Understanding and addressing exposure inequality in ambient air pollution

Yuzhou Wang

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

Julian D. Marshall, Chair

Timothy V. Larson

Lianne Sheppard

Program Authorized to Offer Degree:

Civil and Environment Engineering

University of Washington

Abstract

Understanding and addressing exposure inequality in ambient air pollution

Yuzhou Wang

Chair of the Supervisory Committee:

Julian D. Marshall

Department of Civil and Environmental Engineering

Air pollution is the largest environmental risk factor in the world, causing $\sim$ 6.5 million deaths per year. However, people don't breathe the same air. Air pollution exposures and associated health impact are unevenly distributed across countries, regions, communities, and individuals. Between-country inequalities in ambient air pollution indicate that people living in low- and middle-income countries disproportionately experience the higher pollution exposures and larger burdens of ambient air pollution. Within-country inequalities, including how ambient pollution levels correlate with socio-demographic attributes, are poorly studied other than in the US and a few other high-income countries. Most literature in the US indicates higher air pollution exposures for people of color and low-income populations. During the past decades, the air has gotten cleaner in the US, however, the disparities persist. While ample studies have documented patterns of air pollution exposure inequality in the US, there is almost no scientific literature that explore possible solutions and policies to eliminate the systemic and long-standing inequalities. In addition, inequality patterns in the US may or may not be applicable to other countries, owing to the different historical, social, economic, political, urbanization characteristics.

This dissertation consists of four original studies of ambient air pollution concentration and

exposure inequality (**Chapters 2 to 5**), plus an introduction to these topics (**Chapter 1**) and a summary of findings with potential implications for future research and for policy (**Chapter 6**). There are three main objectives: (1) quantify spatial sources of ambient air pollution (**Chapter 2**); (2) explore exposure inequality patterns in countries other than US (specifically, in China; **Chapter 3**); and (3) investigate policies/approaches to address the persistent exposure inequalities in the US (**Chapters 4 and 5**). I focus on two important air pollutants: fine particulate matter ($PM_{2.5}$), and nitrogen dioxide (NO_2). The approaches to estimate air pollution concentration include both empirical and mechanistic air quality models.

Chapter 1 provides relevant background on air pollution distribution and exposure inequality, highlights the motivation, objectives, approaches, and structure of this dissertation.

Chapter 2 quantifies the spatial sources of ambient NO_2 and $PM_{2.5}$ concentrations. I develop a readily scalable algorithm based on "spatial-increment" to decompose air pollution concentrations into four spatial components: long-range, mid-range, neighborhood, and near-source. I apply the algorithm to the 2010-2015 annual-averaged concentrations from empirical predictions for all census blocks in the contiguous US. I find that NO_2 is of urban origin and varies by urbanicity; $\sim$90% of the concentration differences are driven by "neighborhood" and "mid-range" components; climate or geographic regions have less effects on the NO_2 concentrations. In contrast, $PM_{2.5}$ is a regional pollutant with a strong secondary component; the concentrations are dominated by "long-range" components (>50% in most geographic regions) and vary at state and regional level; urbanicity has modest effects on $PM_{2.5}$ concentrations and minor effects on concentration differences.

Chapter 3 quantifies the relationship between ambient air pollution exposure and socioeconomic status (SES) in China. I combine estimated year 2015 annual-average ambient levels of NO_2 and $PM_{2.5}$ from empirical models with national demographic information, which is derived from both China Health and Retirement Longitudinal Study (CHARLS) cohort and gridded gross domestic product (GDP). I find that in contrast to the typical patterns for the US, ambient air pollution concentrations in China are higher for higher-SES populations and communities than for lower-SES populations, and higher for long-standing urban residents than for rural-to-urban migrant populations. The positive relationship holds for different SES measurements (individual SES score, community-averaged SES score, gridded GDP per capita), in rural and urban locations,

across geographic regions, across a wide range of spatial resolution from 1-100 km, and for modeled vs. measured pollution concentrations. Exposure inequalities are higher for NO_2 than $PM_{2.5}$. My findings are consistent with the idea that in China's current industrialization and urbanization stage, economic development is positively correlated with both SES and air pollution levels.

Chapters 4 and 5 investigates approaches and policy scenarios to address US national racial-ethnic inequalities of air pollution exposure. I use InMAP (Intervention Model for Air Pollution) Source-Receptor Matrix (ISRM) to predict how changes in emissions impact annual-average $PM_{2.5}$ concentrations and exposure inequalities. **Chapter 4** investigates three emission-reduction approaches, and compare their optimal ability to address both average exposure for the overall population and racial-ethnic exposure inequalities. I find that US national inequalities in exposure can be eliminated with minor emission-reductions (optimal: $\sim$1% of total emissions) if targeting specific locations. In contrast, achieving that outcome using existing regulatory strategies would require eliminating essentially all emissions (if targeting specific economic sectors) or is not possible (if requiring urban regions to meet concentration standards). In addition, there is no tradeoff between reducing overall average and reducing national inequalities; rather, the approach that does the best for reducing national inequalities (i.e., location-specific strategies) also does as well as or better than the other two approach (i.e., sectors-specific; meeting concentration standards) for reducing overall averages.

Chapter 5 expands the investigation in **Chapter 4** to a specific government environmental justice (EJ) policy – the Biden Administration's Justice40 Initiative ("J40"), which is a general policy to address environmental injustice. Climate and Economic Justice Screening Tool (CEJST) is the signature element and ongoing approach of J40. I investigate whether emission-reductions brought about by CEJST/J40 investments will eliminate disparities in $PM_{2.5}$ exposure by race-ethnicity and other attributes in 20 years, through comparing a Business As Usual ("BAU") scenario against two scenarios wherein CEJST-identified locations ("J40 communities") experience accelerated emission-reductions. BAU simply continues historical rates of emissions and emission-changes into the future; in the two CEJST scenarios, I double or quadruple emission reduction in J40 communities, relative to BAU. I find that under BAU scenario, disparities remain in place. The two accelerated scenarios will only eliminate inequalities for J40 communities and for low-income populations in 20 years; yet

they do not reduce relative disparities by race-ethnicity. The results indicate that additional and more targeted actions, beyond CEJST/J40, will be needed to end racial-ethnic exposure disparities in the next decades.

Chapter 6 summarizes the findings of the studies presented in **Chapters 2 to 5** and discusses implications for future research and policy. Overall, these studies provide new knowledge and insights for air pollution exposure inequalities by: (1) extending the EJ knowledge to China – indicating a consistent positive relationship between SES and ambient air pollution exposures; and (2) advancing the EJ knowledge in the US – implying a possible general solution to address the persistent racial-ethnic inequalities in the US, and providing a useful example of regulatory impact analyses for assessing the effectiveness of the EJ policies.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I would not have been able to reach this milestone without the support of many people. My deepest appreciation goes to my advisor and role model, Dr. Julian D. Marshall, for his continued support and guidance in my research. I am grateful to him for his patience, enthusiasm, immense knowledge, insightful suggestions on my research, and invaluable career advice. I've learned a lot from him how to be a better researcher and am always inspired by his personality. I could not have imagined having a better advisor for my PhD study. I sincerely thank my committee members, Dr. Timothy V. Larson and Dr. Lianne Sheppard, and Graduate School Representative, Dr. Alex Turner, for their continued support and constructive advice on my dissertation.

I also want to thank my collaborators on my research papers: Joshua S. Apte, Jason D. Hill, Cesunica E. Ivey, Regan F. Patterson, Allen L. Robinson, Christopher W. Tessum, Dana Johnson, Esther Min, Rachel Morello-Frosch, Yafeng Wang, Hao Xu, Yaohui Zhao, Matthew Bechle, Sun-Young Kim, Peter J. Adams, Spyros N. Pandis, C. Arden Pope III, Lianne Sheppard, and Adam A. Szpiro, for their creative ideas, insightful discussions, and great contributions to my research papers. They are all excellent researchers. I really learned a lot from them and have been lucky to have them collaborating on my work.

I would also like to thank my colleagues, former colleagues, and friends in the Marshall Research Group: Jiawen Liu, Bujin Bekbulat, Lara Clark, Matthew Bechle, Maninder Thind, Christopher Tessum, Fang Guo, Hao Xu, Ying Gao, Beidi Diao, Songlin Yang, and Yifan Wen for creating a healthy and supportive environment in the group and making my graduate student life a lot more fun. I would also like to thank my colleagues Bujin Bekbulat and Libby Koolik for our collaboration on developing new InMAP/ISRM and making training sessions for InMAP/ISRM.

I would also like to acknowledge the funding agencies that funded my research work: Center for Air, Climate, and Energy Solution (CACES), which was supported under Assistance Agreement Number R835873 awarded by the United States Environmental Protection Agency; and Bay Area Air Quality Management District.

Finally, I could not express my appreciation enough to my family. I am always grateful to my parents for their unconditional love and support. Special thanks go to my husband, Zhuoran Zhang. He is always supportive, and has accompanied me through so many happiness moments and challenging periods during my PhD life.

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

This dissertation explores inequalities in exposure to ambient air pollution, and investigates possible solutions to address those inequalities. This chapter provides background information on these topics, highlights the motivation and objectives of my research works, introduces general approaches applied in, and outlines the structure of remaining chapters.

1.1 Background

1.1.1 Air pollution and its distribution

Air pollution is the largest environmental risk factor in the world, causing $\sim$ 6.5 million deaths per year ($\sim$ 4 million are from ambient particulate matter [PM]) [1–3]. Air pollution is associated with many acute and chronic health impacts, including respiratory disease, cardiovascular disease, cancers, chronic obstructive pulmonary disease, adverse pregnancy outcomes, cognitive impairment and neurological diseases, and others [1, 2, 4].

Sources of air pollution consist of "natural" (e.g., biogenic sources, dust) and anthropogenic sources, which include point (e.g., power plants, industrial sites), area (e.g., agricultural fields), and mobile sources (e.g., gasoline vehicles). Air pollution can be either directly emitted to the atmosphere from an emission source (primary air pollutants, e.g., primary PM, black carbon [BC], sulphur dioxide [SO_2]), or formed in the atmosphere from precursor gases through chemical reactions and microphysical processes (secondary air pollutants, e.g., ozone [O_3], secondary PM, nitrogen dioxide [NO_2]).

Air pollution concentrations in the environment, which can vary greatly within a region or a city, reflect spatial patterns in emissions, physical and chemical processes governing production and removal of pollutants, and turbulent advection and dispersion that drives transport and dilution. Concentrations of ambient air pollution vary on length scales from meters, to 100s or 1000s of km, affected by time, location, and pollutant. The spatial variability of air pollution concentrations at a location can reveal information on pollution sources. For example, spatially homogenous

concentrations suggest dominance of regional sources and/or secondary pollutants (e.g., sulfate fine PM [PM $\leq$2.5 μm in aerodynamic diameter, "$PM_{2.5}$"] formed from SO_2 emissions from coal power plants, ammonium $PM_{2.5}$ formed from ammonia [NH_3] emissions from agricultural fields), whereas heterogeneities on short length-scales suggest influence by local emission sources (e.g., nitrogen oxides [NO_x] emissions from traffic in an urban high-way, BC emissions from a brick production factory).

1.1.2 Air pollution exposure inequality

People don't breathe the same air. Air pollution exposures and associated health impacts are unevenly distributed across countries, regions, communities, and individuals from different socio-demographic groups. Between-country disparities in ambient air pollution indicate that people living in low- and middle-income countries (especially in East and South Asia) disproportionately experience the higher pollution exposures and larger health burdens (89% of the premature deaths occur in these areas) of ambient air pollution [1, 5, 6]. Within-country disparities, including how ambient pollution levels correlate with socio-demographic attributes, are poorly studied other than in the United States and a few other high-income countries. As described next, most studies for the US indicates higher pollution levels in disadvantaged communities with less political power to resist polluting land-uses [7–12]; however, this pattern may not be applicable to other countries which have different historical, social, economic, and industrialization characteristics [13–15].

Air pollution exposure inequality patterns are influenced by the spatial scales of air pollution and population data due to spatial aggregation problem [16–19]. For example, a coarser-resolution analysis may smooth out the spatial clustering of population and pollution, causing modeled pollution to be artificially (and incorrectly) diluted away from nearby populations, which likely results in the underestimation of inequalities among population groups [16, 17].

1.1.3 Air pollution exposure inequality in the US

The studies about inequalities in air pollution exposure in the US began in the 1970s [20,21]. In the 1990s, following the rise of the environmental justice movement, an emerging literature began documenting the air pollution disparities by race and income [12, 20, 22].

The US Environmental Protection Agency (EPA) defines environmental justice (EJ) as "the fair treatment and meaningful involvement of all people regardless of race, color, national origin,

or income, with respect to the development, implementation, and enforcement of environmental laws, regulations, and policies." [23] That definition focuses mainly on process, i.e., on the treatment of people and their involvement in regulatory and governmental processes. Another aspect, in addition to process, is outcomes (e.g., concentrations, risk). My work focuses on outcomes (concentrations). In this dissertation, EJ refers to avoiding a situation in which a specific social group is disproportionately affected by environmental hazards.

The vast majority of the existing literature for the US indicates higher air pollution levels for communities of color and populations with lower socioeconomic status (SES) [7–12, 24], a pattern that holds for nearly all emission sources and all U.S. states ([10, 25, 26]). Disparities by race-ethnicity are larger than, and distinct from, those by income [8, 10, 20, 27, 28]. Exposure disparities often co-occur with disparities in other social stressors, which can exacerbate community vulnerability and susceptibility [29].

During the past decades, the Clean Air Act has been effective at reducing average concentrations and absolute exposure disparities [30], but relative disparities are generally ignored and have persisted [8, 10, 28, 31–33]. Disadvantaged communities (especially the people of color communities) with the largest air pollution burdens decades ago still experience the highest air pollution exposures today [32].

Racial-ethnic inequalities in US ambient air pollution and subsequent exposures are attributable to market dynamics, sociopolitical reasons, and racist planning [34]. The market dynamics explanations include two parts: (1) new pollution facilities tend to be located in the places where lands are cheap and industrial labor pools are nearby, and those locations are more likely to be near the communities of color; (2) racial-ethnic and socioeconomic compositions of the nearby neighborhoods may also significantly change after the facility has been sited, which further exacerbates the inequalities [34]. Sociopolitical explanations include that the disadvantaged communities lack social capital and political power/access, and thus are less able to stop locally undesirable land-uses such as highways, industrial facilities, and other sources of pollution [9, 11, 20, 25, 31, 35]. In addition to these two explanations, another important cause is racist planning, including (1) historical, race-based housing segregation and (2) racist land-use practices [24, 35–44], which has led to the disproportionate siting of multiple pollution sources in communities of color and low-income communities [28, 35, 37, 38, 45, 46]. This pattern continues today: on average, polluting land-uses are more likely to be located near communities with less political power to resist their placement [42].

The existing literature has documented exposure inequities [7–11, 27, 28, 32, 33, 47] and investigated the impacts on inequities of emission changes for specific sources (e.g., industrial facilities, electricity generation units, on-road traffic, etc. [48–57]) or locations [12, 58–62]. However, there is almost no scientific literature on how to eliminate air pollution exposure inequalities [25, 63]. That aspect is an important research gap in the literature [64, 65].

1.1.4 Air pollution exposure inequality in China and other countries

While disparities are well documented in the US, limited knowledge exists for China or other low- or middle-income countries [9], where 91% of the premature deaths from outdoor air pollution occur (26% are in China) [1,6]. From 1990 to 2013, the population-weighted average ambient $PM_{2.5}$ concentration in China increased 38% (from 39 $\mu g/m^3$ to 55 $\mu g/m^3$), which is almost twice the global averaged value (20.4%) [5]. Less than 20% of the Chinese population lives in cities that meet the national annual $PM_{2.5}$ standard (the GB 3095-2012 standard; 35 $\mu g/m^3$) [66], and none lives in cities that meet the World Health Organization's (WHO) annual guideline (5 $\mu g/m^3$) [5, 67–69].

As mentioned above, the environmental inequality patterns in the US are attributed in part to past and present racial discrimination and race- and class-based market dynamics. However, those theories developed in the US may not be applicable for China and elsewhere owing to the different historical, social, economic, political, urbanization, and industrialization characteristics [13–15]. For example, studies have found that environmental inequality patterns in European countries imply a mixed relationship and are different from the US, owing to the lower extent of social segregation and the greater tendency (relative to in the United States) for lower-SES groups to live in the outskirts of the city [70]. Urban expansion patterns [71], local housing policies [72], and discrete choices between benefits/amenities and negative aspects of the environment [73] are also found to influence the exposure inequality patterns in Europe. One study in India proposed an environmental Kuznets curve (EKC; i.e., a theory that environmental degradation first rises and then falls with an increasing per capita income [74–76]) -like pattern (i.e., a inverted U-shaped relationship): the low-SES populations are more likely to be exposed to coal pollution; however, the extremely marginalized communities are less exposed to coal pollution because they are excluded from both negative and positive externalities of industrial development [77].

Previous studies in China have investigated air pollution inequality with respect to specific sources, including industrial emissions [15,78,79], emissions from electric vehicles [56], and household

consumption [80]. Those studies mostly indicate that people with higher income tend to have higher health burdens from air pollution, which is reverse of the patterns in the US. Studies on ambient concentrations [13, 81–83] have mostly focused on a single city (mostly for big cities such as Beijing and Shenzhen) or yielded inconclusive results [84–86]. The lacking of national and comprehensive studies in China is in part attributable to the lack of fine-resolution demographic data (e.g., publicly available census data for China is at the county level [average per county: 0.4 million people; 2,642 km^2]) and pollution concentration data, which limits the potential for some census-based demographic analyses of air pollution in China — analyses that, in the United States, are common.

The lack of comprehensive research on environmental inequality in China and other low- and middle-income countries, including theories and national empirical studies of SES and ambient air pollution, is another gap in the literature [9].

1.2 Goals and objectives

As mentioned in the previous sections, there are two important research gaps in the literature on air pollution exposure inequality:

- Most of the studies are for the US. There is limited knowledge for low- and middle-income countries, where there are higher air pollution levels and larger health burdens. The patterns and theories in the US may or may not be applicable to those countries, owing to the different historical, social, economic, political, urbanization characteristics.

- Literature in the US mostly documents the city-wide or national patterns of racial-ethnic exposure inequality, yet there is almost no scientific literature investigating possible solutions and policies to eliminate the systemic and long-standing inequalities in air pollution.

This dissertation aims to address these two knowledge gaps in the environmental inequality literature by: (1) advancing the understanding of air pollution exposure patterns in a country other than the US (specifically, in China); and (2) investigating possible approaches/policies to address the persistent inequalities in exposure to ambient air pollution in the US.

More specifically, the main objectives are as follows:

1. Objective #1: Quantify the spatial sources of ambient NO_2 and $PM_{2.5}$ for the contiguous US (**Chapter 2**).

2. Objective #2: Quantify the relationships between ambient NO_2 and $PM_{2.5}$ air pollution exposures and SES in China (**Chapter 3**).

3. Objective #3: Investigate policies to address US national racial-ethnic disparities in exposure to $PM_{2.5}$ air pollution (**Chapters 4 and 5**).

Objective #1 focuses on identifying and the spatial patterns of two important criteria air pollutants and quantifying their spatial sources. The broader goal of this objective is to understand where these two air pollutants come from and to target the emission sources that cause the pollution and the exposure inequalities. Objective #2 focuses on understanding patterns of air pollution exposure inequality in China; the broader goal is to shedding new light on the understanding of the environmental inequality patterns outside the US. Objective # 3 focuses on policies and technology scenarios to address the exposure inequalities in the US; the broader goal is to inform a national conversation about addressing environmental injustice in the US.

I focus on NO_2 and $PM_{2.5}$ because they are important and widely tracked. Both are criteria pollutants defined by US EPA, and both have important health effects associated with exposures (e.g., [4, 87–89]). To quantify the air pollution concentrations and population exposures, I employ both empirical (for NO_2 and $PM_{2.5}$) and mechanistic (for $PM_{2.5}$) air quality modeling approaches. The details of air quality modeling approaches are described next.

1.3 Approach

1.3.1 Land-use regression models

Land-use regression (LUR) models are empirical models that use statistical techniques (e.g., regression models) to combine (1) in situ measurement of air pollution concentrations and (2) geographic variables. LUR models aim to predict concentrations at non-measurement locations. Geographic variables often include variables corresponding to emission sources (e.g., population density, traffic network, land-use type) and reflecting the dispersion conditions (e.g., meteorology, elevation, land-cover) [68]. Some recent LURs also incorporate microscale variables capturing the hyper-local features (e.g., street view imagery, points of interest, local climate zones) [90] and air pollution estimates from mechanistic models [91, 92], satellite-based measurements such as Aerosol Optical Depth (AOD) (for $PM_{2.5}$) and tropospheric NO_2 column abundance [68, 93, 94] and mobile

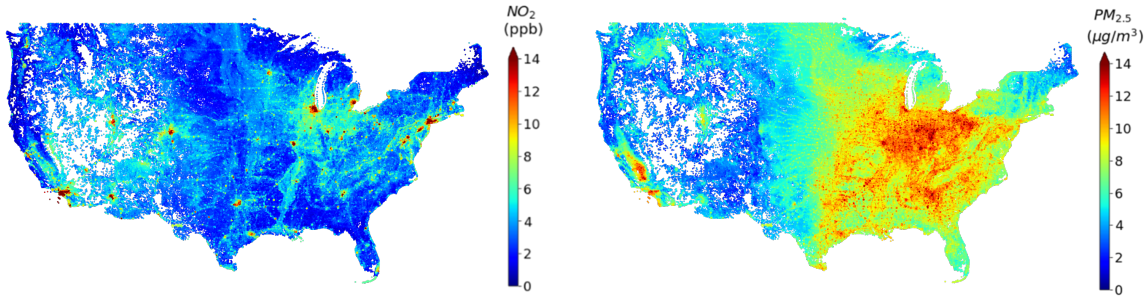

Figure 1.1: Annual-average NO_2 and $PM_{2.5}$ predictions from year-2010 US national LUR models

sampling [95–97]. LURs improve the accuracy of air pollution exposure estimations and are able to capture the within-urban variability of ambient air pollutants at high spatial (e.g., $\leq$1 km) (and, in some case, temporal [e.g., daily]) resolution. LURs have been applied over broad geographic regions for many locations, ranging from local [95–97], to regional/national [68, 90, 93, 98–102], and to global [103] scales.

In this dissertation, I employ the predictions from national LURs in both US (**Chapter 2**) and China (**Chapter 3**) to estimate the ambient concentrations and exposures of NO_2 and $PM_{2.5}$ air pollution [68,93]. Both models incorporate ground-level monitoring data; satellite-derived estimates of ambient concentrations and land-cover; land-uses and other geographic datasets (such as traffic, meteorological data, etc.) in a universal Kriging framework. The national LUR models in the US predict annual average concentrations at census block level (average size: $\sim$0.73 km^2 land area for blocks in the contiguous US; $\sim$53 people for the blocks with non-zero population) for six air pollutants ($PM_{2.5}$, NO_2, O_3, SO_2, PM_{10}, CO) and multiple-years (1979–2015) [93]. The data is freely available online (`https://www.caces.us`). Fig. 1.1 shows the year-2010 national prediction results for NO_2 and $PM_{2.5}$ concentrations for the contiguous US. The national LUR models in China predict annual-average concentrations at 1 km^2 resolution (9.6 million cells) for years 2014 and 2015 [68]. Fig. 1.2 shows the year-2015 national prediction results for ambient NO_2 and $PM_{2.5}$ concentrations in China. The models for US and China have good predictive performance (10-fold cross-validation R^2: 0.84 [NO_2], 0.85 [$PM_{2.5}$] for year-2010 US predictions; 0.78 [NO_2], 0.89 [$PM_{2.5}$] for year-2015 China predictions).

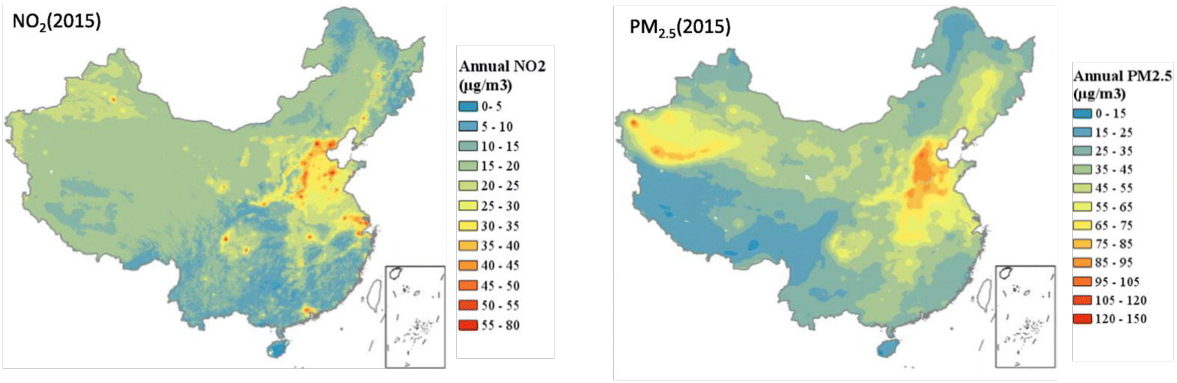

Figure 1.2: Annual-average NO_2 and $PM_{2.5}$ predictions from year-2015 China national LUR models. This figure is adapted from the graphical abstract of Xu et al. 2019 [68].

1.3.2 Chemical transport models

Chemical transport models (CTMs) are three-dimensional (3-D) mechanistic models which use mass balance principles and account for emissions, transport and dispersion by winds, chemical transformations and atmospheric removal processes. CTMs include Eulerian and Lagrangian models. Eulerian CTMs are "box" models simulating the species concentrations using fixed 3-D computational cells, and are commonly used for $PM_{2.5}$. Examples of Eulerian CTMs include Comprehensive Air Quality Model with Extensions (CAMx) [104], the Community Multiscale Air Quality model (CMAQ) [105], the Weather Research and Forecasting model coupled with Chemistry (WRF-Chem) [106], the Goddard Earth Observing System model coupled with Chemistry (GEOS-Chem) [107]. Lagrangian CTMs simulate the production and motion of parcels of air using a moving frame of reference. CTMs represent state-of-the-science air quality models and provide the most robust mechanistic-model estimates of $PM_{2.5}$ air pollution concentrations. However, CTMs are often complicated and cumbersome to use (i.e., computationally expensive; time-consuming to learn and to use; coarse spatial resolution).

To improve the availability and accessibility of CTMs, a set of reduced-complexity models (RCMs) have been developed, which employ more simplified representations of the complex physical and chemical mechanisms that link emissions to concentrations [108]. RCMs often are used as a type of integrated assessment models (IAMs). Although the concentration predictions from RCMs are not as accurate as the golden-standard full-form CTMs, the reduced computational costs (often

in magnitudes) for RCMs make it possible to conduct a series of sensitivity analyses, estimate uncertainties, and explore policy and technology scenarios.

The commonly used RCMs for $PM_{2.5}$ in the US include Intervention Model for Air Pollution (InMAP) [109]; Air Pollution Emission Experiments and Policy Analysis (APEEP) [110] and the updated versions, called AP2 and AP3 [111, 112]; and Estimating Air pollution Social Impact Using Regression (EASIUR) [113]. InMAP leverages pre-processed physical and chemical information from the full-form CTM to calculate annual average marginal changes in primary (i.e., directly-emitted) $PM_{2.5}$ and secondary (chemically formed) $PM_{2.5}$ (NO_x, sulfur oxides [SO_x], NH_3, and volatile organic compound [VOC]) caused by marginal changes in emissions. InMAP employs a variable spatial resolution based on the function of the gradient in population density and pollutant concentration; the resolution is higher in urban areas and lower in rural locations and in high altitude [109]. APEEP employs Gaussian modeling augmented with reduced-form chemistry modules and maps the emission with $PM_{2.5}$ concentration through source-receptor matrices for each $PM_{2.5}$ precursors; the spatial resolution is at the county level [108]. EASIUR employs a regression-based approximation to CTM output to estimate the marginal social costs for four $PM_{2.5}$ species (VOC is not included); the spatial resolution is 36×36 km^2 [108], although current research aims to improve spatial precision.

In addition to the three RCMs mentioned above, there are many other RCMs/IAMs for the US and other regions/countries that connect emissions of local air pollution to $PM_{2.5}$ concentrations, exposures, health effects, environmental effects, or monetary costs. Other IAMs for the US include HYSPLIT with Average Dispersion (HyADS) model which estimates long-term changes in $PM_{2.5}$ associated with SO_2 emissions from coal power plants [114]; and EPA's source apportionment benefit-per-ton tool (SA-BPT) which estimates nationally aggregated monetized health benefits from $PM_{2.5}$ precursor emission reductions for specific sectors of economy [115]. IAMs for other regions/countries include Greenhouse Gas–Air Pollution Interactions and Synergies (GAINS) for European and Asian countries [116, 117], Screening for High Emission Reduction Potentials for Air Quality (SHERPA) for European cities [118], Air Benefit and Cost and Attainment Assessment System (ABaCAS) for China [119], and FAst Scenario Screening Tool (TM5-FASST) for global $PM_{2.5}$ health impacts estimation [120].

In this dissertation, I employ the InMAP source-receptor matrix (ISRM) [121], which is derived from InMAP by running it thousands of times, to estimate the $PM_{2.5}$ concentration changes from

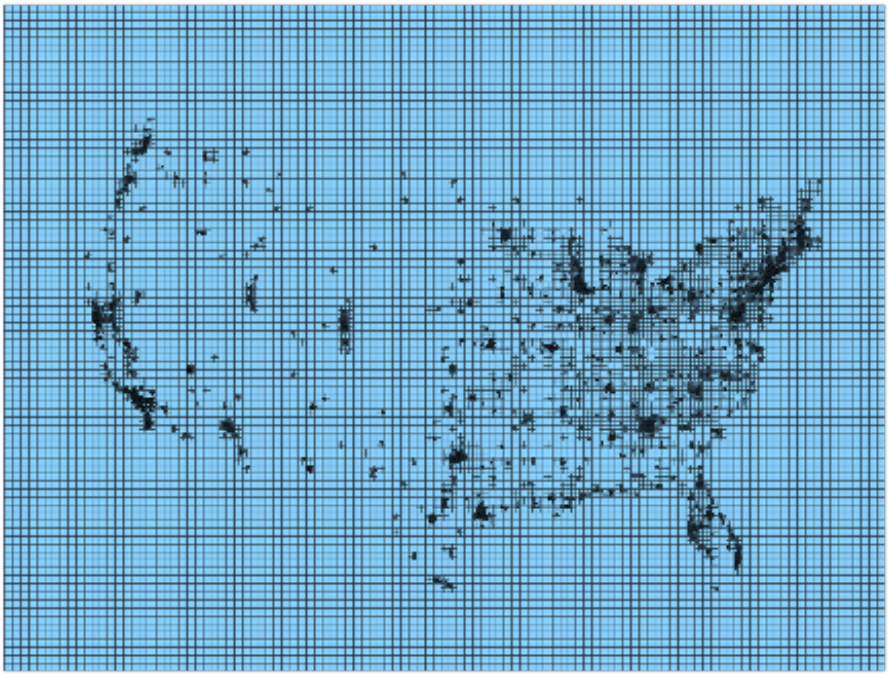

Figure 1.3: Map of ISRM grid cells

emissions changes in the US (**Chapters 4 and 5**). ISRM relates emissions in any one location to the concentration changes in all other locations. ISRM separately tracks 5 pollutants (primary $PM_{2.5}$, NO_x, SO_x, NH_3, and VOC) and 3 stack heights (ground level [<57 m]; low stack height [57 to 379 m]; and high stack height [>379 m]). Each layer of the ISRM consists of 52,411 grid cells for the contiguous US, with a variable spatial scale ranging from 1 km × 1 km in densely-populated urban centers to 48 km × 48 km in sparsely-populated rural areas (see Fig. 1.3). The population-weighted average size of a grid cell is 2.6 km in Urban Areas and 22.6 km in non-Urban Areas (13.2 km overall). ISRM is publicly available online [122].

1.4 Structure

Table 1.1 summarizes study design characteristics for the four studies (chapters) included in this dissertation. As for study location, **Chapters 2, 4 and 5** focus on US at the national level; **Chapter 3** focuses on China at the national level. **Chapter 2** investigates spatial pattern of air pollution concentration; **Chapters 3 and 4** investigate air pollution exposure inequality.

Chapter 5 investigates exposure and health inequality. **Chapter 3** quantifies exposure inequality patterns in China; **Chapters 4 and 5** explore policies to address the racial-ethnic inequality in the US. Regarding which air pollutants and air quality models, **Chapters 2 and 3** employ empirical models (LURs) to estimate levels of NO_2 and $PM_{2.5}$; **Chapters 4 and 5** employ mechanistic models (ISRM) to estimate the $PM_{2.5}$ concentration changes from emission changes.

Table 1.1: Summary of study design characteristics for each Chapter.

Chapter	Location		Topic			Air pollutant		Method	
	US	China	Concentration	Exposure inequality	Health inequality	NO_2	$PM_{2.5}$	Empirical model	Mechanistic model
2	X		X			X	X	X	
3		X		X		X	X	X	
4	X		X	X			X		X
5	X			X	X		X		X

The remaining chapters in this dissertation are organized as follows. **Chapter 2** quantifies the spatial sources of ambient NO_2 and $PM_{2.5}$ concentration. I develop a readily scalable algorithm based on "spatial-increment" to decompose air pollution concentrations into four spatial components: long-range, mid-range, neighborhood, and near-source. I apply the algorithm to the annual-average concentration predictions from empirical models for all census blocks in the contiguous US, from year-2010 to year-2015.

Chapter 3 quantifies the relationships between ambient air pollution exposures and SES in China. I employ year-2015 annual-average NO_2 and $PM_{2.5}$ concentration estimations from empirical models and national demographic data, which is derived from both China Health and Retirement Longitudinal Study (CHARLS) cohort and gridded gross national product (GDP).

Chapters 4 and 5 investigate approaches and policy scenarios to address the national racial-ethnic inequalities in the US. I use ISRM in both chapters to predict how changes in emissions impact annual-average $PM_{2.5}$ concentrations and exposure inequalities. **Chapter 4** explores three broad approaches of emission reductions: 1) "location"-specific approach; 2) "sector"-specific approach; and 3) "NAAQS-like" approach. "Sector" and "NAAQS-like" mirror aspects of current air quality regulations, "location" would be a new regulatory approach.

Chapter 5 expands the results of **Chapter 4** to an ongoing government EJ policy – the Biden Administration's Justice40 Initiative ("J40"), which is a general policy to address environmental injustice. Climate and Economic Justice Screening Tool (CEJST) is the signature element and the ongoing approach of J40. I investigate whether emission-reductions brought about by CEJST/J40 investments will eliminate disparities in $PM_{2.5}$ exposure by race-ethnicity and other attributes in 20 years, through comparing a Business As Usual ("BAU") scenario against two scenarios wherein CEJST-identified locations ("J40 communities") experience accelerated emission-reductions.

Lastly, **Chapter 6** summarizes findings, implications, and contributions, and suggests directions for future investigations on these topics.

Chapter 2

SPATIAL DECOMPOSITION OF AMBIENT AIR POLLUTION IN THE UNITED STATES

Material in this chapter was previously published in the following article: Y. Wang, *et al.*, Spatial decomposition analysis of NO_2 and $PM_{2.5}$ air pollution in the united states, *Atmos. Environ.* **241**, 117470 (2020).

2.1 Summary

Length scales for spatial variability of air pollution concentrations depend on the pollutant and the location. In this Chapter, I develop a readily scalable algorithm based on "spatial-increment", to decompose the air pollution concentration into four spatial components: long-range, mid-range, neighborhood, and near-source. I apply the algorithm to annual-average concentrations of outdoor NO_2 and $PM_{2.5}$ for all census blocks in the contiguous US. For NO_2, "neighborhood" and "mid-range" components dominate both within-city and between-city concentration differences (both components are ∼5-fold larger in large urbanized areas than rural areas). For $PM_{2.5}$, the "long-range" component dominates; this component varies by region (e.g., is three times greater in the Midwest [7 $\mu g/m^3$] than in the West [2.3 $\mu g/m^3$]), whereas variation by urban area size is relatively minor. This chapter provides the first nation-level fine-scale decomposed pollution surfaces to date; this dataset is publicly available. Results can be used to estimate, at least to a zeroth order, the contribution of sources at different distances from the receptor to the annual average pollution in a location of interest.

2.2 Introduction

Outdoor concentrations of air pollutants vary on length scales from meters, to 100s or 1000s of km. Concentrations in the environment reflect spatial patterns in emissions, physical and chemical processes governing production and removal of pollutants, and turbulent advection and dispersion that drives transport and dilution. The degree of spatial heterogeneity varies by time, location, and pollutant.

Here, I use spatial decomposition to investigate how annual-average concentrations across the United States vary at length scales ranging from under 1 km to over 100 km. A motivation is that spatial patterns at a location may reveal information on source contributions: spatially homogenous concentrations suggest dominance of regional sources and/or secondary pollutants (e.g., sulfate $PM_{2.5}$ formed from SO_2 emissions from coal power plants), whereas heterogeneities on short length-scales suggest influence by local emission sources (e.g., NO_x emissions from traffic in an urban area).

The literature has quantified spatial components of air pollution concentrations (e.g., local, urban, long-range components) based on source-receptor modeling and concentration increments. The modeling approach [123–125] often involves comparing results from a chemical transport model (CTM) under different scenarios (e.g., in the model, switching on/off local traffic, total urban emissions, or upwind power plants) [126, 127]. This method is typically used for a single location (a specific city or region); application to many locations is computationally expensive [128]. Recently developed source-apportionment algorithms, such as Particulate Source Apportionment Technology (PSAT) and Integrated Source Apportionment Model (ISAM), running in parallel with CTMs have reduced this cost by a factor of 10–20, but they still can be applied to only a few tens of receptors [129–132]. Some integrated assessment tools (e.g., Air Benefit and Cost and Attainment Assessment System (ABaCAS), Greenhouse Gas–Air Pollution Interactions and Synergies (GAINS), the Intervention Model for Air Pollution (InMAP) Source-Receptor Matrix (ISRM), Screening for High Emission Reduction Potentials for Air Quality (SHERPA), FAst Scenario Screening Tool (TM5-FASST)) retrieve source-appointment relationships through calculated emission-concentration sensitivities of full CTMs, and can be efficiently applied to hundreds of receptors in a larger domain [109, 118, 120, 121, 133]. Those and other modeling approaches could potentially be applied to questions considered here. Limitations of the modeling approaches include (1) problems of nonlinearity (or non-additivity) for secondary pollutants [134–136], (2) limitations of input data (e.g., uncertainty in the emission inventory), and (3) limited spatial resolution for national-scale simulations [16, 133, 137].

Concentration-based approaches use spatial concentration increments, typically from field measurements, to infer spatial or source contributions [126, 138–146]. Lenschow (2001) utilized such a method to estimate the local, urban, and regional background concentrations of PM_{10} for Berlin by subtracting the roadside, urban background, and regional background measurement [143]. Ap-

proaches based on the analysis of temporally -resolved pollutant concentrations have also been used [126, 138, 139, 142, 146]. These methods are limited by the availability of the necessary measurements and have previously been applied to a single or a few locations at a time. Only a few studies have considered spatial decomposition for a broader region relying mainly on empirical modeling prediction surfaces and concentration increment approaches. Antonelli et al. (2017) used an image decomposition method, called wavelet decomposition, to decompose a 1 km $\times$ 1 km spatial surface of year 2003–2008 daily $PM_{2.5}$ predictions in the New England region of the US into three spatial components, determined through visual inspection of the decomposed surfaces [147]. Beelen et al. (2009) estimated regional background, rural and urban concentration surfaces separately using universal kriging from rural and urban background sites, and covariates representing their respective spatial scales and sources, to produce 1 km $\times$ 1 km composite EU-wide maps of year 2001 annual-averaged NO_2, PM_{10} and O_3 concentrations [148].

Here, I explore a readily scalable algorithm to spatially decompose ambient air pollution concentrations using spatial increments, and apply this method to a national fine-scale dataset of outdoor NO_2 and $PM_{2.5}$ predictions. To my knowledge, this Chapter provides the first national high-resolution, spatially decomposed air pollution surfaces for the US. The results, which are publicly available, provide useful information for the contiguous US regarding the contribution of air pollution sources at spatial scales from local to regional.

2.3 Materials and methods

$PM_{2.5}$ and NO_2 concentrations employed here are outdoor annual-average predicted concentrations at all (n $\approx$ 6 million) census block centroids with non-zero population in the contiguous US. The concentration predictions are derived from the Center for Air, Climate, and Energy Solutions (CACES) empirical models [93]. The predictions incorporate satellite-derived estimates of ambient concentrations and land-cover, land-use and other geographic datasets, and ground-level monitoring data in a universal Kriging framework. Publicly-available predictions are for six air pollutants and multiple-years (1979–2015) (`www.caces.us`). I mainly report decomposition analyses for year-2010 in the main results; results for other years as sensitivity tests).

I decompose annual-average concentrations into four spatial components: "long-range", "mid-

range", "neighborhood", and "near-source":

$$C_i^{\text{long-range}} = \min\{C_j | d_{i,j} \leq 100km\}$$
$$C_i^{\text{mid-range}} = \min\{C_j | d_{i,j} \leq 10km\} - \min\{C_j | d_{i,j} \leq 100km\}$$
$$C_i^{\text{neighborhood}} = \min\{C_j | d_{i,j} \leq 1km\} - \min\{C_j | d_{i,j} \leq 10km\}$$
$$C_i^{\text{near-source}} = C_i - \min\{C_j | d_{i,j} \leq 1km\}$$

Here, C_i is the model-predicted concentration at location i; the four superscripts on C_i ("long-range", "mid-range", "neighborhood", "near-source") are the four spatial components considered here; $d_{i,j}$ is the distance between two locations i and j; and $\min\{C_j | d_{i,j} \leq$ "x"km$\}$ represents the minimum concentrations from all block centroids within an "x" km circular buffer of location i (where "x" takes the values above: 1, 10, and 100 km).

The 1, 10 and 100 km length scales are commonly used for neighborhood, city, and regional scales [149, 150]. This approach is straightforward, intuitive, and computationally demanding, but feasible to implement for a national dataset, across multiple pollutants and years. While these four components are inherently defined by the spatial increments of predicted concentrations, they may provide insight into source contribution. The spatial decomposition algorithm is illustrated in Fig. 2.1. "Long-range" (>100 km) likely represents regional background and long-range transport. "Mid-range" (10–100 km) likely represents, e.g., urbanized areas, and agricultural regions. "Neighborhood" (1–10 km) likely represents localized sources such as commercial districts, industrial areas, and intersections of major highways. "Near-source" (<1 km) likely represents hyper-local enhancements (e.g., roadways).

I apply the spatial decompositions to each of the approximately 6 million census block centroids in the contiguous US and analyze results nationally, by state, region, and urbanicity. Urbanicity levels are defined in the 2010 Census [151] based on population: urbanized areas have 50,000 or more people (I subdivided them into small/medium/large urbanized areas, by population tertiles); urban clusters have 2500–50,000 people; rural areas are all remaining census blocks. Rural areas contain 59 million people; urban clusters (n = 3087) contain 29 million people; small urban areas (n = 440) contain 75 million people; medium urban areas (n = 47) contain 75 million people; large urban areas (n = 10) contain 73 million people.

To further explore urban-scale patterns, I selected six cities to investigate more closely: New York, NY (year-2010 urban area population: 18.4 million); Los Angeles, CA (12.1 million); Seattle,

WA (3.1 million); Minneapolis, MN (3.1 million); Spokane, WA (0.5 million); Tuscaloosa, AL (0.1 million). These cities were selected to represent a range of sizes, pollution sources, and geographies (e.g., region of the US, climate, distance-to-coast, and regional economy) across the US. For each city, I conduct the following analyses to understand urban-scale variability. Following Novotny et al. (2011) [98], I consider concentrations along transect lines across the urban center. Here, I apply the decompositions on the points at each 10-m interval along the transect lines, using empirical-model predictions from the nearest block centroid as an approximation of the total concentrations at each 10-m interval. As a sensitivity analysis to explore the robustness of the approach, I compare results for multiple years and for several transects per city (12 transects passing through the city center, at 15° intervals). Also, I compare results with different buffer radii to provide insight on the choice of buffer lengths.

To understand the relationship between the model structure of the prediction surface and the spatial decomposition components, I calculate how much each empirical regression component (i.e., Kriging and each independent variable) contributes to the year-2010 spatial decomposition results for both pollutants. For each block centroid prediction and for both pollutants, I separate the without-Kriging concentration predictions (i.e., the ultimate prediction results minus the Kriging adjustment values), and then apply the same spatial decomposition algorithm to the without-Kriging predictions. The contributions of the Kriging to the four spatial components are calculated by the population-weighted averages of the concentration differences between initial and without-Kriging decompositions. For other regression components, I run the no-intercept multiple regressions of without-Kriging decomposed concentrations on the model-selected independent variables for total concentration estimation. The contributions of each independent variable are calculated as the population-weighted-average product of the variable values and coefficients; the contributions are then aggregated to nine categories according to the variable type. Following Kim et al. (2020) [93], the categories are traffic, urban land-use, rural land-use, population, elevation, emission, imperviousness, vegetation, and satellite.

2.4 Results

2.4.1 Within-urban decomposition results

Figx. 2.2 and 2.3 illustrate, for a case-study city (Seattle), the results of the spatial decomposition approach developed here. Across a ∼100-km urban transect (Fig. 2.2), the "long-range" values are almost constant. That result is expected; those values likely represent the regional background. For Seattle, "long-range" values account for 21% (for NO_2) and 38% ($PM_{2.5}$) of the average predictions. The "mid-range" increments are low at the urban edge and increase gradually approaching the urban center at approximately 50 km along the line, presumably reflecting the effects of aggregated urban emissions in the urban center. The "neighborhood" and "near-source" values vary at much shorter spatial scales, with multiple peaks along the transect, presumably reflecting localized emission sources. Spatial variability is greater for NO_2 than $PM_{2.5}$, and the "near-source" component is larger for NO_2 (15%) than $PM_{2.5}$ (7%).

Fig. 2.3 shows the block centroid decomposed concentrations for Seattle (see Fig. 2.4 for 1, 10, and 100 km buffer minimum concentrations). "Near-source" concentrations of NO_2 reflect the major road network (see Fig. 2.5). For $PM_{2.5}$, "near-source" concentrations are relatively low (<1 $\mu g/m^3$), with no clear spatial patterns. The "neighborhood" component has several hot spots for both pollutants (e.g., downtowns, intersections of major highways).

The transect-line decompositions for different transect directions (Figs. 2.6 to 2.12), different years (Fig. 2.13 and 2.14), and block-level decomposition maps for the additional five cities (Figs. 2.15 to 2.19) imply consistent results. For each case-study city, the decomposition results are generally stable over different years and transects, and spatial patterns for each of the components are in general consistent with expectations. For example, Fig. 2.14 captures the elevated concentrations and contributions of "long-range" $PM_{2.5}$ component in Spokane (and to a lesser degree Seattle) in 2015 owing to the large wildfire season in northeastern Washington State [152]. The decompositions for different cities reveal broadly similar spatial patterns of variability, though the partitions of spatial components differ. For example, for both pollutants, the "mid-range" component (likely representing elevated regional urban concentrations) is typically less pronounced for the smaller cities (Spokane, Tuscaloosa), whereas for bigger cities (New York City, Los Angeles) this component often has a larger contribution and remains elevated throughout the urban portion of the transect, only decreasing when the transects extend into unpopulated areas such as mountainous

areas and open water. Relative to the other cities, the "long-range" contributions are comparatively smaller for $PM_{2.5}$ in Seattle and for NO_2 in Los Angeles.

2.4.2 Decomposition results by state, region and urbanicity

After applying the spatial decomposition to all block centroids in the US (Fig. 2.20), I calculate the corresponding population weighted averages by state (Fig. 2.21). For NO_2, high levels of the "neighborhood" and "mid-range" components are found in states with large cities (e.g., Los Angeles, New York, Chicago). For $PM_{2.5}$, the overall concentration and the "long-range" component are higher in the East than in the West.

Fig. 2.22 summarizes the total and decomposed concentrations by urban area size and geographic region. For NO_2, concentrations are ∼3-fold higher in large urban areas (average: 13 ppb) than rural areas (4 ppb). "Neighborhood" and "mid-range" components (both about 4 ppb) dominate that concentration difference; "long-range" components are nearly the same in absolute terms (<1 ppb difference) across urban/rural groups. In relative terms, the "long-range" component is 21% overall (16% [large urban areas], 41% [rural]; Fig. 2.23). The "near-source" component is 13% overall (12% [large urban areas], 7% [rural]). The differences in decomposed NO_2 concentrations among regions are mostly driven by regional differences amongst the ten largest urbanized areas (see Fig. 2.24). The decomposition patterns suggest that NO_2 is of urban origin and varies locally, and that climate or geographic differences have less effects on the NO_2 concentrations. The mid-range and neighborhood components are greater for the large urbanized areas than for other locations (see Fig. 2.22), consistent with larger cities having comparatively denser traffic and more urban emissions.

In contrast, geographic regions dominate the differences of decomposed $PM_{2.5}$ concentrations, mainly in the "long-range" component. That results likely reflects that $PM_{2.5}$ is a regional pollutant with a strong secondary component. The "long-range" concentration in the Midwest (7 $\mu g/m^3$) is almost three times than in the West (2.3 $\mu g/m^3$); it contributes 30% of the total $PM_{2.5}$ in the West, versus > 50% in the remaining three regions. I hypothesize that these regional differences mainly reflect differences in emissions and emissions density (e.g., traffic, coal-fired power plants, agriculture), but other aspects (e.g., topography, climate, meteorology) may also play a role. The $PM_{2.5}$ concentration differences by urban area size are relatively minor compared with regional differences. $PM_{2.5}$ concentrations are higher for large urban areas than for small urban areas, and

urban concentrations are larger than rural concentrations, but the differences are relatively modest (< 2 $\mu g/m^3$); those differences are mainly driven by the increase of "mid-range" concentrations. "Long-range" components dominate ($\geq 50\%$) $PM_{2.5}$ concentrations in both rural and urban areas.

2.4.3 Contributions of empirical regression terms

Fig. 2.25 summarizes contributions of each empirical regression term (traffic, land-use, satellite data, Kriging, etc.) for the overall pollution concentration to the four decomposed concentrations. A version of Fig. 2.25, but without the satellite contributions, is in Fig. 2.26. The goal of this analysis is to find out how the underlying model structure contributes to the spatial decomposition patterns, not to make causal inference about mechanisms. The results are aggregated to all US blocks, urban blocks, and rural blocks, separately. For both pollutants, satellite concentration estimates dominate the "long-range" component. For NO_2 in urban blocks, two variables that strongly contribute, in addition to satellite estimates, are vegetation (a "negative" contribution: more vegetation corresponds to lower concentrations of NO_2) for the "mid-range" component and imperviousness for the "neighborhood" and "near-source" components. For $PM_{2.5}$, satellite estimates dominate; in addition, impervious surfaces also contribute mainly for the "neighborhood" and "mid-range" components. Land-use types (both rural and urban land-uses) play a most important role for "near-source" $PM_{2.5}$. In comparison, for rural blocks, the contributions of imperviousness are negligible for both pollutants.

2.5 Discussions and conclusions

My results reveal within-city and national patterns of the decomposed concentrations. NO_2 is an air pollutant of predominantly urban origin: differences in mean concentrations are mainly dominated by "neighborhood" and "mid-range" components, and concentrations vary by urbanicity level. $PM_{2.5}$ is dominated by "long-range" transport; concentrations vary at a state and regional level. Those findings are consistent with previous research and understandings [153–155]. My results provide new and useful quantification of the spatial patterns of these two pollutants, on a consistent basis with high spatial precision and throughout the US.

One limitation of my results is the use of model-estimated concentrations; those models, while exhibiting good predictive performance (R^2: 0.84 [NO_2], 0.85 [$PM_{2.5}$]; root-mean-square error: 2.2 ppb [NO_2], 1.2 $\mu g/m^3$ [$PM_{2.5}$]), are inherently less variable than the actual concentrations. The

models are trained on regulatory monitoring data typically designed to capture ambient concentrations. In addition, the empirical models represent national patterns and average relationships between land-use and pollution concentrations; they are unable to predict hot-spots caused by atypical conditions. In some locations, near source contributions will be underestimated. Besides, concentrations from empirical models do not provide quantification of source contribution or transport trajectories, thus I cannot make firm causal statements about mechanisms.

Another aspect of the decompositions is that they are performed on the census block centroid locations. An advantage of the census blocks over equal-distanced grids is that blocks are dense in populated areas where people live and pollution gradients are typically largest. Each block centroid has, on average, 50 additional block centroids within its 1-km circular buffer. However, 578,971 (9.3%) and 938 (0.015%) blocks centroids have no neighbors within 1 and 10 km buffers, respectively. Those blocks represent 5.1% [no 1-km neighboring block] and 0.007% [no 10-km neighboring block] of the contiguous US population. This outcome generally occurs in sparsely populated regions (Fig. 2.27). Those regions would likely have few local sources but the true average local contribution would be nonzero. However, in my approach, the local component is, by definition, zero for locations with the minimum concentration within 1 km (which includes all locations with no 1-km neighboring Blocks). Since the true value is nonzero but the estimated value is zero, I am underestimating (Fig. 2.28). Because the local contributions are expected to be small, the amount of underestimation is likely small in absolute terms; however, other techniques would be needed to quantify the "local" component in these locations.

An additional aspect of this research that is important for interpreting the results is the use of fixed radii (1, 10, and 100 km) to calculate the four spatial components. Fixed radii reflect an approach that is easy to understand and straightforward to apply nationally. However, because of differences in size and shape of built-up areas, the impact of the radii may vary across the US. For example, the "mid-range" component employs a 10 km buffer, yet city-size varies. In the 2010 US Census [151], the mean area of an urban area is 462 km^2 (equivalent radius for a circular layout: 12 km), with an interquartile range of 112 km^2 (equivalent radius: 6 km) to 413 km^2 (11 km) and an overall range of 25.7 km^2 (3 km) to 8936 km^2 (53 km). Thus, the 10 km buffer can be smaller or larger than an urban area. The values used here (1, 10, 100 km) were selected because they represent orders of magnitude and because in preliminary investigations they seemed to capture the components better than alternative values (Figs. 2.29 and 2.30). Moreover, the use of a circular

buffer assumes equal contribution from all directions and ignores meteorology and topography, which could lead to mischaracterization in some locations. For example, the national block-level "long-range" NO_2 and $PM_{2.5}$ maps (Fig. 2.20) exhibit blotchy circular features, which are typically driven by several "clean" points that have much lower concentrations than surroundings, or by sharp gradients in these areas and few points (such as mountainous regions). The blotchy features diminish with aggregation (e.g., to state level [Fig. 2.21]).

Strengths of this study include national application of a straightforward approach for spatial decomposition; shedding new light by applying my method at fine resolution but across a broad domain; and investigating results for the whole US and separately by city, state, region, and level of urbanicity. The spatial increment and moving window approach are easy to understand and computationally feasible for the resolution and domain considered here. Weaknesses include that the empirical model is imperfect (e.g., nonzero error and bias; outlier concentrations may be underestimated), the empirical model is predicted only at centroids, and that the approach does not shed light on mechanisms.

