Space-Based Evaluations of Urban Carbon Dioxide Emissions: Applications, Methods, and Policy Implications

by

Emily Grace Yang

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Climate and Space Sciences and Engineering)
in the University of Michigan
2022

Doctoral Committee:

Associate Professor Eric A. Kort, Chair
Assistant Professor Sara Hughes
Associate Professor Gretchen Keppel-Aleks
Professor Allison L. Steiner

“Efforts will lie, but will never be in vain.” – Yuzuru Hanyu

Emily Grace Yang

egyang@umich.edu

ORCID iD: 0000-0001-5195-9279

Dedication

This dissertation is dedicated to my loved ones, who have supported me through thick and thin.

Acknowledgements

They say it takes a village to raise a child, and I am of the opinion that it also takes a village to help a Ph.D. student graduate. I am so very grateful to my village that has been there to support me academically, physically, and emotionally—I couldn't have gotten this far without them.

Many thanks to my advisor, Eric Kort, for guiding me through this Ph.D. journey. I am thankful that Eric has always been there to help me become a better scientist, while also encouraging me in my more policy-oriented pursuits. I am also very grateful to have gotten to learn from and be supported by a whole host of brilliant faculty members in the CLaSP department and beyond at UMich. Thank you especially to my committee members Allison Steiner, Gretchen Keppel-Aleks, and Sara Hughes for their insights throughout this process.

I am lucky to have been able to collaborate with amazing scientists like Tom Oda, John Lin, and Dien Wu in the OCO-2/3 science team. It has also been wonderful to get to work with other members of the Kort group. A special shout-out goes to Martín, Mackenzie, and John for their support in the early days of my Ph.D. journey; my time at UMich might have turned out very differently without their help and encouragement. Many thanks to Genny as well, for always being very helpful and kind when I have been in need of some additional assistance (which has been fairly frequent!).

I am honored to have been able to interact with many members of this University community who helped to make this journey possible. Thank you to MUSE, Climate Blue, GEO, and GradSWE for working hard to make the campus community and beyond a better place; I feel

very fortunate to have been able to contribute to these groups' efforts and learn from the talented students in these groups. I am also grateful to the rest of the students of CLaSP for being a great community to be a part of, and am thankful for all of GUStO's efforts to make the department a more welcoming and inclusive space for its students. STPP has also been an integral part of my grad school experience, and I am thankful for them helping me learn and grow in my policy pursuits. And thank you to RMF for being such a lovely community and resource during my time here. Additionally, there are many folks throughout the University that I have relied on for issues big and small, and so I want to acknowledge the work of the staff, including IT, administrators, facilities, and grounds, as well as medical personnel and counselors.

I wouldn't imagine being able to complete this process without the support of my loved ones. Physics House has been such a very important part of my time in Ann Arbor, and I am so grateful to have been able to take part in such a wonderful community. Zach, Joe, Glenn, Peter, and Charlotte have really made Ann Arbor feel like home. Being coopted into our physics friend group has been an absolute blast, and I'm glad to have been able to play board games, watch movies, and overall geek out with all of you. I'm also so very thankful for my other friends, family, and mentors who have supported me from near or far. And thank you to my furry friends KB and Spooky for the unconditional love and snuggles.

I would also like to thank those who have helped keep me afloat in the past year or so when life decided to throw some curveballs at me. I wouldn't have been able to make it to this point without the support of my therapist, my physical therapists, and my massage therapists who have all been there to care for me.

Thank you all so much for being a part of my village!

Table of Contents

List of Tables

List of Figures

Abstract

Humans are perturbing the carbon cycle, primarily through the emission of carbon dioxide (CO_2) from the combustion of fossil fuels, resulting in a warming climate and uncertainties in how the carbon cycle will respond to these changes. Cities play an outsized role in contributing to climate change, as they are the source of the majority of fossil CO_2 emissions. Cities are also emerging as leaders in the fight against climate change, through the implementation of urban climate action programs, with ambitious emissions reductions targets, policy plans to achieve these targets, and self-calculated estimates of emissions tracked in self-reported inventories (SRIs). As space-based technology for observing CO_2 and co-located proxy species such as nitrogen dioxide (NO_2) improve, satellites are opening up many pathways for the estimation of urban CO_2 emissions. This dissertation interrogates the utility of space-based assessments of urban CO_2 emissions and their implications, using observations of CO_2 from the Orbiting Carbon Observatory-2 (OCO-2) and the Orbiting Carbon Observatory-3 (OCO-3) in its urban-focused Snapshot Area Map (SAM) mode, as well as observations of NO_2 from the TROPOspheric Monitoring Instrument (TROPOMI).

Applications, methods, and policy implications of assessments of urban CO_2 emissions with the use of satellite observations are explored in this work. First, global emissions inventory representations of five Middle Eastern cities' fossil CO_2 emissions are evaluated using comparisons between OCO-2 observations and simulations using the global emissions inventories in combination with the column version of the Stochastic Time-Inverted Lagrangian Transport (X-STILT) model. These comparisons provide optimum inventory scaling factors that

suggest that the inventory representations are underestimating the Middle Eastern cities' emissions. These results are found to be insensitive to the spatial distribution of the inventories. Next, empirical relationships between SAM CO_2 and TROPOMI NO_2 observations for three cities around the world are derived and applied to observed NO_2 fields to generate NO_2-derived CO_2 fields (NDCFs). Leveraging the greater availability and quality of the TROPOMI NO_2 observations, the NDCF method demonstrates relatively small methodological uncertainties when taken in aggregate, and shows a capacity to estimate emissions at a subannual timescale. The last study in this dissertation evaluates the ability of satellite observation-based estimates of CO_2 to assess urban climate action programs in the U.S. Application of results from previous studies using OCO-2 observations as well as extrapolating these results to the use of OCO-3 SAMs suggest that these satellite-based estimates have some policy relevance. They are shown to be able to evaluate the accuracy of SRIs and assess progress toward long-term emissions reduction targets, but more improvements in methods need to be made to track emissions year-to-year and to estimate sectoral emissions. A lack of compatibility between the design of the climate action plans and what observations can feasibly be used to estimate is found, which limits the ability of satellite-based emissions estimates to achieve policy-relevant urban CO_2 assessments.

Chapter 1 Introduction

1.1 The Changing Carbon Cycle

Analogous to the water cycle, the global carbon cycle refers to the process by which carbon is cycled between four main reservoirs: the atmosphere, the oceans (including the water and the marine biosphere), terrestrial land (including soil and the terrestrial biosphere), and the lithosphere (Ussiri & Lal, 2017). The biogeochemical cycling of carbon is quite complex, as it involves all living things on this planet, inorganic carbon reservoirs, and the links between them (Holmén, 1992). Carbon is constantly being exchanged within and between different carbon reservoirs, transforming between different molecular forms as it cycles, through a wide variety of physical, chemical, and biological processes (Ussiri & Lal, 2017). A number of these processes are shown in Figure 1.1, which details the faster components of the carbon cycle ("The Carbon Cycle," 2011).

Carbon primarily exists in the atmosphere in the form of carbon dioxide (CO_2) and methane (CH_4) (Ussiri & Lal, 2017). In this dissertation, we primarily focus on CO_2. There are various natural CO_2 sinks that take it out of the atmosphere, including the biosphere through photosynthesis, chemical weathering by silicate rocks over millions of years, and dissolution into the ocean resulting in a lowering of the pH (Caldeira & Wickett, 2003; Ussiri & Lal, 2017). However, humans are perturbing the carbon cycle by emitting CO_2 into the atmosphere at a rate of approximately 10 gigatons of carbon per year (Archer, 2010). The majority of this carbon entering the atmosphere is being transferred from the lithosphere due to combustion of fossil fuels and cement manufacturing, with additional carbon transferring from the biosphere due to

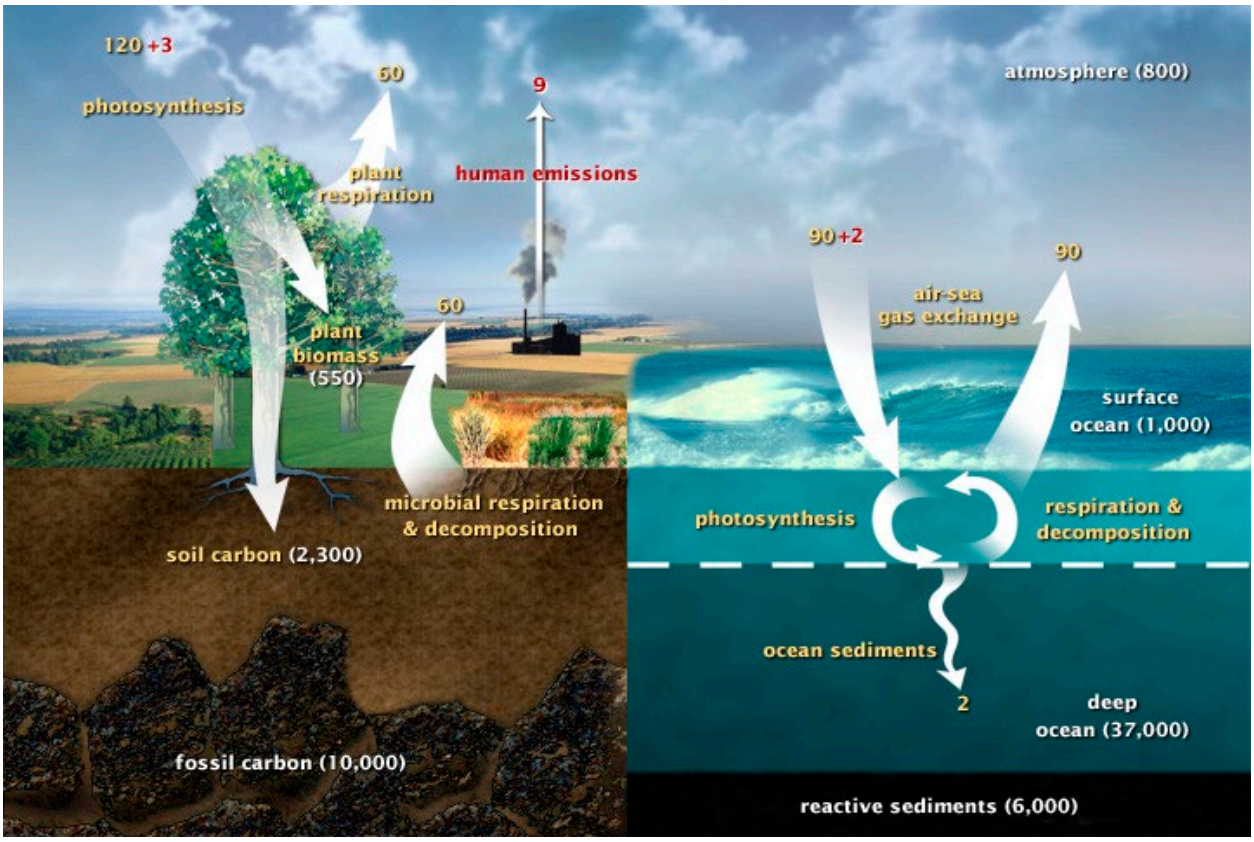

Figure 1.1 The fast carbon cycle, showing the movement of carbon between land, atmosphere, and oceans. Yellow numbers represent natural fluxes, while red are human contributions, all in gigatons of carbon per year. White numbers show stored carbon. Reproduced from "The Carbon Cycle" (2011).

destruction of land biomass (Archer, 2010). This anthropogenic perturbation of carbon into the atmosphere is the main contributor to human-induced climate change, a phenomenon involving the warming of the atmosphere, ocean, and land, all over the world (Masson-Delmotte et al., 2021).

1.1.1 Climate Change and Anthropogenic Emissions

1.1.1.1 Fossil Fuel CO_2 Emissions, Climate Change, and the Carbon Cycle

When discussing climate change, CO_2 is the most important greenhouse gas (GHG) in terms of its contribution to warming. Prior to the industrial revolution, the CO_2 mixing ratio was 280 ppm (IPCC, 2014); since then, this concentration has broken 420 ppm ("The Keeling

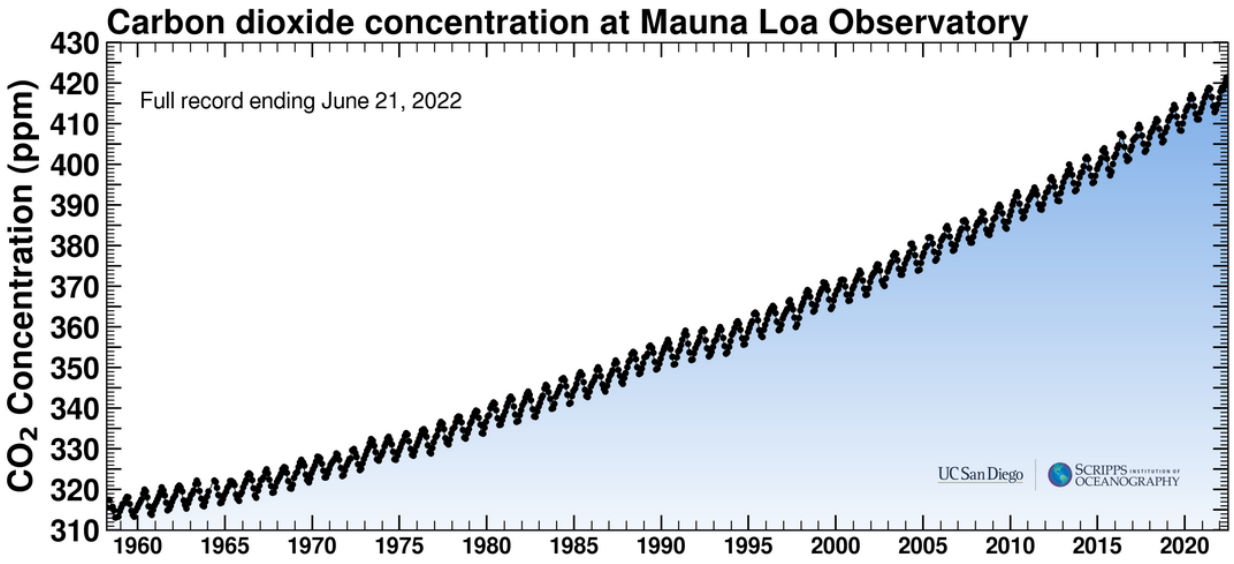

Figure 1.2 The CO_2 concentration in the air as measured daily since 1958 at Mauna Loa, Hawaii. Reproduced from "The Keeling Curve" (2022).

Curve," 2022). Figure 1.2 shows concentrations of global CO_2, as measured in Mauna Loa, Hawaii, taken daily from 1958 to the time of this writing. As we can see in the figure, not only is the mixing ratio of CO_2 rising, but the rate of growth is also increasing, with the latest concentration measurement amounting to 420.42 ppm as of June 21, 2022. This increase in CO_2 concentrations in the atmosphere can be linked to global temperature increases. In Figure 1.3, global surface temperatures are shown to be increasing at the present, an abrupt change from a period of relatively stable temperatures. This is the result of increased anthropogenic emissions of GHGs like CO_2 into the atmosphere, which trap radiation from the surface of the Earth and radiate it back to the surface and cause warming (Jacob, 1999).

As we can see in Figure 1.4, the majority of anthropogenic CO_2 can be attributed to the burning of fossil fuels, both cumulatively and in the period of 2010-2019. Specifically for this time period, 86% of the total emissions (E_{FOS}+E_{LUC}) were derived from fossil CO_2 emissions (E_{FOS}), and 14% were from land-use change (E_{LUC}) (Friedlingstein et al., 2020). Total emissions were subdivided between the atmosphere (46%), ocean (23%), and land (31%), with an

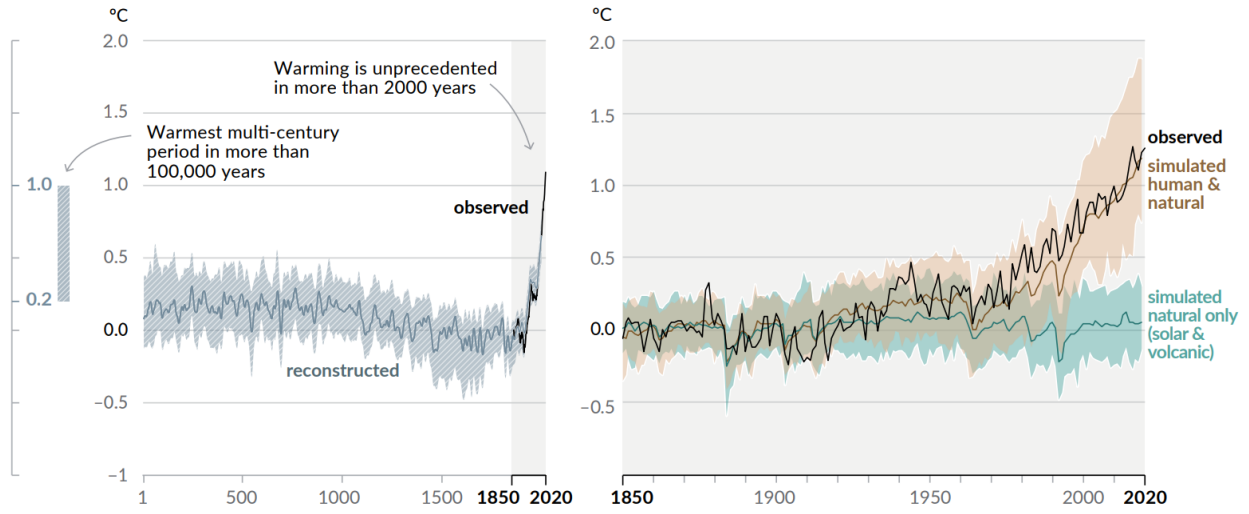

Figure 1.3 Left: change in global surface temperature (decadal average), reconstructed years 1-2000 and observed years 1850-2020. Right: change in global surface temperature (annual average), as observed (black), simulated using human & natural factors (brown), and simulated using only natural factors, during the years 1850-2020. Reproduced from Masson-Delmotte et al. (2021).

unattributed budget imbalance (-1%) (Friedlingstein et al., 2020). Note that, as shown in the figure, there are uncertainties attributed to each source and sink.

Given that the majority of anthropogenic CO_2 emissions are attributed to the combustion of fossil fuels, it is important for us to understand these emissions, for the sake of climate mitigation and adaptation, and for the sake of predicting how carbon sinks will behave in the future. Tracking fossil CO_2 emissions allows us to set accurate baseline values of emissions against which to evaluate climate policies, and to monitor the efficacy of climate mitigation policies (Ciais et al., 2014). However, understanding these emissions is also important in a more indirect way. When assessing global carbon budgets, the carbon cycle scientific community has typically assumed fossil CO_2 emissions to be a known quantity, studying other components of the carbon cycle like the land and ocean sinks by subtracting off fossil emissions (Hutyra et al., 2014). With such assumptions, it is important for the fossil emissions to be as well constrained as possible. Understanding the different components of the carbon cycle is necessary to make accurate future projections about the carbon cycle and our climate. For example, changes in the

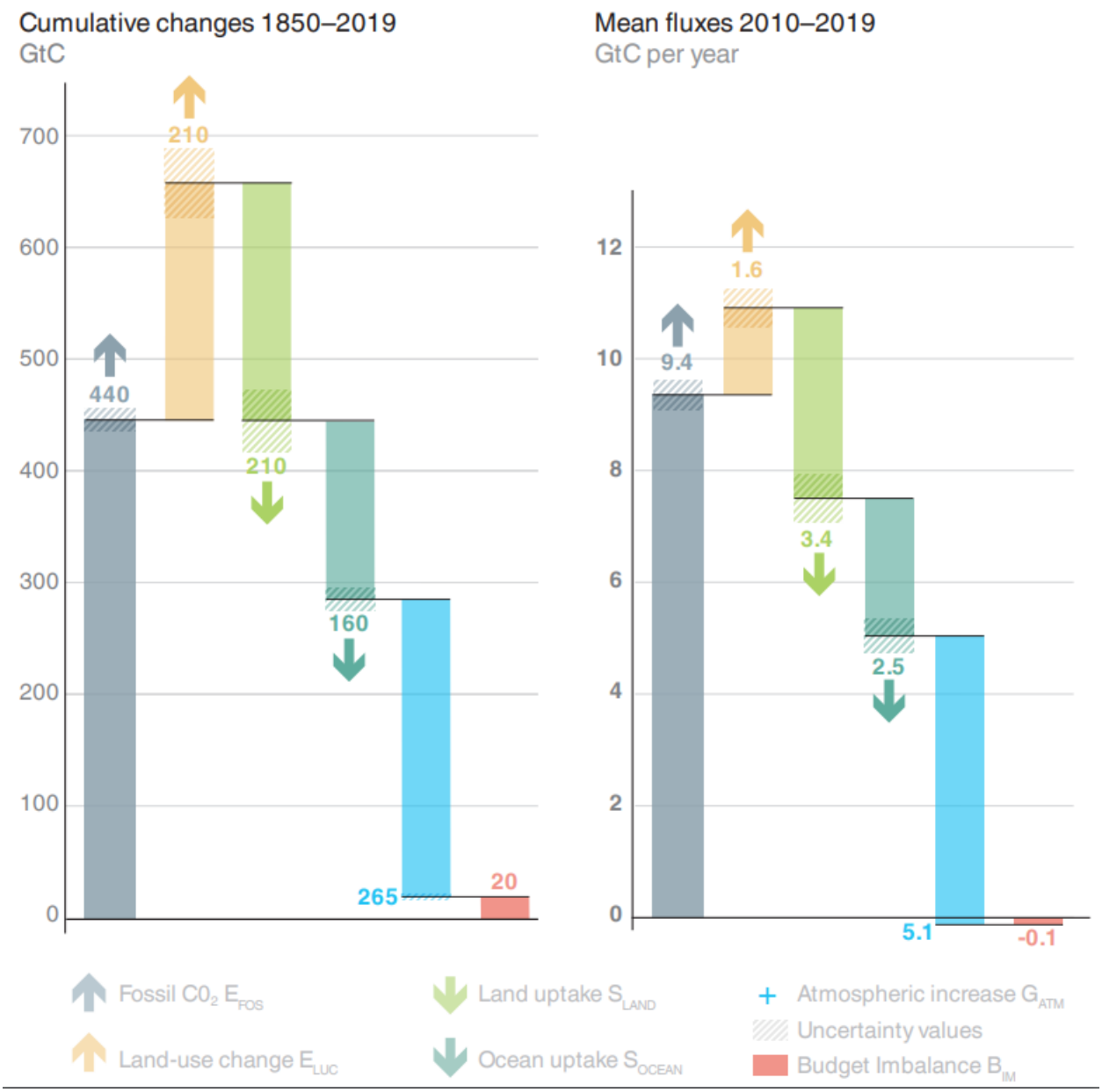

Figure 1.4 Anthropogenic carbon flows, showing cumulative changes during the period of 1850-2019 and mean fluxes during the period of 2010-2019. The sink resulting from cement carbonation is included in E_{FOS}. Reproduced from Friedlingstein et al. (2020).

role of the ocean as a global carbon sink could have enormous consequences for GHG management, making the understanding of ocean fluxes and how they respond to the climate vital (Ciais et al., 2014). The future role of the land biosphere is also an unknown, as we do not understand the mechanisms controlling land uptake well enough to project how it will behave in the future (Archer, 2010). By the end of this century, it is quite possible that warming may turn

the ocean and soils into carbon sources instead, amplifying rather than mitigating the effects of climate change (Archer, 2010); we need to better understand the processes controlling them to predict how they will respond to warming. Better understanding these mechanisms controlling the carbon cycle would also allow us to predict the impacts of GHG management strategies (Ciais et al., 2014). Accordingly, it is incredibly important to constrain fossil CO_2 emissions not just to track them and their direct effect on the climate, but also to allow us to understand the role of the rest of the components of the carbon cycle in the future.

1.1.1.2 NOx and their Link to Fossil CO_2

While not GHGs themselves, the gases collectively referred to as NO_x play an important role in the quantification of fossil CO_2 emissions. Nitrogen monoxide (NO) is a gas that is formed during combustion processes, including fossil fuel combustion, as a result of high temperatures combining molecular nitrogen with molecular oxygen at very high temperatures (Griffin et al., 2019). NO then rapidly oxidizes into nitrogen dioxide (NO_2). These two gases rapidly cycle between each other, and are often considered together as a family known as NO_x. NO_x has a relatively short lifetime within a plume (a few hours), making tropospheric NO_x concentrations correlate strongly with local emissions (Goldberg, Lu, Streets, et al., 2019; Griffin et al., 2019).

Because we can readily observe NO_2 with existing instruments, observing this gas could help us in better understanding fossil CO_2 emissions. Both NO_2 and CO_2 share many sources that are co-located in time and space, suggesting that observing NO_2 could help to constrain fossil emissions and monitor temporal changes in CO_2 emissions (Berezin et al., 2013; Konovalov et al., 2016). Additionally, using NO_2 as a proxy for fossil CO_2 could help to disentangle anthropogenic and biogenic sources of CO_2, allowing for reliable attribution of fossil CO_2

emissions (Ciais et al., 2014). We leverage the relationship between NO_2 and fossil CO_2 in Chapter 3 of this dissertation.

1.2 Cities and their Role in Climate Change

Cities play an incredibly important role in the story of climate change, and they are the focus of the three studies in this dissertation. Despite covering less than 5% of the surface of the globe, cities contribute a disproportionate amount to climate change, with the majority of fossil fuel CO_2 emissions attributed to them (van der Heijden, 2019; International Energy Agency, 2008). At present, more than half of the global population lives in cities, and this is projected to increase to 68% by the middle of the century (United Nations, Department of Economic and Social Affairs, Population Division, 2019). Following the increase in urban population, the global-scale uncertainty of fossil CO_2 has also been increasing with time, largely driven by the overall increase in emissions by developing countries that are quickly urbanizing (Hutyra et al., 2014). Accordingly, constraining urban emissions is vital to constraining fossil CO_2 emissions.

While cities are an epicenter for many of the activities and behavior that contribute to climate change, they are also where the consequences of climate change will be experienced most severely (van der Heijden, 2019). For example, climate change is predicted to increase the frequency of urban heat island events, in which air temperatures in cities rise more than in surrounding areas and cause locally negative impacts to human health, the economy, and the environment (Corburn, 2009). Additionally, the concentration of population and economic impact in cities suggests that a climate change event would have a major impact if it hit a city (van der Heijden, 2019).

Cities are also responding to climate change in the face of stalling national and international climate policy, emerging as leaders in both mitigation and adaptation to climate

change (Hughes, 2019; Rosenzweig et al., 2010). While cities have a disproportionate contribution to climate change, they also play a significant role in global environmental governance. Cities are uniquely positioned to play this role because of their location at the interface of local action and both national and international level climate change adaptation and mitigation commitments, along with their local understanding of the tradeoffs and synergies between adaptation and mitigation (Reckien et al., 2018). They have been forming transnational networks to tackle climate change at the international level, such as the C40 Cities Climate Leadership Group, the Covenant of Mayors, and the ICLEI – Local Governments for Sustainability's Cities for Climate Protection (CCP) program (Lee, 2013). Thousands of cities have been actively engaging in global climate change issues within their own jurisdictions and setting ambitious GHG emission reduction targets, implementing climate action plans to achieve those targets (Hughes, 2019; Lee, 2013). These climate action plans are documents prepared at the city level that contain policies relevant to climate change mitigation and/or adaptation (Reckien et al., 2018). Chapter 4 of this dissertation specifically delves into urban climate action plans in the U.S. and their associated programs, which also include the self-estimation of urban GHG emissions and reporting these estimates in what are referred to as self-reported inventories (SRIs).

1.3 Monitoring the Earth's Changing Systems with Satellite Remote Sensing

Humans are changing the world that we live in, and development over the last several decades in satellite-based technology has allowed us to monitor these changes via space-based observations. These satellites rely on remote sensing technology, which involves detecting and measuring radiation at different wavelengths either reflected or emitted from distant objects or materials, allowing for the identification, classification, and quantification of the objects or

materials that have interacted with that radiation (Kansakar & Hossain, 2016). Earth Observation (EO) satellites have remote sensing instrumentation that allow for measurements and mapping of Earth's physical, chemical, and biological systems. In the span of 1960 to 2030, there will have been hundreds of EO satellite launched into space to monitor our planet and its systems (Guo et al., 2015). Such satellites are monitoring components of the Earth to track environmental conditions that are tied to human health and safety, providing data that can be used in science-based decision-making on critical topics like disaster management, the global environment, and natural resource management (Kansakar & Hossain, 2016; Seltenrich, 2014). By delivering data on how the Earth's systems respond to both natural and anthropogenic changes, EO satellite missions can directly benefit society by giving us the tools we need to make the Earth a more sustainable place to live (Kansakar & Hossain, 2016).

In particular, EO satellites are allowing us to study the carbon cycle and climate change. The United Nations Framework Convention on Climate Change has listed 34 Essential Climate Variables that require contributions from EO satellites; these variables lie within the atmospheric, oceanic, and terrestrial domains, and include atmospheric composition, temperatures, wind speeds, acidity, and biodiversity—all aspects that can be observed from space (Guo et al., 2015). Climate change is a major global issue that transcends political boundaries, and remote sensing from space can help the globalized community achieve common goals by providing data resources that can be accessed and shared by all, including developing countries that may be more resource-limited (Kansakar & Hossain, 2016). As a result, satellite technology is opening up many doors for the study of climate change and the carbon cycle, and it will continue to do so as more EO satellites are launched into space.

The main chapters of this dissertation interrogate various aspects of climate change and the carbon cycle through the lens of satellite observations of CO_2 and NO_2, and the following subsections describe some of these satellites in more detail.

1.3.1 Satellites that Observe CO_2

Satellite missions to observe CO_2 from space have increased our opportunities to understand climate change and the carbon cycle. Observations of CO_2 concentrations from satellites are valuable for the estimation of CO_2 emissions because of the large density of observations provided by satellites, which would reduce some of the uncertainties associated with sparse sampling from other types of instruments (Ciais et al., 2014). Such space-based observations are playing an important role in monitoring the status and progress of compliance with international emission reduction agreements, such as the Montreal Protocol and the upcoming Global Stocktake as part of the Paris Agreement (Kuze et al., 2022).

Space-based remote sensing of CO_2 captures an entire emission plume vertically and horizontally from the top of the atmosphere (Kuze et al., 2022). Satellites retrieve XCO_2, which is the column-averaged dry air mole fraction of CO_2 (Crisp, 2015). In order to estimate XCO_2, satellites use high-resolution spectroscopic observations of reflected sunlight, measuring at three spectral bands around 0.765 μm (O_2 A band), 1.61 μm (weak CO_2 band), and 2.06 μm (strong CO_2 band) (Crisp, 2015). These measurements are then inputted into a retrieval algorithm to estimate XCO_2 along with other atmospheric and surface state properties such as surface albedos and aerosol properties (Crisp, 2015; Wu et al., 2019). The first satellite instruments to use this measurement approach are the European Space Agency (ESA) EnviSat SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY (SCIAMACHY) and the Japanese Greenhouse Gases Observing Satellite (GOSAT) Thermal And Near infrared Sensor for carbon

Observation Fourier Transform Spectrometer (TANSO-FTS), which have helped to advance carbon cycle science at scientific and policy-relevant scales (Crisp, 2015; Kuze et al., 2022). There have been a number of more recent CO_2-observing space-based instruments, including the Orbiting Carbon Observatory-2 (OCO-2), which is designed to measure atmospheric CO_2 with the accuracy, coverage, and resolution needed to detect sinks and sources of CO_2 all over the world (Crisp, 2015). Japan's second GHG observing satellite, GOSAT-2, was launched in 2018, also increasing the accuracy and spatial density of observations as compared to GOSAT (Suto et al., 2021). Finally, the installation of the Orbiting Carbon Observatory-3 (OCO-3) instrument on the International Space Station has opened up new potential pathways for urban CO_2 emissions studies to move forward, with observations during different times of the day (instead of only the afternoon as with OCO-2), and the new Snapshot Area Map (SAM) mode that can capture the entire urban plume (Eldering et al., 2019). Given that satellite-based observations of CO_2 suffer from limitations such as loss of data due to clouds and observations limited to certain periods of the day (Ciais et al., 2014), OCO-3 and upcoming missions such as Japan's Global Observing SATellite for Greenhous gases and Water cycle (GOSAT-GW) and Geostationary Carbon Cycle Observatory (GeoCarb) will provide even more data for the improved understanding of our changing carbon cycle (Kuze et al., 2022; Moore III et al., 2018).

1.3.2 Satellites that Observe NO_2

Satellites are able to measure NO_2 concentrations in the atmosphere because of the gas's unique high frequency spectral features within the 400-500 nm wavelength region (Goldberg, Lu, Streets, et al., 2019). Over the past two decades, our ability to observe near-surface concentrations of air pollutants like NO_2 from space has grown dramatically (Griffin et al., 2019). The first space-based spectrometer to observe NO_2 was the Global Ozone Monitoring

Experiment (GOME), which was first launched in 1995 and operated for 16 years (Coldewey-Egbers et al., 2018; Griffin et al., 2019). The SCIAMACHY instrument, which also measured other gases including CO_2, and the GOME-2 instrument, improved upon the capabilities of GOME (Griffin et al., 2019). Until a few years ago, the Ozone Monitoring Instrument (OMI) was the standard for NO_2 satellite observations, as it was able to resolve pollutant distributions at the urban scale and infer emissions from large point sources with an observational resolution of up to 13 km × 24 km (Griffin et al., 2019). Building on the legacy of the previous ultraviolet-visible space-based spectrometers, the latest remote sensing spectrometer to observe NO_2 is the TROPOspheric Monitoring Instrument (TROPOMI), aboard the European Space Agency's Sentinel-5 Precurser satellite, (Goldberg, Lu, Streets, et al., 2019; Griffin et al., 2019). Launched in October 2017, TROPOMI has unprecedented spatial resolution and data product quality (Griffin et al., 2019). As of 2019, TROPOMI has been able to provide observations with a resolution of up to 5.5 km × 3.5 km (S. Liu et al., 2021), allowing for greatly improved capabilities in observing NO_2 at the urban scale.

1.4 Overview of Dissertation

This dissertation explores different facets of the study of urban fossil CO_2 emissions using satellite remote sensing observations. Chapter 2 focuses on space-based assessment of urban emissions estimates represented by three global CO_2 emissions inventories, for five cities in the Middle East. Using observations from OCO-2 in combination with a Lagrangian particle dispersion model to account for atmospheric transport and link emissions to observations, we evaluate the three inventory representations of afternoon emissions. We optimize emissions inventory estimates with the OCO-2 observations, demonstrating that the inventory representations of afternoon emissions are too low. We also discuss the impact of the prior

inventory spatial distributions on our results, suggesting that we can make robust quantitative assessments of urban emission magnitudes without requiring high-resolution gridded inventories.

In Chapter 3, we explore the potential of using NO_2 satellite data in combination with available CO_2 satellite data to assess urban CO_2 emissions, with a focus on three different cities around the world. In this study, we calculate empirical relationships between satellite observations of CO_2 and NO_2, using these relationships to generate NO_2-derived CO_2 fields from which CO_2 emissions can be estimated. First establishing the method with simulations, we then use actual satellite observations from OCO-3 and TROPOMI with our method to estimate the cities' CO_2 emissions over time. We show monthly variations in urban CO_2 emissions estimates that are comparable to monthly emissions inventory estimates. We demonstrate how powerful the much larger quantity of NO_2 observations is, suggesting that our method can provide a valuable observational constraint on urban CO_2 emissions with the use of proxy data.

