil fuel emissions inventories

Two emission products were used in our study: a gridded emission inventory (Garcia-Reynoso et al., 2018; Rodriguez-Zas et al., 2021. https://github.com/JoseAgustin/emis_2016) developed at the National Autonomous University of Mexico (UNAM), and the Open-source Data Inventory for Anthropogenic CO_2 (ODIAC) (Oda et al., 2018). The UNAM hourly national emission inventory of Mexico covers the entire country at 3-km resolution while the 1-km resolution product is only available for specific subdomains, including the MCMA. This local emission inventory provides CO_2 emissions, as well as other species including CO and NO_x, for three sectors: area sources (residential/commercial), mobile sources (traffic), and point sources (industrial). In our simulation, we defined one independent tracer per sector to quantify individually the sectoral contributions on the observed atmospheric CO_2 enhancements.

The ODIAC emission product provides global emission maps of fossil fuel CO_2. ODIAC downscales national CO_2 emissions from the Carbon Dioxide Information Analysis Center (CDIAC; Boden et al., 2017) to sub-national levels by disaggregating the emissions according to night light data from the Visible Infrared Imaging Radiometer Suite (VIIRS) on the Suomi National Polar-orbiting Partnership satellite (e.g., Elvidge et al., 2013). Combined with temporal scaling factors (Nassar et al., 2013), ODIAC emissions are available at 1×1 km resolution and hourly (including monthly, weekly and daily temporal profiles).

Anthropogenic emissions over the parent domain (d01) are interpolated from ODIAC, including areas outside Mexico and shipping over the ocean. We note, however, that shipping emissions remain low compared to the UNAM inventory. For our second domain at 5-km resolution, we applied the UNAM emission inventory as well as ODIAC in another tracer as a sensitivity test. To conserve the total budget, all the emissions are interpolated into our domains by applying a mass-conserving technique (xESMF - conservative method, https://doi.org/10.5281/zenodo.1134365).

Figure 3.1 shows the difference of the two emission products at the grid covering UNAM campus after interpolation. Due to the

characteristics of nightlight and the temporal scaling factors, ODIAC varies much less than UNAM, the daily amplitude is only 20 tons, one fifth of that for UNAM, which is nearly 100 tons. For the weekly pattern, ODIAC assumes that every weekday is homogeneous; Saturday is 5-6% less than weekdays and Sunday is 5-6% less than Saturday. UNAM, on the contrary, is based on a traffic database and Saturday has a higher emission than most weekdays.

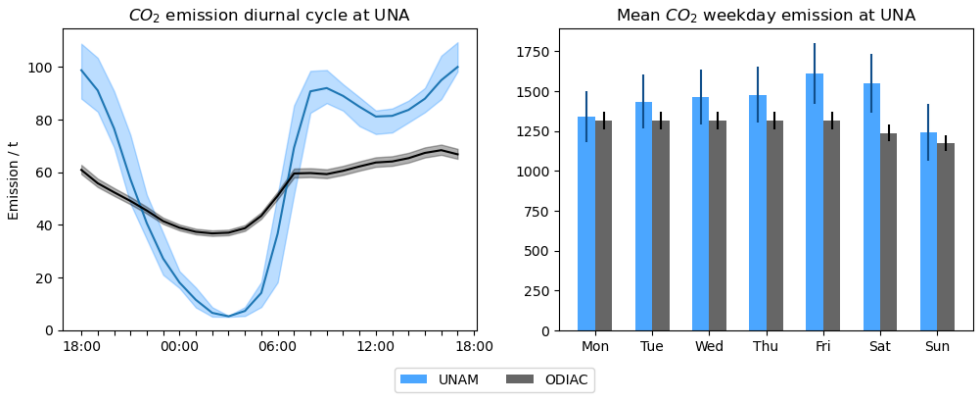


Figure 3.1 *Diurnal cycle (left panels) and weekly cycle (right panels) of fossil fuel CO_2 emissions at the urban station UNA in MCMA for the UNAM and ODIAC emissions products in 2018, with error bars.*

CO_2 emissions from mobile sources remain nearly constant all year long (about 3 Mt per month), but emissions from residential and industrial sources show large seasonal variations with larger values for the month of January (11.93 Mt) compared to May (9.10 Mt). Emissions for the month of July increase to 10.00 Mt due to the energy demand from air conditioning systems (**Figure 3.2**).

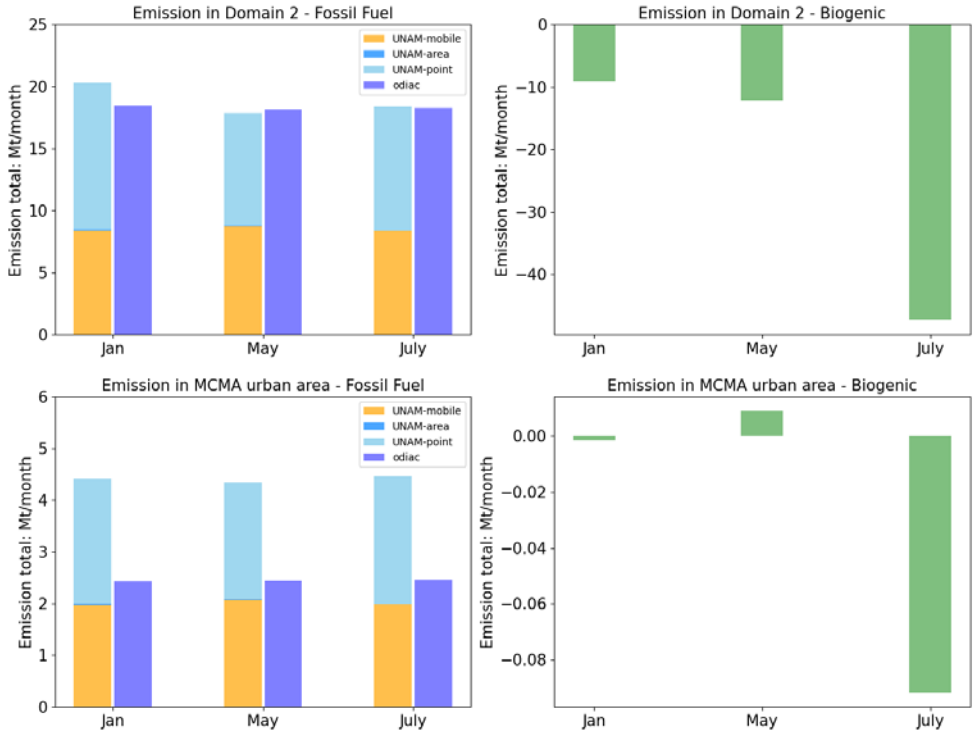


Figure 3.2 *Total fossil fuel CO_2 emissions over the domain 2 (upper panels) and over the urban area of MCMA (lower panels) for the UNAM and ODIAC emissions products (left panels), and biogenic fluxes from the CASA biogeochemical model (right panels) for the months of January, May, and July 2018.*

Total emissions over domain 2 from UNAM and ODIAC are generally consistent, but their spatial distributions differ (**Figure 3.3**). Although both ODIAC and UNAM grid values follow a lognormal distribution in the same interval (**Figure 3.3** bc), ODIAC shows a larger number of low-value grid points (near-zero values). Due to the spatial disaggregation technique, point source emissions in ODIAC tend to be more distributed spatially compared to UNAM, where point sources are geo-located by their exact coordinates. For our application, the ODIAC emissions show the same distribution at 15 km resolution, with consistent country-scale emissions. When it comes to city-scale, ODIAC underestimate the emissions by 43%-45% compared to UNAM, similar to other cities in North America (Chen et al., 2020).

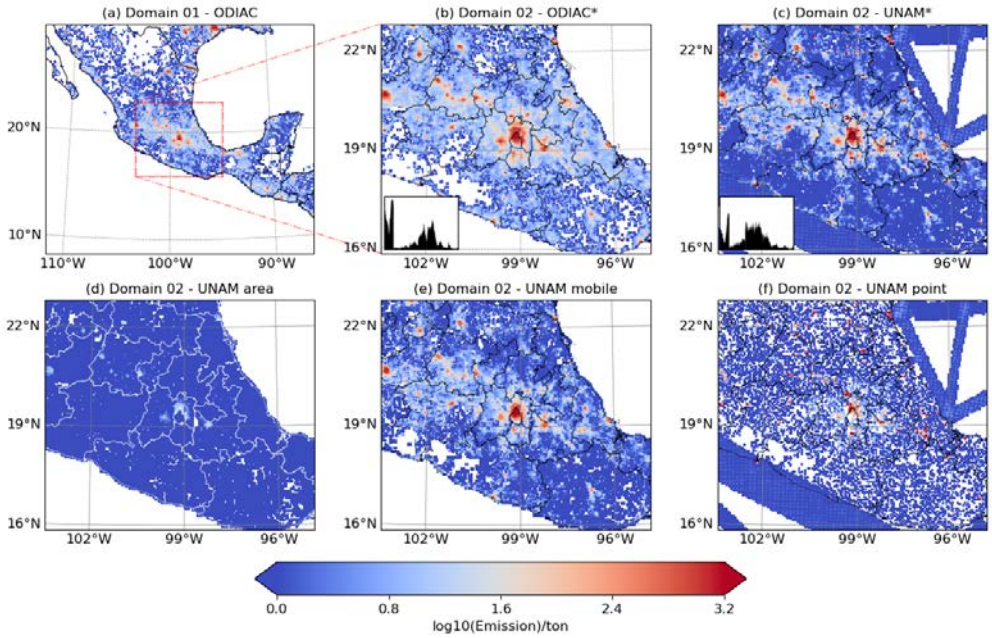


Figure 3.3 *Maps of total fossil fuel CO_2 emissions over (a) domain 1 and domain 2 from (b) ODIAC and (c) the UNAM inventory on 4 January 2018 (UTC time, logarithmic scale, tons). Sectors of UNAM are shown on lower panels: (d) area sources, (e) mobile sources and (f) point sources.*

* Frequency distributions (black histograms) are shown over domain 2 to illustrate the larger ratio of near-zero pixels in ODIAC and the presence of large peaks in the UNAM inventory.

3.1.2 Biogenic CO_2 fluxes: optimized CASA simulation

The biogenic fluxes in MCMA were simulated by the Carnegie Ames Stanford Approach (CASA) model (Zhou et al., 2020) and are also interpolated to our domains by xESMF - conservative. The CASA simulation was optimized by generating an ensemble of perturbed parameters, including the maximum light use efficiency, the optimal temperature of photosynthesis, and the temperature response of respiration. These perturbed parameters were constrained by selecting the best configuration in comparison to AmeriFlux eddy-covariance flux data. Monthly Gross Primary Productivity (GPP) and total ecosystem respiration (Re) were downscaled to 3-hourly resolution based on 3-hourly air temperature and shortwave downward radiation from the North American Regional Reanalysis (Mesinger et al., 2006), as described in Olsen and Randerson (2004). GPP and Re were summed

into Net Ecosystem Exchange (NEE) at 5-km resolution. **Figure 3.4** shows the difference between dry/wet seasons, with a maximum in carbon uptake during the wet season (growing season). In **Figure 3.4**, the phenology of tropical forests (wet and dry) on the east coast of Mexico causes an early start of the growing season (May) while the central and western regions (temperate to semi-arid climates) show negative values (carbon uptake) later in the year (July).

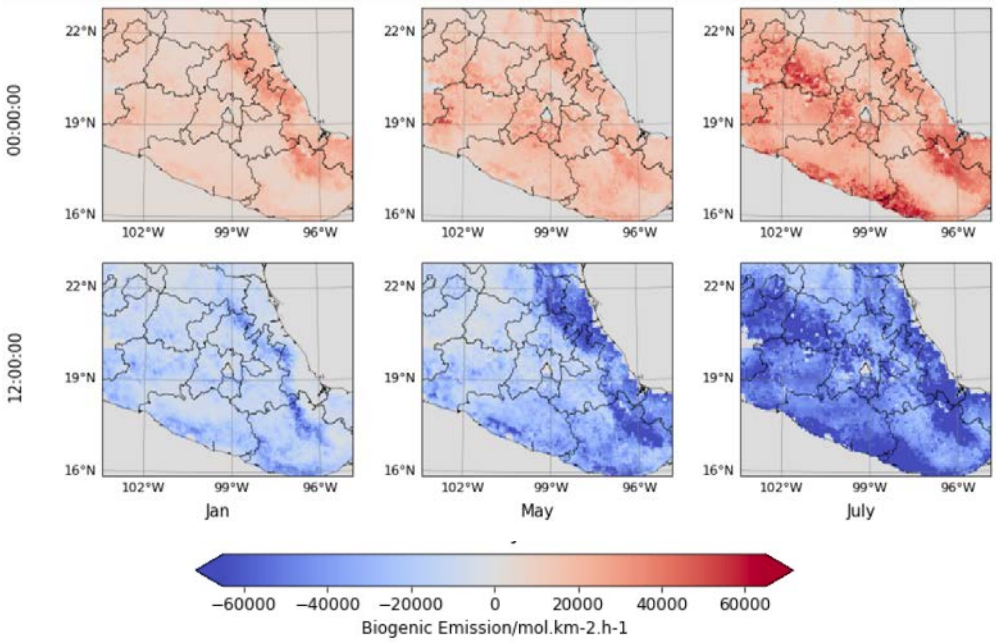


Figure 3.4 *Midday (12:00 UTC-6) and midnight (00:00 UTC-6) biogenic monthly mean fluxes from the CASA biogeochemical model (Zhou et al., 2020) over the WRF simulation domain (5-km resolution) for the months of January (left column), May (middle column) and July (right column) in mol/km^2/hr.*

We note here that the central region of our domain, corresponding to Mexico City, shows very low NEE values caused by the dense urban area (sparse vegetation). The biogenic fluxes in the MCMA differ significantly from the surrounding areas, further examined by separating the biogenic contribution in our modeling system.

3.1.3 Atmospheric CO_2 boundary conditions

The influence of distant sources and sinks was simulated by coupling the optimized CO_2 concentrations from the global model CarbonTracker 2019 (Peters et al., 2007; Jacobson et al., 2020) to our

WRF domain, similar to Feng et al. (2019) over North America. The coupling scheme involves several steps including pressure correction due to terrain height differences, interpolation of the coarse resolution CO_2 fields to the WRF grid, and mass conservation over the total CO_2 column, as described in Butler et al. (2020). As shown in **Figure 3.5**, the CO_2 boundary inflow at our domain boundaries varies at daily and sub-daily timescales (3-hourly in our case) especially over land (northern bound) where air masses from North America flow southward into our domain. We evaluate the CO_2 boundary inflow by comparison to measurements collected at ALZ (mountain site), only occasionally influenced by the city emissions from the valley (cf. section 953.3.1.1). We also note here that we couple the 3D fields from the CarbonTracker global models to represent the vertical gradients between the surface (2,200 m a.s.l.) and the mountain tops (4,000 m a.s.l.).

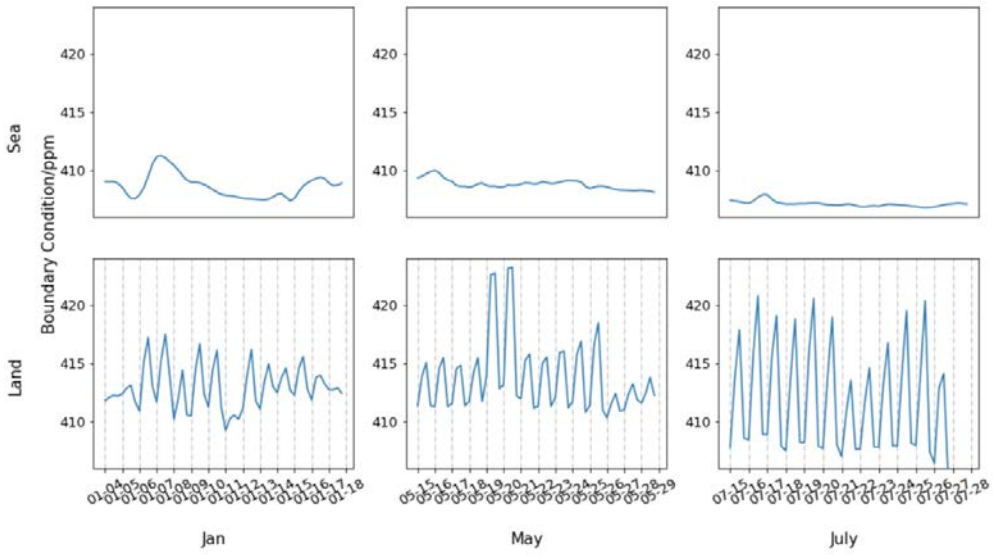


Figure 3.5 *Time series of the CO_2 boundary concentrations at the southern (sea) and northern (land) model boundaries from the CarbonTracker inversion system, in UTC time (local time=UTC-6) for January, May, and July 2018.*

3.2 CO_2 OBSERVATION NETWORK

3.2.1.1 Atmospheric CO_2 measurements

Two high-accuracy analyzers measuring continuously atmospheric

CO_2 concentrations were installed in 2014 at the UNAM university campus (UNA) and at the Altzomoni altitude station (ALZ). The two instruments are cavity ring-down spectrometers (model G2401) manufactured by Picarro (Crosson et al., 2008; Rella et al., 2013). Such instruments can measure every few seconds the atmospheric concentrations of CO_2, CH_4, CO and H_2O, with precision compatible with the WMO/GAW requirements (WMO/IAEA 2020). These instruments are now widely used in the community, and have been subject to numerous evaluations documenting their repeatability, drift, and sensitivity to parameters such as temperature and humidity (Chen et al., 2013; Welp et al., 2013; Yver Kwok et al., 2015; 2021). The two analyzers began to be calibrated at the end of 2018 with a calibration scale consisting of three cylinders prepared by NOAA/ESRL, and traceable to the international WMO scale. This reference scale is now installed at the ALZ. A second scale, prepared at LSCE as part of the MERCI-CO_2 project and also traceable to the WMO scale, is installed at the measurement site on the UNAM campus. The instrument calibration made at the end of 2018 was propagated backwards over the whole year 2018, inducing an estimated uncertainty of approximately 0.1 ppm over 2018, according to typical drift for such instruments (Yver-Kwok et al. (2015)). Another major source of uncertainty is related to the water vapor correction, which has to be done in order to express the results as a mole fraction in dry air. Considering the humidity levels observed at the two stations, Gonzalez del Castillo et al. (2022) estimated an uncertainty of 0.2 ppm and 0.05 ppm respectively for UNA and ALZ. The measurement protocols implemented at both sites are described in detail by Gonzalez del Castillo et al. (2022). CO_2 concentrations are available for the two measuring stations UNA and ALZ in 2018 without any data gap.

The two stations are located in very different environments, with a high exposure to urban emissions for UNA and a priori low at the mountain site (ALZ) at 4,000 m a.s.l. This results in significant differences in the diurnal and seasonal cycles (Gonzalez del Castillo et al., 2022). Thus, the daily peak-to-peak amplitude of CO_2 is on average around 35 ppm at UNA and 5 to 7 ppm at ALZ. It has been shown that the main driver of the diurnal variability at UNA is the development of the atmospheric boundary layer. The daily maximum is reached between 6 and 7 a.m. (local time), and a regular decrease is observed until around 4 p.m. due

to the development of the atmospheric boundary layer leading to a dilution of the compounds emitted at the surface, like CO_2, in a larger volume. CO_2 then accumulates during the night as the atmospheric boundary layer decreases associated with greater vertical stability. In the case of ALZ, Gonzales del Castillo et al. (2022) explained that the diurnal cycle resulted from a combination of the influences of boundary layer dynamics and biospheric fluxes. They also assume that the increase in concentration observed in the afternoon is probably linked to the uplift of polluted air masses from Mexico City and other surrounding urban areas.

3.2.1.2 Atmospheric CO_2 column mixing ratios

In addition to the CO_2 surface measurements, CO_2 total column measurements were performed at the ALZ and UNA stations using solar absorption FTIR spectroscopy. The ALZ station was equipped in 2012 with a high resolution FTIR spectrometer (model IFS-120/5HR, Bruker), which continuously provides vertical column densities of atmospheric trace gases officially contributing to the Network for the Detection of Atmospheric Composition Change (NDACC). The dry air column-averaged mole fractions of CO_2 (XCO_2) used in this study were calculated from the high resolution spectra (0.02 cm^{-1}) measured with KBr beamsplitter, using the PROFFIT 9.6 code (Hase et al., 2004) and adopting the CO_2 and O_2 retrieval methods described in Baylon et al., (2017). Several post-process quality filters were applied to discard data affected by clouds, volcanic ash or low signal, based on both spectra and retrieval quality indicators (signal to noise ratio, relative RMS, wavenumber shift, etc.) and statistical criteria (standard errors). The UNA station was equipped in 2016 by a mobile low resolution (0.5 cm^{-1}) EM27/SUN spectrometer, providing continuous total column measurements of CO_2, CO and CH_4. Low-resolution spectra are processed following the COllaborative Carbon Column Observing Network (COCCON) processing chain (PREPROCESS and PROFFAST-version 01-07-2018 codes) developed by the Karlsruhe Institute of Technology (KIT) and freely available (https://www.imk-asf.kit.edu/english/COCCON.php). A full description of the analytical procedure is given in Sha et al. (2020), Frey et al. (2021), and Pollard et al. (2022). Calibration coefficients and ILS determined by the KIT (https://www.imk-asf.kit.edu/downloads/Coccon/2019-11-

08_Instrument-Calibration.pdf for the used PROFFAST version) were applied according to the COCCON recommendations. We used the GGG2014 version of TCCON meteorological data and priors (daily MAPs files), downloaded from the CalTech server. Intraday ground pressure data were taken from the co-located RUOA pressure sensor measurements (https://www.ruoa.unam.mx/). UNA and ALZ XCO_2 data were inter-calibrated using data obtained from several weeks of ALZ side-by-side measurements. Noisy outliers were filtered out according to a 3-sigma criterion.

3.3 RESULTS

3.3.1 Atmospheric CO_2 contributions from fossil fuel and biogenic sources and sinks

3.3.1.1 Atmospheric in situ CO_2 concentrations

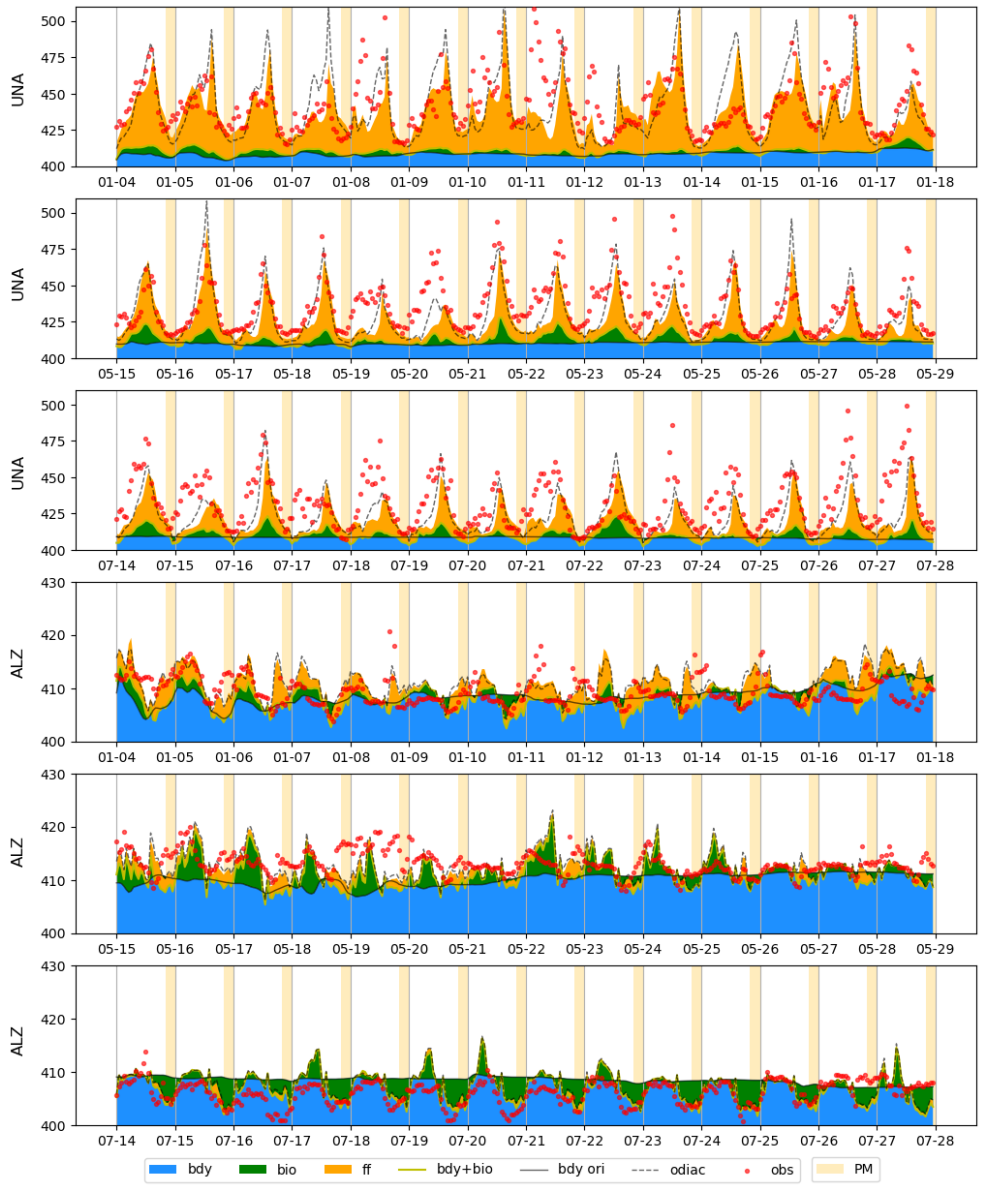


Figure 3.6 *Observed and modeled in situ CO_2 concentrations (in ppm) at both UNA and ALZ stations for the months of January, May, and July 2018, in UTC time. The modeled concentrations were separated into three components: boundary inflow (bdy - in blue), biogenic fluxes (bio*

- in green) and fossil fuel emissions (ff - in orange). The contributions from the boundary conditions are indicated by the gray line and the combined contribution of boundaries and biogenic fluxes (bdy+bio) by the light yellow line. The CO_2 concentrations simulated by ODIAC emission products are shown in gray dotted lines.

Figure 3.5 shows the comparison between the simulated and the observed in situ CO_2 concentrations at both sites (UNA and ALZ). Modeled concentrations are decomposed into three components: the background concentrations originating from the CarbonTracker inversion system (in blue), the fossil fuel contributions from Mexico City and from the entire country (in orange), and the biogenic sources and sinks caused by surface fluxes from the CASA model (in green). The WRF-Chem modeled concentrations tend to be under-estimated during the first hours of the night (about 10 ppm), revealing a time lag in the accumulation of CO_2 in stable conditions. At both stations, the model-data differences are about -0.31 ppm (UNA) and -0.71 ppm (ALZ) during the afternoon hours (cf. **Table 3.1**). While the UNA station is influenced by large fossil fuel contributions (10 ppm during the afternoon hours), the ALZ station shows relatively low contributions (1.8 ppm). The biogenic contributions vary across the two stations (2.0 and -0.27 ppm), indicating that future optimizations (flux inversions) will need to adjust separately the biogenic fluxes. The spatial distribution of biogenic CO_2 fluxes is presented in the discussion section, due to differences in ecosystem types between the Mexico City basin and in the surrounding mountains. We note here a mismatch at ALZ during the first week of May (15 to 20 May 2018), possibly caused by incorrect boundary conditions.

Unit: ppm		Jan		May		July	
		UNA	**ALZ**	**UNA**	**ALZ**	**UNA**	**ALZ**
Mean Error	**AM**	2.25	0.97	-2.51	-0.37	-0.80	1.67
	PM	1.93	0.83	-2.86	-2.91	0.00	-0.04
Mean	**AM**	21.63	2.60	9.04	2.32	8.77	2.69

Absolute Error	**PM**	5.70	2.63	3.54	3.05	3.88	1.75

Table 3.1 *Mean errors and mean absolute errors in atmospheric CO_2 concentrations at 16-48 meters a.g.l. for AM (7:00-12:00 local time UTC-6) and PM (13:00-17:00 local time UTC-6) in the months of January (left column), May (middle), and July (right) of 2018.*

We show the observed and modeled in situ CO_2 concentration mean diurnal cycles at both sites in **Figure 3.6**. The lack of accumulation at night at UNA results in a time lag between the modeled and observed nighttime peaks of about 2 hours. During the afternoon hours (14:00-18:00 local time), the WRF-Chem CO_2 concentrations capture the temporal variations and the absolute values at UNA (413-424 ppm, MAE of 3-5 ppm). Due to its remote location, the ALZ station shows a reduced diurnal cycle amplitude, simulated by WRF-Chem, but mostly driven by the boundary CO_2 concentrations from CarbonTracker (404-410 ppm, MAE of 2-3 ppm). We confirm here the ability of our modeling system to reproduce the urban enhancements and the large-scale boundary conditions at monthly timescale.

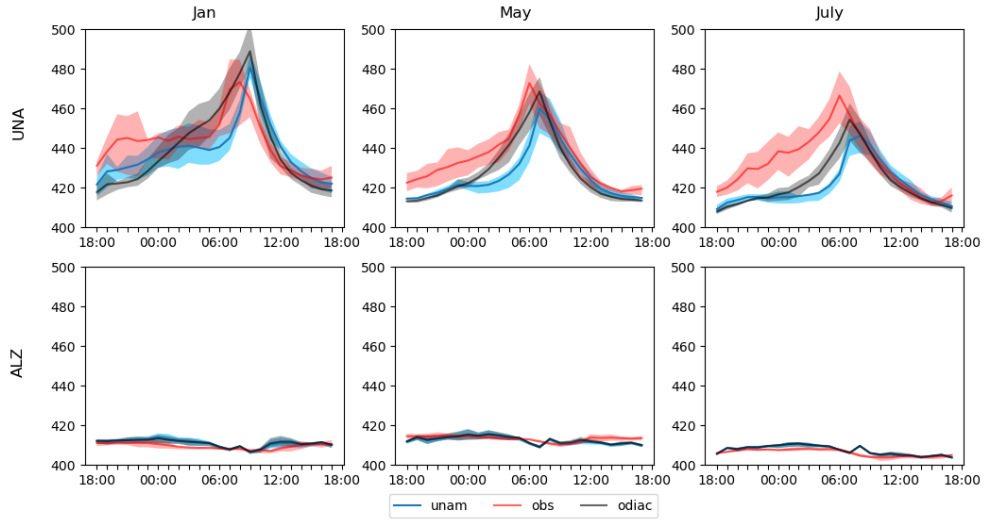


Figure 3.7 *Mean diurnal variations in atmospheric CO_2 concentrations at UNA (upper row) and ALZ (lower row) stations observed (in pink) and simulated by WRF-Chem with UNAM emission inventory (in blue) and with ODIAC (in black), for the months of January (left), May (middle), and July (right) of the year 2018, shown in local time.*

For reference, the concentrations simulated by ODIAC emission inventory is also shown in **Figure 3.5** and **Figure 3.6**. Though the total amount of ODIAC in MCMA is much smaller (c.f. section 3.1.2.1) than UNAM, the concentrations at the urban station UNA are not very low due to local emission and diffusion conditions. The simulation with ODIAC is often higher than UNAM at night, just as the emission shows. After the emission becomes obviously lower than UNAM at daytime, ODIAC simulation is only slightly lower than that of UNAM. The mean absolute errors from ODIAC are -9%-38% higher than those from UNAM (shown in **Figure 3.6**), and the mean errors are sometimes 4 times larger, but these errors are acceptable compared to the difference in emissions most of the time. It means ODIAC has the potential to be included in the future inversion for sensitivity tests.

3.3.1.2 Atmospheric CO_2 column concentrations

We show in **Figure 3.7** the model performances to simulate XCO_2 at both stations. Because EM27 instruments only measure during daytime clear-sky conditions, we focus on the afternoon hours. We note that, due to poor weather conditions, there are nearly no afternoon FTIR measurements complying with the data quality requirements during July 2018. The influence of local sources and sinks is reduced by the dilution of PBL signals within the entire column of air, but their influence remains the main driver of the observed day-to-day variability at UNA. WRF-Chem tends to over-estimate XCO_2 concentrations in January (1.01 ppm) and under-estimate XCO_2 in May (0.53 ppm; or 0.15 ppm if the abnormal period 15-17 May was excluded). The over-estimation of XCO_2 concentrations in January does not coincide with in situ CO_2 concentration differences, possibly caused by a positive bias in background concentrations rather than by local emissions. The day-to-day variations were better captured in May (R=0.81, except for the abnormal values in 15-17 May 2018) compared to January (R=0.63) (cf. **Figure 3.9**). Overall at UNA, the MAE for XCO_2 is about 1 ppm (1.22 ppm in January and 0.97 in May), indicating that background concentrations should be optimized in future inversions. Regarding the period 15-17 May 2018 during which the wind direction shifted from a northerly flow to a westerly flow, WRF-Chem failed at capturing the increase in XCO_2 (**Figure 3.9**). In general, the lack of data in the rainy season impairs our ability to quantify the XCO_2 model

errors at ALZ and UNA but the absence of fossil fuel signals at ALZ indicates that the local circulation from the valley to the mountain tops is absent during summer months.

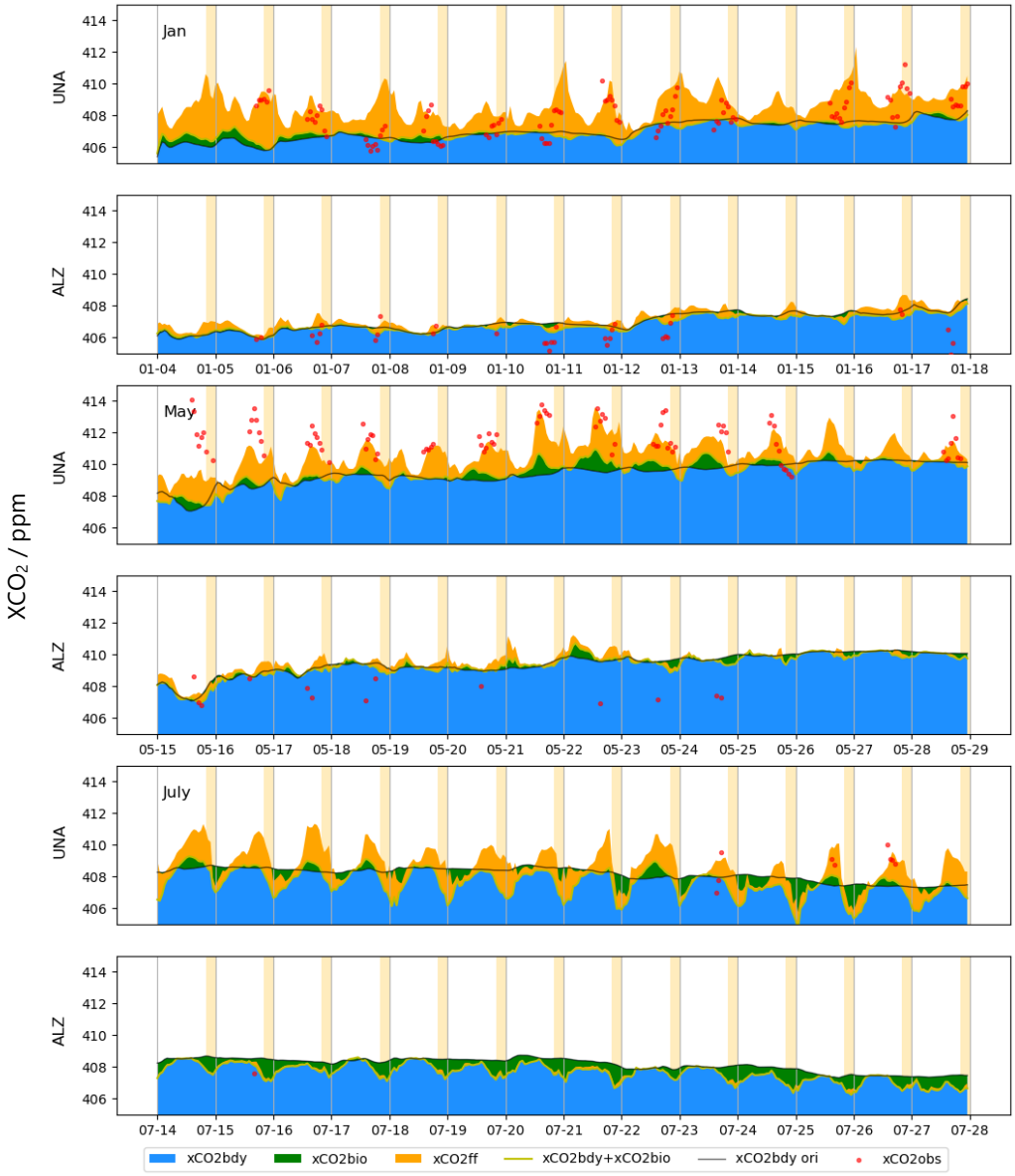


Figure 3.8 *Observed and modeled Atmospheric XCO_2 concentrations (in ppm) at both UNA and ALZ stations for the months of January, May, and July 2018, in UTC time. The modeled concentrations were separated into three components: boundary inflow (bdy - in blue), biogenic fluxes (bio - in green) and fossil fuel emissions from both simulation domains 1 and 2 (ff - in orange). The contributions from the boundary conditions are*

indicated by the gray line and the combined contribution of boundaries and biogenic fluxes (XCO_2bdy+XCO_2bio) by the light yellow line.

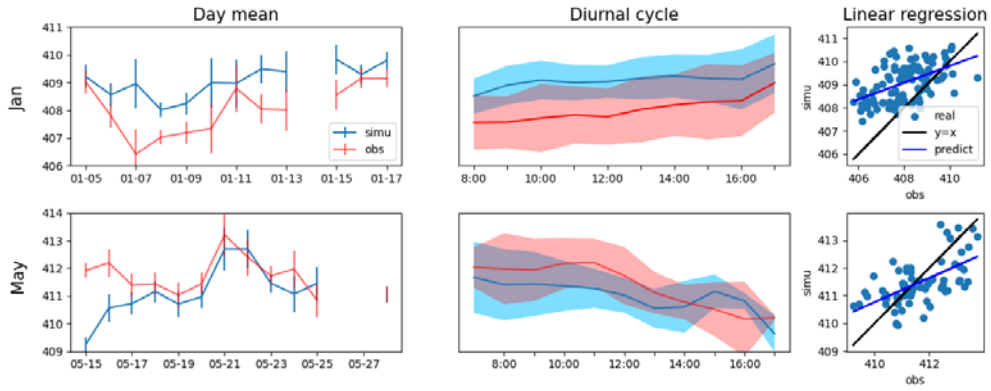


Figure 3.9 *Daily mean XCO_2 concentrations (left panels) and mean XCO_2 diurnal cycles (middle panels, local time UTC-6) at the UNA station, observed by the EM-27/SUN instrument and simulated by WRF-Chem, for the months of January and May 2018. Observed and simulated XCO_2 concentrations (right panels), with the corresponding linear regressions (blue lines, abnormal period May 15-17 excluded) at UNA for the same months.*

3.4 Conclusions and Discussion

In this chapter we evaluated the performances of the WRF-Chem model to simulate the CO_2 and XCO_2 concentrations over MCMA. Regarding surface CO_2 concentrations, the model captured most of the observed variations in the afternoon, with a mean bias of -0.52 ppm and a MAE of 3.42 ppm. WRF-Chem tends to under-estimate the peak values in the morning, showing a 1-hour time lag in the simulation of CO_2 accumulation patterns at night. Thus, only the afternoon data will be assimilated in our future inversion. The model also prompts us that the fossil fuel signals at our background station (ALZ) are low enough (1.8 ppm) to represent the background CO_2 concentrations, but the biogenic flux patterns differ between the background (ALZ) and the downtown (UNA) stations. We conclude here that the gradients between the two stations (UNA-ALZ) are sufficient to detect and to quantify the city emissions in our future inversion, but the biogenic fluxes should also be optimized separately. Except for the large mismatch in wind direction over the period 15-17 May 2018, XCO_2 day-to-day trends were captured by the model during both seasons.

The mean bias was around 1.00 ppm for January and -0.53 ppm for May, most of which was driven by the wind direction shift on 15-17 May 2018, indicating boundary conditions should also be included in future inversions of XCO_2.

3.4.1 Seasonality of fossil fuel contributions

We examine in greater details the spatial and temporal distributions of fossil fuel signatures in our modeled CO_2 concentrations from the mobile sector in **Figure 3.8**. Mobile CO_2 enhancements concentrate in the city center, representing up to 90% of the modeled local enhancements. In the southern part of the city, north to the UNA station, the contribution reaches its maximum value while the northern suburbs are dominated by industrial and energy production sources. The Tula power plant and additional surrounding industries generate a second plume north of Mexico City, where mobile sources have a minor role in the observed enhancements. In terms of emissions, the seasonality of mobile sources remains low (around 4%). The largest mobile signal originates from the dense road network within CDMX. Several satellite cities show significant signals - Toluca de Lerdo to the west, Tizayuka and Pachuca de Soto to the northeast, and Ixtapaluca to the southeast (plumes in **Figure 3.10**, bottom row). While the absolute contribution from mobile sources varies with seasons, the magnitude can be explained by the local atmospheric dynamics. The strong convective activity during the rainy season mixes air masses from the boundary layer and the free Troposphere (convective mixing through updraft and downdraft). As a result, the highest daily-averaged mobile contribution during the rainy season (12 ppm) is half of that during the dry season (22 ppm). We conclude that, with only two stations, future atmospheric inversions would be limited to constraining the emissions from traffic, while missing an important fraction of the energy production sources. Future stations should be deployed in the northern part of MCMA, the outflow of the urban basin. Additionally, a refinement of the fossil fuel emissions inventory (possibly divided into sectors of activities or SNAP's) would be needed to determine the sectoral trends in more detail.

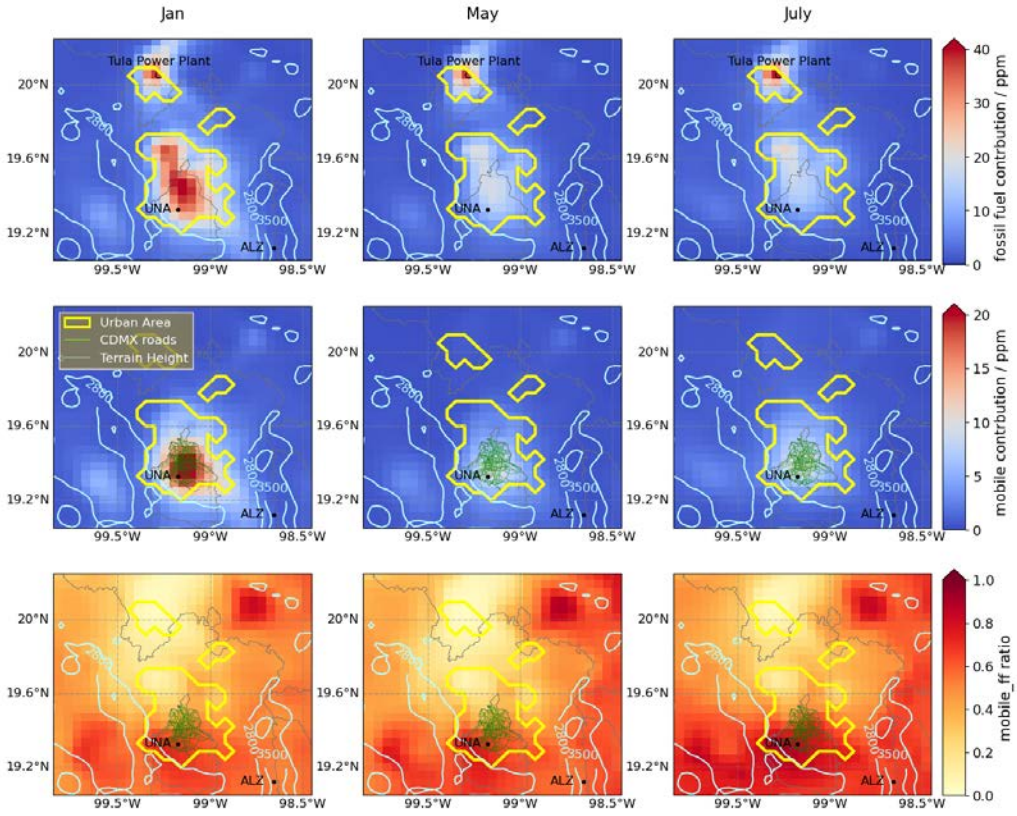


Figure 3.10 *Atmospheric CO_2 concentration absolute contributions from fossil fuel sources (top row), from mobile sources (middle row), and relative contribution of mobile sources to total fossil fuel (bottom row) in the PBL during the dry (left) and rainy (middle and right columns) seasons for the year 2018, in ppm. Topographic contours are indicated in light blue, administrative boundaries in gray, the urban area of MCMA in yellow, and the road network in CDMX in green.*

3.4.2 Biospheric contribution variability

In our study, we used the biogenic flux contribution from the CASA biogeochemical model, optimized using eddy-covariance flux observations from the Ameriflux network (Zhou et al., 2020). While the optimization helps to constrain the model parameters, mediterranean and mountainous ecosystems lack flux measurement stations to constrain the model parameters. However, the use of disturbance forest maps and previous evaluation of CASA suggest that the simulated flux gradients should be representative of the area, with grassland and cropland in the valley north of the city, and forested land in the mountains. We discuss here how the ecosystem diversity, hence

their resulting CO_2 fluxes, translate into CO_2 spatial gradients in our simulated concentrations.