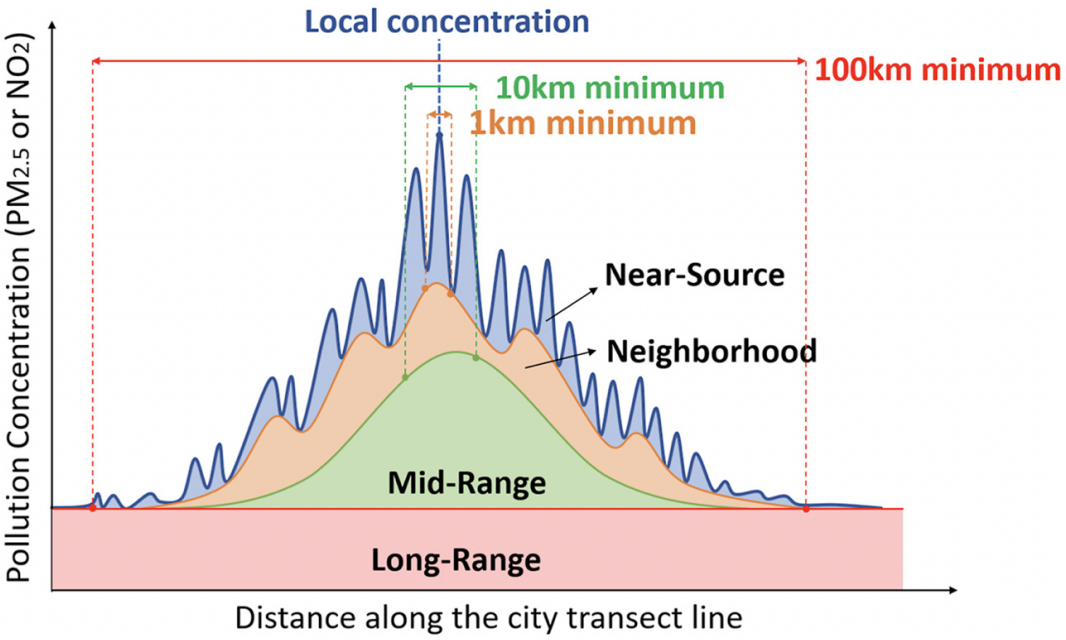

Figure 2.1: A graphic illustration of the spatial decomposition algorithm.

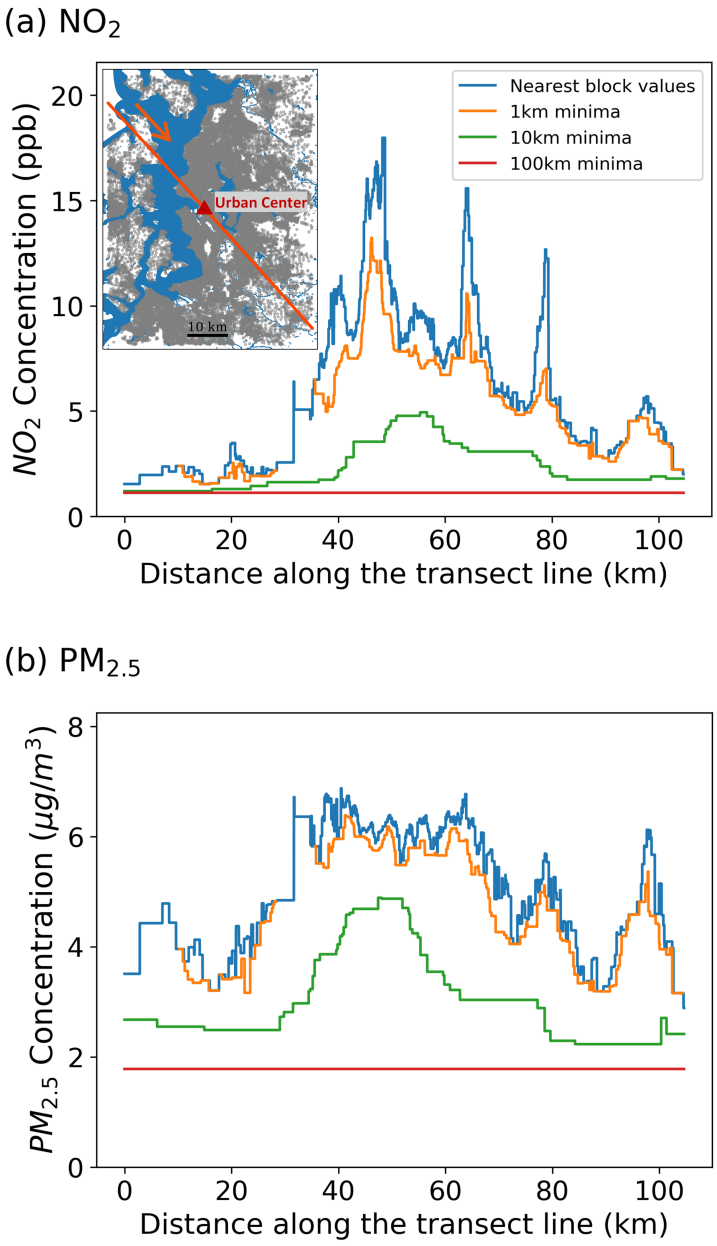

Figure 2.2: Decomposition results for (a) NO_2 and (b) $PM_{2.5}$ along a transect line from Northwest to Southeast across the center of Seattle. The inset map in the upper left of (a) shows the Seattle census block centroids (shown in grey) and transect line.

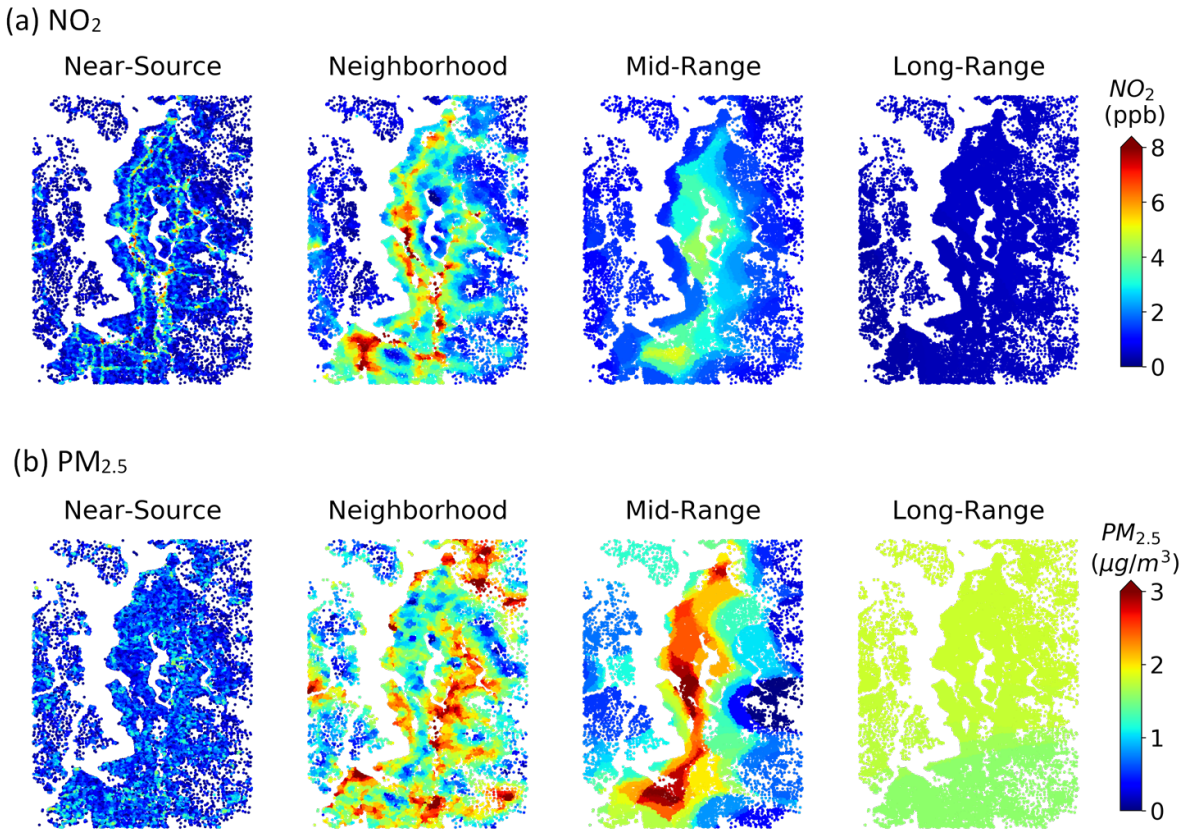

Figure 2.3: The block level decomposed concentration maps of (a) NO_2 and (b) $PM_{2.5}$ for Seattle.

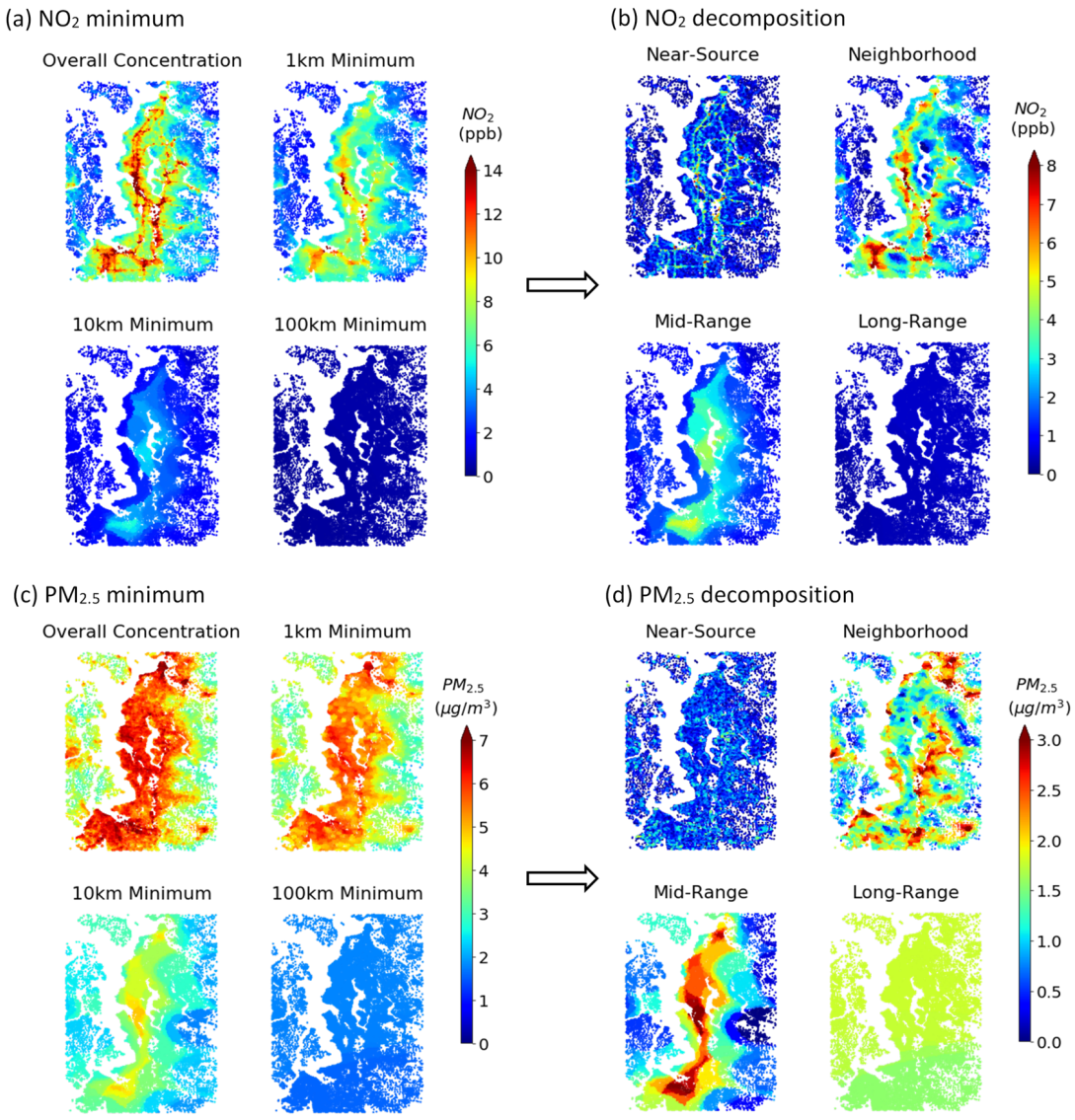

Figure 2.4: The buffer minimum plots and decomposition plots of NO_2 and $PM_{2.5}$ for Seattle area: (a) Predicted NO_2 annual average concentration and corresponding 1, 10, and 100 km buffer minimum maps; (b) Decomposition maps of NO_2; (c) Predicted $PM_{2.5}$ annual average concentration and corresponding 1, 10, and 100 km buffer minimum maps; (d) Decomposition maps of $PM_{2.5}$.

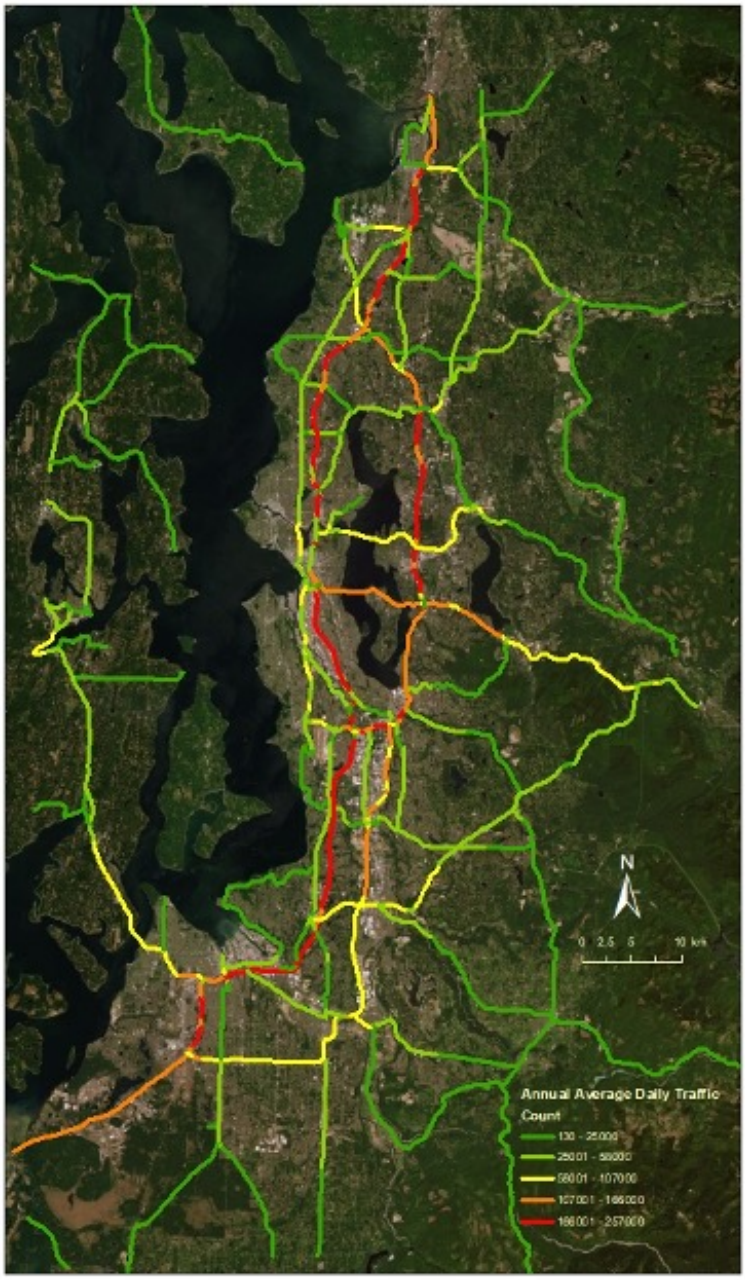

Figure 2.5: 2010 Annual average daily traffic counts for state highway in Seattle urban area. Data is from Washington State Department of Transportation - Statewide Travel and Collision Data Office, can be downloaded freely at: `www.wsdot.wa.gov/mapsdata/travel/traveldata.htm`

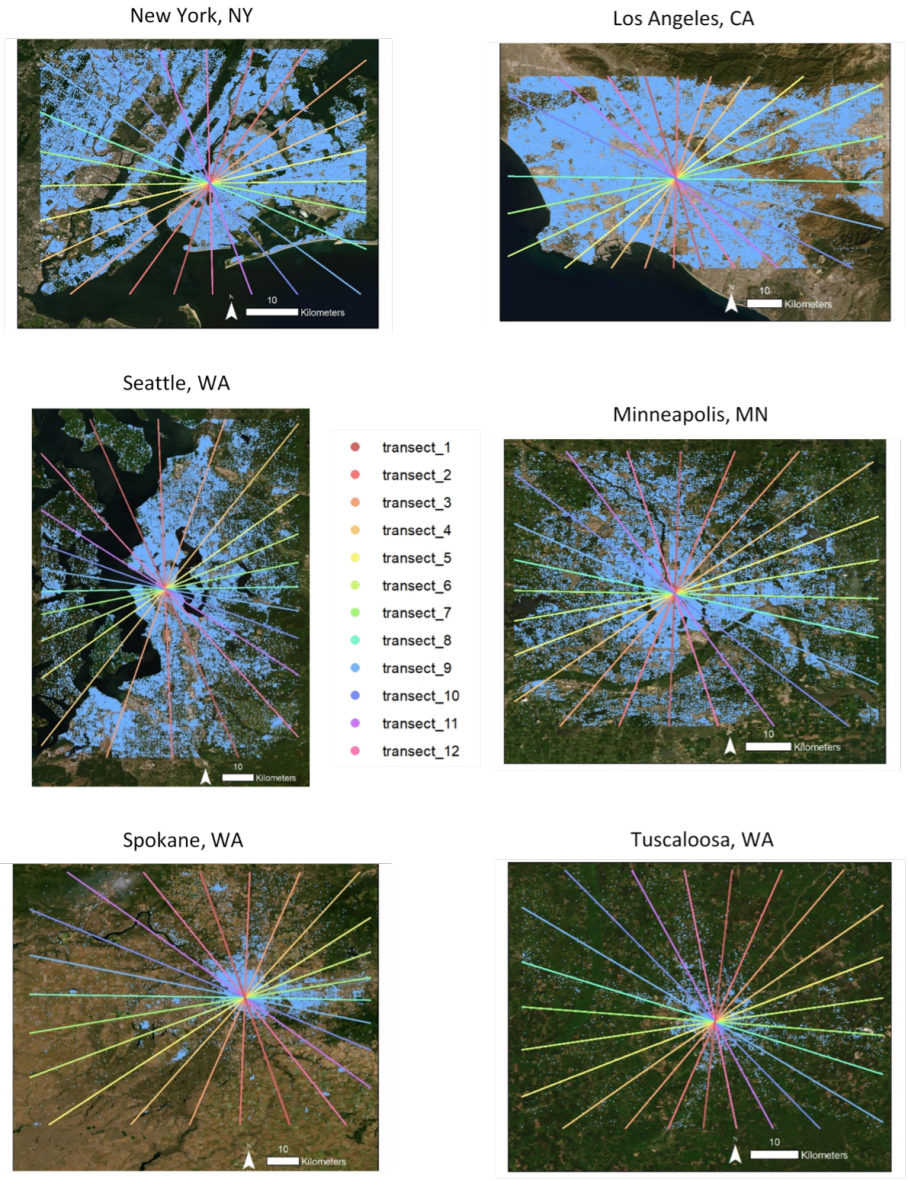

Figure 2.6: Maps of block centroids (shown in blue), multiple transect lines (colored, at each 15°, 12 lines in total for each city; the transects are directed all from West to East), and satellite images (background) for the six cities: New York, NY; Los Angeles, CA; Minneapolis, MN; Seattle, WA; Spokane, WA; Tuscaloosa, AL.

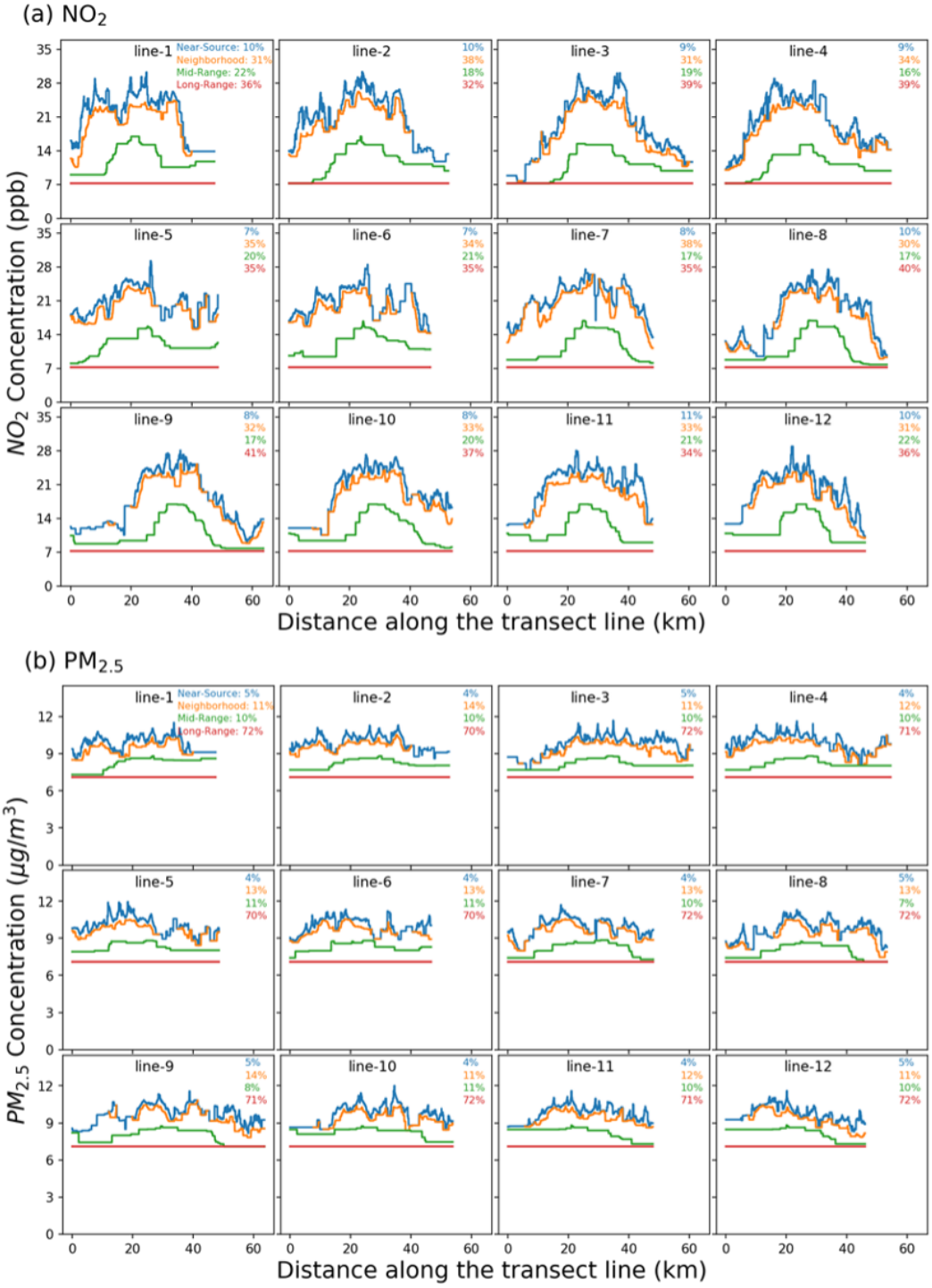

Figure 2.7: Decomposition results and average contributions of four components (upper right of each subplot) for (a) NO_2 and (b) $PM_{2.5}$ along 12 transect lines in New York, NY (see Fig. 2.6 for transect line locations)

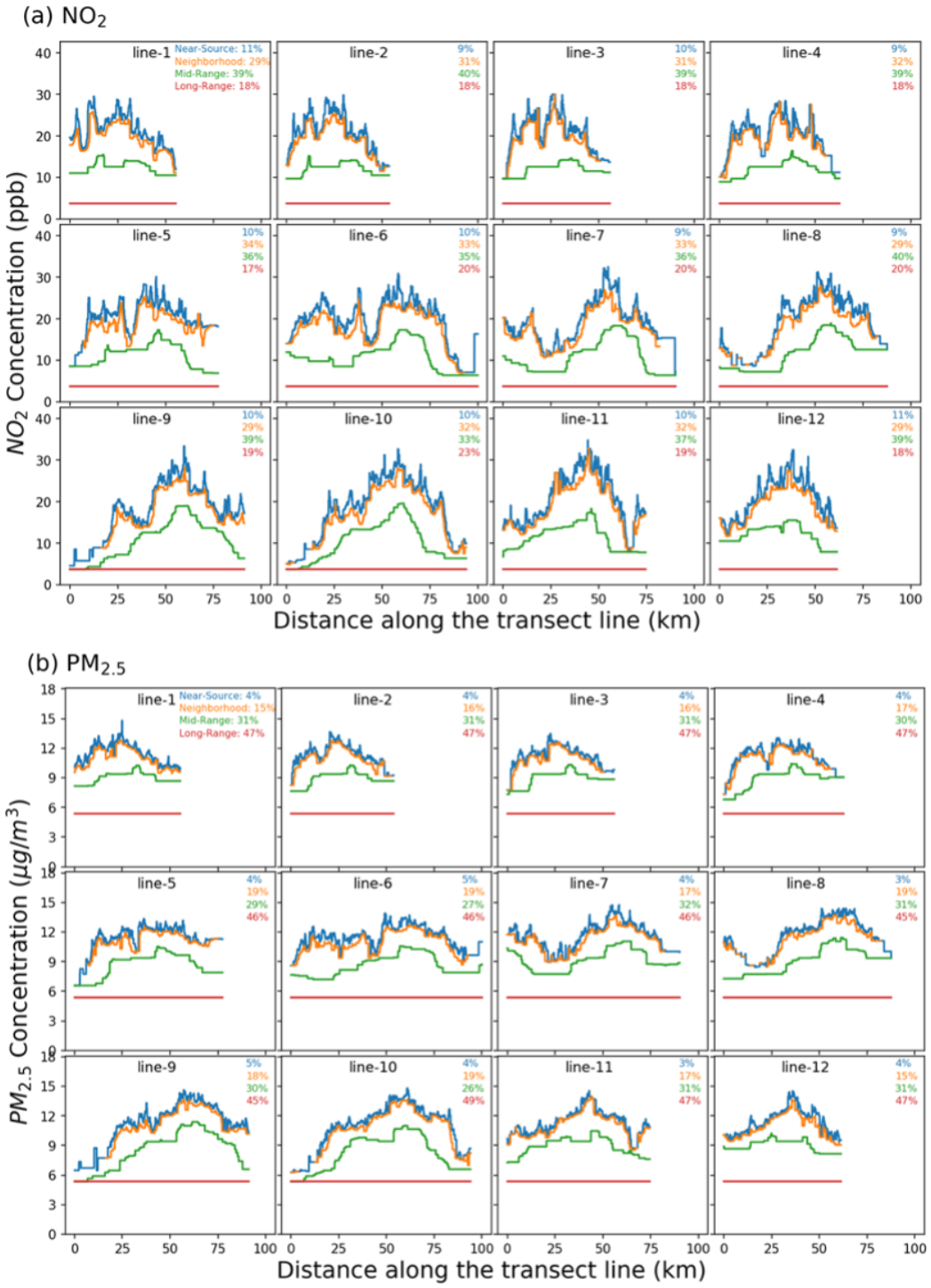

Figure 2.8: Decomposition results and average contributions of four components (upper right of each subplot) for (a) NO_2 and (b) $PM_{2.5}$ along 12 transect lines in Los Angeles, CA (see Fig. 2.6 for transect line locations)

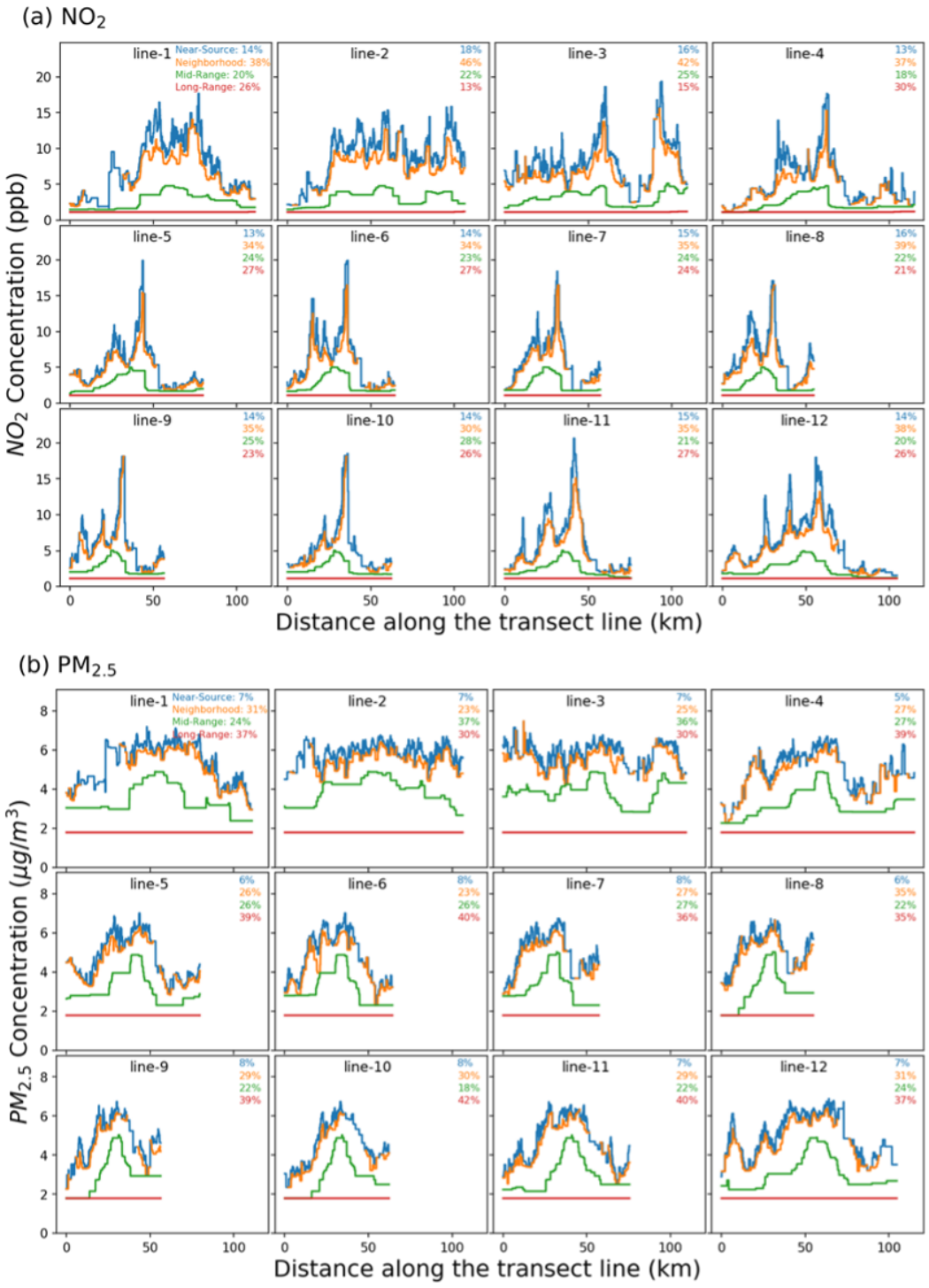

Figure 2.9: Decomposition results and average contributions of four components (upper right of each subplot) for (a) NO_2 and (b) $PM_{2.5}$ along 12 transect lines in Seattle, WA (see Fig. 2.6 for transect line locations)

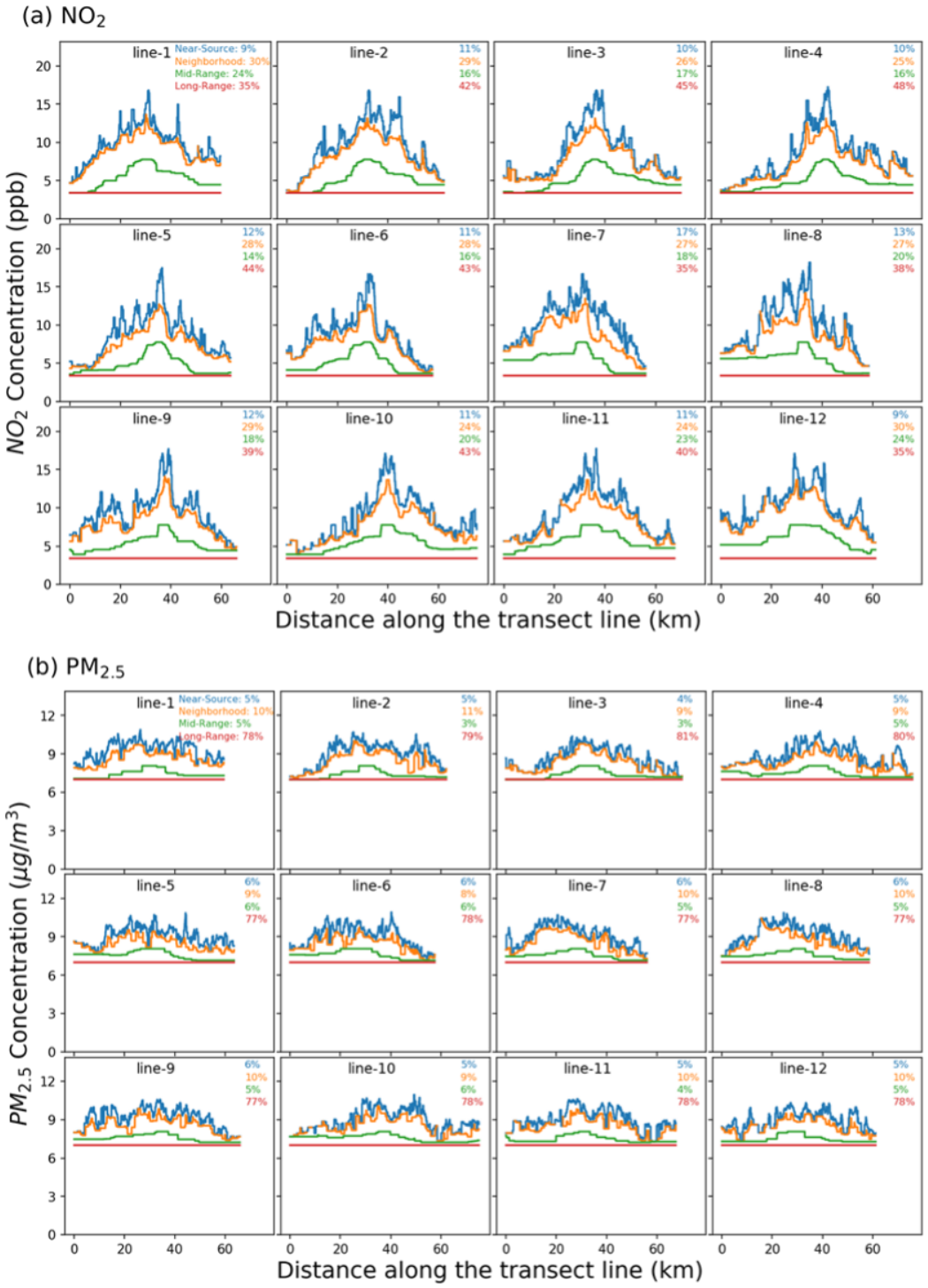

Figure 2.10: Decomposition results and average contributions of four components (upper right of each subplot) for (a) NO_2 and (b) $PM_{2.5}$ along 12 transect lines in Minneapolis, MN (see Fig. 2.6 for transect line locations)

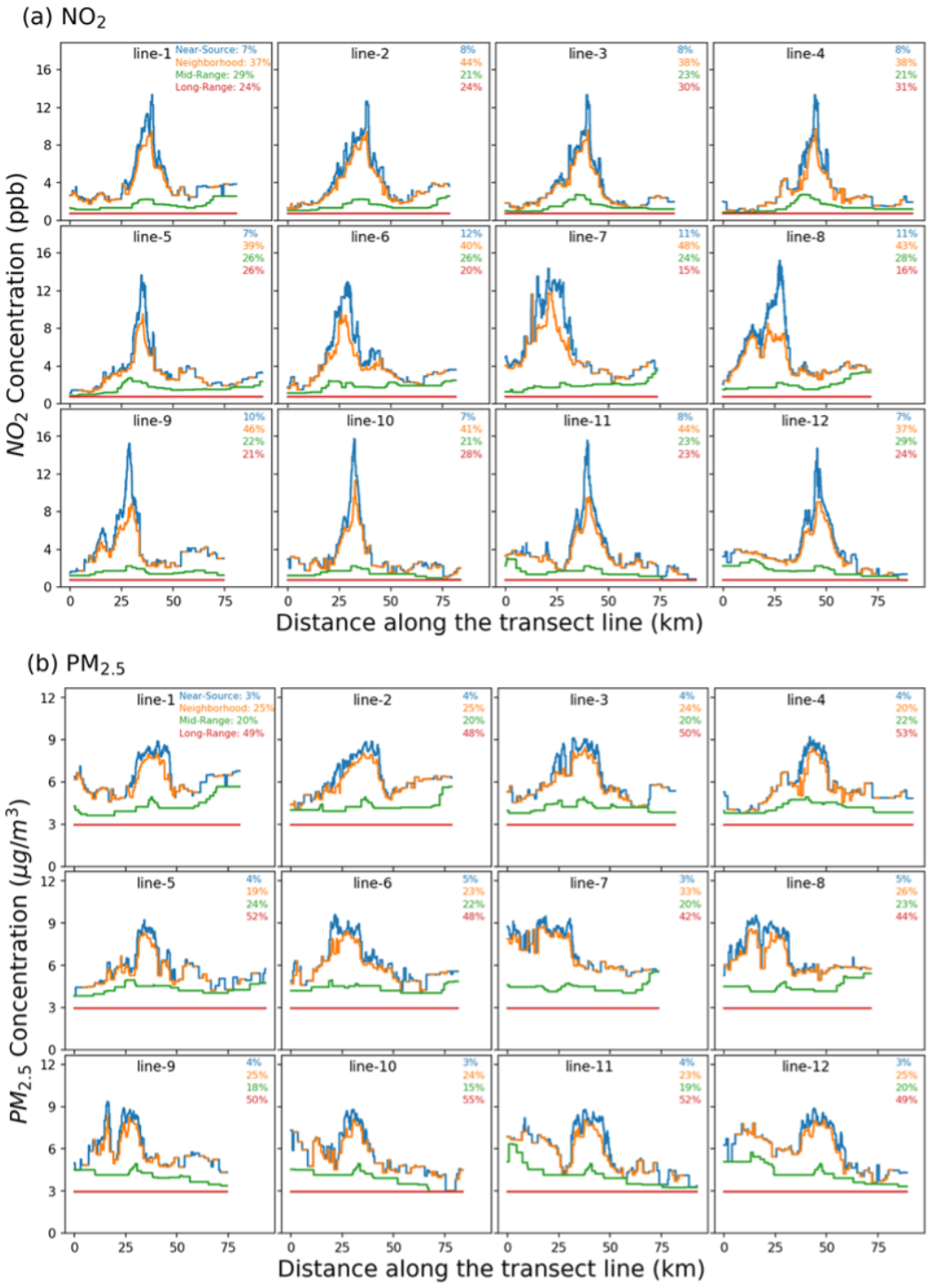

Figure 2.11: Decomposition results and average contributions of four components (upper right of each subplot) for (a) NO_2 and (b) $PM_{2.5}$ along 12 transect lines in Spokane, WA (see Fig. 2.6 for transect line locations)

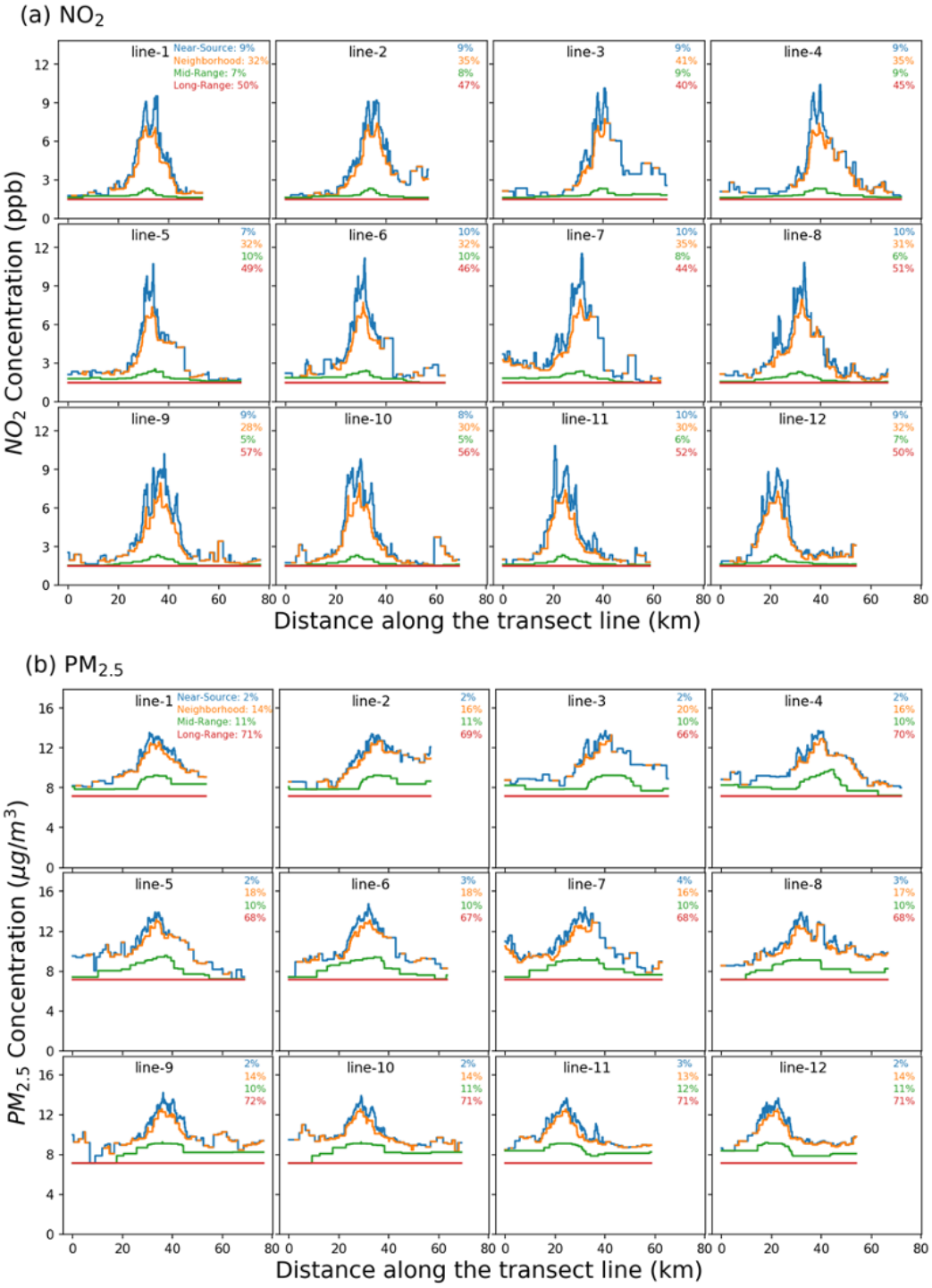

Figure 2.12: Decomposition results and average contributions of four components (upper right of each subplot) for (a) NO_2 and (b) $PM_{2.5}$ along 12 transect lines in Tuscaloosa, AL (see Fig. 2.6 for transect line locations)

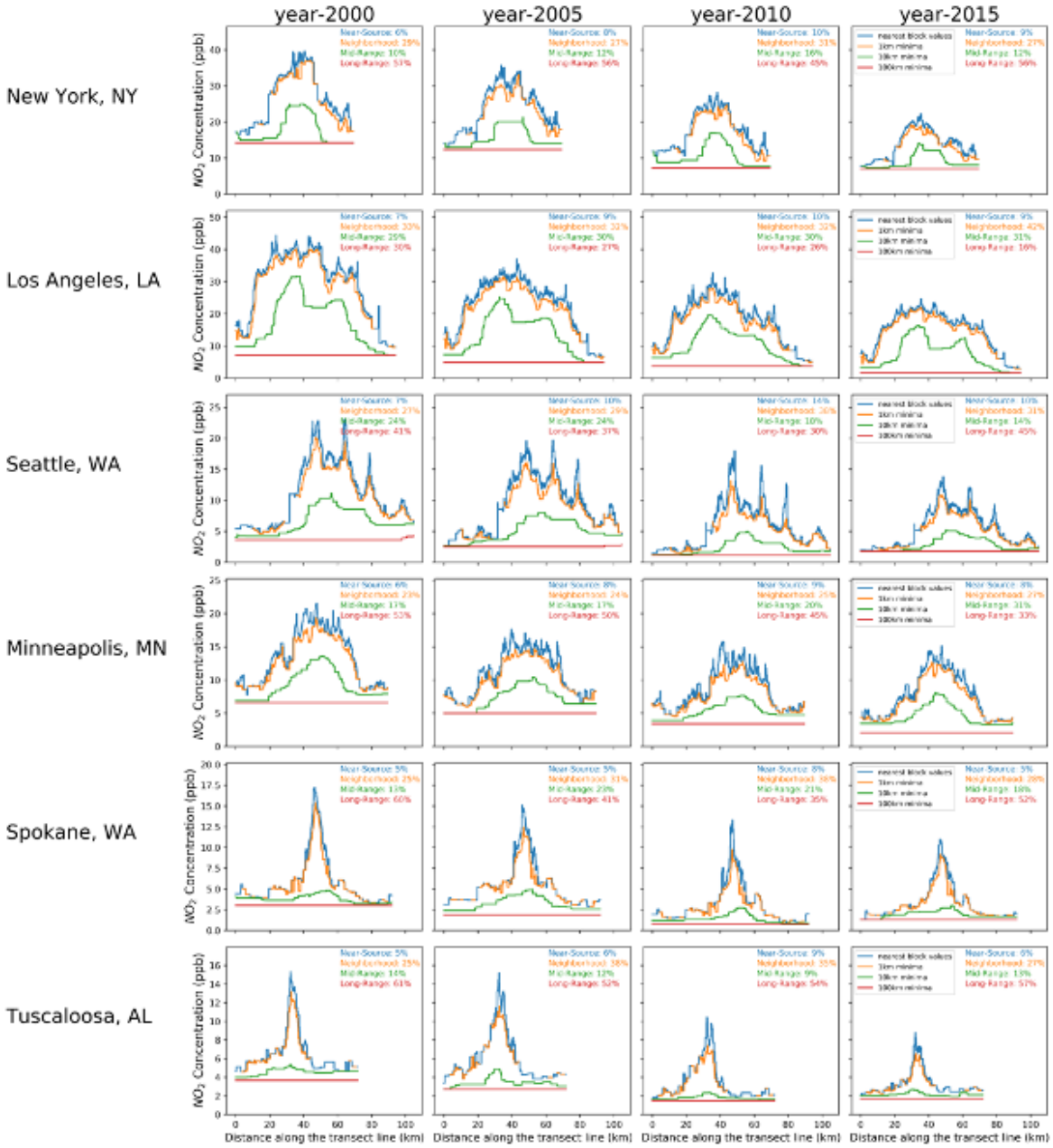

Figure 2.13: Multiple years (2000, 2005, 2010, 2015) NO_2 transect line decomposition plots and average contributions of four components (upper right of each subplot) for six cities.

Figure 2.14: Multiple years (2000, 2005, 2010, 2015) $PM_{2.5}$ transect line decomposition plots and average contributions of four components (upper right of each subplot) for six cities.

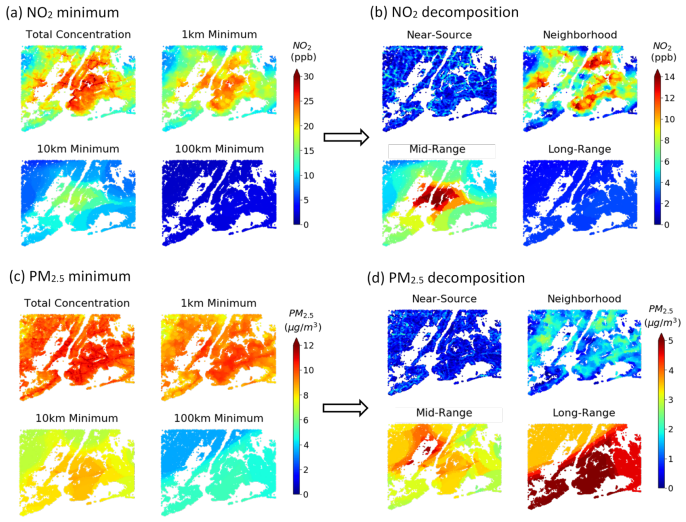

Figure 2.15: The buffer minimum plots and decomposition plots of NO_2 and $PM_{2.5}$ for New York area: (a) Predicted NO_2 annual average concentration and corresponding 1, 10, and 100 km buffer minimum maps; (b) Decomposition maps of NO_2; (c) Predicted $PM_{2.5}$ annual average concentration and corresponding 1, 10, and 100 km buffer minimum maps; (d) Decomposition maps of $PM_{2.5}$.

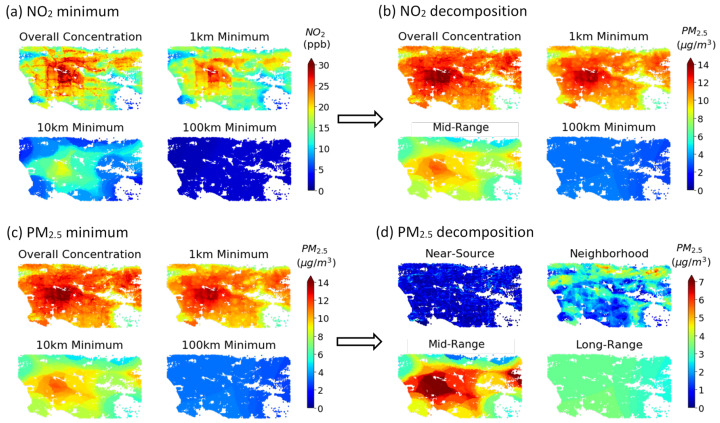

Figure 2.16: The buffer minimum plots and decomposition plots of NO_2 and $PM_{2.5}$ for Los Angeles area: (a) Predicted NO_2 annual average concentration and corresponding 1, 10, and 100 km buffer minimum maps; (b) Decomposition maps of NO_2; (c) Predicted $PM_{2.5}$ annual average concentration and corresponding 1, 10, and 100 km buffer minimum maps; (d) Decomposition maps of $PM_{2.5}$.

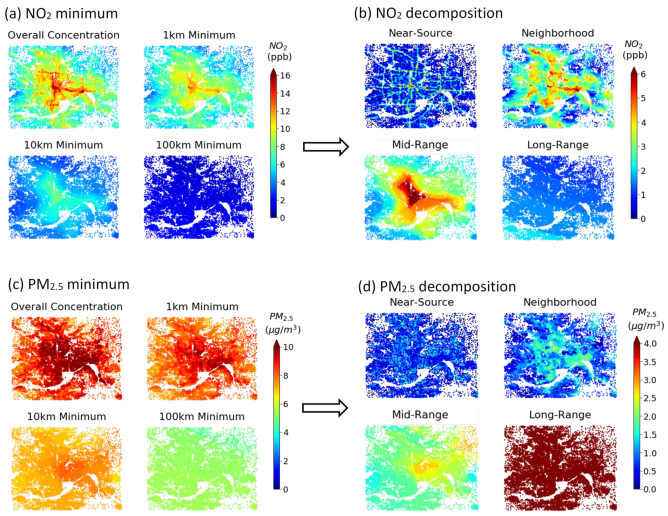

Figure 2.17: The buffer minimum plots and decomposition plots of NO_2 and $PM_{2.5}$ for Minneapolis area: (a) Predicted NO_2 annual average concentration and corresponding 1, 10, and 100 km buffer minimum maps; (b) Decomposition maps of NO_2; (c) Predicted $PM_{2.5}$ annual average concentration and corresponding 1, 10, and 100 km buffer minimum maps; (d) Decomposition maps of $PM_{2.5}$.

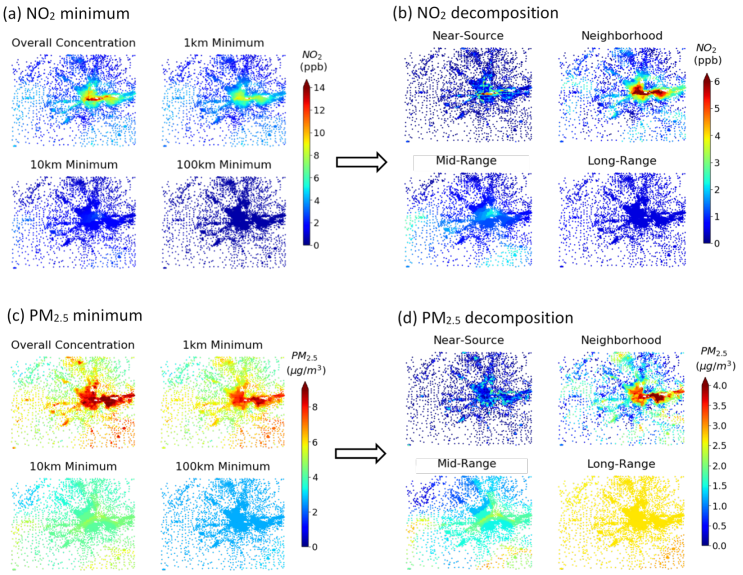

Figure 2.18: The buffer minimum plots and decomposition plots of NO_2 and $PM_{2.5}$ for Spokane area: (a) Predicted NO_2 annual average concentration and corresponding 1, 10, and 100 km buffer minimum maps; (b) Decomposition maps of NO_2; (c) Predicted $PM_{2.5}$ annual average concentration and corresponding 1, 10, and 100 km buffer minimum maps; (d) Decomposition maps of $PM_{2.5}$.

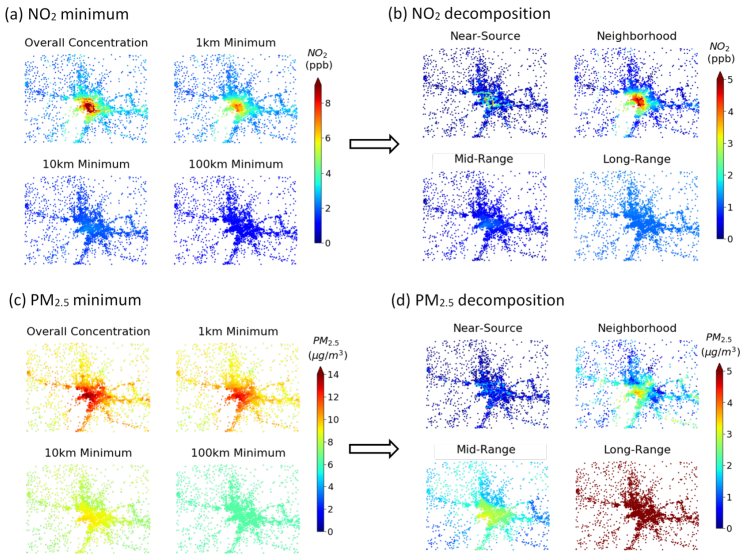

Figure 2.19: The buffer minimum plots and decomposition plots of NO_2 and $PM_{2.5}$ for Tuscaloosa area: (a) Predicted NO_2 annual average concentration and corresponding 1, 10, and 100 km buffer minimum maps; (b) Decomposition maps of NO_2; (c) Predicted $PM_{2.5}$ annual average concentration and corresponding 1, 10, and 100 km buffer minimum maps; (d) Decomposition maps of $PM_{2.5}$.