Chapter 4 discusses an evaluation of the ability of satellite CO_2 observations to assess the effectiveness of urban climate action programs. With an initial focus on the city of Los Angeles, we apply results from previous studies as well as derive approximate uncertainties associated with CO_2 emissions estimates based on observations from the latest satellite technology to evaluate various facets of the city's climate action program. We explore the utility of satellite-based emissions estimates for evaluating the accuracy of the city's reported emissions, assessing overarching emissions targets, tracking year-to-year emissions reductions, tracking specific policies in their urban climate action plan, and evaluating sectoral emissions. We then apply our findings to other cities' climate action programs around the U.S. We discuss short- and long-term options for improving the ability of satellite observations of CO_2 to provide policy-relevant assessments of emissions.

Finally, in Chapter 5, we draw general conclusions from this dissertation and discuss future research directions.

Chapter 2 Using Space-Based Observations and Lagrangian Modeling to Evaluate Urban Carbon Dioxide Emissions in the Middle East

Abstract

Improved observational understanding of urban CO_2 emissions, a large and dynamic global source of fossil CO_2, can provide essential insights for both carbon cycle science and mitigation decision making. Here we compare three distinct global CO_2 emissions inventory representations of urban CO_2 emissions for five Middle Eastern cities (Riyadh, Mecca, Tabuk, Jeddah, and Baghdad) and use independent satellite observations from the Orbiting Carbon Observatory-2 (OCO-2) satellite to evaluate the inventory representations of afternoon emissions. We use the column version of the Stochastic Time-Inverted Lagrangian Transport (X-STILT) model to account for atmospheric transport and link emissions to observations. We compare XCO_2 simulations with observations to determine optimum inventory scaling factors. Applying these factors, we find that the average summed emissions for all five cities are 100 MtC y^{-1} (50-151, 90% CI), which is 2.0 (1.0, 3.0) times the average prior inventory magnitudes. The total adjustment of the emissions of these cities comes out to ~7% (0%, 14%) of total Middle Eastern emissions (~700 MtC y^{-1}). We find our results to be insensitive to the prior spatial distributions in inventories of the cities' emissions, facilitating robust quantitative assessments of urban emission magnitudes without accurate high-resolution gridded inventories.

Plain Language Summary

Carbon dioxide (CO_2) emitted from burning fossil fuels is the most important contributor to climate change and is changing the Earth's carbon cycle. Most of these emissions can be linked to cities. Since cities around the world are quickly changing and growing, especially in developing countries, it is important to better understand the extent of urban CO_2 emissions to understand how the climate and carbon cycle will change. In this study, we evaluate three global emissions inventories—modeled estimates of CO_2 emissions on a grid spanning the globe—and their ability to capture the afternoon emissions of five Middle Eastern cities (Riyadh, Mecca, Tabuk, Jeddah, and Baghdad). This assessment relies on comparing simulations using the inventories with observations from the Orbiting Carbon Observatory-2 (OCO-2) satellite. Based on these comparisons, we see that the inventory representations have underestimated afternoon emissions of the five studied cities, and that the level of underestimation is a substantial portion of total Middle Eastern emissions. Our results are unaffected by the differing spatial patterns of emissions from different inventories. This work demonstrates the ability to use satellites to evaluate sub-national emissions, a valuable advance for both science and policy issues relating to climate change and the carbon cycle.

2.1 Introduction

Before the Industrial Revolution, the CO_2 mixing ratio was 280 ppm (IPCC, 2014); since then, it has surpassed 400 ppm and continues to rise (Betts et al., 2016), mainly due to emissions from burning fossil fuels (Stocker et al., 2013). To predict future climatic trends, it is important for us to know how each part of the carbon cycle is responding to this significant perturbation to CO_2 concentrations. However, rising uncertainties in fossil fuel CO_2 ($FFCO_2$) emissions due to overall increases in absolute emissions and increases in the proportion of emissions from

developing regions with fewer constraints are limiting our understanding of the extent and implications of this large and growing perturbation.

A first step to constraining such uncertainties is to focus on cities, as the majority of $FFCO_2$ emissions derive from urban areas (International Energy Agency, 2008). These in-city emissions are mostly linked to automobile tailpipe emissions, industrial emissions, and home heating; emissions also come from electricity generation from power plants that use fossil fuels, though such emissions may take place outside of the city itself (Lin et al., 2018). Presently, more than half of the global population resides in cities (United Nations, Department of Economic and Social Affairs, Population Division, 2015). Quantifying urban emissions is also becoming more important as cities pledge to take greenhouse gas mitigation actions (Rosenzweig et al., 2010; Gurney et al., 2015; Bodansky, 2016) and require a means of assessing those actions. As the developing world rapidly urbanizes, it is becoming critical to quantify and constrain the $FFCO_2$ emissions in cities, to inform such urban mitigation policy issues as well as to understand the carbon cycle science implications of the $FFCO_2$ perturbation.

Global $FFCO_2$ gridded emissions inventories are valuable tools for carbon cycle science and mitigation policies, but have room for improvement with respect to local- and urban-scale representations of emissions. These emissions inventories use proxies to disaggregate national-level emissions statistics onto a fine-scale grid (Oda & Maksyutov, 2011; Oda et al., 2019)—but at finer spatial scales, the proxy approach becomes insufficient to characterize the spatial distribution of emissions sources (Oda et al., 2019). For example, Gurney et al. (2019) found city-level differences between inventories for four U.S. cities of up to about 20%. Meanwhile, in their study on different inventories for the northeastern United States, Gately and Hutyra (2017) found that more than one-fourth of the grid cells in urban areas at 0.1 x 0.1° resolution had

relative differences of 100-300%. In that study, the researchers determined that existing global-scale $FFCO_2$ inventories are unsuitable for urban emissions monitoring, as they are unable to describe the underlying spatiotemporal patterns of the activities from which urban emissions are sourced (Gately & Hutyra, 2017). Lacking in the spatiotemporal resolution needed to capture the variability in local-level emissions and exhibiting large uncertainties, global $FFCO_2$ emissions inventories need improvement not just for local-scale issues, but also global-scale ones. For example, such large uncertainties are a problem for the accuracy of global carbon budgets that are determined by considering fossil fuel CO_2 emissions to be a known quantity (Hutyra et al., 2014). Additionally, errors in the spatiotemporal distribution or magnitude of $FFCO_2$ fluxes have been shown to propagate to remaining flux components of carbon inversion studies (Gurney et al., 2005; see also Hutyra et al., 2014). Uncertainties in $FFCO_2$ emissions in global inventory representations have also been shown to be a dominating factor in global CO_2 atmospheric inversion model spread and to limit the assessment of regional scale terrestrial fluxes and ocean-land partitioning (Gaubert et al., 2019). For the sake of fine-resolution emissions inventories to be useful at many scales, it is thus a critical task to evaluate and improve the inventory models, and to optimize model representations of smaller scale emissions.

Top-down, space-based observations of CO_2 concentrations are a strong contender for evaluating emissions inventories at a local scale, especially in regions with few or no on-ground observations. In recent years, space-based technological advancements have been evolving the study of urban- and local-scale emissions. Satellites have opened the doors to global, high-resolution measurements of XCO_2, the column-averaged dry-air mole fraction of CO_2. A number of satellites observing XCO_2 have already been launched, such as the Japanese Greenhouse Gases Observing SATellite (GOSAT) and its successor GOSAT-2, the American Orbiting

Carbon Observatory-2 (OCO-2) and its successor (OCO-3), and the Chinese TanSat. A number of studies have already been conducted that have utilized the observations of some of these satellites to better understand local sources of CO_2 such as large cities and power plants (Kort et al., 2012; Hakkarainen et al., 2016; Nassar et al., 2017; Wu et al., 2018; Hedelius et al., 2018). That body of work—and in particular that of Wu et al. (2018), which presents a method of extracting urban XCO_2 signals from OCO-2 and evaluates this method with a case study of Riyadh, Saudi Arabia—serves as the foundation for the work of this present study.

In this study, we evaluate global $FFCO_2$ emissions inventory representations of urban CO_2 emissions in the Middle East, and assess the ability of satellite observations to inform this evaluation. Middle Eastern cities serve as optimal target study domains. Limited cloud cover increases data density, and reduced biospheric signals simplifies the analysis. Additionally, these cities provide a good example of locations with few ground-based measurements making space-based approaches the only observation-based method that could be applied. To complete our assessment, we focus on three global $FFCO_2$ emissions inventories: the Fossil Fuel Data Assimilation System (FFDAS), the Open-source Data Inventory for Anthropogenic CO_2 (ODIAC), and the Emission Database for Global Atmospheric Research (EDGAR). We examine the differences in the spatial distribution and magnitudes of these inventories at the urban scale for cities in the Middle East, focusing on Riyadh, Saudi Arabia but also analyzing Saudi Arabian cities Mecca, Tabuk, and Jeddah, as well as the city of Baghdad, Iraq. We next quantify the relationship between observations of XCO_2 from the OCO-2 satellite and simulations of XCO_2 using the column version of the Stochastic Time-Inverted Lagrangian Transport (X-STILT) model coupled with Global Data Assimilation System (GDAS) reanalysis products and the three inventories. Using this comparison between top-down observations and bottom-up simulations of

XCO_2, we calculate scaling factors to adjust the flux magnitudes of the inventory estimates of city emissions. We then discuss the implications of our estimates on a regional scale, and on the ability of space-based observations to quantify urban emissions.

2.2 Data and Methods

2.2.1 Emissions Data and Methods

2.2.1.1 Global $FFCO_2$ Emissions Inventories

In this study, we evaluate three of the most widely used global $FFCO_2$ emissions inventories: FFDAS, ODIAC, and EDGAR. Figure 2.1 shows the three inventories' representations of the Middle East. Each of these inventories is formed using an approach that combines top-down elements—distribution of emissions in space by using spatial proxies like population or nighttime lights—with bottom-up elements—summing up individual fuel consumption or emissions sources to estimate total emissions at a larger scale (Hutchins et al., 2017). However, the inventories are all gridded using different mechanisms, with variations in energy statistics used, sectors included, and methods of downscaling emissions. These differences contribute to notable discrepancies between the inventories' aggregated emissions totals and spatial distributions. Table 2.1 summarizes the key information about the inventories.

FFDAS is a global product with a spatial resolution of 0.1 x 0.1°. This inventory downscales national emissions statistics reported by the International Energy Agency (IEA), distributing emissions by constraining the Kaya identity, which uses population, GDP, energy intensity, and carbon intensity as multiplicative factors to determine emissions (Rayner et al., 2010; Asefi-Najafabady et al., 2014; Gately & Hutyra, 2017; Hutchins et al., 2017). Constraints are placed on the Kaya identity using satellite nightlights, population density information, and

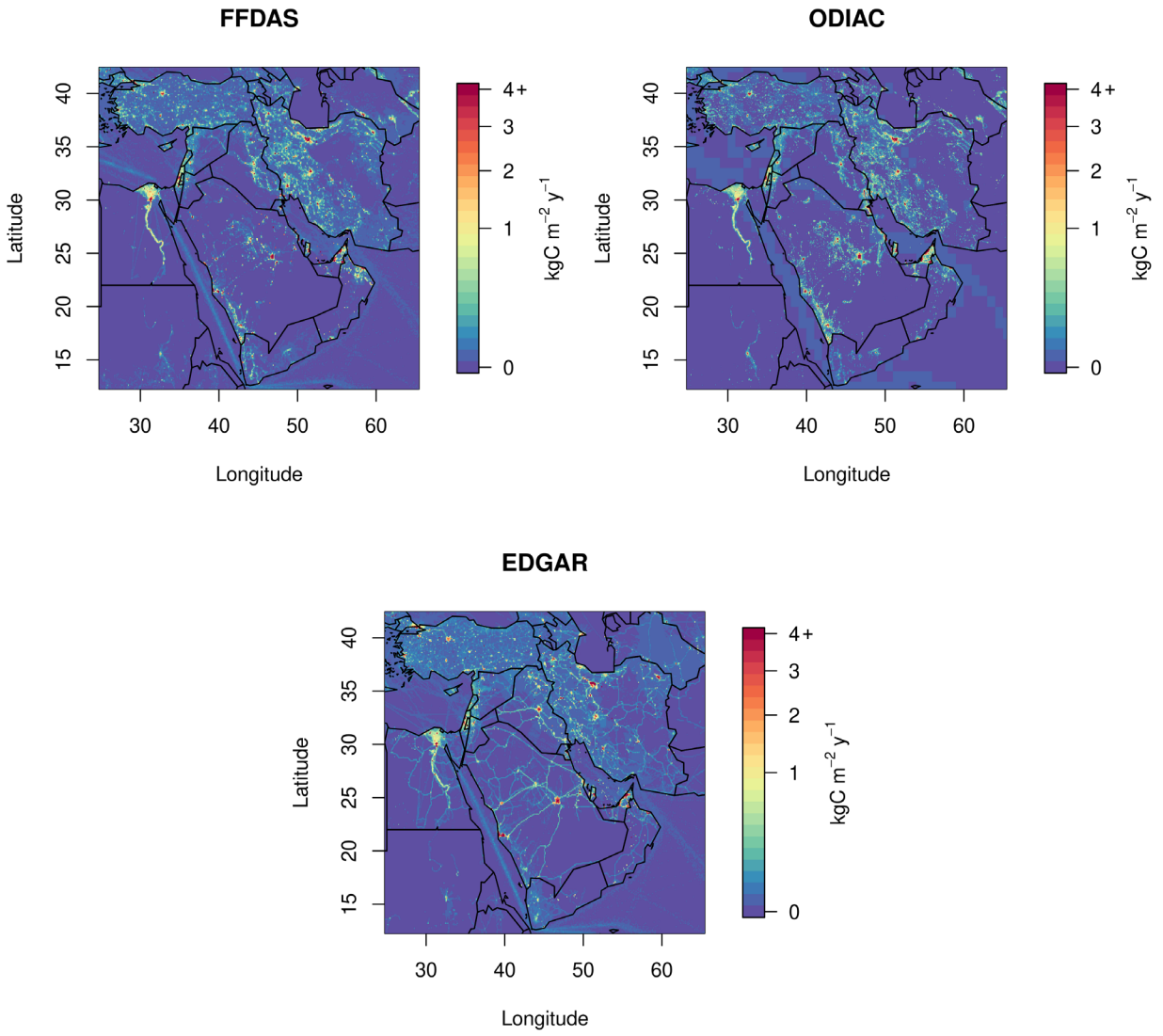

Figure 2.1 Global fossil fuel CO_2 emissions inventory representations of the Middle East: FFDAS, ODIAC, and EDGAR, shown with a square-root scale. The three representations differ in both spatial distribution and magnitude of emissions. Note that all inventories are shown at their native resolutions, with ODIAC having land emissions at a resolution of 1 x 1 km and international aviation and marine bunker emissions at 1 x 1°.

	FFDAS	*ODIAC*	*EDGAR*
Version	2014b (beta)	ODIAC2017	4.3.2
Year(s) used	2014	2014-2016	2012
Resolution	0.1 x 0.1° Hourly/annual	1 x 1 km Monthly	0.1 x 0.1° Annual
Global total	8.9 PgC y^{-1}	9.9 PgC y^{-1}	9.5 PgC y^{-1}
Middle Eastern total	697 MtC y^{-1}	789 MtC y^{-1}	722 MtC y^{-1}
Sectors or categories included	IEA sectors: energy generation, manufacturing, industrial, and transportation, others including residential, commercial, agriculture, and fishing	CDIAC fuel types (liquid, gas, solid, cement production, gas flaring, and international aviation and marine bunkers); re-categorized as point source, nonpoint source, cement production, gas flare, and international aviation and marine bunkers	IPCC sectors: energy, fugitive, industrial processes, solvents and products use, agriculture, waste, and other (emissions due to fossil fuel fires)

Table 2.1 Key information about the global gridded $FFCO_2$ emissions inventories used in this study.

power plant data (publicly disclosed or from the World Electric Power Plants (WEPP) database). Sectors that are included in FFDAS are based on sectors from which IEA produces emissions statistics, including energy generation, industrial, and transportation (land transport native to FFDAS, aviation and shipping added in from EDGAR), as well as other sectors such as residential, commercial, agriculture, and fishing (Asefi-Najafabady et al., 2014; Rayner et al., 2010; Hutchins et al., 2017). Emissions due to cement production and gas flaring are excluded from this product (Asefi-Najafabady et al., 2014). In this study, we use the 2014b (beta) dataset, an early version of the FFDAS v2.0 dataset. This dataset is at both hourly and yearly temporal resolutions and provides emissions data for the year 2014. For 2014, total global emissions in this set come out to 8.9 PgC.

ODIAC is a monthly global product that distributes $FFCO_2$ emissions with a spatial resolution of approximately 1 x 1 km (0.008333 x 0.008333°) (Oda et al., 2018). In the

ODIAC2017 version, the inventory begins with national emissions estimates separated by fuel type (liquid, gas, solid, cement production, gas flaring, and international aviation and marine bunkers) as opposed to emissions sectors, from the Carbon Dioxide Information Analysis Center (CDIAC) at the Oak Ridge National Laboratory (ORNL). These emissions data are then re-categorized into the following categories: point source, nonpoint source, cement production, gas flare, and international aviation and marine bunkers. For the more recent years that are not included in the CDIAC estimates (2015-2016), the 2014 CDIAC emissions estimates are extrapolated using more recent BP global fuel statistical data (Oda et al., 2018; BP, 2017). These emissions statistics are then spatially distributed with multiple spatial proxies such as the Carbon Monitoring for Action (CARMA) database for point sources, nighttime light data collected by Defense Meteorological Satellite Program (DMSP) satellites for nonpoint sources, and ship and aircraft fleet tracks. The emissions are also distributed temporally using seasonality derived from the CDIAC monthly gridded emissions, resulting in a monthly gridded product (Oda et al., 2018). We note that ODIAC makes available at a resolution of 1 x 1° international aviation and marine bunkers emissions information, which is recorded by CDIAC/ORNL but not included in CDIAC's gridded emissions data products. In this study, we use the ODIAC2017 version for the years 2014-2016 (Oda & Maksyutov, 2015), summing the 1 x 1 km land emissions dataset with the 1 x 1° international aviation and marine bunkers data. For the year 2014, total global emissions in this dataset are 9.9 PgC.

EDGAR is an annual global emissions inventory with a resolution of 0.1 x 0.1°, and it relies on international activity data provided by sources like the IEA and emissions factors to determine country-specific CO_2 emissions (Janssens-Maenhout et al., 2019; Crippa et al., 2018; Olivier & Janssens-Maenhout, 2015; European Commission Joint Research Centre, 2017).

National sector totals for emissions are then distributed using a number of different proxies, including location of energy and manufacturing facilities, road networks, shipping routes, population density, and agricultural land use (European Commission Joint Research Centre, 2017; Janssens-Maenhout et al., 2013). EDGAR includes emissions sectors as defined by the Intergovernmental Panel on Climate Change (IPCC): energy (including international aviation and shipping emissions), fugitive, industrial processes (including cement production), solvents and products use, agriculture, waste, and other (emissions due to fossil fuel fires) (Janssens-Maenhout et al., 2019). Gridded $FFCO_2$ emissions for version 4.3.2 are available for years 1970-2012 (Janssens-Maenhout et al., 2019); we use only the most recent year (2012) in this study. For 2012, the total global emissions magnitude for this dataset is 9.5 PgC.

2.2.1.2 Inventory Resolutions, Years, and Domains

Direct comparisons between the three inventories necessitate the inventories to share common spatiotemporal dimensions. For such direct comparisons, we sum the 2014 ODIAC monthly emissions to form a single annual gridded product, and aggregate it from its native spatial resolution to 0.1 x 0.1°. We also use the annual 2014 FFDAS dataset and the 2012 EDGAR dataset. For comparisons with observations, we use ODIAC both in its native resolution and in the aggregated resolution. While the years do not match perfectly, we choose 2014 for ODIAC and FFDAS due to the availability of both inventory and OCO-2 data. We choose to match FFDAS and ODIAC, and use the most recent year (2012) for EDGAR, under the assumption that EDGAR's emissions representations would not change in distribution much between the two years, with mostly changes in magnitude. Regarding the change in magnitude from 2012 to 2014, there is some uncertainty; while Friedlingstein et al. (2014) find that global CO_2 emissions grow 2.5% per year on average, the developing world may experience a different

rate of emissions growth. However, the growth rate in population for the cities of interest are approximately 2-4% per year (for 2004 to 2010 for Saudi Arabian cities and 2009 to 2018 for Baghdad) (Brinkhoff, 2018), suggesting small emissions growth as well. These changes are expected to have minimal effects on the results of this study.

We choose separate domains of interest for each studied city: square domains that range from 0.4 x 0.4° in size to 0.7 x 0.7° to encompass the cities of interest (Figure 2.2). For each city, a first-pass domain is chosen by examining Google Earth to approximate the spatial boundaries of the cities. After an approximate square domain is chosen, a sensitivity analysis is done using the emissions inventories, either by reducing or expanding the domain by increments of 0.1° on each side of the domain to determine at which size the total emissions within the chosen domain changes the least.

Note that the chosen domains do not correspond exactly to the administrative boundaries of the cities themselves. Emissions sources such as nearby suburbs that would not necessarily be part of the cities themselves could be included. However, given that it would be quite difficult to distinguish OCO-2 XCO_2 enhancements as having come from either a city or its neighboring suburb, we choose to set the boundaries of the domains as squares that hold the greatest emissions relative to the rest of the neighboring area.

2.2.1.3 Analytical Methods for Direct Comparisons Between Inventories

To evaluate the differing urban representations of the emissions as modeled by the inventories, we conduct direct comparisons between them using methods to discriminate between the magnitudes and the spatial distributions of the inventories. Some of these methods were also conducted by Hutchins et al. (2017) in their study focusing on the continental U.S. and Gately and Hutyra (2017) in their study of the northeast U.S. We sum up the emissions over the

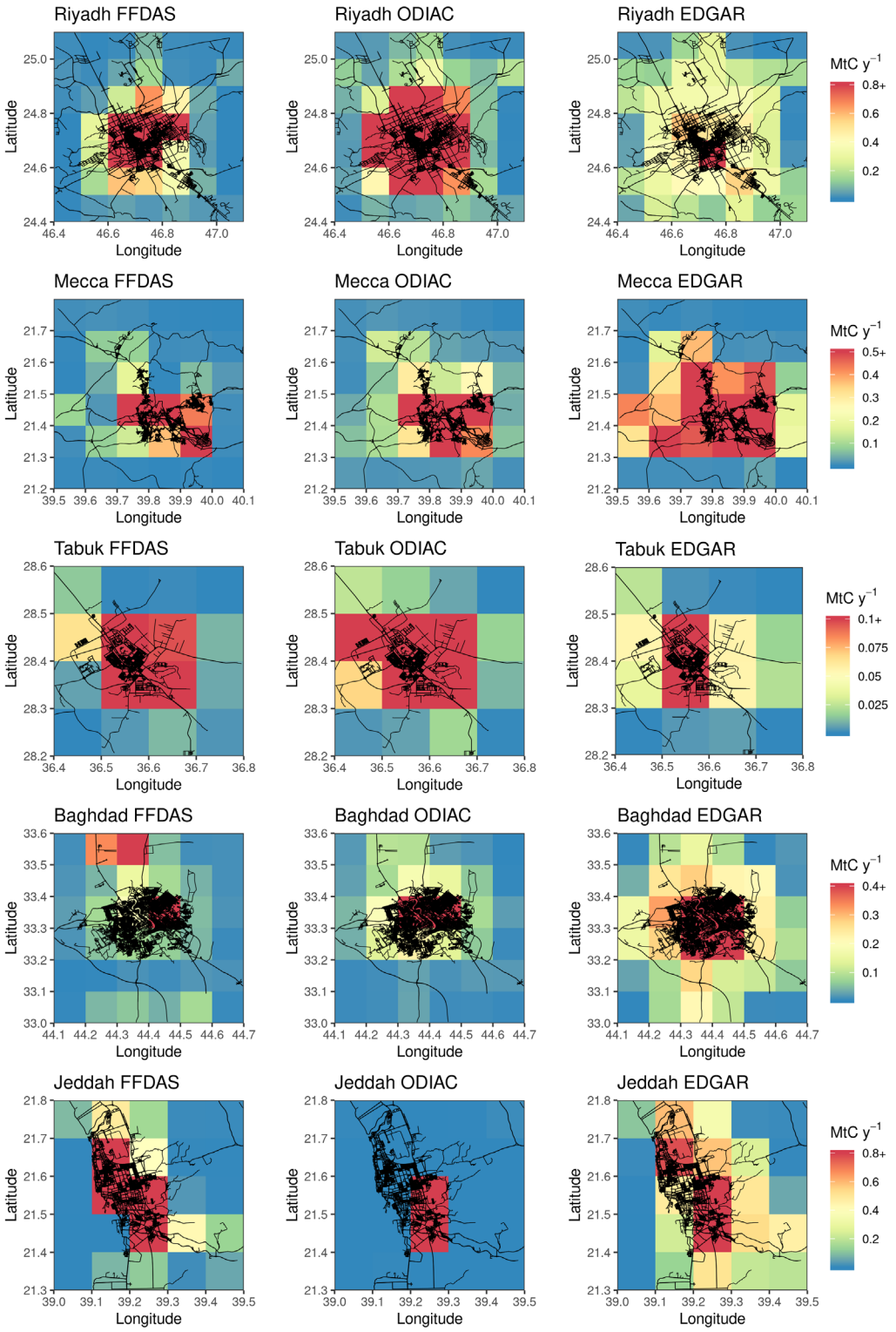

Figure 2.2 Emissions representations of each city of interest for each inventory at a spatial resolution of 0.1 x 0.1° (colors) and roads in that domain (black). At the urban scale, these representations show more clear differences in the spatial distribution and magnitudes of the emissions than at the regional scale. Note that the ODIAC representation for Jeddah has an error due to a mismatch between two gas flare nightlight data sources at that particular location (see Section 2.6.2, Supporting Information); we proceed in this paper by treating it as though there is no error to understand how our methods handle the mismatch.

domains of interest and compare their magnitudes. The emissions of the three inventories are plotted over each domain of interest on a linear scale and compared visually. The correlation between maps is also determined by finding the Kendall rank correlation coefficients between pairs of inventory representations of each city. We also generate cumulative emissions curves by arranging emissions for each inventory in ascending order, taking the cumulative sums of the emissions, and plotting these cumulative sums against each other.

2.2.2 Atmospheric Data and Methods

2.2.2.1 OCO-2 XCO_2 Data and Preprocessing

We use bias-corrected OCO-2 Lite, version 8 data from September 2014 to April 2017 (O'Dell et al., 2018). Overpasses are initially chosen by proximity to the cities of interest; the original net cast for overpasses is quite wide for exploratory purposes, such that any observations within an 8 x 8° box surrounding the location coordinates of the city of interest are selected. These are further filtered by location of observations, to ensure that there are observations present near the city center (at least within 50-75 km away) as well as further out for the purposes of background calculation (within 200-350 km away). Then, overpasses are screened visually for sufficient number, distribution (no or few gaps), and density of observations. The data are then preprocessed by finding the medians of the up to eight cross-track footprints at each measurement time step, so as to remove outliers. Then a block average of the medians is taken in bins of 0.1° by latitude. Most of the overpasses are observed in Nadir Mode, though 7 are in Glint Mode. For Riyadh, Mecca, Tabuk, Baghdad, and Jeddah, we use 11, 6, 3, 3, and 3 overpasses, respectively. The OCO-2 data are not filtered by warn level; see Section 2.6.1 (Supporting Information) for details. Table 2.5 details the 26 overpasses used in this study and

Figure 2.7 maps the overpasses in relation to the cities (in Section 2.6.3, Supporting Figures and Table).

2.2.2.2 Simulations of XCO_2

We simulate XCO_2 with the inventories such that the simulated XCO_2 is directly comparable to the values observed by OCO-2; in this way we can quantify the relationship between the observations and the models. Most of the steps used to simulate XCO_2 follow the methodology of the "X-STILT" approach described by Wu et al. (2018), which utilizes the STILT model to simulate urban XCO_2 values. Here we summarize our process to simulate XCO_2, noting any deviations in the methodology from the X-STILT formulation.

2.2.2.2.1 Summary of Approach

We simulate XCO_2 values along the OCO-2 tracks of interest at a resolution of 0.1° by latitude. The basic process of these simulations at each along-track location involves using GDAS reanalysis products at 0.5 x 0.5° resolution to drive the STILT model, which releases air parcels from each of *n* prescribed column receptor locations backward in time and calculates a "footprint." Footprints are a measure of the sensitivity of the mixing ratio to surface fluxes and are an indicator of the upstream influences of the chosen receptor locations; they are in units of ppm/(μmole/m^2/s). The footprints are then convolved with fluxes given by the three emissions inventories to determine an enhancement at that column receptor location. Combining these enhancements with a background value and a biospheric enhancement gives a full, simulated concentration $CO_{2.sim,n}$ at the *n*th level in the column (Section 2.2.2.2.2). These concentrations are then weighted with the satellite's averaging kernel profiles to determine a single column value of XCO_2 ($XCO_{2.sim.ak}$) that is directly comparable to an OCO-2 XCO_2 value (Section 2.2.2.2.3).

2.2.2.2.2 Simulating CO_2 Concentrations at each *n*th Level

To reduce computational time, we differentiate between simulating CO_2 concentrations at each *n*th level either above a height designated as *MAXAGL*, or below or at this height. We define *MAXAGL* to be the maximum release height of air parcels with STILT. For this study, we use a *MAXAGL* of 6 km for wintertime overpasses, following Wu et al. (2018), and 10 km for non-wintertime overpasses.

At each level at or below *MAXAGL*, we simulate a value of $CO_{2.sim,n}$ by following the basic approach described in 2.2.2.1. Using STILT, we release 100 air parcels at each column receptor 48 hours backward in time, with our column receptors located at every 100 m up from the surface to 3 km for wintertime overpasses (5 km for non-winter), then every 500 m up to the 6 km. Non-wintertime overpasses additionally have receptors located every 1 km up to the 10 km *MAXAGL*. This results in a footprint for each column receptor to be convolved with the emissions inventories to produce enhancements.

To capture the temporal variations of emissions, we use the Temporal Improvements for Modeling Emissions by Scaling (TIMES) scaling factors in combination with the inventory fluxes (Nassar et al., 2013). The TIMES scaling factors can be combined with monthly $FFCO_2$ emissions inventories to account for both diurnal and within-week variability, with 24 hourly grids capturing the diurnal variations and seven daily grids capturing the weekly variations.

The STILT model simulations span several years to match observations from late 2014 to mid-2017, whereas the emissions inventories do not have the same availability. Since FFDAS has an hourly temporal resolution for just the year 2014, we aggregate the data up to monthly emissions by day of week and hour, such that we have 24 grids per day of week in each month to maintain both the diurnal and weekly cycles internal to the 2014 data set. (Note that these

internal diurnal and weekly cycles are derived through application of the TIMES scaling factors.) For ODIAC, we use years 2014-2016 to correspond with the appropriate years' simulations, and use the 2016 grids for the 2017 simulations. As these are monthly grids, we use both the diurnal and weekly TIMES scaling factors along with the ODIAC datasets, which we keep in their native spatial resolutions of 1 x 1 km and also aggregate up to 0.1 x 0.1°. For EDGAR, the most recent year available is 2012; we use this annual file along with both the diurnal and weekly TIMES scaling factors. Each of these adjusted emissions inventories are then convolved with footprints to find the fossil fuel contribution to $CO_{2.sim,n}$, or the enhancements.

We also simulate the natural contribution to $CO_{2.sim,n}$ by using biospheric flux data from CarbonTracker-NearRealTime (CT-NRT) v2016 and v2017. These biospheric fluxes are convolved with footprints in a similar fashion to the inventory $FFCO_2$ fluxes. Fluxes from oceanic and fire-related sources are deemed negligible in comparison to the anthropogenic and biospheric fluxes for this domain and not included.

The boundary conditions to add to the natural and fossil fuel enhancements are determined using the trajectory endpoint technique. CO_2 concentrations sourced from CT-NRT 3-D mole fractions of CO_2 that correspond with the endpoints of the STILT trajectories serve as the boundary conditions. These are then combined with the enhancements to find the full $CO_{2.sim,n}$ value at each level. This technique differs from that suggested by Wu et al. (2018). Above *MAXAGL*, each simulated level *n* is defined by the OCO-2 retrieval levels. $CO_{2.sim,n}$ is found by finding the CT-NRT 3-D mole fractions of CO_2 corresponding to the OCO-2 levels.

2.2.2.2.3 Simulating XCO_2 to be Directly Comparable to OCO-2 Observations

The OCO-2 sensor has sensitivities to CO_2 that are different at varying heights of the atmosphere, which are characterized by the satellite's averaging kernel profiles. OCO-2 thus

retrieves XCO_2 values by using these averaging kernel profiles, which are the product of the satellite's normalized averaging kernel profiles (AK_{norm}) and pressure weighting function (PW), to find the relative weight between the observed "true" profile and the *a priori* profile ($CO_{2,ap}$) (Wu et al., 2018). To make a direct, 1:1 comparison between a simulated profile and the OCO-2 retrieved profile, we thus use the same OCO-2 weighting functions, with the simulated, unweighted column serving as the "true" profile. Our simulated profile weighted with the averaging kernel, $XCO_{2.sim.ak}$, is thus determined as follows, adapted from (O'Dell et al., 2012) and reproduced from Wu et al. (2018):

$$XCO_{2.sim.ak} = \sum_{n=1}^{nlevel}\left\{AK_{norm,n} \times PW_n \times CO_{2.sim,n} + \left(I - AK_{norm,n}\right) \times PW_n \times CO_{2.ap,n}\right\}, \quad (2.1)$$

where I is the identity vector and n is the X-STILT release level. Since the X-STILT levels do not match the 20 levels prescribed by OCO-2, we linearly interpolated AK_{norm}, PW, and $CO_{2.ap}$ values from the OCO-2 levels to the X-STILT levels. Note that these values undergo the preprocessing described in Section 2.2.2.1.

2.2.2.3 Further Data Filtering

After the overpasses are simulated with our X-STILT model, they are filtered by leveraging the STILT output to calculate footprint (influence) values from the city of interest; they pass through the filter if any modeled point at 100 m AGL in the simulated overpass pass a threshold for footprint values (average of 0.01 ppm/(μmole/m^2/s)). In this way, overpasses only pass the filter and are analyzed if they indicate influence from within the city of interest.

Footprint-filtered overpasses are then further filtered by returning to the observations. These remaining overpasses are filtered by their viability for differentiating enhancements from what we refer to as the tail values, which are XCO_2 values that are not enhanced and help define the background specific to each overpass when averaged. In part, this differentiation is done

through visual inspection, as well as by selecting overpasses whose enhanced values on average exceeded the tail values by at least 0.2 ppm.

2.2.2.4 Determining Directly Comparable OCO-2 and Simulated Enhancements

Our interest in both the simulations and the observations of OCO-2 are the enhancements due to fossil fuel combustion within the cities of interest. Thus, our main points of comparison are between the simulated fossil enhancements and observed enhancements, which are determined by subtracting the simulated boundary conditions and natural contributions of XCO_2 from both the OCO-2 observations ($XCO_{2.obs}$) and the fully simulated $XCO_{2.sim}$ from Section 2.2.2.2.3. (Note that subtracting the boundary conditions and natural contributions from $XCO_{2.sim}$ is not exactly the same as solely convolving the footprints with the emissions inventory fluxes; such a convolution ignores the contribution to $XCO_{2.sim}$ from the averaging kernel and *a priori* profiles as described in Section 2.2.2.2.3.)