As shown in **Figure 3.9**, in the early morning, the valley of Mexico remains positively-influenced by ecosystems (ecosystem respiration dominates) due to late exposure to incoming radiation compared to the surrounding mountains. Mid-afternoon hours tend to show no-to-little spatial gradients across our simulation domain, corresponding to a homogeneous uptake of carbon. At night, the largest release of CO_2 from respiration takes place in the valleys, west and north of Mexico City. The positive fluxes remain low at both ALZ (above treeline) and at UNA (within the dense urban area).

Despite the fact that the vegetation tends to be more active during the rainy season, we show no significant increase in the biogenic contribution at UNA nor ALZ (**Figure 3.11**), consistent with previous studies (Sánchez-León et al., 2016). During the dry season, the lack of water can prevent vegetation from growing, while the rainy and cloudy conditions reduce the sunlight during the rainy season. Across the urban area (including UNA), vegetation coverage remains low where population density is high. ALZ lies at high altitude (4000 m a.s.l.), above the treeline, hence without strong carbon uptake. In July, ALZ experiences a colder and drier climate than UNA, as shown in **Figure 3.12**. The photosynthesis is higher than that in May, but the largest absorption is only 5 ppm and the net contribution over two weeks is about -1.5 ppm. Overall, we conclude here that both sites correspond to similar biogenic influences during daytime, while early morning fluxes differ significantly.

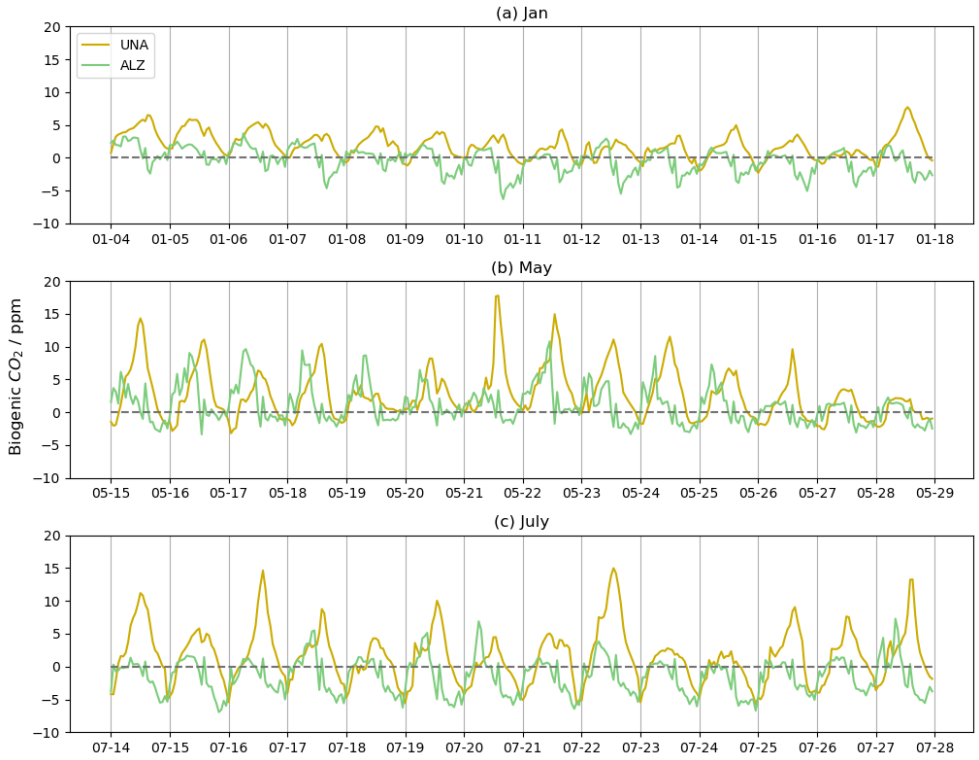


Figure 3.11 *Modeled time series of biogenic contribution to surface CO_2 (in ppm) at UNA and ALZ for the months of (a) January, (b) May and (c) July of 2018, in UTC time (local time=UTC-6).*

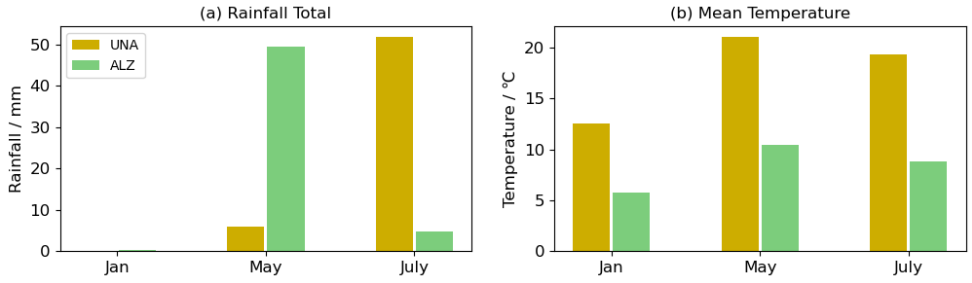


Figure 3.12 *(a) 2-week rainfall and (b) 2-week mean temperature at UNA and ALZ in January, May and July of 2018.*

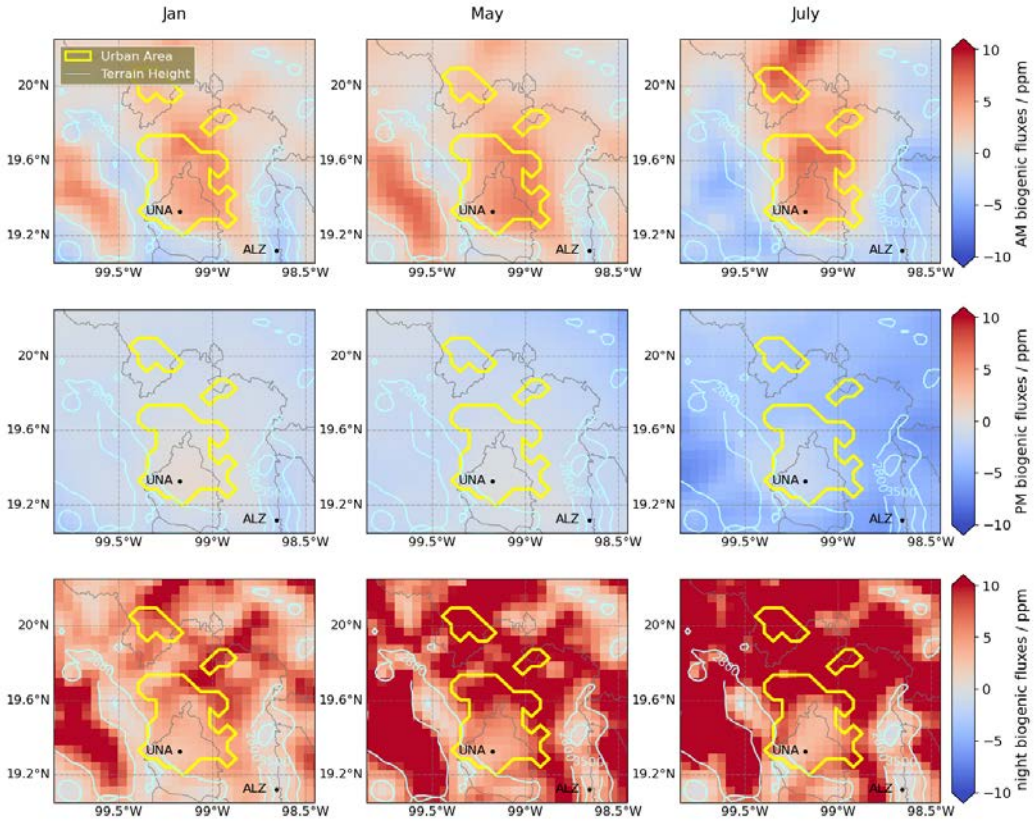


Figure 3.13 *Mean biogenic contribution to surface CO_2 of the three months in (top) AM: 8:00-12:00 local time (middle) PM: 13:00-17:00 local time (bottom) Night: 23:00-4:00 local time, with administrative boundaries (gray lines), high fossil fuel emission urban area over MCMA (yellow lines), two observation station (black point) and terrain height contours (in light blue).*

3.4.3 Optimal network designs for CO_2 inversions

Based on our modeled results, we discuss here the potential use of UNA and ALZ measurements to constrain the city emissions, and we examine potential measurement sites to enhance the current network. We analyze the spatial distribution of background concentrations to monitor large-scale boundary concentrations, and the spatial distribution of fossil fuel signals to maximize the city's signals in our future inversion system. Although human activities vary seasonally, such as heating in winters (not common in the city) and air conditioning in summers, the distribution of fossil fuel signals remains relatively stable, concentrated around the city and the Tula power plant.

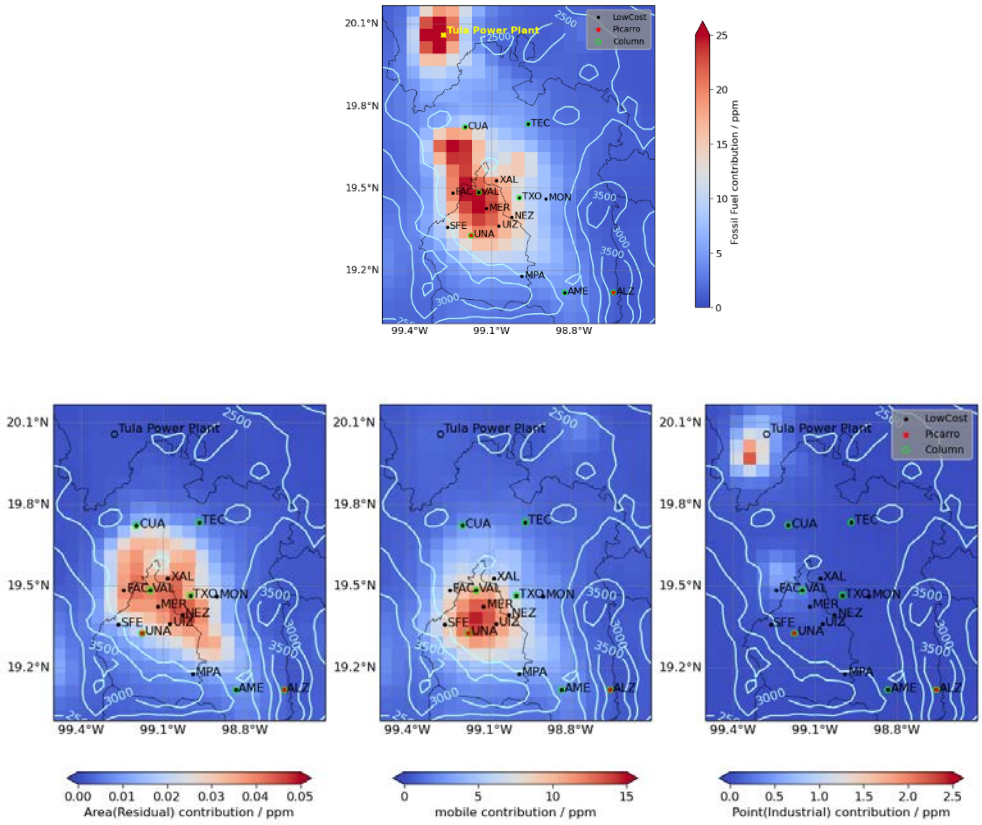


Figure 3.14 *Atmospheric fossil fuel CO_2 concentrations (50 meters agl) simulated by WRF and averaged over the three months of 2018 (top row), with current and potential MERCI-CO_2 observation station locations and elevation contour lines (in light blue). Different emission sectors, area (residential), mobile and point (industrial) are shown in the bottom row.*

The current downtown station (UNA) is located near the city center while ALZ is located further away from the emitting area, at a much higher altitude. A recently-deployed column station (VAL) is located in the city center, co-located with the modeled CO_2 concentration maximum, while CUA, TEC, TXO and AME are on the edge of urban areas, able to capture the background conditions under various wind conditions. For additional measurement locations, the largest unobserved sources are located in the north of MCMA, around and from the Tula power plant where no station has been identified yet. An additional station able to monitor the northern part of CDMX would allow us to monitor emissions from non-transportation sources

(industry and residential sectors), possibly CUA and FAC. Considering our existing stations, UNA is located near the maximum fossil fuel contributions (in terms of CO_2 concentrations) while ALZ observes little to no fossil fuel signals.

To assess the quality of the background station (ALZ), and determine additional locations for complementing our background conditions, we calculated the Spearman correlation coefficient between the modeled background values of UNA with each grid in and around MCMA, similar to Feng et al. (2016). **Figure 3.11** shows the correlation map including (top row) and excluding (bottom row) biogenic fluxes, primarily to determine if biogenic fluxes should be included in our future inversion system as unknowns, or assumed to be directly observed by our background station. We clearly see that biogenic fluxes significantly decrease correlations between the valley and UNA, suggesting that biogenic fluxes are not observed and should be optimized separately. When excluding the biogenic influence, correlations remain high at 0.87 in January, 0.91 in May, and 0.95 in July.

We conclude here that ALZ remains an optimal location for observing the background conditions when excluding the biogenic influence. Potential locations highly-correlated with the UNA background and biogenic fluxes are located north of the city, possibly impacted by the Tula power plant plume in January and May (northerly flow). Therefore, we conclude that a second background station would allow us to evaluate our modeled background concentrations but, based on our current modeling system, would be redundant with the ALZ station.

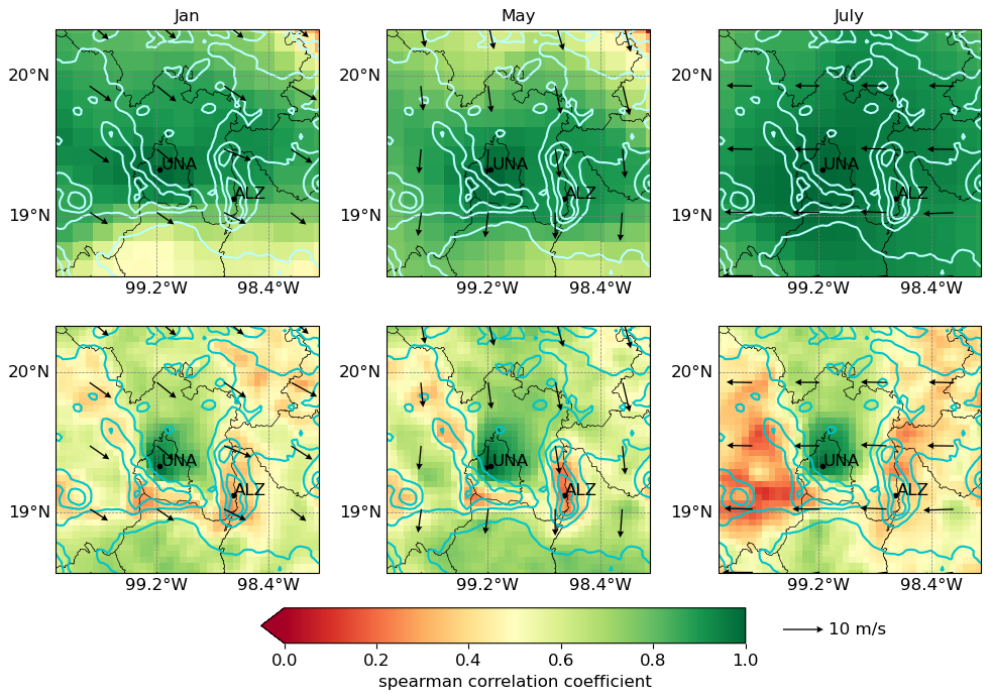


Figure 3.15 *Spearman correlation coefficient maps of background (top) and background + biogenic (bottom) contributions for the three selected months of 2018, with tropospheric wind direction at 3.5km, administrative boundaries, the location of two stations and terrain height contours.*

Considering the various prevailing wind directions (shown in **Figure 3.11**), the fossil fuel signal of ALZ is almost zero in July, but remains 3-5 ppm (**Figure 3.5**) in January and May. If we are going to build another background station, in order to figure out the best position for a background station for MCMA or furtherly for other cities, that is, the place whose background correlation to the urban stations is the highest and the fossil fuel signal is the lowest, we defined an index combining the fossil fuel contribution to local CO_2 concentrations and the background correlation. Considering the range of relative fossil fuel contribution (0% - 30%) and correlation coefficient (0.4 - 1.0), I normalized them to 0.0 - 1.0 first. Then the background index is defined as:

$$background\ index = \frac{corr - ff_{ratio}}{corr + ff_{ratio}}$$

where corr stands for the normalized spearman correlation coefficient

with the urban stations (when there are several, take the average of them, or the weighted average based on the distance from every urban station), and ff_ratio stands for the normalized fossil fuel relative contribution to the CO_2 concentrations of this grid.

This index is defined between 0 and 1. The higher background index means the higher correlation with urban stations and the lower fossil fuel signals. The map of background indexes over MCMA and its surrounding areas in different seasons are shown in **Figure 3.12**. The green areas are the locations suitable for a background station. As **Figure 3.12** shows, ALZ is a relatively good background station. Although it is slightly affected by fossil fuel signals in January and May, there is no obviously better location. Thanks to the basin terrains, the fossil fuel signal at ALZ is no more than 5 ppm, even when the wind blows from the city to ALZ in January.

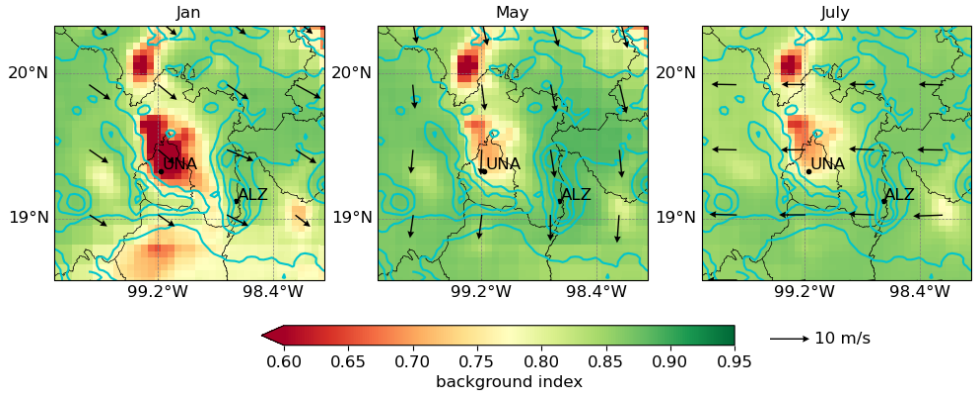


Figure 3.16 *"Background index" maps of MCMA and surrounding areas for the three selected months of 2018, with tropospheric wind direction at 3.5km, administrative boundaries, the location of two stations and terrain height contours.*

Another ideal location for background stations could be to the northeast of UNA, with a higher background index, or to the west of VAL, which is much closer to the city stations and located upwind of the urban area in January and May.

The maps of the background for surface CO_2 (upper row) and XCO_2 (lower row) are shown in **Figure 3.13**. There is no significant difference between surface CO2 and xCO_2 background concentrations in January and May. In July, the background spatial distribution is different in the south, but none of our background station plans concern the south,

where several satellite cities are located. Additionally, the difference is as small as 0.14 ppm, which could be definitely ignored. All the inversion plans for surface CO_2 will also work on XCO_2.

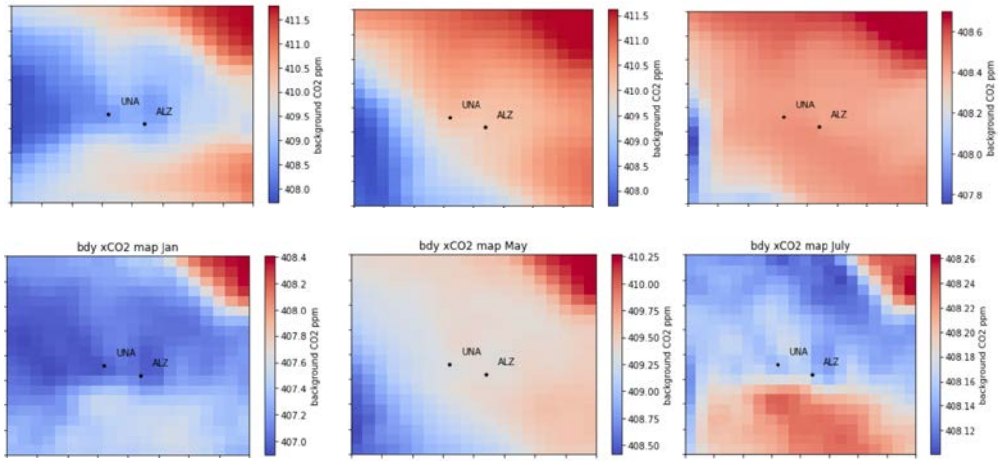


Figure 3.17 *Background map of surface CO_2 (upper row) and XCO_2 (lower row) over MCMA in January, May and July. Observation network is shown in black points.*

REFERENCE

Baylon, J. L., Stremme, W., Grutter, M., Hase, F., & Blumenstock, T. (2017). Background CO 2 levels and error analysis from ground-based solar absorption IR measurements in central Mexico. Atmospheric Measurement Techniques, 10(7), 2425-2434.

Boden, T. A., Marland, G., and Andres, R. J. (2017). Global, Regional, and National Fossil-Fuel CO_2 Emissions, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., USA, https://doi.org/10.3334/CDIAC/00001_V2017.

Butler, M. P., Lauvaux, T., Feng, S., Liu, J., Bowman, K. W., & Davis, K. J. (2020). Atmospheric simulations of total column CO_2 mole fractions from global to mesoscale within the carbon monitoring system flux inversion framework. Atmosphere, 11(8), 787.

Chen, H., Karion, A., Rella, C. W., Winderlich, J., Gerbig, C., Filges, A., Newberger, T., Sweeney, C., and Tans, P. P. (2013). Accurate

measurements of carbon monoxide in humid air using the cavity ring-down spectroscopy (CRDS) technique, Atmospheric Measurement Techniques, 6, 1031–1040. https://doi.org/10.5194/amt-6-1031-2013

Chen, J., Zhao, F., Zeng, N., & Oda, T. (2020). Comparing a global high-resolution downscaled fossil fuel CO_2 emission dataset to local inventory-based estimates over 14 global cities. Carbon Balance and Management, 15(1), 1-15.

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.

Elvidge, C. D., Baugh, K. E., Zhizhin, M., and Hsu, F.-C. (2013). Why VIIRS data are superior to DMSP for mapping nighttime lights, Proceedings of the Asia-Pacific Advanced Network, 35, 62–69, https://doi.org/10.7125/apan.35.7.

Feng, S., Lauvaux, T., Davis, K. J., Keller, K., Zhou, Y., Williams, C., ... & Baker, I. (2019). Seasonal characteristics of model uncertainties from biogenic fluxes, transport, and large-scale boundary inflow in atmospheric CO_2 simulations over North America. Journal of Geophysical Research: Atmospheres, 124(24), 14325-14346.

Frey, M. M., Hase, F., Blumenstock, T., Dubravica, D., Groß, J., Göttsche, F., ... & Pollard, D. F. (2021). Long-term column-averaged greenhouse gas observations using a COCCON spectrometer at the high-surface-albedo site in Gobabeb, Namibia. Atmospheric Measurement Techniques, 14(9), 5887-5911.

García-Reynoso, JA, Mar-Morales, BE, & Ruiz-Suárez, LG (2018). Spatial, temporal and speciation distribution model of the Mexican emissions inventory (base year 2008) for use in air quality modeling (DiETE). International Journal of Environmental Pollution, 34 (4), 635-649. https://doi.org/10.20937/rica.2018.34.04.07.

Hase, F., Hannigan, J. W., Coffey, M. T., Goldman, A., Höpfner, M., Jones, N. B., ... & Wood, S. W. (2004). Intercomparison of retrieval codes used for the analysis of high-resolution, ground-based FTIR measurements. Journal of Quantitative Spectroscopy and Radiative Transfer, 87(1), 25-52.

Jacobson, A. R., Schuldt, K. N., Miller, J. B., Oda, T., Tans, P., Arlyn Andrews, Mund, J., Ott, L., Collatz, G. J., Aalto, T., Afshar, S., Aikin, K., Aoki, S., Apadula, F., Baier, B., Bergamaschi, P., Beyersdorf, A., Biraud, S. C., Bollenbacher, A., … Miroslaw Zimnoch. (2020). CarbonTracker CT2019B. NOAA Global Monitoring Laboratory. https://doi.org/10.25925/20201008

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., et al. (2006). North American regional reanalysis. Bulletin of the American Meteorological Society, 87(3), 343–360. https://doi.org/10.1175/bams-87-3-343

Nassar, R., Napier-Linton, L., Gurney, K. R., Andres, R. J., Oda, T., Vogel, F. R., & Deng, F. (2013). Improving the temporal and spatial distribution of CO_2 emissions from global fossil fuel emission data sets. Journal of Geophysical Research: Atmospheres, 118(2), 917-933.

Olsen, S. C., & Randerson, J. T. (2004). Differences between surface and column atmospheric CO_2 and implications for carbon cycle research. Journal of Geophysical Research, 109(D2). https://doi.org/10.1029/2003JD003968

Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., ... & Tans, P. P. (2007). An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker. Proceedings of the National Academy of Sciences, 104(48), 18925-18930.

Pollard, D. F., Hase, F., Sha, M. K., Dubravica, D., Alberti, C., & Smale, D. (2022). Retrievals of X CO 2, X CH 4 and X CO from portable, near-infrared Fourier transform spectrometer solar observations in Antarctica. Earth System Science Data, 14(12), 5427-5437.

Rella, C. W., Chen, H., Andrews, A. E., Filges, A., Gerbig, C., Hatakka, J., ... & Zellweger, C. (2013). High accuracy measurements of dry mole fractions of carbon dioxide and methane in humid air. Atmospheric Measurement Techniques, 6(3), 837-860.

Rodríguez Zas, J. A., & García Reynoso, J. A. (2021). Actualización del inventario nacional de emisiones de 2013 para la modelación de la calidad del aire en el centro de México. Revista Internacional de

Contaminación Ambiental, 37, 463–487.

Sha, M. K., De Mazière, M., Notholt, J., Blumenstock, T., Chen, H., Dehn, A., ... & Weidmann, D. (2020). Intercomparison of low-and high-resolution infrared spectrometers for ground-based solar remote sensing measurements of total column concentrations of CO 2, CH 4, and CO. Atmospheric Measurement Techniques, 13(9), 4791-4839.

Welp, L. R., Keeling, R. F., Weiss, R. F., Paplawsky, W., and Heckman, S.: Design and performance of a Nafion dryer for continuous operation at CO_2 and CH4 air monitoring sites, Atmos. Meas. Tech., 6, 1217–1226, https://doi.org/10.5194/amt-6-1217-2013, 2013.

WMO (2020), 20th WMO/IAEA Meeting on Carbon Dioxide, Other Greenhouse Gases and Related Tracers Measurement Techniques (GGMT-2019), Jeju Island, South Korea, GAW Rep. No. 255, Geneva, Switzerland.

xESMF: Universal Regridder for Geospatial Data, https://github.com/pangeo-data/xESMF (last access: 24 Jan 2022), https://doi.org/10.5281/zenodo.4294774.

Yver Kwok, C., Laurent, O., Guemri, A., Philippon, C., Wastine, B., Rella, C. W., ... & Ramonet, M. (2015). Comprehensive laboratory and field testing of cavity ring-down spectroscopy analyzers measuring H 2 O, CO 2, CH 4 and CO. Atmospheric Measurement Techniques, 8(9), 3867-3892.

Yver-Kwok, C., Philippon, C., Bergamaschi, P., Biermann, T., Calzolari, F., Chen, H., ... & Wyss, S. (2021). Evaluation and optimization of ICOS atmosphere station data as part of the labeling process. Atmospheric Measurement Techniques, 14(1), 89-116.

Zhou, Y., C. A. Williams, T. Lauvaux, K. J. Davis, S. Feng, I. Baker, S. Denning, & Y. Wei. (2020). A multiyear gridded data ensemble of surface biogenic carbon fluxes for North America: Evaluation and analysis of results. Journal of Geophysical Research: Biogeosciences, 125(2), e2019JG005314. https://doi.org/10.1029/2019JG005314

SUPPLEMENTARY INFORMATION

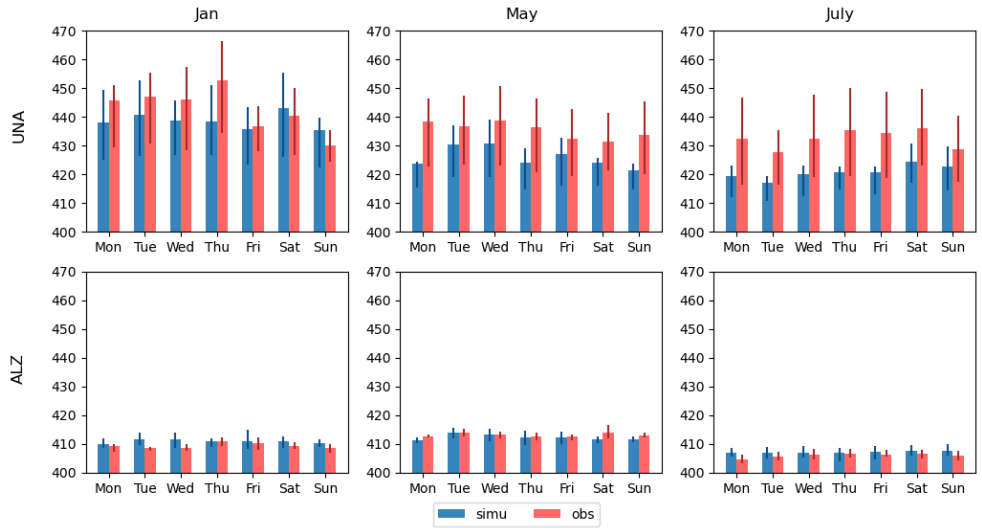


Figure 3.18 *Averaged afternoon CO_2 concentrations (in ppm) with interquartile range for the different days of the week at UNA (upper row) and ALZ (lower row) for the months of January (left column), May (middle), and July (right) of 2018.*

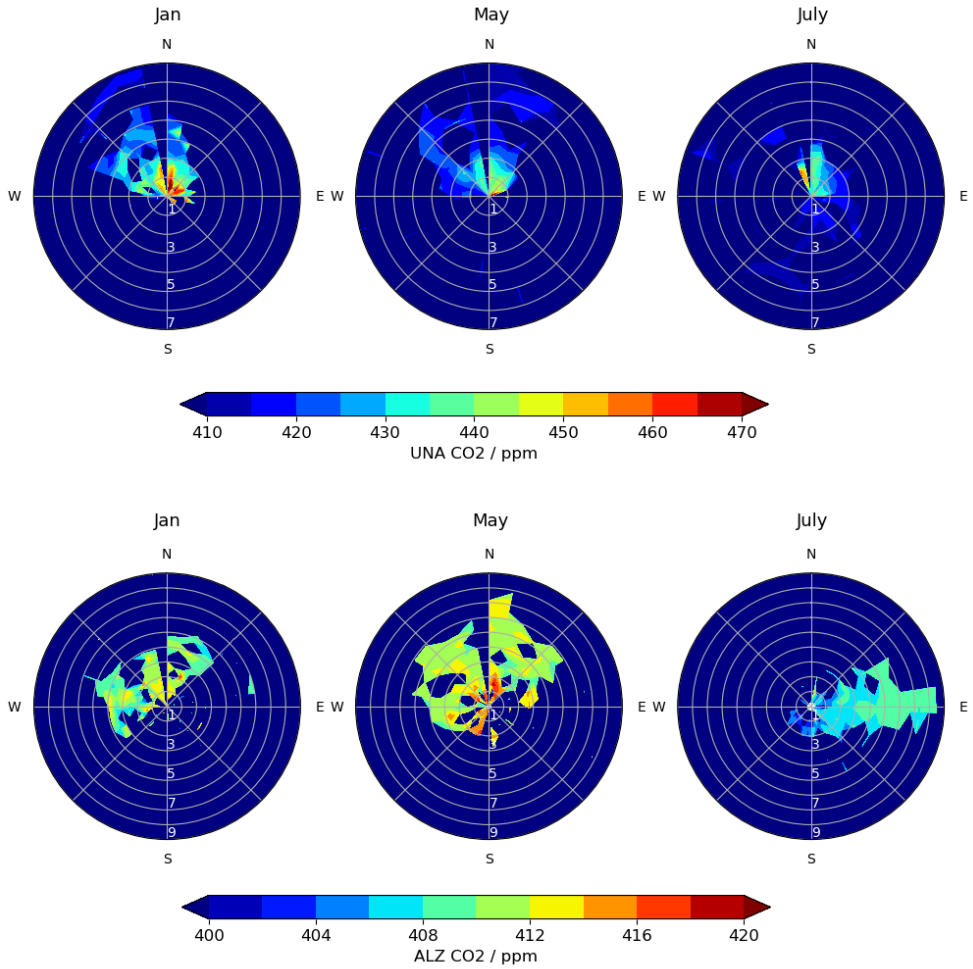


Figure 3.19 *Wind roses of in situ CO_2 concentrations at UNA and ALZ for 2-week period in the month January (left column), May (middle column) and July (right column) of 2018.*

4 CHAPTER IV : BAYESIAN ESTIMATE OF CO_2 EMISSIONS FROM THE MEXICO CITY METROPOLITAN AREA BASED ON 1-YEAR ATMOSPHERIC MEASUREMENTS AND INVERSION MODELING

SUMMARY

The comparisons of CO_2 simulations and observations detailed in Chapter 3 validated the performances of our modeling system under various climate conditions. Based on the same configurations, a 1-year WRF-Chem simulation was performed for the purpose of optimizing the estimates of anthropogenic CO_2 emissions in MCMA through inversions. The main objective of this chapter is to attribute atmospheric CO_2 signatures to fossil fuel combustion, reduce the uncertainties in the prior emissions estimates, with the long-term goal of evaluating the impact of carbon emission reduction strategies in MCMA.

To accomplish our objectives, we performed a 1-year simulation including tracers of fossil fuel signals from UNAM and ODIAC emission inventories, biogenic fluxes from CASA (Net ecosystem exchange, NEE) and background concentrations, as well as carbon monoxide (CO) in and out of MCMA. Fossil fuel signals are separated into 4 time windows: night (00:00-05:00), AM (06:00-13:00), PM (14:00-17:00), and evening (18:00-23:00). Each time window is further divided into the signals in and out of MCMA. Tracers are also included for different emission sectors, including area (residential), mobile (transportation), point (industrial and others). Station-to-station concentration gradients are assimilated with observed gradients to make Bayesian inferences of CO_2 surface fluxes. Errors in the inversion system are defined by observation errors, transport errors (in terms of model-measurement misfit), day covariances and time window covariances.

We show here that various inversion configurations reduced the prior uncertainties of the UNAM emission inventory product in MCMA by over 50%. Uncertainties during the daytime, in the dry season and in MCMA's fossil fuel emissions are most reduced. The posterior estimate

with the lowest gradient error is 61.7 Mt/yr, indicating a 13% underestimation by UNAM and a 6% overestimation by SEDEMA statistics. The months with the most emissions are June to October, which are almost the entire rainy season. The CO_2 emission in the dry season is slightly overestimated by UNAM and the emission in the rainy season is underestimated. There is an abnormal biogenic flux in June, which is improved by our inversion system. Although the underestimation of ODIAC is partly corrected, the error of posterior gradients driven by ODIAC exhibit twice as those driven by UNAM.

We tested data screening of gradients based on statistical outliers or large deviations of observation in our inversions. Both data screening methods improved the posterior gradient mismatch with observations, with the large-observation-deviation exclusion plan providing the lowest mismatch. Block sizes also impact the performance of inversions through day covariances. 1-day, 5-day, 30-day and all-day blocks are tested in our inversion system. The 1-day block provides daily emission scaling factors without the information of day covariances, at the cost of increased posterior uncertainty. The 30-day block provides better estimates than the 5-day, but when the block size increases to 365, the inversion system no longer provides reasonable emission fluxes.

The abstract of this chapter has been presented at AGU 22 and the full-text will shortly be edited and submitted as a manuscript.

Bayesian Estimate of CO_2 Emissions from the Mexico City Metropolitan Area Based on 1-year Atmospheric Measurements and Inversion Modeling

Y XU[1], T. Lauvaux[2], JA García Reynoso[3], M. Grutter[3], J. Lian[1, 4], N. Taquet[3], W. Stremme[3], M.Ramonet[1]

[1] Laboratoire des Sciences du Climat et de l'Environnement (LSCE), IPSL, UMR CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France

[2] Groupe de Spectrométrie Moléculaire et Atmosphérique (GSMA), Université de Reims-Champagne Ardenne, UMR CNRS 7331, Reims, France

[3] Instituto de Ciencias de la Atmósfera y Cambio Climático, Universidad Nacional Autónoma de México, Mexico City, Mexico

[4] Origins S.A.S., Suez Group, Tour CB21, Paris, France

ABSTRACT

Cities contribute more than 70% of global fossil fuel CO_2 emissions. During the past decades, cities and local governments have implemented local climate policies to mitigate their CO_2 emissions. But these policies rely on bottom-up estimated emissions inventories from energy consumption data and statistical records, prone to large uncertainties. Top-down estimation based on Bayesian inversion approaches offers a complementary solution to reduce these uncertainties. Mexico City, located in the Valley of Mexico, is one of the largest cities in the world (22 million inhabitants). Thanks to a French-Mexican collaboration (Mexico City's Regional Carbon Impacts - MERCI-CO_2), a network of CO_2 in-situ (Cavity Ring-Down Spectrometer, CRDS - Picarro) and ground-based column (solar-absorption Fourier Transform InfraRed, FTIR - EM27/Sun) sensors have been deployed across the Mexico City metropolitan area. A series of atmospheric CO_2 simulations between April 2018 and March 2019 were performed with the Weather Research and Forecasting model coupled with Chemistry (WRF-Chem). The mean absolute errors for UNA and ALZ in PM during the

simulation year are 4.95 ppm and 2.54 ppm and the mean errors are -1.56 ppm and 0.22 ppm. The simulations, together with the surface CO_2 and XCO_2 measurements, provide the opportunity to produce top-down estimation of CO_2 emissions in and around Mexico City. Through the assimilation of CO_2 concentration gradients between the urban (UNA) and the rural station (ALZ), the inversion system produced CO_2 emission estimates from different sources - background, anthropogenic (including residential, transportation and industrial) and biogenic. The median of the posterior emissions in domain 2 (the central Mexico) estimated by various inversion configurations is 104.2 Mt/yr in the dry season and 145 Mt/yr in the rainy season, indicating a 1.2% reduction and a 16% increase from the prior UNAM 1-km CO_2 emission inventory. The inversions reduced the errors by 10% - 60%. The uncertainties in MCMA, during daytime and in the dry season were reduced most.

4.1 INTRODUCTION

Cities are reported to account for more than 70% of the global fossil fuel carbon emissions (Seto et al., 2014). With the global process of urbanization, they are playing a more and more important role in mitigating CO_2 emissions. Major cities all over the world are taking measures to control their carbon emissions. To reach the goal of carbon neutrality by 2050, the City of Paris aims to promote an 80 % reduction in the local carbon footprint compared to 2004 levels, winning the UN Global Climate Action Awards for Climate Leaders in 2021 (UN Climate Change, 2021). In New York City (NYC), over 125 participants of the NYC Carbon Challenge pledge have cut their annual carbon emissions by more than 600,000 t and 13 leading participants are projected to achieve carbon neutrality within their NYC portfolios by 2030 (MOCEJ, 2022). The mayor of London has set a series of policies on heating and transportation for London to be net zero carbon by 2030, too (London Government, 2022). Tokyo sets its goal of net zero CO_2 emissions by 2050 under ''Zero Emission Tokyo Strategy '' at the U20 Tokyo Mayors Summit in 2019 (TMG, 2019).

In order to get a complete knowledge of the emission patterns and evaluate the effectiveness of actions on climate change in the cities, many cities have compiled their own high-resolution emission

inventories, e.g. Hestia at building/street scale for Los Angeles, Indianapolis, Salt Lake City and DC-Baltimore (Gurney et al., 2012), Origins.Earth (https://www.origins.earth) for Ile-de-France (covering the Greater Paris Metropolis), London Atmospheric Emissions Inventory (LAEI) (LAEI, 2021). However, these emission inventories rely on self-reported statistical data, prone to large uncertainties and inconsistencies, especially under-estimation (Gately et al., 2017; Gurney et al., 2021).

Top-down Bayesian inversion offers a solution for the uncertainties of emission inventories in megacities. It assimilated the concentration gradients from observation and simulation to adjust the prior emissions and reduce the uncertainties (Bréon et al., 2015). Lauvaux et al. (2016) found the posterior emissions in Indianapolis over the 2012–2013 dormant season is 20% higher than the prior Hestia emissions. The inversion in Paris decreases the annual emissions from 51.9 to 37.4±2.1 Mt CO_2/yr (Staufer et al., 2016). The uncertainties in Cape Town emission inventories have been reduced by 28.0% (2.3 - 50.5%) after inversion (Nickless et al., 2018).

Additional trace species could be used to separate the attribution of emission sectors of atmospheric CO_2. Carbon monoxide (CO) is known as a co-emitted species with CO_2 in fossil fuel combustion. It is widely used as a trace species of CO_2 source analysis, because it is easy to measure and there have been numerous studies indicating the sources and sinks of atmospheric CO (Wong et al., 2015;.Nathan et al., 2018; Nalini et al., 2022).

The French-Mexican project **Me**xico City's **R**egional **C**arbon **I**mpacts (MERCI-CO_2) aims to reduce the uncertainties of CO_2 emission in Mexico City. Mexico City (Ciudad de México, CDMX), located in the Valley of Mexico, is one of the largest megacities in the world. The Mexico City Metropolitan Area (MCMA, or Greater Mexico City) has a population of about 21.6 million inhabitants, ranking fifth in the world, and projected to increase by 11.6% from 2018 to 2030 (UN, 2018). As population grows, the annual CO_2 emissions in MCMA have risen at an average rate of 9.5% per year during six consecutive years, reaching 66.0 Mt in 2018. The main source of CO_2 emission in MCMA is transportation, responsible for around 60% of the total CO_2 anthropogenic emissions (SEDEMA, 2012; SEDEMA, 2021). Rapid

urbanization leads to rapid motor vehicle growth. Almost half of motor vehicles registered for circulation in the country run in MCMA. According to the official statistics, the number of vehicles (including public service vehicles) has reached 14.7 million by the end of 2020, which has increased by 102% during the last decade. (INEGI, 2021). Due to the topography of the enclosed basin, it's not conducive to atmospheric dispersion, leading to the accumulation of CO_2.