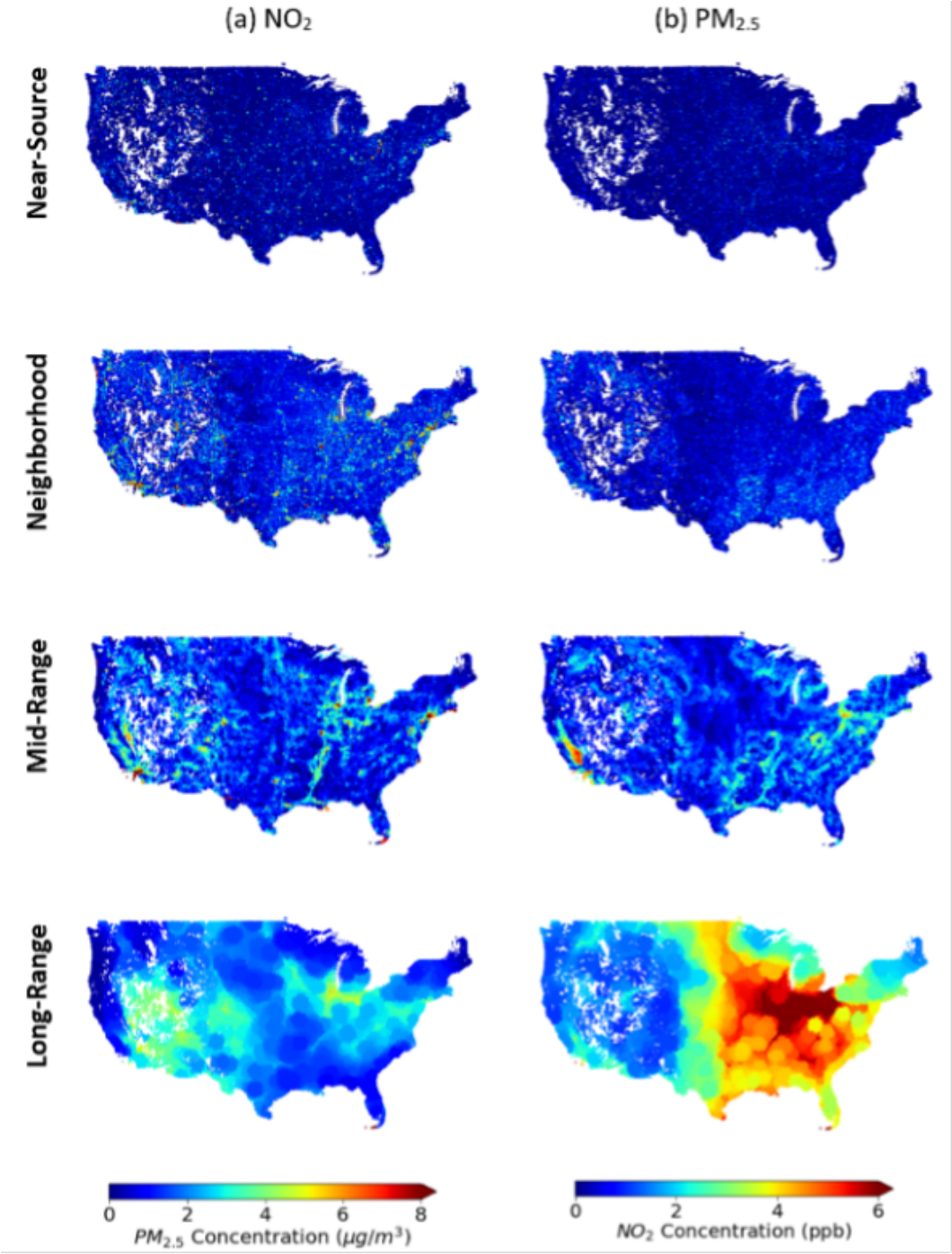

Figure 2.20: Decomposition maps of: (a) NO_2 and (b) $PM_{2.5}$ for all Census block centroids in the contiguous US.

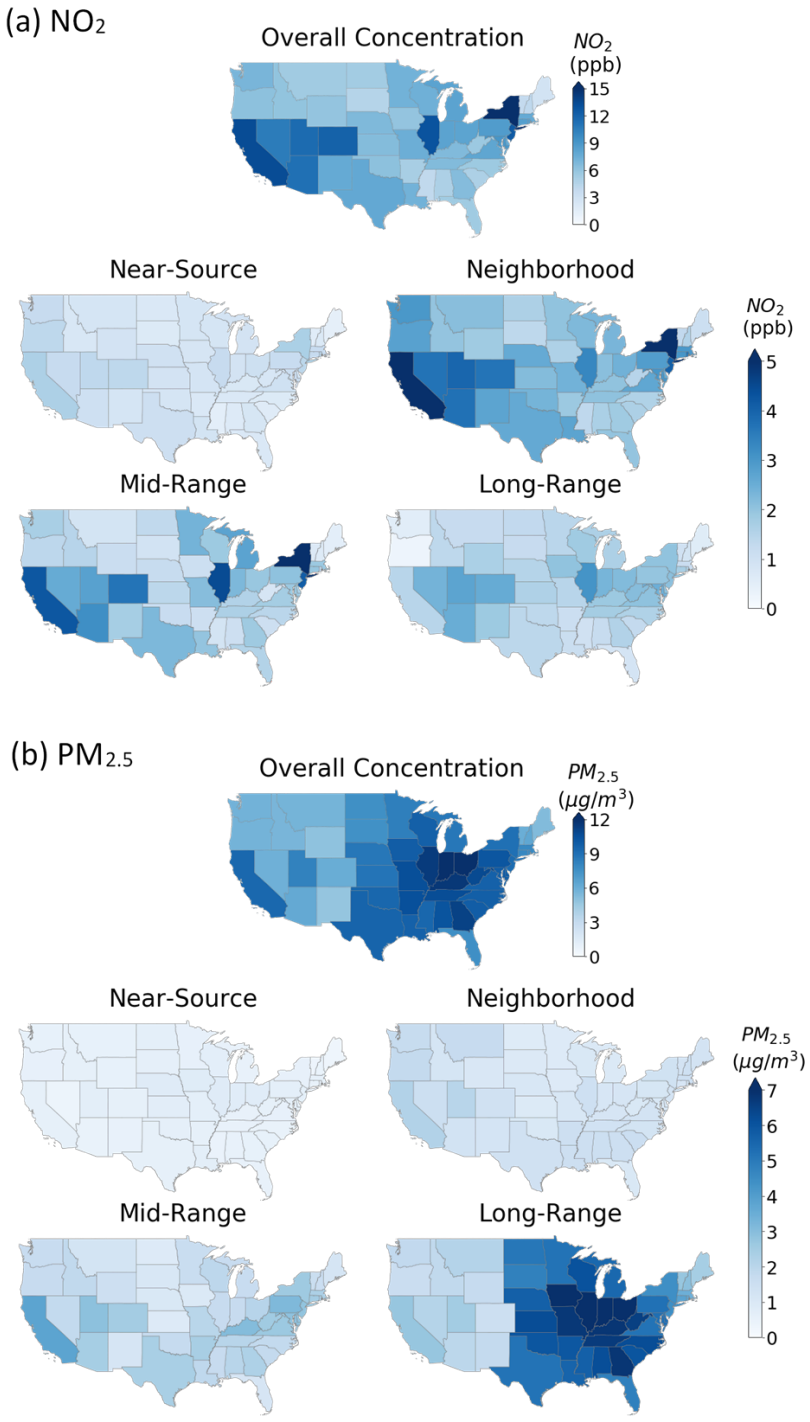

Figure 2.21: Population-weighted averages of overall and decomposed concentrations for (a) NO_2 and (b) $PM_{2.5}$ at ∼6 million census blocks by state.

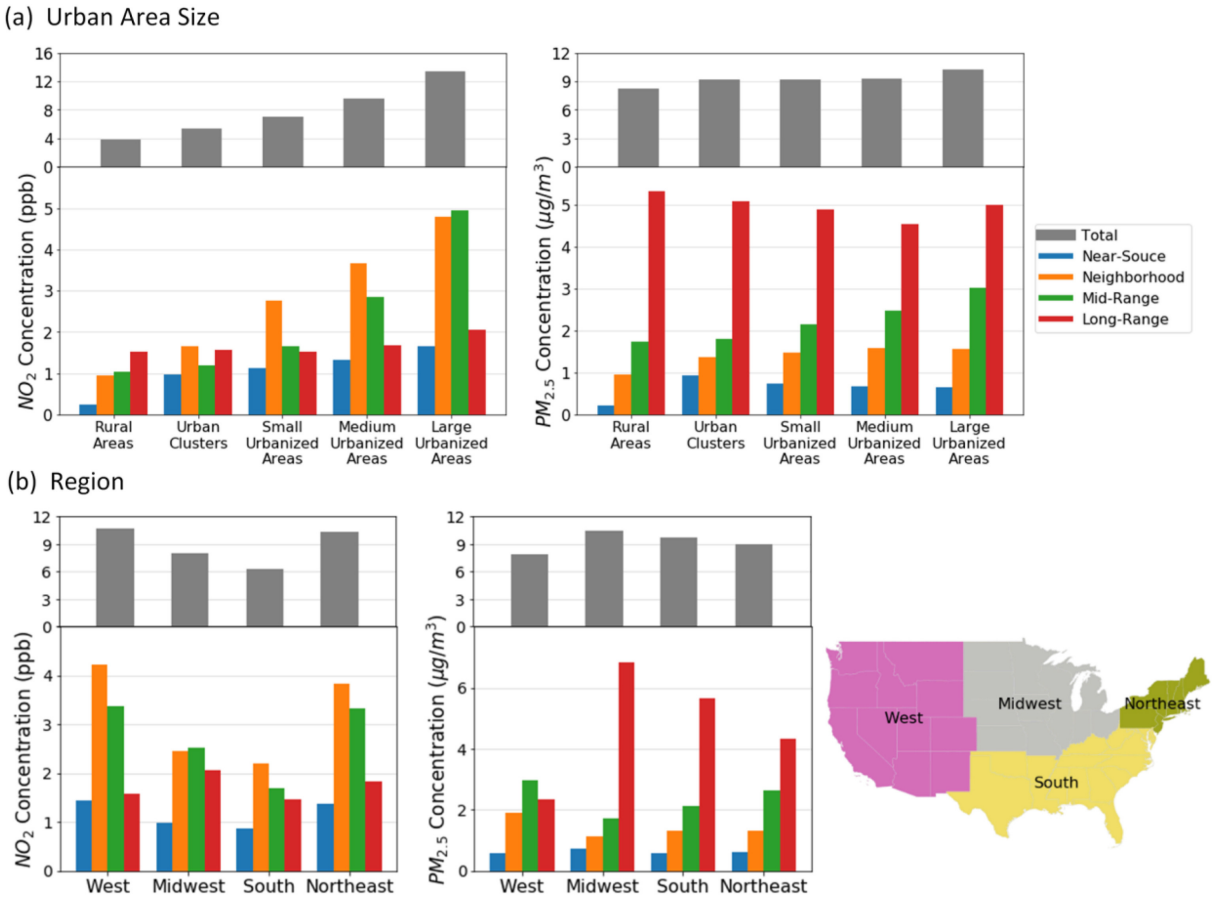

Figure 2.22: Population-weighted averages of overall (grey bars) and decomposed concentrations (colored bars) for NO_2 and $PM_{2.5}$ at $\sim$ 6 million census blocks by (A) urban area size; (B) US region.

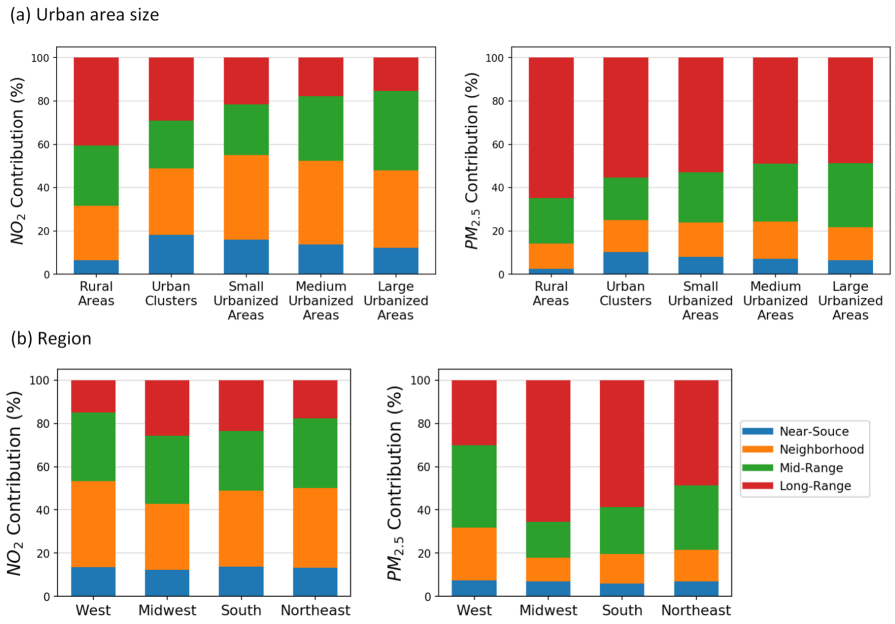

Figure 2.23: Decomposed components contributions for NO_2 and $PM_{2.5}$ by (a) urban area size and (b) geographic region.

Figure 2.24: Decomposed concentrations for NO_2 by the interaction of urban area size and geographic region (20 categories in total).

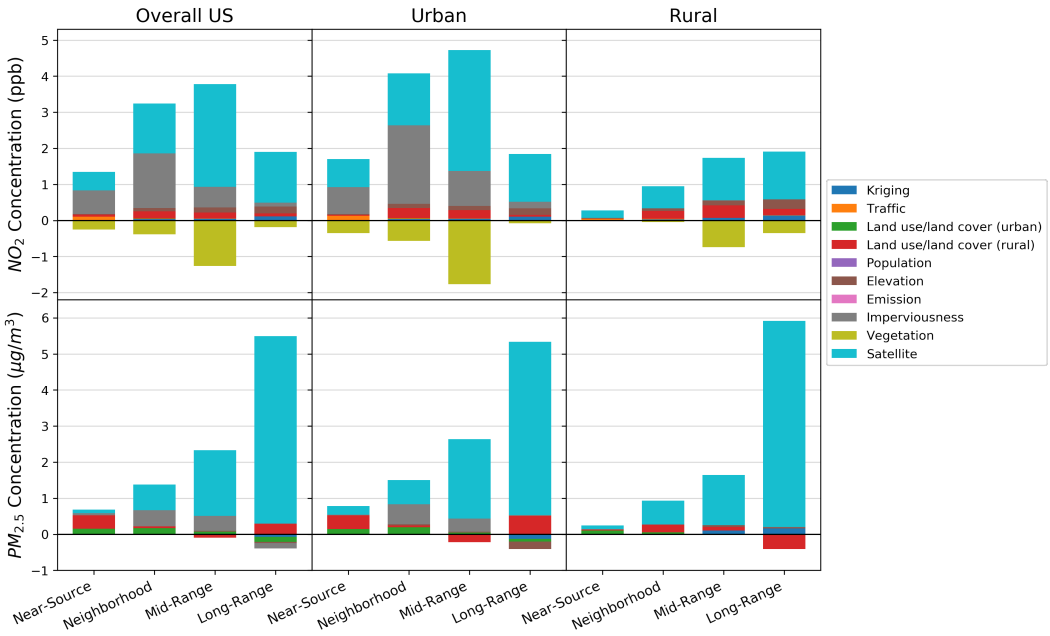

Figure 2.25: Contributions of each empirical regression term for the overall pollution concentration to the four decomposed concentrations for NO_2 and $PM_{2.5}$ for the overall US, all urban blocks and all rural blocks.

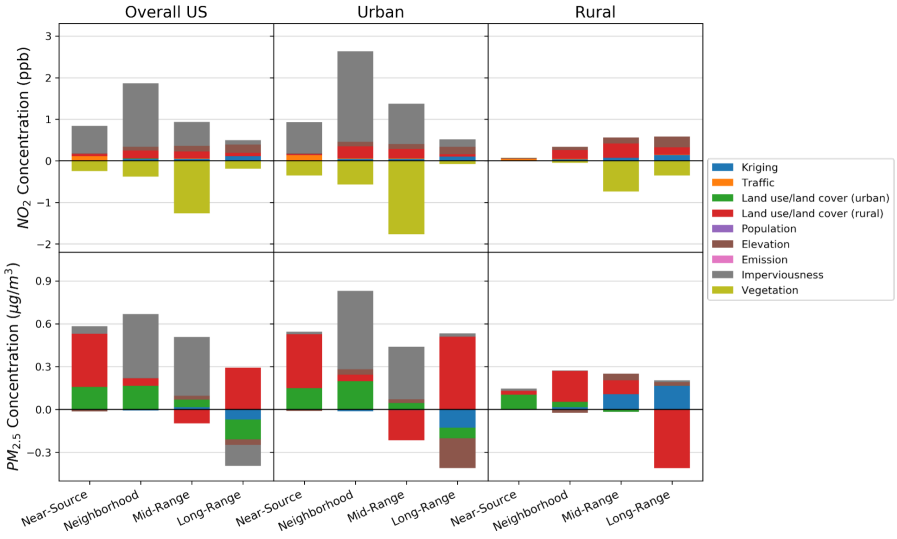

Figure 2.26: Contributions of each empirical regression term for the overall pollution concentration to the four decomposed concentration for NO_2 and $PM_{2.5}$, without satellite variables, for the overall US, all urban blocks, and all rural blocks.

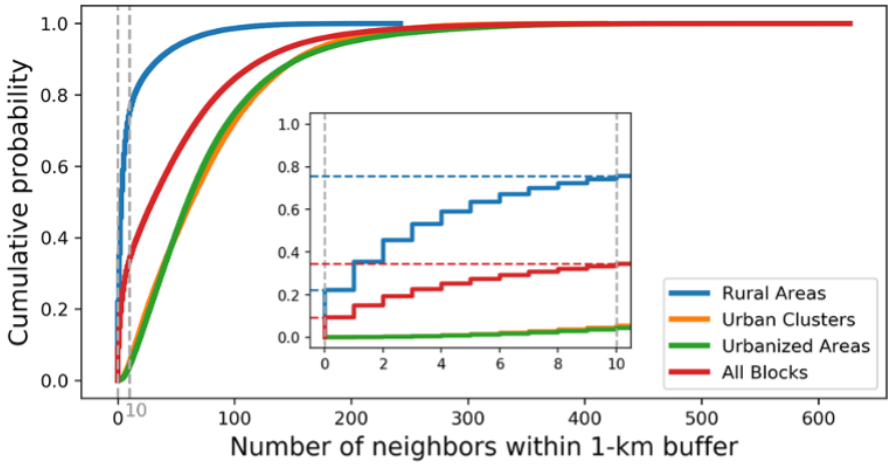

Figure 2.27: Cumulative probability of number of neighbors within 1-km buffer in rural areas, urban clusters and urbanized areas.

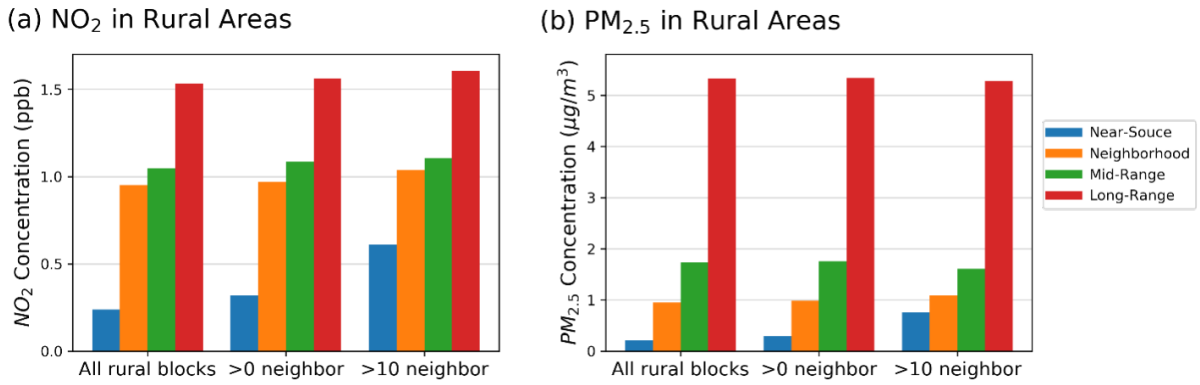

Figure 2.28: Population-weighted average decomposed concentrations for (a) NO_2 and (b) $PM_{2.5}$ for: (1) All block centroids in rural areas; (2) block centroids in rural areas with at least one nearby block centroid within 1-km circular buffers; and (3) block centroids in rural areas with at least ten nearby block centroids within 1-km circular buffers. The two plots show that after adding some thresholds (>0 neighbor, >10 neighbor) to the block centroids in the rural areas, the weighted averages of "near-source" increased to 2∼3-fold of the no-threshold values for both two pollutants. While we cannot discern whether increases are caused by the underestimation from our methods in sparsely populated regions or by the removing of locations with inherently less short-range (or local) sources, the "near-source" values in the restricted estimates (>0 neighbor, >10 neighbor) for rural areas are similar to or lower than for urban clusters (Fig. 2.22), and likely reflect an upper bound on the "near-source" contribution in rural locations.

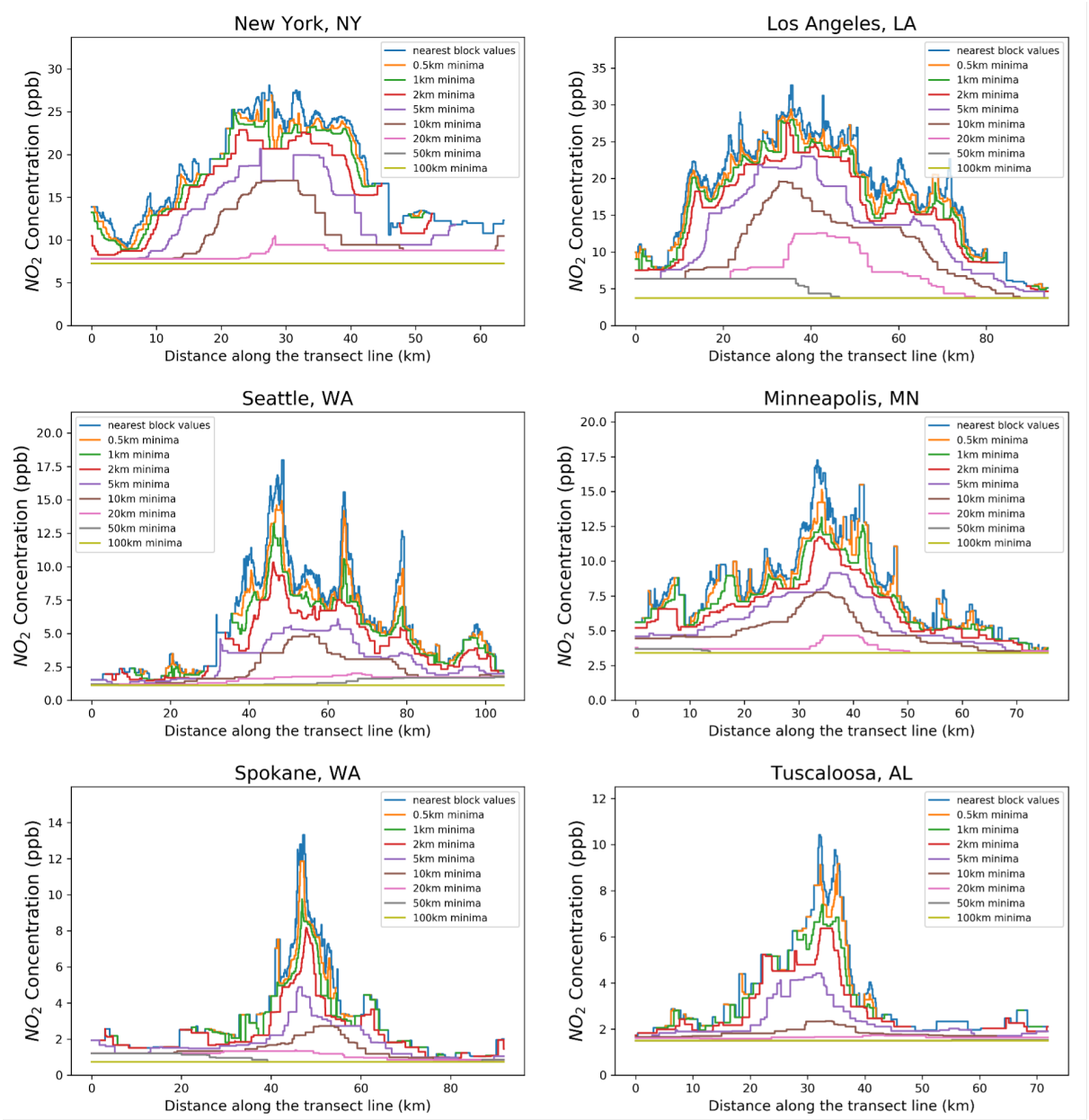

Figure 2.29: NO_2 decomposition results along transect line on multiple buffer radii (0.5 km, 1 km, 2 km, 5 km, 10 km, 20 km, 50 km, 100 km) for six cities.

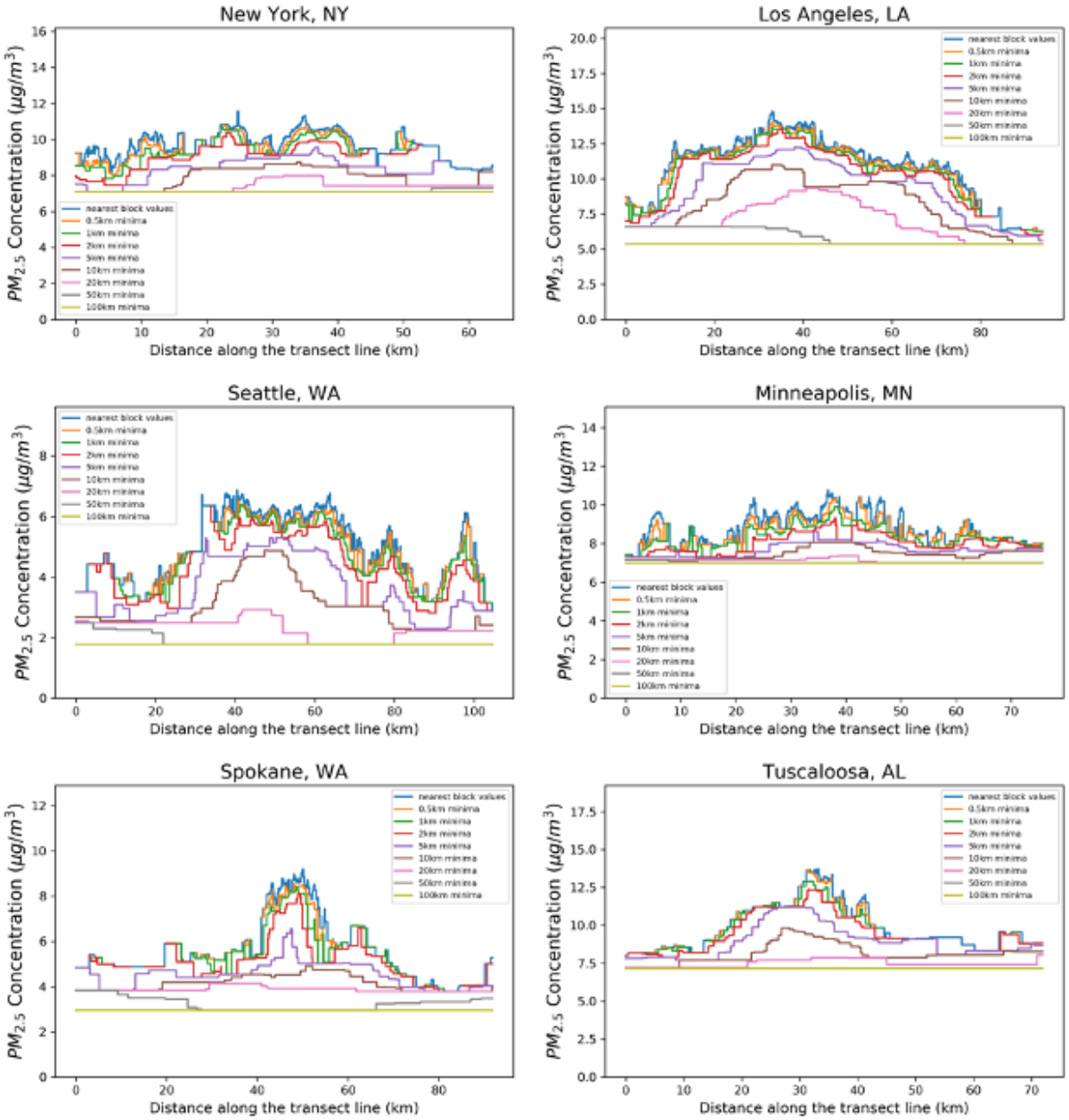

Figure 2.30: $PM_{2.5}$ decomposition results along transect line on multiple buffer radii (0.5 km, 1 km, 2 km, 5 km, 10 km, 20 km, 50 km, 100 km) for six cities.

Chapter 3

AMBIENT AIR POLLUTION AND SOCIOECONOMIC STATUS IN CHINA

Material in this chapter was previously published in the following article: Y. Wang, Y. Wang, H. Xu, Y. Zhao, J. D. Marshall, Ambient air pollution and socioeconomic status in China, *Environ. Health Perspect.* **130**, 067001 (2022).

3.1 Summary

Air pollution disparities by socioeconomic status (SES) are well documented for the United States, with most literature indicating an inverse relationship (i.e., higher concentrations for lower-SES populations). Few studies exist for China, a country accounting for 26% of global premature deaths from ambient air pollution. The objective of this chapter is to quantify the relationship between ambient air pollution exposures and SES in China. To do so, I combined estimated year 2015 annual-average ambient levels of NO_2 and $PM_{2.5}$ from empirical models with national demographic information, which are derived from both gridded gross national product (GDP) and China Health and Retirement Longitudinal Study (CHARLS) cohort. The results indicate that ambient NO_2 and $PM_{2.5}$ levels were higher for higher-SES populations than for lower-SES population, higher for long-standing urban residents than for rural-to-urban migrant populations, and higher for the majority ethnic group (Han) than for the average across nine minority groups. For three SES measurements (individual SES score, community-averaged SES score, gridded GDP per capita), a 1-interquartile range higher SES corresponded to higher concentrations of 6-9 $\mu g/m^3$ NO_2 and 3-6 $\mu g/m^3$ $PM_{2.5}$; average concentrations for the highest and lowest 20th percentile of SES differed by 41–89% for NO_2 and 12–25% for $PM_{2.5}$. This pattern held in rural and urban locations, across geographic regions, across a wide range of spatial resolution, and for modeled vs. measured pollution concentrations. My findings are consistent with the idea that in China's current industrialization and urbanization stage, economic development is correlated with both SES and air pollution. To my knowledge, this chapter provides the most comprehensive picture to date of ambient air pollution disparities in China; the results differ dramatically from results and from

theories to explain conditions in the United States.

3.2 Introduction

Ambient air pollution causes ∼4 million deaths per year [1–3], yet the disease burden is not evenly distributed across individuals, communities, countries, regions, or demographic groups. Between-country disparities in ambient air pollution are well documented [1,5]; in contrast, within-country disparities, including how ambient pollution levels correlate with SES and other demographic attributes, are poorly studied other than in the United States and a few other high-income countries. Much of the existing literature documents disparities for specific locations (e.g., in a specific city). A smaller body of literature documents disparities nationwide for the United States [7, 8, 10, 156, 157]. The vast majority of the literature indicates, for North America, higher pollution levels for low-SES than for high-SES communities and individuals [7–12]; evidence for European countries is more limited and suggests a mixed relationship [9, 70]. Limited knowledge exists for China or other low- or middle-income countries [9], where 91% of the premature deaths from outdoor air pollution occur (26% are in China) [1, 6]. Less than 20% of the Chinese population lives in cities that meet the national annual $PM_{2.5}$ standard (the GB 3095-2012 standard; 35 $\mu g/m^3$) [66], and none live in cities meeting the World Health Organization's (WHO) annual guideline (5 $\mu g/m^3$) [67–69]. Publicly available census data for China are relatively coarse [county level; average per county: 0.4 million people; 2,642 km^2]; this limits the potential for some census-based demographic/air pollution analyses in China that, in the United States, are common.

The environmental inequality patterns in the United States are generally explained in part by past and present racial discrimination and race- and class-based market dynamics: Low-SES communities lack social capital and political power and access and are therefore less able to stop locally undesirable land-uses such as highways, industry, and other sources of pollution [9, 11, 20, 25, 31, 35]. Longstanding residential segregation by race and income adds to the potential for disparities [33]. Racism by private individuals (e.g., racial covenants on the sale of property) and systemically by private companies (e.g., banks' lending practices) and by local, state, and national government (e.g., redlining) in the past and present has supported and accelerated exposure disparities [35,38,158]. The net result in the United States is disproportionate air pollution burdens for low-SES communities and individuals, with well-documented theory to explain those disparities.

Yet, theories developed in the United States may not be applicable for China and elsewhere

owing to the different historical, social, economic, political, urbanization, and industrialization characteristics [13–15]. For example, others have found that environmental inequality patterns in European countries are different from the United States owing to the lower extent of social segregation and the greater tendency (relative to in the United States) for lower-SES groups to live on the outskirts of the city [70]. Urban expansion patterns [71], local housing policy [72], and discrete choice between benefits/amenities and negative aspects of the environment [73] are also found to influence the exposure inequality patterns in Europe. A study in India proposed an environmental Kuznets curve [EKC; i.e., a theory that environmental degradation first rises and then falls with an increasing per capita income [74–76]] -like pattern (i.e., a U-shaped relationship), where marginalized communities are excluded from negative and positive externalities of industrial development [77].

As the world's largest developing economy, China has in recent decades experienced rapid urbanization [159,160], increased income disparities [161,162], and widespread rural-to-urban migration [160,163]. Secondary (manufacturing) industry represents a large proportion of gross national product (GDP) in China (in 2018, 41% in China vs. 19% in the United States); cities or places with more industry and manufacturing jobs tend to be of higher income, on average, and potentially also experience more air pollution [164,165]. In addition, with comparatively (relative to in the United States, on average) larger city size and population density, more-centralized urban structure, and poorer traffic conditions [166, 167], people in China tend to live (on average, relative to in the United States) closer to their workplace [168]. Those strong aspects in regional and urban form reflect differing housing preferences and civil infrastructure (e.g., transportation systems) [169]; they also reflect differences in history in China relative to the United States (e.g., China did not experience "White flight"—an exodus of affluent White people in the United States, largely starting in the 1950s and 1960s, from urban to suburban communities [38]). The net result is, in China, a comparatively higher concentration of high-SES population in urban centers (often with higher air pollution levels) and lower-SES population in outskirts and suburban areas [13,170,171]. These many conditions support the hypothesis that, unlike in the United States, in China air pollution exposures might be higher for high-SES than for low-SES populations.

Previous studies in China have investigated air pollution inequality with respect to specific sources, including industrial emissions [15,78,79], emissions from electric vehicles [56], and household consumption [80]. Studies on ambient concentrations [13, 81–83] have focused on a single city or

yielded inconclusive results [84–86]. The lack of comprehensive research on environmental inequality in China, including theories and national empirical studies of SES and ambient air pollution, are important gaps in the literature.

In this Chapter, I put forward an alternative hypothesis (i.e., distinct from the EKC theory, described above) that, in China, higher-SES populations are more exposed to ambient air pollution than are lower-SES populations. I test the hypothesis using ambient annual-average NO_2 and $PM_{2.5}$ levels in China. To test my hypothesis, I conduct multiple national investigations of air pollution disparities by SES, including for individual- and area-level measures of SES, based on ambient air pollution levels for two pollutants: NO_2 and $PM_{2.5}$.

3.3 Materials and methods

3.3.1 Individual SES characteristics and migration status

To obtain spatially explicit demographic information on a group of individuals, I employ data from the 2015 China Health and Retirement Longitudinal Study (CHARLS) [172]. CHARLS is an interview-based nationally representative survey of people in China $\geq$ 45 years of age and their spouses (n=21,095; see [173] for CHARLS locations). All participants in CHARLS gave informed consent, and the protocol was approved by the ethical review committee at Peking University. I obtained each individual's highest education attainment (hereafter, education), occupation, and household annual per capita income in yuan (hereafter, income) from the survey. Education was classified as a) illiterate, b) sishu/home school and below, c) elementary school, d) middle school, or e) high school and above. Occupation was categorized according to people's lifetime profession instead of the current working status (given that some people were retired): Individuals who have only done agricultural work (have never done nonagricultural work for >10 d) in their lifetime were classified as engaging in agricultural work; those who have ever done nonagricultural jobs for >10 d (no matter whether they also have done agricultural work) were classified as engaging in nonagricultural work. Income was calculated for each household (shared by the sample individuals on the same household) and is presented as the net post-tax income. The household total annual income includes a) household members' wages, bonus incomes, or pensions, and b) household agricultural, self-employed activities, public transfer, and other types of transferred income (e.g., from parents, children, relatives). Household per capita income was calculated by dividing the

household total annual income by family size.

In addition to the three SES variables that I mainly focused on (education, occupation, income), I also investigated each individual's ethnicity (sometimes also translated as "nationality," but referring to group identification, not citizenship) and household per capita living expenditure (hereafter, expenditure). According to the 2020 census [174], China is officially composed of 56 ethnic groups, with Han as the dominant group (91.1% of the total population) and the other 55 groups as minorities (combined, 8.9% of the total population). Clusters of people who are ethnic minorities are found in the bordering northwest (e.g., Uygur), north (e.g., Mongol), northeast (e.g., Manchu), south (e.g., Zhuang), and southwest (e.g., Tibet), with some in the central interior (e.g., Hui). In the CHARLS survey, there are 10 ethnic categories: Han (92.3% of CHARLS respondents), Hui (0.5%), Zhuang (1.0%), Uyghur (0.5%), Yi (0.5%), Tibet (0.9%), Miao (0.6%), Mongol (1.1%), Dai (0.2%), and other (2.6%) [175]. Expenditure was calculated for each household (household annual total expenditure divided by family size) using the CHARLS survey. Total expenditure included weekly expenditure on food and restaurants; monthly expenditure on communication, utilities, transportation, household and personal items, entertainment, and housekeepers; and yearly expenditure on clothing, traveling, heating, furniture, education, medical, fitness and beauty, automobile, taxes and donations, and others.

I separately considered each individual's rural/urban migration status by combining each respondent's household location (urban districts [urban] vs. rural counties [rural] within the prefecture-level cities) and hukou status. Hukou is the official household registry system that records a citizen's location-of-origin and determines local residence rights, such as medical care, unemployment benefits, school enrollments, and public housing [69, 163]. Importantly, hukou is a) typically unchanged during a person's life even if that person moves and b) assigned as nonagricultural (i.e., urban) or agricultural (i.e., rural). Since 2014, the Chinese government has in some cases initiated the canceling of the hukou system; although the binary hukou classification still dominates in most places, some people's hukou is "unified residence," i.e., a single/combined category that sidesteps the urban/rural distinction. Survey respondents who indicated unified residence hukou (n=346, 1.6%) or no hukou (n=31, 0.1%) have been excluded here. Because rural-to-urban migrants cannot enjoy many social benefits (e.g., medical care), the hukou system discourages that migration and also puts such migrants at a disadvantage relative to urban residents with a nonagricultural hukou [176–180]. I employed three rural/urban migration categories: urban resident (i.e., with a

nonagricultural hukou); rural-to-urban migrant (i.e., an urban resident with an agricultural hukou); and rural resident (i.e., with an agricultural hukou). An urban-to-rural migrant—i.e., someone who lives in a rural area but with a nonagricultural hukou—is uncommon (<3% of the data set) and not included in this analysis.

I also considered population density at the individual's household location as a control variable in my analysis. Population density was determined at the community level, using the population count in the community divided by total area of the community. The community information was collected from the CHARLS 2011 community survey [181].

3.3.2 Constructing three SES matrices

As described next, I obtained three SES matrices to represent demographic conditions at the individual and areal levels in China, and then compared the results across methods. Using multiple independent approaches to quantifying SES sheds additional light on the questions considered and informs whether results are robust to the methods employed.

1. An individual-level SES was derived from CHARLS data using a standardized SES score reflecting education, occupation, and income; the approach reflects factor analysis of mixed data (FAMD), which is a principal component analysis method applicable to data sets with both quantitative and qualitative variables [182, 183]. I use the "MissMDA" package in R to handle missing values (see above for category classifications), then "FactoMineR" to perform FAMD: education and occupation as categorical variables, income (log-transformed) as quantitative. An individual's SES (n=21,095) is defined here as the normalized scores of the first FAMD dimension (explaining 30% of the total variance; see Fig. 2.1 and Table 3.1 for weights and contributions for income and each category of education and occupation). A higher score represents a higher SES.

2. An area-level SES was derived from CHARLS data by aggregating individual SES to the community-level average. "Community" is a formally defined geographic unit, used in the census and for mail systems; it is the smallest de facto administration level in China. As defined in the 2011 (baseline year) CHARLS survey, CHARLS individuals are from 450 communities, including 237 rural communities [also termed villages; located in rural counties] and 213 urban communities (i.e., located in urban districts).

3. A second area-level SES reflects per capita GDP (1-km^2 resolution). These data were derived by combining year 2015 national gridded GDP predictions based on nighttime lights and population images [184] (1-km^2 resolution) and the WorldPop year 2015 population density data set [185] (30 arc seconds resolution, or $\sim$1km^2 at the equator). I resampled the population density data set to match the coordinates of the gridded GDP data set using a bilinear interpolation approach and then divided the gridded GDP by the resampled population density to calculate the GDP per capita. To investigate the air pollution disparities at multiple spatial resolutions, I also resampled the raster layer of GDP per capita and averaged it to different grid sizes (2, 5, 10, 20, 50, and 100km; after removing grid cells with zero or missing values, the number of grid cells for the 1-km grid and each of those six grid sizes were n=3,861,463, n=1,058,301, n=204,565, n=61,863, n=18,985, n=3,953, and n=1,220, respectively).

As mentioned above, one of the motivations for using three distinct approaches was to avoid the spatial coarseness of census data. Other motivations included shedding deeper light on the question and informing whether the result is robust to the different methods employed. Each of these three demographic data sets was then combined with ambient air pollution concentrations, as described next.

3.3.3 NO_2 *and* $PM_{2.5}$ *ambient concentrations*

For all three SES data sets, I employed as my ambient air pollution metric the annual-average concentrations of two pollutants (NO_2 and $PM_{2.5}$). Pollution concentrations were derived from national empirical models for China [68]. The models incorporated monitoring data, satellite observations, universal kriging, and predictor data, such as land-use, traffic, and meteorological data (Table 3.2). Model predictions were year 2015 annual-average concentrations at 1-km^2 resolution [9.6 million cells; 10-fold cross-validation R^2: 0.78 (NO_2), 0.89 ($PM_{2.5}$)] [68]. In sensitivity analyses (below), I used the monitoring data directly [i.e., excluding the empirical models of [68]].

For the individual-level CHARLS analyses, concentrations were estimated using the empirical models for that individual's residential location (longitude and latitude recorded during the interview). For community-level CHARLS analyses, concentrations were estimated using the average individual-level concentration within each community. For GDP per capita, I resampled the GDP per capita data set to the same grid as the pollution data using bilinear interpolation and then

matched the two data sets to derive GDP per capita on the 1-km pollution grid. The CHARLS demographic data, GDP per capita data, and pollution estimation data from the empirical models all were for year 2015. For sensitivity analyses regarding spatial resolution, I resampled the pollution data into the 2-, 5-, 10-, 20-, 50-, and 100-km grids, as described above for GDP per capita.

3.3.4 Quantifying air pollution disparities by each SES factor

I focused on NO_2 and $PM_{2.5}$. Those two pollutants were selected because they are important and widely tracked (both are criteria pollutants); both have important health effects associated with exposures (e.g., [4, 87–89]; and, importantly, publicly available national models exist for the two pollutants [68]. Investigating two pollutants provides additional testing of the hypothesis investigated.

For an individual's migration status and each factor that contributes to SES (i.e., education, occupation, and income), I determined relationships with ambient concentrations, considering both unadjusted and fully adjusted effects. Unadjusted effects were determined by comparing the means and quantiles (10th, 25th, 50th, 75th, and 90th) of NO_2 and $PM_{2.5}$ concentrations for the samples in each migration and SES group (here, I grouped income using quintiles; any missing group for each variable was also included in my analysis). Specifically, for migration status and SES factors, I calculated the maximum differences (both absolute and relative differences, with absolute values) in NO_2 and $PM_{2.5}$ concentrations between the group averages and the population averages. I also determined the unadjusted effects for population density at household locations (grouped into tertiles).

Adjusted effects were determined using two sets of regression models: One set comprised multivariate ordinary least squares regression models that included only individual-level factors (migration status, education, occupation, income, and age) as independent variables; the other set comprised multilevel mixed-effects linear regression models that also included prefecture-city (usually including both urban districts and rural counties) random intercepts to control for city clustering effects. In the multilevel mixed-effect models, I also included population density at the household locations as a control variable. In both models, migration status, education, and occupation were treated as categorical variables. To achieve normality, income data were log-transformed and then standardized to the whole study population; parameter estimates referred to a 1-unit increase in the z-score for the log of income. All regressions were run only on individuals with complete data

[n=15,197; 72% of the CHARLS cohort; incomplete data (n=5,898; 28%) were excluded from the regressions]; results are reported in terms of best-estimate values and 95% confidence intervals (CIs). To investigate potential bias due to missing data, I performed sensitivity tests using the multiple imputation approach [186], which creates several imputed data sets by replacing missing values with imputed values and combining the results obtained from each of them. I used the mice package in R to generate five imputed data sets using probable means methods and then calculated the pooled regression results of the five imputed data sets for both individual-level multivariate regression models and prefecture-city random intercepts regression models.

In sensitivity tests, I also investigated the unadjusted and adjusted relationships between ethnicity and household per capita living expenditure and ambient concentrations. For ethnicity, the unadjusted effects were determined for the 10 ethnic groups; for the adjusted effects, I created a dummy variable for ethnic minorities (i.e., individuals who are not Han), and included it in individual-level multivariate regression models and multilevel mixed-effects regression models with prefecture-city random intercepts. For expenditure, the unadjusted effects were determined for each quintile; for the adjusted effects, I included the log-transformed and standardized (to the whole study population) expenditure in both individual-level and multilevel models.

3.3.5 Metrics quantifying air pollution disparities by individual SES score, community SES score, and gridded GDP per capita

I quantified disparities using two metrics. The first metric involved linear regressions of concentration on SES score or GDP per capita (log scale). Here, I calculated the regression slopes multiplied by the interquartile ranges (IQRs) of SES (SES score, log of GDP per capita) to quantify the air pollution disparities between high- and low-SES groups. The second metric involved the (absolute and relative) disparity in mean NO_2 and $PM_{2.5}$ concentrations between the population with the highest 20% and the lowest 20% SES (population weighted for GDP per capita). To better visualize the disparity patterns for each subgroup of the total population (see below for subgroup classifications), I also calculated the mean pollution concentration and mean SES (SES score, log of GDP per capita) for each 10% of the subsample.

For the individual-level SES score, disparities were quantified on the overall populations as well as separately for three values for migration status. In sensitivity analyses, I separately quantified disparities in five geographic regions in China (Fig. 2.2); this approach reflects geographic

variability as well as, implicitly, economic development and climate conditions. For the community-averaged SES score, I quantified the disparities across all communities and by urban communities and rural villages separately. For gridded GDP per capita, I quantified the disparities on overall grids and by urban/rural grid cells. Urban/rural cells were defined according to the spatial cities of China in 2015 from Beijing City Lab [187] using community boundaries and urban built-up areas. The urban/rural classifications were done using the mask function in the Python "rasterio" package (mask 1-km gridded GDP using urban definition shapefiles). Grid cells a) with the centers inside the urban definition shapefiles or b) selected by Bresenham's line algorithm (which determines which points on a two-dimensional raster should be selected to form a close approximation to a straight line between two given points) were designated as urban; all other grid cells were defined as rural. To investigate whether the results were robust to the specific urban/rural definition employed, I also conducted sensitivity tests using three alternative urban/rural land classifications: a) township population density in year 2010 [188], b) a neighborhood-level vector cellular automata model (this model reproduces global patterns and behavior from local interactions of cells, representing the cells as a collection of interconnected irregular geographic objects) based on density, neighborhood condition, and other spatial variables for year 2012 [189], and c) a 300-m resolution global land cover map (GLOBCOVER) data set [190], which is based on satellite images. All four urban definition shapefiles were downloaded from Beijing City Lab [191]. In my approaches, individuals/communities/grids were included in either urban or rural areas.

In addition, separately, I conducted sensitivity analyses quantifying the disparities using six alternative spatial resolutions from 2 to 100 km (vs. 1 km in the main analysis). The motivation was to see whether environmental inequality patterns differed across spatial scales of analysis.

3.3.6 Sensitivity test using monitored concentrations

To test the robustness of results using modeled concentrations, as a sensitivity analysis I employed monitoring data directly (i.e., excluding the empirical model). By definition, this analysis was restricted to locations with a monitor (i.e., the 1-km grid cells containing a monitor; n=1,466, see Fig. 2.3 for the spatial distribution of the monitor locations). The population-weighted average concentrations (in micrograms per meter cubed) for NO_2 and $PM_{2.5}$ (37 and 55 $\mu g/m^3$, respectively) at monitor locations were higher than the national averages (28 and 53 $\mu g/m^3$), especially for NO_2. However, the GDP per capita (population weighted; in yuan) at monitor locations (mean=16,847;

median=13,408) was lower than the population average (mean=59,080; median=24,691). Below, I give separate results for all monitor locations (n=1,466), urban monitor locations (n=1,076; 73% of monitors), and rural monitor locations (n=390; 27%). Comparing the modeled vs. monitored concentrations, the model performance was similar in both urban and rural locations for both two pollutants [regression R^2: 0.88 (NO_2), 0.95 ($PM_{2.5}$) for urban monitor locations, and 0.89 (NO_2), 0.95 ($PM_{2.5}$) for rural monitor locations].

3.4 Results

3.4.1 Air pollution exposures by individual's SES

Based on the individual data (CHARLS cohort), average ambient air pollution concentrations at home locations were 24 $\mu g/m^3$ (IQR: 18–31) for NO_2 and 51 $\mu g/m^3$ (38–60) for $PM_{2.5}$. Those values are several times higher than the WHO annual mean guidelines for NO_2 (10 $\mu g/m^3$) and $PM_{2.5}$ (5 $\mu g/m^3$) [6]. The proportion of the population exceeding the WHO guideline was 97% for NO_2 and 100% for $PM_{2.5}$.

Consideration of individuals' urban migrant status and SES groups (Fig. 2.4; Table 3.3) revealed the following. For NO_2, mean concentrations for rural residents (22 $\mu g/m^3$) and rural-to-urban migrants (26 $\mu g/m^3$) were lower than for urban residents (31 $\mu g/m^3$); for $PM_{2.5}$, mean concentrations were nearly identical for rural residents (50 $\mu g/m^3$) and rural-to-urban migrants (also 50 $\mu g/m^3$; a <2% difference), which were slightly lower than for urban residents (55 $\mu g/m^3$). In univariate consideration of all three SES variables (education, occupation, and income), NO_2 and $PM_{2.5}$ concentrations were generally higher for higher-SES groups (i.e., higher education, non-agricultural work, higher income) than for lower-SES groups, which is consistent with the hypothesis described in the "Introduction" section. The maximum concentration disparity percentage (with absolute value) between group means and the population average was 28% for migrant status and 12–18% for SES groups, for NO_2; for $PM_{2.5}$, analogous relative disparities were 8% and 3–7%. NO_2 and $PM_{2.5}$ concentrations were both higher for individuals in middle- and high-population density communities and lower for low population density communities (Table 3.3). Air pollution disparities by SES variables generally held even controlling for migrant status (Fig. 2.5). In individual-level fully adjusted models and city random intercept regression models (Table 3.4), the generally positive relationships still held (except for education in the city random intercept model

for $PM_{2.5}$, which had nonsignificant negative slopes for higher education groups), and many but not all SES variables were statistically significant. In multiple imputation models (Table 3.5), the positive relationships generally held; the regression coefficients for both individual-level models and multilevel models did not change substantially (differences were generally <10%).

Results by ethnicity (Tables 3.6 and 3.7) indicate that, relative to the population-average exposure for NO_2 and $PM_{2.5}$ (24 and 51 $\mu g/m^3$, respectively), average exposures were nearly the same for people who are Han (25 and 52 $\mu g/m^3$, respectively; that finding was expected because the overall population is 92% Han), higher for people who are Hui (39 and 67 $\mu g/m^3$, respectively) or Uyghur (58 $\mu g/m^3$ for $PM_{2.5}$), and lower for people in each of the remaining seven groups [and, for Uyghur for NO_2 (21 $\mu g/m^3$)]. (Average SES scores were higher-than-average for people who are Hui or Uyghur, and lower-than-average for the remaining seven minority groups; thus, results by ethnicity were generally consistent with the finding above that average NO_2 and $PM_{2.5}$ exposures in China were higher than average for upper-SES individuals.) Considering all people in any one of the nine ethnic minority groups, average exposures were 25% lower (NO_2) and 24% lower ($PM_{2.5}$) than the overall population average. For the adjusted effects models, including ethnic minority status as a dummy (i.e., binary) variable yielded a negative coefficient for both pollutants in the individual-level fully adjusted models and a nonsignificant coefficient in the city random intercept models. That finding implies that the lower ambient concentrations for ethnic minorities are mainly due to the concentration differences in the cities they live (i.e., a between-city, rather than a within-city, effect).

The unadjusted results by expenditure (Table 3.8) indicate that NO_2 concentration was higher for individuals with higher expenditure (23 vs. 26 $\mu g/m^3$, respectively, for the 20% of individuals with the lowest and highest expenditure); there was no significant relationship between $PM_{2.5}$ concentration and expenditure (50–51 $\mu g/m^3$ for all quintiles of expenditure); the mean SES scores were higher for individuals with higher expenditure. For the adjusted effects models (Table 3.7), expenditure had a slightly positive coefficient for NO_2 in the city random intercept model, a negative coefficient for $PM_{2.5}$ in the individual-level model, and nonsignificant coefficients in the two other models ($PM_{2.5}$, city-random-intercept; NO_2 individual-level model). The generally positive relationships for urban migration status and the three major SES variables (education, occupation, and income) still held after controlling for ethnicity and expenditure.

3.4.2 Air pollution exposures by three SES matrices

The three SES approaches (Fig. 2.6) revealed a consistent pattern: Ambient air pollution concentrations at home locations were higher for high-SES than for low-SES individuals or areas across methods, pollutants, and urban–rural status, again suggesting that the findings here are generally consistent with the hypothesis offered in the "Introduction" section. According to linear regression models of concentrations vs. SES, at 1-km resolution, a 1-IQR higher SES corresponded to a higher concentration of 5.6 μg/m^3 NO_2 (95% CI: 5.4, 5.9), 3.5 μg/m^3 $PM_{2.5}$ (95% CI: 3.0, 3.9) for individual-level data (Table 3.9); 9.4 μg/m^3 NO_2 (95% CI: 7.9, 10.8), 6.0 μg/m^3 $PM_{2.5}$ (95% CI: 3.3, 8.8) for community-level data (Table 3.10); and 7.3 μg/m^3 NO_2 (95% CI: 7.1, 7.4), 4.1 μg/m^3 $PM_{2.5}$ (95% CI: 4.1, 4.2) for GDP per capita (log scale) (Table 3.11). Those results, rounded and presented as a range, suggest that a 1-IQR higher SES is associated with a range of 6–9 μg/m^3 NO_2 and 3–6 μg/m^3 $PM_{2.5}$ higher concentration of air pollution (expressed as a percentage of the population-weighted mean exposure: 23–39% NO_2, 7–12% $PM_{2.5}$). Results by subregion (Fig. 2.7, Table 3.12) were generally consistent with Fig. 2.6, with a small number of exceptions ($PM_{2.5}$ in central and northern China).

