NORTHWESTERN UNIVERSITY

An Analysis of Urban Air Quality and Health Impacts Using High Resolution Chemical Transport Model Simulations

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Earth and Planetary Sciences

By

Anastasia Montgomery

EVANSTON, ILLINOIS

September 2023

Abstract

Air pollution is a pervasive environmental issue that has significant impacts on human health. Urban areas are particularly susceptible to high levels of air pollution due to concentrated emissions and populations. Urban air pollution has been linked to a range of adverse health effects, including respiratory diseases, cardiovascular diseases, and increased mortality rates. Pollution arises from the release of chemicals into the atmosphere, largely stemming from industrial activities, transportation, and energy generation. Understanding the sources, transport, and effects of air pollutants is essential, though limitations in computational capability and data availability has long hindered the identification of high-impact hotspots at health relevant scales. Recent developments into chemical transport models and non-regulatory monitoring techniques now play a crucial role, as these methods contribute to the identification of pollutant levels, the identification of vulnerable populations, and the support of evidence-based decision-making for effective pollution control strategies.

This dissertation investigates air quality over the Midwestern United states, centered over Lake Michigan-Chicago, using the latest developments in air quality modeling and observational tools. This research employs novel geospatial statistics and analytical methods to investigate concentrations of health-hazardous pollutants, namely nitrogen dioxide (NO_2), ozone (O_3), and fine particulate matter ($PM_{2.5}$).

Chapter 2 presents a high-resolution simulation of air pollution over the Southern Lake Michigan-Chicago region. Chemical transport models (CTMs) allow researchers to simulate how pollutants disperse and transform in the atmosphere. High-resolution simulations may enhance the accuracy of the model and increase its usefulness for public health assessments and policymaking. In this chapter, I examine the performance of a 1.3 km and 4 km simulation of pollutant concentrations over a four-month period (August 2018, October 2018, January 2019, and April 2019) in the Southern Lake Michigan-Chicago region. The 1.3 km simulation exhibits slightly better performance compared to the 4 km simulation. The study reveals distinct urban and rural patterns of pollution, with urban areas experiencing significantly

higher concentrations of NO_2 and $PM_{2.5}$ (20% – 60% higher than rural areas), while O_3 is simulated to be lower in urban areas (-6% compared to rural areas). Furthermore, the simulation highlights significant disparities in pollutant concentrations across neighborhoods in Chicago, with features such as highways contributing to substantial variations in pollution levels. This simulation provides valuable insights into the O_3 chemistry regime in Chicago, finding that the O_3 regime is transitional and VOC-limited, depending on the month of study. Overall, this research contributes to a better understanding of air pollution dynamics in the Southern Lake Michigan-Chicago region, shedding light on the spatial distribution of pollutants and their underlying chemistry.

Chapter 3 presents a hotspot analysis of Chicago air quality by using three novel air quality products: a low-cost sensor network, observations from a satellite instrument, and a chemical transport model. The study addresses the lack of intraurban data validation by assessing the spatial agreement of air pollution patterns across multiple high-resolution datasets and applies the hotspot analysis to make recommendations for researchers and policymakers. I apply a hotspot clustering algorithm, Getis Ord G_i^*, to identify areas of agreement and disagreement among the data products. The analysis reveals a Consensus hotspot on the West side of Chicago, indicating elevated pollution levels across different data products, wind directions, and seasons. This hotspot, predominantly inhabited by Hispanic and Latino people, requires urgent intervention as an environmental justice priority. Additionally, a medium-agreement hotspot identified by the low-cost sensors and satellite (i.e., Observational hotspot) highlights the need for additional regulatory monitoring in the affected community. Furthermore, a highway hotspot shows variations in NO_2 concentrations near recessed and elevated highways, a feature not captured in model simulations. These findings provide insights into areas of high pollution exposure, underscores the importance of targeted interventions, and recommends additional development of monitoring tools for improved air quality management in Chicago. By integrating this hotspot approach into air quality management frameworks, policymakers can develop targeted interventions and implement sustainable practices to mitigate the effects of air pollution on vulnerable populations.

In Chapter 4, the WRF-CMAQ simulation is used to investigate air pollution from an environmental justice perspective in Chicago. By integrating socioeconomic characteristics with air quality data, the study aims to uncover disparities in exposure and the contributing factors to environmental injustice. This research is the first to examine intraurban air pollution using a high-resolution CTM, addressing the limitations of coarser models that overlook significant pollution sources like highways within cities. This study is also the first comprehensive analysis to consider multiple pollutants, health outcomes, and their connection to environmental justice issues in Chicago. Findings reveal that pollutants and demographics exhibit high spatial variability, with no strong linear relationship between pollutants and racial, ethnic, or economic demographics. Although average exposure disparities in race and ethnicity are relatively small compared to baseline health rates, the Black population consistently experiences significantly higher rates of mortality, asthma, and pediatric asthma hospitalizations related to pollution. Racial and ethnic exposure disparities persist across income levels, indicating that income does not alter the relationship between pollution exposure or health outcomes. To address these inequalities, policies should consider both impact and exposure, as areas of health impact may not align with areas of high exposure. The study highlights the importance of incorporating health and exposure information in addressing air pollution injustices and advocates for equitable solutions.

In summary, this PhD research is the first to apply a 1.3 km high-resolution CTM to study Chicago air pollution. By applying this high-resolution data, this research also presents frameworks to isolate areas of outsized exposure and analyze dataset disagreement through hotspot analysis. Further, the data generated in this research is applied to analyze the link of exposure and public health at urban scales. By advancing CTM research, developing new methods of evaluation, and pioneering methods to investigate pollution inequalities at intraurban scales, this research ultimately supports the development of evidence-based pollution analysis and effective controls.

Acknowledgments

I sit in disbelief at the summit of this accomplishment. Reflecting on the many years of this journey, I am overcome with gratitude for all the people who have supported me through this. First and foremost, I want to thank my advisor, Dr. Daniel Horton, for his guidance, patience, and kindness. I am deeply grateful for your invaluable insights and your mentorship. I have learned so much from you and I hope to continue to make you proud.

I would also like to thank my committee, Dr. Jordan Schnell, Dr. Yarrow Axford, and Dr. Patricia Beddows. Jordan, thank you for hundreds of hours of training, I could not be where I am today without you. Yarrow and Trish, thank you for your unwavering support and believing in me. I want to thank my other mentors, Dr. Tracey Holloway and Dr. Madeleine Daepp. Tracey, you were the first person to believe in me and I am forever grateful. Madeleine, you have made me a better researcher.

Thank you to the CCRG lab, including Dr. Sara Camilleri, Victoria Lang, Dr. Howard Chen, Amy Rogin, Lucy Yang, Maxime Visa, Grace Hauser Chen, Dr. Irene Crisologo, and Dr. Ryan Hauser. The work we have done together has been fun, rewarding, and inspiring. I am honored to have been on the same team as you and I am excited for our future collaboration.

I also want to thank the rest of the Earth and Planetary Sciences Department. Thank you to the graduate students for making work fun. Thank you to Dr. Suzan van der Lee and Dr. Rosemary Bush for your lessons and encouragement. Thank you to Dr. Matt Hurtgen for your commitment to bettering our department and making me feel heard. I also want to extend a special thank you to my friends, Dr. Matt Selensky and Floyd Nichols, whose friendship has been a wonderful light in our dark office.

Most importantly, thank you to my friends and family. Thank you, Mom and Dad, for encouraging me to pursue happiness. Thank you to my sister, Alona, for everything. To the rest of my friends, thank you for your time, support, and your amazing talent to distract me. I am forever grateful for your reminders to celebrate the wins.

List of Abbreviations

EPA – Environmental Protection Agency
CAA – Clean Air Act
NAAQS – National Ambient Air Quality Standards
WRF-CMAQ – Weather Research Forecast and Community Multiscale Air Quality Model
SMOKE – Sparse Matrix Operating Kernel of Emissions
CTM – Chemical Transport Model
NO_2 – Nitrogen dioxide
O_3 – Ozone
$MDA8O_3$ – Maximum daily 8-hour average O_3
VOCs – Volatile organic compounds
CO_2 – Carbon dioxide
SO_2 – Sulfur dioxide
$PM_{2.5}$ – Particulate matter under 2.5 microns in diameter
CO – Carbon monoxide
EJ – Environmental justice

Dedication

For my friends, my family, my city,

and the air we share

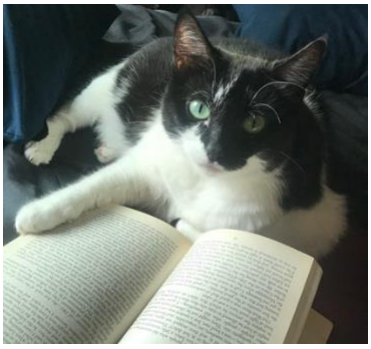


Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Worldwide, air pollution contributes to 3 – 7 million premature deaths annually (Pozzer et al., 2023; Vohra et al., 2021). This figure outpaces deaths from tobacco, vehicle collisions, and AIDS (WHO, 2023), resulting in air pollution contributing to the deaths of 1 in 5 people annually (Vohra et al., 2021). This environmental phenomenon further exacerbates existing social inequalities by contributing to outsized pollution exposure to minority communities (Clark et al., 2014; Demetillo et al., 2021; Goodkind et al., 2019; Nguyen & Marshall, 2018a; Tessum et al., 2021) and unjust health outcomes (Alexeeff et al., 2018; Buonocore et al., 2023; Castillo et al., 2021). These issues come to a head in urban areas, where a high density of population intermingles with a high density of emissions. Today, half of the global population lives in an urbanized area, with the figure projected to 68% by 2050, so it is anticipated that a larger portion of the global population will encounter urban pollutants in the near future (UN, 2019).

With the harms borne out of urban air pollution, the field of atmospheric chemistry was pulled from a theoretical to an applied science relatively quickly. The discovery of carbon dioxide (CO_2) and molecular oxygen came in the late 18th century, with atmospheric measurements taken nearly 100 years later (Brimblecombe, 1998). From 1920 - 1940, pioneering atmospheric chemists began to hypothesize the role of ozone (O_3) in the structure of the atmosphere (Gotz et al., 1934) and its mechanism for formation (Bates & Nicolet, 1950; Brewer, 1949; Chapman, 1930). After a century of increasing industrialization and emissions, air pollution became a public health crisis. In the United States, a smog event in Donora, PA in 1948 caused 20 – 40 mortalities and sustained health impacts (Townsend, 1950). A few years later, a smog event London, England in 1952 was directly linked to 400 – 1,000 mortalities over the course of 4 days (Wilkins, 1954).

After outcry to lethal smog events, the US government responded by enacting the Air Pollution Control Act in 1955. This law allocated federal funds to air pollution research (Moran, 1954) and was

followed by the first iteration of the Clean Air Act (CAA) to establish the regulation of air pollutants (Act, 1963). The CAA was updated in 1970, providing a major shift in the regulation to authorize the Environmental Protection Agency (EPA) to regulate criteria pollutants by setting National Air Quality Standards (NAAQS, Act, 1970). With the power of the CAA and EPA, a national monitoring network was established to provide continuous ambient monitoring of pollutants, officially cementing the mitigation of air pollution as a national priority.

The NAAQS regulate the 6 criteria pollutants: ground-level ozone (O_3), nitrogen dioxide (NO_2), fine particulate matter (PM_{10}, $PM_{2.5}$), carbon monoxide (CO), sulfur oxides (SOx, SO_2), and lead (Pb). The CAA mandates that concentrations for the NAAQS are set to protect human health with adequate margin of safety (Act, 1970). With the monitoring capabilities to enforce NAAQS, the evidence linking air pollution and negative health outcomes have grown stronger. Here, I highlight some of the pathways and detrimental health effects of O_3, NO_2, and $PM_{2.5}$. First, O_3 is a highly reactive chemical which interacts and dissolves into the lung tissues (Mudway, 2000). As such, long-term and short-term exposure to O_3 is associated with increased risk of asthma, bronchitis, inflammation, hospitalizations, and premature mortality (Anderson et al., 2013; Chen et al., 2007; Hoek et al., 2013). Second, NO_2 is also a respiratory irritant which interacts and damages the lung tissue (Persinger et al., 2002). Long-term and short-term exposure to NO_2 is associated with increased risk of asthma, hospitalizations, respiratory infections, and premature mortality (Huangfu & Atkinson, 2020). Finally, $PM_{2.5}$ is associated with cardiovascular and respiratory diseases, as $PM_{2.5}$ can be deposited into the lungs and absorbed into the blood stream (Sørenseng et al., 2003). As such, exposure to $PM_{2.5}$ is linked to increased risk of asthma, heart attacks, hospitalizations, decreased lung function, and premature mortality (M. Franklin et al., 2007).

After the initial thrust of health-related air pollution research borne out of EPA regulations, the modern age of air quality sciences began with the advent of computers. New applications of computational power included remote-sensing tools which were onboarded onto satellite payloads. In

1978, the first space-borne instrument to measure atmospheric makeup, the Total Ozone Monitoring Spectrometer (TOMS), was launched by the National Aeronautics and Space Administration (NASA). This geo-orbiting instrument took daily snapshots of O_3 and SO_2 in the atmosphere at a spatial resolution of 1092 x 1108 km (Bowman & Krueger, 1985). The instrument was launched to collect data for meteorological, air pollution, and oceanographic scientists. Notably, TOMS was instrumental in identifying the stratospheric O_3 hole over the Antarctic (Arrigo, 2003; Margitan et al., 1995). This discovery spurred international cooperation to reduce the use of chlorofluorocarbons (CFCs), the chemical compound which was found to deplete the O_3 concentrations (Rowland, 1990). With the success of TOMS, dozens of remote-sensing instruments have been launched into space (Martin, 2008), including those which focus on tropospheric pollutants to support CTM assessment and human health assessments.

Further computational development spurred innovation into numerical modeling tools for application in atmospheric sciences. The first models of atmospheric flows were developed in the late 1890s, named the three primitive equations: 1) continuity, to represent the conservation of mass, 2) conservation of momentum, to describe hydrodynamical flow on the surface of a sphere, and 3) thermal energy equations, to reflect heat sources in the system (Jacobs, 1999). These primitive equations form the basis of numerical weather modeling today. The first general circulation model (GCM) to incorporate oceanic and atmospheric processes was developed by the National Oceanic and Atmospheric Administration (NOAA) in the late 1960s (NOAA, 2022), which was followed by the first global, 3-dimensional model of tropospheric chemistry (Zimmerman, 1984). To develop global circulation models with chemistry, GCMs incorporate the primitive equations with additional variables to capture the fluxes, deposition, and chemical production/loss of atmospheric constituents. As computers have continued to increase in power, modern dynamic chemical transport models (CTMs) include detailed chemistry, advection, hydrology, and sub-grid processes (Byun & Schere, 2006; Henze et al., 2007; Wong et al., 2011). Modern CTMs are used to inform increasingly complex earth-system dynamic problems and are considered the state-of-the science method to address questions in atmospheric chemistry.

The latest advancements in CTMs continue to exploit the computational efficiency gains of the modern age. One expanded capacity due to higher computational power are higher-resolution model simulations. Operational CTMs used to create air quality forecasts run at a horizontal grid resolution between 36 x 36 km^2 (21 x 21 mi^2) or 12 x 12 km (7.4 x 7.4 mi^2) (Kang et al., 2010). Higher-resolution 'hindcasts', which do not need to be quickly spun up for operational use, have been downscaled to approximately 1 x 1 km^2 resolution (Torres-Vazquez et al., 2022). Higher resolution air quality simulations are better at predicting air quality compared to their coarse counterparts (Fountoukis et al., 2013; Gan et al., 2016) and provide enhanced spatiotemporal resolution to better reflect exposure and health disparities (Mohegh et al., 2021; Thompson & Selin, 2012).

Finally, the latest advancement in computational networks have benefitted new monitoring of air pollution through the internet of things (IoT), the umbrella term to describe the physical sensors and software providing communication about the environment (Suresh et al., 2014). Originally developed to enable the remote monitoring of objects, urban planners and city officials have taken advantage of the IoT networking capabilities to monitor traffic, heat, and pollution for use in policy development (Theodoridis et al., 2013). Decreasing costs in sensor parts and faster internet connectivity have supported the proliferation of low-cost air quality sensor networks (Kumar et al., 2015). These monitoring networks have been supported by local governments and executed through partnerships with scientists (Catlett et al., 2017), non-profit organizations (Peters et al., 2022), and corporations (Daepp et al., 2022). Some low-cost sensor projects have no central governance, such as PurpleAir, which is an internet-connected monitoring network driven by citizen scientists (Barkjohn et al., 2021). Low-cost sensor networks are particularly attractive for cities, as the networks can provide real-time information about the air pollution at spatial resolutions which satellite instruments and operational forecasts do not currently match.

Regardless, ground-level pollution measurements from EPA reference-grade monitors are the gold-standard method to capture the 'nose-level' exposure of pollutants for policy and health impact studies, even though these measurements are spatially and (occasionally) temporally limited. By

combining reference-grade measurements with supplemental information from satellite instruments, low-cost sensors, and air quality models, higher-resolution air quality information can be achieved and applied for use in policy-making and health-impact assessments (Ryan & LeMasters, 2007).

The push for higher-resolution air quality data continues to be motivated by the original motivation for the CAA: the devastating health impacts of air pollution. With better information on health impacts and environmental emissions, clusters of cancer (Thun & Sinks, 2004) and asthma (Etzel, 2003) have been directly linked to local emission sources. In the US, areas experiencing high rates of cancer, asthma, and cardiovascular disease are also overwhelmingly people of color (POC, the people who identify as non-White Hispanic and Latino, Black, Asian, Pacific Islander, and Indigenous). Poor health incidence rates in POC communities are in part are due to lower health care access (Makri & Stilianakis, 2008) and co-morbidities (Seposo et al., 2020; Sinden & Stockley, 2010; Sørensen et al., 2023) which elevate health risks , however, these elevated risk factors are influenced themselves by systemic racism (Feagin & Bennefield, 2014; Johnson, 2020). Systemic racism is the way in which polices and practices in a society or organization support a continued unfair treatment based on race and support an unfair advantage to others (Carmichael & Hamilton, 1967; Paradies et al., 2015). As systemic racism influences health outcomes, scholars have also hypothesized that systemic racism also influences the kind of environment which POC have access to live in. For example, a discriminatory federal mortgage appraisal policy by the Home Owners Lending Corporation (HOLC) limited the housing options of POC in the 1940s by declaring high POC neighborhoods as unworthy of investment, a practice called redlining (Nelson et al., 2021). Current air pollution concentrations in historically redlined areas show that redlined areas 1) have high POC populations and 2) experience higher $PM_{2.5}$ concentrations (Lane et al., 2022). This example encapsulates that regardless of racist intent, the environmental inequalities disproportionally effect POC: numerous studies have found that POC are more exposed to polluting emission sources (Goodkind et al., 2019; Tessum et al., 2019, 2021), experience higher pollution exposure levels (Clark et al., 2022a; Morello-Frosch & Jesdale, 2006), and experience more health impacts due to air pollution

(Chambliss et al., 2021; Southerland et al., 2021). The environmental justice (EJ) movement has come as a response to the systemic environmental racism and socioeconomic disenfranchisement. The EJ movement calls for the liberation of POC from the institutional factors which cause POC higher environmental risks (Mohai et al., 2009).

A key tenant of EJ is to address the injustice caused by environmental risks through equal enforcement of current policies and the development and enforcement of policies which lessen inequalities. Previous studies on racial exposure inequalities look at national-scale disparities (Clark et al., 2014, 2022a; Morello-Frosch & Jesdale, 2006; Tessum et al., 2021), though urban areas are more polluted and have more POC populations (Pew Research, 2018). National exposure disparity studies will control for 'urban' vs. 'rural' demographic differences by computing inequalities relative to average state or county-level demographic exposures, but a single county or state may include urban, suburban, and rural development levels. Few studies quantify the systemic air pollution exposure disparities within city-limits themselves. Studies which leverage municipal boundaries may better quantify intraurban disparities, as this boundary can inform local city policy to mitigate the urban emission sources which contribute to intraurban disparities, targeting the local factors which contribute to inequalities and expanding analysis beyond the national trend of higher pollution in urban areas. With the development of high-fidelity and high-resolution air quality datasets, air pollution monitoring capabilities have caught up to the demand for EJ policies to directly target urban disparities in unequal pollution exposure and impacts.

In this dissertation, I use a high-resolution chemical transport model to analyze air quality over the Southern-Lake Michigan-Chicago region (Figure 1.1). The following experiments in this dissertation use a coupled Weather Research Forecast Community Multiscale Air Quality Modeling System (WRF-CMAQ, Wong et al., 2011) simulation to analyze criteria pollutants. WRF-CMAQ creates numerical simulations of air quality by using information about the emissions, chemistry, and physics of the atmosphere. The model relies on first principles to predict concentrations of airborne gases, particles, and their compositions. To set up our model framework, custom meteorologically informed emissions based

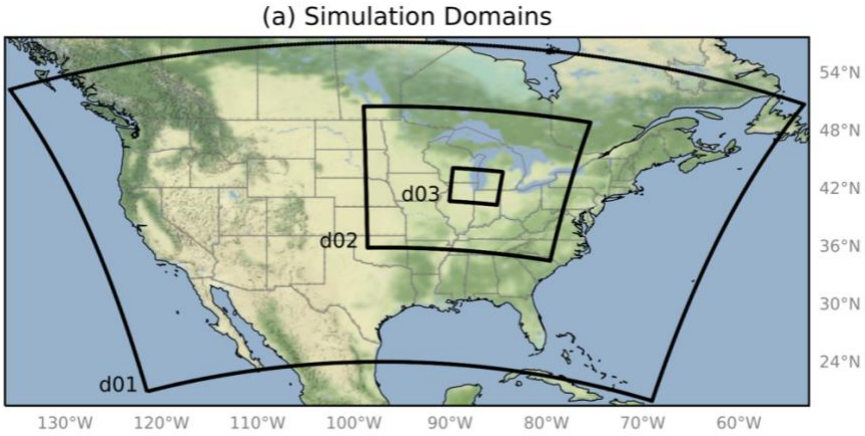


Figure 1.1 Model set up for WRF-CMAQ simulation described in this dissertation. (a) Shows the nested simulation configuration, with the d03 (1.3 km) domain centered around Chicago. For additional information on model configuration, see Chapter 2.

on the 2016 NEI (Eyth et al., 2019) were processed using the Sparse Matrix Operating Kernel of Emissions (SMOKE, B.H. Baek & Seppanen, 2018) and meteorology from a standalone WRF simulation. The spatial surrogates used to set up emissions at the 1.3 km simulation were provided by the Lake Michigan Air Directors Consortium (LADCO, 2019) and the 4 km surrogates were provided by CMAS (CMAS, 2018). With these customized inputs, the following experiments are the first to investigate air quality in Chicago using a high resolution CTM.

This dissertation is comprised of five chapters, with an introduction (1), three results chapters (2-4) and a conclusion (5). Each of the results chapter in this dissertation uses the WRF-CMAQ model with the aim to 1) characterize and validate urban air pollution spatial heterogeneity, 2) intercompare CTMs output with low-cost sensor measurements and satellite observations and 3) use the model to gain insights on how air pollution interacts with populations and impacts health. Appendices 1 – 3 provide supplemental information for Chapters 2 – 4, while Appendices 4 – 8 provide references and the abstracts of projects which were published over the course of the PhD to which I contributed to.

In Chapter 2 of this dissertation, we apply the WRF-CMAQ model to characterize air quality in the Southern Lake Michigan region. This study addresses a gap in the literature: few high-resolution (~1 km x 1 km horizontal resolution) characterizations of urban air quality had been performed (as of

publication, only Torres-Vazquez et al., 2022 over the New York City Metropolitan area). This research is the first to apply a CTM at this spatial resolution in Chicago, elucidating how enhanced horizontal resolution can alter model performance and how the finer simulation of pollutants can be more representative of emission sources, which is particularly important for near-highway neighborhoods. The chapter focuses on validating the simulation of key pollutants, namely nitrogen dioxide (NO_2), ozone (O_3), and particulate matter ($PM_{2.5}$), and then analyzing the output in the context of the urban environment. Chicago, a city in nonattainment for O_3 levels, has long been studied to identify the dominant pathways for O_3 formation, and in part, this study is motivated by conflicting chemical regimes identified for the formation of O_3 over Chicago. O_3 is formed from the emission of primary pollutants (VOCs and NOx) and the interaction with the environment (radiation and water vapor). The regime of O_3 formation, be it NOx or VOC-limited, has been widely studied through there is no consensus to which chemical regime Chicago follows (T. Foley et al., 2011; Jin et al., 2017; Jing & Goldberg, 2022). Further, given the high-resolution nature of this study, this study is motivated by being the first to be able to characterize neighborhood-scale variation in pollutants within Chicago using a CTM. This study was published in the *Journal of Geophysical Research: Atmospheres* in March 2023.

The final takeaways of Chapter 2 are that:

2.1 The 1.3 km WRF-CMAQ simulation outperforms the 4 km simulation in simulating most criteria pollutants.

2.2 Simulated spatial pollutant patterns show distinct urban-rural footprints, with urban NO_2 and $PM_{2.5}$ 20% – 60% higher than rural, and urban O_3 6% lower.

2.3 Pollutants within Chicago are simulated to have substantial concentration disparities across neighborhoods.

2.4 Fine-scale emitting features, such as highways, are more apparent in the fine-resolution simulation which results in higher concentrations of NO_2 and $PM_{2.5}$ in near-highway neighborhoods.

2.5 The O_3 chemical regime in Chicago is NOx-limited in August 2018, VOC-limited in October 2018, January 2019, and April 2019.

In Chapter 3 of this dissertation, we build off the validation and characterization of Chicago air pollution using WRF-CMAQ by using NO_2 observations from a low-cost sensor network and satellite instrument. This study addresses the growing research gap about data validation at intraurban scales, as few reference monitors are available to test the spatial validity of air pollution patterns from multiple high-resolution datasets. Previous studies have used a combination of low-cost sensors, satellite observations, and/or CTM output to create high-resolution analysis of urban air quality (Cordova et al., 2021; Ryan & LeMasters, 2007), though few studies assess the data validity of each tool. I use a hotspot clustering algorithm, Getis Ord Gi* (Getis & Ord, 2010) to assess agreement across data products. To do this, I run the hotspot clustering algorithm for each air quality data product and identify overlapping hotspots. In identifying areas with 3/3 hotspot overlap (i.e.: Consensus hotspot), I highlight the area in Chicago which has strong evidence for elevated pollution and would therefore need intervention most urgently. In areas with 2/3 dataset agreement (i.e.: Medium-Agreement Hotspots), I investigate the underlying reasons which may contribute to dataset mismatch and suggest additional development into each tool. This study is under review in *Environmental Research Letters*, as of July 2023.

The final takeaways of Chapter 3 are that:

3.1 By analyzing the area of intersection across Getis-Ord Gi* hotspots, we can identify areas with high data agreement (Consensus hotspot) for intervention, or medium agreement, to target the underlying reasons for air quality product disagreement.

3.2 While the spatial patterns of NO_2 identified by each data product shows different formations, the Consensus hotspot on the West side of Chicago is consistently identified across products, wind directions, and seasons.

3.3 With high agreement across air quality data products, we find that between 332,000 – 501,000 Chicago residents are exposed to significantly elevated levels of NO_2 pollution,

with this area being majority Hispanic and Latino, further highlighting this area for intervention as an EJ priority.

3.4 The medium-agreement hotspot identified in the low-cost sensors and satellite (Observational Hotspot) but not model highlights an EJ community which could be better served with additional regulatory monitoring.

3.5 The medium-agreement hotspot over the highway (Highway hotspot) highlights how sensors near recessed highways report in higher NO_2, while sensors near elevated highways report higher NO_2 in August 2022, which is a feature not captured in model simulations.

In Chapter 4, I use the WRF-CMAQ simulation to analyze air pollution through an environmental justice lens. By integrating the socioeconomic characteristics of Chicago with the air quality data from the simulation, I aim to uncover potential disparities in exposure and the underlying factors contributing to environmental injustice. This study is the first to use a high-resolution CTM to investigate air pollution at the intraurban scale, which is important as Chapter 2 identifies how emissions sources, such as highways, are significant pollution sources in Chicago, and these sources are smoothed out by coarser models. This study is also the first to comprehensively look at multiple pollutants, multiple health outcomes, and connect the impacts of intraurban air pollution to environmental justice issues in Chicago. We utilize this analysis to emphasize that incorporating health and exposure information is a more effective approach to addressing air pollution injustices compared to policies that focus on only one aspect. Overall, by employing an environmental justice perspective, I aim to shed light on the intersectionality of social and environmental issues, ultimately advocating for more equitable and sustainable solutions to air pollution challenges. This study is being prepared for submission.

The final takeaways of Chapter 4 are that:

4.1 The pollutants and demographics have high spatial variability, but no pollutant has a strong linear relationship to any racial, ethnic, or economic demographic, so areas of high pollutants affect multiple demographics and incomes.

4.2 The average exposure disparities across Chicago are relatively small compared to the baseline health incidence rates. As such, we find small inequalities in exposure with respect to race and ethnicity: compared to the Chicago average, the Hispanic and Latino population are exposed to 2 – 4% higher NO_2 and $PM_{2.5}$ and the Black population is exposed to 1% higher $MDAO_3$. In contrast, the Black population consistently experiences 30 – 45% higher mortality, asthma, and pediatric asthma hospitalization rates. As such, the Black population experiences the worst pollution-related health outcomes, with 38 – 40% higher rates of NO_2, $PM_{2.5}$, and $MDAO_3$-attributable mortality, asthma, and pediatric asthma hospitalization rates.

4.3 At every income level, we find the same racial and ethnic exposure disparities, indicating that income does not change the relationship between pollution exposure. The Hispanic and Latino population experience higher NO_2 and $PM_{2.5}$ exposure at all income levels, the Black population experiences higher $MDAO_3$ exposure at all income levels. Pollution-attributable health outcomes decrease slightly with higher incomes across all racial and ethnic groups, though the racial and ethnic segregation across health outcomes exist at all income levels with the Black population experiencing the worst health outcomes.

4.4 Given that areas of health impact are not necessarily located in areas of high exposure, I suggest developing policies which balance impact and exposure to best reduce inequalities in both spheres. I show that by reducing pollution in areas with the highest health incidence rates results in increased exposure inequalities for the Hispanic and Latino population, though greatly benefits the Black population. In reducing pollution in areas with the highest concentrations, the attributable health inequalities for the Black

population increase but the exposure inequalities decrease. By isolating areas with the highest exposure and mortality rates, we decrease inequalities in exposure for the Hispanic and Latino population and decrease health impacts in the Black community, though to lesser degree than isolating one outcome.

To summarize, this dissertation research is the first to apply a high-resolution WRF-CMAQ model at the neighborhood scale over Chicago to investigate the model validity, chemical speciation, and multi-species impacts of exposure. This research also provides new methodological frameworks to process and validate high-resolution air quality data products when regulatory monitoring is not available for validation. Moreover, by applying this new data to address environmental injustices, we uncover the nuanced exposure and health disparities caused by Chicago air pollution and develop new recommendations on how to address these injustices.

Chapter 2

Simulation of Neighborhood-Scale Air Quality with Two-Way Coupled WRF-CMAQ Over Southern Lake Michigan-Chicago Region

This chapter is published as: Montgomery, A., Schnell, J. L., Adelman, Z., Janssen, M., & Horton, D. E. (2023). Simulation of neighborhood-scale air quality with two-way coupled WRF-CMAQ over southern Lake Michigan-Chicago region. Journal of Geophysical Research: Atmospheres, 128, e2022JD037942. https://doi. org/10.1029/2022JD037942

Abstract

The southern Lake Michigan region of the United States, home to Chicago, Milwaukee, and other densely populated Midwestern cities, frequently experiences high pollutant episodes with unevenly distributed exposure and health burdens. Using the two-way coupled Weather Research Forecast and Community Multiscale Air Quality Model (WRF-CMAQ), we investigate criteria pollutants over a southern Lake Michigan domain using 1.3 and 4 km resolution hindcast simulations. We assess WRF-CMAQ's performance using data from the National Climatic Data Center and EPA Air Quality System. Our 1.3 km simulation slightly improves on the 4 km simulation's meteorological and chemical performance while also resolving key details in areas of high exposure and impact, i.e., urban environments. At 1.3 km, we find that most air quality-relevant meteorological components of WRF-CMAQ perform at or above community benchmarks. WRF-CMAQ's chemical performance also largely meets community standards, with substantial nuance depending on the performance metric and component assessed. For example, hourly simulated NO_2 and O_3 are highly correlated with observations ($r > 0.6$) while $PM_{2.5}$ is less so ($r = 0.4$). Similarly, hourly simulated NO_2 and $PM_{2.5}$ have low biases (<10%), whereas O_3 biases are larger (>30%). Simulated spatial pollutant patterns show distinct urban-rural footprints, with urban NO_2 and $PM_{2.5}$ 20-60% higher than rural, and urban O_3 6% lower. We use our 1.3 km simulations to resolve high-pollution areas within individual urban neighborhoods and

characterize seasonal changes in O_3 regimes across tight spatial gradients. Our findings demonstrate both the benefits and limitations of high-resolution simulations, particularly over urban settings.

Plain Language Summary

In this study we use an air quality model to simulate air pollution at very fine spatial scales over a central midwestern U.S. domain that includes Chicago, IL and Milwaukee, WI. We assess our model's performance relative to meteorological and air quality observations and then characterize the spatial patterns of modeled pollutants. We find large differences in air pollution between urban and rural settings. Because our model operates at fine spatial scales, we are also able to discuss differences in air pollution in different neighborhoods in individual cities. In Chicago, we find elevated pollution near highways and in south and west side neighborhoods, findings that are consistent with previous reports of disparate air quality related health impacts.

2.1 Introduction

Exposure to poor air quality in the U.S. has been found to exacerbate respiratory diseases (Kurt et al., 2016), drive disparate health burdens in racial minority populations (Jbaily et al., 2022; Tessum et al., 2021), and contribute to ~100,000 premature deaths annually (Goodkind et al., 2019). Given the substantial public health burden associated with exposure to poor air quality, it is essential to resolve pollutant concentrations at high spatiotemporal resolutions. Pollutant exposure in high population settings, i.e., urban environments, can vary widely, which can contribute to disparities in health outcomes on a neighborhood-by-neighborhood basis across individual cities (Alexeeff et al., 2018; O'Leary & Lemke, 2014; Southerland et al., 2021). Determining the relationship between heterogeneous pollutant exposure and disparate health effects is challenging given observational constraints. For example, regulatory-grade air quality monitoring stations are relatively sparse and therefore spatial coverage is limited, particularly in urban settings. Observing platforms that do have better spatial coverage, e.g., remote sensed satellite observations, often have temporal limitations such as making only one observation

a day in the case of polar orbiters (Penn & Holloway, 2020). Given the need to resolve pollutants across impact-relevant scales (Clark et al., 2022b), researchers often turn to physics- and chemistry-based Chemical Transport Models (CTMs) which allow for the spatial heterogeneity of pollutants to be estimated at high temporal resolutions in areas that are otherwise unmonitored (J. Hu et al., 2019).

State-of-the-science CTMs resolve pollutants at ever-increasing spatiotemporal scales. For example, the Community Multiscale Air Quality modeling system (CMAQ; Byun & Schere, 2006) was developed by the U.S. Environmental Protection Agency (EPA) to study the complex interactions of pollutants and meteorology and increase our understanding of atmospheric processes. Over time, the spatial resolution of CTMs like CMAQ has increased as computational costs decrease and spatially-defined inputs are resolved at finer scales (Gan et al., 2016). Higher resolution CTM studies have the potential to simulate more accurate meteorology, emissions, and pollutant concentrations than coarser resolution models (Fountoukis et al., 2013; Gan et al., 2016; Torres-Vazquez et al., 2022). However, some simulated meteorological and chemical variables may show lower model performance at finer resolutions because of incomplete characterizations of complex terrain and limitations in the planetary boundary layer formation (Tran et al., 2018; Zhang et al., 2014). Notably, epidemiological studies have identified the use of high-fidelity, high-resolution air quality characterizations as better at capturing pollution-related health impacts (Jiang & Yoo, 2018; Southerland et al., 2021; Thompson et al., 2014).

Given the potential benefits of high resolution CTM studies, here we utilize WRF-CMAQ to characterize pollutant concentrations over a southern Lake Michigan domain, a region in the central midwestern U.S., which includes the major population centers of Chicago, IL and Milwaukee, WI. Previous modeling studies have focused on this region due to the atmospheric complexities associated with Lake Michigan and high O_3 pollution in the region (Abdi-Oskouei et al., 2020; Dye et al., 1995; T. Foley et al., 2011). By and large, air quality in this region has been improving due to emission controls and the outsourcing of industry and manufacturing (Jing et al., 2014). However, pockets of poor air

quality persist, particularly in and downwind of urban centers like Chicago, which has been in EPA 8-hour O_3 National Ambient Air Quality Standards (NAAQS) non-attainment status since 2004 (EPA, 2022).

From a regional perspective, poor air quality in Midwestern summers is often associated with warm stagnant air masses (Jing et al., 2017; Schnell & Prather, 2017; Tai et al., 2010), while near-surface winter pollution is largely restricted to particulate matter accumulations associated with temperature inversions (Tran et al., 2018). However, at local scales, local geography, meteorology, and emissions often play a synergistic role. For example, in Chicago, the coastal geography, micro-meteorology, and high-emitting urban footprint combine to create an active atmospheric regime that often facilitates accumulation of primary pollutants and/or the precursors of secondary pollutants. Indeed, Chicago's O_3 NAAQS nonattainment status is a direct result of interacting emissions, geography, and meteorology – particularly the interaction of precursor emissions with Lake Michigan's lake breeze. Because the formation of O_3 is generally dependent on the ratio of precursor emissions, i.e., nitrogen oxides (NOx) and volatile organic compounds (VOCs), the EPA has restricted NOx emissions (EPA, 2019). However, previous studies have found that Chicago is in a transitional or VOC-limited regime (Duncan et al., 2014; Jin et al., 2017; Jing et al., 2014) – suggesting a limitation to the efficacy of emissions controls that only consider NOx and do not also reduce VOCs.

Similar to O_3, the concentration of NO_2 depends on meteorological factors such as winds and temperature (Harkey et al., 2015), however given the relatively short lifetime of NO_2, proximity to emission sources also plays a substantial role, particularly in the development of intra-urban heterogeneities. In satellite analyses, Chicago appears as a large source of NO_2 pollution to the greater Midwest (Goldberg et al., 2021), a factor that contributes to the formation and elevated concentration of downwind/rural O_3. In addition, NO_2 can be a precursor to $PM_{2.5}$ formation through the oxidation of NO_2 to nitrate. While Chicago is currently in compliance with $PM_{2.5}$ standards, $PM_{2.5}$ has previously been

found to be elevated in comparison with rural areas due to the confluence of transportation, energy generation, and industrial emissions, as well as atmospheric transport, and secondary formation processes (Zhang et al., 2014).

Given the above complexities, we hypothesize that the characterization of air quality in the southern Lake Michigan-Chicago region would benefit from resolving fine scale interactions and impacts of local geography on atmospheric chemistry and meteorology. As such, here we use a high-spatial resolution numerical model that includes atmospheric meteorology, chemistry, and components of their interactions and feedbacks, i.e., the two way-coupled Weather Research and Forecasting-Community Multiscale Air Quality modeling system (WRF-CMAQ). These high-resolution simulations are made possible by the Lake Michigan Air Directors Consortium (LADCO) spatial surrogate dataset (LADCO, 2022), which defines the allocation of county-level emission information to a 1.3 km grid. We characterize our air quality estimates as "neighborhood-scale", because our 1.3 km simulations have an average of ~5 grid cells for each of the 77 community areas that comprise Chicago (Chicago Data Portal, 2022). LADCO spatial surrogates are used in the Sparse Matrix Operating Kernel of Emissions (SMOKE) processing system (B.H. Baek & Seppanen, 2018) with the U.S. EPA 2016 Beta modeling platform (Eyth et al., 2019) to produce emission data for our 1.3 km grid. We use this emissions dataset in WRF-CMAQ to simulate 4 months representative of the 4 meteorological seasons and characterize pollutant concentrations over a central-Midwestern and Chicago-centric domain.

2.2 Methods

2.2.1 CTM Simulations and Domains

We performed CTM simulations using the two-way coupled Community Multi-scale Air Quality (CMAQ, v5.2; Byun & Schere, 2006) and Weather Research and Forecasting (WRF, v3.8; Skamarock et al., 2008) modeling system (WRF-CMAQ; Wong et al., 2012). The two-way configuration of WRF-CMAQ allows feedbacks between simulated aerosols and WRF's shortwave radiation scheme. To

perform WRF-CMAQ simulations, we follow the methodology of Wong et al. (2012): (1) we produce dynamically downscaled meteorology with stand-alone WRF simulations, (2) we then use the stand-alone WRF output to create meteorologically-informed emissions data using the Sparse Matrix Operating Kernel of Emissions (SMOKE), and lastly (3) we run the coupled WRF-CMAQ model, incorporating the meteorologically-informed SMOKE emissions data.

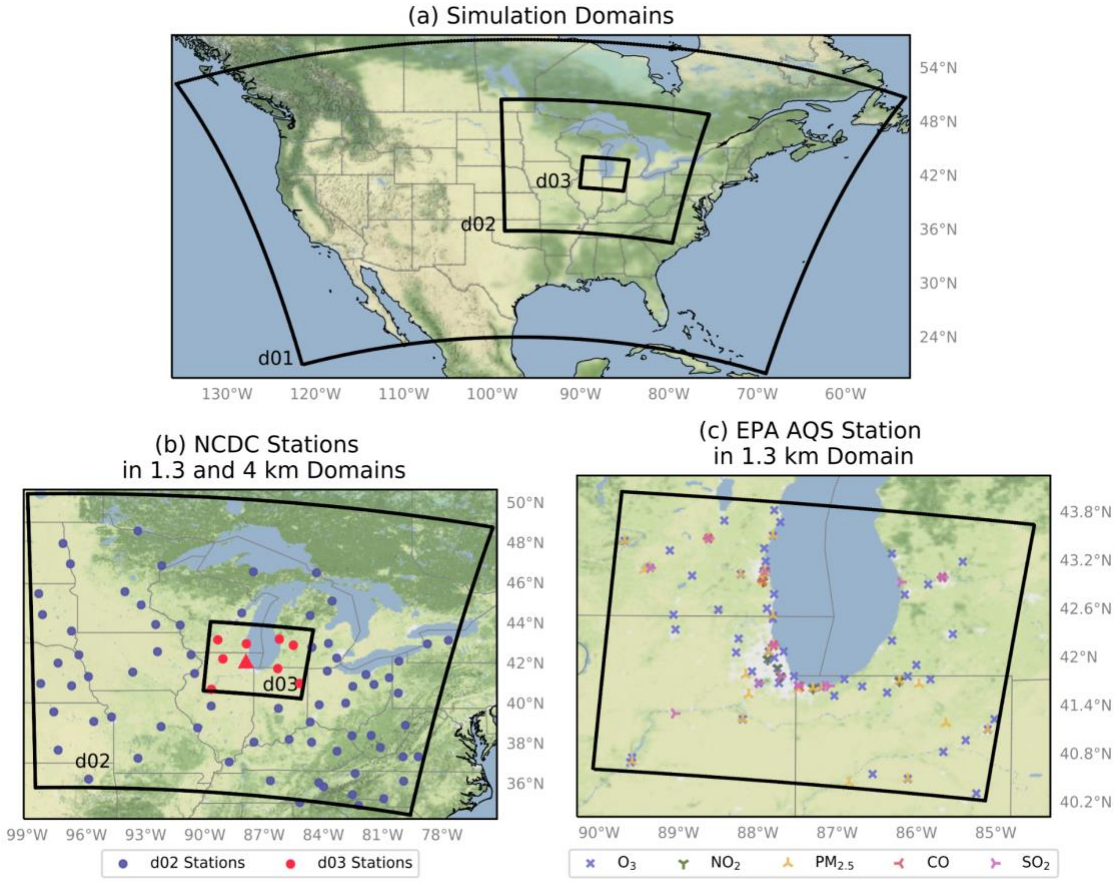


Figure 2.1 Nested model domains and observation sites. (a) Spatial footprint of nested model domains. We perform stand-alone WRF simulations at 12 km horizontal resolution in the d01 domain. Dynamically downscaled two-way coupled WRF-CMAQ simulations are performed in the d02 and d03 domains at 4 km and 1.3 km horizontal resolutions. (b) We use Local Climatological (LCD) Data from the National Climatic Data Center (NCDC) for meteorological validation of the 4 km (d02, purple) and 1.3 km (d03, pink) domains. The station located at Chicago-O'Hare is indicated by a triangle in the center of the d03 domain in panel (b). (c) EPA Air Quality System (AQS) stations are used to validate simulated pollutants. AQS stations are colored by chemical species and several AQS sensors for different chemical species are co-located, thus overlap on the map.

To generate boundary and initial conditions and facilitate the production of meteorologically-informed emissions data, we first perform a stand-alone WRF simulations to generate three-dimensional

meteorology in nested domains with 12 km (CONUS; d01), 4 km (Midwest; d02), and 1.3 km (southern Lake Michigan; d03) resolutions (Figure 2.1a). We use a 10-day spin-up period and simulate four months – August 2018, October 2018, January 2019, and April 2019 – using a 60, 20, and 6-second timestep for the 12, 4, and 1.3 km domains respectively. To allow soil moisture and soil temperature variables to reach a state of statistical equilibrium with observational constraints, we turn on the soil moisture initialization option during the 10 day spin-up (Pleim & Xiu, 2003). We run WRF with 35 vertical layers from the surface to 30 hPa with a lowest model level thickness of ~20 m. Initial conditions and 3 hourly lateral boundary conditions for the 12 km domain are sourced from the North Atlantic Regional Reanalysis (NARR; Mesinger et al., 2006). Simulated WRF meteorology is nudged toward reanalysis using Four-Dimensional Data Assimilation (FDDA) above the boundary layer, using nudging coefficients from LADCO (2022) and Otte (2008) for temperature and wind (3×10^{-4} s^{-1}, 1×10-5 s^{-1} , and 1×10-6 s^{-1} for 12 km, 4 km, and 1.3 km) and the water vapor mixing ratio (1×10^{-5} s^{-1}, 1×10^{-6} s^{-1}, and 1×10^{-7} s^{-1} for 12 km, 4 km, and 1.3 km. We incorporate the land cover product from the National Land Cover Database (NLCD; Dewitz, 2021) at a 9 arc-second resolution. For the WRF physics options, we select the Morrison 2-moment microphysics scheme (Morrison et al., 2009), version 2 of the Kain-Fritsch (KF2) cumulus cloud parameterization for the 12- and 4-km simulations (Kain, 2004),the Asymmetric Convective Model version 2 (ACM2) for the planetary boundary layer (Pleim, 2007) and the Pleim-Xiu land surface model (Xiu & Pleim, 2001) with soil moisture and temperature nudging (Pleim & Gilliam, 2009; Pleim & Xiu, 2003). We use the Rapid Radiative Transfer Model for GCMs (RRTMG; Clough et al., 2005) for both our shortwave and longwave radiation schemes.

To create 4 km and 1.3 km emissions inputs for use in WRF-CMAQ, we processed the EPA's 2016 Beta emissions modeling platform with the SMOKE software. We process the 2016v7.2 National Emissions Inventory (Eyth et al., 2019) using the 2016 SMOKE Beta Platform, relying on 4 km spatial surrogates provided by CMAS (CMAS, 2022) and 1.3 km spatial surrogates provided by LADCO (LADCO, 2022). The spatial surrogates map county-level emissions inventories to model grid cells by

using the geographic attributes of the modeling area (such as population, industry, and economic activity). Meteorological conditions are a key factor in determining various vehicle emission processes such as tailpipe exhaust, evaporative processes, brake and tire wear, idling, cold starts (wherein an engine has been at rest for 12+ hrs), and hoteling (wherein long-haul trucks idle for extended rest periods) (SMOKE v4.5 User's Manual, 2017). As such, we integrate the 2018 and 2019 stand-alone WRF-simulated meteorology into the MOVES version developed for the 2016 beta platform, using the 2016 MOVES activity data (Eyth et al., 2019). We use SMOKE to create emissions for the on-road, point, and nonpoint sectors, with meteorology affecting the point, onroad, and some non-point sector processes. We calculate biogenic emissions (BEIS), lightning NOx emissions, and windblown dust "inline" during the coupled WRF-CMAQ simulation. Within the point sector, we use 2018 and 2019 CEM data to integrate the reported energy-generating unit (EGU) data.

To ultimately simulate atmospheric pollutants, we run the two-way coupled version of WRF-CMAQ at 4 km and 1.3 km. We first run coupled WRF-CMAQ over the 4 km domain using the meteorological boundary conditions from the 12 km stand-alone WRF simulation, nudging from NARR (3-hourly resolution) and chemical boundary and initial conditions from CAM-Chem (Emmons et al., 2020; The CESM2 Development Team, 2019), with chemical variables mapped in Table S1.1 . We run the coupled 4 km simulation with an 18-second time step and 10-minute radiation time step, with CMAQ coupled every 8 WRF steps. We then use the 4 km output as meteorological and chemical boundary conditions for the 1.33 km WRF-CMAQ simulation. For the 1.33 km WRF-CMAQ simulation, we run WRF with a 6 second time step and 5-minute radiation time step, with CMAQ coupled every 8 WRF steps. Both the 4 km and 1.3 km simulations integrate the Carbon Bond Mechanism version 6 and aerosol module version 6 with aqueous chemistry (cb6r3_ae6_aq) to create atmospheric constituents. For the two-way WRF-CMAQ 4 km and 1.3 km simulations, we applied FDDA coefficients to keep the model simulation closer to observations above the boundary layer, using default nudging coefficients for temperature and wind (3×10^{-4} s^{-1}) and the water vapor mixing ratio (1×10^{-5} s^{-1}).

2.3 Results

To present our two-way coupled WRF-CMAQ simulations and highlight their ability to resolve neighborhood-scale air quality, we begin with evaluations of the model's domain-wide meteorological and chemical performance across temporal scales. We then highlight the utility of simulations that resolve air quality within individual neighborhoods by conducting an in-depth analysis of intra-urban air quality by characterizing pollutant heterogeneities across Chicago, IL and their interactions with local meteorological features, infrastructure, emissions sources, and the temporal distribution of emissions. Lastly, we assess the benefits, and in some cases disbenefits, of higher spatial resolution for model-observation fidelity performance.

We begin by comparing model-simulated air quality and meteorological data from our highest resolution 1.3 km domain (d03; Figure 2.1) to ground-based observations. We evaluate model performance for each simulated month. Our air quality performance evaluation primarily focuses on O_3, NO_2, and $PM_{2.5}$, although other EPA criteria pollutants (i.e., CO and SO_2) are also discussed. We evaluate model fidelity to meteorological and air pollutant observations using the following performance metrics: mean observation (μ_d), mean prediction (μ_p), normalized mean bias (*NMB*), normalized mean error (*NME*), correlation coefficients (*r*), mean error (*ME*), mean bias (*MB*), and root mean squared error (*RMSE*) as defined in Table S1.2. By normalizing model-simulated variables, the statistical performance of our simulations can be compared to similar model simulations performed over locations with different meteorology, emission profiles, and chemical regimes.

2.3.1 WRF-CMAQ Meteorological Performance

To assess the performance of the two-way coupled WRF-CMAQ meteorological output over the 1.3 km domain, we compare model simulated variables to ground-based measurements of meteorological conditions. We use hourly observational data from METAR stations aggregated by the National Climatic Data Center (NCDC) (Figure 2.1b). We focus on 2-m temperature (T2) and relative humidity (RH) at 2

m, and wind speed (WS) and wind direction (WD) at 10 m, each of which is important to the fate and transport of atmospheric pollutants. We evaluated model fidelity using the performance recommendations outlined in Table S1.3 (Emery et al. 2001). Model-observation comparisons occur where WRF grid cells contain NCDC stations (Figure 2.1b). The 1.3 km domain contains 10 NCDC stations, which allows for model-observation comparison and assessment at 0.01% of the simulation grid cells (90,720 total). We also assess the model's meteorological performance within Chicago city limits, which has a single NCDC station.

In Table 1.1, we summarize the model's 1.3 km domain (d03) hourly meteorological performance against observations. We assess diurnal meteorological performance in Figure S1.1. For each month, WRF-CMAQ simulated T2, WD, and WS meet the correlation performance criteria suggested by Emery et al., 2001 (Table S1.3). Emery et al. (2001) do not make RH performance recommendations. WRF performance is best when simulating T2; model-station agreements have low biases and errors, though January 2019 and August 2018 have slightly higher biases than Emery et al.'s suggested benchmark (Table 1.1). Model simulations have a consistent warm bias across seasons, with the highest biases in August 2018 and January 2019 (MB = 0.8 °C), and highest mean errors in January 2019 and April 2019 (ME = 1.9 °C). Simulated RH is also highly correlated with observations ($r > 0.70$), with the highest bias and error in April 2019 (MB = 5.1 %, ME = 11.3). Simulated wind speeds are biased low in each season but meet MB benchmark criteria in January 2019 and April 2019. The lowest WRF performance is shown by WD, which only meets suggested MB criteria for April 2019. Model simulated WD and station measurements are highly correlated ($r > 0.5$), except for August 2018 ($r = 0.3$). The simulation is wetter ($MB < 6\%$) and warmer ($MB < 0.8$ °C) than observations for all months, except for August 2018, where RH is biased low (MB = -5.7%).

Table 2.1 Comparison of two-way coupled WRF-CMAQ simulated hourly meteorological variables with NCDC observations for 1.3 km (d03) simulations. The average observed value is noted as μd, while the predicted value is noted as μp. *Indicates performance outside of Emery et al. (2001) suggested benchmarks (Table S1.3).

Var	Month	μ_d	μ_p	*MB*	*ME*	*RMSE*	*r*
T2 (°C)	08/2018	23.2	24.0	0.8*	1.8	2.4	0.9
	10/2018	10.8	11.1	0.3	1.8	2.3	0.9
	01/2019	-5.8	-5.0	0.8*	1.9	2.5	1.0
	04/2019	9.2	9.2	0.0	1.9	2.5	0.9
	Average	9.4	9.8	0.5	1.9	2.4	0.9
RH (%)	08/2018	76.6	71.0	-5.7	11.1	14.4	0.7
	10/2018	74.8	76.6	1.9	12.2	15.4	0.7
	01/2019	74.6	78.6	4.0	9.6	11.7	0.7
	04/2019	66.1	71.9	5.8	12.4	16.1	0.8
	Average	73.0	74.5	1.5	11.3	14.4	0.7
WS (m/s)	08/2018	6.9	6.1	-0.7*	2.9	3.7*	0.6
	10/2018	8.8	7.3	-1.5*	3.2	4.1*	0.7
	01/2019	10.2	10.1	-0.1	3.4	4.7*	0.7
	04/2019	10.6	8.3	-2.3*	3.8	4.8*	0.7
	Average	9.1	8.0	6.7	22.3	35.4	0.6
WD (°)	08/2018	166.0	196.7	30.7*	78.8*	127.8	0.3
	10/2018	190.6	207.7	17.1*	53.6*	102.2	0.5
	01/2019	192.8	208.5	15.7*	41.4*	89.5	0.6
	04/2019	166.8	171.0	4.3	52.2*	99.4	0.6
	Average	179.1	196.0	17.0	56.5	104.7	0.5

Within Chicago city limits, there is one NCDC meteorological station, located ~16 km inland from Lake Michigan at O'Hare International Airport on the northwestern edge of the city (denoted by a triangle in Figure 2.1b). Model performance in comparison to the O'Hare meteorological station is shown in Table S1.4 and Figure S1.2. Similar to the full 1.3 km domain comparison, meteorology in the model grid cell that contains O'Hare shows high correlations with NCDC observations for T2, WS, and RH.

Also, like the full domain comparison, WRF-simulated T2 has the highest correlation with observations, while WD correlations are lowest. The ME and RMSE are higher for the O'Hare grid cell for WS and RH than for the full 1.3 km domain comparison, but lower for T2 and WD. Unlike the full 1.3 km domain, simulated T2 is cooler than observations at O'Hare (-0.2 °C), though the RH biases are similar to the domain average (Table S1.4).