The city has the objective of reducing at least 10% of CO_2 emissions by 2024 compared to 2018. This goal has been half-reached by strategies on tree planting, water resource management, zero waste, public transportation promotion, adoption of standards for motor vehicles and clean energies (Government Headquarters of Mexico City, 2021). For 2021-2030, MCMA is continuing to reduce 65.2 Mt CO_2 emissions, mainly on traffic sources (SEDEMA, 2021), and approaching net zero emissions after 2030 (Government Headquarters of Mexico City, 2022).

Thanks to MERCI-CO_2, a dense network of CO_2 measurements has been deployed in MCMA. In this study, we estimated anthropogenic and biogenic CO_2 emissions in MCMA through a Bayesian inversion system based on the observations from this network and a 1-year atmospheric transport simulation by the Weather Research and Forecasting model coupled with Chemistry (WRF-Chem). The simulated concentration gradients are assimilated with those from the observation network, to optimize the prior local carbon emissions, and to constrain the uncertainties. Additionally, CO/CO_2 ratios are involved to validate the contribution of traffic sources.

4.2 METHODS

4.2.1 CO_2 observation network

Two high-precision Cavity Ring-Down Spectrometers (CRDS) (model G2401) commercialized by Picarro (Crosson et al., 2008; Rella et al., 2013) were installed in MCMA and started continuous measurements in mid-2014. The urban station UNA is located on the campus of National Autonomous University of Mexico (*Universidad Nacional Autónoma de México*, UNAM), which is 2280 m a.s.l, while the background station ALZ is at the top of the Altzomoni mountain, which is about 4000 m a.s.l.

The CRDS instruments were calibrated at the end of 2018, with a calibration scale consisting of three cylinders prepared by NOAA/ESRL, and traceable to the international WMO scale. The calibration was propagated backwards to provide measurements of CO_2, CH_4, CO and H_2O every several seconds, with an estimated CO_2 uncertainty of approximately 0.1 ppm over the whole year 2018. Due to the local humidity levels, the uncertainties related to water vapor correction is estimated as 0.2 ppm for UNA and 0.05 ppm for ALZ.

Apart from surface CO_2 measurement, there are also dry air column-averaged mole fractions of CO_2 (XCO_2) measurements performed at the two stations, as **Figure 4.1** shows. Both stations were equipped with the solar absorption Fourier Transform Infrared (FTIR) spectroscopy. The ALZ station was equipped with a high resolution (0.02 cm^{-1}) FTIR spectrometer (model IFS-120/5HR, Bruker) in 2012 and the UNA station was equipped in 2016 by a mobile low resolution (0.5 cm^{-1}) EM27/SUN spectrometer in 2016, providing continuous total column measurements of CO_2, CO and CH_4. Inter-calibration was applied to the XCO_2 measurements from both stations, with the side-by-side measurements during several weeks at ALZ by a 3-sigma criterion.

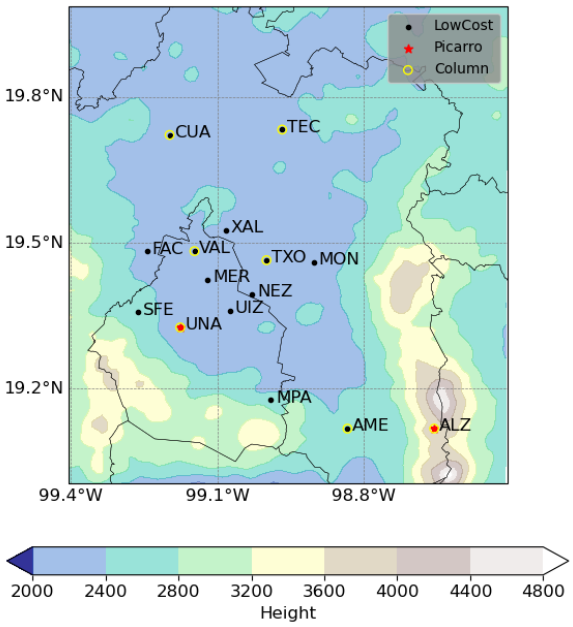


Figure 4.1 *MERCI-CO_2 measurement network (current and future stations) with topography and administrative divisions in MCMA.*

The network shown in **Figure 4.1** includes 7 column sensors that started to collect XCO_2 measurements from September, 2020 and 13 future low-cost medium-precise (LCMP) sensors developed by LSCE. The inversion based on the column sensors (October 2020 to April 2021) was made by Ke et al. (2023) and was compared with my inversions (April 2018 to March 2019.) on the overlapped months.

4.2.2 Model and prior emissions

The atmospheric transport model applied in this study is WRF-Chem (Skamarock et al., 2008, Grell et al., 2005) version 3.9, using the Planet Boundary Layer (PBL) scheme Mellor–Yamada–Nakanishi–Niino (MYNN) (Nakanishi and Niino, 2009), coupled with the MYNN atmospheric surface layer scheme and the single layer urban canopy model UCM. WRF-Chem was configured with one-way two-nested domains, whose resolutions are 15 km and 5 km (shown in **Figure 3.1**). There are 49 vertical layers in each domain, the first of which is around 16 meters above ground level and the second is around 45 meters. Such configurations have been evaluated over the MCMA in various seasons in the last two chapters.

The local emission inventory are developed at National Autonomous University of Mexico (*Universidad Nacional Autónoma de México*, UNAM) (Garcia-Reynoso et al., 2018; Rodriguez-Zas et al., 2021, https://github.com/JoseAgustin/emis_2016), based on Mexico's National Emissions Inventory data (SEMARNAT, 2021). The hourly high-resolution UNAM emission product provides the gridded emission inventories for CO, CO_2, NOx and other species in two resolutions: 3-km for national scale and 1-km for several states (*estados*) and regions, including MCMA. The 3-km version is interpolated to my domain 2 by a python package xESMF, using the mass-conserving method "conservative" (xESMF, https://doi.org/10.5281/zenodo.1134365). The UNAM emission products have three source sectors: area (residential), mobile (transportation) and point (industrial and others). The emission factor ratio of CO and CO_2 for the mobile and point sectors are 0.05 and 0.006, while that for the area sector varies a lot. As the area only contributed less than 2% in total emissions, we set the CO/CO_2 emission ratio for mobile and non-mobile sources as 0.05 and 0.006.

For the larger 15-km domain which exceeds the boundary of Mexico,

a widely-used high-resolution global fossil fuel emission map product based on nightlight satellite data, the Open-source Data Inventory for Anthropogenic CO_2 (ODIAC) (Oda et al., 2018), is interpolated spatially from 1-km map to our 15-km domain by xESMF-conservative and temporally with scaling factors for weekdays and diurnal cycle (Nassar et al., 2013). The same product is also interpolated to the 5-km domain 2, for the sensitive test of emission products compared to the UNAM emission inventory.

The biogenic CO_2 flux used in WRF-Chem is the 3-hourly Carnegie Ames Stanford Approach (CASA) model (Zhou et al., 2020), linear interpolated to hourly. The CASA model had been optimized by eddy-covariance flux data and the optimized configuration of Net Ecosystem Exchange (NEE) was also interpolated to our domains by xESMF - conservative.

Instead of the default fixed boundary condition used in CO simulation, the influence of CO_2 background concentration was simulated with the global model CarbonTracker 2019B (Peters et al., 2007; Jacobson et al., 2020). The 3°×2° CO_2 mole fractions from CarbonTracker were interpolated to our domain 1 in the mass-conserved way Feng et al. (2019) used over North America, with pressure correction.

Due to the data availability of CarbonTracker 2019B, our simulations and inversions are limited to the period until 2019-03-30. We ran a one-year WRF-Chem simulation for a one-year inversion in MCMA, from 2018-03-30 00:00:00 to 2019-03-30 00:00:00, UTC time.

4.2.3 Inversion configuration

Bréon et al. (2015) developed a linear Bayesian inversion method to adjust the prior knowledge of anthropogenic and biogenic CO_2 fluxes. The method relies on an atmospheric transport model with boundary conditions and the CO_2 concentration measurements. It defines an emission scaling vector $\boldsymbol{x}$, which gathers all the scaling factors for different emission variables. $\boldsymbol{x}_A$ is the posterior estimates, while $\boldsymbol{x}_B$ is the prior estimates independent from the observations and the model operator $\mathbf{H}$. The modeled gradient $\mathbf{H}\boldsymbol{x}_B$ is compared with the observed CO_2 gradient vector $\mathbf{y}$. To optimize the scaling factors, we take advantage of the prior uncertainty covariance matrix $\mathbf{B}$, the linear emission-to-model operator $\mathbf{H}$, and the observation error covariance

matrix **R**. The optimal solution is (Enting, 2002; Tarantola, 2005):

$$\boldsymbol{x_A} = \boldsymbol{x_B} + (\boldsymbol{B^{-1}} + \boldsymbol{H^T R^{-1} H})^{-1} \boldsymbol{H^T R^{-1}} (\boldsymbol{y} - \boldsymbol{H x_B})$$

and the posterior uncertainty covariance matrix of $\boldsymbol{x}_A$ is

$$A = (B^{-1} + H^T R^{-1} H)^{-1}$$

Due to the large dimensions of the inverse problem, we divided our inversion over the 1-year time period into 5-day blocks. Inversions are made for every block, and generate a group of scaling factors for different variables (for time windows, sources, masks, etc.) of CO_2. Only the gradients in the afternoon (local time 14:00-17:00, when the PBL is assumed to be well mixed) are assimilated in the inversion system. Now that the MCMA lies in the Valley of Mexico and the atmosphere is usually well mixed in the basin, wind direction is not taken as a criterion of data.

4.2.3.1 Control vector $\boldsymbol{x}$ and the prior estimate of emission flux $\boldsymbol{x_B}$

The control vector x contains the scaling factor of the boundary concentrations, total CO_2 emissions in each time period (the length of the time periods depend on various plans), and biogenic fluxes. All the variables are optimized in the inversions.

In order to specify the CO_2 emission in different time windows, sectors and MCMA / non-MCMA areas, a series of tracers and masks are applied in the model. There are 4 time windows for each day, divided by local time: AM (06:00-13:00), PM (14:00-17:00), evening (18:00-23:00), and night (00:00-05:00). The PM period starts from 14:00 instead of 12:00, because the mixing layer in MCMA is not stable in the early-afternoon (Osibanjo et al., 2022). The three sectors from UNAM emission inventory, area (residential), mobile (transportation) and point (industrial and others), are traced separately to identify the source sector attribution. An urban mask was applied to separate the emissions from MCMA and other areas. The urban mask (shown as the pink area in **Figure 2.1**, right panel) is not the administrative boundary of MCMA, but defined by the fossil fuel emission area with a threshold of 10^5 mol/km^2/h. It covers the northern part of Mexico City, the western part of Mexico State, and the industrial areas to the northwest of MCMA, including a large power plant (Tula power plant) to the north of Mexico City. There are also some smaller cities in the southern part of the domain that exhibit a large fossil fuel signal, e.g. Toluca de Lerdo

and Puebla Zaragoza, but they are more than 50 km away from the MCMA and their emission won't have significant impacts on the atmospheric CO_2 gradients. Therefore, they are not included in the urban mask.

The configuration of reference inversion (Plan 0) contains only one fossil fuel variable in ***x***, which means it creates one scaling factor for the 5-day block. Plan 1 expands to 5 variables, which contains the scaling factors for each day. Plan 2 includes the four time windows on the basis of Plan 1. Plan 3 includes the three factors on the basis of Plan 1. For Plan 1-3, the ''a'' after the number means the sensitivity test on emission inventories (UNAM vs ODIAC), ''b'' means the sensitivity test on block sizes (5 days vs 1 days) and ''c'' means urban area mask of MCMA was applied in the inversion. Plan 4 is the same as Plan 1 but for CO, with a fixed background mole fraction of 90 ppb (based on the minimum value of CO observation, similar to Linán-Abanto et al., 2021), which is consistent with the global background value 50-120 ppb (WMO, 2000). Plan 4c is also Plan 4 with the MCMA mask. Since CO is simulated with , Plan 4a attempted to improve the background concentration, because it is shown that the background of the two stations may have a difference of around 10 ppb. The prior scaling factors are put as zeros.

4.2.3.2 *Measurement gradients y*

y contains the measured gradients during PM (local time 14:00-17:00) that are assimilated with the simulations to constrain the prior emissions. Measurements at the two stations cover the whole year, with a data availability of 99.93 %. Measurements between UTC time 2018-12-12 15:00:00 to 2018-12-17 19:00:00 are excluded from our study, because there was a wild fire causing unusual fluxes (c.f section 4.3). The original measurements are averaged to hourly, consistent with the model output. The standard deviation of the measurements in one hour is keeped as a part of the observation errors (matrix **R**), which is averagely 3.72 ppm for UNA and 0.64 ppm for ALZ.

4.2.3.3 Model operator H

The linear operator **H** is the link between emissions and measurements, including the process of the spatial-temporal distribution of emission, the atmospheric transport, and the sampling from WRF-Chem:

$$x \mapsto y = H_{samp} H_{trans} H_{emi-map} x$$

in which H_{samp} and H_{trans} remain the same, and $H_{emi-map}$ is tested for two different emission products UNAM and ODIAC in Plan 0a. Different from Bréon et al. (2015), I include the boundary conditions in **H** because it is also affected by the transport and sampling process.

4.2.3.4 Prior uncertainty covariance matrix B

The prior relative uncertainty of the variables for anthropogenic emissions in one inversion block (5 days) is assumed as 20%, similar to Staufer et al. (2016). The uncertainties matrix **B** is divided into 3 independent blocks: boundary conditions (bdy), anthropogenic fossil fuel emissions (ff) and biogenic fluxes (bio). The prior relative error of bdy is set as 1% and that of bio is set as 50%, based on our confidence to the performance of CarbonTracker and CASA in MCMA.

Inside the ff block, the prior covariances between days and time windows are set up as in Bréon et al. (2015) and Staufer et al. (2016). The distribution of the correlations between consecutive days are defined by an exponentially-decaying function. In our study, the error correlation extends up to 7 days. That means the correlation of error for the concentrations can last for 7 days at most. The current error of concentration is independent of the error a week before and a week after. Correlations between successive time windows are set as 1, 0.4, 0.2, and 0.4, the same with Staufer et al. (2016). That is, the same time window of two days has 100 % of the covariance defined as above; the neighboring time windows have 40 %, and the alternate time windows only have 20 %. **Figure 4.2** shows the distribution of uncertainty covariance matrix B with covariance of days (Plan 1) and covariance between time windows as well as days (Plan 2).

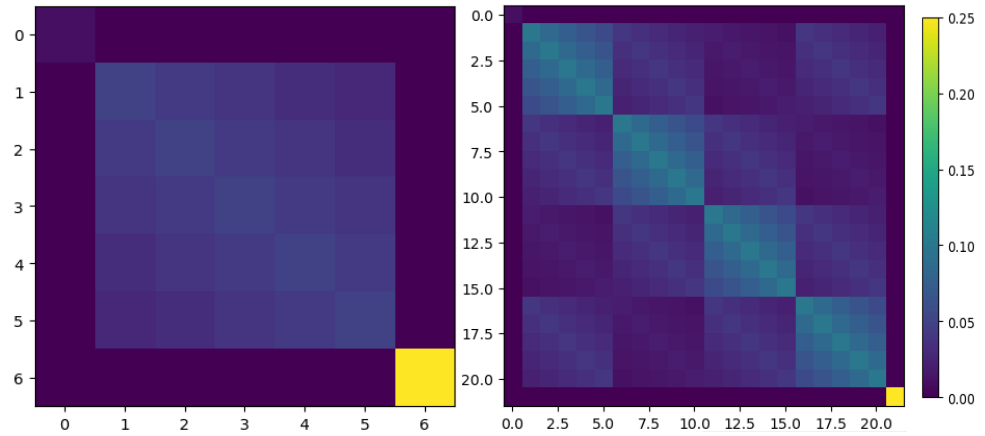


Figure 4.2 *Distribution of uncertainty covariance matrix* ***B*** *with covariance of days (Plan 1, left) and covariance between time windows as well as days (Plan 2, right).*

4.2.3.5 Observation error covariance matrix R

As we have already pointed out, the standard errors of the hourly measurements are a part of the observation error covariance. Other parts include model error, which is represented by the model-measurement differences on gradients. The model-measurement gradient PM average is around 2 ppm in our study, similar to the 3-ppm value Bréon et al. (2015) and Staufer et al. (2016) used. Owing to the fact that the MCMA measurements vary more than those in Paris, this model-measurement gradient mismatch is on the same scale with the standard errors of measurements. We combined them to form the observation error covariance matrix R. The uncertainties are assumed independent. **R** is thus a diagonal matrix.

As for CO, the model-measure misfit is 8.88 ppb, much smaller than the measured uncertainties is 66.56 ppb. We took the measured uncertainties to build the observation error covariance **R**.

4.3 RESULTS

The one-year simulation was compared with the observation, as shown in **Figure 4.3**. Our simulation reproduced the CO_2 concentrations from April 2018 to March 2019. It captured the minimum values in late May 2018, late September 2018, mid-December 2018 and January to March 2019. However, the model underestimated the CO_2 mole fractions at UNA. Though the simulations reach the peaks at the same time with

the observations, these peaks were usually underestimated in the prior modeled concentrations, e.g. the peak values in April 2018. Almost all the observations above 430 ppm were missed in the simulation. We also noticed this persistent underestimation in April and May 2018 no longer exists in 2019. The mean absolute errors for UNA and ALZ in PM during the simulation year are 4.95 ppm and 2.54 ppm and the mean errors are -1.56 ppm and 0.22 ppm, which are sufficient for inversion.

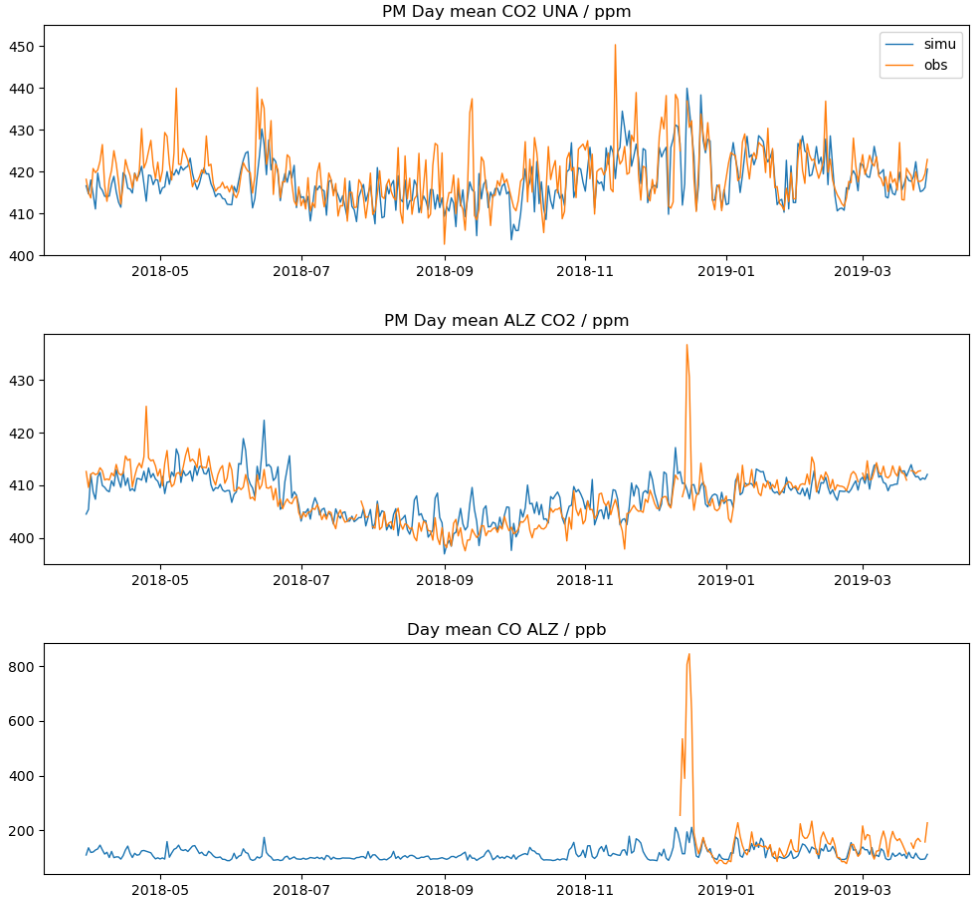


Figure 4.3 *Time series of CO_2 daymean during PM at UNA and ALZ, with CO daymean at ALZ, in the one-year simulation from 2018-03-30 to 2019-03-29.*

We noticed there was an extreme high peak value in December 2018 at ALZ, which is not captured by the model. We then checked the corresponding CO measures. There is also a peak of CO at the same time. The CO/CO_2 ratio at the peak is about 4%, indicating the source

of the peak is a very efficient burning, most likely a wildfire. The period with this unexpected combustion source from UTC time 2018-12-12 15:00:00 to 2018-12-17 19:00:00 is thus excluded in our inversions.

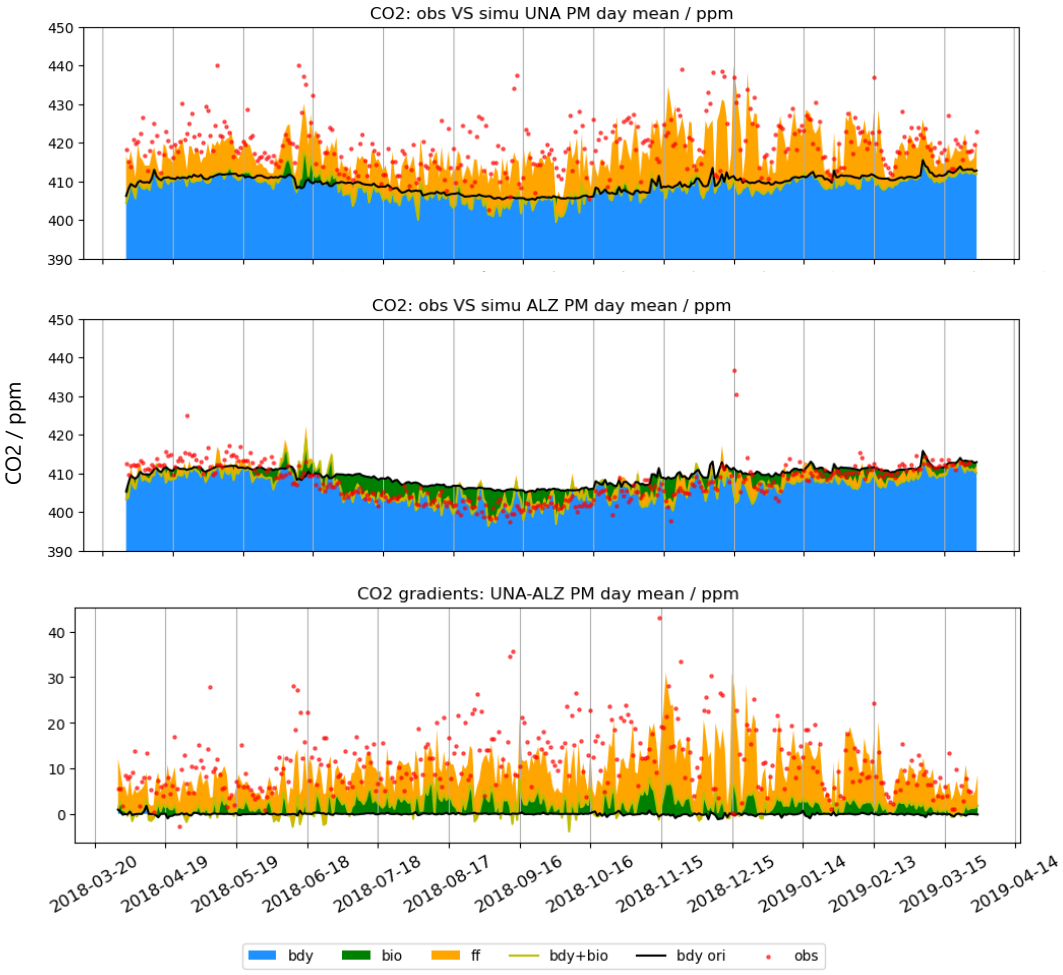


Figure 4.4 *Sources of CO_2 concentrations and gradients (UNA-ALZ) daymean during the one-year simulation from 2018-03-30 to 2019-03-29. The modeled concentrations were separated into three components: boundary inflow (bdy - in blue), biogenic fluxes (bio - in green) and fossil fuel emissions (ff - in orange). The contributions from the boundary conditions are indicated by the black line and the combined contribution of boundaries and biogenic fluxes (bdy+bio) by the yellow line.*

In addition to time series and errors, the modeled components of atmospheric CO_2 are analyzed as well, shown in **Figure 4.4**. Consistent with the conclusions we made on the background station, the background CO_2 concentration at ALZ is almost identical with that of

UNA during PM, and the fossil fuel signal at ALZ is as low as 0.63 ppm during the rainy season and 1.37 ppm during the dry season. The gradients of UNA-ALZ are not affected by boundary conditions, and mainly driven by the fossil fuel signal in the urban area as well as the difference of biogenic fluxes.

4.3.1 Inversion results

Generally, our inversions reduced the errors of prior emission inventories by 10% - 60%. The errors in MCMA, of the dry season and during PM were reduced the most. According to the climate records, we use 2018-05-28 as the start of the rainy season, and 2018-10-10 as the start of the dry season. The errors are listed in **Table 4.1**. Plan 2 series (with covariances between days and time windows) performs the best across all error metrics used for evaluation.

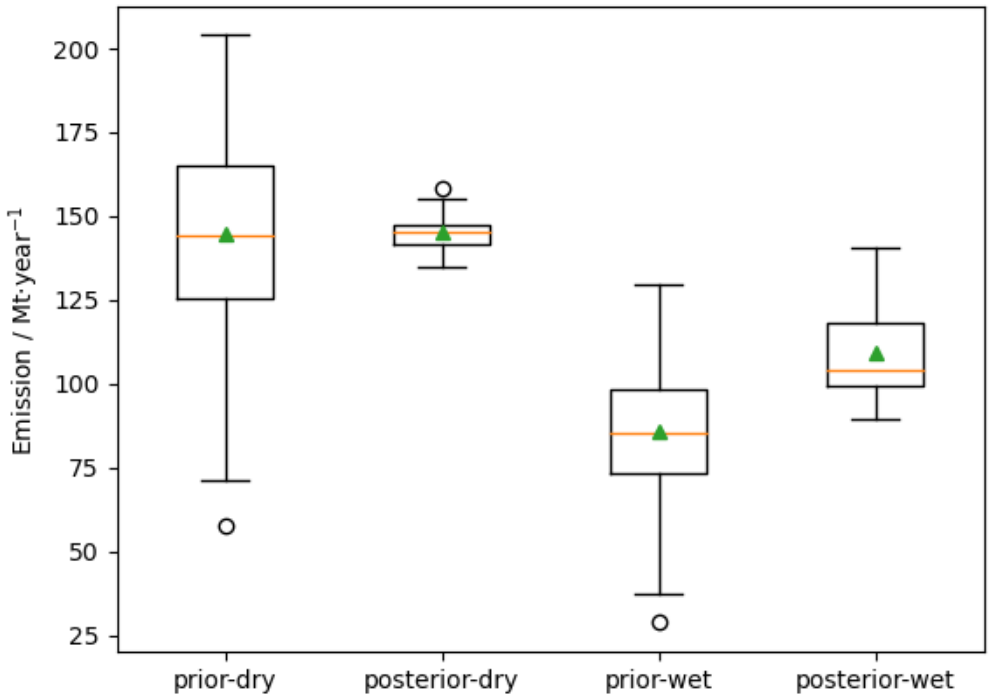


Figure 4.5 *Boxplot of the prior and posterior emissions of the 12 inversion plans of anthropogenic CO_2 in domain 2 during dry season and wet season. Medians are shown in orange lines and means are shown in green triangles.*

In terms of median, the UNAM emission inventory product overestimated the emissions of the dry season by 1.29 Mt, and underestimated the emissions of the rainy season by 19.95 Mt. The

details of the prior and posterior emissions and reduced errors of the inversions are shown in **Table 4.2** and **Figure 4.5**. Plan 0b (block size=1 day) has the lowest errors among all Plan 0 series (shown in **Table 4.1**), but the uncertainties shown in **Table 4.2** are the highest.

Mt/y	Bias				StDev				RMSE				R^2			
No.	**dry**		**wet**		**dry**		**wet**		**dry**		**wet**		**dry**		**wet**	
	prior	**post**	**prior**	**post**	**prior**	**post**	**prior**	**post**	**prior**	**post**	**prior**	**post**	**prior**	**post**	**prior**	**post**
0	-1.09	-1.60	-3.31	-2.41	6.92	5.85	8.25	6.49	7.01	6.07	8.89	6.92	0.27	0.43	0.01	0.22
0a	-3.47	-2.38	-5.35	-2.81	6.76	6.10	8.10	6.77	7.60	6.55	9.71	7.33	0.25	0.39	0.00	0.18
0b	0.33	-1.06	-1.90	-1.99	8.03	4.46	8.46	5.32	8.04	4.58	8.67	5.68	0.32	0.74	0.11	0.53
0c	-1.02	-2.09	-3.29	-2.19	6.93	5.85	8.25	6.45	7.01	6.21	8.88	6.81	0.27	0.44	0.01	0.25
0d	-0.59	-1.61	-2.41	-1.83	6.29	5.15	6.61	5.38	6.31	5.40	7.04	5.69	0.20	0.33	0.04	0.19
1	-1.09	-1.50	-3.31	-2.10	6.92	4.63	8.25	5.78	7.01	4.86	8.89	6.15	0.27	0.63	0.01	0.36
1c	-1.02	-1.50	-3.29	-2.13	6.93	4.62	8.25	5.80	7.01	4.86	8.88	6.18	0.27	0.63	0.01	0.35
2	-1.02	-1.22	-3.29	-1.82	6.93	4.28	8.25	5.28	7.01	4.45	8.88	5.59	0.27	0.68	0.01	0.45
2c	-1.02	-1.23	-3.29	-1.86	6.93	4.31	8.25	5.31	7.01	4.48	8.88	5.63	0.27	0.68	0.01	0.45
3	-1.02	-1.55	-3.45	-2.24	6.92	4.69	8.25	5.90	7.00	4.94	8.94	6.31	0.27	0.62	0.01	0.33
3c	-1.02	-1.55	-3.45	-2.28	6.92	4.70	8.25	5.93	7.00	4.95	8.94	6.35	0.27	0.62	0.01	0.33

Table 4.1 *Mean Bias, Standard Deviation (StDev), Root Mean Square Error (RMSE) and coefficient of determination (R^2) of the prior and posterior simulation of the CO_2 concentration gradients (UNA-ALZ) from various inversion Plans in dry and wet seasons. Every variable is separately colored with the green-yellow-red colormap, in which green stands for small errors and red stands for large errors.*

No	Emission	Test	Prior - dry / Mt	Posterior - dry / Mt	Prior - wet / Mt	Posterior - wet / Mt	Reduced error
0	UNAM	ff+bio+bdy	146.25	148.36	84.22	113.55	62.2 %
0a	ODIAC	0-emission	133.67	157.94	80.08	140.19	51.8 %
0b	UNAM	0-block size	146.25	141.67	84.22	100.28	37.7 %
0c	UNAM	0-mask	146.25	134.74	84.22	98.68	30.2 % (71.4%)*
0d	UNAM 5%-95%	0-extreme values	146.25	141.45	84.22	117.34	59.8 %
1	UNAM	day covariance	146.25	155.18	84.22	122.43	58.1 %
1c	UNAM	1-mask	146.25	147.44	84.22	96.58	25.1 % (57.8%)*
2	UNAM	time window	146.25	146.92	84.22	112.14	41.1 %
2c	UNAM	2-mask	146.25	145.64	84.22	100.76	23.4 % (48.5%)*
3	UNAM	sectors	146.68**	144.96	83.07**	104.17	20.1 %
3c	UNAM	3-mask	146.68**	144.97	83.07**	89.58	10.1 % (11.8%)*
4	UNAM CO	CO	1.97	1.88	1.21	1.21	13.9 %
4a	UNAM CO	4-bdy err	1.97	1.88	1.21	1.21	13.5 %
4c	UNAM CO	4-mask	1.97	1.88	1.21	1.21	26.3 % (31.8%)*

Table 4.2 *Summary of the results for all inversions, with prior and posterior emissions in domain 2 during both seasons and reduced uncertainties (including variances and covariances) of fossil fuel emissions.*

* Percentage values in brackets are the reduced errors in the MCMA urban mask.

** Due to the random errors of the emission program, the sum of the three sectors has a minor discrepancy to total fossil fuel emissions in other plans.

4.3.1.1 The reference inversion

The reference inversion (Plan 0) is done with basic configurations: all the data in one block are treated as a whole. It generated only one fossil fuel scaling factor with all the 20 gradients in one block, leading to a lower uncertainty but a higher misfit. As shown in **Figure 4.6**, Plan 0 assimilated the peak in December 2018 and corrected many of the model-measurement misfits, but failed to improve the peak in May 2018 and the mismatch in September to October 2018. The performance of other inversions are compared with Plan 0 in the sensitivity tests (c.f section 4.4.1).

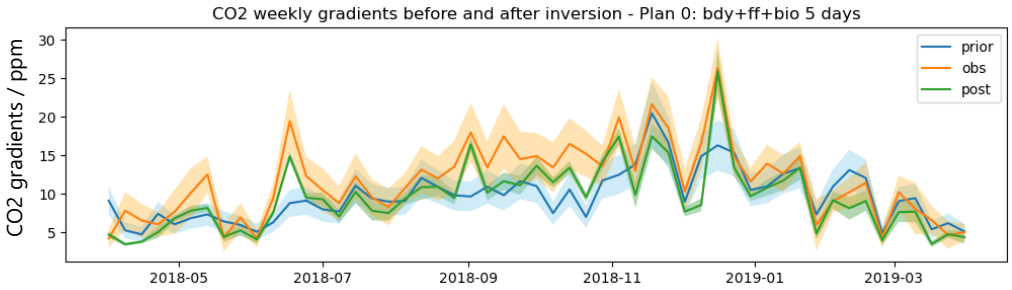


Figure 4.6 *The weekly average of CO_2 gradients (UNA-ALZ) before and after inversion Plan 0 compared to observations, with shade of InterQuartile Range.*

On the other hand, we take the best plan, Plan 2c with day covariances and time window covariances as well as MCMA mask, to investigate the seasonal variation, MCMA and non-MCMA contribution, and diurnal cycle of CO_2 emissions in central Mexico (Domain 2) and in MCMA. Unless otherwise specified, the following analysis is based on Plan 2c.

4.3.1.2 Seasonal variation

Figure 4.7 shows the emission fluxes adjusted by scaling factors. The uncertainty of fossil fuel fluxes in MCMA is reduced by around 50%. The yearly emission nearly remains constant, while the temporal distribution is improved. The posterior emission during the dry season of 2018 in MCMA (31.8 Mt) is in agreement with the emission estimation in 2020 by Ke et al. (2023, *in preparation*) using XCO_2 inversion (31.6 Mt). There are several emission peaks during the rainy season, probably related to drying, dehumidification and other needs. MCMA is warm (6 - 22 ℃) in the coldest month of the year (January).

Winter heating doesn't seem to increase CO_2 emissions.

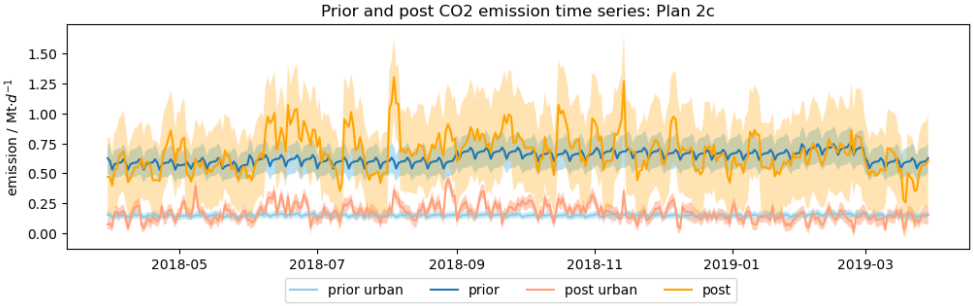


Figure 4.7 *Prior and posterior CO_2 daily emissions in domain 2 and in MCMA from inversion Plan 2c (with day covariances and time window covariances, as well as MCMA mask), with uncertainties.*

We summed the emissions in months in **Figure 4.8**. From October to February, UNAM emission inventory overestimated the emissions in central Mexico, while from March to September, it underestimated the emissions. These overestimates and underestimates are more severe in MCMA. From June to October, UNAM underestimates nearly 30% of the monthly emissions. Emissions in the rainy season are higher than those of the dry season, and were underestimated more than those of the dry season.

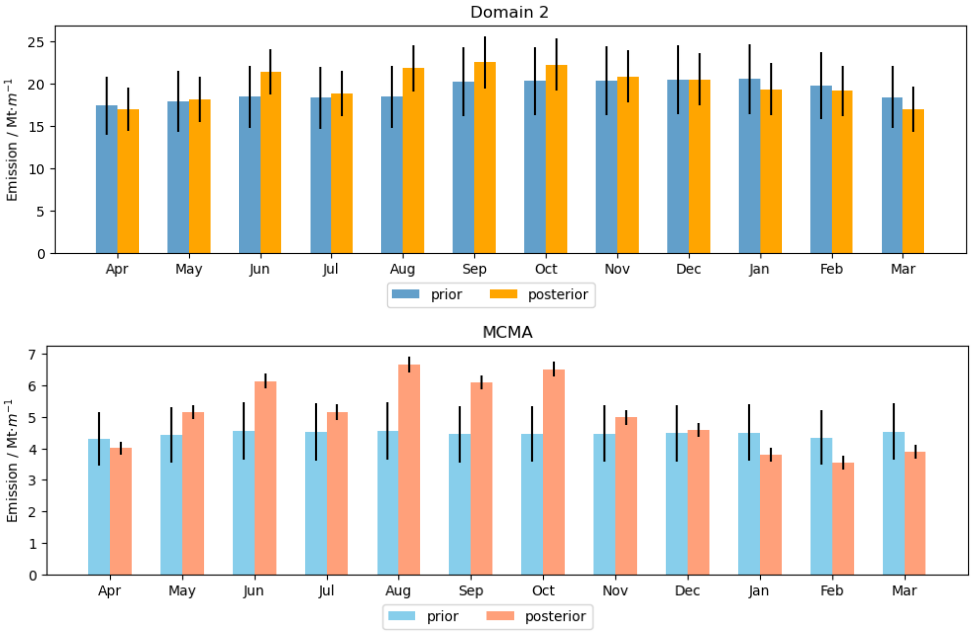


Figure 4.8 *Prior and posterior CO_2 monthly emissions in domain 2 and in MCMA from inversion Plan 2c (with day covariances and time window*

covariances, as well as MCMA mask).

4.3.1.3 MCMA and non-MCMA

The total of CO_2 emissions in MCMA has increased from 53.62 Mt/yr to 61.74 Mt/yr after inversion Plan 2c, more closely related to the SEDEMA emission report of 66.0 Mt/yr. As shown in **Figure 4.9**, the uncertainty of emissions in MCMA is reduced by almost 50%, even though the uncertainty out of MCMA is only reduced by about 3%.

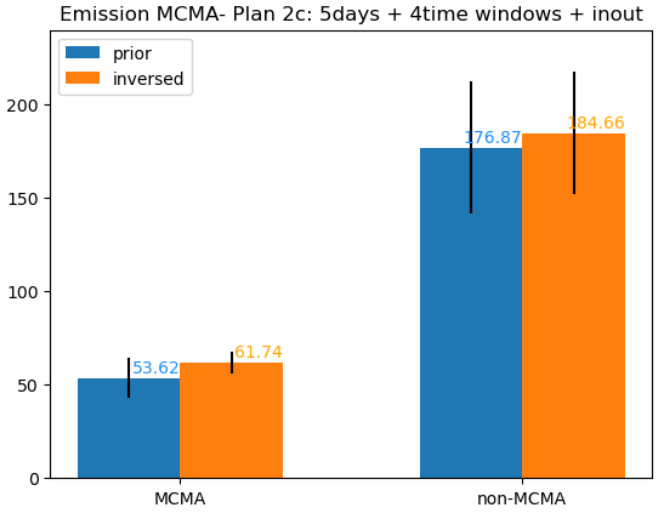


Figure 4.9 *CO_2 yearly emission (in Mt) in and out of MCMA urban mask before and after inversion Plan 2c, with uncertainty bar.*

4.3.1.4 Time windows

Since our inversion was performed only with data from the afternoon, the emission uncertainties of PM are improved most. **Figure 4.10** shows the reduced error by emission Plan 2c. The matrix for non-MCMA areas (top left) almost remains the same, indicating the non-MCMA areas have not improved much. On the other hand, the matrix for MCMA shows significant improvement.

Among the four blocks for evening, night, AM and PM, the first two are not improved much either (around 12% - 16% errors reduced), given that the transports and the concentration patterns are very different after sunset. Conversely, AM and PM emission uncertainties are reduced by 75.4% and 85.1%. More specifically, the UNAM emission inventory underestimated the fluxes during morning time by 4.9 %, and the fluxes in the afternoon by 18.6%, as shown in **Figure**

4.11. There was not much optimization in emissions in the evening and at night.

The posterior diurnal and weekly cycles of CO_2 emissions are shown in **Figure 4.12**. The emissions and errors out of MCMA hardly changed, while the emissions on weekdays and in the afternoon increased. We note that there is a sharp increase in PM in the diurnal cycle. In theory, there should not be a plateau in daytime emissions. It is exactly the time window "PM" in our inversion system, suggesting an underestimation of evening rush hours in UNAM.