Disparities were larger for community-level than for individual-level data (Tables 3.9 and 3.10). For example, average concentrations for the highest and lowest 20th percentage of SES differed by 41% (NO_2) and 12% ($PM_{2.5}$) for individual-level CHARLS data, compared with 89% (NO_2) and 25% ($PM_{2.5}$) for community-aggregated CHARLS data. This finding reflects in part that within- and between-community variabilities of SES were similar in magnitude, whereas variability in modeled concentrations was less within-community than between-community (Table 3.13). Area-level data reflect the average of the community, and averaging reduces the overall variability more for SES than for ambient concentrations. Air pollution concentrations and SES were higher in urban than rural areas. The positive concentration–SES relationships ($p< 2 \times 10^{-16}$) held even when using different definitions of urban/rural land (Table 3.11 and Fig. 2.8).

3.4.3 Sensitivity analysis

My core result—the positive relationships between ambient concentrations and SES—persisted even for dramatically different spatial units of analysis. Specifically, when averaging GDP per capita and air pollution concentrations across grid sizes from 1 to 100km (i.e., from 1 to 10,000

km^2), slopes of best-fit lines (Figs. 2.9 and 2.10) remained positive ($p< 1 \times 10^{-10}$). In general, disparities decreased with coarser resolution. This result is likely because the coarser resolution smooths out the spatial clustering of SES (Fig. 2.11); similar findings were noted in the United States for modeled $PM_{2.5}$ concentrations [16] and, generally, for empirical models of NO_2 and $PM_{2.5}$ [17].

My results were robust to the use of monitored rather than modeled pollution concentrations (Tables 3.14 and 3.15). Specifically, the core result (positive relationship between SES and concentrations) generally held when using monitored or modeled concentrations at all monitored locations and at rural monitor locations; the results were not significant at urban monitored locations (after controlling for population density, the results for NO_2 became significant at urban monitored locations). For both pollutants, quantifications of that relationship (the relative concentration differences between highest and lowest 20% SES) differed <3% between monitored vs. measured data (Tables 3.14 and 3.15). That agreement in part reflects the strength of the models employed [R2: 0.78 (NO_2), 0.89 ($PM_{2.5}$); root-mean-square-error: 5.9 $\mu g/m^3$ (NO_2), 6.3 $\mu g/m^3$ ($PM_{2.5}$)].

3.5 Discussions

Across several analyses, I found a positive relationship between SES and ambient air pollution in China: On average, NO_2 and $PM_{2.5}$ concentrations are higher than average for higher-SES populations and lower than average for lower-SES, rural-to-urban migrant, and ethnic minority populations. These findings are remarkably robust, holding for urban and rural locations, across nearly all geographic subregions within China, for three measures of SES (one individual measure, two areal measures), for modeled vs. measured pollution concentrations, for multiple definitions of urban vs. rural, and across a 100-fold range (in terms of area: 10,000-fold range) of spatial resolutions. Findings here are consistent with the hypothesis described in the "Introduction" section, that ambient air pollution exposures are higher for higher-SES than for low-SES individuals and areas.

Prior findings on air pollution and demographics in China (e.g., [84–86, 192]) have generally been inconclusive; in some cases, they propose an EKC relationship (e.g., for urban areas; [193]). (In general, this prior work employed census demographic data.) In contrast, findings here are consistent and monotonic.

Disparities are larger for NO_2 than for $PM_{2.5}$ (Tables 3.10 to 3.13), likely reflecting the greater

spatial variability for NO_2 than $PM_{2.5}$. NO_2 is a traffic-related pollutant of predominantly urban origin and is influenced by local emission sources; $PM_{2.5}$ has long-range and secondary components and thus varies at a regional level [153, 194]. The larger disparities for NO_2 than for $PM_{2.5}$ are consistent with results from other countries [7, 8, 10, 195], likely reflecting spatial aspects of the two pollutants [i.e., greater spatial variability for NO_2 than for $PM_{2.5}$, owing to differences in the emission sources and atmospheric chemistry (e.g., the strong regional secondary component for $PM_{2.5}$)], which held in other countries too [93, 194].

I mainly explored rural-to-urban migration status and three SES measures in my study; in addition, I also considered the effects of ethnicity and of family living expenditures. Other unmeasured factors (e.g., industrial structure, housing price, and hukou policies in different cities and regions) may have also influenced my results. For example, on average, cities that rely more on secondary industry (i.e., manufacturing) may tend to have more industrial pollution ($PM_{2.5}$) than cities that mainly rely on tertiary industry (i.e., services), so even at the same income/GDP per capita level, the ambient pollution for their residents are different. Future work could explore those factors that are unmeasured here.

The positive pollution–SES relationship in China is generally the reverse of patterns typically reported in the literature, which in part reflects that the existing literature focuses mostly on the United States and, to a lesser extent, other high-income countries. This difference in results likely reflects the different cultures, economic and political systems, history, demographics, level of industrialization, and urban structure in China vs. the United States. Several examples of such are described in the "Introduction" section. For example, differences in urban characteristics tend to increase the proportion of a) high-SES populations that live close to the urban center and b) low-SES populations that live in the suburbs and exurbs [13, 170, 171].

My findings likely have implications for other locations besides China. Specifically, for countries that are (like China) relatively homogenous racially and with economic development that is relatively recent (past decades), results here suggest that exposures may follow a different pattern than has been observed for the United States. Future work can usefully explore patterns in other contexts to develop the underlying framework and look at changes over time. For example, without further evidence, I would be hesitant to ascribe my hypothesis to other points in time. If future economic development and urbanization is concurrent with improvements in air pollution—especially in high- and upper-middle-income counties in China [196]—then current patterns of environmental

inequality may change over time. In recent years, with economic reform increasing public pressures for clean production, increased public awareness of air pollution, and more complete and rigid environmental protection laws, both air pollution level and GDP growth rate are decreasing in China. Industries are moving from urban areas to rural/suburbs and from major cities to surrounding cities [180]; these patterns may, in the future, weaken or shift the relationship between SES and air pollution that is reported here. In addition, with continued rural-to-urban migration, lower-SES populations may increasingly move to the heavily polluted megacities (e.g., Beijing, Shanghai, Guangzhou) for work, which could also weaken the positive pollution–SES relationship. Environmental justice theory in the United States [158, 197] and elsewhere [198, 199] highlights the concept of sacrifice zones—locations that lack political power and receive disproportionately high environmental risks. This is a critical aspect of how relationships here might shift over time that will depend in part on inequities in political power and whether sacrifice zones become more prevalent.

Limitations of my methods include the following. First, I investigated ambient concentrations at home locations rather than personal exposures. This approach, which is common in the literature, omits factors such as mobility (e.g., travel for work and recreation), occupational exposure, and near-source exposures (e.g., environmental tobacco smoke; time spent on-roadways; indoor use of solid fuels for cooking, heating, or lighting) [200–212]. Future work could shed important additional insight by considering those factors. [Some of the factors are already well studied, e.g., indoor use of solid fuels for cooking generally happens in lower-income, not in higher-income, households [213, 214].] Second, two of the three SES metrics are derived from CHARLS, which is a cohort of individuals $\geq$45 years of age that is representative of that age group but is imperfectly representative of the overall (i.e., all ages) Chinese population (Table 3.16). (The third SES metric, derived from gridded GDP data, does not face this limitation.) Third, my methods do not reveal which emission sources drive the concentrations and inequalities; future work could usefully investigate this question [56, 193, 215, 216]. Fourth, my methods do not employ a spatial regression approach; future work could shed additional insight by accounting for the spatial clustering of air pollution [217]. Fifth, I studied conditions in China and not in other countries. Future studies could usefully address this point by, for example, using a unified definition of SES and exposures to compare across multiple countries.

This chapter describes a framework of air pollution inequality in China and uses several approaches to testify my hypothesis. Strengths of my approach include use of multiple data sets

and lines of evidence, including multiple approaches to estimating SES; investigation nationally, comprehensively, and for several subpopulations (e.g., subnational region; migrant status) using consistent methods; comparing across two pollutants, with good spatial precision; use of modeled as well as measured concentrations; use of global population density data and cohort data on demographics and home-location, which improves the spatial precision of the demographic information and decreases the spatial aggregation problem [16–19]; and sensitivity analyses that consider multiple perturbations to the core analyses [e.g., a 100-fold (by area, a 10,000-fold) change in spatial resolution; use of different definitions of urban vs. rural]. The multiple approaches and sensitivity analyses reveal similar core findings, which adds confidence regarding the robustness of findings.

3.6 Conclusions

This study finds that for two important air pollutants (NO_2 and $PM_{2.5}$), average ambient exposures in China are higher for higher-SES than for lower-SES populations and higher for long-standing urban residents than for rural-to-urban migrant populations. The results are robust to using multiple individual- and areal-level SES approaches, in both rural and urban locations, across geographies, across a 100-fold range of spatial resolution, and for monitored vs. modeled ambient concentrations. The disparity for NO_2 is larger than for $PM_{2.5}$.

My results are consistent with my hypothesis, which is the opposite of most of the existing environmental justice literature; findings here likely reflect correlations among economic development, SES, and pollution (i.e., economic development increases SES and pollution levels). The findings may have implications for locations outside of China, especially for other low- and middle-income countries.

Table 3.1: Coordinates and contributions of SES variables on SES score (first dimension of factor analysis of mixed data [FAMD]).

SES variable	Category	Coordinate on dimension 1	Contribution on dimension 1 (%)
Education	Illiterate	-0.45	8.3
	Did not finish elementary school but can read	-0.13	0.7
	Elementary school	-0.03	0.0
	Middle school	0.30	3.7
	High school and above	0.42	7.3
Occupation	Agricultural work	-0.92	34.7
	Nonagricultural work	0.92	34.7
Income	Log of household per capita income	0.51	10.7

Table 3.2: Spatial resolutions of predictors in the land-use regression models of NO_2 and $PM_{2.5}$ concentrations.

Predictor		Spatial resolution
Monitoring data		–
Meteorological data		0.25° (around 25 km)
Satellite data		0.1° (around 10 km)
Geographic data	Road network	Collected using 100 m to 10 km buffers
	Land cover	30 m resolution; collected using 300m to 30km buffers
	Point of interest	Collected using 100 m to 50 km buffers
	Fire count	Collected using 5 km to 100 km buffers
	Elevation	1 km
	Population density	30 second (around 1 km)
	Vegetation	0.05° (around 500 m)

Table 3.3: Mean (with interquartile range [IQR]) NO_2 and $PM_{2.5}$ concentrations ($\mu g/m^3$) in each SES group and maximum difference (absolute [$\mu g/m^3$] and relative [%], with absolute value) from population averages.

		NO_2		$PM_{2.5}$	
	Number	Group mean with IQR	Maximum difference from population average	Group mean with IQR	Maximum difference from population average
Migration					
Rural resident	12,533	21.8 (16.3-27.8)		49.7 (35.5-59.6)	
Rural-to-urban migrant	3,860	25.6 (18.2-31.8)	6.8 (28%)	50.3 (36.7-58.8)	4.2 (8%)
Urban resident	3,519	31.2 (21.9-39.6)		54.9 (42.8-62.1)	
Missing	1,183	28.7 (20.4-35.1)		50.2 (39.2-57.3)	
Education					
Illiterate	5,044	22 (16.0-28.1)		48.9 (34.5-59.2)	
Did not finish elementary school but can read	3,374	22.7 (16.8-28.3)		48.2 (36.0-56.9)	
Elementary school	4,016	23.9 (17.9-30.0)	4.5 (18%)	50.6 (37.8-58.8)	3.4 (7%)
Middle school	3,988	26.6 (19.0-32.5)		53.5 (39.8-62.4)	
High school and above	2,367	28.9 (20.6-34.8)		54.2 (41.1-62.4)	
Missing	2,306	24.8 (18.2-30.9)		50.2 (37.8-58.8)	
Occupation					
Agricultural work	9,197	21.6 (16.3-27.4)	2.8 (12%)	49.3 (35.5-58.9)	1.4 (3%)
Nonagricultural work	11,454	26.7 (19.0-32.6)		51.9 (38.8-59.9)	
Missing	444	25.2 (18.2-31.9)		50.2 (35.4-60.5)	
Household per capita income (Yuan)					
<2,000	3,663	22.1 (16.2-28.1)		49.1 (34.8-59.3)	
2,000-5,500	3,441	23.2 (17.4-29.3)		50.6 (36.8-60.0)	
5,500-12,000	3,669	24.2 (18.0-30.5)	3.5 (14%)	51.1 (38.1-59.7)	2.1 (4%)
12,000-24,000	3,529	26.1 (18.8-32.2)		51.5 (39.1-59.3)	
<24,000	3,294	27.9 (19.8-34.1)		52.8 (41.3-60.0)	
Missing	3,499	23.4 (17.5-28.9)		49.4 (36.0-58.4)	
Population density					
Low population density	6,600	20.2 (14.9-23.1)		46.7 (34.2-57.1)	
Middle population density	6,639	24.5 (19.2-30.9)	4.8 (19%)	53.3 (40.5-63.3)	4.0 (8%)
High population density	6,736	29.2 (20.6-36.7)		52.6 (38.8-59.9)	
Missing	1,120	20.6 (13.4-27.9)		47.8 (29.9-56.4)	

Table 3.4: Increase in mean NO_2 ($\mu g/m^3$) and $PM_{2.5}$ ($\mu g/m^3$) concentrations in relation to migration condition and SES factors.

	NO_2			$PM_{2.5}$		
	Individual-level-factor only model [a]	Random intercepts model [b]	Random intercepts model [c]	Individual-level-factor only model [a]	Random intercepts model [b]	Random intercepts model [c]
Migration – Rural resident						
Rural-to-urban migrant	3.3 (3.0, 3.7)	2.6 (2.4, 2.8)	1.4 (1.1, 1.6)	0.1 (-0.6, 0.8)	2.9 (2.6, 3.2)	2 (1.7, 2.3)
Urban resident	7.3 (6.8, 7.7)	5.4 (5.2, 5.7)	3.1 (2.8, 3.3)	3.1 (2.3, 4.0)	3.9 (3.5, 4.2)	2.3 (2.0, 2.7)
Education – Illiterate						
Did not finish elementary school but can read	-0.1 (-0.5, 0.3)	0 (-0.2, 0.2)	0 (-0.1. 0.3)	-1 (-1.8, -0.2)	-0.2 (-0.5, 0.0)	0 (-0.2, 0.2)
Elementary school	0.8 (0.4, 1.2)	-0.1 (-0.3, 0.1)	-0.1 (-0.3, 0.1)	1.3 (0.5, 2.1)	-0.3 (-0.6, 0.0)	-0.1 (-0.4, 0.1)
Middle school	2.4 (2.0, 2.9)	0.1 (-0.2, 0.3)	0.1 (-0.1,0.4)	4.1 (3.3, 5.0)	-0.4 (-0.7, -0.1)	-0.1 (-0.3, 0.2)
High school and above	2.1 (1.5, 2.6)	0.1 (-0.2, 0.3)	0.1 (-0.2,0.4)	3.4 (2.6, 4.4)	-0.7 (-1.1, -0.4)	-0.2 (-0.5, 0.1)
Occupation – Agricultural work						
Nonagricultural work	2.1 (1.8, 2.4)	1.1 (1.0, 1.3)	1.1 (0.9, 1.2)	0.4 (-0.2, 1.0)	0.1 (-0.1, 0.3)	0.2 (0.0, 0.4)
Household per capita income [d]	0.7 (0.5, 0.8)	0.2 (0.1, 0.3)	0.2 (0.1, 0.3)	0.6 (0.3, 0.9)	0.1 (-0.0, 0.2)	0.2 (0.1, 0.3)
Log of population density			0.7 (0.7, 0.8)			0.3 (0.2, 0.3)

[a] The model only considers individual's age, migration condition, and socioeconomic characters.

[b] The model controls the individual's characters, and adds random intercepts for each prefecture-level city.

[c] The model controls the individual's characters, population density, and adds random intercepts for each prefecture-level city.

[d] Household per capita income is first log-transformed and then transformed into z-score. Parameter estimate refers to a 1-unit increase in the z-score for log value of household per capita income.

Table 3.5: Multiple imputation regressions (pooled results) for mean NO_2 ($\mu g/m^3$) and $PM_{2.5}$ ($\mu g/m^3$) concentrations in relation to migration condition and SES factors.

	NO_2		$PM_{2.5}$	
	Individual-level-factor only model [a]	Random intercepts model [b]	Individual-level-factor only model [a]	Random intercepts model [b]
Migration – Rural resident				
Rural-to-urban migrant	3.1 (2.8, 3.5)	1.4 (1.2, 1.6)	0 (-0.6, 0.6)	2.1 (1.9, 2.4)
Urban resident	7.2 (6.8, 7.6)	3 (2.7, 3.3)	2.6 (1.9, 3.4)	2.7 (2.4, 3.0)
Education – Illiterate				
Did not finish elementary school but can read	0.1 (-0.3, 0.5)	0.1 (-0.1, 0.3)	-0.4 (1.1, 0.4)	-0.2 (-0.4, 0.2)
Elementary school	0.9 (0.5, 1.3)	0 (-0.3, 0.2)	1.7 (1.0, 2.4)	-0.3 (-0.5, -0.1)
Middle school	2.3 (1.9, 2.7)	0.2 (-0.1, 0.4)	4.1 (3.4, 4.8)	-0.4 (-0.6, -0.1)
High school and above	2.3 (1.8, 2.8)	0.1 (-0.1, 0.4)	3.7 (2.8, 4.6)	-0.6 (-0.9, -0.3)
Occupation – Agricultural work				
Nonagricultural work	2 (1.8, 2.3)	0.9 (0.8, 1.1)	0.6 (0.1, 1.1)	0 (-0.2, 0.2)
Household per capita income [c]	0.7 (0.5, 0.9)	0.2 (0.1, 0.3)	0.6 (0.3, 0.9)	0.1 (0.0, 0.2)
Log of population density		0.8 (0.7, 0.9)		0.3 (0.3, 0.4)

[a] The model only considers individual's age, migration condition, and socioeconomic characters.

[b] The model controls the individual's characters, population density, and adds random intercepts for each prefecture-level city.

[c] Household per capita income is first log-transformed and then transformed into z-score. Parameter estimate refers to a 1-unit increase in the z-score for log value of household per capita income.

Table 3.6: Mean (with IQR) NO_2 and $PM_{2.5}$ concentrations ($\mu g/m^3$), and mean SES scores in each ethnicity group.

Ethnicity group	Number	Group mean with IQR		Mean standardize SES score
		NO_2 concentration ($\mu g/m^3$)	$PM_{2.5}$ concentration ($\mu g/m^3$)	
Han	14,936	24.7 (18.2-30.8)	52.0 (39.1-60.3)	-0.1
Hui	80	39.1 (32.2-44.3)	67.4 (55.9-72.5)	0.7
Zhuang	155	12.3 (7.6-11.7)	36.0 (34.1-35.2)	-0.5
Uyghur	79	21.1 (20.5-21.2)	57.6 (57.4-57.9)	0.3
Yi	77	14.8 (12.9-16.6)	27.1 (23.2-27.8)	-0.8
Tibet	146	14.8 (12.1-16.1)	28.1 (24.2-24.5)	-0.7
Miao	94	14.3 (11.9-15.3)	30.9 (31.1-31.7)	-0.8
Mongol	173	17.3 (15.1-18.1)	34.9 (32.2-37.3)	-0.2
Dai	34	15.3 (15.1-15.7)	29.9 (29.8-29.9)	-0.7
Other	413	20.0 (15.2-21.7)	42.6 (32.3-53.1)	-0.2
Ethnicity Minorities	1,251	18.4 (12.7-21.0)	39.3 (29.9-46.5)	-0.3
Missing	4,908	25.3 (18.2-31.9)	49.7 (36.8-58.3)	0.3

Table 3.7: Increase in mean NO_2 ($\mu g/m^3$) and $PM_{2.5}$ ($\mu g/m^3$) concentrations in relation to migration condition, SES factors, ethnicity, and living expenditure.

	NO_2		$PM_{2.5}$	
	Individual-level-factor only model [a]	Random intercepts model [b]	Individual-level-factor only model [a]	Random intercepts model [b]
Migration – Rural resident				
Rural-to-urban migrant	3.2 (2.9, 3.6)	1.4 (1.2, 1.6)	0.1 (-0.6, 0.8)	2.1 (1.8, 2.4)
Urban resident	7.4 (6.9, 7.8)	3 (2.7, 3.3)	3.7 (2.8, 4.6)	2.3 (2.0, 2.7)
Education – Illiterate				
Did not finish elementary school but can read	-0.2 (-0.6, 0.3)	0.1 (-0.1, 0.3)	-1.2 (-2.1, -0.4)	0 (-0.3, 0.2)
Elementary school	0.6 (0.1, 1.0)	0 (-0.3, 0.2)	0.9 (0.1, 1.7)	-0.2 (-0.5, 0.0)
Middle school	2.2 (1.7, 2.7)	0.2 (-0.1, 0.4)	3.9 (3.0, 4.8)	-0.1 (-0.4, 0.2)
High school and above	1.9 (1.3, 2.5)	0.1 (-0.1, 0.4)	3.4 (2.3, 4.5)	-0.1 (-0.5, 0.2)
Occupation – Agricultural work				
Nonagricultural work	1.9 (1.6, 2.3)	0.9 (0.8, 1.1)	0.4 (-0.2, 1.0)	0.2 (0.1, 0.4)
Household per capita income [c]	0.5 (0.4, 0.7)	0.2 (0.1, 0.3)	0.6 (0.3, 0.9)	0.2 (0.1, 0.3)
Ethnicity – Han ethnicity				
Ethnicity minorities	-5.4 (-6.0,-4.8)	0.1 (-0.3, 0.4)	-11.5 (-12.7, -10.4)	0.1 (-0.3, 0.6)
Household per capita living expenditure [c]	0 (-0.2, 0.2)	0.1 (0.1, 0.2)	-1.1 (-1.4, -0.8)	0 (-0.1, 0.1)
Log of population density		0.8 (0.7, 0.9)		0.3 (0.2, 0.4)

[a] The model only considers individual's age, migration condition, and socioeconomic characters.

[b] The model controls the individual's characters, population density, and adds random intercepts for each prefecture-level city.

[c] Household per capita income and household per capita living expenditure are first log-transformed and then transformed into z-score. Parameter estimates refer to a 1-unit increase in the z-score for log value of household per capita income/living expenditure.

Table 3.8: Mean (with IQR) NO_2 and $PM_{2.5}$ concentrations ($\mu g/m^3$), and mean SES scores in each household per capita living expenditure group (by quintile).

Household per capita living expenditure quintile (Yuan)	Number	Group mean with IQR		Mean standardize SES score
		NO_2 concentration ($\mu g/m^3$)	$PM_{2.5}$ concentration ($\mu g/m^3$)	
<4,500	4,021	23.3 (17.4-29.9)	51.8 (37.3-62.2)	-0.4
4,500-8,000	3,917	23.5 (17.6-29.5)	50.8 (37.8-59.7)	-0.2
8,000-12,500	3,939	24.2 (18.0-30.5)	50.6 (37.7-59.3)	0
12,500-22,500	4,286	25.0 (18.2-30.9)	50.3 (38.1-58.3)	0.2
>22,500	3,969	26.3 (18.4-32.3)	50.5 (38.0-58.8)	0.3
Missing	4,908	24.0 (16.7-30.5)	49.2 (34.3-59.2)	-0.1

Table 3.9: Individual level relationships between pollution concentration and combined SES score.

		Group mean concentration ($\mu g/m^3$)		Relative concentration difference between highest 20% and lowest 20% SES (%)	Linear regression line slope with 95% CI [a]	Regression slope × IQR [b]
		Lowest 20% SES	Highest 20% SES			
NO_2	Overall	21	30	41	3.0 (2.9, 3.2)	5.6 (5.4, 5.9)
	Rural resident	20	24	17	1.3 (1.1, 1.4)	2.2 (1.9, 2.4)
	Rural-to-urban migrant	23	28	25	2.4 (2.1, 2.7)	4.1 (3.6, 4.6)
	Urban resident	28	34	23	3.4 (2.8, 3.9)	2.3 (1.9, 2.6)
$PM_{2.5}$	Overall	48	54	12	1.9 (1.6, 2.1)	3.5 (3.0, 3.9)
	Rural resident	47	52	10	1.4 (1.1, 1.8)	2.4 (1.9, 3.0)
	Rural-to-urban migrant	49	53	9	1.5 (1.0, 2.1)	2.6 (1.7, 3.6)
	Urban resident	54	58	3	0.3 (-0.6, 1.3)	0.2 (-0.4, 0.9)

[a] The unit of regression slope is $\mu g/m^3/\log(\text{yuan})$

[b] Calculated by linear regression slope times interquartile range of individual SES score. The unit is $\mu g/m^3$.

Table 3.10: Community level relationships between pollution concentration and combined SES score.

		Overall	Rural	Urban
Community area	Median (km^2)	6.4	13	2
	IQR (km^2)	1.7 - 20	5.6 - 30	0.8 - 6.0
Mean NO_2 concentration	Lowest 20% SES ($\mu g/m^3$)	19	18	22
	Highest 20% SES ($\mu g/m^3$)	36	24	41
	Relative difference (%)	89	38	91
	Linear regression line slope with 95% CI [a]	8.6 (7.2, 9.9)	7.0 (4.2, 9.7)	10.5 (7.9, 13.1)
	Regression slope × IQR [b]	9.4 (7.9, 10.8)	3.2 (1.9, 4.4)	7.6 (5.7, 9.5)
Mean $PM_{2.5}$ concentration	Lowest 20% SES ($\mu g/m^3$)	44	42	48
	Highest 20% SES ($\mu g/m^3$)	55	52	60
	Relative difference (%)	25	24	25
	Linear regression line slope with 95% CI [a]	5.5 (3.0, 8.1)	11.2 (4.3, 18.1)	6.1 (1.1, 11.1)
	Regression slope × IQR [b]	6.0 (3.3, 8.8)	5.1 (2.0, 8.2)	4.4 (0.8, 8.0)

[a] The unit of regression slope is $\mu g/m^3/\log(\text{yuan})$

[b] Calculated by linear regression slope times interquartile range of individual SES score. The unit is $\mu g/m^3$.

Table 3.11: Pollution concentration versus gridded GDP per capita in 1 km grid cells (population weighted).

		Group mean concentration ($\mu g/m^3$)		Relative concentration difference between highest 20% and lowest 20% GDP per capita (%)	Linear regression line slope with 95% CI [a]	Regression slope * IQR [b]
		Lowest 20% GDP per capita	Highest 20% GDP per capita			
NO_2	Overall	17	26	48	3.0 (2.9, 3.0)	7.3 (7.1, 7.4)
	Rural land	17	24	44	1.8 (1.8, 1.9)	4.4 (4.3, 4.5)
	Urban land	32	34	8	1.3 (1.1, 1.4)	2.0 (1.7, 2.2)
$PM_{2.5}$	Overall	43	51	19	1.7 (1.7, 1.7)	4.1 (4.1, 4.2)
	Rural land	42	51	21	1.5 (1.5, 1.5)	3.6 (3.6, 3.6)
	Urban land	56	57	3	0.9 (0.7, 1.1)	1.4 (1.1, 1.7)

[a] The unit of regression slope is $\mu g/m^3/\log(\text{yuan})$

[b] Calculated by linear regression slope times interquartile range of individual SES score. The unit is $\mu g/m^3$.

Table 3.12: Summary of socioeconomic characteristics, pollution concentration and disparities by region.

Region			West	Central	Southeast	North	Northeast
	Mean income (Yuan)		13,466	14,273	20,388	16,676	17,036
	Mean SES score		-0.21	-0.01	0.19	0.14	0.24
	Migration percentage (%)		17	17	30	15	8
$PM_{2.5}$	Mean concentration ($\mu g/m^3$)		40	58	43	72	52
	Disparities between highest 20% versus lowest 20% SES	Absolute ($\mu g/m^3$)	5 (38, 43)	-0.2 (57.5, 57.3)	8 (39, 47)	3 (71, 74)	0.6 (50.9, 51.5)
		Relative (%)	13	0	20	4	1
NO_2	Mean concentration ($\mu g/m^3$)		20	25	24	34	26
	Disparities between highest 20% versus lowest 20% SES	Absolute ($\mu g/m^3$)	7 (17, 24)	4 (24, 28)	11 (19, 30)	9 (30, 39)	11 (20, 31)
		Relative (%)	40	16	55	31	52

Table 3.13: Within- and between-community variabilities of NO_2 concentration, $PM_{2.5}$ concentration and SES score.

	Within-community IQR [a]	Between-community IQR [b]	Ratio of within- and between-community IQR
NO_2 ($\mu g/m^3$)	0.08	14	0.006
$PM_{2.5}$ ($\mu g/m^3$)	0.03	22	0.001
SES score	1.3	1.5	0.9

[a] Calculated by the median of IQRs for each community.

[b] Calculated by the IQR of medians for each community.

Table 3.14: Disparities of concentration versus gridded GDP per capita in 1 km grid cells using monitored and predicted concentrations at monitoring site locations.

			Monitoring site locations	
			Monitored	Predicted
NO_2	Regression slope of pollution versus GDP per capita		1.1 (0.6, 1.5)	1.0 (0.6, 1.4)
	Mean concentration ($\mu g/m^3$)	Lowest 20% GDP per capita group	28	28
		Highest 20% GDP per capita group	32	32
		Relative difference	14%	13%
$PM_{2.5}$	Regression slope of pollution versus GDP per capita		0.8 (0.1, 1.4)	0.7 (0.1, 1.3)
	Mean concentration ($\mu g/m^3$)	Lowest 20% GDP per capita group	50	50
		Highest 20% GDP per capita group	53	52
		Relative difference	5%	4%

Table 3.15: Disparities of concentration versus gridded GDP per capita in 1 km grid cells using monitored and predicted concentrations at urban and rural monitoring site locations.

			Urban monitored locations		Rural monitored locations	
			Monitored	Predicted	Monitored	Predicted
NO_2	Regression slope of pollution versus GDP per capita		-0.1 (-0.7, 0.5)	-0.2 (-0.7, 0.4)	1.3 (0.7, 1.9)	1.2 (0.7, 1.7)
	Mean concentration ($\mu g/m^3$)	Lowest 20% GDP per capita group	33	33	23	23
		Highest 20% GDP per capita group	32	32	32	32
		Relative difference	-3%	-3%	43%	40%
$PM_{2.5}$	Regression slope of pollution versus GDP per capita		-0.6 (-1.5, 0.3)	-0.5 (-1.4, 0.4)	1.0 (0.1, 1.9)	0.8 (-0.1, 1.6)
	Mean concentration ($\mu g/m^3$)	Lowest 20% GDP per capita group	56	55	43	44
		Highest 20% GDP per capita group	53	52	53	52
		Relative difference	-6%	-5%	22%	19%

Table 3.16: Comparison of socioeconomic characters of CHARLS 2015 cohort sample and national census data [a].

	National data	National data (>45 year old)	CHARLS sample
Illiterate (%)	6	12	26
Elementary school (%)	26	36	39
Middle school (%)	38	35	21
High school and above (%)	30	17	14
Agricultural work (%)	37	54	45
Household per capita income (Yuan)	21,996	–	15,697

[a] National education and occupation data are from *1% Population Sample Survey in 2015*; national household per capita income is from *China Statistical Year book 2015*.

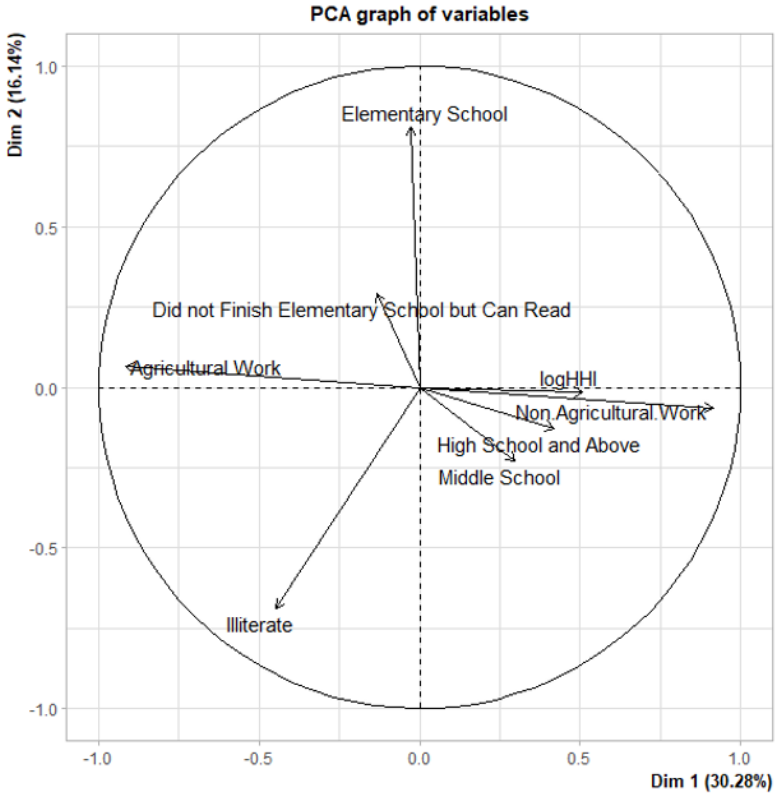

Figure 3.1: PCA analysis of socioeconomic factors.

Figure 3.2: Region classification in China.

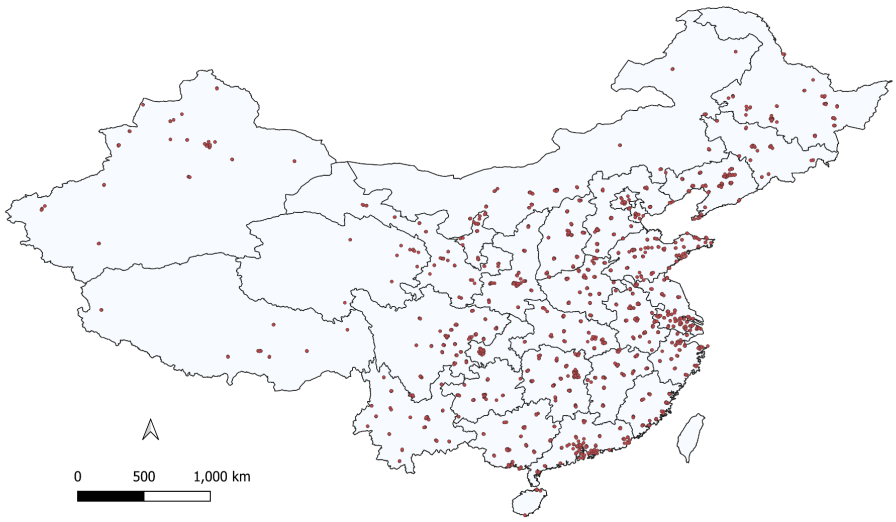

Figure 3.3: Spatial distribution of the monitor locations.

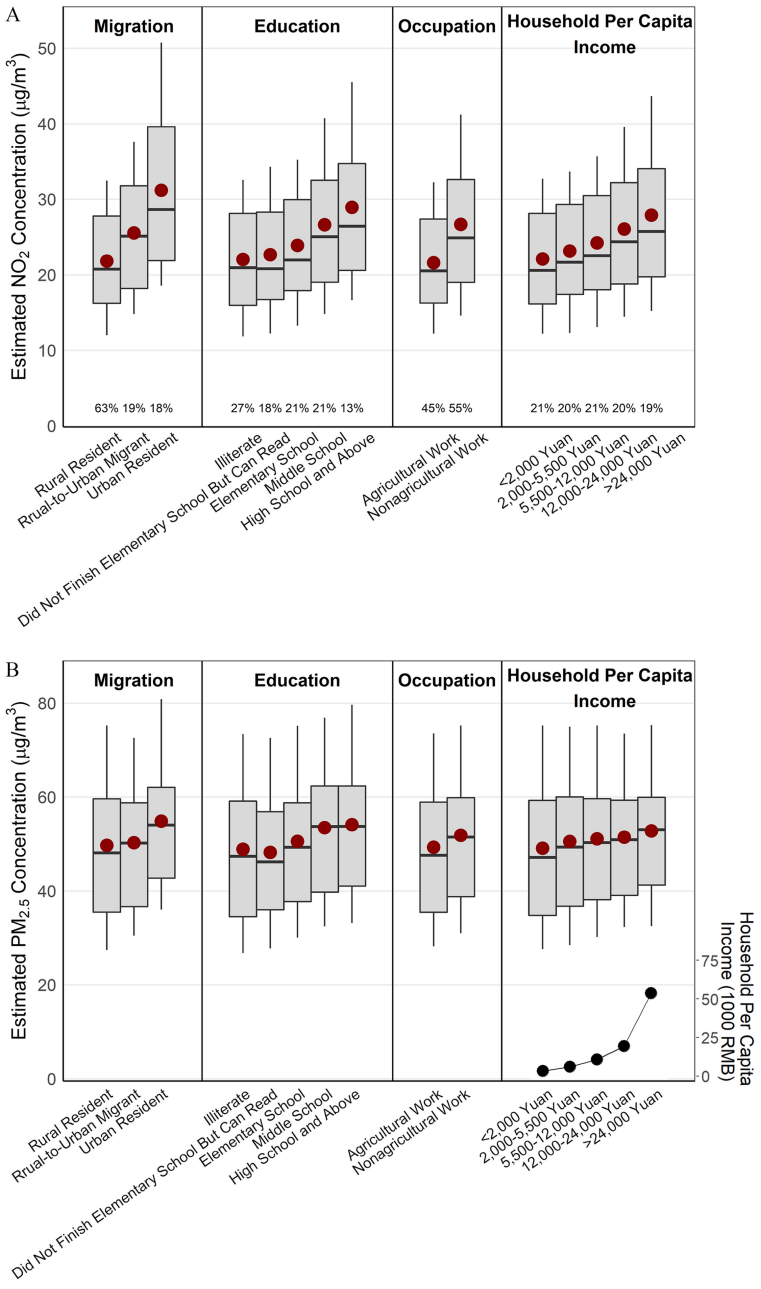

Figure 3.4: Estimated ambient (A) NO_2 and (B) $PM_{2.5}$ concentration by individual's rural-to-urban migration status, education, occupation, and income quintile. Box and whiskers indicate the 10th, 25th, 50th, 75th, and 90th percentiles and the mean (red circle). Income levels are displayed in the lower right of (B). The percentage numbers of individuals in each subgroup are annotated at the bottom of (A).

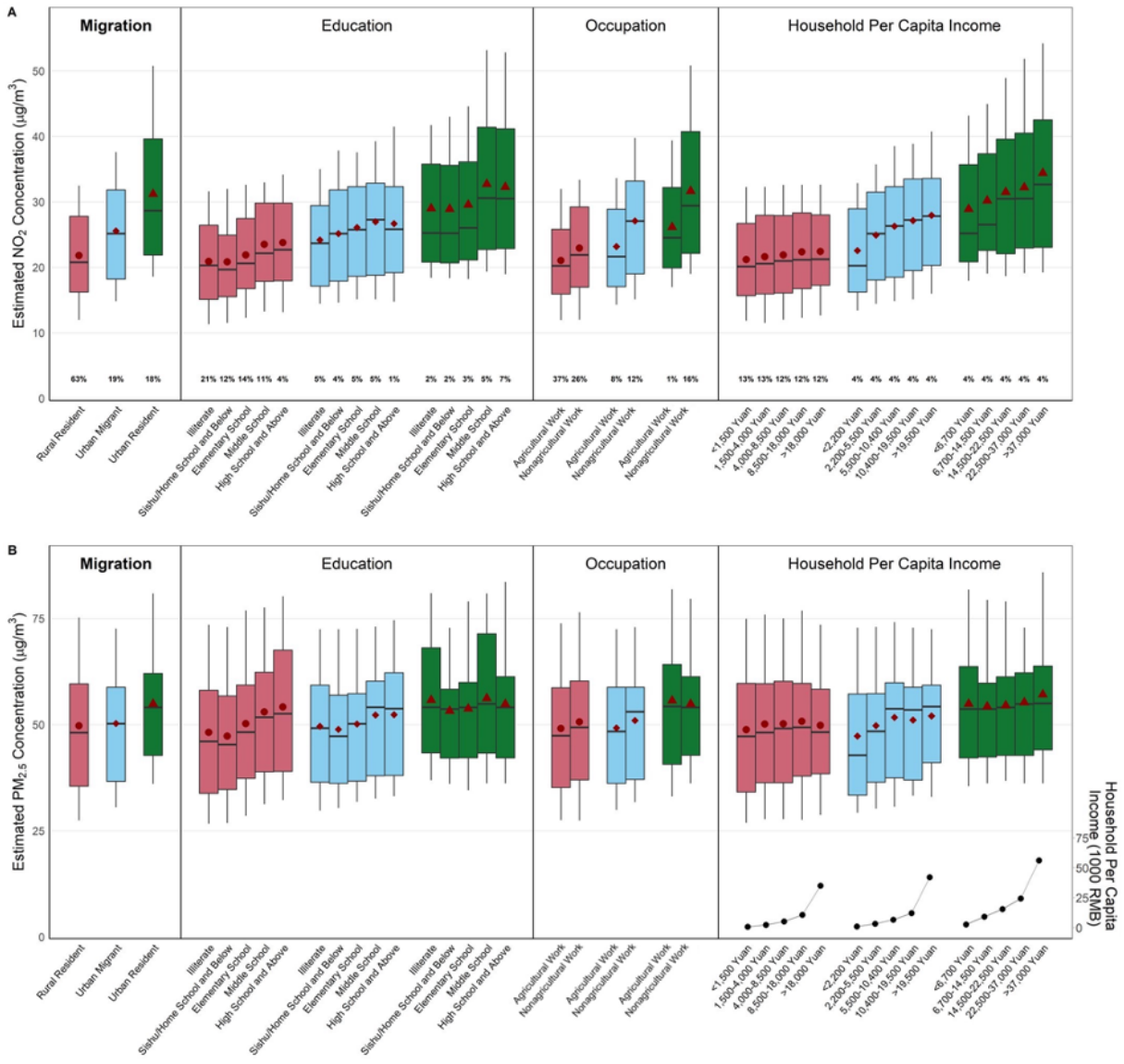

Figure 3.5: Estimated outdoor (A) NO_2 and (B) $PM_{2.5}$ concentration by the individual's highest educational attainment, occupation, and household per capita income quintile for each migration group. Box-and-whiskers indicate the 10th, 25th, 50th, 75th, and 90th percentile concentrations in each migration and SES group, the points indicate the mean concentration, and the percentage numbers in (A) indicate the population percentages of the samples in each group. The household per capita income quintiles are calculated based on the cohort samples, and the average household per capita income levels (1000 RMB) for each income group are shown in the lower right of (B).

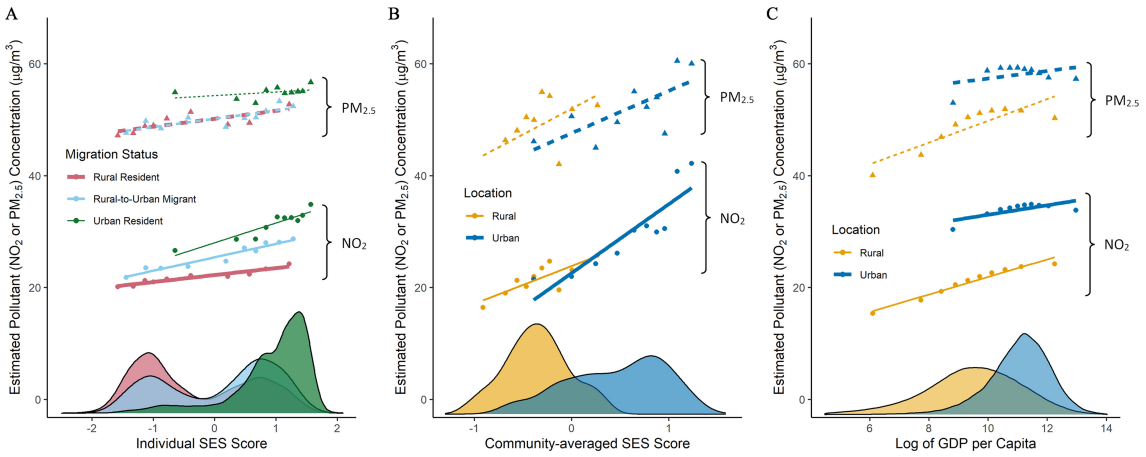

Figure 3.6: Relationship between SES and ambient NO_2 and $PM_{2.5}$ concentrations, based on (A) individual data, (B) areal data derived by aggregating the individual data to the community-level, and (C) areal data derived from national gridded GDP and world population density data sets. Data are plotted by urban–rural status, reflecting available data for individual data (A), three groups (rural resident, urban resident, and rural-to-urban migrant); for areal data (B,C), two groups (rural, urban). SES values reflect available data: (A) individual SES, (B) community-averaged SES, and (C) GDP per capita. Each plotted point represents the mean pollution concentration for 10% of the subsample. For example, in (A), the left-most red point represents the 10% of the rural residents with the lowest standardized SES score, and the right-most blue point represents the 10% of the rural-to-urban migrants with the highest standardized SES score. Plots also display best-fit lines and kernel densities. All of the best-fit lines have a positive slope, indicating that in all cases considered, higher SES is correlated with higher concentrations of ambient air pollution.

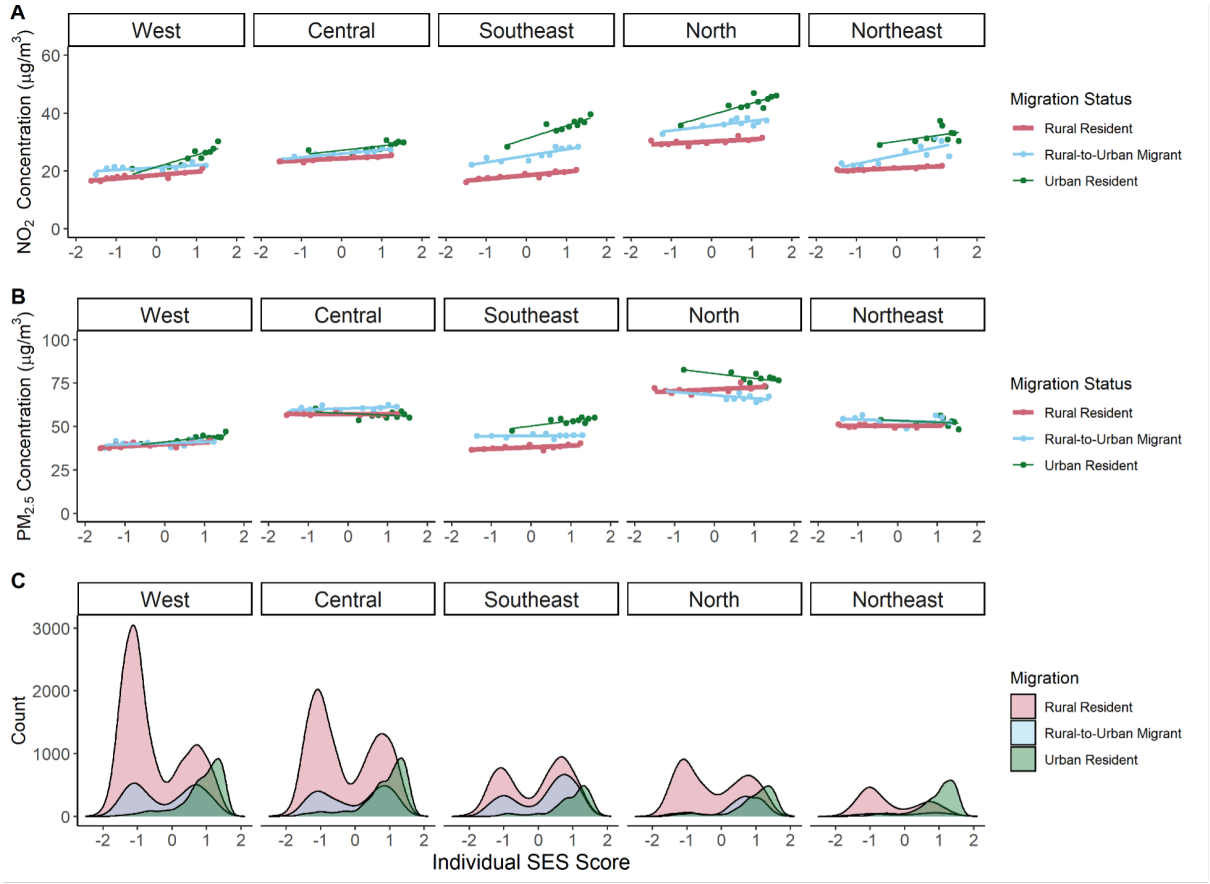

Figure 3.7: Relationship between (A) NO_2 and (B) $PM_{2.5}$ concentration versus combined SES score in each geographic region. Each point in the subplots represents the mean pollution concentration for 5% of the samples binned by standardized SES scores; each line represents the linear regression lines of concentration versus standardized SES score or gridded GDP for each group. (C) Density plots (counts) of standardized SES scores.

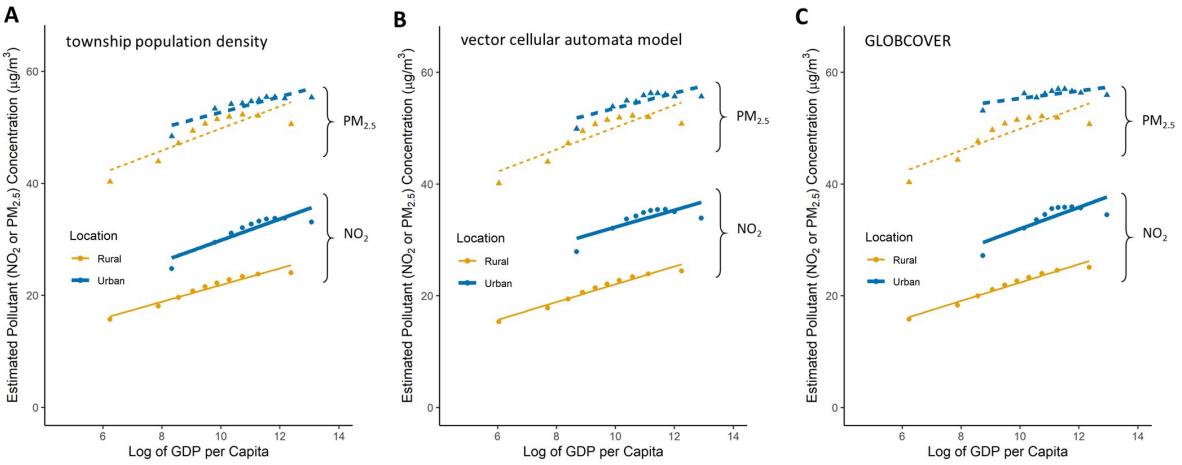

Figure 3.8: Relationship between average NO_2 and $PM_{2.5}$ exposures and SES, based on national gridded GDP and world population density datasets. Urban and Rural land are defined by (A) township-level population density in year-2010 [188], (B) a neighborhood-level vector cellular automata model based on density, neighborhood condition, and other spatial variables for year-2012 [189], (C) 300 m resolution GLOBCOVER dataset [190]. The shapefiles of all three classifications are available at Beijing City Lab (`www.beijingcitylab.com`). This figure is analogous to Fig. 3.6C in the main text, but with different urban-rural definitions. For all the best-fit lines, the slopes are positive ($p<2e\text{-}16$)

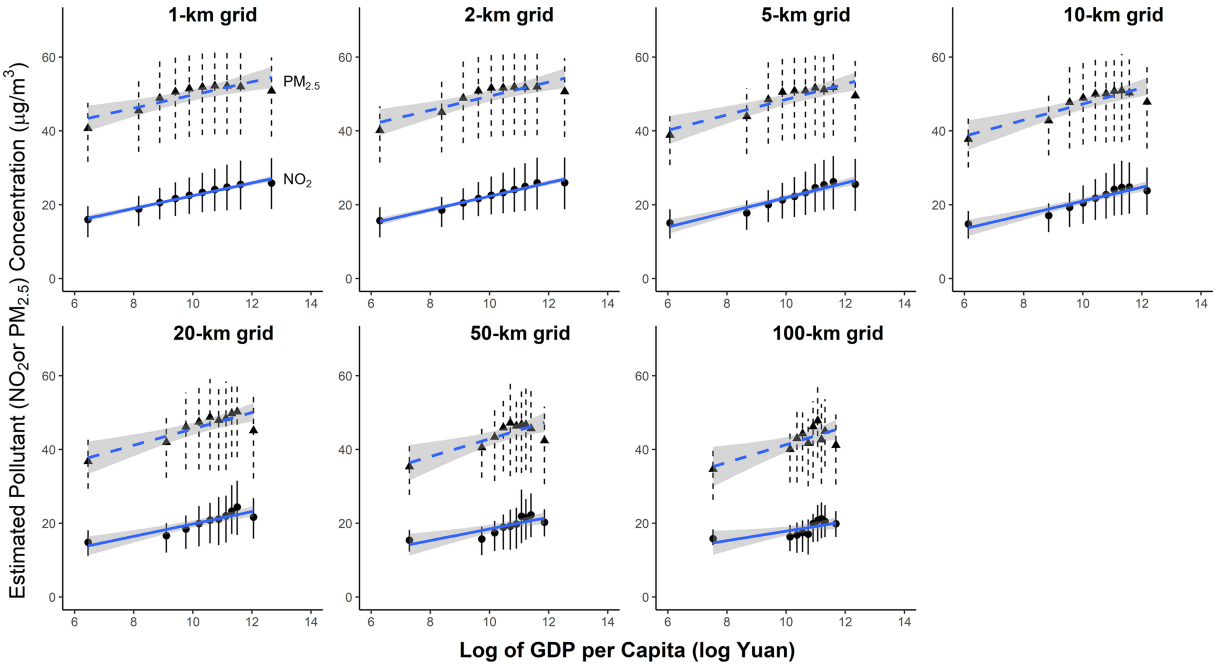

Figure 3.9: Relationship between pollution concentration and log GDP per capita, by grid cell size. Each point shows the mean log GDP per capita and mean pollution concentration of every 10% of population; each segment shows the IQR of pollution concentration. Best-fit lines by pollutant are shown in each plot.