An additional constant value that corrects the bias between the CarbonTracker-derived background and the observed background is also either added to or subtracted from the full XCO_2 values to determine final enhancements. This bias correction term is determined as follows. For Riyadh, a swath of 3° of latitudinal points for each observed overpass of interest is examined, differentiating between the enhanced points and the other points, the tails of the overpass, by means of expert judgment and comparison with the simulated enhancements from the city of interest. The other cities undergo a similar process, except not all overpasses use 3° of data. The other cities have enhancements from other nearby cities, artifacts to be avoided, or their own large-scale variability, so the length of the tails may be shorter for non-Riyadh overpasses to only capture the local background. Next, for all cities, the tails are averaged to a

single constant. This constant serves as a bias-correction term between the CarbonTracker-derived background and the OCO-2 observed background.

This methodology for determining the background differs from that described in Wu et al. (2018). However, we choose to use this methodology for two reasons: computation time savings and the ability to capture upwind variability using CarbonTracker. Regarding the former reason, Wu et al. (2018) opts for a method that requires a plume defined through forward-modeling with STILT that is computationally expensive, especially given the number of overpasses analyzed in this study. Regarding the latter reason, we use CarbonTracker to define our boundary conditions in this study to fully capture variability upwind of the urban plume. Any biases between the OCO-2 observations and the CarbonTracker-derived background are then corrected for using observations with little local influence. In this way, we account for large-scale variation in XCO_2, while also correcting biases using the observations.

2.2.2.5 Quantifying the Relationship between the Modeled and Observed Enhancements

We evaluate the relationship between the simulated and observed enhancements for each overpass of interest by generating cumulative enhancement curves, which is a modified area-under-the-curve technique. In this technique, the data points that are enhanced for each day or in aggregate are placed in ascending order, and their cumulative sums are plotted against each other, in a similar fashion as to the cumulative emissions curves described in Section 2.2.1.3. This method allows for the examination of the spatial distribution of enhancements while simultaneously latitudinally integrating the areas under the curves generated by the enhancements. Each modeled curve is also scaled such that the total summed enhancements for that particular overpass was equivalent to that of the corresponding OCO-2 sum, to make scaled cumulative enhancement plots.

The scaling factors used to scale the cumulative enhancement plots are combined using a bootstrap method to determine a single, mean scaling factor that relates the modeled enhancements to the observed ones—one scaling factor for each inventory per city. The bootstrap method is as follows. For each overpass analyzed, there is a set of scaling factors (with one scaling factor per inventory representation) and a single associated uncertainty. Each scaling factor and its corresponding uncertainty is assumed to have a Gaussian distribution, with the scaling factor as the mean and the uncertainty as the standard deviation. It is resampled based on this distribution, outputting a single resampled scaling factor. All of the resampled scaling factors for an inventory per city further undergo a bootstrap resampling, using the mean statistic. This dual set of resamplings is repeated one million times for each inventory per city. The mean of the vector of one million resamplings is the mean scaling factor for each inventory per city. Meanwhile, the 5th and 95th percentiles serve as the bounds of the 90% confidence interval.

The mean scaling factors are then multiplied by the corresponding prior emissions magnitude, in order to scale the prior inventory estimates by the OCO-2 observations.

2.2.2.6 Evaluation of Uncertainties

We determine uncertainties for both the modeled and OCO-2 integrated enhancements. Combining these two sources of uncertainty also results in an uncertainty for the scaling factors and for our estimates of $FFCO_2$ emissions for each city.

For the model integrated enhancements, we leverage the 33% fractional uncertainty over five overpasses due to horizontal and vertical transport evaluated by Wu et al. (2018), which was estimated by quantifying the effect of the inclusion in the STILT model of a wind error component derived from radiosonde observations (horizontal) and calculating the root-mean-squared errors between enhancements found with different rescaled planetary boundary layer

heights (vertical). To find an average fractional uncertainty per overpass, we multiply this value by the square root of 5, resulting in a ~74% fractional uncertainty to describe the transport uncertainty for a single overpass's latitudinally integrated modeled enhancements. While their study focused on Riyadh, we apply this uncertainty to all of the Middle Eastern cities in the study. This transport uncertainty is the only one we use for the modeled integrated enhancements. Since the emissions inventories do not have their own uncertainties and this study attempts to constrain those very emissions, we do not include a measure of the inventory uncertainties for the modeled integrated enhancements.

For the OCO-2 integrated enhancements, we first evaluate the uncertainties for each binned XCO_2 value at each 0.1° of latitude. We approximate this by considering the standard deviation of the medians in each bin.

The uncertainties related to the background also are included, since the OCO-2 enhancements are determined by subtracting the background from the total XCO_2. Contributions to the background uncertainty come from both the OCO-2 tails and from CarbonTracker. The tails' uncertainty contribution is determined by first only using the bin-level standard deviations that correspond only to the tails. Those values are added in quadrature and divided by the number of values used to find the average tail values. Meanwhile, the CarbonTracker contribution to the uncertainty is calculated by finding the standard deviation of the CarbonTracker-derived values (both boundary condition and biospheric influence) within the enhanced region of each overpass.

The contributions to the OCO-2 uncertainties (OCO-2 spread, tail spread, and CarbonTracker spread) are then added in quadrature, resulting in uncertainty values corresponding to each binned XCO_2 enhancement at each 0.1° of latitude. The combined binned

uncertainties are then also added in quadrature, resulting in an uncertainty corresponding to the OCO-2 enhancements for a single overpass. A fractional OCO-2 uncertainty is found by dividing this single uncertainty by the integrated enhancements for each overpass.

To determine the uncertainties for each scaling factor corresponding to a single overpass, the 74% transport uncertainty and the fractional OCO-2 uncertainties (generally substantially smaller than the transport uncertainties) are combined by taking the square root of the sum of the two squared fractional uncertainties.

The overpass-level uncertainties are then used as inputs into the bootstrap method described in Section 2.2.2.5 to determine 90% confidence intervals for the mean scaling factors.

2.3 Results

2.3.1 Inventory Representations of the Middle East in Emissions-Space

On a regional scale, we determine that the three studied emissions inventories represent the Middle East quite differently, from the perspective of both spatial distributions of emissions and magnitudes of emissions. In Figure 2.1, we show inventory representations of a domain that approximates the Middle East. A visual appraisal of these representations suggests that the way cities, roads, and transport are distributed throughout the Middle East vary from inventory to inventory. The total emissions magnitudes for this domain varies between inventories as well. For FFDAS, ODIAC, and EDGAR, respectively, the emissions represented in this domain come out to 697, 789, and 722 MtC y^{-1}.

2.3.2 Inventory Representations of Urban Centers in the Middle East

Emissions representations of each city of interest for each inventory at 0.1 x 0.1° resolution are displayed in Figure 2.2, with emissions scaling linearly. Note that the displayed

scales are the same for each representation for a city, though the maximum values vary for each city.

Figure 2.2 illustrates that the emissions within the chosen domains are distributed quite differently, to the point that it is not immediately apparent that different inventories are even representing the same city, apart from overlapping of the highest-valued grid cells in each representation. Same-city patterns across the inventories are not apparent in this form. A sense of these distributions can be gained by determining the Kendall rank correlations between pairs of inventories for each city, as shown in Table 2.2. Apart from those correlation coefficients associated with Jeddah's representation by ODIAC, which has an error with its representation (see Section 2.6.2, Supporting Information), the remaining correlation coefficients suggest that the same-city inventory representations are correlated with each other to varying degrees without being identical.

The magnitudes of the emissions within each representation, i.e. the sums of the emissions in each displayed domain, are listed in Table 2.3. Other than in the case of Jeddah, the FFDAS and ODIAC summed magnitudes for all other cities are more similar to each other than with EDGAR.

The maximum values in each domain are also listed in Table 2.3. For each city, the maximum valued grid cell among the three inventories is almost always represented by FFDAS. Those maximum values are often significantly higher than those of the other inventories, sometimes multiple times higher. This suggests that for FFDAS, large point sources carry more weight than they do for the other two inventories, whether that is due to the method of distribution of emissions by proxy, or due to large point source data containing much higher values for FFDAS than those used by the other inventories.

	FFDAS-ODIAC	*ODIAC-EDGAR*	*EDGAR-FFDAS*
Riyadh	0.81	0.81	0.78
Mecca	0.70	0.82	0.69
Tabuk	0.77	0.85	0.78
Baghdad	0.64	0.67	0.51
Jeddah	-0.010*	-0.15*	0.71

Table 2.2 Kendall rank correlation coefficients between pairs of inventories for each studied city. For all paired inventory domains other than those including the ODIAC representation of Jeddah (labeled with *), the inventory representations are correlated with each other.

	Pop. (million people)	*Emissions Sums (MtC y^{-1})*			*Maximum Values (MtC y^{-1})*		
		FFDAS	*ODIAC*	*EDGAR*	*FFDAS*	*ODIAC*	*EDGAR*
Riyadh	5.2	29.0	28.4	18.2	17.6	6.36	7.26
Mecca	1.5	5.81	6.54	7.89	3.10	1.88	0.616
Tabuk	0.4	2.48	2.37	0.548	1.65	0.837	0.137
Baghdad	6.7	4.74	3.60	7.00	1.91	1.26	1.38
Jeddah	3.4	13.4	4.08*	19.4	5.74	2.73*	8.97

Table 2.3 City populations (April 2010 census data for Saudi Arabian cities, October 2009 estimate for Baghdad) (Brinkhoff, 2018), along with, the total sums (magnitudes) and maximum values of CO_2 emissions within each city's domain, for the three inventories. (Jeddah values marked with * indicate the error in their representation by ODIAC; see Section 2.6.2, Supporting Information.)

In Figure 2.3, we show cumulative emissions curves for each inventory and city. These curves reaffirm that the total magnitudes of emissions for FFDAS and ODIAC tend to be more similar than they are for EDGAR. These figures also give us a further understanding of the distribution of the emissions for each inventory representation of a city. For each city (other than Jeddah), the FFDAS and ODIAC curves are more rounded out than the EDGAR curves, suggesting that more of the total emissions in the domain of interest are due to fewer and higher-emission grid cells. This also means that EDGAR's urban representations distribute emissions more evenly than the other two inventories. This could be attributed to EDGAR distributing on-road emissions more homogeneously than the other inventories, which was discussed by Gately and Hutyra (2017) for an earlier version of EDGAR. Regarding the greater similarity in distribution between the FFDAS and ODIAC curves, this could be due to their shared use of

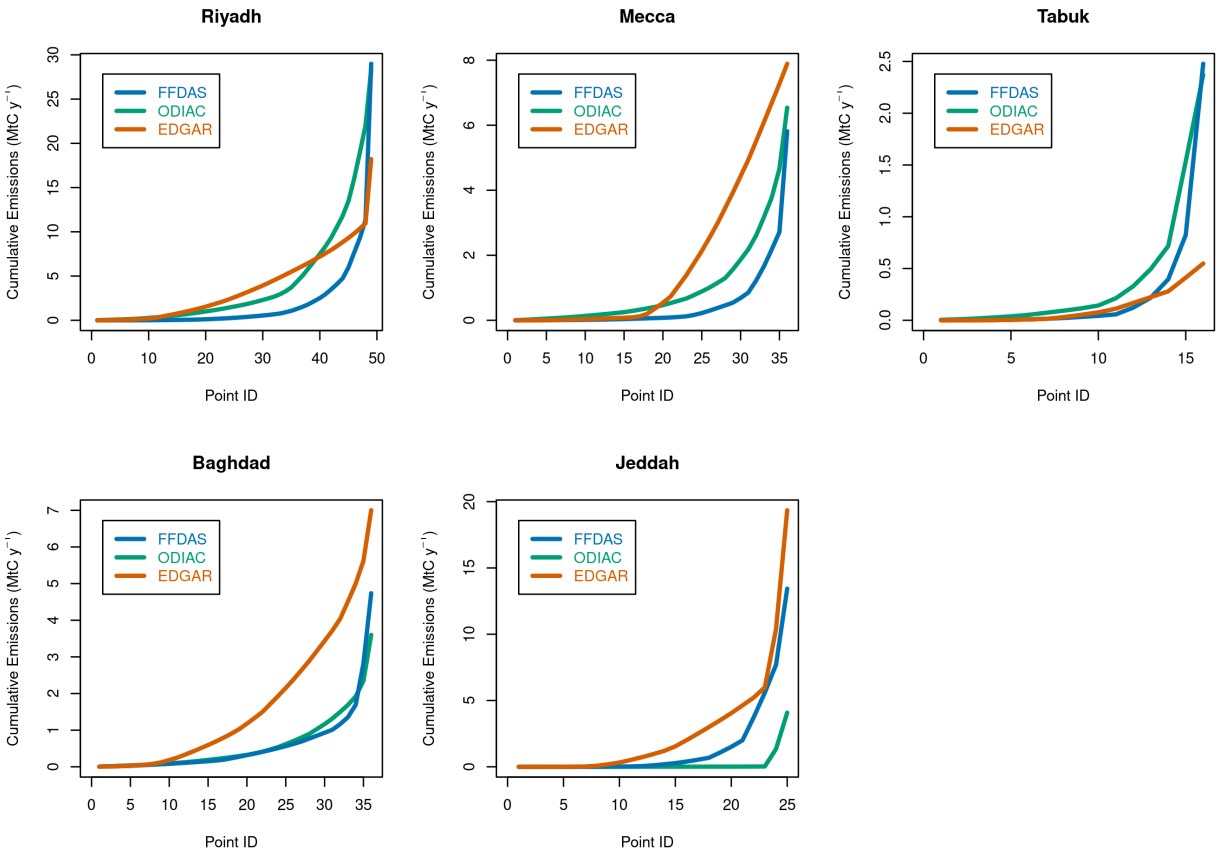

Figure 2.3 Cumulative emissions curves for each inventory and city. These figures simultaneously represent magnitudes and spatial distributions of emissions for each city and inventory. The total emissions magnitudes are represented by the rightmost point, whereas the distributions are represented by the shape of the curves, with straighter curves being more evenly spread than those that are more rounded out. Based on these curves, the EDGAR representations of the cities of interest differ most from the other inventories, with more variant magnitudes and more evenly spread emissions.

nighttime lights as a proxy for distributing emissions. In any case, all curves show EDGAR as having the most evenly spread emissions, suggesting that all studied cities follow similar patterns of emissions distributions for the same inventories, in spite of relative differences in emissions magnitudes.

2.3.3 Simulated Atmospheric Concentrations of Urban Centers in the Middle East

The magnitude and spatial distribution differences between the different inventories across the cities of interest are found not only in their emissions, but also in their respective

simulations of concentrations. As an example to show the manifestations of these differences into concentration-space, Figure 2.4 shows the enhancements for four different observed and modeled overpasses, for three cities and all on different dates. Note that we have included ODIAC in its native resolution ("ODIAC" in green) and at the resolution aggregated to match the other inventories ("Agg ODIAC" in pink). In these sample overpasses, all of the modeled overpasses capture the urban plumes, though the distribution and the magnitudes of the captured enhancements vary by inventory used in the model. While these modeled overpasses' attributes are not representative of all modeled overpasses for their respective cities, they do show how the differences in emissions magnitudes and distributions have manifested themselves into concentration-space on these particular days due to atmospheric transport. A similar figure to Figure 2.4 but with observational uncertainties included is in the Supporting Information as Figure 2.8. Corresponding observed and simulated total XCO_2 plots are found in Figure 2.9.

2.3.4 Comparison between Observed and Simulated Enhancements

When comparing the simulated enhancements by latitude to the observed ones, such as in Figure 2.4, we continue to see that the magnitude and distribution of enhancements in the models differ from what OCO-2 has observed. In this figure, we show enhancements vs. latitude, where the black squares are the binned observed enhancements and the other colored points show the different modeled enhancements. In certain cases, the magnitudes of the enhancements seem fairly consistent between the observations and models (e.g., Riyadh 2016-02-16); in others, they are quite different (e.g., Mecca 2016-03-22), which we attribute to daily variations in emissions that are not captured by the temporal resolution of the models. However, the models capture the urban plumes that the satellite has observed, suggesting that there is a fidelity to the models. There are also cases in which the observed and the modeled plumes are latitudinally shifted from

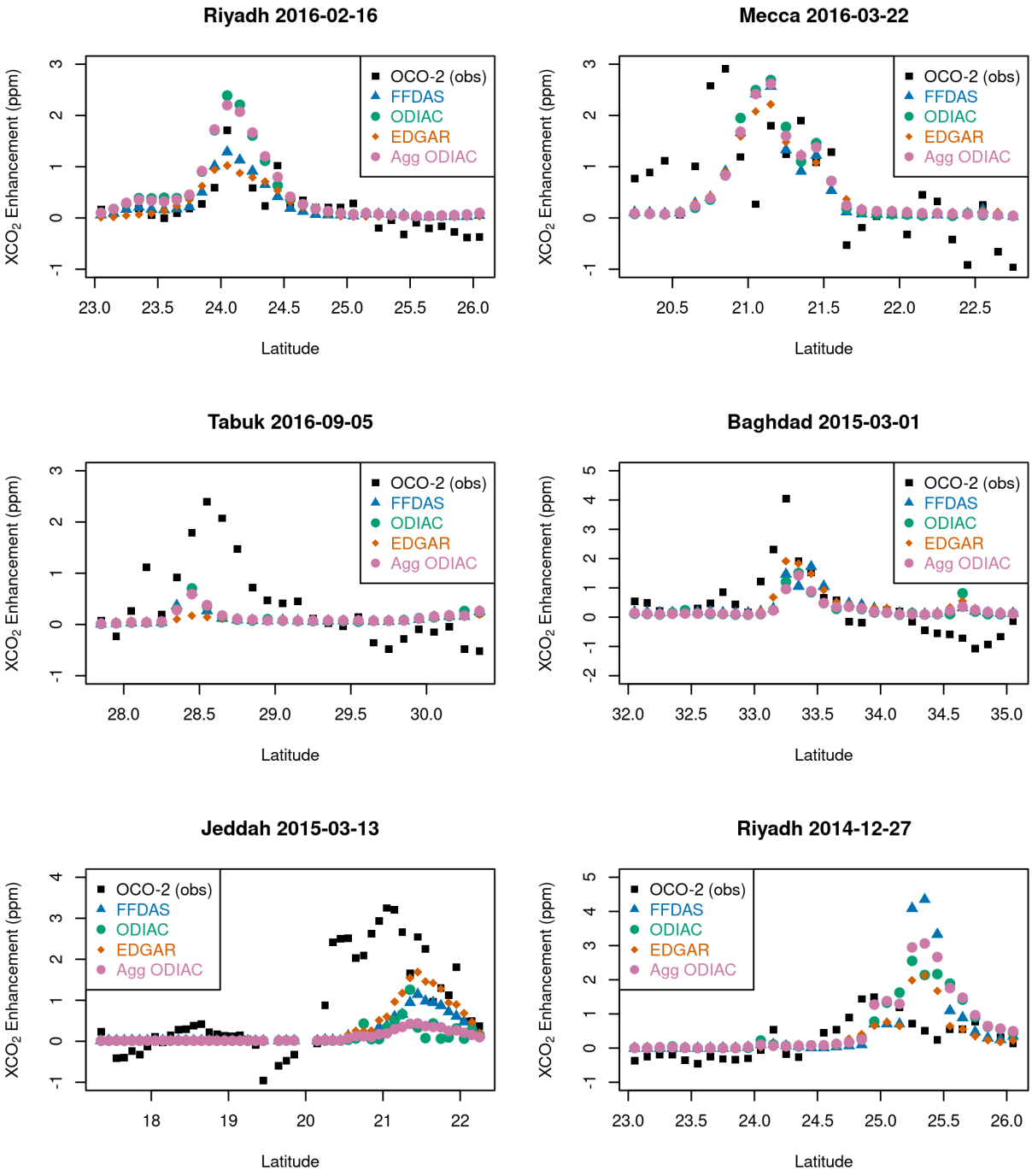

Figure 2.4 Enhancements of sample observed (black) and modeled (other colors) overpasses for different cities and days. Note that we have included ODIAC in its native resolution ("ODIAC" in green) and at the resolution aggregated to match the other inventories ("Agg ODIAC" in pink). All modeled overpasses capture the urban plume depicted in the observations. Differences in magnitudes and spatial distributions of the emissions manifest themselves in the differing representations of the enhancements. The sample overpasses for Jeddah on March 13, 2015 and Riyadh on December 27, 2014 depict latitudinal shifts in the urban plume as a result of transport errors that offset the location of the plumes. Our integral method of comparison between the observed and modeled enhancements is not inhibited by these latitudinal shifts.

each other, as in the overpass of Riyadh 2014-12-27. This latitudinal shift, which is likely due to transport errors, demonstrates the model's ability to capture the urban plume, even when transport errors offset the plume's location from the satellite-observed plume.

The differing distributions of enhancements and latitudinal shifts in the plumes between the simulations and the observations lend themselves to an integral method of comparing enhancements. This process is captured in the cumulative enhancement curves shown in the four unscaled examples (left) in Figure 2.5, which correspond to four of the same overpasses shown in Figure 2.4. The cumulative sum, or magnitude of the enhancements for each overpass are captured by the rightmost points on each curve, visually representing the differences in magnitude between the different models and the observation for each respective overpass.

In the scaled cumulative enhancement curves of Figure 2.5 (right), the summed modeled enhancements are all scaled to match those of the observed enhancements for each overpass. These curves make the differences in distribution of enhancements within the urban plume clearer. Despite the distributions of the enhancements still differing, this method normalizes the magnitudes of the modeled enhancements to those of the OCO-2 observations, resulting in overpass-specific scaling factors for each inventory, for each studied day. For the overpasses in Figure 2.5, we have scaling factors ranging from 0.50 (aggregated ODIAC simulation of overpass for Riyadh 2014-12-27) to 2.77 (aggregated ODIAC simulation of overpass for Baghdad 2015-03-01). Scaling factors vary from day to day for a specific city.

All overpass-level scaling factors are combined into a singular scaling factor for each inventory and city with the bootstrap method discussed in Section 2.2.2.5. These mean scaling factors are listed in Table 2.4 with their 90% confidence intervals. These scaling factors are determined by comparison with midday satellite observations, meaning we correct the whole

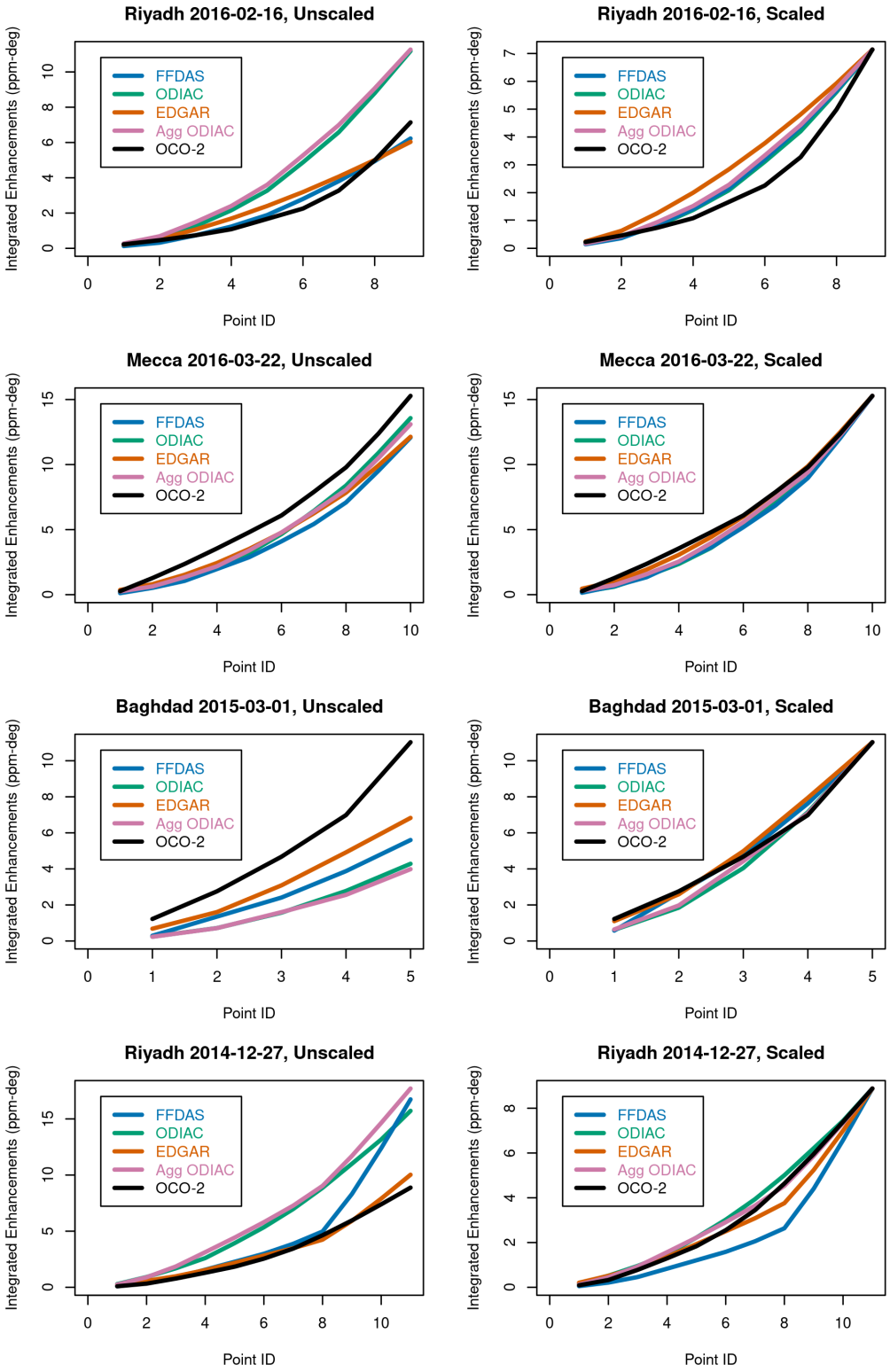

Figure 2.5 Cumulative enhancement curves, unscaled (left) and scaled (right), corresponding to the several of the overpasses shown in Figure 2.4. The unscaled curves illustrate magnitude differences between the models and observations, while the scaled illustrate the spatial differences in the enhancements. The unscaled modeled curves on the left are scaled on the right to match the integral of the OCO-2 enhancements. We use those scaling factors to quantify the relationship between the modeled and observed enhancements.

	FFDAS	*ODIAC*	*EDGAR*	*Agg. ODIAC*
Riyadh	1.1 (0.5, 1.8)	1.2 (0.5, 1.9)	1.8 (1.0, 2.6)	1.1 (0.4, 1.8)
Mecca	1.7 (0.7, 2.7)	1.7 (0.6, 2.8)	1.4 (0.5, 2.4)	1.6 (0.6, 2.6)
Tabuk	3.4 (0.9, 5.8)	3.1 (0.7, 5.4)	8.7 (4.6, 12.7)	3.1 (0.7, 5.4)
Baghdad	3.0 (1.3, 4.9)	3.0 (1.6, 4.4)	2.2 (0.8, 3.7)	4.0 (1.8, 6.4)
Jeddah	3.3 (2.1, 4.4)	6.7 (5.5, 7.9)	1.9 (0.9, 2.9)	6.6 (4.9, 8.5)

Table 2.4 Mean scaling factors that represent the relationship between the inventory modeled enhancements for each city and the satellite observations. Their respective 90% confidence intervals are in parentheses.

inventory based on midday comparisons, an approach that relies on the diurnal model representation being accurate.

2.3.5 Estimates of Urban Emissions in the Middle East

The mean scaling factors multiplied with their respective inventories result in new, scaled estimates of urban emissions in the Middle East. These estimates are shown in the first five panels of Figure 2.6. In these panels, for each city, the colored bars represent the inventory magnitudes, while the gray bars represent the corresponding scaled emissions estimates, with 90% confidence intervals. Though some of the confidence intervals are quite large, we find that many of the scaled inventory estimates and their corresponding confidence intervals are either higher than the prior inventory emissions magnitudes or overlap only slightly. This suggests that, as compared to the OCO-2 observations, the global emissions inventories are underestimating several of the urban representations of afternoon emissions. This is the case for the EDGAR representation of Riyadh, and all prior inventory representations of Tabuk, Baghdad and Jeddah. The mean scaled estimates all exceed the prior inventory estimates for all cities.

As discussed in Section 2.3.2, all of the prior inventory representations of each city demonstrate varying spatial distributions and magnitudes of emissions, with no specific measure of uncertainty. However, our work has determined emissions estimates for each city that converge across inventories and are indistinguishable within the confines of the confidence intervals. Based on a limited number of observations, we have thus found emissions magnitude

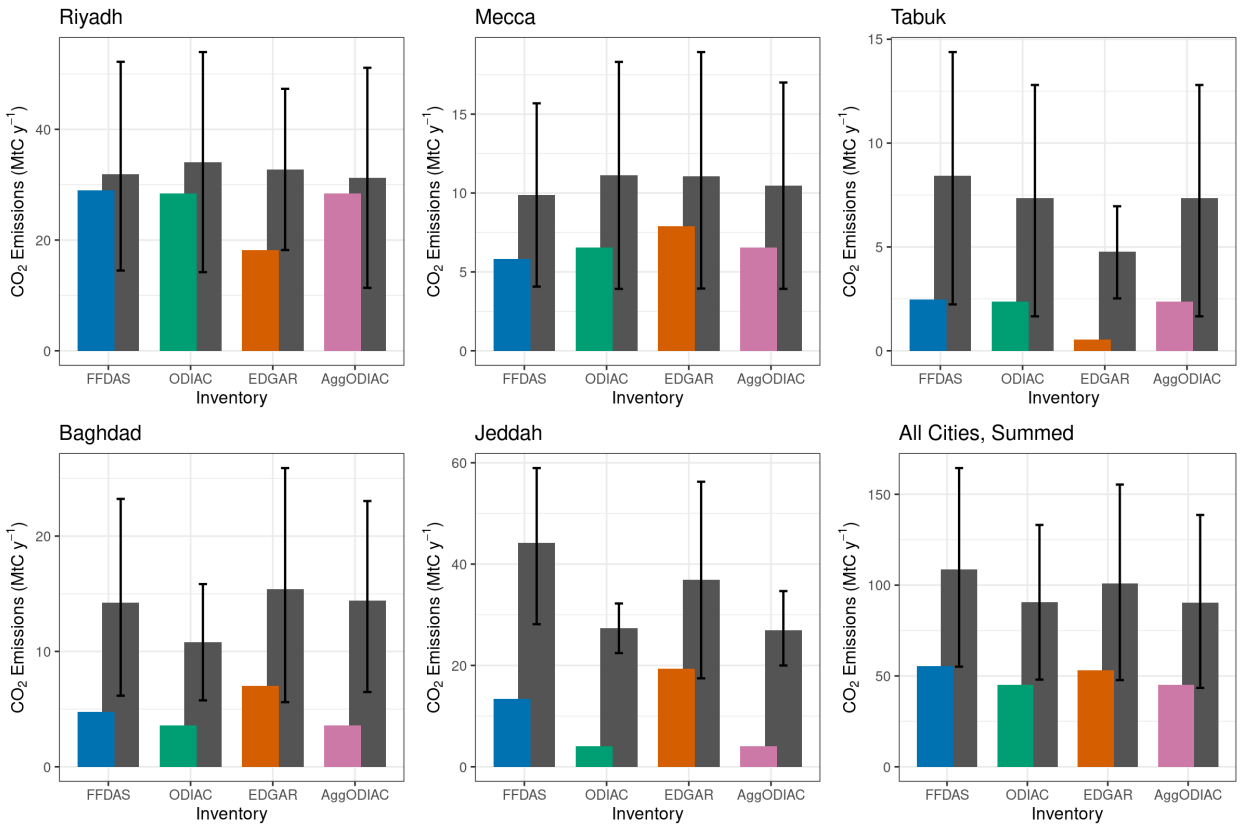

Figure 2.6 Prior (colored) and scaled (gray) emissions magnitude estimates for our five cities of interest: Riyadh, Mecca, Tabuk, Baghdad, and Jeddah, as well as the sums of the emissions of all five cities. The black lines on the gray bars represent the 90% confidence intervals. When taken in aggregate, the prior emissions magnitudes underestimate emissions as compared to those scaled by our emissions scaling factors.

estimates for each city that are insensitive to the distributions of the prior inventory emissions and that adjust the magnitudes of the emissions, in spite of the differences between the priors. This is a surprisingly powerful result, especially given the significance placed on the accuracy of high-resolution inventories for previous ground-based inversion studies, (e.g., Lauvaux et al., 2016; Oda et al., 2017). While the space-based column measurements of CO_2 concentrations have not given us information to quantify the spatial distribution of emissions, it has allowed us to bypass the differences in spatial distributions by providing an integral constraint to quantify urban magnitudes of CO_2 emissions.

Sums of the different emissions estimates for each inventory (prior and scaled with confidence intervals) across the five cities are shown at the bottom right in Figure 2.6. As shown

in this figure, our scaled estimates, with minimal or no overlap between our 90% confidence intervals and the prior emissions, demonstrate larger emissions than those estimated by the unscaled inventories. In other words, the five cities are, when summed, underestimated by their prior inventory representations as compared to their observation-scaled counterparts.

2.4 Discussion

2.4.1 Implications for Middle Eastern Emissions and their Representations

Our scaled emissions and their respective 90% confidence intervals constrain emissions of five Middle Eastern cities. We find that for Riyadh, EDGAR representations underestimate afternoon emissions as compared to what our observations suggest. Moreover, for Tabuk, Baghdad, and Jeddah, almost all prior inventory representations underestimate afternoon emissions. After summing emissions for all five cities, prior estimates for all inventories overlap minimally or not at all with lower bounds of the scaled emissions confidence intervals. Thus, we can highlight that, in sum, inventory representations across all five cities underestimate afternoon emissions. If prior and scaled summed emissions are averaged across three inventories, the average scaled summed emissions of all cities are 100 MtC y^{-1} (50, 151), which is 2.0 (1.0, 3.0) times greater than the average prior summed emissions.

Our scaled estimates have regional significance. The entire Middle East as a whole emits approximately 700 MtC y^{-1} (Boden et al., 2016). The difference between the average scaled and prior summed emissions for our five cities is 49 (-1, 100) MtC y^{-1}, which is equivalent to ~7% (0%, 14%) of total Middle Eastern emissions. Meanwhile, the emissions of Saudi Arabia and Iraq are approximately 200 MtC y^{-1} (Boden et al., 2016). Prior and scaled summed emissions come out to ~26% and 50% (25%, 76%) of the Saudi Arabian and Iraqi total, respectively. The difference between the average scaled and prior summed emissions for the five studied cities

comes out to ~25% (-1%, 50%) of the sum of the two countries' total emissions. However, further work at a national or regional scale is needed to determine whether the underestimation in these urban emissions representations suggest either underestimation in the national totals or if our estimated discrepancy is more a result of misallocation of the emissions throughout the studied countries.

This work demonstrates the ability to use satellite observations to constrain urban emissions of cities of a variety of sizes. Given the much smaller magnitude of emissions of Tabuk as compared to Riyadh, for example, it is a surprising result that we can observe XCO_2 enhancements from a city like Tabuk from space (as in Figure 2.10 in Supporting Information). It does appear that with appropriate isolation from our strong XCO_2 signals, and with sufficient concentration of emissions and favorable atmospheric dynamics, even relatively small cities (with populations of hundreds of thousands) can be identified and have emissions quantified.