2.3.2 WRF-CMAQ Pollutant Simulation Performance

To assess WRF-CMAQ pollutant simulation performance, we compare model simulated criteria pollutant concentrations to measurements of NO_2, O_3, $PM_{2.5}$, SO_2 and CO from EPA Air Quality System (AQS) monitoring stations, which report hourly observations of each pollutant. Model-observation comparisons occur where WRF-CMAQ grid cells contain EPA AQS stations (Figure 2.1c). The number of EPA monitoring stations changes within the 1.3 km domain depending on the season with 115 total stations in August 2018, 115 in October 2018, 73 in January 2019, and 119 in April 2019. Over our simulation period, there were a maximum of 10 NO_2, 67 O_3, 29 $PM_{2.5}$, 8 SO_2, and 5 CO monitors, though some stations occasionally drop offline during our simulation period. Notably, the number of O_3 monitors drops from 67 to 23 in January 2019. We filter EPA station data by removing observations with negative concentration values and quality assurance (QA) flags.

We compare model simulated pollutants to AQS observations over several different time scales, including 4-monthly or annualized mean, monthly mean, daily mean, hourly mean, and daily maximum (Figure 2.2). We calculate the annualized mean by averaging across our simulated months of August 2018, October 2018, January 2019, and April 2019. In the following, we provide quantitative assessments of each criteria pollutant across different temporal periods to provide context for model performance on both fine (hourly, daily, and daily maximum) and coarse scales (monthly and annual), with diurnal profiles available in the supplement (Figure S1.3). In general, the model and observations have better agreement with longer time-averaging slices, e.g., lower biases and errors. To provide greater context for our model performance, we follow EPA recommendations (Dennis et al., 2010) and compare the

performance of WRF-CMAQ over our 1.3 km domain to previously published CTM studies (Table S1.5). Our comparisons use fine-scale, domain-agnostic studies (<4 km horizontal resolution), or coarser-scale (>4 km horizontal resolution) studies focused on Chicago, the Midwest, or the Great Lakes region, i.e., studies with similar model domains. We select studies that use the same statistical metrics as in Table S1.1, simulate time periods after the year 2000, and integrate a similar CTM (WRF-CMAQ or WRF-Chem). We do not focus on other benchmark studies that use coarser and/or older versions of CTMs or emissions models (Emery et al., 2017; Simon et al., 2012).

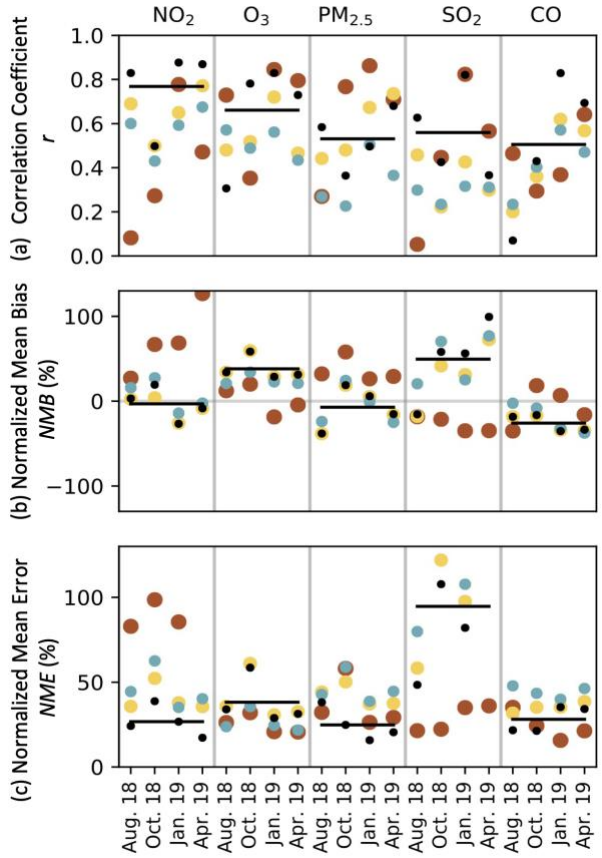


Figure 2.2 Model performance relative to observations of 5 pollutants over 5 different time-averaging periods. The periods are: hourly (red), daily (yellow), daily maximum (blue), and monthly (black). The average of the four simulation months, each drawn from a different meteorological quarter, i.e., the annualized performance is depicted as a black horizontal line. Each panel summarizes a key performance metric: (a) correlation coefficient (pearson r), (b) normalized mean bias (NMB), (c) and normalized mean error (NME).

Table 2.2 WRF-CMAQ performance metrics for hourly simulated 1.3 km (d03) pollutants as compared to EPA AQS station observations. The average observed value is noted as μ_d, while the predicted value is noted as μ_p.

		1.3 km Domain Performance				
Var	**Month**	μ_d	μ_p	***NMB*%**	***NME*%**	***r***
NO_2	08/18	10.38	10.69	2.98	55.76	0.59
	10/18	10.76	11.05	2.64	66.08	0.47
	01/19	13.13	9.59	-26.95	45.45	0.62
	04/19	11.19	10.57	-5.56	51.21	0.63
	Average	11.37	10.48	-6.72	54.62	0.57
O_3	08/18	30.28	40.25	32.92	40.04	0.69
	10/18	20.38	32.12	57.66	62.55	0.58
	01/19	24.61	31.47	27.88	34.10	0.65
	04/19	36.33	47.35	30.35	34.46	0.61
	Average	27.90	37.80	37.20	42.79	0.63
$PM_{2.5}$	08/18	12.12	7.49	-38.21	54.61	0.25
	10/18	6.78	7.89	16.43	67.18	0.35
	01/19	9.42	9.83	4.39	50.87	0.52
	04/19	7.60	6.26	-17.62	53.94	0.51
	Average	8.98	7.87	-8.76	56.65	0.41
SO_2	08/18	0.76	1.41	87.23	169.66	0.21
	10/18	0.83	1.13	37.14	139.52	0.11
	01/19	0.99	1.18	19.68	110.15	0.21
	04/19	0.79	1.25	57.57	152.94	0.12
	Average	0.76	0.96	25.33	116.39	0.18
CO	08/18	250.72	204.53	-18.42	43.25	0.24
	10/18	229.61	188.81	-17.77	46.99	0.31
	01/19	284.08	186.90	-34.21	40.55	0.46
	04/19	281.43	183.51	-34.79	44.70	0.40
	Average	261.46	190.94	-26.30	43.87	0.35

Of all pollutants, we find that WRF-CMAQ-simulated NO_2 is closest to the observations with low *NMB* and high correlations across months (Figure 2.2). The annualized average hourly correlation of NO_2 is high (r = 0.6; Figure 2, Table 1.2), while its bias is low (*NMB* < -7%; Figure 2.2, Table 2.2). NO_2 model-observation correlations are generally greater than 0.6 regardless of temporal assessment scale,

except for October 2018 (Figure 2.2). We find slight high biases in model simulated hourly NO_2 in August (*NMB* = 3%) and October 2018 (*NMB* = 3%) and low biases in January (*NMB* = -27%) and April 2019 (*NMB* = -6%) (Table 2.2). When compared to previously published WRF-CMAQ studies with different domains/resolutions, our NO_2 simulation performance exceeds *NMBs* and correlations reported by Bickford et al. (2014), Harkey et al. (2015), and Vijayaraghavan et al. (2009).

Model-simulated O_3 is high relative to observations, with limited variation across seasons (Figure 2.2 and Table 2.2). We find that the annualized average correlation of simulated hourly O_3 is high (r = 0.6), but that the annualized *NMB* (38%) and *NME* (42%) are high. The highest *NME* for O_3 occurs in our October 2018 simulation (58%), which corresponds with the highest *NME* for NO2 (64%). The lowest *NMB* and *NME* are found in January 2019 (27%, 34%), which has the lowest concentrations of O_3. Compared to other studies in the Great Lakes region, our biases and errors are higher than those of Bickford et al. (2013), who ran WRF-CMAQ without two-way coupling, and Abdi-Ouskouei et al. (2020), who used WRF-Chem. Other similar CTMs studies report O_3 biases similar to those reported here (Odman et al., 2019; S. Pan et al., 2017; Qin et al., 2019; Travis et al., 2016; Zhang et al., 2014). Our high O_3 bias is mainly driven by an over prediction of simulated O_3 concentrations during periods of low observed O_3, particularly at night (Figure S3). During warm "ozone season" months when observed O_3 is high (O_3 > 60 ppb), our *NMB* is negative (-5.4% and -7.2% for August and April) and *NME* are less than 17% (Table S1.6). When our model performance evaluation is limited to hours when observed O_3 concentrations are greater than the 50%ile value, average annualized *NMB* is reduced to ~25% (Table S1.6). Lastly, and further confirming WRF-CMAQ's challenges with capturing low O_3 concentrations, the *NMB* in our model-simulated daily maximum 8-hr running average O_3 ($MDAO_3$) is ~27% when annualized, ~25% in O_3 season months, and only ±2% when $MDAO_3$ is greater than 60 ppb (Table S1.6).

Unlike model simulated O_3, our simulated hourly $PM_{2.5}$ concentrations have low biases and low correlations (Figure 2.2 and Table 2.2). The annualized average correlation of hourly $PM_{2.5}$ is 0.4, with

NMB of -10% and *NME* of 56%. August 2018 hourly $PM_{2.5}$ has the largest bias (-38%) and lowest correlation (r = 0.25), while the highest *NME* (67%) and highest positive *NMB* (16%) are found in October 2018. Within the Great Lakes region, we find that our model-observation agreement for $PM_{2.5}$ has higher correlations and similar *NME*s to Bickford et al., 2013. Compared to other WRF-CMAQ studies within the continental U.S., our $PM_{2.5}$ *NME* is lower than Hogrefe et al., 2015 but higher than Liu et al., 2010 and Wang et al., 2021. Our $PM_{2.5}$ *NMB* and MB are similar to Hogrefe et al. (2015) and Wang et al. (2021), but lower than Liu et al. 2010 and Torres-Vazquez et al., 2022.

The agreement of our model simulated SO_2 and CO compared to the AQS observations were the lowest of the 5 criteria pollutants (Figure 2.2 and Table 2.2). Annualized average correlation of SO_2 is 0.18, with *NMB* = 25% and *NME* = 116%. Annual average correlation of CO is 0.35, with *NMB* = -26% and *NME* = 44%. Few previous WRF-CMAQ studies report their performance of SO_2 and CO. Compared to those that do, our simulation of SO_2 had lower *NMB*s, higher *NME*s, and lower correlations (Bickford et al., 2014; Campbell et al., 2019).

2.3.3 Domain-wide Characterization of WRF-CMAQ Simulated Pollutants

In Figure 2.3, we show monthly-average simulated NO_2, O_3, and $PM_{2.5}$ concentrations over the 1.3 km domain (d03) for each season (Figure S1.4 shows model results overlaid with station observations). Simulating pollutants at a 1.3 km spatial resolution facilitates the characterization of distinct urban-rural patterns, the influence of Lake Michigan on regional O_3 distributions, pollutant hotspots over highway corridors, stationary emitting sources, and urban centers. We distinguish urban areas from rural by classifying the most populated census tracts (>95%ile) from the American Community Survey 2018 (Manson, Steven et al., 2022) as urban, and all other areas as rural (Figure S1.5). In the following, we individually discuss domain-wide analyses of each pollutant and then highlight the model's characterization of pollutants within the city of Chicago.

The simulated NO_2 concentrations largely track high-population areas and highway corridors (Figure 2.3a-e). In all seasons, the interstate highway system that connects population centers is highlighted by the interconnected system of roadways with elevated NO_2 concentrations. The lowest NO_2 concentrations in our domain are simulated over northernmost and easternmost portions of Lake Michigan, in areas distant from emissions sources. Likewise, rural areas distant from roadways have low NO_2 concentrations. We find that the average annual urban concentration of NO_2 in our domain is simulated to be 3.5 ppb (59.8%) higher than average concentrations in rural portions of our domain (Figure 2.3 & Table S1.7). Across seasons, domain-wide NO_2 concentrations correspond with the magnitude of simulated NOx emissions (Figure S1.6). In January 2019, domain average NO_2 concentrations are highest (μ = 3.1 ppb; Figure 2.3c), which corresponds to our highest simulated NOx emissions ($\mu_{NOx,January}$ = 2.6 x 10^{-3} g/s). The lowest domain average NO_2 concentrations occur in April 2019 (μ = 2.4 ppb; Figure 2.3d), which co-occurs with low NOx emissions ($\mu_{NOx,april}$ = 2.5 x 10^{-3} g/s).

In contrast to NO_2, O_3 concentrations are simulated to be relatively low over urban areas and roadways, with low concentrations over individual highway corridors apparent (Figure 2.3f-j). Across seasons, the spatial distribution of O_3 concentrations is relatively consistent, however the magnitude of O_3 concentrations varies by season. Of the four months that we simulate, we find that the highest O_3 concentrations occur in April 2019, (μ = 49.4 ppb; Figure 2.3i), while the lowest concentrations are simulated in January 2019 (μ = 33.9 ppb; Figure 2.3h). Over simulated warm season months (i.e., April 2019 and August 2018), we find that domain-wide O_3 concentrations are ~1.5 times higher than cool-season concentrations. Comparing concentrations across the urban-rural divide (Figure S1.5), we find that O_3 over urban areas is simulated to be ~3.3 ppb (9.4%) lower than over rural areas, with the greatest urban-rural difference in cool season months (Table 2.3).

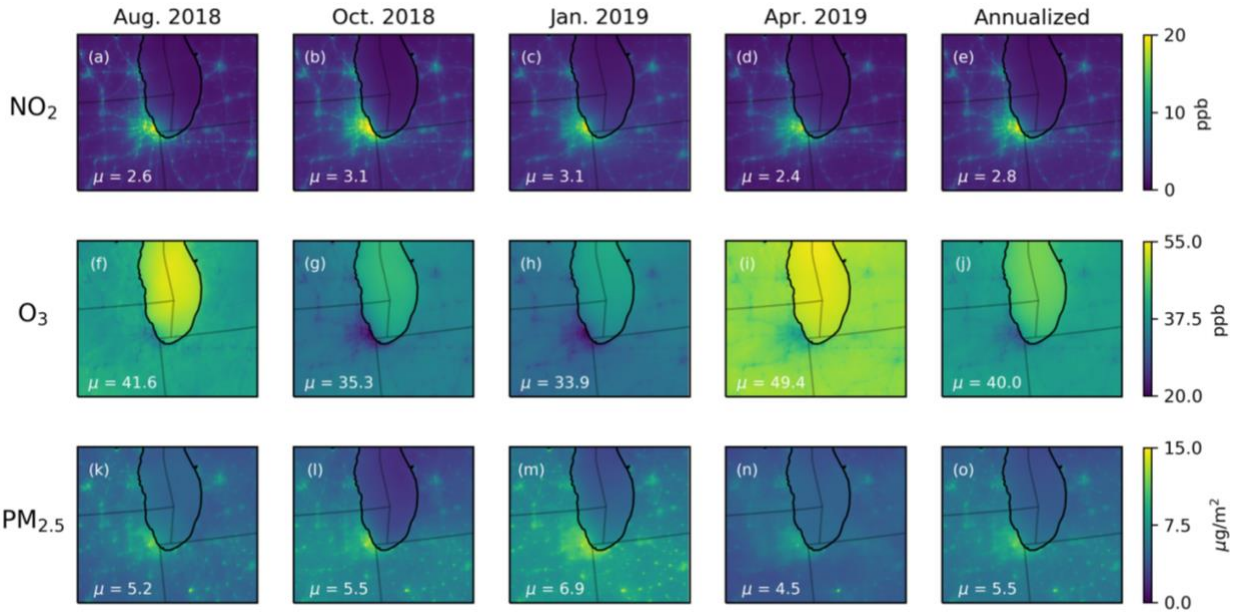


Figure 2.3 Monthly and annualized average WRF-CMAQ simulated NO_2 (a-e), O_3 (f-j) and $PM_{2.5}$ (k-o) for the 1.3 km simulation domain (d03). From left-to-right each column presents August 2018, October 2018, January 2019, and April 2019, with the domain-average (μ) concentrations annotated in the lower left of each panel. The right-most column provides the annualized average. Note: the colorbar for O_3 concentrations begins at 20 ppb.

Across all seasons, the highest simulated concentration of O_3 in the 1.3 km domain occurs over Lake Michigan (Figure 2.3f-j). The simulation of elevated over-lake O_3 concentrations is consistent with previous Lake Michigan observation and modeling campaigns (Doak et al., 2021; Dye et al., 1995; Foley et al., 2011), and similar to other studies focused on inland bodies of water (e.g., Chesapeake Bay, Goldberg et al., 2014). Elevated O_3 over Lake Michigan is thought to be dependent on the circulation of primary pollutants from land to lake via the lake-breeze. Elevated O_3 over the lake occurs through the following idealized sequence of events: (1) In the morning land-based emissions (O_3 precursors) are transported over the lake by a land-breeze, which combine with shipping emissions, and are trapped within a shallow boundary layer. (2) As the day warms, the land-lake temperature gradient weakens, and the land breeze dwindles. Fewer NO_x emissions are transported to the lake. (3) As sunlight increases, photochemical production of O_3 over the lake is enhanced. Due to the low number of depositional pathways over the lake, O_3 accumulates. (4) On days where a lake breeze forms, O_3 is advected inland,

often to areas where the original precursor emissions did not originate. Lake breeze effects are primarily a warm season phenomenon, however, the low concentration of over-lake depositional pathways also contributes to elevated cool season O_3 concentrations (Figure 2.3f-j) (Doak et al., 2021; Dye et al., 1995). In our simulations, land-lake O_3 concentration differences are greatest in August 2018, as the average concentration of O_3 over land (39.8 ppb) is 11 ppb lower than the average O_3 concentration over the lake (50.4 ppb).

Domain-wide simulated $PM_{2.5}$ concentrations show greater spatial smoothness, such that elevated $PM_{2.5}$ hot spots have a more diffusive footprint compared to NO_2 hot spots (Figure 2.3k-o). Across months, the spatial pattern of simulated $PM_{2.5}$ concentrations is relatively consistent and largely tied to the location of emission sources. Despite consistent spatial patterns across months, the relative magnitude of $PM_{2.5}$ concentrations is influenced by meteorological conditions (e.g., boundary layer height and wind speeds), the magnitude of seasonal primary PM emissions, and secondary PM pollutant formation reactions. The domain-wide average concentration of $PM_{2.5}$ peaks in January 2019 (μ = 6.9 μ/m^3) and is lowest in April 2019 (μ = 4.5 μ/m^3), which mirrors the pattern of emissions of PM and its precursors (Figure S1.6). Both stationary and mobile sources of $PM_{2.5}$ typically co-emit NO_x emissions, as such simulated $PM_{2.5}$ hotspots tend to co-occur with NO_2 hotspots over urban areas, highways, and stationary sources (Figure 2.3a-e and k-o). However, compared to NO_2, $PM_{2.5}$ concentrations are more spatially diffuse due to the longer atmospheric lifetime of PM species (Laughner & Cohen, 2019; Pinto et al., 2004). Further, the urban-rural concentration disparity is lower, i.e., $PM_{2.5}$ concentrations are 22% higher in urban v. rural areas, compared to 60% for NO_2 (Table S1.7). Simulated grid cells with the highest concentration of $PM_{2.5}$ occur outside of urban areas and are primarily associated with emissions from industrial and manufacturing point sources.

2.3.3 Neighborhood-Scale Characterization of WRF-CMAQ Simulated Pollutants

Our domain-wide analysis demonstrates the ability of the 1.3 km WRF-CMAQ simulations to characterize differences in urban-rural regimes and identify pollutant hotspots, however it does not highlight the ability of the model to resolve and characterize neighborhood-scale air quality. To demonstrate this ability, we provide an in-depth analysis of a sub-region of the 1.3 km modeling domain, i.e., the city of Chicago (Figure 2.4a), a city with ~2.7 million people (American Community Survey, 2018). At the 1.3 km model resolution, there are on average ~5 grid cells in each of the 77 neighborhoods that comprise the city (Figure 2.4). Chicago sits close to the center of our 1.3 km domain, and in Figure 2.3 is identifiable as both an NO_2 and $PM_{2.5}$ hotspot at the southwest corner of Lake Michigan. Major sources of emissions within Chicago include transportation, industry, and buildings. The city has 6 major interstate highways (I-290, I-294, I-90, I-94, I-55, I-57) that loosely outline the City's lakeside central business district or "Loop". There are two airports within City limits, O'Hare in the northwest and Midway in the south central. Most industrial activities occur on the west and southwest sides of the city.

In Figure 2.4b-d, we show the city's average annualized concentrations of NO_2, O_3, and $PM_{2.5.}$ Simulated concentrations of criteria pollutants are higher within Chicago than the domain average, so Figure 2.4 has a different color bar than that used in Figure 2.3. For all pollutants, the 1.3 km WRF-CMAQ simulations reveal substantial spatial heterogeneity amongst neighborhoods. Spatial gradients are particularly substantial for simulated NO_2, with concentrations in some neighborhoods double those in others. The predominant spatial pattern of simulated pollutants strongly corresponds to the interstate highway system (Figure 2.4a), although the highest concentrations of NO_2 and $PM_{2.5}$ are simulated where the lowest concentrations of O_3 are simulated. O_3 concentrations are elevated in neighborhoods that abut Lake Michigan, consistent with the influence of the lake breeze, and in neighborhoods without interstate highways. In the following paragraphs we discuss annualized pollutant pattern details across Chicago, as well as patterns of individual months. We determine 'highest' or 'lowest' values of pollutant concentrations by computing the 95%ile or 5%ile values, respectively.

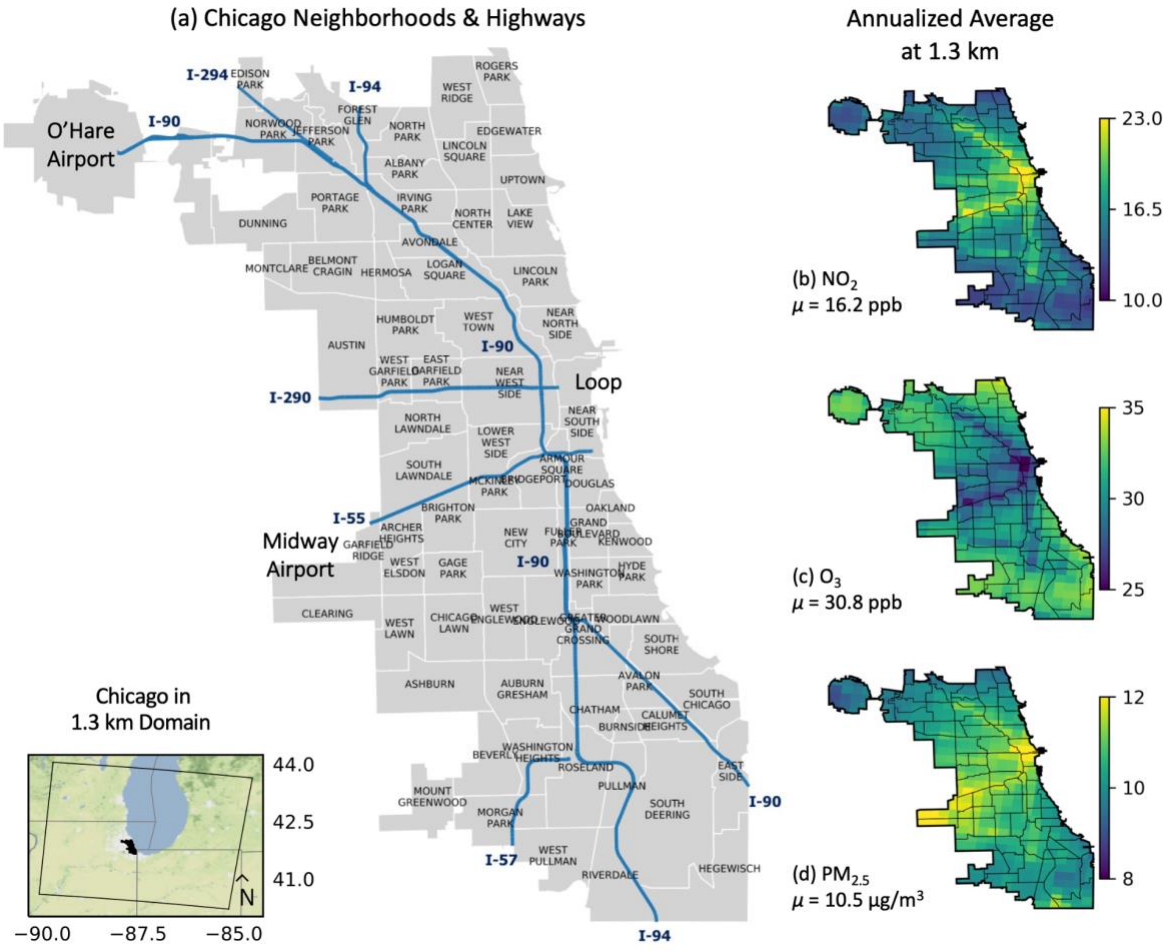


Figure 2.4 Chicago geography and simulated pollutants. (a) Neighborhoods and highways within the city of Chicago. Inset at lower left depicts the location of Chicago (black) within the 1.3 km domain. In the right column, we provide annualized WRF-CMAQ simulated (b) NO_2, (c) O_3, and (d) $PM_{2.5}$ concentrations over Chicago and average concentrations within city limits (μ).

In Figure 2.4b-d, we show the city's average annualized concentrations of NO_2, O_3, and $PM_{2.5}$. Simulated concentrations of criteria pollutants are higher within Chicago than the domain average, so Figure 4 has a different color bar than that used in Figure 2.3. For all pollutants, the 1.3 km WRF-CMAQ simulations reveal substantial spatial heterogeneity amongst neighborhoods. Spatial gradients are particularly substantial for simulated NO_2, with concentrations in some neighborhoods double those in others. The predominant spatial pattern of simulated pollutants strongly corresponds to the interstate highway system (Figure 2.4a), although the highest concentrations of NO_2 and $PM_{2.5}$ are simulated where

the lowest concentrations of O_3 are simulated. O_3 concentrations are elevated in neighborhoods that abut Lake Michigan, consistent with the influence of the lake breeze, and in neighborhoods without interstate highways. In the following paragraphs we discuss annualized pollutant pattern details across Chicago, as well as patterns of individual months. We determine 'highest' or 'lowest' values of pollutant concentrations by computing the 95%ile or 5%ile values, respectively.

The highest annualized concentrations of NO_2 (where $\mu_{annual} > 19.5$ ppb) are simulated on the West side of Chicago and in the Loop, where highways are prevalent (e.g., I-90, I-290, and I-55) and simulated NO_x emissions are high (Figure S1.6). The lowest NO_2 concentrations (where $\mu_{annual} < 11.3$ ppb) are simulated in the lake-front neighborhoods, with the exception of those in the Loop. Lakefront neighborhoods are east of the main interstate highways, where lower NO_x emissions and ventilation contribute to the relatively low simulated NO_2 (Figure S1.6). Across seasons, NO_2 concentrations within Chicago remain highest over the 3 most-trafficked inter-state highways: I-290, I-90, and I-94 (Figure 2.4b). Simulated NO_2 concentrations are lowest in April 2019 (μ = 14.2 ppb) and highest in October 2018 (μ = 18.7 ppb), although the greatest NO_2 bias was also found in October 2018. Compared to the full model domain (μ_{domain} = 2.4-3.2 ppb; Figure 2.3a-d), average NO_2 over Chicago is nearly 5 times higher ($\mu_{Chicago}$ = 14.2-18.7 ppb) across seasons (Figure 2.5a-d). Unlike the NOx-NO_2 emission-concentration relationship found over the full domain, we find that average NO_2 concentrations over Chicago do not correspond with emission magnitude differences, as Chicago NOx emissions are lowest in August 2018 and highest in January 2019 (Figure S1.6), while NO_2 concentrations are highest in October 2018 and lowest in April 2019. We discuss this disconnect further in our discussion of model performance in section 2.4.1.

Simulated annualized and individual month O_3 concentrations within Chicago tend to be the spatial inverse of simulated NO_2 concentrations (Figures 2.4 & 2.5). The lowest concentrations ($\mu_{annual} <$ 28.5 ppb) of O_3 are simulated on the West side of the city, near the interstates. These locations are also

simulated to have the highest NO_2 concentrations, i.e., O_3 is suppressed via titration by NO. O_3 concentrations are highest in the warm months, with August 2018 (μ = 35.7 ppb) and April 2019 concentrations (μ = 41.1 ppb) nearly double October 2018 (μ = 23.3 ppb) and January 2019 concentrations (μ = 22.7 ppb). Average annualized O_3 concentrations in Chicago (μ_{annual} = 30.7 ppb) are simulated to be significantly lower than the domain average because of the lake reservoir of O_3 (μ = 40.0 ppb). Even when land-only O_3 concentrations are isolated, Chicago has concentrations that are slightly lower than the rest of the full model domain ($\mu_{land,domain,annual}$ = 38.8 ppb). Warm-season O_3 is highest near Northern lake-front neighborhoods, which are distant from the major interstates, have low NO_x emissions (Figure S1.6), and subject to lake breeze advection of the reservoir of O_3 over Lake Michigan. In the cool months, O_3 concentrations are simulated to be highest on the western edges of the city. However, the cooler months have a lower range of O_3 concentrations (±8.9 ppb) than warmer months (±15.5 ppb).

Annualized $PM_{2.5}$ concentrations in Chicago correspond well with the spatial patterns of the interstate system, though the $PM_{2.5}$ footprint is spatially more extensive than that of NO_2 (Figure 2.4). $PM_{2.5}$ concentrations in Chicago are simulated to be 2 times higher than the average concentration of the full model domain ($\mu_{domain,annual}$ = 5.5 $\mu g/m^3$, $\mu_{Chicago,annual}$ = 10.2 $\mu g/m^3$). $PM_{2.5}$ concentrations peak on the west side of Chicago near Midway airport and the intersection of I-290 and I-55 with I-90 (μ = 12 – 13 $\mu g/m^3$). The lowest concentrations of $PM_{2.5}$ occur on the lakefront (μ = 8 – 10 $\mu g/m^3$). Similar to O_3, Chicago $PM_{2.5}$ levels show strong seasonal variations, though the simulated concentrations of $PM_{2.5}$ are highest in the seasons when O_3 is lowest. As such, simulated $PM_{2.5}$ peaks in the cooler months (Figure 5j-k) and is lowest in April 2019 (μ = 8 $\mu g/m^3$). Areas of high $PM_{2.5}$ in Chicago are consistent across seasons, in particular on the west side of the city and within the Loop.

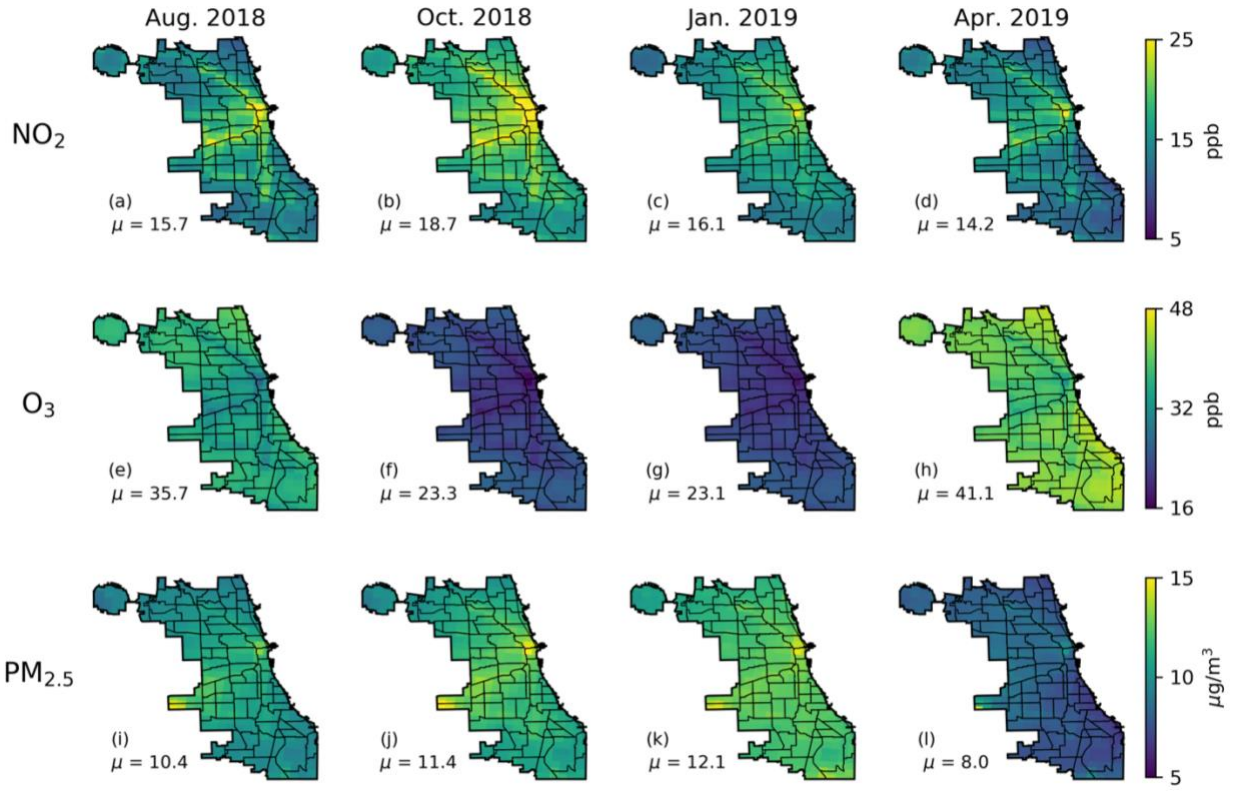


Figure 2.5 Monthly average NO_2 (a-d), O_3 (e-h) and $PM_{2.5}$ (i-l) concentrations over Chicago, as simulated in the 1.3 km domain. Columns depict August 2018, October 2018, January 2019, and April 2019 simulations. Chicago-average concentrations are annotated (μ).

2.3.4 Benefits and Disbenefits of Increased Model Resolution

Our nested modeling framework facilitates assessment of the potential benefits and/or disbenefits of attempting to resolve neighborhood scale meteorology and air pollutants with a CTM. That is, given that our methodology simulates atmospheric chemistry and meteorology at both 4 km and 1.3 km resolutions, an assessment of increased spatial resolution on model performance is possible. However, our chosen methodology does not provide a pure spatial resolution sensitivity analysis, i.e., while the underlying emissions data of each simulation is the same (NEI 2016v1), the meteorologically informed emissions are slightly different due to differences in the 4 km and 1.3 km WRF simulations. Further, we compare stations to grid cells in each domain, and the 1.3 km grid cells cover a smaller area than the 4 km grid cell. Despite this imperfect sensitivity analysis, we compare model performance at both resolutions. We restrict our comparison to model performance over the 1.3 km domain. As such, we use the same 10

NCDC meteorological observing stations and the ~125 EPA AQS stations shown in Figure 2.1 and discussed in the WRF-CMAQ simulation comparison. After this performance comparison, we investigate spatial changes in simulated pollutants over Chicago at the different resolutions.

For the simulated meteorology, we find that the increase in resolution benefits the performance of WS and WD, has no influence on RH, and has disbenefits for T2 (Table 2.1 & Table S1.8). Higher resolution provides the greatest benefit to simulated WD as *MB* and *ME* are lower at 1.3 km and correlations are higher. In contrast, we find that observed T2 is better captured in the 4 km simulation, as lower bias and error and higher correlations are found compared to the 1.3 km simulation. Within Chicago (with just one NCDC station) the 1.3 and 4 km simulations perform similarly to their domain-wide performance, with WD showing slightly lower biases in the 4 km domain (Table S1.9).

For simulated pollutant concentrations we find higher model-observation correlations in the higher resolution (1.3 km) simulation, though this improvement is coupled with increased *ME* (Table S10). The 1.3 km simulation showed higher correlations than the 4 km simulation for all criteria pollutants, though this increase was marginal ($\Delta r < 0.1$, Table S1.10). The 1.3 km simulation of NO_2 has a closer agreement to the EPA stations, but this comes with slightly higher normalized errors (+0.25%; Table S1.10). On average, the 1.3 km simulation *NMB* was lower than that of the 4 km model simulation for each season, which came at a trade-off, as *NME* was only lower in the 1.3 km simulation in April 2019. The correlation between AQS observation and model outputs for NO_2 were similar for the 1.3 km and 4 km simulations. The 1.3 km simulation lowered the *NMB* by 8% in August 2018 and January 2019, with marginal bias improvement in April 2019 and January 2019. In contrast, the 1.3 km simulation of O_3 showed higher *NMB* and *NME* than the 4 km simulation. For $PM_{2.5}$, we find that the 1.3 km resolution simulation has a lower *NMB* than the 4 km simulation for 3 out of 4 seasons, but the *NME* is marginally higher (0.9%) in the higher resolution simulation. Simulated SO_2 showed the largest improvement with finer model resolution ($\Delta NMB = 10\%$, $\Delta r = 0.03$), however this was also the pollutant with the lowest

performance in both the 1.3 km and 4 km domains. CO had slightly better performance in the 4 km domain, as the *NME* and *NMB* were 0.5% to 1% higher, respectively, in 1.3 km simulation.

These meteorological and pollutant performance analyses are limited due to the low number of sensors relative to the number of grid cells simulated in our modeling domain (125 EPA stations and 10 NCDC stations vs 90,720 grid cells). In addition, the finer resolution creates opportunities for local-scale meteorological processes to influence agreement, particularly for pollutants which are not well mixed in the atmosphere (Zhang et al., 2014). The measurement-prediction relationship can be greatly influenced by model grid cell size, plumes, and wind speeds and directions. For example, while the relative amount of SO_2 simulated in a plume may be correct, due to the increase in the number of grid cells in a higher resolution simulation, the potential for an erroneously simulated wind direction to adversely influence model grid cell-observation fidelity increases.

While model performance when assessed against very limited station observations demonstrates improvement (albeit marginal) when the model resolution increases from 4 km to 1.3 km, we note that a key advantage of our higher resolution simulation is the ability to characterize neighborhood-scale air quality, particularly within intra-urban environments. As Chicago is the densest metropolitan area in our modeling domain, we focus our analysis on air quality differences between the 1.3 and 4 km simulations within city limits. We find that over Chicago the 1.3 km simulation has higher average NO_2, $PM_{2.5}$ and O_3 concentrations than the 4 km simulation (Figure 2.6) . Differences in pollutant concentrations at different model resolutions can be caused by a number of factors, including differences in the underlying emissions data, differences in simulated micro-scale meteorology, and/or the nuances of emission-chemistry interactions at the grid cell level. In the following, we provide examples of each.

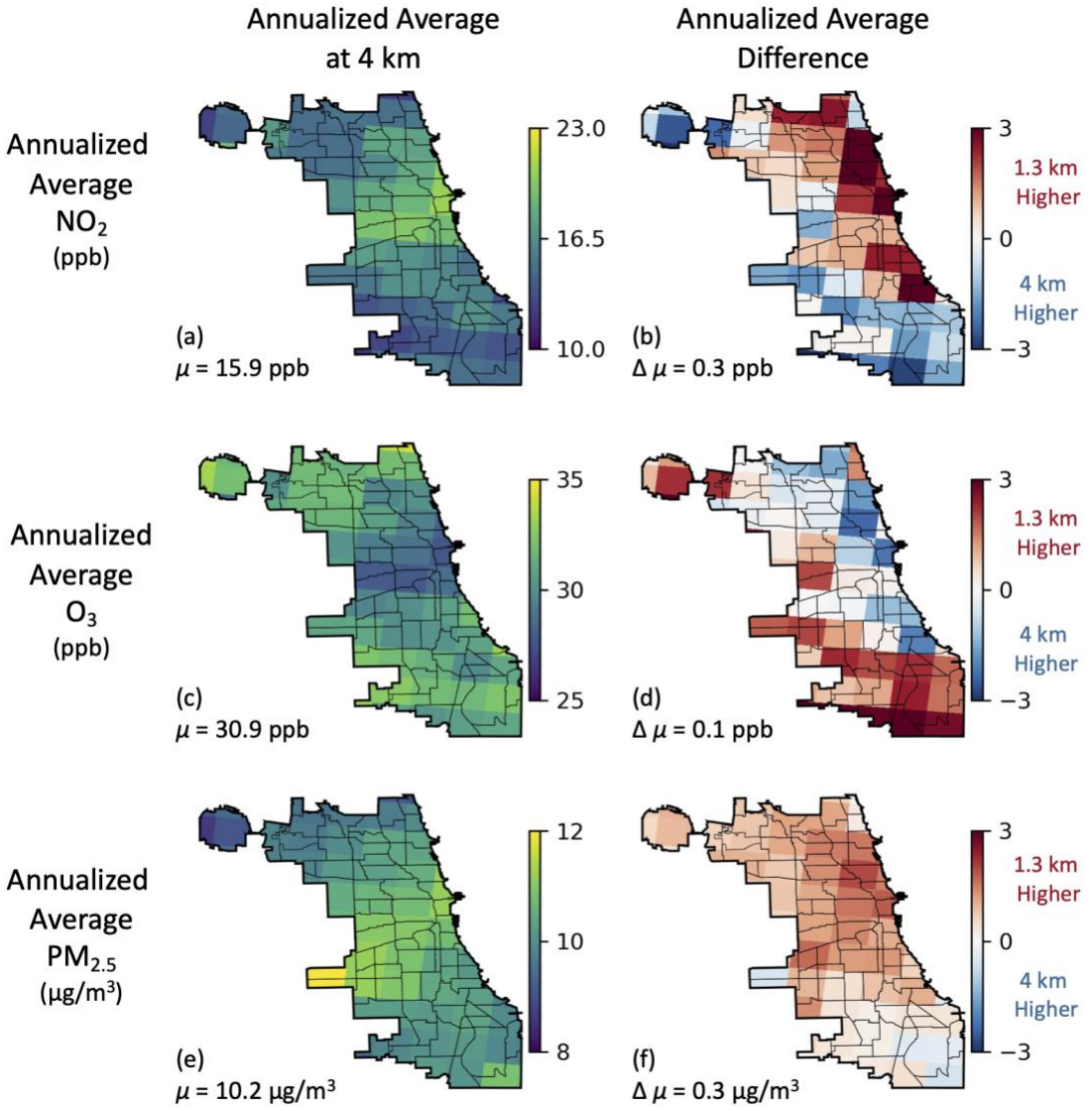


Figure 2.6 Average annualized NO_2 (a), O_3 (c), and $PM_{2.5}$ (e) within Chicago city limits as simulated in the 4 km simulation. (b,d,f). Relative pollutant differences between 1.3 km and 4 km resolution simulations. Average Chicago concentrations (μ) and average differences between model resolutions ($\Delta\mu$) are annotated.

The overall higher concentrations of NO_2 and $PM_{2.5}$ simulated over Chicago at 1.3 km are in part due to differences in the underlying meteorologically-informed emissions. In the case of NO_2, the difference in domain average emissions over the 1.3 km domain and the corresponding subset of the 4 km domain is 0, as the net flux of emissions is the same between the two. However, isolating subsets of the

domains does not result in the same balance, as shown in the Chicago subset of emissions (Figure S6) and resulting pollutant concentration differences (Figure 2.6). Emissions are different between the 1.3 km and 4 km domains because the WRF-simulated temperatures and relative humidity are different at the grid cell level which affects the emissions processing (Figure S1.7). These differences are particularly apparent near the lake shore, where the level of detail afforded by the 1.3 km grid (Figure S1.8) results in notable emission differences (Figure S1.7). Specifically, we find that over Chicago, emission rates of NO_x, organic and elemental carbon, and VOCs are 7.5%, 6.2%, and 1.5% higher in the 1.3 km emissions relative to the 4 km emissions. These higher emission rates, in part, explain the higher average concentration of pollutants over Chicago. For example, elevated NO_x and PM species emissions over the city center and highways (Figure S1.7), contribute to higher NO_2 and $PM_{2.5}$ concentrations in the 1.3 km simulation (Figure 2.6). For O_3, it is likely that higher relative NOx emissions at 1.3 km suppress O_3 concentrations via enhanced NO titration compared to the 4 km simulation.

Despite the above, we find that differences in emissions do not always result in corresponding and co-located changes in pollutant concentrations. As an example, we highlight near-lake grid cells on the north side of Chicago where 1.3 km NOx emissions are lower than at 4 km, but the resulting NO_2 pollutant concentrations are higher (Figure S1.7 g-j). Similarly, on the south side of Chicago we note grid cells with higher NOx emissions at 1.3 km, but lower NO_2 concentrations than simulated at 4 km. This finding is notable because it suggests that fine-scale differences in simulated meteorology drive nonlinear outcomes that affect the production and destruction of pollutants, and is indicative of the value added by fine-scale, fully-coupled CTMs. This finding also suggests caution should be taken when assessing health and justice outcomes at neighborhood scales, particularly if the methods employed do not include dynamic meteorological processes.

2.3.5 O_3 Regimes over the Domain

One notable feature of our 1.3 km simulation is the ability to characterize the O_3 regime at fine spatial resolutions. Localized formation of tropospheric O_3 is a nonlinear process that depends on the relative abundances of precursor emissions, the transport of O_3 and other precursor emissions from upwind areas, and the scale and magnitude of local sinks. Despite this complexity, O_3 production environments are often simplified as either NOx- or VOC-limited regimes (Sillman et al., 1990; Kleinman, 1994). An area is considered "NOx-limited" when VOCs are more available than NO_2, and as such, O_3 production is limited by the radical termination of NO_2 by OH. O_3 production is "VOC-limited" when NO_x is abundant, and O_3 production is limited by the availability of peroxy radicals from VOC oxidation (Schroeder et al., 2017). To determine if areas are NOx- or VOC-limited, VOC concentrations can be compared to NOx concentrations (Ashok & Barrett, 2016). The resultant ratio serves as a proxy to describe the chemical loss of HO_2 + RO_2 (LROx) over the chemical loss of NO_x (LNOx) (Schroeder et al., 2017). Given limited surface observations, researchers often turn to columnar ratios of HCHO to NO_2 from remotely sensed instruments like TropOMI to provide spatially continuous estimates of O_3 regimes (Ashok & Barrett, 2016). While there are not definitive VOC:NOx or HCHO:NO_2 ratio values to delineate whether an area is NO_x- or VOC-limited, it is generally accepted that very low ratios (e.g., < ~1) indicate an area is VOC-limited, very high ratios (e.g., > ~2 for column, >~7 for surface over Chicago) indicate an area is NOx-limited, and values between the high and low range are considered "transitional" (Ashok & Barrett, 2016; Jin et al., 2017). Despite the above context, regime thresholds are uncertain and can be influenced by a number of factors, including local meteorology and VOC species composition (*Rethinking the Ozone Problem in Urban and Regional Air Pollution*, 1991; Seinfeld & Pandis, 2016).

Several previous studies using a variety of methods have attempted to characterize the Chicago region's O_3 regime. These studies have arrived at different conclusions over the years, including some that have found Chicago to be NO_x-limited (Laughner & Cohen, 2019), VOC-limited (Blanchard et al., 2008; Koplitz et al., 2022), or in a transitional state (Jin et al., 2020; Jing & Goldberg, 2022). In Figure 2.7 and Table S1.12, we provide WRF-CMAQ simulated annualized daytime (7 am - 7 pm) surface

VOC:NO_x and column HCHO:NO_2 ratios. We find substantial detail and heterogeneity in simulated surface-level VOC:NO_x ratios (Figure 2.7a), although the simulated column average HCHO:NO_2 ratio shows similar spatial patterns (Figure 2.7b). The more diffuse spatial gradients in the column are due to the integration of the chemicals through the simulated troposphere. The greatest difference between the surface-level and column ratios are shown in the finer characterization of highways and population centers, which show high NOx saturation at the surface (Figure 2.7). Both the domain average annualized VOC:NO_x and HCHO:NO_2 ratios are higher over the full 1.3 km domain than over Chicago, indicating that while the full domain is NO_x limited, the annualized city regime is VOC-limited to transitional. To provide finer temporal detail, we characterize the O_3 regime for each simulated month. Similar to Ashok & Barrett (2016), we find that the surface and column ratios change with the season (Figure S1.9, Figure S1.10). Domain-average surface ratios range from ~7.0-18.5 suggesting a NOx-limited regime in all months (Figure S1.9), while column average ratios range from ~0.6 to 4.3, indicative of a more seasonally dynamic environment (Figure S11.0). Over Chicago, the surface ratios range from ~3.0-4.6, firmly in the transitional regime (Figure S1.9), while column ratios span ~0.2-1.3 indicative of a VOC-limited to transitional regime (Figure S1.10). Given the diversity of previous findings and fundamental uncertainties in regime thresholds, it is perhaps not unexpected that our regime determinations are somewhat sensitive to methodological choice. Despite these uncertainties, the detail afforded by the 1.3 km surface ratio plots is notable (Figures 2.7a & S1.9), both for its contrast with column-based characterizations and its potential in O_3 abatement investigations. We find that the simulated surface O_3 ratio has a spatial gradient that changes over relatively short distances. The consequence of this finding, and its effect on policy design for O_3 precursor control alludes to the complexity of the system and the benefits of resolving atmospheric chemistry and pollutants at the neighborhood-scale.

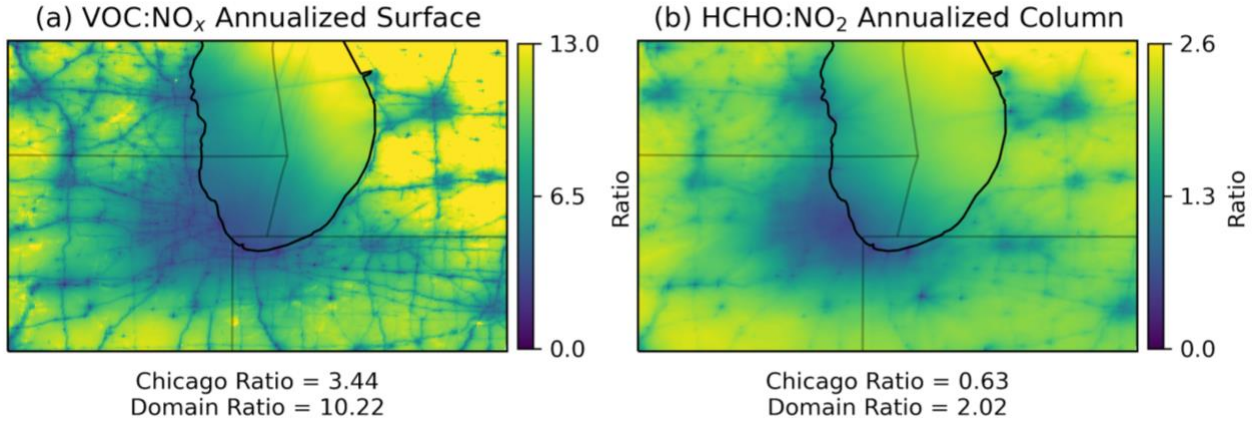


Figure 2.7 Daytime (7 am - 7 pm) annualized average of the (a) VOC:NOx surface ratio and (b) HCHO:NO_2 column ratios over the 1.3 km domain.

4 Discussion

In the above we present the first neighborhood-scale (1.3 km) two-way coupled WRF-CMAQ simulations focused on the southern Lake Michigan-Chicago region. We perform hindcast simulations of individual months from each season and assess the model's performance against meteorological and pollutant station observations, as well as against coarser resolution (4 km) simulations. Below we summarize our results and discuss notable findings and experimental caveats.

2.4.1 Summary of Model Performance

In our WRF-CMAQ simulations, we show that the WRF-simulated meteorological variables WD, WS, and T2 meet performance criteria suggested by Emery et al. (2001). The lowest performing simulated meteorological variable is WD, a variable that models have historically struggled to reproduce with high fidelity and which has previously been shown to be sensitive to model resolution, boundary layer parameterization, and land cover schemes (Carvalho et al., 2012). In our simulations, we find that WRF-CMAQ best-captures observed WD in January 2019 and October 2018, but struggles in August 2018 and April 2019, likely due to more diffuse warm season winds, similar to findings presented in Zhang et al. (2014). In addition, recent WRF simulations have demonstrated the influence of different lake temperature datasets on meteorology, particularly air temperatures and convection, in domains near to

Lake Michigan (Wang et al., 2022). Future work should assess the role of lake temperatures, and uncertainties therein, on the simulation of pollutants.

In the CMAQ portion of our two-way coupled WRF-CMAQ simulation, NO_2 concentrations show the best performance when compared against EPA criteria pollutant AQS station observations, including high correlations and low biases and errors. However, we note that over Chicago (Figure 2.6) and at EPA stations (Table 2.2), the highest concentration of NO_2 is simulated in October 2018, in contrast with NOx emissions which peak in January 2019 (Figure S1.6). This mismatch is likely a consequence of model biases. In January 2019, the NO_2 bias over the domain and at the nearest Chicago AQS Station (17-031-3103) is highest and negative (*NMB* = -26%). The relatively large negative January NO_2 bias is likely driven by an anomalously large negative bias in simulated RH compared to other months (i.e., domain-average RH mean bias is -7.5% and Chicago RH mean bias is -15%). The January NO_2 negative bias is found to be largest in the early morning hours, which is coincident with the simulated RH bias maximum (Figures S1.1-S1.3). Previous studies have demonstrated that lower RHs are associated with lower concentrations of NO_2 (Harkey et al., 2015). It is also notable that October 2018, the month with the highest simulated concentration of NO_2 over Chicago, and lowest, but positive *NMB* (2.6% over all stations, 4.2% at station 17-031-3103), shows the largest night-time buildup of NO_2 and O_3 (Figure S1.3), which is likely caused by low night-time titration (Sharma, et al., 2017). Low night-time titration has previously been demonstrated to be a consequence of weak vertical mixing in the ACM2 PBL scheme (Zhao et al., 2018).

Similar to NO_2, we find high correlations between observed and simulated O_3 concentrations, which reflects the strong anticorrelated relationship between O_3 and NO_2. However, unlike NO_2, O_3 is biased high across all seasons, with the largest biases occurring at night (except for January 2019). Previous studies have found that applying the ACM2 PBL scheme may contribute to excess nighttime vertical mixing, which can dilute NO concentrations at the surface and reduce O_3 titration as well as

transport ozone from the residual layer to the surface, both of which contribute to higher nighttime O_3 concentrations (J. Hu et al., 2016; Sharma et al., 2017). Transport from boundary conditions likely contribute to the high bias as well. We find that surface-level O_3 in CAM-Chem is consistently biased high, particularly at night (Table S13) which would contribute excess O_3 transport into our d02 (4km) simulation domain. Previous studies have found that an overestimation of NOx emissions can contribute to excess O_3 (Qin et al., 2019; Travis et al., 2016), however in our simulations NO_2 is biased low or negative on average (Table 2.2). Future studies could address the simulated O_3 biases by running sensitivities which target boundary layer physics (X.-M. Hu et al., 2012) or by modifying the urban mixing parameterizations, such as percentage of urban area (PURB) (Wang et al., 2021).

Compared to NO_2 and O_3, simulated $PM_{2.5}$ has a lower correlation with EPA AQS station observations, though our model performance is comparable to results reported by many previous studies (e.g., Wang et al. 2021 and Torres-Vazquez et al., 2022). Given that SO_2 contributes to secondary PM formation, the relatively poor performance of WRF-CMAQ simulated SO_2 likely influences the $PM_{2.5}$ model-observation agreement. Previous studies have reported that model-station agreement of $PM_{2.5}$ can be strongly influenced by wind direction, wind speed, transport, and emissions inventories (Hughes et al., 2021; Zhang et al., 2014) and it is likely that these factors also play a role here. For instance, in the results presented here, we employ MOVES2014a which does not account for emissions from off-network idling of vehicles. MOVES3, released in 2020 (Eyth, 2021), does include these processes, which may be critical for more accurate simulation of PM, particularly in high heavy-duty truck trafficked warehouse environments common within urban settings.