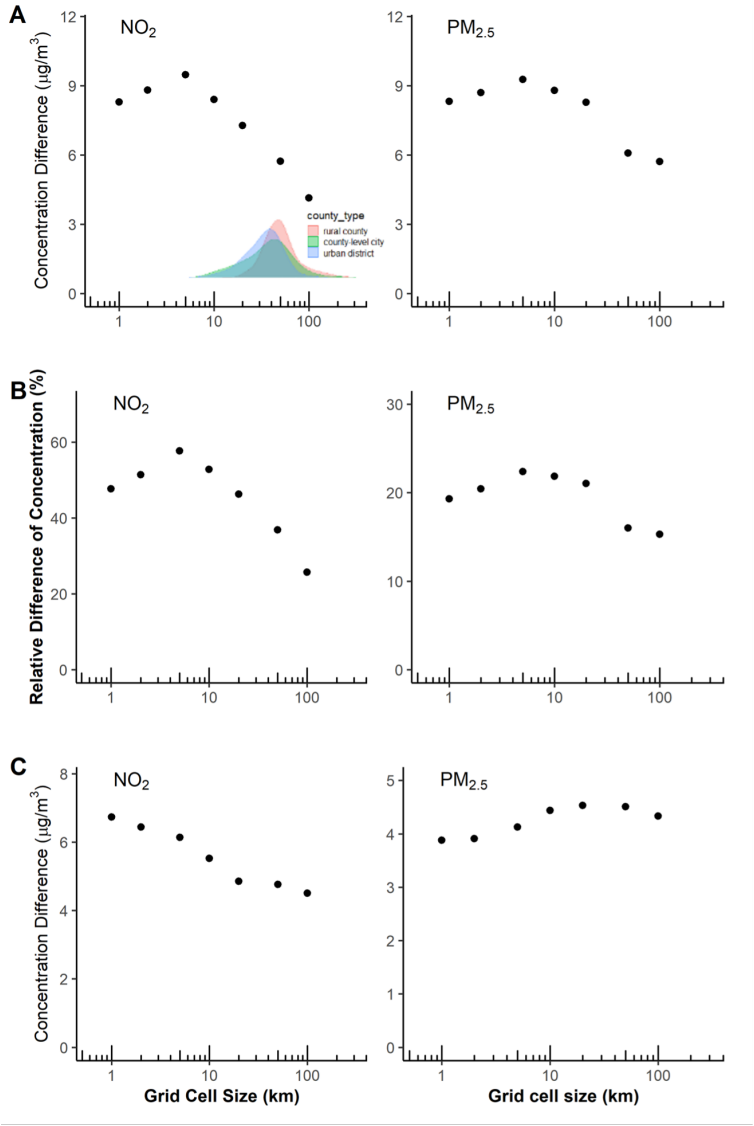

Figure 3.10: Effect of grid cell size on the concentration-GDP relationships: (A) Concentration difference between populations in highest 20% of GDP per capita locations and lowest 20% of GDP per capita locations; (B) Relative difference between mean pollution concentrations in highest 20% of GDP per capita and lowest 20% of GDP per capita; (C) Concentration change per 1-IQR increase in log of GDP per capita, calculated by regression slope of pollution concentration versus log of GDP per capita × IQR of log GDP per capita. All concentrations and quantiles are population-weighted. The distributions of county size for rural counties, county-level cities, and urban districts are shown in (A). For all regression slopes in (B), p<1e-10.

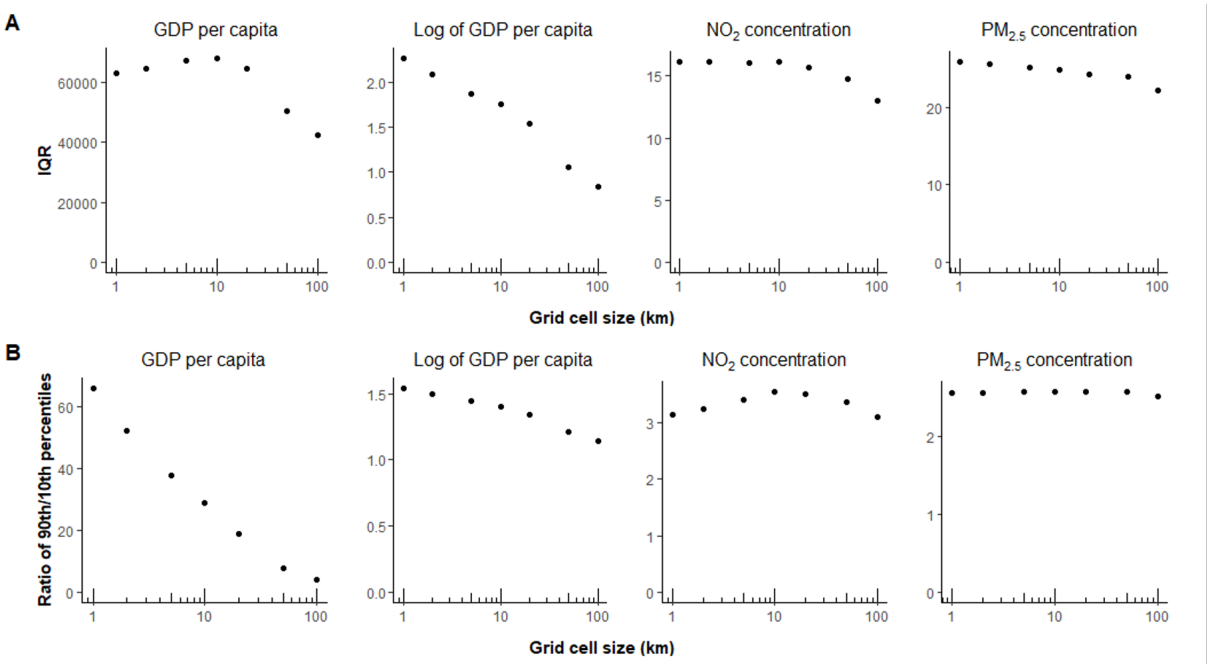

Figure 3.11: (A) IQR and (B) ratio of 90th versus 10th percentiles of GDP per capita, log GDP per capita, NO_2 concentration and $PM_{2.5}$ concentration by grid cell size.

Chapter 4

ADDRESSING US NATIONAL RACIAL-ETHNIC $PM_{2.5}$ EXPOSURE INEQUALITY

Material in this chapter was previously published in the following article: Y. Wang, *et al.*, Location-specific strategies for eliminating US national racial-ethnic $PM_{2.5}$ exposure inequality, *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2205548119 (2022).

4.1 Summary

Air pollution levels in the US have decreased dramatically over the past decades, yet national racial-ethnic exposure disparities persist. For ambient $PM_{2.5}$, I investigate three emission-reduction approaches, and compare their optimal ability to address two goals: (1) reduce the overall population average exposure ("overall average") and (2) reduce the difference in the average exposure for most-exposed racial-ethnic group versus for the overall population ("national inequalities"). I show that national inequalities in exposure can be eliminated with minor emission-reductions (optimal: $\sim$1% of total emissions) if they target specific locations. In contrast, achieving that outcome using existing regulatory strategies would require eliminating essentially all emissions (if targeting specific economic sectors) or is not possible (if requiring urban regions to meet concentration standards). Lastly, I do not find a tradeoff between the two goals (i.e., reducing overall average; reducing national inequalities); rather, the approach that does the best for reducing national inequalities (i.e., location-specific strategies) also does as well as or better than the other two approach (i.e., sectors-specific; meeting concentration standards) for reducing overall averages. Overall, my findings suggest that incorporating location-specific emissions reductions into the US air quality regulatory framework (1) is crucial for eliminating long-standing national average exposure disparities by race-ethnicity and (2) can benefit overall average exposures as much or more as the sector-specific and concentration-standards approaches.

4.2 Introduction

The Clean Air Act has dramatically reduced outdoor air pollution levels in the US, with (during 1990-2020) aggregate benefits exceeding costs ∼30-to-1 ($2t vs $65b) [218]. Important regulatory strategies include the National Ambient Air Quality Standards (NAAQS) and sector-specific emission-reduction technology requirements (e.g., Best Achievable Control Technology [BACT] standards). However, exposure inequalities persist [8, 10, 28, 31–33]. Disparities by race-ethnicity are larger than, and distinct from, those by income [8, 10, 20, 27, 28]. Racial-ethnic inequalities in US ambient air pollution and subsequent exposures are attributable in part to racist planning, including historical, race-based housing segregation and land-use practices [24,36–43]. Environmental racism scholars have suggested that strategies and policies for eliminating disparities will be most effective when racial-ethnic injustices are centered and directly addressed [44, 219–221].

The existing literature documents exposure inequities [7–11, 27, 28, 32, 33, 47] and investigates the impacts on inequities of emission changes for specific sources (e.g., [48–57]) or locations [12, 58–62, 222]. However, the scientific literature has not investigated how to eliminate national racial-ethnic inequalities in air pollution, nor what level of emission-reduction would be required to do so [25].

I examine three potential approaches to reduce or eliminate national exposure inequalities: (1) location-specific emission reductions (hereafter, "location"), (2) sector-specific emission reductions ("sector"; analogous to BACT-type approaches), and (3) requiring regions to meet a concentration standard ("NAAQS-like"). Approaches #2 and #3 mirror aspects of current regulations; approach #1 would be a new regulatory approach. I find that the "location" approach is by far the most effective (can eliminate national disparities with only small absolute emission-reductions); the "sector" approach is poor (can reduce disparities but requires substantially larger emission-reductions; cannot eliminate disparities except by eliminating nearly all emissions); "NAAQS-like" is the least effective (does not eliminate disparities). The "location" approach is also the strongest of the three for reducing population-average exposures.

4.3 Materials and methods

4.3.1 Emission data and the source-receptor matrix

I use the InMAP (Intervention Model for Air Pollution) Source-Receptor Matrix (ISRM) [121] to estimate, for the contiguous US, concentrations caused by anthropogenic emission. Specifically, the ISRM, which is simulated using InMAP [109], provides an estimate of the isolated impact of a 1-t emission change at any source location upon $PM_{2.5}$ concentration at each receptor locations.

Emission data, which are from the 2014 US EPA National Emission Inventory v1, are grouped into 14 source sectors, five pollutants, three emission stack heights (ground level: < 57 m; low stack height: 57 - 379 m; high stack height: > 379 m), and are allocated to the individual grid cells (52,411 in total) of the ISRM; see [25] for details. The 14 sectors are: (1) agriculture, (2) coal electricity utility, (3) non-coal electricity utility, (4) commercial cooking, (5) construction, (6) diesel heavy-duty vehicle, (7) gasoline light-duty vehicle, (8) industrial, (9) road dust, (10) residential gas combustion, (11) residential wood combustion, (12) residential others, (13) off-highway vehicle and equipment, and (14) miscellaneous. Emissions from biogenic, wildfire, and international sources are not investigated. The five pollutants are: primary $PM_{2.5}$ and four precursors of secondary $PM_{2.5}$: nitrogen oxides [NO_x], sulfur oxides [SO_x], ammonia [NH_3], and volatile organic compounds [VOC]. The spatial resolution of ISRM grid cells ranges from 1 km $\times$ 1 km (e.g., in densely-populated urban centers) to 48 km $\times$ 48 km (e.g., in remote, sparsely populated regions). The average (i.e., unweighted average) spatial resolution is 13.2 km overall (2.6 km in Urban Areas, 22.6 km in non-Urban Areas); the population-weighted average spatial resolution is 12.2 km overall (3.2 km [Urban], 20.9 [non-Urban]). The ISRM has separate matrices for the five pollutant types and three stack heights. The ISRM and emission data are freely available from `zenodo.org/record/2589760` and `zenodo.org/record/5831940`, respectively.

4.3.2 Calculating national exposure inequalities by race-ethnicity and social vulnerability

I calculate the total $PM_{2.5}$ concentration from all source-locations for each grid cell. National population-weighted average concentrations overall and for the race-ethnicity and social vulnerability subgroups are calculated as straightforward weighted averages: the sum of the multiplied values of population count and concentration for each grid cell divided by the total corresponding population. Race-ethnicity is classified according to the 2010 Census at block level. I focus

here on the four largest racial-ethnic groups: non-Hispanic White (63% of the total population; herein, "White"); Latino or Hispanic of any race (17% herein, "Hispanic"); non-Hispanic Black or African American (12%; herein, "Black"); and non-Hispanic Asian and Pacific Islander (5%; herein, "Asian"). Populations who identify as American Indian, as another race, or as multiracial are grouped here as "Other" (3%).

To identify vulnerable populations, I employ the Centers for Disease Control and Prevention / Agency for Toxic Substances and Disease Registry's (CDC/ATSDR's) Social Vulnerability Index (SVI, `www.atsdr.cdc.gov/placeandhealth/svi/index.html`). SVI indicates, by Census tract, the population ability to prevent human suffering and economic loss in a natural or human-caused disaster, or a disease outbreak. SVI is calculated from 15 social factors across four themes: socioeconomic status, household composition, race/ethnicity/language, and housing/transportation. CDC defines the 10% of Census tracts (excluding zero-population tracts) with the highest SVIs as "high-vulnerability" (HV) and remaining tracks as non- high vulnerability (non-HV); here, I label grid cells overlapping (ArcGIS command: INTERSECT) those tracts as "HV" (9.5% of the total population) and remaining grid cells as "non-HV" (90.5%). For additional sensitivity analyses, I also used two alternative definitions for HV locations: (1) the Census tracts in the highest 20% CDC SVIs (19.1% of the total population; in contrast, the main definition employs the top 10% of Census tracts); and (2) the Census tracts in the highest 10% of the $PM_{2.5}$ EJ index defined by EPA's EJScreen (Environmental Justice Screening and Mapping Tool, `www.epa.gov/ejscreen/environmental-justice-indexes-ejscreen`). $PM_{2.5}$ exposure disparity by race-ethnicity is calculated as the population-weighted average concentration for the most-exposed racial-ethnic group minus the population-weighted average concentration for the total population. In the modeled "initial conditions" (i.e., without any hypothetical/simulated emission-reduction), the most-exposed racial-ethnic group is Black. During the emission reduction procedure (see below), the most-exposed racial-ethnic group can be any group. $PM_{2.5}$ exposure inequality for HV is calculated as the population-weighted average concentration for the HV or the non-HV (the higher concentration between the two groups) minus the population average for the total population. I mainly focus on the absolute term of inequality ($\mu g/m^3$); I also report the relative (i.e., percent) change in inequality during the reduction procedure.

4.3.3 Calculating within-region and within-urban exposure inequalities

To calculate the within-region exposure inequalities, I use the geographic region classification by regional offices of US EPA. There are 10 EPA regions (`www.epa.gov/aboutepa/regional-and-geographic-offices`). I calculate the exposure inequalities within each of the 10 regions first, and then calculate the national average (population-weighted) of within-region inequalities.

The within-urban exposure inequalities are calculated based on the same idea. I use Core-Based Statistical Area (CBSA) as defined by US census (`www2.census.gov/geo/tiger/TIGER2019/CBSA/`). CBSA is the geographic unit that EPA uses for the National Ambient Air Quality Standards (NAAQS), and refers to both metropolitan statistical areas and micropolitan statistical areas. There are 894 CBSAs in the contiguous US including 379 metropolitan statistical areas and 515 micropolitan areas.

4.3.4 Emission reduction scenarios and the optimization methods

In order to perform optimization, I first calculate the $PM_{2.5}$ concentrations from every location (i.e., each grid cell; n=52,411 in total) & pollutant ("location" approach, 239,348 combinations [i.e., 52,411 locations; 5 pollutants; of the 262,055 maximum possible location-pollutant pairings, 22,707 have zero emissions and so are not considered here as an opportunity for emission-reduction.] in total) and from every sector & pollutant ("sector" approach, 61 combinations [i.e., 14 sectors, 5 pollutants; of the 70 maximum possible sector-pollutant pairings, 9 have zero emissions and are excluded here.] in total) for the total population and for each sub-population (race-ethnicity groups, HV). I then calculate the marginal concentration difference between each sub-population and total population, per one tonne emission change of each by-sector or by-location combination:

$$MCD_{i,j} = (C_{i,j} - C_{i,P})/E_i$$

Here, i represents a certain combination of sector & pollutant or location & pollutant (e.g., NO_x from road dust, SO_x from grid-cell #100, etc.); j represents a certain sub-population (e.g., Hispanic population, HV, etc.); P represents the total population; $C_{i,j}$ is the population-weighted average $PM_{2.5}$ concentration from source/location i for group j; $C_{i,P}$ is the population-weighted average $PM_{2.5}$ concentration from source/location i for the total population; E_i is the total emission from source/location i; $MCD_{i,j}$ is the marginal concentration difference between group i and the population average, per one tonne emission from source/location j.

The optimization determines the order in which emissions will be reduced, and models the exposure impacts as those emission-reductions occur. The optimization is performed using the following algorithm:

1. Determine the most-exposed racial-ethnic group or the more-exposed social vulnerability group (j_0) at the initial (i.e., zero emission-reduction) condition, and then determine the combination (i_0) with the largest marginal concentration difference (MCD) for that subgroup.

2. Reduce all emissions from i_0, and then calculate the remaining concentrations for the total population and for all subgroups.

3. Repeat step #1, for the new conditions: Determine the most-exposed racial-ethnic group or the most-exposed social vulnerability group (j_1) for the updated concentrations, and then find out the combination (i_1; from the remaining available options) with the largest MCD for that subgroup.

4. Repeat step #2, for the new conditions: Remove all emissions from i_1, and then calculate the remaining concentrations for the total population and for all subgroups.

5. Repeat steps #3 and #4 until the emissions from all combinations have been reduced to zero. The concentrations for the total population and all subgroups also reach zero at this point, and the optimization ends.

Via the steps above, I simulate the local optimum solution to minimize the concentration inequality.

4.3.5 Regional and urban -level optimizations

In addition to optimization to reduce the overall national disparities (see above), I also explore the optimization methods for reducing disparities at regional and urban scales. I assume that each region or Urban Area reduces emissions in order to optimally reduce the inequalities within that region/Urban Area. I use the ten EPA regions and US Urban Areas defined by the US census (`www2.census.gov/geo/tiger/TIGER2018/UAC`). I consider only Urban Areas (as defined in

the US Census: population: 50,000 or more) with more than 20 ISRM grid cells within their geographic boundaries (n = 171). (If Urban Areas are divided into small/medium/large based on national population tertiles, as described in [194], my approach provides 10 large, 44 medium, and 117 small Urban Areas.) I explore both optimization approaches ("location"; "sector").

4.3.6 "NAAQS-like" methods

For the "NAAQS-like" approach, I set CBSAs as the emission reduction unit and set one concentration target for all CBSAs. The NAAQS-like goal is that the concentrations in each grid cell in a CBSA doesn't exceed the concentration target (i.e., the NAAQS – National Ambient Air Quality Standard). I separately employ multiple concentration targets, including 10, 9, 8, 7, 6, and 5 $\mu g/m^3$ (the recently updated WHO guideline for annual average $PM_{2.5}$ concentration is 5 $\mu g/m^3$). For this approach, the emissions reduction algorithm is:

1. Check each CBSA to see whether the maximum concentration within the CBSA is lower than the concentration target.

2. For those CBSAs that exceed the concentration target, reduce the emission for all sources and locations within the CBSA by 1% of the total emission (i.e., proportional to sources and locations), and calculate the updated grid-cell concentrations for each CBSA.

3. Repeat step 1: Check each CBSA to see whether the updated concentrations reach the target.

4. Repeat step 2: For those CBSAs that still exceed the target, reduce all emissions within the CBSA by 1% of the remaining emission or by 1 tonne/year (use the larger value between the two), and calculate the updated grid-cell concentrations for each CBSA. The minimum reduction unit is one tonne/year.

5. Repeat the steps 3) and 4) until all CBSAs have reached the target (the NAAQS) or reached zero emission.

The rationale for some of the specific steps in this optimization are as follows: For some CBSAs, even if I reduce all the emission within their boundary, the maximum concentration in that CBSA will still exceed the target (the NAAQS). For those cases, I avoid the optimization entering an

infinite loop via these steps described above: (1) having a small emission reduction (1 t/y) as the minimum reduction per optimization step, and (2) halting the optimization if all emissions being considered have become zero.

4.4 Results

The base (no emission reductions) model predicts the population-average $PM_{2.5}$ concentration (units: $\mu g/m^3$) is 7.0; for racial-ethnic sub-populations it is 6.5 (White), 8.5 (Black), 7.7 (Hispanic), 8.0 (Asian), and 6.6 (other). The concentration disparity (which the "location" and "sector" approaches aim to optimally reduce) is 1.4 $\mu g/m^3$ (20%) for non-Hispanic Black (herein "Black"); a value that is consistent with empirical analyses [10, 28].

The results reveal that the "location" approach is substantially more effective and more efficient in reducing concentration disparities for racial-ethnic groups and population average concentration than "sector" or "NAAQS-like" approaches (Fig. 4.1). "Effective" refers to successfully eliminating disparities; "efficient" refers to the reduction in disparity per unit reduction in emissions. The "location" approach is the most effective of the three approaches in part because it is so much more efficient. For example, it requires 28-fold less emission-reduction to achieve a 50% reduction in racial-ethnic concentration disparities compared to the "sector" approach (0.04 MT/y (relative value: 0.1% of total national emissions) for "location" versus 1.2 (4%) for "sector"). To reduce the disparity by 90% and 99%, respectively, the analogous emission reductions are 54-fold different (0.2 (0.7%) for "location" versus 12 (40%) for "sector") and 83-fold different (0.4 (1.2%) versus 30 (99%)).

If optimizing to reduce overall population-average concentrations (rather than disparities), the "location" and "sector" approaches provide similar improvements initially (Fig. 4.1: for the first ∼3 MT/y in emission reductions, both approaches reduce the average concentration to ∼4 $\mu g/m^3$), after which the "location" approach is comparatively more effective (reduces the concentration quicker). Historically, total emissions of the five pollutants declined ∼13 MT (27%) from 2010 (46 MT) to 2020 (33 MT), an annual change of ∼1.3 MT (3.2%) [223].

"NAAQS-like" does not eliminate disparities and is much less efficient than "location" and "sector" (Fig. 4.1), i.e., it requires dramatically greater emission reductions to achieve comparable reductions in concentration disparities. For example, having all US urban areas meet a NAAQS-like concentration standard of 6 $\mu g/m^3$ (much lower than the current NAAQS [12 $\mu g/m^3$]) would require

9.1 MT/y of emission reduction (30.% of total national emissions). This would reduce the population average concentration dramatically, to 3.5 $\mu g/m^3$; but, it would only lower, not eliminate, racial-ethnic disparities (to 0.6 $\mu g/m^3$ (17%)). My finding that the "NAAQS-like" and "sector" approaches cannot eliminate racial-ethnic disparities even after substantial emission reductions is consistent with recent historical analyses [10, 28, 32, 35].

As described in Appendix I, I conducted several sensitivity analyses, such as considering relative, rather than absolute, inequality; urban or regional, rather than national, disparities; disparities for government-defined high-vulnerability locations, rather than racial-ethnic disparities; and, average concentrations, rather than disparities (see Appendix I; Figs. 4.2 to 4.8). For "NAAQS-like", I considered several values for the standard (Fig. 4.9). In each case, the conclusion still holds, that the "location" approach is by far the most efficient approach. None of the NAAQS-like scenarios eliminate disparities.

The "location" approach (Fig. 4.10A) prioritizes urban emission reductions in the Midwest and Southeast where there are clusters of (i) emission sources and (ii) the most-exposed racial-ethnic group. For the national-level optimization, higher prioritization for emission-reductions is consistently correlated in univariate (Fig. 4.11) and multivariate (Table 4.1) analyses with lower income, higher density of people and of emissions, higher percentage of non-White, higher percentage of Black, and greater levels of segregation. In contrast, the relationship with percentage of Hispanic and Asian varies across models (see Appendix II). The spatial extent of high-priority locations varies by urban area (Figs. 4.10, 4.12 and 4.13, e.g., is comparatively large in Los Angeles and Atlanta [>50% of the area is in the top 50% reduction priorities] and comparatively small in New York, Boston, Miami, Dallas, and Philadelphia [<20% of the area is in the top 50% reduction priorities]). In addition, the high-priority location(s) in an urban area can be in one cluster (e.g., Philadelphia, Chicago, Washington, DC, Atlanta) or many clusters (e.g., New York, Houston, Dallas, Boston). These between-urban differences likely reflect differences in within-urban patterns in emissions and racial-ethnic segregation [224–226].

For the "location" approach, comparing the base-case with the sensitivity analysis optimizing reduction of within-urban (rather than national) disparities (Fig. 4.14A, Figs. 4.15 and 4.16), the latter has more spatial variability than the former, but both approaches identify similar locations for emission reduction. This suggests that identifying and reducing emissions in those high-impact locations would in many cases reduce both urban-scale and national-scale disparities.

The "sector" approach (Fig. 4.14B, Figs. 4.17 to 4.19) prioritizes emission reduction from commercial cooking, road dust, residential gas combustion, and construction. Those sources have the largest marginal benefits and so are ranked highly by my optimization. However, those four sectors have relatively small absolute emissions and therefore only modestly reduce the total disparities (reduce 0.4 $\mu g/m^3$ [31%] of disparity in total, Fig. 4.18). Industrial has the largest total emissions (9 MT/y [30%]) and also the largest contribution to inequality (0.3 $\mu g/m^3$ [25%]). However, the marginal improvement per MT/y is comparatively low, emphasizing that for industrial emissions, large emissions reductions are needed to achieve large benefits to disparities. Industry contains many types of sources; exposure disparities may be more sensitive to specific types of industrial sources than the industrial sector as a whole.

Comparing the five types of emissions that contribute to $PM_{2.5}$ – "primary" (directly-emitted) $PM_{2.5}$, and four precursor species that can form secondary $PM_{2.5}$ – primary $PM_{2.5}$ and NH_3 have the highest reduction priorities for all sectors. Reduction of primary $PM_{2.5}$ emissions causes the largest inequality reduction for most of the sectors (account for 57% of the total disparity). For addressing $PM_{2.5}$ disparities, the precursors VOC and NO_x have the lowest reduction priorities for all sectors.

4.5 Discussions and conclusions

My analysis provides insight for general, archetypal emission reduction approaches. Limitations of my approach include the following. I do not consider several important factors, including emission reduction costs, technologies, and enforcement. Additionally, this study focused on ambient concentrations, which are related but distinct from individual-level exposures and risks; indoor air pollution, micro-environments, and individual's mobility can also contribute to exposure disparities [60,62,227–229], and background rates of disease can modulate disparities by concentration [230]. I focused on disparities in $PM_{2.5}$ because most of the health damages from air pollution are attributed to $PM_{2.5}$ [1,231]. I do not explore other air pollutants, climate effects, and toxicities. My approach employs national-level source-receptor matrix and emission inventories, which include documented uncertainties [25, 109, 121]. My results are robust at national and aggregate level; for investigating specific locations, I recommend additional (local) data and analyses, including local air quality data, local emission sources, and source-receptor matrix at finer resolution. My location analysis was performed at, on average, a 3.2-km scale in Urban Areas; future work can investi-

gate whether greater spatial resolution provides additional efficiencies [16]. A more-recent emission inventory is available (2017 is available; I used 2014); I do not expect that my core conclusions would change for an inventory that is 3 years different. Finally, I grouped sources into 14 sectors; there may be additional insight from considering more granular categorization, especially for the industrial sector.

Exposure disparities are a legacy of race-based planning [20, 35, 37, 38, 41, 41, 42, 44]. They reflect systematic discrimination, housing segregation, and segregation in the proximity to pollution sources: for almost every sector of economy, at various spatial scales, and persisting for decades [20, 25, 46]. My study highlights the need for a fundamentally different framework for national air quality regulation in the US that involves location-focused emission reductions in order to address national racial-ethnic exposure disparities. That framework would help accelerate efforts to redress the harms caused by environmental racism.

My findings can inform national action (e.g., implementation of the Biden-Harris Administration's Justice40 Initiative [232, 233]) and emerging state and local environmental justice laws to identify overburdened communities and develop community emissions reduction plans [234–236]. My study supports the long-standing request from environmental justice communities and local organizations for location-specific solutions that center overburdened communities [37, 158, 237–239]. My results also support putting in safeguards to address the potential for pollution trading (e.g., GHG-focused cap-and-trade) to exacerbate pollution inequities, especially for already overburdened communities [240–242].

My results can help inform *where* to specifically target emission reductions, but more research is needed. Future work can further explore how various location-specific emission-reduction strategies – framed at a specific spatial scale (e.g., regional, state, urban, neighborhood) and incorporating hyperlocal emission sources, community characteristics, and context (e.g., historic and contemporary zoning, planning policies, engagement of community groups) – can swiftly achieve benefits across heavily impacted communities. Because $PM_{2.5}$ includes primary and secondary components (i.e., is emitted and also can form in the atmosphere), the most efficient locations to reduce emissions might be in a community or might be upwind. The present work considers national average inequalities (i.e., the difference in the average exposure for most-exposed racial-ethnic group versus for the overall population); future work should investigate other aspects of inequalities (also considering, e.g., exposure distributions, not just averages; other demographic attributes, not just race-ethnicity; and

other geographic units) as well as broader aspects of achieving environmental justice (e.g., remedies for past harms). Tools are urgently needed that can (1) connect local and national decisions (e.g., planning and zoning changes, infrastructure investments, provision of public services such as mass transit, and household installation of solar panels) with benefits to highly-impacted communities, and (2) support new and innovative approaches to environmental improvement (e.g., low-emission zones, reflecting cumulative burdens).

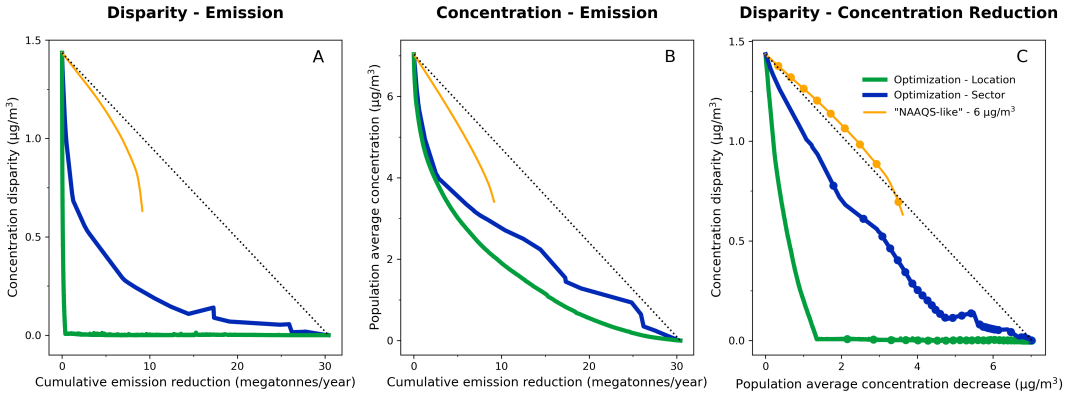

Figure 4.1: $PM_{2.5}$ exposure disparity and concentration reduction curves. (A) Concentration disparity between the most-exposed racial-ethnic group (Asian, Black, Hispanic, White, or Other) and the population average (y-axis) versus cumulative emission reduction (x-axis). (B) Population average concentration versus cumulative emission reduction. (C) Concentration disparity versus population averaged concentration (i.e., the y-axis values from the left and middle panels). For each panel, current conditions are the left side (i.e., "do nothing" at x=0) and a complete (100%) emission-reduction is the right side (i.e., achieving zero emissions: lower-right, at x=30.4 MT/y). Each panel compares three approaches to emission-reduction: "location" (green line), "sector" (blue line), and "NAAQS-like" (i.e., employing a concentration standard, here 6 $\mu g/m^3$; orange line). An "equal reduction" approach, where all emissions are reduced proportionately, would be a straight line (black dotted line). The "location" approach (green line) can eliminate national disparities with modest total emission reductions, whereas with the other two approaches, national disparities remain even after substantial emission reductions (see left and right panels). The "location" approach also does as well as or better than the other two approaches, for population average concentration (see middle and right panels)

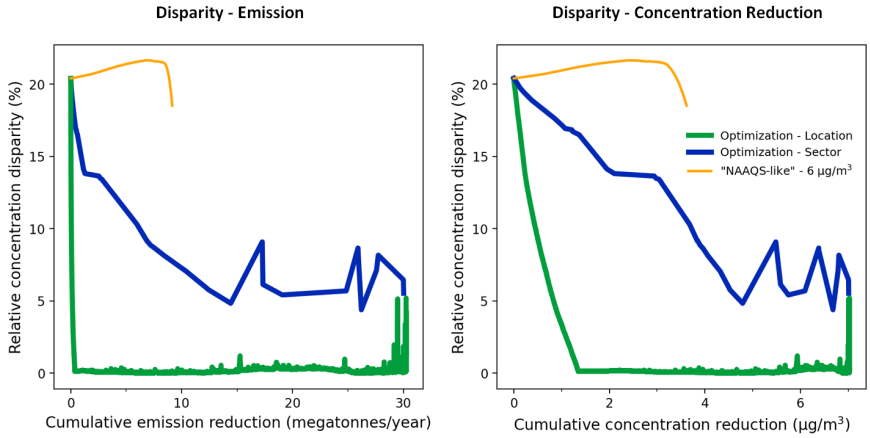

Figure 4.2: Relative $PM_{2.5}$ exposure disparity changes with emission reduction and concentration reduction.

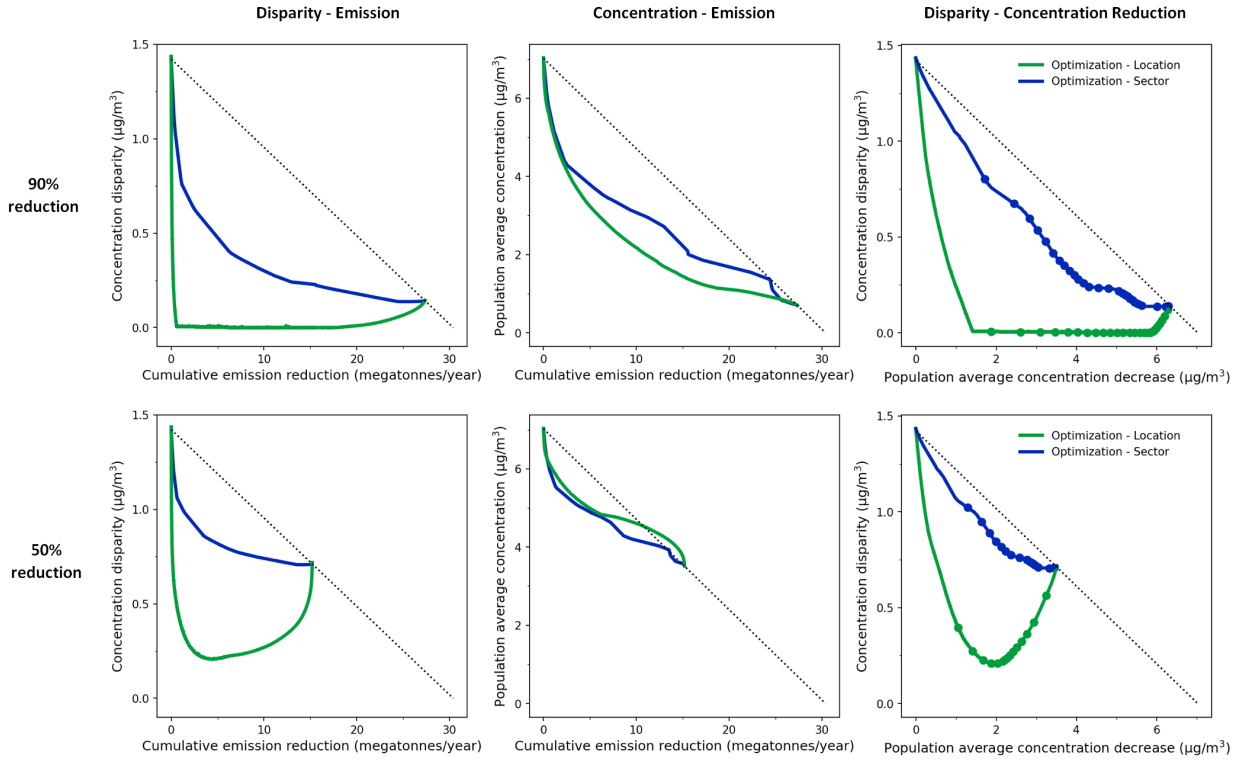

Figure 4.3: $PM_{2.5}$ exposure disparity and concentration reduction curves for the alternative conditions of (90% and 50%) emission reduction. Where a line trends upward (i.e., has a positive slope), any emission reduction would increase the exposure disparity (between the most-exposed racial-ethnic groups and the population average); here, the optimization procedure priorities emission-reductions with the lowest marginal increase in the exposure disparity.

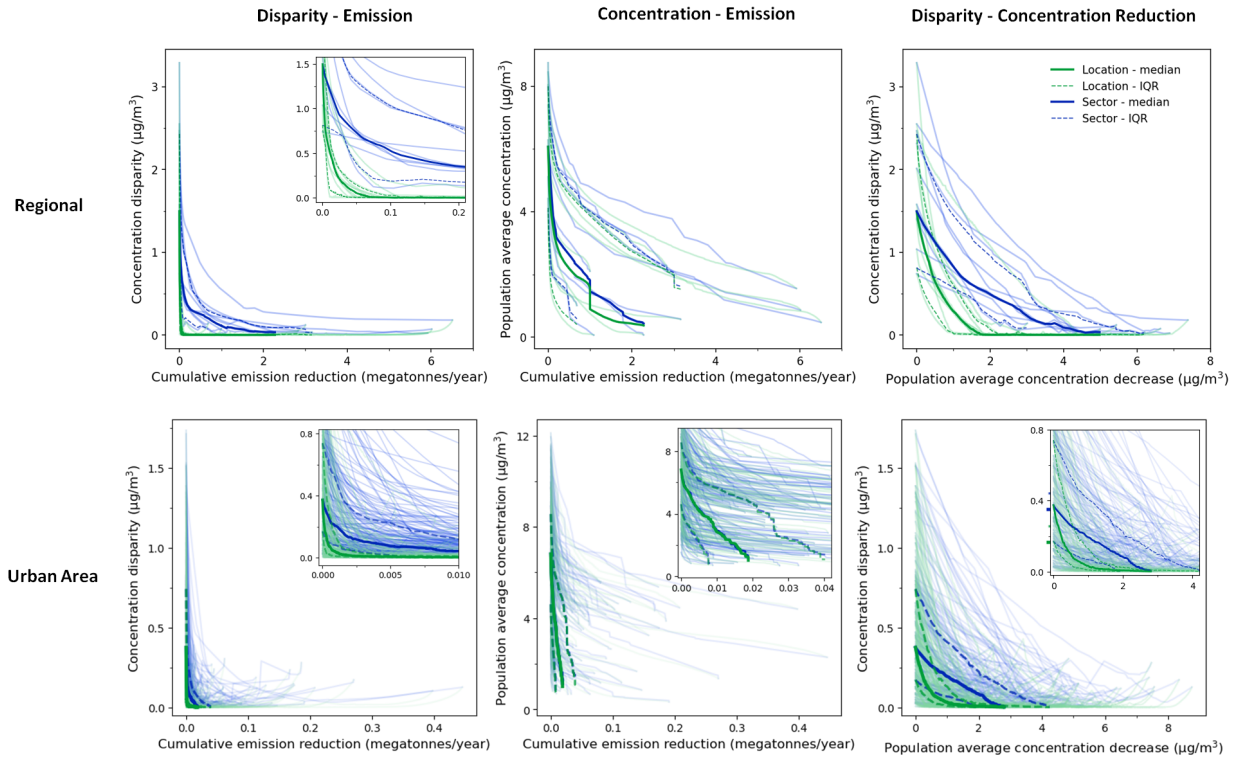

Figure 4.4: $PM_{2.5}$ exposure disparity and concentration reduction curves. Top row: within-region results, reflecting each region's emission-reductions to optimally reduce disparities in that region. Each light-color line reflects one US EPA region (n=10); median and interquartile range (IQR) are dark-color lines. Bottom row: within-urban results, reflecting each Urban Area's emission-reductions to optimally reduce disparities in that Urban Area [UA]. Each light-color line reflects one UA (n=171); median and IQR are dark-color lines. Some panels display zoom-in results in a sub-panel. For both rows, the location-based approach eliminates racial-ethnic disparities in exposure well before the source-based approach (i.e., the green line is below the blue line). For example, at the regionally level the location-based approach rapidly reduces disparities to zero; the source-based approach does not.

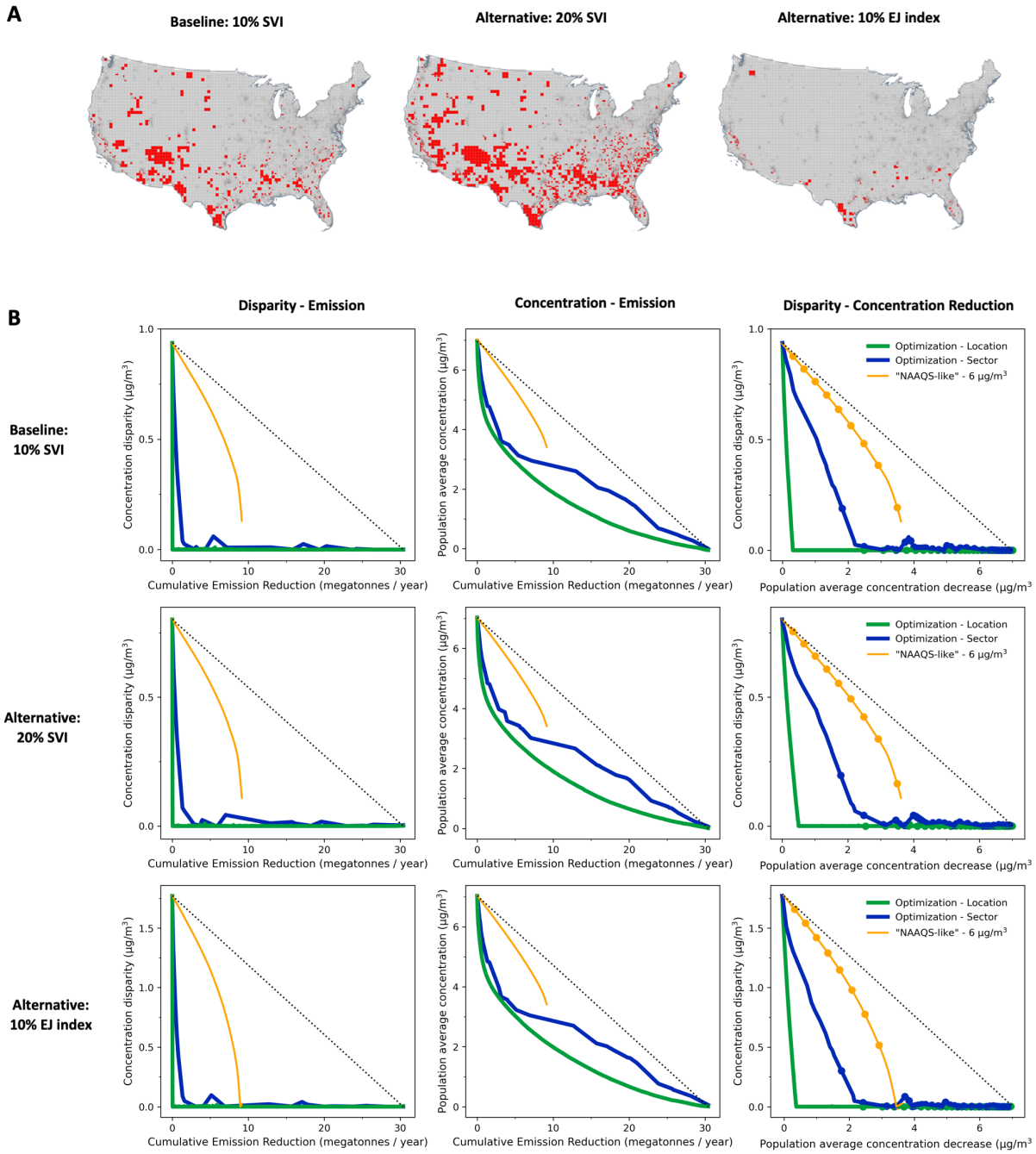

Figure 4.5: Distribution map of "high vulnerability" locations, and $PM_{2.5}$ exposure disparity and concentration reduction curves for HV locations.

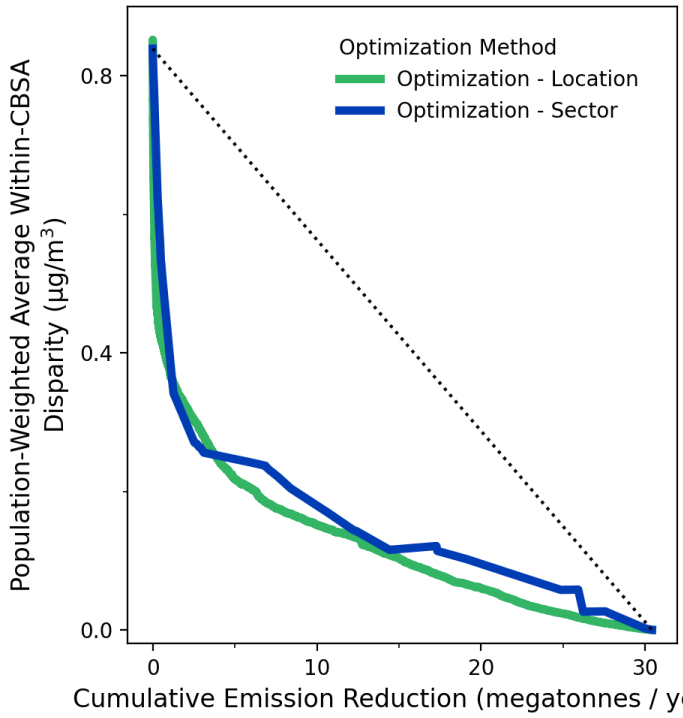

Figure 4.6: Urban disparity reduction curves for the two optimization approaches.

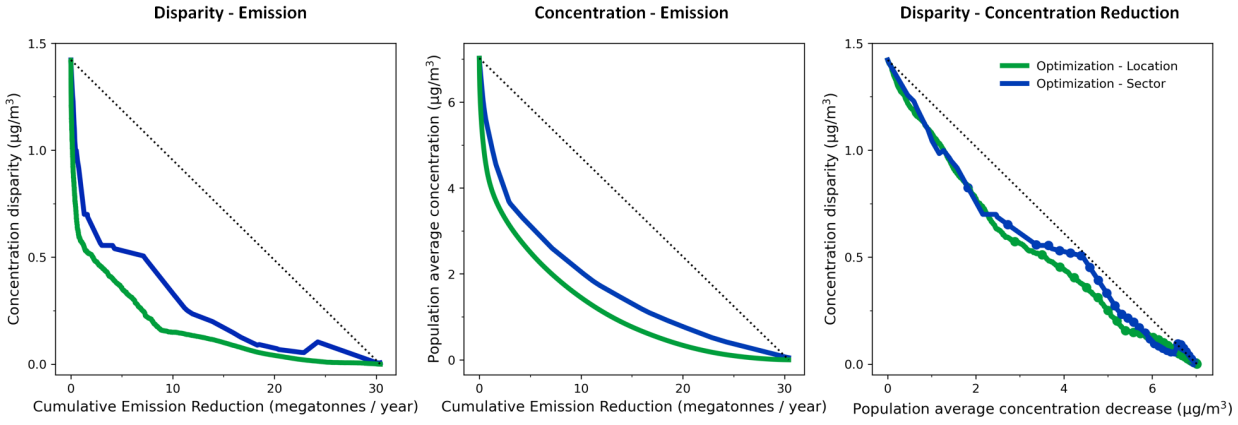

Figure 4.7: $PM_{2.5}$ exposure disparity and concentration reduction curves reflecting optimization to reduce average exposure concentration.

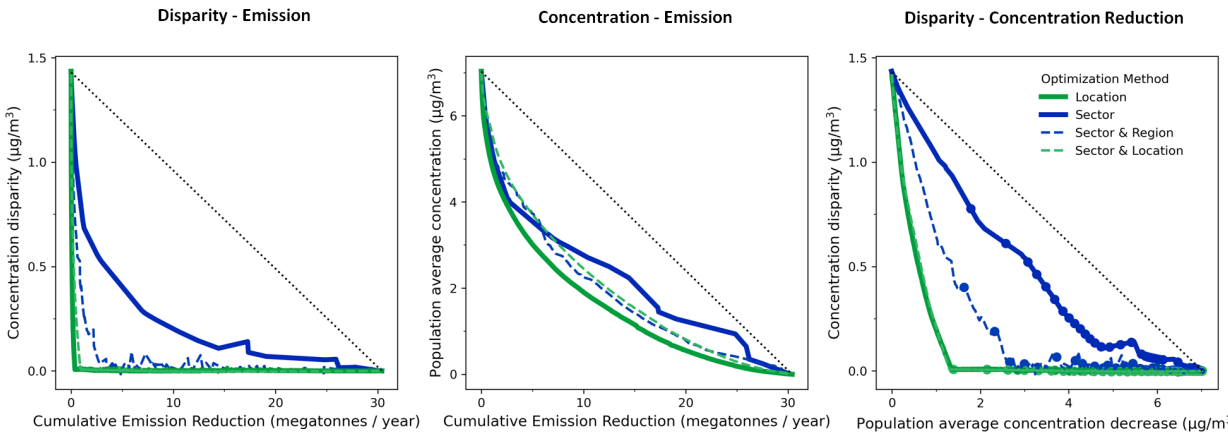

Figure 4.8: $PM_{2.5}$ exposure disparity and concentration reduction curves, comparing four approaches to emission-reduction: optimization by sector (blue line, same as Fig. 4.1), optimization by sector and geographic regions (blue dash line), optimization by location (green line, same as Fig. 4.1), and optimization by sector and location (green dash line).

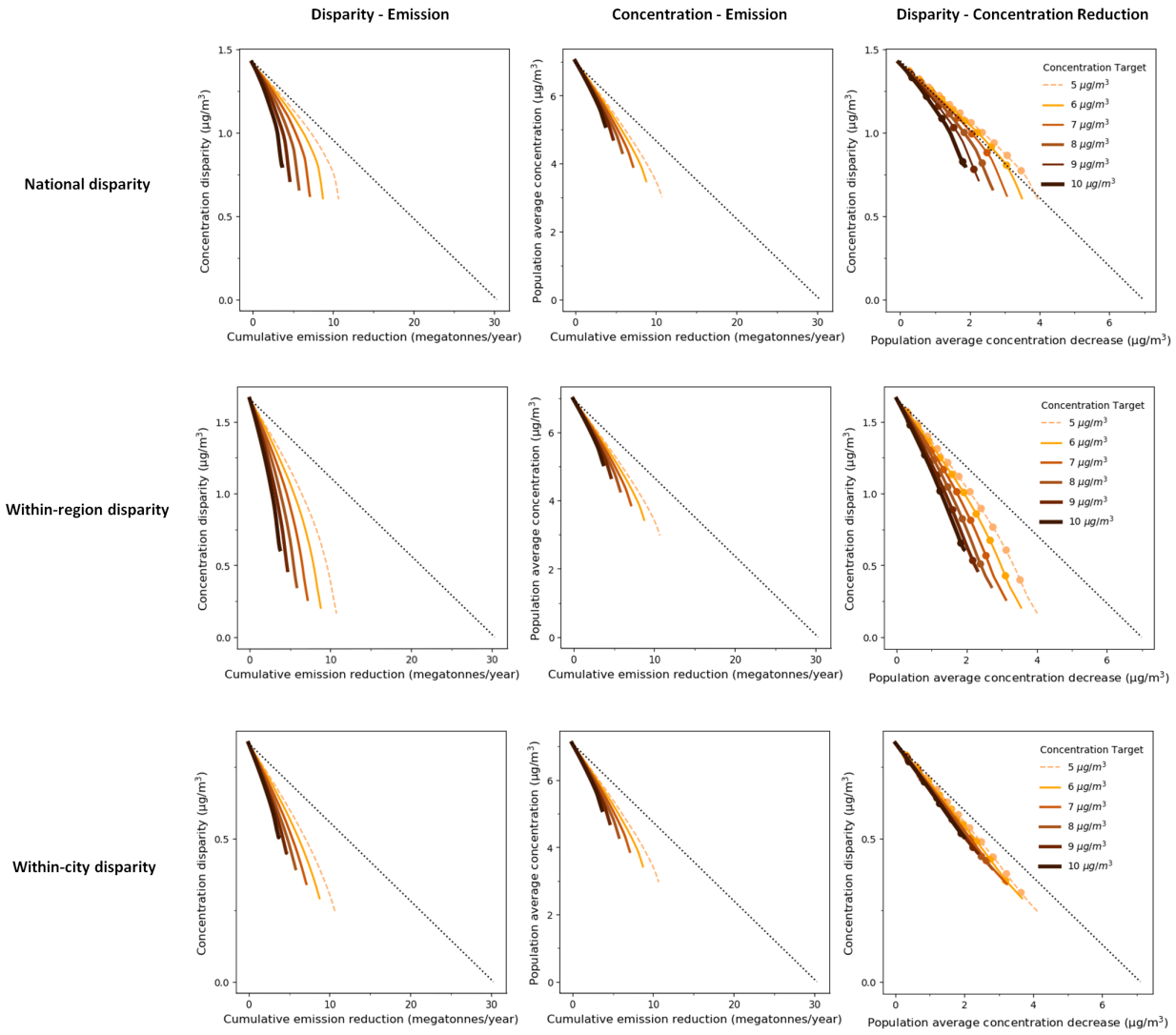

Figure 4.9: $PM_{2.5}$ exposure disparity and concentration reduction curves for "NAAQS-like" approach. Rows and columns are analogous to Fig. 4.1. Here, each CBSA reduces emissions inside that CBSA to meet the concentration target (5, 6, 7, 8, 9, or 10 $\mu g/m^3$); the figure shows the resulting disparities and concentrations (top row: nationally; middle row: by regional; bottom row: by CBSA). None of the scenarios investigated here result in disparities reaching zero.

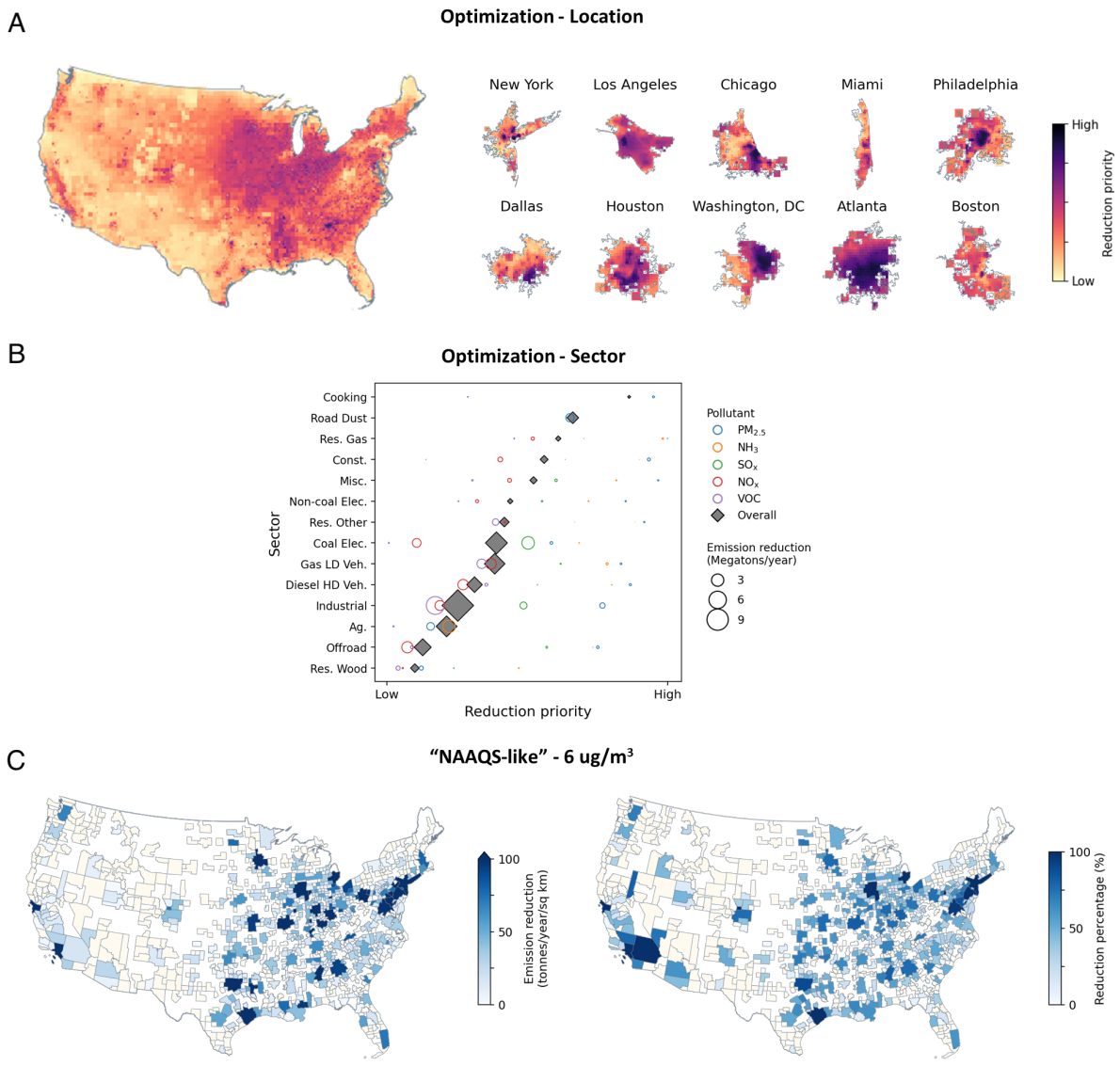

Figure 4.10: Emission reductions for the three approaches: (A) by location (i.e., corresponding to the green lines, Fig. 4.1), (B) by sector (corresponding to blue lines, Fig. 4.1), and (C) "NAAQS-like" (orange lines, Fig. 4.1). Panel (A) displays national results (left) and zoomed-in results for 10 large Area (right). Spatial units displayed in panel (C) are Core-Based Statistical Areas (CBSAs), the geographic unit for NAAQS evaluation. The three approaches offer fundamentally different ways of formulating and prioritizing emission reductions.