This study provides an objective evaluation of the emission downscale approaches used by the gridded emissions inventories. For the five studied cities, EDGAR's prior estimates overlap substantially with the confidence intervals of our scaled emissions estimates only for Mecca. For every other city, the EDGAR representations have emissions lower than or minimally overlapping with the lower bound of the confidence intervals we estimate. As such, EDGAR does not seem to accurately capture urban emissions of these studied cities. Why is this? EDGAR tends to distribute emissions across the urban domain more than the other two inventories. Of the studied cities, Riyadh is the largest, and EDGAR's emissions estimate is significantly lower than those of other inventories. It is possible that the distributed nature of how EDGAR grids emissions is less able to capture the high emissions intensity of such a large city than alternate gridding methods that rely on proxies such as nightlights.

FFDAS and ODIAC (and its aggregated version) both also appear to underestimate emissions in comparison with the OCO-2 data, though their emission magnitudes are within the large confidence intervals of three of our optimized estimates of individual cities. These two inventories share the use of nighttime lights as a proxy for emissions distribution, and this work is somewhat suggestive this may be a useful proxy for global representation of urban domains, but further work is required to further parse out their similarities, differences, and reasons for their underestimation of the studied urban emissions.

Our work also has the potential to identify larger errors in inventory representations of cities. For example, the ODIAC representation of Jeddah has an error in it due to a mismatch between two nightlight data sources (one for distributed emissions and the other for gas flare emissions, see Oda et al. 2018) at that particular location. This error was identified with our methods. We have not corrected for the spatial distribution error in this manuscript, but we have estimated the magnitude of emissions for Jeddah that can be used to take into account when correcting ODIAC for future versions of the inventory.

2.4.2 The Spatial Resolution of the Prior Emissions

When making direct comparisons between the emissions inventories themselves, we scale ODIAC to match the spatial resolution of the other inventories. However, in its native spatial resolution, ODIAC has 12 times higher resolution in each spatial dimension. In order to assess the effect of using the two different spatial resolutions, all modeled results include both native resolution results ("ODIAC") and results in which the resolution is aggregated up to 0.1 x 0.1° ("Agg ODIAC"). We can see in plots such as Figure 2.4 that the differences in spatial resolution do affect the individual modeled enhancement values, as the points for the two ODIAC resolutions do not perfectly overlap with each other. However, within the scope of our

study, the differences in spatial resolution do not greatly affect the outcome of our scaled emissions estimates. For each studied city, the scaled estimates across the different ODIAC resolutions provide convergent estimates. Thus, our estimates are insensitive to both the differences in the studied resolutions as well as the spatial distributions in the priors.

2.4.3 Limitations and their Implications

The OCO-2 satellite follows a sun-synchronous orbit, meaning that all the observations used in this study are temporally limited to local afternoon. Consequently, we cannot assess the diurnal cycle of our results, and our results are dependent on the accuracy of the diurnal model representation by the TIMES scaling factors. Our results thus suggest an emissions underestimation by the modeled inventory representations and their corresponding parts—namely, the convolution of the inventories themselves and the diurnal model representations. (Note that the FFDAS product has its own internal diurnal cycle embedded into the product based on the TIMES factors.) We cannot disentangle the TIMES factors from the inventory representations, which suggests that this underestimation could be a result of the inventories having too-low emissions, the TIMES diurnal cycle not fully capturing afternoon emissions, or a combination of these two factors.

Given the widespread usage of the TIMES factors for capturing diurnal cycles, it is possible that there is a systematic bias across many model representations of diurnality of emissions. The TIMES factors are potentially unable to fully correct for the lack of diurnality in the emissions inventories, particularly at the urban scale, a scale for which these factors were not designed. While this potential bias cannot be tested using OCO-2 data alone, as is done in this study, observations from multiple times of day from other satellites such as the recently launched OCO-3 could allow for assessments of the diurnal cycle represented by the TIMES factors.

Moreover, there is great potential for carbon-observing satellites in geostationary orbit to assess the diurnal cycle of CO_2 emissions, as they would be able to make multiple observations of the same locations throughout the day. For example, the upcoming Geostationary Carbon Cycle Observatory (GeoCARB) mission, which will continuously monitor column concentrations of CO_2, CH_4, and CO in the Americas throughout the day from geostationary orbit, is an important step forward in this respect (Moore III et al., 2018).

Assessments of the emissions magnitudes of inventory representations are also limited by the signal-to-noise ratio (SNR) of the OCO-2 observations. The methods used in this paper to analyze observed overpasses rely on high SNR and optimal conditions; it is possible that overpasses with higher SNR simultaneously also have stronger signals, suggesting that our results may be biased toward higher scaling factors and thus suggesting greater underestimation by the modeled representations of afternoon emissions. This potential bias highlights the limitations of not just this study but for other studies involving space-based observations for the use of urban emissions assessments. The use of space-based observations of XCO_2 can be hindered by a number of factors, such as clouds, high albedo, topography, and wind conditions; for regions at urban scales, these factors in turn limit the types of studies for which the observations can be used. Given these limitations, our study does not seek to quantify absolute emissions in the studied cities, but rather assess the inventory representations' ability to capture the variations in urban emissions that are suggested by the observations used in this study. The underestimation by these inventory representations that we have evaluated thus suggest that the inventories can be improved to better capture such variations. The larger observed swaths available at urban scales in both the target and snapshot area mapping (SAM) modes of OCO-3

will provide for much more data at urban scales, and could lessen some of the limitations associated with OCO-2 data (Eldering et al., 2019).

2.5 Summary and Conclusions

Global $FFCO_2$ emissions inventories need to be evaluated and improved at the urban-scale for the sake of carbon cycle science and urban mitigation policies. In this study, we evaluate global $FFCO_2$ emissions inventory representations of afternoon urban CO_2 emissions in the Middle East, and assess the ability of satellite observations to inform this evaluation. We find the relationship between top-down satellite observations and bottom-up simulations of XCO_2 to calculate scaling factors to adjust the prior inventory estimates of five Middle Eastern cities' emissions. Based on our findings, we estimate that the sum of the studied cities' scaled emissions are on average 2.0 (1.0, 3.0) times the prior inventory magnitudes. The underestimation of these five cities' emissions by the inventories comes out to ~7% (0%, 14%) of total Middle Eastern emissions. Our results are insensitive to the spatial differences in the inventory representations of the cities' emissions, facilitating robust quantitative assessments of urban emissions inventory representations. This is in contrast with atmospheric inverse analyses carried out with ground-based observations, which would be heavily dependent on the spatial distribution of the prior inventories. Using space-based XCO_2 observations allows us to constrain urban emissions in a fashion not previously available to us, enabling evaluation of inventories and downscaling methods at subnational scales. These results are based only on a few overpasses for each city; with additional observations, urban emissions could be even further constrained, and more cities could be included in a future study. In particular, the recent launch of the OCO-3 is especially exciting: it opens up new potential pathways for urban CO_2 emissions studies to move forward, with observations during different times of the day (instead of only the afternoon as with OCO-

2), and the new SAM mode that could capture the entire urban plume (Eldering et al., 2019). The upcoming launch of the geostationary GeoCARB mission will also help to fill in some of the observational gaps of the sun-synchronous OCO-2 and OCO-3 satellites by mapping column concentrations of CO_2 in the Americas with multiple measurements of the same locations every day. Constraining urban emissions is critically important, and satellites are paving the way for such work to be done.

2.6 Supporting Information

This section contains supplementary information, including a supporting description of data processing (Section 2.6.1), details on an error in ODIAC's representation of the city of Jeddah (Section 2.6.2), a figure and table illustrating information about the overpasses used in this study (Figure 2.7 and Table 2.5), figures illustrating additional detail to Figure 2.4 above (Figures 2.8 and 2.9), and a figure supporting the discussion in Section 2.4.1 on our ability to observe enhancements from relatively small cities (Figure 2.10).

2.6.1 Impact of Data Screening with Warn Levels on Scaling Factors

To maximize usable data, we conduct a sensitivity study to assess what level of warn level filtering to apply to the OCO-2 data used in this study. Since we have the densest observational dataset for Riyadh, we use this as a case study. We consider the impact of using data from warn levels 0, 2, 4, and 5 for the 9 overpasses that retain sufficient data density for evaluation when screened at level 0. There is a trade-off between including potentially biased data points (warn level 5) and reducing the data volume such to limit statistics (warn level 0). We find that our scaling factors are not statistically different irrespective of warn level used for Riyadh. At level 0, 2, and 4 with 90% confidence intervals, our scaling factors for FFDAS,

ODIAC, EDGAR, and aggregated ODIAC are, respectively: 1.2 (0.5, 1.9), 1.3 (0.5, 2.1), 1.9 (1.0, 2.8), and 1.2 (0.5, 1.9). At level 5, the scaling factors are all the same other than the EDGAR scaling factor, 1.9 (1.0, 2.9), which has a slight increase in its upper bound. The minimal effect of warn level screening on the scaling factors may in part be due to the differential nature of our urban analysis, and/or that we average individual soundings together and consider the variability in the soundings averaged as an uncertainty in our analysis, both of which might mitigate the inclusion of skewed soundings. Given the results of this Riyadh study, we report results using the most generous inclusion of data (warn level 5) in this manuscript.

2.6.2 ODIAC Representation Error for the City of Jeddah

ODIAC's representation error of Jeddah's emissions is due to a mismatch between two gas flare nightlight data sources at that particular location. ODIAC uses multiple nighttime light data sources to distribute emissions. Generally speaking, nightlight intensity is used as a proxy for the magnitude of human activities, and correspondingly, CO_2 emissions. However, the pixels associated with gas flares are very bright and thus do not serve as an appropriate proxy for emissions. These bright pixels are thus removed. Gas flares are still included in the ODIAC data product, however; a modified version of NOAA's gas flare nightlight data is used to distribute gas flare emissions country by country. At Jeddah's location, the emissions representation corresponds to a difference between the two gas flare nightlight products. This is a model error in ODIAC and is fixed in the latest version of the product.

2.6.3 Supporting Figures and Table

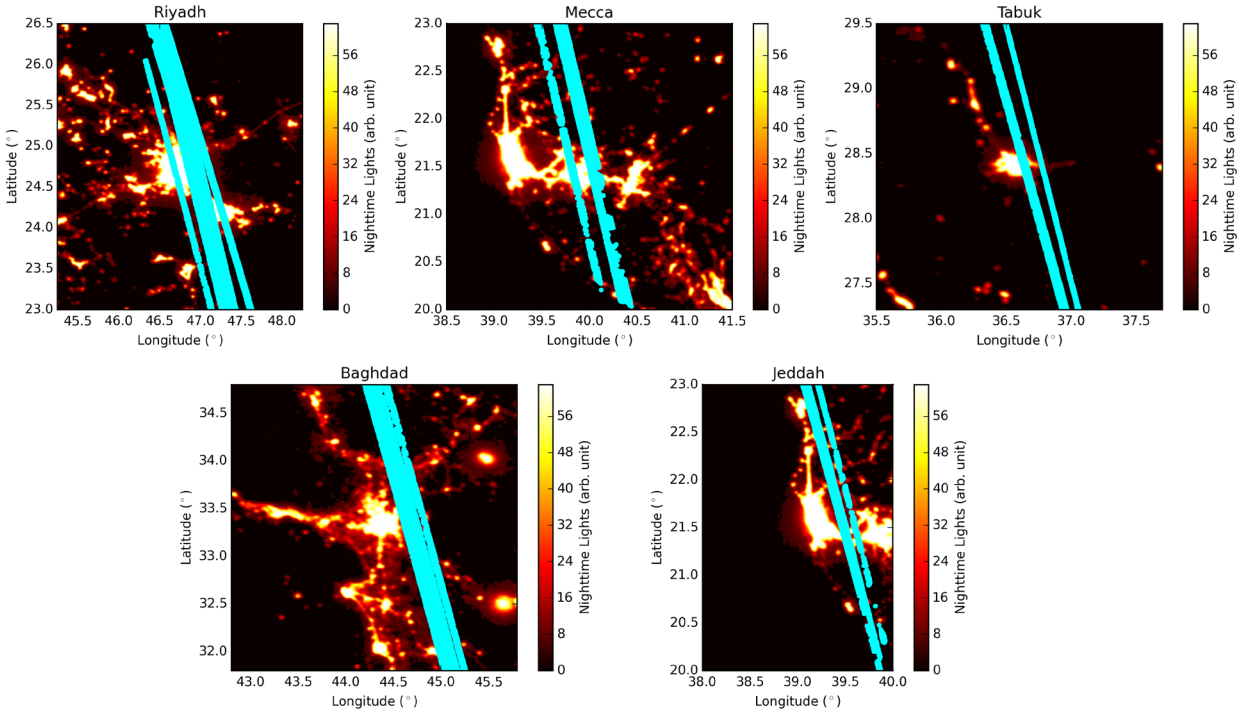

Figure 2.7 All studied overpasses (light blue), overlaid upon the cities of interest, which represented by nighttime lights. The cities of interest are centered within each subplot; in particular, Mecca is the central high-intensity unit of the three shown in its corresponding subplot.

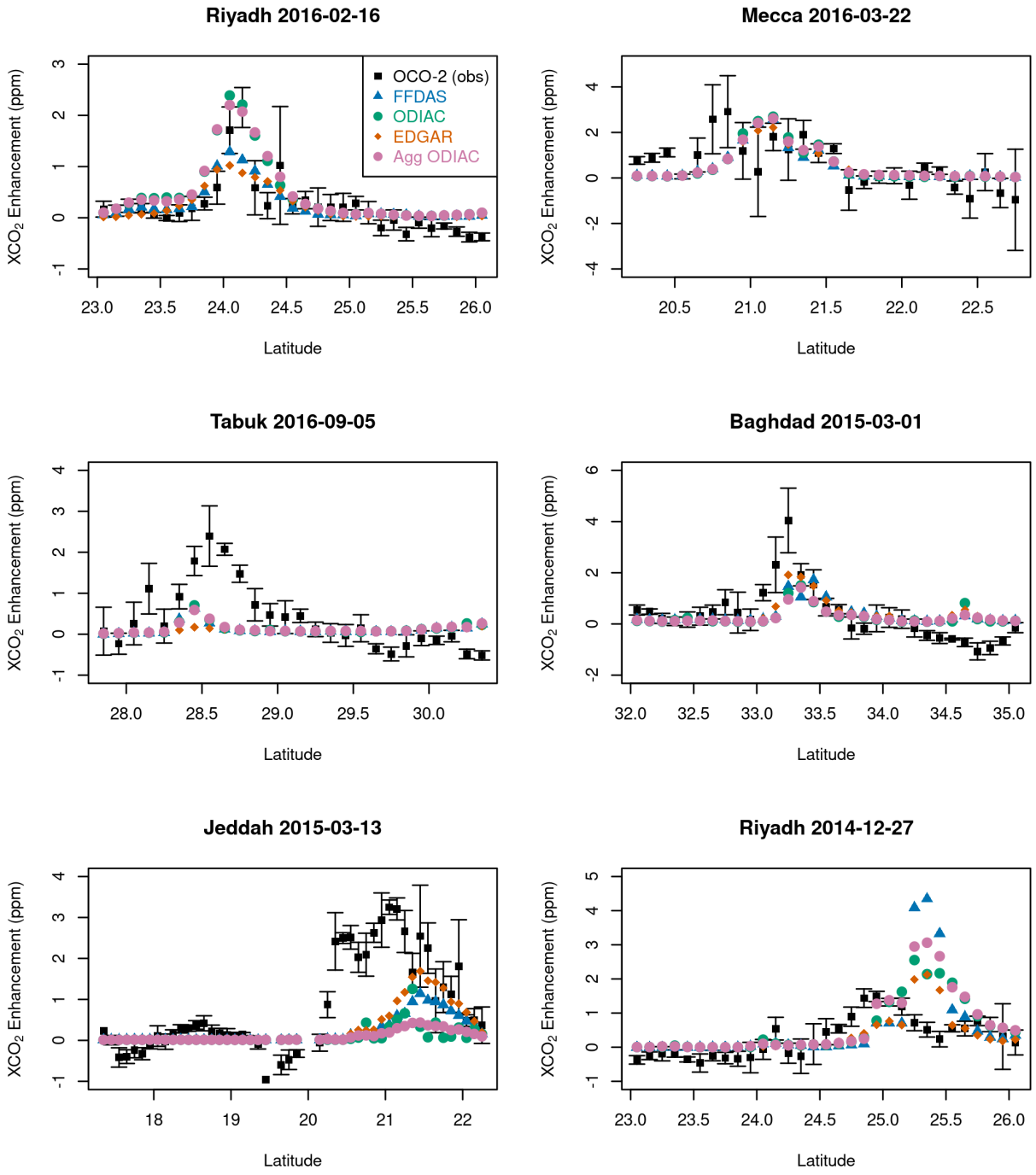

Figure 2.8 Enhancements of sample observed and modeled overpasses for different cities and days, as in Figure 2.4 but with uncertainties on the observations. The uncertainties are derived from the spread of the OCO-2 data, with each uncertainty value representing the standard deviation of the observed points whose medians are used in each plotted observed point.

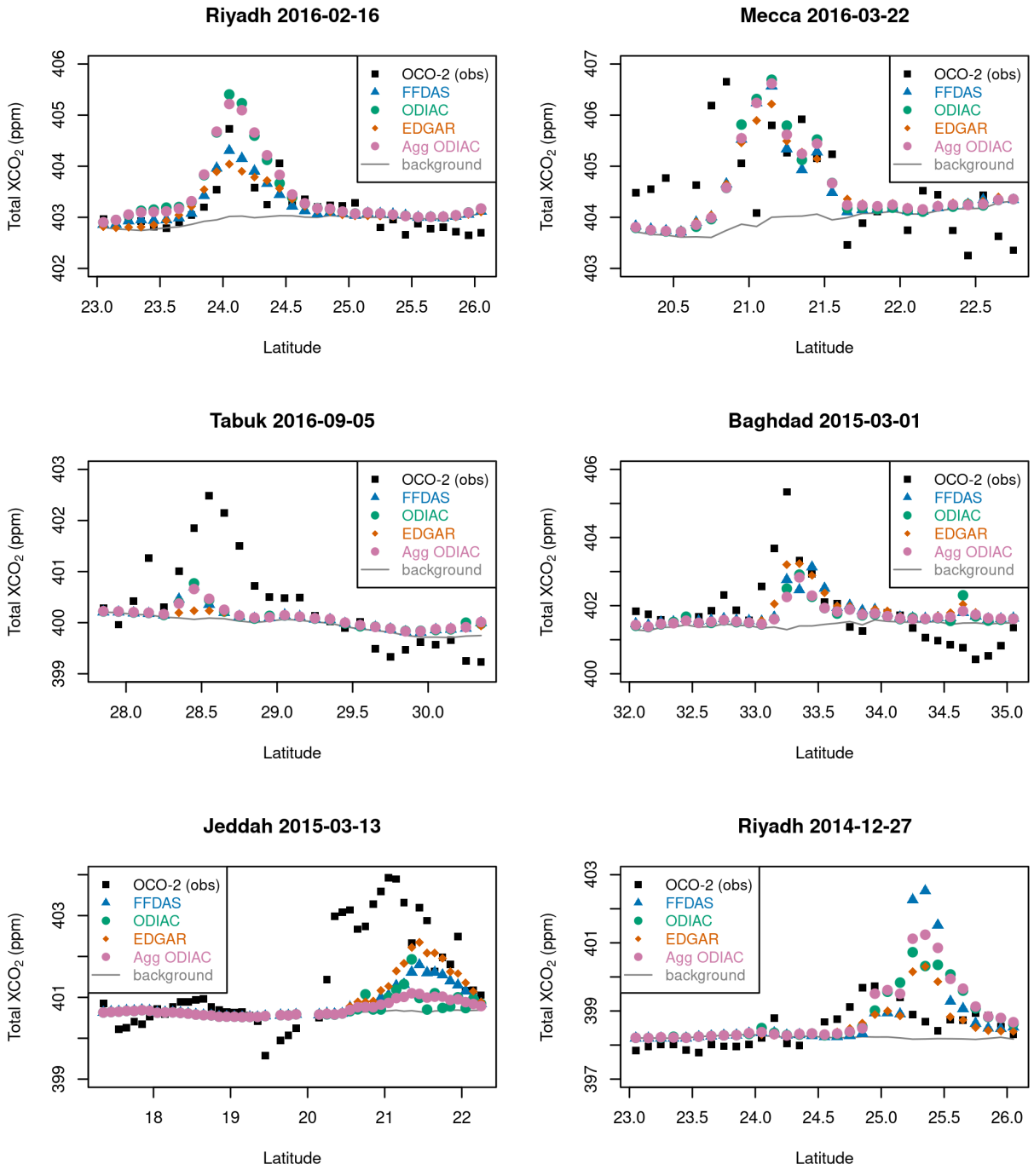

Figure 2.9 Total XCO_2 for sample observed (black) and modeled (other colors) overpasses for different cities and days, matching those of Figure 2.4.

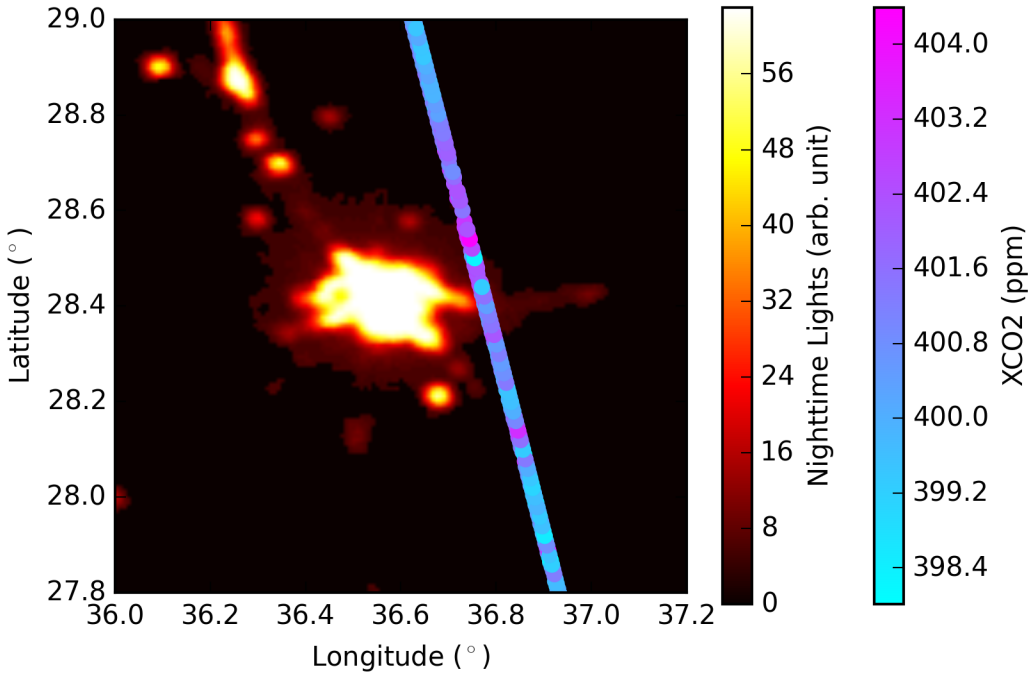

Figure 2.10 OCO-2 observations of XCO_2 on 2016-09-05 overlaid on gridded nighttime lights of the city of Tabuk, Saudi Arabia.

City	*Date*	*Hour (UTC)*	*Observation Mode*
Riyadh	2014-12-27	10	Nadir
Riyadh	2014-12-29	10	Glint
Riyadh	2015-01-28	10	Nadir
Riyadh	2015-03-01	10	Nadir
Riyadh	2015-08-17	10	Glint
Riyadh	2015-09-09	10	Nadir
Riyadh	2015-12-16	10	Glint
Riyadh	2016-02-16	10	Nadir
Riyadh	2016-05-22	10	Nadir
Riyadh	2016-07-25	10	Nadir
Riyadh	2016-10-29	10	Nadir
Mecca	2014-09-25	11	Nadir
Mecca	2014-11-12	11	Nadir
Mecca	2015-11-15	11	Nadir
Mecca	2016-02-19	11	Nadir
Mecca	2016-03-22	11	Nadir
Mecca	2017-04-10	11	Nadir
Tabuk	2015-05-23	11	Glint
Tabuk	2016-09-05	11	Nadir
Tabuk	2017-04-17	11	Nadir
Baghdad	2015-03-01	10	Nadir
Baghdad	2015-05-20	10	Nadir
Baghdad	2015-11-12	10	Nadir
Jeddah	2015-03-13	11	Glint
Jeddah	2015-04-14	11	Glint
Jeddah	2016-02-28	11	Glint

Table 2.5 Information on all overpasses used in this study: city, date, hour (UTC) at which observations were made, and observation mode.

2.7 Acknowledgments, Samples, and Data

Funding for this study was provided by the National Aeronautics and Space Administration (NASA) grants (NNX15AI42G, NNX14AM76G, NNX15AI40G & 80NSSC19K0092) and the National Science Foundation Graduate Research Fellowship Grant No. DGE 1256260. The authors thank Dr. Genevieve Plant for assistance in error calculations, and the OCO-2 science team for general feedback. The OCO-2 data are hosted at the OCO-2 data archive maintained at the NASA Goddard Earth Science Data and Information Services Center, and can be found at: https://disc.gsfc.nasa.gov/datasets?project=OCO. The most recent version of FFDAS (v2.2) data can be found at: http://ffdas.rc.nau.edu/Data.html. The ODIAC

emission data product is archived at the Center for Global Environmental Research, National Institute for Environmental Studies (NIES), Japan (http://db.cger.nies.go.jp/dataset/ODIAC/DL_odiac2017.html). EDGAR is provided and archived by the European Commission, Joint Research Centre (JRC)/Netherlands Environmental Assessment Agency (PBL) (http://edgar.jrc.ec.europa.eu/overview.php?v=432_GHG&SECURE=123). Derived data and model products can be found at the University of Michigan Deep Blue archive (https://doi.org/10.7302/rbjy-7e98).

Chapter 3 Using Space-Based CO_2 and NO_2 Observations to Estimate Urban CO_2 Emissions

This chapter is awaiting submission to *Journal of Geophysical Research: Atmospheres*. It was coauthored with Eric A. Kort, Lesley Ott, Tomohiro Oda, and John C. Lin.

Abstract

As the majority of fossil fuel CO_2 emissions derive from cities, the use of novel techniques to leverage available satellite observations of CO_2 and proxy species to constrain urban CO_2 is of great importance. In this study, we seek to empirically determine relationships between satellite observations of CO_2 and the proxy species NO_2, applying these relationships to NO_2 fields to generate NO_2-derived CO_2 fields (NDCFs) from which CO_2 emissions can be estimated. We first establish this method using simulations of CO_2 and NO_2 for the cities of Buenos Aires, Melbourne, and Mexico City, finding that the method is viable throughout the year. For the same three cities, we next calculate empirical relationships (slopes) between co-located observations of NO_2 from the TROPOspheric Monitoring Instrument (TROPOMI) and Snapshot Area Mode (SAM) observations of CO_2 from Orbiting Carbon Observatory-3 (OCO-3). Applying varying combinations of slopes to generate NDCFs, we evaluate methodological uncertainties for each slope application method and use a simple mass balance method to estimate CO_2 emissions from NDCFs. We demonstrate monthly urban CO_2 emissions estimates that are comparable to emissions inventory estimates. We additionally prove the utility of our method by demonstrating how large uncertainties at a grid cell level (equivalent to ~1-3 ppm)

can be reduced substantially when aggregating emissions estimates from NDCFs generated from all NO_2 swaths (about 1-6%). Rather than rely on prior knowledge of emission ratios, our method circumvents such assumptions and provides a valuable observational constraint on urban CO_2 emissions.

3.1 Introduction

As the global carbon dioxide (CO_2) mixing ratio rises, it is becoming increasingly important to assess fossil CO_2 emissions from cities. At present, the majority of fossil CO_2 emissions entering the atmosphere derive from urban areas, resulting mostly from automobile tailpipe emissions, industrial emissions, home heating, and fossil fuel-based power plants, though the lattermost emissions may take place away from the urban center itself (International Energy Agency, 2008; Lin et al., 2018). As of 2018, 55% of the global population resided in urban areas, and this proportion is projected to increase to 68% by 2050 (United Nations, Department of Economic and Social Affairs, Population Division, 2019). Not only are urban fossil CO_2 emissions expected to play an increasingly important role in the global carbon cycle, but also their quantification is becoming a higher priority as cities pledge to reduce emissions (Bodansky, 2016; Gurney et al., 2015; Rosenzweig et al., 2010) and require a means of assessing progress toward pledged goals (Lauvaux et al., 2020; Mueller et al., 2021). However, there are many gaps in our understanding of urban carbon emissions, along with high uncertainties, due in large part to limited data collected at the urban scale (Gately & Hutyra, 2017; Hutyra et al., 2014). In more recent years, there has been a push to fill this gap using space-based observations of CO_2 to better understand urban emissions (Kort et al., 2012; Hakkarainen et al., 2016; Hedelius et al., 2018; Wu et al., 2018; Park et al., 2020; Wu et al., 2020; Yang et al., 2020; Ye et al., 2020; Kiel et al., 2021; Lei et al., 2021; Kuze et al., 2022).

Observation of gaseous species that are co-emitted with fossil CO_2 during combustion has the potential to provide much-needed information to fill the urban carbon cycle knowledge gaps. These proxy species, such as the nitrogen oxides family (NO_x) and carbon monoxide (CO), share many sources that are co-located in time and space with those of CO_2, and thus can be used to help constrain fossil CO_2 emissions and monitor temporal changes in emissions (Berezin et al., 2013; Konovalov et al., 2016). In particular, observations of nitrogen dioxide (NO_2), one of the gases in the NO_x family, have the potential to greatly increase our understanding of urban carbon emissions for a variety of reasons. Using NO_2 as a proxy for fossil CO_2 allows for the quantification of anthropogenic emissions and disentanglement from biogenic sources of CO_2 (Ciais et al., 2014; Goldberg, Lu, Oda, et al., 2019). Additionally, the relatively short lifetime of NO_2 means that observed NO_2 concentrations are less affected by long-range transport than CO_2 or longer-lived proxies like CO, and lend themselves to more robust interpretation in a local setting (Berezin et al., 2013; Goldberg, Lu, Oda, et al., 2019). NO_2 observations are also less sensitive to clouds than those of CO_2, and quantification of NO_2 emissions is less affected by background variability (Kuhlmann et al., 2019).

Space-based NO_2 observations in particular have the potential to improve the quantification of CO_2 emissions. Satellite observations of NO_2 have been used in combination with NO_x:CO_2 inventory ratios to indirectly estimate CO_2 emissions. For example, Berezin et al. (2013) used NO_2 observations from the Global Ozone Monitoring Experiment (GOME) and the SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY (SCIAMACHY) space-based instruments to infer NO_x emissions, then used NO_x:CO_2 emission conversion factors defined by the Emission Database for Global Atmospheric Research (EDGAR) and the Regional Emission Inventory in Asia (REAS) to estimate multiannual trends of CO_2 emissions in China.

Konovalov et al. (2016) also made regional estimates of CO_2 emissions in Western Europe using satellite observations of both proxy species NO_2 and CO from the Ozone Monitoring Instrument (OMI) and the Infrared Atmospheric Sounding Interferometer (IASI) in combination with EDGAR emission ratios, finding that NO_2 observations had the potential to provide much stronger constraints to the total annual fossil CO_2 emissions in the study region than those of CO. More recent studies have used similar methods to constrain CO_2 emissions at smaller scales, with Goldberg et al. (2019) using OMI NO_2 observations and NO_x:CO_2 ratios from the U.S. EPA to infer CO_2 emissions from eight U.S. megacities, while Liu et al. (2020) indirectly estimated CO_2 emissions from eight U.S. power plants by combining OMI NO_2 observations with the NO_x:CO_2 relationship reported by continuous emissions monitoring system (CEMS) programs. There are limitations to studies that rely on prior NO_x:CO_2 ratios, however, as this relationship changes over time.

There have also been studies that have leveraged satellite NO_2 observations to provide additional information to existing satellite CO_2 observations. In this vein, Reuter et al. (2019) used NO_2 observations from the TROPOspheric Monitoring Instrument (TROPOMI) to adjust wind direction and constrain the shape of the CO_2 plume observed by the Orbiting Carbon Observatory-2 (OCO-2) to make CO_2 flux estimates for six case studies of various cities, power plants, and wildfires. Kiel et al. (2021) also related CO_2 observations of Los Angeles from the Orbiting Carbon Observatory-3 (OCO-3) instrument to TROPOMI NO_2 observations to ensure that the observed CO_2 distributions showed true signals derived from local emissions. Finally, there has been a study in which the authors derived a mean NO_x:CO_2 ratio from co-located TROPOMI and OCO-2 observations of the Matimba power plant in South Africa, then applied the ratio to annual NO_x emissions derived from TROPOMI to make CO_2 flux estimates that were

consistent with emissions inventory estimates from the Open-source Data Inventory for Anthropogenic CO_2 (ODIAC) (Hakkarainen et al., 2021).

In this study, we demonstrate a method for estimating urban CO_2 emissions using NO_2-derived CO_2 fields (NDCFs) empirically derived from space-based observations of urban CO_2 and NO_2, and we assess this method using simulated and observational data for the cities of Buenos Aires, Melbourne, and Mexico City. Rather than rely on prior knowledge of NO_x:CO_2 emission ratios, as did Berezin et al. (2013), Konovalov et al. (2016), Goldberg et al. (2019), and Liu et al. (2020), our method circumvents such assumptions and depends instead on finding empirical relationships between observations of column-averaged dry-air mole fraction of CO_2 (XCO_2) from OCO-3 and tropospheric NO_2 observations from TROPOMI. These relationships can then be applied to TROPOMI NO_2 data to generate NDCFs for time periods that have no co-located CO_2 observations. Inverse methods can then be used on the NO_2-derived CO_2 fields to estimate CO_2 emissions. We evaluate the feasibility of this method using co-located Goddard Earth Observing System Composition Forecast (GEOS-CF) simulations of XCO_2 and NO_2, and we demonstrate the method with observations from OCO-3 and TROPOMI. We further assess mass balance estimates of CO_2 emissions based on the NDCFs, and discuss uncertainties associated with this method. Given the relative abundance and resolution of space-based NO_2 observations compared to those of CO_2, we explore the potential our method has on the constraint of urban CO_2 emissions.

3.2 Data and Methods

The goal of this study is to demonstrate a method of generating urban NO_2-derived CO_2 fields (NDCFs) from empirical relationships between observations of CO_2 and NO_2, and use these NDCFs to estimate emissions of CO_2. We first establish this method using simulations of

these gases, as our pool of available observations is limited. We then find empirical relationships from available co-located satellite observations of CO_2 and NO_2 to explore the utility of the method and emissions estimates based on the method. Our areas of focus are the following three cities: Buenos Aires, Argentina; Melbourne, Australia; and Mexico City, Mexico.

3.2.1 Simulation Data and Methods

3.2.1.1 Simulation Data and Pre-processing

We use hourly simulations of CO_2 and NO_2 from the GEOS-CF system focused on our cities of interest (Keller et al., 2021). The GEOS-CF system is a 0.25° global constituent prediction system that combines the GEOS weather and aerosol modeling system with the GEOS-Chem chemistry module, simulating various atmospheric constituents such as CO_2 and NO_2 in the past, as well as five-day forecasts of said constituents. The NO_2 data are publically available, while the CO_2 simulations are not yet publically available.