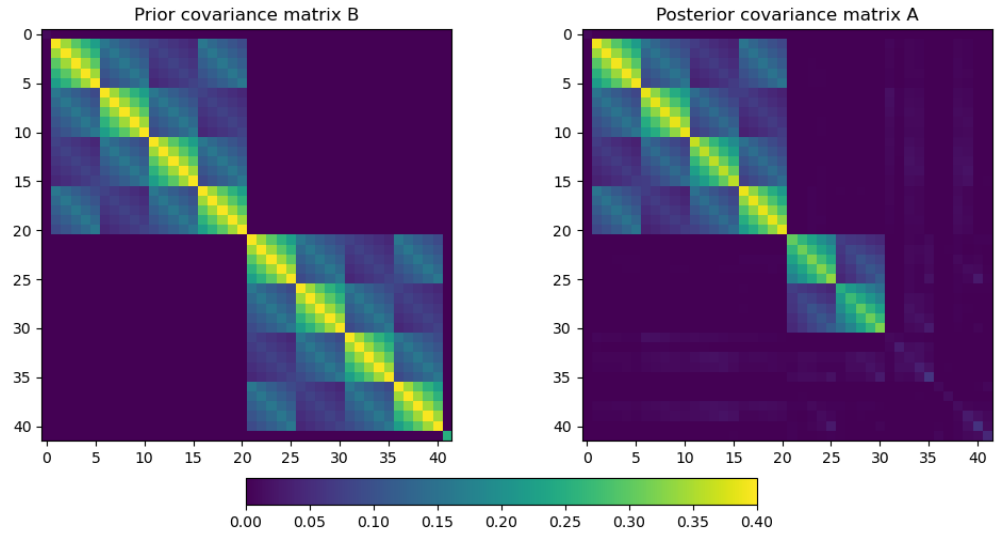


Figure 4.10 *Prior and posterior uncertainty covariance matrix of Plan 2c (for the last block from 2019-03-25 to 2019-03-30). The top-left 20×20 matrix is for non-MCMA and the bottom-right one is for MCMA. In each 20×20 matrix, the diagonal is corresponding to the 20 time windows in the 5-day block. The four blocks are the 5 days' evening, night, AM and PM, from top-left to bottom-right.*

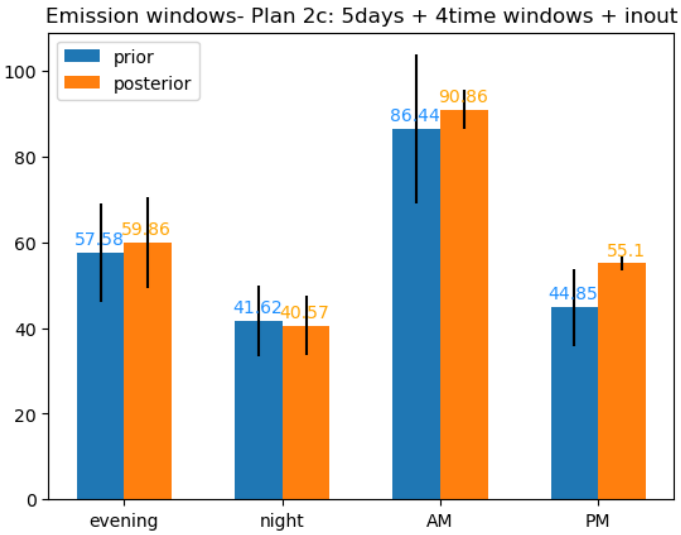


Figure 4.11 *CO_2 yearly emission (in Mt) of the four time windows (evening, night, AM, PM) before and after inversion Plan 2c, with uncertainty bar.*

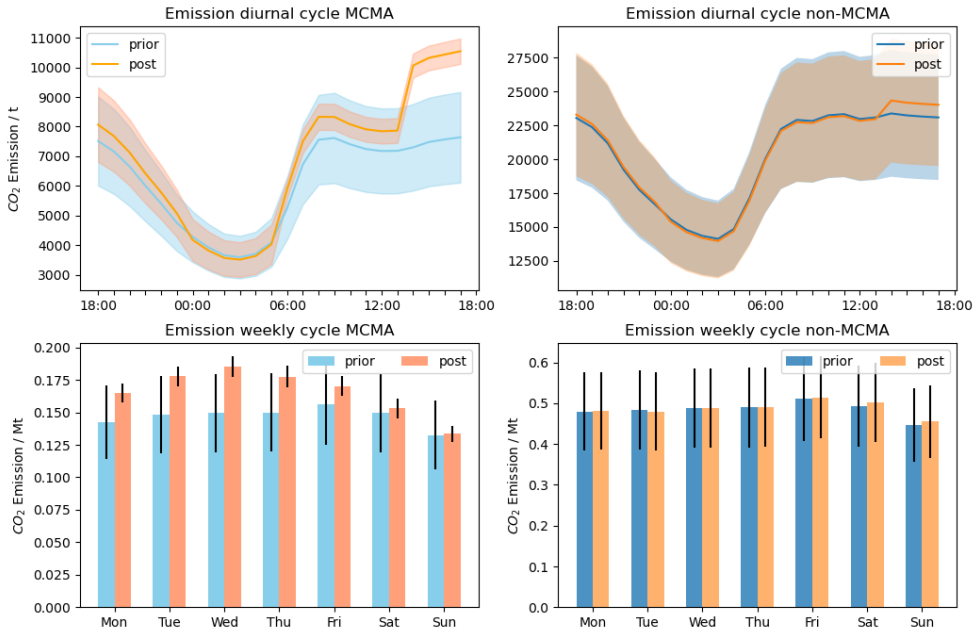


Figure 4.12 *Prior and Posterior averaged CO_2 emission diurnal and weekly cycles in and out of MCMA by inversion Plan 2c.*

4.3.2 Sectoral and biogenic contributions

4.3.2.1 Bioflux

There was an abnormal peak in the biogenic fluxes from CASA in June.

According to the seasonal variation of vegetation in Mexico, June in the rainy season is the time when plant growth is relatively active. The biogenic flux should be negative, as in May and July.

Our inversions corrected this large bias. Plan 2c (The top panel in **Figure 4.13**) decreased the NEE fluxes in June, but they are still positive and different from May and July. Plan 0 series (excluded 0b), Plan 1 series and Plan 3 series are better, with a posterior flux behavior similar to May and June, but they also reduced the intake of CO_2 in the growing season (the prior negative fluxes in July). Plan 0b differs greatly from the others, manifesting as severe fluctuations in posterior emissions, due to the small sample size and high uncertainty. Regardless of the circumstances, the abnormal prior biogenic fluxes led to more uncertainty in the inversions. If the NEE could be calibrated, the inversion on fossil fuel fluxes would be probably improved.

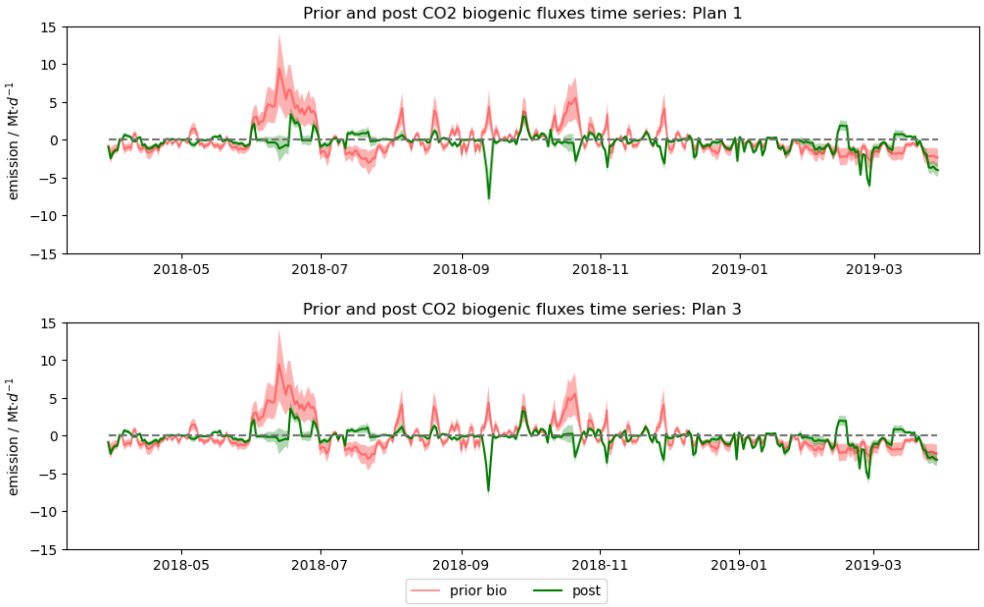


Figure 4.13 *Prior and posterior CO_2 daily biogenic fluxes in domain 2 from inversion Plan 2c (with day covariances and time window covariances, as well as MCMA mask), Plan 0, Plan 0b, Plan 1 and Plan 3, with uncertainties.*

4.3.2.2 Mobile contribution

Plan 3c, with sector tracers and the MCMA urban mask, is used for the emission sector attributions in and out of MCMA. According to **Figure 3.2**, the emissions from area sources of UNAM emission inventory are more than two orders of magnitude smaller than other sectors. This involves the definition of area sources and point sources. There are 8 vertical layers in UNAM emission inventory, and all sources above the first layer are divided into point sources. Here, we examined the mobile source, which is believed only in the ground layer.

Figure 4.14 shows the prior and posterior emissions of mobile sources in and out of MCMA urban mask. The year total of mobile source emissions was increased from 23.93 Mt to 29.44 Mt, about 47.7% of total emissions. SEDEMA reports that mobile sources contribute 57% - 65% of yearly emission in MCMA administrative boundaries, which should be around 40 Mt. In order to verify whether the 10 Mt underestimation is caused by the difference of the MCMA urban mask and MCMA administrative boundary, or it is the systematic bias in the inversion, the CO/CO_2 ratio is calculated with prior and posterior

concentrations to figure out the attribution of mobile sources (c.f. 4.3.3).

As expected, the mobile source is relatively stable, its variability being noticeable in the weekly cycle. However, our posterior mobile emissions show an amplitude of approximately 50%, e.g. in May 2018 shown in **Figure 4.14**. Moreover, the uncertainties in MCMA are only reduced by 11.8% and the uncertainty of the mobile source in MCMA is only reduced by 5%.

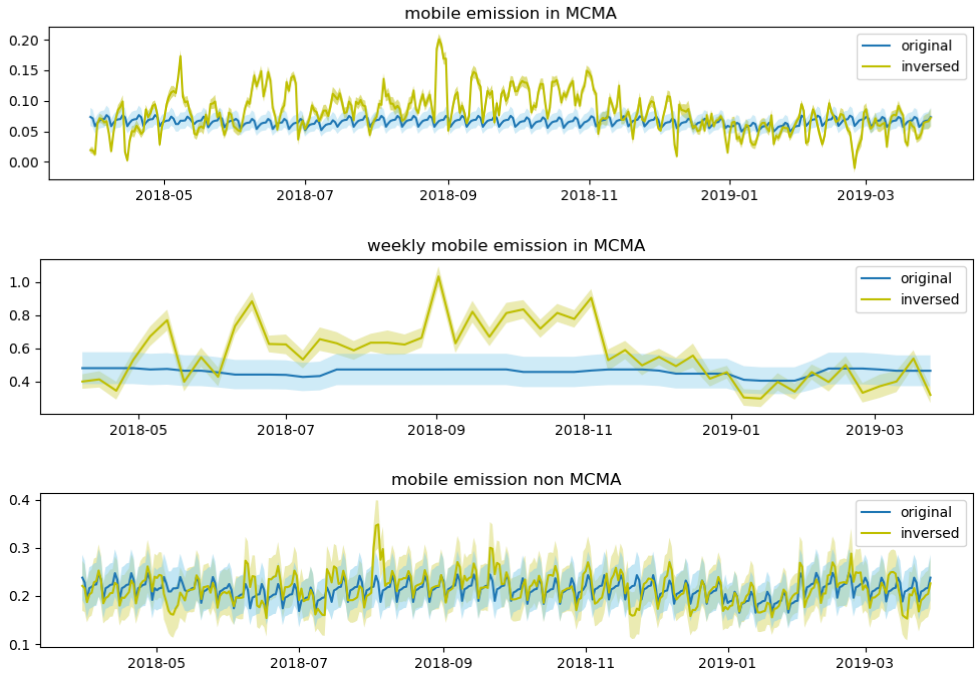


Figure 4.14 *Prior and posterior CO_2 daily and weekly emissions (unit: Mt/d) in MCMA and daily emissions in non-MCMA area from inversion Plan 3c (with day covariances and emission sector tracers as well as MCMA mask), with uncertainties.*

4.3.3 CO/CO_2 ratio

The observation records of CO started in mid-December 2018. After the aforementioned wildfire was excluded, there remained 100 days for inversion until March 30, 2019. The model-observation mismatch in CO gradients is smaller than that of CO_2, and most of them are fixed in the inversion.

We introduced a 50% high uncertainty of background conditions in Plan 4b, since the prior background concentration 90 ppb causes a 10-

ppb overestimation on the minimums at ALZ (as shown in **Figure 4.15**). Unexpectedly, Plan 4b didn't have significant improvement on the inversion. The difference between the two posterior concentrations is only -16 ppb to 15 ppb in hourly data, and not obvious enough to be shown in the daymean in **Figure 4.16**. A 10 ppb difference is rather negligible compared to the UNA-ALZ gradients up to 500 ppb.

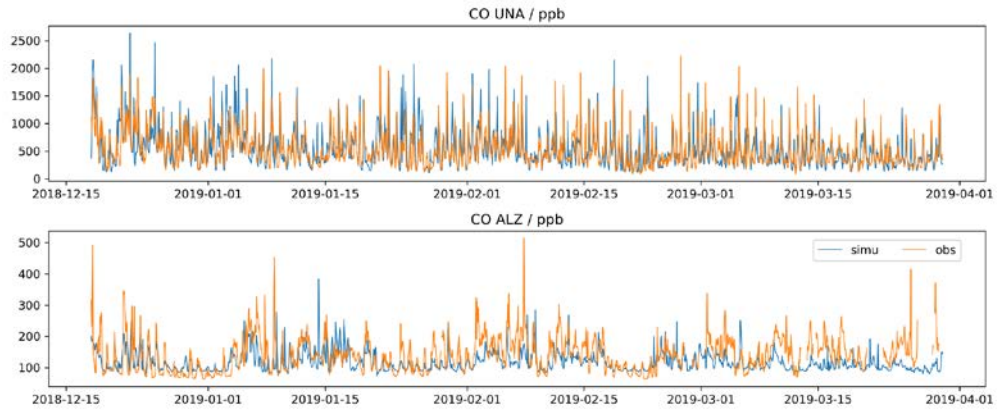


Figure 4.15 *Time series of CO observation and simulation from our one-year modeling.*

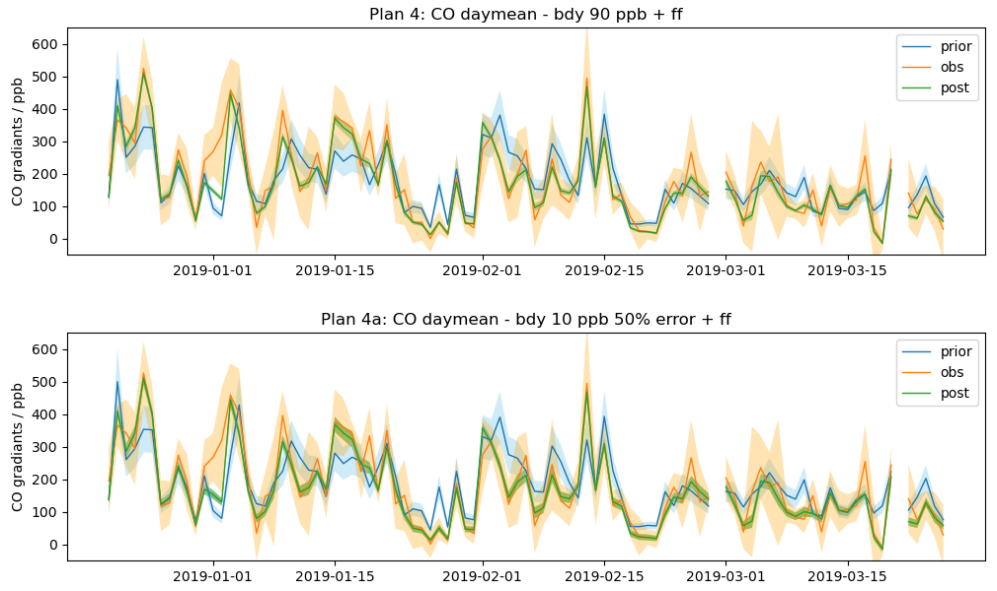


Figure 4.16 *The daymean of CO gradients (UNA-ALZ) before and after inversion Plan 4 and 4a compared to observations, with shade of InterQuartile Range.*

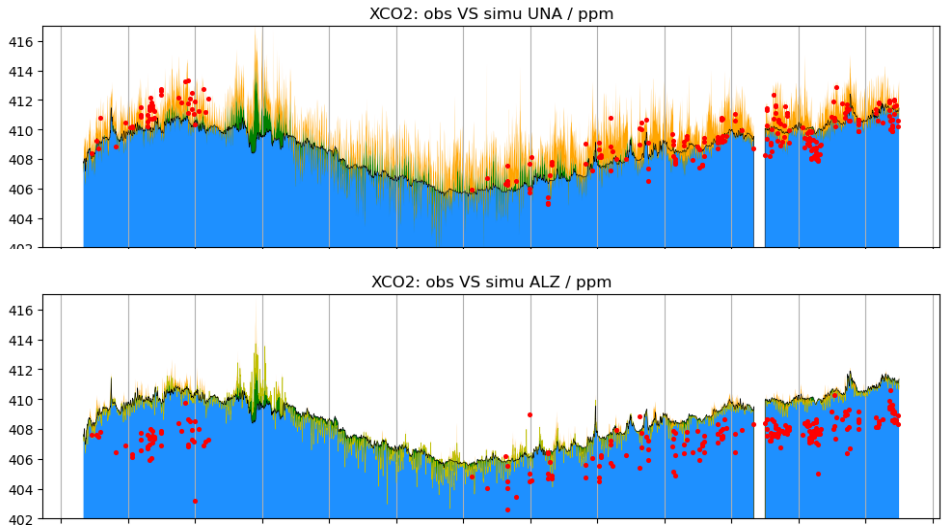
Plan 4c optimized the CO fluxes in and out of MCMA. The prior CO

emission in MCMA is 0.14 Mt/yr and the posterior is 0.11 Mt/yr. Compared with the prior and posterior CO_2 gradients in MCMA mentioned in section 4.3.2.2, the CO/CO_2 ratio has decreased from 21.35 to 17.02. The mobile contribution to CO_2 emissions, as inferred from the CO/CO_2 ratio using an emission factor ratio of 50 and 6 ppb/ppm for mobile and non-mobile sources, has decreased from 34.9% to 25.05%. This value is not only much lower than the 57% - 65% reported by SEDEMA, but also lower than the 47.7% contribution from mobile sources reported by UNAM.

The failure of CO/CO_2 ratio inversion indicates that there may be some random error in the CO or CO_2 inversion. A more robust solution is to perturb the emissions to create high-engine and low-engine scenarios, and see if the end of inversions converges, similar to Nathan et al. (2018), Lauvaux et al. (2020) and Nalini et al. (2022).

4.3.4 XCO_2

XCO_2 mixing ratio was driven from the WRF-Chem. **Figure 4.17** shows the agreement between the simulations and observations. The mean error is 0.95 ppm and the mean absolute error is 1.96 ppm. When it comes to the PM period, the errors are -1.58 ppm (mean error) and 4.44 ppm (mean absolute error). Contrary to surface CO_2, the errors are larger during PM and at ALZ. Except for that, the data availability of XCO_2 is poor. There is no available PM observation from March to September after data screening for quality control, mainly due to the weather conditions.



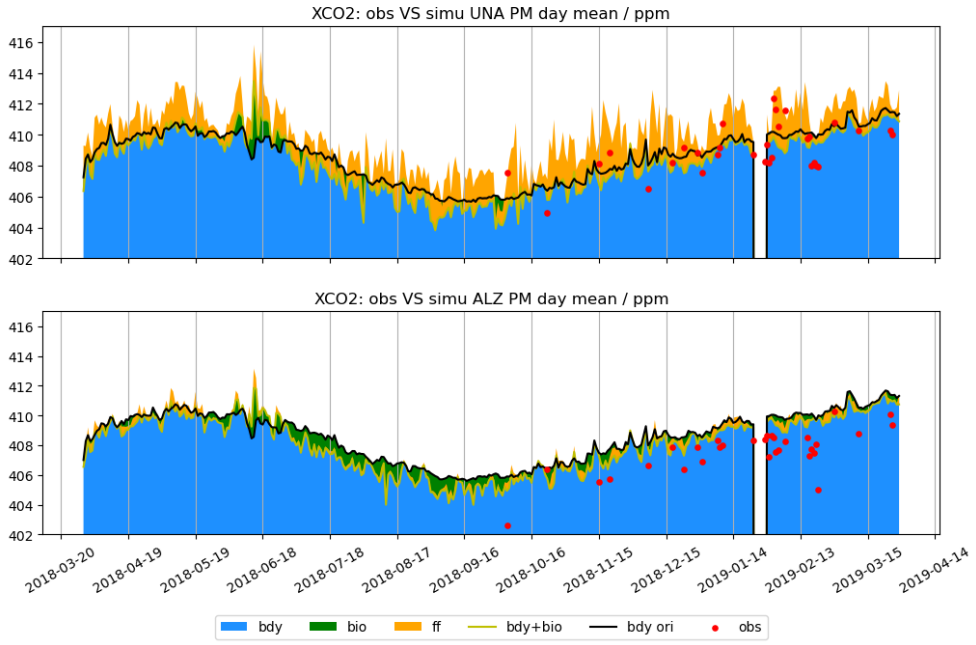


Figure 4.17 *Observed and modeled XCO_2 concentration time series (upper row) and PM daymean (lower row) in ppm at both UNA and ALZ stations for the simulated year 2018-03-30 to 2019-03-30, in UTC time. The modeled concentrations were separated into three components: boundary inflow (bdy - in blue), biogenic fluxes (bio - in green) and fossil fuel emissions (ff - in orange). The contributions from the boundary conditions are indicated by the gray line and the combined contribution of boundaries and biogenic fluxes (bdy+bio) by the light yellow line.*

Table 4.3 shows the data availability of the XCO_2 measurements during PM time. The dataset of UNA is larger than ALZ, but our inversion based on the urban-rural gradients requires data when UNA and ALZ are both available. Even more unfortunate is there are some times when ALZ is available but UNA is not. It worsens the lack of gradients. Even at the most available months (January to March), the days with available gradients are less than one-third of the month. I chose the best available month (February) as an example to test the potential constraint of XCO_2 data in our one-year run. There are 11 records available in February, and they are from 11 different days (February 1-3, 6, 16-21 and 28). Almost all of them are measured at local time 14:00 (UTC time 20:00).

UTCdatetime	UNA	UNAstdev	ALZ	ALZstdev	UNA-ALZ

2018-04	20	20	2	2	0
2018-05	27	27	0	0	0
2018-06	10	10	0	0	0
2018-07	0	0	0	0	0
2018-08	2	2	0	0	0
2018-09	5	5	2	1	0
2018-10	6	6	2	1	2
2018-11	18	18	7	2	2
2018-12	35	35	13	9	5
2019-01	53	53	8	7	8
2019-02	72	72	13	9	11
2019-03	56	56	9	5	8

Table 4.3 *Monthly data availability of XCO_2 hourly observation in the afternoon between 2018-03-30 to 2019-03-29 at UNA and ALZ with standard deviations.*

When applying the inversion plan 0b (block size=1d) to this limited data set, most gradients do not improve after inversion (shown in **Figure 4.18**). The small sample size leads to high uncertainties. Owing to the lack of observations in the rainy season, the emissions during the rainy season haven't improved. The dry season is better, with an adjustment on the total emission from 146.25 Mt to 139.41 Mt, consistent with Plan 0b and most plans on surface concentrations. However, the uncertainties are not really reduced because of the small sample size, as shown in **Figure 4.19**.

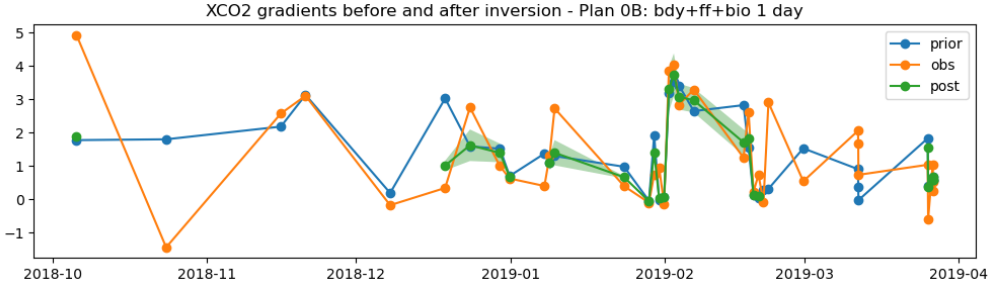


Figure 4.18 *XCO_2 UNA-ALZ mixing ratio gradients before and after inversion with Plan 0b, (the 1-day block plan)*

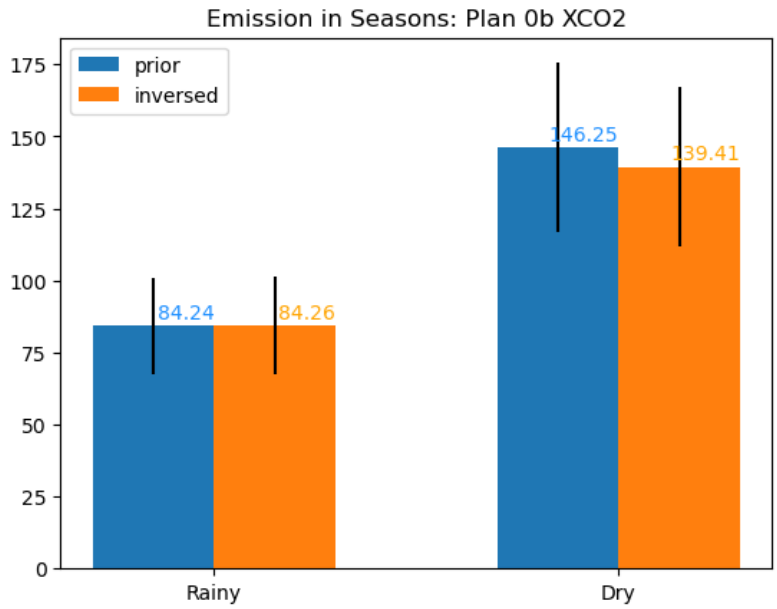


Figure 4.19 *CO_2 yearly emission (in Mt) of the two seasons before and after inversion by XCO_2 Plan 0b, with uncertainty bar.*

4.4 Discussion

4.4.1 Sensitivity experiments

4.4.1.1 Data screening

We combined the standard deviations of observations with the model-measurement mismatch for **R**, which is 2 ppm in our study and 3 ppm in Bréon et al. (2015) and Staufer et al. (2016). The standard deviations of our hourly observation gradients vary from 0.27 ppm to 38.26 ppm. If we add a criteria that the standard deviations of both stations are less than 3 ppm, the size of the database would drop by half.

With this observation Stdev filter, the prior mismatch as well as the prior observation differs, as shown in **Figure 4.20**. The reduced error remains consistent, while the first large mismatch in mid-September disappeared, with the peak in observation turned to a trough. It indicates this observation peak and this mismatch is probably caused by a low-quality observation period.

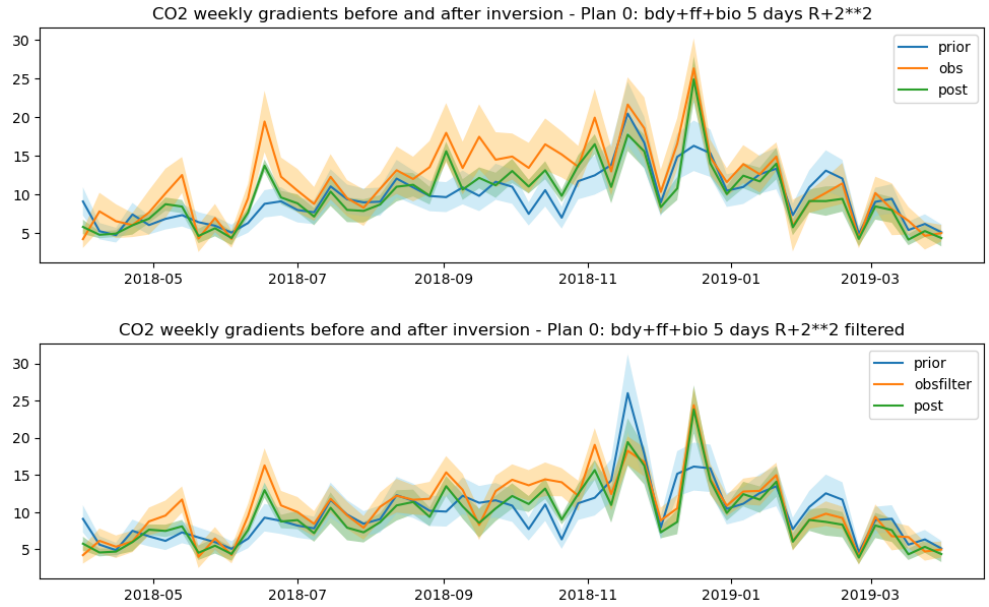


Figure 4.20 *The weekly average of CO_2 gradients (UNA-ALZ) before and after inversion Plan 0 compared to observations, with and without filter on StDev of hourly observations, with shade of InterQuartile Range.*

Despite the optimization of the gradients, the total emission is also adjusted. The emission in the dry and wet season has changed from 113.55 Mt/yr and 148.36 Mt/yr to 105.83 Mt/yr and 141.20 Mt/yr, much closer to the median of all inversions, which is 104.17 Mt/yr and 144.96 Mt/yr.

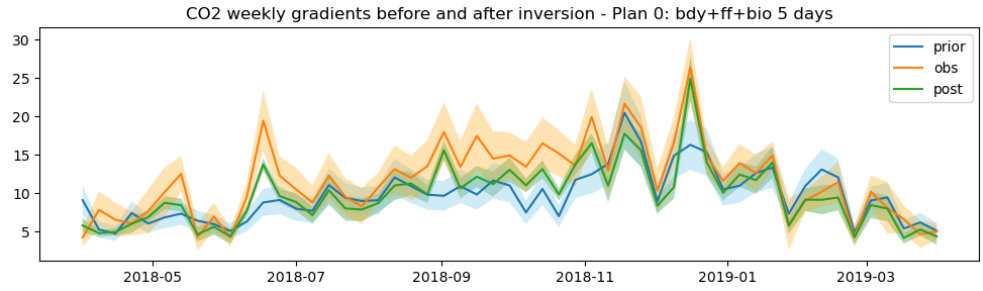
Another data screening is made in Plan 0d, which removed the gradients beyond 5% - 95%. **Table 4.4** shows the error comparison between the two data screening with reference Plan 0. Plan 0d removed 10% gradients and this observation filter (we call it Plan 0obs) removed nearly 50% observations. Plan 0obs significantly improved the errors of model-measurement mismatch.

Mt/y	Bias				StDev				RMSE				R^2			
No.	dry		wet		dry		wet		dry		wet		dry		wet	
	prior	post	prior	post	prior	post	prior	post	prior	post	prior	post	prior	post	prior	post
0	-1.09	-1.60	-3.31	-2.41	6.92	5.85	8.25	6.49	7.01	6.07	8.89	6.92	0.27	0.43	0.01	0.22
0obs	-0.02	-0.92	-1.61	-1.54	5.70	4.39	6.60	5.13	5.70	4.48	6.79	5.36	0.37	0.57	0.06	0.25
0d	-0.59	-1.61	-2.41	-1.83	6.29	5.15	6.61	5.38	6.31	5.40	7.04	5.69	0.20	0.33	0.04	0.19

Table 4.4 *Mean Bias, Standard Deviation (StDev), Root Mean Square Error (RMSE) and coefficient of determination (R^2) of the prior and posterior simulation of the CO_2 concentration gradients (UNA-ALZ) from various data screening Plans in dry and wet seasons.*

4.4.1.2 Prior emissions - ODIAC

As **Figure 4.21** shows, the prior gradients of ODIAC are much smaller than UNAM. Its prior emissions are also 17 Mt/yr less than UNAM. Nevertheless, their posterior gradients are consistent. Our inversion system is able to correct the 17 Mt/yr bias, which is around 7% of total emissions. However, ODIAC is 40% lower than UNAM in the total emission in MCMA (c.f. **Figure 3.2**), the center of domain 2. Night light does not fully represent human activities, especially in mega-cities. Our inversion without footprints cannot deal with this problem unless a city mask is also applied to ODIAC.



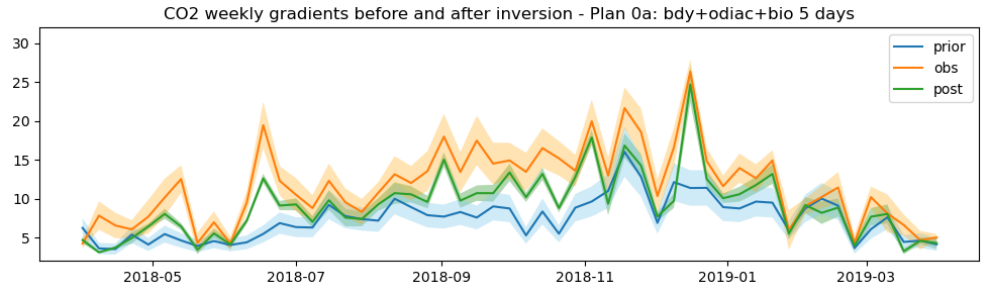
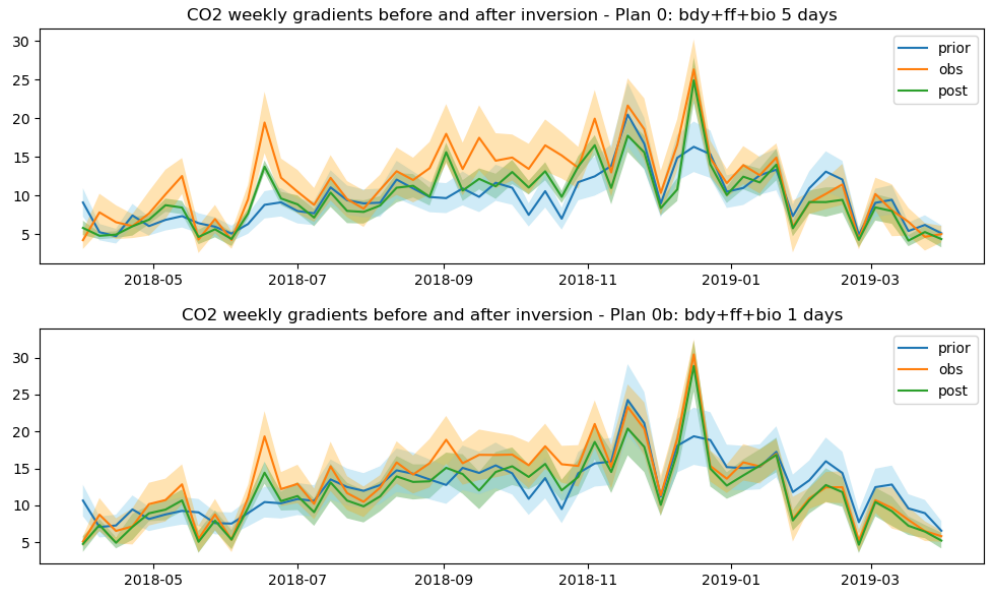


Figure 4.21 *The weekly average of CO_2 gradients (UNA-ALZ) before and after inversion Plan 0 and Plan 0a compared to observations, with shade of InterQuartile Range.*

4.4.1.3 Block sizes

Plan 0b tested a smaller block for 1 day and the posterior gradients and emissions are shown in **Figure 4.22**. The small sample size causes high uncertainties without the impact of day covariances and time window covariances, but the uncertainty could be double-validated by the reference Plan 0. The posterior weekly variations are similar, and Plan 0b captured the trends of CO_2 gradients better, especially in 2019. Unlike Plan 0, which uses only one factor for 5 days, the co-application of Plan 0 and Plan 0b provides detailed daily emissions and fluxes similar to Plan 1 series, but with only the fossil fuel tracer. It could be an optimal solution when the dataset is large and the memory becomes insufficient to apply day covariances.



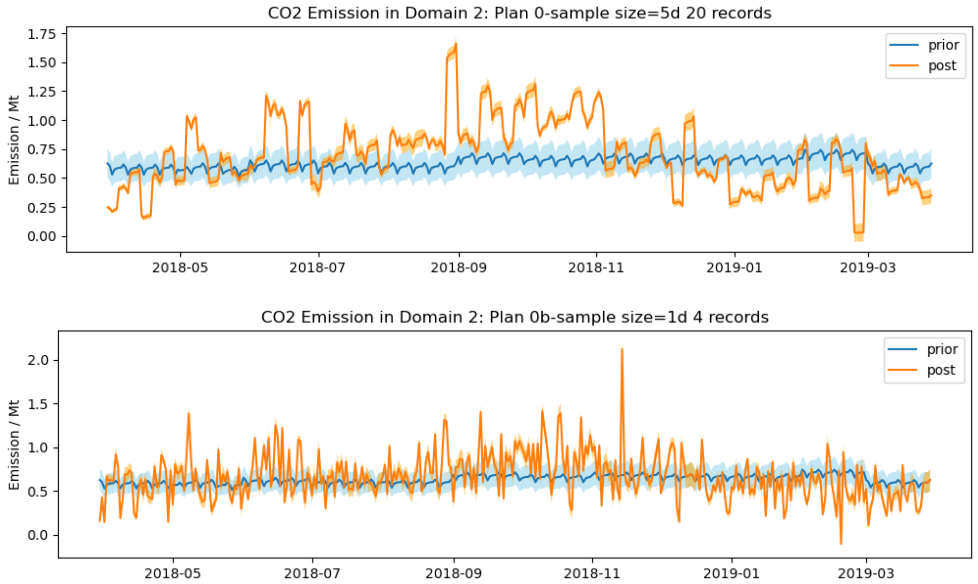


Figure 4.22 *The weekly average of CO_2 gradients (UNA-ALZ) and the daily emission fluxes before and after inversion Plan 0 and Plan 0b compared to observations, with shade of InterQuartile Range.*

In the other way, the length of a 5-day block is convenient for tests, but not as reasonable as a longer block, e.g. 30-day blocks in Bréon et al. (2015) and Staufer et al. (2016).

Here I applied a novel whole-block plan, Plan 2cb, with all the 365 days in one block. It is based on Plan 2c, and tested with block sizes. Instead of 20% uncertainty for 5 days, I use 20% uncertainty for a year. The posterior gradients in **Figure 4.23** are almost perfectly following the observation; the reduced uncertainty is 78.7% in MCMA; the posterior emission fluxes in **Figure 4.24** are 145.64 Mt/yr in the dry season and 100.76 in the wet season, similar to the median of all plans. The only problem is that the time series of scaling factors is not reasonable: they are sometimes negative and sometimes enormous. This may be related to improper prior uncertainty matrix **B**. The current system is not suitable for a whole-block inversion.

The 30-day block plan, Plan 2cb30, which keeps all other configurations in Plan 2c with 30-day blocks, was also tested. The posterior is not as similar to the observations as Plan 2cb, but it did improve the assimilation of the peaks in May 2018 and February 2019. The reduced error in MCMA is 47.4%, close to that of Plan 2c. Its posterior emissions are also close to Plan 2c, but the sharp peaks (e.g.

in early August and around November) are smoothed, with a greater likelihood of being closer to the true situation. The posterior total in the rainy season (106.35 Mt/yr) is close to the median of all plans and Plan 2c, while that of the dry season is reduced to 136.98 Mt/yr.

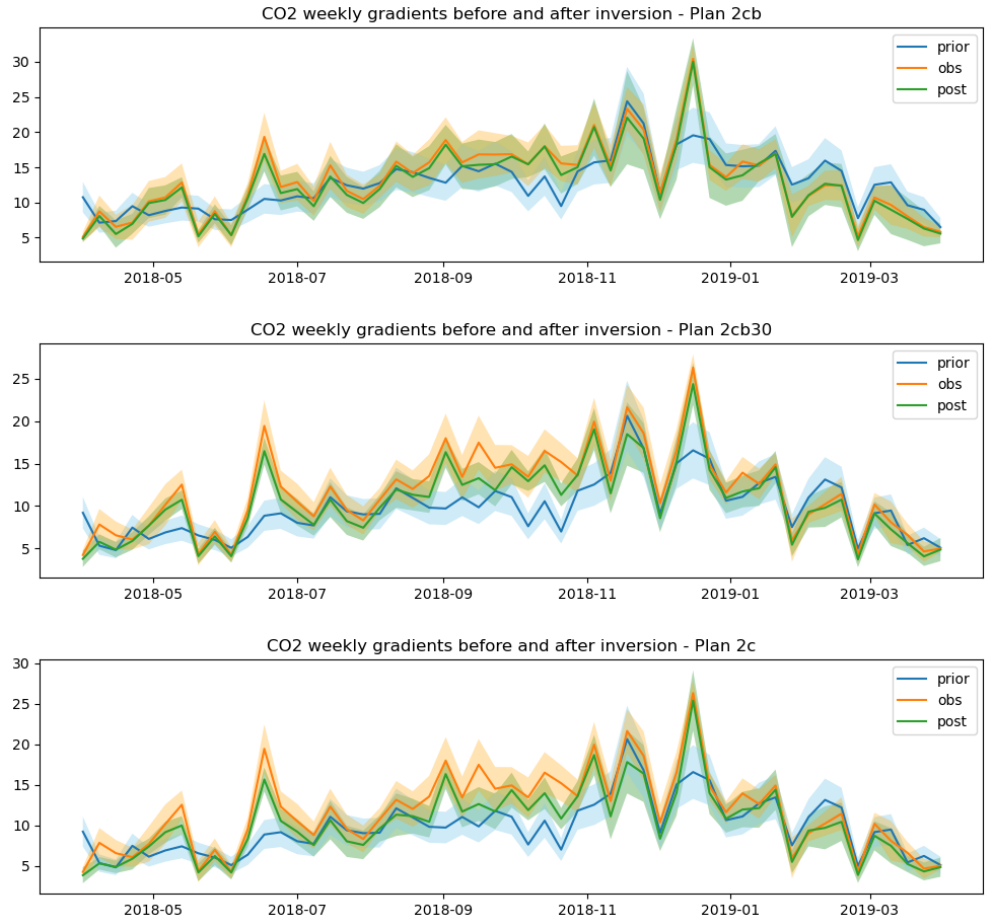


Figure 4.23 *The weekly average of CO_2 gradients (UNA-ALZ) before and after inversion of Plan 2cb (whole-block), Plan 2cb30 (30-day blocks) and Plan 2c (5-day blocks) compared to observations, with shade of InterQuartile Range.*

Figure 4.24 *The daily emission fluxes before and after inversion of Plan 2cb (whole-block) , Plan 2cb30 (30-day block) and the reference Plan 2c, compared to observations, with shade of InterQuartile Range.*

4.4.2 Limitations and perspectives

Due to the limitation in the size of the observation network in 2018 and 2019, our inversions only adjusted and reduced the uncertainties of the total emission in MCMA. We were not able to use any explicit tower footprints with a Lagrangian model in this study, and the peak-to-peak amplitude of posterior daily emissions is much larger than expected. With the deployment of the column sensors and LCMP sensors, the lack of data should be improved in the future study.

For the convenience of computation, my inversion plans are almost all done with the block size as 5-days. As a matter of fact, 30-day blocks (Bréon et al., 2015; Staufer et al., 2016) and whole-block are more widely adapted and reasonable. A larger block would reduce the loss

of unconsidered correlations between consecutive days divided into different blocks, and provide larger . The whole-block inversion system would be a supplementary solution as a comparison to the current system. The reason why a whole-block inversion led to an untrustable result remains to be solved.

Sensitivity tests could also be made on alternative background concentration products, biogenic flux products, transport models, emission products or various prior emission amounts. An alternative biogenic flux product (e.g. the diagnostic biosphere Vegetation Photosynthesis and Respiration Model, VPRM) (Mahadevan et al., 2008; Ahmadov et al., 2007, 2009) would provide a reference to the abnormal fluxes in June. If other products keep the robustness of the inversion system, we would be able to overcome the limits of time period according to the current products, and move to a period with more available XCO_2 measurements. Perturbation of prior emissions could test the robustness of the inversion system. In Chapter 2 and 3 we noticed the wind of July is overestimated, leading to a better dilution condition. Thus, the underestimation of the rainy season may be caused by dilution instead of underestimated inversion. Another operator $\mathbf{H}_{trans}$ would diagnose this underestimation.

This study attempted to constrain the traffic sector attribution, by means of the inversion of CO/CO_2 ratio, and with separate emission sectors. Although there is prior information of emission sectors, the current observation network is not providing extra information on mobile sources, with the urban station located in the UNAM campus. A mobile station located near the main roads would enhance our ability to analyze the mobile sources with prior mobile emissions and CO/CO_2 ratio. In the same way, extra stations deployed near Tula power plant would improve the optimization of industrial sources. The CO/CO_2 ratio inversion in this Chapter was rudimentary and susceptible to random biases. Perturbations with high-mobile and low-mobile scenarios would help to increase the stability of the inversion and to identify the contribution of the mobile sector.

REFERENCE

Bréon, F. M., Broquet, G., Puygrenier, V., Chevallier, F., Xueref-Remy, I., Ramonet, M., ... & Ciais, P. (2015). An attempt at estimating Paris

area CO 2 emissions from atmospheric concentration measurements. Atmospheric Chemistry and Physics, 15(4), 1707-1724.

CLARS-FTS from mount Wilson, California. Atmospheric Chemistry and Physics, 15(1), 241–252. https://doi.org/10.5194/acp-15-241-2015

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.

Feng, S., Lauvaux, T., Davis, K. J., Keller, K., Zhou, Y., Williams, C., ... & Baker, I. (2019). Seasonal characteristics of model uncertainties from biogenic fluxes, transport, and large-scale boundary inflow in atmospheric CO_2 simulations over North America. Journal of Geophysical Research: Atmospheres, 124(24), 14325-14346.

García-Reynoso, JA, Mar-Morales, BE, & Ruiz-Suárez, LG (2018). Spatial, temporal and speciation distribution model of the Mexican emissions inventory (base year 2008) for use in air quality modeling (DiETE). International Journal of Environmental Pollution, 34 (4), 635-649. https://doi.org/10.20937/rica.2018.34.04.07.

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

Government Headquarters of Mexico City: Head of Government participates in the Summit of Leaders on Climate 2021, https://www.cgaai.cdmx.gob.mx/comunicacion/nota/participa-jefa-de-gobierno-en-cumbre-de-lideres-sobre-el-clima-2021 (last access: 29 December 2022), 22 April 2021.