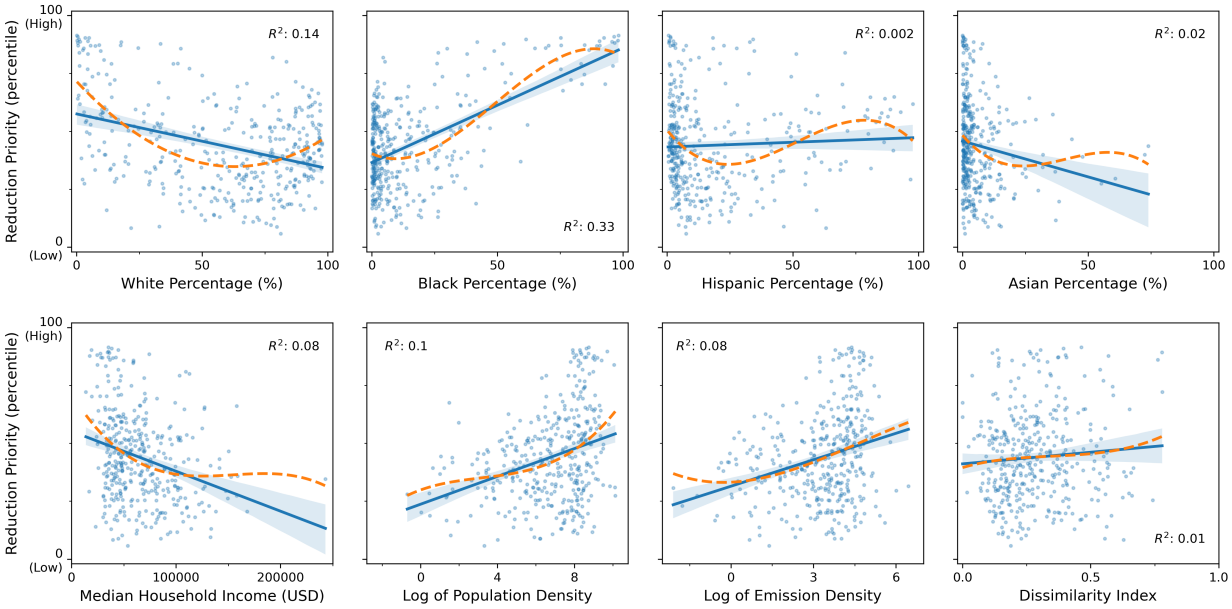

Figure 4.11: Scatter plots with best fit line (blue lines) and spline smoothing line (orange lines; order = 3) of reduction priority versus racial-ethnic composition, household income, population density, emission density, and racial segregation index in the location (grid cell). Reduction priority is converted to 0–100 scale: 100 represents highest priority, 0 represents lowest priority. Points, best-fit lines, and regression R-squares are for the 1% random sub-sample of all the locations with none missing value, non-zero emissions, and non-zero populations (n = 398). Population density and emission density are at log-scale. The unit of population density is log of persons per square kilometer; the unit of emission density is log of tonnes per square kilometer.

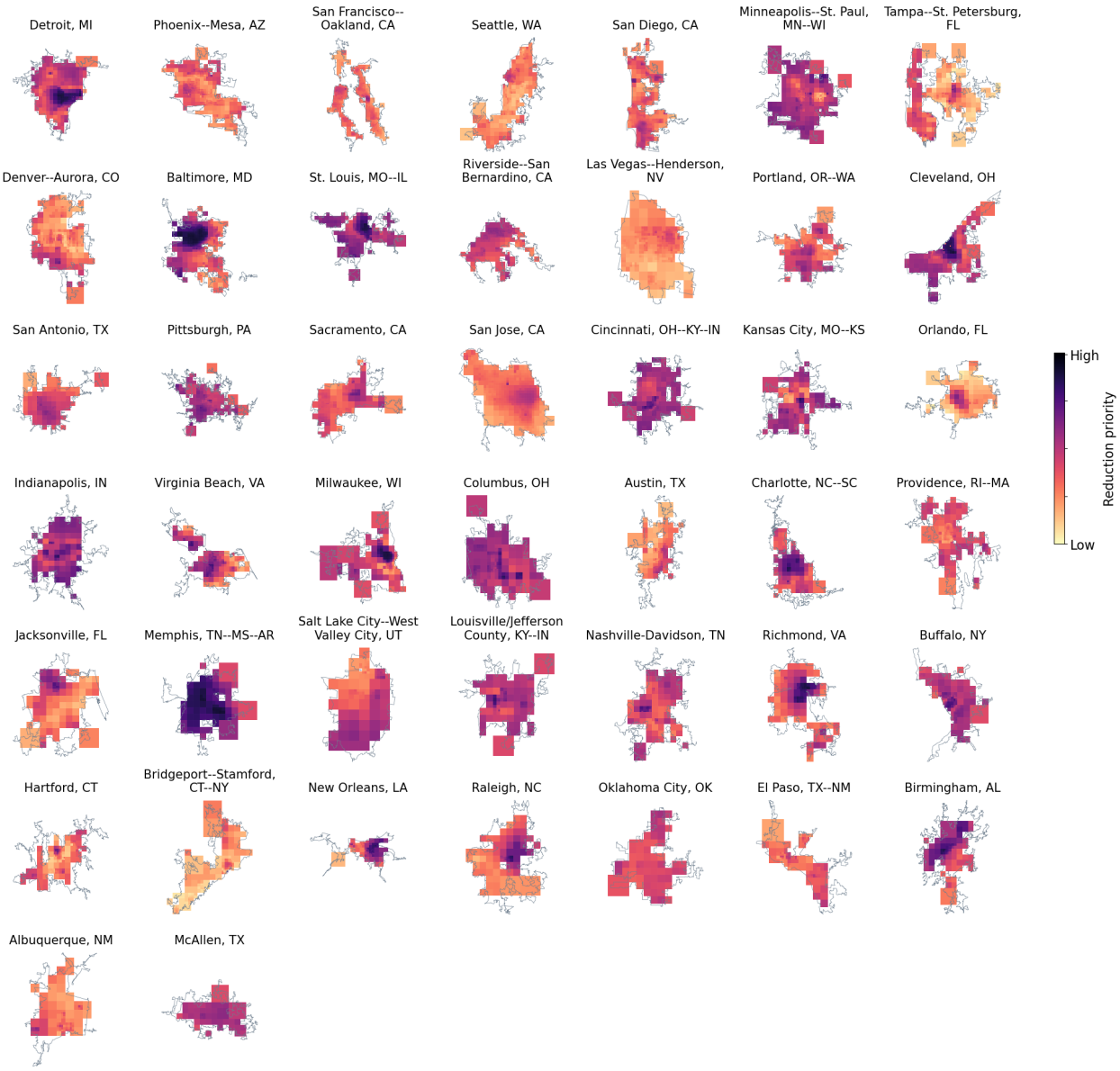

Figure 4.12: Reduction priority maps for optimization by location methods for 44 medium Urban Areas.

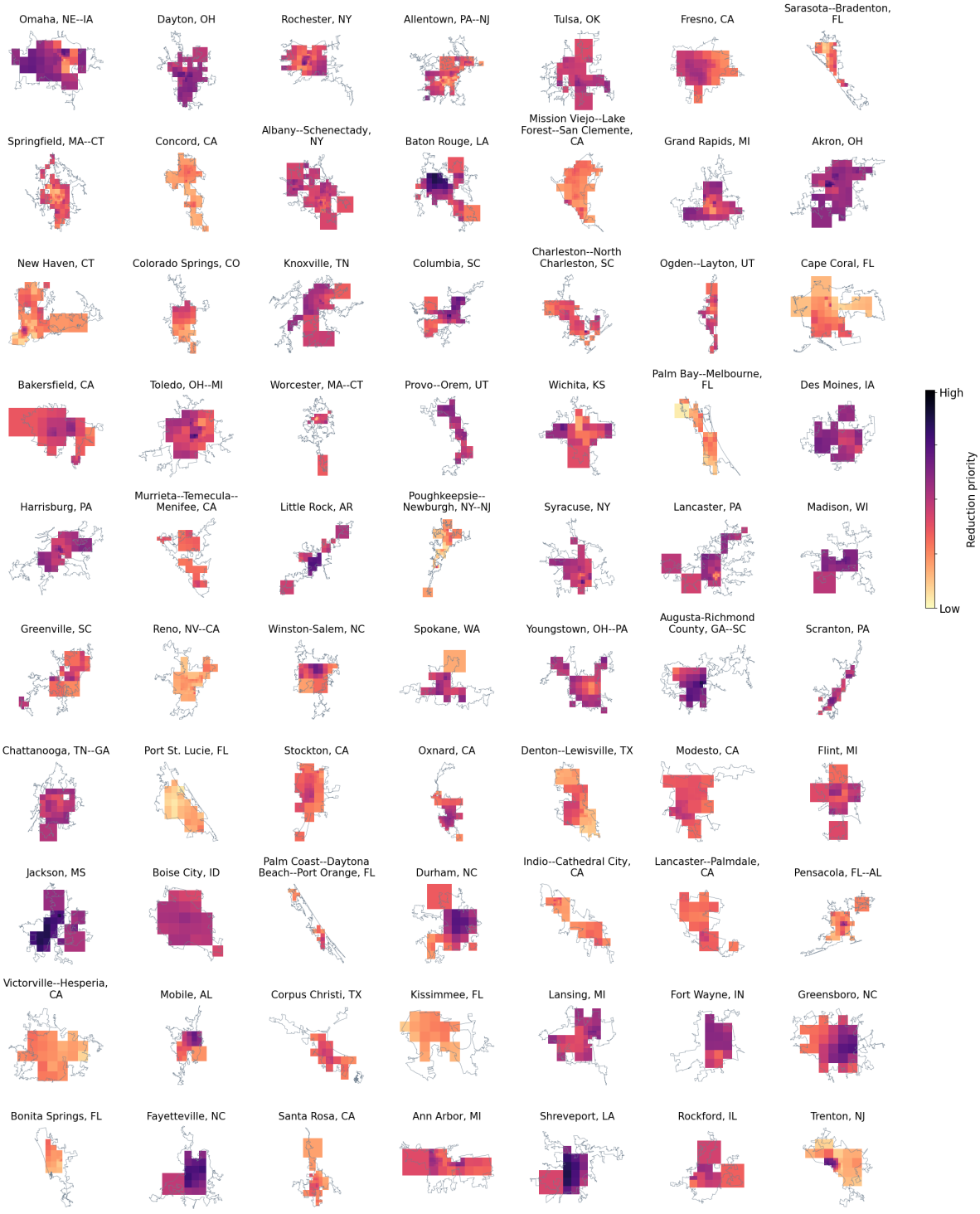

Figure 4.13: Reduction priority maps for optimization by location methods for 70 (out of 381) small Urban Areas.

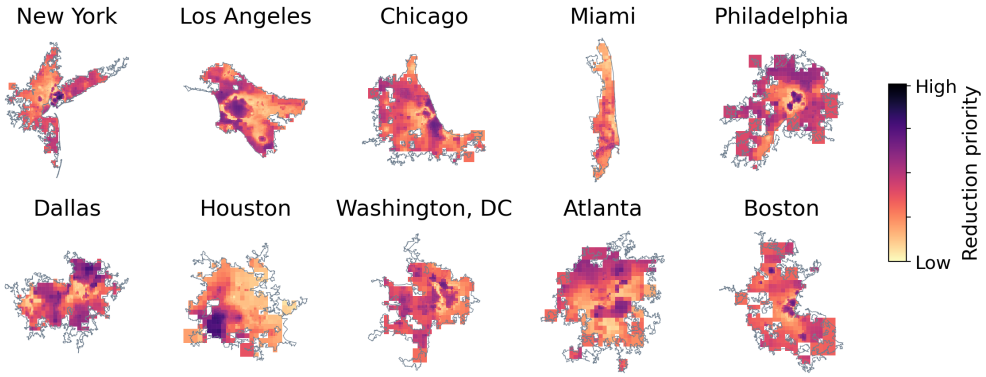

Figure 4.14: Reduction priority maps for urban-level optimization by location methods for 10 large Urban Areas.

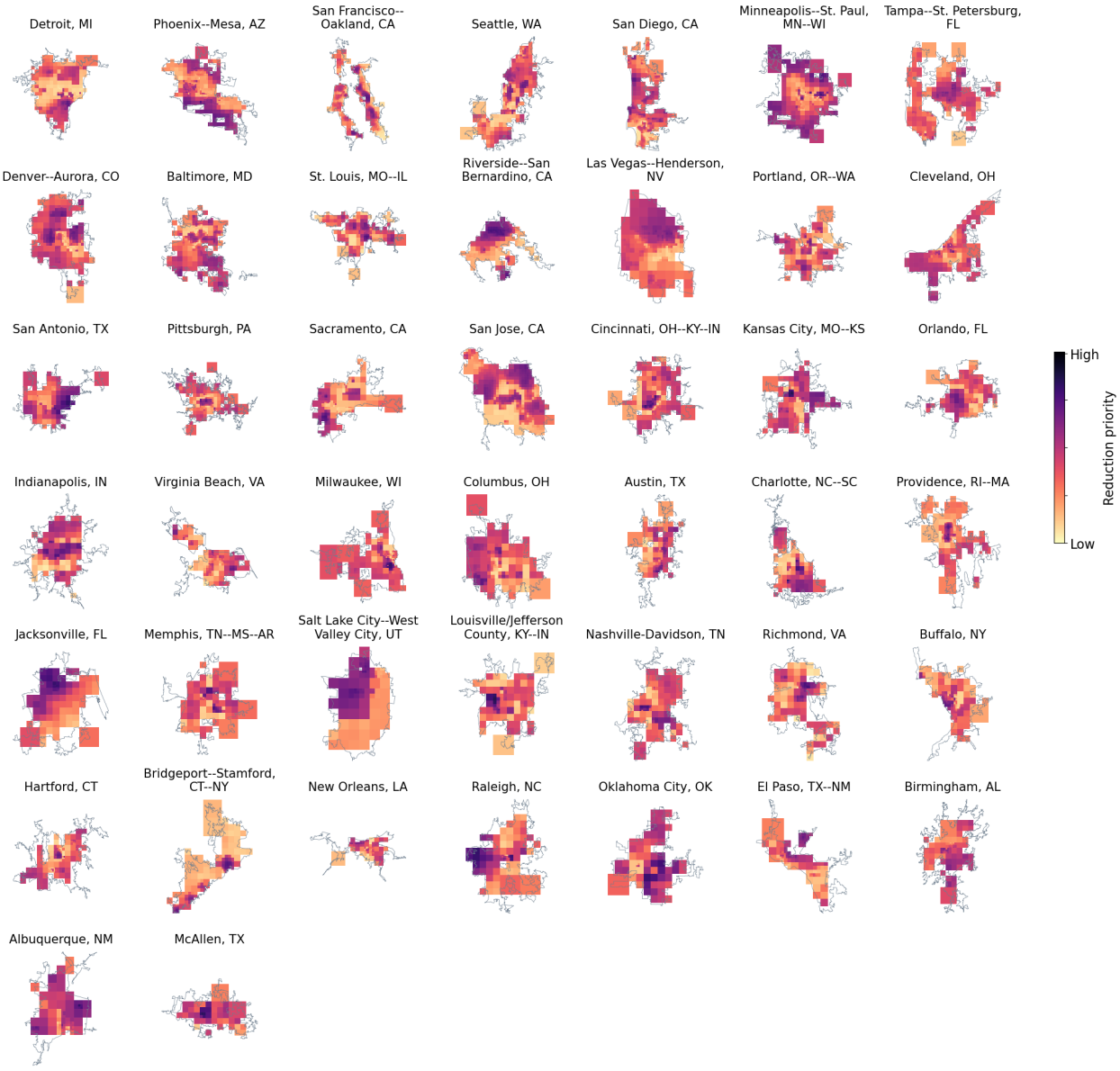

Figure 4.15: Reduction priority maps for urban-level optimization by location methods for 44 medium Urban Areas.

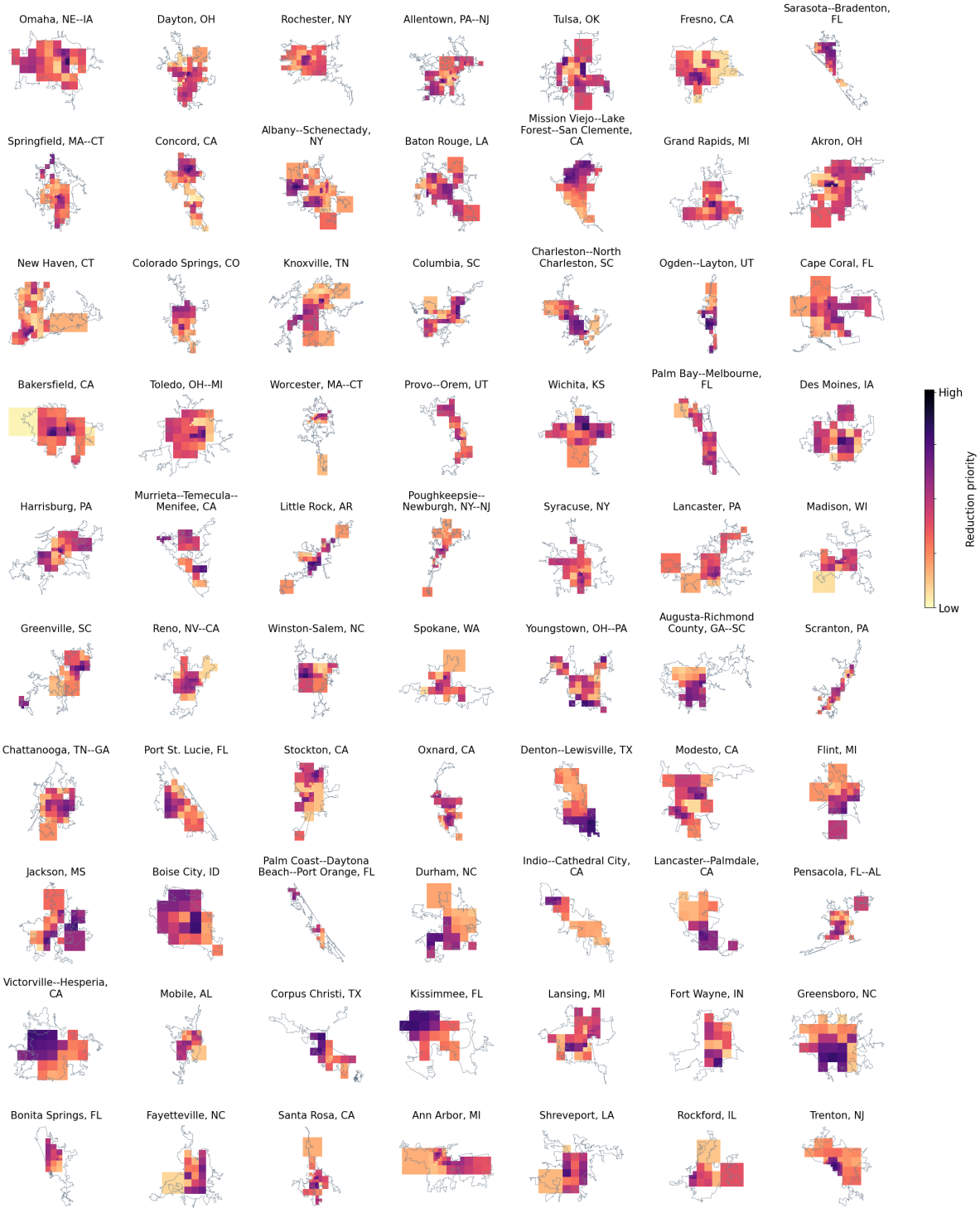

Figure 4.16: Reduction priority maps for urban-level optimization by location methods for 70 (out of 381) small Urban Areas.

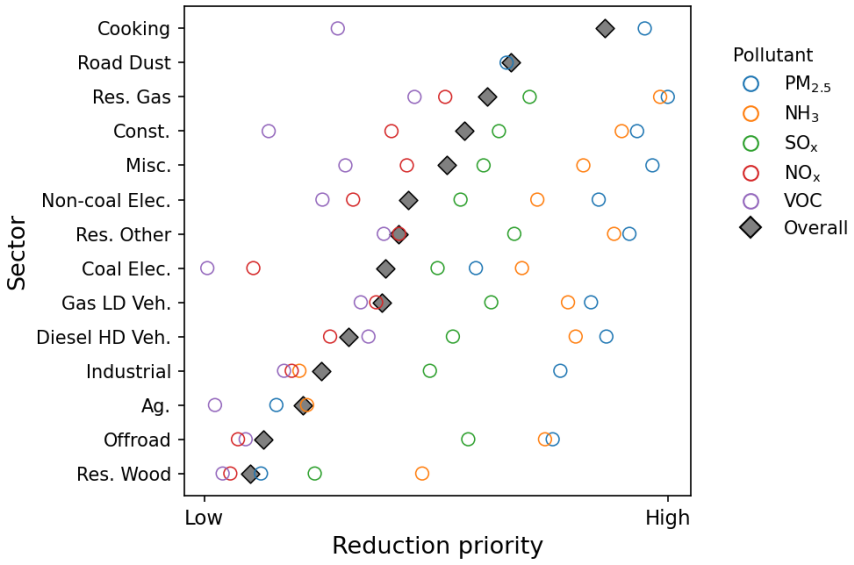

Figure 4.17: Emission reduction priority for optimization by sector method. The plot is an alternative version of Fig. 4.10b-left, where the icons are equally size (so they are more easily visible) instead of sized proportionately to emissions.

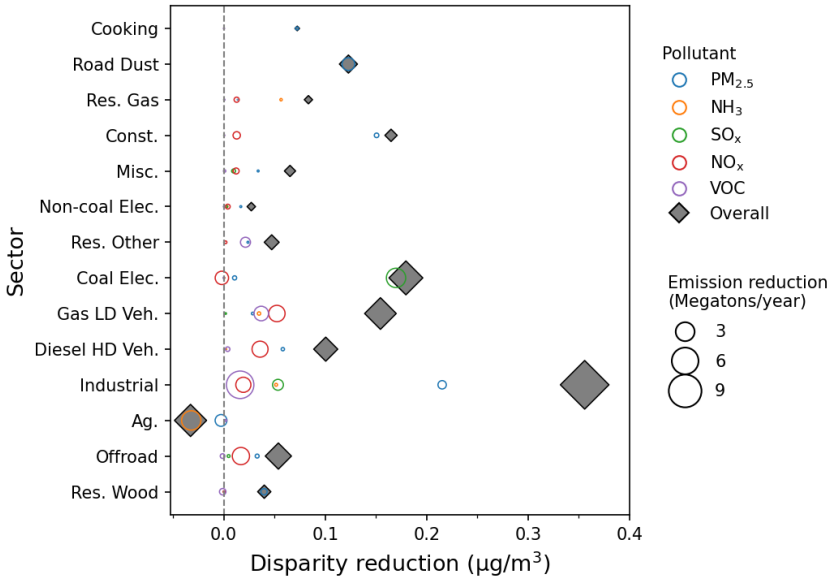

Figure 4.18: Disparity reduction for optimization by sector method. The icons sizes are proportionately to emissions.

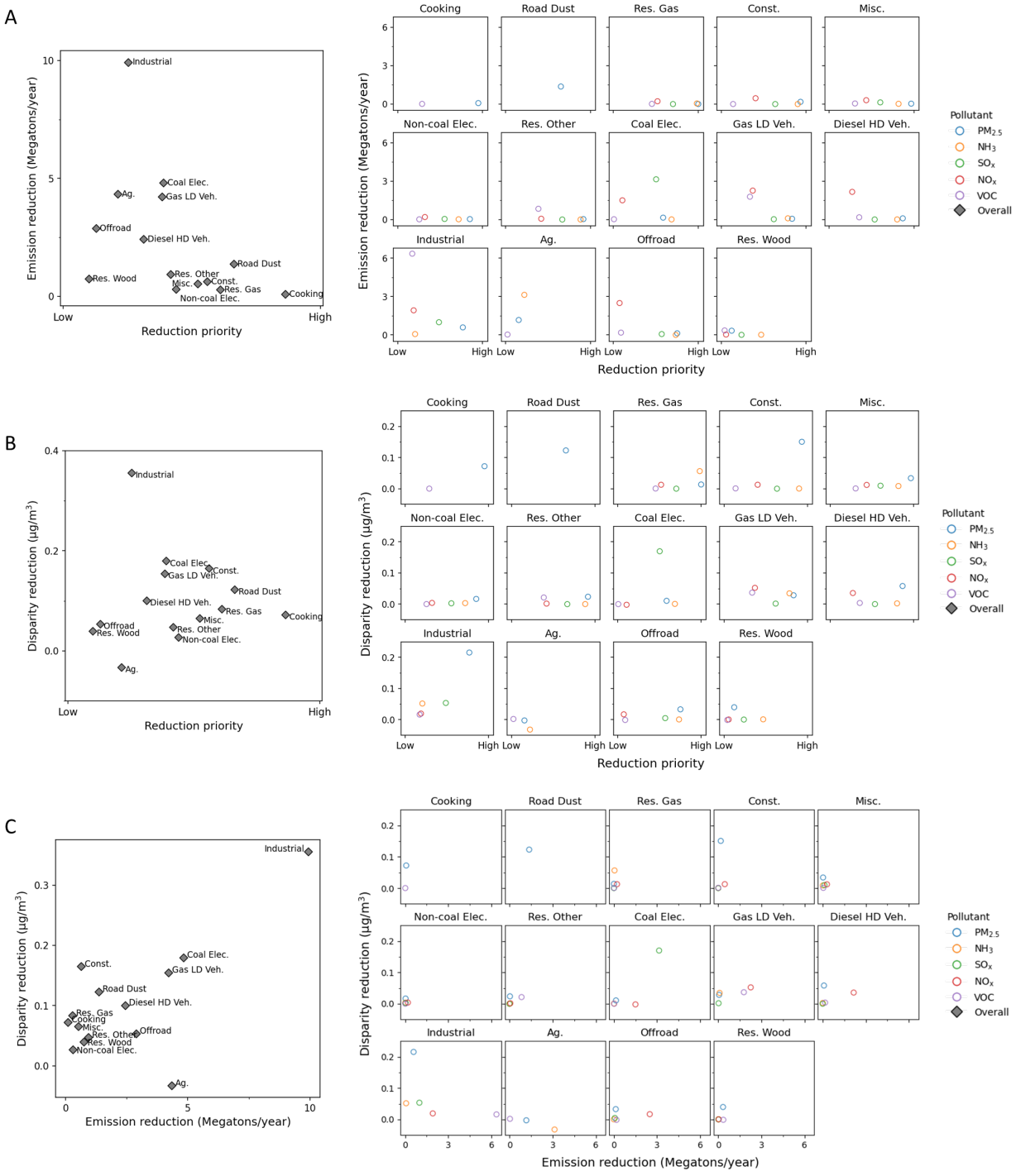

Figure 4.19: Emission reduction priority, emission reduced and disparity reduced for optimization by sector method. This figure* is an alternative version of Fig. 4.10b.

Table 4.1: Multiple linear regression results for reduction priority in relation to racial-ethnic composition, household income, population density, emission density, racial segregation index, and interaction term between racial-ethnic composition and racial segregation in the location (grid cell).

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12	Model 13
Percentage of non-White population		2.6 (2.3, 2.9)				2.9 (2.6, 3.2)				1.3 (0.6, 2.0)			
Percentage of Black population			4.2 (4.1, 4.3)				4.4 (4.3, 4.5)				6.9 (6.6, 7.2)		
Percentage of Hispanic population				-2.5 (-2.6, -2.3)				-1.4 (-1.6, -1.3)				3.5 (3.1, 3.9)	
Percentage of Asian population					-1.5 (-1.6, -1.4)				-1.0 (-1.1, -0.9)				0.5 (0.3, 0.7)
Median household income	-4.5 (-4.7, -4.3)	-3.9 (-4.1, -3.7)	-2.1 (-2.3, -1.9)	-5.2 (-5.4, -5.0)	-3.6 (-3.8, -3.4)	-2.3 (-2.5, -2.1)	-1.8 (-2.0, -1.7)	-3.8 (-4.0, -3.6)	-2.3 (-2.5, -2.1)	-2.4 (-2.6, -2.2)	-1.8 (-2.0, -1.6)	-3.5 (-3.7, -3.3)	-2.2 (-2.4, -1.9)
Log of population density	7.5 (7.2, 7.8)	7.0 (6.7, 7.3)	6.5 (6.2, 6.8)	8.1 (7.8, 8.4)	8.0 (7.7, 8.3)	8.7 (8.4, 9.0)	8.4 (8.1, 8.7)	10.1 (9.8, 10.4)	10.2 (9.9, 10.6)	8.6 (8.3, 8.9)	8.6 (8.3, 9.0)	10.3 (10.0, 10.6)	10.5 (10.1, 10.8)
Log of emission density	2.8 (2.4, 3.1)	2.3 (2.0, 2.6)	1.4 (1.2, 1.7)	3.2 (2.9, 3.5)	3.2 (2.9, 3.5)	2.9 (2.6, 3.2)	1.5 (1.2, 1.8)	3.6 (3.3, 3.9)	3.4 (3.1, 3.7)	2.8 (2.5, 3.2)	1.6 (1.3, 2.0)	3.8 (3.5, 4.1)	3.6 (3.2, 3.9)
Racial segregation index (non-White/White)						5.4 (5.2, 5.7)				6.5 (6.0, 6.9)			
Racial segregation index (Black/White)							3.6 (3.4, 3.9)				4.9 (4.6, 5.2)		
Racial segregation index (Hispanic/White)								5.6 (5.4, 5.9)				8.1 (7.7, 8.4)	
Racial segregation index (Asian/White)									5.3 (4.9, 5.7)				6.2 (5.8, 6.6)
Percentage of non-White population * Racial segregation index (non-White/White)										1.9 (1.2, 2.7)			
Percentage of Black population * Racial segregation index (Black/White)											-2.0 (-2.2, -1.7)		
Percentage of Hispanic population * Racial segregation index (Hispanic/White)												-4.0 (-4.3, -3.7)	
Percentage of Asian population * Racial segregation index (Asian/White)													-1.7 (-1.9, -1.5)
Number of observations	41,439	41,439	41,439	41,439	41,439	41,070	39,427	40,757	39,354	41,070	39,427	40,757	39,354
Adjusted R-square	0.21	0.21	0.32	0.22	0.22	0.24	0.32	0.25	0.22	0.24	0.32	0.26	0.22

Notes: The table reports the multiple linear regression results for reduction priority in relation to the location's characteristics. Model 1 only includes income, population density, and emission density (three "baseline" variables); Models 2-5 include all the three "baseline" variables as well as racial-ethnic composition; Models 6-9 include three "baseline" variables, racial-ethnic composition, and segregation index; Models 10-13 include three "baseline" variables, racial-ethnic composition, segregation index, and interaction of racial-ethnic composition and segregation index. Models 2, 6, and 10 are for the total non-White population; Models 3-5, 7-9, and 11-13 are for specified racial-ethnic groups (Black, Hispanic, or Asian). Reduction priority at each location is calculated by the emission weighted average priorities for the five pollutants, and is normalized to 0–100 scale (percentile; 100 represents highest priority, 0 represents lowest priority). All the predictor variables are normalized to the interquartile range (IQR), so the slopes represent the increase in disparity reduction priority (percentile) in relation to per IQR increase in each characteristics. For all the models, the p-values for all the slopes are less than 0.05.

4.6 Appendix I: Sensitivity analyses comparing three emissions-reduction approaches

In addition to the base case for the three approaches, I also perform extensive sensitivity analyses.

To compare the "location" and "sector" approaches, I (a) considered relative inequality rather than absolute inequality (Fig. 4.2); (b) used two alternative reduction scenarios (only reducing 50% or 90% of each emission location or sector, such that emissions are not completely eliminated; Fig. 4.3); (c) optimized to reduce inequality in each geographic region or in each Urban Area instead of nationally (Fig. 4.4); (d) reduced inequality for HV instead of racial-ethnic groups (Fig. 4.5); (e) considered urban disparities instead of national disparities (Fig. 4.6); (f) optimized to reduce average concentrations rather than concentration-disparities (Fig. 4.7); and (g) optimized location-sectors and sector-regions (i.e., modified the "sector" approach to make it more similar to the "location" approach; Fig. 4.8).

For sensitivity analysis [b] (i.e., only partial emission-reductions), I apply 90% (in another case, 50%) emission reductions for each location-pollutant or sector-pollutant parings. In these two scenarios, the emission reductions end once 10% (or, 50%) of the original total concentrations and concentration inequalities has been reached, respectively. If no further emission reduction could reduce the disparity (i.e., if the marginal concentration differences become negative for all the rest of parings), one could decide to halt the optimization at that point. (Further emission-reductions beyond that point will increase racial-ethnic disparities; the reason is that further emission-reductions reduce concentrations more for less-exposed groups than for more-exposed group.) However, to shed additional light on these scenarios, my optimization algorithm instead continues its simulation: it simulates the local optimum to keep the disparity low and non-negative (see green lines for 50% reduction scenario in Fig. 4.3: the line slopes upward after 5 MT/y emission reduction). Compared with the main approach, partial emission-reductions require more emission reduction amounts to reach the same disparity reductions.

For sensitivity analysis [c] (optimize regionally and locally instead of nationally), I reduce the emissions within each EPA region or Urban Area, separately. For the regional optimization, the median emission reductions to reduce 90% (from 1.5 to 0.15 $\mu g/m^3$) of the median regional disparity are 0.03 and 0.8 MT/y (a 27-fold difference) for "location" and "sector" approaches, respectively. For the local optimization, the median disparity is 0.4 $\mu g/m^3$; to reduce 90% of the

disparity, the required emission reductions for the two approached are 0.001 and 0.01 MT/y (a 11-fold difference), for "location" and "sector" approaches, respectively. Considering large, medium, and small urban areas (UAs) separately for the local optimization, the median disparities are 1.0, 0.6, and 0.2 $\mu g/m^3$ (relative disparities: 11%, 9%, and 3%), respectively. (Large/medium/small UAs are defined following [194], as population tertiles: n=10 large UAs, population >4m; n=44 medium UAs, population 728k – 4m; n=177 small UAs, population <728k; restricted to UAs with >20 ISRM grid cells.) To reduce 90% of the disparity, the required emission reductions for "location" and "sector" approaches are 0.003 and 0.06 MT/y (a 17-fold difference) for large urban areas, 0.002 and 0.03 MT/y (a 14-fold difference) for medium urban areas, and 0.0003 and 0.003 MT/y (a 12-fold difference) for small urban areas. Thus, my findings (the greater efficiency of the "location" approach relative to the "sector" approach) are consistent across large, medium, and small urban areas.

For sensitivity analysis [d] (social vulnerability), the population-average modeled $PM_{2.5}$ concentration is 7.9 $\mu g/m^3$ for HV locations and 6.9 $\mu g/m^3$ for non-HV locations (using the main definition for HV: 10% of CDC's SVI). The estimated disparity for HV is 0.9 $\mu g/m^3$ (13%) (i.e., the average $PM_{2.5}$ level for HV locations versus the population average). Using two alternative definitions for HV locations (20% of CDC's SVI; 10% $PM_{2.5}$ EJ index in EJScreen), the disparities for HV are 0.8 $\mu g/m^3$ (11%) and 1.8 $\mu g/m^3$ (25%), respectively. The overall reduction efficiency for HV is higher compared with the racial-ethnic disparities, using all three HV definitions. For all the three HV definition, "location" approach is much more efficient than "sector" approach in reducing the disparity for HV locations. Both approaches (i.e., "location" and "sector") reduce disparities by 90% or more, at less than 1.5 MT/y emission-reduction. In summary, the optimization can dramatically reduce disparities at far less emission-reduction for HV than by race-ethnicity; in all sensitivity analyses (as is also true for the main case), the "location" approach is more efficient than "sector" approach at reducing disparities.

For sensitivity analysis [e] (i.e., urban disparities), I explore the within-CBSA disparity changes for the baseline national optimizations. The results indicate that "location" approach is slightly more efficient than "sector" approach at most of the emission reduction levels, and the required emission reductions to reduce 50% (from 0.83 to 0.41 $\mu g/m^3$) of the within-urban disparities are 0.7 MT/y and 1.3 MT/y for "location" and "sector" approaches, separately.

For sensitivity analysis [f], I adjust the optimization metric in "location" and "sector" ap-

proaches to reduce average concentrations rather than concentration-disparities (Fig. 4.7). The goal of this alternative optimization changes from addressing/minimizing exposure disparity to maximize the health benefit for the total population. Results from [f] also indicate "location" approach is more efficient (here, at reducing population-average exposure concentrations) than "sector" approach.

For sensitivity analysis [g], I explore two alternative emission reduction steps: sector & region & pollutant combinations (the "sector and geographic region" approach) and sector & location combinations (the "sector and location" approach). The "sector and geographic region" approach adds EPA region of the emission sources as a further dimension in the reduction steps compared with the "sector" approach, and has 595 combinations (i.e., 14 sectors; 5 pollutants; 10 EPA regions; of the 700 maximum possible sector-pollutant-region pairings, 105 have zero emissions and so are not considered here as an opportunity for emission-reduction.) in total; the "sector and location" approach changes the pollutant type to source sector compared with the "location" approach, and has 509,128 combinations (i.e., 14 sectors; 52,411 locations; of the 733,754 maximum possible sector-location parings, 172,215 have zero emissions.) in total. Results from [g] (Fig. 4.8) indicate that adding additional dimension of geographic region improve the efficiency of "sector" approach; the efficiency of the approach combining sector and location (i.e., "sector" approach modified to be similar to "location" approach) is almost the same as "location" approach. This finding reflects that this sensitivity analysis modifies the "sector" approach by adding information on location (the EPA region), i.e., it is partially a "sector" - "location" hybrid method.

To examine the NAAQS-like scenario, I varied the NAAQS-like concentration standard (specifically 5, 6, 7, 8, 9, and 10 $\mu g/m^3$; Fig. 4.9) and, as a sensitivity analysis, considered urban and regional disparities rather than national disparities (Fig. 4.9, middle and bottom rows). None of the NAAQS-like scenarios explored eliminate the national, regional, and urban disparities.

4.7 Appendix II: Relationship between reduction priority for the "location" approach and grid cell characteristics

To determine the relationships between grid cell characteristics and emission reduction priority for the "location" approach, I employ both unadjusted (univariate) and adjusted (multivariate) analyses. The grid cell characteristics included in the analysis are racial-ethnic composition, median household income, population density, pollution emission density, and racial segregation index.

Racial-ethnic segregation is represented by dissimilarity index at each location (grid cell), which measures the percentage of the White population in an ISRM grid which would have to change census block to equalize the racial distribution between White and non-White (or a specified non-White, e.g., Hispanic) population groups across all blocks in the grid cell. The formula of segregation index is: $D_i = 0.5 \sum_{j=1}^{n} \left| \frac{w_{i,j}}{W_i} - \frac{n_{i,j}}{N_i} \right|$, where D_i is the dissimilarity index in ith ISRM grid; W_i is the total White population in ISRM grid i; N_i is the total non-White population in ISRM grid i; $w_{i,j}$ represents the White population in jth census block that within the boundary of ISRM grid i; $n_{i,j}$ represents the non-White population in jth census block that within the boundary of ISRM grid i.

For the unadjusted (univariate) analyses (Fig. 4.11), White percentage, Asian percentage, and median household income are negatively related with reduction priority and statistically significant ($p<0.05$). Black percentage, population density, emission density, and segregation index have positive relationships with reduction priority and statistically significant ($p<0.05$). Hispanic percentage is positively related with emission reduction priority, but the relationship is not statistically significant ($p = 0.32$). This result implies that, in general and averaged across the country, to optimally reduce disparities one would target emission-reductions in locations that have higher values for Black percentage, population density, emission density, and segregation. (An analogous result holds for Hispanic percentage, but the relationship is "noisier" (has more scatter).)

For the adjusted (multivariate) analyses, I employ four groups of multiple linear regression models (Table 4.1; 13 models in total). The first group (model 1) has three independent variables: income, population density, and emission density (the "baseline" variables). The second group (models 2-5) has the three "baseline" variables, plus racial-ethnic compositions. The third group (models 6-9) has the second-group variables, plus segregation indexes. The fourth group (models 10-13) has the third-group variables, plus an interaction term between racial-ethnic composition and segregation index. The second, third, and fourth groups each contain four regression models: one for the combined non-White population and one for each of the three specified groups (Black, Hispanic, Asian).

In all of the regression models (and, consistent with the univariate analyses), population density and emission density have positive slopes ($p<0.001$) and median household income has negative slopes ($p<0.001$).

The slopes (i.e., the beta coefficients in the regression models) of racial-ethnic composition and

segregation index have different patterns across racial-ethnic groups (non-White; Black; Hispanic; Asian). For the non-White group, both non-White percentage and segregation index have positive slopes in all the models (models 2, 6, and 10 for non-White percentage; models 6 and 10 for non-White/White segregation index). The interaction term between non-White percentage and segregation index (model 10) has a slight positive value, which indicates that with an increase of segregation level, the positive slope for non-White percentage becomes slightly steeper. The patterns (models 3, 7, and 11) for Black population are generally the same as combined non-White group. The only difference is that the interaction term of Black percentage and segregation index is negative (model 11), which indicates that with an increase of segregation level, the positive slope for Black percentage become less steep. Regression models for Hispanic (models 4, 8, and 12) and Asian (models 5, 9, and 13) groups have similar patterns, which are different from the non-White group and the Black group. In the regression models without interactions (models 4, 5, 8, and 9), Hispanic & Asian percentages have negative coefficients; segregation indexes have positive coefficients. However, in the models with interaction terms (models 12 and 13), the slopes of Hispanic & Asian percentages become positive, and the interaction terms are negative. The results for Hispanic and Asian population indicate that at a zero segregation levels, Hispanic and Asian percentages both have positive slopes. With an increase of segregation levels, the positive slopes become less steep, which eventually flip to be negative; at the average segregation levels, the slopes for Hispanic and Asian percentages are both negative (model 8 and 9). The p values for all the slopes in all 13 models are less than 0.001.

Chapter 5

HOW WOULD EMISSION-REDUCTIONS IN JUSTICE40 LOCATIONS REDUCE RACIAL-ETHNIC DISPARITIES IN AIR POLLUTION EXPOSURE?

5.1 Summary

During the past decades, air quality in the US has improved dramatically, however, disparities persist. Exposure disparities by race-ethnicity are larger than, and independent of, disparities by income or other socioeconomic measures. The Biden Administration's Climate and Economic Justice Screening Tool (CEJST) is the signature element and the ongoing approach of the White House Justice40 Initiative ("J40"), which is a general policy to address environmental injustice. To explore whether emission-reductions brought about by CEJST/J40 investments will eliminate disparities in $PM_{2.5}$ exposure by race-ethnicity and other attributes. I compare a Business As Usual ("BAU") scenario against two scenarios wherein CEJST-identified locations ("J40 communities") experience accelerated emission-reductions. BAU simply continues historical rates of emissions and emission-changes into the future; in the two CEJST scenarios, I double or quadruple emission reduction in J40 communities, relative to BAU. I use InMAP to predict how changes in emissions would impact annual-average $PM_{2.5}$ concentrations and concentration-disparities. Under the BAU scenario, disparities remain in place. Racial-ethnic disparities remain larger than disparities for J40 communities and for low-income populations. The CEJST scenarios eliminate absolute and relative disparities for J40 communities and for low-income populations in 20 years; however, they do not reduce relative disparities by race-ethnicity. Results from my analysis indicate that additional and more targeted actions, beyond CEJST/J40, will be needed to end racial-ethnic exposure disparities in the next decades. In the bigger picture, This chapter argues that Regulatory Impact Analysis (RIA) for air pollution must quantify whether and how policies/tools will not only affect air quality overall, but also reduce absolute and relative disparities, and that regulatory approaches are needed that are specifically designed to remove disparities.

5.2 Introduction

Air pollution is one of the most significant environmental risk factors in the U.S., causing an estimated ∼100,000 annual premature deaths, which corresponds to billions of dollars of health damage per day. While there have been substantial improvements in air quality in recent decades, disparities in air pollution exposure have been remarkably persistent [10, 28, 32, 243].

The largest air pollution disparities are by race-ethnicity, which represent a major environmental injustice. Disparities by other attributes (e.g., income, age, education) are relevant but are much smaller than (and statistically distinct from) disparities by race-ethnicity [10, 28]. Disparities by race exist in every U.S. state, nearly all emission categories, and across multiple decades [10, 25, 26]. The underlying cause is racist policy, planning, and actions [35, 37, 38, 45].

What viable approaches could continue historical improvements in air quality and also eliminate racial-ethnic and other exposure disparities? Here, I investigate potential impacts on air pollution and air pollution disparities of the Climate and Economic Justice Screening Tool (CEJST), which is being used to guide agency decision-making for the J40, a cornerstone of the Biden Administration's effort to address environmental injustice. The goal of J40 is "that 40 percent of the overall benefits of certain Federal investments flow to disadvantaged communities that are marginalized, underserved, and overburdened by pollution" (`www.whitehouse.gov/environmentaljustice/justice40`). J40 communities are identified by CEJST, and although billions of dollars are being allocated based on CEJST, its potential impact on environmental disparities remains largely unstudied.

My analyses indicate that CEJST will only modestly reduce air quality disparities, and in some cases will increase them. This highlights the broader problem of insufficient investigation of how existing or proposed policies will affect environmental justice outcomes. Addressing inequalities in exposure to air pollution will require – as a necessary but insufficient step – measuring and reporting exposure disparities, and explicitly testing how proposed regulatory actions will affect disparities.

To illustrate the type of analysis needed to evaluate the effect of policies on exposure disparities, I investigate here how application of CEJST to guide targeted emission-reductions for $PM_{2.5}$ air pollution could affect exposure disparities. Specifically, I predict annual-average $PM_{2.5}$ concentrations throughout the contiguous US based on the emissions of each chemical species of $PM_{2.5}$ and its precursors from each sector of the economy. I consider three future emission scenarios: (1)

BAU without J40, and scenarios where J40 leads to, in the J40 communities, a (2) doubling or (3) quadrupling of emission-reduction rates, relative to BAU. BAU simply continues historical rates of emissions and emission-changes into the future; it represents the expected future effects of historical air pollution regulatory approaches. For all three scenarios, non-J40 communities remain at BAU. J40 communities are Census Tracts that CEJST identifies as facing multiple environmental hazards and social vulnerabilities (see Materials and Methods for details). The "doubling" and "quadrupling" scenarios represent examples of aggressive or very aggressive emission-reductions in J40 communities.

I analyzed the impact of these scenarios on human exposure using a reduced-complexity chemical transport model called InMAP that simulates the fate and transport of anthropogenic emissions to predict how changes in emissions would alter $PM_{2.5}$ concentrations and concentration-disparities. $PM_{2.5}$ – which causes deaths by increasing rates of heart attack, stroke, lung cancer, respiratory infections, and more – is responsible for most of the deaths and monetized health impacts of ambient air pollution. I consider disparities for four groups: (i) people living in J40 communities; (ii) low-income people (i.e., people in households with incomes at or below two times the poverty level); (iii) people of color (i.e., all people except non-Hispanic (NH) Whites); and, (iv) the most-exposed racial-ethnic group of the four groups considered (NH White, NH Black, NH Asian, and Hispanic). InMAP simulates both primary (i.e., directly-emitted) and secondary (i.e., forms in the atmosphere from chemical reactions) $PM_{2.5}$, provides national coverage at high spatial resolution (as small as 1 km in urban centers), and is computationally efficient. I also estimate health risks from $PM_{2.5}$ (see Materials and Methods for details).

5.3 Materials and Methods

5.3.1 Source-receptor matrix

I use a publicly available source-receptor matrix (ISRM) that is derived from the Intervention Model for Air Pollution (InMAP) to estimate $PM_{2.5}$ concentrations for the contiguous U.S. from all anthropogenic emission sources [109, 121]. InMAP is a reduced complexity model that tracks the transport of emissions and their physical and chemical processes. ISRM relates changes in emissions at any one location to concentration changes at all other locations. The ISRM has separate layers for the 5 pollutants (primary $PM_{2.5}$ and four precursors of secondary $PM_{2.5}$: nitrogen oxides [NO_x],

sulfur oxides [SO_x], ammonia [NH_3], and volatile organic compounds [VOCs]) and 3 stack heights (ground level: <57 m; low stack height: 57 to 379 m; high stack height: >379 m). Each layer of ISRM consists of 52,411 grid cells for the contiguous us, with a variable spatial scale ranging from 1 km × 1 km in densely-populated urban centers to 48 km × 48 km in sparsely-populated rural areas. The population-weighted average size of a grid cell is 2.6 km in Urban Areas and 22.6 km in non-Urban Areas.

5.3.2 Emission data and historical emissions reduction rates

I use 2014 US EPA National Emission Inventory (NEI) v1 as the baseline emission data, and consider all the available NEIs from 1990 to 2014 (a total of 12 NEIs; `www.epa.gov/air-emissions-inventories/national-emissions-inventory-nei`) to calculate the historical emissions reduction rates. All the NEIs are grouped into 14 source sectors and 5 pollutants, and the 2014 NEI is further allocated to the individual ISRM grid cell and each stack height layer. The 14 source sectors are: 1) agriculture ("Ag"); 2) coal electricity utility ("Coal Elec."); 3) noncoal electricity utility ("Non-coal Elec."); 4) commercial cooking ("Cooking"); 5) construction ("Const."); 6) diesel heavy-duty vehicle ("Diesel HD Veh."); 7) gasoline light-duty vehicle ("Gas LD Veh."); 8) industrial; 9) road dust ("Road Dst"); 10) residential gas combustion ("Res. Gas"); 11) residential wood combustion ("Res. Wood"); 12) residential others ("Res. Other"); 13) off-highway vehicle and equipment ("Offroad"); and 14) miscellaneous ("Misc."). These sectors are aggregated from Environmental Protection Agency "Source Classification Codes" (SCCs) except biogenic, wildfire, and international sources [244].

The historical emission-reduction rates are calculated by sectors and pollutants (61 combinations in total [i.e., 14 sectors, 5 pollutants; of the $14 \times 5 = 70$ maximum possible sector–pollutant pairings, 9 have zero emissions and are excluded here]). For each sector and pollutant, the annual-average emissions change rate is calculated as the slope of the best-fit line for the linear regression of emission values versus years. In the main chapter, I run the regressions for all available NEIs from 1990 to 2014; as a sensitivity test, I also run the regressions for the recent 3 NEIs (2008, 2011, 2014) to estimate emission trends in the recent $\sim$ decade. The zero and missing values are excluded from the regression models. Most of the 61 sector-pollutant combinations are decreasing over time (i.e., a positive emissions reduction rate); only a small number are the reverse (positive slope; negative emissions-reduction rate).

5.3.3 Calculating national exposure concentrations and disparities by J40, race-ethnicity, and income

I calculate the $PM_{2.5}$ concentrations from all anthropogenic emissions for all ISRM grid cells. The mean $PM_{2.5}$ exposure for the overall population and by demographics (J40, race-ethnicity, and income) are calculated as population-weighted averages and across all the grids. The approach to population-weighted averages is the conventional approach: I first calculate the sum of the multiplied values of concentration and population count (for overall population and for each subgroup) for each grid cell, and then divide that value by the sum of the population count.

J40 locations are defined using the Climate and Economic Justice Screening Tool (CEJST) from the Council on Environmental Quality [245]. CEJST highlights disadvantaged census tracts across the U.S. as those which are overburdened and underserved; per the J40, these communities should receive 40% of the overall benefits of programs included in the Biden–Harris Administration's J40. The census tracts are identified as disadvantaged by CEJST, v1.0, if they (1) meet the threshold for at least one of the tool's categories of burden (8 categories in total: climate change, energy, health, housing, legacy pollution, transportation, water and wastewater, and workforce development) and corresponding economic indicator (income or education), or (2) are on the lands of a Federally Recognized Tribe. The disadvantaged communities defined by CEJST v1.0 includes 37% of the 2010 census tracts (excluding zero-population tracts) and 34% of the U.S. population. Here, I label the grid cells that overlap with the J40 communities (criteria: $> 50\%$ of the population or land area in the grid cell is within the J40 census tracts) as the J40 communities (34% of the total population), and the remaining grid cells (76% of the total population) as non-J40 communities.

Race-ethnicity data are from the 2010 Census, at the block level. I merge the block level race-ethnicity to the grids through intersection of block centroids and ISRM grid cells. I employ five racial-ethnic groups: non-Hispanic White (63% of the total population; hereafter, "White"); Latino or Hispanic of any race (17%; hereafter, "Hispanic"); non-Hispanic Black or African American (12%; hereafter, "Black"); non-Hispanic Asian and Pacific Islander (5%; hereafter, "Asian"); and POC (37%), which are defined as everyone except the NH White population (i.e., "POC" includes Hispanic, Black, Asian, American Indian or Alaska Native, some other races, or mixed races).

Income data are from the EPA's EJScreen (Environmental Justice Screening and Mapping Tool; `www.epa.gov/ejscreen/overview-socioeconomic-indicators-ejscreen`) at census

block group level. I merge the EJScreen data with ISRM grid cells, and then define the "low-income" population by multiplying the low-income percentages (i.e., percentages of population with household income less than or equal to twice the federal poverty level) with population counts in the grids (32% of the total population); the remained of the population is defined as "non-low-income" (68%).

$PM_{2.5}$ exposure inequalities are calculated for 4 subpopulations: 1) J40 communities; 2) POC populations; 3) most-exposed race-ethnicity (in the baseline year, the most-exposed race-ethnicity is Black; this changes during the simulation, as the emissions decrease); and 4) low-income populations. For each subpopulation, the exposure inequality is calculated by the mean exposure in that subgroup minus the mean exposure for the overall population. I report exposure disparity in both absolute ($\mu g/m^3$) and relative (i.e., percent) terms.