For this study, we use concurrent and co-located hourly CO_2 (dry air column-averaged mixing ratio) and tropospheric column NO_2 (tropospheric vertical column density) simulations for our three cities of interest. We use data from each season (referred to as their northern hemisphere season in this sentence only), with one week's worth of data for spring (2019-04-14 to 2019-04-20), summer (2019-07-14 to 2019-07-20), and fall (2019-10-13 to 2019-10-19), and four weeks' worth of data for winter (2020-01-12 to 2020-02-08). Note that there are a few gaps in the northern hemisphere winter CO_2 data, with six hourly simulations missing from the first week and one missing from each of the following weeks. Figure 3.1a and b show an example of a set of co-located CO_2 and NO_2 simulations for Buenos Aires in April 2019.

We choose separate domains of interest for each city of interest, focusing on the area emitting CO_2 with the highest intensity. These domains are chosen by comparing the spatial

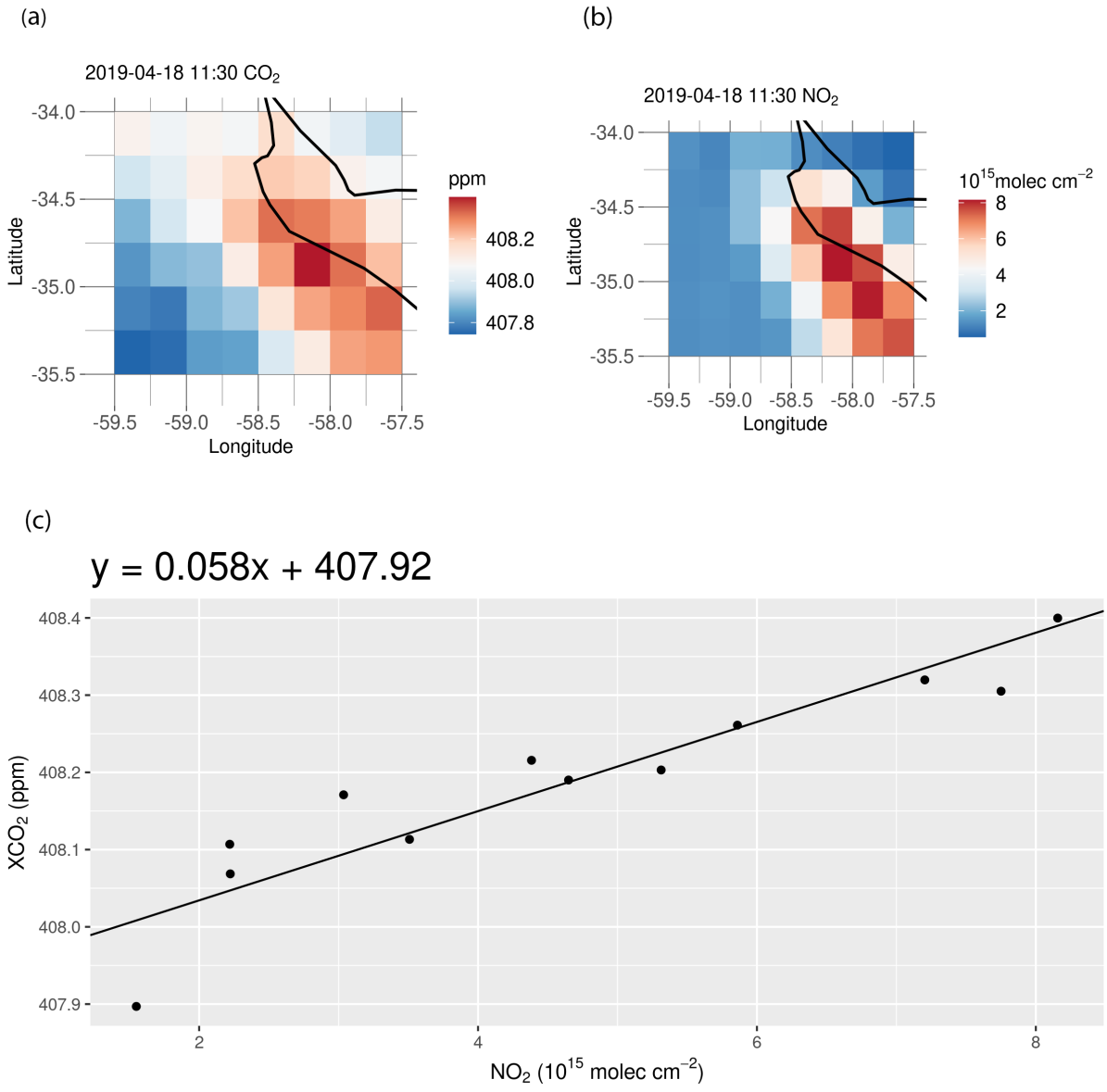

Figure 3.1 An example of a set of hourly, co-located GEOS-CF simulations of (a) CO_2 and (b) NO_2, and (c) the relationship between the two, in April 2019 in Buenos Aires.

boundaries of the cities shown on Google Earth with the areas of highest carbon emission represented by the Open-Data Inventory for Anthropogenic Carbon dioxide (ODIAC) (Oda et al., 2018). The domain for Buenos Aires is 1.0° × 0.75° in size, and the domains for Melbourne and Mexico City are both 0.75° × 0.75° squares.

Once domains are chosen, we filter the simulations. We only use hours during which daylight is present for each location. Additionally, we filter out co-located sets of simulations with the highest 95th percentile of NO_2 values greater than 1.5×10^{16} molecules cm^{-2}, as we have observed nonlinearity in the CO_2-NO_2 relationship when such high values are present. We then calculate the CO_2-NO_2 slopes and intercepts for daytime simulations (see Figure 3.1c) and filter for correlation (at least 0.1), R^2 (at least 0.1), and slope (must be positive). The simulations are further screened via visual inspection to ensure clarity of plumes.

3.2.1.2 Simulating NO_2-Derived CO_2 Fields, Evaluating Simulated Uncertainties, and Assessing Sensitivity to Seasonal Changes

To generate NDCFs, we multiply single slopes with an NO_2 field from an hour different from which the slopes are derived. The field generated in such a manner represents an NO_2-derived CO_2 enhancement field. When the intercept associated with the slope is added to the enhancements, the resulting field is an NO_2-derived CO_2 concentration field. Note that these NDCFs represent total CO_2, with the empirical slopes capturing all impacts from the biosphere or chemistry.

To estimate uncertainties from the simulations, we group simulations by city and season, and we generate concentration fields by applying each calculated slope and intercept in a season to every other NO_2 simulation in that season. Next, we find the relative difference between the individual grid values of the NDCFs and those of the "true" CO_2 fields that temporally correspond to the NO_2 fields from which the NDCFs are derived. The standard deviation of the gridded relative differences serves as a measure of the uncertainty of each simulated NDCF grid cell.

We test the sensitivity of the NDCFs to seasonal changes in biosphere activity and NO_2 chemical lifetime by assessing the seasonality of the uncertainties.

3.2.2 Observation Data and Methods

3.2.2.1 Satellite Data and Pre-processing

For our CO_2 observations (dry air column-averaged mixing ratio), we use bias-corrected OCO-3 Level 2 Lite, version 10.4r data in Snapshot Area Map (SAM) mode (Eldering et al., 2019). These SAMs are fine-scale spatial samplings (with footprint size of 4 km^2 or smaller) of large contiguous areas that generally target urban centers. We filter the SAMs so that they have at least 1000 soundings each, with reasonable data density near the city center of interest, and all used data points are screened by quality flag. Additionally, if a CO_2 field demonstrates enhancements on the edge of a SAM far from and unrelated to the urban center, the enhanced edge is masked out.

For our NO_2 observations, we use Level 2 version 1 observations of tropospheric NO_2 tropospheric vertical column density (with resolution of 5.5 km × 3.5 km) from the TROPOMI instrument onboard the Sentinel-5P satellite (Veefkind et al., 2012). We define the "source area domain" as the same domains used in the GEOS-CF simulations (Section 3.2.1.1). We define the "surrounding domain" as a larger area surrounding the source area domain, generally an additional 0.5° box around. The TROPOMI swaths are filtered similarly to those used by Huber et al. (2020). Only pixels with "flag_value" greater than or equal to 0.75 are used. Next, we only use swaths with at least 30 pixels in the source area domain and 30 pixels in the surrounding domain. Finally, the number of passing pixels must be at least 25% the total number of pre-filtered pixels. As with the GEOS-CF simulations, we also filter out swaths whose highest 95th percentile of NO_2 values are greater than 1.5×10^{16} molecules cm^{-2}.

As an additional filter, we plot co-located SAMs and TROPOMI swaths and visually inspect them to ensure plume clarity. We use available SAMs and TROPOMI swaths from the end of 2019 through mid-2021.

3.2.2.2 Defining Empirical Relationships between CO_2 and NO_2 Observations and Generating NDCFs

As with the GEOS-CF simulations, we calculate slopes and intercepts between co-located and semi-concurrent observations. Figure 3.2 shows an example of this process. Because the satellites pass over the cities of interest at different times, we must account for the time difference. We spatially shift the NO_2 field by increments of 0.01° in each direction up to 0.3° and calculate the correlation after each shift between the CO_2 values, which are averaged to fit the corresponding TROPOMI grid, and the NO_2 values, which are spatially masked to match the outline of the SAM. If the highest correlation found is less than 0.2, that co-located set of observations is filtered out. Otherwise, the shift corresponding to the highest correlation is then applied to the NO_2 field, and the slope and intercept are calculated. Additionally, co-located sets can be filtered out through visual inspection if there are no clear plumes.

Once the empirical relationships have been found between the co-located CO_2 and NO_2 fields, NDCFs can be generated. We generate NDCFs from NO_2 swaths that have corresponding SAMs (e.g., Figure 3.2e), and we also generate NDCFs from the large quantity of NO_2 swaths that do not have corresponding SAMs. For when we have co-located observations, we use a circular application of slopes, i.e., we apply the same-day slopes and intercepts to the NO_2 field to make direct comparisons with the corresponding SAM. For the non-co-located NDCFs, we apply slopes and intercepts with the following four different methods. 1) Single Slope: one slope at a time is applied to all NO_2 fields temporally following a set of co-located observations. 2)

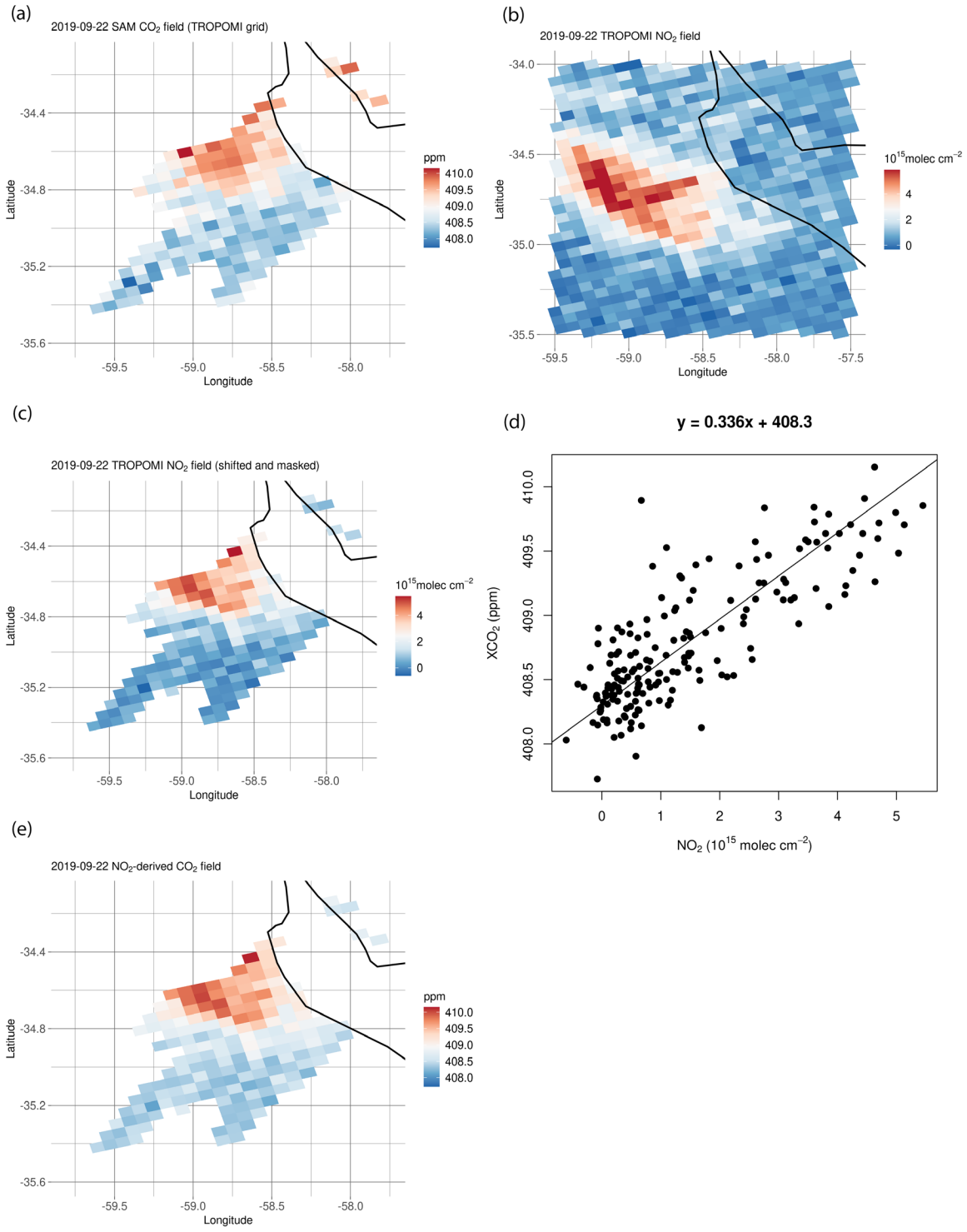

Figure 3.2 (a) An example of a SAM in Buenos Aires. The values have been averaged to match the corresponding TROPOMI grid. (b) The corresponding TROPOMI NO_2 field from the same day, observed about one hour after the SAM. (c) The same NO_2 field shifted spatially to have the highest correlation with the CO_2 data and masked to match the SAM. (d) Scatterplot of CO_2 and NO_2 data from plots (a) and (c) from which the empirical relationship can be found. (e) NDCF generated using relationship from (d) applied to the field in (c).

Average Slope: all available slopes are averaged together and applied to every NO_2 field. 3) Single Slope with Baseline: similar to the single slope method, but each slope is also averaged with the average slope value, to capture temporal variations but dampen noise. 4) Rolling Averages: averages of multiple slopes are applied to subsequent NO_2 fields. For Buenos Aires and Mexico City, we use rolling averages of two, three, four, and five slopes. For Melbourne, we only use a rolling average of two slopes, as the number of Melbourne SAMs that pass our filters is only three.

3.2.2.3 Estimating CO_2 Emissions with a Simple Mass Balance Method

To estimate CO_2 emissions, we use a simple mass balance method similar to that described by Buchwitz et al. (2017). In their study, Buchwitz et al. demonstrated a fast method of emissions estimation for time-averaged satellite observations of methane. We apply their method to SAMs and NDCFs with a few changes to account for the different application. Note that we use this emissions estimate method to be able to fairly quickly generate a large quantity of order-of-magnitude emissions estimates with some temporal variations rather than precise emissions estimates; this mass balance method is not a part of the NDCF method we are describing in this study.

The mass balance method is based on a box model, in which the "box" described is the CO_2 source region, e.g., the domain of the city of interest. The emissions E associated with the source region are given as

$$E = CF \times \Delta C \times V/L, \quad (3.1)$$

where CF is a correction factor, ΔC is the average enhancement in the source region as a vertical column density, V is the wind speed, and L is the length of the box. Regarding CF and V, Buchwitz et al. used $CF = 2$ and a constant V of 1.1 m s^{-1}, which were derived from comparing

their estimates to a prior. As we are not using time-averaged observations, we use 10 m wind speeds from the ERA5 reanalysis data (Hersbach et al., 2020), averaging the wind speeds over the domain. With wind speeds of approximately 3-6 m s^{-1}, we find from a comparison with inventory data that $CF = 1/3$ is sensible for our study. For ΔC, we first find the average concentration in the source region and subtract off a background value, which we determine to be the average value of the lowest 20th percentile of values within the bounds of the surrounding domain. We then convert this enhancement from units of ppm to kg m^{-2} using a simple conversion factor of 0.01 kg m^{-2} per 1 ppm of enhancement, as described by Kuhlmann et al. (2020). Regarding L, we use an average of the ERA5 wind direction in the source region to determine whether the winds are flowing in a more north-south direction or east-west direction. Depending on this direction, we use the corresponding length of the domain as L.

We use the mass balance method to estimate CO_2 emissions from SAMs, corresponding co-located NDCFs, and NDCFs generated with the four slope application methods described in Section 3.2.2.2.2. We also make comparisons between our emissions estimates and EDGAR monthly sector-specific gridmaps of CO_2 emissions for the year 2015 (Crippa et al., 2020). We combine emissions from the four sectors that produce the most emissions in cities: power industry, manufacturing, road transport, and buildings.

3.2.2.4 Evaluating Methodological Uncertainties

Similar to how we determine simulation uncertainties in Section 3.2.1.2, we use the spread of the observation-derived data to estimate methodological uncertainties that capture the uncertainties associated with the slope calculations and with the assumptions behind the method itself. Grouping the co-located data by city, we generate separate sets of NDCFs by applying our four slope application methods to every NO_2 field in the group, with the exception of fields

whose slopes were used in the slope calculation. Specifically for the Average Slope method, we find an average of every slope except one and apply the average to that corresponding NO_2 field, and rotate through all available co-located datasets. Once the separate sets of NDCFs are generated, we again calculate the standard deviation of relative differences of individual grid cells between SAMs and the corresponding NDCFs generated with the different methods. These standard deviations serve as a measure of the methodological uncertainty that can be propagated through to NDCF emissions estimates. The methodological uncertainty is derived relative to concentration, as opposed to enhancement, in order to ensure that there is some uncertainty regardless of the level of enhancement.

Note that we do not include an evaluation of uncertainties from the mass balance method. This study assesses the utility of the NDCF method and not the mass balance method, and uncertainties associated with the mass balance method would be large enough to mask the effect of the NDCFs on their own.

3.3 Results

3.3.1 Establishment of the NDCF Method with GEOS-CF Simulations

Generating simulation-based NDCFs with reasonable levels of error and low bias in comparison to the "true" CO_2 simulations suggests that the NDCF method can be reasonably applied to observations. Following the methods described in Section 3.2.1.2, we generate simulation-based NDCFs with application of a single slope for each city and season, find the relative differences between the NDCFs and their "true" counterparts, and use the spread of the distribution of these relative differences at grid cell-level as a measure of the simulation-based uncertainty. An example distribution of the relative differences for Mexico City (July) is found in Figure 3.3. For this example, the standard deviation of the relative differences is 0.24% of the

total concentration of a grid cell, which is equivalent to 0.97 ppm if we assume a concentration of 410 ppm. There is low bias associated with this example because the Gaussian distribution corresponding to the standard deviation of the data and the Gaussian distribution fit to the data are almost the same. This is the same for all other seasons and cities as well.

Figure 3.4 shows the spread of the relative differences for all cities and seasons. The simulation-based uncertainties are at fairly reasonable levels, ranging between 0.11 and 0.97 ppm if we assume a concentration of 410 ppm. With this figure, we also assess the impact of seasonal effects—biospheric activity and NO_2 chemistry—on our method. The NDCF method involves generating total CO_2, which empirically captures seasonal effects in the slopes used. Accordingly, we would expect to see larger uncertainties associated with certain times of year if

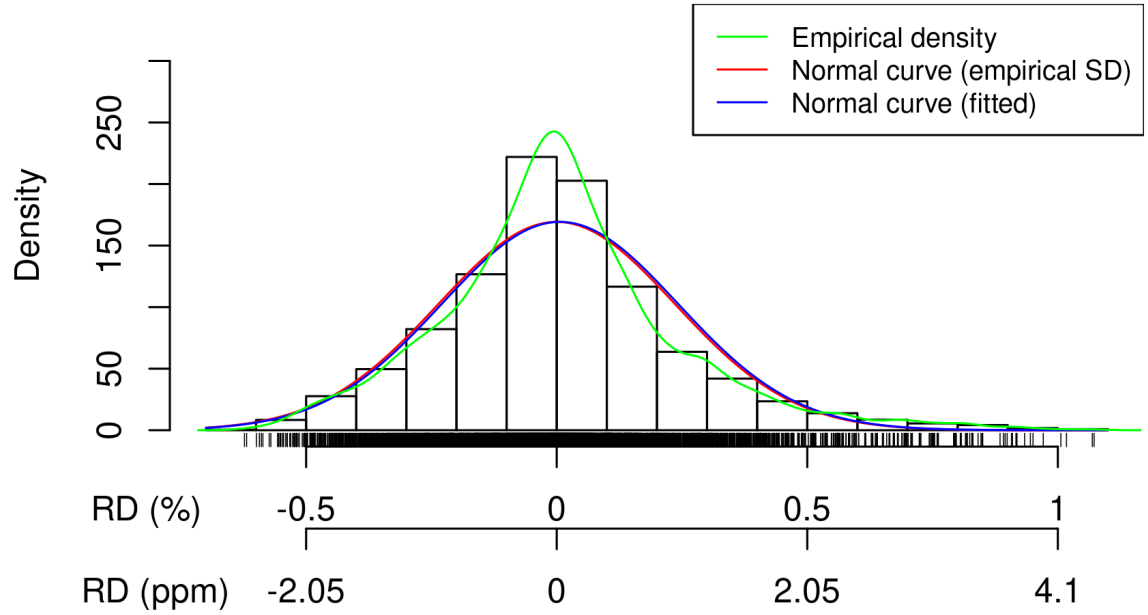

Figure 3.3 An example of the distribution of grid cell-level relative differences (RD) for Mexico City July (summer). The green line shows the empirical density curve, the red line shows the Gaussian curve using the standard deviation of the relative differences, and the blue curve fits a Gaussian to the data. Because the red and blue curves are very similar to each other, there is low bias. The standard deviation of the relative differences for this example is 0.24% of the total concentration of a grid cell, which would be equivalent to 0.97 ppm if we assume a concentration of 410 ppm.

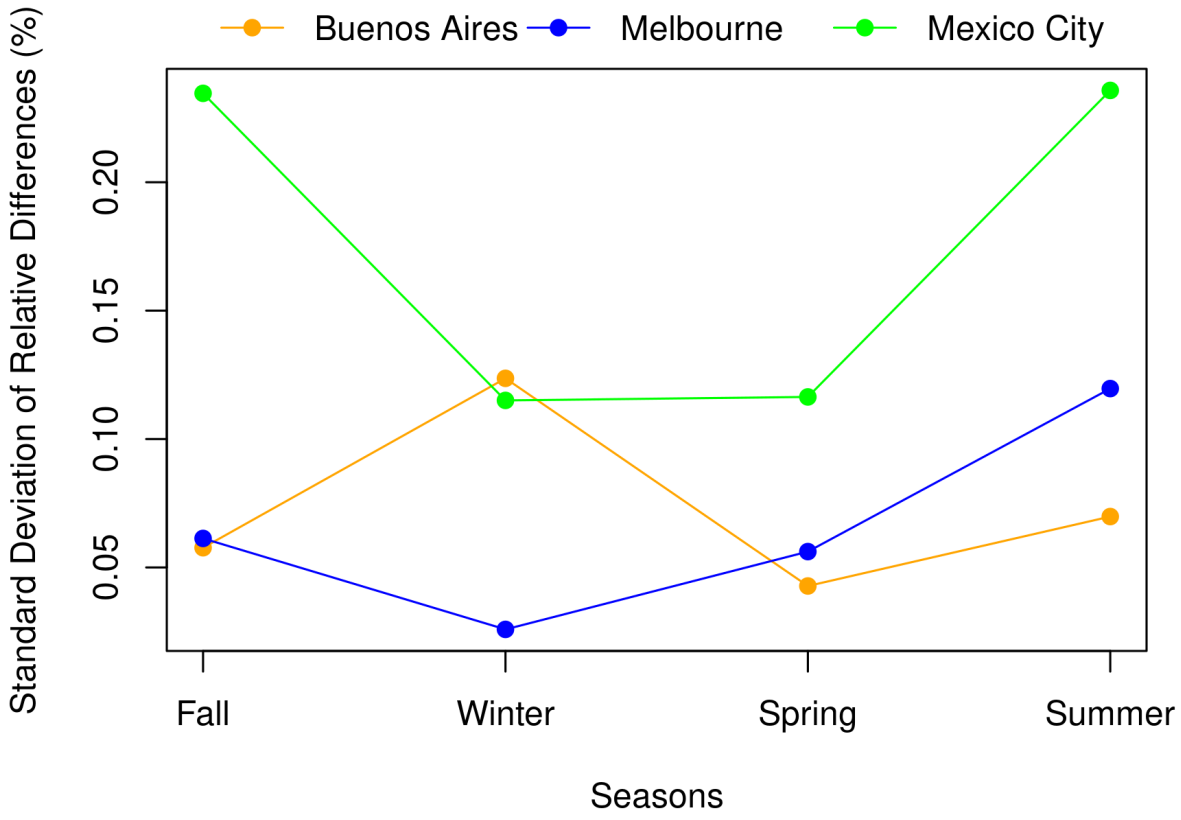

Figure 3.4 A seasonal comparison of the simulation uncertainties. Note that these are for the time period of April 2019 to February 2020, and that Mexico City's values have been reordered to follow the seasonal pattern of the other cities as opposed to their temporal order. This plot shows no clear seasonal dependence across the cities, suggesting that changes in biospheric activity and NO_2 chemical lifetime due to season are not primary drivers of variations in uncertainty.

these seasonal effects were a dominant source of error—a systematic bias. Meanwhile, chemical or biological effects that are not strongly correlated between the gases would just impart additional noise to our uncertainties. While there are variations between months, Figure 3.4 shows no clear seasonal pattern when comparing across the cities. This suggests that, while there is inherent noise driving the variations in uncertainty across seasons, seasonal changes in biospheric activity and NO_2 chemistry are not the primary drivers of those variations. Thus, these seasonal effects do not appear to be a dominating source of error.

3.3.2 Methodological Uncertainties in Observation-based NDCFs

Distributions of the relative differences between observation-based NDCFs and their corresponding SAMs, for each slope application method, are shown in Figures 3.5-3.7, with each figure corresponding to a different city. The standard deviations of these relative differences (listed in Table 3.1) represent the methodological uncertainties corresponding to their respective slope application methods and cities. These uncertainties at a grid cell-level are higher than those associated with the simulations. Assuming 410 ppm as the concentration for a grid cell, these result in uncertainties ranging between about 1.5 and 3 ppm. We expect higher uncertainties due to observational noise.

For Buenos Aires and Mexico City (Figures 3.5 and 3.7), as expected, we see that uncertainties are highest for the Single Slope method (panel (a) in both figures). Averaging with a baseline (c) or averaging all available slopes together (b) reduces uncertainties by dampening noise associated with the use of slopes that may be less representative. For Mexico City (Figure 3.7), we also observe the expected behavior: uncertainties are reduced as more slopes are averaged together (Figure 3.7d-g). For Buenos Aires (Figure 3.5d-g), we see the opposite behavior, though the uncertainties are fairly close to each other for these four cases of rolling averages. It appears that the particular slopes that are available for Buenos Aires generate greater relative differences when rolling averages are taken; we discuss this further regarding the mass balance estimates in Section 3.3.3.2. As for Melbourne (Figure 3.6), which has fewer co-located observations to derive slopes from, there are a few things of note. First, likely due to the limited number of observations, we see higher bias than for the other cities. Second, the Single Slope case (Figure 3.6a) demonstrates higher uncertainties than the cases involving the application of Average Slopes and Single Slope with Baseline (Figures 3.6b and c, respectively). Third, while

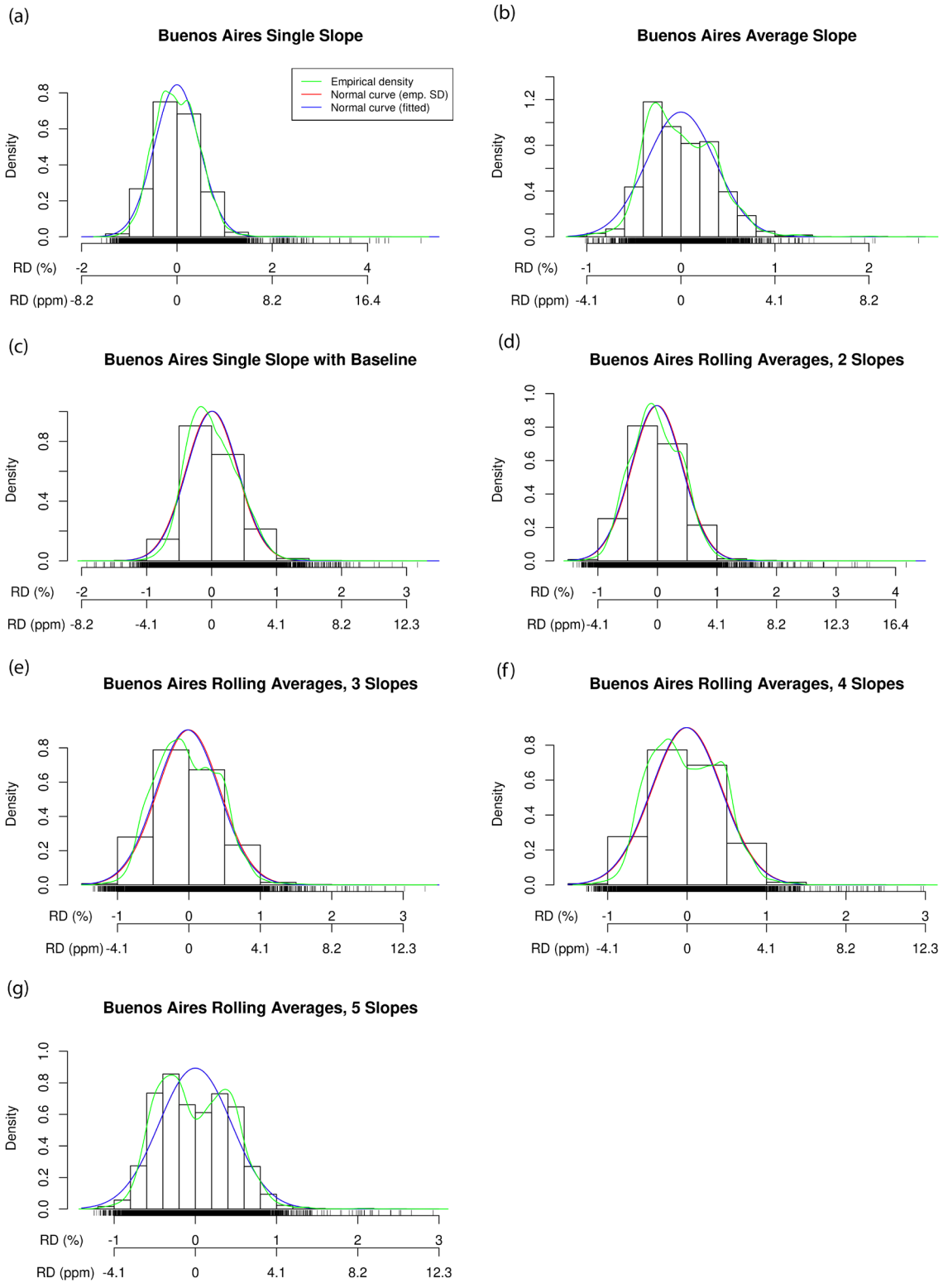

Figure 3.5 Distributions of relative differences, as in Figure 3.3, but for Buenos Aires observation-based NDCFs derived using the different slope application methods: (a) Single Slope, (b) Average Slope, (c) Single Slope with Baseline, (d) Rolling Averages (two slopes), (e) Rolling Averages (three slopes), (f) Rolling Averages (four slopes), and (g) Rolling Averages (five slopes).

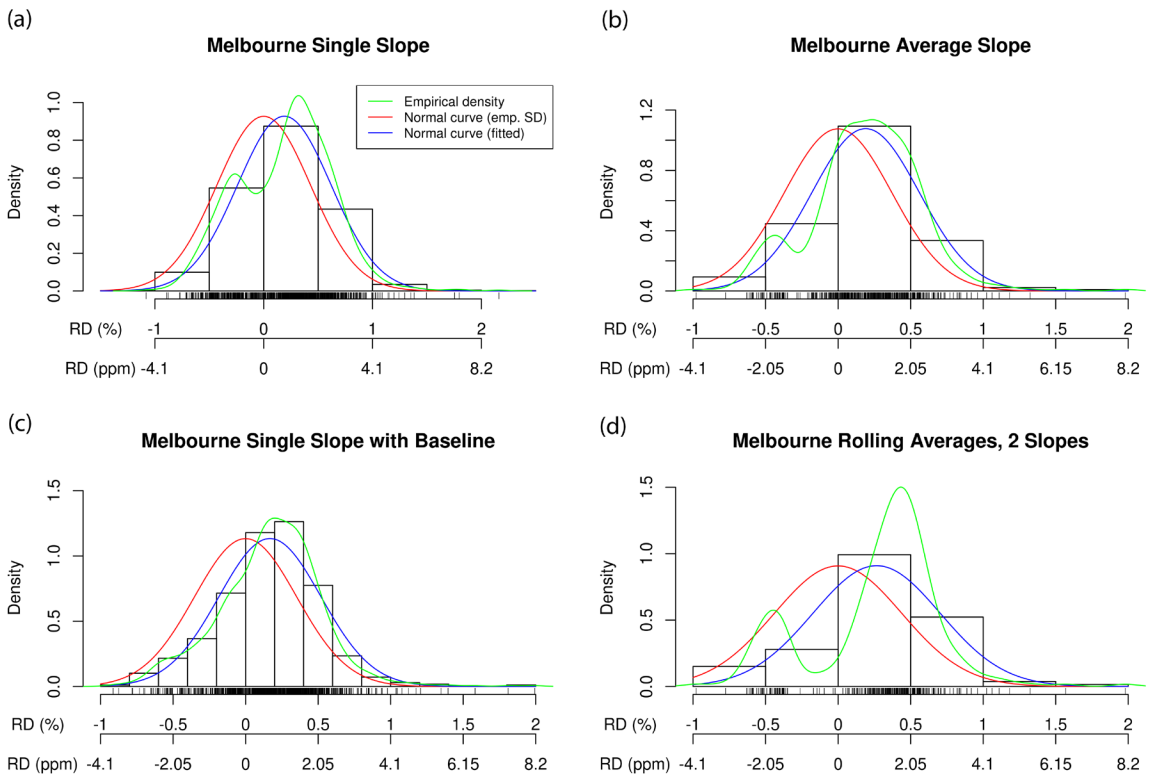

Figure 3.6 Same subplots as Figure 3.5 (a-d), but for Melbourne. Note that the low number of SAM observations limits the Rolling Averages method to two slopes only.