Since the two-way coupled WRF-CMAQ methodology employs nested domains of increasing spatial resolution, we take the opportunity to discuss differences, advantages, and disadvantages of neighborhood-scale (1.3 km) simulations versus those performed at coarser resolutions (4 km). By and large, when model results are assessed against meteorological and pollutant station observations, we find

incremental performance improvements at the higher simulation resolution. We do note a few occurrences of slightly degraded model-observation fidelity at higher-resolution (e.g., T2 and O_3), but primarily find that higher resolution simulations marginally improve hindcast simulations of both meteorology and atmospheric chemistry, like previous thematically similar studies (e.g., Torres-Vazquez et al., 2022).

We note that our 1.3 km to 4 km simulation comparison is not a pure resolution-focused sensitivity experiment. For example, MOVES emissions processing influences on-road sector emissions, and due to differences between the 1.3 km and 4 km WRF-simulated meteorology, on road emissions differ over roadways (Figure S1.6). However, total emissions are the same within the 1.3 km domain subset of the 4 km domain. Despite emissions differences of 1-6% over Chicago (Figure S1.6), pollutant concentrations differ by only 1-2% between the 1.3 km and 4 km simulations (Figure 2.6). Ultimately, we find that the most valuable feature of increasing model resolution comes from the finer characterization of emission sources, meteorology, and subsequent pollutant concentrations. To quantitatively demonstrate differences in pollutant gradients between model resolutions, we plot the spatial variogram of the model at EPA AQS observation sites and across random points in the model domain (Figure S1.11). We use the methods described in Marzban & Sandgathe, 2009 and Touma et al., 2018, which show that variograms can be used to identify atmospheric features in model simulations and validate model simulations with limited point-observations. Here we use variograms to (a) validate observed spatial pollutant gradients, (b) compare spatial gradients at observation points to simulated pollutants at different model resolutions, and (c) assess differences in simulated spatial variability at different model resolutions. We note that the observation-based variogram is noisy due to the unequal placement of stations over the domain and the low number of stations (n = 9, 63, 12 for NO_2, O_3, and $PM_{2.5}$) relative to the size of the domain (n = 90,720) (Figure S10a). At EPA observation sites, we find that the spatial variances of simulated NO_2 and O_3 are closer to observed variances at fine scales (<50 km) in the 1.3 km resolution simulation (Figure S1.10b). We also compute a “synthetic” variogram (as described in the Supplemental Discussion) computing the variogram at 75 random grid cells across the domain to grid cells within a 50 km radius.

We use this analysis to show that at short distances (<70 km), the 1.3 km simulation has higher pollutant variances across the domain (Figure S1.10c). As such, we show that at finer spatial scales (Figure 2.7) and across the domain (<70 km, Figure S1.11), the spatial gradients achieved in the 1.3 km output are measurably steeper and have greater fidelity to EPA observations. The ability to resolve air pollution gradients at neighborhood-scale resolutions, using physics- and chemistry-based numerical models, is critical for a number of air quality-relevant disciplines, and continued efforts should be made to both improve model performance and apply these tools to fundamental research queries in the fields of atmospheric chemistry, health, policy, and environmental justice.

2.4.2 Summary of Chicago Pollutant Findings

Characterizing neighborhood-scale spatial heterogeneities in pollutant concentrations over urban settings, such as Chicago, is critical for better understanding health impacts and constraining the contribution of pollutants to inequitable impacts across population subgroups. In our simulations, we find that Chicago has 2 to 5 times higher NO_2 and $PM_{2.5}$ concentrations than neighboring rural areas (Figure 2.2), and within city limits annualized pollutant concentrations between neighborhoods can vary by a factor of 1.8 (Figure 2.4). To highlight the utility of high resolution spatially resolved model simulations and intra-city pollutant differences, we analyze zonally averaged annualized pollutant concentrations from Chicago's western suburbs to Lake Michigan in the east (Figure 2.8). In both 1.3 km and 4 km simulations, pollutants over this zonal swath display a distinct west-to-east profile, with NO_2 and $PM_{2.5}$ peaking over the core of the city, and relatively high O_3 concentrations over the lake and in the western suburbs. In general, zonal patterns of O_3 concentrations are the inverse of simulated NO_2 concentrations, which is consistent with the NOx-saturated regime identified in the core of the city (Figure 2.7). This inverse pattern is replicated when comparing $PM_{2.5}$ to O_3, albeit with lesser fidelity. In addition to elevated NO_2 over the city center, a western peak is simulated near O'Hare International Airport (Figure 2.8a). Despite the inverse NO_2-$PM_{2.5}$-O_3 zonal pattern, some sections of the city do see relatively high co-occurring concentrations of NO_2, O_3, and $PM_{2.5}$. For example, Chicago's west side, near -87.8°W, has

relatively high concentrations of each pollutant (Figure 2.8a), likely due to the confluence of highways and industrial areas. The 4 km output shows 0.5 ppb NO_2 and -0.5 ppb lower O_3 in the suburbs (west of -87.8W), though the 1.3 km simulation shows slightly higher NO_2, O_3, and $PM_{2.5}$ over Chicago. Results such as these suggest that summarizing city-wide air quality using limited observations could be problematic as it does not capture the substantial spatial heterogeneity.

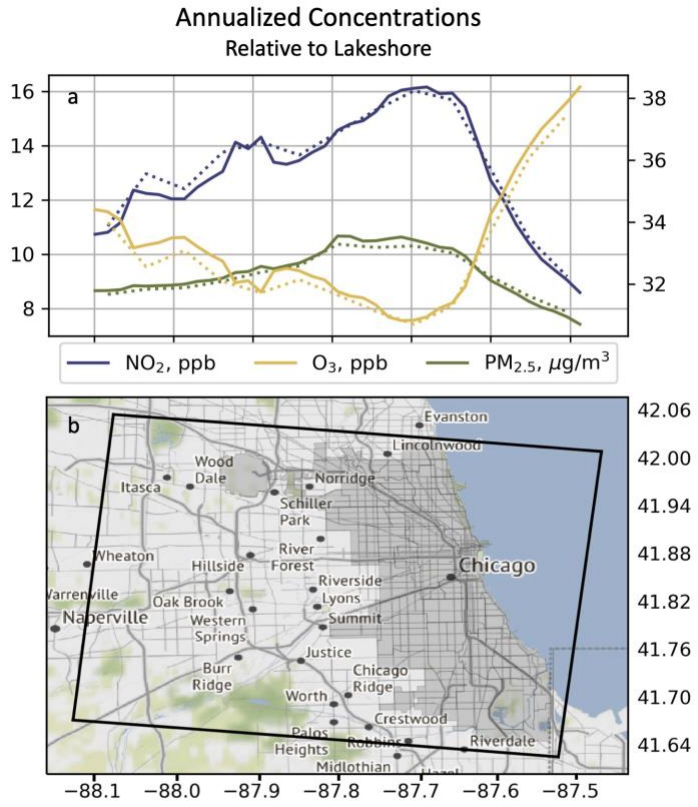


Figure 2.8 (a) Zonal average of pollutants over Chicago region, from suburbs to city center to lake. Mean pollutant concentrations of NO_2, $PM_{2.5}$, and O_3 are provided in and (b) the footprint of the averaging domain is depicted around Chicago. Solid line shows the 1.3 km output, while the dotted line shows the 4 km output. Concentrations are plotted across longitudes and stretch from the western suburbs to Lake Michigan in the east.

Neighborhood-scale simulations may also prove useful for health, environmental justice, and targeted abatement strategy investigations. Our finding of elevated west side pollutants (Figure 2.4 & Figure 2.7) is consistent with previous health-focused work that has identified elevated clusters of air

quality related diseases on Chicago's west side (Gupta et al., 2008). Beyond identifying locations of elevated pollutants, high-resolution CTMs may be used to design and test effective mitigation and abatement strategies. For example, given Chicago's current EPA O_3 non-attainment status, designing strategies that effectively target O_3 precursors requires foreknowledge of the City and region's chemical regimes (Figure 2.7). Simulations capable of characterizing neighborhood-scale O_3 and background VOC:NOx conditions can provide critical insights into amelioration strategies and can be used to predict and/or avert unintended consequences, such as localized O_3 increases due to NOx or VOC control measures. High resolution CTMs are particularly suited to this task, given their ability to constrain changes in secondary pollutants, a capability not available in reduced complexity models.

2.4.3 Limitations

Despite the promising neighborhood-scale results reported here, there are several caveats to bear in mind when considering our two-way coupled WRF-CMAQ results. Chief amongst these considerations is the use of four individual months and their annualized means to characterize the region's air quality and atmospheric chemistry regime. Neighborhood-scale CTM simulations are computationally expensive, which has limited our ability to simulate full seasons or multiple years. Previous studies have demonstrated that internal meteorological variability can have profound consequences on pollutant concentrations (Fiore et al., 2022; Garcia-Menendez et al., 2017), and this facet should be remembered when considering our results. A key example from our study is the high O_3 concentration in our April 2019 simulation. Typically, O_3 in this region peaks in July, however April of 2019 (our chosen simulation month) had higher O_3 concentrations than the typical summer O_3 season. A second key consideration of our study is that the EPA air quality monitoring system was not designed with CTM-validation in mind. AQS sensors are relatively sparse and very often not within urban settings. As such, we use EPA data here, but advocate for the use of hyper-local observing networks to operationally monitor neighborhood-scale air quality and perform model validation.

Further, we note limitations with our model configuration. To perform two-way WRF-CMAQ simulations, we use the default nudging coefficients for the FDDA process above the boundary layer to encourage model-station agreement. This choice has been shown to dampen fine-scale meteorological processes and alter aerosol feedbacks in higher resolution simulations, i.e., 4-12 km (Gan et al., 2016; Hogrefe et al., 2015). Future studies should explore the effects of nudging at even finer neighborhood-scale resolutions.

5 Conclusions

In the above, we present the first neighborhood-scale two-way coupled WRF-CMAQ simulations to be performed over a Chicago-centric southern Lake Michigan domain. Both the meteorological and chemical components of our model largely perform at or above recommended standards. We note that our 1.3 km simulation incrementally outperforms our 4 km simulation with respect to most air quality-relevant meteorological variables. In terms of chemical performance, we observe that the 1.3 km simulation outperforms the 4 km simulation with respect to grid cell-to-observation station comparisons for NO_2, O_3, $PM_{2.5}$, and CO concentrations. SO_2 is the only pollutant that showed higher model-observation fidelity at the coarser model resolution, but this was also the chemical with the lowest model-station agreement at both the 1.3 km and 4 km resolutions. Consideration of these performance assessments should be tempered by knowledge that both meteorological and pollutant observing networks allow for model-to-observation comparisons at a maximum of 0.1% of simulated grid cells.

Neighborhood-scale, 1.3 km simulations, are made possible by spatial surrogates curated for the region by LADCO. These surrogates facilitate the simulation of fine-scale features and processes, none more evident than the effect of resolving on-road emissions within urban settings, where we simulate anomalously high roadway-adjacent NO_2 and $PM_{2.5}$ concentrations, and anomalously low O_3 concentrations. Over our full simulation domain, we find that the highest concentrations of O_3 are found over Lake Michigan during warm season months, where concentrations are simulated to be a full 30% higher than the domain average. In the largest urban area simulated in our domain, Chicago, IL, we find

that concentrations of NO_2 are five times higher than the domain average, $PM_{2.5}$ three times higher, and O_3 slightly lower. We also note spatiotemporal O_3 regime variability within the full model domain, where simulated surface and column average VOC:NOx and HCHO:NO_2 ratios differ by season and location. Over the full domain, surface conditions are simulated to be NOx-limited, however over Chicago, conditions are simulated to be transitional – with column average ratios adding further nuance. Likewise over Chicago, our higher resolution simulations show higher average concentrations of NO_2 and $PM_{2.5}$ than our coarser model simulations, suggesting that coarser models may underestimate exposure to these pollutants and their associated health impacts. Lastly, within Chicago city limits, we find that pollutants can vary by a factor of ~2 between neighborhoods, a finding potentially corroborated by observed inequitable health outcomes.

Acknowledgments

We acknowledge support from the U.S. National Science Foundation Grant CBET-1848683 to DEH; McCormick Center for Engineering Sustainability and Resilience seed grant to DEH; The Ubben Program for Carbon and Climate Science postdoctoral fellowship to JLS; and The Data Science fellowship from the National Science Foundation Research Traineeship and Northwestern Integrated Data-Driven Discovery in Earth and Astrophysical Sciences to AM. We also extend thanks to 3 reviewers for substantive feedback on our study and D. Touma for variogram assistance.

Chapter 3

Intraurban NO_2 Hotspot Detection across Multiple Air Quality Products

This chapter is in review as: Montgomery, A., Daepp, M. I. G., Abdin, M., Choudhury, P., Malvar, S., Counts, S., Horton D. E. (2023). Intraurban NO_2 Hotspot Detection across Multiple Air Quality Products. Environmental Research Letters

Abstract

High-resolution air quality data products have the potential to help quantify inequitable environmental exposures over space and across time by enabling the identification of hotspots, or areas that consistently experience elevated pollution levels relative to their surroundings. However, when different high-resolution data products identify different hotspots, the spatial sparsity of "gold-standard" regulatory observations leaves researchers, regulators, and concerned citizens without a means to differentiate signal from noise. This study compares NO_2 hotspots detected within the city of Chicago, IL, USA using three distinct high-resolution (1.3 km) air quality products: (1) an interpolated surface from Microsoft Research's Project Eclipse – a dense network of over 100 low-cost sensors; (2) a two-way coupled WRF-CMAQ simulation; and (3) a down-sampled surface using TropOMI satellite instrument observations. We use the Getis-Ord G_i^* statistic to identify hotspots of NO_2 and stratify results into high-, medium-, and low-agreement hotspots, including one consensus hotspot detected in all three datasets. Interrogating medium- and low-agreement hotspots offers insights into dataset discrepancies, such as sensor placement, model physics, data retrieval caveats, and the potential for missing emission inventories. When treated as complements rather than substitutes, our work demonstrates that novel air quality products can enable researchers to address discrepancies in data products and can help regulators evaluate confidence in policy relevant insights.

3.1 Introduction

Disparate exposures to ambient air pollution contribute to racial and economic inequities in disease burdens (Hajat et al., 2015; Tessum et al., 2021). Despite the considerable progress of the U.S.

Clean Air Act in contributing to population-wide reductions in air pollution exposure (Currie & Walker, 2019), neighborhoods with the highest historic pollution exposure remain subject to relatively higher pollution levels (Colmer et al., 2020). Although both regional and intraurban gradients contribute to observed differences in exposures across population subgroups (Chambliss et al., 2021), localized hotspots—areas that consistently experience elevated air pollution levels relative to their surroundings—can identify important and often modifiable local emissions sources (Hajat et al., 2015; Clark et al., 2014; Chambliss et al., 2021). Contemporary U.S. zoning policies ensure that land use and emissions sources are clustered in space (Hirt, 2015), and local regulations like cumulative impacts ordinances (Lee, 2020) or federal policies like JUSTICE40 (The White House, 2021) increasingly seek to target areas where marginalized communities experience disproportionate pollution burdens. Yet sparse regulatory monitoring systems, such as the US EPA AQS network, are designed for regional monitoring and cannot provide insights on neighborhood-scale variations in pollutants or exposure. As such, better ways of identifying hotspots are needed to ensure that mitigation efforts serve the most affected communities. There are many pollutants in the urban air; among these is nitrogen dioxide (NO_2), which is implicated in premature mortality (Song et al., 2023) and morbidities including asthma (Hansel et al., 2008; Liu et al., 2019), impaired prenatal development (Huang et al., 2015; van den Hooven et al., 2012), and cardiovascular disease (Stieb et al., 2020). NO_2 also contributes to the formation of other health-hazardous pollutants and can exhibit steep intraurban spatial gradients (Chambliss et al., 2021). NO_2 concentrations are highest near emissions sources like power plants (Liu et al., 2016) and roadways (Karner et al., 2010), with concentrations accumulating during stagnant meteorological conditions (Goldberg et al., 2020) or during high-emitting periods like morning rush hour (Zhang & Batterman, 2013). Due to the pollutant's short lifetime, NO_2 hotspots can be difficult to observe unless monitoring instruments are collecting data when and where NO_2 is emitted.

To address the need for routine monitoring of NO_2 at intraurban scales, researchers have taken a variety of approaches including: using high-density, low-cost sensor networks (Jain et al., 2021; Weissert

et al., 2020), conducting high-resolution chemical transport modeling (Di et al., 2019), and over-sampling satellite imagery (Goldberg et al., 2021; Dressel et al., 2022). The resulting data products have the potential to improve the characterization of NO_2 in urban environments, but they are also subject to important limitations with respect to spatial and temporal data availability or data validity.

Dense citywide sensing networks provide insights in real time, capturing important diurnal and spatial variation in NO_2 concentrations. To enable sampling at the scales needed for adequate coverage, sensors must be affordable and easily replaceable—but ensuring low cost requires important tradeoffs with respect to accuracy (Morawska et al., 2018; Larkin et al., 2017). Even with dense coverage, gaps remain, and thus researchers must apply interpolation schemes that contribute additional uncertainty (Schneider et al., 2017; Gressent et al., 2020).

Chemical transport models (CTMs) exploit scientific understandings of chemical and physical processes to provide spatially and temporally fine-scale estimates of air quality, so calibration and validation of these models can lead to new insights, in addition to reducing errors. However, outside of large-scale airborne monitoring campaigns (e.g., Abdi-Oskouei et al. 2020; Torres-Vazquez et al. 2022), researchers are limited in their ability to adequately validate and calibrate CTMs at fine spatial scales. The most commonly available validation data are spatially-limited regulatory networks or temporally-limited satellite observations (Wong et al., 2012; Kuhlmann et al., 2015), and thus researchers face challenges when attempting to routinely calibrate CTMs at fine spatiotemporal scales. Reduced complexity models (RCMs) can also produce high spatial resolution air quality products by simplifying or forgoing chemistry and physics, but instead leveraging statistical or machine learning relationships between emitting activities and pollutants (Wang et al., 2023; Burke et al., 2021; Tessum et al., 2021). These RCMs reduce chemical and meteorological output to achieve their gains and are shown to be less accurate in simulating secondary pollutants compared to CTMs, but these products are likewise challenged to validate their output on the scales at which they resolve pollutants.

Like CTMs, satellite observations can produce air pollution estimates across geographies. To create intraurban estimates of NO_2, satellite data products are oversampled (i.e., averaged from native-resolution observations to a regularly spaced grid) to provide high-resolution insights (Goldberg et al., 2021; Dressel et al., 2022). A limitation of satellite observations is that it measures the entire vertical column, rather than the "nose-level" pollution that affects human health. Moreover, existing observing systems provide just one or two daily snapshots, which excludes important diurnal variation (Penn & Holloway, 2020), as well as large swaths of data due to cloud cover and other meteorological conditions (Van Geffen et al., 2020). However, new geostationary satellites promise to ameliorate some of these issues, indicating these new data sources are likely to play a critical role in diversified intraurban observatories. In this paper, we apply a hotspot detection algorithm to identify clusters of elevated NO_2 pollution in each of three state-of-the-science datasets. We refer to post-processed data from each product as a surface, i.e., (1) a machine learning-interpolated surface built from a dense, citywide low-cost sensing network, (2) a high-resolution CTM simulated surface, and (3) an over-sampled satellite-observation derived surface. We compare the location, extent, and temporal persistence of hotspots across data surfaces, and identify intra-urban areas of high-, medium-, and low-agreement. We find one high-agreement Consensus hotspot and interrogate disagreements between datasets for additional insights. Through this work, we offer a method to amplify signal and better understand noise across multiple novel high resolution air quality data products.

3.2 Study Domain

Our study focuses on Chicago, Illinois, which is located on the southwest coast of Lake Michigan in the central United States (Figure 3.1). Our motivation to focus on Chicago is two-fold: (1) Chicago is a large source of NO_x emissions in the Great Lakes Region, contributing to elevated NO_2 and secondary pollutant concentrations such as O_3, both locally and regionally. (2) Chicago also has several innovative data sets for neighborhood-scale air pollution research, including a dense, citywide low-cost sensing

configuration (Daepp et al., 2022) and a recently validated, high-resolution WRF-CMAQ simulation configuration (Montgomery et al., 2023).

Given seasonal differences in emissions and meteorology, we analyzed data from one warm month and one cool month (August 2021 and February 2022). Chicago's August 2021 mean temperature was +1.7°C warmer than average (NWS) and the Eclipse sensing network, which was deployed in July 2021, had largely stabilized, with relatively few devices needing to be relocated or replaced after initial adjustments during the first month of the deployment (Daepp et al., 2022). In winter, we selected February 2022 due to its low average temperatures (-0.8°C from climate normal, NWS) as well as evidence of similar or fewer Eclipse sensors with missing data (n = 89) compared to other cool months (n = 92 for December, 82 for January).

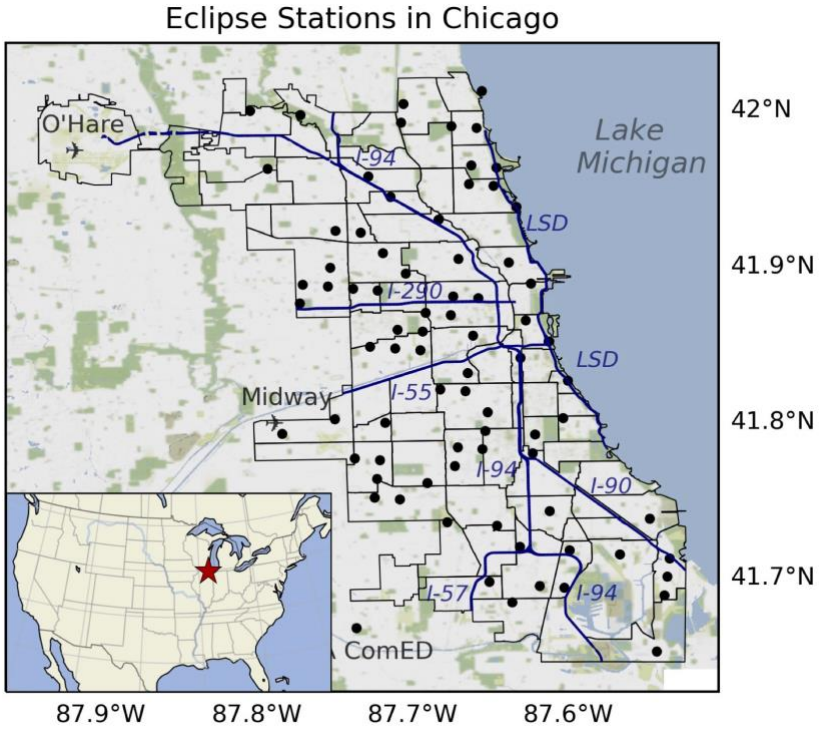


Figure 3.1 The city of Chicago, IL with major geographic features such as airports, highways (navy), and neighborhood boundaries (black) delineated. The locations of the NO_2 Eclipse sensors used in this study are marked as black circles. In the lower left corner, Chicago is marked as a red star within the map of the United States.

3.3 Methods

3.3.1 Air Quality Data Products

To create the Eclipse Network surface, we obtain ground-level measurements of NO_2 from the Microsoft Research Eclipse sensor network. The Eclipse network is comprised of over 100 low-cost air quality sensors around Chicago, deployed through a collaboration between Microsoft Research and JCDecaux Chicago— the local subsidiary of the world's largest outdoor advertising agency, JCDecaux SA— which maintains over 1,000 bus shelters across all geographic sectors of the city. The sensing hardware and network design are described in (Daepp et al., 2022). Devices were allocated to 80 sites using a stratified random sampling approach following Matte et al. (2013) and 26 additional sites recommended by local and community partner organizations. Because low-cost sensors are subject to error and noise, the network additionally included co-locations with EPA regulatory monitors (Clements et al., 2022). We used the ongoing co-location data to develop a calibration algorithm that improved accuracy relative to gold-standard EPA regulatory monitoring data (Table C1). To create a 1.3 x 1.3 km-gridded high-resolution daily and monthly gridded product for this study, we tested several machine-learning and geostatistical methods for spatial data interpolation and selected the best performing approach, a random forest (RF) model (Table C2). Further details on the calibration approach and the interpolation methods can be found in Appendix C.

To create the CTM data surface, we run the two-way coupled Community Multi-scale Air Quality (CMAQ, v5.2; Byun & Schere, 2006) and Weather Research and Forecasting (WRF, v3.8; Skamarock et al., 2008) modeling system (WRF-CMAQ, Wong et al. 2012). We created a custom emissions dataset using the EPA 2017 Sparse Matrix Operating Kernel of Emissions (SMOKE) Modeling product (Eyth et al., 2019), with which we downscaled the National Emissions Dataset using 1.3 km spatial surrogates from LADCO (LADCO, 2020). Further configuration and validation of our August 2021 and February 2022 WRF-CMAQ are described in Appendix A.

To create the TropOMI surface, we obtained geospatially continuous NO_2 observations from the Tropospheric Monitoring Instrument (TropOMI) aboard the Sentinel-5p satellite. The Sentinel-5p satellite

is a geo-orbiting satellite, so the TropOMI instrument provides daily retrievals of atmospheric species at 13:30 local time with a nadir resolution of 5.5 x 3.5 km for each grid cell. We used the L2 NO_2 product, processed by the Royal Netherlands Meteorological Institute (KNMI), which applies the DOMINO algorithm to convert Level-1b irradiance measurements in the 405 – 465 nm range into NO_2 vertical column density (VCD) (Van Geffen et al., 2020). To create a cross comparable product against our other air quality data sets, we regridded the L2 daily observations of NO_2 from TropOMI to the 1.3 km x 1.3 km WRF-CMAQ grid. Full satellite processing information and ground-based comparison is available in Appendix B.

3.3.2 Land Use and Social Characteristics of Hotspots

To provide additional context and discussion for the detected hotspots, we analyze land use characteristics and socioeconomic variables that prior research has shown are commonly associated with NO_2 pollution (Larkin et al., 2017). We assess hotspot relationships with zoning footprints, highway locations, and traffic speeds using data from the Chicago Data Portal (https://data.cityofchicago.org/); greenspace using MODIS normalized difference vegetation index (NDVI) obtained from the Planetary Computer (https://planetarycomputer.microsoft.com/); and census tract-level income and demographic data from the 2016-2020 American Community Family Survey (Manson et al., 2022). We regrid datasets using area-weighted averages (zoning, traffic, and socioeconomic data) or bilinear interpolation (NDVI). To assess the relationship of each characteristic to each hotspot, we compare the distribution of characteristics within hotspots to the city. Relationships are considered statistically robust at the 95th percentile confidence level ($p < 0.05$) when their t-test with Bonferroni adjustment for multiple testing yields $p_{Bonferroni} < 0.0167$.

3.3.3 Hotspot Detection

We define NO_2 hotspots using the Getis-Ord G^*_i statistic (hereafter, G^*_i), which is used to identify areas where significantly high or low values are spatially clustered (Getis & Ord, 1992; Ord & Getis, 1995). The G^*_i statistic uses z-scores to identify areas where a grid cell and its neighbors' values are significantly

higher or lower than would be expected if values were distributed randomly across space. In this study, we focus on areas of high NO_2 concentrations and assign hotspot status when significance testing exceeds the 95% confidence level ($p < 0.05$). Because we cross-compare 3 data products, we alter our significance screen with a Bonferroni adjustment for multiple testing such that $p_{Bonferroni} < 0.0167$. We conduct our G^*_i analysis using the Python package ESDA (Rey & Anselin, 2007). To compute the G_i* statistic for a given grid cell *i*:

$$G_i^* = \frac{\sum_{j=1}^{n} w_{i,j}x_j - \bar{X}\ \sum_{j=1}^{n} w_{i,j}}{S\sqrt{\frac{\left[n\ \sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j}\right)^2\right]}{n-1}}} \tag{1}$$

where x_j is the NO_2 concentration for grid cell *j* and *n* is the total number of grid cells. We calculate the spatial weight between grid cells *i* and *j* using the Queen's Contiguity method ($w_{ij} = 1$ if two cells are adjacent and 0 otherwise). Moreover,

$$\bar{X} = \frac{\sum_{j=1}^{n} x_j}{n} \tag{2}$$

and

$$S = \sqrt{\frac{\sum_{j=1}^{n} x_j^2}{n} - \left(\bar{X}\right)^2} \tag{3}$$

We compute the G^*_i statistic on each high-resolution data product's NO_2 surface to identify hotspots for each month considered (Figure S2.1). We then assess agreement across product surfaces (Figure S2.2). We define high-agreement hotspots as areas whose grid cells meet G^*_i hotspot criteria in all 3 datasets. Medium-agreement hotspots are areas whose grid cells meet hotspot criteria in 2 of the 3 surfaces, and low-agreement hotspots meet criteria in just one. We conduct this agreement analysis for each surface-pairing and each month (Figure 3.4). As an additional meteorological robustness screen, we test the effect of wind direction on identified hotspots by determining daily average wind direction, and then binning days according to quadrant, i.e., northeast, southeast, southwest, and northwest. For each data product, we

then average daily NO_2 concentrations and apply the G^*_i statistic to determine if wind direction substantially influences a hotspot's location or spatial extent (Figures S2.5, S2.6).

3.4 Results

3.4.1 Comparison of High-Resolution NO_2 Surfaces

To demonstrate the relative abundance of NO_2 concentrations across the city and differences therein, we begin by normalizing each NO_2 data product independently (Figure 3.2). We also compute the weighted area average of the normalized surfaces (μ) to demonstrate the relative difference in concentration magnitudes between products. Each surface provides a distinct spatial pattern of NO_2 pollution over Chicago, although pollutant patterns are not necessarily congruous. All data surfaces show relatively high NO_2 on the western edge of the city. However, over the remainder of the city elevated NO_2 concentration footprints do not consistently overlap across the three datasets. Our Eclipse surface has the greatest normalized city-wide mean concentration, with distinct areas of elevated NO_2 whose locations differ across months (Figure 3.2a,d). In contrast, the WRF-CMAQ NO_2 surface has lesser month-to-month variation, with elevated concentrations largely coincident with highways and the city center, a.k.a., "The Loop" (Figure 3.2b,e). Lastly, the TropOMI surface indicates elevated NO_2 concentrations on the west side of the city during both months, but few elevated areas nearer Lake Michigan (Figure 3.2c,f).

To quantitatively assess dataset similarity, we conduct pairwise grid cell-to-grid cell comparisons and compute Pearson correlation coefficients (r) and mean biases (mb) between surfaces. The datasets show positive linear relationships across surfaces, with correlation coefficients ranging from 0.1 to 0.7 (Figure S2.1). Eclipse-derived and WRF-CMAQ surfaces are consistently positively correlated ($r \geq 0.5$), though concentration estimates from Eclipse are on average 0.7 ppb higher in February and 4.8 ppb lower in August (Figure S2.1c). Consistency between the TropOMI surface and other surfaces varies by month (Figure S2.1a, b). The TropOMI-derived surface is strongly positively correlated with the other data products in August ($r = 0.6$), but in February correlations are considerably less ($r = 0.1$ for WRF-CMAQ and $r = 0.3$ for Eclipse). From a mean bias perspective, TropOMI concentration magnitudes show greater

similarity to Eclipse than to WRF-CMAQ (Figure S2.1a, b), although it should be reiterated that TropOMI captures columnar average NO_2, rather than the nose-level conditions captured by Eclipse and WRF-CMAQ.

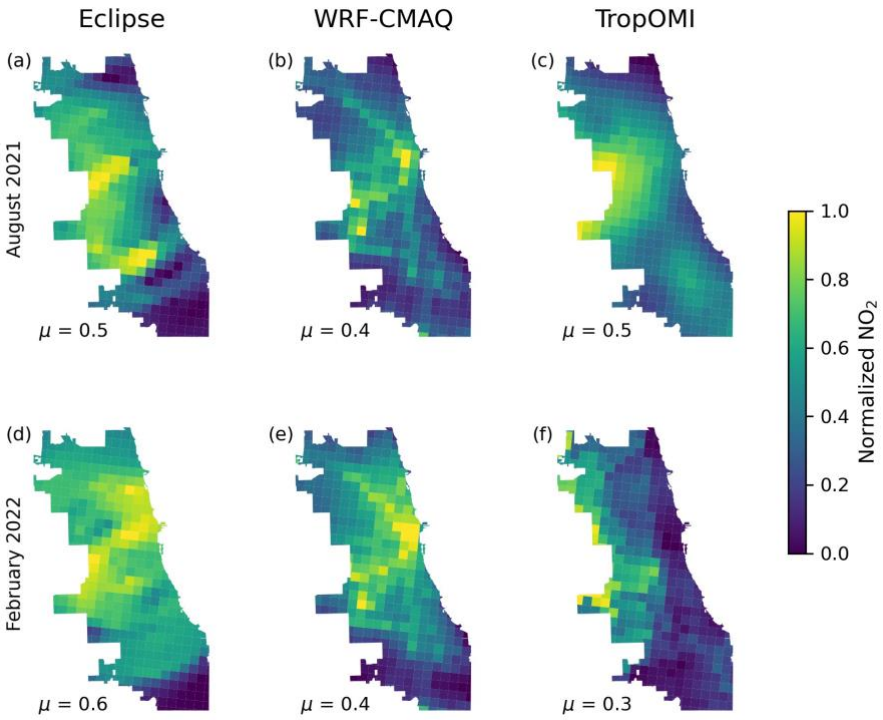


Figure 3.2 Normalized NO_2 concentrations for 1-month average August 2021 and February 2022 from the (a,d) interpolated Eclipse sensor network, (b,e) WRF-CMAQ simulations, and (c,f) TropOMI satellite observations. In the lower left, we provide the weighted city-wide area average (μ) of normalized NO2 concentrations for each data product.

3.4.2 Identification of Hotspots

Given the comparisons above, we conclude that while each high-resolution surface shows distinct intra-urban variation, concentration estimates have varying levels of consistency at a grid-cell level. To identify high-impact areas of agreement despite the datasets' differences, we apply the G^*_i hotspot identification statistic to each data surface (Figure S2.2) and assess levels of agreement between datasets according to overlapping hotspot footprints (Figure S2.3). We primarily focus our results on high- to medium-agreement hotspots (i.e., 3/3 and 2/3 datasets in agreement), but discuss low-agreement hotspots (1/3) where appropriate.

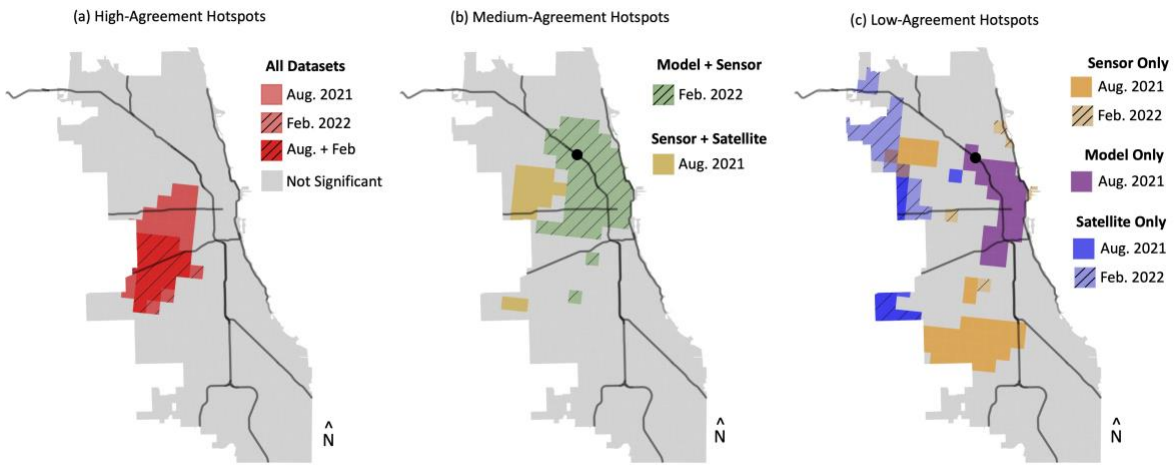


Figure 3.2. High-, medium-, and low-agreement NO2 concentration hotspots. Agreement is based on the number of datasets with overlapping footprints that meet G∗i hotspot criteria, i.e., 3/3 = high, 2/3 = medium, and 1/3 = low. Datasets include the Eclipse sensor network, WRF-CMAQ model, and TropOMI satellite. The black dot in panels (b, c) is the location of the I-90/94 Kennedy Expressway EPA monitor discussed in Figure 3.4.

3.4.3 Consensus Hotspot

We identify one high-agreement region of the city – a corridor on the west-central margin of the city – in which all three data products indicate statistically significant high NO_2 (Figure 3.3a). This Consensus hotspot is found in both months considered, but the spatial footprint of the August hotspot is twice the area of the February hotspot due to a contraction of the TropOMI February hotspot (Figure S2.2). For each dataset, the average NO_2 concentrations that comprise the identified hotspot range from 16 to 31% higher than the Chicago area average in August and 18 to 22% higher in February (Table S2.1, S2.2). We find that this Consensus hotspot is robust to wind direction, as G^*_i criteria are met in all three datasets regardless of wind direction or month of consideration (Figures S2.4, S2.5).

3.4.4 Medium-Agreement Hotspots

While there is high agreement amongst datasets on the Consensus hotspot depicted in Figure 3.3a, the datasets differ on its spatial extent. In the Eclipse and TropOMI datasets, we find an adjoining medium-agreement hotspot that extends further north and west of the Consensus hotspot (Figure 3.3b; yellow grid cells). Given the two underlying data sources, we refer to this area as an Observational hotspot. We note

that the Observational hotspot is only found in August. In the August observational datasets, we observe NO_2 concentrations that are 14 to 27% higher than the city-wide average (Table S1). While NO_2 concentrations in this area of our August WRF-CMAQ simulation do not meet G_i^* hotspot criteria, they are modestly (4%), but insignificantly, higher than the city average (Table S1, S2). Due to the missing data in TropOMI coverage on a daily basis, we are unable to systematically analyze the persistence of TropOMI hotspots across wind directions.

A second medium-agreement hotspot is found in February and is identified in both the WRF-CMAQ and Eclipse datasets (Figure 3.3b, green hatched grid cells). As this area straddles DuSable Lake Shore Drive (LSD) and I-94 and I-290 (consult Figure 2.1 for roadway labels), we refer to it as the Highway hotspot. Grid cells in the Highway hotspot have NO_2 concentrations that are 26% greater than the Chicago average (Table S2.1). The February Highway hotspot meets G_i^* criteria regardless of wind direction (Figures S2.4, S2.5). In addition, we note that in August, WRF-CMAQ simulates an NO_2 hotspot with a similar spatial footprint, with concentrations that are 35% higher than the city-wide average (Figure 3.3c; purple grid cells). However, G_i^* hotspot criteria are not met in either observational dataset in August, thus the area only attains low-agreement hotspot status in our analysis. We note that this area includes Eclipse and EPA monitors fortuitously located in close proximity, allowing us to explore potential reasons for this model-observation disagreement below.

3.4.5 Highway Hotspot Interrogation

To identify factors that contribute to the Highway hotspot disagreement between WRF-CMAQ and Eclipse, we compare diurnal WRF-CMAQ and Eclipse variability to observations from an on-road EPA monitor on the hotspot's northern edge (Figures 3.5a, b). The EPA monitor is located on an elevated stretch of the I-90/94 Kennedy Expressway and is fortuitously located less than 150 m from an Eclipse monitor (Figures 3.4c, 5c). Hourly time-series data at this location in February 2022 reveal relatively strong agreement between the EPA sensor observations and both WRF-CMAQ (NMB = -1.1%) and Eclipse data (NMB = -8.1%) (Table S2.2). However, product agreement with the EPA sensor is lesser in

August 2021. In August, both Eclipse and WRF-CMAQ have higher magnitude biases compared to February, with Eclipse bias higher and negative (NMB = -32.3%) and WRF-CMAQ bias higher but positive (NMB = 39.6%) (Figure 3.5a, Table S2.3). We note that WRF-CMAQ's bias is largely attributable to the model's nighttime bias (NMB = 65.1% v. 16.0% during the day; Table S2.3), a finding discussed in Montgomery et al. (2023) and Zhao et al. (2019) related to low model-simulated nighttime titration due to weak vertical mixing in our chosen planetary boundary layer scheme. In contrast, Eclipse's biases are similar at night (NMB = -30.8%) and during the day (NMB = -34.3%) (Table S2.3). We explore Eclipse's consistent bias below.

One potential explanation for the difference in NO_2 hotspot classifications at this location is related to the placement of the Eclipse sensor relative to the primary pollution source i.e., highway traffic. Previous work has demonstrated that pollution measurements near highways can be substantially influenced by both the distance of the sensor from the highway and the sensor height (Salmond et al., 2013; Gilbert et al., 2003). At this location, the EPA sensor is on an elevated highway 7 m above ground level, while the Eclipse sensor is 150 m distant and 2.4 m above ground level on a nearby bus station (Figure 3.4c). Given previous reports of sensor distance/height impacts on pollutant concentration measurements, we explore distance/height relationships between highways and Eclipse sensors across our city-wide network. In Chicago, uncovered Class-1 roadways, i.e., highways with heavy-duty vehicle traffic, exist at three elevations: ground-level (g.l.), below-grade open cut (a.k.a. recessed, 3 m below g.l.), and elevated (4 m above g.l.) (City of Chicago). Only one Class 1 highway in the city is mostly (> 90% length) at ground-level, DuSable Lake Shore Drive, however, heavy-duty vehicle and commercial traffic are restricted. We therefore exclude the ground-level highway from this analysis. We first examine the relationship between Eclipse sensors within 2.5 km of Class-1 highways and monthly average NO_2 concentrations. Within 2.5 km of Class-1 roads, there are 20 sensors near recessed highways and 15 near elevated highways (Figure S2.6a). We find no robust relationships ($r < 0.05$) between an Eclipse sensor's

distance from a highway and its average reported monthly NO_2 concentration regardless of highway elevation or month (Figure S2.6 b,c). Next, we assess the relationship between highway elevation and

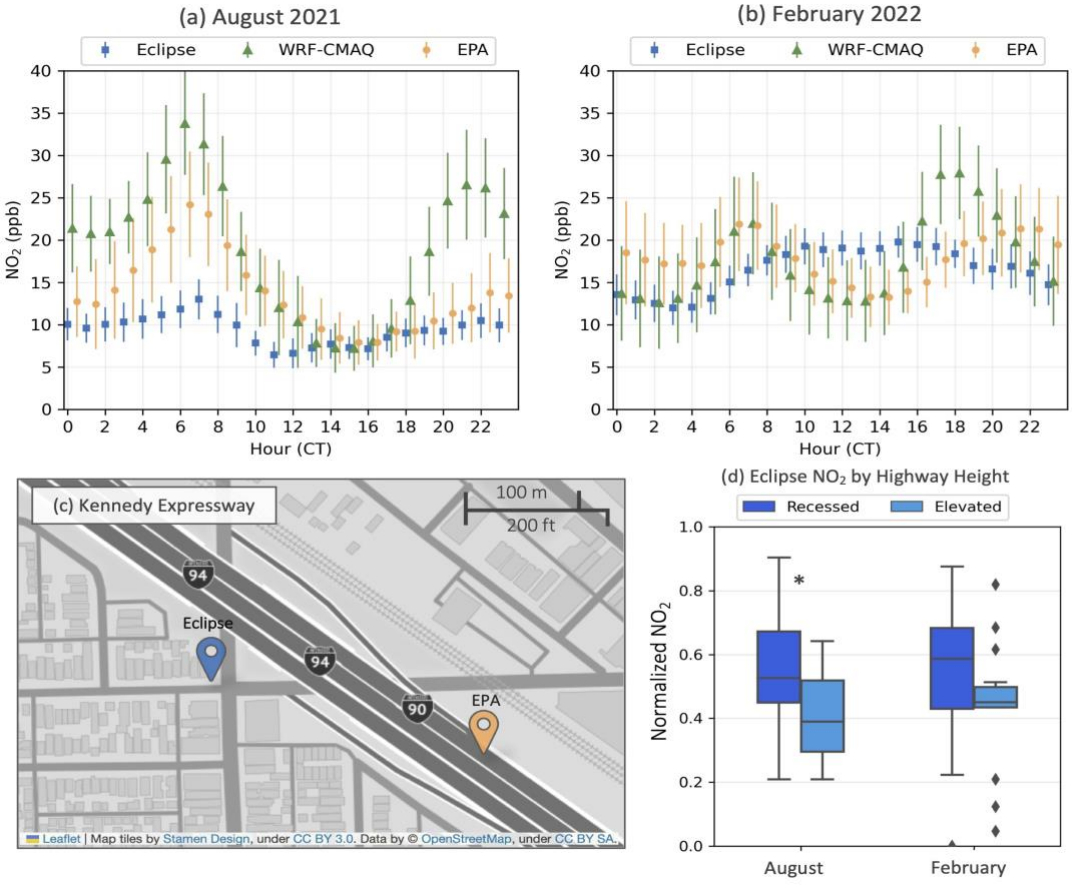


Figure 3.3 Highway hotspot discrepancy investigation. (a) August 2021 and (b) February 2022 average diurnal NO_2 concentration variations from an Eclipse sensor(blue), an EPA sensor (orange), and the co-located WRF-CMAQ grid-cell (green). (c) Precise locations of the EPA and Eclipse sensors relative to the I-90/94 Kennedy Expressway. The Eclipse sensor is at ground-level (i.e., mounted on a bus stop enclosure) while the EPA sensor is on an elevated highway 7 m above ground level. In (d) we compare normalized NO_2 concentrations from all near-highway (< 2.5 km) Eclipse sensors (n = 35), grouped by highway heights, i.e., recessed (n = 20) or elevated (n = 15). The asterisk indicates statistically robust distribution differences while diamonds show outliers. Co-located CMAQ grid cell data are presented in Figure S2.1, S2.2.

Eclipse reported NO_2 concentrations. We find that in August, mean NO_2 concentrations at Eclipse sensors near elevated highways are 12% lower in magnitude (-1.2 ppb) than concentrations reported by sensors near recessed highways (Figure 3.3d), a difference that is statistically robust (p = 0.03, Table S2.4). A similar, although muted and not significant pattern holds in February (13%, -1.9 ppb, p = 0.19, Table

S2.4). We hypothesize that systemic differences in NO_2 concentrations reported by Eclipse sensors, such as differences driven by highway elevation, contribute to the high negative Eclipse bias we find at the co-located EPA and Eclipse sensors near the I-90/94 Kennedy Expressway (Figure 3.3c), and speculate that this systemic bias may be the reason that Eclipse data near this location does not meet G^*_i criteria in August.

Highway elevations are not explicitly modeled in the WRF-CMAQ. We note that this model limitation likely contributes to the NO_2 concentration differences noted between the Eclipse sensors and WRF-CMAQ simulations. We perform a highway elevation analysis with WRF-CMAQ (Figure S2.7), like the Eclipse analysis reported in Figure 3.3d. That is, we bin model grid cells that contain the Eclipse sensors assigned to the recessed and elevated highway categories used above. We find small (3-5%) and insignificant ($p > 0.05$, Table S2.4) differences in mean NO_2 concentrations for both highway elevations in WRF-CMAQ. Given the lack of differentiated highway elevations in WRF-CMAQ, the lack of simulated NO_2 concentration differences near highways is expected, however this WRF-CMAQ-Eclipse contrast may help explain some of the model-sensor mismatch inherent to dataset comparisons and hyperlocal topographies.

3.4.6 Land-use and Social Characteristics of Hotspots

In the high-agreement Consensus hotspot, we find that compared to the city-wide average, this area is characterized by significantly higher industrial zoning (2.5 times higher), lower greenness (NDVI, -23%), and a significantly high proportion of Hispanic or Latino residents (2 times higher) relative to the Chicago average. No other attributes pass our robustness screening. The population of the Consensus hotspot consists of 332,000 – 501,000 people, depending on the month. We do note that the Consensus hotspot has a non-significant though high population density (+12%), with lower income (-9%) and lower public assistance (-5%) per capita, though not robustly different than the city on average (Table S1, S2).

In the medium-agreement Observational hotspot, we find no significant relationships with land use characteristics. In contrast with the high-confidence hotspot, the Observational hotspot has more

residential zoning (+17%), but similar industrial zoning to the city-wide average. Given that emissions data in WRF-CMAQ is partially determined by land-use characteristics, it is perhaps not surprising that NO_2 concentrations are not higher in this area in the model output; if the Observational evidence reflects a true signal, there may be an additional emissions source not currently captured in the emissions data. Notably, the medium-agreement Observational hotspot has a significantly higher Black population relative to the other hotspots and to the Chicago average (i.e., 3.2 times higher). The population of the Observational hotspot is approximately 58,000 people. Other nonsignificant attributes include a high population density (+34%), lower income (-20%) and higher public assistance (-13%) per capita, though not significantly different than the city on average (Table S2.1, S2.2).

In the medium-agreement Highway hotspot, the grid cells share just one robust land use characteristic with the Consensus hotspot, low NDVI (-25%). The Highway hotspot does not contain significant industrial zoning; instead, compared to the city-wide average it has significantly higher population density (+63%), more highway coverage (3.3 times higher), higher commercial zoning (+25%) and lower residential zoning (-50%) (Table S2.1, S2.2). Due to the high population density, this area also has a high population (244,000 – 1.1 million). Importantly and somewhat obviously, the Highway hotspot has significantly more traffic than the city-wide average, with 61% more arterials and 3.4 times the average bus speeds. The income and assistance in this hotspot are significant, as the Highway hotspot corresponds to an affluent part of the city (with income per capita nearly 3 times higher than the city average), though there is also high public assistance in this area (+50% public assistance, Table S2.1, S2.2).

3.6 Discussion

Each of our identified hotspots can offer insights for researchers and regulators. To provide additional context and discussion for the detected hotspots, in the discussion we include land use characteristics and socioeconomic variables that prior research has shown are commonly associated with NO_2 pollution (Larkin et al., 2017).

First, the high-agreement Consensus hotspot identifies a large, contiguous region with high NO_2 concentrations relative to the city-wide average. The Consensus hotspot affects between 332,000 – 501,000 people (Table S1, S2). The Consensus hotspot is evident across different wind directions and the affected area has significantly higher industrial zoning and low greenness – factors commonly associated with higher NO_x emissions. Taken together, this evidence suggests that local sources, rather than regional transport, contribute to the elevated NO_2 concentrations. Further, the Consensus hotspot comprises an area that is majority Hispanic or Latino (54%; Table S1, S2). Although the estimates from each data product, separately, are subject to concerns regarding potential sources of bias and noise, the consistency of results across data sources as well as the urgency of EJ-related health inequities suggests that this area should be prioritized for clean air interventions.

Second, the medium-agreement Observational hotspot may identify an area where model simulations could be improved, whether through emissions inputs or model physics. In the Observational hotspot, land-use characteristics do not indicate a source of high-emissions. However, since both Observational datasets support its presence, an emission source may be missing in the underlying emissions data. Previous studies have used Observational datasets to constrain NO_x emissions (Goldberg et al., 2022) and to identify specific NO_2 emission sources (Georgoulias et al., 2020; Zhang et al., 2022). However, low data coverage from TropOMI could impact the identification of the Observational hotspot, particularly in the winter when meteorological conditions are not conducive to TropOMI observations. This data scarcity highlights the utility of upcoming remote sensing technologies like TEMPO (Naeger et al., 2021) that will provide higher spatiotemporal coverage, which in turn could better identify hotspots and help constrain NO_x emission sources. Beyond emission uncertainties, the choice of model physics and parameterizations in WRF-CMAQ can bias simulated NO_2 concentrations due to poorly simulated meteorological processes and/or challenges associated with urban settings (Gilliam et al., 2015; Montgomery et al., 2023; Pleim et al., 2014). The Observational hotspot identifies an intra-urban area wherein the causes of model-observation mismatch should be thoroughly investigated, so as to determine

whether a NO_x emission source is missing or if a modification of model physics better captures the observed build-up of NO_2.

Since the Observational hotspot is comprised of twice the average Black population compared to the city average (Table S2.1, S2.2) and sits at the northern edge of the high-agreement Consensus hotspot, determining the validity of the Observational hotspot is an important question with environmental justice and regulatory implications. The affected area has relatively more Black residents, lower incomes, and more residents receiving direct public assistance compared to the city average or the other identified hotspots, and thus may again indicate an area where excess emissions constitute an environmental justice burden. Given only medium agreement amongst air quality data products, this area would be well-served by additional routine monitoring specifically for NO_2 or a mobile monitoring campaign that could better evaluate the hotspot's "true" bounds under a variety of ambient conditions (Chambliss et al., 2021). Third, the medium-agreement Highway hotspot highlights an area where both the model and Eclipse data indicate an NO_2 hotspot associated with significantly higher traffic for 1 of the 2 months considered (Table S2.1, S2.2). In August, the Eclipse sensor network does not identify this hotspot. We show that in August, sensors placed below elevated highways report less NO_2 relative to their counterparts near recessed highways (Figure 3.4). While the low-cost sensor network has consistent placement with respect to height and location at ground-level, the placement of sensors near high-emitting sources like highways were not standardized, which complicates the model-sensor comparison. This finding is not discussed in the literature when creating high-density, intra-urban sensor-model comparisons, and this study highlights the importance of this hyperlocal interface for model-sensor comparison. However, in February, we identify a hotspot in the Eclipse data in this area. We note that in February, differences between elevated and recessed Eclipse NO_2 concentrations are less pronounced than in August, potentially related to the lower boundary layer and lesser dilution in the cool season, a phenomenon that is apparent in both Eclipse and CMAQ data (Figure 3.2d,e).

Whether the observed difference in NO_2 across elevated and recessed highways reflects an actual difference in exposure depends on the research question: for regulators seeking to quantify traffic emissions, our findings highlight the importance of placing sensors at the level of elevation at which emissions occur. But for public health researchers and practitioners who are concerned with the pollution levels where people breathe, ground-level sensing captures a meaningful difference in the adverse effects associated with highway heights and shows that CTM output may need to be adjusted for use in hyperlocal pollutant exposure and health impact quantification. This finding further shows the benefit of comparatively evaluating multiple different air quality data products for producing new and valuable insights.

The work reported here is subject to several important limitations. First, we examine just two months of data due to the computationally intensive requirements of running WRF-CMAQ; although we chose these months to be representative of one summer and one winter month based on meteorological conditions, further investigation is needed to evaluate the persistence of observed hotspots over other periods. Our ability to identify hotspots in either Observational dataset is further limited by missing data, particularly in the winter. Eclipse sensors, which are solar-powered reported proportionally fewer days at fewer locations in February compared to August. Likewise, February TropOMI retrievals had significantly fewer valid grid cells (-60%) than the August retrieval, with no significant clustering in valid grid cells (Fig. B1(b)). This highlights the utility of using CTMs to fill gaps in observational networks. However, new geostationary satellites will mitigate some data coverage issues by enhancing the number of retrievals per day. As such, our hotspot identification research could be used as a foundation for future work preparing for new data sources and the role they could play in diversified intra-urban air quality characterization efforts. Second, the boundaries of the hotspots in each of the different affected areas are sensitive to the interpolation scheme. As shown in Appendix B and C, the interpolation of TropOMI and Eclipse to the 1.3 km grid do not integrate additional meteorological or land-use information that influences the spatial heterogeneity of NO_2. Future work could further enhance the spatial

representativeness of the observational data products by explicitly considering these characteristics (Jain et al., 2021; Yu & Liu, 2021). Given that the interpolation schemes affect the representation of the data on the grid, changing interpolation schemes would affect the resulting clustering output of the model. Finally, we examined data for just one city; however, our method would easily generalize to other cities as dense, urban-scale sensing networks continue to proliferate.

3.7 Conclusion

In this paper, we construct high-resolution surfaces estimating NO_2 pollution across a major U.S. city from three different data products. Although all three datasets exhibit positive correlations, associations are subject to noise. However, when we apply a hotspot detection algorithm to each of these three surfaces, we identify a region where all three data products show significantly elevated NO_2 concentrations in both summer and winter months, suggesting with high confidence the presence of a large contiguous area with elevated NO_2 pollution. We estimate that this hotspot affects as many as 501,000 people, who are exposed to NO_2 concentrations that are 16 to 32% higher than the city-wide average. Moreover, this Consensus hotspot is evident regardless of wind direction suggesting a need to interrogate contributions from local sources.