Government Headquarters of Mexico City: The "Mexico partnership for Net Zero Cities" project is launched in Mexico City with the support of the United States Government, https://www.cgaai.cdmx.gob.mx/comunicacion/nota/lanzan-en-la-ciudad-de-mexico-el-proyecto-mexico-partnership-net-zero-cities-con-el-apoyo-del-gobierno-de-estados-unidos (last access: 29 December 2022), 07 December 2022.

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., & Eder, B. (2005). Fully coupled “online” chemistry

within the WRF model. Atmospheric environment, 39(37), 6957-6975.

Gurney, K. R., Liang, J., Roest, G., Song, Y., Mueller, K., & Lauvaux, T. (2021). Under-reporting of greenhouse gas emissions in US cities. Nature communications, 12(1), 1-7.

Gurney, K. R., Razlivanov, I., Song, Y., Zhou, Y., Benes, B., & Abdul-Massih, M. (2012). Quantification of fossil fuel CO_2 emissions on the building/street scale for a large US city. Environmental science & technology, 46(21), 12194-12202.

INEGI (Instituto Nacional de Estadística y Geografía, Mexico National Institute of Statistics and Geography): Registered Motor Vehicles in Circulation, available at: https://www.inegi.org.mx/programas/vehiculosmotor/ (last access: 29 December 2022), 2021.

Jacobson, A. R., Schuldt, K. N., Miller, J. B., Oda, T., Tans, P., Andrews, A., ... & Moore, F. (2020). CarbonTracker CT2019B, NOAA Global Monitoring Laboratory.

LAEI (London Atmospheric Emissions Inventory): London Datastore, available at: https://data.london.gov.uk/dataset/london-atmospheric-emissions-inventory--laei--2019 (last access: 29 December 2022), 2021

Lauvaux, T., Miles, N. L., Deng, A., Richardson, S. J., Cambaliza, M. O., Davis, K. J., ... & Wu, K. (2016). High-resolution atmospheric inversion of urban CO_2 emissions during the dormant season of the Indianapolis Flux Experiment (INFLUX). Journal of Geophysical Research: Atmospheres, 121(10), 5213-5236.

London Government: Pathways to Net Zero Carbon by 2030, available at: https://www.london.gov.uk/programmes-and-strategies/environment-and-climate-change/climate-change/zero-carbon-london/pathways-net-zero-carbon-2030 (last access: 29 December 2022)

MOCEJ (NYC Mayor's Office of Climate and Environmental Justice): NYC Carbon Challenge - Reducing greenhouse gas emissions. https://climate.cityofnewyork.us/initiatives/nyc-carbon-challenge/ (last access: 29 December 2022)

Nakanishi, M., & Niino, H. (2009). Development of an improved turbulence closure model for the atmospheric boundary layer. Journal of the Meteorological Society of Japan. Ser. II, 87(5), 895-912.

Nalini, K., Lauvaux, T., Abdallah, C., Lian, J., Ciais, P., Utard, H., et al. (2022). High-resolution Lagrangian inverse modeling of CO_2 emissions over the Paris region during the first 2020 lockdown period. Journal of Geophysical Research: Atmospheres, 127, e2021JD036032. https://doi.org/10.1029/2021JD036032

Nassar, R., Napier-Linton, L., Gurney, K. R., Andres, R. J., Oda, T., Vogel, F. R., & Deng, F. (2013). Improving the temporal and spatial distribution of CO_2 emissions from global fossil fuel emission data sets. Journal of Geophysical Research: Atmospheres, 118(2), 917-933.

Nathan, B. J., Lauvaux, T., Turnbull, J. C., Richardson, S. J., Miles, N. L., & Gurney, K. R. (2018). Source sector attribution of CO_2 emissions using an urban CO/CO_2 Bayesian inversion system. Journal of Geophysical Research: Atmospheres, 123, 13,611–13,621. https://doi.org/10.1029/2018JD029231

Nickless, A., Rayner, P. J., Engelbrecht, F., Brunke, E.-G., Erni, B., and Scholes, R. J.: Estimates of CO_2 fluxes over the city of Cape Town, South Africa, through Bayesian inverse modelling, Atmos. Chem. Phys., 18, 4765–4801, https://doi.org/10.5194/acp-18-4765-2018, 2018.

Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., ... & Tans, P. P. (2007). An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker. Proceedings of the National Academy of Sciences, 104(48), 18925-18930.

Rella, C. W., Chen, H., Andrews, A. E., Filges, A., Gerbig, C., Hatakka, J., ... & Zellweger, C. (2013). High accuracy measurements of dry mole fractions of carbon dioxide and methane in humid air. Atmospheric Measurement Techniques, 6(3), 837-860.

Rodríguez Zas, J. A., & García Reynoso, J. A. (2021). Actualización del inventario nacional de emisiones de 2013 para la modelación de la

calidad del aire en el centro de México. Revista Internacional de Contaminación Ambiental, 37, 463–487.

SEDEMA (Secretaría del Medio Ambiente de la Ciudad de México, Secretary of the Environment of Mexico City): Contaminantes y de efecto invernadero 2012. Secretaría del Medio Ambiente de la Ciudad de México, available at: http://www.aire.cdmx.gob.mx/descargas/publicaciones/flippingbook/inventario-emisioneszmvm2012/inventario-emisioneszmvm2012.pdf

SEDEMA: Inventario de emisiones de la Zona Metropolitana del Valle de México 2018. Contaminantes criterio, tóxicos y compuestos de efecto invernadero. Secretaría del Medio Ambiente de la Ciudad de México (datos preliminares), 2021.

Seto, K. C., Dhakal, S., Bigio, A., Blanco, H., Delgado, G. C., Dewar, D., Huang, L., Inaba, A., Kansal, A., Lwasa, S., McMahon, J., Müller, D. B., Murakami, J., Nagendra, H., and Ramaswami, A.: Climate Change 2014: Mitigation of Climate Change. IPCC Working Group III Contribution to AR5. Chap. 12 - Human settlements, infrastructure and spatial planning. Cambridge University Press, Cambridge, UK and New York, NY, USA, 2014.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D., Duda, M. G., ... Powers, J. G. (2008). A Description of the Advanced Research WRF Version 3 (No. NCAR/TN-475+STR). University Corporation for Atmospheric Research. doi:10.5065/D68S4MVH.

Staufer, J., Broquet, G., Bréon, F. M., Puygrenier, V., Chevallier, F., Xueref-Rémy, I., ... & Ciais, P. (2016). The first 1-year-long estimate of the Paris region fossil fuel CO 2 emissions based on atmospheric inversion. Atmospheric Chemistry and Physics, 16(22), 14703-14726.

TMG (Tokyo Metropolitan Government): TMG announces the "Zero Emission Tokyo Strategy" for contributing to the world's net-zero CO_2 emissions by 2050, https://www.kankyo.metro.tokyo.lg.jp/en/about_us/zero_emission_tokyo/strategy.html (last access: 29 December 2022), 2019.

Tomohiro Oda, Shamil Maksyutov (2015), ODIAC Fossil Fuel CO_2 Emissions Dataset (Version name : ODIAC2019), Center for Global

Environmental Research, National Institute for Environmental Studies, doi:10.17595/20170411.001. (Reference date: 2022/05/22)

UN (United Nation) Climate Change (6 October 2021). UN Global Climate Action Awards: Winners Unveiled Today. UN Climate Press Release (Bonn, Germany). https://unfccc.int/news/un-global-climate-action-awards-winners-unveiled-today

United Nations, Department of Economic and Social Affairs, Population Division (2018). The World's Cities in 2018—Data Booklet (ST/ESA/ SER.A/417).

Wong, K. W., Fu, D., Pongetti, T. J., Newman, S., Kort, E. A., Duren, R., et al. (2015). Mapping CH 4: CO 2 ratios in Los Angeles with

xESMF: Universal Regridder for Geospatial Data, https://github.com/pangeo-data/xESMF (last access: 24 Jan 2022), https://doi.org/10.5281/zenodo.4294774.

Zhou, Y., C. A. Williams, T. Lauvaux, K. J. Davis, S. Feng, I. Baker, S. Denning, & Y. Wei. (2020). A multiyear gridded data ensemble of surface biogenic carbon fluxes for North America: Evaluation and analysis of results. Journal of Geophysical Research: Biogeosciences, 125(2), e2019JG005314. https://doi.org/10.1029/2019JG005314

5 CHAPTER 5 CONCLUSIONS AND PERSPECTIVES

5.1 CONCLUSIONS

The interpretation of CO_2 measurements in this study is mainly driven by an urban atmospheric CO_2 Bayesian inversion system that assimilates the modeled concentrations with surface and column measurements from the MERCI-CO_2 observation network deployed in Mexico-City. The system generates posterior estimates of CO_2 emission inventories in MCMA based on the measures from UNA and ALZ, the urban and rural monitoring stations, over a one year period starting from March 2018. The inversions successfully reduced uncertainties in prior emission inventories developed by Dr. García-Reynoso in UNAM. Additionally, the thesis introduces a robust atmospheric simulation system that not only works in the CO_2 inversion system, but also has the potential to be associated with other species, including GHG (e.g. CH4) and other trace gasses (e.g. NOx and CO).

In the following sections, I conclude and discuss three sub-topics presented in this study. The first one is about the evaluation of model performance on meteorological parameters related to CO_2 transportation and mixing conditions. The second part regards the reproduction of CO_2 concentrations in MCMA and their spatial and temporal variability, including a discussion of the quality of the background station ALZ, its background concentration patterns, and fossil fuel signals. The last and key part of the study focuses on the optimization of anthropogenic and biogenic emissions by a one-year simulation, along with an attempt to quantify the attribution of emission sectors.

5.1.1 The performance of WRF modeling system

Our modeling system is developed on WRF-Chem version 3.9, coupled with the single layer urban canopy model UCM. For the purpose of verifying that the model replicates the transport conditions of CO_2 and of optimizing the WRF configuration, various meteorological parameters, including wind speed, wind direction, vertical profile of horizontal wind, temperature and PBL heights, were compared and

evaluated with the observations for a 2-week period in January 2018. Sensitivity tests were conducted on various physical options, domain settings, urban canopy models, model inputs, nudging plans and other configuration options in Chapter 2.

According to the sensitivity tests, meteorological drivers and urban canopy models significantly reduced the mismatch between the simulations and the observations on temperature, wind and CO_2 concentrations. Simple grid nudging, on the contrary, did not improve the performance of the model. Some PBL schemes would cause extremely high wind in mountainous areas. A background concentration product is necessary for CO_2 simulation. Notably, the meso-resolution 5-km domain showed little difference from the high-resolution 1-km domain in the mixing conditions of CO_2 during the afternoon time. In fact, its performance was even better in the vertical profile of horizontal wind speeds.

An optimal configuration was determined based on the sensitivity tests. The configuration of one-way double nesting 5-km meso-resolution modeling was then tested under various climate conditions in the dry season (January), the wet season (July) and the transition between the two seasons (May). Each case was carried out within 4 weeks, the first two of which are reserved for background concentration spinning up from a distance and the last two are utilized for evaluation. The meteorological evaluation of this configuration in three months in Chapter 2 demonstrates that our modeling system sufficiently reconstructs atmospheric transportation conditions in MCMA, when compared to similar studies in other cities.. The urban area inside the Mexico valley is better reproduced than the rural area in the mountainous south west for most of the parameters. However, wind speeds were usually slightly overestimated in the rainy season and wind directions tended to turn clockwise in one of the 26 stations in the dry season and in the northeast area of the city in the rainy season. The former one abnormal station is possibly caused by the relatively low resolution, which couldn't represent the complex urban building structure. The modeled horizontal wind speed error is around 0.5 m/s compared to observed values under the height of 2000 m, covering most of the mixing layer. The bias is 50% larger in May, but still under 1 m/s in PBL. I combined the two variables (“Mixing layer height” for

AM and “Residual layer height” for PM) driven from the Lidar at UNAM to assess the model performance on PBL heights. The modeled PBL heights are consistent with the combined Lidar observation in both January and May (July is lacking data due to weather reasons).

Regarding CO_2 simulation in Chapter 3, the model captured the trends of atmospheric CO_2 concentrations during the afternoon, but failed to capture the accumulation at night, which correlates with the decrease of PBL heights and high vertical stability. Due to this reason, there is a one-hour time lag between the simulated and observed concentration peaks.

Alternative emission products did not bring about a considerable change in the CO_2 mole fractions during the afternoon. Despite a difference of approximately 40% in total emission in MCMA and a difference of 30% during daytime in the grid where UNA is located, ODIAC didn't make strong differences on the concentrations during daytime. It even decreased the model-measurement mismatch at night owing to the high nighttime emission based on nighttime light data.

The multi-layer urban canopy model BEP tends to create a bimodal distribution in the diurnal distribution, though the extra peaks at night did not impact the model performance in the daytime. BEP reduced the model bias on daytime CO_2 concentrations, but did not improve the mean absolute errors and even increased them in PM.

There are fewer measurements available for the total columns XCO_2, but the discrepancy between the model and measurement is larger compared to the surface measurements CO_2. About one-third of the observations are available in January and 27% in May at UNA, including morning observations. Measures at ALZ are even less, indicating difficulty in the retrieval of the observations. No XCO_2 measures are available during the afternoon period in July. Besides, the background concentration in May at ALZ is overestimated from May 17, while that at UNA seems to be underestimated before May 17. The same background mismatch existed at UNA on January 7-10. The realistic background concentration patterns of XCO_2 at the two stations may differ. Furthermore, there may be biases in the model in simulations of CO_2 in the upper atmosphere.

5.1.2 Characteristics of surface atmospheric CO_2 concentrations and gradients in MCMA

Due to the exposure to the local urban emissions the standard deviation of CO_2 measurements taken within a one-hour period could be as high as 38 ppm. For comparison with model results in Chapter 3, we used hourly averages of the measurements.

The urban station UNA is located in the southern part of the city, where the primary source of CO_2 emissions is associated with the traffic sector. The typical CO_2 concentrations at UNA in April 2018 to March 2019 ranged from 420 ppm to 470 ppm (with peaks up to 500 ppm), including 5 ppm to 60 ppm attributed to fossil fuel signals, according to the WRF simulations. These concentrations are high compared to the global average (415 ppm, WMO 2022) and other megacities in the world, e.g. Paris (390 - 430 ppm in 2016, Lian et al., 2019), due to the accumulation of pollutants in the Mexico-City basin.

As expected, the concentrations of CO_2 are the lowest on Sundays in January and July, when the traffic is reduced, but not in May. The seasonal difference of weekly cycles are probably related to human activities: more trips on Sundays in May due to the warm weather than in January and the less rainfall than in July. The average diurnal cycle for UNA is unimodal, peaking at 8:00 in AM in the dry season and 6:00 AM in the rainy season (depending on sunrise time). However, in reality the CO_2 diurnal cycle follows a bimodal distribution in about 50% of the days in January and 29% in May and July. The second peak at night is much smaller and its timing varies a lot. Therefore, it was smoothed in the averaged diurnal cycle.

At the background station ALZ, the diurnal cycle exhibits a distinctive pattern. As it is located away from the urban center, the CO_2 concentration variations at ALZ are mainly driven by biogenic fluxes. The averaged diurnal cycle keeps at around 410 ppm in January and May, and at around 405 ppm in July, with a daily amplitude of less than 5 ppm. It reaches its minimum before midday, and then rebound slightly due to the fossil fuel signal transported from MCMA basin in January and May. In July, the prevailing wind direction prevents the transportation of fossil fuel signals to ALZ.

Our modeling system enables me to break down the sources of local

CO_2 concentrations into background concentrations (also known as boundary conditions), fossil fuel signals and biogenic fluxes. Specifically, the UNAM emission inventory product divides the fossil fuel signals into three sectors: area (residential), mobile (transportation) and point (industrial and other sources). The traffic sources are mainly concentrated in the southern part of MCMA while the industrial sources are concentrated in the north. The residential sources, though negligible compared to the other two sectors, are concentrated in and to the west of MCMA, which almost coincides with the MCMA urban mask that I defined based on anthropogenic emissions.

The modeled contribution of CO_2 fossil fuel signal decreased by 50% during the rainy season, but this decrease is somewhat insignificant in observations. The difference between the maximum of the averaged diurnal cycle of January and July is about 34 ppm in the model, but only 7 ppm in the observations. This implies that the emission of July may be underestimated, or the model transportation error in the rainy season is larger than in the dry season.

Based on the knowledge gained from the model, the network design of MERCI-CO_2 is refined. I suggested several locations for urban stations to identify the contribution of large industrial sources and the fossil fuel signals in northern MCMA, and the quality of our current background station ALZ has been assessed. An ideal background station should have a background concentration highly correlated to the urban station, and a low fossil fuel signal. ALZ shows a low level of fossil fuel signals according to our modeling system. I then examined the background concentration maps to verify whether the background at ALZ represents the UNA background pattern, and whether there is a better background location. As we have assumed, whatever the prevailing wind direction is, the ALZ background concentration covaries with that of UNA, thanks to the basin terrains. The background concentrations are about 410 ppm on average, and vary more in the dry season than in the rainy season.

However, the biogenic flux patterns of the two stations are strikingly divergent. In the morning, the MCMA remains positive, dominated by ecosystem respiration, while the ALZ has started photosynthesis and carbon uptake. In the afternoon with well mixed air, the distinction diminishes, but the NEE of UNA is still higher than ALZ. The divergent

biogenic patterns require separate optimization of biogenic fluxes in the inversion in Chapter 4.

5.1.3 Estimation and uncertainty of emissions in MCMA

With the model configurations evaluated in Chapters 2 and 3, a one-year simulation was performed for inversion purposes. The simulation period commenced on March 16, 2018 with the first two weeks dedicated to spinning up the background concentration. Boundary conditions, biogenic fluxes, alternative emission products, various time windows, urban masks, emission sectors and carbon monoxide were simulated in separate tracers. The modeled concentration components confirmed the conclusion from Chapter 3 that the gradients of UNA and ALZ are driven by fossil fuel signals and biogenic fluxes throughout the year, with their background concentrations being nearly identical. The errors of the simulation are -1.56 ppm for the urban station UNA and 0.22 ppm for the rural station ALZ, which are sufficient for inversion.

There are two typical approaches to long-term Bayesian inversions: first, aggregating the data over a long period and then performing the inversion, and second, performing the inversions over shorter periods and then aggregating the results. Both approaches are tested in Chapter 4, but I used the second approach in most tests for the ease of debugging within a single block. Our inversion system involves three basic variables, background concentration, fossil fuel signal, and biogenic flux. The fossil fuel signal variable can be decomposed into different time windows, MCMA or non-MCMA areas and emission sectors in our one-year simulation starting from March 30, 2018. Each inversion block consists of five consecutive days.

Four series of CO_2 inversion plans were carried out in Chapter 4. Plan 0 series aggregates the 5 days in each block and creates one scaling factor based on the assimilation between the simulation and observation of the gradients, while Plan 1 series specifies the scaling factor for each day with an assumption of a day-to-day correlation up to 7 days. Plan 2 series takes advantage of not only the day-to-day correlation, but also the correlations between time windows. They assimilated the CO_2 concentration gradient of each time window each day. Plan 3 series considered the contributions of emission sectors based on Plan 1 series. Results suggest that UNAM emission

inventories overestimate the emissions in the dry season slightly, and underestimate in the rainy season by about 20%. The posterior gradient errors are reduced to less than 2 ppm in most of the year, except for September and October. Abnormal biogenic fluxes in June were fixed as well. Plan 2 series, which have the most complete covariance information produced the lowest posterior error and reduced the uncertainty during the daytime by more than 75% and in MCMA by 50%-70%. The posterior year total of MCMA is 61.7 Mt, which is 8 Mt less than UNAM gridded emission inventory and 4 Mt greater than the SEDEMA statistics. Overall, it seems the inversion plans improved the accuracy of emission estimates and provided valuable insights into the sources of CO_2 in the MCMA region.

I conducted multiple sensitivity tests across all plans, including test a for an alternative emission product (ODIAC), test b for block sizes, test c for MCMA urban mask, and test d for data screening. Although ODIAC showed similar performances to UNAM in the afternoon, its prior and posterior errors in gradients were more than 2 times of UNAM. Despite this, the inversion effectively corrected the 7% underestimation of ODIAC in domain 2. Small block sizes provided daily estimates without the assumption of day covariances, resulting in lower posterior flux errors and higher uncertainties. A medium sized block with sufficient day covariances, e.g. 30 days, corrected small mismatches in gradients. However, the first-aggregated-and-then-inversed approach, equivalent to a whole block of 365 days, successfully reproduced the gradients but failed to create a reasonable posterior emission time series. Owing to the location of the stations, our observation provides more information about MCMA than non-MCMA area. Therefore, the utilization of the MCMA urban mask significantly optimized the uncertainty in MCMA.

Two different data filtering strategies were applied to the inversion. Plan 0d kept observation gradients distributed between the 5th and the 95th percentile, while Plan 0obs filtered the hourly data with a standard deviation larger than 3 ppm, resulting in a 50% reduction of available data. After various data screening, the prior gradient mismatches between the model and measurements are completely different. Plan 0obs with the lowest prior bias did not reduce the bias. However, it improved the RMSE and R-square the most, especially in

the dry season.

Additionally, there are Plan 4 series, which assimilated the gradients of the tracer gas CO and CO/CO_2 ratio to constrain the CO_2 emissions from the mobile sector. The mobile source attribution from UNAM emissions, WRF modeled concentrations driven by UNAM emission, CO/CO_2 ratio and SEDEMA in MCMA are 47.7%, 34.9%, 25.1%, and around 60%, respectively, which are totally different from each other. The identification on the attribution of mobile sources in MCMA should be improved in future studies.

However, The same plans on XCO_2 were ineffective in achieving their intended goals, due to the lack of qualified measurements. Even during the month of February when more data are available (11 available records in February 1-3, 6, 16-21 and 28), there were no significant improvements in the posterior concentration gradients.

5.2 PROSPECTIVES

There are several aspects for improvement in future studies of MERCI-CO_2 beyond the knowledge gained in the previous Chapters.

One potential improvement is to test optimal parameters for the urban canopy model BEP, which has been proved to improve the model performance in urban-scale CO_2 simulations (Lian et al., 2019). Although our WRF coupled to BEP with default parameters on the structure of city buildings is susceptible to model instability and did not show significant improvements, testing optimal parameters would have a great potential to enhance the model performance. Observation nudging is another widely used option to improve urban-scale WRF simulations (Lian et al., 2018), which assimilates the WMO observations of meteorological parameters (e.g. wind, temperature and relative humidity) in the domains with the model to fix the drift that occurs during the model's execution. Although the attempt to apply observation nudging in our simulation system was hindered by data quality issues, developing alternative inputs for OBSGRID could reduce the model-measure mismatch on meteorological conditions. Spectral nudging, which requires particular PBL schemes that haven't been tested in our study, could be an option for the future modeling if needed.

Using alternative CO_2 products for boundary conditions, e.g. CAMS global atmospheric CO_2 inversion product (Chevallier, 2017), could expand the simulation to observation-intensive periods and improve data availability. The deployment of a network of 10 mid-cost medium-precision sensors of CO_2 in MCMA has been delayed by Covid-19, limiting the optimization of the spatial distributions. Nevertheless, our modeling system using MYJ is capable of TKE (Turbulent Kinetic Energy) and could be utilized for footprint once the network is completed in 2022-2023. Future sensitivity tests of inversions could include transport models (Staufer et al., 2016; Pisso, 2019), perturbation on total emission amounts (Staufer et al., 2016) and prior errors, etc.

Both the attempts on sector separated inversion and CO/CO_2 ratio didn't manage to optimize the sector attribution of mobile. The posterior proportions of mobile sources present notable deviation and none of them is consistent with prior knowledge. For the inversion on sector emission, an additional observation station near to main roads would provide additional information on traffic emissions. Sensitivity tests on perturbation of prior sector emissions would verify the robustness of the mobile source inversion. For the tracer gas CO, it is evident that there is a bias on the current inversions. It may be related to underestimated removal of atmospheric CO by oxidation with HOx radicals. I also note there are distinct differences in the CO and CO_2 ratio in the emissions and in the modeled concentrations. Scenario simulations on high and low mobile contribution would help solve this problem.

REFERENCE

WMO. (2022). Greenhouse Gas Bulletin: The State of Greenhouse Gases in the Atmosphere Based on Global Observations through 2021.

Lian, J., Bréon, F. M., Broquet, G., Zaccheo, T. S., Dobler, J., Ramonet, M., ... & Ciais, P. (2019). Analysis of temporal and spatial variability of atmospheric CO_2 concentration within Paris from the GreenLITE™ laser imaging experiment. Atmos. Chem. Phys, 19, 13809-13825.

Lian, J., Wu, L., Bréon, F. M., Broquet, G., Vautard, R., Zaccheo, T. S., ... & Ciais, P. (2018). Evaluation of the WRF-UCM mesoscale model and ECMWF global operational forecasts over the Paris region in the prospect of tracer atmospheric transport modeling. Elementa:

Science of the Anthropocene, 6.

Chevallier, F., Description of the CO_2 inversion production chain. CAMS deliverable CAMS73_2015SC2_D73.1.5.5_201703_CO_2 inversion production chain_v1, available at: http://atmosphere.copernicus.eu/, 2017.

Pisso, I., Patra, P., Takigawa, M. et al. Assessing Lagrangian inverse modelling of urban anthropogenic CO_2 fluxes using in situ aircraft and ground-based measurements in the Tokyo area. Carbon Balance Manage 14, 6 (2019). https://doi.org/10.1186/s13021-019-0118-8.

LIST OF ABBREVIATIONS

Abbreviations	Full description
ACES	The Anthropogenic Carbon Emissions System
ACO	Acolman SEDEMA station
AGU	American Geoscience Union
AJM	Ajusco Medio SEDEMA station
AJU	Ajusco SEDEMA station
ALZ	Altzomoni MERCI-CO_2 site
AM	Morning (Ante Meridiem)
AME	AMECAMECA_HSAN MERCI-CO_2 site
ANR	French Agence Nationale de la Recherche
BAU	Business-as-usual
BEP	Building Environment Parameterization
BJU	Benito Juarez SEDEMA station
CAMS	Copernicus Atmosphere Monitoring Service
CASA	Carnegie Ames Stanford Approach
CCA	Centro de Cienciasde la Atmósfera
CCAM	Conformal Cubic Atmospheric Model
CDIAC	Carbon Dioxide Information Analysis Center
CDMX	Ciudad de México
CH4	Methane
CHO	Chalco SEDEMA station
CMAQ	Community Multiscale Air Quality modeling system
CO	Carbon monoxide
CO_2	Carbon dioxide
COCCON	COllaborative Carbon Column Observing Network
CONACYT	Mexican Consejo Nacional de Ciencia y Tecnología
COP	Conference of the Parties
CRDS	Cavity Ring-Down Spectroscopy
CT	Carbon Tracker
CUA	Cuajimalpa SEDEMA station
CUT	Cuautitlán SEDEMA station
DARTE	Database of Road Transportation Emissions
ECMWF	European Centre for Medium-Range Weather Forecasts
EDGAR	Emissions Database for Global Atmospheric Research
EGU	European Geoscience Union
EPA	US Environmental Protection Agency
ERA-5	ECMWF reanalysis for global climate and weather - 5th

	generation
ERPA	Emission Reductions Payment Agreements
ESRL	NOAA Earth System Research Laboratories
FAC	FES Acatlán SEDEMA station
FFDAS	Fossil Fuel Data Assimilation System
FLEXPART	FLEXible PARTicle dispersion model
FTIR	Fourier Transform Infrared
GAM	Gustavo A. Madero SEDEMA statoin
GAW	WMO Global Atmosphere Watch Programme
GGMT	Greenhouse Gases and Related Measurement Techniques
GHG	Greenhouse Gases
GMAO	Global Modeling and Assimilation Office
GPP	Gross Primary Productivity
GSFC	Goddard Space Flight Center
GSMA	Groupe de Spectrométrie Moléculaire et Atmosphérique
HGM	Hospital General de México SEDEMA station
HRRR	High-Resolution Rapid Refresh model
HYSPLIT	Hybrid Single-Particle Lagrangian Integrated Trajectory model
IAEA	WMO International Atomic Energy Agency
ICOS	Integrated Carbon Observation System
ILS	Instrument Landing System
INEGI	Instituto Nacional de Estadística y Geografía
INFLUX	Indianapolis Flux Experiment
INN	Investigaciones Nucleares SEDEMA station
IPCC	Intergovernmental Panel on Climate Change
KIT	Karlsruhe Institute of Technology
LAA	Lab. de Análisis Ambiental SEDEMA station
LAEI	London Atmospheric Emissions Inventory
LCMP	low-cost medium-precise
LMD	Laboratoire de Météorologie Dynamique
LPDM	Lagrangian particle dispersion model
MAE	Mean absolute error
MAP	Mean Areal Precipitation
MCMA	Mexico City Metropolitan Area
ME	Mean error
MER	Merced SEDEMA station
MERCI-CO_2	Mexico City's Regional Carbon Impacts
MGH	Miguel Hidalgo SEDEMA station
MMM	NCAR Mesoscale and Microscale Meteorology Laboratory
MOCEJ	NYC Mayor's Office of Climate and Environmental Justice

MON	Montecillo SEDEMA station
MPA	Milpa Alta SEDEMA station
MYJ	Mellor–Yamada–Janjić PBL scheme
MYNN	Mellor–Yamada–Nakanishi–Niino PBL scheme
NAM	The North American Mesoscale Model
NCAR	National Center for Atmospheric Research
NDACC	Network for the Detection of Atmospheric Composition Change
NEE	Net ecosystem exchange
NEZ	Nezahualcóyotl SEDEMA station
NOx	Nitrogen oxides
NOAA	National Oceanic and Atmospheric Administration
NYC	New York City
OCO	Orbiting Carbon Observatory Satellite
ODIAC	Open-source Data Inventory for Anthropogenic CO_2
OMM	Organisation Mondiale de la Météorologie
PBL	Planet Boundary Layer
PED	Pedregal SEDEMA station
PM	Afternoon (Post Meridiem)
Re	Ecosystem respiration
RMS	Root mean square
RMSE	Root mean square error
RUOA	Red Universitaria de Observatorios Atmosféricos
SAG	San Agustín SEDEMA station
SEDEMA	Secretary of the Environment of Mexico City
SFE	Santa Fe SEDEMA station
SLUCM	Single-Layer urban canopy Model
SNAP	Selected Nomenclature for Air Pollution
SRI	Self-Reported Inventories
STILT	Stochastic Time-Inverted Lagrangian Transport model
TAH	Tláhuac SEDEMA station
TCCON	Total Carbon Column Observing Network
TEC	TECAMAC MERCI-CO_2 site
TKE	Turbulent kinetic energy
TLA	Tlalnepantla SEDEMA station
TMG	Tokyo Metropolitan Government
TROPOMI	The TROPOspheric Monitoring Instrument
TXO	TEXO MERCI-CO_2 site
UAX	UAM Xochimilco SEDEMA station
UCM	Single layer Urban Canopy Model
UIZ	UAM Iztapalapa SEDEMA station

UNA	UNAM MERCI-CO_2 site
UNAM	Universidad Nacional Autónoma de México
UNFCCC	United Nations Framework Convention on Climate Change
URBACT	URBan ACTion III Programme
UTC	Coordinated Universal Time
VAL	VALLEJO MERCI-CO_2 site
VIF	Villa de las Flores SEDEMA station
VIIRS	Visible Infrared Imaging Radiometer Suite
VPRM	Vegetation Photosynthesis and Respiration Model
WMO	World Meteorological Organization
WPS	WRF Preprocessing System
WRF	Weather Research and Forecasting Model
WRF-Chem	Weather Research and Forecasting model with Chemistry
WRI	World Resources Institute
XAL	Xalostoc SEDEMA station
XCO_2	Dry air column-averaged mole fractions of CO_2
xESMF	A Python package based on xarray and all ESMF regridding algorithms
YSU	Yonsei University scheme
ZMVM	Zone métropolitaine de la vallée de Mexico

Appendix A : The submitted version of Chapter 2 & Chapter 3

Evaluation of atmospheric CO_2 simulation over the Mexico City metropolitan area by WRF-Chem

Y. Xu[a], T. Lauvaux[a,b], M. Grutter[c], N. Taquet[c], JA. García-Reynoso[c], O.Laurent[a], M. Lopez[a], J. Lian[a], X. Lin[a], W. Stremme[c], M. Ramonet[a]

[a] Laboratoire des Sciences du Climat et de l'Environnement (LSCE), IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France
[b] Groupe de Spectrométrie Moléculaire et Atmosphérique (GSMA), Université de Reims Champagne Ardenne, UMR CNRS 7331, Reims, France
[c] Centro de Ciencias de la Atmósfera, Universidad Nacional Autónoma de México, México

Abstract

Atmospheric inversions have been utilized across multiple cities to improve self-reported emission inventories, but the atmospheric dynamics over large metropolitan areas located in complex topography remain challenging for atmospheric transport models. Here, we evaluate the performance of an atmospheric CO_2 modeling system and assess the optimal deployment of measurement sites to capture the whole-city emissions. As part of the French-Mexican project "Mexico City's Regional Carbon Impacts (MERCI-CO_2)", two observation stations of surface and two column CO_2 sensors were deployed in the Mexico City metropolitan area. We show that our modeling system is able to capture the local atmospheric dynamics with WRF at high resolution. The mean error in wind speed is 0.69 m/s and the relative error in PBL is 6.4%. The model also simulates CO_2 gradients observed by in situ and column sensors during dry and rainy seasons. The relative error of CO_2 surface mixing ratio in the afternoon is -4.3%. The current network is deemed sufficient to capture city emissions (as a whole - not with intra-urban information) and to define the background conditions. Mountains are ideal to capture the background but the biogenic signals are responsible for large CO_2 spatial gradients and temporal variations. Therefore, we conclude that biogenic fluxes must be optimized within future inversion systems. However, we demonstrate how stations in the valley and mountain are better simulated compared to those on the basin slopes, a deterrent to the deployment of sites in the suburbs of Mexico.

Keywords: urban emissions, greenhouse gas measurements, atmospheric modeling, Mexico City, WRF-chem, network design

• Introduction

Atmospheric greenhouse gas (GHG) concentrations have significantly increased at global scale since the industrial revolution. Carbon dioxide (CO_2), the main contributor to climate change, has seen its concentration reaching 414.7 ppm in 2021 (NOAA, 2022) compared to its pre-industrial level of 260-270 ppm (Wigley, 1983). Among the major sources of CO_2, emissions from cities represent more than 70% of global fossil fuel CO_2 emissions (IPCC, WG2, 2014). Countries and sub-national governments around the world are implementing policies to mitigate CO_2 emissions (IPCC, WG3, 2014). The European Climate Law sets a target to reduce the net GHG emissions of the European Union by at least 55%

by 2030 compared to 1990 levels[5]. At the sub-national level, cities have also pledged to reduce their emissions, organized into large international consortiums. About 100 large metropolitan areas have joined the C40 City consortium, a network of mayors advocating for climate policies in line with the 1.5°C ambition of the Paris Agreement[6]. In parallel, more than 12,000 cities from 142 countries joined the Global Covenant of Mayors, committed to reduce their annual GHG emissions by 1.9 Gt CO_2 in 2030, compared to a business-as-usual (BAU) trajectory[7].

Accounting of GHG emissions is performed following varied established protocols resulting in uncertainties and inconsistencies across Self-Reported Inventories (SRI) of CO_2 emissions (Chen et al., 2019; Mueller et al., 2021; Gurney et al., 2021). Other non-standard approaches have also been developed to estimate city-scale emissions using spatial and temporal disaggregation methods of national-scale emissions, based on satellite night light data, a proxy of human activities, in combination with activity data from power plants and large industries (Oda et al., 2015, Olivier et al., 2015). Both methods (inventories and spatial disaggregation) are prone to large uncertainties, impairing the ability to verify the effectiveness of emission reduction policies over time. Common protocols, recent and detailed local activity data, and independent verification remain essential to improve current fossil fuel CO_2 estimates (Mueller et al., 2021). More recently, atmospheric methods (inverse modeling) have been developed and applied to monitor CO_2 emissions from megacities, e.g. Paris (Bréon et al., 2015; Staufer et al., 2016), Indianapolis (Lauvaux et al., 2016; 2020), Cape Town (Nickless et al., 2018), Washington D.C. and Los Angeles (Feng et al., 2016; Yadav et al., 2021). However, atmospheric approaches require modeling systems affected by transport errors, especially over complex terrain and near coastal areas, where the local dynamics vary rapidly at sub-kilometric scales (Ye et al., 2020). Atmospheric inversion systems also rely on the quality of the emissions inventories, often delayed by several years, increasing the uncertainties in fast-growing metropolitan areas. The assessment of atmospheric models remains key to producing reliable and accurate estimates, taking advantage of existing meteorological measurement networks at the ground and collected during aircraft campaigns (Feng et al., 2016; Deng et al., 2017; Lian et al., 2018; Lopez-Coto et al., 2020a).

The Mexico City Metropolitan Area (MCMA), or Greater Mexico City (incl. La Ciudad de México, CDMX) is one of the largest megacities in the world. As the capital city of Mexico, its population of about 22 million inhabitants (UN, 2018) ranks fifth in the world with a population growth rate of about 0.9% per year. Motor vehicle ownership has grown by 30% from 2014 to 2020 (INEGI Mexico, 2021), leading to a 25% increase in CO_2 emissions. The Secretaría del Medio Ambiente de la Ciudad de México (SEDEMA, https://www.sedema.cdmx.gob.mx/) updates an inventory of pollutant emissions, including CO_2, every 2 years, with a delay of about 2 years. The latest reported inventory of GHG emissions of Mexico City is available for the year 2018 (SEDEMA, 2021). The total CO_2 emissions for 2018 amount to 66 Mt CO_2 in the MCMA, with about 85% from the transport and industry sectors. This represents a 50% increase in estimated CO_2 emissions compared to the inventory for 2014 (SEDEMA, 2016), mainly due to a 60% increase in the mobile sources sector. However, methodological changes and additional activity data explain the strong increase in only four years. In 2015, Mexico committed to reduce by

[5] https://ec.europa.eu/clima/eu-action/european-green-deal/european-climate-law_en

[6] https://www.c40.org/about-c40/

[7] https://www.globalcovenantofmayors.org/who-we-are/

22% its greenhouse gas emissions by 2030 and to achieve carbon neutrality in Mexico City by 2050 as stated in its Energy Transition Law.

The French-Mexican project “Mexico City’s Regional Carbon Impacts (MERCI-CO_2)”, funded by the French *Agence Nationale de la Recherche* (ANR) and the Mexican *Consejo Nacional de Ciencia y Tecnología* (CONACYT), enabled the strengthening of measuring atmospheric CO_2 concentrations over the MCMA by the deployment of additional instrumentation. The project was successful to incorporate three types of CO_2 measurements in and around Mexico City. First, two high-precision spectrometers collect hourly CO_2, CH_4, and CO concentration measurements (Gonzalez del Castillo et al., 2022) calibrated on international measurement scales. Second, about eight CO_2 sensors of intermediate cost and precision are to be deployed around the area. Those mid–cost instruments are based on the development done by Arzoumanian et al. (2019). The periods of confinement due to COVID-19 in Mexico City have delayed the deployment of these instruments, initially planned for 2020. And finally, two measurements of total columns of CO_2, CH_4, and CO are being collected, one from the mountain tops (Altzomoni station, ALZ; 19.12°N, 98.66°W) and a second downtown (UNAM university campus, UNA; 19.33°N, 99.18°W). The objective of these surface and column measurements is to characterize the atmospheric CO_2 enhancements caused by anthropogenic emissions from the MCMA. To quantify the physical relationship between surface emissions from urban sources and atmospheric concentrations measured by the various instruments, an atmospheric transport model is required. This atmospheric approach comes up against several obstacles and requires, on the one hand, sufficient measurement density to characterize the evolution of concentrations as a function of exposure to emissions; and on the other hand an atmospheric circulation model capable of representing horizontal and vertical mixing in an urban environment (Feng et al., 2016). The application of atmospheric top-down inversion methods to CO_2 on a regional scale must also face specific difficulties, such as the strong overlap of atmospheric signals related to anthropogenic and biospheric emissions; as well as atmospheric signals related to local/regional emissions from those on a larger spatial scale (Schuh et al., 2021).

As part of MERCI-CO_2, an atmospheric CO_2 modeling system was configured over the MCMA and its surroundings. The location of Mexico-City spreading through a wide valley surrounded by mountains can prove to be advantageous in favoring the accumulation of CO_2 in the planetary boundary layer (PBL) and therefore generating relatively high atmospheric gradients, more easily detectable by the analyzers. On the other hand, the complex topography around the basin represents a major difficulty for the simulation of atmospheric transport. In this study, we present the meteorological evaluation of atmospheric CO_2 simulations using the Weather Research and Forecasting model (WRF, Skamarock et al., 2008) adapted to the regional scale, and the results of a simulation of the atmospheric concentrations of CO_2 (WRF-Chem, Grell et al., 2005). The reliability of the model was tested based on wind speed, wind direction, PBL heights and temperature monitored at several locations disseminated in and around the city. We then compared the simulated surface and column CO_2 concentrations to the available observations in Mexico-City.

• Methods

o Atmospheric modeling system

The atmospheric simulations were conducted with WRF-Chem version 3.9 (Skamarock et al., 2008, Grell et al., 2005). The atmospheric turbulence in the PBL was simulated using the level-2.5 closure scheme Mellor–Yamada–Nakanishi–Niino (MYNN) (Nakanishi and Niino, 2009), coupled to the

MYNN atmospheric surface layer scheme and the single-layer urban canopy model (UCM, Tewari et al., 2007; Chen et al., 2011). MYNN was selected based on previous studies that demonstrated its ability to reproduce convective boundary layers (Feng et al., 2016; Deng et al., 2017; Lian et al., 2018) and possibly nighttime boundary layer dynamics (Lopez-Coto et al., 2020b). Our simulation was performed over three months (January, May and July 2018) to examine three typical climatic conditions over Mexico City: the dry season, the rainy season, and the transition between the two. During the month of July, recurring extreme weather events decrease the availability of observations, especially for Lidar and FTIR column measurements (XCO_2). In every month's run, the first 2 weeks are used for spinning up CO_2 concentrations from distant areas (background variations), and the last 2 weeks are used for evaluation and analysis. The three comparison periods of 2018 are 4-17 January, 15-28 May, and 14-27 July (UTC time).

WRF-Chem was configured with one-way two-nested domains for the Valley of Mexico City, with the spatial resolutions of 15 km and 5 km (Figure 1). The larger domain (domain 1, d01) covers almost the entire territory of Mexico and several neighboring countries; the smaller domain (domain 2, d02) covers the MCMA as well as several surrounding states, to include the whole Valley of Mexico. There are 49 vertical layers in each domain, the first of which is around 16 meters above ground level. Due to the complex terrain of Mexico City's valley, the model would be very unstable with finer vertical layers near the ground.

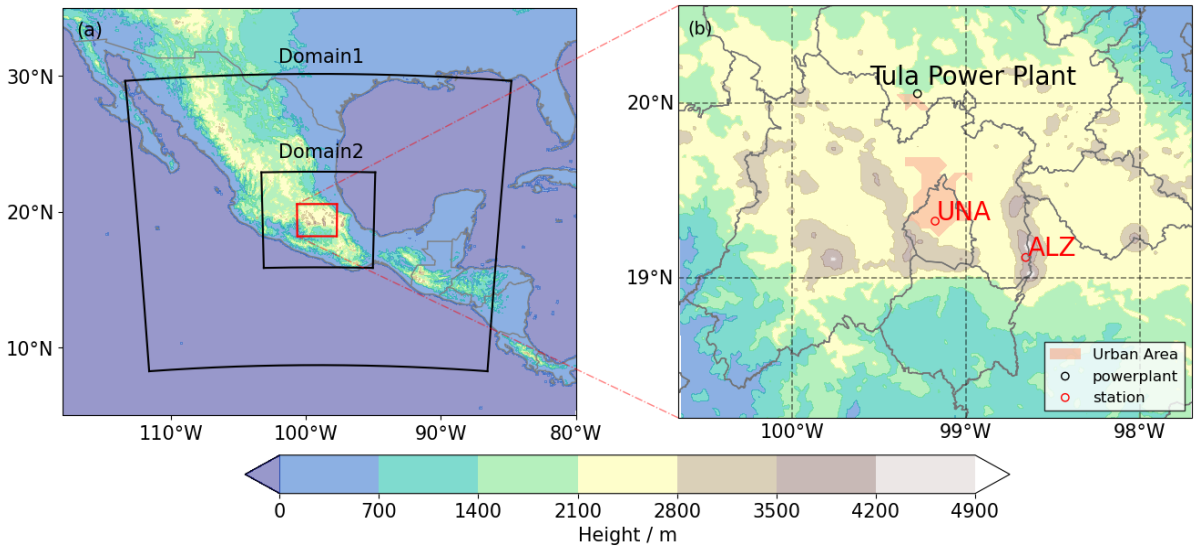


Figure 1: Simulation domains (a) of the two nested domains at 15-km (Domain 1) and 5-km (Domain 2) resolutions with their corresponding terrain heights (in meters) and (b) cropped area of Mexico City and its surroundings including the CO_2 observation stations for surface and total columns measurements (red circles), a large power plant nearby (black circle), the urban area of MCMA (in pink), the administrative boundaries of Mexico City (in gray) and the model topography in meters.