5.3.4 Calculating health impact from $PM_{2.5}$ exposure

To calculate the health impact from $PM_{2.5}$ exposure, I assume the all-cause mortality Hazard Ratio (HR) per 10 $\mu g/m^3$ exposure to $PM_{2.5}$ is 1.08 [246] The HRs at different $PM_{2.5}$ levels are estimated using the log-linear model: $HR = \exp(\beta \times C)$, where C is the population-weighted average $PM_{2.5}$ concentration, β is the linear coefficient $= \ln(1.08)/10 = 0.0077$ $m^3/\mu g$. The national crude mortality rate is 724.6 deaths per 100,000 standard US population in 2014 [247]. The attributable risk is calculated as HR minus 1.0. For example, the mortality risk at a $PM_{2.5}$ concentration of 10 $\mu g/m^3$ would be calculated as:

$$
\begin{aligned}
\text{mortality risk} &= \text{crude mortality rate} \times [exp(\beta \times C) - 1] \\
&= 724.6 \text{ deaths per 100,000 people per year} \times [\exp(0.0077 m^3/\mu g \times 10 \mu g/m^3) - 1] \\
&= 724.6 \times [\exp(0.077) - 1] \text{ deaths per 100,000 people per year} \\
&= 724.6 \times (1.08 - 1) \text{ deaths per 100,000 people per year} \\
&= 724.6 \times 0.08 \text{ deaths per 100,000 people per year} \\
&= 58 \text{ deaths per 100,000 people per year}
\end{aligned}
$$

The health impact per change in $PM_{2.5}$ concentration deviates slightly from strictly linear. For example, the health-risk reduction attributable to a 1-$\mu g/m^3$ reduction in $PM_{2.5}$ concentration is, if starting at a concentration of 3 $\mu g/m^3$ (i.e., the benefit from the concentration dropping from 3 to 2 $\mu g/m^3$), 5.47 fewer deaths per 100,000 people per year; that value is 5% larger than if the

same 1-μg/m^3 reduction occurs starting at a concentrations of 9 μg/m^3 (i.e., the benefit from the concentration dropping from 9 to 8 μg/m^3 is 5.22 fewer deaths per 100,000 per year). In Fig. 5.1 in the main article, the right-axis scale reflects that fact, i.e., the left-axis (the concentration) is a linear scale, but tick-marks on the right-axis (health risk) deviate slightly from linear. In Fig. 5.2 (upper row; right-axis), calculated health risks are based on the population average concentration in year 20. (If the right-axis in Fig. 5.2 were instead based on current-year concentrations, the calculated risks would be slightly [$\sim$3%] larger.)

5.3.5 Emissions-reduction scenarios

I consider three emissions-reduction scenarios, each of which consider the next 20 years (baseline emission year: 2014): the "business as usual (BAU)" scenario and two scenarios where emission-reductions in J40 communities are double or quadruple the BAU rate. In the BAU scenario, each sector and pollutant follow the historical yearly emission reduction rate (see Figs. 5.3 to 5.9) until the emissions in that sector and pollutant have been reduced to zero. In the "doubling" scenario, the emission reduction rates in the J40 communities are doubled (if the historical emissions-reduction rates are positive [i.e., the emissions have decreasing trends]) or are set to zero (if the historical emissions-reduction rates are negative [i.e., the emissions have increasing trends]); the emissions-reduction rates keep the same as BAU scenarios in non-J40 communities. The "quadrupling" scenario is directly analogous to the "doubling" scenario but with quadrupling instead of doubling.

5.4 Results

In the current state (i.e., Year 0, prior to applying any new emissions reductions), InMAP predicts that people of color in the U.S. are exposed to $PM_{2.5}$ concentrations that are $\sim$14% higher-than-average. Disparities by race-ethnicity are much larger than disparities by J40 or low-income -status [10]. Among people-of-color groups, the Black population is most-exposed to $PM_{2.5}$ ($\sim$20% higher-than-average). As emission-reductions occur, concentrations for Black populations decrease, and Asian people soon become most-exposed; however, concentrations remain higher-than-average for Black, Hispanic, and Asian populations.

As expected, for all three scenarios, all demographic groups experience cleaner air in the future. However, two key findings emerge with respect to exposure disparities.

First, under BAU, disparities remain in place. As emission-reductions occur, $PM_{2.5}$ concentra-

tions decrease at slightly different rates for different groups (for example, InMap predicts that Asian people will soon become the most exposed group). However, concentrations remain higher-than-average for Black, Hispanic, and Asian populations (Fig. 5.1). In addition, racial-ethnic disparities persist and remain much larger than disparities between J40 and non-J40 communities, and between low-income and non-low-income households. This finding underscores that new emission reduction / regulatory strategies, deviating from historical approaches (i.e., BAU), are needed to reduce and eliminate observed exposure disparities.

Second, the two scenarios with enhanced emission-reductions in J40 communities eliminate absolute and relative disparities for J40 communities and for low-income populations, yet they do not reduce relative disparities by race-ethnicity (although they do decrease absolute disparities) (Fig. 5.2, Figs. 5.11 and 5.12). In fact, both J40 scenarios end up increasing the relative exposure disparity for the most exposed racial-ethnic group (Fig. 5.2g), relative to present-day and the BAU future. This outcome could be interpreted as undermining the goal of eliminating exposure disparities by race.

My findings are robust to several sensitivity analyses, including considering alternative methods and outcomes (see Appendix I). Results from the sensitivity analyses indicate that only with enhanced emission-reduction in communities of color will both absolute and relative racial-ethnic disparities in $PM_{2.5}$ air pollution be reduced.

5.5 Discussions

This study quantifies, for the first time, the regulatory impacts of an ongoing and signature national-wide environmental justice policy of the Biden Administration (J40/CEJST). My analysis highlights the importance and necessity of Regulatory Impact Assessments (RIA), that quantify the impacts on overburdened/disadvantaged communities, for all environmental justice policies (including J40/CEJST); this study also provide an example of conducting such RIA for $PM_{2.5}$, using new tools. The model employed has good agreement with a comprehensive model (WRF-Chem) (e.g., for predictions of total $PM_{2.5}$ changes, the population-weighted mean fractional bias = -17%; the population-weighted R^2 = 0.90) [109] and high spatial resolution (average spatial resolution: 2.6 km in Urban Areas; 22.6 km in non-Urban Areas [13.2 km overall]). Further work is needed to incorporate cumulative burdens (i.e., not just $PM_{2.5}$) and equity aspects of health risk assessment, such as differences in susceptibility, baseline disease rates, and access to healthcare;

to test and improve high-resolution reduced-complexity models for all criteria pollutants and for air toxics; and to identify the most effective policies and strategies for achieving location-specific emission-reductions (e.g., low-emission zone, enhanced enforcement, modifications to the permitting process, financial incentives, or other, potentially novel, approaches). Impacts for other pollutants, such as NO_2, ultrafine particles, and air toxics, may (relative to $PM_{2.5}$) be more localized; have greater disparities, including disproportionately affecting communities of color; and be more of concern to people living near those emission sources.

Failure of J40/CEJST to reduce the largest source of air pollution disparities (i.e., those by race-ethnicity), would significantly undermine the Biden Administration's environmental justice goals.

Why might CEJST-driven actions be ineffective in addressing air pollution disparities? There are many factors, including the paucity of efforts to test the potential impacts of proposed strategies and tools on air pollution disparities, and whether, overall, racial-ethnic disparities in exposure can be eliminated in a reasonable time-frame (e.g., in a decade or less). My analysis highlights that this sort of testing and evaluation is essential and that new tools such as InMAP enable such analyses [25, 248, 249].

Another contributing factor regarding CEJST-driven actions is, compared with the national average, J40 communities are composed of only modestly higher proportions of people of color (especially Black, Hispanic, and Indigenous populations) and low-income populations (Fig. 5.10). This reflects political decisions and, in part, concern regarding potential legal challenges if a policy or tool explicitly includes race as a factor in allocating federal funds. (J40/CEJST do not include race as a factor.)

Present-day air pollution disparities reflect decades of racist policy and practice [35, 38, 42]. Since race-based actions created this problem, solving it will likely be more difficult if the government does not consider, or bars itself from considering, information about the racial makeup of communities as part of its decision-making and action.

Regulatory approaches are needed that are specifically designed to remove race-based and socioeconomic disparities. For example, RIA for air pollution should quantify whether and how relevant policies will not only affect air quality overall, but also reduce absolute and relative exposure disparities. Regulatory information about air pollution – for example, EPA "Status and Trends" reports – should quantify disparities or exposures for overburdened communities. While politically

challenging, emission reduction efforts must also address disparities by race-ethnicity if we wish to uphold everyone's right to breathe clean air.

Overall, J40 aims to address many issues, not just $PM_{2.5}$. There will be myriad benefits from J40, including via investments in communities and building resilience to climate change. At the same time, J40 communities, as defined by CEJST, comprise ∼34% of the U.S. population. The goal of delivering 40% of benefits to 34% of the population represents a modest deviation from an exactly proportional share of the benefits; problems identified above lie not with the J40 goal, but with the CEJST tool shaping how J40 is carried out. Results from my analysis indicate that additional and more targeted actions, beyond J40 and the current CEJST, will be needed to end racial-ethnic exposure disparities in air quality. Other policies, including at the state level (e.g., in California, New Jersey, and Washington), also aim to address environmental disparities. However, the effectiveness of these other policies at reducing disparities should also be evaluated as we have done here for J40 and CEJST.

5.6 Conclusions

Conclusions from this work include the following. First, the US Environmental Protection Agency and other entities can and should quantify how proposed programs, regulations and decision-making tools would impact environmental justice outcomes, especially air pollution exposure disparities by race-ethnicity. If possible, this analysis should be undertaken while such initiatives and tools are being developed. This type of national analysis previously would have been difficult or impossible to do, because of computation-cost and spatial-resolution limitations of most air quality models; however, recently-developed air quality models facilitate this type of analysis by being faster and easier to carry out, and by providing national coverage at much higher spatial resolution than many conventional models.

Second, my prior work indicates that there are solutions available that will eliminate disparities of concern, within a reasonable time frame [63]. More work is needed to understand these approaches and then to design and implement regulations to achieve more equitable outcomes.

Finally, current approaches in the Clean Air Act have been effective at reducing average concentrations and absolute exposure disparities [30], but relative disparities are generally ignored and have persisted [10, 28, 32, 63, 243]. If, as my results suggest, neither BAU nor CEJST can eliminate the racial disparities disparities in environmental hazard exposures, then other regulatory strategies

and revised tools are needed.

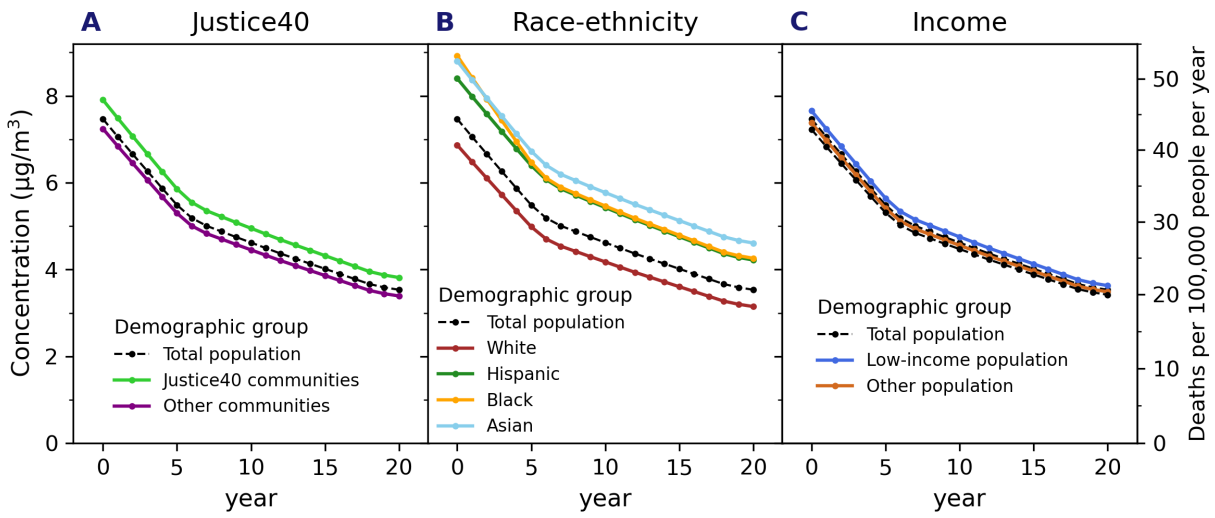

Figure 5.1: Predicted population-weighted average $PM_{2.5}$ concentration (left y-axis) and attributable health risks (right y-axis) for the "business as usual" (BAU) scenario, disaggregated by (A) J40 community status, (B) racial-ethnic group, and (C) low-income status. In each panel, the overall population average is the black dashed line. Although all groups are predicted to experience cleaner air, inequalities are predicted to persist (i.e., in each panel, lines are approximately parallel). Disparities in $PM_{2.5}$ concentrations (in each panel, the spread among the lines) are larger by race-ethnicity than by J40 community status or low-income status; average $PM_{2.5}$ concentrations are larger for people of color than for J40 communities and for low-income people. The three racial groups (White, Black, Asian) refer to non-Hispanic ethnicity.

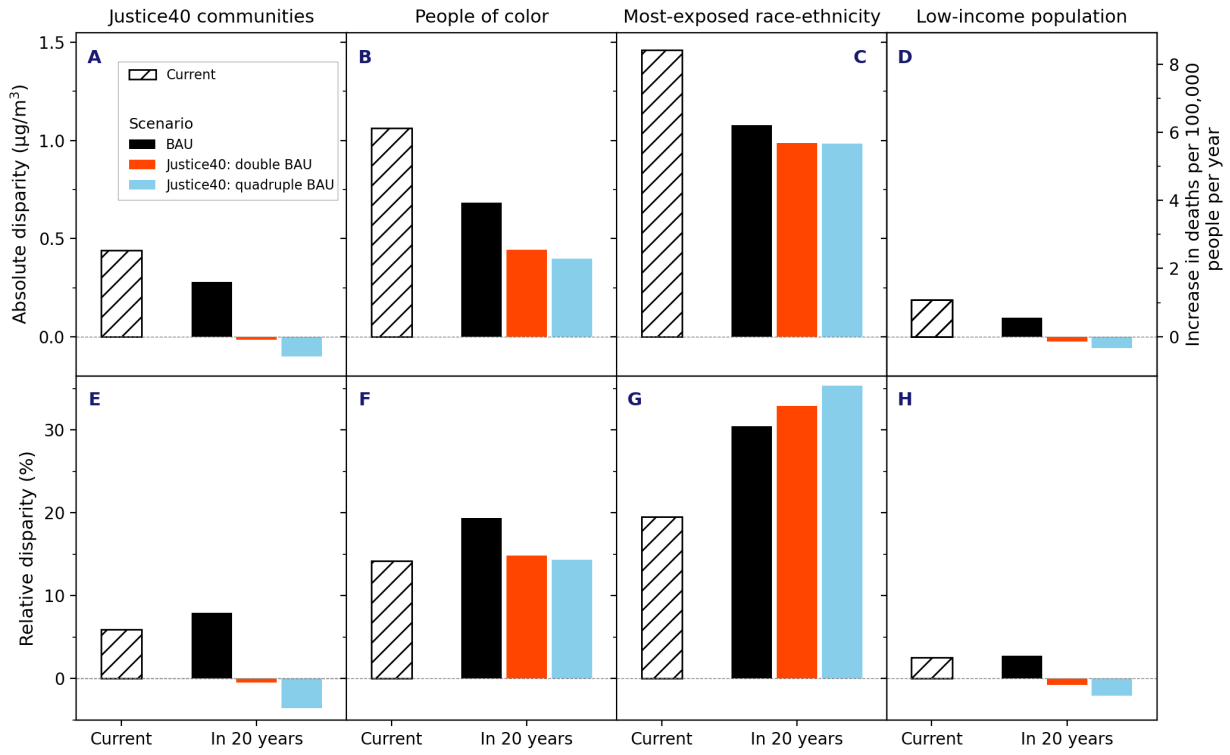

Figure 5.2: $PM_{2.5}$ and health risk disparities for the three future scenarios: BAU (business as usual; continuing historical emission-reduction trends without J40), and cases where emission-reductions in J40 communities are double or quadruple the BAU rate. $PM_{2.5}$ disparities (left y-axis values) are relative to the population-average $PM_{2.5}$ concentration, in absolute terms (top row; A to D) or in relative terms (bottom row; E to H). The columns represent population groups: (A and E) J40 communities; (B and F) people of color (i.e., all racial-ethnic groups except non-Hispanic White); (C and G) most-exposed racial-ethnic group (Black people currently; Asian people in Year 20); (D and H) low-income population. Zero disparity (y=0, horizontal line) would mean that the average exposure for that group is equal to the population average; negative y-axis values indicate that average concentrations are below the population average. The BAU scenario (black bar) does not bring disparities to zero. Both J40 scenarios (red and blue bars) eliminate or reverse disparities (i.e., leads to disparities at or below zero) for the left and right columns but not for the middle two columns. That result suggests that aggressive emission-reductions in J40 communities may eliminate higher-than-average air pollution for J40 communities (left column) and for low-income populations (right panel) but not for race-ethnicity (middle two columns).

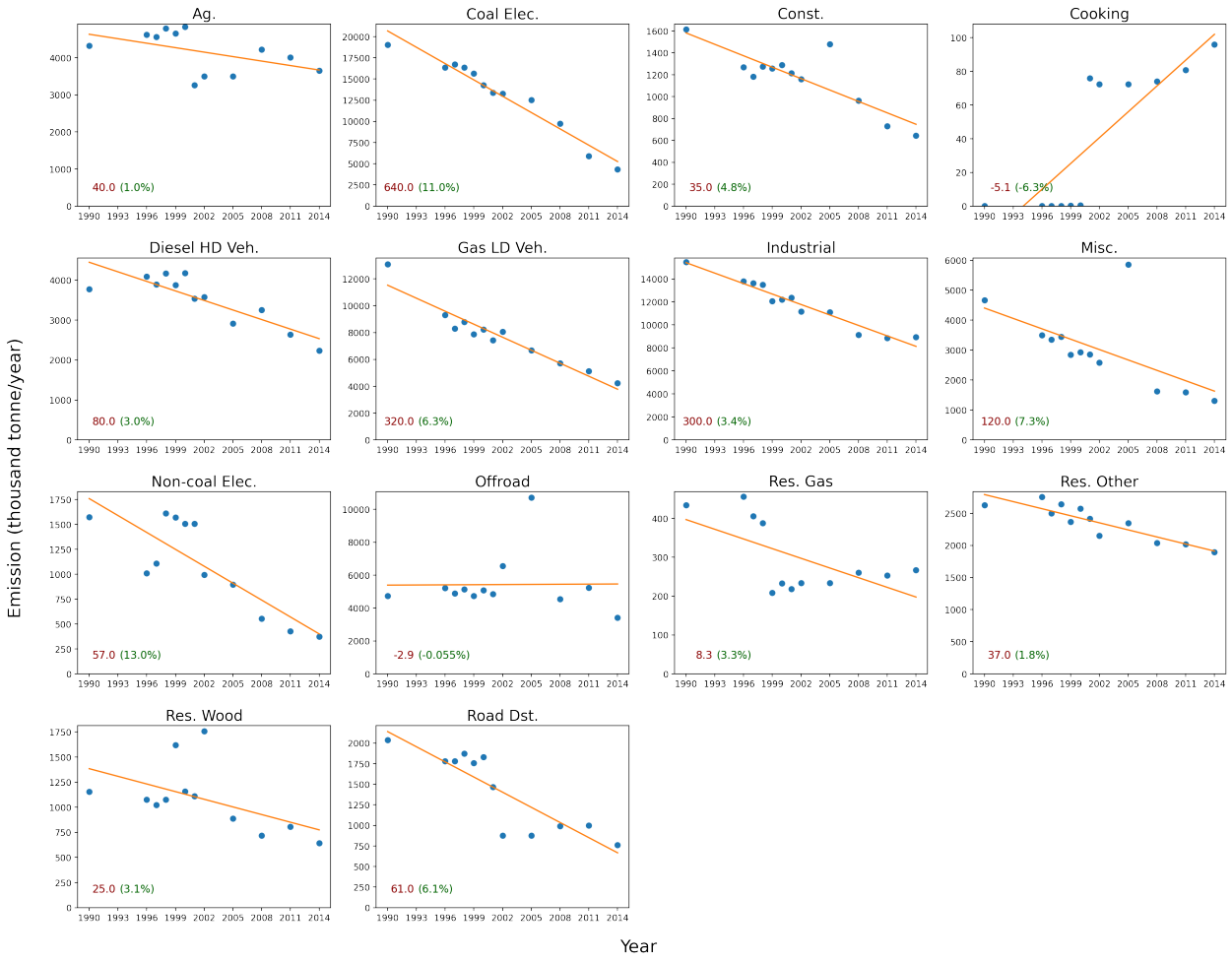

Figure 5.3: Historical emission reduction from 1990-2014 for each sector of economy. The text on the lower left corner of each subplot shows the average emission reduction values in absolute (thousand tonne/year; red text) and relative (relative to the year-2014 emissions; green text in parentheses).

Figure 5.4: Historical emission reduction for primary $PM_{2.5}$ from 1990-2014 for each sector of economy. The text on the lower left corner of each subplot shows the average emission reduction values in absolute (thousand tonne/year; red text) and relative (relative to the year-2014 emissions; green text in parentheses).

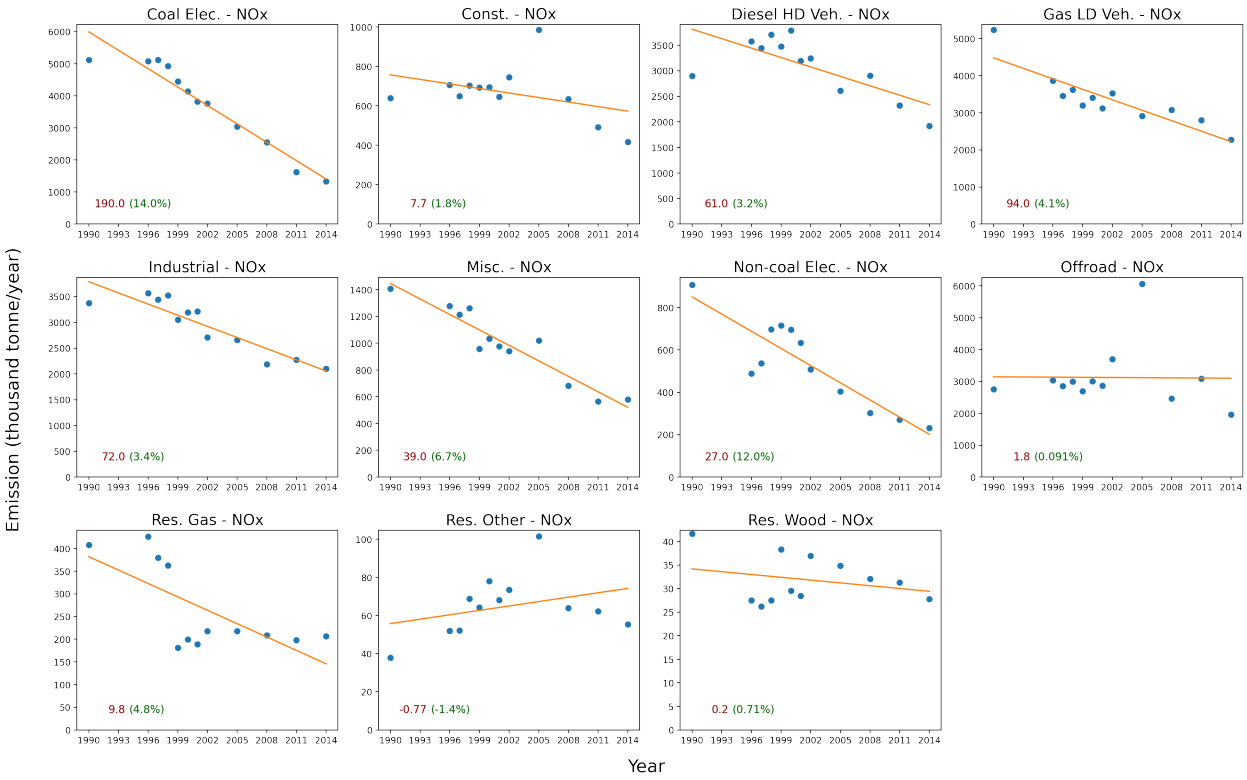

Figure 5.5: Historical emission reduction for NO_x from 1990-2014 for each sector of the economy. The text on the lower left corner of each subplot shows the average emission reduction values in absolute (thousand tonne/year; red text) and relative (relative to the year-2014 emissions; green text in parentheses in parentheses).

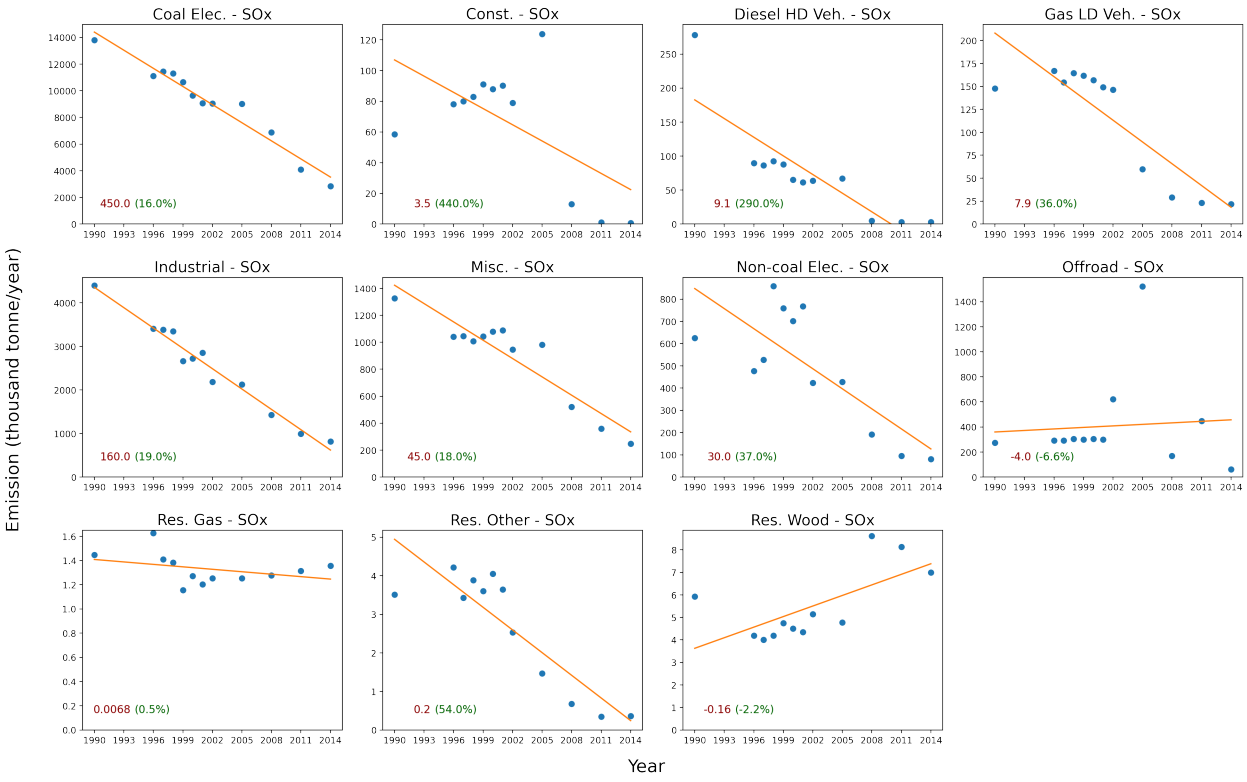

Figure 5.6: Historical emission reduction for SO_x from 1990-2014 for each sector of the economy. The text on the lower left corner of each subplot shows the average emission reduction values in absolute (thousand tonne/year; red text) and relative (relative to the year-2014 emissions; green text in parentheses).

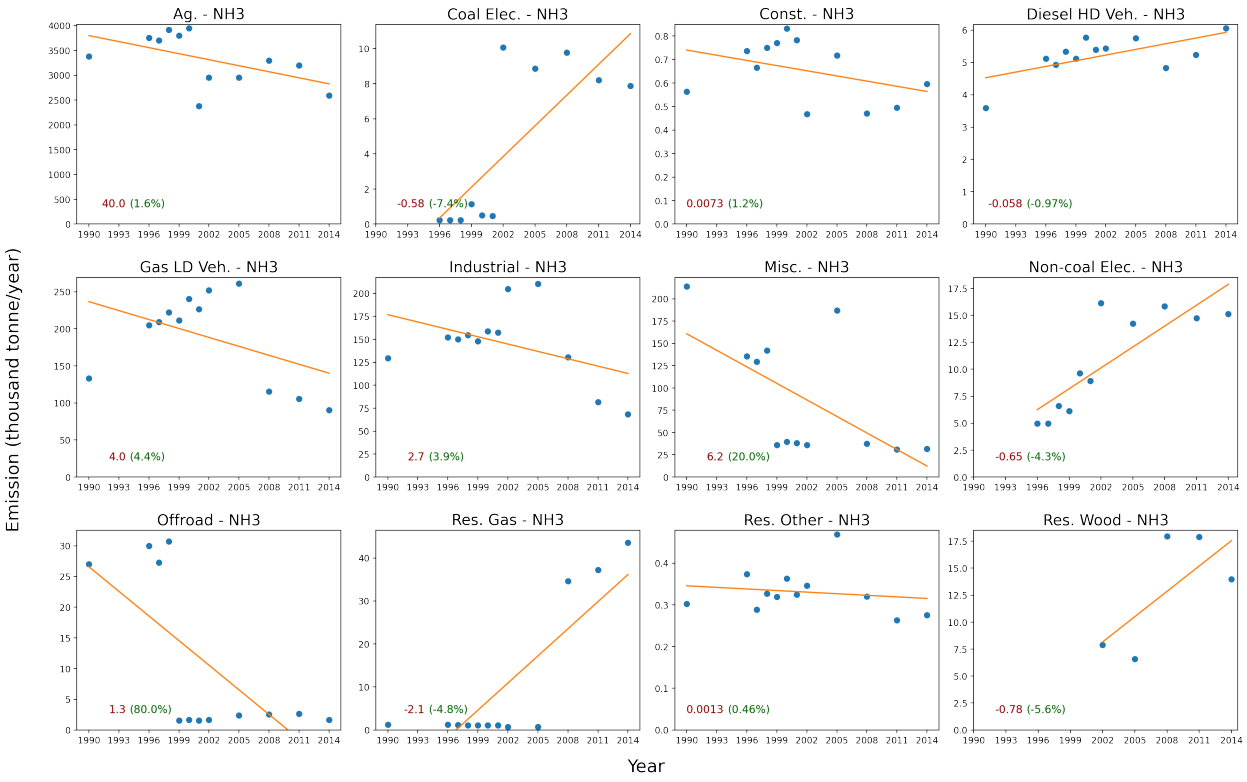

Figure 5.7: Historical emission reduction for NH_3 from 1990-2014 for each sector of the economy. The text on the lower left corner of each subplot shows the average emission reduction values in absolute (thousand tonne/year; red text) and relative (relative to the year-2014 emissions; green text in parentheses).

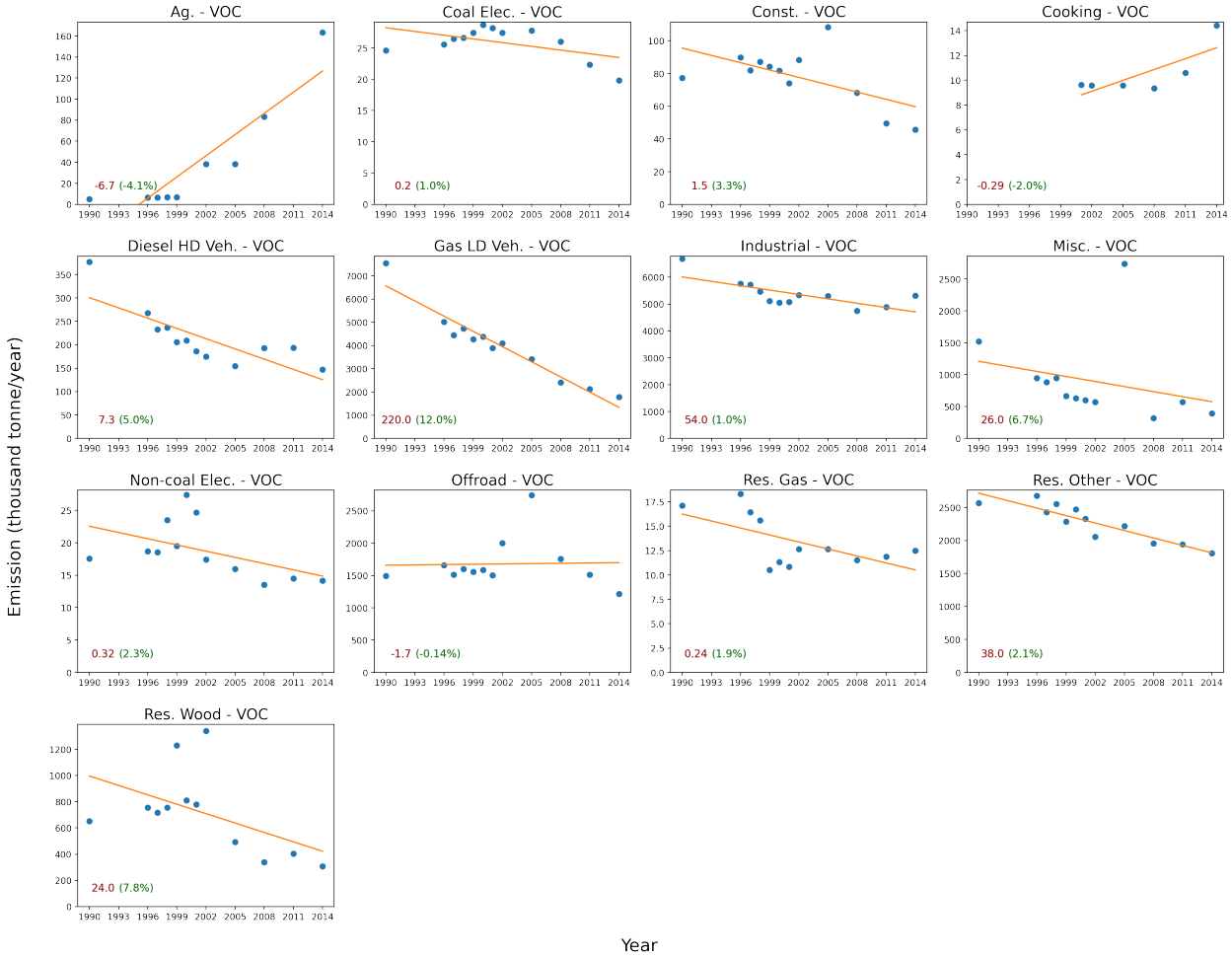

Figure 5.8: Historical emission reduction for VOC from 1990-2014 for each sector of the economy. The text on the lower left corner of each subplot shows the average emission reduction values in absolute (thousand tonne/year; red text) and relative (relative to the year-2014 emissions; green text in parentheses).

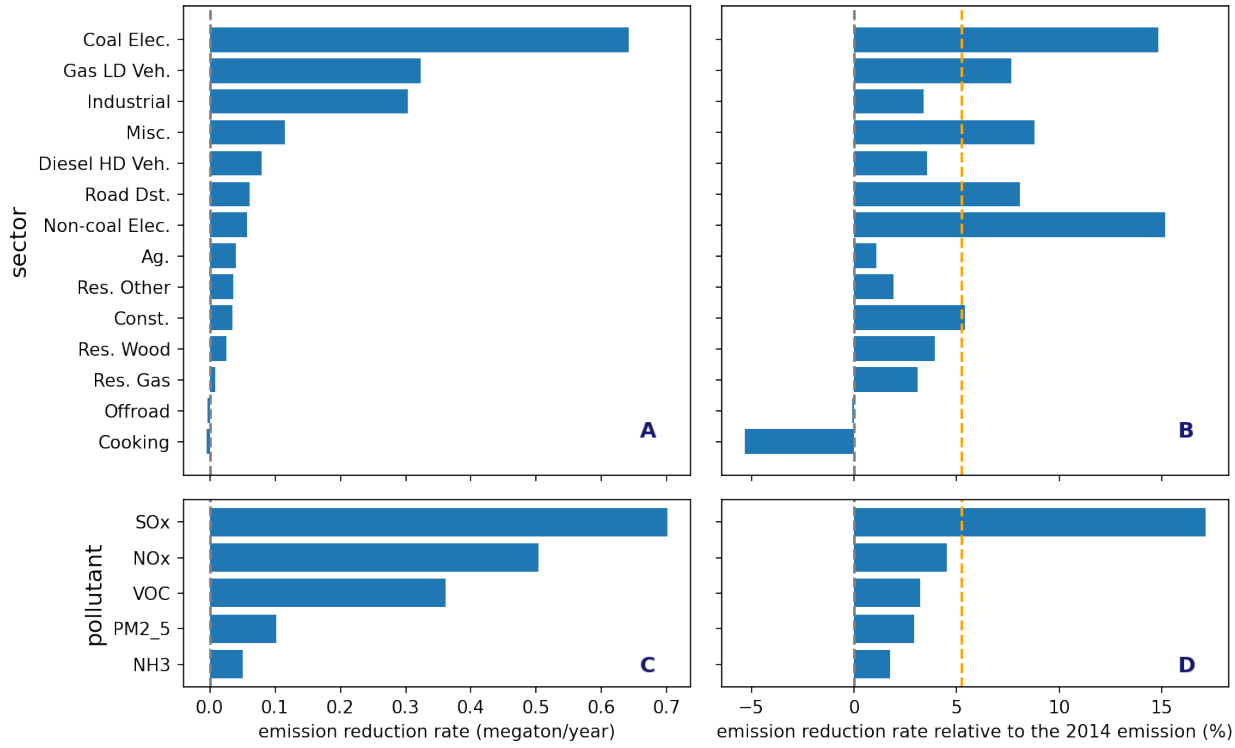

Figure 5.9: Absolute (right panels) and relative (left panels) emission reduction rate from 1990 to 2014 by sector of the economy (first row) and by pollutant (second row).

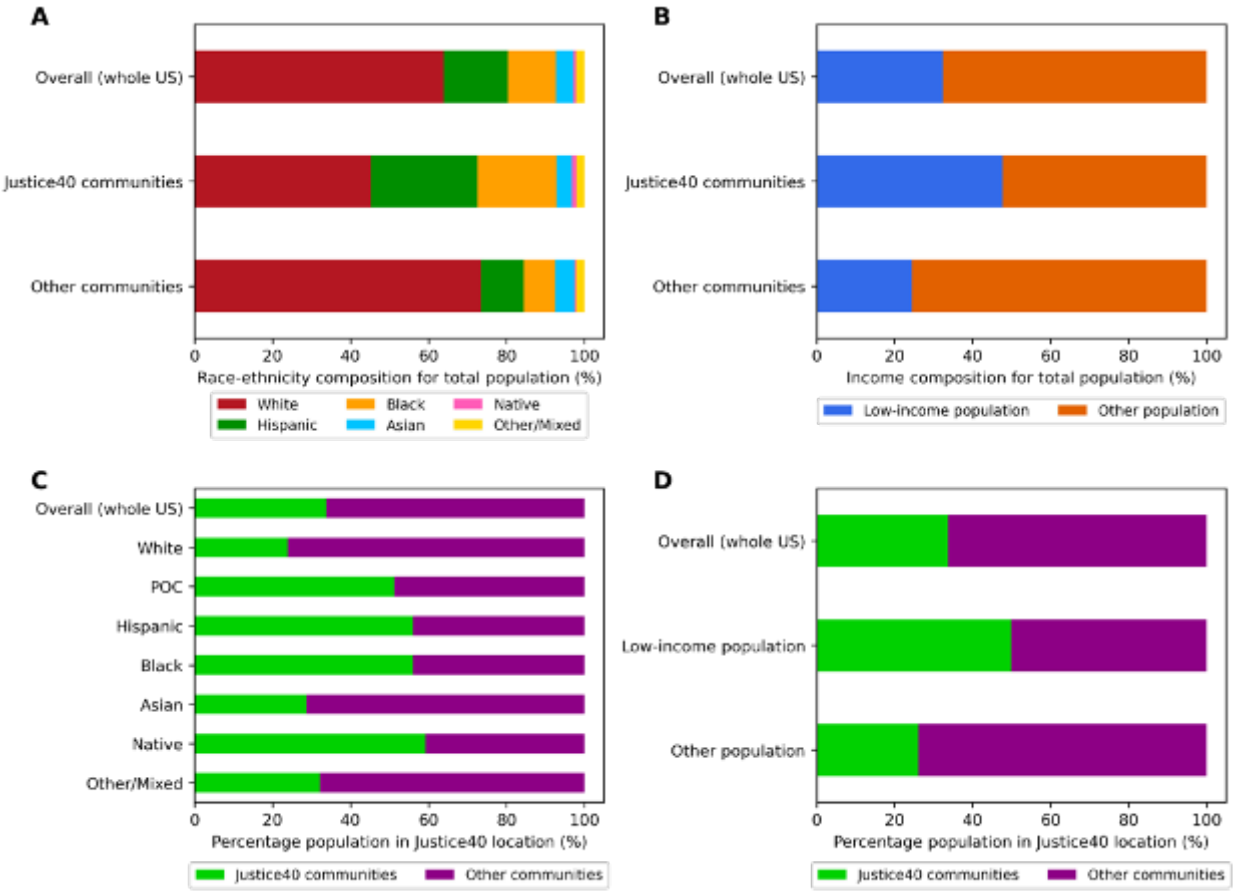

Figure 5.10: Demographic composition for J40 and non-J40 locations. (A) Race-ethnicity composition by J40 community status. (B) Income composition by J40 community status. (C) Percentage of population in J40 communities by race-ethnicity. (D) Percentage of population in J40 communities by income. POC (people of color) refers to all people who are not "non-Hispanic White".

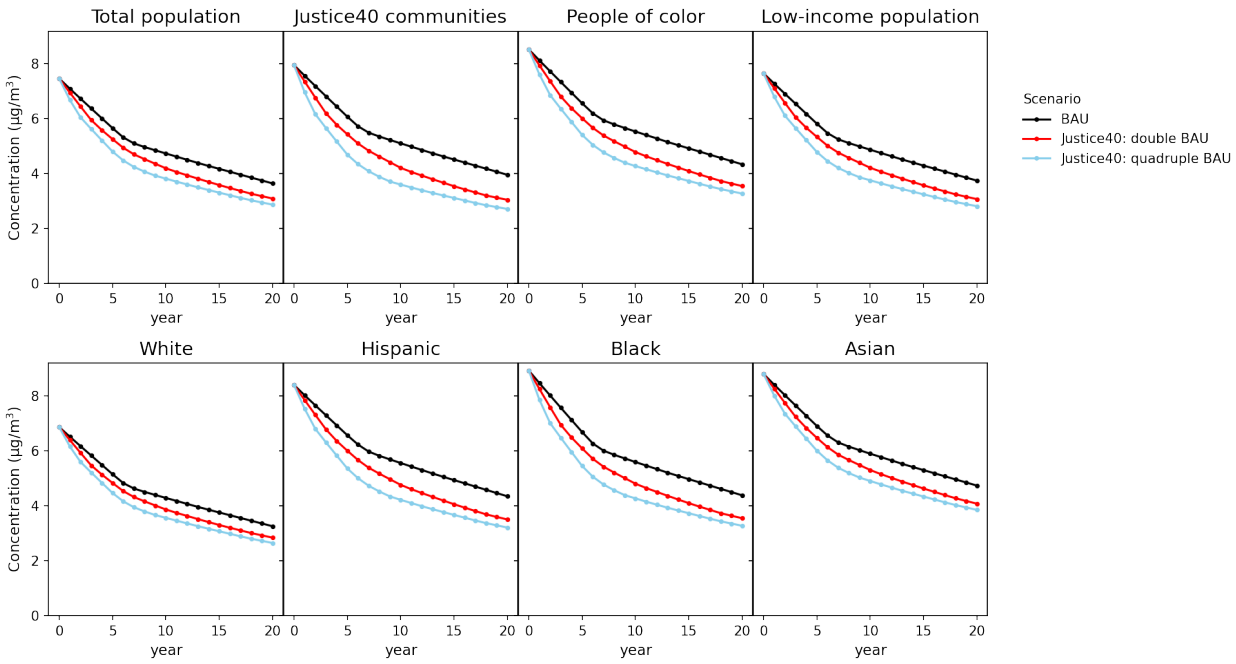

Figure 5.11: $PM_{2.5}$ concentration predictions in 20 years by demographic group for each scenario. Demographic groups include total population, J40 location, people of color (i.e., all groups except non-Hispanic White), low-income population, and the 4 largest race-ethnicity groups (White, Hispanic, Black, and Asian). Scenario are (1) BAU, (2) double the emission-reductions in J40 locations; (3) quadruple the emission-reductions in J40 locations.

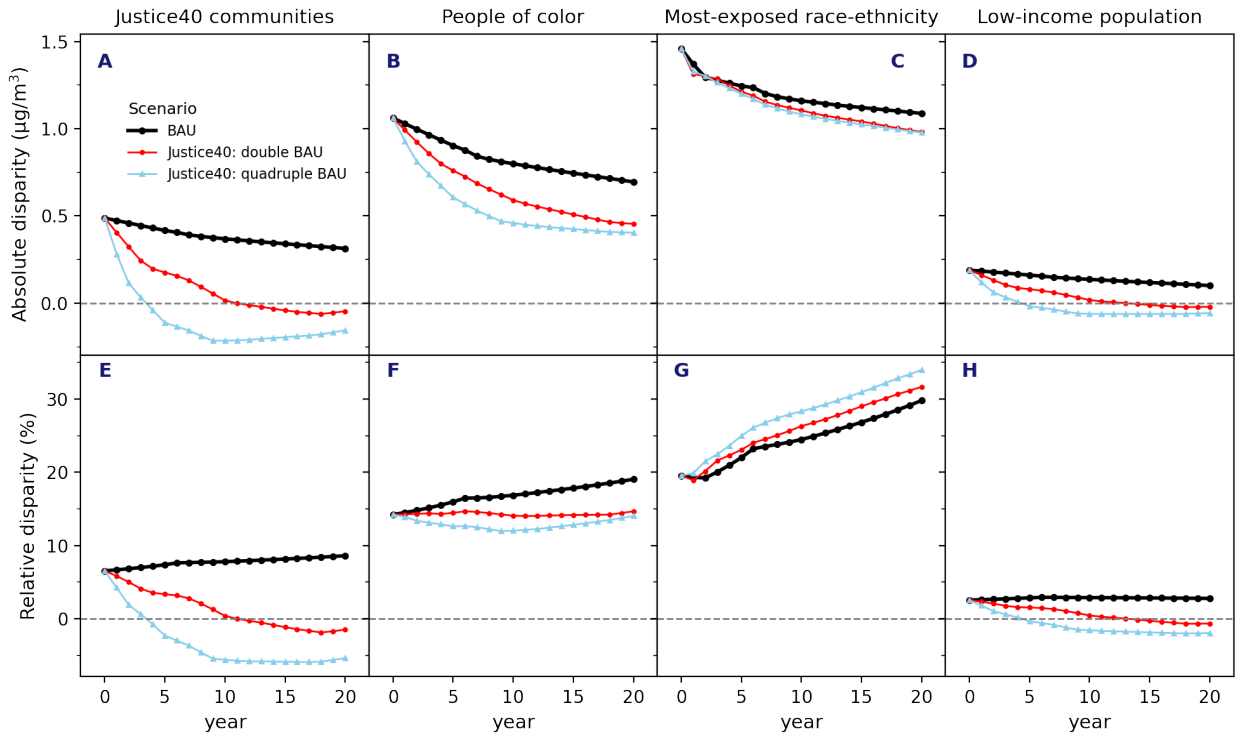

Figure 5.12: $PM_{2.5}$ disparity in 20 years for three possible futures: BAU (business as usual; continuation of historical emission-reduction trends), and two cases where emission-reductions in J40 communities are double the BAU rate or quadruple the BAU rate. Disparities (y-axis values) are relative to the population-average $PM_{2.5}$ concentration, in absolute terms (top row; A to D) or in relative terms (bottom row; E to H). The columns represent population groups: (A and E) J40 communities; (B and F) people of color (i.e., all racial-ethnic groups except non-Hispanic White); (C and G) most-exposed racial-ethnic group; (D and H) low-income population. Zero disparity (y=0, horizontal dashed line) would mean exposures for that group are equal to the population average; negative values indicate average concentrations are below the population average. The BAU policy (black line) does not bring disparities to zero. "Double" and "quadruple" cases (red and blue lines) eliminate or reverse disparities (i.e., leads to disparities at or below zero) for the left and right columns but not for the middle two columns. That suggests that doubling or quadrupling emission-reductions in J40 communities may address higher-than-average air pollution for J40 communities (left column) and for low-income populations (right panel) but not for race-ethnicity (middle two columns).

5.7 Appendix I: Sensitivity analyses

In addition to three emissions-reduction scenarios I explore in the main chapter, I also conduct several sensitivity analyses, including considering alternative methods and outcomes. Sensitivity analyses I consider include: (a) considering the distributions rather than only average values; (b) using "locations" rather than "subpopulations" as the unit of analysis; (c) cases with "rebound", meaning that the additional emission-reductions in J40 communities (i.e., any reductions greater than BAU) are shifted to non-J40 locations rather than "disappearing"; and (d) an alternative approach to predicting future emissions. (e) To compare against the BAU results here, which use a mechanistic model (InMAP), I separately investigated straightforward projections of concentrations and concentration-disparities using an empirical model. (f) Lastly, I conducted several sensitivity analyses with enhanced emission-reductions in communities defined by various indices from EJScreen, instead of J40 communities. (As described below, these six sensitivity analyses are illustrated further in Figs. 5.13 to 5.20.)

For sensitivity test [a] (considering distribution), the distributions of $PM_{2.5}$ concentrations are calculated as the weighted quantiles (10th, 25th, 50th, 75th, and 90th) in each demographic group (J40, race-ethnicity, and income) (Fig. 5.13).

For sensitivity test [b] (demographics in the location), I group the grid cells into quintiles (population weighted) by percentage of people of color (hereafter, POC) and percentage of low-income populations, and then calculate the population-weighted average $PM_{2.5}$ concentrations for each quintile (Fig. 5.14).

For sensitivity test [c] (cases with "rebound"), I consider two alternative scenarios which shift the emissions from the J40 communities to other communities (Fig. 5.15). In the "doubling with rebound" and "quadrupling with rebound" scenarios, the emissions-reduction rates for J40 communities are the same as the "doubling" and "quadrupling" scenarios, respectively; however, for other communities, the emissions-reduction rates are decreased to offset the increased emissions in J40 communities (i.e., for "with rebound", the overall national total emission reduction rates by sectors and pollutants are the same as BAU).

For sensitivity test [d], I calculate the historical emissions-reduction rates using the NEIs from the recent $\sim$ one decade (2008, 2011, 2014), rather than using the NEIs from 1990-2014 (Fig. 5.16).). I explore the three scenarios (BAU, "doubling", and "quadrupling") using the alternative

historical rates.

For sensitivity test [e] (using empirical models to predict BAU), I employ national empirical models to predict $PM_{2.5}$ exposures in the next 20 years (Figs. 5.17 and 5.18).). The 1999-2014 annual-average $PM_{2.5}$ concentrations at Census block group level are obtained from the publicly available Center for Air, Climate, and Energy Solutions (CACES) empirical models, also called land-use regression models, for the contiguous United States (`www.caces.us/data`). I merge the empirical prediction data with the ISRM grids, and then calculate the population weighted average $PM_{2.5}$ concentrations for the overall population and by demographic group. For each group, I calculate the yearly concentration change rates (in relative form) from 1999 to 2014 as the averaged value of concentration change percentage each year relative to the previous year:

$$RR_i = \frac{1}{15} \times \sum_{i=2000}^{2014} (\frac{C_{year,i} - C_{year-1,i}}{C_{year-1,i}} \times 100\%)$$

Here, i represent a certain population group; RR_i represents concentration change rate for group i; $C_{year,i}$ and $C_{year-1,i}$ represent $PM_{2.5}$ exposure concentrations in a given year for group i. I then use the relative change rates to predict the $PM_{2.5}$ exposures for each demographic group in 20 years.

For sensitivity test [f] (EJScreen), I also employ several scenarios with accelerated emission reduction rates in three alternative sets of locations (instead of J40 locations) (Figs. 5.19 and 5.20). All three alternative indicators are defined by EJScreen: 1) percentages of POC populations; 2) $PM_{2.5}$ concentrations; and 3) percentages of low-income populations. For each indicator, I apply enhanced emissions-reduction rates ("double" and "double with rebound") in 30% of the locations with highest scores (e.g., highest POC percentages).

The results from sensitivity tests [a] - [e] (Figs. 5.13 to 5.20) are consistent with the main paper (Figs. 5.1 and 5.2): (1) following historical rates will not address the disparities; (2) disparities by race-ethnicity are larger than by J40 locations; (3) enhanced emissions-reductions in J40 communities will only eliminate disparities in J40 communities and for low income populations, but not for racial-ethnic disparities (only reduce disparities in absolute terms, but not in relative terms). Results from sensitivity tests [f] (Figs. 5.19 and 5.20) indicate that only the scenarios with enhanced emission-reduction in communities of color will reduce both absolute and relative racial-ethnic disparities.

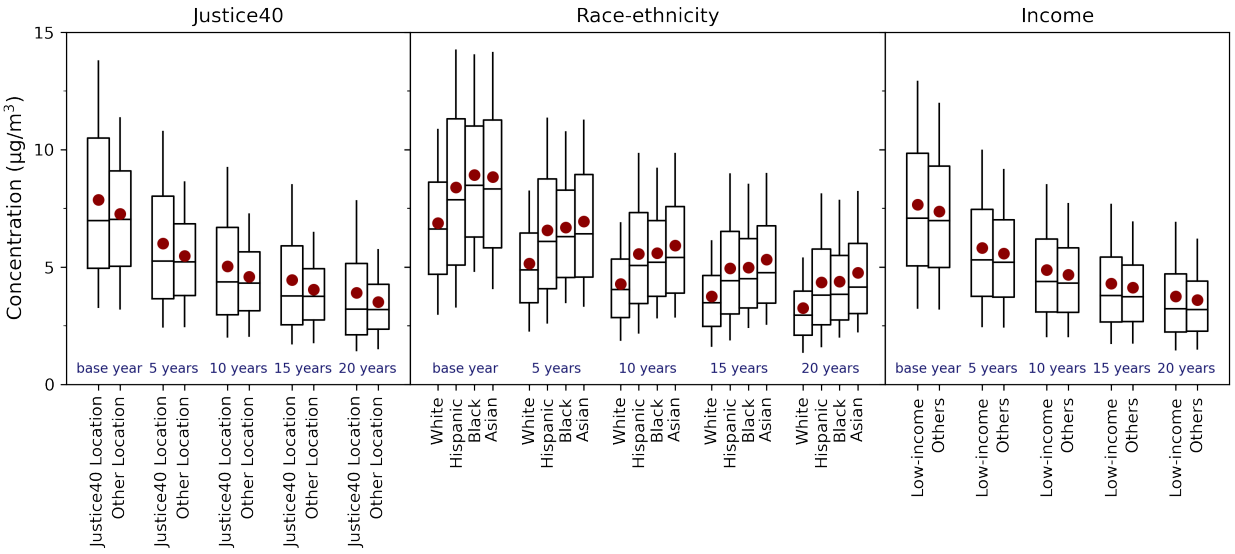

Figure 5.13: Distribution of $PM_{2.5}$ exposure by demographics (J40 community status, race-ethnicity, and income). In each panel, box-and-whisker represents the 10th, 25th, 50th, 75th, and 90th percentiles, and the red circle represents the population-weighted mean.