We also identify two regions where hotspots are detected in either the model-derived or the remote sensing and ground sensor-derived surfaces. While disagreement across data products limits our confidence in the use of these regions for targeted interventions, further interrogation of product differences suggest clear strategies to improve our confidence in each dataset. Future work could adapt our approach to detect areas of concern systematically and automatically, either with high confidence — indicating the need for targeted intervention — or with lower confidence indicating priority areas for expanded monitoring and evaluation. Through this work, we show how multiple novel high-resolution data products can act as complementary components of a diversified urban monitoring and modeling framework.

Acknowledgements

All authors acknowledge funding and support from Microsoft Research. D.E.H. acknowledges support from US National Science Foundation grant CBET-1848683. The authors thank A. Anyachebelu and J. Liu for their feedback on early versions of this work. We are also grateful to C. Needham Jr, L. Story, D. Gehring, G. Jancke, T. Werner, R. Mansour, and S. Mudd for Eclipse sensor network design and support. In addition, we are grateful to Chicago's community-based organizations for their insights and expertise, as well as P. Banerjee, M. Grazioli, G. Brussel, N. Clochard-Bossuet, P. Rehus, and the JCDecaux leadership and maintenance team for their support.

Chapter 4

Exposure and health disparities in Chicago Air Pollution

Abstract

Exposure to ambient air pollutants contributes to hundreds of thousands of deaths over the US, with harms concentrated in urban areas. Here, we use neighborhood-scale (~1 km) CTM-based estimates of air quality to quantify air pollution exposure and health outcome inequalities over Chicago, IL using three pollutants (NO_2, $PM_{2.5}$, and $MDA8O_3$) and three health outcomes (premature morality, asthma, and pediatric asthma hospitalizations). We perform this analysis with an environmental justice framework, so our analysis is performed focusing on the 4 major racial and ethnic groups in Chicago: Non-Hispanic White, Black, Hispanic and Latino, and Asian populations. We look to 1) quantify exposure inequalities, 2) quantify the inequalities in pollution-attributable health outcomes, and 3) develop a framework to address exposure and health inequalities. We find that exposure inequalities differ depending on the pollutant, with no population subgroup consistently unequally impacted. The Hispanic/Latino population have the highest population-weighted average exposure to NO_2 and $PM_{2.5}$ (+4% compared to Chicago average), while Black and White populations have the highest population-weighted average exposure to $MDA8O_3$ (+1-2% compared to Chicago average). In analyzing population-weighted exposure across household incomes, we find that increasing incomes do not change the average population-weighted concentration of pollutants across racial and ethnic groups. In contrast, we find that health outcomes from pollution exposure consistently and negatively impact Chicago's Black and low-income populations, mainly driven by baseline health incidence rates. Increasing household incomes decrease the pollution-attributable mortality across all racial and ethnic groups, even though exposure is similar across incomes, further highlighting the importance of baseline health incidence rates. As exposure and health impacts are not always collocated, we pose integrating social, health, and exposure information to best target pollution in Chicago to promote more widespread and equitable outcomes.

4.1 Introduction

Air pollution is a major environmental and public health concern worldwide, particularly in urban areas where high levels of pollutants such as nitrogen dioxide (NO_2), fine particulate matter ($PM_{2.5}$), and ozone (O_3) interact with high density populations. Exposure to these pollutants can lead to a range of adverse health effects, including respiratory (Hoek et al., 2013) and cardiovascular diseases (B. A. Franklin et al., 2015), as well as premature mortality (M. Franklin et al., 2007). Due to the known health impacts of pollution, regulations have been set in place to limit pollution emissions, with an estimated 230,000 lives saved from the Clean Air Act (U.S. EPA, 2011).

However, air pollution-attributable health impacts are not equally distributed across the US population. Children and elders are particularly susceptible to the negative health outcomes of air pollution (Zheng et al., 2015). Additional external social factors exacerbate health outcomes, which result in racial and ethnic minorities experiencing poor health outcomes relative to the White population (Southerland et al., 2021). These health outcome inequalities are partially explained by age, lifestyle, and comorbidities (Sørensen et al., 2023), though there is increasing evidence of environmental contributors. Previous research shows that in the US, racial minorities and low-income people are overexposed to pollution relative to the White population (Clark et al., 2014; Morello-Frosch & Jesdale, 2006; Paolella et al., 2018; Tessum et al., 2021). Systemic factors such as historical patterns of segregation, discriminatory housing policies, and the siting of industrial facilities and highways have contributed to the unequal distribution of pollution and its associated health risks (Tessum et al., 2021). Environmental justice (EJ) is a critical framework that can be used to address the disproportionate burden of air pollution on marginalized and vulnerable communities.

The majority of federal monitoring is set up to analyze pollution on a regional scale, but there is a growing body of evidence highlighting the systemic inequalities in long-term air pollution exposure at local scales (Southerland et al., 2022). High resolution air pollution data is necessary to capture the true

breadth of exposure and health inequalities, as otherwise, health outcomes are underestimated (Mohegh et al., 2021; Gan et al., 2016; Paolella et al., 2018). However, urban exposure and pollution studies are limited by the spatial scale of air pollution data (Gardner-Frolick et al., 2022), as well as baseline disease incidence rates (Southerland et al., 2021).

To move EJ forward on a local scale, high-resolution air pollution data and high-resolution public health data should be combined. By isolating city boundaries, this analysis is more representative of how jurisdictions operate and can focus on city-wide practices that enhance unequal pollution outcomes (Lane et al., 2022). Further, this removes the suburban and urban demographic bias, which influences the exposure inequality calculation (Clark).

In this study, we focus on Chicago. Chicago has a long history of segregation (Sandoval, 2011) and environmental hazards which disproportionately affect marginalized communities due to the resident's proximity to industrial facilities, transportation infrastructure, and waste disposal sites (EPA, 2022; Baden & Coursey, 2002). Further, there are inequitable health outcomes related to baseline health incidence over Chicago, particularly for the Black population (R. S. Gupta et al., 2008). Previous national-scale studies have shown that in Chicago, there is outsized air pollution exposure for people of color regardless of race/ethnicity, independently of income (Goodkind et al., 2019; Kerr et al., 2021; Tessum et al., 2021), though these studies use a broader definition of Chicago (the metropolitan statistical area). Local, shorter-term studies of Chicago pollution have shown inequalities for $PM_{2.5}$ are greatest for the Hispanic/Latinx community (Esie et al., 2022). Furthermore, no studies have computed the Chicago marginalized community's health burden due to NO_2, $PM_{2.5}$, or O_3 using high-resolution air quality data.

Our study aims to quantify the inequalities in air pollution exposure and air pollution-attributable health impacts, as the inequalities relate to race, ethnicity, and incomes. By computing inequalities in exposure and health impacts across these demographics, we can analyze the EJ implications of pollution and pollutant-outcome patterns in Chicago. We will further advance the research into health inequalities by focusing on multiple pollutants and multiple health outcomes, including mortality, adult asthma

incidence rates, and pediatric asthma hospitalizations. By analyzing outcomes for adults and children, we can further highlight outcomes across multiple vulnerable demographic populations.

To perform this analysis, we use a high-resolution WRF-CMAQ simulation and census-level baseline health incidence rates. We calculate the attributable mortality, asthma, and pediatric asthma hospitalizations from NO_2, $PM_{2.5}$, and maximum-daily 8-hour O_3 ($MDA8O_3$) over Chicago. We examine the exposure inequalities through a socioeconomic and racial and ethnic lens. We further use the inequality analysis to develop a hypothetical framework to target pollutants in an equitable fashion to maximize social and health impacts. By identifying and analyzing these disparate outcomes, we seek to shed light on the complex interplay between air pollution exposure, demographic factors, and health, ultimately highlighting the need for targeted interventions and policy measures to address environmental injustices and promote equitable health outcomes for all population subgroups.

4.2. Data and Methods

4.2.1 Air Quality Data

We use simulated annualized air pollution data from two-way coupled WRF-CMAQ 1.3 km simulations whose set up, performance, and validation are described in full in Montgomery et al (2023). We use 1.3 km output from a coupled, two-way WRF-CMAQ computational fluid dynamics model simulation (CMAQ, v5.2, Byun & Schere, 2006; WRFv3.8, Skamarock et al., 2008; WRF-CMAQ, Wong et al., 2012). We create high-resolution input emissions by using the Sparse Matrix Operating Kernel of Emissions (SMOKE) Modeling Platform (2016beta version, Eyth et al., 2019), using 1.3 km spatial surrogates from LADCO (LADCO, 2022). The spatial surrogates enable the downscaling of coarser land-use and activity information, such as vehicle activity. To briefly describe the model simulation, we performed the simulations based off the Wong et al. (2012) methodology, which has three steps. First, we run a stand-alone WRF simulation to generate meteorology for emissions processing and boundary conditions at 12 km, 4 km, and 1. 3 km, using boundary conditions from the North American Regional Reanalysis (NARR) dataset for the 12 km domain. Second, we use the 4 km and 1.3 km WRF output to

process emissions in SMOKE, creating 4 and 1.3 km emissions. Finally, we run the coupled WRF-CMAQ model at 4 km to create meteorological and chemical boundaries for the 1.3 km simulation, using WACCM for chemical boundary conditions and the 12 km WRF run for meteorological boundaries. Finally, to create the 1.3 km two-way WRF-CMAQ output, we use the 1.3 km emissions from SMOKE and the 4-km WRF-CMAQ output as chemical and meteorological conditions. We isolate the data from the 1.3 km domain to Chicago to perform this analysis.

To create an annualized air quality data set, we run the model for August 2018, October 2018, January 2019, and April 2019. We average the 4-months of data to create a representative 'annualized' average concentration of pollutants. We isolate NO_2, $PM_{2.5}$, and maximum 8-hour average daily ozone ($MDAO_3$) concentrations from the CTM output. We then regrid these pollutants from the 1.3 km native grid to the census-tract by using area-weighting,

$$Pollutant_{ct} = \frac{\sum_1^n (x_n * area_n)}{\sum_1^n area_n}$$

where the pollution value in a census tract ($Pollutant_{ct}$) depends on the pollution concentration (x_n), the area of intersection between the census tract and the 1.3 km pixel ($area_n$), across the number of grid-cells of intersection (n).

4. 2.2 Population and income data

We use demographic information from the American Family Survey, using the 2015-2019 5-year average, downloaded from IPUMS NHGIS (Manson, Steven et al., 2022). We separate the racial and ethnic groups over Chicago for this analysis as follows: non-Hispanic White (White), non-Hispanic Black (Black), and non-Hispanic Asian (Asian) populations. We combine all races of Hispanic and Latino populations to identify the Hispanic/Latino population. All other racial and ethnic groups are combined under Other. To analyze income, we use the median per capita income and household income stratified by race from the 2015-2019 American Family Survey (Manson, Steven et al., 2022). We classify census tracts with the lowest quintile (<20%ile) of median income in Chicago as low-income, while the census tracts with the highest quintile (>80%ile) median income are categorized as high-income.

4.2.3 Health Metrics and Calculations

Table 4.1. Baseline health incidence rates used in this study.

Health Outcome	Pollutant	$\boldsymbol{\beta}$ (CI)	Age Group	Study
All-Cause Mortality	NO_2	1.040 (1.011 – 1.069)	30 – 99	Atkinson and Butland, 2018
	$PM_{2.5}$	1.060 (1.040 – 1.080)	30 – 99	Krewski et al., 2009
	$MDAO_3$	1.020 (1.010 – 1.040)	30 – 99	Turner et al., 2019
Asthma	NO_2	1.110 (1.060 – 1.260)	18 – 99	Anderson et al., 2013
	$PM_{2.5}$	1.160 (0.980 – 1.370)	18 – 99	Anderson et al., 2013
	$MDAO_3$	1.009 (1.006 – 1.011)	30 – 99	Li et al., 2019
Pediatric Asthma Hospitalizations	NO_2	1.018 (1.014 - 1.022)	1 – 18	Orellano et al., 2017
	$PM_{2.5}$	1.048 (1.028 - 1.067)	1 – 18	Lim et al., 2016
	$MDAO_3$	1.011 (1.007 – 1.014)	1 – 18	Zheng et al., 2015

We derive the attributable NO_2, $PM_{2.5}$, and O_3 mortality and morbidity outcomes using epidemiologically derived health response functions (Table 4.1). We use all-cause baseline mortality rates from Industrial Economic, Incorporated with rates for each 5-year age group (IEc 2010-2015) (USALEEP, 2022) derived from the USALEEP abridged life tables for use in national health analysis (Arias et al., 2018), asthma-incidence rates from the Illinois Department of Public Health (IDPH, 2017), and pediatric asthma emergency room (ER) visits from the Illinois Department of Public Health (IDPH, 2017). We calculate health outcomes for each census tract (CT) by calculating the attributable fraction (AF, 1) and resulting health outcomes (HI, 2):

$$AF = 1 - \exp(-\beta x_{ct})$$

$$HI_{ct} = BIR_{ct} * POP_{ct} * AF_{ct}$$

The attributable fraction calculates the relationship between the pollutant (x_{ct}) in a census tract and the health outcome, as related by the β coefficient. β values are determined by epidemiological studies which quantify the relationship between air pollution concentrations and specified health outcome, with the values outlined in Table 1. Values reported in 10 μg/m^3 are converted to ppb equivalents at standard temperature and pressure. The health impact is then computed by taking the baseline incidence rate (BIR)

of the health outcome, the population of the census tract which is affected by β, and the AF. To create the health impact rates, we divide the health impact in a census tract (HI_{ct}) by the number of people in the census tract (Pop_{ct}), such that:

$$HIR_{ct} = \frac{HI_{ct}}{Pop_{ct}}$$

4.2.4 Spatial Variability and Attribute Correlations

As both demographics characteristics and pollutants are highly spatially variable, we quantify the quartile coefficient of variation (QCV) of the attributes.

$$QCV = \frac{x_{75\%ile} - x_{25\%ile}}{\bar{x}}$$

Where $x_{75\%ile}$ and $x_{25\%ile}$ is the 75%ile and 25%Ile value of the attribute x, with $\bar{x}$ representing the average values of the attribute x over Chicago. The QVC allows for us to compare variation across attributes with different units, so larger QVCs indicate more variation of that attribute across the city, which we interpret as a greater degree of segregation of values.

We further analyze the correlation (Pearson's r) between pollutant concentrations and social characteristics to quantify the strength of association across variables. Highly correlated variables indicate a strong linear relationship between demographics and outcomes, we would consider this relationship as a basis for an EJ intervention.

4.2.5 Exposure Inequality

We use two metrics to quantify the exposure and health disparities across Chicago. First, we compute the population-weighted average exposure to create the average disparity across Chicago, such that:

$$Population - Weighted\ Average = \frac{\sum_1^n (x_n * population_n)}{\sum_1^n populaction_n}$$

where the pollutant (x_n) in a census tract is multiplied against the total population or population subgroup in a census tract ($population_n$) over Chicago (n = 756 census tracts). This provides one value to identify the average exposure of a demographic group over all of Chicago, but cannot identify spatial patterns.

To calculate the inequalities of attributes (e.g., pollution, BMR) of any specific demographic group, we calculate the mean difference of the group from the Chicago average, such that:

$$Inequality = \frac{x_d - \bar{x}}{\bar{x}}$$

Where x_d is the population-weighted average value of a demographic group, and $\bar{x}$ is the Chicago average of that attribute. We use this inequality value to make assessments about EJ, such that the largest inequalities across demographics in exposure or pollution-related health impacts are identified as an EJ issue.

Second, we bin census tracts into quintiles using the concentration of each pollutant so that we can analyze the average demographic makeup for areas with increasing pollutant concentrations. This allows for us to look at nonlinear outcomes of inequalities and highlight how disparities change across distributions. For each quintile, we perform descriptive statistics (e.g.: calculating the mean and interquartile range). We then identify how social and demographic characteristics vary across pollutant groupings. We also test the inverse by grouping census tracts into quintiles based off the demographic makeup to test how pollution concentrations change with increasing racial and ethnic proportions.

4.3. Results

4.3.1 Pollution and sociodemographic spatial patterns over Chicago

In this section, we will discuss the specific spatial patterns of pollutants and demographics over Chicago. The patterns of pollution (Figure 4.1 a-c) and patterns of demographics (Figure 4.1 d-h) have high numerical variability (Figure 4.1i). In general, the pollution patterns are not strongly linearly associated with demographic patterns, as shown by the weak correlation across pollutants and demographics ($r < 0.3$, Supplementary Figure 3.1). The primary focus of this analysis is to quantify the segregated spatial patterns and analyze the concurrence of demographic and pollution patterns.

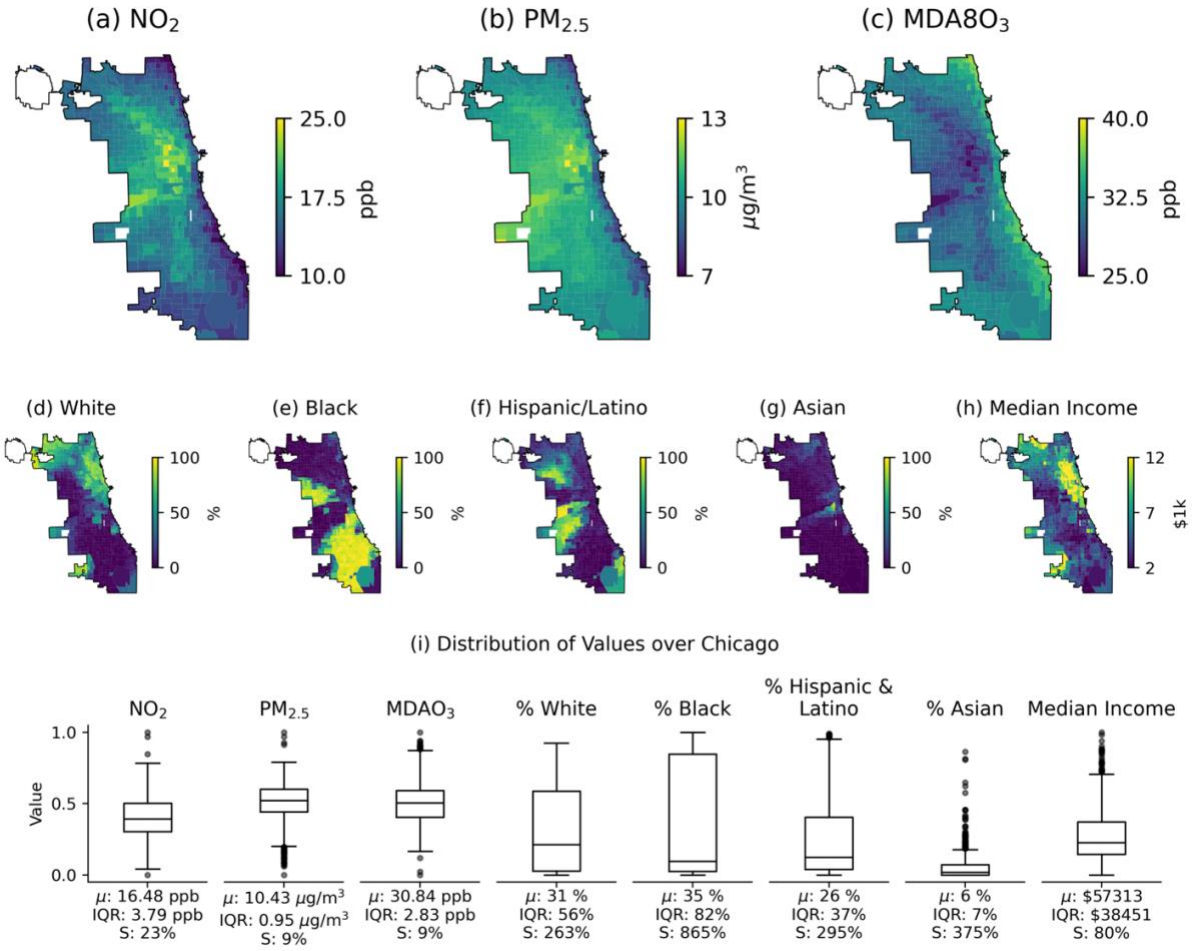


Figure 4.1 Annualized air pollutant concentrations (a-c), proportion of racial groups (d-g), median household income (h), and distribution of values (i) over Chicago. Note, in (i), NO_2, $PM_{2.5}$, $MDA8O_3$, and median income are normalized to for comparison across all values.

The spatial patterns of pollutants mirror locations of major emissions sources in Chicago, with NO_2 and $PM_{2.5}$ high over highways and industrial areas while $MDAO_3$ is depleted. The average NO_2 concentration is simulated to be 16.4 ppb, $PM_{2.5}$ is simulated to be 10.4 ppb, and $MDA8O_3$ is simulated to be 30.8 ppb. For NO_2, concentrations are highest over highways and in the central and western parts of the city (Figure 4.1a). The spatial gradient for NO_2 is steep, as the interquartile range (IQR) for NO_2 is 3.7 ppb with a QVC of 23%. $PM_{2.5}$ concentrations are simulated to be more spatially diffuse, though concentrated in the center of the city and West side (Figure 4.1b). $PM_{2.5}$ has an IQR of 1.0 μg/m^3, with a spread of 9%. $MDA8O_3$ concentrations are highest near the lakeshore and lowest along the highways (Figure 4.1c). The IQR of $MDA8_{O3}$ is 2.8 ppb, with a QCV of 9%. Compared to NO_2, $PM_{2.5}$ and

$MDA8O_3$ is more spatially diffuse. The highest concentrations of $PM_{2.5}$ coincide with the highest NO_2 concentrations, while $MDAO_3$ is highest outside of those regions.

The spatial patterns of racial and ethnic groups over Chicago are distinct for each racial and ethnic groups, with few areas of overlap (Figure 4.1d-g), highlighting the high degree of segregation in the city. For the major racial and ethnic groups of Chicago, the largest demographic population are non-Hispanic White (33%), Black (29%), Hispanic and Latino (28%), and Asian (7%). The most spatially-varying racial and ethnic group is the Black population (865%), which indicates the strength of separation of the Black population within Black census tracts. In contrast, the White population (with the lowest QVC = 263%) is more spread across the city. Spatially, the highest proportion of White residents are located along the Northern lakefront and near O'Hare airport (Figure 4.1d). The highest proportion of Black residents occur on the South and Northwest side of the city (Figure 4.1e). The highest proportion of the Hispanic and Latino population is located on the West side of Chicago, with pockets in the North and South of the city as well (Figure 4.1f). The highest proportion of the Asian population in a small area in the center of the Chicago (Figure 4.1g). Other races and ethnicities account for less than <2% of Chicago population, so our analysis focuses on the White, Black, Hispanic and Latino, and Asian populations.

In analyzing the income distribution over Chicago, we find that income, like the racial and ethnic clustering, is similarly spatially variable. The highest incomes are clustered in the central and Northeast regions of the city, with lower incomes on the West and South sides of the city.

In comparing the socioeconomic characteristics to the pollution patterns, we show that there is higher spatial variability in social characteristics than in pollutants. Comparing the highest QVC across pollutants, racial and ethnic groups, and median incomes, we see that NO_2 is far more spatially diffuse (23%) than income (93%) and any racial or ethnic group (263% - 865%). Most obviously, the spatial patterns of pollutants show where the highest pollution emitting sources are, but no single demographic group perfectly mirrors that pattern. Through this analysis, we highlight high concentrations of NO_2 and $PM_{2.5}$ on the West side of Chicago and in the central city, which have high Hispanic and Latino

populations. However, given the weak correlations between census-level pollution and demographic, additional analysis must be done to identify inequalities for EJ, which is done in the following sections.

4.3.2 Average exposure and health rates over Chicago

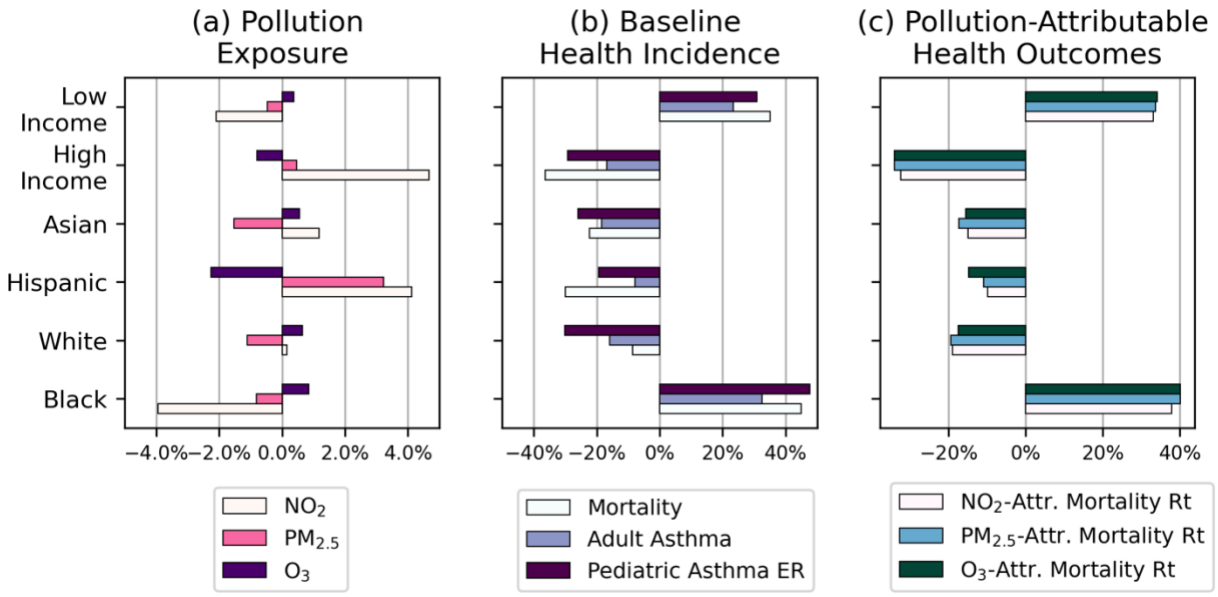


Figure 4.2 Differences (%) compared to Chicago average of population-weighted average exposure (a: NO_2, $PM_{2.5}$, O_3) and the baseline health incidence rates (b: baseline mortality rates (BMR), asthma incidence, pediatric asthma ER visits) and (c) pollutant attributable health outcomes.

In this section, we will examine the inequalities across population-weighted average concentrations, baseline health incidence rates, and pollution-attributable health outcomes across race, ethnicity, and income. With high inequalities, we will interpret this group and health or exposure outcome as a priority for EJ intervention. To highlight inequalities, we will calculate the population-weighted average values for each attribute and compare it to the Chicago average across subgroups. The primary focus of our analysis is to identify which population group experiences the greatest and smallest disparities.

4.3.3 Pollution Exposure Disparities

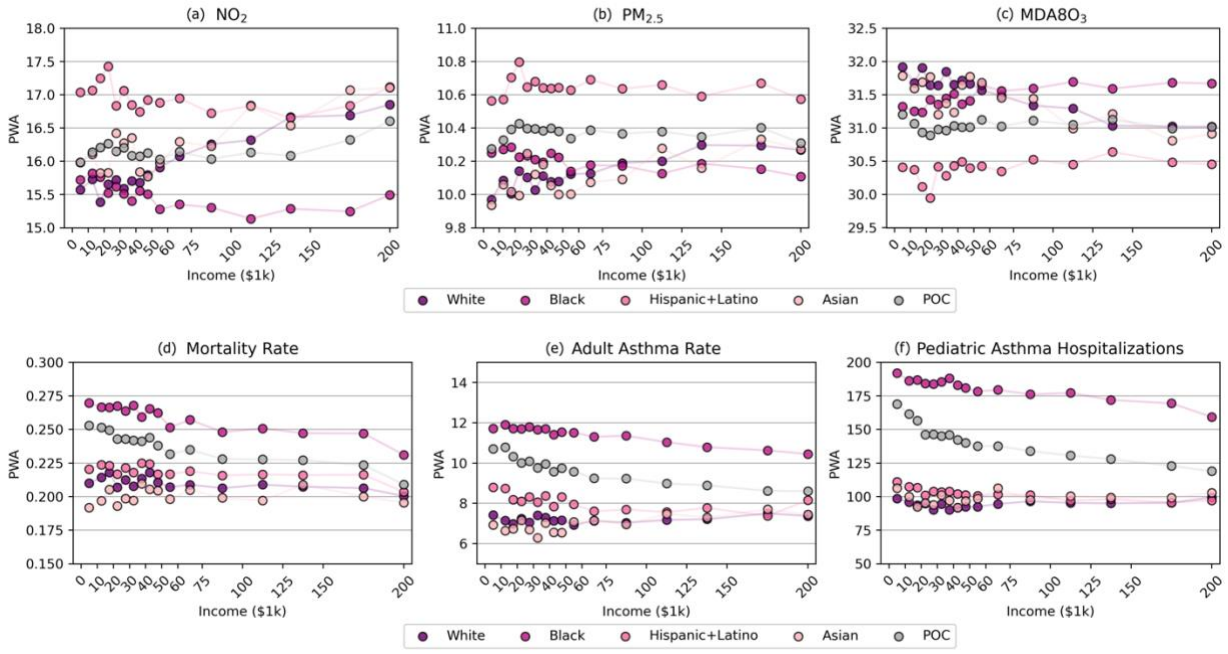


Figure 4.3 Population-weighted average exposure to pollutants (a-c) and health incidence (d-f) across incomes stratified by racial and ethnic groups.

To analyze the inequalities in exposure across racial and ethnic groups, we compute the population weighted average exposure to NO_2, $PM_{2.5}$, and $MDA8O_3$ for each major racial and ethnic groups. We find that the average population-weighted exposure across racial and ethnic groups vary by less than 5% for NO_2, $PM_{2.5}$, and $MDA8O_3$. For NO_2, the highest population-weighted average exposure over racial and ethnic groups occurs with the Hispanic & Latino population (4% greater than Chicago population weighted average). The lowest NO_2 exposure occurs with the Black population (-4%). For $PM_{2.5}$, the highest population-weighted average exposure occurs with the Hispanic/Latino population (3%), while the lowest exposure is with the Asian population (-2%). For $MDA8O_3$, the highest population-weighted average exposure occurs with the Black population (1%), while the lowest exposure occurs in the Hispanic & Latino population (-2%). As such, we find that there are small (±5%) exposure inequalities across racial and ethnic groups, with no single racial or ethnic group experiencing outsized exposure disparity for NO_2, $PM_{2.5}$, and $MDA8O_3$.

To analyze the inequalities in exposure across high-income and low-income neighborhoods, we compute the population weighted average exposure to NO_2, $PM_{2.5}$, and $MDA8O_3$ for the high- and low-

income areas in Chicago. The high-income area has the highest NO_2 exposure and slightly above-average $PM_{2.5}$ exposure (4.5%, 0.5% higher than Chicago average, Figure 4.2a). This area experiences low NO_2 and $PM_{2.5}$ exposure (-2%, -0.5% from Chicago average), and slightly elevated $MDA8O_3$ exposure (0.5% from Chicago average, Figure 4.2a). To summarize, we find that the highest income areas are exposed to higher NO_2 and $PM_{2.5}$ while the lower income areas are exposed to higher $MDA8O_3$.

To analyze the effect of income on exposure across racial and ethnic groups, we further compute the population weighted average exposure across household incomes stratified by race and ethnicity (Figure 4.3a-c). We find that the Hispanic and Latino population experience the highest exposure to $PM_{2.5}$ and lowest $MDA8O_3$ at every income level (Figure 4.3a,c). Further, the Hispanic and Latino population experience higher NO_2 until the income of $150,000 (Figure 4.3b). No other demographics experience consistently elevated or depressed exposure across incomes, the exposure disparity depends on the pollutant and the median household income. For example, in identifying the population which experiences the highest $MDA8O_3$ across incomes, we show that the Black population is simulated to be exposed to the highest relative $MDA8O_3$ concentration at incomes higher than $60k, though below that income threshold, the White and Asian population experiences the highest $MDA8O_3$ (Figure 4.3c). Further, identifying which population subgroup experiences the lowest NO_2, we find that at low incomes (<$20k), the White population is estimated to experience the lowest NO_2, though above $20k, the Black population is estimated to experience the lowest NO_2 (Figure 4.3a). In general, there are no robust trends in increasing incomes and population weighted average exposure across subgroups, meaning that there are no robust differences in exposure at low and high incomes across population subgroups.

To further test the effect of income on exposure across racial an ethnic groups, we separate the White population and compare to a combination of Black, Hispanic and Latino, and Asian population together (i.e., people of color of POC, as done in Tessum et al., 2021). Through this recombination, we find that POCs consistently experience higher NO_2 and $PM_{2.5}$ exposure than the White population, regardless of income (Figure 4.2a,b, grey). We also find that the white population is simulated to be

exposed to the highest levels of $MDA8O_3$ across all household incomes (Figure 4.2c). As shown in the racial and ethnic breakdown across exposures, the higher exposure to NO_2 and $PM_{2.5}$ is driven by the outsized pollutant exposure by the Hispanic and Latino population.

To summarize the results of the exposure analysis, we find that over Chicago, there are small (+-5%) exposure disparities across racial and ethnic groups, though no single racial or ethnic group shows a consistent exposure inequality for NO_2, $PM_{2.5}$, and $MDA8O_3$. The Hispanic and Latino population is exposed to the highest NO_2 (+4%) and $PM_{2.5}$ (+3%), while the Black population is exposed to the highest $MDA8O_3$ (1%). In high-income areas, these areas are simulated to have higher NO_2 and $PM_{2.5}$, while low-income areas are simulated to have higher $MDA8O_3$. However, household income does not vastly change exposure inequalities for NO_2 and $PM_{2.5}$ across race, so as shown in Chicago average concentrations, the Hispanic and Latino population is exposed to the highest $PM_{2.5}$, highest NO_2 (up to $150k), and lowest $MDA8O_3$. As such, this analysis highlights the outsized exposure inequality to NO_2 and $PM_{2.5}$ for the Hispanic and Latino population in Chicago, indicating that these pollutants should be prioritized for intervention to lower exposures in areas of high Hispanic and Latino population.

4.3.4 Baseline Health Incidence Inequalities

To analyze the inequalities in baseline health incidence rates across racial and ethnic groups, we compute the population weighted average incidence rates of mortality, adult asthma, and pediatric asthma hospitalizations (Figure 4.2b). The highest baseline health incidence rates are less heterogeneous across population subgroups, as Chicago's Black population experiences the highest population-weighted baseline mortality rates (+45%), asthma incidence rates (+32%), and pediatric asthma hospitalizations rates (+48%, Figure 4.2b). As the low end, the disparities vary across subgroups: The Hispanic and Latino population have the lowest average mortality rate, the Asian population has the lowest adult asthma incidence rate (-18%), and the White population has the lowest pediatric asthma ER hospitalizations (-30%, Figure 4.2b). In contrast to exposure, baseline health rates show greater disparities across racial and

ethnic groups than the average pollution exposure (±5% vs. ±50%), though the Black population consistently experiences the highest mortality, asthma, and pediatric asthma ER incidence rates.

To analyze the inequalities in exposure across high-income and low-income neighborhoods, we compute the population weighted average incidence rates of mortality rates, adult asthma rates, and pediatric asthma hospitalization rates over high-income and low-income areas. The low-income area experiences higher-than-average baseline health incidence rates of mortality, adult asthma, and pediatric asthma hospitalizations (22% – 37%). The high-income area experiences lower-than-average baseline health incidence rates, with the high-income area having lower mortality rates than any racial or ethnic group (-18% – -38%).

To analyze the effect of income on baseline health incidence rates across racial and ethnic groups, we compute the population weighted average exposure across household incomes stratified by race and ethnicity (Figure 4.3d-c). At higher incomes, all racial and ethnic groups experience lower mortality rates, asthma incidence rates, and pediatric asthma hospitalizations than the low-income counterparts. At all incomes, the Black population experiences higher mortality rates, adult asthma rates, and pediatric asthma hospitalization rates.

To summarize, the baseline health incidence rates have high inequalities that mainly are driven by disparate health incidence of the Black population. The Black population has the highest rates of mortality, asthma, and pediatric asthma hospitalizations (32 – 45%). The lowest baseline health incidence rates depend on the racial or ethnic group. With increasing income, all racial and ethnic groups show lower rates of mortality, asthma, and pediatric asthma hospitalizations, though the Black population has the highest health incidence rates across all incomes.

4.3.5 Pollution-Attributable Health Outcomes and Inequalities

To quantify the total mortality burden from Chicago pollution exposure, we compute the attributable all-cause mortality for NO_2, $PM_{2.5}$, and $MDA8O_3$. We find that the highest health-attributable mortality comes from $MDAO_3$ concentrations. We estimate that the attributable mortality in Chicago is

17 (8 – 33, 5%ile to 95%ile range) for NO_2, 12 (4 - 16) for $PM_{2.5}$, and 24 (6 - 33) for $MDA8O_3$ (Table 3.2). Spatially, the highest pollution-attributable mortality for all pollutants is located on the South side of Chicago (Supplementary Figure 3.2). Further, the Black population contributes to highest proportion of the total NO_2- $PM_{2.5}$, and MDA O_3-attributable mortality (41%, Supplementary Table 3.1, Figure 4.2c). The lowest average attributable mortality rates for all pollutants occurs in the White population (Figure 4.2c, Supplementary Table 3.1).

To quantify the total adult asthma burden from Chicago pollution exposure, we compute the attributable adult asthma for NO_2, $PM_{2.5}$, and $MDA8O_3$. The highest adult asthma is attributed to $MDA8O_3$ exposure. We compute 7,500 (4100 – 16,500) for NO_2, 12,500 (0 – 26,100) for $PM_{2.5}$, and 31,800 (0 – 55,500) for $MDA8O_3$ (Table 4.2). Spatially, the highest pollution-attributable asthma for all pollutants is located on the South and West side of Chicago (Supplementary Figure 3.3). The Black population has the highest rates of attributable asthma for all pollutants (6.3 – 18.2 per 1000 adults), accounting for 43% of total asthma cases (Supplementary Table 3.2). The lowest attributable asthma rates occurs in the Asian population (Supplementary Table 3.2).

To quantify the total pediatric asthma hospitalization burden from Chicago pollution exposure, we compute the attributable pediatric asthma hospitalization for NO_2, $PM_{2.5}$, and $MDA8O_3$. The highest pediatric asthma hospitalization burden is from $MDA8O_3$. We compute 13 (13 - 13) for NO_2, 16 (16 – 16) for $PM_{2.5}$, and 45 (44 – 45) for $MDAO_3$ pediatric asthma hospitalizations(Table 4.2). Spatially, the highest pollution-attributable asthma for all pollutants is located on the South and West side of Chicago (Supplementary Figure 3.4). The Black population has the highest rates of attributable pediatric asthma hospitalizations (1.2 – 4.2 per 10,000 children), accounting for 46% of total pediatric asthma hospitalizations (Supplementary Table 3.3). The lowest attributable pediatric asthma hospitalization rate occurs in the Asian population (0.1 – 0.3 per 10,000 children, Supplementary Table 3.2).

To summarize, the pollution-attributable health impacts have high inequalities that mainly are driven by disparate health outcomes in the Black population. The Black population has the highest

attributable of mortality, asthma, and pediatric asthma hospitalization for all three pollutants. The high-income areas have the lowest pollution-attributable health outcome, and the low-income areas have the highest pollution-attributable health outcomes. Even with increasing incomes, the Black population consistently experiences the highest pollution-attributable health outcomes. From this health analysis, we identify the Black population as an EJ priority to reduce pollution-attributable health outcomes.

Table 4.2 Total attributable mortality, asthma, and pediatric asthma hospitalizations over Chicago. Rates for mortality and hospitalizations are per 100k, rates for asthma are per 10k. The 5%-95%ile confidence intervals are included in parenthesis.

	Total Cases per Year over Chicago			Average Rate over Chicago		
Pollutant	NO_2	$PM_{2.5}$	$MDAO_3$	NO_2	$PM_{2.5}$	$MDAO_3$
Mortality	17 (8, 33)	12 (4, 16)	24 (6, 33)	1 (0.6, 2.1)	0.8 (0.3, 1)	1.5 (0.4, 2.2)
Asthma	7466 (4163, 16533)	12541 (0, 26110)	31780 (0, 55492)	18 (10, 40)	30 (0, 63)	78 (0, 135)
Pediatric Asthma ER	13 (12.5, 12.6)	16 (15.4, 16)	45 (44.3, 44.6)	2.3 (2.3, 2.4)	2.9 (2.8, 2.9)	8.2 (8.1, 8.2)

4.3.6 Proportional demographic exposure and health outcomes

To understand the variations in pollutant exposure among different population subgroups in Chicago, we conducted a detailed analysis of the proportional demographic exposure to increasing levels of pollutants. Broadly, the exposure (Figure 4.4a-c) and pollution-attributable health outcomes (Figure 4.4d-f) do not result in similar proportional demographic outcomes. We focus on mortality in the main text (Figure 4.4d-f), but similar health outcomes are observed in asthma and pediatric asthma ER hospitalizations (Supplementary Figure 3.2). By using the proportional exposure, we can analyze broad trends of exposure or health impacts given the demographics, so an increasing demographic group with increasing pollutant would indicate that this group is an EJ priority.

Figure 4.4 Exposure across demographic groups, grouped by levels of (a) NO_2, (b) $PM_{2.5}$, and (c) $MDAO_3$. Mortality related to each pollutant, grouped by levels of mortality rates. The lowest exposure or mortality rates occur in Q1, while the highest occur in Q5.

The exposure and health effects of NO_2 and $PM_{2.5}$ affect racial and ethnic groups in similar proportions. As the average concentration of NO_2 or $PM_{2.5}$ increases in an area (Figure 4.4 a,b), there is a higher proportion of Hispanic and Latino population and a lower proportion of Black population. However, the Hispanic and Latino populations only make up the majority of population in areas with the third quintile (Q3, 10.4 – 10.7 ppb) of NO_2 concentration and the fourth quintile (Q4, 0.8 – 1.2 μg/m^3) of $PM_{2.5}$ concentration. In areas with the highest NO_2 levels (Q5, 18.9 – 27.9 ppb in Figure 4.4a), the largest demographic group by fraction (i.e., proportional demographic group) are White people (40%), while in areas with the lowest NO_2 concentrations (Q1, 8.8 – 14.2 ppb in Figure 4.3a), the largest demographic group is the Black population (80%). The mortality effects related to NO_2 and $PM_{2.5}$ do not mirror the exposure patterns: in areas with the highest NO_2- or $PM_{2.5}$-attributable mortality, the Black population is

overwhelmingly represented (80%), whereas in areas with low NO_2- or $PM_{2.5}$-attributable mortality, the White population is disproportionately represented (62%, Figure 4.4a,b and Figure 4.3d,e).

For $MDA8O_3$, both the highest and lowest $MDA8O_3$ areas have similar proportions of White populations (Figure 4.4c,f). However, increasing concentrations of $MDA8O_3$ have increasing proportion of Black population and decreasing proportion of Hispanic and Latino populations. The proportional demographic share of mortality follows a similar pattern to that observed for mortality related to NO_2 and $PM_{2.5}$: as $MDA8O_3$ mortality increases, there is a corresponding increase in the proportional representation of the Black population.

The overarching pattern we observe is that areas with high pollution do not necessarily result in similarly high health outcomes (Supplementary Figure 3.4-3.6, Supplementary Table 3.4). This is because the highest health impacts are calculated to occur in areas of poorest health, as opposed to areas of highest exposure. Areas with the highest NO_2 and $PM_{2.5}$ have, on average, 20 - 25% lower BMR, 4 – 8% lower asthma incidence rates, and 1 – 4% lower pediatric asthma hospitalizations (Supplementary Table 3.4). In contrast, areas with highest $MDA8O_3$ have high mortality (12% higher than Chicago average) and high pediatric asthma hospitalizations (4% higher than Chicago average), though lower adult asthma incidence (-8%). This phenomenon of opposing exposure and underlying health incidence rates result in the non-linear outcomes of exposure and impact. As such, the EJ priority for this analysis shows that targeting high-pollution areas for the purpose of health may not always result in the highest health outcomes.

4.3.7 Combining exposure and health incidence rates to target pollution

As shown in the exposure and health analysis, the exposure and health inequalities are not necessarily always collocated. In this section, we test a hypothetical policy outcome which attempts to lessen inequalities by halving pollution in areas which target the pollutants in areas of high pollution (NO_2, $PM_{2.5}$, or $MDAO_3$) and in areas of high underlying mortality rates. By halving the pollution in each target area, we show how different policies would contribute to lessening or increasing inequalities of pollution exposure and attributable health outcomes.

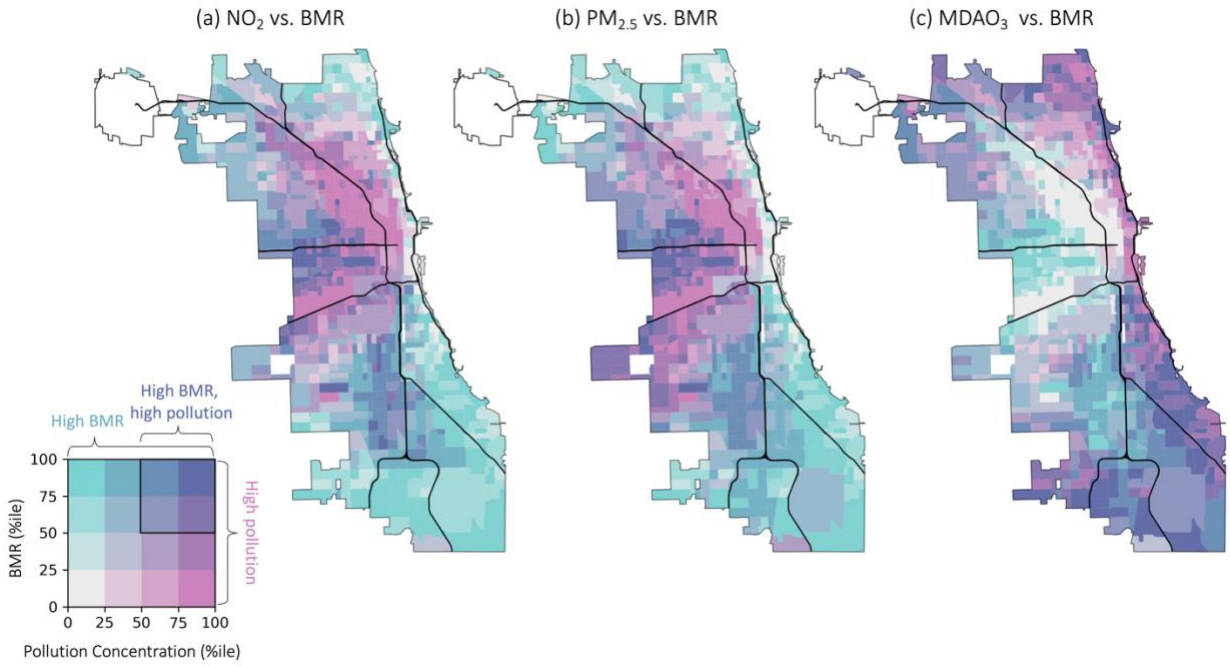


Figure 4.5 Baseline mortality compared to (a) NO_2, (b) $PM_{2.5}$, and (c) $MDAO_3$. Areas with high BMR and high pollution (>50%ile) are in purple, areas with high BMR are blue – purple (>75%ile), areas with high pollution are pink – purple (>75%ile).

To do this, we isolate areas with high pollution, high BMR, or high BMR and high pollution, and calculate the Chicago-wide implications of halving the concentration of pollutants in each area. We present the areas of inequality in Figure 4.5, which shows a bivariate plot with pollution concentrations and baseline mortality to highlight the underlying health data which informs the health outcomes. The areas of high baseline mortality and high pollutants are purple (with >50%ile BMR + >50%ile pollution), areas with high exposure are pink (>75%ile pollution), and high mortality are blue (>75%ile BMR). Each area of interest has roughly the same number of census tracts (~200) and cover a similar area (~200 km^2). By reducing pollution by half in the high BMR and high exposure areas (Purple, Figure 4.5), we reduce exposure by 4 – 7% and mortalities 6 – 12% over Chicago (Supplementary Table 3.5). By decreasing pollution by half in the high pollution and high BMR areas, we reduce annual mortality by 6 – 12% over Chicago (Supplementary Table 3.6). Further, average population exposure is decreased by 4 – 7% (Supplementary Table 3.6). Regardless of pollutant, reducing pollution in this area benefits the Black,

Hispanic and Latino, and Low-income populations the most. As such, we alleviate some exposure inequalities, as the Hispanic, Latino and Black populations reduce their average NO_2 and $PM_{2.5}$ exposure by 7 – 17% (Supplementary Table 3.5). Even though there are no areas of overlap between areas of high $MDAO_3$ and areas of high NO_2 or $PM_{2.5}$, the benefits of this policy outcome still overlap, as the highest BMR occur in areas with high proportions of Black residents, and this policy outcome targets populations with overburdened exposure, e.g., the Hispanic and Latino population (on average 50% and 30%, respectively; Table 3.6).

Reducing pollution in areas of high pollution (Pink, Figure 5) results in greater reductions for population-weighted average concentrations for pollutants (9 – 12%, Supplementary Table 3.7 – 3.8), though a smaller reduction in mortality outcomes (7 – 9%, Supplementary Table 3.7 – 3.8). By decreasing pollution by half in the high pollution areas, we reduce mortalities by 6 – 10% (Supplementary Table 3.8). Further, the population-weighted average exposure over Chicago results in an average reduction of 10 – 12% in concentration (Supplementary Table 3.8). By targeting only high pollution areas for $PM_{2.5}$ and NO_2, the Hispanic and Latino population decrease exposure most (17 – 19%, Supplementary Table 3.7), though the Black population reduces exposure the least (8 – 9%), resulting in relative exposure and mortality inequalities to increase for the Black population for NO_2 and $PM_{2.5}$. For $MDA8O_3$, reducing $MDA8O_3$ by half in the highest $MDA8O_3$ areas result in the Black and Asian populations to reduce their exposure the most (18 -- 19%, Supplementary Table 3.7), though Black mortality from $MDA8O_3$ is still high (17% above Chicago average). The modest change in mortality from this policy outcome is because the areas of high NO_2 and $PM_{2.5}$ areas have BMR that are 20 – 25% lower than the Chicago average, which results in the lower health benefits than what is shown in targeting the high BMR and high pollution areas (8 – 11% vs. 7 – 8%, Supplementary Table 3.6 vs. 3.8).

Reducing pollution in areas of high BMR (Blue, Figure 4.5) results in the greatest reductions of mortality (16 – 18%), and moderate changes in pollutants (8 – 10%, Supplementary Table 3.9 – 3.10). By decreasing pollution by half in the high pollution areas, we reduce mortalities by 16 – 18%

(Supplementary Table 3.10). Further, the population-weighted average exposure over Chicago results an average reduction of 8 – 10% in concentration (Supplementary Table 3.10). By targeting pollution in areas of high BMR, the the Black population experiences the greatest change in NO_2, $PM_{2.5}$, and $MDA8O_3$ (-25%) and mortality (-30%). This results in exposure inequalities increasing for all other subgroups, though only the mortality increases significantly for the Asian population (6% for all three pollutants, Supplementary Table 3.9). The high BMR area is simulated to have NO_2 and $PM_{2.5}$ that is 5% and 1% lower than the Chicago average, though $MDA8O_3$ is 1% higher. Because pollution is already lower than the Chicago average in this area, there may be further challenges in reducing pollution.

In summary, targeting high BMR and high pollution areas result in different outcomes with respect to health and exposure equity. By targeting areas with both high BMR and high pollution, we show that there are more equitable health outcomes across the most exposed and most health-impacted groups. When targeting only exposure or health, we reduce the average exposure or health more than the combined policy, but there are equity tradeoffs that deepen inequalities for the groups who are affected.

4.4 Discussion

We estimated the exposure and health inequalities due to concentrations of NO_2, O_3, and $PM_{2.5}$ over Chicago. By analyzing the underlying the spatial distribution of both health and exposure outcomes, we show that inequalities can be characterized and addressed through a variety of methods.

First, we show that air pollution exposure varies within Chicago across racial and income groups. No single group experiences outsized pollution exposure for all three pollutants analyzed, so inequalities in exposure depend on the pollutant (Figure 4.2). Further, exposure does not change significantly for any population subgroup as income increases (Figure 4.3). Given this, we see the highest exposure to NO_2 and $PM_{2.5}$ affect the Hispanic and Latino population and high-income population (Figure 4.2). Proportionally, areas of high NO_2 and high $PM_{2.5}$ are majority White and Hispanic and Latino, while areas of low NO_2 have high Black populations and areas of low $PM_{2.5}$ have high White populations (Figure 4.3). This is similar to other studies, which show high NO_2 affecting Black and Hispanic communities (Kerr et al.,

2021), however this contrasts to other studies (Jbaily et al., 2022) which show high $PM_{2.5}$ affects low-income communities more in Chicago. Some of these differences are due to different definitions of Chicago (city-limit vs. MSA) and exposure comparison (national vs. intraurban). Further, we find that areas of high (>80%ile) $MDAO_3$ are mostly White with significant Black populations (Figure 4.2), while areas of low (<20%ile) $MDAO_3$ are mostly White and Hispanic and Latino.

Second, while air pollution patterns are spatially distinct, the air pollution related health outcomes can be similar across the city, regardless of pollutant. This is due to the contribution of baseline health incidence rates, which vary to a larger degree than the pollution exposure (Figure 4.1). This is exemplified by NO_2 and O_3, which have opposite spatial patterns over Chicago (Figure 4.1a vs c), but the spatial patterns in mortality, adult asthma, and pediatric asthma hospitalizations are similar (Supplementary Figures 3.2 – 3.4). Similarly, regardless of pollutant, there is an over-representation of the Black population in experiencing the worst health outcomes (Figure 4.3), even though simulated pollution exposure is highest for Hispanic and Latino populations for NO_2 and $PM_{2.5}$ (Figure 4.2). We also note that when organizing across proportional racial groups, there is no strong relationship between increasing proportional demographics and increased pollution exposure (Supplementary Figure 3.9), e.g. areas with majority Black populations experience similar levels of pollutants as areas with majority White, Hispanic, or Asian populations (Supplementary Figure 3.9a vs. b).

As such, while disproportionate air pollution exposure exacerbates pollution inequalities in the city, baseline health rates show greater racial and ethnic disparities than the average pollution exposure (Figure 4.2). The baseline health disparities highlight the complex interplay of socioeconomic factors, systemic inequalities, and environmental conditions that contribute to health inequities. Factors such as access to healthcare, quality of housing, educational opportunities, and employment play significant roles in shaping baseline health rates (Feinglass et al., 2007). Therefore, addressing pollution disparities alone may not be sufficient in achieving health equity, and comprehensive strategies that address both the social

determinants of health and environmental justice are crucial for reducing disparities and promoting equitable health outcomes for all communities (Jbaily et al., 2022).

That exposure and the resulting health impacts are not necessarily collocated highlights the issue with addressing pollution vs. pollution related harms. Policies which target elevated pollution may consider remediating pollution over the highest pollution areas though this may not result in high health benefits or result in a more equitable distribution of pollution. Previous studies have shown that integrating social demographic information and exposure can alleviate exposure inequalities (Goodkind et al., 2019; Nguyen & Marshall, 2018b; Wang et al., 2022), though these studies do not consider the baseline health incidence of the population. Policies which target areas with high baseline health incidence rates will provide significant health reductions, though these areas may already experience relatively low pollution. As such, when developing policy outcomes which targets the health outcomes of elevated pollutants, consideration of the underlying health may better support targeted policies which alleviate pollution exposure inequalities and their resulting health impacts. As previous studies have noted, both spatial resolution of air quality data and health data affects the outcomes of equity and health studies, though the underlying health characteristics are constituting a greater share of health outcomes (Southerland et al., 2021).