The meteorological drivers used as boundary and initial conditions are the 6-hourly ECMWF Reanalysis version 5 (ERA-5, Hersbach et al., 2017), at 0.25° × 0.25° resolution. Due to the lack of available high-quality meteorological observations in and around Mexico City, the meteorological simulations were performed without data assimilation. However, several surface stations, not referenced at the World Meteorological Organization (WMO), were used to evaluate the model performances.

o Surface CO_2 emissions and boundary conditions

▪ Fossil fuel emissions inventories

Two emission products were used in our study: a gridded emission inventory (Garcia-Reynoso et al., 2018; Rodriguez-Zas et al., 2021. https://github.com/JoseAgustin/emis_2016) developed at the National Autonomous University of Mexico (UNAM), and the Open-source Data Inventory for Anthropogenic CO_2 (ODIAC) (Oda et al., 2018). The UNAM hourly national emission inventory of Mexico covers the entire country at 3-km resolution while the 1-km resolution product is only available for specific subdomains, including the MCMA. This local emission inventory provides CO_2 emissions, as well as other species including CO and NO_x, for three sectors: area sources (residential/commercial), mobile sources (traffic), and point sources (industrial). In our simulation, we defined one independent tracer per sector to quantify individually the sectoral contributions on the observed atmospheric CO_2 enhancements.

The ODIAC emission product provides global emission maps of fossil fuel CO_2. ODIAC downscales national CO_2 emissions from the Carbon Dioxide Information Analysis Center (CDIAC; Boden et al., 2017) to sub-national levels by disaggregating the emissions according to night light data from the Visible Infrared Imaging Radiometer Suite (VIIRS) on the Suomi National Polar-orbiting Partnership satellite (e.g., Elvidge et al., 2013). Combined with temporal scaling factors (Nassar et al., 2013), ODIAC emissions are available at 1×1 km resolution and hourly (including monthly, weekly and daily temporal profiles).

Anthropogenic emissions over the parent domain (d01) are interpolated from ODIAC, including areas outside Mexico and shipping over the ocean. We note, however, that shipping emissions remain low compared to the UNAM inventory. For our second domain at 5-km resolution, we applied the UNAM emission inventory as well as ODIAC in another tracer as a sensitivity test. To conserve the total budget, all the emissions are interpolated into our domains by applying a mass-conserving technique (xESMF - conservative method, https://doi.org/10.5281/zenodo.1134365). CO_2 emissions from mobile sources remain nearly constant all year long (about 3 Mt per month), but emissions from residential and industrial sources show large seasonal variations with larger values for the month of January (11.93 Mt) compared to May (9.10 Mt). Emissions for the month of July increase to 10.00 Mt due to the energy demand from air conditioning systems (Figure 2).

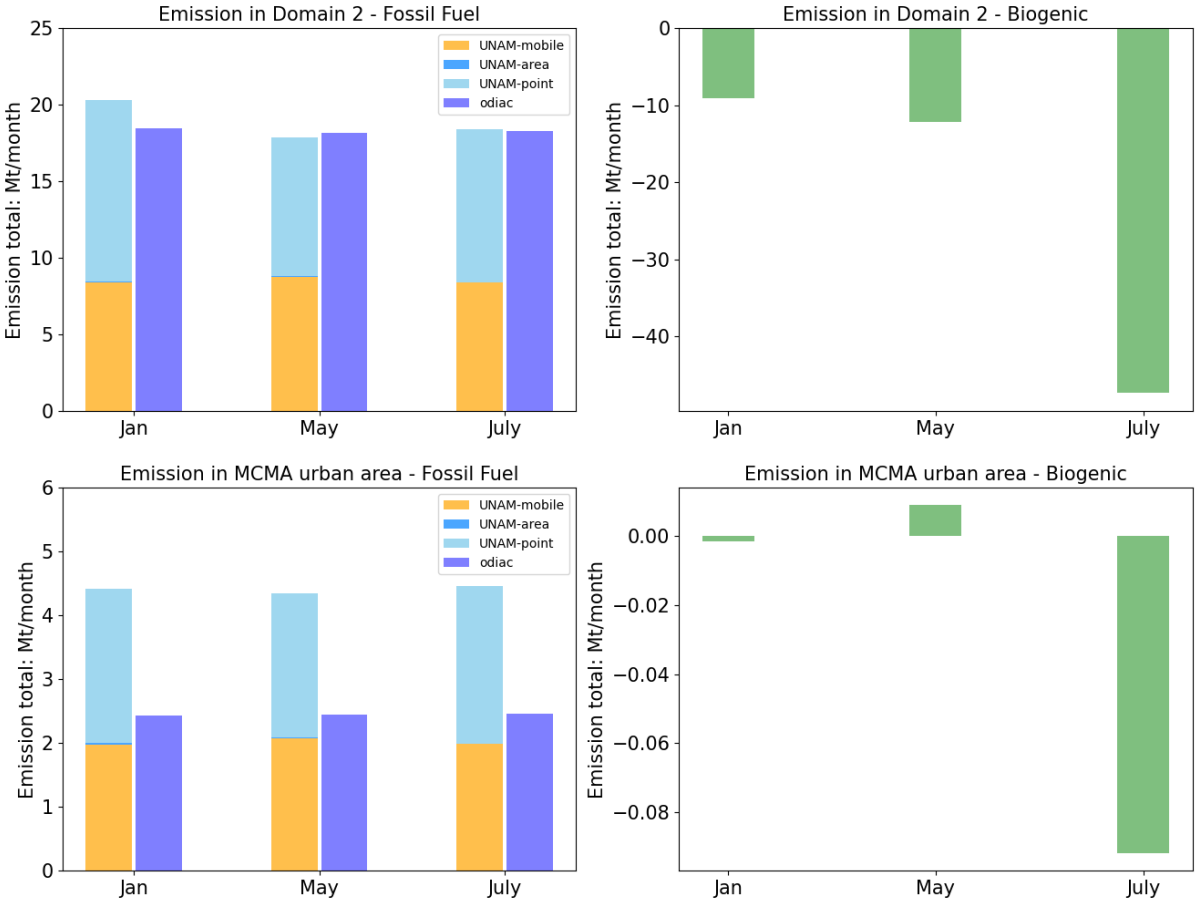


Figure 2: Total fossil fuel CO_2 emissions over the domain 2 (upper panels) and over the urban area of MCMA (lower panels) for the UNAM and ODIAC emissions products (left panels), and biogenic fluxes from the CASA biogeochemical model (right panels) for the months of January, May, and July 2018.

Total emissions over domain 2 from UNAM and ODIAC are generally consistent, but their spatial distributions differ (Figure 3). Although both ODIAC and UNAM grid values follow a lognormal distribution in the same interval (Figure 3 bc), ODIAC shows a larger number of low-value grid points (near-zero values). Due to the spatial disaggregation technique, point source emissions in ODIAC tend to be more distributed spatially compared to UNAM, where point sources are geo-located by their exact coordinates. For our application, the ODIAC emissions show the same distribution at 15 km resolution, with consistent country-scale emissions. When it comes to city-scale, ODIAC underestimate the emissions by 43%-45% compared to UNAM, similar to other cities in North America (Chen et al., 2020).

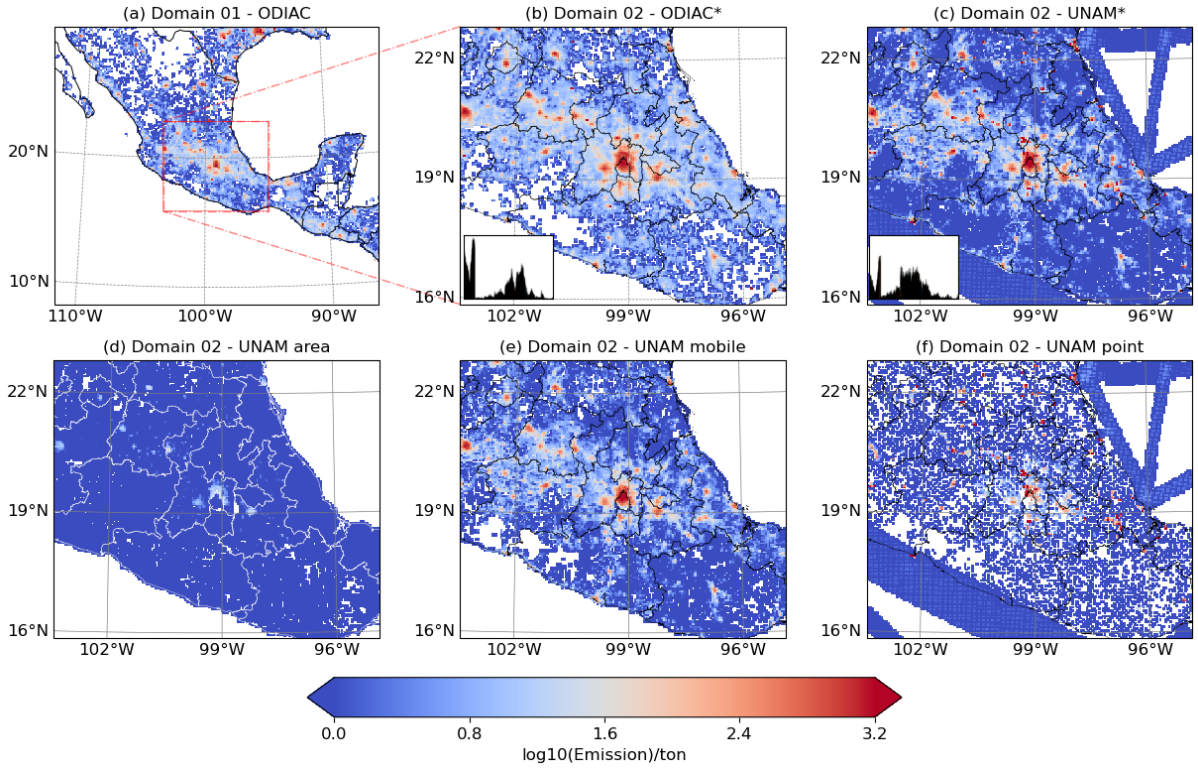


Figure 3: Maps of total fossil fuel CO_2 emissions over (a) domain 1 and domain 2 from (b) ODIAC and (c) the UNAM inventory on 4 January 2018 (UTC time, logarithmic scale, tons). Sectors of UNAM are shown on lower panels: (d) area sources, (e) mobile sources and (f) point sources.

* Frequency distributions (black histograms) are shown over domain 2 to illustrate the larger ratio of near-zero pixels in ODIAC and the presence of large peaks in the UNAM inventory.

▪ Biogenic CO_2 fluxes: optimized CASA simulation

The biogenic fluxes in MCMA were simulated by the Carnegie Ames Stanford Approach (CASA) model (Zhou et al., 2020) and are also interpolated to our domains by xESMF - conservative. The CASA simulation was optimized by generating an ensemble of perturbed parameters, including the maximum light use efficiency, the optimal temperature of photosynthesis, and the temperature response of respiration. These perturbed parameters were constrained by selecting the best configuration in comparison to AmeriFlux eddy-covariance flux data. Monthly Gross Primary Productivity (GPP) and total ecosystem respiration (Re) were downscaled to 3-hourly resolution based on 3-hourly air temperature and shortwave downward radiation from the North American Regional Reanalysis (Mesinger et al., 2006), as described in Olsen and Randerson (2004). GPP and Re were summed into Net Ecosystem Exchange (NEE) at 5-km resolution. Figure 4 shows the difference between dry/wet seasons, with a maximum in carbon uptake during the wet season (growing season). In Figure 4, the phenology of tropical forests (wet and dry) on the east coast of Mexico causes an early start of the growing season (May) while the central and western regions (temperate to semi-arid climates) show negative values (carbon uptake) later in the year (July).

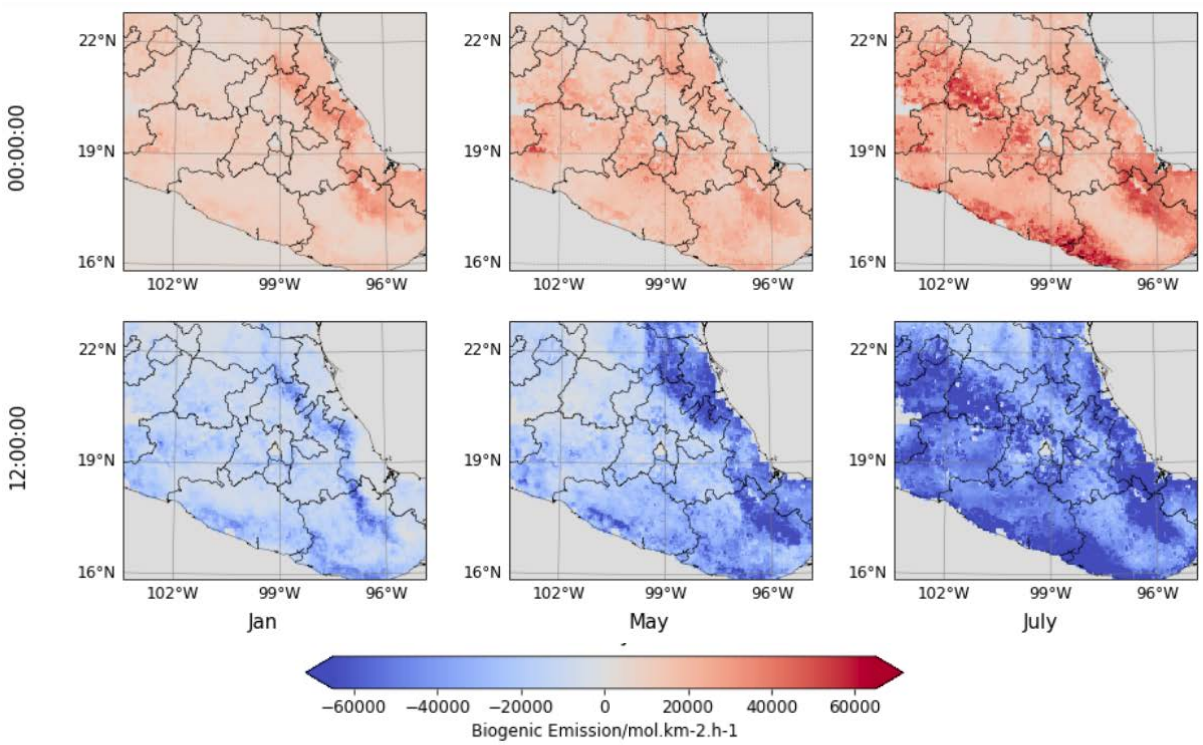


Figure 4: Midday (12:00 UTC-6) and midnight (00:00 UTC-6) biogenic monthly mean fluxes from the CASA biogeochemical model (Zhou et al., 2020) over the WRF simulation domain (5-km resolution) for the months of January (left column), May (middle column) and July (right column) in mol/km^2/hr.

We note here that the central region of our domain, corresponding to Mexico City, shows very low NEE values caused by the dense urban area (sparse vegetation). The biogenic fluxes in the MCMA differ significantly from the surrounding areas, further examined by separating the biogenic contribution in our modeling system.

- ## Atmospheric CO_2 boundary conditions

The influence of distant sources and sinks was simulated by coupling the optimized CO_2 concentrations from the global model CarbonTracker 2019 (Peters et al., 2007; Jacobson et al., 2020) to our WRF domain, similar to Feng et al. (2019) over North America. The coupling scheme involves several steps including pressure correction due to terrain height differences, interpolation of the coarse resolution CO_2 fields to the WRF grid, and mass conservation over the total CO_2 column, as described in Butler et al. (2020). As shown in Figure S1, the CO_2 boundary inflow at our domain boundaries varies at daily and sub-daily timescales (3-hourly in our case) especially over land (northern bound) where air masses from North America flow southward into our domain. We evaluate the CO_2 boundary inflow by comparison to measurements collected at ALZ (mountain site), only occasionally influenced by the city emissions from the valley (cf. section 3.2). We also note here that we couple the 3D fields from the CarbonTracker global models to represent the vertical gradients between the surface (2,200 m a.s.l.) and the mountain tops (4,000 m a.s.l.).

o Meteorological and CO_2 observations

▪ Meteorological surface network

We evaluated the WRF meteorological performances using measurements from the air quality network installed and operated by SEDEMA. Horizontal wind speed and direction, temperature, pressure, and relative humidity were measured across 26 surface stations located in and around Mexico City, publicly available from SEDEMA[8]. The 26 stations are listed in table 1, including the station's names, their geographic coordinates, and the data availability for the year 2018. In previous reports, we found specific comments on possible disturbances affecting wind measurements at three sites. At Merced (MER), located near downtown Mexico City, a three-story school located to the south of the station blocks the air flow from that direction, in addition to an elevated metro railway to the west. At Xalostoc (XAL), the inlet location is partially blocked by a building on one side of the shelter. Finally, at Cuajimalpa (CUA), the site is located on a two-story building with clear air flow in three out of four directions. There is a line of trees on one side of the shelter that may slightly impact wind direction and wind speed measurements, but the meteorological station is at the same altitude or slightly higher than the treeline. In conclusion, most of the SEDEMA meteorological observations match international standards thanks to rigorous data quality control.

ID	Station	Entity	Latitude	Longitude	Altitude	Sampling height/m	Since	% data in 2018
ACO	Acolman	Mexico State	19.6355	-98.9120	2198	13.3	2011	84.35%
AJM	Ajusco Medio	CDMX	19.2722	-99.2077	2548	15	2015	96.21%
AJU	Ajusco	CDMX	19.1543	-99.1626	2942	4	2015	68.29%
BJU	Benito Juarez	CDMX	19.3717	-99.1591	2257	5.5	2015	82.19%
CHO	Chalco	Mexico State	19.2670	-98.8861	2253	7.3	2011	87.03%
CUA	Cuajimalpa	CDMX	19.3653	-99.2917	2704	10.6	2000	98.05%
CUT	Cuautitlán	Mexico State	19.7222	-99.1986	2263	4.5	2012	64.93%
FAC	FES Acatlán	Mexico State	19.4825	-99.2435	2299	10.8	1986	61.37%
GAM	Gustavo A. Madero	CDMX	19.4828	-99.0947	2242	6	2015	98.86%
HGM	Hospital General de México	CDMX	19.4116	-99.1522	2234	15	2012	86.94%
INN	Investigaciones Nucleares	Mexico State	19.2920	-99.3805	3080	4	2015	3.84%
LAA	Lab. de Análisis Ambiental	CDMX	19.4838	-99.1473	2255	5.1	2016	97.59%
MER	Merced	CDMX	19.4246	-99.1196	2245	17.2	1986	98.85%
MGH	Miguel Hidalgo	CDMX	19.4041	-99.2027	2327	11.5	2015	99.77%
MON	Montecillo	Mexico State	19.4604	-98.9029	2252	3.5	2000	97.24%
MPA	Milpa Alta	CDMX	19.1769	-98.9902	2592	8	2016	72.74%
NEZ	Nezahualcóyotl	Mexico State	19.3937	-99.0282	2235	12.1	2011	99.47%
PED	Pedregal	CDMX	19.3252	-99.2041	2326	11.5	1986	96.05%
SAG	San Agustín	Mexico State	19.5330	-99.0303	2241	8.6	1986	78.31%
SFE	Santa Fe	CDMX	19.3574	-99.2629	2599	5.5	2012	84.87%
TAH	Tláhuac	CDMX	19.2465	-99.0106	2297	3.6	2000	75.79%
TLA	Tlalnepantla	Mexico State	19.5291	-99.2046	2311	7	1986	92.36%
UAX	UAM Xochimilco	CDMX	19.3044	-99.1036	2246	18.7	2015	99.71%
UIZ	UAM Iztapalapa	CDMX	19.3608	-99.0739	2221	19.6	2014	64.29%

[8] http://www.aire.cdmx.gob.mx/default.php?opc=%27aKBhnmI=%27&opcion=Zw==

VIF	Villa de las Flores	Mexico State	19.6582	-99.0966	2242	4.5	2000	86.16%
XAL	Xalostoc	Mexico State	19.5260	-99.0824	2160	30	1986	96.72%

Table 1: Coordinates, heights and first year of data acquisition of the SEDEMA meteorological stations in and around Mexico City.

- ## Lidar profiling

The Doppler Lidar located at UNA (Wind Cube 100, Leosphere) was used for wind profile and mixing- and residual-layer height observations. The instruments measure with 50 m vertical resolution backscattered light from a pulsed laser in the infrared that is transmitted to the sky at four different directions and Doppler-shifted, 15º from the zenith, from which a 3-dimensional wind field is reconstructed. Mixing and residual-layer heights are calculated from the aerosol backscattering signal (carrier-to-noise ratio) and the variance of the vertical velocity and reported by the manufacturer. Since January 2017, data is continuously collected by this instrument and stored in the RUOA repository (Red Universitaria de Observatorios Atmosféricos, http://www.ruoa.unam.mx) and the results have been evaluated against other data sources (Burgos-Cuevas et al 2022).

- ## Atmospheric CO_2 measurements

Two high-accuracy analyzers measuring continuously atmospheric CO_2 concentrations were installed in 2014 at the UNAM university campus (UNA) and at the Altzomoni altitude station (ALZ). The two instruments are cavity ring-down spectrometers (model G2401) manufactured by Picarro (Crosson et al., 2008; Rella et al., 2013). Such instruments can measure every few seconds the atmospheric concentrations of CO_2, CH_4, CO and H_2O, with precision compatible with the WMO/GAW requirements (WMO/IAEA 2020). These instruments are now widely used in the community, and have been subject to numerous evaluations documenting their repeatability, drift, and sensitivity to parameters such as temperature and humidity (Chen et al., 2013; Welp et al., 2013; Yver Kwok et al., 2015; 2021). The two analyzers began to be calibrated at the end of 2018 with a calibration scale consisting of three cylinders prepared by NOAA/ESRL, and traceable to the international WMO scale. This reference scale is now installed at the ALZ. A second scale, prepared at LSCE as part of the MERCI-CO_2 project and also traceable to the WMO scale, is installed at the measurement site on the UNAM campus. The instrument calibration made at the end of 2018 was propagated backwards over the whole year 2018, inducing an estimated uncertainty of approximately 0.1 ppm over 2018, according to typical drift for such instruments (Yver-Kwok et al. (2015)). Another major source of uncertainty is related to the water vapor correction, which has to be done in order to express the results as a mole fraction in dry air. Considering the humidity levels observed at the two stations, Gonzalez del Castillo et al. (2022) estimated an uncertainty of 0.2 ppm and 0.05 ppm respectively for UNA and ALZ. The measurement protocols implemented at both sites are described in detail by Gonzalez del Castillo et al. (2022). CO_2 concentrations are available for the two measuring stations UNA and ALZ in 2018 without any data gap.

The two stations are located in very different environments, with a high exposure to urban emissions for UNA and a priori low at the mountain site (ALZ) at 4,000 m a.s.l. This results in significant differences in the diurnal and seasonal cycles (Gonzalez del Castillo et al., 2022). Thus, the daily peak-to-peak amplitude of CO_2 is on average around 35 ppm at UNA and 5 to 7 ppm at ALZ. It has been shown that the main driver of the diurnal variability at UNA is the development of the atmospheric boundary

layer. The daily maximum is reached between 6 and 7 a.m. (local time), and a regular decrease is observed until around 4 p.m. due to the development of the atmospheric boundary layer leading to a dilution of the compounds emitted at the surface, like CO_2, in a larger volume. CO_2 then accumulates during the night as the atmospheric boundary layer decreases associated with greater vertical stability. In the case of ALZ, Gonzales del Castillo et al. (2022) explained that the diurnal cycle resulted from a combination of the influences of boundary layer dynamics and biospheric fluxes. They also assume that the increase in concentration observed in the afternoon is probably linked to the uplift of polluted air masses from Mexico City and other surrounding urban areas.

▪ Atmospheric CO_2 column mixing ratios

In addition to the CO_2 surface measurements, CO_2 total column measurements were performed at the ALZ and UNA stations using solar absorption FTIR spectroscopy. The ALZ station was equipped in 2012 with a high resolution FTIR spectrometer (model IFS-120/5HR, Bruker), which continuously provides vertical column densities of atmospheric trace gases officially contributing to the Network for the Detection of Atmospheric Composition Change (NDACC). The dry air column-averaged mole fractions of CO_2 (XCO_2) used in this study were calculated from the high resolution spectra (0.02 cm^{-1}) measured with KBr beamsplitter, using the PROFFIT 9.6 code (Hase et al., 2004) and adopting the CO_2 and O_2 retrieval methods described in Baylon et al., (2017). Several post-process quality filters were applied to discard data affected by clouds, volcanic ash or low signal, based on both spectra and retrieval quality indicators (signal to noise ratio, relative RMS, wavenumber shift, etc.) and statistical criteria (standard errors). The UNA station was equipped in 2016 by a mobile low resolution (0.5 cm^{-1}) EM27/SUN spectrometer, providing continuous total column measurements of CO_2, CO and CH_4. Low-resolution spectra are processed following the COllaborative Carbon Column Observing Network (COCCON) processing chain (PREPROCESS and PROFFAST-version 01-07-2018 codes) developed by the Karlsruhe Institute of Technology (KIT) and freely available (https://www.imk-asf.kit.edu/english/COCCON.php). A full description of the analytical procedure is given in Sha et al. (2020), Frey et al. (2021), and Pollard et al. (2022). Calibration coefficients and ILS determined by the KIT (https://www.imk-asf.kit.edu/downloads/Coccon/2019-11-08_Instrument-Calibration.pdf for the used PROFFAST version) were applied according to the COCCON recommendations. We used the GGG2014 version of TCCON meteorological data and priors (daily MAPs files), downloaded from the CalTech server. Intraday ground pressure data were taken from the co-located RUOA pressure sensor measurements (https://www.ruoa.unam.mx/). UNA and ALZ XCO_2 data were inter-calibrated using data obtained from several weeks of ALZ side-by-side measurements. Noisy outliers were filtered out according to a 3-sigma criterion.

• Results

o Meteorological performances

▪ Surface meteorological evaluation

Model simulations were performed over three months (January, May, and July), representing three contrasted seasons in the region of Mexico-City (Jáuregui, 2002). January is the coldest month of the year with temperatures between 6 and 22°C, with low precipitation, during the dry season. During the 2 weeks of January 2018, the average temperature of Mexico City, calculated from the 26 weather stations,

varies from 2.33 to 21.32°C with a mean of 11.66 ± 4.48 °C (Figure S2). Conversely, the month of May is the hottest month of the year in Mexico City (13 to 26°C), with a mean temperature of 20.01 ± 3.42 °C recorded in 2018 (Figure S2). The month of July is characterized by lower temperatures compared to May (18.37 ± 3.05 °C in July 2018), but with higher precipitation rates in the city (52 mm) corresponding to the rainy season.

The model succeeds in simulating most of the seasonal, synoptic and diurnal variability of the average temperature deduced from the network of 26 SEDEMA stations (Figure S2). In May and July, the mean daily biases are respectively equal to 0.61°C and 0.53°C, but the peak-to-peak amplitude is underestimated by 3.03 and 2.26°C. For most of the days, night minimums are overestimated by about 2°C and daily maximums are underestimated by about 0.4-0.8°C. On the contrary, the average temperatures of the 26 stations over the month of January are underestimated by about 0.33°C. This positive anomaly may be related to residential and commercial heating in winter. Correlation coefficients for the near-surface temperature (3.5 to 30 m a.g.l.) between observation and simulation are very high (0.95, 0.97 and 0.92 for January, May and July separately). The fit of our regression model (R-square) is larger than 0.85. The simulated diurnal cycle in May is highly correlated with observations (R=0.93) simulated, including the sharp drop after sunset on clear nights and the smooth variations on cloudy nights.

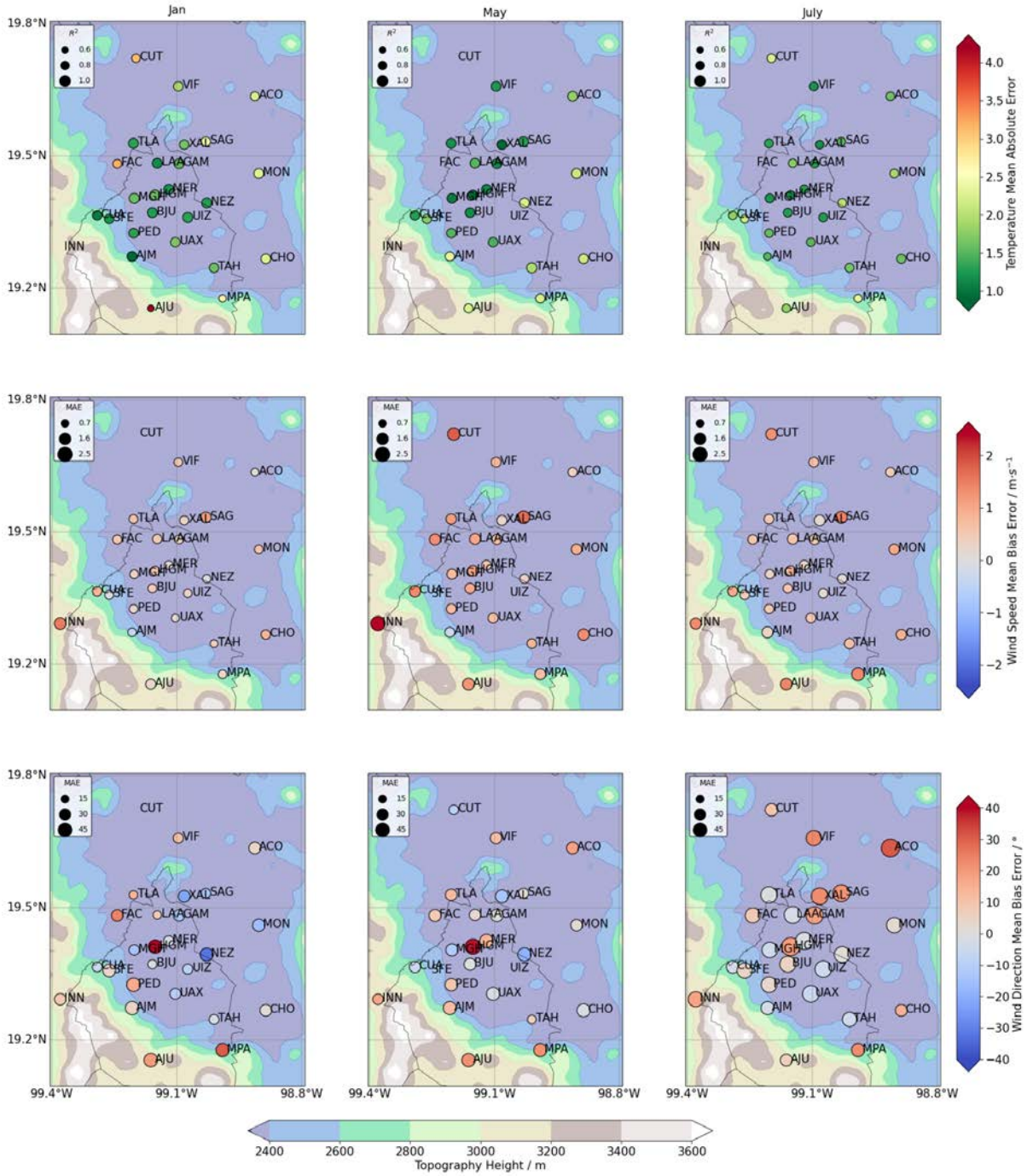


Figure 5. Mean Absolute Error (MAE) in first-level temperature (0-16m; upper row) and Mean Error in wind speed (middle row) and wind direction (bottom row) for the months of January (left column), May (middle column) and July (right column) of the year 2018.

Observed and simulated surface wind speeds were compared at the 26 meteorological stations. The average Mean Absolute Errors (MAEs) for wind speed vary from 0.72 to 1.67 m/s across the station network. For January, May and July, the network-wide MAE of the 26 stations is 0.98, 1.40 and 1.29 m/s, respectively. As shown in Figure 5 and Figure S3, mountain stations (e.g. INN, AJU and MPA) show degraded performances (1.5-2.5 m/s overestimated). Urban stations correspond to a mean bias of less than 1 m/s and a regression fit (R-square) larger than 0.5. Surface wind directions vary across stations with MAE's between 50° to 70° (cf. Figure S4). Unlike wind speed, the simulated wind directions vary across the basin independently of the topography, possibly due to the misrepresentation of

building heights and the use of default parameters in the urban canopy model. Overall, meteorological performances are similar to previous urban studies without data assimilation (Feng et al., 2016; Lian et al., 2018).

- Vertical mixing and horizontal wind evaluation

We compared the vertical wind measurements collected by the Lidar deployed at UNA (South city-side) to our simulations. The scanning elevation of the wind reconstruction covers a 15-degree angle from the zenith (75° to 90°; every scan contains 4 directions). The average wind speed from the 4 directions is used for the model evaluation of the wind speed. As for the model, the approximate height above the surface was calculated using the geopotential height:

$$Height = (PH + PHB)/g - \mathrm{H}GT$$

in which *PH*, *PHB*, *HGT* and g stand for perturbation geopotential height, base-state geopotential height, terrain height and gravitational acceleration, respectively.

Two vertical wind error profiles from the Lidar and the model are shown in Figure 6 to illustrate our model-data performances (left panel), in addition to the MAEs from 0 to 4,000 m above ground level (right panel). The modeled vertical gradients in wind profiles are consistent with observations. Since the sample heights of the model and the Lidar are different, MAE was calculated based on the observation records and the simulation wind speed (49 layers) interpolated to observation layer heights (99 heights). MAEs are less than 2 m/s in both cases. In January, the mean error and mean absolute errors in the horizontal mean wind are -0.36 m/s and 1.28 m/s respectively, with standard deviations of 0.9 m/s and 0.6 m/s. In May, -0.72 m/s and 2.00 m/s, with standard deviations of 1.7 m/s and 1.1 m/s. Both errors increase for layers higher than 2000 m, which corresponds to the maximum observing elevation of the Lidar instrument.

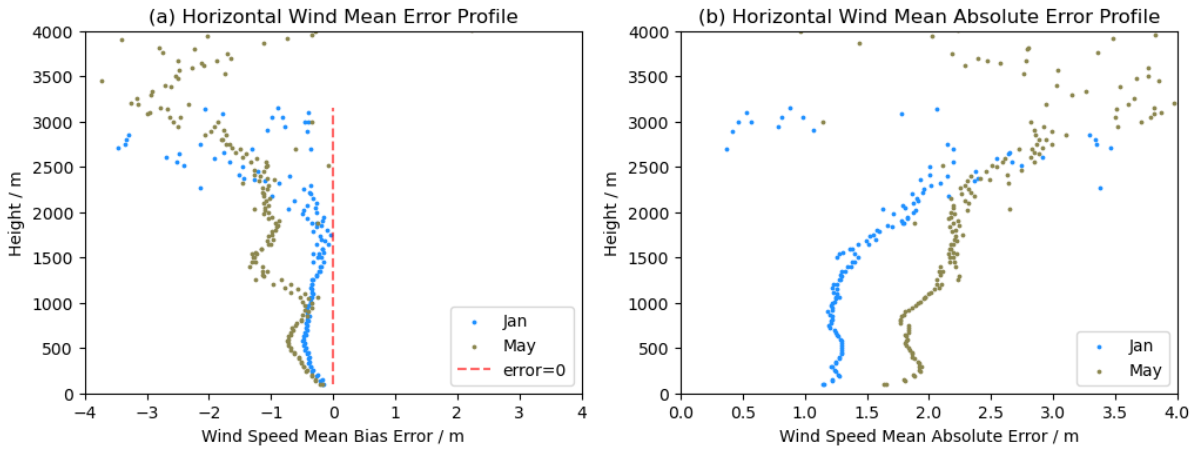


Figure 6. Vertical profile of 2-week mean horizontal wind speed errors at UNA in the afternoon: (a) mean error profile (b) mean absolute error profile.

We have also used the PBL heights deduced from the Lidar installed at UNA. Two atmospheric layer heights are diagnosed from raw Lidar data: a mixing layer height and a residual layer height. As shown in Figure 7, mixing layer heights diagnosed by the Lidar correspond to the top of the mixed layers but fail to capture the maximum elevations deduced from the residual heights during the afternoons. Conversely, the residual layer heights correspond to the top of the mixed layer during the afternoon, while measuring the actual residual layer at night. To avoid a negative bias in Lidar measurements during

daytime, we combined the two products to derive the most representative height of the mixed layer during daytime. The mixing layer diagnostic is used in the morning (local time 8:00-13:00) while residual layer heights are used to evaluate the mixed layer height in the afternoon (local time 14:00-18:00). To evaluate the temporal variations of the modeled PBL heights, we performed a linear regression for the months of January and May, with estimated R-square values of 0.68 and 0.64, respectively. The mean bias during morning hours is about 15 m in January and about 280 m in May. During the afternoon hours, the bias reaches -580 m (in January) and 110 m (in May). These values are similar to the study of Feng et al. (2016) over the basin of Los Angeles using aircraft PBL heights. During the afternoons, the modeled PBL heights are usually lower than the Lidar measurements during the month of January. We note here that no Lidar measurement is available for the month of July 2018.

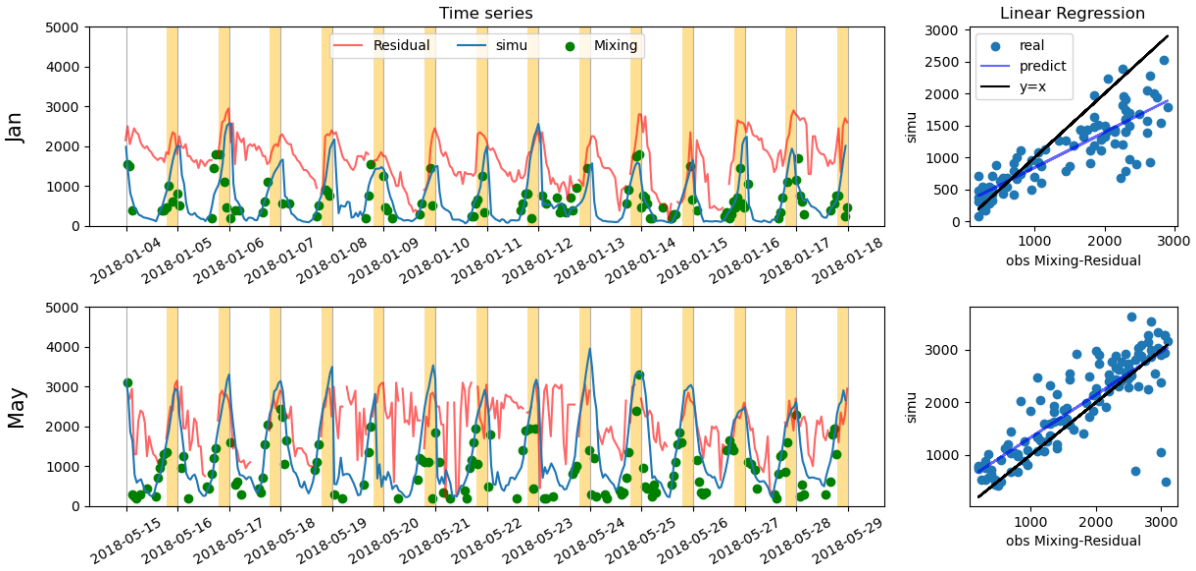


Figure 7. Time series and linear regression of PBL height (in meters) simulation (in blue) and observation ("residual layer" heights in red lines and mixing layer heights in green dots) in UNA (time in UTC, golden background stands for local time UTC-6 afternoon: 14:00-18:00)

- Atmospheric CO_2 contributions from fossil fuel and biogenic sources and sinks

  - Atmospheric in situ CO_2 concentrations

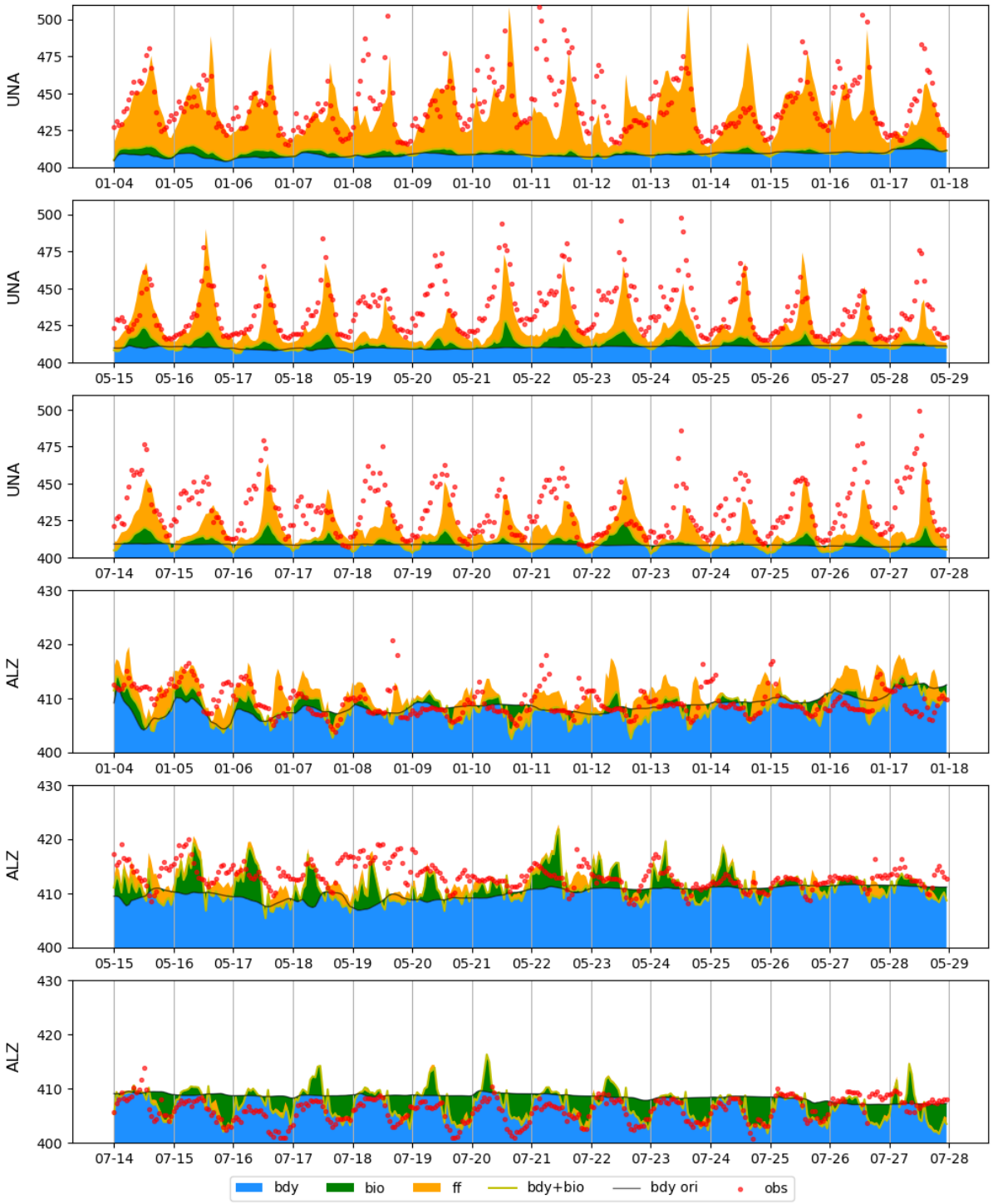


Figure 8. Observed and modeled in situ CO_2 concentrations (in ppm) at both UNA and ALZ stations for the months of January, May, and July 2018, in UTC time. The modeled concentrations were separated into three components: boundary inflow (bdy - in blue), biogenic fluxes (bio - in green) and fossil fuel emissions (ff - in orange). The contributions from the boundary conditions are indicated by the

gray line and the combined contribution of boundaries and biogenic fluxes (bdy+bio) by the light yellow line.