	Buenos Aires	*Melbourne*	*Mexico City*
Single Slope	0.47% (1.94 ppm)	0.43% (1.76 ppm)	0.79% (3.26 ppm)
Average Slopes	0.37% (1.50 ppm)	0.37% (1.52 ppm)	0.59% (2.44 ppm)
Single Slope with Baseline	0.40% (1.63 ppm)	0.35% (1.44 ppm)	0.75% (3.07 ppm)
Rolling Averages (2 slopes)	0.430% (1.76 ppm)	0.44% (1.80 ppm)	0.78% (3.20 ppm)
Rolling Averages (3 slopes)	0.441% (1.81 ppm)	-	0.75% (3.09 ppm)
Rolling Averages (4 slopes)	0.443% (1.82 ppm)	-	0.72% (2.94 ppm)
Rolling Averages (5 slopes)	0.447% (1.83 ppm)	-	0.65% (2.68 ppm)

Table 3.1 Methodological uncertainties associated with each slope application method

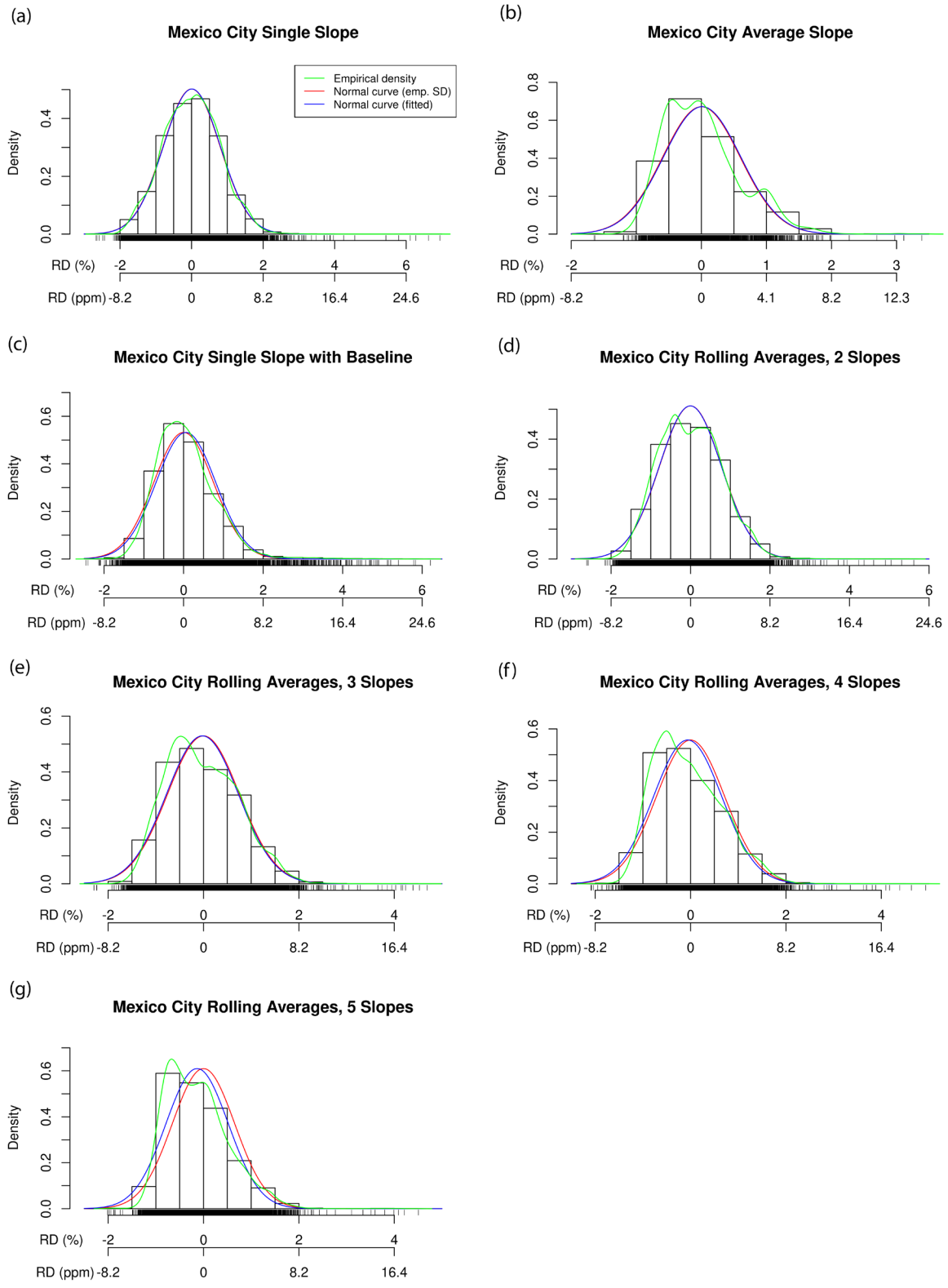

Figure 3.7 Same subplots as Figure 3.5, but for Mexico City.

the Single Slope (Figure 3.6a) and Rolling Averages of two slopes (Figure 3.6d) cases show very similar uncertainties, the latter actually has slightly higher uncertainty. We expect that this is also a peculiarity of the low number of observations used. However, the uncertainties are generally lower for Melbourne than the other cities, as a result of fairly similar slopes used.

3.3.3 Emissions Estimates from Observation-based NDCFs

3.3.3.1 Emissions Estimates from Co-located Observations: SAMs and NDCFs

Figure 3.8 shows CO_2 emissions estimates from the co-located observations for all three cities. In black, we show the mass balance estimates of SAMs, and in purple, we show the estimates of corresponding NDCFs derived from circular slope application. A comparison between SAM and NDCF emissions estimates shows them to be within the same order of magnitude, and many of the dates presented in Buenos Aires are quite similar to each other (Figure 3.8a). However, we do also expect differences due to the methodological uncertainties we have already discussed in Section 3.3.2, as well as uncertainties due to the mass balance method itself. Estimates of NDCF emissions being in the same order of magnitude as SAM emissions is a reasonable demonstration of effectiveness given the large uncertainties expected.

3.3.3.2 Emissions Estimates from NDCFs Generated from Remaining Available TROPOMI NO_2 Fields

Time series of mass balance estimates of NDCFs generated from all available TROPOMI NO_2 swaths, excluding those that are used in the co-located analysis and slope calculation, are shown in Figures 3.9-3.11, with each figure corresponding to a different city. The black dots represent individual mass balance estimates, and the purple dots represent the value of the slope applied to subsequent dates for all slope application methods other than Average Slope. For each

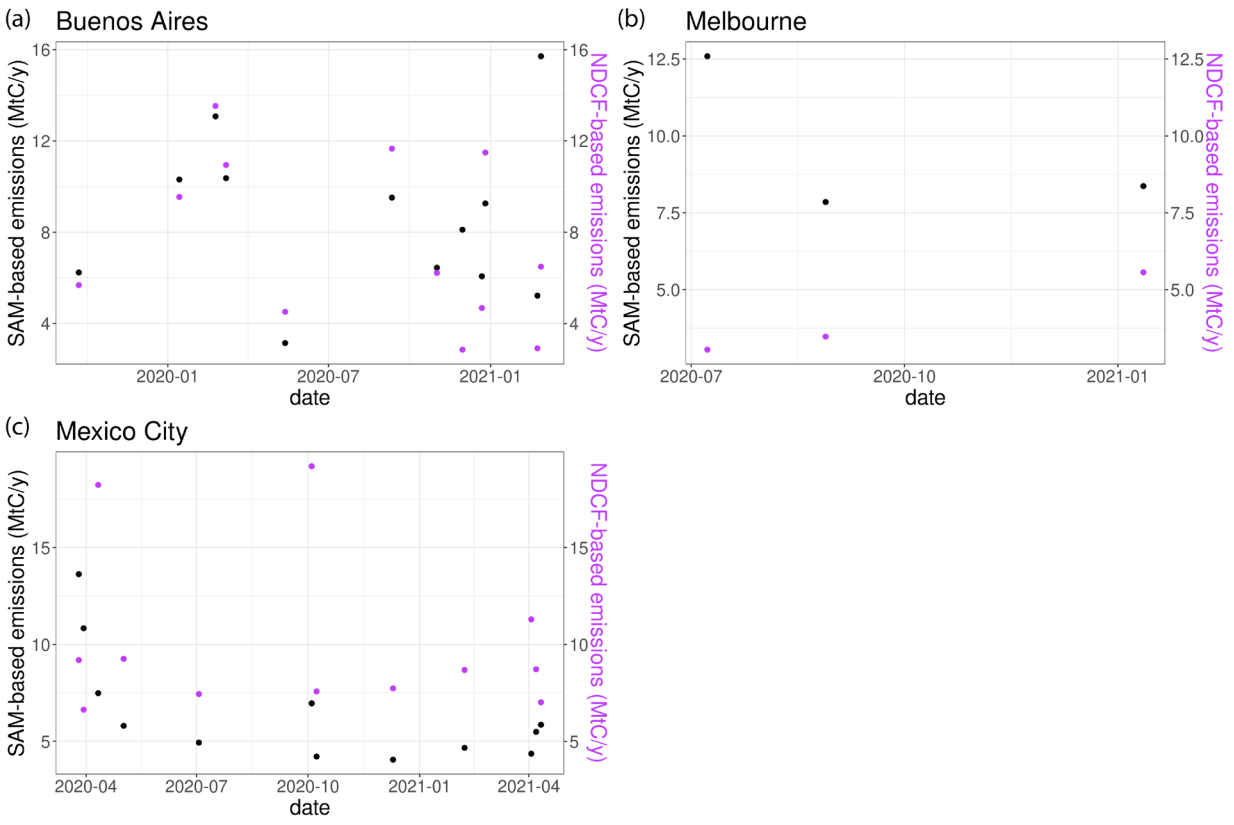

Figure 3.8 Mass balance estimates of CO_2 emissions from co-located observations, with the estimates of SAMs in black and the corresponding NDCFs in purple, for (a) Buenos Aires, (b) Melbourne, and (c) Mexico City.

city, there are variations in the emissions estimates, especially depending on the value of the slopes used. This is particularly apparent in panel (a) of each of Figures 3.9-3.11, which shows the Single Slope method. High or low slopes that are not representative and that are not averaged with any other values result in large periods of time with enhanced or suppressed emissions estimates that do not appear reasonable. This issue is potentially compounded with the Rolling Averages method if multiples of these types of slopes occur consecutively, which possibly explains the increased uncertainties for Buenos Aires with the Rolling Averages method discussed in Section 3.3.2. Despite these variations across methods, we observe that many of the time series suggest a seasonal cycle, with higher emissions in the winter than in the summer. These subannual variations are more common in cases not using the Single Slope method. These other methods, which average more than one slope together, demonstrate more physically

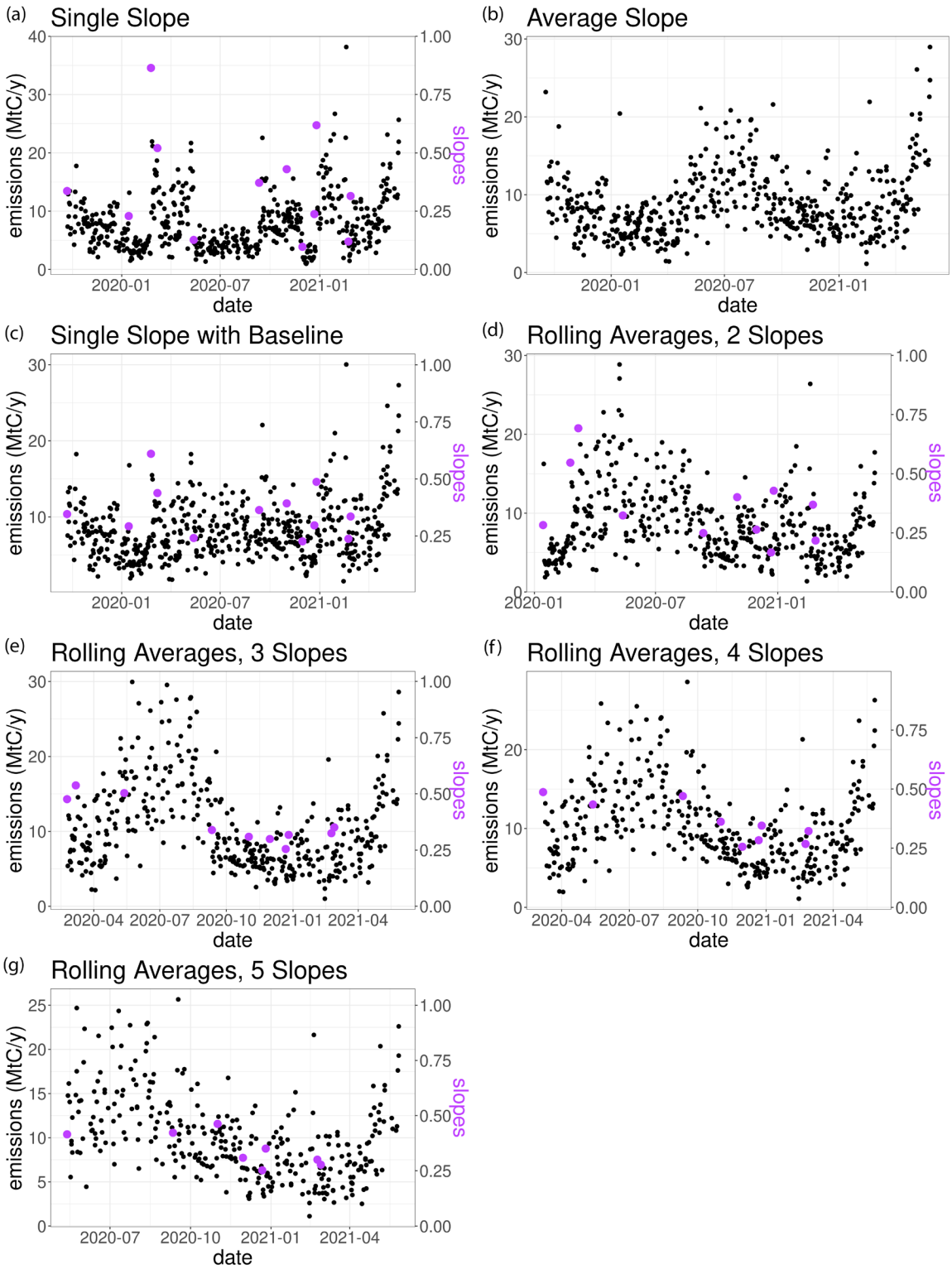

Figure 3.9 Time series of mass balance estimates of Buenos Aires CO_2, calculated from NDCFs derived from the different slope application methods: (a) Single Slope, (b) Average Slope, (c) Single Slope with Baseline, (d) Rolling Averages (two slopes), (e) Rolling Averages (three slopes), (f) Rolling Averages (four slopes), and (g) Rolling Averages (five slopes). For all subplots other than (b), we show the slope used to generate all subsequent NDCFs in purple.

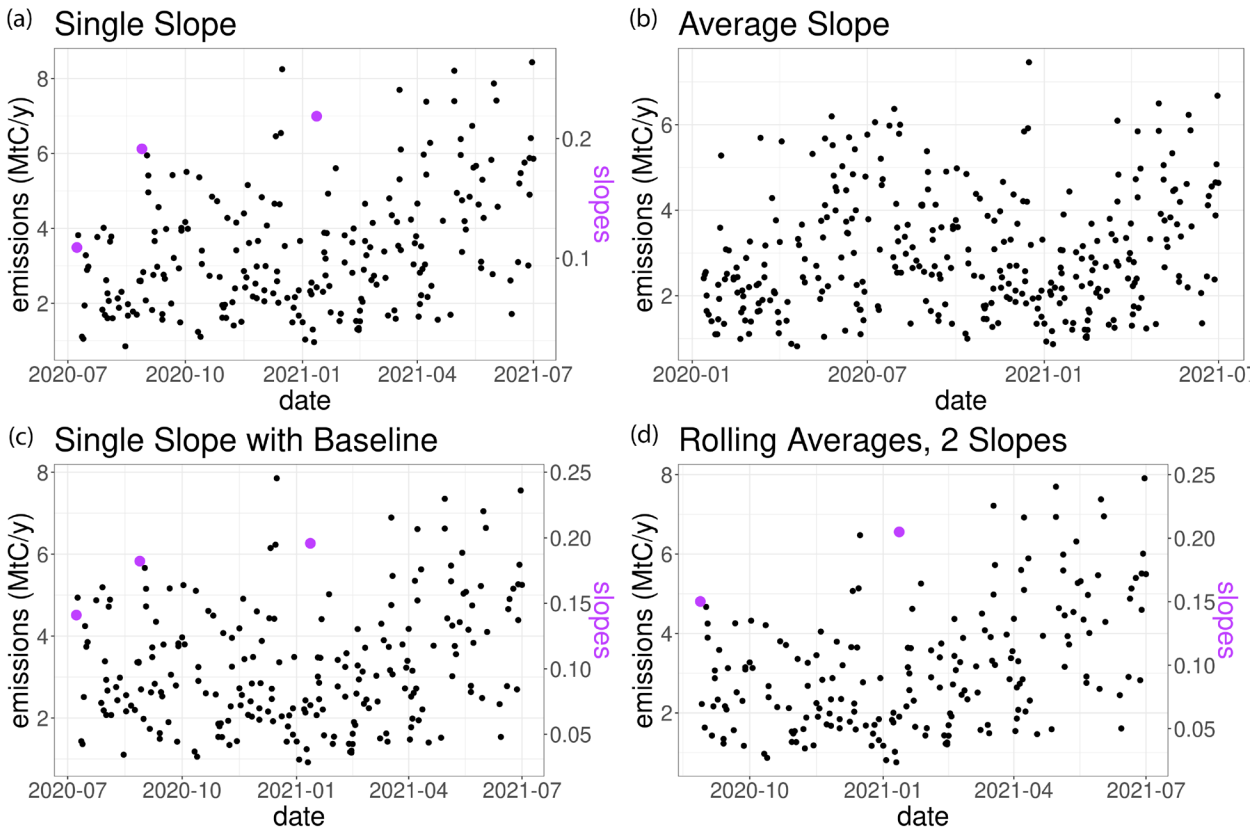

Figure 3.10 Same subplots as Figure 3.9 (a-d), but for Melbourne. Note that the low number of SAM observations limits the Rolling Averages method to two slopes only.

reasonable temporal variation. In particular, the Single Slope with Baseline method captures some of the temporality of the Single Slope method but dampens some of the noise from slopes that may not be representative of the entire period to which those slopes are applied. The Rolling Averages methods show relatively similar results for each city across the differing numbers of slopes used to generate the NDCFs. However, since the Rolling Averages method applies averaged slopes to NO_2 fields following the date of the latest slope used in the average, there is a greater loss of NO_2 data to be used to generate NDCFs with a greater number of slopes averaged.

3.3.3.3 Emissions Estimates Aggregated Over Time

While uncertainties for individual NDCF emissions estimates are large, the value of the NDCF method becomes apparent in aggregate. Figures 3.12-3.14 show monthly averages of the

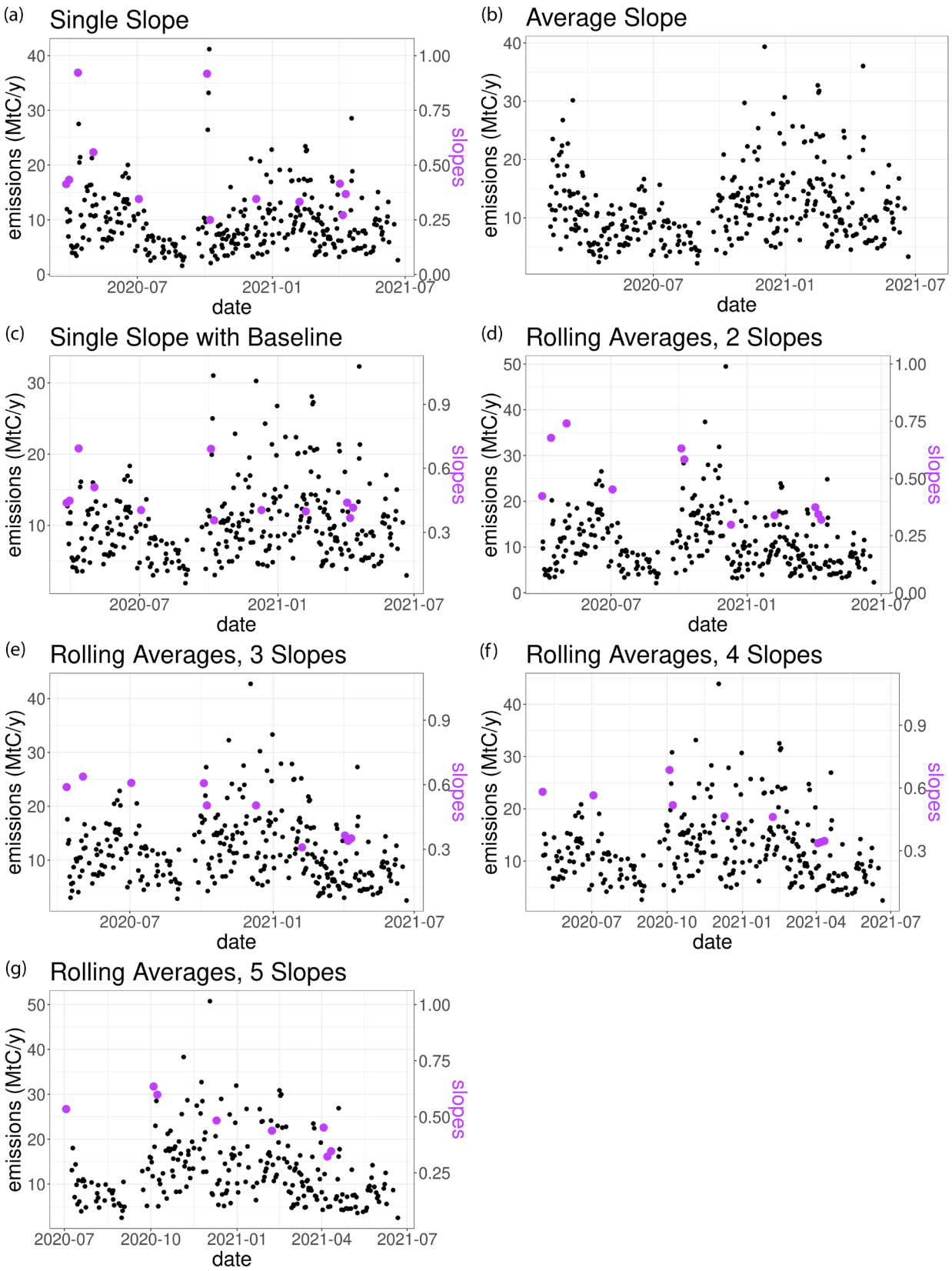

Figure 3.11 Same subplots as Figure 3.9, but for Mexico City.

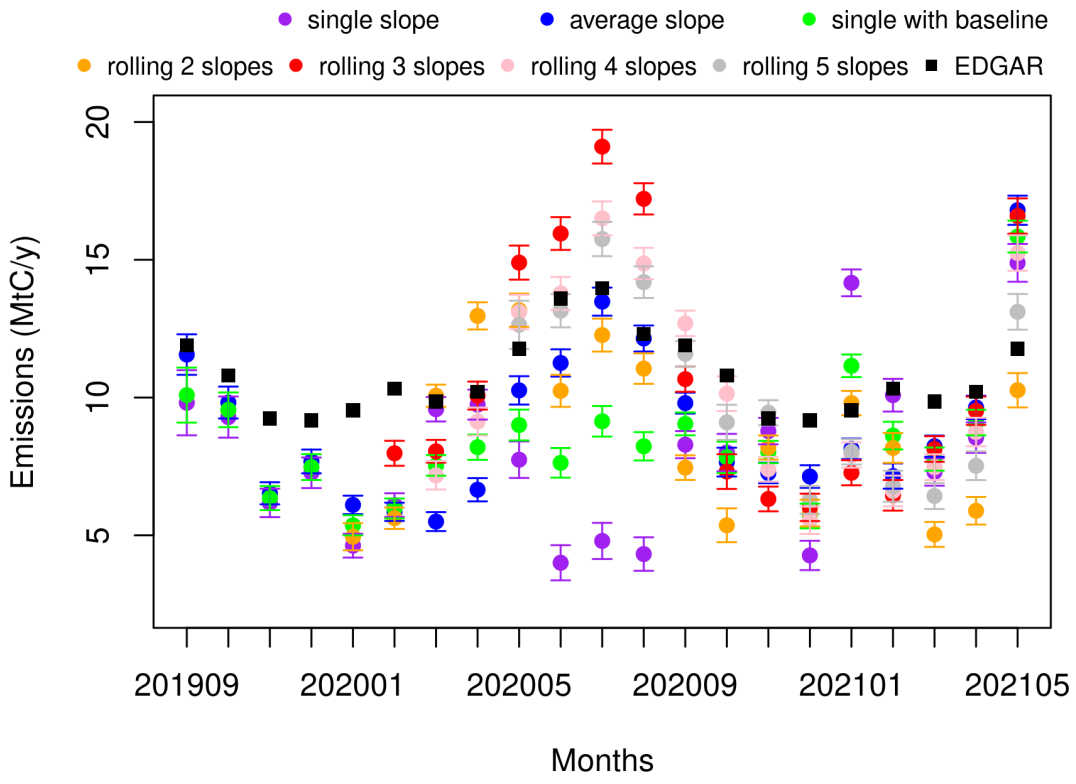

Figure 3.12 Mass balance estimates of Buenos Aires CO_2 averaged by month, with propagated uncertainties, corresponding to the different slope application methods in different colors (see legend). EDGAR monthly data for the year 2015 are shown in black squares, with multiples of the same month showing repeated values. Each of these EDGAR points represents the sum of emissions from the four sectors of interest: power industry, manufacturing, road transport, and buildings.

mass balance estimates, with methodological uncertainties (excluding mass balance method uncertainties). These figures show all of the slope application methods together, along with the 2015 EDGAR monthly and sectoral emissions estimates (with monthly values repeated across multiple years). In these figures, the monthly variations of the emissions estimates are quite apparent. These variations are in line with what is reported by EDGAR. Such monthly-scale estimates and variations have not previously been demonstrated by satellite estimates of CO_2 emissions. With the very large number of NO_2 swaths available for the generation of NDCFs, we can reduce uncertainty substantially in aggregate and capture unprecedented temporality in emissions estimates.

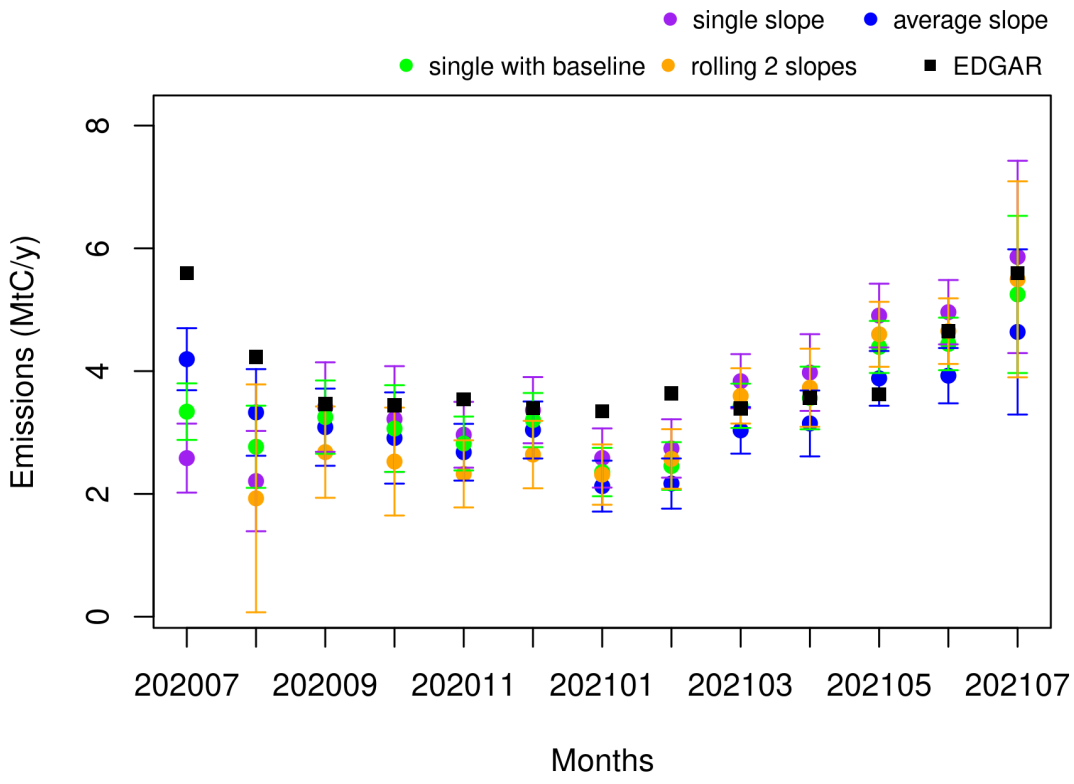

Figure 3.13 Same as Figure 3.12 but for Melbourne. Note that the low number of SAM observations limits the Rolling Averages method to two slopes only.

Along with monthly estimates, we can average every mass balance estimate for a city and slope application method to estimate overall emissions over the time period of interest. Table 3.2 lists these estimates and their corresponding methodological uncertainties (excluding mass balance method uncertainties). Across methods, the emissions estimates are fairly similar to each other, with greater variation across methods with the inclusion of more unrepresentative slopes. Methodological uncertainties at this level of aggregation are quite low (ranging between about 1-6% of the estimates themselves), given the large number of NDCFs used in each of the averages. The much greater availability of TROPOMI swaths as opposed to SAMs demonstrates the utility of the NDCF method.

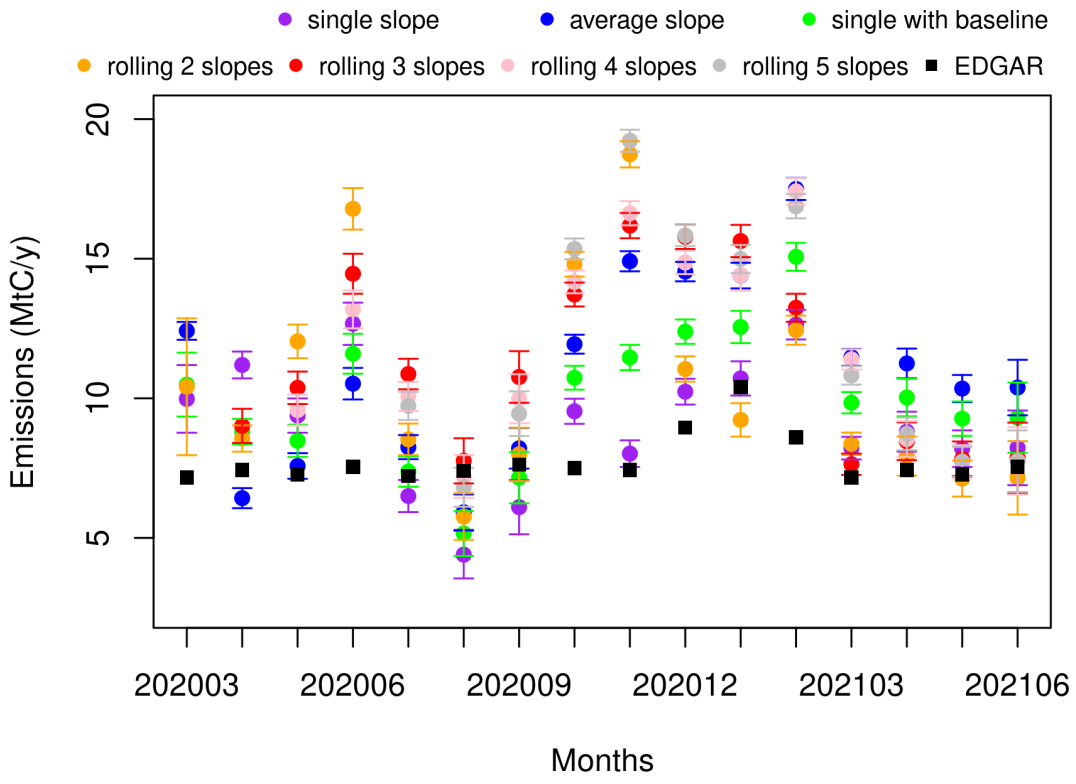

Figure 3.14 Same as Figure 3.12, but for Mexico City.

	Buenos Aires	*Melbourne*	*Mexico City*
Single Slope	7.82 ± 0.12	3.37 ± 0.18	9.33 ± 0.15
Average Slopes	8.63 ± 0.10	2.97 ± 0.12	11.70 ± 0.10
Single Slope with Baseline	8.21 ± 0.10	3.21 ± 0.14	10.35 ± 0.14
Rolling Averages (2 slopes)	8.52 ± 0.12	3.09 ± 0.19	10.98 ± 0.15
Rolling Averages (3 slopes)	10.49 ± 0.13	-	11.67 ± 0.15
Rolling Averages (4 slopes)	10.18 ± 0.14	-	12.31 ± 0.15
Rolling Averages (5 slopes)	9.97 ± 0.15	-	12.83 ± 0.14
EDGAR 2015	11.06	3.83	7.88

Table 3.2 Averaged emissions estimates in MtC y^{-1} across slope application methods and cities. EDGAR annual emissions estimates for the year 2015, for the four sectors addressed in this study, are also presented at the bottom.

3.4 Discussion

Given the much greater availability of TROPOMI NO_2 data than OCO-3 SAM CO_2 data, the application of empirical relationships between CO_2 and NO_2 to the TROPOMI fields to generate NDCFs opens up many possibilities for assessing urban CO_2 emissions subannually and over time. Our study demonstrates monthly assessments of urban CO_2 emissions that have previously not been done before with satellite data.

Some CO_2-only studies have previously opted for wintertime data only to avoid influence from the biosphere (e.g., Wu et al., 2018), while NO_2 observation-based studies of CO_2 often used observations from only warm months due to the shorter NO_2 lifetime (e.g., Goldberg et al., 2019). As discussed in Section 3.3.1, we would expect systematic biases in our calculated slopes and uncertainties if these seasonal effects were the main sources of error, but there is no evidence that this is the case. Accordingly, using the atmospheric correlation between NO_2 and CO_2 to generate NDCFs allows us to fairly simply estimate CO_2 emissions throughout the year.

Single-day estimates of urban CO_2 emissions with the NDCF method would result in high methodological uncertainties, but aggregating emissions estimates from the large quantity of NDCFs that can be generated decreases uncertainties substantially. With the hundreds of NO_2 swaths taken over our study periods, our aggregated emissions estimates correspond to methodological uncertainties of only about 1-6% of the emissions estimates themselves (excluding uncertainties related to the mass balance method). Thus, we do not recommend the use of the NDCF method on a single-day basis but rather in aggregate, whether to assess subannual trends or to find an annual emissions estimate. Additionally, while a simple mass balance method is used in this study to demonstrate the utility of the NDCF method, CO_2

emissions can also be estimated from NDCFs using methods that are more rigorous to calculate more precise emissions estimates.

The NDCF method is shown to be a fairly powerful tool for assessment of time-averaged CO_2 emissions, though limited observational data of CO_2 may cause difficulties. In our study, we use a very limited number of SAMs (three) for Melbourne, due to limited overpasses and our data filtration. With only a small number of slopes to work with for this city, our mass balance emissions estimates are still reasonable as compared to EDGAR, as the slopes passing the filter appear to be fairly representative. Slope filtration is particularly important for a case like this, as a single unrepresentative slope could skew emissions estimates substantially.