Our study is impacted by several limitations. First, baseline health incidence rates greatly modify the health outcomes, so there is a need for high-resolution, high fidelity health data to aid in the development of targeted interventions for exposure and heath disparities, as discussed in (Alexeeff et al., 2018; Nguyen & Marshall, 2018b). Second, the attributable mortality changes when applying different relative risks, which we try to mitigate by using meta-analysis values for the relative risks. Third, our study was limited by available health data and studies, for example, the RR for NO_2 and $PM_{2.5}$ asthma incidence in Anderson et al., 2013 are reported for all ages, but we do not have asthma incidence rates for pediatric asthma, so we limit our analysis to adult-only. Thirdly, we use CTM model output, which has biases and is limited with observations at the intraurban scale to provide ground-truthing. There are

known O_3 biases in the WRF-CMAQ model, in our simulation we have a high MDA8O_3 bias (27%, as shown in Montgomery et al., 2023), which may cause the reported MDA8O_3 -attributable outcomes to be higher than observed. This is important, because MDA8O_3 also shows the highest attributable health outcomes.

4.5 Conclusion

With high resolution air pollution output, we find that pollution exposure across demographic groups can vary widely and impact low-income and minority communities within Chicago. Combined with even higher disparities in baseline health incidence rates, the pollution-related health impacts result in disparities larger than what is found with just exposure. As such, simply relying on air pollution exposure or baseline health incidence rates cannot sufficiently identify areas of pollution impacts. In combining heath and exposure information and targeting neighborhoods with poor health incidence and high pollution exposure, we show that there can be more equitable exposure and health outcomes across the city.

Future work may use this framework to analyze exposure and health disparities within a city to analyze how to best craft exposure disparities. Further, we show how there is a need for more detailed exposure assessments and the integration of social factors in exposure modeling and health data for developing policy outcomes to mediate health-related harms.

Chapter 5

Conclusions

The establishment of environmental laws, advancements in technology, and the transition to cleaner energy sources have gradually improved urban air quality in many cities. The research presented in this dissertation encapsulate the latest advancements of applying chemical transport models for use at intraurban scales. However, the complex interplay between urbanization, industrial activities, transportation, and climate change continues to pose challenges, necessitating ongoing research and innovative solutions to ensure sustainable and healthy cities for future generations.

In the following paragraphs, I will discuss ongoing developments which relate to and advance the research established in this dissertation. Finally, I will discuss questions which are motivated by outstanding questions in this dissertation and are inspired by the advancements in the field.

Continued development into chemical transport modeling holds great promise in further improving our understanding of urban air quality dynamics. Currently, chemical transport models (CTMs) play a crucial role in simulating and predicting the dispersion of pollutants in urban environments, but future work should improve the efficiency and accuracy of CTMs. As described in Chapter 2, we use novel spatial surrogates and brute computational power to create a high-resolution simulation of air quality over the Southern-Lake Michigan region, which complements research which use reduced complexity models to create similar results for use in health studies (K. M. Foley et al., 2014; Tessum et al., 2017). A benefit of the reduced complexity model is shown from the lower computational costs, resulting in longer periods of study (>1 year) and engendering confidence in results for long-term exposure.

Other CTM developments have increased the resource intensiveness and complexity of the model. Large eddy simulations (LES) use mathematical models to simulate turbulence, using direct numerical simulations of the Navier-Stokes equation to solve for fluid flow. This method is significantly

more computationally expensive than a CTM, as CTMs decrease computational costs by relying on parametrizations to simulate fluid flow. The most recent advancement in fine-resolution modeling are CTM-LES, which have been applied to model pollutant build-up in urban canyons (Chan & Butler, 2021), canopies (Clifton et al., 2022), and urban areas (Wang et al., 2023), though these studies typically simulate a short time periods (24-hours). CTM-LES are better at simulating atmospheric mixing and concentrations of pollutants near emission sources. However, the application of operational CTM-LES to create serviceable forecasts are limited by the computational costs.

An important advancement that can increase efficiencies in CTMs, CTM-LES and reduced complexity modeling comes from advancements in physical modeling techniques utilizing machine learning (ML), artificial intelligence (AI), and neural networks (NN). These approaches have the potential to revolutionize CTMs by reducing computation expenses, increasing accuracies, and allowing for more extensive studies in attribution, health impact, and policy (Cabaneros et al., 2019; Cordova et al., 2021; Zhong et al., 2023). As it stands, the research outlined in Chapter 2 currently encapsulates the state-of-the-science air quality simulations, and these data are used to train NN models that are not based in first principles (Schneider et al., 2017; Schultz et al., 2021). Research into hybrid modeling approaches which incorporate physical process models (CTMs, CTM-LES) with ML/AI have been able to improve the simulation of flows in atmospheric boundary layer (McCandless et al., 2022) and increase the computational efficiency of CTMs (Reichstein et al., 2019; Schultz et al., 2021). With proper application of ML and AI, these modeling enhancements will enable researchers to quickly investigate complex interactions between pollutant sources, atmospheric conditions, and urban morphology. With these advancements, we can enhance our ability to develop targeted and effective mitigation strategies for improving air quality in cities.

In addition to the advancements in CTMs, the future of intraurban pollution monitoring is poised for significant progress. These sensor networks, coupled with advances in data analysis and integration, enable real-time monitoring of pollutant levels at finer spatial scales within urban areas (Kumar et al.,

2015). Furthermore, the emergence of geostationary satellite instruments presents an unprecedented opportunity for intraurban monitoring and the identification of spatial pollution patterns at new temporal scales (Zoogman et al., 2017). Future work may use a combination of low-cost sensors and satellites to enhance air quality alerts (P. Gupta et al., 2018), particularly as wildfires grow more frequent and models cannot capture missing emissions. However, as more data becomes available, it will be crucial to address the challenges of data mismatch and integrate heterogeneous datasets effectively. The research conducted in Chapter 3 provides valuable insights into the development of methodologies and tools to handle such data discrepancies, ensuring the reliability and usefulness of future intraurban pollution monitoring systems.

An additional benefit to continue to refine satellite data products and low-cost sensors is that areas lacking reference-grade monitors will increase information and this advancement will support environmental justice (EJ) on a global scale. While the context of the EJ movement in this dissertation has focused on US environmental racism, the global EJ movement has highlighted how countries such as the US has benefitted from exploiting less-developed regions and contributed directly to their environmental degradation. Currently, most attributable air pollution mortalities are concentrated in the global south (Southerland et al., 2022; Vohra et al., 2021). In the future, the global south and Africa are poised to experience substantial population growth, increased urbanization, and increased pollutant-attributable mortality (Chowdhury et al., 2018; Yang et al., 2022). However, despite these trends, these regions still have significantly fewer monitoring resources compared to more developed regions like the US. One way for the US to support global EJ movements is to develop and freely share technological advancements in satellite data products and low-cost sensors. This path benefits the US and global community further by improving emissions inventories (Geng et al., 2017), informing climate policy (Finer et al., 2018; G. Pan et al., 2021), and by improving CTM simulations (Murray et al., 2012) in areas which are understudied. Further, future research should continue to establish the critical link between air quality and public health, particularly in the context of climate change and changing emissions. Climate change poses a significant

threat to air quality, as rising temperatures and changing weather patterns can exacerbate pollution episodes (Horton et al., 2014) and introduce new challenges in controlling pollutants. Moreover, changing demographics, such as urbanization and population shifts, will lead to variations in health outcomes even if air quality improvements are achieved (Yang et al., 2022). As highlighted in Chapter 4, the underlying health of the population significantly contributes to inequalities for minority racial and ethnic and low-income groups, with this effect apparent regardless of pollutant species or exposure. High-fidelity health data is essential for accurately assessing the impact of air pollution on human health and developing effective policies to safeguard public well-being. By leveraging comprehensive health data, policymakers can make informed decisions to mitigate health risks associated with air pollution and promote equitable health outcomes for all members of society. This necessitates the development of country-wide (or global) high-resolution health datasets.

Given the direction of the field, I will close this dissertation by discussing the outstanding questions that this research inspires and a brief description of how to address the questions.

(i) What inputs should be modified optimize the WRF-CMAQ configuration over Chicago? In Chapter 2, I hypothesize that some of the model-observation mismatch can be attributed to poor boundary layer simulation. To address this, the physics in the model could benefit from integrating observed lake surface temperatures (like Abdi-Oskouei et al., 2020) and modifying urban parameterizations, such as land-use and buildings (De La Paz et al., 2016; Xu & Chen, 2021). In Chapter 3, I also identify regions of northwest Chicago which have elevated in NO_2 in observations but not simulations. In conjunction with complaints about increased warehousing (Waddell, 2021), input emissions can be modified to consider off-network idling at warehouse locations, thus perhaps better reflecting the changing emitting behaviors [project currently underway in collaboration with V. Lang].

(ii) What are the benefits of running a CTM vs. CTM-LES vs. reduced complexity model vs. NN? Beyond obvious computational tradeoffs, there are no studies which provide direct comparison of output across model types. In part this is limited by data validation, but continued advancements in

observations and computation makes this a question worth revisiting. Using a Chicago case study during the Eclipse campaign could provide interesting information about surface-level concentrations, though a full air campaign would be the best benchmark for column performance.

(iii) Is Chicago unique in the pattern observed for EJ exposure vs. health impacts? In Chapter 4, we find that Chicago has smaller exposure disparities across racial and ethnic groups than the exposure disparities reported by national-scale studies, though there are several methodological differences that make direct comparisons difficult. Future work should create national-scale high resolution CTM simulation and analyze the EJ implications of exposure vs. impact disparities across all major cities. By clarifying where health outcomes are driven by exposure vs. baseline health incidence, then better policies can be enacted to address the racist outcomes of environmental pollution.

(iv) How do we structure climate policy outcomes to maximize exposure, health, and justice benefits across the US? There are many concerns about changing our current infrastructure into greener alternatives, such as localized pollution increases from increased vehicle electrification, environmental degradation due to raw materials, and the cost associated with changing anything. Applying CTMs with life-cycle analysis can address these concerns and test out numerous policies. With increased computational efficiencies, many policy iterations could be performed, and policies could be selected which maximize benefits.

In conclusion, by advancing CTM development to become more efficient and accurate, we can facilitate more comprehensive studies on pollutant dispersion in high-impact urban areas. Additionally, the use of low-cost sensor networks and geostationary satellites holds great potential for enhanced intraurban pollution monitoring and spatial pattern identification. However, it is essential to address data mismatch challenges and develop robust methodologies for data integration. Lastly, a deeper understanding of the intricate relationship between air quality and health, supported by high-quality health data, is crucial for developing policies that prioritize public health and promote sustainable and healthy cities for future generations.

Thank you for reading this dissertation, fellow graduate students, my committee, and 4 other people.

References

Abdi-Oskouei, M., Carmichael, G., Christiansen, M., Ferrada, G., Roozitalab, B., Sobhani, N., et al. (2020). Sensitivity of Meteorological Skill to Selection of WRF-Chem Physical Parameterizations and Impact on Ozone Prediction During the Lake Michigan Ozone Study (LMOS). *Journal of Geophysical Research: Atmospheres*, *125*(5). https://doi.org/10.1029/2019JD031971

Act, C. A. (1963). Public Law 88-206 (pp. 392–401). Presented at the 88th Congress, 1st Session.

Act, C. A. (1970). Public Law 91-604 (Vol. 30). Presented at the 91st Congress, 2ndsession.

Alexeeff, S. E., Roy, A., Shan, J., Liu, X., Messier, K., Apte, J. S., et al. (2018). High-resolution mapping of traffic related air pollution with Google street view cars and incidence of cardiovascular events within neighborhoods in Oakland, CA. *Environmental Health*, *17*(1), 38. https://doi.org/10.1186/s12940-018-0382-1

Anderson, H. R., Favarato, G., & Atkinson, R. W. (2013). Long-term exposure to air pollution and the incidence of asthma: meta-analysis of cohort studies. *Air Quality, Atmosphere & Health*, *6*(1), 47–56. https://doi.org/10.1007/s11869-011-0144-5

Arrigo, K. R. (2003). Impact of a deep ozone hole on Southern Ocean primary production. *Journal of Geophysical Research*, *108*(C5), 3154. https://doi.org/10.1029/2001JC001226

Ashok, A., & Barrett, S. R. H. (2016). Adjoint-based computation of U.S. nationwide ozone exposure isopleths. *Atmospheric Environment*, *133*, 68–80. https://doi.org/10.1016/j.atmosenv.2016.03.025

Baden, B. M., & Coursey, D. L. (2002). The locality of waste sites within the city of Chicago: a demographic, social, and economic analysis. *Resource and Energy Economics*, *24*(1–2), 53–93. https://doi.org/10.1016/S0928-7655(01)00060-4

Barkjohn, K. K., Gantt, B., & Clements, A. L. (2021). Development and application of a United States-wide correction for PM 2.5 data collected with the PurpleAir sensor. *Atmospheric Measurement Techniques*, *14*(6), 4617–4637.

Bates, D. R., & Nicolet, M. (1950). The photochemistry of atmospheric water vapor. *Journal of Geophysical Research*, *55*(3), 301–327. https://doi.org/10.1029/JZ055i003p00301

B.H. Baek, & Seppanen, C. (2018, July 25). Bokhaeng/Smoke: Smoke V4.5 Public Release (April 2017). Zenodo. https://doi.org/10.5281/ZENODO.1321280

Bickford, E., Holloway, T., Karambelas, A., Johnston, M., Adams, T., Janssen, M., & Moberg, C. (2014). Emissions and Air Quality Impacts of Truck-to-Rail Freight Modal Shifts in the Midwestern United States. *Environmental Science & Technology*, *48*(1), 446–454. https://doi.org/10.1021/es4016102

Blanchard, C. L., Tanenbaum, S., & Lawson, D. R. (2008). Differences between Weekday and Weekend Air Pollutant Levels in Atlanta; Baltimore; Chicago; Dallas–Fort Worth; Denver; Houston; New York; Phoenix; Washington, DC; and Surrounding Areas. *Journal of the Air & Waste Management Association*, *58*(12), 1598–1615. https://doi.org/10.3155/1047-3289.58.12.1598

Bowman, K. P., & Krueger, A. J. (1985). A global climatology of total ozone from the Nimbus 7 total ozone mapping spectrometer. *Journal of Geophysical Research: Atmospheres*, *90*(D5), 7967–7976.

Brewer, A. W. (1949). Evidence for a world circulation provided by the measurements of helium and water vapour distribution in the stratosphere. *Quarterly Journal of the Royal Meteorological Society*, *75*(326), 351–363. https://doi.org/10.1002/qj.49707532603

Brimblecombe, P. (1998). History of urban air pollution. *Urban Air Pollution—European Aspects*, 7–20.

Buonocore, J. J., Reka, S., Yang, D., Chang, C., Roy, A., Thompson, T., et al. (2023). Air pollution and health impacts of oil & gas production in the United States. *Environmental Research: Health*, *1*(2), 021006. https://doi.org/10.1088/2752-5309/acc886

Byun, D., & Schere, K. L. (2006). Review of the Governing Equations, Computational Algorithms, and Other Components of the Models-3 Community Multiscale Air Quality (CMAQ) Modeling System. *Applied Mechanics Reviews*, *59*(2), 51. https://doi.org/10.1115/1.2128636

Cabaneros, S. M., Calautit, J. K., & Hughes, B. R. (2019). A review of artificial neural network models for ambient air pollution prediction. *Environmental Modelling & Software*, *119*, 285–304. https://doi.org/10.1016/j.envsoft.2019.06.014

Campbell, P. C., Bash, J. O., & Spero, T. L. (2019). Updates to the Noah Land Surface Model in WRF-CMAQ to Improve Simulated Meteorology, Air Quality, and Deposition. *Journal of Advances in Modeling Earth Systems*, *11*(1), 231–256. https://doi.org/10.1029/2018MS001422

Carmichael, S., & Hamilton, C. V. (1967). *Black power: the politics of liberation in America* (Vintage ed). New York: Vintage Books.

Castillo, M. D., Kinney, P. L., Southerland, V., Arno, C. A., Crawford, K., van Donkelaar, A., et al. (2021). Estimating Intra-Urban Inequities in PM $_{2.5}$ -Attributable Health Impacts: A Case Study for Washington, DC. *GeoHealth*, *5*(11). https://doi.org/10.1029/2021GH000431

Catlett, C. E., Beckman, P. H., Sankaran, R., & Galvin, K. K. (2017). Array of things: a scientific research instrument in the public way: platform design and early lessons learned (pp. 26–33). Presented at the Proceedings of the 2nd international workshop on science of smart city operations and platforms engineering.

Chambliss, S. E., Pinon, C. P. R., Messier, K. P., LaFranchi, B., Upperman, C. R., Lunden, M. M., et al. (2021). Local- and regional-scale racial and ethnic disparities in air pollution determined by long-term mobile monitoring. *Proceedings of the National Academy of Sciences*, *118*(37), e2109249118. https://doi.org/10.1073/pnas.2109249118

Chan, E. C., & Butler, T. M. (2021). urbanChemFoam 1.0: large-eddy simulation of non-stationary chemical transport of traffic emissions in an idealized street canyon. *Geoscientific Model Development*, *14*(7), 4555–4572. https://doi.org/10.5194/gmd-14-4555-2021

Chapman, S. (1930). On ozone and atomic oxygen in the upper atmosphere. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *10*(64), 369–383. https://doi.org/10.1080/14786443009461588

Chen, T.-M., Kuschner, W. G., Gokhale, J., & Shofer, S. (2007). Outdoor air pollution: ozone health effects. *The American Journal of the Medical Sciences*, *333*(4), 244–248.

Chowdhury, S., Dey, S., & Smith, K. R. (2018). Ambient PM2.5 exposure and expected premature mortality to 2100 in India under climate change scenarios. *Nature Communications*, *9*(1), 318. https://doi.org/10.1038/s41467-017-02755-y

Clark, L. P., Millet, D. B., & Marshall, J. D. (2014). National Patterns in Environmental Injustice and Inequality: Outdoor NO2 Air Pollution in the United States. *PLoS ONE*, *9*(4), e94431. https://doi.org/10.1371/journal.pone.0094431

Clark, L. P., Harris, M. H., Apte, J. S., & Marshall, J. D. (2022a). National and Intraurban Air Pollution Exposure Disparity Estimates in the United States: Impact of Data-Aggregation Spatial Scale. *Environmental Science & Technology Letters*, *9*(9), 786–791. https://doi.org/10.1021/acs.estlett.2c00403

Clark, L. P., Harris, M. H., Apte, J. S., & Marshall, J. D. (2022b). National and Intraurban Air Pollution Exposure Disparity Estimates in the United States: Impact of Data-Aggregation Spatial Scale. *Environmental Science & Technology Letters*, acs.estlett.2c00403. https://doi.org/10.1021/acs.estlett.2c00403

Clifton, O. E., Patton, E. G., Wang, S., Barth, M., Orlando, J., & Schwantes, R. H. (2022). Large Eddy Simulation for Investigating Coupled Forest Canopy and Turbulence Influences on Atmospheric Chemistry. *Journal of Advances in Modeling Earth Systems*, *14*(10). https://doi.org/10.1029/2022MS003078

Clough, S. A., Shephard, M. W., Mlawer, E. J., Delamere, J. S., Iacono, M. J., Cady-Pereira, K., et al. (2005). Atmospheric radiative transfer modeling: a summary of the AER codes. *Journal of*

Quantitative Spectroscopy and Radiative Transfer, *91*(2), 233–244. https://doi.org/10.1016/j.jqsrt.2004.05.058

Cordova, C. H., Portocarrero, M. N. L., Salas, R., Torres, R., Rodrigues, P. C., & López-Gonzales, J. L. (2021). Air quality assessment and pollution forecasting using artificial neural networks in Metropolitan Lima-Peru. *Scientific Reports*, *11*(1), 24232. https://doi.org/10.1038/s41598-021-03650-9

Daepp, M. I. G., Cabral, A., Ranganathan, V., Iyer, V., Counts, S., Johns, P., et al. (2022). Eclipse: An End-to-End Platform for Low-Cost, Hyperlocal Environmental Sensing in Cities. In *2022 21st ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)* (pp. 28–40). Milano, Italy: IEEE. https://doi.org/10.1109/IPSN54338.2022.00010

De La Paz, D., Borge, R., & Martilli, A. (2016). Assessment of a high resolution annual WRF-BEP/CMAQ simulation for the urban area of Madrid (Spain). *Atmospheric Environment*, *144*, 282–296. https://doi.org/10.1016/j.atmosenv.2016.08.082

Demetillo, M. A. G., Harkins, C., McDonald, B. C., Chodrow, P. S., Sun, K., & Pusede, S. E. (2021). Space-Based Observational Constraints on NO_2 Air Pollution Inequality From Diesel Traffic in Major US Cities. *Geophysical Research Letters*, *48*(17). https://doi.org/10.1029/2021GL094333

Dennis, R., Fox, T., Fuentes, M., Gilliland, A., Hanna, S., Hogrefe, C., et al. (2010). A framework for evaluating regional-scale numerical photochemical modeling systems. *Environmental Fluid Mechanics*, *10*(4), 471–489. https://doi.org/10.1007/s10652-009-9163-2

Dewitz, J. (2021). National Land Cover Database (NLCD) 2019 Products [Data set]. U.S. Geological Survey. https://doi.org/10.5066/P9KZCM54

Doak, A. G., Christiansen, M. B., Alwe, H. D., Bertram, T. H., Carmichael, G., Cleary, P., et al. (2021). Characterization of ground-based atmospheric pollution and meteorology sampling stations during the Lake Michigan Ozone Study 2017. *Journal of the Air & Waste Management Association*, *71*(7), 866–889. https://doi.org/10.1080/10962247.2021.1900000

Duncan, B. N., Prados, A. I., Lamsal, L. N., Liu, Y., Streets, D. G., Gupta, P., et al. (2014). Satellite data of atmospheric pollution for U.S. air quality applications: Examples of applications, summary of data end-user resources, answers to FAQs, and common mistakes to avoid. *Atmospheric Environment*, *94*, 647–662. https://doi.org/10.1016/j.atmosenv.2014.05.061

Dye, T. S., Roberts, P. T., & Korc, M. E. (1995). Observations of Transport Processes for Ozone and Ozone Precursors during the 1991 Lake Michigan Ozone Study. *Journal of Applied Meteorology*, *34*(8), 1877–1889. https://doi.org/10.1175/1520-0450(1995)034<1877:OOTPFO>2.0.CO;2

Emery, C., Liu, Z., Russell, A. G., Odman, M. T., Yarwood, G., & Kumar, N. (2017). Recommendations on statistics and benchmarks to assess photochemical model performance. *Journal of the Air & Waste Management Association*, *67*(5), 582–598. https://doi.org/10.1080/10962247.2016.1265027

Emmons, L. K., Schwantes, R. H., Orlando, J. J., Tyndall, G., Kinnison, D., Lamarque, J., et al. (2020). The Chemistry Mechanism in the Community Earth System Model Version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*(4). https://doi.org/10.1029/2019MS001882

EPA. (2022). *Current Nonattainment Counties for All Criteria Pollutants*. Retrieved from https://www3.epa.gov/airquality/greenbook/ancl.html

Esie, P., Daepp, M. I. G., Roseway, A., & Counts, S. (2022). Neighborhood Composition and Air Pollution in Chicago: Monitoring Inequities With a Dense, Low-Cost Sensing Network, 2021. *American Journal of Public Health*, *112*(12), 1765–1773. https://doi.org/10.2105/AJPH.2022.307068

Etzel, R. A. (2003). How environmental exposures influence the development and exacerbation of asthma. *Pediatrics*, *112*(Supplement_1), 233–239.

Eyth, A. (2021). National Emission Inventory (NEI) 2016 modeling platform version 2 [Data set]. UNC Dataverse. https://doi.org/10.15139/S3/SAXVSF

Eyth, A., Vukovich, J., Farkas, C., & Strum, M. (2019, September). Technical Support Document (TSD): Preparation of Emissions Inventories for the Version 7.2 2016 North American Emissions Modeling Platform. U.S. Environmental Protection Agency. Retrieved from https://www.epa.gov/sites/default/files/2019-09/documents/2016v7.2_regionalhaze_emismod_tsd_508.pdf

Feagin, J., & Bennefield, Z. (2014). Systemic racism and U.S. health care. *Social Science & Medicine*, *103*, 7–14. https://doi.org/10.1016/j.socscimed.2013.09.006

Feinglass, J., Lin, S., Thompson, J., Sudano, J., Dunlop, D., Song, J., & Baker, D. W. (2007). Baseline Health, Socioeconomic Status, and 10-Year Mortality Among Older Middle-Aged Americans: Findings From the Health and Retirement Study, 1992-2002. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *62*(4), S209–S217. https://doi.org/10.1093/geronb/62.4.S209

Finer, M., Novoa, S., Weisse, M. J., Petersen, R., Mascaro, J., Souto, T., et al. (2018). Combating deforestation: From satellite to intervention. *Science*, *360*(6395), 1303–1305. https://doi.org/10.1126/science.aat1203

Fiore, A. M., Hancock, S. E., Lamarque, J.-F., Correa, G. P., Chang, K.-L., Ru, M., et al. (2022). Understanding recent tropospheric ozone trends in the context of large internal variability: a new perspective from chemistry-climate model ensembles. *Environmental Research: Climate*, *1*(2), 025008. https://doi.org/10.1088/2752-5295/ac9cc2

Foley, K. M., Napelenok, S. L., Jang, C., Phillips, S., Hubbell, B. J., & Fulcher, C. M. (2014). Two reduced form air quality modeling techniques for rapidly calculating pollutant mitigation potential across many sources, locations and precursor emission types. *Atmospheric Environment*, *98*, 283–289. https://doi.org/10.1016/j.atmosenv.2014.08.046

Foley, T., Betterton, E. A., Robert Jacko, P. E., & Hillery, J. (2011). Lake Michigan air quality: The 1994–2003 LADCO Aircraft Project (LAP). *Atmospheric Environment*, *45*(18), 3192–3202. https://doi.org/10.1016/j.atmosenv.2011.02.033

Fountoukis, C., Koraj, Dh., Denier van der Gon, H. A. C., Charalampidis, P. E., Pilinis, C., & Pandis, S. N. (2013). Impact of grid resolution on the predicted fine PM by a regional 3-D chemical transport model. *Atmospheric Environment*, *68*, 24–32. https://doi.org/10.1016/j.atmosenv.2012.11.008

Franklin, B. A., Brook, R., & Arden Pope, C. (2015). Air Pollution and Cardiovascular Disease. *Current Problems in Cardiology*, *40*(5), 207–238. https://doi.org/10.1016/j.cpcardiol.2015.01.003

Franklin, M., Zeka, A., & Schwartz, J. (2007). Association between PM2.5 and all-cause and specific-cause mortality in 27 US communities. *Journal of Exposure Science & Environmental Epidemiology*, *17*(3), 279–287. https://doi.org/10.1038/sj.jes.7500530

Gan, C.-M., Hogrefe, C., Mathur, R., Pleim, J., Xing, J., Wong, D., et al. (2016). Assessment of the effects of horizontal grid resolution on long-term air quality trends using coupled WRF-CMAQ simulations. *Atmospheric Environment*, *132*, 207–216. https://doi.org/10.1016/j.atmosenv.2016.02.036

Garcia-Menendez, F., Monier, E., & Selin, N. E. (2017). The role of natural variability in projections of climate change impacts on U.S. ozone pollution: NATURAL VARIABILITY IN OZONE PROJECTIONS. *Geophysical Research Letters*, *44*(6), 2911–2921. https://doi.org/10.1002/2016GL071565

Gardner-Frolick, R., Boyd, D., & Giang, A. (2022). Selecting Data Analytic and Modeling Methods to Support Air Pollution and Environmental Justice Investigations: A Critical Review and Guidance Framework. *Environmental Science & Technology*, *56*(5), 2843–2860. https://doi.org/10.1021/acs.est.1c01739

Geng, G., Zhang, Q., Martin, R. V., Lin, J., Huo, H., Zheng, B., et al. (2017). Impact of spatial proxies on the representation of bottom-up emission inventories: A satellite-based analysis. *Atmospheric Chemistry and Physics*, *17*(6), 4131–4145. https://doi.org/10.5194/acp-17-4131-2017

Getis, A., & Ord, J. K. (2010). The Analysis of Spatial Association by Use of Distance Statistics. *Geographical Analysis*, *24*(3), 189–206. https://doi.org/10.1111/j.1538-4632.1992.tb00261.x

Goldberg, D. L., Loughner, C. P., Tzortziou, M., Stehr, J. W., Pickering, K. E., Marufu, L. T., & Dickerson, R. R. (2014). Higher surface ozone concentrations over the Chesapeake Bay than over the adjacent land: Observations and models from the DISCOVER-AQ and CBODAQ campaigns. *Atmospheric Environment*, *84*, 9–19. https://doi.org/10.1016/j.atmosenv.2013.11.008

Goldberg, D. L., Anenberg, S. C., Kerr, G. H., Mohegh, A., Lu, Z., & Streets, D. G. (2021). TROPOMI NO_2 in the United States: A Detailed Look at the Annual Averages, Weekly Cycles, Effects of Temperature, and Correlation With Surface NO_2 Concentrations. *Earth's Future*, *9*(4). https://doi.org/10.1029/2020EF001665

Goodkind, A. L., Tessum, C. W., Coggins, J. S., Hill, J. D., & Marshall, J. D. (2019). Fine-scale damage estimates of particulate matter air pollution reveal opportunities for location-specific mitigation of emissions. *Proceedings of the National Academy of Sciences*, *116*(18), 8775–8780. https://doi.org/10.1073/pnas.1816102116

Gotz, F. W. P., Meetham, A. R., & Dobson, G. M. B. (1934). The vertical distribution of ozone in the atmosphere. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, *145*(855), 416–446. https://doi.org/10.1098/rspa.1934.0109

Gupta, P., Doraiswamy, P., Levy, R., Pikelnaya, O., Maibach, J., Feenstra, B., et al. (2018). Impact of California Fires on Local and Regional Air Quality: The Role of a Low-Cost Sensor Network and Satellite Observations. *GeoHealth*, *2*(6), 172–181. https://doi.org/10.1029/2018GH000136

Gupta, R. S., Zhang, X., Sharp, L. K., Shannon, J. J., & Weiss, K. B. (2008). Geographic variability in childhood asthma prevalence in Chicago. *Journal of Allergy and Clinical Immunology*, *121*(3), 639-645.e1. https://doi.org/10.1016/j.jaci.2007.11.036

Harkey, M., Holloway, T., Oberman, J., & Scotty, E. (2015). An evaluation of CMAQ NO_2 using observed chemistry-meteorology correlations: CMAQ EVALUATION WITH NO_2 - METOROLOGY CORRELATIONS. *Journal of Geophysical Research: Atmospheres*, *120*(22), 11,775-11,797. https://doi.org/10.1002/2015JD023316

Henze, D. K., Hakami, A., & Seinfeld, J. H. (2007). Development of the adjoint of GEOS-Chem. *Atmospheric Chemistry and Physics*, *7*(9), 2413–2433. https://doi.org/10.5194/acp-7-2413-2007

Hoek, G., Krishnan, R. M., Beelen, R., Peters, A., Ostro, B., Brunekreef, B., & Kaufman, J. D. (2013). Long-term air pollution exposure and cardio- respiratory mortality: a review. *Environmental Health*, *12*(1), 43. https://doi.org/10.1186/1476-069X-12-43

Hogrefe, C., Pouliot, G., Wong, D., Torian, A., Roselle, S., Pleim, J., & Mathur, R. (2015). Annual application and evaluation of the online coupled WRF–CMAQ system over North America under AQMEII phase 2. *Atmospheric Environment*, *115*, 683–694. https://doi.org/10.1016/j.atmosenv.2014.12.034

Horton, D. E., Skinner, C. B., Singh, D., & Diffenbaugh, N. S. (2014). Occurrence and persistence of future atmospheric stagnation events. *Nature Climate Change*, *4*(8), 698–703. https://doi.org/10.1038/nclimate2272

Hu, J., Chen, J., Ying, Q., & Zhang, H. (2016). One-year simulation of ozone and particulate matter in Chinausing WRF/CMAQ modeling system. *Atmospheric Chemistry and Physics*, *16*(16), 10333–10350. https://doi.org/10.5194/acp-16-10333-2016

Hu, J., Ostro, B., Zhang, H., Ying, Q., & Kleeman, M. J. (2019). Using Chemical Transport Model Predictions To Improve Exposure Assessment of $PM_{2.5}$ Constituents. *Environmental Science & Technology Letters*, *6*(8), 456–461. https://doi.org/10.1021/acs.estlett.9b00396

Hu, X.-M., Doughty, D. C., Sanchez, K. J., Joseph, E., & Fuentes, J. D. (2012). Ozone variability in the atmospheric boundary layer in Maryland and its implications for vertical transport model. *Atmospheric Environment*, *46*, 354–364. https://doi.org/10.1016/j.atmosenv.2011.09.054

Huangfu, P., & Atkinson, R. (2020). Long-term exposure to NO2 and O3 and all-cause and respiratory mortality: A systematic review and meta-analysis. *Environment International*, *144*, 105998. https://doi.org/10.1016/j.envint.2020.105998

Hughes, D. D., Christiansen, M. B., Milani, A., Vermeuel, M. P., Novak, G. A., Alwe, H. D., et al. (2021). PM2.5 chemistry, organosulfates, and secondary organic aerosol during the 2017 Lake Michigan Ozone Study. *Atmospheric Environment*, *244*, 117939. https://doi.org/10.1016/j.atmosenv.2020.117939

Jbaily, A., Zhou, X., Liu, J., Lee, T.-H., Kamareddine, L., Verguet, S., & Dominici, F. (2022). Air pollution exposure disparities across US population and income groups. *Nature*, *601*(7892), 228–233. https://doi.org/10.1038/s41586-021-04190-y

Jiang, X., & Yoo, E. (2018). The importance of spatial resolutions of Community Multiscale Air Quality (CMAQ) models on health impact assessment. *Science of The Total Environment*, *627*, 1528–1543. https://doi.org/10.1016/j.scitotenv.2018.01.228

Jin, X., Fiore, A. M., Murray, L. T., Valin, L. C., Lamsal, L. N., Duncan, B., et al. (2017). Evaluating a Space-Based Indicator of Surface Ozone-NO $_x$ -VOC Sensitivity Over Midlatitude Source Regions and Application to Decadal Trends: Space-Based Indicator of O $_3$ Sensitivity. *Journal of Geophysical Research: Atmospheres*, *122*(19), 10,439-10,461. https://doi.org/10.1002/2017JD026720

Jin, X., Fiore, A., Boersma, K. F., Smedt, I. D., & Valin, L. (2020). Inferring Changes in Summertime Surface Ozone–NO $_x$ –VOC Chemistry over U.S. Urban Areas from Two Decades of Satellite and Ground-Based Observations. *Environmental Science & Technology*, *54*(11), 6518–6529. https://doi.org/10.1021/acs.est.9b07785

Jing, P., & Goldberg, D. L. (2022). Influence of conducive weather on ozone in the presence of reduced NOx emissions: A case study in Chicago during the 2020 lockdowns. *Atmospheric Pollution Research*, *13*(2), 101313. https://doi.org/10.1016/j.apr.2021.101313

Jing, P., Lu, Z., Xing, J., Streets, D. G., Tan, Q., O'Brien, T., & Kamberos, J. (2014). Response of the summertime ground-level ozone trend in the Chicago area to emission controls and temperature changes, 2005–2013. *Atmospheric Environment*, *99*, 630–640. https://doi.org/10.1016/j.atmosenv.2014.10.035

Jing, P., Lu, Z., & Steiner, A. L. (2017). The ozone-climate penalty in the Midwestern U.S. *Atmospheric Environment*, *170*, 130–142. https://doi.org/10.1016/j.atmosenv.2017.09.038

Johnson, T. J. (2020). Intersection of Bias, Structural Racism, and Social Determinants With Health Care Inequities. *Pediatrics*, *146*(2), e2020003657. https://doi.org/10.1542/peds.2020-003657

Kain, J. S. (2004). The Kain–Fritsch Convective Parameterization: An Update. *Journal of Applied Meteorology*, *43*(1), 170–181. https://doi.org/10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2

Kang, D., Mathur, R., & Rao, S. T. (2010). Real-time bias-adjusted O3 and PM2. 5 air quality index forecasts and their performance evaluations over the continental United States. *Atmospheric Environment*, *44*(18), 2203–2212.

Kerr, G. H., Goldberg, D. L., & Anenberg, S. C. (2021). COVID-19 pandemic reveals persistent disparities in nitrogen dioxide pollution. *Proceedings of the National Academy of Sciences*, *118*(30), e2022409118. https://doi.org/10.1073/pnas.2022409118

Koplitz, S., Simon, H., Henderson, B., Liljegren, J., Tonnesen, G., Whitehill, A., & Wells, B. (2022). Changes in Ozone Chemical Sensitivity in the United States from 2007 to 2016. *ACS Environmental Au*, *2*(3), 206–222. https://doi.org/10.1021/acsenvironau.1c00029

Kumar, P., Morawska, L., Martani, C., Biskos, G., Neophytou, M., Di Sabatino, S., et al. (2015). The rise of low-cost sensing for managing air pollution in cities. *Environment International*, *75*, 199–205. https://doi.org/10.1016/j.envint.2014.11.019

Kurt, O. K., Zhang, J., & Pinkerton, K. E. (2016). Pulmonary health effects of air pollution: *Current Opinion in Pulmonary Medicine*, *22*(2), 138–143. https://doi.org/10.1097/MCP.0000000000000248

LADCO. (2022, July 27). Lake Michigan Air Directors Consortium (LADCO) Technical Support Document. Retrieved from https://www.ladco.org/technical/ladco-internal/ladco-projects/ladco-2015-o3-naaqs-moderate-area-sip-technical-support-document/

Lane, H. M., Morello-Frosch, R., Marshall, J. D., & Apte, J. S. (2022). Historical Redlining Is Associated with Present-Day Air Pollution Disparities in U.S. Cities. *Environmental Science & Technology Letters*, *9*(4), 345–350. https://doi.org/10.1021/acs.estlett.1c01012

Laughner, J. L., & Cohen, R. C. (2019). Direct observation of changing NO_x lifetime in North American cities. *Science*, *366*(6466), 723–727. https://doi.org/10.1126/science.aax6832

Li, X., Chen, Q., Zheng, X., Li, Y., Han, M., Liu, T., et al. (2019). Effects of ambient ozone concentrations with different averaging times on asthma exacerbations: A meta-analysis. *Science of The Total Environment*, *691*, 549–561. https://doi.org/10.1016/j.scitotenv.2019.06.382

Liu, X.-H., Zhang, Y., Olsen, K. M., Wang, W.-X., Do, B. A., & Bridgers, G. M. (2010). Responses of future air quality to emission controls over North Carolina, Part I: Model evaluation for current-year simulations. *Atmospheric Environment*, *44*(20), 2443–2456. https://doi.org/10.1016/j.atmosenv.2010.04.002

Makri, A., & Stilianakis, N. I. (2008). Vulnerability to air pollution health effects. *International Journal of Hygiene and Environmental Health*, *211*(3–4), 326–336. https://doi.org/10.1016/j.ijheh.2007.06.005

Manson, Steven, Schroeder, Jonathan, Van Riper, David, Kugler, Tracy, & Ruggles, Steven. (2022). National Historical Geographic Information System: Version 17.0 (Version 17.0) [Data set]. Minneapolis, MN: IPUMS. https://doi.org/10.18128/D050.V17.0

Margitan, J. J., Barnes, R. A., Brothers, G. B., Butler, J., Burris, J., Connor, B. J., et al. (1995). Stratospheric Ozone Intercomparison Campaign (STOIC) 1989: Overview. *Journal of Geophysical Research*, *100*(D5), 9193. https://doi.org/10.1029/95JD00509

Martin, R. V. (2008). Satellite remote sensing of surface air quality. *Atmospheric Environment*, *42*(34), 7823–7843.

Marzban, C., & Sandgathe, S. (2009). Verification with Variograms. *Weather and Forecasting*, *24*(4), 1102–1120. https://doi.org/10.1175/2009WAF2222122.1

McCandless, T., Gagne, D. J., Kosović, B., Haupt, S. E., Yang, B., Becker, C., & Schreck, J. (2022). Machine Learning for Improving Surface-Layer-Flux Estimates. *Boundary-Layer Meteorology*, *185*(2), 199–228. https://doi.org/10.1007/s10546-022-00727-4

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

Mohai, P., Pellow, D., & Roberts, J. T. (2009). Environmental Justice. *Annual Review of Environment and Resources*, *34*(1), 405–430. https://doi.org/10.1146/annurev-environ-082508-094348

Mohegh, A., Goldberg, D., Achakulwisut, P., & Anenberg, S. C. (2021). Sensitivity of estimated NO_2-attributable pediatric asthma incidence to grid resolution and urbanicity. *Environmental Research Letters*, *16*(1), 014019. https://doi.org/10.1088/1748-9326/abce25

Moran, G. R. (1954). Air Pollution Control Act and Its Administration. *Rutgers L. Rev.*, *9*, 640.

Morello-Frosch, R., & Jesdale, B. M. (2006). Separate and Unequal: Residential Segregation and Estimated Cancer Risks Associated with Ambient Air Toxics in U.S. Metropolitan Areas. *Environmental Health Perspectives*, *114*(3), 386–393. https://doi.org/10.1289/ehp.8500

Mudway, I. (2000). Ozone and the lung: a sensitive issue. *Molecular Aspects of Medicine*, *21*(1–2), 1–48. https://doi.org/10.1016/S0098-2997(00)00003-0

Murray, L. T., Jacob, D. J., Logan, J. A., Hudman, R. C., & Koshak, W. J. (2012). Optimized regional and interannual variability of lightning in a global chemical transport model constrained by

LIS/OTD satellite data: IAV of lightning constrained by LIS/OTD. *Journal of Geophysical Research: Atmospheres*, *117*(D20). https://doi.org/10.1029/2012JD017934

Nelson, R. K., Winling, L., Marciano, R., Connolly, N., & Ayers, E. L. (2021). Mapping inequality. *American Panorama*.

Nguyen, N. P., & Marshall, J. D. (2018a). Impact, efficiency, inequality, and injustice of urban air pollution: variability by emission location. *Environmental Research Letters*, *13*(2), 024002. https://doi.org/10.1088/1748-9326/aa9cb5

Nguyen, N. P., & Marshall, J. D. (2018b). Impact, efficiency, inequality, and injustice of urban air pollution: variability by emission location. *Environmental Research Letters*, *13*(2), 024002. https://doi.org/10.1088/1748-9326/aa9cb5

Odman, M. T., White, A. T., Doty, K., McNider, R. T., Pour-Biazar, A., Qin, M., et al. (2019). Examination of Nudging Schemes in the Simulation of Meteorology for Use in Air Quality Experiments: Application in the Great Lakes Region. *Journal of Applied Meteorology and Climatology*, *58*(11), 2421–2436. https://doi.org/10.1175/JAMC-D-18-0206.1

O'Leary, B. F., & Lemke, L. D. (2014). Modeling spatiotemporal variability of intra-urban air pollutants in Detroit: A pragmatic approach. *Atmospheric Environment*, *94*, 417–427. https://doi.org/10.1016/j.atmosenv.2014.05.010

Otte, T. L. (2008). The Impact of Nudging in the Meteorological Model for Retrospective Air Quality Simulations. Part I: Evaluation against National Observation Networks. *Journal of Applied Meteorology and Climatology*, *47*(7), 1853–1867. https://doi.org/10.1175/2007JAMC1790.1

Pan, G., Xu, Y., & Ma, J. (2021). The potential of CO2 satellite monitoring for climate governance: A review. *Journal of Environmental Management*, *277*, 111423. https://doi.org/10.1016/j.jenvman.2020.111423

Pan, S., Choi, Y., Roy, A., & Jeon, W. (2017). Allocating emissions to 4 km and 1 km horizontal spatial resolutions and its impact on simulated NOx and O3 in Houston, TX. *Atmospheric Environment*, *164*, 398–415. https://doi.org/10.1016/j.atmosenv.2017.06.026

Paolella, D. A., Tessum, C. W., Adams, P. J., Apte, J. S., Chambliss, S., Hill, J., et al. (2018). Effect of Model Spatial Resolution on Estimates of Fine Particulate Matter Exposure and Exposure Disparities in the United States. *Environmental Science & Technology Letters*, *5*(7), 436–441. https://doi.org/10.1021/acs.estlett.8b00279

Paradies, Y., Ben, J., Denson, N., Elias, A., Priest, N., Pieterse, A., et al. (2015). Racism as a Determinant of Health: A Systematic Review and Meta-Analysis. *PLOS ONE*, *10*(9), e0138511. https://doi.org/10.1371/journal.pone.0138511

Penn, E., & Holloway, T. (2020). Evaluating current satellite capability to observe diurnal change in nitrogen oxides in preparation for geostationary satellite missions. *Environmental Research Letters*, *15*(3), 034038. https://doi.org/10.1088/1748-9326/ab6b36

Persinger, R. L., Poynter, M. E., Ckless, K., & Janssen-Heininger, Y. M. (2002). Molecular mechanisms of nitrogen dioxide induced epithelial injury in the lung. *Molecular and Cellular Biochemistry*, *234*, 71–80.

Peters, D. R., Popoola, O. A., Jones, R. L., Martin, N. A., Mills, J., Fonseca, E. R., et al. (2022). Evaluating uncertainty in sensor networks for urban air pollution insights. *Atmospheric Measurement Techniques*, *15*(2), 321–334.

Pinto, J. P., Lefohn, A. S., & Shadwick, D. S. (2004). Spatial Variability of PM $_{2.5}$ in Urban Areas in the United States. *Journal of the Air & Waste Management Association*, *54*(4), 440–449. https://doi.org/10.1080/10473289.2004.10470919

Pleim, J. E. (2007). A Combined Local and Nonlocal Closure Model for the Atmospheric Boundary Layer. Part I: Model Description and Testing. *Journal of Applied Meteorology and Climatology*, *46*(9), 1383–1395. https://doi.org/10.1175/JAM2539.1

Pleim, J. E., & Gilliam, R. (2009). An Indirect Data Assimilation Scheme for Deep Soil Temperature in the Pleim–Xiu Land Surface Model. *Journal of Applied Meteorology and Climatology*, *48*(7), 1362–1376. https://doi.org/10.1175/2009JAMC2053.1

Pleim, J. E., & Xiu, A. (2003). Development of a Land Surface Model. Part II: Data Assimilation. *Journal of Applied Meteorology*, *42*(12), 1811–1822. https://doi.org/10.1175/1520-0450(2003)042<1811:DOALSM>2.0.CO;2

Pozzer, A., Anenberg, S. C., Dey, S., Haines, A., Lelieveld, J., & Chowdhury, S. (2023). Mortality Attributable to Ambient Air Pollution: A Review of Global Estimates. *GeoHealth*, *7*(1), e2022GH000711. https://doi.org/10.1029/2022GH000711

Qin, M., Yu, H., Hu, Y., Russell, A. G., Odman, M. T., Doty, K., et al. (2019). Improving ozone simulations in the Great Lakes Region: The role of emissions, chemistry, and dry deposition. *Atmospheric Environment*, *202*, 167–179. https://doi.org/10.1016/j.atmosenv.2019.01.025

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, *566*(7743), 195–204. https://doi.org/10.1038/s41586-019-0912-1

Rethinking the Ozone Problem in Urban and Regional Air Pollution. (1991) (p. 1889). Washington, D.C.: National Academies Press. https://doi.org/10.17226/1889

Rowland, F. S. (1990). Stratospheric ozone depletion by chlorofluorocarbons. *Ambio*, 281–292.

Ryan, P. H., & LeMasters, G. K. (2007). A Review of Land-use Regression Models for Characterizing Intraurban Air Pollution Exposure. *Inhalation Toxicology*, *19*(sup1), 127–133. https://doi.org/10.1080/08958370701495998

Sandoval, J. S. O. (2011). Neighborhood Diversity and Segregation in the Chicago Metropolitan Region, 1980-2000. *Urban Geography*, *32*(5), 609–640. https://doi.org/10.2747/0272-3638.32.5.609

Schneider, T., Lan, S., Stuart, A., & Teixeira, J. (2017). Earth System Modeling 2.0: A Blueprint for Models That Learn From Observations and Targeted High-Resolution Simulations. *Geophysical Research Letters*, *44*(24). https://doi.org/10.1002/2017GL076101

Schnell, J. L., & Prather, M. J. (2017). Co-occurrence of extremes in surface ozone, particulate matter, and temperature over eastern North America. *Proceedings of the National Academy of Sciences*, *114*(11), 2854–2859. https://doi.org/10.1073/pnas.1614453114

Schultz, M. G., Betancourt, C., Gong, B., Kleinert, F., Langguth, M., Leufen, L. H., et al. (2021). Can deep learning beat numerical weather prediction? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *379*(2194), 20200097. https://doi.org/10.1098/rsta.2020.0097

Seinfeld, J. H., & Pandis, S. N. (2016). *Atmospheric chemistry and physics: from air pollution to climate change* (Third edition). Hoboken, New Jersey: John Wiley & Sons.

Seposo, X., Ueda, K., Sugata, S., Yoshino, A., & Takami, A. (2020). Short-term effects of air pollution on daily single- and co-morbidity cardiorespiratory outpatient visits. *Science of The Total Environment*, *729*, 138934. https://doi.org/10.1016/j.scitotenv.2020.138934

Sharma, S., Sharma, P., & Khare, M. (2017). Photo-chemical transport modelling of tropospheric ozone: A review. *Atmospheric Environment*, *159*, 34–54. https://doi.org/10.1016/j.atmosenv.2017.03.047

Simon, H., Baker, K. R., & Phillips, S. (2012). Compilation and interpretation of photochemical model performance statistics published between 2006 and 2012. *Atmospheric Environment*, *61*, 124–139. https://doi.org/10.1016/j.atmosenv.2012.07.012

Sinden, N. J., & Stockley, R. A. (2010). Systemic inflammation and comorbidity in COPD: a result of "overspill" of inflammatory mediators from the lungs? Review of the evidence. *Thorax*, *65*(10), 930–936. https://doi.org/10.1136/thx.2009.130260

Skamarock, W., Klemp, J., Dudhia, J., Gill, D., Barker, D., Wang, W., et al. (2008). *A Description of the Advanced Research WRF Version 3* [Application/pdf] (p. 1002 KB). UCAR/NCAR. https://doi.org/10.5065/D68S4MVH

SMOKE v4.5 User's Manual. (2017, April 30). The institute for the Environment - The University of North Carolina at Chapel Hill. Retrieved from https://www.cmascenter.org/smoke/documentation/4.5/manual_smokev45.pdf

Sørensen, M., Daneshvar, B., Hansen, M., Dragsted, L. O., Hertel, O., Knudsen, L., & Loft, S. (2003). Personal PM2. 5 exposure and markers of oxidative stress in blood. *Environmental Health Perspectives*, *111*(2), 161–166.

Sørensen, M., Poulsen, A. H., Hvidtfeldt, U. A., Christensen, J. H., Brandt, J., Frohn, L. M., et al. (2023). Effects of Sociodemographic Characteristics, Comorbidity, and Coexposures on the Association between Air Pollution and Type 2 Diabetes: A Nationwide Cohort Study. *Environmental Health Perspectives*, *131*(2), 027008. https://doi.org/10.1289/EHP11347

Southerland, V. A., Anenberg, S. C., Harris, M., Apte, J., Hystad, P., van Donkelaar, A., et al. (2021). Assessing the Distribution of Air Pollution Health Risks within Cities: A Neighborhood-Scale Analysis Leveraging High-Resolution Data Sets in the Bay Area, California. *Environmental Health Perspectives*, *129*(3), EHP7679, 037006. https://doi.org/10.1289/EHP7679

Southerland, V. A., Brauer, M., Mohegh, A., Hammer, M. S., van Donkelaar, A., Martin, R. V., et al. (2022). Global urban temporal trends in fine particulate matter (PM2·5) and attributable health burdens: estimates from global datasets. *The Lancet Planetary Health*, *6*(2), e139–e146. https://doi.org/10.1016/S2542-5196(21)00350-8

Suresh, P., Daniel, J. V., Parthasarathy, V., & Aswathy, R. (2014). A state of the art review on the Internet of Things (IoT) history, technology and fields of deployment (pp. 1–8). Presented at the 2014 International conference on science engineering and management research (ICSEMR), IEEE.

Tai, A. P. K., Mickley, L. J., & Jacob, D. J. (2010). Correlations between fine particulate matter (PM2.5) and meteorological variables in the United States: Implications for the sensitivity of PM2.5 to climate change. *Atmospheric Environment*, *44*(32), 3976–3984. https://doi.org/10.1016/j.atmosenv.2010.06.060

Tessum, C. W., Hill, J. D., & Marshall, J. D. (2017). InMAP: A model for air pollution interventions. *PLOS ONE*, *12*(4), e0176131. https://doi.org/10.1371/journal.pone.0176131

Tessum, C. W., Apte, J. S., Goodkind, A. L., Muller, N. Z., Mullins, K. A., Paolella, D. A., et al. (2019). Inequity in consumption of goods and services adds to racial–ethnic disparities in air pollution exposure. *Proceedings of the National Academy of Sciences*, *116*(13), 6001–6006. https://doi.org/10.1073/pnas.1818859116

Tessum, C. W., Paolella, D. A., Chambliss, S. E., Apte, J. S., Hill, J. D., & Marshall, J. D. (2021). $PM_{2.5}$ polluters disproportionately and systemically affect people of color in the United States. *Science Advances*, *7*(18), eabf4491. https://doi.org/10.1126/sciadv.abf4491

The CESM2 Development Team. (2019). CESM2.1/CAM-chem Instantaneous Output for Boundary Conditions [Data set]. UCAR/NCAR - Atmospheric Chemistry Observations and Modeling Laboratory. https://doi.org/10.5065/NMP7-EP60

Theodoridis, E., Mylonas, G., & Chatzigiannakis, I. (2013). Developing an iot smart city framework (pp. 1–6). Presented at the IISA 2013, IEEE.

Thompson, T. M., & Selin, N. E. (2012). Influence of air quality model resolution on uncertainty associated with health impacts. *Atmospheric Chemistry and Physics*, *12*(20), 9753–9762. https://doi.org/10.5194/acp-12-9753-2012

Thompson, T. M., Saari, R. K., & Selin, N. E. (2014). Air quality resolution for health impact assessment: influence of regional characteristics. *Atmospheric Chemistry and Physics*, *14*(2), 969–978. https://doi.org/10.5194/acp-14-969-2014

Thun, M. J., & Sinks, T. (2004). Understanding cancer clusters. *CA: A Cancer Journal for Clinicians*, *54*(5), 273–280.

Torres-Vazquez, A., Pleim, J., Gilliam, R., & Pouliot, G. (2022). Performance Evaluation of the Meteorology and Air Quality Conditions From Multiscale WRF-CMAQ Simulations for the Long

Island Sound Tropospheric Ozone Study (LISTOS). *Journal of Geophysical Research: Atmospheres*, *127*(5). https://doi.org/10.1029/2021JD035890

Touma, D., Michalak, A. M., Swain, D. L., & Diffenbaugh, N. S. (2018). Characterizing the Spatial Scales of Extreme Daily Precipitation in the United States. *Journal of Climate*, *31*(19), 8023–8037. https://doi.org/10.1175/JCLI-D-18-0019.1

Townsend, J. G. (1950). Investigation of the smog incident in Donora, Pa., and vicinity. *American Journal of Public Health and the Nations Health*, *40*(2), 183–189.