Figure 8 shows the comparison between the simulated and the observed in situ CO_2 concentrations at both sites (UNA and ALZ). Modeled concentrations are decomposed into three components: the background concentrations originating from the CarbonTracker inversion system (in blue), the fossil fuel contributions from Mexico City and from the entire country (in orange), and the biogenic sources and sinks caused by surface fluxes from the CASA model (in green). The WRF-Chem modeled concentrations tend to be under-estimated during the first hours of the night (about 10 ppm), revealing a time lag in the accumulation of CO_2 in stable conditions. At both stations, the model-data differences are about -0.31 ppm (UNA) and -0.71 ppm (ALZ) during the afternoon hours (cf. Table S1). While the UNA station is influenced by large fossil fuel contributions (10 ppm during the afternoon hours), the ALZ station shows relatively low contributions (1.8 ppm). The biogenic contributions vary across the two stations (2.0 and -0.27 ppm), indicating that future optimizations (flux inversions) will need to adjust separately the biogenic fluxes. The spatial distribution of biogenic CO_2 fluxes is presented in the discussion section, due to differences in ecosystem types between the Mexico City basin and in the surrounding mountains. We note here a mismatch at ALZ during the first week of May (15 to 20 May 2018), possibly caused by incorrect boundary conditions.

We show the observed and modeled in situ CO_2 concentration mean diurnal cycles at both sites in Figure 9. The lack of accumulation at night at UNA results in a time lag between the modeled and observed nighttime peaks of about 2 hours. During the afternoon hours (14:00-18:00 local time), the WRF-Chem CO_2 concentrations capture the temporal variations and the absolute values at UNA (413-424 ppm, MAE of 3-5 ppm). Due to its remote location, the ALZ station shows a reduced diurnal cycle amplitude, simulated by WRF-Chem, but mostly driven by the boundary CO_2 concentrations from CarbonTracker (404-410 ppm, MAE of 2-3 ppm). We confirm here the ability of our modeling system to reproduce the urban enhancements and the large-scale boundary conditions at monthly timescale.

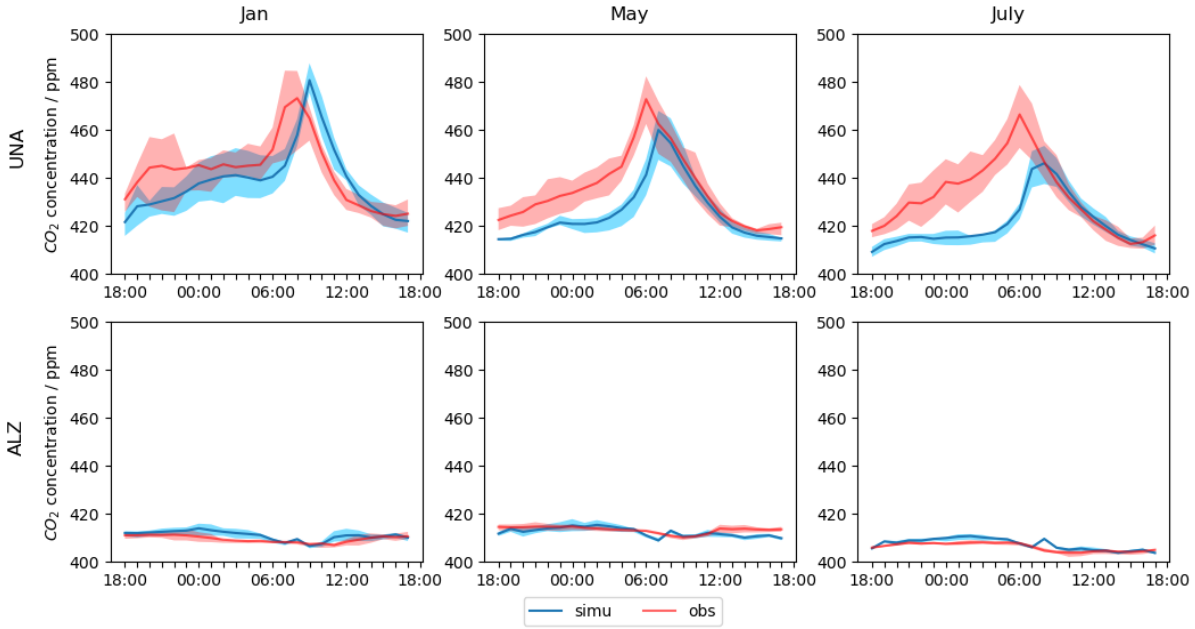


Figure 9. Mean diurnal variations in atmospheric CO_2 concentrations at UNA (upper row) and ALZ (lower row) stations observed (in pink) and simulated by WRF-Chem (in blue) for the months of January (left), May (middle), and July (right) of the year 2018, shown in local time.

- Atmospheric CO_2 column concentrations

We show in Figure 10 the model performances to simulate XCO_2 at both stations. Because EM27 instruments only measure during daytime clear-sky conditions, we focus on the afternoon hours. We note that, due to poor weather conditions, there are nearly no afternoon FTIR measurements complying with the data quality requirements during July 2018. The influence of local sources and sinks is reduced by the dilution of PBL signals within the entire column of air, but their influence remains the main driver of the observed day-to-day variability at UNA. WRF-Chem tends to over-estimate XCO_2 concentrations in January (1.01 ppm) and under-estimate XCO_2 in May (0.53 ppm; or 0.15 ppm if the abnormal period 15-17 May was excluded). The over-estimation of XCO_2 concentrations in January does not coincide with in situ CO_2 concentration differences, possibly caused by a positive bias in background concentrations rather than by local emissions. The day-to-day variations were better captured in May (R=0.81, except for the abnormal values in 15-17 May 2018) compared to January (R=0.63) (cf. Figure S6). Overall at UNA, the MAE for XCO_2 is about 1 ppm (1.22 ppm in January and 0.97 in May), indicating that background concentrations should be optimized in future inversions. Regarding the period 15-17 May 2018 during which the wind direction shifted from a northerly flow to a westerly flow, WRF-Chem failed at capturing the increase in XCO_2 (Figure S6). In general, the lack of data in the rainy season impairs our ability to quantify the XCO_2 model errors at ALZ and UNA but the absence of fossil fuel signals at ALZ indicates that the local circulation from the valley to the mountain tops is absent during summer months.

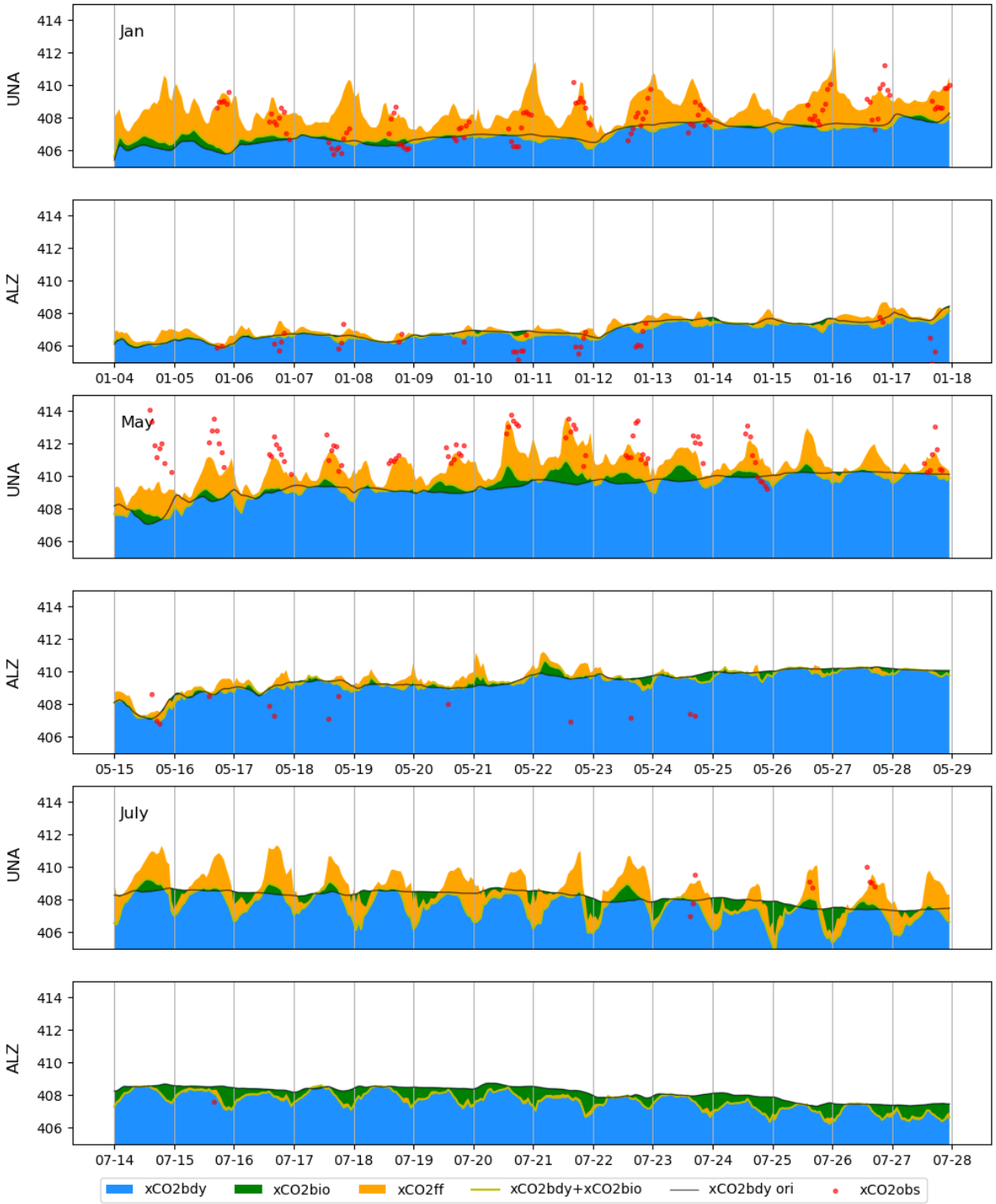


Figure 10. Observed and modeled Atmospheric column CO_2 concentrations (in ppm) at both UNA and ALZ stations for the months of January, May, and July 2018, in UTC time. The modeled concentrations were separated into three components: boundary inflow (bdy - in blue), biogenic fluxes (bio - in green) and fossil fuel emissions from both simulation domains 1 and 2 (ff - in orange). The contributions from the boundary conditions are indicated by the gray line and the combined contribution of boundaries and biogenic fluxes (XCO_2bdy+XCO_2bio) by the light yellow line.

• Conclusions and Discussion

In this paper we evaluated the performances of the WRF-Chem model to simulate the CO_2 and XCO_2 concentrations over MCMA. The meteorological conditions (surface temperature, wind speed and wind direction) were evaluated across the SEDEMA observation network of 26 stations. Within and around the city, the Mean Absolute Errors (MAE's) in wind speed are around 1 m/s, while the wind speed errors were 1.5-2.5 m/s over-estimated by the model in the mountainous areas outside the city. The regression model of measured and modeled PBL heights at UNA are significant with R-square values larger than 0.64 for both seasons, while the mean errors in horizontal wind speed at UNA are all less than 1 m/s within the Planetary Boundary Layer (PBL), confirming our simulation captures the local dynamics across the basin. Regarding CO_2 concentrations, the model captured most of the observed variations in the afternoon, with a mean bias of -0.52 ppm and a MAE of 3.42 ppm. WRF-Chem tends to under-estimate the peak values in the morning, showing a 1-hour time lag in the simulation of CO_2 accumulation patterns at night. Thus, only the afternoon data will be assimilated in our future inversion. The model also prompts us that the fossil fuel signals at our background station (ALZ) are low enough (1.8 ppm) to represent the background CO_2 concentrations, but the biogenic flux patterns differ between the background (ALZ) and the downtown (UNA) stations. We conclude here that the gradients between the two stations (UNA-ALZ) are sufficient to detect and to quantify the city emissions in our future inversion, but the biogenic fluxes should also be optimized separately. Except for the large mismatch in wind direction over the period 15-17 May 2018, XCO_2 day-to-day trends were captured by the model during both seasons. The mean bias was around 1.00 ppm for January and -0.53 ppm for May, most of which was driven by the wind direction shift on 15-17 May 2018, indicating boundary conditions should also be included in future inversions of XCO_2.

o Seasonality of fossil fuel contributions

We examine in greater details the spatial and temporal distributions of fossil fuel signatures in our modeled CO_2 concentrations from the mobile sector in Figure 11. Mobile CO_2 enhancements concentrate in the city center, representing up to 90% of the modeled local enhancements. In the southern part of the city, north to the UNA station, the contribution reaches its maximum value while the northern suburbs are dominated by industrial and energy production sources. The Tula power plant and additional surrounding industries generate a second plume north of Mexico City, where mobile sources have a minor role in the observed enhancements. In terms of emissions, the seasonality of mobile sources remains low (around 4%). The largest mobile signal originates from the dense road network within CDMX. Several satellite cities show significant signals - Toluca de Lerdo to the west, Tizayuka and Pachuca de Soto to the northeast, and Ixtapaluca to the southeast (plumes in Figure 13, bottom row). While the absolute contribution from mobile sources varies with seasons, the magnitude can be explained by the local atmospheric dynamics. The strong convective activity during the rainy season mixes air masses from the boundary layer and the free Troposphere (convective mixing through updraft and downdraft). As a result, the highest daily-averaged mobile contribution during the rainy season (12 ppm) is half of that during the dry season (22 ppm). We conclude that, with only two stations, future atmospheric inversions would be limited to constraining the emissions from traffic, while missing an important fraction of the energy production sources. Future stations should be deployed in the northern part of MCMA, the outflow of the urban basin. Additionally, a refinement of the fossil fuel emissions inventory (possibly divided into sectors of activities or SNAP's) would be needed to determine the sectoral trends in more detail.

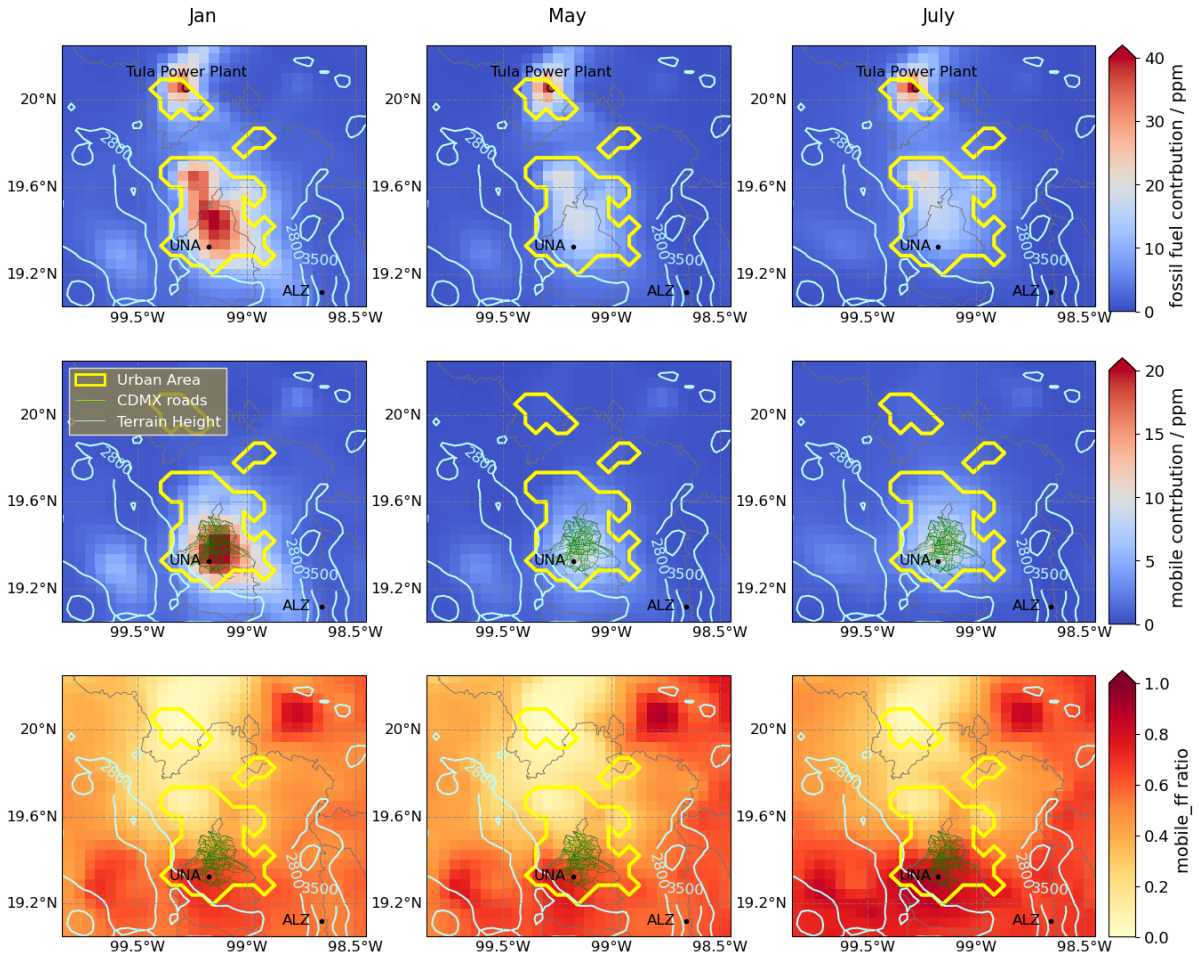


Figure 11. Atmospheric CO_2 concentration absolute contributions from fossil fuel sources (top row), from mobile sources (middle row), and relative contribution of mobile sources to total fossil fuel (bottom row) in the PBL during the dry (left) and rainy (middle and right columns) seasons for the year 2018, in ppm. Topographic contours are indicated in light blue, administrative boundaries in gray, the urban area of MCMA in yellow, and the road network in CDMX in green.

o Biospheric contribution variability

In our study, we used the biogenic flux contribution from the CASA biogeochemical model, optimized using eddy-covariance flux observations from the Ameriflux network (Zhou et al., 2020). While the optimization helps to constrain the model parameters, mediterranean and mountainous ecosystems lack flux measurement stations to constrain the model parameters. However, the use of disturbance forest maps and previous evaluation of CASA suggest that the simulated flux gradients should be representative of the area, with grassland and cropland in the valley north of the city, and forested land in the mountains. We discuss here how the ecosystem diversity, hence their resulting CO_2 fluxes, translate into CO_2 spatial gradients in our simulated concentrations.

As shown in Figure 12, in the early morning, the valley of Mexico remains positively-influenced by ecosystems (ecosystem respiration dominates) due to late exposure to incoming radiation compared to the surrounding mountains. Mid-afternoon hours tend to show no-to-little spatial gradients across our simulation domain, corresponding to a homogeneous uptake of carbon. At night, the largest release of

CO_2 from respiration takes place in the valleys, west and north of Mexico City. The positive fluxes remain low at both ALZ (above treeline) and at UNA (within the dense urban area).

Despite the fact that the vegetation tends to be more active during the rainy season, we show no significant increase in the biogenic contribution at UNA nor ALZ (Figure S8), consistent with previous studies (Sánchez-León et al., 2016). During the dry season, the lack of water can prevent vegetation from growing, while the rainy and cloudy conditions reduce the sunlight during the rainy season. Across the urban area (including UNA), vegetation coverage remains low where population density is high. ALZ lies at high altitude (4000 m a.s.l.), above the treeline, hence without strong carbon uptake. In July, ALZ experiences a colder and drier climate than UNA, as shown in Figure S9. The photosynthesis is higher than that in May, but the largest absorption is only 5 ppm and the net contribution over two weeks is about -1.5 ppm. Overall, we conclude here that both sites correspond to similar biogenic influences during daytime, while early morning fluxes differ significantly.

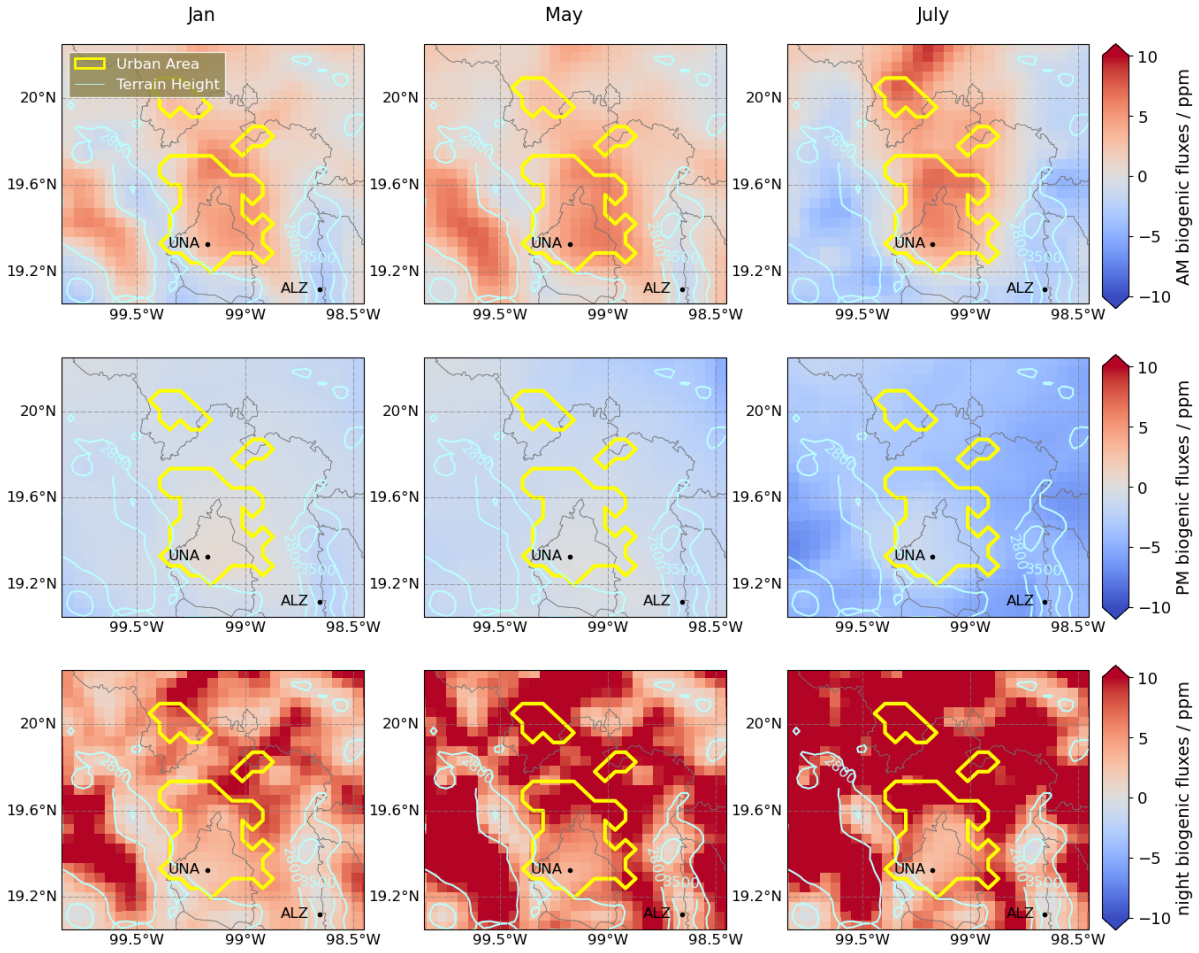


Figure 12. Mean biogenic contribution to surface CO_2 of the three months in (top) AM: 8:00-12:00 local time (middle) PM: 13:00-17:00 local time (bottom) Night: 23:00-4:00 local time

o Optimal network designs for CO_2 inversions

Based on our modeled results, we discuss here the potential use of UNA and ALZ measurements to constrain the city emissions, and we examine potential measurement sites to enhance the current network. We analyze the spatial distribution of background concentrations to monitor large-scale boundary concentrations, and the spatial distribution of fossil fuel signals to maximize the city's signals in our

future inversion system. Although human activities vary seasonally, such as heating in winters (not common in the city) and air conditioning in summers, the distribution of fossil fuel signals remains relatively stable, concentrated around the city and the Tula power plant.

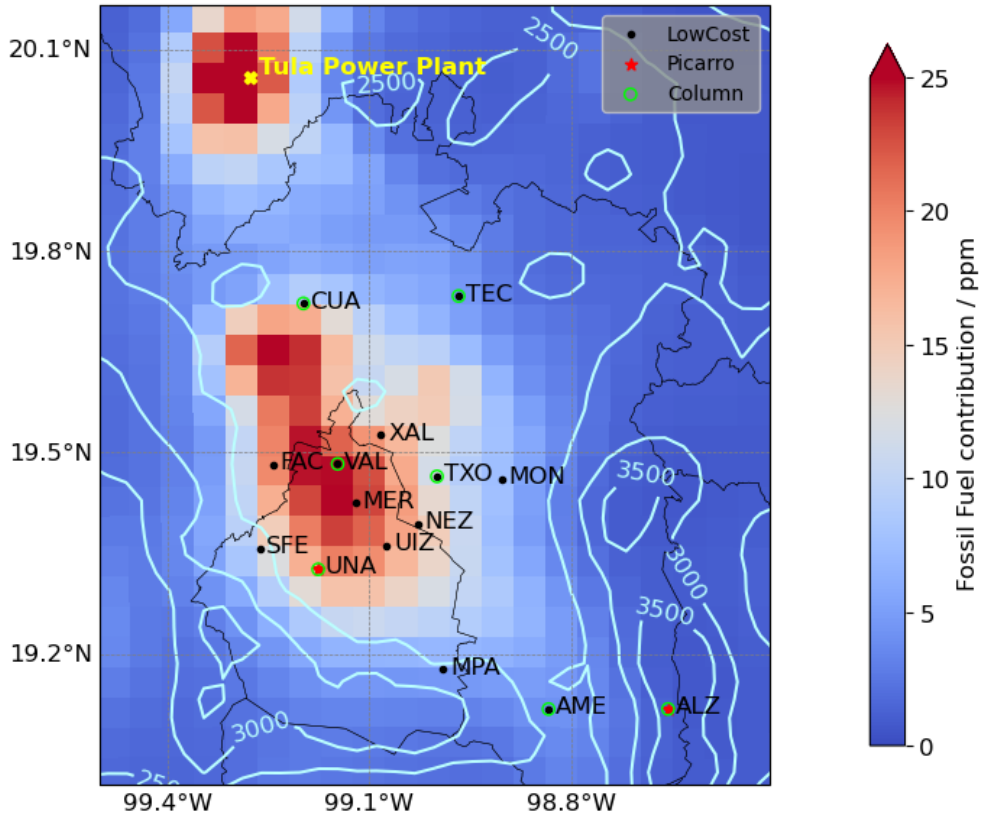


Figure 13. Atmospheric fossil fuel CO_2 concentrations (50 meters agl) simulated by WRF and averaged over the three months of 2018, with current and potential MERCI-CO_2 observation station locations and elevation contour lines (in light blue).

The current downtown station (UNA) is located near the city center while ALZ is located further away from the emitting area, at a much higher altitude. A recently-deployed column station (VAL) is located in the city center, co-located with the modeled CO_2 concentration maximum, while CUA, TEC, TXO and AME are on the edge of urban areas, able to capture the background conditions under various wind conditions. For additional measurement locations, the largest unobserved sources are located in the north of MCMA, around and from the Tula power plant where no station has been identified yet. An additional station able to monitor the northern part of CDMX would allow us to monitor emissions from non-transportation sources (industry and residential sectors), possibly CUA and FAC. Considering our existing stations, UNA is located near the maximum fossil fuel contributions (in terms of CO_2 concentrations) while ALZ observes little to no fossil fuel signals.

To assess the quality of the background station (ALZ), and determine additional locations for complementing our background conditions, we calculated the Spearman correlation coefficient between the modeled background values of UNA with each grid in and around MCMA, similar to Feng et al. (2016). Figure 14 shows the correlation map including (top row) and excluding (bottom row) biogenic fluxes, primarily to determine if biogenic fluxes should be included in our future inversion system as unknowns, or assumed to be directly observed by our background station. We clearly see that biogenic fluxes significantly decrease correlations between the valley and UNA, suggesting that biogenic fluxes are not

observed and should be optimized separately. When excluding the biogenic influence, correlations remain high at 0.87 in January, 0.91 in May, and 0.95 in July.

We conclude here that ALZ remains an optimal location for observing the background conditions when excluding the biogenic influence. Potential locations highly-correlated with the UNA background and biogenic fluxes are located north of the city, possibly impacted by the Tula power plant plume in January and May (northerly flow). Therefore, we conclude that a second background station would allow us to evaluate our modeled background concentrations but, based on our current modeling system, would be redundant with the ALZ station.

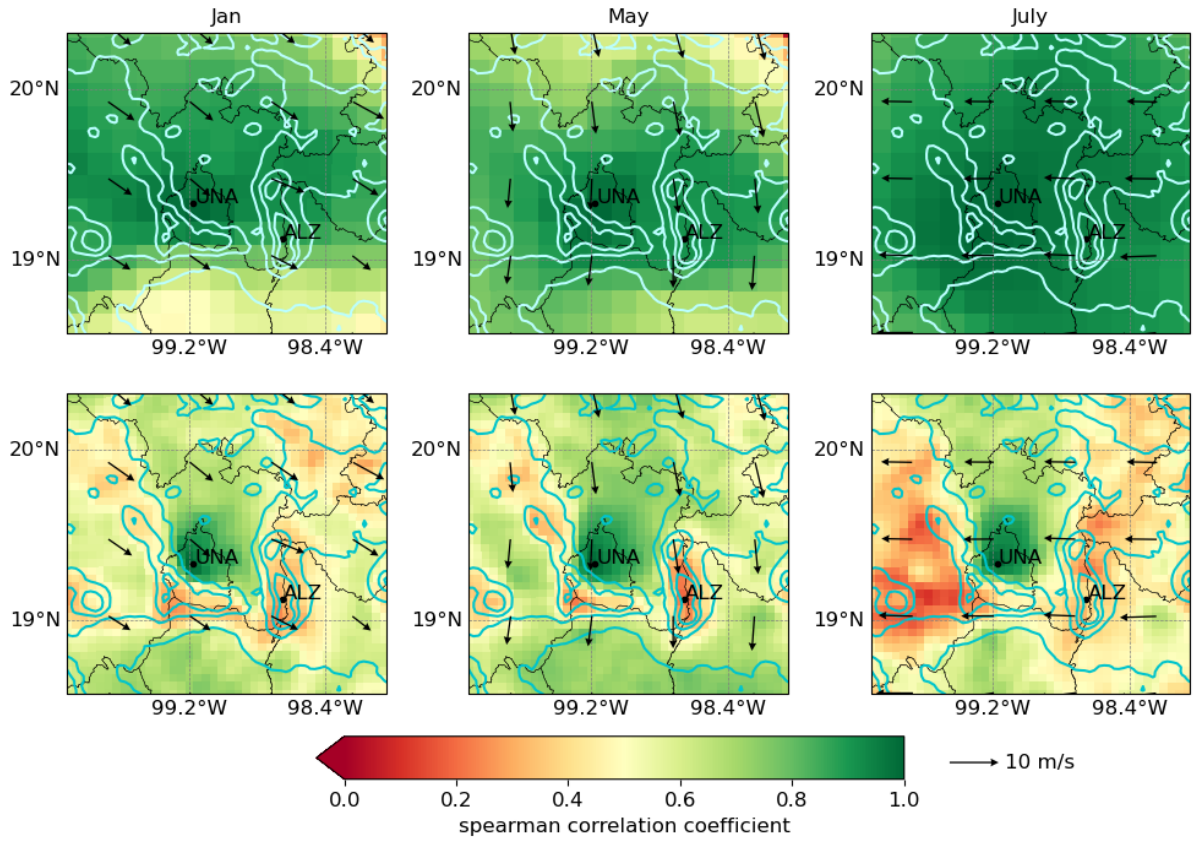


Figure 14. Spearman correlation coefficient maps of background (top) and background + biogenic (bottom) contributions with troposphere wind direction at 3.5km for the three selected months of 2018.

Acknowledgement

Financial support as part of the French-Mexican project MERCI-CO_2 through grants ANR (ANR-17-CE04-0013-01) and CONACYT (290589) are acknowledged. The RUOA network (www.ruoa.unam.mx) provided surface CO_2 and wind profiling data. Special thanks to Alejandro Bezanilla and Omar Lopez for instrument operation and technical assistance. T. Lauvaux was supported by the project CIUDAD (Make Our Planet Great Again program) and the Chair Fellowship CASAL (French Ministry of Research - CNRS).

Reference

Arzoumanian, E., Vogel, F. R., Bastos, A., Gaynullin, B., Laurent, O., Ramonet, M., & Ciais, P. (2019). Characterization of a commercial lower-cost medium-precision non-dispersive infrared sensor for atmospheric CO_2 monitoring in urban areas. Atmospheric Measurement Techniques, 12(5), 2665-2677.

Baylon, J. L., Stremme, W., Grutter, M., Hase, F., & Blumenstock, T. (2017). Background CO_2 levels and error analysis from ground-based solar absorption IR measurements in central Mexico. Atmospheric Measurement Techniques, 10(7), 2425-2434.

Boden, T. A., Marland, G., and Andres, R. J. (2017). Global, Regional, and National Fossil-Fuel CO_2 Emissions, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., USA, https://doi.org/10.3334/CDIAC/00001_V2017.

Bréon, F. M., Broquet, G., Puygrenier, V., Chevallier, F., Xueref-Remy, I., Ramonet, M., ... & Ciais, P. (2015). An attempt at estimating Paris area CO_2 emissions from atmospheric concentration measurements. Atmospheric Chemistry and Physics, 15(4), 1707-1724.

Burgos-Cuevas, A., Magaldi, A., Adams, D.K. et al. Boundary Layer Height Characteristics in Mexico City from Two Remote Sensing Techniques. Boundary-Layer Meteorol (2022). https://doi.org/10.1007/s10546-022-00759-w

Butler, M. P., Lauvaux, T., Feng, S., Liu, J., Bowman, K. W., & Davis, K. J. (2020). Atmospheric simulations of total column CO_2 mole fractions from global to mesoscale within the carbon monitoring system flux inversion framework. Atmosphere, 11(8), 787.

Chen, F, Kusaka, H, Bornstein, R, Ching, J, Grimmond, CSB, Grossman-Clarke, S, Loridan, T, Manning, KW, Martilli, A, Miao, S, Sailor, D, Salamanca, FP, Taha, H, Tewari, M, Wang, X, Wyszogrodzki, AA and Zhang, C. (2011). The integrated WRF/urban modelling system: development, evaluation, and applications to urban environmental problems. International Journal of Climatology 31(2): 273– 288. DOI: https://doi.org/10.1002/joc.2158

Chen, G., Shan, Y., Hu, Y., Tong, K., Wiedmann, T., Ramaswami, A., ... & Wang, Y. (2019). Review on city-level carbon accounting. Environmental science & technology, 53(10), 5545-5558.

Chen, H., Karion, A., Rella, C. W., Winderlich, J., Gerbig, C., Filges, A., Newberger, T., Sweeney, C., and Tans, P. P. (2013). Accurate measurements of carbon monoxide in humid air using the cavity ring-down spectroscopy (CRDS) technique, Atmospheric Measurement Techniques, 6, 1031–1040. https://doi.org/10.5194/amt-6-1031-2013

Chen, J., Zhao, F., Zeng, N., & Oda, T. (2020). Comparing a global high-resolution downscaled fossil fuel CO_2 emission dataset to local inventory-based estimates over 14 global cities. Carbon Balance and Management, 15(1), 1-15.

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.

Deng, A., Lauvaux, T., Davis, K. J., Gaudet, B. J., Miles, N., Richardson, S. J., ... & Gurney, K. R. (2017). Toward reduced transport errors in a high resolution urban CO_2 inversion system. Elementa: Science of the Anthropocene, 5.

Elvidge, C. D., Baugh, K. E., Zhizhin, M., and Hsu, F.-C. (2013). Why VIIRS data are superior to DMSP for mapping nighttime lights, Proceedings of the Asia-Pacific Advanced Network, 35, 62–69, https://doi.org/10.7125/apan.35.7.

Feng, S., Lauvaux, T., Davis, K. J., Keller, K., Zhou, Y., Williams, C., ... & Baker, I. (2019). Seasonal characteristics of model uncertainties from biogenic fluxes, transport, and large-scale boundary inflow in atmospheric CO_2 simulations over North America. Journal of Geophysical Research: Atmospheres, 124(24), 14325-14346.

Feng, S., Lauvaux, T., Newman, S., Rao, P., Ahmadov, R., Deng, A., ... & Yung, Y. L. (2016). Los Angeles megacity: a high-resolution land–atmosphere modeling system for urban CO_2 emissions. Atmospheric Chemistry and Physics, 16(14), 9019-9045.

Frey, M. M., Hase, F., Blumenstock, T., Dubravica, D., Groß, J., Göttsche, F., ... & Pollard, D. F. (2021). Long-term column-averaged greenhouse gas observations using a COCCON spectrometer at the high-surface-albedo site in Gobabeb, Namibia. Atmospheric Measurement Techniques, 14(9), 5887-5911.

García-Reynoso, JA, Mar-Morales, BE, & Ruiz-Suárez, LG (2018). Spatial, temporal and speciation distribution model of the Mexican emissions inventory (base year 2008) for use in air quality modeling (DiETE). International Journal of Environmental Pollution, 34 (4), 635-649. https://doi.org/10.20937/rica.2018.34.04.07.

González del Castillo, E., Taquet, N., Bezanilla, A., Stremme, W., Ramonet, M., Laurent, O., ... & Grutter, M. (2022). CO_2 variability in the Mexico City region from in situ measurements at an urban and a background site. Atmosphere , 35 (2), 377-393.

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., & Eder, B. (2005). Fully coupled “online” chemistry within the WRF model. Atmospheric environment, 39(37), 6957-6975.

Gurney, K. R., Liang, J., Roest, G., Song, Y., Mueller, K., & Lauvaux, T. (2021). Under-reporting of greenhouse gas emissions in US cities. Nature communications, 12(1), 553.

Hase, F., Hannigan, J. W., Coffey, M. T., Goldman, A., Höpfner, M., Jones, N. B., ... & Wood, S. W. (2004). Intercomparison of retrieval codes used for the analysis of high-resolution, ground-based FTIR measurements. Journal of Quantitative Spectroscopy and Radiative Transfer, 87(1), 25-52.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J-N. (2017): Complete ERA5 from 1979: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service (C3S) Data Store (CDS). (Accessed on 21-Jan-2021)

INEGI (Instituto Nacional de Estadística y Geografía, Mexico National Institute of Statistics and Geography) Mexico (2021). Registered Motor Vehicles in Circulation, available at: https://www.inegi.org.mx/programas/vehiculosmotor/ (last access: 29 December 2022)

Jacobson, A. R., Schuldt, K. N., Miller, J. B., Oda, T., Tans, P., Arlyn Andrews, Mund, J., Ott, L., Collatz, G. J., Aalto, T., Afshar, S., Aikin, K., Aoki, S., Apadula, F., Baier, B., Bergamaschi, P., Beyersdorf, A., Biraud, S. C., Bollenbacher, A., … Miroslaw Zimnoch. (2020). CarbonTracker CT2019B. NOAA Global Monitoring Laboratory. https://doi.org/10.25925/20201008

Lauvaux, T., Gurney, K. R., Miles, N. L., Davis, K. J., Richardson, S. J., Deng, A., ... & Turnbull, J. (2020). Policy-relevant assessment of urban CO_2 emissions. Environmental Science & Technology, 54(16), 10237-10245.

Lauvaux, T., Miles, N. L., Deng, A., Richardson, S. J., Cambaliza, M. O., Davis, K. J., ... & Wu, K. (2016). High-resolution atmospheric inversion of urban CO_2 emissions during the dormant season of the Indianapolis Flux Experiment (INFLUX). Journal of Geophysical Research: Atmospheres, 121(10), 5213-5236.

Lian, J., Wu, L., Bréon, F. M., Broquet, G., Vautard, R., Zaccheo, T. S., ... & Ciais, P. (2018). Evaluation of the WRF-UCM mesoscale model and ECMWF global operational forecasts over the Paris region in the prospect of tracer atmospheric transport modeling. Elementa: Science of the Anthropocene, 6.

Lopez-Coto, I., Hicks, M., Karion, A., Sakai, R. K., Demoz, B., Prasad, K., & Whetstone, J. (2020). Assessment of planetary boundary layer parameterizations and urban heat island comparison: Impacts and implications for tracer transport. Journal of applied meteorology and climatology, 59(10), 1637-1653. https://doi.org/10.1175/JAMC-D-19-0168.1

Lopez-Coto, I., Ren, X., Salmon, O. E., Karion, A., Shepson, P. B., Dickerson, R. R., ... & Whetstone, J. R. (2020). Wintertime CO_2, CH4, and CO emissions estimation for the Washington, DC–Baltimore metropolitan area using an inverse modeling technique. Environmental science & technology, 54(5), 2606-2614.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., et al. (2006). North American regional reanalysis. Bulletin of the American Meteorological Society, 87(3), 343–360. https://doi.org/10.1175/bams-87-3-343

Mueller, K. L., Lauvaux, T., Gurney, K. R., Roest, G., Ghosh, S., Gourdji, S. M., ... & Whetstone, J. (2021). An emerging GHG estimation approach can help cities achieve their climate and sustainability goals. Environmental Research Letters, 16(8), 084003.

Nakanishi, M., & Niino, H. (2009). Development of an improved turbulence closure model for the atmospheric boundary layer. Journal of the Meteorological Society of Japan. Ser. II, 87(5), 895-912.

Nassar, R., Napier-Linton, L., Gurney, K. R., Andres, R. J., Oda, T., Vogel, F. R., & Deng, F. (2013). Improving the temporal and spatial distribution of CO_2 emissions from global fossil fuel emission data sets. Journal of Geophysical Research: Atmospheres, 118(2), 917-933.

Nickless, A., Rayner, P. J., Engelbrecht, F., Brunke, E. G., Erni, B., & Scholes, R. J. (2018). Estimates of CO_2 fluxes over the City of Cape Town, South Africa, through Bayesian inverse modelling. Atmospheric Chemistry and Physics, 18(7), 4765-4801. https://doi.org/10.5194/acp-18-4765-2018

NOAA, https://gml.noaa.gov/ccgg/trends/global.html (last access: 2022/12/29)

Olivier, J.G.J., Janssens-Maenhout, G., Muntean, M. and Peters, J.A.H.W. (2015) Trends in global CO_2 emissions: 2015 Report. PBL Netherlands Environmental Assessment Agency, The Hague; European

Commission, Joint Research Centre (JRC), Institute for Environment and Sustainability (IES). Internet: http://edgar.jrc.ec.europa.eu/news_docs/jrc-2015-trends-in-global-CO_2-emissions-2015-report-98184.pdf. JRC report 98184/ PBL report 1803, 2015.

Olsen, S. C., & Randerson, J. T. (2004). Differences between surface and column atmospheric CO_2 and implications for carbon cycle re- search. Journal of Geophysical Research, 109(D2). https://doi.org/10.1029/2003JD003968

Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., ... & Tans, P. P. (2007). An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker. Proceedings of the National Academy of Sciences, 104(48), 18925-18930.

Pollard, D. F., Hase, F., Sha, M. K., Dubravica, D., Alberti, C., & Smale, D. (2022). Retrievals of X CO_2, X CH 4 and X CO from portable, near-infrared Fourier transform spectrometer solar observations in Antarctica. Earth System Science Data, 14(12), 5427-5437.

Rella, C. W., Chen, H., Andrews, A. E., Filges, A., Gerbig, C., Hatakka, J., ... & Zellweger, C. (2013). High accuracy measurements of dry mole fractions of carbon dioxide and methane in humid air. Atmospheric Measurement Techniques, 6(3), 837-860.