Of the four slope application methods we explore in this study, we find that the Single Slope case is the most susceptible to physically unrealistic temporal fluctuations, whereas the other three result in more reasonable temporality in emissions estimates. The Average Slopes and Rolling Averages cases demonstrate monthly variations in emissions that we expect based on EDGAR (see Figures 3.12 - 3.14) and are much better candidates for use in future studies. The Single Slope with Baseline case also captures these temporal variations for Melbourne and Mexico City, but this is not demonstrated as much for Buenos Aires (Figures 3.9-3.14). This particular case seems to capture the temporality of slopes while dampening some of the potential noise associated with these slopes. Limited co-located observational data hampers our ability to understand how slopes change over time and evaluate how well the Single Slope and Single Slope with Baseline methods work. If higher frequencies of co-located observations were to be available, this would likely increase the utility of the Single Slope and Single Slope with Baseline methods, though with a high enough frequency it would be more sensible to use the CO_2 observations themselves rather than the NDCF method. Given the current state of data

availability and the results of this study, using a slope application method that averages at least two slopes together seems to work well in aggregate. Overall, the use of the vast quantities of available NO_2 data to generate NDCFs and facilitate the study of urban CO_2 emissions is quite powerful.

3.5 Summary and Conclusions

Using proxy species such as NO_2 has proven to be a useful tool for the estimation of urban CO_2 emissions. In this study, we demonstrate a method of utilizing empirical relationships between satellite observations of CO_2 and NO_2 and applying these relationships to NO_2 fields to generate NO_2-derived CO_2 fields, or NDCFs. We first establish this method with GEOS-CF simulations of these gases for the cities of Buenos Aires, Melbourne, and Mexico City, and then we evaluate the utility of the method on observations of CO_2 from OCO-3 in SAM mode and NO_2 from TROPOMI. We find methodological uncertainties from the spread of relative differences at the grid cell level between the NDCFs and SAMs ranging between 0.35% (1.44 ppm assuming a 410 ppm concentration) and 0.79% (3.26 ppm). We also use a simple mass balance method to estimate emissions from NDCFs generated using four different slope application methods: Single Slope, Average Slopes, Single Slope with Baseline, and Rolling Averages. Slope application methods that average more than one slope together are more effective. We make monthly-scale emissions estimates that are comparable to monthly sectoral EDGAR emissions inventory estimates from 2015, an unprecedented temporal scale achieved by space-based urban CO_2 emissions estimates. We also find that, due to the large number of TROPOMI NO_2 swaths available for the generation of NDCFs, our methodological uncertainties are drastically reduced when the emissions estimates are aggregated by month or longer. While the current availability of CO_2 satellite data limits the amount of information we can extract from

such data on its own, the use of this data in conjunction with more abundant proxy data creates many possibilities for future studies to better understand and constrain urban CO_2 emissions.

Chapter 4 Evaluation of the Utility of Space-Based Observations of CO_2 for Assessing Urban Climate Action Program Effectiveness

Abstract

Cities play an important role in tackling climate change. However, there are questions as to whether the greenhouse gas (GHG) emissions estimates reported in their self-reported inventories (SRIs) are accurate enough to track the effectiveness of the policies in their climate action plans. Here, we evaluate the ability of satellite observation-derived estimates of carbon dioxide (CO_2) to assess the effectiveness of urban climate action programs in the U.S. First focusing on Los Angeles (L.A.), we apply results from previous studies as well as derive approximate uncertainties associated with CO_2 emissions estimates based on observations in Snapshot Area Map (SAM) mode from the Orbiting Carbon Observatory-3 (OCO-3). We determine that satellite-based emissions estimates can help inform SRI emissions estimates, but they cannot be used to distinguish between year-to-year Scope 1 emissions. We find a greater utility of SAM-based emissions estimates for assessing L.A.'s overarching GHG emissions targets, and we explore the ability of these estimates to track specific policies in their urban climate action plan and sectoral emissions. We then apply our findings to other urban climate action plans in the U.S. Overall, we find a lack of compatibility between the designs of climate action plans and the current scientific capabilities associated with satellite observations of CO_2, and we discuss options for improving the ability of satellite observations of CO_2 to provide policy-relevant assessments of emissions.

4.1 Introduction

Cities play an inextricable role in the wicked problem of climate change. Despite covering less than 5% of the area of the globe, cities are the source of the majority of anthropogenic global greenhouse gas (GHG) emissions, an outsized contribution to climate change (van der Heijden, 2019; International Energy Agency, 2008). While contributing heavily to climate change, cities are also the places where the negative impacts of climate change will be felt the most severely (van der Heijden, 2019). Along with this, cities are increasingly becoming important actors in efforts to both mitigate and adapt to climate change (Rosenzweig et al., 2010).

While climate policy is making fairly minimal progress at the national and international levels, thousands of city governments around the world have committed to reducing their GHG emissions to mitigate climate change (Hughes, 2019). Many of these cities are joining transnational city networks like the C40 Climate Leadership Group and Local Governments for Sustainability (referred to as ICLEI). These cities have set ambitious targets to reduce their GHG emissions over the next few decades. The ability of these cities to achieve their emissions reductions goals lies in the effectiveness of their climate action plans, a set of policies designed to ultimately reduce the city's GHG emissions (Deetjen et al., 2018). Deetjen et al. (2018) reviewed 29 urban climate plans in the U.S. and found that cities can implement five broad climate change mitigation strategies in their climate action plans: 1) shifting transportation modes, 2) reducing building energy consumption, 3) reducing power sector emissions, 4) improving public utilities and green spaces, and 5) addressing regional impacts.

In order to assess the success of these strategies, cities estimate their own GHG emissions and report them in self-reported inventories (SRIs) (Mueller et al., 2021). These SRIs can contain

emissions estimates across three scopes: Scope 1, which come from sources located within the city boundary; Scope 2, which derive from the use of grid-supplied electricity; and Scope 3, which come from goods and services from outside the city boundary but result from actions occurring within the city boundary (LA Sanitation & Environment, 2022; Mueller et al., 2021). However, estimating GHG emissions is complex, and city governments may not have the resources to account for their emissions accurately, or they may be politically motivated to avoid disclosing all of their emissions (Hughes, 2019). It has been found that, on average, cities under-report their GHG emissions by 18.3% (Gurney et al., 2021), calling into question the basis of their climate action plans and associated progress toward plan targets.

While the cities themselves may have difficulties in tracking their own emissions, there is an increasing interest among the scientific community in using observations of carbon dioxide (CO_2) to make policy-relevant assessments of CO_2 emissions (e.g., Lauvaux et al., 2020; Mueller et al., 2021). These researchers use a combination of CO_2 observations and bottom-up inventory information to evaluate Scope 1 GHG emissions, which can readily be linked to atmospheric observations and are more straightforward to estimate than other scopes (Mueller et al., 2021). Discrepancies in Scope 1 emissions estimates between the SRIs and the observation-based estimates, then, imply discrepancies between other types of reported and actual emissions.

In recent years, satellites have opened the doors to high-resolution, temporally frequent, globally available observations of CO_2. There have been a number of studies that have used satellite observations to assess urban CO_2 emissions (e.g., Kort et al., 2012; Hakkarainen et al., 2016; Hedelius et al., 2018; Wu et al., 2018; Park et al., 2020; Wu et al., 2020; Yang et al., 2020; Ye et al., 2020; Kiel et al., 2021; Lei et al., 2021; Kuze et al., 2022). Extensions of these space-based studies could lead to CO_2 emissions estimates that could be useful for the assessment of

SRIs and urban climate action plans. In particular, Los Angeles (L.A.) has been the focus of a number of space-based studies (e.g., Hedelius et al., 2018; Kiel et al., 2021; Kort et al., 2012; Roten et al., in press; Wu et al., in press; Ye et al., 2020), making this megacity a valuable test case for the utility of satellite observation-based emissions estimates in evaluations of urban policies.

In this chapter, we seek to answer the following question: What can satellite observations tell us about the effectiveness of urban climate action programs? First focusing on the city of L.A., we evaluate how well current satellite-based observations and related techniques can assess Scope 1 emissions of L.A.—those reported in SRIs, as well as how they change over time. We then evaluate the utility of satellite assessments of overarching GHG emissions reduction targets in L.A.'s climate action plan, as well as the more specific policies and sectors of their plan. We then discuss how our findings can be applied to other urban climate plans in the U.S., and what types of improvements can be made to achieve a stronger overlap between what urban policy makers need and what scientific experts in space-based CO_2 emissions estimates can provide.

4.2 Satellite Assessments of L.A.'s Climate Action Plan

4.2.1 Overview of L.A.'s Climate Action Program

In 2019, L.A. released their climate action plan, entitled *L.A.'s Green New Deal*, which is the first four-year update to their 2015 *Sustainable City pLAn* (https://plan.lamayor.org/). (Hereafter their climate action plan is referred to as the "pLAn.") This pLAn is governed by four key principles: 1) a commitment to the Paris Climate Agreement and to scientifically driven strategies for achieving carbon neutrality, zero waste, and zero wasted water; 2) the delivery of environmental justice and equity, 3) a green economy accessible to all Angelenos, and 4) leadership by example to drive positive change. These key principles that drive the pLAn are

interdisciplinary by nature and focus on many aspects other than the reduction of GHG emissions alone, and the pLAn itself covers many goals that overlap with Sustainable Development Goals (Mayor Eric Garcetti, 2019).

Regarding GHG emissions, the pLAn sets the overarching goals of cutting emissions by 50% by 2025 as compared to their 1990 baseline, 73% by 2035, and 100% (carbon neutral) by 2050 (Mayor Eric Garcetti, 2019). In order for the city to accomplish these goals, the plan has many targets organized into 13 focus areas: Renewable Energy, Local Water, Clean and Healthy Buildings, Housing & Development, Mobility and Public Transit, Zero Emission Vehicles, Industrial Emissions and Air Quality Monitoring, Waste and Resource Recovery, Food Systems, Urban Ecosystems and Resilience, Environmental Justice, Prosperity and Green Jobs, and Lead By Example (objectives related to L.A. municipal functions). The combination of these grouped goals all contribute to reducing emissions overall by the overarching goals listed above, as well as reducing emissions separated by sector, as reproduced in Figure 4.1. Note that the targets proposed in the current pLAn would not allow the city to achieve carbon neutrality by 2050; new technology or carbon negative projects would be required to bridge the gap between the current targets and the goal of carbon neutrality.

To track progress toward their emissions goals, the city of L.A. is using an annual GHG inventory as part of their climate action program (LA Sanitation & Environment, 2022). This SRI contains the city's GHG emissions estimates based on the Global Protocol for Community-Scale GHG Emission Inventories (GPC) for a "Basic" reporting level, which includes Scope 1 emissions from fuel use in building, transport, and industry; Scope 2 emissions from energy consumption from the grid; and Scope 1 and 3 emissions from waste from within the boundary of the city (Mayor Eric Garcetti, 2019). Table 4.1 reproduces the emissions estimates in the

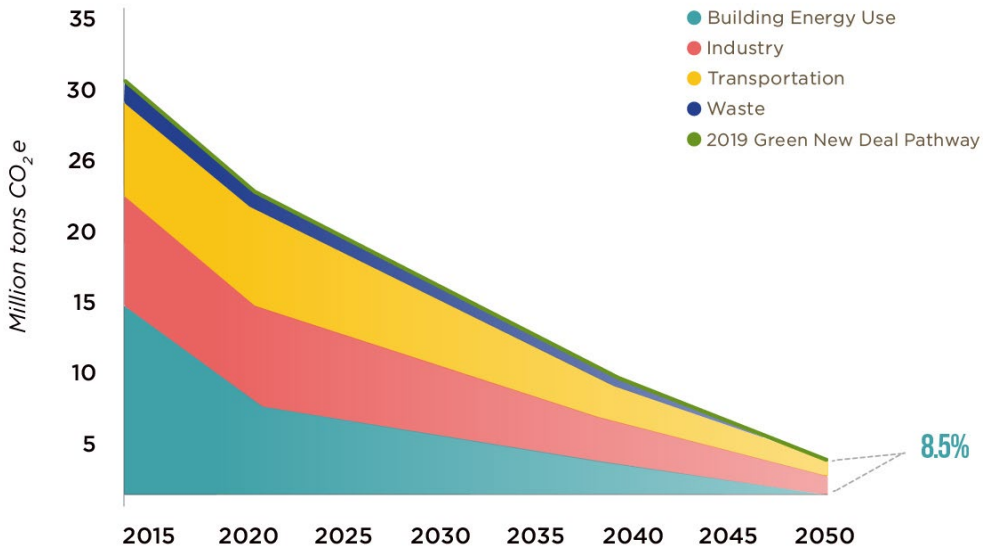

Figure 4.1 L.A. GHG emissions by sector. The targets in the current pLAn would reduce emissions down to 8.5% by 2050. Future iterations of the pLAn will aim to close this gap by using carbon negative projects or new technology. Figure reproduced from https://plan.lamayor.org/.

	1990	*2014*	*2015*	*2016*	*2017*	*2018*	*2019*	*2020*	*1990 vs. 2020 percent change*
Stationary Energy	26.0	23.8	23.1	19.6	19.6	19.1	17.7	16.8	-35%
Transportation	5.6	4.7	4.3	4.1	3.9	3.9	3.9	3.0	-47%
Waste	1.2	1.1	1.2	1.2	1.3	1.3	1.3	1.3	12%
Total Emissions	32.8	29.6	28.6	24.9	24.8	24.3	22.9	21.1	-36%

Table 4.1 L.A.'s GHG emissions inventory data in $MtCO_2e$, showing the 1990 baseline and annual estimated emissions from 2014 to 2020, separated by sectors. Reproduced from LA Sanitation & Environment (2022).

inventory by sector, in units of megatonnes CO_2 equivalent ($MtCO_2e$). L.A.'s SRI estimates are further broken down by subsectors (e.g., types of buildings, types of transportation, etc.), but not by scope; many of the subsectors would produce emissions associated with multiple scopes at once. Note that these reported emissions are self-calculated. Gurney et al. (2021) showed that L.A. under-reported its emissions by more than 50%, suggesting significant flaws in their accounting.

4.2.2 Ability of Satellite Observations to Track L.A.'s Scope 1 Emissions and Assess Overarching Emissions Goals

In this section, we evaluate the usefulness of satellite observation-based CO_2 emissions estimates in assessing L.A.'s SRI, Scope 1 emissions, and overarching emissions targets listed in the pLAn. We first examine the results of previous space-based studies that estimated urban CO_2 emissions and determine what similar results for L.A. would mean in terms of their climate action program. Because of the dearth of such studies, we then make an approximate estimation of uncertainties associated with the latest satellite technology and techniques to evaluate what these methods have to offer in terms of policy relevance for the city's emissions overall and associated targets.

There are only a few studies that have used satellite observations to estimate urban CO_2 emissions with a full uncertainty analysis. We take two existing examples and discuss their implications on the use of satellite-based assessments of L.A.'s climate action program.

In Chapter 2 of this dissertation, Yang et al. (2020) used satellite observations to estimate urban CO_2 emissions in the Middle East. Despite having only a small number of Orbiting Carbon Observatory-2 (OCO-2) overpasses to work with and thus large uncertainties for each city, often greater than 50%, the researchers were able to inform the adjustment of inventory magnitudes for these Middle Eastern cities, finding that most of the inventory representations of emissions magnitudes for those cities were underestimated as compared to what the observations suggested. Given that Gurney et al. (2021) found that L.A. was underestimating their emissions by more than 50%, this suggests that even a study like that of Yang et al. (2020), which used a very simple integral method and had limited numbers of observed overpasses to work with (between 3 and 11 for the different cities studied), could be useful in informing the emissions of

L.A. and assessing the accuracy of the SRI Scope 1 emissions estimates. This is especially the case because the uncertainty associated with such methods would decrease with an increasing number of overpasses, and L.A. is observed quite frequently by satellite compared to other cities. We expect Scope 1 emissions to be about half the total GHG emissions from a city (Mueller et al., 2021). Accordingly, assessments like those done in Chapter 2 but for the city of L.A. could quite feasibly estimate L.A.'s Scope 1 emissions with enough certainty to confirm whether half the SRI's reported emissions are vastly underestimated, as implied in the study by Gurney et al. (2021).

A more sophisticated method was used by Hedelius et al. (2018) to estimate CO_2 emissions of the South Coast Air Basin (SoCAB), which includes L.A. In their study, they used 29 OCO-2 overpasses and a Bayesian inversion over the time period of September 2014 to August 2016 to estimate SoCAB CO_2 emissions, and found a fractional uncertainty of 25%. While this uncertainty level is much smaller than that found by Yang et al. (2020), likely mostly attributed to the much larger quantity of OCO-2 overpasses used, a 25% fractional uncertainty on an emissions estimate is still quite large. Certainly, L.A.'s SRI could be assessed in the same way as if we applied the methodologies used by Yang et al. (2020), but is there greater policy relevance associated with emissions estimates with an uncertainty of 25%? If we assume a straight-line trend downward from the year 2014 (the first year tracked in L.A.'s SRI) to 2050, there would need to be an approximately 3% drop in emissions per year, on average, to achieve carbon neutrality. As such, a 25% fractional uncertainty on an emissions estimate is insufficient for tracking annual emissions changes, as year-to-year emissions estimates would be indistinguishable from each other with that level of uncertainty. Satellite-based emissions estimates of CO_2 with a fractional uncertainty of 25% should be able to assess progress toward

the pLAn's overarching GHG emissions goals, however. If we continue to assume that half of the total GHG emissions are Scope 1, L.A.'s goals for 2025 and 2035 equate to 8.2 and 4.428 $MtCO_2$ in Scope 1 emissions. With a 25% uncertainty attached to emissions of that magnitude, these goals are distinguishable from each other. Thus, these satellite observations can be used to assess longer-term progress toward L.A.'s overarching GHG emissions goals, even if they cannot be used to evaluate emissions trends at an annual pace.

What if the latest satellite technology, in place of OCO-2, were used for the sake of assessing L.A.'s climate action program? We seek to get a sense of how well Orbiting Carbon Observatory-3 (OCO-3) Snapshot Area Map (SAM) data can constrain L.A. emissions, instead of the OCO-2 data used by Hedelius et al. (2018). The SAM observation mode was specifically designed to capture anthropogenic emissions, allowing observational data to be collected over large contiguous areas (approximately 80×80 km^2) with a single overpass (Kiel et al., 2021). As such, SAMs are the state of the art in terms of satellite observations of urban CO_2, and we aim to evaluate the utility of these observations for L.A.'s climate action program.

The uncertainties associated with the study conducted by Hedelius et al. (2018) were calculated by adding in quadrature uncertainties of five different types: random uncertainties (5%), uncertainty due to choice of initial values in the inversion process (10%), uncertainties in the prior flux (5%), uncertainties from observations and the boundary condition (10%), and uncertainties due to model winds (20%). When we filter L.A. SAMs to those with at least 1000 soundings, there are 37 in the period of February 2020 to May 2022. As Hedelius et al. (2018) used 29 overpasses, we would expect a decrease in the total uncertainty due to an increased number of observations. If we simply adjust the total 25% uncertainty to account for the differing number of observations by multiplying it by the square root of 29 and dividing by the square root

of 37, we find a reduction in overall uncertainty to 22.6%. If we used OCO-3 SAMs instead of OCO-2 overpasses, we would additionally expect reductions in the uncertainties from observations and the boundary condition due to the increased spatial coverage from SAMs. In the unrealistic case that the use of SAMs rather than OCO-2 overpasses would reduce such observational and boundary condition uncertainties to 0%, the total fractional uncertainty would drop to 20.8%, which is not a large enough reduction to change the policy relevance of these types of emissions estimates. For the purposes of further discussion, we estimate that the use of SAMs in place of OCO-2 observations would cut the observational and boundary condition uncertainties in half, leading to a total fractional uncertainty of 21.2%, a value that is still too large for annual emissions tracking.

In order to actually distinguish year-to-year emissions estimates of Scope 1 emissions, uncertainties would need to be much smaller. Examining Table 4.1, we can calculate that the mean difference in total emissions from year to year is 1.42 $MtCO_2e$, meaning Scope 1 emissions would, on average, be expected to decrease by about 0.7 $MtCO_2$ per year. This suggests that an uncertainty of about 0.35 $MtCO_2$ (~3% fractional uncertainty) could be policy-relevant. If we were to reduce our 21.2% value uncertainty to 3% by only increasing the number of observations, we would need 1848 SAMs. This is not particularly feasible for an annual estimate. In other words, the current state of satellite technology and the associated components required to convert observations of concentrations into emissions is not useful for year-to-year emissions tracking for L.A. Improvements in all aspects of the estimation process—number of overpasses, instrument accuracy, background calculation, prior accuracy, and transport modeling accuracy—would need to occur before this type of emissions tracking could be made with satellites.

Essentially, we find that satellite observations of CO_2 and associated techniques are broadly useful for diagnosing inconsistencies in L.A.'s SRI estimates, as well as for assessing progress toward the pLAn's overarching emissions reduction goals. However, assessments at finer temporal scales of a year or less would require improvements on all inputs into inversion schemes.

4.2.3 Satellite Observation-based Assessments of Specific Targets and Sectoral Emissions Delineated in the pLAn

Under each of its 13 focus areas, the pLAn has specific targets. For example, one of the targets under the Clean and Healthy Buildings focus area is to reduce building energy use per square foot for all building types 22% by 2025, 34% by 2035, and 44% by 2050 as compared to a 2015 baseline (Mayor Eric Garcetti, 2019). Such a target mixes scopes, as buildings use a combination of electricity from the grid (Scope 2) and other energy sources like natural gas that release emissions on-site when burned in furnaces for heating, for example (Scope 1). Overlapping scopes makes such targets like this difficult to measure via satellite observation without further parsing. Even when targets do not overlap in scope, such as the goal to reduce industrial emissions by 38% by 2035 and 82% by 2050, there are difficulties in observing such targets, as satellites are able to observe the integral of all emissions in an area without separating by focus area. Moreover, some of the targets that are listed in the pLAn do not even have GHG accounting associated with it, such as for Local Water goals, which include a statement that the L.A. Department of Water and Power is at present developing protocols to monitor GHG emissions related to water management; without existing accounting, satellite observations cannot offer a useful constraint on policy changes. Essentially, specific targets in the pLAn are

not particularly amenable to CO_2 satellite-based assessment, as they have been designed for practitioners without bearing in mind compatibility with assessment via observations.

Tracking L.A.'s sectoral emissions, rather than specific targets, would potentially be feasible by satellite, though this has never done before. Figure 4.1 shows the emissions trends related to sectors that the pLAn aims to achieve. If L.A. separated these trends out by scope, we have reason to believe that the sectoral, Scope 1 trends could be tracked on a multiple-year timescale in the same way that whole-city Scope 1 emissions can be tracked, and the SRI's sectoral emissions estimates can be evaluated. As mentioned above, using only satellite observations of CO_2 would not be able to separate out sectoral emissions, as satellites see the integral of all CO_2 in an area. However, combining satellite observations of CO_2 with additional *a priori* information could allow for tracking Scope 1 emissions by sector.

Nathan et al. (2018) attempted to use flask-based measurements of different atmospheric trace gases alone to attribute sector-level emissions in Indianapolis, but they were unable to do so successfully due to the spatial overlap of almost every source sector. They found that the use of additional information, such as inventories of trace gases at the sector level, would be necessary to accomplish this task. Lauvaux et al. (2020) achieved policy-relevant, sector-specific assessments (of grouped stationary vs. non-stationary sectors) of Indianapolis CO_2 emissions using a combination of a dense network of atmospheric sensors and a very high-resolution space/time-explicit bottom-up, sector-specific emissions inventory called Hestia (Gurney et al., 2012). Hestia is only available for four cities, including L.A. While this has yet to be done, the methods used by Lauvaux et al. (2020) could be used while substituting the observations from the network of atmospheric sensors with observations from SAMs to produce sector-specific emissions estimates. Mueller et al. (2021) also suggested that inversions utilizing observations in

combination with bottom-up, sector-explicit inventories with the granularity of Hestia could help city planners to detect the largest-emitting buildings as priorities to audit for achieving goals in their climate action plans. Meanwhile, Wu et al. (in press) used SAMs and TROPOMI carbon monoxide observations in combination with an urban land cover classification dataset to study sector-specific combustion efficiencies—the first satellite-based, intra-urban, sector-specific assessment that we know of. The researchers stated that the use of more sophisticated approaches could constrain sector-specific emissions of the two gases. Note that the urban land cover classification dataset used by Wu et al. is only available for a few cities around the world, including L.A.

In summary, various studies have suggested that L.A.'s CO_2 emissions at a sectoral level could be constrained in a policy-relevant manner using a combination of satellite observations and either a very high-resolution, sector-specific, bottom-up inventory like Hestia or an urban land cover classification dataset. However, this work has never previously been done, and would be an involved process requiring data that is available for L.A. and only a few other locations. Additionally, for sector-level emissions estimates to be useful, L.A.'s SRI would need to explicitly parse out Scope 1 emissions for each sector.

4.3 Looking Beyond L.A.

4.3.1 Application to U.S. Urban Climate Action Plans

In this section, we review the climate action plans of four other cities in the U.S. (Ann Arbor, MI; Portland, OR; Boston, MA; and Austin, TX) and assess similarities and differences with each other and with that of L.A. We then make assessments about the ability of satellite observations of CO_2 to assess progress on climate action plans overall.

Ann Arbor's Living Carbon Neutrality Plan is a document whose fourth and latest version was released in April 2020 (Ann Arbor Office of Sustainability & Innovations, 2020). The main goal of this plan is to achieve carbon neutrality by 2030. The plan is comprised of seven overarching strategies: 1) power the grid with 100% renewable energy, 2) switch appliances to electric, 3) improve energy efficiency, 4) reduce vehicle miles traveled by at least 50%, 5) change the use, reuse, and disposal of materials, 6) enhance resilience, and 7) other. These strategies are explicitly acknowledged as being linked with each other. 44 actions are associated with these strategies, with quantified GHG emissions reductions associated with many of the actions. While scopes are not explicitly mentioned, the plan implies that emissions that are the focus of being cut are associated with Scopes 1 and 2. The plan separates energy use and emissions goals for different sectors: Residential Electricity, Commercial Electricity, University of Michigan Electricity, Grid Losses, Residential Natural Gas, Commercial Natural Gas, University of Michigan Natural Gas, Fugitive Emissions, Transportation, Residential Solid Waste, and Commercial Solid Waste.

Portland's 2015 *Climate Action Plan* has overarching goals of reducing GHG emissions by 40% of their 1990 baseline by 2030 and 80% by 2050 (City of Portland, 2015). Carbon emissions in the plan are organized into three sectors: building energy, transportation, and waste disposal. There are nine focus areas (Buildings and Energy; Urban Form and Transportation; Consumption and Solid Waste; Food and Agriculture; Urban Forest, Natural Systems, and Carbon Sequestration; Climate Change Preparation; Community Engagement, Outreach, and Education; Local Government Operations; and Implementation), some of which overlap with sectors and some of which are not explicitly or quantitatively tied to GHG emissions. Each focus area has objectives and associated specific targeted actions. In order to track emissions, Portland

has sector-based and consumption-based SRIs to approximate emissions associated with Portland's activities, whether they take place within or outside of the boundaries of the city itself (all scopes).

Boston's 2019 update to its climate action plan lists goals of reducing emissions by 50% of its 2005 baseline by 2030, and 100% (carbon neutral) by 2050 (Mayor Martin J. Walsh, 2019). The plan relies on accounting for the following sectors: commercial/industrial/large residential, transportation, small residential, fugitive emissions, and wastewater treatment. The straight-line decrease in emissions from 2005 to 2050 amounts to a 2.2% decrease in emissions per year. The plan has actions and targets for three areas: buildings, transportation, and energy supply. While there are specific goals for each area, some are quantifiable in terms of GHG emissions, and others are not. No scopes are mentioned in the plan, and many specific goals overlap in scope, making measurement via satellite difficult (e.g., energy efficiency in buildings).

The city of Austin is aiming to achieve net-zero GHG emissions by 2040 in its plan released in 2021 (City of Austin, 2021). The plan's associated SRI covers Scope 1 and Scope 2 emissions separated by sector: energy in buildings, transportation, industrial, refrigerants, and waste. There are five climate equity goal areas in the plan: sustainable buildings, transportation and land use, transportation electrification, food and product consumption, and natural systems. For each goal area, there are specific strategies, some more quantifiable than others, and many overlapping in terms of emissions scopes. Some strategies also require additional evaluation of energy usage before specific plans can be made. Many strategies lack clear, quantifiable metrics for assessment of success. Carbon offsets are included as part of the plan, with the goal of reducing direct emissions as much as possible but using carbon offsets for any remaining emissions to achieve carbon neutrality by 2040.

What is apparent from reviewing these four urban climate action plans along with L.A.'s pLAn is that they are not standardized at all, and are instead fit to the needs, values, and capabilities of the cities themselves. While this is not surprising, it does make it very difficult for a standardized procedure to measure GHG emissions to be useful. Additionally, for all five of these cities, many of the targets and goal areas are very much interconnected, without clear boundaries between said targets regarding emissions; instead, planned actions often cover multiple scopes at once, without necessarily having a clearly quantified metric for success. Without delineating out Scope 1, 2, and 3 GHG emissions reduction goals, satellite observations of CO_2 cannot be used to assess the impact of specific policy actions. Based on the organization of these climate action plans, it seems that satellite observations would instead be useful for the purposes of evaluating the accuracy of SRIs and assessing multi-year progress toward overarching emissions goals. We expect that, while the assumption of Scope 1 emissions comprising half the emissions of a city is not a perfect one, it can serve as a reasonable approximation for the sake of tracking progress toward overarching emissions goals, in the absence of SRIs better delineated by scopes.

While all five cities separate out emissions by sector, the sectors they choose to separate by are not standardized either. As done by Lauvaux et al. (2020), sectors can be combined into several groupings to make observation-based assessment simpler and more standardized. However, grouping sectors together would make satellite assessments of emissions useful on a broader scale as opposed to relevant for the assessment of specific, targeted policies. Moreover, Hestia, the sector-specific, high-resolution, bottom-up inventory used by Lauvaux et al. (2020) in their study, is only available for four cities in the U.S. Lauvaux et al. (2020) asserted that the quality and amount of information required to produce the Hestia product for Indianapolis was

"unprecedented" and would be "challenging to reproduce for many cities across the world," suggesting that the kind of bottom-up inventories available for other cities would only be able to provide whole-city policy-relevant assessments, rather than sector-specific assessments. As such, inventories like Hestia need to be generated for other cities to have the same type of analysis done at the sector level.

There are some additional complications involved in assessing urban climate action plans around the U.S. using satellite observation-based emissions estimates. For one, not all cities have the same type and quality of satellite data available. SAMs are the state of the art in terms of urban satellite CO_2 observations, but there are no SAMs taken for a smaller city like Ann Arbor, MI. There are OCO-2 overpasses that could potentially be used to generate emissions estimates, but the much smaller density of observations (10 km wide as opposed to ~80 km × 80 km SAMs) coupled with the smaller CO_2 signals from a smaller city makes satellite assessment of a city like Ann Arbor much more difficult. Along with this, there are general complications associated with all cities, regardless of size. First, emissions in an area that a satellite can observe are not limited to administrated boundaries of a city, so care must be taken to ensure that emissions estimates are limited to the city in question in order to be policy-relevant. Second, observed reductions in emissions cannot necessarily be attributed to the actions of a city government alone, as urban GHG emissions can be influenced by policies implemented at other levels of government (Hughes, 2019). This suggests that satellite observations of CO_2 are better fit for the role of assessing more overarching goals than specific policies.

Overall, it seems that using satellites like OCO-3 to estimate CO_2 emissions from cities has some utility for assessing broad goals in urban climate action plans, but lacks the ability to assess progress toward specific targets without more clearly defined and quantifiable goals in the

plans, separated by scope. If cities want to better validate their SRIs, track progress of their plans, and assess sector-level progress with observation-based emissions estimates, they need to more clearly define scopes and create measurable targets. On the observation side, *a priori* inventories like Hestia need to be made available to potentially achieve sector-specific emissions estimates to compare with sector-specific SRI values. Along with development of or improvements to *a priori* inventories, satellite technology itself as well as transport modeling need to undergo improvements to decrease the gap between current capabilities and policy-relevant constraints to emissions at the city level.

4.3.2 Further Bridging the Science-Policy Gap

Given the underestimated emissions reported in SRIs (Gurney et al., 2021) and the strong interest in the scientific community to provide policy-relevant CO_2 emissions estimates (e.g., Gurney et al., 2020; Lauvaux et al., 2020; Mueller et al., 2021), it seems that both urban decision makers and scientists would benefit from a framework that allows for scientific assessments to provide more useful products for decision makers. How can we bridge the gap between CO_2 satellite studies and urban climate action plans? What is evident in this study is that there is a lack of compatibility between what the climate action plans aim to achieve and what emissions estimation techniques based on current satellite technology have the capability to provide.

Following a model of co-production of science and policy from the beginning helps to prevent mismatches such as have been discussed here (Lemos & Morehouse, 2005). This model suggests that communication and iteration between scientists and those who wish to use scientific results is key. This iteration is grounded in three components: 1) sustained stakeholder interaction, 2) usable science, and 3) interdisciplinarity. Stakeholder interaction refers to how much stakeholders are involved in aspects of research, such as in helping to define the problem

and research questions, choosing methods, actually conducting research and analysis, etc. In our case, this would refer to the involvement of urban planners and policy makers throughout every step of satellite observation-based work, from the planning stages (even in discussions of future satellite technology needs), to the asking of science questions, to the analysis of results. Usable science refers to science that results in knowledge that meets the needs of constituents. In our case, this would refer to satellite observation-based work that provides results that are immediately understandable and applicable to urban policy needs. Interdisciplinarity refers to scientists from different disciplines working together to solve problems that go beyond the scope of a single discipline. In our case, this would mean atmospheric scientists working with social and political scientists to define research questions, conduct research, and analyze results.

Achieving this level of co-production between science and policy is not easy, but science intermediaries such as boundary organizations, consultancies, and extension agencies can help connect the limited number of scientists to the many decision makers that would benefit from scientific information (Ultee et al., 2018). In their commentary on three cases of science intermediaries and their role in sea-level adaptation, Ultee et al. (2018) found that, while efforts to engage scientists with intermediaries can be resource intensive and context-dependent, intermediaries have the power to increase the benefits of science and engineering in preventing the negative effects of sea-level rise. While their commentary focused on sea-level rise and adaptation, we can extrapolate their findings to our own case, which would suggest that scientific intermediaries can play an important role in connecting researchers in the satellite and atmospheric science communities to urban planners seeking to track and reduce GHG emissions in their cities.

4.4 Summary and Conclusions

In this chapter, we aim to understand how well satellites can assess the different facets of urban climate action programs. With an initial focus on the city of L.A., we assess the level of uncertainty associated with CO_2 emissions estimates of the city using satellite observations. Previous studies suggest that OCO-2 satellite observations are able to help evaluate the accuracy of L.A.'s SRI, which serves as the basis for pLAn targets and is fundamental to the tracking of pLAn policies. The level of uncertainty associated with satellite-based emissions estimates is too high for us to track emissions from year to year. However, this level of uncertainty is sufficient for distinguishing between emissions that would be estimated during the target years of 2025, 2035, and 2050, suggesting that satellite observations can be useful for assessing the overarching targets of the pLAn. In a thought experiment replacing OCO-2 observations with SAMs, we roughly estimate a fractional uncertainty of 21.2% on emissions estimates, which is still too large to be able to distinguish between annual Scope 1 GHG emissions; improvements on all fronts of the inversion process need to be made before year-to-year emissions tracking can be accomplished. Recent studies suggest that, while specific policies in the pLAn cannot be assessed for their impact on GHG emissions, space-based emissions estimates could be used to assess L.A.'s sectoral level emissions over multi-year timespans. A review of multiple urban climate plans around the U.S. suggests that none of these plans parse out scopes well, making it difficult for space-based assessments of emissions to be policy-relevant. We find that there is an overall lack of compatibility between the climate action plans and the utility of satellite observation-based emissions estimates. There are improvements to be made on both the science and policy sides, such as increased availability of *a priori* datasets and clearer distinctions between scopes in climate action plan policies. Additionally, utilizing a model of science-policy

co-production, perhaps facilitated by science intermediaries, could, in the long-term, improve the policy relevance of space-based assessments of urban CO_2 emissions.