Tran, T., Tran, H., Mansfield, M., Lyman, S., & Crosman, E. (2018). Four dimensional data assimilation (FDDA) impacts on WRF performance in simulating inversion layer structure and distributions of CMAQ-simulated winter ozone concentrations in Uintah Basin. *Atmospheric Environment*, *177*, 75–92. https://doi.org/10.1016/j.atmosenv.2018.01.012

Travis, K. R., Jacob, D. J., Fisher, J. A., Kim, P. S., Marais, E. A., Zhu, L., et al. (2016). Why do models overestimate surface ozone in the Southeast United States? *Atmospheric Chemistry and Physics*, *16*(21), 13561–13577. https://doi.org/10.5194/acp-16-13561-2016

U.S. EPA. (2011, April). The Benefits and Costs of the Clean Air Act from 1990 to 2020. Retrieved from https://www.epa.gov/sites/default/files/2015-07/documents/fullreport_rev_a.pdf

Vohra, K., Vodonos, A., Schwartz, J., Marais, E. A., Sulprizio, M. P., & Mickley, L. J. (2021). Global mortality from outdoor fine particle pollution generated by fossil fuel combustion: Results from GEOS-Chem. *Environmental Research*, *195*, 110754. https://doi.org/10.1016/j.envres.2021.110754

Waddell, K. (2021, December 9). When Amazon Expands, These Communities Pay the Price. *Consumer Reports*. Retrieved from https://www.consumerreports.org/corporate-accountability/when-amazon-expands-these-communities-pay-the-price-a2554249208/

Wang, J., Xue, P., Pringle, W., Yang, Z., & Qian, Y. (2022). Impacts of Lake Surface Temperature on the Summer Climate Over the Great Lakes Region. *Journal of Geophysical Research: Atmospheres*, *127*(11). https://doi.org/10.1029/2021JD036231

Wang, K., Zhang, Y., Yu, S., Wong, D. C., Pleim, J., Mathur, R., et al. (2021). A comparative study of two-way and offline coupled WRF v3.4 and CMAQ v5.0.2 over the contiguous US: performance evaluation and impacts of chemistry–meteorology feedbacks on air quality. *Geoscientific Model Development*, *14*(11), 7189–7221. https://doi.org/10.5194/gmd-14-7189-2021

Wang, K., Tong, Y., Gao, J., Zhang, X., Zuo, P., Wang, C., et al. (2021). Pinpointing optimized air quality model performance over the Beijing-Tianjin-Hebei region: Mosaic approach. *Atmospheric Pollution Research*, *12*(11), 101207. https://doi.org/10.1016/j.apr.2021.101207

Wang, Y., Ma, Y.-F., Muñoz-Esparza, D., Dai, J., Li, C. W. Y., Lichtig, P., et al. (2023). Coupled mesoscale–microscale modeling of air quality in a polluted city using WRF-LES-Chem. *Atmospheric Chemistry and Physics*, *23*(10), 5905–5927. https://doi.org/10.5194/acp-23-5905-2023

Wang, Y., Apte, J. S., Hill, J. D., Ivey, C. E., Patterson, R. F., Robinson, A. L., et al. (2022). Location-specific strategies for eliminating US national racial-ethnic PM2.5 exposure inequality. *Proceedings of the National Academy of Sciences*, *119*(44), e2205548119. https://doi.org/10.1073/pnas.2205548119

Wilkins, E. T. (1954). Air pollution aspects of the London fog of December 1952. *Quarterly Journal of the Royal Meteorological Society*, *80*(344), 267–271. https://doi.org/10.1002/qj.49708034420

Wong, D. C., Pleim, J., Mathur, R., Binkowski, F., Otte, T., Gilliam, R., et al. (2011). *WRF-CMAQ two-way coupled system with aerosol feedback: software development and preliminary results* (preprint). Atmospheric Sciences. https://doi.org/10.5194/gmdd-4-2417-2011

Wong, D. C., Pleim, J., Mathur, R., Binkowski, F., Otte, T., Gilliam, R., et al. (2012). WRF-CMAQ two-way coupled system with aerosol feedback: software development and preliminary results. *Geoscientific Model Development*, *5*(2), 299–312. https://doi.org/10.5194/gmd-5-299-2012

Xiu, A., & Pleim, J. E. (2001). Development of a Land Surface Model. Part I: Application in a Mesoscale Meteorological Model. *Journal of Applied Meteorology*, *40*(2), 192–209. https://doi.org/10.1175/1520-0450(2001)040<0192:DOALSM>2.0.CO;2

Xu, H., & Chen, H. (2021). Impact of urban morphology on the spatial and temporal distribution of PM2.5 concentration: A numerical simulation with WRF/CMAQ model in Wuhan, China. *Journal of Environmental Management*, *290*, 112427. https://doi.org/10.1016/j.jenvman.2021.112427

Yang, H., Huang, X., Westervelt, D. M., Horowitz, L., & Peng, W. (2022). Socio-demographic factors shaping the future global health burden from air pollution. *Nature Sustainability*, *6*(1), 58–68. https://doi.org/10.1038/s41893-022-00976-8

Zhang, H., Chen, G., Hu, J., Chen, S.-H., Wiedinmyer, C., Kleeman, M., & Ying, Q. (2014). Evaluation of a seven-year air quality simulation using the Weather Research and Forecasting (WRF)/Community Multiscale Air Quality (CMAQ) models in the eastern United States. *Science of The Total Environment*, *473–474*, 275–285. https://doi.org/10.1016/j.scitotenv.2013.11.121

Zheng, X., Ding, H., Jiang, L., Chen, S., Zheng, J., Qiu, M., et al. (2015). Association between Air Pollutants and Asthma Emergency Room Visits and Hospital Admissions in Time Series Studies: A Systematic Review and Meta-Analysis. *PLOS ONE*, *10*(9), e0138146. https://doi.org/10.1371/journal.pone.0138146

Zhong, X., Ma, Z., Yao, Y., Xu, L., Wu, Y., & Wang, Z. (2023). WRF–ML v1.0: a bridge between WRF v4.3 and machine learning parameterizations and its application to atmospheric radiative transfer. *Geoscientific Model Development*, *16*(1), 199–209. https://doi.org/10.5194/gmd-16-199-2023

Zoogman, P., Liu, X., Suleiman, R. M., Pennington, W. F., Flittner, D. E., Al-Saadi, J. A., et al. (2017). Tropospheric emissions: Monitoring of pollution (TEMPO). *Journal of Quantitative Spectroscopy and Radiative Transfer*, *186*, 17–39. https://doi.org/10.1016/j.jqsrt.2016.05.008

Appendix 1

Supplement for Chapter 2

Supporting Information for:

Simulation of Neighborhood-Scale Air Quality with two-way coupled WRF-CMAQ over Southern Lake Michigan-Chicago Region

Contents of Appendix 1:

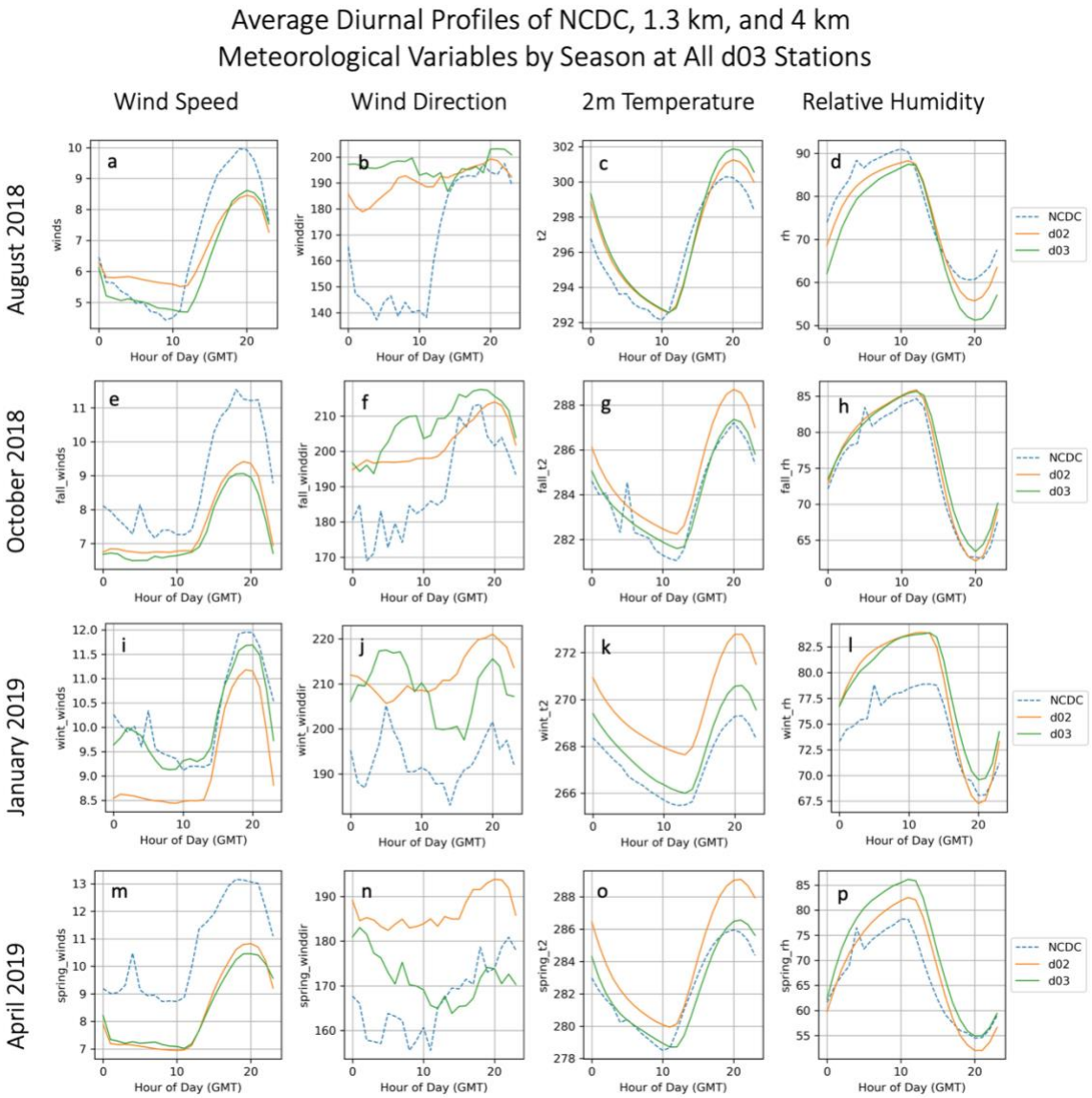


Figure S1.1. Diurnal profiles of meteorological variables averaged over all 10 NCDC stations or nearest pixels in the 1.3 km domain for the 1.3 km (d03, green), 4 km (d02, orange), and observations (NCDC, blue dash).

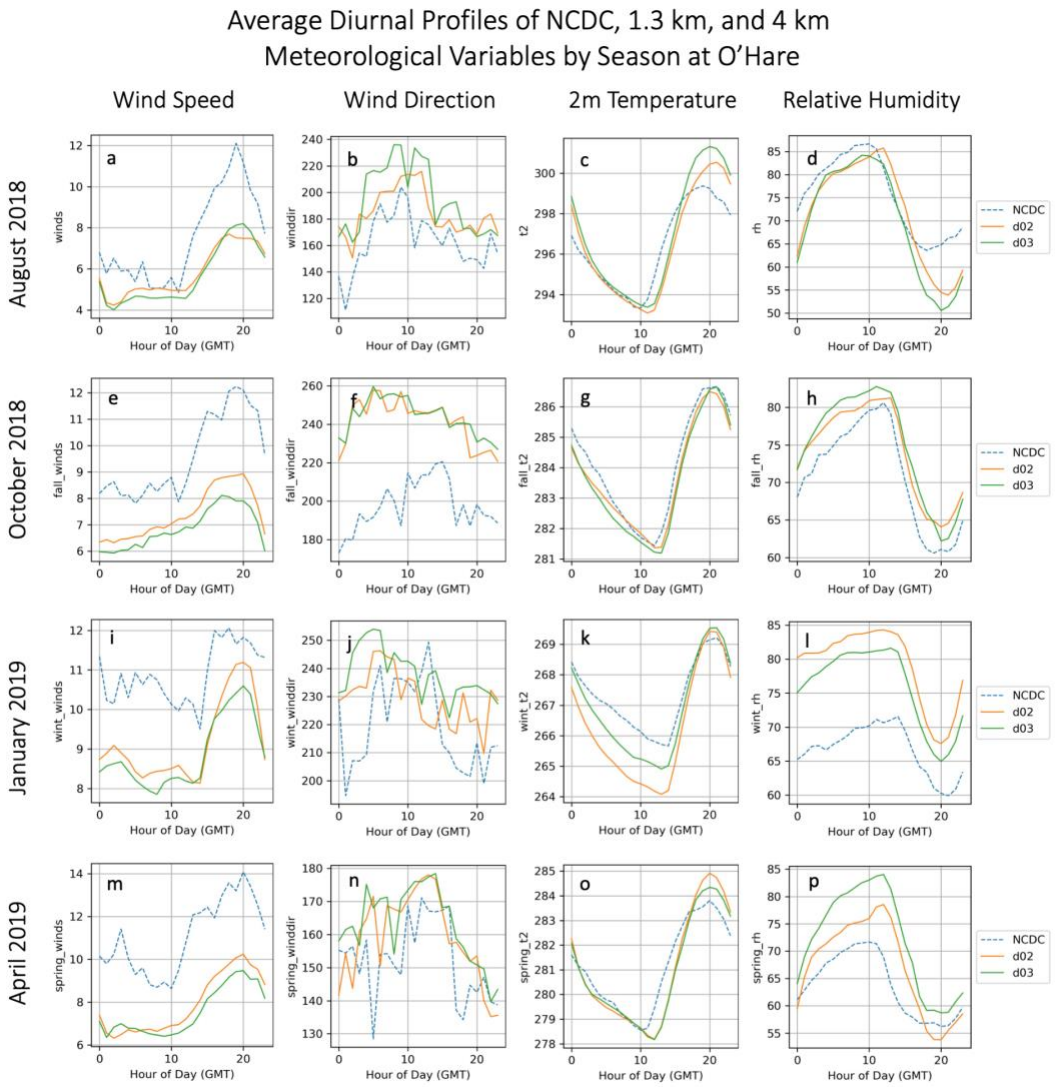


Figure S1.2. Diurnal profiles of meteorological variables at the O'Hare NCDC station or nearest pixel for the 1.3 km (d03, green), 4 km (d02, orange), and observations (NCDC, blue dash).

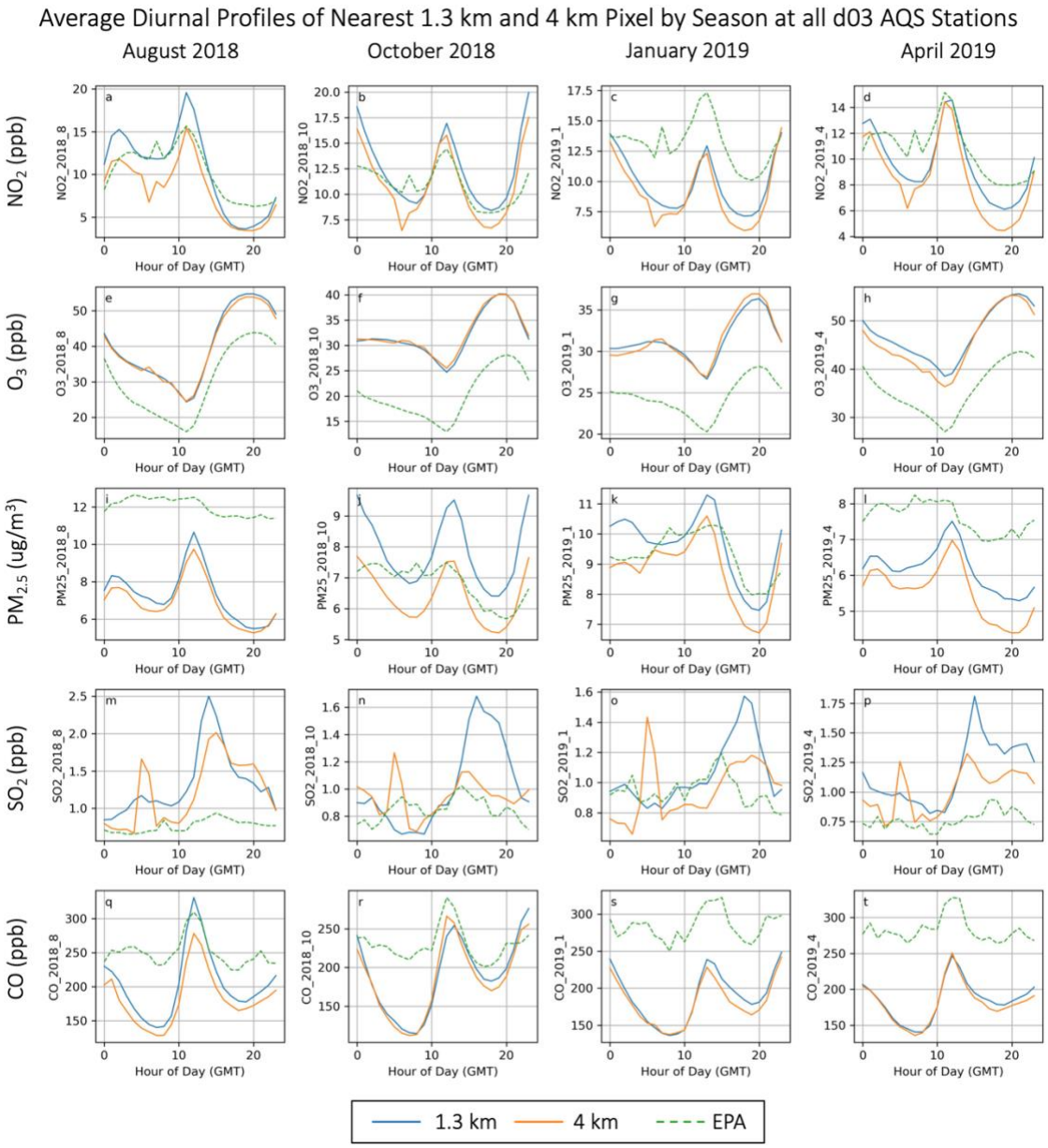


Figure S1.3. CMAQ (blue for 1.3 km and orange for 4 km) and EPA (green dash) average hourly estimates of pollutants to show the diurnal prediction of pollutants as compared to observations.

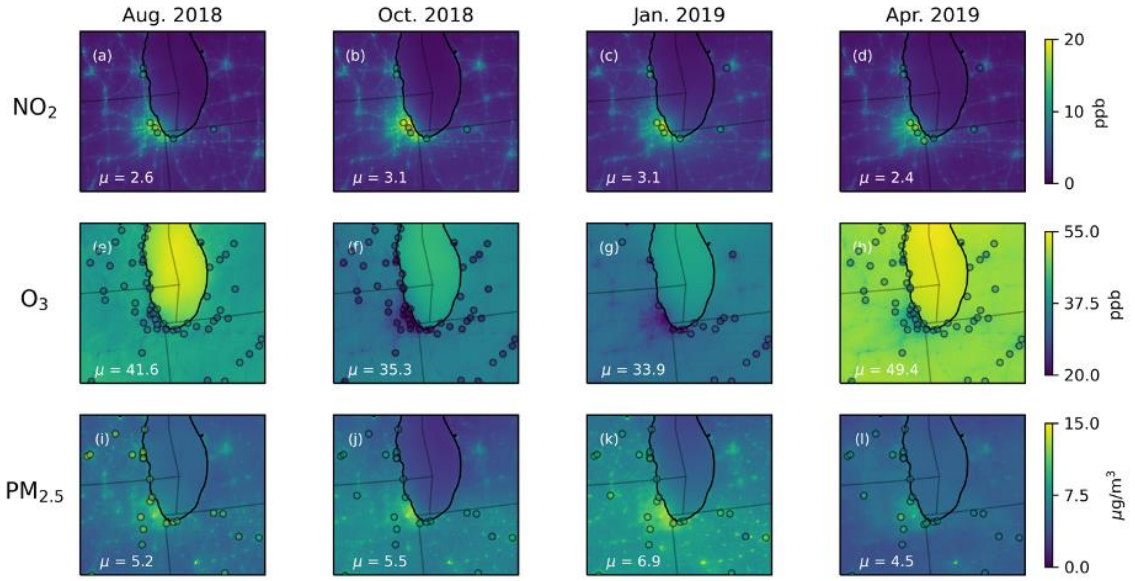


Figure S1.4. Monthly WRF-CMAQ simulation with EPA AQS observations overlaid.

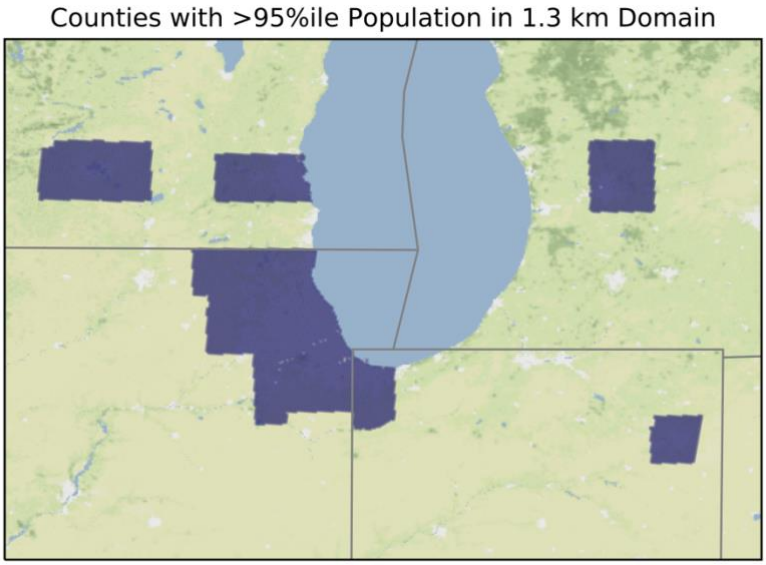


Figure S1.5. Shaded areas highlight urban zones in the urban vs. rural analysis. Urban counties were chosen by identifying areas with the highest (>95%ile) number of residents within the 4 states in the 1.3 km domain. Population was taken from the 2018 American Community Survey. The shaded area represents the area mask used to calculate the values for urban areas in Table 4, and land-based areas outside of the shade are considered rural. Counties defined as urban include: Cook County, IL, DuPage County, IL, Kane County, IL, Lake County, IL, McHenry County, IL, Will County, IL, Allen County, IN, Lake County, IN, Marion County, IN, Genesee County, MI, Kent County, MI, Macomb County, MI, Oakland County, MI, Washtenaw County, MI, Wayne County, MI, Dane County, WI, Milwaukee County, WI, Waukesha County, WI.

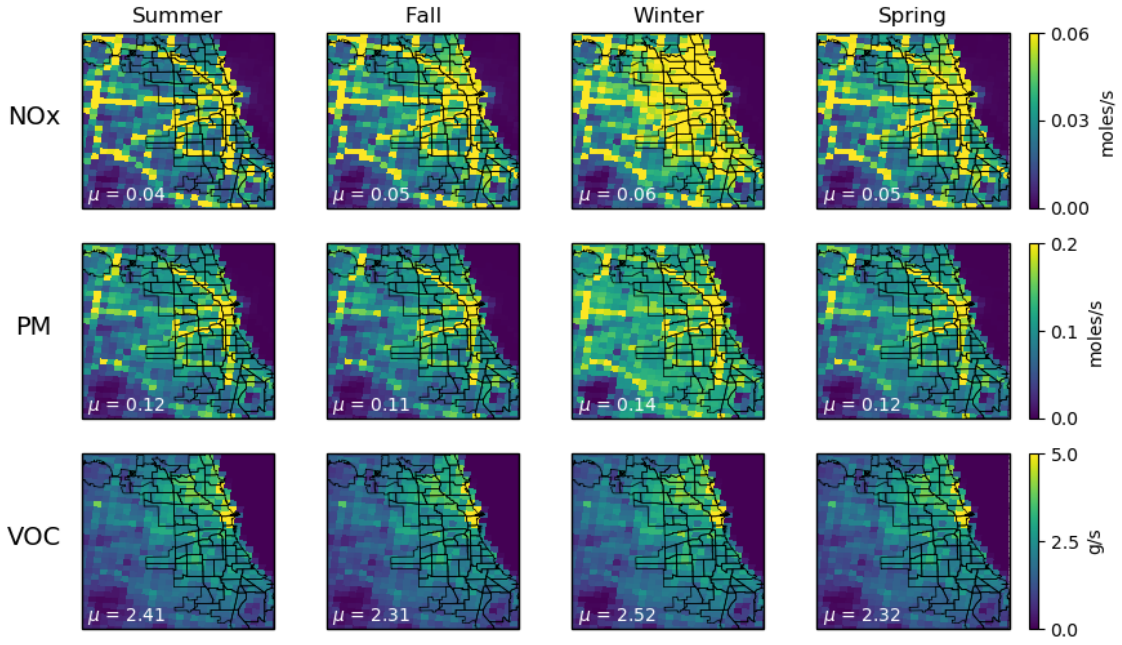


Figure S1.6. Seasonal emissions of NOx, PM (EC + OC), and VOCs over Chicago region, with Chicago neighborhoods outlined in black. Average emission rates (μ; in g/s or moles/s) within Chicago city limits are annotated on each panel.

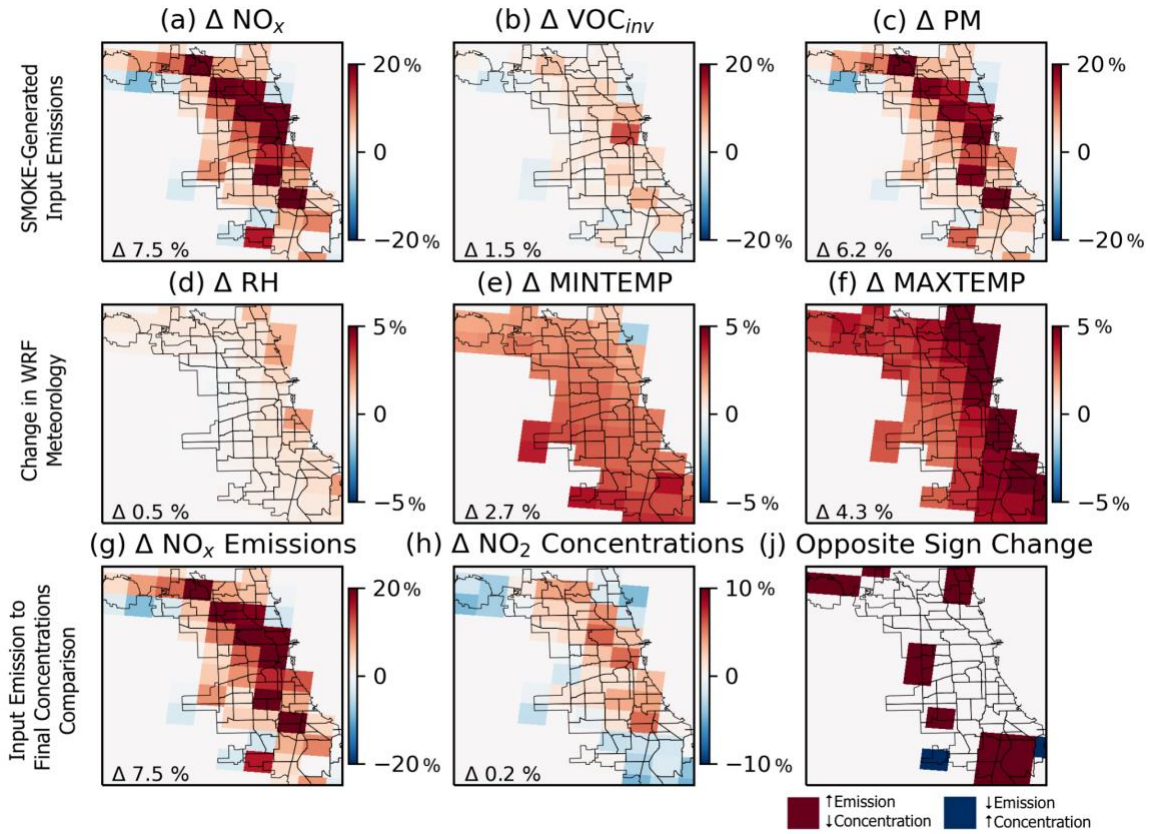


Figure S1.7. Difference (%) in annualized emissions of (a) NOx, (b) VOCs, (c) PM (EC+OC) and WRF meteorological variables (d-f) between 1.3 km and 4 km simulations. The 1.3 km emissions dataset is scaled up to the 4 km grid to create a cell-by-cell comparison. The differences in emissions are due to the finer-scale characterization of emission sources and meteorological processing, rather than differences in the spatial surrogates. As shown by example for NOx and NO2, the difference in input emissions (g) in the 1.3 km and 4 km datasets do not necessarily result in similar changes in concentrations (h) in the resulting WRF-CMAQ output, with grid cell mismatches highlighted in (j).

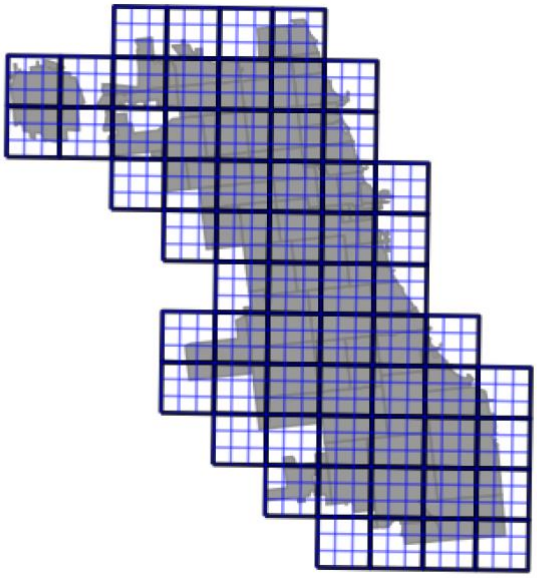


Figure S1.8. Mask of Chicago grid cells, with grid cells outlined over Chicago (gray), showing 4 km (black) and 1.3 km (blue) grid cells.

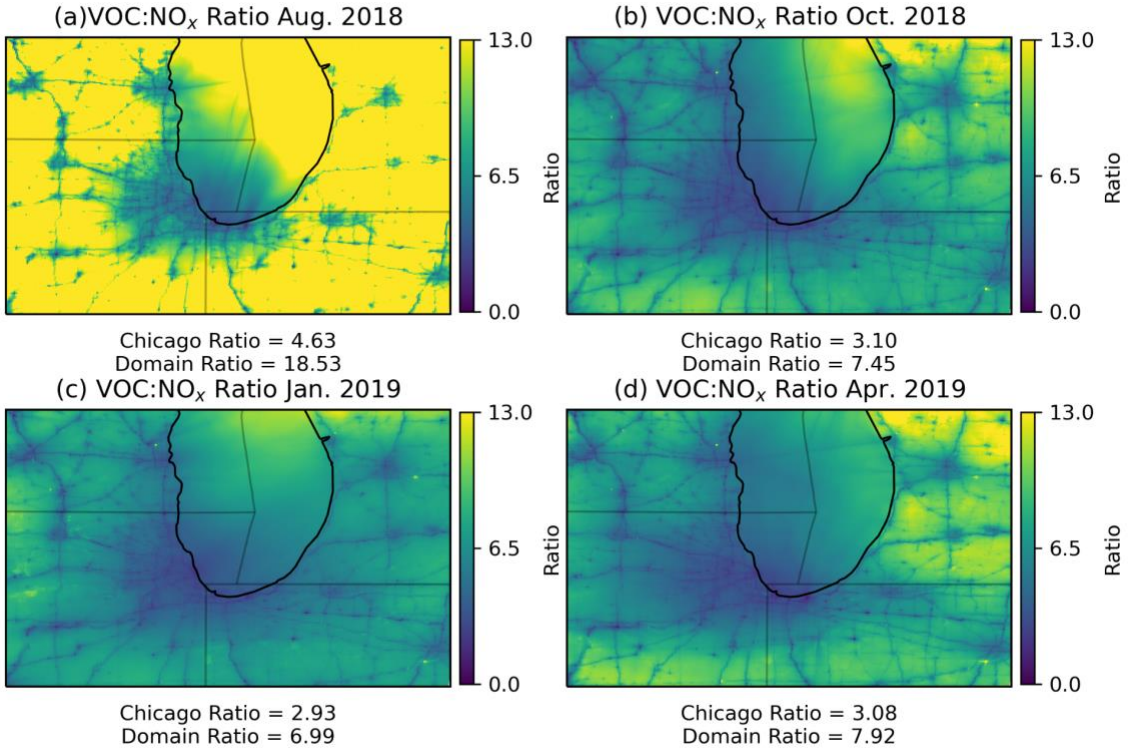


Figure S1.9. Daytime VOC:NOx surface ratios for each month of simulation.

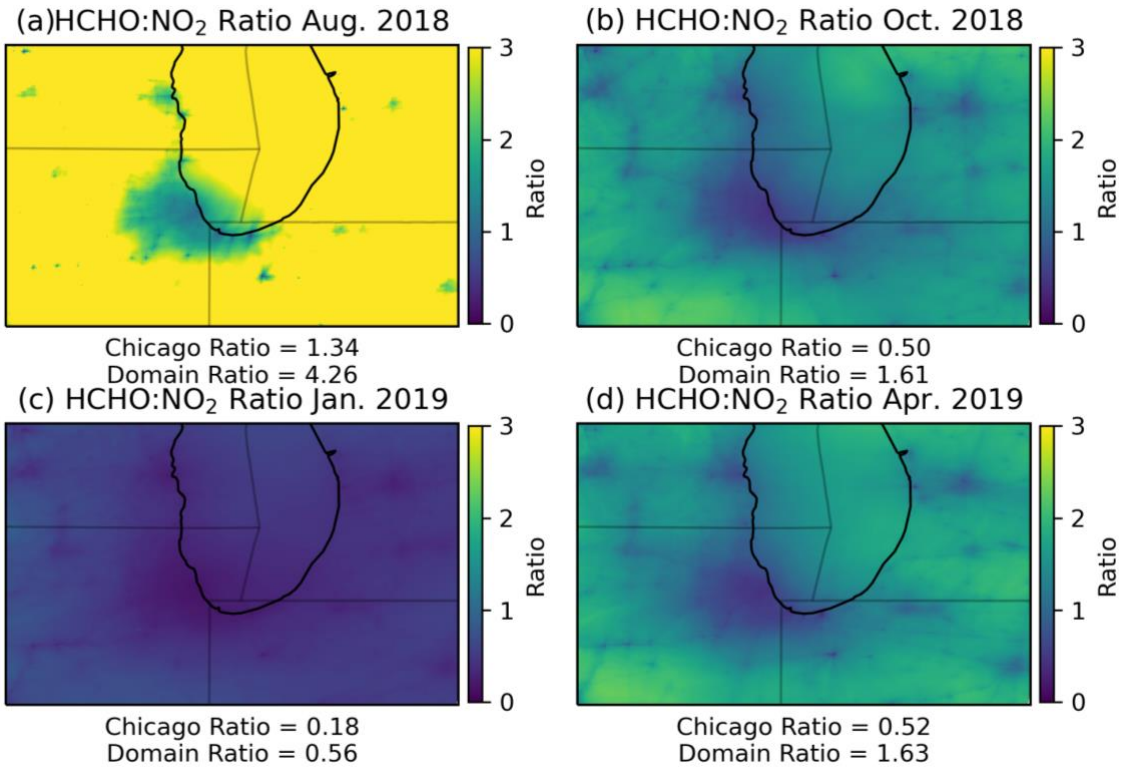


Figure S1.10. Daytime HCHO:NO_2 column ratios for each month of simulation.

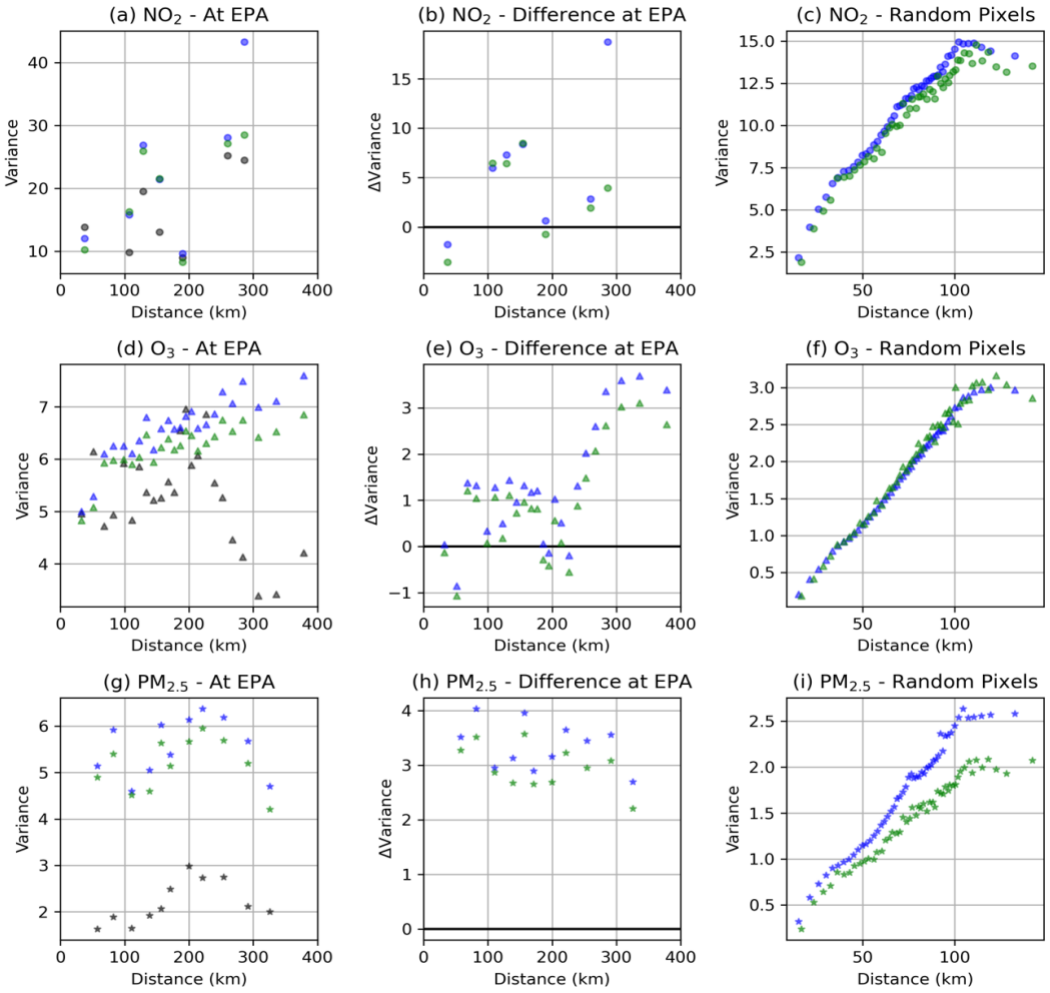


Figure S1.11. We compute the empirical variogram by using the methods-of-moments estimator (Marzban & Sandgathe, 2008)**.** We compute the variogram function at NO_2, O_3, and $PM_{2.5}$ EPA AQS observation sites (a,d,g) and the nearest grid-cell output at each model resolution. By subtracting the EPA AQS variogram from the model variogram (b,e,h), we identify which model resolution better captures the observed variance. We then create a synthetic variogram to expand the analysis beyond the limitations of the EPA observation sites (c,f,i). To compute the synthetic variogram, we randomly sample 75 grid cells over the 1.3 km domain and the overlapping 4 km grid cells. We compute the variogram for the surrounding pixels within a 150 km radius. This analysis does not provide verification, but instead provides information on the relative variance across model resolutions, and as such, the relative steepness of spatial gradients across the two resolutions. When one model resolution shows higher variogram values than the other at the same distance $\boldsymbol{y}$, we interpret the spatial gradients to be steeper in that model.

Table S1.1. Boundary condition chemical species mapping from CAM-CHEM chemical species to CMAQ chemical species.

CMAQ Species	CAM-Chem Species	Mozart2CMAQ Factors
NO	NO	1
NO2	NO2	1
O3	O3	1
H2O2	H2O2	1
N2O5	N2O5	1
HNO3	HNO3	1
PNA	HO2NO2	1
PAN	PAN	1
OPAN	MPAN	1
CO	CO	1
PAR	C3H6,BIGENE,BIGALK,HYAC,MEK	1.0, 1.0, 5.0, 2.0, 3.0
OLE	C3H6,BIGENE	1.0, 0.5
IOLE	BIGENE	0.5
FORM	CH2O	1
ALD2	CH3CHO	1
MGLY	CH3COCHO	1
ETHA	C2H6	1
ETH	C2H4	1
ETHY	C2H2	1
PRPA	C3H8	1
ACET	CH3COCH3	1
ETOH	C2H5OH	1
MEOH	CH3OH	1
MEPX	CH3OOH	1
FACD	HCOOH	1
AACD	CH3COOH	1
KET	HYAC,MEK	1.0, 1.0
ISOP	ISOP	1

CMAQ Species	CAM-Chem Species	Mozart2CMAQ Factors
ISPD	MACR,MVK	1.0, 1.0
TERP	MTERP	1
TOL	TOLUENE	1
XYL	XYLENES	1
BENZ	BENZENE	1
CRES	CRESOL,PHENOL	1.0, 1.0
SO2	SO2	1
NH3	NH3	1
MECN	CH3CN	1
HCN	HCN	1
TOLA	TOLUENE	1
XYLA	XYLENES	1
BNZA	BENZENE	1
ISP	ISOP	1
TRP	MTERP	1
CH4	CH4	1
DMS	DMS	1
GLY	GLYOXAL	1
ASO4J	so4_a1,so4_a2	0.30208, 0.30208
ASO4K	NH4	1.61111
ANO3J	soa3_a1,soa3_a2,soa4_a1,soa4_a2,soa5_a1,soa5_a2	0.22308 (for all)
ANO3K	soa1_a1,soa1_a2,soa2_a1,soa2_a2	0.22308 (for all)
ANH4J	pom_a1,pom_a4	0.13182, 0.13182
ATRP1J	bc_a1,bc_a4	0.29, 0.29
ATRP2J	dst_a1,dst_a2	0.29, 0.29
APNCOMJ	dst_a3	0.29
APOCJ	ncl_a1,ncl_a2	1.26087, 1.26087
AECJ	ncl_a1,ncl_a2	0.82857, 0.82857

Table S1.2. Definition of performance metrics used in this study.

Metrics	***Definition***
Mean (μ)	$\mu = \frac{1}{n}\sum_{j=1}^{n} P_j$
Mean Bias (*MB*)	$MB = \frac{1}{n}\sum_{1}^{n} (P_j - O_j)$
Mean Error (*ME*)	$ME = \frac{1}{n}\sum_{1}^{n} \lvert P_j - O_j \rvert$
Root Mean Squared Error (*RMSE*)	$RMSE = \sqrt{\sum_{1}^{n} \frac{(P_j - O_j)^2}{n}}$
Normalized Mean Bias (*NMB*)	$NMB = \frac{\sum_1^n (P_j - O_j)}{\sum_1^n O_j}$
Normalized Mean Error (*NME*)	$NME = \frac{\sum_1^n \lvert P_j - O_j \rvert}{\sum_1^n O_j}$
Correlation Coefficient (*r*)	$r = \frac{\sum [(P_j - \underline{P}) \times (O_j - \underline{O})]}{\sqrt{\sum (P_j - \underline{P})^2 \times \sum (O_j - \underline{O})^2}} \times 100$

Note: Subscript *j* represents the pairing of *N* observations (*O*) and predictions (*P*) by site and time.

Table S1.3. Emery et al. 2001 recommendations for WRF-CMAQ meteorological performance

	Unit	***MB***	***RMSE***	***ME***	***IOA***
T	°C	±0.5	n/a	±2	>0.7
WS	m/s	±0.5	2	n/a	>0.6
WD	°	±10	n/a	±30	n/a
RH	n/a	n/a	n/a	n/a	n/a

Table S1.4. Comparison of two-way coupled WRF-CMAQ simulated hourly meteorological variables with O'Hare Meteorological station (n = 1) for 1.3 km (d03) and 4 km (d02) grid cells. The average observed value is μd, while the predicted value is as μp.

		O'Hare	1.3 km					4 km				
	Month	μ_d	μ_p	*MB*	*ME*	*RMSE*	*r*	μ_p	*MB*	*ME*	*RMSE*	*r*
T2 (°C)	08/2018	23.4	23.9	0.5	1.9	2.5*	0.8	296.5	0.1	1.7	2.2*	0.8
	10/2018	11	10.5	-0.5	1.7	2.1*	0.9	283.7	-0.3	1.7	2.1*	0.9
	01/2019	-5.7	-6.1	-0.4	1.5	1.9	1.0	266.3	-1.1*	1.7	2.2*	1.0
	04/2019	8.1	7.9	-0.2	1.8	2.5*	0.9	281.0	-0.1	1.8	2.5*	0.9
	Average	9.2	9	-0.2	1.7	2.3	0.9	281.9	-0.1	1.7	2.3	0.9
RH (%)	08/2018	75.3	69.8	-5.5	11.7	15.8	0.7	71.7	-3.6	11.0	14.9	0.7
	10/2018	71.0	74.8	3.8	12.5	15.5	0.7	74.1	3.1	13.1	16.1	0.7
	01/2019	66.7	75.9	9.3	12.2	13.9	0.7	79.1	12.4	14.4	16.2	0.7
	04/2019	63.6	72.0	8.4	13.4	16.9	0.8	67.1	3.4	12.9	16.0	0.8
	Average	69.1	73.1	4.0	12.5	15.5	0.7	73.0	3.8	12.9	15.8	0.7
WS (m/s)	08/2018	7.5	5.7	-1.9*	3.2	4.1*	0.6	5.8	-1.7*	3.2	4.2*	0.5
	10/2018	9.6	6.9	-2.7*	3.7	4.7*	0.6	7.4	-2.2*	3.3	4.2*	0.7
	01/2019	10.9	8.9	-2.0*	3.0	4.1*	0.8	9.2	-1.7*	2.9	3.9*	0.8
	04/2019	11.1	7.5	-3.5*	4.1	5.1*	0.8	7.9	-3.2*	3.9	4.9*	0.8
	Average	9.7	7.2	-2.5	3.5	4.5	0.6	7.6	-2.2	3.3	4.3	0.7
WD (°)	08/2018	162.8	194.5	31.7*	78.9	128.1*	0.3	184.8	22.0*	74.3	122.4*	0.3
	10/2018	197.5	244.0	46.4*	61.1	111.9*	0.3	241.4	43.9*	58.2	107.8*	0.4
	01/2019	219.4	238.2	18.7*	33.2	76.8*	0.7	229.6	10.2	34.5	77.3*	0.7
	04/2019	152.4	163.2	10.7	45.4	88.8*	0.7	159.1	6.6	45.8	88.2*	0.6
	Average	183.0	209.9	26.9	54.7	101.4	0.5	203.7	8.4	53.2	98.9	0.5

Table S1.5. Studies for the WRF-CMAQ comparison and reported metrics of performance. Values are pulled from tables, text, or figures in the published works. Star (*) denotes daytime O_3 analysis.

Study	Pub. Year	Area	Length	Res.	Var	*MB*	*RMSE*	*NMB (%)*	*NME (%)*	*r*
Zhang et al., 2014	2014	Eastern US, Chicago	7-yr	12, 4 km	O_3	-	-	<15	<35	-
Abdi-Oskouei et al., 2020*	2020	Great Lakes Region	Summer	4 km	O_3	-4.4	12.3 – 14.1	-8 – -4.4	-	-
Qin et al, 2019	2019	Great Lakes Region	1 mo., July	12, 4 km	MDA O_3	1.1	10.6	2.0	15.1	-
Odman et al., 2019*	2019	Midwest	July	12, 4 km	MDA O_3	6	-	-	-	-
Torres-Vazquez et al., 2021	2021	NYC-Long Island	Summer	12, 4, 1.3 km	O_3	0.2	11.2	-	-	0.7 – 0.8
					NO_x	6.8	21	-	-	0.5 – 0.6
					$PM_{2.5}$	3.7	8.8	-	-	0.5 – 0.6
Lawal et al., 2022	2022	Atlanta, GA	1 mo., Summer	24, 4 km	NO_x	-	1.5	45	-	0.9
Hogrefe et al., 2015	2015	CONUS	1-yr	12 km	$PM_{2.5}$	-	-	1.4 – 59.8	34.1 – 81.4	0.45 – 0.69
Pan et al., 2017	2017	Houston, TX	2-weeks	4 km, 1 km	NOx	-0.6 – 0.1	-	-	-	0.6 – 0.7
					O_3	10 – 11.4	-	-	-	0.8 – 0.9
Campbell et al., 2018	2018	Eastern US, CONUS	Summer	12 km	O_3	-	-	53 – 60	49 – 53	0.7
					SO_2	-	-	59 – 62	69 – 72	0.3
Liu et al., 2010	2010	North Carolina	2 mo., Summer Winter	12, 4 km	$PM_{2.5}$	-	-	-48.4 – 34.9	34.9 – 49.2	-
Harkey et. al, 2015	2015	CONUS	Winter, Summer	36 km	NO_2	-	-	-	-	0.12 – 0.24
Bickford et al., 2014	2013	Midwest	2 mo., Summer & Winter	12 km	NO_2	-	-	-25.5 – -3.1	52.3 – 44.8	0.5
					SO_2	-	-	-48.8	65-80	0.3 – 0.4
					$PM_{2.5}$	-	-	-39.3 – -34.3	43.4 – 46.1	0.6
					O_3	-	-	7.4	18	0.7

Table S1.6. O_3 performance in the 1.3 km domain (d03) assessed against EPA AQS station observations using different O_3 thresholds. The average observed value is noted as μ_d, while the predicted value is noted as μ_p.

O_3 Filter	**Month**	***μ_d***	***μ_p***	***NMB%***	***NME%***	***r***
1-hr O_3 > 60 ppb	**Aug. 2018**	67.7	64.1	-5.4	16.8	0.2
	Apr. 2019	63.0	58.4	-7.2	8.9	0.1
8-hr MDAO_3 > 60 ppb	**Aug. 2018**	65.9	66.7	1.2	12.9	0.2
	Apr. 2019*	61.0	59.5	-2.5	2.7	-
1-hr O_3 > (50%ile Observed O_3)	**Aug. 2018**	35.9	42.6	18.7	21.4	0.2
	Oct. 2019	24.5	34.5	40.9	42.9	0.4
	Jan. 2019	29.7	35.3	19.0	21.6	0.2
	Apr. 2019	40.8	49.8	22.1	23.1	0.4
	Average	32.7	40.6	25.2	27.3	0.3
8-hr MDAO_3	**Aug. 2018**	42.8	53.6	25.1	26.8	0.5
	Oct. 2019	28.0	39.0	39.3	40.8	0.4
	Jan. 2019	29.7	37.3	25.4	27.0	0.6
	Apr. 2019	44.2	55.2	24.8	25.6	0.4
	Average	36.2	46.3	28.7	30.1	0.5

Table S1.7. Differences in urban–rural concentrations of pollutants over the 1.3 km domain (d03). Differences are described using absolute difference (Δ) or percent (%). Visual depiction of urban vs. rural is provided in Figure S2.

		Aug. 2018		**Oct. 2018**		**Jan. 2019**		**April 2019**		**Average**	
	units	Δ	%	Δ	%	Δ	%	Δ	%	Δ	%
NO_2	ppb	3.3	60.2	4.2	61.7	3.4	54.9	3.3	62.5	3.5	59.8
O_3	ppb	-2.5	6.4	-4.1	-13.1	-3.7	-12.2	-2.8	-5.9	-3.3	-9.4
$PM_{2.5}$	ug/m^3	1.5	23.1	1.6	22.8	1.9	22.1	1.1	20.4	1.5	22.1
SO_2	ppb	0.2	31.9	0.1	23.8	0.1	22.4	0.1	36.5	0.1	28.6
CO	ppb	42.3	25.9	46.5	31.6	35.2	22.1	34.9	22.5	39.7	25.5

Table S1.8 Comparison of two-way coupled WRF-CMAQ simulated hourly meteorological variables with NCDC observations for 4 km (d02) simulations using only NCDC stations within the 1.3 km domain. The average predicted value is noted as μ_p, while the average observed value is noted as μ_d. *Denotes model performance within Emery et al., 2001 guidelines.

Var	Month	μ_d	μ_p	*MB*	*ME*	*RMSE*	*r*
T2 (°C)	08/2018	23.2	23.4	0.3	1.7	2.2	0.9
	10/2018	10.8	11.0	0.2	1.7	2.1	0.9
	01/2019	-5.8	-4.9	0.9*	2.1*	2.7	1.0
	04/2019	9.2	9.3	0.1	1.9	2.5	0.9
	Average	9.4	9.7	0.4	1.9	2.4	0.9
RH (%)	08/2018	76.6	73.8	-2.8	10.4	13.7	0.7
	10/2018	74.8	78.3	3.5	12.0	15.2	0.7
	01/2019	74.6	81.2	6.5	10.6	13.0	0.7
	04/2019	66.1	66.5	0.5	13.2	16.6	0.8
	Average	73.0	75.0	1.9	11.6	14.6	0.7
WS (m/s)	08/2018	6.9	6.0	-0.9*	3.0	3.8*	0.6
	10/2018	8.8	7.5	-1.3*	2.9	3.8*	0.7
	01/2019	10.2	10.4	0.2	3.6	5.0*	0.6
	04/2019	10.6	8.5	-2.1*	3.5	4.5*	0.7
	Average	9.1	8.1	-1.0	3.3	4.3	0.7
WD (°)	08/2018	166.0	191.4	25.4*	75.7*	123.4	0.3
	10/2018	190.6	204.4	13.8*	50.9*	97.6	0.5
	01/2019	192.8	208.5	15.7*	43.2*	91.7	0.6
	04/2019	166.8	171.8	5.0	50.2*	97.3	0.6
	Average	179.1	194.0	15.0	55.0	102.5	0.5

Table S1.9. Comparison of two-way coupled WRF-CMAQ simulated hourly meteorological variables with O'Hare NCDC station for 1.3 km (d03) and 4 km (d02) grid cells. The average observed value is noted as μ_d, while the predicted value is noted as μ_p. *Denotes model performance within Emery et al., 2001 guidelines.

		O'Hare	**1.3 km Domain**					**4 km Domain**				
	Month	μ_d	μ_p	*MB*	*ME*	*RMSE*	*r*	μ_p	*MB*	*ME*	*RMSE*	*r*
T2 (°C)	08/18	296.4	296.9	0.5	1.9	2.5*	0.8	296.5	0.1	1.7	2.2*	0.8
	10/18	284.0	283.5	-0.5	1.7	2.1*	0.9	283.7	-0.3	1.7	2.1*	0.9
	01/19	267.3	266.9	-0.4	1.5	1.9	1.0	266.3	-1.1*	1.7	2.2*	1.0
	04/19	281.1	280.9	-0.2	1.8	2.5*	0.9	281.0	-0.1	1.8	2.5*	0.9
	Avg.	282.2	282.0	-0.2	1.7	2.3	0.9	281.9	-0.1	1.7	2.3	0.9
RH (%)	08/18	75.3	69.8	-5.5	11.7	15.8	0.7	71.7	-3.6	11.0	14.9	0.7
	10/18	71.0	74.8	3.8	12.5	15.5	0.7	74.1	3.1	13.1	16.1	0.7
	01/19	66.7	75.9	9.3	12.2	13.9	0.7	79.1	12.4	14.4	16.2	0.7
	04/19	63.6	72	8.4	13.4	16.9	0.8	67.1	3.4	12.9	16.0	0.8
	Avg.	69.1	73.1	4.0	12.5	15.5	0.7	73.0	3.8	12.9	15.8	0.7
WS (m/s)	08/18	7.5	5.7	-1.9*	3.2	4.1*	0.6	5.8	-1.7*	3.2	4.2*	0.5
	10/18	9.6	6.9	-2.7*	3.7	4.7*	0.6	7.4	-2.2*	3.3	4.2*	0.7
	01/19	10.9	8.9	-2.0*	3.0	4.1*	0.8	9.2	-1.7*	2.9	3.9*	0.8
	04/19	11.1	7.5	-3.5*	4.1	5.1*	0.8	7.9	-3.2*	3.9	4.9*	0.8
	Avg.	9.7	7.2	-2.5	3.5	4.5	0.6	7.6	-2.2	3.3	4.3	0.7
WD (°)	08/18	162.8	194.5	31.7*	78.9	128.1*	0.3	184.8	22.0*	74.3	122.4*	0.3
	10/18	197.5	244.0	46.4*	61.1	111.9*	0.3	241.4	43.9*	58.2	107.8*	0.4
	01/19	219.4	238.2	18.7*	33.2	76.8*	0.7	229.6	10.2	34.5	77.3*	0.7
	04/19	152.4	163.2	10.7	45.4	88.8*	0.7	159.1	6.6	45.8	88.2*	0.6
	Avg.	183.0	209.9	26.9	54.7	101.4	0.5	203.7	8.4	53.2	98.9	0.5

Table S1.10. Performance of 4 km simulation over the 1.3 km (d03) domain, compared to the 1.3 km simulation performance using EPA AQS station data as the observational standard. Positive values in the Δ columns indicate that the 1.3 km simulation has higher fidelity to observations than the 4 km simulation. Values in the Δμ and Δr column are the subtracted differences.