Rodríguez Zas, J. A., & García Reynoso, J. A. (2021). Actualización del inventario nacional de emisiones de 2013 para la modelación de la calidad del aire en el centro de México. Revista Internacional de Contaminación Ambiental, 37, 463–487.

Sánchez-León, E., Castro, T., Peralta, Ó., Álvarez-Ospina, H., Espinosa, M. D. L. L., & Martínez-Arroyo, A. (2016). Estimated carbon dioxide exchange for three native species in an ecological reserve of Mexico City. Atmósfera, 29(3), 189-196.

Schuh, A. E., Otte, M., Lauvaux, T., & Oda, T. (2021). Far-field biogenic and anthropogenic emissions as a dominant source of variability in local urban carbon budgets: A global high-resolution model study with implications for satellite remote sensing. Remote Sensing of Environment, 262, 112473.

SEDEMA (2016). Inventario de Emisiones de la CDMX 2014. Dirección General de Calidad del Aire, Dirección de Proyectos de Calidad del Aire. Ciudad de México (http://www.aire.cdmx.gob.mx/descargas/publicaciones/flippingbook/inventario-emisiones-cdmx2014-2/mobile/index.html)

SEDEMA (2021). Inventario de Emisiones de la Zona Metropolitana del Valle de Méxi$CO_2$018. Dirección General de Calidad del Aire, Dirección de Proyectos de Calidad del Aire. Ciudad de México. Agosto (http://www.aire.cdmx.gob.mx/descargas/publicaciones/flippingbook/inventario-emisiones-cdmx-2018/Inventario-de-emisiones-cdmx-2018.pdf)

Seto, K., Bigio, A., Blanco, H., Delgado, G. C., Dewar, D., Huang, L., Inaba, A., Kansal, A., Lwasa, S., McMahon, J., Müller, D. B., Murakami, J., Nagendra, H., & Ramaswami, A. (2015). Mitigation of climate change. In C. B. Field et al. (Eds.), Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (pp. 361–409). United Kingdom and New York, NY: Cambridge University Press.

Sha, M. K., De Mazière, M., Notholt, J., Blumenstock, T., Chen, H., Dehn, A., ... & Weidmann, D. (2020). Intercomparison of low-and high-resolution infrared spectrometers for ground-based solar remote sensing measurements of total column concentrations of CO_2, CH 4, and CO. Atmospheric Measurement Techniques, 13(9), 4791-4839.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D., Duda, M. G., … Powers, J. G. (2008). A Description of the Advanced Research WRF Version 3 (No. NCAR/TN-475+STR). University Corporation for Atmospheric Research. doi:10.5065/D68S4MVH

Somanathan E., T. Sterner, T. Sugiyama, D. Chimanikire, N.K. Dubash, J. Essandoh-Yeddu, S. Fifita, L. Goulder, A. Jaffe, X. Labandeira, S. Managi, C. Mitchell, J. P. Montero, F. Teng, and T. Zylicz, 2014: National and Sub-national Policies and Institutions. 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. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel and J.C. Minx (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Staufer, J., Broquet, G., Bréon, F. M., Puygrenier, V., Chevallier, F., Xueref-Rémy, I., ... & Ciais, P. (2016). The first 1-year-long estimate of the Paris region fossil fuel CO_2 emissions based on atmospheric inversion. Atmospheric Chemistry and Physics, 16(22), 14703-14726.

Tewari, M, Chen, F, Kusaka, H and Miao, S. (2007). Coupled WRF/Unified Noah/urban-canopy modeling system. NCAR WRF Documentation 122: 1–22. NCAR, Boulder

Tomohiro Oda, Shamil Maksyutov (2015), ODIAC Fossil Fuel CO_2 Emissions Dataset (Version name : ODIAC2019), Center for Global Environmental Research, National Institute for Environmental Studies, doi:10.17595/20170411.001. (Reference date : 2022/05/22)

United Nations, Department of Economic and Social Affairs, Population Division (2018). The World's Cities in 2018—Data Booklet (ST/ESA/ SER.A/417).

Welp, L. R., Keeling, R. F., Weiss, R. F., Paplawsky, W., and Heckman, S.: Design and performance of a Nafion dryer for continuous operation at CO_2 and CH4 air monitoring sites, Atmos. Meas. Tech., 6, 1217–1226, https://doi.org/10.5194/amt-6-1217-2013, 2013.

Wigley, T. (1983). The pre-industrial carbon dioxide level. Climatic change, 5(4), 315-320.

WMO (2020), 20th WMO/IAEA Meeting on Carbon Dioxide, Other Greenhouse Gases and Related Tracers Measurement Techniques (GGMT-2019), Jeju Island, South Korea, GAW Rep. No. 255, Geneva, Switzerland.

xESMF: Universal Regridder for Geospatial Data, https://github.com/pangeo-data/xESMF (last access: 24 Jan 2022), https://doi.org/10.5281/zenodo.4294774.

Yadav, V., Ghosh, S., Mueller, K., Karion, A., Roest, G., Gourdji, S. M., ... & Whetstone, J. (2021). The impact of COVID-19 on CO_2 emissions in the Los Angeles and Washington DC/Baltimore metropolitan areas. Geophysical research letters, 48(11), e2021GL092744.

Ye, X., Lauvaux, T., Kort, E. A., Oda, T., Feng, S., Lin, J. C., ... & Wu, D. (2020). Constraining fossil fuel CO_2 emissions from urban area using OCO-2 observations of total column CO_2. Journal of Geophysical Research: Atmospheres, 125(8), e2019JD030528.

Yver Kwok, C., Laurent, O., Guemri, A., Philippon, C., Wastine, B., Rella, C. W., ... & Ramonet, M. (2015). Comprehensive laboratory and field testing of cavity ring-down spectroscopy analyzers measuring H 2 O, CO_2, CH 4 and CO. Atmospheric Measurement Techniques, 8(9), 3867-3892.

Yver-Kwok, C., Philippon, C., Bergamaschi, P., Biermann, T., Calzolari, F., Chen, H., ... & Wyss, S. (2021). Evaluation and optimization of ICOS atmosphere station data as part of the labeling process. Atmospheric Measurement Techniques, 14(1), 89-116.

Zhou, Y., C. A. Williams, T. Lauvaux, K. J. Davis, S. Feng, I. Baker, S. Denning, & Y. Wei. (2020). A multiyear gridded data ensemble of surface biogenic carbon fluxes for North America: Evaluation and analysis of results. Journal of Geophysical Research: Biogeosciences, 125(2), e2019JG005314. https://doi.org/10.1029/2019JG005314

Supplementary Information: Evaluation of atmospheric CO_2 simulation over Mexico City metropolitan area by WRF-chem

Y. Xu[1], T. Lauvaux[2], M. Grutter[3], N. Taquet[3], JA. García-Reynoso[3], O.Laurent[1], M. Lopez[1], J. Lian[1], X. Lin[1], W. Stremme[3], M. Ramonet[1]

[1] Laboratoire des Sciences du Climat et de l'Environnement (LSCE), IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France

[2] Groupe de Spectrométrie Moléculaire et Atmosphérique (GSMA), Université de Reims Champagne Ardenne, UMR CNRS 7331, Reims, France

[3] Centro de Ciencias de la Atmósfera, Universidad Nacional Autónoma de México, México

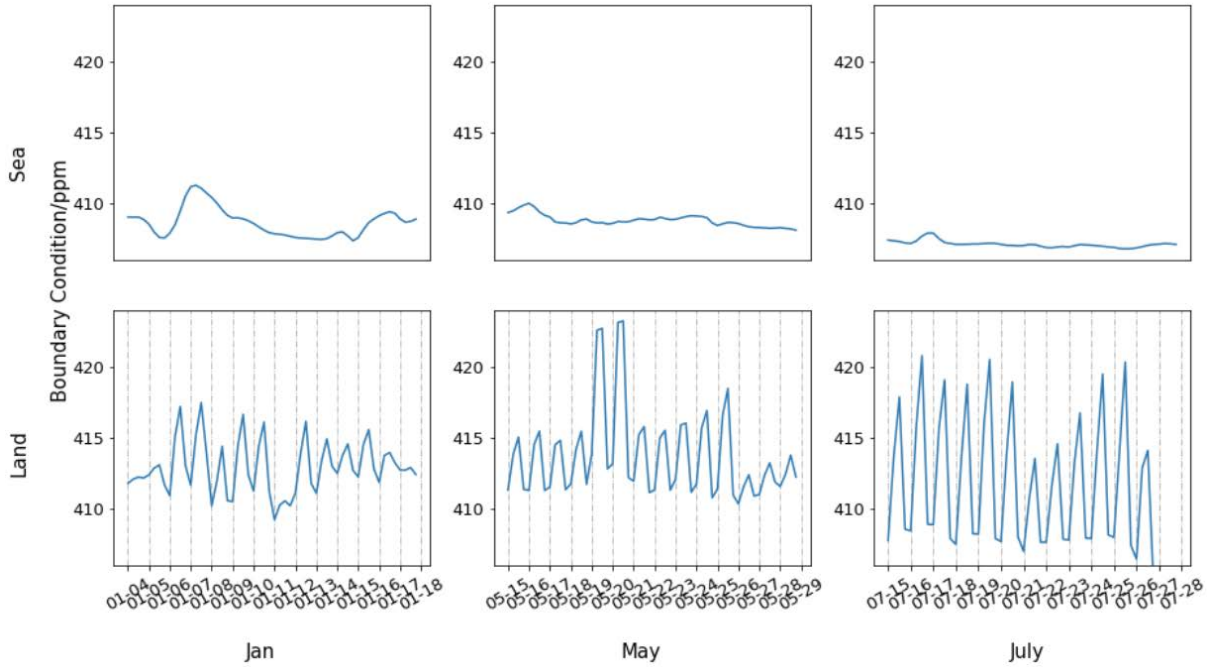


Figure S1. Time series of the CO_2 boundary concentrations at the southern (sea) and northern (land) model boundaries from the CarbonTracker inversion system, in UTC time (local time=UTC-6) for January, May, and July 2018.

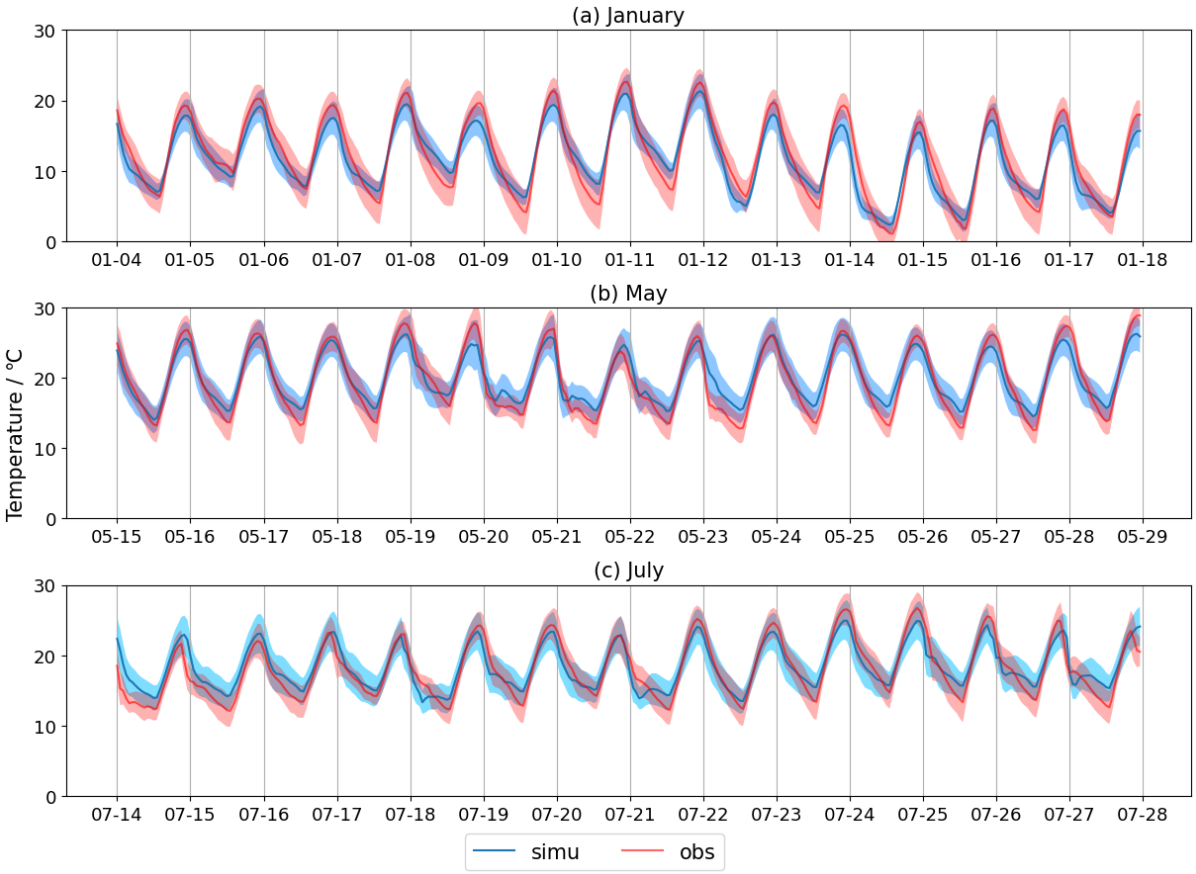


Figure S2. Time series of the mean hourly temperature over the 26 meteorological stations in and around Mexico City, in UTC time (local time=UTC-6), for the months of (a) January, (b) May, and (c) July of 2018.

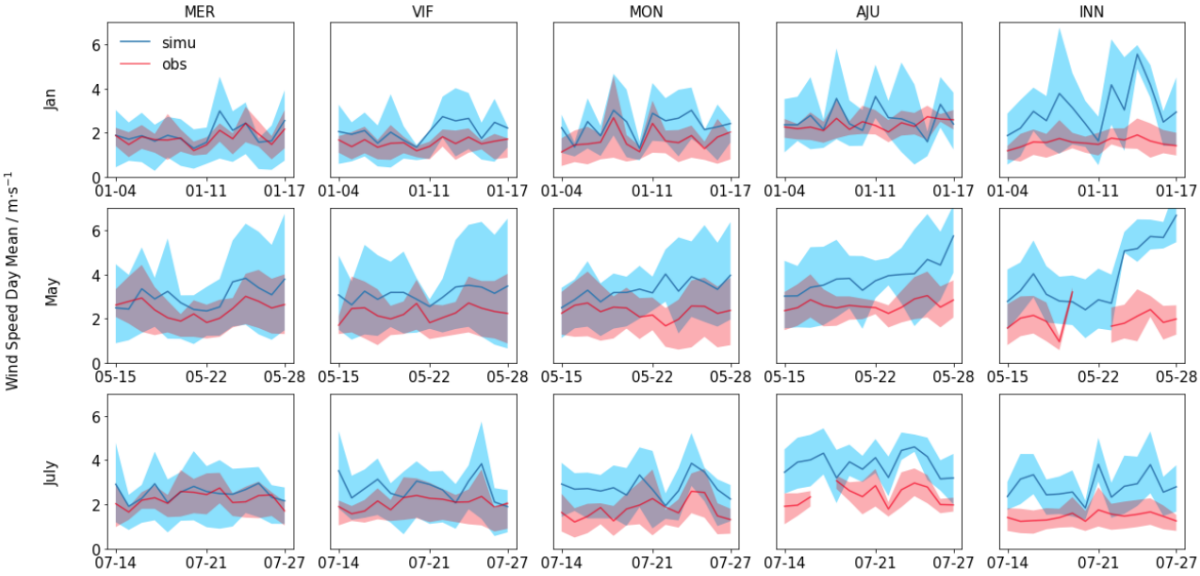


Figure S3. Wind speed daymean with interquartile range in January, May and July 2018 at 5 of the 26 stations: MER in the city center, VIF in the north, MON in the east, AJU in the south (mountains), and INN in the west (mountains), in UTC time (local time=UTC-6).

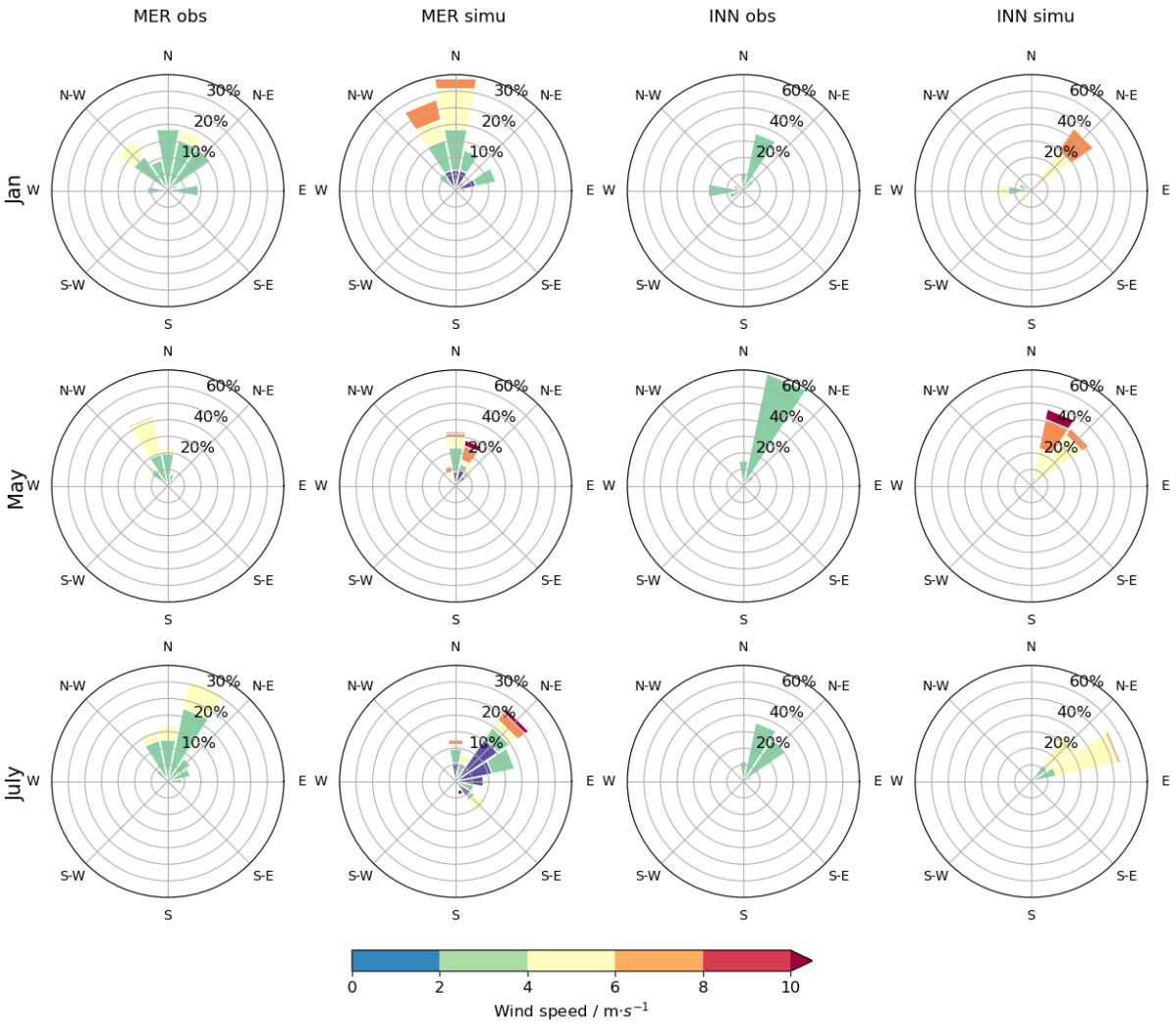


Figure S4. Simulated and observed wind roses during daytime hours (07:00-17:00 local time, UTC-6) at two different meteorological stations: MER (in the city center) and INN (in surrounding mountains) over the 3 comparison period in January, May and July 2018, when observed wind speed > 2m/s.

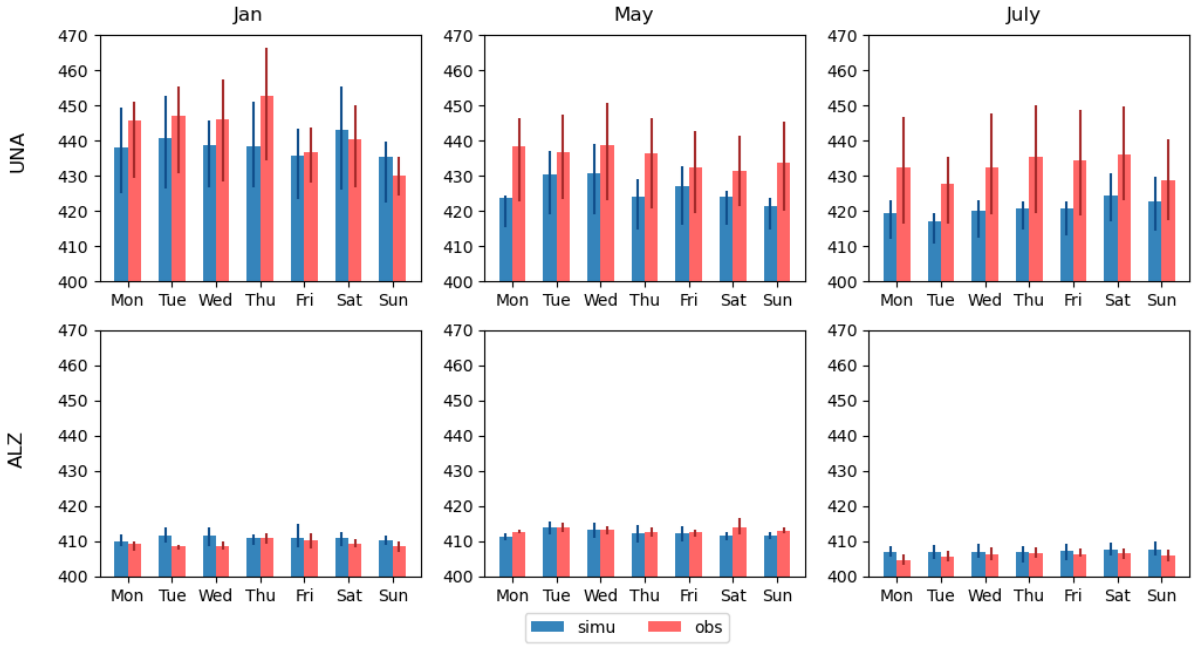


Figure S5. Averaged afternoon CO_2 concentrations (in ppm) with interquartile range for the different days of the week at UNA (upper row) and ALZ (lower row) for the months of January (left column), May (middle), and July (right) of 2018.

Unit: ppm		Jan		May		July	
		UNA	**ALZ**	**UNA**	**ALZ**	**UNA**	**ALZ**
Mean Error	**AM**	2.25	0.97	-2.51	-0.37	-0.80	1.67
	PM	1.93	0.83	-2.86	-2.91	0.00	-0.04
Mean Absolute Error	**AM**	21.63	2.60	9.04	2.32	8.77	2.69
	PM	5.70	2.63	3.54	3.05	3.88	1.75

Table S1. Mean errors and mean absolute errors in atmospheric CO_2 concentrations at 16-48 meters a.g.l. for AM (7:00-12:00 local time UTC-6) and PM (13:00-17:00 local time UTC-6) in the months of January (left column), May (middle), and July (right) of 2018.

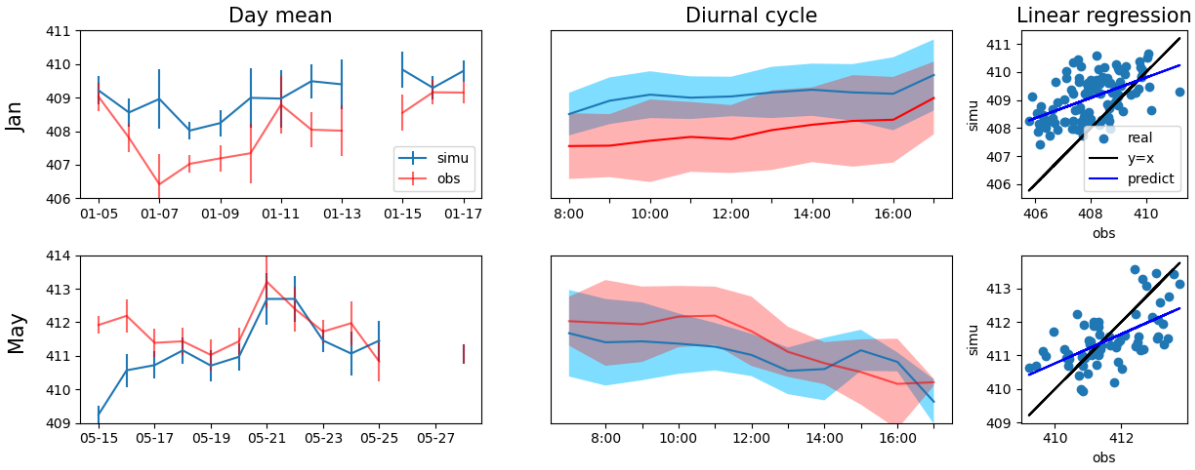


Figure S6. Daily mean XCO_2 concentrations (left panels) and mean XCO_2 diurnal cycles (middle panels, local time UTC-6) at the UNA station, observed by the EM-27/SUN instrument and simulated by WRF-Chem, for the months of January and May 2018. Observed and simulated XCO_2 concentrations (right panels), with the corresponding linear regressions (blue lines, abnormal period May 15-17 excluded) at UNA for the same months.

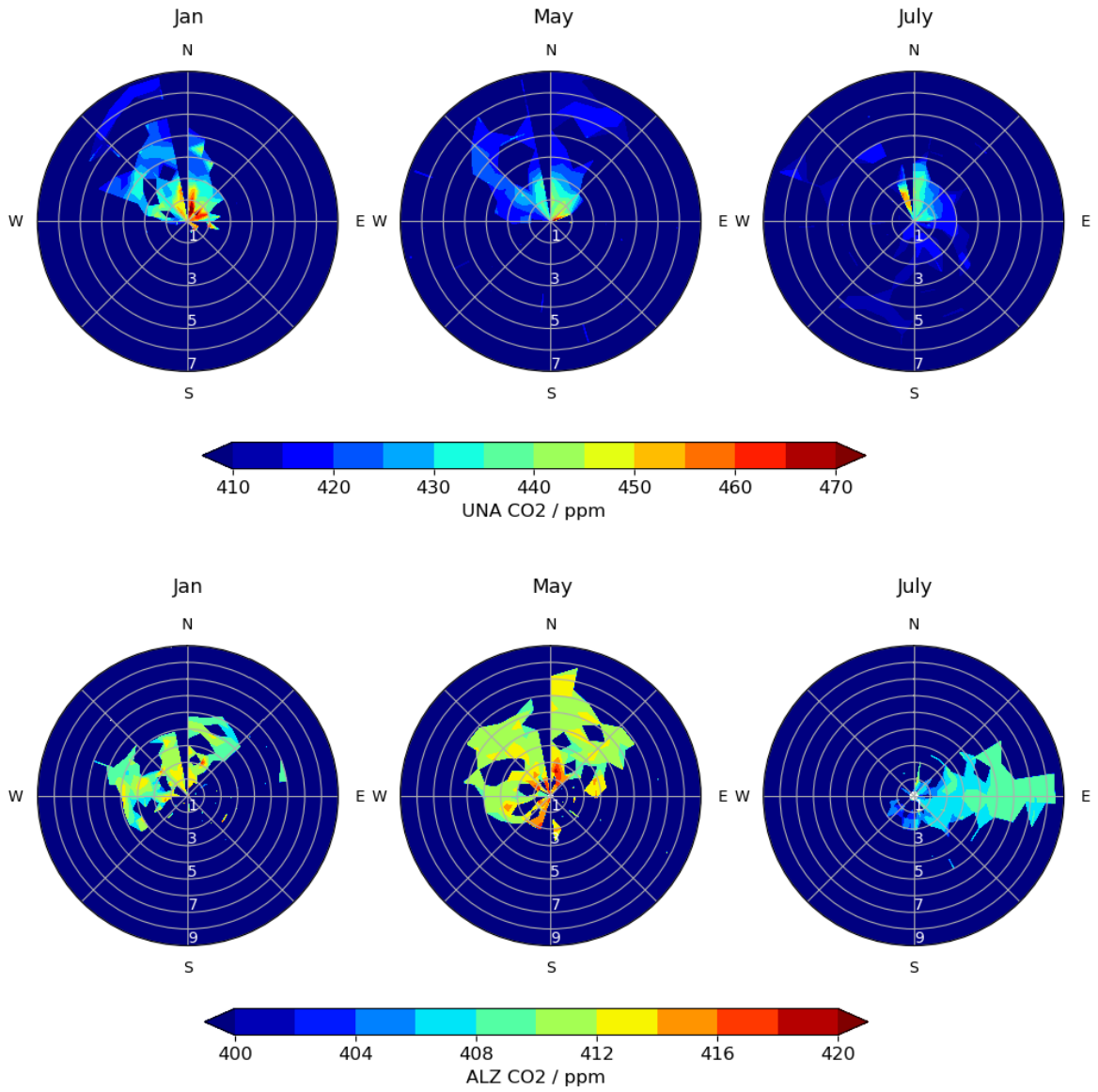


Figure S7. Wind roses of in situ CO_2 concentrations at UNA and ALZ for 2-week period in the month January (left column), May (middle column) and July (right column) of 2018.

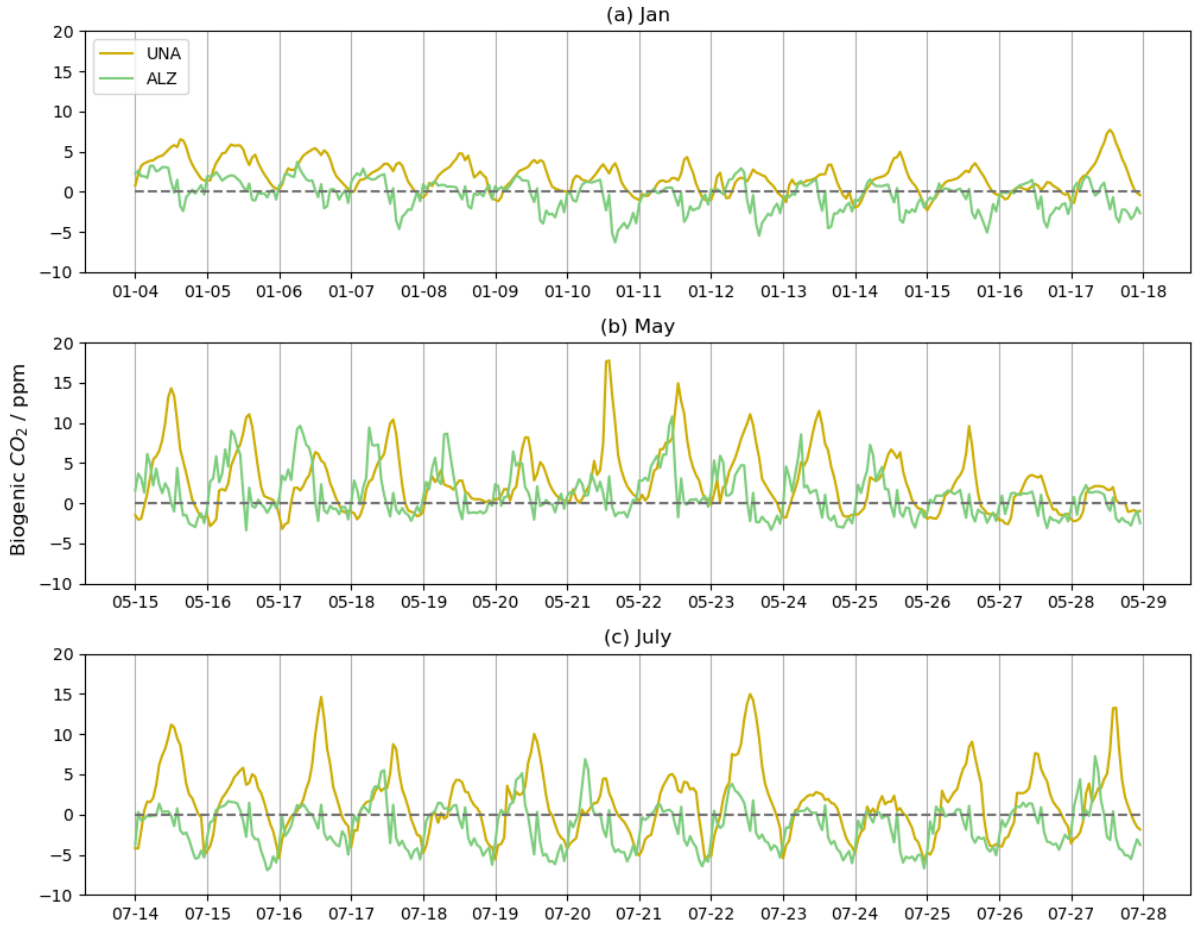


Figure S8. Modeled time series of biogenic contribution to surface CO_2 (in ppm) at UNA and ALZ for the months of (a) January, (b) May and (c) July of 2018, in UTC time (local time=UTC-6).

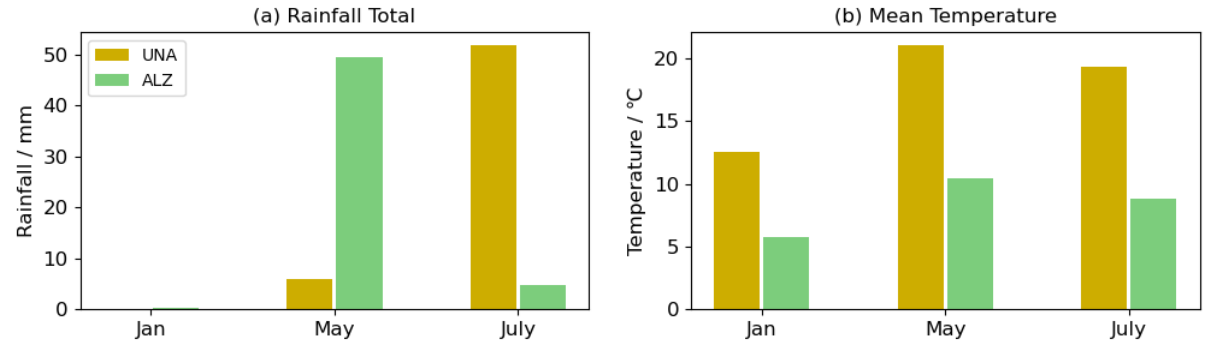


Figure S9. (a) 2-week rainfall and (b) 2-week mean temperature at UNA and ALZ in January, May and July of 2018.

APPENDIX B : ABSTRACT AT EGU 2021

Interpretation of Atmospheric CO_2 measurements in Mexico City

Yang XU1, Michel Ramonet1, Thomas Lauvaux1, Jinghui Lian1, Francois-Marie Bréon1, Philippe Ciais1, Michel Grutter de la Mora2, Agustin Garcia2

1Laboratoire des Sciences du Climat et de l'environment (LSCE-IPSL, CEA-CNRS-UVSQ), France

2Centro de Cienciasde la Atmósfera, Universidad Nacional Autónoma de México (CCA, UNAM), Mexico

The French-Mexican project Mexico City's Regional Carbon Impacts (MERCI-CO_2) is building a CO_2 observation network in the Metropolitan Zone of the Valley of Mexico (ZMVM). The project investigates the atmospheric signals generated by the city's emissions on total column and surface measurements, aiming at reducing the uncertainties of CO_2 emissions in ZMVM and evaluating the effects of policies that had been implemented by the city authorities.

A nested high-resolution atmospheric transport simulation based on the Weather Research and Forecasting model coupled with Chemistry (WRF-Chem) is performed to analyze the observed CO_2 mixing ratios during dry and wet seasons over Mexico City and its vicinity. Both anthropogenic emissions (UNAM 1-km fossil fuel emissions) and biogenic fluxes (CASA 5-km simulations) are taken into account. The model configuration, with a horizontal resolution of 1km and using the Single-Layer urban canopy Model (SLUCM), has been evaluated over two weeks in January 2018 using meteorological measurements from 26 stations set by the Air Quality Agency of Mexico City (Secretary of the Environment of Mexico City - SEDEMA). The reconstruction of meteorological conditions in the urban area shows better performances than sub-urban and mountainous areas. Due to the complex topography, wind speeds in mountain areas are 2-3 m/s over estimated and wind direction simulations in some stations are 90° deflected, especially in southern mountains.

Two high-precision CO_2 analyzers deployed in urban and rural areas of Mexico City are used to evaluate the WRF CO_2 1-km simulations. The model reproduced the diurnal cycle of CO_2 mixing ratios at the background station but under-estimates the nighttime accumulation at the urban station. Mean absolute errors of CO_2 concentrations range from 6.5 ppm (background

station) to 27.1 ppm (urban station), mostly driven by the elevated nocturnal enhancements (up to 500ppm at UNAM station). Based on this analysis, we demonstrate the challenges and potential of mesoscale modeling over complex topography, and the potential use of mid-cost sensors to constrain the urban GHG emissions of Mexico City.

APPENDIX C : ABSTRACT AT AGU 2022

Estimate of fossil fuel CO_2 Emissions from the Mexico City Metropolitan Area Based on 1 year of Atmospheric Measurements and Inversion Modeling

Y XU1, T. Lauvaux2, M.Ramonet1, JA García Reynoso3, M. Grutter3, J. Lian1, 4

1 Laboratoire des Sciences du Climat et de l'Environnement (LSCE), IPSL, UMR CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France

2 Groupe de Spectrométrie Moléculaire et Atmosphérique (GSMA), Université de Reims-Champagne Ardenne, UMR CNRS 7331, Reims, France

3 Instituto de Ciencias de la Atmósfera y Cambio Climático, Universidad Nacional Autónoma de México, Mexico City, Mexico

4 Origins S.A.S., Suez Group, Tour CB21, Paris, France

Megacities contribute more than 70% of global fossil fuel CO_2 emissions. During the past decades, cities and local governments have implemented local climate policies to mitigate their CO_2 emissions. But these policies rely on bottom-up emissions inventories from energy consumption data and statistical records, prone to large uncertainties. Top-down estimation based on Bayesian inversion approaches offers a complementary solution to reduce these uncertainties. Mexico City, located in the Valley of Mexico, is one of the largest cities in the world (22 million inhabitants). Thanks to a French-Mexican collaboration, (Mexico City's Regional Carbon Impacts - MERCI-CO_2), a large network of CO_2 in-situ (Cavity Ring-Down Spectrometer, CRDS - Picarro) and ground-based column (solar-absorption Fourier transform infrared, FTIR - EM27/Sun) sensors have been deployed across the Mexico City metropolitan area. A series of atmospheric CO_2 simulations between April 2018 and March 2019 were performed with the Weather Research and Forecasting model coupled with Chemistry (WRF-Chem). These, together with the surface CO_2 and XCO_2 measurements, provide the opportunity to produce top-down estimation of CO_2 emissions in and around Mexico City. Through the assimilation of CO_2 concentration gradients between the urban (UNAM) and the rural station (ALZ), the inversion system produced CO_2 emissions estimates from different sources - background, anthropogenic (including residential, transportation and industrial) and biogenic. We investigate the discrepancies between the UNAM 1-km CO_2 emission inventory and our inverse estimates over a year, , and we determine

the optimal characteristics of a future network able to monitor Mexico's CO_2 emissions combining additional in situ and column sensors over the region.

APPENDIX D : RESUME SUBSTANTIEL EN FRANÇAIS

Les villes sont responsables de plus de 70 % des émissions mondiales de CO_2 et jouent donc un rôle important dans les actions visant à atténuer le changement climatique. Les municipalités et les gouvernements non-étatiques se sont engagés à réduire les émissions de CO_2 urbaines, et se rapprocher ainsi de la neutralité carbone d'ici 2050. Afin d'évaluer l'impact des Plans Climat, des inventaires d'émissions de gaz à effet de serre sont établis à l'échelle de la ville . Mais ces inventaires d'émissions basés sur les données de consommation et de production d'énergie sont sujets à de grandes incertitudes. La modélisation par inversion atmosphérique offre une solution complémentaire capable de réduire les incertitudes. Elle combine la modélisation du transport atmosphérique et les mesures de concentrations en gaz à effet de serre pour affiner les estimations des émissions issues des inventaires.

La zone métropolitaine de Mexico (MCMA) est l'une des plus grandes mégalopoles du monde. Les émissions annuelles de dioxyde de carbone (CO_2) dans la MCMA sont passées de 42,1 millions de tonnes à 66,0 millions de tonnes entre 2012 et 2018. Le gouvernement mexicain a prévu de réduire les émissions de CO_2 de 65,2 millions de tonnes au cours de la période 2021-2030. Afin d'évaluer quantitativement les stratégies de réduction des émissions de CO_2, un projet franco-mexicain intitulé "Impacts du carbone dans la région de Mexico" (MERCI-CO_2) a déployé un réseau d'instruments d'observation du CO_2 in situ et en colonne dans la région de Mexico. Les gradients de concentration de CO_2 sont assimilés dans un système d'inversion basé sur le modèle de transport atmosphérique WRF-Chem pour améliorer l'estimation préalable des émissions de CO_2 à l'intérieur et à l'extérieur de la MCMA.

Au cours de cette thèse, le modèle atmosphérique a été tout d'abord évalué sous diverses configurations, y compris en utilisant divers forçages aux bords, différents domaines de simulation, différents schémas physiques et dynamiques. Plusieurs variables météorologiques ont été prises en compte pour comparer les simulations et les observations, en utilisant des données collectées aux

stations météorologiques de la région mais également issues d'un instrument LiDar ainsi que de radiosondes de l'Organisation Mondiale de la Météorologie (OMM). Afin de quantifier les erreurs les plus impactantes pour la simulation des concentrations en CO_2 atmosphérique, le travail de thèse s'est focalisé sur les erreurs de simulation des températures de l'air, des vitesses et directions du vent et les hauteurs de mélange (couche limite atmosphérique) . Ces tests de sensibilité ont permis de définir la configuration optimale du système de modélisation utilisé dans les chapitres suivants.

Dans un second temps, des cartes de concentration de CO_2 au-dessus de la MCMA sur trois périodes caractéristiques de la région (janvier, mai et juillet) ont été reconstruites par le modèle WRF-Chem à la résolution de 5 km, en utilisant deux inventaires d'émissions: inventaire local préparé par les chercheurs de l'UNAM et un inventaired'émissions globales, ODIAC, ajusté par des facteurs d'échelle temporels. L'évaluation des simulations de CO_2 a été réalisée sur la base de mesures in-situ de CO_2 (CRDS) et de mesures de colonne (XCO_2) par FTIR. Parallèlement à cette évaluation, nous avons également analysé les distributions temporelle et spatiale des signaux de CO_2, notamment le cycle diurne, la variation hebdomadaire et la variation saisonnière, ainsi que la zone impactée par les flux anthropiques et la variation des flux biogéniques. Sur la base de cette analyse, nous avons évalué le potentiel de quantification du réseau actuel, y compris l'emplacement de la station de fond, tout en étudiant les emplacements potentiels de nouvelles stations.

Pour conclure ce travail de thèse, une inversion sur une année complète a été réalisée sur la MCMA du 30 mars 2018 au 30 mars 2019. En se basant sur l'assimilation des gradients entre la station urbaine (campus de l'UNAM) à la station rurale (Altzomoni), l'inversion a ajusté les émissions anthropiques issues des deux inventaires ainsi que les flux biogéniques du modèle Carnegie Ames Stanford Approach (CASA). Les concentrations de fond, représentées par les concentrations en CO_2 issues du modèle CarbonTracker 2019B sont restées les mêmes. Un ensemble de plusieurs inversions a été réalisé sur l'année pour mieux quantifier les incertitudes des émissions en faisant varier les covariances d'erreur temporelles, en variant la fenêtre d'assimilation, en séparant les secteurs d'activités (trafic), en

appliquant un masque sur la MCMA, et en filtrant les données de concentrations en CO_2. Cet ensemble nous permet d'améliorer les performances de l'inversion et de spécifier l'impact du signal urbain et des différentes composantes du système. Un derniere expérience nous a permis d'étudier l'impact de l'assimilation des concentrations en monoxyde de carbone (CO), dont les observations débutent en décembre 2018.