Chapter 5 Conclusions

5.1 Summary of Findings

This dissertation explores multiple facets of the evaluation of urban CO_2 emissions through the lens of space-based observations. It is incredibly important for us to constrain urban emissions, as the majority of fossil CO_2 emissions are derived from cities. For the sake of better monitoring the human contributions of climate change, evaluating plans for mitigating climate change, and understanding the different carbon cycle components' responses to both natural and anthropogenic changes, we need to improve our quantification of urban CO_2 emissions—and space-based observations can help us do this.

In Chapter 2, we evaluated global fossil CO_2 emissions inventory representations of afternoon urban CO_2 emissions in the Middle East, and assessed the ability of satellite observations to inform this evaluation. By finding relationships between top-down satellite observations and bottom-up simulations of XCO_2, we calculated scaling factors for adjusting prior inventory estimates of five Middle Eastern cities' emissions. We estimated that the sum of the cities' scaled emissions were on average 2.0 (1.0, 3.0) times the prior inventory magnitudes, suggesting that the underestimation of the cities' emissions amounted to ~7% (0%, 14%) of total emissions in the Middle East. These results were insensitive to spatial differences in the prior inventory representations of the studied cities' emissions, allowing for robust evaluations of urban emissions inventory representations.

Chapter 3 of this dissertation focused on the use of satellite observations of the proxy species NO_2 to estimate urban CO_2 emissions. We demonstrated a method applying empirical

relationships between satellite observations of CO_2 and NO_2 to NO_2 fields, generating NO_2-derived CO_2 fields (NDCFs). First establishing the method with GEOS-CF simulations for three cities around the world, we then evaluated the method using SAMs and TROPOMI observations. We calculated methodological uncertainties at the grid cell level between 0.35% (1.44 ppm assuming a 410 ppm concentration) and 0.79% (3.26 ppm). We used a simple mass balance method to estimate emissions from NDCFs generated using four different slope application methods, finding that slope application methods that averaged more than one slope together were more effective. We were able to produce CO_2 emissions estimates over time as well as reduce methodological uncertainties through aggregations of estimates, both facilitated by the much greater availability of NO_2 observations than those of CO_2.

In Chapter 4, we interrogated the utility of satellite-based CO_2 emissions in the assessment of the different components of urban climate action programs in the U.S. Applying the results of previous studies as well as a thought experiment on uncertainties associated with emissions estimates using the latest satellite technology, we found that space-based urban CO_2 emissions estimates can help evaluate the accuracy of SRIs and assess progress in attaining overarching emissions targets, but these uncertainties are too large to track annual emissions progress. We also explored the feasibility of satellite assessments of emissions at the intra-urban scale, determining that such assessments would not be able to track progress of specific policies in climate action plans. However, with some additional advancement in methodologies, they have the potential to provide policy-relevant estimates of sectoral emissions. We found that climate action plans are not designed to be compatible with assessment via satellite-based emissions estimates, especially because scopes are not parsed out well in the plans. We explored improvements that could be made on both the science and policy fronts to improve this

compatibility, recommending a model of science-policy co-production as a means of attaining more policy-relevant space-based evaluations of urban CO_2 emissions in the long term.

5.2 Further Research Directions

As space-based technology for the observation of CO_2 and proxy species continues to improve, more avenues of further research can be explored. The latest in urban CO_2 observational technology from space, OCO-3 SAMs, have opened up many pathways for the assessment of urban emissions, with measurements taken over much larger areas than used to be available, and during different times of day. Accordingly, we see many possibilities for future studies focusing on intra-urban gradients of emissions, assessed throughout the day. More work that expands upon sectoral emissions in particular, through the combined use of SAMs and satellite observations of co-located proxy species, would be fruitful. Future CO_2-observing satellite missions will only grow such opportunities. The European Copernicus program's planned constellation of multiple polar orbiting satellites, the Copernicus CO_2 Monitoring Mission (CO2M), is particularly exciting, as it will provide co-located measurements of CO_2 and NO_2 from the same satellites, allowing for improved sectoral emissions estimates at the urban scale (Kaminski et al., 2022).

Additionally, we can see the application of NDCFs providing policy-relevant assessments of CO_2 emissions with further improvements of the method. By providing large amounts of data previously unavailable while only using CO_2 observations, NDCFs have the potential to deliver CO_2 emissions assessments at a finer temporal scale than previously possible. Using the NDCF method in combination with more sophisticated inversion methods beyond the scope of the work in this dissertation could lead to the tracking of CO_2 emissions estimates over annual or even subannual scales, facilitating finer evaluation of progress toward urban climate action plan goals.

Finally, there is more work to be done on improving the utility of satellite-based emissions estimates of CO_2 for use by decision makers and practitioners. Improved interdisciplinarity, along with the expansion of science intermediary programs, can help to facilitate this. We see this as a vital step toward making our science more useful in the face of climate change—one of the most complex problems of our time—which cannot be solved without cooperation and collaboration across multiple fields and disciplines.

Bibliography

Ann Arbor Office of Sustainability & Innovations. (2020). *Ann Arbor's Living Carbon Neutrality Plan*. Retrieved from https://www.a2gov.org/departments/sustainability/Documents/A2Zero%20Climate%20Action%20Plan%20_4.0.pdf

Archer, D. (2010). *The Global Carbon Cycle*. Princeton, NJ: Princeton University Press.

Asefi-Najafabady, S., Rayner, P. J., Gurney, K. R., McRobert, A., Song, Y., Coltin, K., et al. (2014). A multiyear, global gridded fossil fuel CO2 emission data product: Evaluation and analysis of results. *Journal of Geophysical Research: Atmospheres*, *119*(17), 2013JD021296. https://doi.org/10.1002/2013JD021296

Berezin, E. V., Konovalov, I. B., Ciais, P., Richter, A., Tao, S., Janssens-Maenhout, G., et al. (2013). Multiannual changes of CO_2 emissions in China: indirect estimates derived from satellite measurements of tropospheric NO_2 columns. *Atmospheric Chemistry and Physics*, *13*(18), 9415–9438. https://doi.org/10.5194/acp-13-9415-2013

Betts, R. A., Jones, C. D., Knight, J. R., Keeling, R. F., & Kennedy, J. J. (2016). El Nino and a record CO2 rise. *Nature Climate Change*, *6*(9), 806–810. https://doi.org/10.1038/nclimate3063

Bodansky, D. (2016). The Paris Climate Change Agreement: A New Hope? *American Journal of International Law*, *110*(2), 288–319. https://doi.org/10.5305/amerjintelaw.110.2.0288

Boden, T., Andres, R., & Marland, G. (2016). Global, Regional, and National Fossil-Fuel CO2 Emissions (1751 - 2013) (V. 2016) [Data set]. Environmental System Science Data Infrastructure for a Virtual Ecosystem; Carbon Dioxide Information Analysis Center (CDIAC), Oak Ridge National Laboratory (ORNL), Oak Ridge, TN (United States). https://doi.org/10.3334/cdiac/00001_v2016

BP. (2017). Statistical Review of World Energy. Retrieved July 19, 2019, from https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html

Brinkhoff, T. (2018, May 21). City Population. Retrieved January 17, 2020, from http://www.citypopulation.de/

Buchwitz, M., Schneising, O., Reuter, M., Heymann, J., Krautwurst, S., Bovensmann, H., et al. (2017). Satellite-derived methane hotspot emission estimates using a fast data-driven

method. *Atmospheric Chemistry and Physics*, *17*(9), 5751–5774. https://doi.org/10.5194/acp-17-5751-2017

Caldeira, K., & Wickett, M. E. (2003). Anthropogenic carbon and ocean pH. *Nature*, *425*(6956), 365–365. https://doi.org/10.1038/425365a

Ciais, P., Dolman, A. J., Bombelli, A., Duren, R., Peregon, A., Rayner, P. J., et al. (2014). Current systematic carbon-cycle observations and the need for implementing a policy-relevant carbon observing system. *Biogeosciences*, *11*(13), 3547–3602. https://doi.org/10.5194/bg-11-3547-2014

City of Austin. (2021). *Austin Climate Equity Plan*. Retrieved from https://www.austintexas.gov/sites/default/files/files/Sustainability/Climate%20Equity%20Plan/Climate%20Plan%20Full%20Document__FINAL.pdf

City of Portland. (2015). *Climate Action Plan*. Retrieved from https://www.portland.gov/sites/default/files/2019-07/cap-2015_june30-2015_web_0.pdf

Coldewey-Egbers, M., Slijkhuis, S., Aberle, B., Loyola, D., & Dehn, A. (2018). The Global Ozone Monitoring Experiment: review of in-flight performance and new reprocessed 1995–2011 level 1 product. *Atmospheric Measurement Techniques*, *11*(9), 5237–5259. https://doi.org/10.5194/amt-11-5237-2018

Corburn, J. (2009). Cities, Climate Change and Urban Heat Island Mitigation: Localising Global Environmental Science. *Urban Studies*, *46*(2), 413–427. https://doi.org/10.1177/0042098008099361

Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., Aardenne, J. A. van, et al. (2018). Gridded emissions of air pollutants for the period 1970–2012 within EDGAR v4.3.2. *Earth System Science Data*, *10*(4), 1987–2013. https://doi.org/10.5194/essd-10-1987-2018

Crippa, M., Solazzo, E., Huang, G., Guizzardi, D., Koffi, E., Muntean, M., et al. (2020). High resolution temporal profiles in the Emissions Database for Global Atmospheric Research. *Scientific Data*, *7*(1), 121. https://doi.org/10.1038/s41597-020-0462-2

Crisp, D. (2015). Measuring atmospheric carbon dioxide from space with the Orbiting Carbon Observatory-2 (OCO-2). In *Proceedings of SPIE* (Vol. 9607, pp. 960702-1-960702–7). https://doi.org/10.1117/12.2187291

Deetjen, T. A., Conger, J. P., Leibowicz, B. D., & Webber, M. E. (2018). Review of climate action plans in 29 major U.S. cities: Comparing current policies to research recommendations. *Sustainable Cities and Society*, *41*, 711–727. https://doi.org/10.1016/j.scs.2018.06.023

Eldering, A., Taylor, T. E., O'Dell, C. W., & Pavlick, R. (2019). The OCO-3 mission: measurement objectives and expected performance based on 1 year of simulated data.

Atmospheric Measurement Techniques, *12*(4), 2341–2370. https://doi.org/10.5194/amt-12-2341-2019

European Commission Joint Research Centre. (2017, October 30). Methodology. Retrieved April 10, 2018, from http://edgar.jrc.ec.europa.eu/methodology.php

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Olsen, A., et al. (2020). Global Carbon Budget 2020. *Earth System Science Data*, *12*(4), 3269–3340. https://doi.org/10.5194/essd-12-3269-2020

Gately, C. K., & Hutyra, L. R. (2017). Large Uncertainties in Urban-Scale Carbon Emissions. *Journal of Geophysical Research: Atmospheres*, *122*(20), 2017JD027359. https://doi.org/10.1002/2017JD027359

Gaubert, B., Stephens, B. B., Basu, S., Chevallier, F., Deng, F., Kort, E. A., et al. (2019). Global atmospheric CO2 inverse models converging on neutral tropical land exchange, but disagreeing on fossil fuel and atmospheric growth rate. *Biogeosciences*, *16*(1), 117–134. https://doi.org/10.5194/bg-16-117-2019

Goldberg, D. L., Lu, Z., Streets, D. G., de Foy, B., Griffin, D., McLinden, C. A., et al. (2019). Enhanced Capabilities of TROPOMI NO2: Estimating NOX from North American Cities and Power Plants. *Environmental Science & Technology*, *53*(21), 12594–12601. https://doi.org/10.1021/acs.est.9b04488

Goldberg, D. L., Lu, Z., Oda, T., Lamsal, L. N., Liu, F., Griffin, D., et al. (2019). Exploiting OMI NO2 satellite observations to infer fossil-fuel CO2 emissions from U.S. megacities. *Science of The Total Environment*, *695*, 133805. https://doi.org/10.1016/j.scitotenv.2019.133805

Griffin, D., Zhao, X., McLinden, C. A., Boersma, F., Bourassa, A., Dammers, E., et al. (2019). High-Resolution Mapping of Nitrogen Dioxide With TROPOMI: First Results and Validation Over the Canadian Oil Sands. *Geophysical Research Letters*, *46*(2), 1049–1060. https://doi.org/10.1029/2018GL081095

Guo, H.-D., Zhang, L., & Zhu, L.-W. (2015). Earth observation big data for climate change research. *Advances in Climate Change Research*, *6*(2), 108–117. https://doi.org/10.1016/j.accre.2015.09.007

Gurney, K. R., Razlivanov, I., Song, Y., Zhou, Y., Benes, B., & Abdul-Massih, M. (2012). Quantification of Fossil Fuel CO2 Emissions on the Building/Street Scale for a Large U.S. City. *Environmental Science & Technology*, *46*(21), 12194–12202. https://doi.org/10.1021/es3011282

Gurney, K. R., Liang, J., O'Keeffe, D., Patarasuk, R., Hutchins, M., Huang, J., et al. (2019). Comparison of Global Downscaled Versus Bottom-Up Fossil Fuel CO2 Emissions at the Urban Scale in Four U.S. Urban Areas. *Journal of Geophysical Research: Atmospheres*, *124*(5), 2823–2840. https://doi.org/10.1029/2018JD028859

Gurney, K. R., Chen, Y., Maki, T., Kawa, S. R., Andrews, A., & Zhu, Z. (2005). Sensitivity of atmospheric CO2 inversions to seasonal and interannual variations in fossil fuel emissions. *Journal of Geophysical Research: Atmospheres*, *110*(D10). https://doi.org/10.1029/2004JD005373

Gurney, K. R., Romero-Lankao, P., Seto, K. C., Hutyra, L. R., Duren, R., Kennedy, C., et al. (2015). Climate change: Track urban emissions on a human scale. *Nature News*, *525*(7568), 179. https://doi.org/10.1038/525179a

Gurney, K. R., Song, Y., Liang, J., & Roest, G. (2020). Toward Accurate, Policy-Relevant Fossil Fuel CO2 Emission Landscapes. *Environmental Science & Technology*, *54*(16), 9896–9907. https://doi.org/10.1021/acs.est.0c01175

Gurney, K. R., Liang, J., Roest, G., Song, Y., Mueller, K., & Lauvaux, T. (2021). Under-reporting of greenhouse gas emissions in U.S. cities. *Nature Communications*, *12*(1), 553. https://doi.org/10.1038/s41467-020-20871-0

Hakkarainen, J., Ialongo, I., & Tamminen, J. (2016). Direct space-based observations of anthropogenic CO2 emission areas from OCO-2. *Geophysical Research Letters*, 2016GL070885. https://doi.org/10.1002/2016GL070885

Hakkarainen, Janne, Szelag, M. E., Ialongo, I., Retscher, C., Oda, T., & Crisp, D. (2021). Analyzing nitrogen oxides to carbon dioxide emission ratios from space: A case study of Matimba Power Station in South Africa. *Atmospheric Environment: X*, 100110. https://doi.org/10.1016/j.aeaoa.2021.100110

Hedelius, J. K., Liu, J., Oda, T., Maksyutov, S., Roehl, C. M., Iraci, L. T., et al. (2018). Southern California megacity CO_2, CH_4, and CO flux estimates using ground- and space-based remote sensing and a Lagrangian model. *Atmospheric Chemistry and Physics*, *18*(22), 16271–16291. https://doi.org/10.5194/acp-18-16271-2018

van der Heijden, J. (2019). Studying urban climate governance: Where to begin, what to look for, and how to make a meaningful contribution to scholarship and practice. *Earth System Governance*, *1*, 100005. https://doi.org/10.1016/j.esg.2019.100005

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Holmén, K. (1992). 11 The Global Carbon Cycle. In S. S. Butcher, R. J. Charlson, G. H. Orians, & G. V. Wolfe (Eds.), *International Geophysics* (Vol. 50, pp. 239–262). Academic Press. https://doi.org/10.1016/S0074-6142(08)62694-7

Huber, D. E., Steiner, A. L., & Kort, E. A. (2020). Daily Cropland Soil NOx Emissions Identified by TROPOMI and SMAP. *Geophysical Research Letters*, *47*(22), e2020GL089949. https://doi.org/10.1029/2020GL089949

Hughes, S. (2019). *Repowering Cities: Governing Climate Change Mitigation in New York City, Los Angeles, and Toronto*. Ithaca, NY: Cornell University Press.

Hutchins, M. G., Colby, J. D., Marland, G., & Marland, E. (2017). A comparison of five high-resolution spatially-explicit, fossil-fuel, carbon dioxide emission inventories for the United States. *Mitigation and Adaptation Strategies for Global Change*, *22*(6), 947–972. https://doi.org/10.1007/s11027-016-9709-9

Hutyra, L. R., Duren, R., Gurney, K. R., Grimm, N., Kort, E. A., Larson, E., & Shrestha, G. (2014). Urbanization and the carbon cycle: Current capabilities and research outlook from the natural sciences perspective. *Earth's Future*, *2*(10), 2014EF000255. https://doi.org/10.1002/2014EF000255

International Energy Agency. (2008). *World Energy Outlook 2008*. International Energy Agency.

IPCC. (2014). *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (p. 151). Geneva, Switzerland: IPCC.

Jacob, D. J. (1999). *Introduction to Atmospheric Chemistry*. Princeton, NJ: Princeton University Press.

Janssens-Maenhout, G., Pagliari, V., Guizzardi, D., & Muntean, M. (2013). *Global emission inventories in the Emission Database for Global Atmospheric Research (EDGAR) - Manual (I). Gridding: EDGAR emissions distribution on global gridmaps* (JRC Technical Reports) (pp. 1–33). Joint Research Centre of the European Commission. Retrieved from https://publications.europa.eu/en/publication-detail/-/publication/75050af2-850c-4b87-ae35-eaa82de96489/language-en

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., et al. (2019). EDGAR v4.3.2 Global Atlas of the three major greenhouse gas emissions for the period 1970–2012. *Earth System Science Data*, *11*(3), 959–1002. https://doi.org/10.5194/essd-11-959-2019

Kaminski, T., Scholze, M., Rayner, P., Houweling, S., Voßbeck, M., Silver, J., et al. (2022). Assessing the Impact of Atmospheric CO2 and NO2 Measurements From Space on Estimating City-Scale Fossil Fuel CO2 Emissions in a Data Assimilation System. *Frontiers in Remote Sensing*, *3*. Retrieved from https://www.frontiersin.org/article/10.3389/frsen.2022.887456

Kansakar, P., & Hossain, F. (2016). A review of applications of satellite earth observation data for global societal benefit and stewardship of planet earth. *Space Policy*, *36*, 46–54. https://doi.org/10.1016/j.spacepol.2016.05.005

Keller, C. A., Knowland, K. E., Duncan, B. N., Liu, J., Anderson, D. C., Das, S., et al. (2021). Description of the NASA GEOS Composition Forecast Modeling System GEOS-CF v1.0. *Journal of Advances in Modeling Earth Systems*, *13*(4), e2020MS002413. https://doi.org/10.1029/2020MS002413

Kiel, M., Eldering, A., Roten, D. D., Lin, J. C., Feng, S., Lei, R., et al. (2021). Urban-focused satellite CO2 observations from the Orbiting Carbon Observatory-3: A first look at the Los Angeles megacity. *Remote Sensing of Environment*, *258*, 112314. https://doi.org/10.1016/j.rse.2021.112314

Konovalov, I. B., Berezin, E. V., Ciais, P., Broquet, G., Zhuravlev, R. V., & Janssens-Maenhout, G. (2016). Estimation of fossil-fuel CO_2 emissions using satellite measurements of "proxy" species. *Atmospheric Chemistry and Physics*, *16*(21), 13509–13540. https://doi.org/10.5194/acp-16-13509-2016

Kort, E. A., Frankenberg, C., Miller, C. E., & Oda, T. (2012). Space-based observations of megacity carbon dioxide. *Geophysical Research Letters*, *39*(17), L17806. https://doi.org/10.1029/2012GL052738

Kuhlmann, G., Broquet, G., Marshall, J., Clément, V., Löscher, A., Meijer, Y., & Brunner, D. (2019). Detectability of CO_2 emission plumes of cities and power plants with the Copernicus Anthropogenic CO_2 Monitoring (CO2M) mission. *Atmospheric Measurement Techniques*, *12*(12), 6695–6719. https://doi.org/10.5194/amt-12-6695-2019

Kuhlmann, G., Brunner, D., Broquet, G., & Meijer, Y. (2020). Quantifying CO_2 emissions of a city with the Copernicus Anthropogenic CO_2 Monitoring satellite mission. *Atmospheric Measurement Techniques*, *13*(12), 6733–6754. https://doi.org/10.5194/amt-13-6733-2020

Kuze, A., Nakamura, Y., Oda, T., Yoshida, J., Kikuchi, N., Kataoka, F., et al. (2022). Examining partial-column density retrieval of lower-tropospheric CO2 from GOSAT target observations over global megacities. *Remote Sensing of Environment*, *273*, 112966. https://doi.org/10.1016/j.rse.2022.112966

LA Sanitation & Environment. (2022). *City of Los Angeles 2020 Community Greenhouse Gas Inventory*. Retrieved from https://www.lacitysan.org/cs/groups/public/documents/document/y250/mdc0/~edisp/cnt074753.pdf

Lauvaux, T., Miles, N. L., Deng, A., Richardson, S. J., Cambaliza, M. O., Davis, K. J., et al. (2016). High-resolution atmospheric inversion of urban CO2 emissions during the dormant season of the Indianapolis Flux Experiment (INFLUX). *Journal of Geophysical Research: Atmospheres*, *121*(10), 5213–5236. https://doi.org/10.1002/2015JD024473

Lauvaux, T., Gurney, K. R., Miles, N. L., Davis, K. J., Richardson, S. J., Deng, A., et al. (2020). Policy-Relevant Assessment of Urban CO2 Emissions. *Environmental Science & Technology*. https://doi.org/10.1021/acs.est.0c00343

Lee, T. (2013). Global Cities and Transnational Climate Change Networks. *Global Environmental Politics*, *13*(1), 108–127. https://doi.org/10.1162/GLEP_a_00156

Lei, R., Feng, S., Danjou, A., Broquet, G., Wu, D., Lin, J. C., et al. (2021). Fossil fuel CO2 emissions over metropolitan areas from space: A multi-model analysis of OCO-2 data

over Lahore, Pakistan. *Remote Sensing of Environment*, *264*, 112625. https://doi.org/10.1016/j.rse.2021.112625

Lemos, M. C., & Morehouse, B. J. (2005). The co-production of science and policy in integrated climate assessments. *Global Environmental Change*, *15*(1), 57–68. https://doi.org/10.1016/j.gloenvcha.2004.09.004

Lin, J. C., Mitchell, L., Crosman, E., Mendoza, D. L., Buchert, M., Bares, R., et al. (2018). CO2 and Carbon Emissions from Cities: Linkages to Air Quality, Socioeconomic Activity, and Stakeholders in the Salt Lake City Urban Area. *Bulletin of the American Meteorological Society*, *99*(11), 2325–2339. https://doi.org/10.1175/BAMS-D-17-0037.1

Liu, F., Duncan, B. N., Krotkov, N. A., Lamsal, L. N., Beirle, S., Griffin, D., et al. (2020). A methodology to constrain carbon dioxide emissions from coal-fired power plants using satellite observations of co-emitted nitrogen dioxide. *Atmospheric Chemistry and Physics*, *20*(1), 99–116. https://doi.org/10.5194/acp-20-99-2020

Liu, S., Valks, P., Pinardi, G., Xu, J., Chan, K. L., Argyrouli, A., et al. (2021). An improved TROPOMI tropospheric NO_2 research product over Europe. *Atmospheric Measurement Techniques*, *14*(11), 7297–7327. https://doi.org/10.5194/amt-14-7297-2021

Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., et al. (Eds.). (2021). Summary for Policymakers. In *Climate Change 2021: The Physical Science Basis*. Cambridge University Press. Retrieved from https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_SPM.pdf

Mayor Eric Garcetti. (2019). *L.A.'s Green New Deal: Sustainable City pLAn*. City of Los Angeles. Retrieved from https://plan.lamayor.org/sites/default/files/pLAn_2019_final.pdf

Mayor Martin J. Walsh. (2019). *City of Boston Climate Action Plan: 2019 Update*. City of Boston. Retrieved from https://www.boston.gov/sites/default/files/embed/file/2019-10/city_of_boston_2019_climate_action_plan_update_4.pdf

Moore III, B., Crowell, S. M. R., Rayner, P. J., Kumer, J., O'Dell, C. W., O'Brien, D., et al. (2018). The Potential of the Geostationary Carbon Cycle Observatory (GeoCarb) to Provide Multi-scale Constraints on the Carbon Cycle in the Americas. *Frontiers in Environmental Science*, *6*. https://doi.org/10.3389/fenvs.2018.00109

Mueller, K. L., Lauvaux, T., Gurney, K. R., Roest, G., Ghosh, S., Gourdji, S. M., et al. (2021). An emerging GHG estimation approach can help cities achieve their climate and sustainability goals. *Environmental Research Letters*, *16*(8), 084003. https://doi.org/10.1088/1748-9326/ac0f25

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 CO2 emissions from global fossil fuel emission data sets. *Journal of Geophysical Research: Atmospheres*, *118*(2), 917–933. https://doi.org/10.1029/2012JD018196

Nassar, R., Hill, T. G., McLinden, C. A., Wunch, D., Jones, D. B. A., & Crisp, D. (2017). Quantifying CO2 Emissions From Individual Power Plants From Space. *Geophysical Research Letters*, *44*(19), 2017GL074702. https://doi.org/10.1002/2017GL074702

Nathan, B., Lauvaux, T., Turnbull, J., & Gurney, K. (2018). Investigations into the use of multi-species measurements for source apportionment of the Indianapolis fossil fuel CO2 signal. *Elementa: Science of the Anthropocene*, *6*, 21. https://doi.org/10.1525/elementa.131

Oda, T., & Maksyutov, S. (2011). A very high-resolution (1 km×1 km) global fossil fuel CO2 emission inventory derived using a point source database and satellite observations of nighttime lights. *Atmos. Chem. Phys.*, *11*(2), 543–556. https://doi.org/10.5194/acp-11-543-2011

Oda, Tomohiro, & Maksyutov, S. (2015). ODIAC Fossil Fuel CO2 Emissions Dataset (ODIAC2017). *Center for Global Environmental Research, National Institute for Environmental Studies*. https://doi.org/10.17595/20170411.001

Oda, Tomohiro, Lauvaux, T., Lu, D., Rao, P., Miles, N. L., Richardson, S. J., & Gurney, K. R. (2017). On the impact of granularity of space-based urban CO_2 emissions in urban atmospheric inversions: A case study for Indianapolis, IN. *Elem Sci Anth*, *5*(0), 28. https://doi.org/10.1525/elementa.146

Oda, Tomohiro, Maksyutov, S., & Andres, R. J. (2018). The Open-source Data Inventory for Anthropogenic CO_2, version 2016 (ODIAC2016): a global monthly fossil fuel CO_2 gridded emissions data product for tracer transport simulations and surface flux inversions. *Earth System Science Data*, *10*(1), 87–107. https://doi.org/10.5194/essd-10-87-2018

Oda, Tomohiro, Bun, R., Kinakh, V., Topylko, P., Halushchak, M., Marland, G., et al. (2019). Errors and uncertainties in a gridded carbon dioxide emissions inventory. *Mitigation and Adaptation Strategies for Global Change*. https://doi.org/10.1007/s11027-019-09877-2

O'Dell, C. W., Connor, B., Bösch, H., O'Brien, D., Frankenberg, C., Castano, R., et al. (2012). The ACOS CO2 retrieval algorithm – Part 1: Description and validation against synthetic observations. *Atmos. Meas. Tech.*, *5*(1), 99–121. https://doi.org/10.5194/amt-5-99-2012

O'Dell, Christopher W., Eldering, A., Wennberg, P. O., Crisp, D., Gunson, M. R., Fisher, B., et al. (2018). Improved retrievals of carbon dioxide from Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm. *Atmospheric Measurement Techniques*, *11*(12), 6539–6576. https://doi.org/10.5194/amt-11-6539-2018

Olivier, J. G. J., & Janssens-Maenhout, G. (2015). *CO2 emissions from fuel combustion Part III: Total Greenhouse Gas Emissions; 2015 Edition* (pp. 1–43). International Energy Agency. Retrieved from http://www.pbl.nl/sites/default/files/cms/publicaties/pbl-2015-iea-co2-emissions-from-fuel-combustion-1971-2013-part-3-total-greenhouse-gas-emissions.pdf

Park, C., Jeong, S., Park, H., Yun, J., & Liu, J. (2020). Evaluation of the Potential Use of Satellite-Derived XCO2 in Detecting CO2 Enhancement in Megacities with Limited Ground Observations: A Case Study in Seoul Using Orbiting Carbon Observatory-2. *Asia-Pacific Journal of Atmospheric Sciences*. https://doi.org/10.1007/s13143-020-00202-5

Rayner, P. J., Raupach, M. R., Paget, M., Peylin, P., & Koffi, E. (2010). A new global gridded data set of CO2 emissions from fossil fuel combustion: Methodology and evaluation. *Journal of Geophysical Research: Atmospheres*, *115*(D19), D19306. https://doi.org/10.1029/2009JD013439

Reckien, D., Salvia, M., Heidrich, O., Church, J. M., Pietrapertosa, F., De Gregorio-Hurtado, S., et al. (2018). How are cities planning to respond to climate change? Assessment of local climate plans from 885 cities in the EU-28. *Journal of Cleaner Production*, *191*, 207–219. https://doi.org/10.1016/j.jclepro.2018.03.220

Reuter, M., Buchwitz, M., Schneising, O., Krautwurst, S., O'Dell, C. W., Richter, A., et al. (2019). Towards monitoring localized CO_2 emissions from space: co-located regional CO_2 and NO_2 enhancements observed by the OCO-2 and S5P satellites. *Atmospheric Chemistry and Physics*, *19*(14), 9371–9383. https://doi.org/10.5194/acp-19-9371-2019

Rosenzweig, C., Solecki, W., Hammer, S. A., & Mehrotra, S. (2010). Cities lead the way in climate-change action. *Nature*, *467*(7318), 909–911. https://doi.org/10.1038/467909a

Roten, D., Lin, J. C., Kunik, L., Mallia, D., Wu, D., Oda, T., & Kort, E. A. (2022). The Information Content of Dense Carbon Dioxide Measurements from Space: A High-Resolution Inversion Approach with Synthetic Data from the OCO-3 Instrument. *Atmospheric Chemistry and Physics Discussions*, 1–43. https://doi.org/10.5194/acp-2022-315

Seltenrich, N. (2014). Remote-Sensing Applications for Environmental Health Research. *Environmental Health Perspectives*, *122*(10), A268–A275. https://doi.org/10.1289/ehp.122-A268

Stocker, T. F., Qin, D., Plattner, G.-K., Alexander, L. V., Allen, S. K., Bindoff, N. L., et al. (2013). Technical Summary. In T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, et al. (Eds.), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 33–115). Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. https://doi.org/10.1017/CBO9781107415324.005

Suto, H., Kataoka, F., Kikuchi, N., Knuteson, R. O., Butz, A., Haun, M., et al. (2021). Thermal and near-infrared sensor for carbon observation Fourier transform spectrometer-2 (TANSO-FTS-2) on the Greenhouse gases Observing SATellite-2 (GOSAT-2) during its first year in orbit. *Atmospheric Measurement Techniques*, *14*(3), 2013–2039. https://doi.org/10.5194/amt-14-2013-2021

The Carbon Cycle. (2011, June 16). [Text.Article]. Retrieved June 19, 2022, from https://earthobservatory.nasa.gov/features/CarbonCycle

The Keeling Curve. (2022). Retrieved June 19, 2022, from https://keelingcurve.ucsd.edu

Ultee, L., Arnott, J. C., Bassis, J., & Lemos, M. C. (2018). From Ice Sheets to Main Streets: Intermediaries Connect Climate Scientists to Coastal Adaptation. *Earth's Future*, *6*(3), 299–304. https://doi.org/10.1002/2018EF000827

United Nations, Department of Economic and Social Affairs, Population Division. (2015). *World Urbanization Prospects: The 2014 Revision*.

United Nations, Department of Economic and Social Affairs, Population Division. (2019). *World Urbanization Prospects: The 2018 Revision (ST/ESA/SER.A/420)*. New York: United Nations.

Ussiri, D. A. N., & Lal, R. (2017). Introduction to Global Carbon Cycling: An Overview of the Global Carbon Cycle. In D. A. N. Ussiri & R. Lal (Eds.), *Carbon Sequestration for Climate Change Mitigation and Adaptation* (pp. 61–76). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-53845-7_3

Veefkind, J. P., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., et al. (2012). TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote Sensing of Environment*, *120*, 70–83. https://doi.org/10.1016/j.rse.2011.09.027

Wu, D., Lin, J. C., Fasoli, B., Oda, T., Ye, X., Lauvaux, T., et al. (2018). A Lagrangian approach towards extracting signals of urban CO_2 emissions from satellite observations of atmospheric column CO_2 (XCO_2): X-Stochastic Time-Inverted Lagrangian Transport model ("X-STILT v1"). *Geoscientific Model Development*, *11*(12), 4843–4871. https://doi.org/10.5194/gmd-11-4843-2018

Wu, D., Lin, J. C., Oda, T., & Kort, E. A. (2020). Space-based quantification of per capita CO2 emissions from cities. *Environmental Research Letters*, *15*(3), 035004. https://doi.org/10.1088/1748-9326/ab68eb

Wu, D., Liu, J., Wennberg, P. O., Palmer, P. I., Nelson, R. R., Kiel, M., & Eldering, A. (2022). Towards sector-based attribution using intra-city variations in satellite-based emission ratios between CO_2 and CO. *Atmospheric Chemistry and Physics Discussions*, 1–32. https://doi.org/10.5194/acp-2021-1029

Wu, L., Hasekamp, O., Hu, H., aan de Brugh, J., Landgraf, J., Butz, A., & Aben, I. (2019). Full-physics carbon dioxide retrievals from the Orbiting Carbon Observatory-2 (OCO-2) satellite by only using the 2.06 µm band. *Atmospheric Measurement Techniques*, *12*(11), 6049–6058. https://doi.org/10.5194/amt-12-6049-2019

Yang, E. G., Kort, E. A., Wu, D., Lin, J. C., Oda, T., Ye, X., & Lauvaux, T. (2020). Using Space-Based Observations and Lagrangian Modeling to Evaluate Urban Carbon Dioxide

Emissions in the Middle East. *Journal of Geophysical Research: Atmospheres*, *125*(7), e2019JD031922. https://doi.org/10.1029/2019JD031922

Ye, X., Lauvaux, T., Kort, E. A., Oda, T., Feng, S., Lin, J. C., et al. (2020). Constraining Fossil Fuel CO2 Emissions From Urban Area Using OCO-2 Observations of Total Column CO2. *Journal of Geophysical Research: Atmospheres*, *125*(8), e2019JD030528. https://doi.org/10.1029/2019JD030528