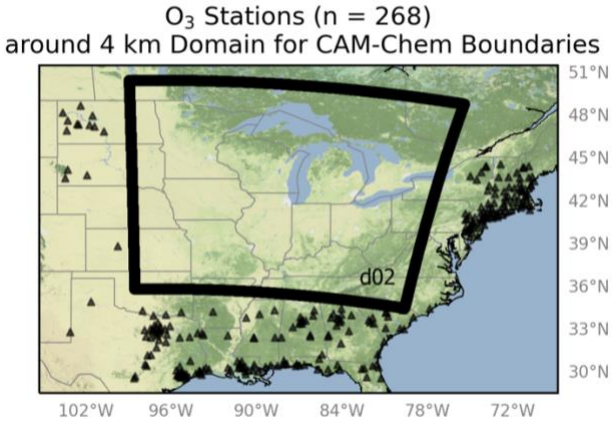
		4 km				**Δ (1.3 km - 4 km)**			
Var	Season	μ_d	*NMB*%	*NME*%	r	$\Delta\mu_{(1.3-4)}$	Δ*NMB*%	Δ*NME*%	Δr
SO_2	08/18	0.7	-1.9	89.6	0.1	0.1	-6.7	-2.8	0.07
	10/18	0.7	43.5	124.2	0.2	0.1	-24.8	-17.4	-0.07
	01/19	1.2	30.3	105.0	0.3	0.1	-10.7	5.1	-0.08
	04/19	1.3	71.7	145.5	0.2	0.0	-0.4	16.1	-0.09
NO_2	08/18	11.3	10.9	54.1	0.6	1.0	-8.9	1.0	0.00
	10/18	11.3	5.3	56.7	0.5	0.2	-1.4	7.6	-0.02
	01/19	10.7	-18.2	43.5	0.6	1.2	-8.8	1.9	0.02
	04/19	10.1	-4.5	53.1	0.6	0.4	-4.0	-1.3	0.02
O_3	08/18	39.7	31.2	39.0	0.7	-0.5	1.6	1.0	0.00
	10/18	31.9	56.8	61.4	0.6	-0.2	0.8	1.1	0.00
	01/19	30.9	25.8	33.0	0.7	-0.5	2.0	1.0	0.00
	04/19	46.9	29.0	33.5	0.6	-0.4	1.3	1.0	0.00
CO	08/18	204.4	-18.4	43.0	0.2	-0.1	0.0	0.2	0.02
	10/18	178.9	-22.0	44.5	0.3	-9.8	4.3	2.4	-0.02
	01/19	188.3	-33.6	40.0	0.5	1.5	-0.5	0.5	0.01
	04/19	180.5	-36.2	45.4	0.4	-3.0	1.5	-0.8	0.03
$PM_{2.5}$	08/18	7.7	-35.8	53.8	0.3	0.5	-3.5	1.1	0.00
	10/18	7.7	14.3	64.6	0.4	-0.1	1.6	2.3	-0.02
	01/19	9.8	4.4	51.7	0.5	0.2	-1.6	-1.2	0.01
	04/19	6.1	-19.1	52.6	0.5	0.1	-0.3	1.4	-0.01

Table S1.11. Comparison of early daytime (8 am – 12 pm) and late night (10 pm – 4 am) concentrations of key pollutants with the normalized mean bias.

		Day Average			**Night Average**			**Day NMB**		**Night NMB**	
	Month	EPA	4km	1.3km	EPA	4km	1.3km	4km	1.3km	4km	1.3km
NO_2	8/18	7.9	5.2	6.2	12.7	9.5	12.8	-34%	-21%	-25%	0%
	10/18	9	8.1	9.9	10.9	9.3	10.3	-10%	9%	-15%	-6%
	1/19	12.4	7.3	8.4	13.1	7.6	8.4	-41%	-32%	-42%	-36%
	4/19	9.2	6	7.8	11.5	8.4	9.1	-35%	-15%	-27%	-20%
O_3	8/18	36.6	46.7	47.7	20.7	31.4	31.5	28%	30%	52%	53%
	10/18	22.8	35.2	34.6	16.8	30.1	29.9	54%	52%	79%	79%
	1/19	24.6	33.1	32.3	23.6	30.5	30.7	35%	31%	29%	30%
	4/19	37.7	49.2	49.3	31.7	40.7	43.1	30%	31%	28%	36%
$PM_{2.5}$	8/18	11.6	6.4	6.9	12.5	6.8	7.2	-44%	-41%	-46%	-42%
	10/18	6.3	5.9	7.3	7.2	6	7.2	-7%	16%	-16%	0%
	1/19	9.5	8.2	9.2	9.7	9.2	9.8	-13%	-3%	-5%	1%
	4/19	7.1	5	5.8	8	5.7	6.3	-29%	-18%	-28%	-21%
SO_2	8/18	0.9	1.8	1.9	0.7	1	1.1	105%	119%	42%	54%
	10/18	0.9	1.1	1.5	0.9	0.9	0.7	13%	61%	1%	-18%
	1/19	1	1.1	1.3	0.9	1	0.9	3%	26%	6%	-2%
	4/19	0.9	1.2	1.5	0.7	0.9	0.9	38%	79%	24%	29%
CO	8/18	237.5	188.6	208.2	246.2	142.6	158.6	-21%	-12%	-42%	-36%
	10/18	215.9	194	204.1	219.3	126.9	128.8	-10%	-5%	-42%	-41%
	1/19	292.5	187.9	204.8	271.5	146.5	146.7	-36%	-30%	-46%	-46%
	4/19	273.1	183	190.9	278.3	150.1	151.5	-33%	-30%	-46%	-46%

Table S1.12. Performance of surface O_3 in the CAM-Chem simulation compared to EPA AQS stations near the edge of the 4 km (d02) domain, shown in the embedded figure.

Locations of Stations:



Temporal Average	Month	Year	EPA	CAM	NMB %	NME %	MB	RMSE	r
6-Hourly	8	2018	27.0	33.1	22.4	47.1	6.1	15.4	0.2
	10	2018	22.2	33.2	49.2	56.3	10.9	15.1	0.4
	1	2019	24.5	36.7	49.8	53.5	12.2	16.2	0.0
	4	2019	35.3	35.2	-0.1	26.3	-0.1	11.8	0.2
Daily	8	2018	26.8	33.1	23.3	23.5	6.3	7.0	0.0
	10	2018	22.0	33.2	50.7	50.7	11.2	11.4	0.3
	1	2019	24.3	36.7	50.8	50.8	12.4	13.2	-0.4
	4	2019	35.2	35.2	0.0	7.1	0.0	3.2	0.0
Daily Max (Daytime)	8	2018	77.5	36.4	-53.0	53.0	-41.1	42.8	0.2
	10	2018	56.2	36.5	-35.1	35.1	-19.7	21.7	-0.2
	1	2019	47.1	46.7	-0.7	10.6	-0.3	6.4	0.0
	4	2019	67.4	39.2	-41.9	41.9	-28.2	29.3	0.2
Daily Min (Night time)	8	2018	0.5	30.8	6491	6491	30.3	30.3	1.0
	10	2018	0.9	30.2	3296	3296	29.3	29.3	0.7
	1	2019	1.8	33.8	1784	1784	32.0	32.1	0.3
	4	2019	1.2	33.5	2655	2655	32.3	32.3	0.1
Monthly	8	2018	26.9	33.1	23.0	28.1	6.2	9.1	-0.4
	10	2018	21.7	33.0	52.2	53.1	11.3	12.3	0.3
	1	2019	24.3	36.7	50.8	50.8	12.4	13.0	0.2
	4	2019	35.2	35.2	0.1	7.7	0.0	3.4	0.2

Appendix 2

Supplement for Chapter 3

Supporting Information for

Intraurban NO_2 Hotspot Detection across Multiple Air Quality Products

Contents of Appendix 2:

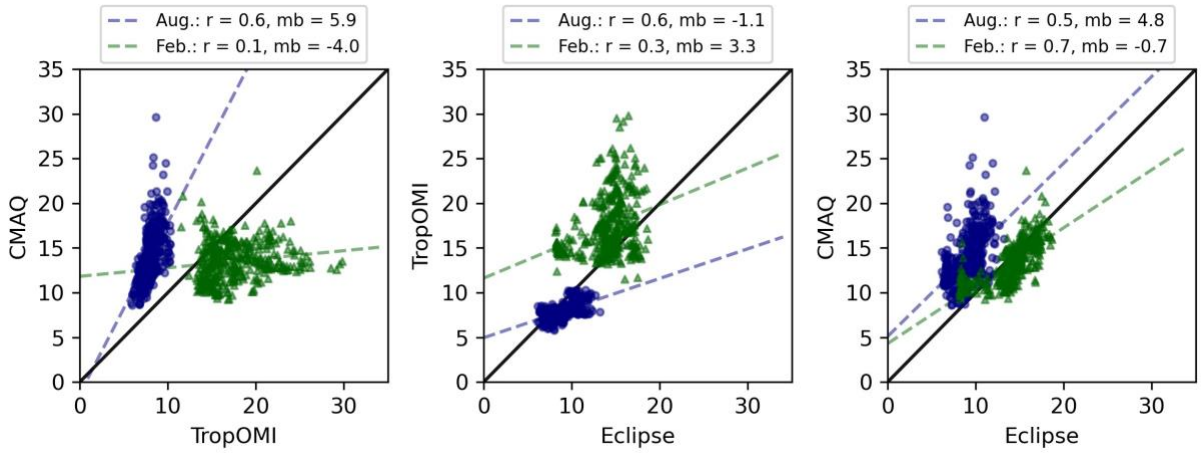


Fig. S2.1. Grid cell to grid cell comparison of NO_2 products from (a) CMAQ and TropOMI, (b) Eclipse and TropOMI, and (c) CMAQ and Eclipse. Monthly August data are shown with blue circles and February with green triangles. The best-fit linear regression is shown as a dashed line (blue for August, green for February), while the 1:1 line is solid black. Note that the 1:1 line is not necessarily the expected relationship between the datasets, particularly for ground-based (Eclipse + CMAQ) to columnar (TropOMI) comparisons. For each month, the Pearson correlation coefficient (r) and mean bias (mb) between datasets are annotated.

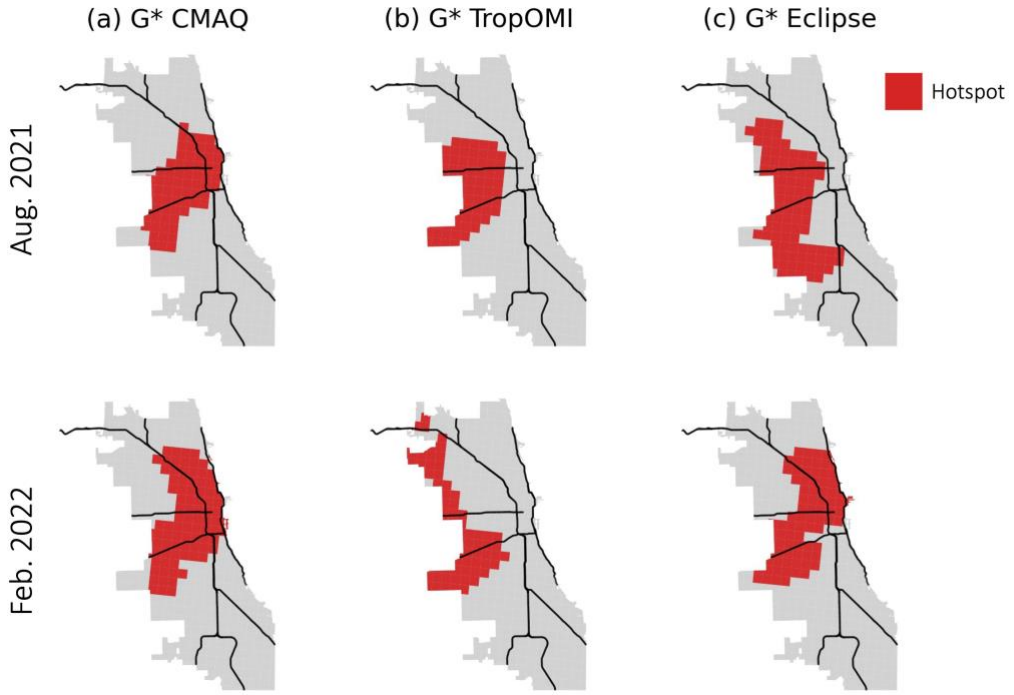


Fig. S2.2. Individual NO_2 hotspot results for (a) WRF-CMAQ, (b) TropOMI, and (c) Eclipse over Chicago for August 2021 and February 2022.

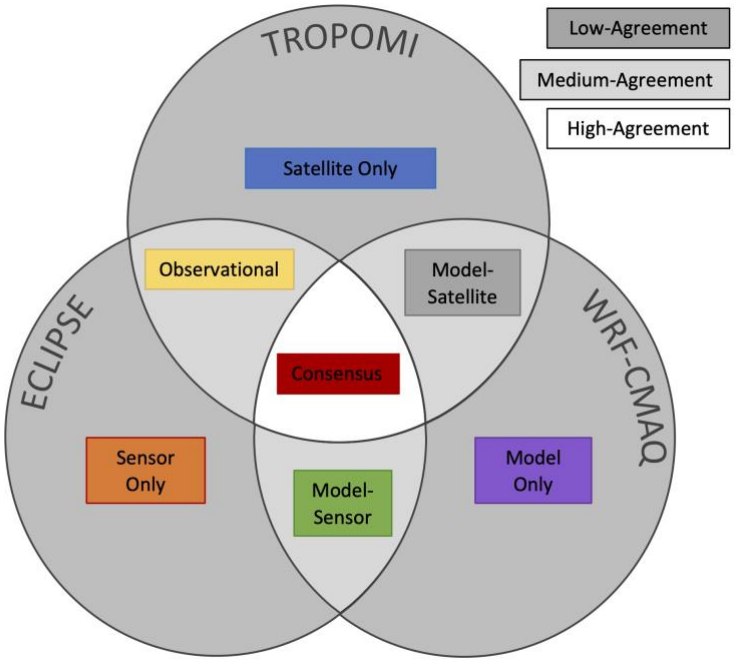


Fig. S2.3. Visual of hotspot intersection definitions for high-agreement hotspot, medium-agreement hotspots (2 out of 3 datasets), and low-agreement hotspots (only 1 dataset identifies hotspots). This follows the color scheme in Figure 1.

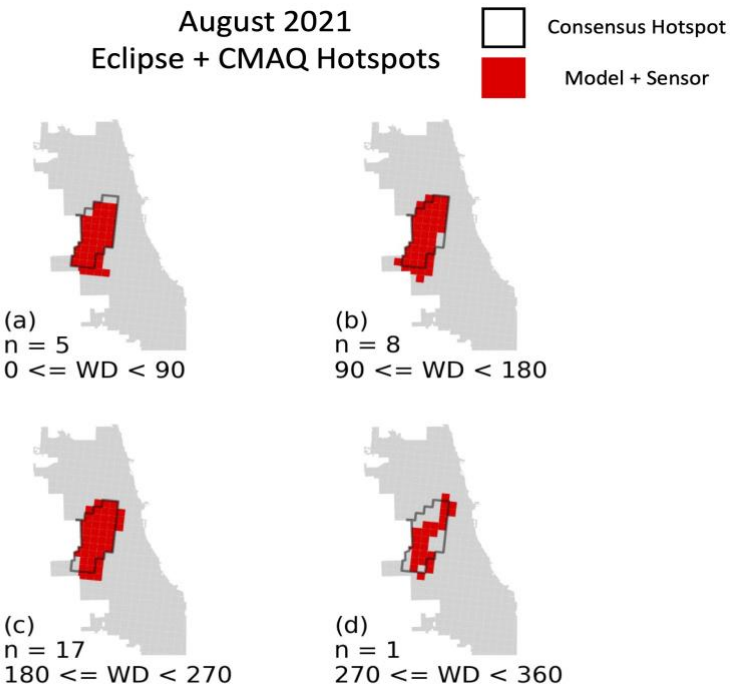


Fig. S2.4. The intersection of WRF-CMAQ and Eclipse hotspots over Chicago given each dominant wind directions for August 2021. The number of days with each dominant wind direction are noted. The extent of the consensus hotspot is outlined in black.

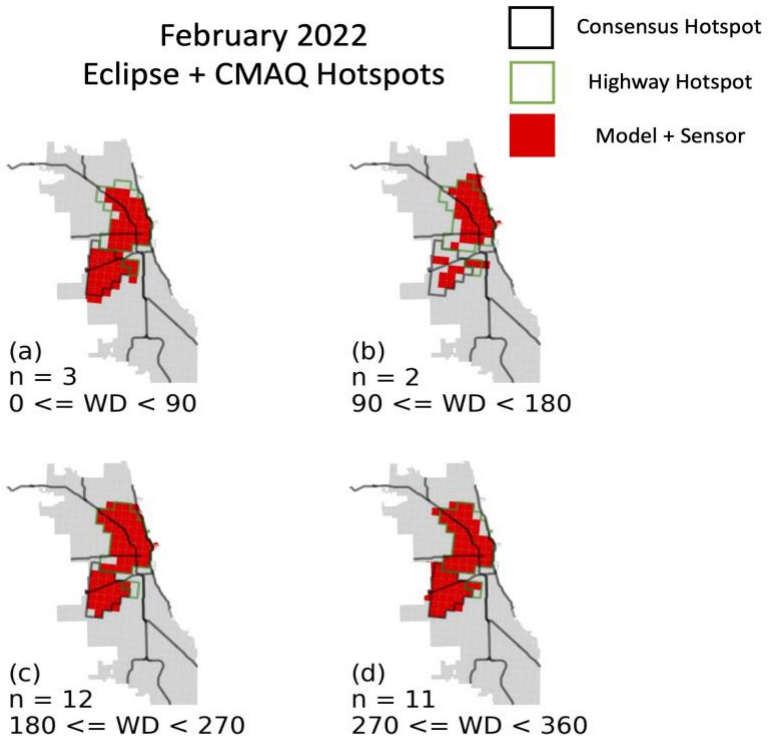


Fig. S2.5. The intersection of WRF-CMAQ and Eclipse hotspots over Chicago given each dominant wind directions for February 2021. The number of days with each dominant wind direction are noted. The extent of the consensus hotspot is outlined in black, the extent of the highway hotspot is in green.

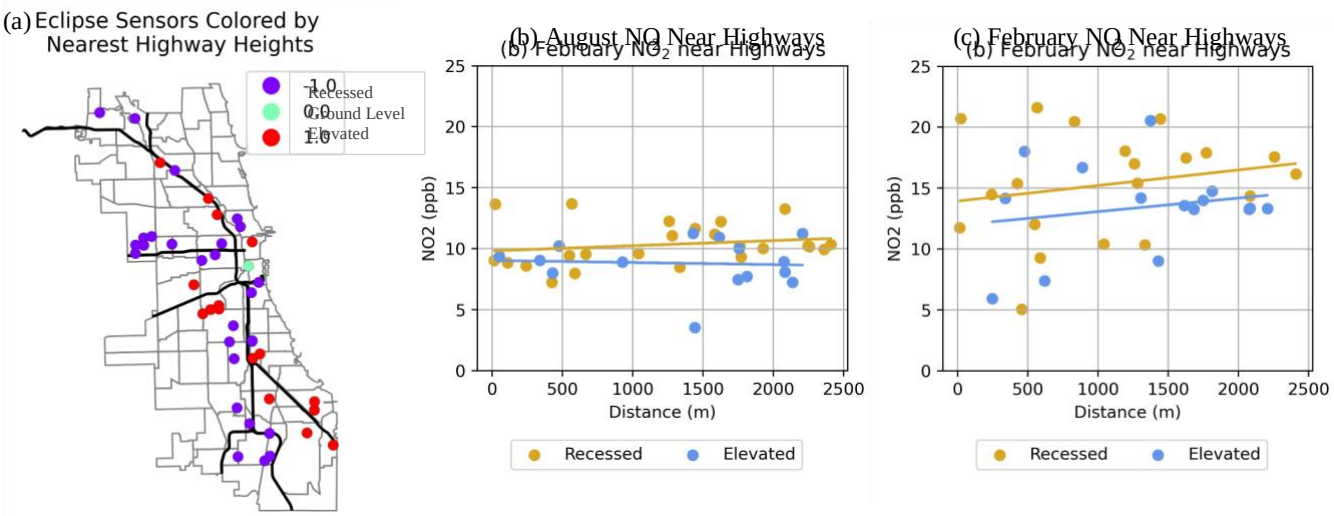


Fig. S2.6. (a) Locations of Eclipse sensors colored by highway heights: Recessed (purple, n = 20), ground-level (green, n = 1), and elevated (red, n = 15). (b) The average NO_2 at near-highway (< 2.5 km) sensors stratified by highway heights for (a) August and (b) February. Sensors are grouped by elevation: yellow shows average Eclipse measurements at stations near recessed highways and blue are sensors near elevated highways. The line of regression comparing average NO_2 to distance are drawn for each highway type.

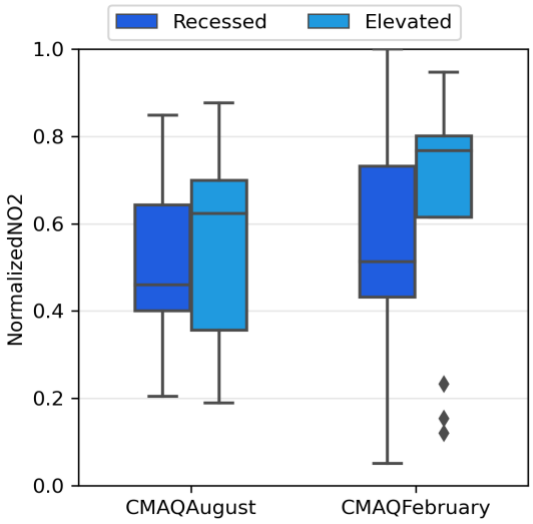


Fig. S2.7. WRF-CMAQ normalized NO_2 concentrations near elevated and recessed highways. Differences are not statistically significant.

Table S2.1. Average values of NO_2 concentration, land-use, and social characteristics within hotspots and for all of Chicago for August 2021. Statistically significant characteristics relative to the Chicago average are indicated by an asterisk (*) and determined using a t-test with Bonferroni adjustment.

Type	Characteristics	Units	Consensus	Observational	Modeled	Chicago
NO_2	Eclipse	ppb	11.0*	10.4*	9.4	9.1
	CMAQ	ppb	18.3*	14.5	18.2*	14.0
	TropOMI	molecules/cm^2	9.3*	9.3*	8.1	8.0
Green Space	ndvi	-	0.3*	0.4	0.3*	0.4
Traffic	Highway	Count	0.4	0.1	0.7*	0.3
	Arterials	Count	20.0	23.4	27.9*	17.3
	Bus Speeds	m/s	6.9	6.2	13.7*	4.0
Zoning	Industrial	Area	0.3*	0.1	0.2	0.1
	Residential	Area	0.4	0.5	0.2*	0.4
	Commercial	Area	0.1	0.1	0.2*	0.1
Demographics	Total Population	People	8385	10004	12202*	7472
	White	People	1485	1089	7288*	2543
	Black	People	1949	4610*	1023	2115
	Hispanic/Latino	People	4488*	4109	1378*	2175
	Asian	People	364	85	2196*	478
Income + Assistance	Income	$ per person	65336	57235	209428*	71828
	Public Assistance	$ per recipient	2748	3270	5945*	2893

Table S2.2. Average values of NO_2 concentration, land-use, and social characteristics within hotspots and for all of Chicago for August 2022. Statistically significant characteristics relative to the Chicago average are indicated by an asterisk (*) and determined using a t-test with Bonferroni adjustment. In comparing this table to August 2021, demographic, traffic, zoning, and income characteristics change because of the extent of the hotspots in February 2022, as opposed to people moving; which contrasts to NDVI and NO_2, which also change due to seasonality.

Type	Characteristics	Units	Consensus	Highway	Chicago
NO_2	Eclipse	ppb	16.8*	15.3	14.2
	CMAQ	ppb	16.4*	15.9*	13.5
	TropOMI	molecules/cm^2	21.4*	16.7	17.5
Green Space	ndvi	-	0.1*	0.1*	0.2
Traffic	Highway	Count	6.4	7.1*	4.1
	Arterials	Count	19.7	27.3*	17.4
	Bus Speeds	m/s	6.4	7.1*	4.0
Zoning	Industrial	Area	0.4*	0.1	0.1
	Residential	Area	0.4	0.4	0.4
	Commercial	Area	0.1	0.1*	0.1
Demographics	Total Population	People	8574	9830*	7472
	White	People	680*	2770*	2543
	Black	People	531	1417*	2115
	Hispanic/Latino	People	6961*	3566	2175
	Asian	People	363	1883*	478
Income + Assistance	Income	$ per person	34668	80610*	71828
	Public Assistance	$ per recipient	2325	3770*	2893

Table S2.3. February 2022 near I-290 Highway. Units are in ppb unless otherwise noted. EPA measurements in the table reflect the average EPA values when compared to observations, so missing observations affect the average EPA measurement.

Data	Time	EPA	Prediction	mb	nmb (%)	rmse	nrmse (%)	r
CMAQ	All Day	17.9	17.7	-0.2	-1.1	10.5	2.5	0.5
	Morning (6:00 – 9:00 CT)	20.5	20.3	-0.2	-0.9	12.6	2.8	0.3
	Afternoon (16:00 – 19:00 CT)	17.7	27.8	10.1	56.9	12.5	3.0	0.8
Eclipse	All Day	18.0	16.5	-1.5	-8.1	9.3	2.2	0.4
	Morning (7:00 – 9:00 CT)	19.9	17.0	-2.9	-14.3	8.5	1.9	0.7
	Afternoon (16:00 – 19:00 CT)	17.9	19.2	1.4	7.6	7.0	1.7	0.2

Table S2.4. August 2021 near I-90 Highway. Units are in ppb unless otherwise noted. EPA measurements in the table reflect the average EPA values when compared to observations, so missing observations affect the average EPA measurement.

Data	Time	EPA	Prediction	mb	nmb (%)	rmse	nrmse (%)	r
CMAQ	All Day	13.7	19.2	5.4	39.6	12.7	3.4	0.5
	Morning (6:00 – 9:00 CT)	21.2	28.9	7.6	36.0	15.1	3.3	0.4
	Afternoon (16:00 – 19:00 CT)	9.3	9.6	0.3	3.1	6.6	2.2	0.4
Eclipse	All Day	13.8	9.4	-4.5	-32.3	10.6	2.8	0.3
	Morning (7:00 – 9:00 CT)	21.2	12.1	-9.1	-42.9	13.0	2.8	0.6
	Afternoon (16:00 – 19:00 CT)	9.2	8.5	-0.7	-7.3	6.0	2.0	-0.1

Table S2.5. Average NO_2 concentrations from Eclipse and overlaid WRF-CMAQ grid cell, grouped by recessed and elevated highways. Units are in ppb unless otherwise noted.

Month	Data	Recessed	Elevated	Difference	Difference (%)	p-value
August 2021	Eclipse	10.3	9.2	1.2	12%	0.03
	WRF-CMAQ	14.8	15.2	-0.4	-3%	0.65
February 2022	Eclipse	15.3	13.4	1.9	13%	0.19
	WRF-CMAQ	14.2	14.9	-0.8	-5%	0.21

Appendix A: WRF-CMAQ Specifications and Model Performance.

WRF-CMAQ Model Specifications.

To create the coupled Weather Research Forecast and Community Multiscale Air Quality (WRF-CMAQ) model simulations, we follow the guidance outlined in Byun & Schere, 2006, such that we 1) run a stand-alone WRFv3.2 simulation, 2) create meteorologically-informed emissions using the 2017 SMOKE platform, and 3) run the coupled WRF-CMAQ model. In this study, we only use the 1.3 km output (Figure A1(b)).

We run standalone WRF with 3 nested domains at a horizontal resolution of 12 km, 4 km, and 1.3 km resolution Figure A1(a). The WRF simulation is run with 35 vertical layers from the surface to 30 hPA. We use the 12 km WRF output to create initial and boundary conditions for the 4 km domain and we use the 4 km model output to create the same for the 1.3 km domain. The WRF physics options are described in Table A1. The initial conditions and 3 hourly lateral boundary conditions for the 12 km domain are sourced from the North Atlantic Regional Reanalysis (NARR; (5)).

We use the 2017 NEI Modeling Platform (3) to create emissions for August 2021 and February 2022. We update the PTEGU and marine shipping sector to integrate the historical reported emissions: for the PTEGU sector, we use the Continuous Emissions Monitoring data (CEM) from the EPA and for the marine shipping sector, we time-shift the most recently available marine shipping emissions data (2019) from the EPA. We use WRF output to create meteorologically informed emissions for the onroad sector using MOVE2014 output in the 2017 platform. To ultimately create emissions, we use the 2017 platform to create emissions for the onroad, point-source, and non-road sectors. We calculate biogenic emissions (BEIS), lightning NOx emissions, and windblown dust inline during the CMAQ simulation.

The final WRF-CMAQ simulations were performed at the 4 km and 1.3 km domain. We use the same WRF physics option as described in Table A1. For the 4 km outer-simulation, Like the standalone WRF simulation, we use NARR (5) for meteorological boundary conditions. We use WACCM (6) for chemical initial and time-evolving boundary conditions. We then use the chemical and meteorological output from

4 km simulation to provide boundaries for the 1.3 km simulation. Both the 4 km and 1.3 km domain use the Carbon Bond Mechanism version 6 and aerosol module version 2016 with aqueous chemistry (cb6r3_ae6_aq) to create atmospheric constituents.

WRF-CMAQ Model Performance.

We analyze the 1.3 km WRF-CMAQ performance over the full and cropped Chicago domain (Figure 1b) and within Chicago (Figure 1). To evaluate the statistical performance of the WRF-CMAQ simulation, we calculate the mean bias (MB), normalized mean bias (NMB), root-mean squared error (RMSE), normalized root-mean squared error (NRMSE), and Pearson's correlation coefficient (r). We calculate the performance of observations (x) to model output (y), with pixel locations indicated by index locations (i,j). Within Chicago, there is only one EPA NO_2 station within city limits, with 2 located near the city (Figure 3.1). The EPA COM-ED station, just south of city limits, is used to calibrate the Eclipse sensor network, but we choose to expand the model performance to the other 2 nearby EPA stations as the EPA COM-ED station does not report NO_2 for August 2021.

Given this, we find that WRF-CMAQ simulated NO_2 in Chicago is biased high in August 2021 (6.1 ppb, 39.8%), but biased low in February 2022 (-0.1 ppb, -0.5%), though the hourly variation is similarly captured in summer and winter (r = 0.5) (Table S2.1). The bias in the August 2021 is driven by nighttime (10 pm - 5 am CST) concentrations of NO_2 ($August_{night}$ MB = 9.2 ppb, NMB = 57%, r = 0.33). For posterity, we include the full 1.3 km modeling domain performance in Table A2.

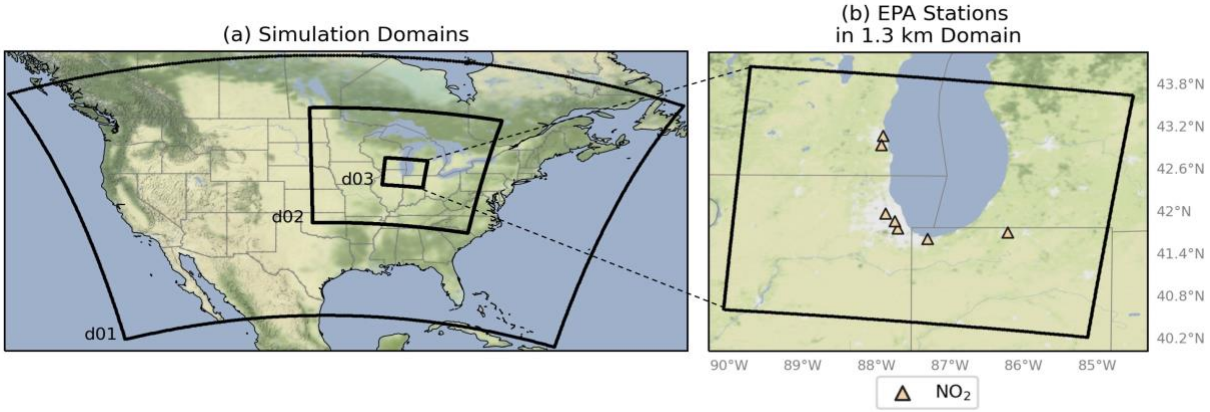


Fig. A1. (a) The nested WRF domains with (b) full domain EPA NO_2 coverage for the 1.3 km simulation.

Table A1. Model physics options

Model	Domain	Time Period	Option	Value
WRF	12 km, 4km, 1.3 km	Full run	Microphysics	Morrison 2-Moment
	12 km, 4 km	Full run	Cumulus Parameterization	Kain-Fritsch v2 (KF2)
	12 km, 4km, 1.3 km	Full run	Planetary Boundary Layer	Asymmetric Convective Model v2
	12 km, 4km, 1.3 km	Full run	Shortwave Radiation	Rapid Radiative Transfer Model for GCMs (RRTMG)
	12 km, 4km, 1.3 km	Full run	Longwave Radiation	Rapid Radiative Transfer Model for GCMs (RRTMG)
	12 km, 4km, 1.3 km	Spin-Up	Soil Moisture	On
	12 km, 4km, 1.3 km	Full run	Temperature Nudging	On
	12 km, 4km, 1.3 km	Full run	FDDA Nudging: Temperature	0.0003
	12 km, 4km, 1.3 km	Full run	FDDA Nudging: Winds	0.0001
	12 km, 4km, 1.3 km	Full run	FDDA Nudging: Water Vapor	0.0001
	12 km, 4km, 1.3 km	Full run	Timestep (s)	60, 20, 6
CMAQ	4 km, 1.3 km	Full run	Radiation Timestep (min)	10, 5
	4 km, 1.3 km	Full run	Coupling (steps)	8

Table A2. Model performance of NO_2 (in ppb) in Chicago and over 1.3 km Domain.

Domain	Stations	Month	*μEPA*	*μCMAQ*	MB	NMB	RMSE	NRMSE	r
Chicago	2	08/21	15.3	21.4	6.1	0.4	14.0	0.9	0.5
	3	02/22	16.5	16.4	-0.1	0.0	11.2	0.7	0.5
d03	7	08/21	11.8	13.7	1.9	0.3	10.8	0.9	0.5
	7	02/22	16.9	14.3	-2.6	-0.2	11.4	0.7	0.5

Appendix B: TropOMI Satellite Processing and Validation

TropOMI Processing

We retrieved daily, level-2 observations of NO_2 from TropOMI using the SentinelSat Python package. In order to create a product which we can directly compare to the interpolated eclipse surface and CMAQ grid, we down-scaled the L2 data (3.5 km x 7 km) to the WRF-CMAQ grid (1.3 km x 1.3 km). We use a three step approach to create the monthly average NO_2 from the L2 retrievals. First, we apply the bilinear interpolation function to regrid the daily retrievals (with pixel area of 3.5 km x 7 km) to the 1.3 km x 1.3 km grid. Second, we mask for missing or low-quality pixels, where $qa_{value} < 0.75$. Finally, we average the daily retrievals to create a monthly average. We repeat this process to create a surface for August 2021 and February 2022. As shown in B1, the number of valid, 1.3 km pixels are significantly higher in August 2021 compared to February 2022, with an average data retention of 78% per pixel in August compared to 16% in February.

We do not create "ground-level" NO_2 from the TropOMI satellite by implementing an averaging factor, and this may contribute to biases in the VCD values within urban areas (7). As our domain for clustering exclusively centers on Chicago, we assume all Chicago pixels would contain the same urban bias, though caution should be taken interpreting the absolute values from the TropOMI satellite for this reason. Further, in testing our interpolation of TropOMI pixels against EPA values, we find a linear relationship between the column and ground-based measurements ($r = 0.3 - 0.7$; Figure B1), which is also supported by previous studies (cite), so we omitted further processing.

TropOMI Product Performance against EPA stations

We then compare the regridded satellite product to EPA stations, as previous studies have used this as a check for spatial representativeness. Due to the limited EPA station coverage in Chicago, we expand the testing set to include the full Eclipse sensor network. As such, we show that the satellite re-gridding algorithm shows a weakly linear relationship with the Eclipse and EPA measurements of NO_2 ($0.3 < r < 0.7$; B1). While the absolute values cannot be directly compared across the TropOMI and

ground-based data sets, the *mb* and *rmse* further highlight the strength of the linearity of the column and ground-based data.

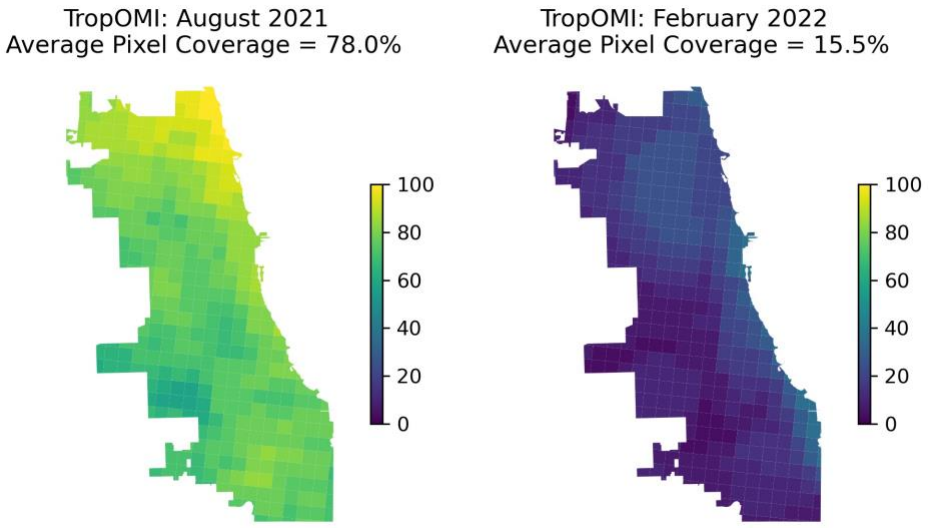


Fig. B1. The (a) August 2021 and (b) February 2022 valid pixel counts over Chicago to show the valid NO_2 retrievals within Chicago.

Table B1. TropOMI column NO_2 compared to EPA and Eclipse ground-based NO_2 measurements.

Date	Observation	N	Ground-Based	TropOMI	r
*August	EPA	2	13.7	6.9	-1
	Eclipse	115	9.4	7.6	0.4
*February	EPA	3	16.8	16.3	0.25
	Eclipse	89	14.7	17.9	0.34

Appendix C: Eclipse Sensor Network Specifications and Interpolation Performance

Hardware and Network Design.

This paper uses data from Project Eclipse, a dense citywide network of low-cost sensing devices ("Eclipse nodes") first deployed in Chicago in July of 2021. Eclipse nodes comprise four electrochemical gas sensors (NO_2, CO, O_3, and SO_2) from the Alphasense B4-series; a Sensirion SPS30 optical particle sensor collecting readings on PM_2.5; and temperature, humidity, and pressure sensors. The citywide network included (1) 80 locations selected across the city following a stratified random sampling approach adapted from the design of the New York City Community Air Survey (8); (2) 23 locations selected by local community-based organizations; and (3) 6 additional locations recommended by local partner organizations based on their observations of gaps in the overall network. An additional nine devices were collocated at three EPA regulatory monitoring stations (see Calibration Section for details). Devices collected and communicated readings every five minutes via an LTE-m network. For further details on the hardware and network design, see (9) and (10).

Quality Assurance/Quality Control

The raw data included 902,133 readings covering 3,363 device-days across 123 locations in August 2021 and 717,274 readings over 2,702 device-days from 106 locations in February 2022. We cleaned and processed the data following a QA/QC approach adapted from (10). First, 5-minute readings were aggregated to hourly averages. We applied a 75% completeness criterion, excluding sensor-hours whenever we observed fewer than 9/12 expected readings (2.7% of sensor-hours in August 2021 and 3.6% of sensor-hours in February 2022). We then aggregated to daily averages, excluding sensor-days with greater than 18/24 expected readings (2.6% of sensor-days in August 2021 and 4.4% of sensor-days in February 2022). We further excluded any devices that, after QA/QC, reported for fewer than 5 days (14% of locations in August 2021 and 2.8% of locations in February 2022). Finally, to exclude malfunctioning sensors, we further implemented a nearest-neighbor check in which we calculated the correlation between daily averages from a given device with the overall daily average values of all

neighbor devices within 5km. After examining the distribution of correlations observed in the data, we excluded a small number of additional devices that diverged from nearby devices so clearly as to suggest a potential malfunction or device issue rather than true hyperlocal variation ($\rho > 0.5$). This cutoff led to the exclusion of N = 1 location in August and N = 10 locations in February. We further exclude Eclipse nodes that were co-located with EPA stations from our training set to ensure that these data do not lead us to overestimate performance in comparison with EPA.

The final data set included 2,923 device-days from 95 locations in August 2021 and 2,073 device-days from 85 locations in February 2022. The relatively higher fraction of short-lived locations in August is largely attributable to communications issues with locations that emerged in the first several months of the deployment and then stabilized; relatively higher hourly and daily data loss was observed in February versus in August, however, because devices switched automatically into power-saving mode when limited sunlight reduced solar charging to below a specified threshold. Throughout the project, incorporation into existing infrastructure maintenance routines and a replace-reset-relocate protocol minimized data missingness (9).

Calibration

As has been well-documented, low-cost sensors are subject to inaccuracies due to drift and interference in field deployments (11). To address these concerns, the Eclipse project developed a calibration function to improve the accuracy of the sensor readings following previous research (12, 13). Briefly, 3 additional devices were co-located at an EPA regulatory monitoring station in Chicago that collected hourly readings of NO_2 using a ThermoScientific monitor (method 074).∗. Using all available data until the period of this study (May 12th, 2021 - February 28th, 2022), we developed a machine learning calibration model as follows.

First, we split the data into training (70%) and test sets (30%). To account for dependencies in the data—both temporal autocorrelation and the similarities between multiple devices at the same location, reading at the same time—we grouped our data by day before conducting the split. Within the training

set, we further used 5-fold cross validation, grouped by day, to train and tune models. We then evaluated model accuracy on our test set.

We considered input features including relative humidity (RH), absolute humidity (AH), temperature in Fahrenheit (TempF), pressure (P), gas readings (NO_2, O_3, CO, SO_2) and fine particulate matter (PM_2.5). We also included the meteorological parameters of wind speed (WS) and wind direction (WD) obtained from the National Oceanic and Atmospheric Administration weather station at Chicago O'Hare airport. Given prior evidence of nonlinear relationships with relative humidity (11), we further assessed whether using exponential transformations, b-splines (BS0, BS1, BS2), or interactions with other variables improved the models' fit. To increase robustness to outliers, gas readings were winsorized at the 0.001 and 0.999 percentiles.

We experimented with a variety of possible models; we then focused on developing a smaller subset that performed well on the initial training data. These included linear regression, random forest regression, gradient boosting regression, and categorical boosting regression. Linear regression, random forests, and gradient boosting are among the most common models used for calibration in prior literature (11); gradient boosting, in particular, is a leading method for learning problems in which data are noisy, features are heterogenous, and relationships are complex. However, because boosting algorithms estimate gradients at each step using the same data points with which the model was built, they can be subject to prediction shift, a bias in the residual errors that affects generalizability on unseen data (14). Categorical boosting uses ordered boosting, which enables the calculation of residuals on a different sample from the data with which each model is trained on, to mitigate the problem of prediction shift (14, 15).

After finding the top regression methods using an initial set of all features, we evaluated all possible input feature combinations to find the best performing model. We then iterated through the regression methods once more to confirm that the top methods remained the best-performing using the new predictors. We investigated additional adjustments to further improve our models. These included scaling, polynomial regression, and hyperparameter tuning; we tested these adjustments using AzureML

automated ML model-training jobs. Results were ultimately comparable with models using default hyperparameters. Finally, we selected and evaluated the calibration model with the best overall performance.

The best-performing model for NO_2 was the categorical boosting regression. Table C1 presents the performance metrics for our model; included variables and their importance are detailed in Figure C1. The NRMSE is slightly higher than thresholds recommended by EPA for hotspot detection (NRMSE = 42% versus 30%) but the relatively low RMSE and high R2 suggest that the readings provide meaningful signal. Also important is the low coefficient of variation (CV), suggesting high inter-device reliability. The CV is high even when data is not calibrated, suggesting that the problems of interference and drift affect different devices in similar ways – bolstering our confidence that a calibration algorithm developed at one location and for one device would generalize to another device located elsewhere. Unfortunately, we were not able to secure a second location, during the study period, at which to validate this hypothesis; however, recent evidence from the Breathe London network similarly offers evidence of reliability in calibration results across locations (16).

Figure C1 shows the relative importance of the parameters included in the final model. Electrochemical gas sensors are highly and non-linearly sensitive to temperature, which is likely a reason that boosting approaches outperform simpler, linear methods (11). As expected, the NO_2 sensor is an important input; however, it is relatively less important than CO, suggesting that the CO sensors—which tend to be more long-lasted that NO_2 sensors (17)—may be driving up accuracy through the detection of combustion-related sources that are correlated with NO_2. The inclusion of the remaining gas sensors and WS further contribute small but meaningful performance improvements. These results highlight the benefits of multi-sensor devices for enabling more robust calibration functions, but we also note that results should be treated with caution – and compared with other sources of data, as is done here – given the importance of proxy and correlational factors as contributors to the calibrated estimate.

Table C1. Evaluation of Calibration Models for NO_2. Metrics are calculated following EPA guidelines for performance protocols and metrics.

	RMSE	NRMSE	R^2	CV	SD
Raw Data	74.5	100%	-2.35	14.7%	10.8%
5-Fold CV	4.3	41.5%	0.70	11.9%	1.2
Test Set	4.58	42%	0.71	11.3%	1.2

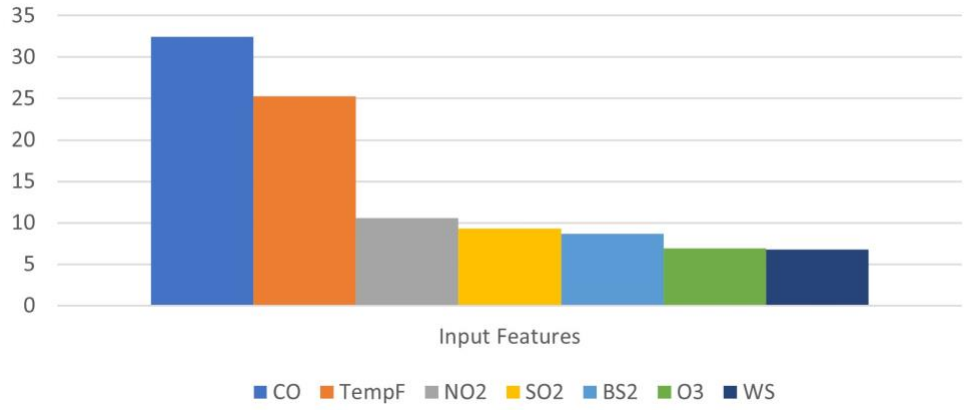


Fig. C1. Importance Score of Input Features of NO_2 Calibration Model

Interpolation

We investigated several methods of interpolating data from the Eclipse node locations to the 1.3km grid. In particular, we compared a commonly-used geostatistical method (inverse distance weighting)† as well as machine learning approaches (random forests and support vector machines). For each method, we modeled the calibrated NO_2 readings as a function of distance only; although researchers also commonly interpolate using more data-intensive methods such as land-use regression, we did not include other predictive variables because we were seeking to evaluate the value of a "sensor-only" approach—which would be more manageable for e.g. a government agency seeking to routinely monitor hotspots, particularly in contexts where land use variables are poor quality or difficult to obtain.

We trained and evaluated our models using a machine learning approach. First, for each month, we split our data by location—allocating approximately 80% of device locations to a training set and 20%

of device locations to the test set. Within the training set, we then used leave-one-out cross validation (LOOCV) to tune the hyperparameters of each model. Because our interest was in the spatial variation, rather than in the temporal variation, we conducted LOOCV for each day and then used the parameters that performed best, on average, across all days (as evidenced by producing the lowest RMSE). We used each model to interpolate all training data to each grid cell for each day, and then averaged across the grid cell estimates to obtain monthly estimates. Finally, we evaluated the models' predictions in comparison with the locations in the test data as well as in comparison with the readings at all EPA stations in Chicago for which gold-standard data were available (One station in August and 2 stations in February).

Table C2 presents the performance of each of the three interpolation models. All models perform relatively better, compared to the test set, in August versus in February with relatively lower RMSE and higher ρ—although NRMSE are similar. The random forests model performs best in comparison with the EPA sites as well as on the test set for August, although the inverse distance weighted model does perform slightly better in predicting test set estimates for February. Notably, all three models produce similar surfaces (Figures C2 and C3). Nevertheless, given the slightly better performance enabled by the machine learning approach, we choose to use random forest interpolation on the complete data set for the analyses in this study.

Table C2. Evaluation of Interpolation Models for Eclipse NO_2.

	Inverse Distance Weighting	Support Vector Machine	Random Forests
August			
Test Data			
RMSE	2.01	1.96	1.70
NRMSE	0.20	0.19	0.17
ρ	0.60	0.56	0.69
EPA			
RMSE	4.13	3.85	3.77
NRMSE	0.30	0.28	0.28
February			
Test Data			
RMSE	2.82	3.27	3.09
NRMSE	0.19	0.22	0.21
ρ	0.43	0.17	0.36
EPA			
RMSE	1.57	2.05	1.42
NRMSE	0.10	0.13	0.09

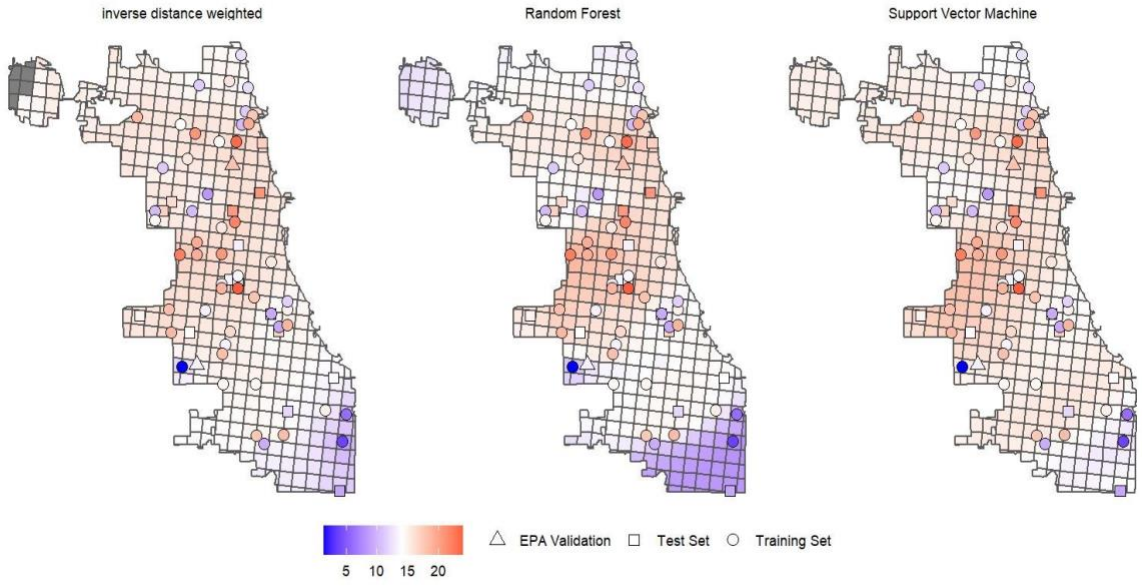


Fig. C3. Monthly average estimated NO_2 by interpolation model for February, 2022

Appendix 3

Supplement for Chapter 4

Supporting Information for

Exposure and health disparities in Chicago Air Pollution

Contents of Appendix 3:

Supplementary Figures and Tables for Main Text

Figures S3.1 – S3.9

Tables S3.1 – S3.10

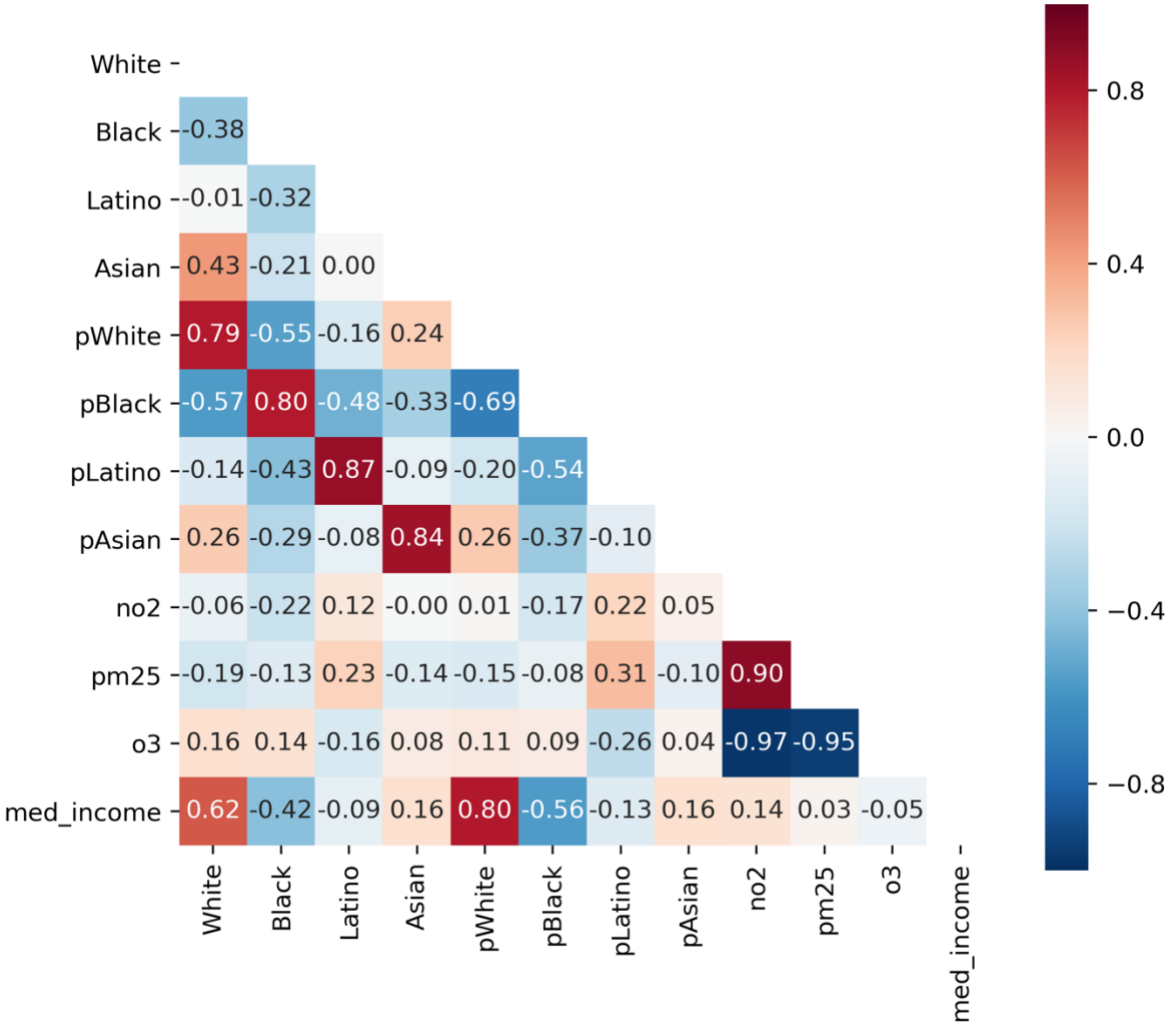


Figure S3.1. Correlation between demographic variables and pollutants.

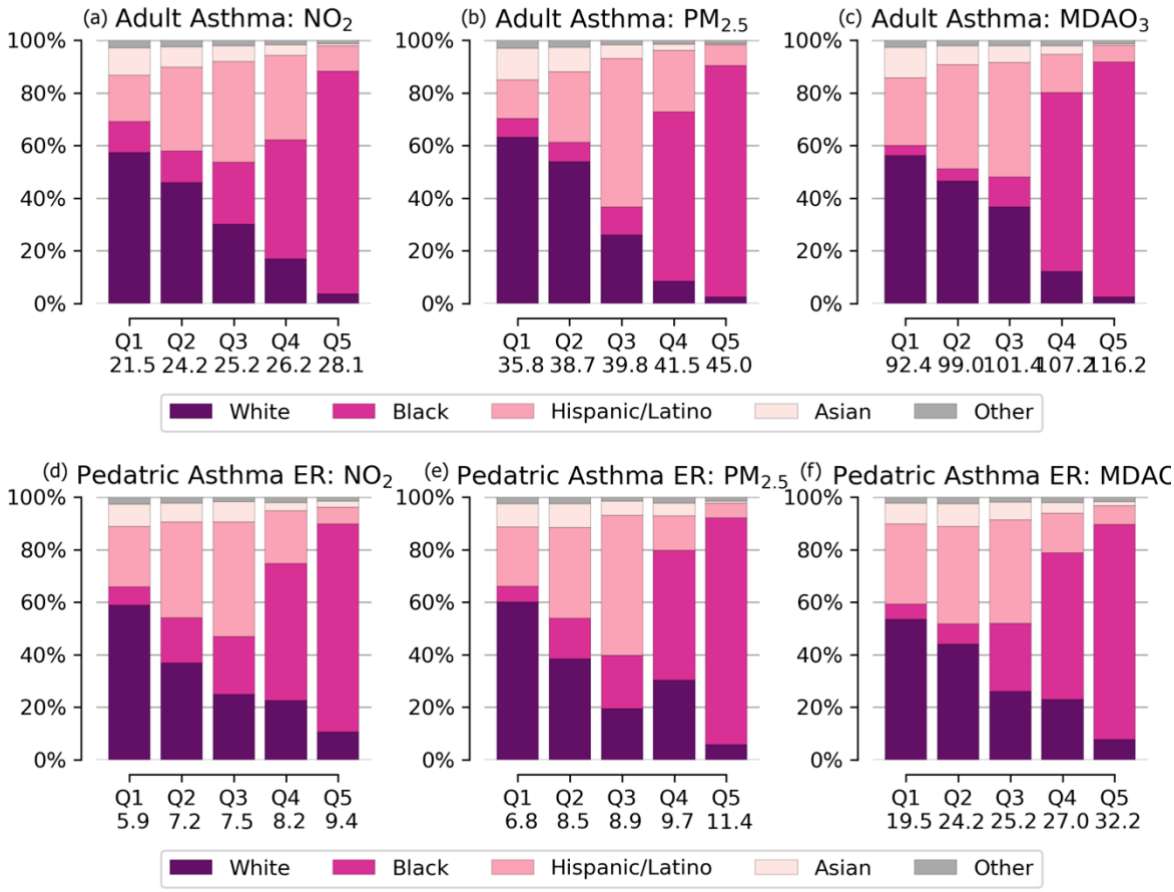


Figure S3.2. Average demographics per quantile of adult asthma (a-c) and pediatric asthma hospitalizations (d-f) attributed to NO_2, $PM_{2.5}$, and $MDAO_3$. Quantile bounds are annotated below the x-axis.

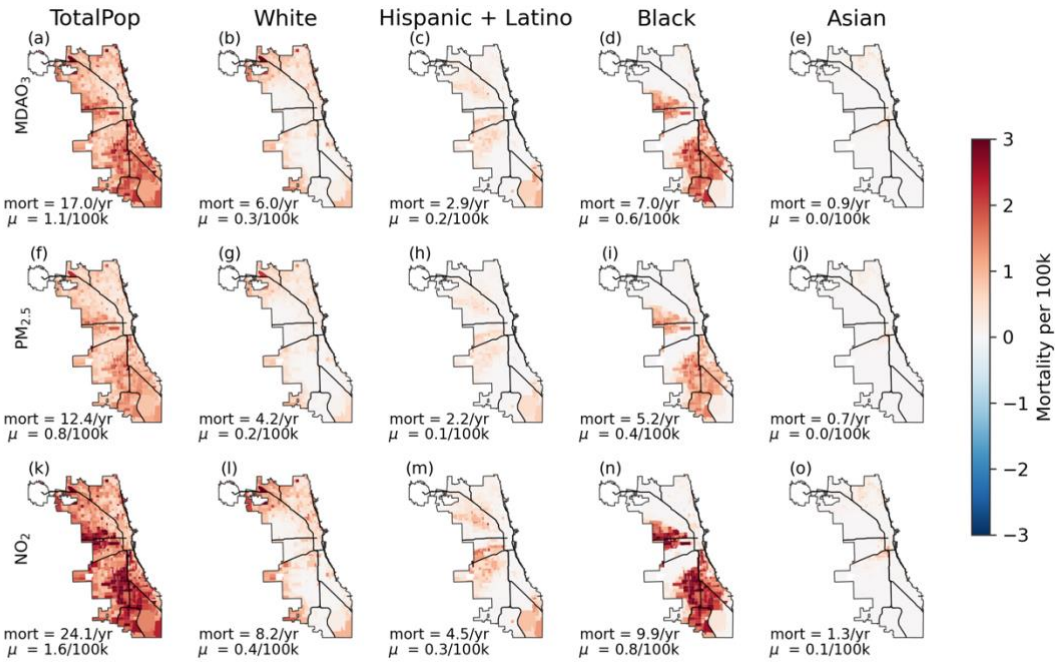


Figure S3.3. Attributable mortality rate per 100k adults from NO_2, $PM_{2.5}$, $MDAO_3$. Annotated are the total attributable cases (cases) and the average rates (μ).

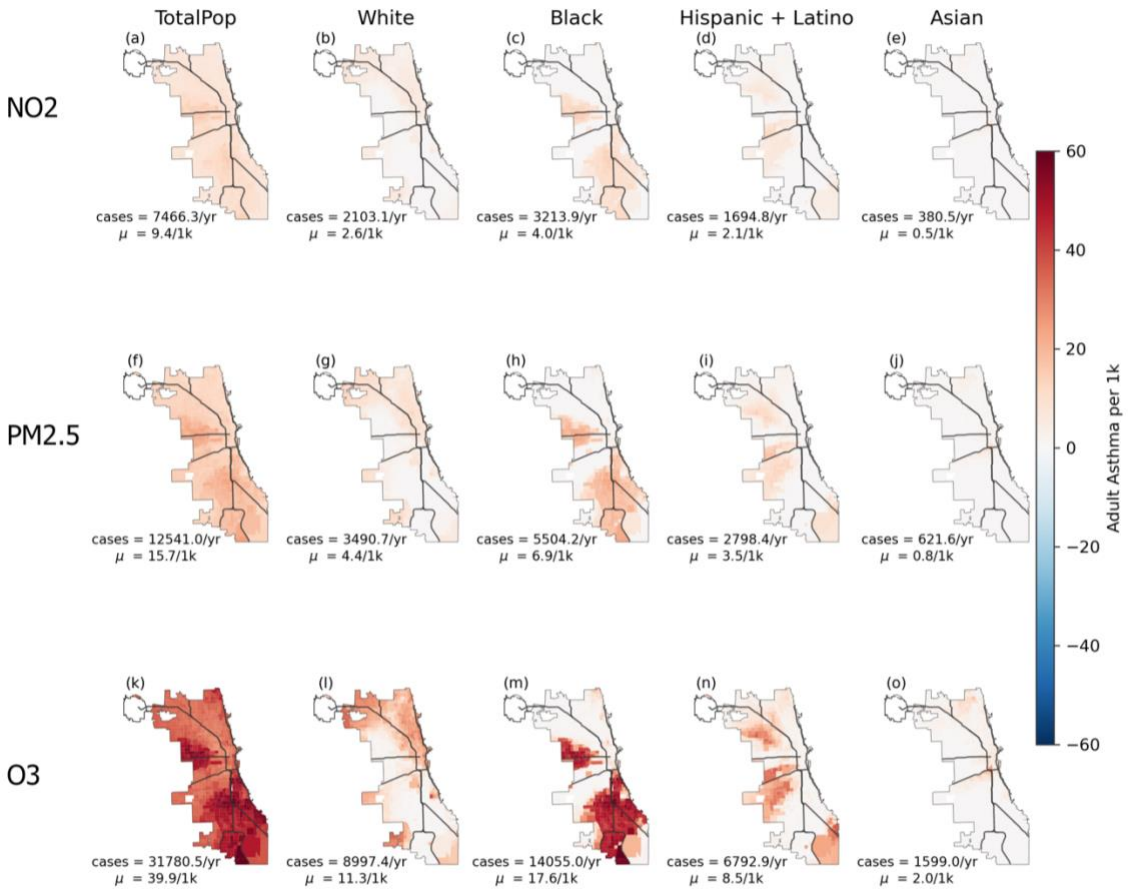


Figure S3.4. Attributable asthma rates from NO_2, $PM_{2.5}$, and $MDAO_3$. $PM_{2.5}$, $MDAO_3$. Annotated are the total attributable cases (cases) and the average rates (μ).

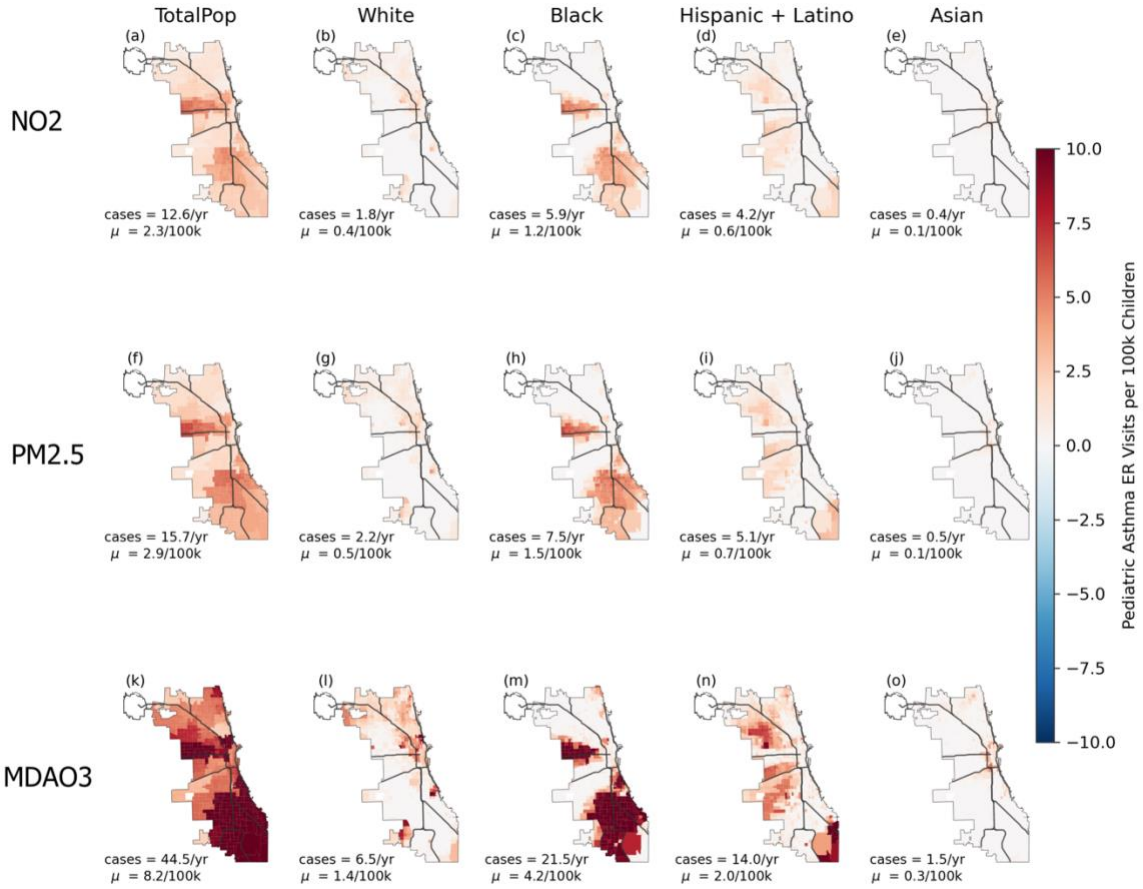


Figure S3.5. Pediatric asthma ER visits as rates per 100k children from NO_2, $PM_{2.5}$, $MDAO_3$. Annotated are the total attributable cases (cases) and the average rates (μ).

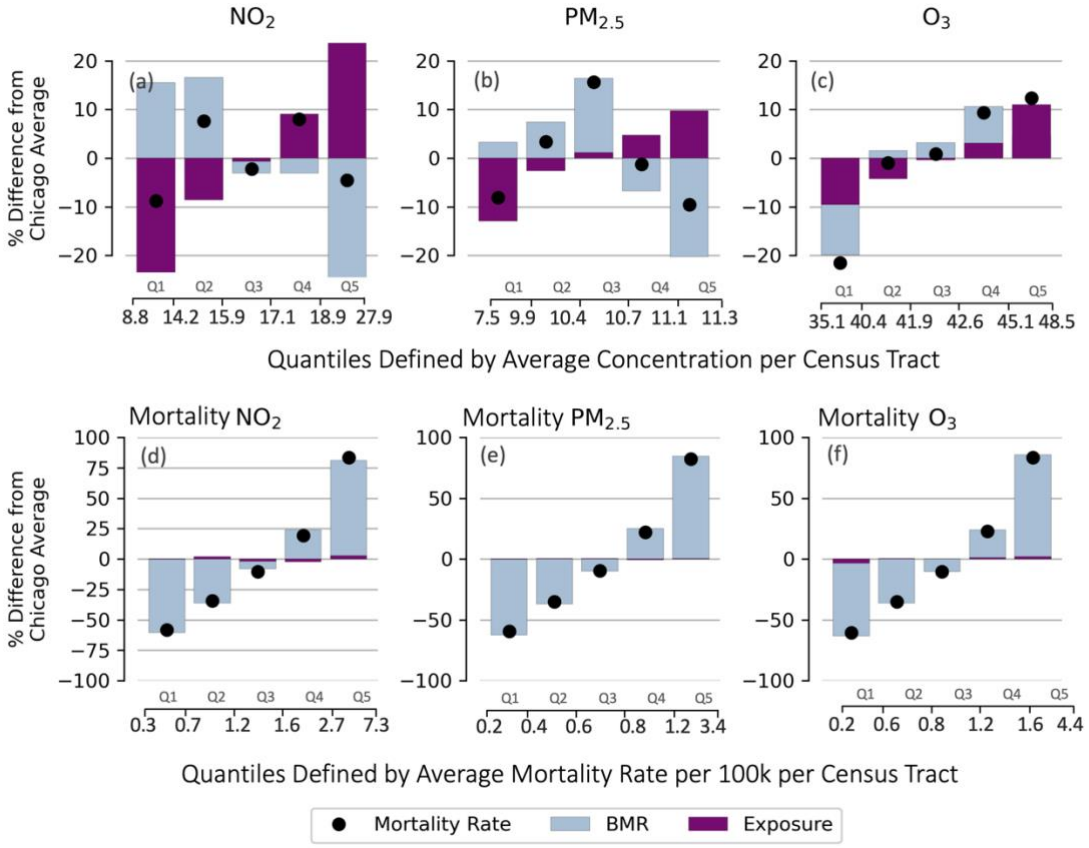


Figure S3.6. Baseline mortality rates (blue) comparted to exposure of pollutants (purple), with attributable mortality rate plotted as a black circle. For a-c, quantiles are ordered by pollution exposure levels (Q1 = lowest pollution, Q5 = highest pollution). For d-f, groups are ordered by baseline mortality rates (BMR), with Q1 having the lowest BMR and Q5 having the highest BMR.

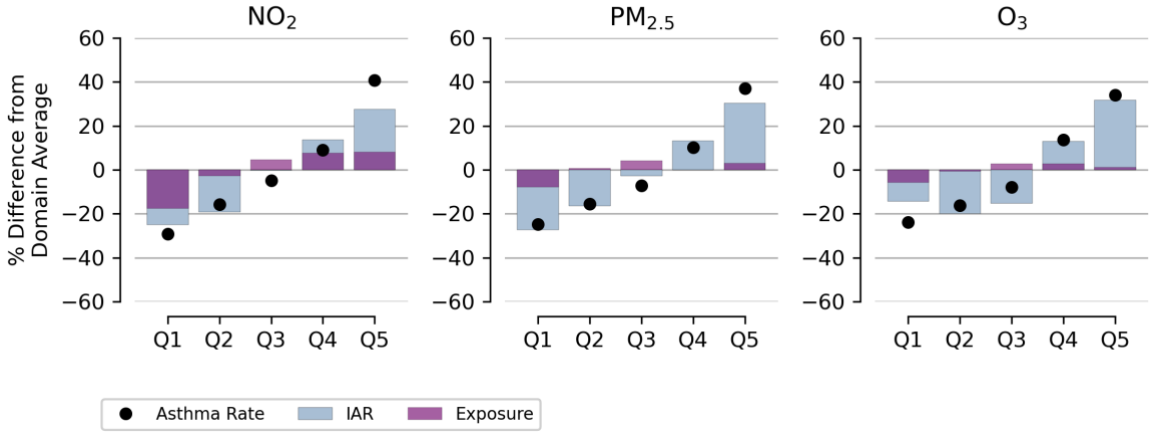


Figure S3.7. Baseline adult asthma incidence rates (blue) compared to exposure of pollutants (purple), with asthma rate plotted as a black circle. Groups are ordered by asthma incidence rates (AIR), with Q1 having the lowest (AIR), and Q5 having the highest AIR.

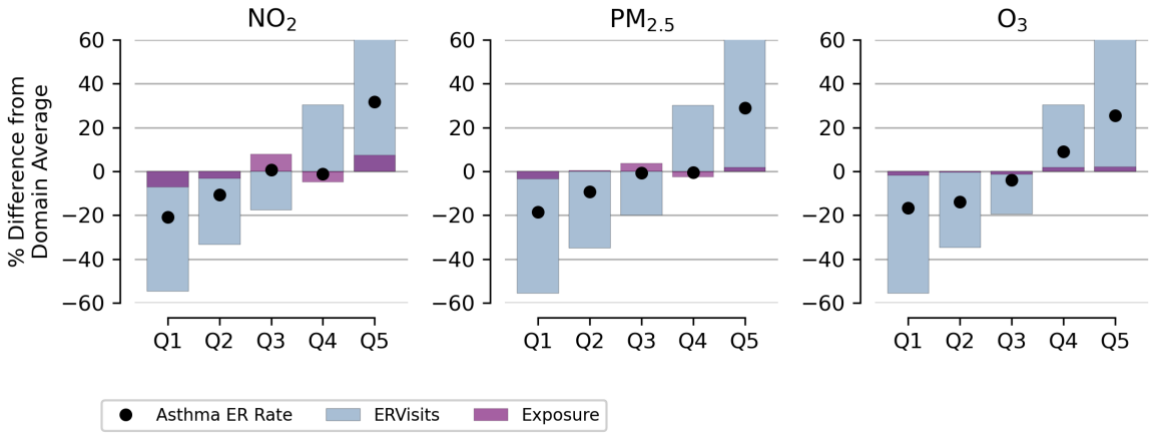


Figure S3.8. Baseline pediatric asthma ER hospitalization rates (blue) compared to exposure of pollutants (purple), with asthma ER hospitalization rates plotted as a black circle. Groups are ordered by pediatric asthma hospitalization rates (ER), with Q1 having the lowest ER and Q5 having the highest ER.

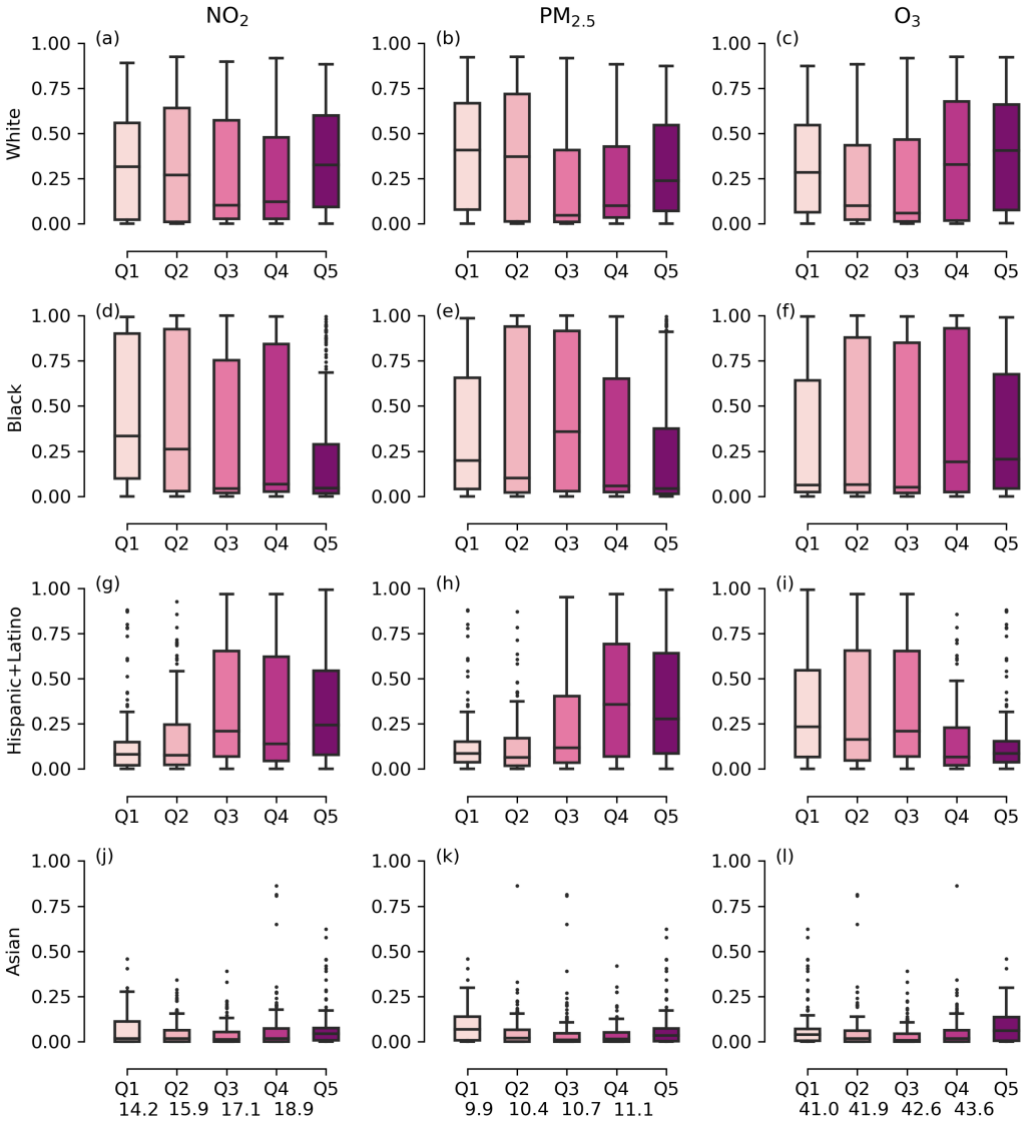


Figure S3.9. Demographic makeup of quintiles increasing pollutants of NO_2, $PM_{2.5}$, and $MDAO_3$. The ranges of pollutant concentrations are annotated at the bottom x axis.

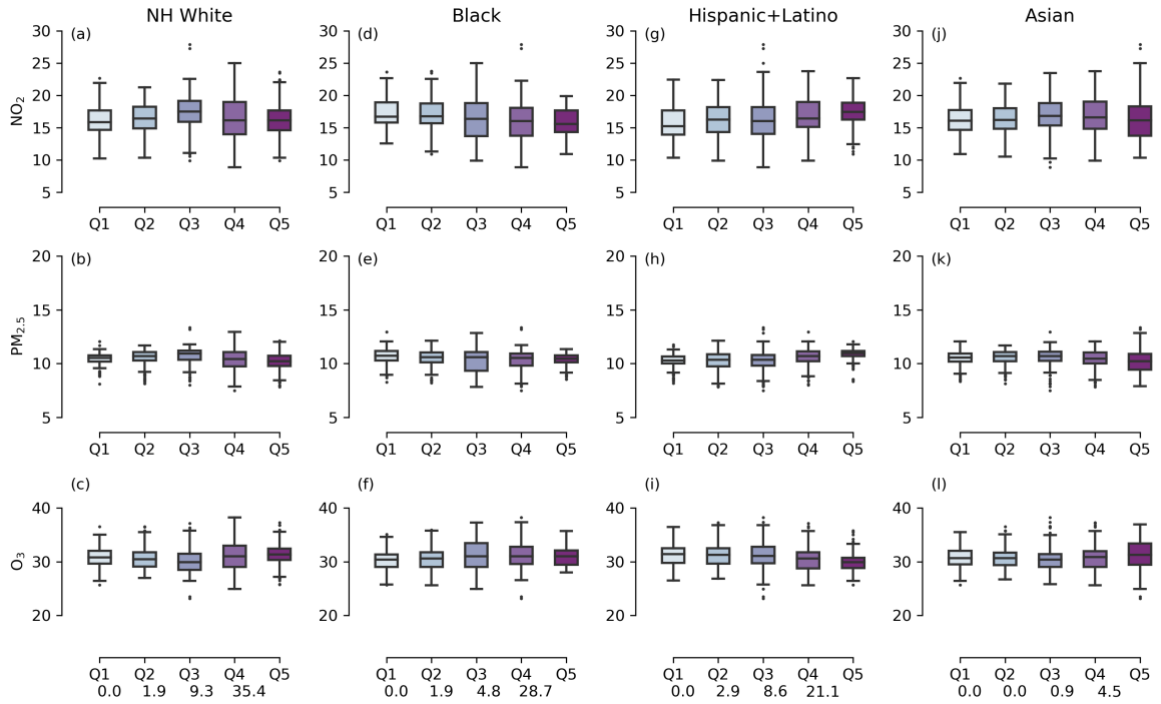


Figure S3.10. Average pollution concentrations organized by increasing proportions of White, Black, Hispanic and Latino, and Asian populations. The ranges of demographics are annotated at the bottom x axis.

Table S3.1. Mortality from pollutants, divided by population subgroup.

Pollutant	Population	Mortality	Mortality (5CI)	Mortality (95CI)	Mortality RT	Mortality RT (5CI)	Mortality RT (95CI)
NO_2	Chicago	17.0	8.7	32.6	1.1	0.6	2.1
	White	6.0	3.0	11.4	0.3	0.2	0.6
	Hispanic/ Latino	2.9	1.5	5.6	0.2	0.1	0.3
	Black	7.0	3.6	13.5	0.6	0.3	1.1
	Asian	0.9	0.5	1.8	0.0	0.0	0.1
	High Income	3.5	0.9	5.1	1.0	0.25	1.43
	Low Income	5.3	1.4	7.5	2.3	0.6	3.3
$PM_{2.5}$	Chicago	12.4	4.2	16.4	0.8	0.3	1.1
	White	4.2	1.4	5.6	0.2	0.1	0.3
	Hispanic/ Latino	2.2	0.8	3.0	0.1	0.0	0.2
	Black	5.2	1.8	6.8	0.4	0.1	0.5
	Asian	0.7	0.2	0.9	0.0	0.0	0.0
	High Income	1.8	0.6	2.4	0.4	0.2	0.7
	Low Income	2.8	0.9	3.7	0.9	0.3	1.2
$MDAO_3$	Chicago	24.1	6.1	33.8	1.6	0.4	2.2
	White	8.2	2.1	11.5	0.4	0.1	0.6
	Hispanic/ Latino	4.5	1.1	6.3	0.3	0.1	0.4
	Black	9.9	2.5	13.9	0.8	0.2	1.1
	Asian	1.3	0.3	1.9	0.1	0.0	0.1
	High Income	2.5	1.3	4.8	0.7	0.4	1.3
	Low Income	3.4	1.7	6.5	1.5	0.8	3.0

Table S3.2. Attributable adult asthma from pollutants, divided by population subgroup.

Pollutant	Population	Asthma	Asthma (5CI)	Asthma (95CI)	Asthma RT	Asthma RT (5CI)	Asthma RT (95CI)
NO_2	Chicago	7466.3	4163.6	16533.0	17.9	10.0	39.6
	White	2103.1	1172.9	4655.3	5.8	3.3	13.0
	Black	3213.9	1791.2	7129.0	6.3	3.5	14.0
	Hispanic/ Latino	1694.8	946.0	3743.4	4.4	2.5	9.8
	Asian	380.5	212.3	841.3	1.1	0.6	2.4
	High Income	2737.3	1527.2	6051.2	7.9	4.3	17.4
	Low Income	2612.5	1456.7	5786.5	12.2	6.8	27.0
$PM_{2.5}$	Chicago	12541.0	0.0	26110.8	30.2	0.0	63.0
	White	3490.7	0.0	7274.9	9.8	0.0	20.4
	Black	5504.2	0.0	11466.3	10.9	0.0	22.7
	Hispanic/ Latino	2798.4	0.0	5810.6	7.4	0.0	15.3
	Asian	621.6	0.0	1295.5	1.8	0.0	3.8
	High Income	4466.2	0.0	9302.1	12.7	0.0	26.6
	Low Income	4498.9	0.0	10425.3	16.8	0.0	35.1
$MDAO_3$	Chicago	31780.5	0.0	55492.0	77.6	0.0	135.3
	White	8997.4	0.0	15678.3	25.6	0.0	44.6
	Black	14055.0	0.0	24517.3	28.2	0.0	49.2
	Hispanic/ Latino	6792.9	0.0	11924.7	18.1	0.0	31.6
	Asian	1599.0	0.0	2788.4	4.7	0.0	8.2
	High Income	11403.0	0.0	19899.0	32.1	0.0	56.1
	Low Income	11121.0	0.0	19434.6	51.4	0.0	89.5

Table S3.3. Pediatric asthma hospitalizations from pollutants, divided by population subgroup.

Pollutant	Population	Pediatric ER	Pediatric ER (5CI)	Pediatric ER (95CI)	Pediatric ER RT	Pediatric ER RT (5CI)	Pediatric ER RT (95CI)
NO_2	Chicago	12.6	12.5	12.6	2.3	2.3	2.4
	White	1.8	1.8	1.8	0.4	0.4	0.4
	Black	5.9	5.9	5.9	1.2	1.2	1.2
	Hispanic/ Latino	4.2	4.1	4.2	0.6	0.6	0.6
	Asian	0.4	0.4	0.4	0.1	0.1	0.1
$PM_{2.5}$	Chicago	15.7	15.4	16.0	2.9	2.8	2.9
	White	2.2	2.2	2.3	0.5	0.5	0.5
	Black	7.5	7.3	7.6	1.5	1.4	1.5
	Hispanic /Latino	5.1	5.0	5.2	0.7	0.7	0.8
	Asian	0.5	0.5	0.5	0.1	0.1	0.1
$MDAO_3$	Chicago	44.5	44.3	44.6	8.2	8.1	8.2
	White	6.5	6.4	6.5	1.4	1.4	1.4
	Black	21.5	21.4	21.6	4.2	4.2	4.2
	Hispanic/ Latino	14.0	14.0	14.1	2.0	2.0	2.0
	Asian	1.5	1.5	1.5	0.3	0.3	0.3

Table S3.4. Difference (%) of average exposure, baseline mortality (BMR), adult asthma incidence rate (AIR), pediatric asthma hospitalization rate (PAIR), and attributable mortality rate, asthma rate, and pediatric asthma hospitalization rate for NO_2, $PM_{2.5}$, and $MDAO_3$ in quantiles, grouped by pollutant levels from lowest (Q1) to highest (Q2).

Pollutant	Quantile	Range	Exposure	BMR	AIR	PAIR	Mortality Rate	Asthma Rate	Ped. ER Rate
NO_2 (ppb)	Q1	8.8 - 14.2	-23.4	15.5	-4.0	11.4	-8.8	-20.4	-14.1
	Q2	14.2 - 15.9	-8.6	16.6	-2.2	4.1	7.6	-4.6	-4.0
	Q3	15.9 - 17.1	-0.7	-3.1	-1.9	-10.8	-2.2	-1.1	-10.6
	Q4	17.1 - 18.9	9.1	-3.1	7.6	-0.4	8.0	10.3	9.6
	Q5	18.9 - 27.9	23.6	-25.9	0.6	-4.4	-4.6	15.9	19.1
$PM_{2.5}$ ($\mu g/m^3$)	Q1	7.5 - 9.9	-13.0	3.2	-8.1	1.2	-8.1	-14.3	-11.8
	Q2	9.9 - 10.4	-2.6	7.5	-5.0	-2.7	3.4	-1.4	-5.1
	Q3	10.4 - 10.7	1.1	16.4	2.2	8.5	15.6	6.9	10.0
	Q4	10.7 - 11.1	4.7	-6.7	10.8	1.2	-1.3	4.6	6.0
	Q5	11.1 - 11.3	9.8	-20.3	-0.6	-8.2	-9.5	4.3	0.9
$MDAO_3$ (ppb)	Q1	35.1 - 40.4	-9.5	-19.9	5.1	0.1	-21.6	-10.2	-9.3
	Q2	40.4 - 41.9	-4.2	1.6	9.1	1.6	-1.0	-1.2	-2.7
	Q3	41.9 - 42.6	-0.3	3.2	-1.3	0.3	0.9	1.8	-0.3
	Q4	42.6 - 45.1	3.1	10.6	-3.9	-6.6	9.4	4.0	-3.7
	Q5	45.1 - 48.5	11.0	4.5	-8.9	4.6	12.4	5.7	15.9

Table S3.5. Exposure and mortality inequalities of the baseline (X_0) and when reducing pollution by half (X_n) in the high pollution, high BMR areas (census tracts with pollution >50%ile & BMR >50%ile). Exposure or mortality inequalities are computed by taking the % difference of the average pollution across Chicago, e.g., positive values mean the value is higher than Chicago average.

		Chicago Population Weighted Averages					Exposure Inequality		Mortality			Mortality Inequality		Mortality Rates		
Pollutant	Group	n	X_0	X_n	Δ (unit)	Δ (%)	X_0 (%)	X_n (%)	X_n	New	Δ (%)	X_0	X_n	X_0	X_n	Δ (%)
$PM_{2.5}$ ($\mu g/m^3$)	Black	168	10.3	8.5	1.8	17.2	-0.8	-9.8	2.3	1.9	18.2	18.2	8.3	0.9	0.8	19.3
	White	168	10.2	9.8	0.4	4.1	-1.1	4.2	1.8	1.7	5.7	-5.9	-0.7	0.5	0.5	6.2
	Hispanic/ Latino	168	10.7	9.6	1.1	10.2	3.3	1.8	1.8	1.5	12.8	-8.7	-10.9	0.6	0.5	14.3
	Asian	168	10.2	9.8	0.4	4	-1.5	3.8	1.9	1.8	4.3	-3.5	3.3	0.5	0.5	5.8
	High Income	6	10.4	10.1	0.3	2.9	0.5	7.2	1.9	1.7	11.4	0.6	-0.3	0.7	0.6	13.3
	Low Income	71	10.4	8.2	2.2	21.4	0.8	-13	1.9	1.7	11.4	0.6	-0.3	0.7	0.6	13.3
$MDAO_3$ (ppb)	Black	228	31.3	23.7	7.6	24.1	0.9	-10.5	8.3	6.2	26.1	18	9.3	3.4	2.5	25.2
	White	228	31.2	27.3	3.9	12.6	0.7	2.8	6.8	5.4	20.3	-3.4	-3.7	1.9	1.5	20.5
	Hispanic/ Latino	228	30.3	27.5	2.8	9.3	-2.2	3.7	6.1	5.2	14.4	-13.8	-7.6	2	1.7	13.3
	Asian	228	31.2	27.6	3.6	11.5	0.6	4	7	5.7	17.9	-0.8	2	2	1.6	17.8
	High Income	29	30.8	27.9	2.9	9.5	-0.8	5	7.1	5.6	20.9	0.5	-0.5	2.4	1.9	20.8
	Low Income	65	30.7	23.8	6.9	22.5	-0.9	-10.3	7.1	5.6	20.9	0.5	-0.5	2.4	1.9	20.8
NO_2 (ppb)	Black	157	15.6	13.1	2.6	16.6	-4.3	-12.7	1.2	1	16.9	15.1	6.2	0.5	0.4	19.4
	White	157	16.3	15.7	0.6	3.7	-0.2	5.1	0.9	0.9	5.1	-5.8	-0.8	0.3	0.3	6
	Hispanic/ Latino	157	17	15.7	1.3	7.7	3.7	4.7	0.9	0.8	9	-8.1	-7.2	0.3	0.3	11
	Asian	157	16.5	15.4	1.1	6.7	0.8	2.9	1	0.9	7.2	-1.3	1.7	0.3	0.3	10.5
	High Income	5	17.1	16.8	0.3	1.6	4.3	12.2	1	0.9	10.1	0	-0.1	0.3	0.3	12.8
	Low Income	70	16.4	12.6	3.8	22.9	0.3	-15.4	1	0.9	10.1	0	-0.1	0.3	0.3	12.8

Table S3.6. Population-weighted average concentrations and mortality changes over Chicago for reducing pollution in half where pollutants are high and BMR is high(>50%ile). Units for NO_2 and $MDAO_3$ are in ppb; $PM_{2.5}$ is $\mu g/m^3$; units for mortality rate is mortality per 100k residents.

	Average in Target Area			Population-weighted Average over Chicago			Avg. Mortality over Chicago			Avg. Mortality Rate over Chicago		
Pollutant	Original	New	Difference	Original	New	Difference	Original	New	Difference	Original	New	Difference
NO_2	18.2	9.1	50%	14.3	13.3	7%	19.6	17.3	12%	1.1	1.0	11%
$PM_{2.5}$	11.0	5.5	50%	9.4	8.9	5%	10.7	9.9	8%	0.6	0.5	9%
$MDAO_3$	32.6	16.3	50%	28.8	27.7	4%	11.2	10.6	6%	0.6	0.6	8%

Table S3.7. Exposure and mortality inequalities of the baseline (Original) and when reducing pollution by half (New) in the high pollution areas (census tracts with pollution >75%ile). Exposure or mortality inequalities are computed by taking the % difference of the average pollution across Chicago.

			Chicago Population Weighted Averages				Exposure Inequality		Mortality			Mortality Inequality		Mortality Rates		
Pollutant	Group	n	X_0	X_n	Δ (unit)	Δ (%)	X_0 (%)	X_n (%)	X_n	New	Δ (%)	X_0	X_n	X_0	X_n	Δ (%)
$PM_{2.5}$ ($\mu g/m^3$)	Black	200	10.3	9.5	0.8	7.8	-0.8	5.3	2.3	2.1	5.5	18.2	23.2	0.9	0.9	6.5
	White	200	10.2	8.9	1.3	12.7	1.1	-0.6	1.8	1.6	8.8	-5.9	-5.3	0.5	0.5	9.6
	Hispanic/ Latino	200	10.7	8.7	2.0	19.0	-3.3	-3.6	1.8	1.5	16.3	-8.7	-15.6	0.6	0.5	17.7
	Asian	200	10.2	8.9	1.3	12.8	1.5	-1.1	1.9	1.7	8.3	-3.5	-2.3	0.5	0.5	10.0
	High Income	43	10.4	8.7	1.7	16.4	-0.5	-3.2	1.9	1.8	9.3	0.6	0.7	0.7	0.6	10.1
	Low Income	43	10.4	9.1	1.4	13.1	-0.8	0.9	1.9	1.8	9.3	0.6	0.7	0.7	0.6	10.1
$MDAO_3$ (ppb)	Black	200	31.3	25.7	5.6	17.9	-0.9	-2.0	8.3	6.8	18.3	18.0	16.9	3.4	2.8	16.3
	White	200	31.2	25.8	5.5	17.5	-0.7	-1.7	6.8	5.5	19.9	-3.4	-6.2	1.9	1.6	18.3
	Hispanic/ Latino	200	30.3	28.2	2.1	7.0	2.2	7.6	6.1	5.6	8.5	-13.8	-4.4	2.0	1.8	7.8
	Asian	200	31.2	25.2	6.0	19.2	-0.6	-3.8	7.0	5.5	22.0	-0.8	-6.3	2.0	1.6	19.5
	High Income	41	30.8	25.7	5.1	16.5	0.8	-2.0	7.1	5.9	17.1	0.5	0.9	2.4	2.0	15.3
	Low Income	33	30.7	26.9	3.8	12.3	0.9	2.7	7.1	5.9	17.1	0.5	0.9	2.4	2.0	15.3
NO_2 (ppb)	Black	200	15.6	14.3	1.4	8.7	4.3	1.5	1.2	1.1	5.3	15.1	19.6	0.5	0.4	7.1
	White	200	16.3	13.9	2.4	14.9	0.2	-1.4	0.9	0.9	8.9	-5.8	-5.9	0.3	0.2	10.1
	Hispanic/ Latino	200	17.0	14.1	2.8	16.6	-3.7	0.5	0.9	0.8	12.2	-8.1	-11.5	0.3	0.3	14.4
	Asian	200	16.5	14.0	2.5	15.1	-0.8	-0.6	1.0	0.9	9.7	-1.3	-2.2	0.3	0.3	11.7
	High Income	52	17.1	13.8	3.3	19.3	-4.3	-2.2	1.0	0.9	8.5	0.0	0.4	0.3	0.3	9.9
	Low Income	45	16.4	14.1	2.3	14.3	-0.3	-0.1	1.0	0.9	8.5	0.0	0.4	0.3	0.3	9.9

Table S3.8. Population-weighted average concentrations and mortality changes over Chicago for reducing pollution in half where pollutants are high (>75%ile). Units for NO_2 and $MDAO_3$ are in ppb; $PM_{2.5}$ is $\mu g/m^3$; units for mortality rate is mortality per 100k residents.

	Average in Target Area			Population-weighted Average over Chicago			Avg. Mortality over Chicago			Avg. Mortality Rate over Chicago		
Pollutant	Original	New	Difference	Original	New	Difference	Original	New	Difference	Original	New	Difference
NO_2	20.0	10.0	50%	14.3	12.9	10%	19.6	17.7	10%	1.1	1.0	8%
$PM_{2.5}$	11.4	5.7	50%	9.4	8.3	12%	10.7	9.8	9%	0.6	0.5	9%
$MDAO_3$	33.9	17.0	50%	28.8	26.1	9%	11.2	10.5	6%	0.6	0.6	7%

Table S3.9. Exposure and mortality inequalities of the baseline (Original) and when reducing pollution by half (New) in the high BMR areas (census tracts with BMR >75%ile). Exposure or mortality inequalities are computed by taking the % difference of the average pollution across Chicago.

			Chicago Population Weighted Averages				**Exposure Inequality**		**Mortality**			**Mortality Inequality**		**Mortality Rates**		
Pollutant	Group	n	X_0	X_n	Δ (unit)	Δ (%)	X_0 (%)	X_n (%)	X_n	New	Δ (%)	X_0	X_n	X_0	X_n	Δ (%)
$PM_{2.5}$ ($\mu g/m^3$)	Black	200	10.3	7.7	2.6	25.3	-0.8	-18.2	2.3	1.6	30.9	18.2	-3.6	0.9	0.6	32.3
	White	200	10.2	9.8	0.4	4.4	-1.1	4.4	1.8	1.6	10.1	-5.9	-0.2	0.5	0.5	10.6
	Hispanic/ Latino	200	10.7	10.3	0.4	3.8	3.3	9.7	1.8	1.6	7.3	-8.7	-0.1	0.6	0.5	7.8
	Asian	200	10.2	9.8	0.4	4.3	-1.5	4.1	1.9	1.7	8.8	-3.5	3.9	0.5	0.5	10.0
	High Income	7	10.4	10.2	0.2	1.8	0.5	8.9	1.9	1.6	16.4	0.6	-0.8	0.7	0.5	18.9
	Low Income	95	10.4	7.5	3.0	28.4	0.8	-20.3	1.9	1.6	16.4	0.6	-0.8	0.7	0.5	18.9
$MDAO_3$ (ppb)	Black	200	31.3	23.4	7.9	25.1	0.9	-16.3	8.3	5.9	29.4	18.0	-1.9	3.4	2.3	30.7
	White	200	31.2	29.7	1.5	4.9	0.7	6.2	6.8	6.1	10.4	-3.4	1.8	1.9	1.7	10.9
	Hispanic/ Latino	200	30.3	29.1	1.3	4.2	-2.2	3.9	6.1	5.6	7.6	-13.8	-6.3	2.0	1.8	8.0
	Asian	200	31.2	29.7	1.5	4.7	0.6	6.3	7.0	6.4	8.9	-0.8	6.3	2.0	1.8	10.1
	High Income	7	30.8	30.2	0.6	2.0	-0.8	7.8	7.1	5.9	16.1	0.5	-0.8	2.4	1.9	18.4
	Low Income	95	30.7	21.9	8.8	28.7	-0.9	-21.8	7.1	5.9	16.1	0.5	-0.8	2.4	1.9	18.4
NO_2 (ppb)	Black	200	15.6	11.7	3.9	25.2	-4.3	-21.3	1.2	0.8	31.0	15.1	-6.7	0.5	0.3	32.6
	White	200	16.3	15.7	0.7	4.1	-0.2	5.2	0.9	0.9	9.6	-5.8	0.0	0.3	0.2	10.2
	Hispanic/ Latino	200	17.0	16.4	0.6	3.6	3.7	9.9	0.9	0.9	6.9	-8.1	0.5	0.3	0.3	7.5
	Asian	200	16.5	15.8	0.7	4.1	0.8	6.2	1.0	0.9	8.4	-1.3	6.2	0.3	0.3	9.7
	High Income	7	17.1	16.8	0.3	1.7	4.3	12.7	1.0	0.8	16.0	0.0	-1.3	0.3	0.3	18.6
	Low Income	95	16.4	11.8	4.6	28.1	0.3	-20.8	1.0	0.8	16.0	0.0	-1.3	0.3	0.3	18.6

Table S3.10. Population-weighted average concentrations and mortality changes over Chicago for reducing pollution in half where BMR is high (>75%ile). Units for NO_2 and $MDAO_3$ are in ppb; $PM_{2.5}$ is $\mu g/m^3$; units for mortality rate is mortality per 100k residents.

	Average in Target Area			Population-weighted Average over Chicago			Avg. Mortality over Chicago			Avg. Mortality Rate over Chicago		
Pollutant	Original	New	Difference	Original	New	Difference	Original	New	Difference	Original	New	Difference
NO_2	15.7	7.8	50%	14.3	13.1	8%	19.6	16.4	16%	1.1	0.9	19%
$PM_{2.5}$	10.3	5.4	50%	9.4	8.5	10%	10.7	8.8	18%	0.6	0.5	21%
$MDAO_3$	31.2	15.6	50%	28.8	26.0	10%	11.2	9.2	18%	0.6	0.5	21%

Appendix 4

Air quality and health implications of electrifying heavy-duty vehicles assessed at equity-relevant neighborhood-scales

Sara Camilleri, Anastasia Montgomery, Maxime Visa, Jordan Schnell, Zac Adelman, Mark Janssen, Emily Grubert, Susan Anenberg, Daniel Horton

Available as: Camilleri, S., Montgomery, A., Visa, M., Schnell, J., Adelman, Z., Janssen, M., Grubert, E., Anenberg, S. and Horton, D., 2023. Air quality and health implications of electrifying heavy-duty vehicles assessed at equity-relevant neighborhood-scales. (2023). *Nature Sustainability* (Accepted, as of 08/01/2023).

Abstract

Heavy-duty vehicles (HDVs) disproportionately contribute to the creation of air pollutants and emission of greenhouse gases – with marginalized populations unequally burdened by the impacts of each. Shifting to non-emitting technologies, like electric HDVs (eHDVs) is underway, however, the associated air quality and health implications have not been resolved at equity-relevant scales. Here, we use a neighborhood-scale (~1km) air quality model to evaluate air pollution, public health, and equity implications of a 30% transition of predominantly diesel HDVs to eHDVs over the region surrounding North America's largest freight hub, Chicago, Illinois. We find decreases in NO_2 and $PM_{2.5}$ but O_3 increases, particularly in urban settings. NO_2 and $PM_{2.5}$ decreases reduce premature deaths/yr by ~580 and ~70, respectively, while O3 increases add ~50 deaths/yr. We find the largest pollutant and health benefits in "least White" communities, highlighting the potential for eHDVs to reduce air pollution and health burdens, especially in marginalized communities.

Appendix 5

Neighborhood-scale air quality, public health, and equity implications of multi-modal vehicle electrification

Maxime, Visa Sara F. Camilleri, **Anastasia Montgomery**, Jordan L. Schnell, Mark Janssen, Zachariah E. Adelman, Susan C. Anenberg, Emily A. Grubert, and Daniel E. Horton.

Available as: Visa, M., Camilleri, S.F., Montgomery, A., Schnell, J.L., Janssen, M., Adelman, Z.E., Anenberg, S.C., Grubert, E.A. and Horton, D.E., 2023. Neighborhood-scale air quality, public health, and equity implications of multi-modal vehicle electrification. *Environmental Research: Infrastructure* (Under review, 08/01/2023).

Abstract

Electric vehicles (EVs) constitute just a fraction of the current U.S. transportation fleet; however, EV market-share is surging. EV adoption reduces on-road transportation greenhouse gas emissions by decoupling transportation services from petroleum, but impacts on air quality and public health depend on the nature and location of vehicle usage and electricity generation. Here, we use a regulatory-grade chemical transport model and an electricity dispatch algorithm to characterize neighborhood-scale (~1 km) air quality and public health benefits and tradeoffs associated with a multi-modal EV transition. We focus on a Chicago-centric regional domain wherein 30% of the on-road transportation fleet is instantaneously electrified and changes in on-road, refueling, and power plant emissions are considered. We find decreases in annual population-weighted domain mean NO_2 (-11.84%) and $PM_{2.5}$ (-2.56%) with concentration reductions of up to-5.1 ppb and-0.97 µg m-3 in urban cores. Conversely, annual population-weighted domain mean $MDA8O_3$ concentrations increase +0.65%, with notable intra-urban changes of up to +2.3 ppb. Despite mixed pollutant concentration outcomes, we find overall positive public health

outcomes, largely driven by NO_2 decreases that produce mortality reductions that are ~5 times greater in census tracts with disproportionately large non-white populations.

Appendix 6

Potential of breadfruit cultivation to contribute to climate-resilient low latitude food systems

Lucy Yang, Nyree Zerega, **Anastasia Montgomery**, Daniel E. Horton

Available as: Yang, L., Zerega, N., Montgomery, A., & Horton, D. E. (2022). Potential of breadfruit cultivation to contribute to climate-resilient low latitude food systems. *PLoS Climate*, *1*(8), e0000062. https://doi.org/10.1371/journal.pclm.0000062

Abstract

The number of people in food crisis around the world is increasing, exacerbated by COVID-19, conflict, and climate change. Major crop yields are projected to decrease in low-latitude regions, making tropical and sub-tropical food systems particularly vulnerable. Increased cultivation of breadfruit (Artocarpus altilis), a neglected and underutilized species (NUS), has the potential to enhance climate resilience and overall sustainability of low-latitude agricultural systems. To better understand breadfruit's cultivation suitability and geographic range in current and future climates, we use breadfruit presence data collected from previous studies and a global citizen science database, and a selection of bioclimactic variables, to build an ensemble of 6 species distribution models that delineate the current climatically viable breadfruit range. We then assess the climatically viable future breadfruit range (2061–2080) under stabilization and high emission scenarios using an ensemble of 8 global circulation model (GCM) projections. The area of suitable breadfruit range within the global tropics and subtropics is projected to decrease ~4.4% in the stabilization scenario and ~4.5% in the high emission scenario. In Southeast Asia and the Pacific Islands, yield quality and consistency show minimal decreases under the high emission scenario, with increases in total suitable area under both. In contrast, in Latin America and the Caribbean, the current suitable

breadfruit range is projected to contract ~10.1–11.5% (stabilization-high emission). Present and future model suitability outputs suggest opportunities to successfully expand breadfruit cultivation over the next decades in sub-Saharan Africa, where food insecurity is coincidentally high. However, in all regions, high emission scenario conditions reduce the overall consistency and quality of breadfruit yields compared to the stabilization scenario. Our results have the potential to inform global food security adaptation planning, highlighting breadfruit as an ideal NUS to incorporate in food security adaptation strategies.

Appendix 7

The COVID-19 lockdowns: a window into the Earth System

Noah S Diffenbaugh, Christopher B Field, Eric A Appel, Ines L Azevedo, Dennis D Baldocchi, Marshall Burke, Jennifer A Burney, Philippe Ciais, Steven J Davis, Arlene M Fiore, Sarah M Fletcher, Thomas W Hertel, Daniel E Horton, Solomon M Hsiang, Robert B Jackson, Xiaomeng Jin, Margaret Levi, David B Lobell, Galen A McKinley, Frances C Moore, **Anastasia Montgomery**, Kari C Nadeau, Diane E Pataki, James T Randerson, Markus Reichstein, Jordan L Schnell, Sonia I Seneviratne, Deepti Singh, Allison L Steiner, Gabrielle Wong-Parodi

Available as: Diffenbaugh, Noah S., et al. "The COVID-19 lockdowns: a window into the Earth System." *Nature Reviews Earth & Environment* (2020): 470-481. https://doi.org/10.1038/s43017-020-0079-1

Abstract

Restrictions to reduce human interaction have helped to avoid greater suffering and death from the COVID-19 pandemic, but have also created socioeconomic hardship. This disruption is unprecedented in the modern era of global observing networks, pervasive sensing and large-scale tracking of human mobility and behaviour, creating a unique test bed for understanding the Earth System. In this Perspective, we hypothesize the immediate and long-term Earth System responses to COVID-19 along two multidisciplinary cascades: energy, emissions, climate and air quality; and poverty, globalization, food and biodiversity. While short-term impacts are dominated by direct effects arising from reduced human activity, longer-lasting impacts are likely to result from cascading effects of the economic recession on global poverty, green investment and human behaviour. These impacts offer the opportunity for novel insight, particularly with the careful deployment of targeted data collection, coordinated model experiments and solution-oriented randomized controlled trials, during and after the pandemic.

Appendix 8

Coronavirus disease 2019 (COVID-19) mortality and neighborhood characteristics in Chicago

Molly Scannell Bryan, Jiehuan Sun, Jyotsna Jagai PhD, Daniel E. Horton, Anastasia Montgomery, Robert Sargis, Maria Argos

Available as: Bryan, M. S., Sun, J., Jagai, J., Horton, D. E., Montgomery, A., Sargis, R., & Argos, M. (2021). Coronavirus disease 2019 (COVID-19) mortality and neighborhood characteristics in Chicago. *Annals of Epidemiology*, 56, 47-54. https://doi.org/10.1016/j.annepidem.2020.10.011

Abstract

Purpose: To describe coronavirus disease 2019 (COVID-19) mortality in Chicago during the spring of 2020 and identify at the census-tract level neighborhood characteristics that were associated with higher COVID-19 mortality rates.

Methods: Using Poisson regression and regularized linear regression (elastic net), we evaluated the association between neighborhood characteristics and COVID-19 mortality rates in Chicago through July 22 (2514 deaths across 795 populated census tracts).

Results: Black residents (31% of the population) accounted for 42% of COVID-19 deaths. Deaths among Hispanic/Latino residents occurred at a younger age (63 years, compared with 71 for white residents). Regarding residential setting, 52% of deaths among white residents occurred inside nursing homes, compared with 35% of deaths among black residents and 17% among Hispanic/Latino residents. Higher COVID-19 mortality was seen in neighborhoods with heightened barriers to social distancing and low health insurance coverage. Neighborhoods with a higher percentage of white and Asian residents had

lower COVID-19 mortality. The associations differed by race, suggesting that neighborhood context may be most tightly linked to COVID-19 mortality among white residents.

Conclusions: We describe communities that may benefit from supportive services and identify traits of communities that may benefit from targeted campaigns for prevention and testing to prevent future deaths from COVID-19.