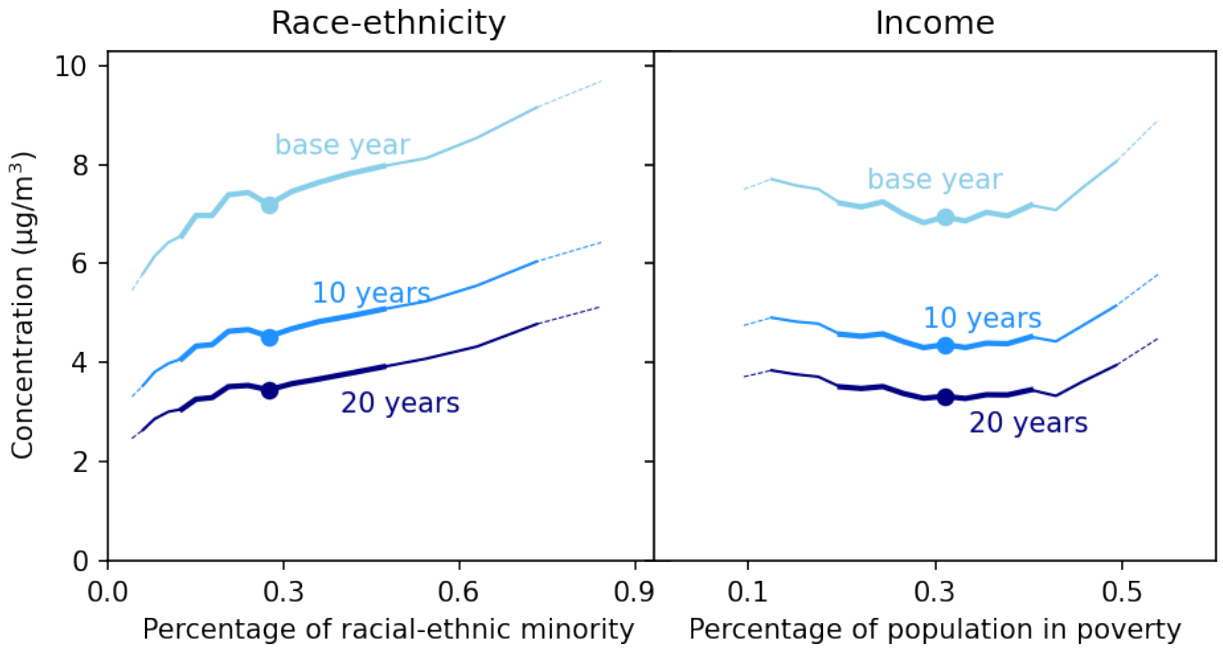

Figure 5.14: Mean $PM_{2.5}$ exposure by demographics of the location. The left panel shows the relationship between $PM_{2.5}$ exposure and percentage of racial-ethnic minority; the right panel shows the relationship between $PM_{2.5}$ exposure and percentage of low-income population. In each panel, the thicker portion of the lines indicate the 5th to 15th quintiles of census block groups, the thin lines indicate the 2nd and 18th quintiles, the dashed lines indicate the 1st to 19th quintiles, and the square icons indicate the median.

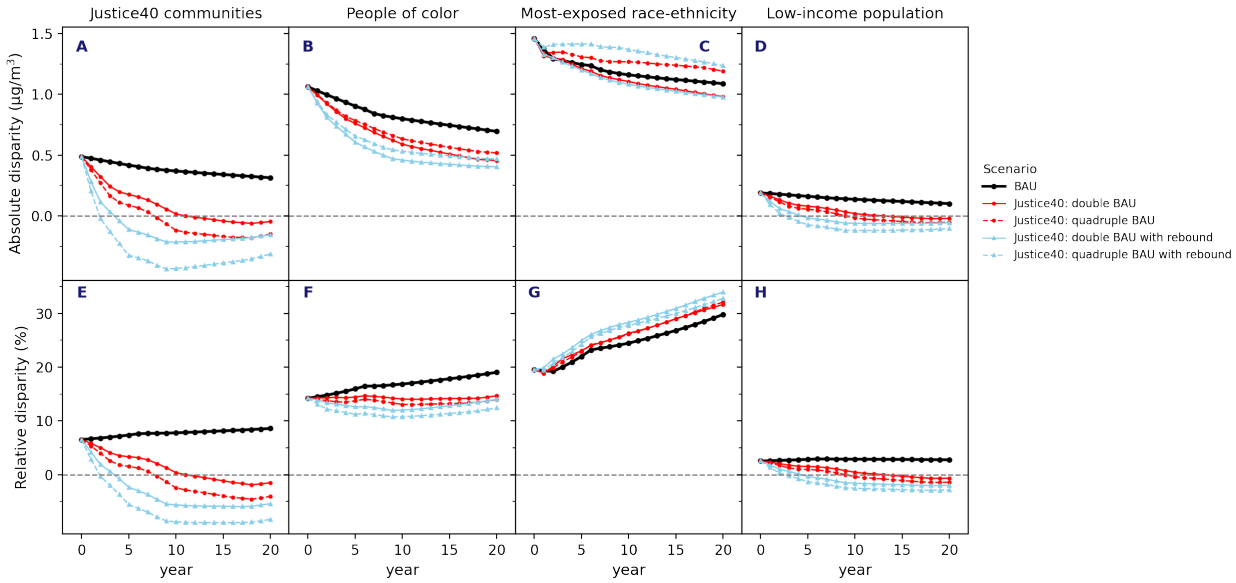

Figure 5.15: $PM_{2.5}$ disparity relative to the population-average, for each emissions scenario for BAU (business as usual), i.e., historical trends of emission reduction occur in all locations, and application of additional emission-reductions in J40 communities. Top row (A to D) absolute disparities in $PM_{2.5}$ levels; bottom row (E to H) relative disparities (% difference relative to the population average). The columns represent population groups: (A and E) J40 communities; (B and F) people of color (i.e., all racial-ethnic groups except non-Hispanic White); (C and G) most exposed racial-ethnic group; (D and H) low-income population. For each panel, the five lines represent the five emissions-reduction scenarios: (1) BAU (business as usual), i.e., historical trends of emission reduction occur in all locations; (2) emission-reductions in J40 communities is double the BAU; (3) similar to #2 but with rebound (i.e., double the emission reductions in J40 locations, with the extra emission-reductions added to non-J40 locations); (4) emission-reductions in J40 locations are quadruple the BAU; (5) similar to #4 but with rebound (i.e., quadruple the emission-reductions in J40 locations, with the extra emission-reductions added to non-J40 locations). The policies considered (red and blue lines) address higher-than-average air pollution for J40 locations (left column) and for low-income populations (right panel) but not for race-ethnicity (middle two columns). Negative values indicate that average concentrations are below the population average.

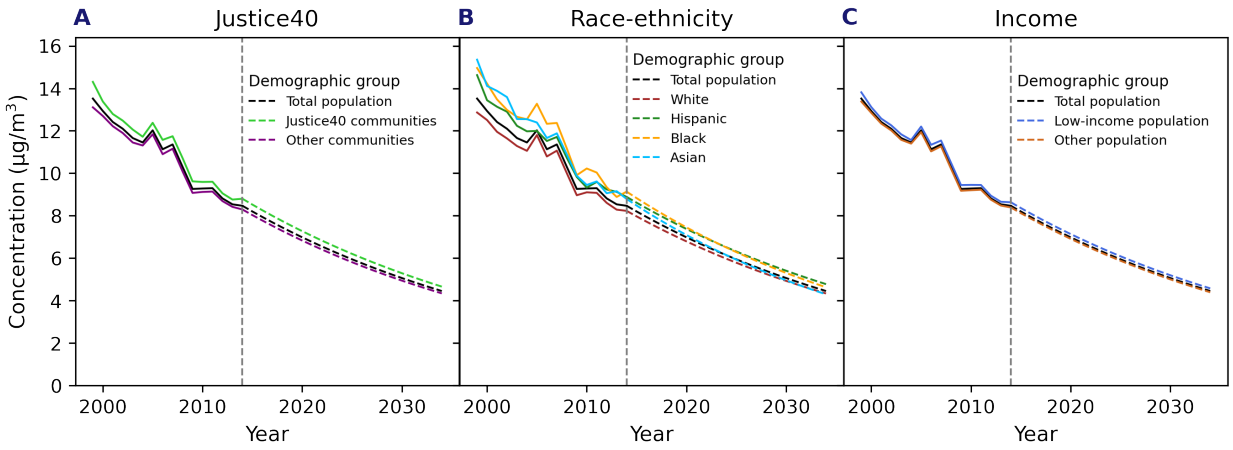

Figure 5.16: $PM_{2.5}$ concentrations and disparities predictions in 20 years for each demographic group. The empirical models provide yearly predictions for years 1999-2014 at Census tract level. The concentrations for years 2015-2024 are predicted using the average yearly concentration reduction percentage from 1999-2014.

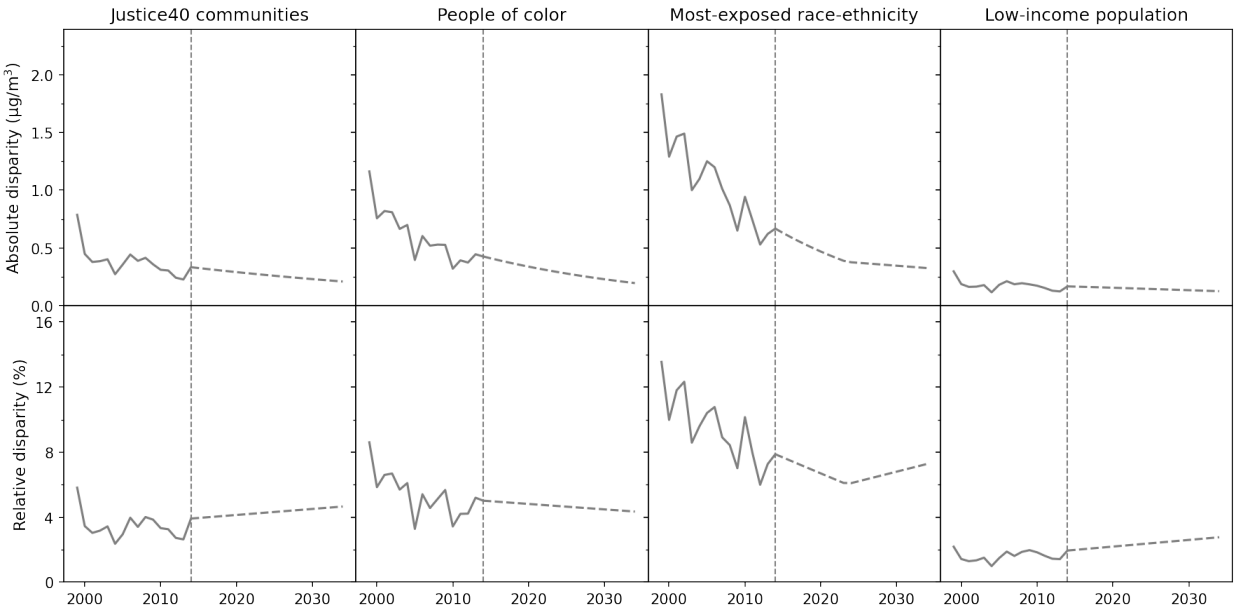

Figure 5.17: Absolute and relative disparities predictions from empirical models. The four panels from left to right represent disparities for J40 locations, for people of color (all other race-ethnicities except non-Hispanic White), for most-exposed race-ethnicity, and for low-income population.

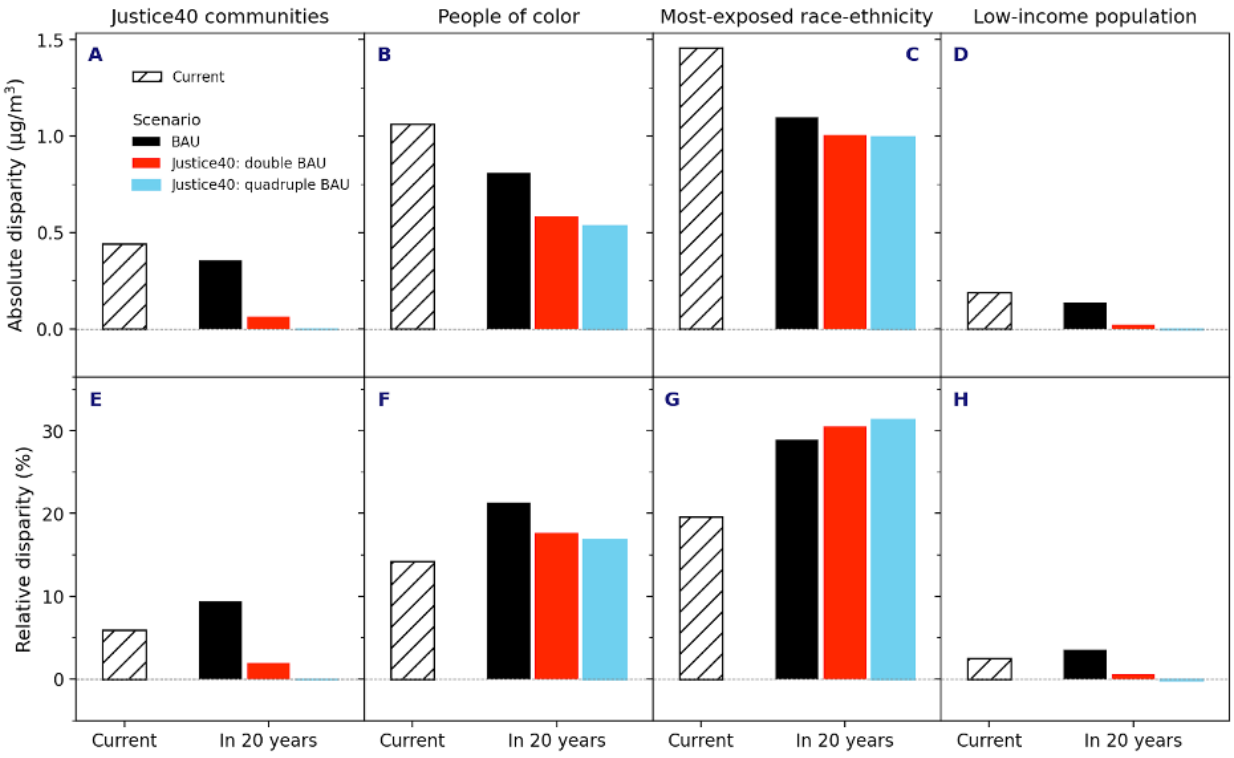

Figure 5.18: Absolute and Relative $PM_{2.5}$ disparities changes in 20 years in three scenarios using historical emission-reduction rates from 2008-2014 (instead of, as in Fig. 5.2, from 1990-2014).

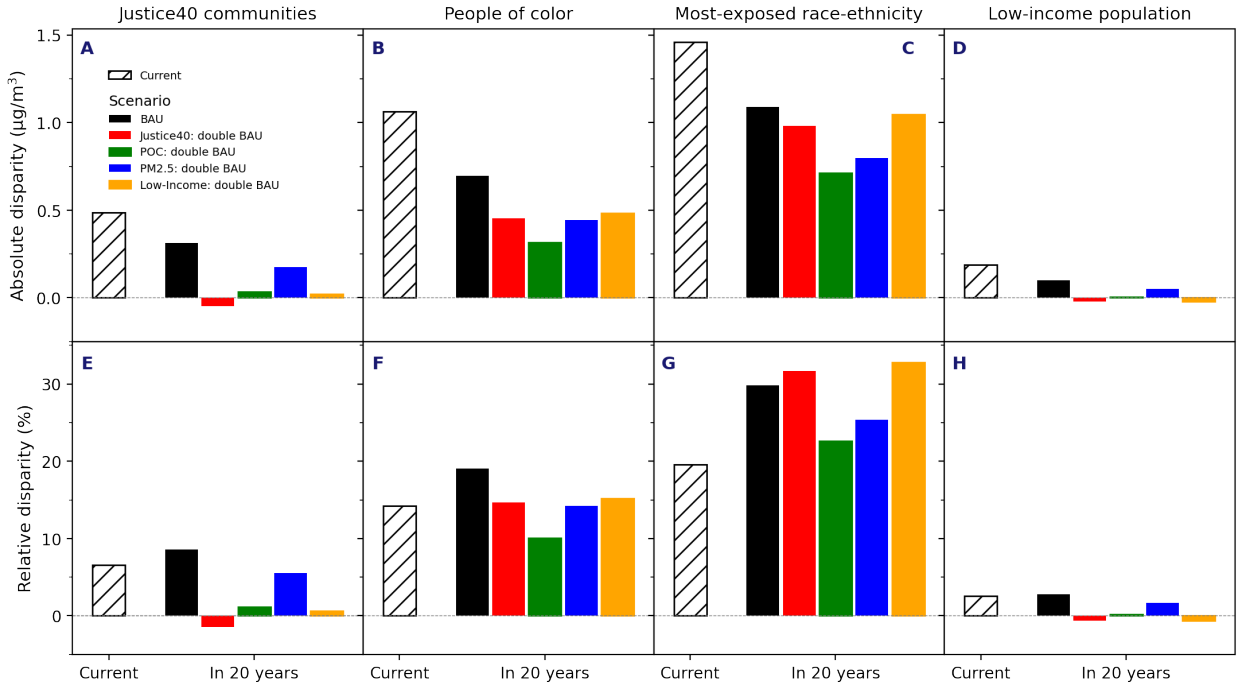

Figure 5.19: Absolute and Relative $PM_{2.5}$ disparities changes in 20 years for alternative enhanced emission-reduction scenarios. Each panel includes five scenarios: the first two scenarios ("BAU" and "J40: double BAU") are the same as Fig. 5.2; the other three scenarios are to define locations nationally where emission-reductions are doubled based on EJScreen tool rather than J40. "POC: double BAU" represents enhanced emission-reductions using percentage of people of color population; "$PM_{2.5}$: double BAU" represents enhanced emission reductions using $PM_{2.5}$ concentration; "low-income: double BAU" represents enhanced emission reductions using percentage of low-income population. Each of the three scenarios define the Census Tract with top 30% scores for the indicators as enhanced emission-reductions.

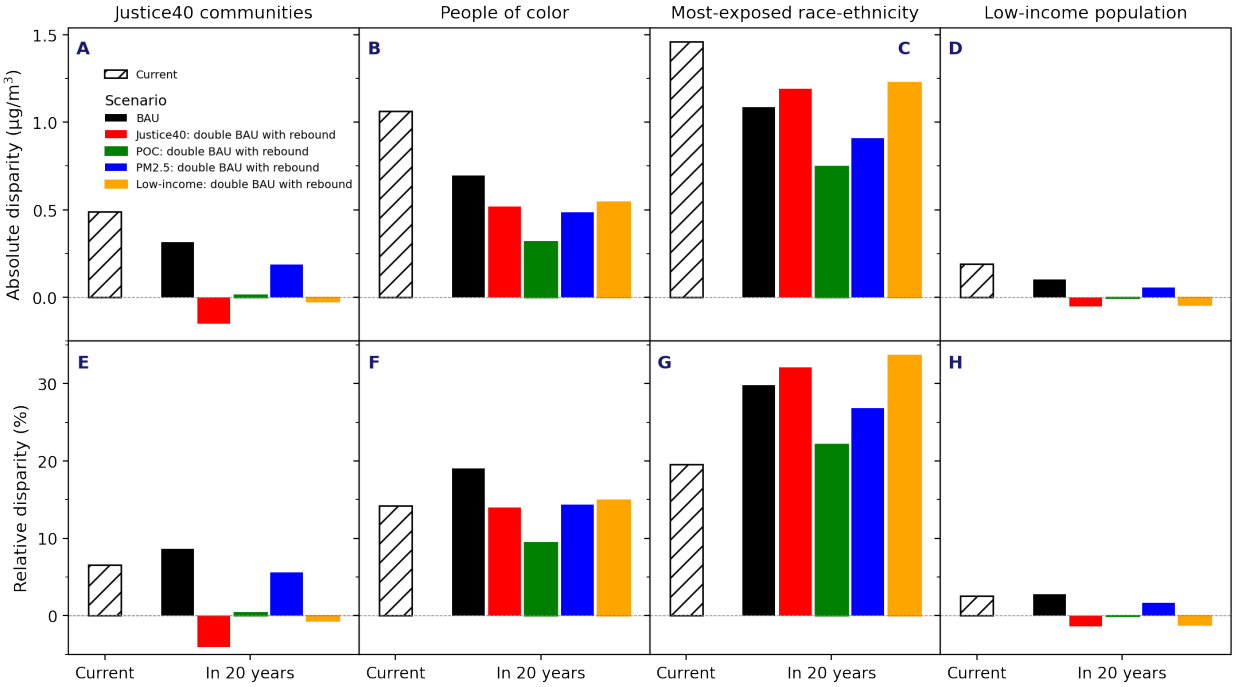

Figure 5.20: Absolute and Relative $PM_{2.5}$ disparities changes in 20 years for alternative doubling emission-reduction with rebound scenarios. The "doubling BAU with rebound" scenarios are analogous to Fig. 5.19 but with the extra emission-reductions in locations with enhanced emission reductions "rebound" to the rest of the locations (i.e., "with rebound", the overall total national emission-reductions each year are the same as BAU).

Chapter 6

CONCLUSIONS

This chapter summarizes the main findings of the studies presented in **Chapters 2 to 5**, describes general limitations of these findings, discusses contributions and potential policy implications, and suggests areas for further investigation.

6.1 Summary of findings

This dissertation investigates air pollution concentration and exposure patterns in US and China, and explores possible approaches/policies to address the persistent US racial-ethnic inequalities in exposure to ambient air pollution. Air pollutants investigated in this dissertation are two important criteria air pollutants: $PM_{2.5}$ and NO_2. Research approaches I employed include both empirical and mechanistic models.

The followings are the main research findings for each of the research objectives (introduced in **Chapter 1**):

- Objective #1: Quantify the spatial sources of ambient NO_2 and $PM_{2.5}$ for the contiguous US (**Chapter 2**).

  I find that NO_2 is of urban origin and varies by urbanicity; climate or geographic regions have less effects on the NO_2 concentrations. For example, concentrations are ~3-fold higher in large urban areas (average: 13 ppb) than rural areas (4 ppb); ~90% of the differences are driven by "neighborhood" and "mid-range" components. The concentration differences by region are minor after controlling for urbanicity. In contrast, $PM_{2.5}$ is a regional pollutant with a strong secondary component; the concentrations are dominated by "long-range" components and vary at state and regional level. For example, "long-range" component contributes to >50% of the total concentration in almost all geographic regions (except in the West [~30%]); "long-range" concentration in the Midwest (7 $\mu g/m^3$) is almost three times than in the West (2.3 $\mu g/m^3$). $PM_{2.5}$ concentrations and differences by urbanicity are modest or minor.

- Objective #2: Quantify the relationships between ambient NO_2 and $PM_{2.5}$ air pollution exposures and socioeconomic status (SES) in China (**Chapter 3**).

  I find that in reverse of the typical patterns in the US, the ambient NO_2 and $PM_{2.5}$ levels in China are higher for higher-SES populations than for lower-SES populations, higher for long-standing urban residents than for rural-to-urban migrant populations, and higher for the majority ethnic group (Han) than for the average across nine minority groups. For three SES measurements (individual SES score, community-averaged SES score, gridded GDP per capita), a 1-interquartile range higher SES corresponds to higher concentrations of 6-9 $\mu g/m^3$ NO_2 and 3-6 $\mu g/m^3$ $PM_{2.5}$; average concentrations for the highest and lowest 20th percentile of SES differ by 41–89% for NO_2 and 12–25% for $PM_{2.5}$. Exposure inequalities are higher for NO_2 than $PM_{2.5}$. The positive relationship (i.e., SES and pollution levels are positively correlated, in China) holds in rural and urban locations, across geographic regions, across a wide range of spatial resolution from 1 km to 100 km, and for modeled vs. measured pollution concentrations. My findings are consistent with the idea that in China's current industrialization and urbanization stage, economic development is positively correlated with both SES and air pollution.

- Objective #3: Investigate policies to address US national racial-ethnic disparities in exposure to $PM_{2.5}$ air pollution (**Chapters 4 and 5**).

  I find that if we have no intervention and simply continue the historical rates of emission reductions, disparities will remain in place. Further, for three broad emission-reduction approaches (**Chapter 4**), I find that US national inequalities in exposure can be eliminated with minor emission-reductions (optimal: $\sim$1% of total emissions) if targeting specific locations. In contrast, achieving that outcome using existing regulatory strategies would require eliminating essentially all emissions (if targeting specific economic sectors) or is not possible (if requiring urban regions to meet concentration standards). In addition, I find no tradeoff between reducing overall average concentration and reducing national inequalities; rather, the approach that does the best for reducing national inequalities (i.e., location-specific strategies) also does as well as or better than the other two approach (i.e., sectors-specific; meeting concentration standards) for reducing overall averages.

As for an ongoing location-based strategies – CEJST/J40 policy from Biden Administration (**Chapter 5**), based on more realistic emission-reduction scenario analyses (unlike the best potential analyses in **Chapter 4**), I find that enhanced (doubled or quadrupled) emission-reductions in J40 communities only eliminated inequalities for J40 communities and for low-income populations in 20 years; yet they do not reduce relative disparities by race-ethnicity (although they do decrease absolute disparities). The results indicate that although CEJST/J40 is a location-based government policy designed to address EJ, it's not just focused on $PM_{2.5}$; in addition, the J40 communities are composed of modestly higher proportions of people of color and low-income populations compared with national averages. Thus, our results suggest that additional and more targeted actions, beyond CEJST/J40, will be needed to end racial-ethnic exposure disparities in the next decades.

Overall, this dissertation finds that patterns of air pollution exposure disparities vary by: (1) countries, based on historical and current contexts; and (2) pollutants, based in part on their spatial sources and patterns of spatial variability. To address the persistent air pollution exposure inequalities in the US, policies need to: (1) incorporate location-specific emissions reductions into the US air quality regulatory framework, and (2) specifically target emission sources that impact overburdened communities.

6.2 Limitations

Specific limitations for each study are discussed in **Chapters 2 to 5**. I summarize here three general limitations for my studies.

First, these studies estimate population exposures using ambient concentrations, which are related but distinct from individual-level exposures and risks. Indoor air pollution, micro-environments, occupational exposure, and individual's mobility can also contribute to exposure disparities [60,62, 205,206,209,210,227,227,229]. In addition, differences in susceptibility, background rates of disease, and access to healthcare can modulate disparities by concentration [230] (**Chapter 5** investigated health disparities, but simply used the same background rates of disease for different population). Future work could shed important additional insight by considering those factors for exposure and health disparities.

Second, my results are based on national air quality models and national demographic data,

which have documented uncertainties but are robust at national and aggregate level. However, for investigating on specific regions or cities, I recommend additional (local) data and analyses, including local air quality data, local emission sources and hot-spots, local policies, climate and seasonal patterns, and, local community characteristics. In addition, EPA NEIs may underestimate the emissions from "super emitters", such as $PM_{2.5}$ from light duty vehicles; the adoption rate of electrical vehicles also varies in different places, which may effect the forecast of future emissions. Future research for specific regions or cities could investigate more on those factors.

Third, my dissertation only investigates two important air pollutants ($PM_{2.5}$, NO_2) because they have important health effects [4, 87–89] and more importantly, there are publicly available models (LURs for $PM_{2.5}$ and NO_2; InMAP for $PM_{2.5}$) for these two pollutants [68, 93, 109]. I do not explore other air pollutants, climate effects, and toxicities, which may have different distribution patterns among populations and contribute to the cumulative burdens.

6.3 Contributions and implications

This dissertation contributes to the literature of air pollution and environmental inequality with new knowledge, insights, approaches, and useful datasets. The results for understanding and addressing EJ also provides policy implications that will be of potential interest to scientists and engineers, policy makers, public health officials, environmental justice communities, and environmental organizations and advocates. The following are contributions and potential policy implications for the works presented in **Chapters 2 to 5** in the aspects of the findings, methods, and data.

Chapter 2 develops an straightforward and readily scalable algorithm for spatial decomposition, based on "spatial-increment" and moving window approach with fixed radii (1, 10, and 100 km). This algorithm is computationally feasible for high spatial resolution and for large domains, and can be easily applied in other countries and for a specific city, state, or region, with flexible choice of radii. The multi-year (2000-2015) spatial decomposition data for annual-averaged ambient $PM_{2.5}$ and NO_2 in contiguous US at both census block and block group levels can be freely available online: `data.mendeley.com/datasets/ckdvx3d3zc/1` (block group level) and `data.mendeley.com/datasets/bz8pdbcvf2/2` (block level). The spatial decomposition data has already been used in an epidemiology study [250] and an EJ study [251] to identify the spatial sources that contribute to the adverse health outcomes and exposure disparities.

Chapter 3 provides new knowledge and insights by extending the EJ research to China. As mentioned in **Chapter 1**, most of the previous EJ studies are for the US and a few high-income countries; few papers investigate China and other middle- and low-income countries, which have the largest health burdens (∼90% of the premature deaths happen in those regions) of ambient air pollution. **Chapter 3** provides the first comprehensive analyses for ambient air pollution exposure inequalities in China, and indicates a positive relationship patterns in China, which is different from the general knowledge for the US. My findings likely have implications for other locations besides China. Specifically, for countries that are (like China) relatively homogenous racially and with economic development that is relatively recent (past decades), my results suggest that exposures may follow a different pattern than has been observed for the US.

Chapters 4 and 5 tackle a broad systemic problem of national importance for US, and, for the first time, identify possible solutions to address the persistent national racial-ethnic inequalities in air pollution exposure. These two studies also employ a new tool (InMAP/ISRM), which is a RCM/IAM, and provide useful examples of using the new tool in Regulatory Impact Analysis.

Chapter 4 indicates the ineffectiveness of the current US air quality regulations, and highlights the importance of incorporating location-specific emissions reductions into the regulatory framework. The results support the long-standing request from environmental justice communities and local organizations for location-specific solutions that center overburdened communities [37, 158, 237–239], and inform national action (e.g., J40/CEJST) and emerging state and local environmental justice laws to identify overburdened communities and develop community emissions reduction plans [234–236].

Chapter 5 indicates the possible failure of current J40/CEJST to reduce racial-ethnic disparity (the largest source of inequalities) in air pollution, and highlights the importance of specifically targeting the communities that are overburdened by pollution and disadvantaged in race-ethnicity. Although J40/CEJST is a location-specific emission-reduction strategies, it doesn't consider racial makeup when defining the J40 communities. Since race-based actions created the inequalities, solving it will likely be more difficult if the government does not consider, or bars itself from considering, information about the racial makeup of communities as part of its decision-making and action. Findings here indicate that to address the inequality, additional policies and revised tools are needed that are specifically designed to remove race-based and socioeconomic disparities. Another factor that causes the ineffectiveness of J40/CEJST is the lacking of Regulatory Impact

Analysis to test the potential impacts of proposed strategies and tools on air pollution disparities. My analyses also highlight the necessity of this sort of testing and evaluation; and the importance of new tools such as InMAP/ISRM to conduct such analyses, by being faster and easier to carry out, and by providing national coverage at much higher spatial resolution than many conventional models [25, 248, 249].

In addition, to make InMAP/ISRM more accessible for other researchers and policy makers to use, I've made some tutorials on using InMAP/ISRM on local computers and on google cloud. Also, the code for data processing, doing analyses and making plots in **Chapter 5** is publicly available through GitHub. I'm also currently involved in developing new InMAP/ISRMs for other countries or specific regions (e.g., India, California).

6.4 Future directions for research

The followings are areas for future research, building directly on the findings and limitations of the studies presented in this dissertation.

(1) Environmental inequality analyses for other pollutants and air toxicities.

This dissertation only investigated exposure inequality patterns for $PM_{2.5}$ and NO_2, and disparity reduction strategies for $PM_{2.5}$. This focus reflects current availability of data and models. Future research could investigate exposure patterns and disparity-reduction strategies for other air pollutants [10], including air toxics [26] and source-apportioned $PM_{2.5}$ [252]. For example, some pollutants (e.g., ultrafine particles) and toxics may be more localized; have greater disparities among populations, including disproportionately affecting communities of color; and be more of concern to people living near those emission sources. Some secondary pollutants (e.g. ozone, SO_2) may be more regional and more reflective of the meteorological factors; thus, different pollutants require different controlling strategies.

(2) Cumulative impacts with climate effects and social stressors.

This dissertation only investigated one environmental risk (exposure to ambient air pollution) among populations. However, exposure disparities to air pollution often co-occur with disparities in other social stressors, such as poverty, racial discrimination, crime, and substance abuse, which can exacerbate community vulnerability and susceptibility [29]. In addition, air pollution and climate change are closely related: many air pollutants are also referred to as climate forcers. For example, different components of $PM_{2.5}$ can have either warming (e.g., black carbon) or cooling (e.g., SO_x)

effects on the climate. Emissions from many sectors, activities, and locations have both health impacts (from air pollution) and climate impacts, and may have some trade-off between the two factors [57]. Future research could shed light on the cumulative impacts with social stressors and climate impacts to better identify the disparities in risks for the overburdened communities, and hopefully will identify "win-win" approaches in dealing with both air pollution and climate change.

(3) Impacts of wildfire emissions to the exposure and health inequalities.

For addressing air pollution exposures and exposure disparities, this dissertation focused on anthropogenic sources, without investigation of wildfires and emissions from biogenic sources. Wildfires have received increasing attention in recent years. For example, one recent study indicates that the contribution of wildfires to the total $PM_{2.5}$ in US has increased from 10% a decade ago to 25% recently; the spatial patterns in ambient smoke exposure do not follow traditional socioeconomic pollution exposure gradients [253]. Other studies also imply that the toxicity of $PM_{2.5}$ from wildfires may be different from other sources [254, 255]. Future research could investigate how those sources contribute to the health impacts, exposure disparities, and investigate possible technologies/policies to address emissions from those sources.

(4) Attributions of historical causes to current-day exposure inequality.

Attributing the historical causes and trends to present-day exposure inequality in the US is another area that further research could investigate. Many studies have developed EJ theories to explain the inequalities in US [24,36–39,43,44]; a few recent studies have investigated the historical causes of exposure inequality, such as the effects of red-lining policies [35, 256], and the facility siting versus "minority move-in" [36]. Future studies could further advance the understanding of historical causes by investigating the historical trends regarding where emission sources are located, historical patterns in road construction/widening, the effects of when were industrial facilities put in on present-day inequalities in proximity to emission sources, or effects of year of home-construction on present-day exposure inequalities.

(5) Longitudinal analyses of environmental inequalities in low- and middle-income countries.

This dissertation investigated the patterns of air pollution exposure inequalities in China. Future work can usefully explore patterns in other countries, look at changes over time, use some unified definitions of SES and exposures to compare patterns across multiple countries, and explore underlying theories to further explain the differences in exposure patterns in different countries. For example, as discussed in **Chapter 3**, if future economic development and urbanization is con-

current with improvements in air pollution—especially in high- and upper-middle-income counties in China [196]—then current patterns of environmental inequality may change over time. In addition, EJ theory in the US [158, 197] and elsewhere [198, 199] highlights the concept of sacrifice zones—locations that lack political power and receive disproportionately high environmental risks. This is a critical aspect of how the current pollution-SES relationships in China might shift over time that will depend in part on inequities in political power and whether sacrifice zones become more prevalent. These conjectures highlight the need for longitudinal analyses of environmental inequalities in low- and middle-income countries.

(6) Strategies to reduce air pollution levels in low- and middle-income countries.

Another future direction for low- and middle-income countries is to investigate possible strategies to reduce the air pollution levels. For example, in year-2015, there are more than 80% of the Chinese population lives in cities with > 35 μg/m^3 annual $PM_{2.5}$ levels [68]. Conducting scenario analyses in low- and middle-income countries are much harder than in US because the detailed, high-resolution, publicly available emission inventories often are lacking (unlike the NEIs in the US); plus, there are less publicly available RCMs/IAMs in those countries. Future research could work on building high-resolution emission inventories incorporating satellite observations [257, 258]. In addition, developing RCMs (e.g. country-specific InMAP [259]) in those regions is urgently needed.

(7) Identifications of personal exposures.

As discussed in limitation section, this dissertation investigates ambient concentrations, which are related to but distinct from individual-level exposures. Future research could shed important additional insight by considering mobility (e.g., travel for work and recreation), occupational exposure, and near-source exposures (e.g., environmental tobacco smoke; time spent on-roadways; indoor use of solid fuels for cooking, heating, or lighting) factors to better identify the personal exposures [200–212]. Some of the factors are already well studied, e.g., indoor use of solid fuels for cooking generally happens in lower-income, not in higher-income, households [213, 214]. Future research also could apply low-cost sensors [260, 261], mobile monitors [262], and wearable/portable monitors [263, 264] to better assess personal exposures.

BIBLIOGRAPHY

[1] A. J. Cohen, *et al.*, Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015, *Lancet* **389**, 1907–1918 (2017).

[2] J. Lelieveld, J. S. Evans, M. Fnais, D. Giannadaki, A. Pozzer, The contribution of outdoor air pollution sources to premature mortality on a global scale, *Nature* **525**, 367–371 (2015).

[3] S. S. Lim, *et al.*, A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010, *Lancet* **380**, 2224–2260 (2012).

[4] C. A. Pope III, *et al.*, Lung cancer and cardiovascular disease mortality associated with ambient air pollution and cigarette smoke: shape of the exposure–response relationships, *Environ. Health Perspect.* **119**, 1616–1621 (2011).

[5] M. Brauer, *et al.*, Ambient air pollution exposure estimation for the global burden of disease 2013, *Environ. Sci. Technol.* **50**, 79–88 (2016).

[6] World Health Organization, (WHO), "Ambient (outdoor) air pollution" (2021). `https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health`. Accessed Oct 8, 2020.

[7] L. P. Clark, D. B. Millet, J. D. Marshall, National patterns in environmental injustice and inequality: outdoor NO_2 air pollution in the United States, *PloS ONE* **9**, e94431 (2014).

[8] L. P. Clark, D. B. Millet, J. D. Marshall, Changes in transportation-related air pollution exposures by race-ethnicity and socioeconomic status: Outdoor nitrogen dioxide in the United States in 2000 and 2010, *Environ. Health Perspect.* **125**, 097012 (2017).

[9] A. Hajat, C. Hsia, M. S. O'Neill, Socioeconomic disparities and air pollution exposure: A global review, *Curr. Environ. Health Rep.* **2**, 440–450 (2015).

[10] J. Liu, *et al.*, Disparities in air pollution exposure in the United States by race/ethnicity and income, 1990–2010, *Environ. Health Perspect.* **129**, 127005 (2021).

[11] J. D. Marshall, K. R. Swor, N. P. Nguyen, Prioritizing environmental justice and equality: diesel emissions in Southern California, *Environ. Sci. Technol.* **48**, 4063–4068 (2014).

[12] J. D. Marshall, Environmental inequality: air pollution exposures in California's South Coast Air Basin, *Atmos. Environ.* **42**, 5499–5503 (2008).

[13] H. Guo, *et al.*, Who are more exposed to $PM_{2.5}$ pollution: A mobile phone data approach, *Environ. Int.* **143**, 105821 (2020).

[14] K. Jian, Environmental justice: can an american discourse make sense in Chinese environmental law, *Temp. J. Sci. Tech. & Envtl. L.* **24**, 253 (2005).

[15] C. Ma, Who bears the environmental burden in China—an analysis of the distribution of industrial pollution sources?, *Ecol. Econ.* **69**, 1869–1876 (2010).

[16] D. A. Paolella, *et al.*, Effect of model spatial resolution on estimates of fine particulate matter exposure and exposure disparities in the United States, *Environ. Sci. Technol. Lett.* **5**, 436–441 (2018).

[17] L. P. Clark, M. H. Harris, J. S. Apte, J. D. Marshall, National and intraurban air pollution exposure disparity estimates in the United States: Impact of data-aggregation spatial scale, *Environ. Sci. Technol. letters* **9**, 786–791 (2022).

[18] W. Bowen, An analytical review of environmental justice research: What do we really know?, *Environ. Manage.* **29**, 3–15 (2002).

[19] J. Maantay, Mapping environmental injustices: pitfalls and potential of geographic information systems in assessing environmental health and equity., *Environ. Health Perspect.* **110**, 161–171 (2002).

[20] R. J. Brulle, D. N. Pellow, Environmental justice: Human health and environmental inequalities, *Annu. Rev. Public Health* **27**, 103–124 (2006).

[21] L. B. Lave, E. P. Seskin, *Air pollution and human health* (Routledge, 2013).

[22] P. Brown, Race, class, and environmental health: a review and systematization of the literature, *Environ. Res.* **69**, 15–30 (1995).

[23] U.S. Environmental Protection Agency, (EPA), "Environmental Justice" (EPA, 2023). `https://www.epa.gov/environmentaljustice`. Accessed Apr 6th, 2023.

[24] R. Morello-Frosch, R. Lopez, The riskscape and the color line: examining the role of segregation in environmental health disparities, *Environ. Res.* **102**, 181–196 (2006).

[25] C. W. Tessum, *et al.*, $PM_{2.5}$ polluters disproportionately and systemically affect people of color in the United States, *Sci. Adv.* **7**, eabf4491 (2021).

[26] R. Morello-Frosch, B. M. Jesdale, Separate and unequal: residential segregation and estimated cancer risks associated with ambient air toxics in US metropolitan areas, *Environ. Health Perspect.* **114**, 386–393 (2006).

[27] A. Rosofsky, J. I. Levy, A. Zanobetti, P. Janulewicz, M. P. Fabian, Temporal trends in air pollution exposure inequality in massachusetts, *Environ. Res.* **161**, 76–86 (2018).

[28] A. Jbaily, *et al.*, Air pollution exposure disparities across US population and income groups, *Nature* **601**, 228–233 (2022).

[29] R. Morello-Frosch, M. Zuk, M. Jerrett, B. Shamasunder, A. D. Kyle, Understanding the cumulative impacts of inequalities in environmental health: implications for policy, *Health Aff.* **30**, 879–887 (2011).

[30] J. Currie, J. Voorheis, R. Walker, What caused racial disparities in particulate exposure to fall? New evidence from the Clean Air Act and satellite-based measures of air quality, *Am. Econ. Rev.* **113**, 71–97 (2023).

[31] C. W. Tessum, *et al.*, Inequity in consumption of goods and services adds to racial–ethnic disparities in air pollution exposure, *Proc. Natl. Acad. Sci. U.S.A* **116**, 6001–6006 (2019).

[32] J. Colmer, I. Hardman, J. Shimshack, J. Voorheis, Disparities in $PM_{2.5}$ air pollution in the United States, *Science* **369**, 575–578 (2020).

[33] S. E. Chambliss, *et al.*, Local-and regional-scale racial and ethnic disparities in air pollution determined by long-term mobile monitoring, *Proc. Natl. Acad. Sci. U.S.A* **118**, e2109249118 (2021).

[34] P. Mohai, D. Pellow, J. T. Roberts, Environmental justice, *Annu. Rev. Environ. Resour.* **34**, 405–430 (2009).

[35] H. M. Lane, R. Morello-Frosch, J. D. Marshall, J. S. Apte, Historical redlining is associated with present-day air pollution disparities in U.S. cities, *Environ. Sci. Technol. Lett.* **In Press** (2022).

[36] M. Pastor, J. Sadd, J. Hipp, Which came first? Toxic facilities, minority move-in, and environmental justice, *J. Urban Aff.* **23**, 1–21 (2001).

[37] R. D. Bullard, *Dumping in Dixie: Race, Class, and Environmental quality* (Routledge, Boulder, CO, 1990).

[38] R. Rothstein, *The Color of Law: A Forgotten History of How Our Government Segregated America* (Liveright Publishing, New York, NY, 2017).

[39] A. H. Whittemore, Racial and class bias in zoning: Rezonings involving heavy commercial and industrial land use in Durham (NC), 1945–2014, *J. Am. Plann. Assoc.* **83**, 235–248 (2017).

[40] L. Pulido, *Environment* (Routledge, 2017), pp. 379–407.

[41] L. Pulido, Geographies of race and ethnicity II: Environmental racism, racial capitalism and state-sanctioned violence, *Prog. Hum. Geogr.* **41**, 524–533 (2017).

[42] B. Bryant, P. Mohai, *Race and the incidence of environmental hazards: A time for discourse* (Routledge, 2019).

[43] R. Lopez, Segregation and black/white differences in exposure to air toxics in 1990., *Environ. Health Perspect.* **110**, 289–295 (2002).

[44] R. D. Bullard, Dismantling environmental racism in the USA, *Local Environ.* **4**, 5–19 (1999).

[45] R. Morello-Frosch, The political economy of environmental discrimination, *Environ. Plann. C Gov. Policy* **20**, 477–496 (2002).

[46] I. Mikati, A. F. Benson, T. J. Luben, J. D. Sacks, J. Richmond-Bryant, Disparities in distribution of particulate matter emission sources by race and poverty status, *Am. J. Public Health* **108**, 480–485 (2018).

[47] M. R. Jones, *et al.*, Race/ethnicity, residential segregation, and exposure to ambient air pollution: The Multi-Ethnic Study of Atherosclerosis (MESA), *Am. J. Public Health* **104**, 2130–2137 (2014).

[48] M. P. Thind, C. W. Tessum, I. L. Azevedo, J. D. Marshall, Fine particulate air pollution from electricity generation in the US: Health impacts by race, income, and geography, *Environ. Sci. Technol.* **53**, 14010–14019 (2019).

[49] R. Shah, *et al.*, Socio-economic disparities in exposure to urban restaurant emissions are larger than for traffic, *Environ. Res. Lett.* **15**, 114039 (2020).

[50] D. Houston, J. Wu, P. Ong, A. Winer, Structural disparities of urban traffic in Southern California: Implications for vehicle-related air pollution exposure in minority and high-poverty neighborhoods, *J. Urban Aff.* **26**, 565–592 (2004).

[51] M. Ash, J. K. Boyce, Racial disparities in pollution exposure and employment at US industrial facilities, *Proc. Natl. Acad. Sci. U.S.A* **115**, 10636–10641 (2018).

[52] G. C. Pratt, M. L. Vadali, D. L. Kvale, K. M. Ellickson, Traffic, air pollution, minority and socio-economic status: Addressing inequities in exposure and risk, *Int. J. Environ. Res. Public Health* **12**, 5355–5372 (2015).

[53] J. Chakraborty, Automobiles, air toxics, and adverse health risks: Environmental inequities in Tampa Bay, Florida, *Ann. Assoc. Am. Geogr.* **99**, 674–697 (2009).

[54] S. Gurram, A. L. Stuart, A. R. Pinjari, Agent-based modeling to estimate exposures to urban air pollution from transportation: Exposure disparities and impacts of high-resolution data, *Comput. Environ. Urban Syst.* **75**, 22–34 (2019).

[55] Y. Li, A. Kumar, Y. Li, M. J. Kleeman, Adoption of low-carbon fuels reduces race/ethnicity disparities in air pollution exposure in California, *Sci. Total Environ.* **834**, 155230 (2022).

[56] S. Ji, *et al.*, Environmental justice aspects of exposure to $PM_{2.5}$ emissions from electric vehicle use in China, *Environ. Sci. Technol.* **49**, 13912–13920 (2015).

[57] M. Thind, C. Tessum, M. J. D, Environmental health, racial-ethnic health-disparity, and climate impacts of freight transport in the United States (2022). In review, Environ. Sci. Technol.

[58] R. F. Patterson, R. A. Harley, Effects of freeway rerouting and boulevard replacement on air pollution exposure and neighborhood attributes, *Int. J. Environ. Res. Public Health* **16**, 4072 (2019).

[59] N. P. Nguyen, J. D. Marshall, Impact, efficiency, inequality, and injustice of urban air pollution: variability by emission location, *Environ. Res. Lett.* **13**, 024002 (2018).

[60] K. Do, *et al.*, A data-driven approach for characterizing community scale air pollution exposure disparities in inland southern California, *J. Aerosol Sci.* **152**, 105704 (2021).

[61] C. Mullen, *et al.*, Patterns of distributive environmental inequity under different $PM_{2.5}$ air pollution scenarios for Salt Lake County public schools, *Environ. Res.* **186**, 109543 (2020).

[62] M. T. Chu, *et al.*, Real-time indoor $PM_{2.5}$ monitoring in an urban cohort: Implications for exposure disparities and source control, *Environ. Res.* **193**, 110561 (2021).

[63] Y. Wang, *et al.*, Location-specific strategies for eliminating US national racial-ethnic $PM_{2.5}$ exposure inequality, *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2205548119 (2022).

[64] J. I. Levy, Invited perspective: Moving from characterizing to addressing racial/ethnic disparities in air pollution exposure, *Environ. Health Perspect.* **129**, 121302 (2021).

[65] G. T. Goldman, Invited perspective: Leveraging air quality research advancements to assess and address exposure inequities, *Environ. Health Perspect.* **131**, 031302 (2023).

[66] China Ministry of Ecology and Environment, (MEE), "Ambient air quality standards (GB 3095—2012)" (2016). `https://www.mee.gov.cn/ywgz/fgbz/bz/bzwb/dqhjbh/dqhjzlbz/201203/t20120302_224165.htm`. [In Chinese] Accessed Feb 13, 2022.

[67] C. Song, *et al.*, Air pollution in China: status and spatiotemporal variations, *Environ. Pollut.* **227**, 334–347 (2017).

[68] H. Xu, *et al.*, National $PM_{2.5}$ and NO_2 exposure models for China based on land use regression, satellite measurements, and universal kriging, *Sci. Total Environ.* **655**, 423–433 (2019).

[69] Z. Zhang, D. J. Treiman, Social origins, hukou conversion, and the wellbeing of urban residents in contemporary China, *Soc. Sci. Res.* **42**, 71–89 (2013).

[70] S. Temam, *et al.*, Socioeconomic position and outdoor nitrogen dioxide (NO_2) exposure in Western Europe: A multi-city analysis, *Environ. Int.* **101**, 117–124 (2017).

[71] G. Cesaroni, *et al.*, Socioeconomic position and health status of people who live near busy roads: the Rome Longitudinal Study (RoLS), *Environ. Health* **9**, 1–12 (2010).

[72] S. Havard, S. Deguen, D. Zmirou-Navier, C. Schillinger, D. Bard, Traffic-related air pollution and socioeconomic status: a spatial autocorrelation study to assess environmental equity on a small-area scale, *Epidemiology* pp. 223–230 (2009).

[73] C. M. Padilla, *et al.*, Air quality and social deprivation in four French metropolitan areas—a localized spatio-temporal environmental inequality analysis, *Environ. Res.* **134**, 315–324 (2014).

[74] M. A. Cole, A. J. Rayner, J. M. Bates, The environmental Kuznets curve: an empirical analysis, *Environ. Dev. Econ.* **2**, 401–416 (1997).

[75] D. I. Stern, The rise and fall of the environmental Kuznets curve, *World Dev.* **32**, 1419–1439 (2004).

[76] P. J. Marcotullio, E. Williams, J. D. Marshall, Faster, sooner, and more simultaneously: how recent road and air transportation CO_2 emission trends in developing countries differ from historic trends in the united states, *J. Environ. Dev.* **14**, 125–148 (2005).

[77] J. Kopas, *et al.*, Environmental justice in india: incidence of air pollution from coal-fired power plants, *Ecol. Econ.* **176**, 106711 (2020).

[78] J. He, S. Ding, D. Liu, Exploring the spatiotemporal pattern of $PM_{2.5}$ distribution and its determinants in Chinese cities based on a multilevel analysis approach, *Sci. Total Environ.* **659**, 1513–1525 (2019).

[79] E. D. Schoolman, C. Ma, Migration, class and environmental inequality: Exposure to pollution in China's Jiangsu Province, *Ecol. Econ.* **75**, 140–151 (2012).

[80] H. Zhao, *et al.*, Inequality of household consumption and air pollution-related deaths in China, *Nat. Commun.* **10**, 1–9 (2019).

[81] G. Huang, W. Zhou, Y. Qian, B. Fisher, Breathing the same air? Socioeconomic disparities in $PM_{2.5}$ exposure and the potential benefits from air filtration, *Sci. Total Environ.* **657**, 619–626 (2019).

[82] V. O. Li, Y. Han, J. C. Lam, Y. Zhu, J. Bacon-Shone, Air pollution and environmental injustice: Are the socially deprived exposed to more $PM_{2.5}$ pollution in hong kong?, *Environ. Sci. Policy* **80**, 53–61 (2018).

[83] X. Zhao, *et al.*, Spatial associations between social groups and ozone air pollution exposure in the Beijing urban area, *Environ. Res.* **164**, 173–183 (2018).

[84] W. Xu, *et al.*, Spatiotemporal variation and socioeconomic drivers of air pollution in China during 2005–2016, *J. Environ. Manage.* **245**, 66–75 (2019).

[85] F. Shen, *et al.*, Temporal variations of six ambient criteria air pollutants from 2015 to 2018, their spatial distributions, health risks and relationships with socioeconomic factors during 2018 in China, *Environ. Int.* **137**, 105556 (2020).

[86] Y. Wang, S. Komonpipat, Revisiting the environmental Kuznets curve of $PM_{2.5}$ concentration: evidence from prefecture-level and above cities of China, *Environ. Sci. Pollut. Res. Int.* **27**, 9336–9348 (2020).

[87] R. Beelen, *et al.*, Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project, *Lancet* **383**, 785–795 (2014).

[88] J. D. Kaufman, *et al.*, Association between air pollution and coronary artery calcification within six metropolitan areas in the USA (the Multi-Ethnic Study of Atherosclerosis and Air Pollution): a longitudinal cohort study, *Lancet* **388**, 696–704 (2016).

[89] F. Lu, *et al.*, Systematic review and meta-analysis of the adverse health effects of ambient $PM_{2.5}$ and PM_{10} pollution in the Chinese population, *Environ. Res.* **136**, 196–204 (2015).

[90] T. Lu, *et al.*, National empirical models of air pollution using microscale measures of the urban environment, *Environ. Sci. Technol.* **55**, 15519–15530 (2021).

[91] K. de Hoogh, *et al.*, Development of West-European $PM_{2.5}$ and NO_2 land use regression models incorporating satellite-derived and chemical transport modelling data, *Environ. Res.* **151**, 1–10 (2016).

[92] M. Wang, *et al.*, Combining land-use regression and chemical transport modeling in a spatiotemporal geostatistical model for ozone and $PM_{2.5}$, *Environ. Sci. Technol.* **50**, 5111–5118 (2016).

[93] S.-Y. Kim, *et al.*, Concentrations of criteria pollutants in the contiguous US, 1979–2015: role of prediction model parsimony in integrated empirical geographic regression, *PloS ONE* **15**, e0228535 (2020).

[94] G. Hoek, Methods for assessing long-term exposures to outdoor air pollutants, *Curr. Environ. Health Rep.* **4**, 450–462 (2017).

[95] E. S. Robinson, *et al.*, Land-use regression modeling of source-resolved fine particulate matter components from mobile sampling, *Environ. Sci. Technol.* **53**, 8925–8937 (2019).

[96] M. Hatzopoulou, *et al.*, Robustness of land-use regression models developed from mobile air pollutant measurements, *Environ. Sci. Technol.* **51**, 3938–3947 (2017).

[97] S. Hankey, J. D. Marshall, Land use regression models of on-road particulate air pollution (particle number, black carbon, $PM_{2.5}$, particle size) using mobile monitoring, *Environ. Sci. Technol.* **49**, 9194–9202 (2015).

[98] E. V. Novotny, M. J. Bechle, D. B. Millet, J. D. Marshall, National satellite-based land-use regression: NO_2 in the United States, *Environ. Sci. Technol.* **45**, 4407–4414 (2011).

[99] L. D. Knibbs, M. G. Hewson, M. J. Bechle, J. D. Marshall, A. G. Barnett, A national satellite-based land-use regression model for air pollution exposure assessment in Australia, *Environ. Res.* **135**, 204–211 (2014).

[100] D. Vienneau, *et al.*, Western European land use regression incorporating satellite-and ground-based measurements of NO_2 and PM_{10}, *Environ. Sci. Technol.* **47**, 13555–13564 (2013).

[101] P. Hystad, *et al.*, Creating national air pollution models for population exposure assessment in Canada, *Environ. Health Perspect.* **119**, 1123–1129 (2011).

[102] Q. He, B. Huang, Satellite-based mapping of daily high-resolution ground $PM_{2.5}$ in China via space-time regression modeling, *Remote Sens. Environ.* **206**, 72–83 (2018).

[103] A. Larkin, *et al.*, Global land use regression model for nitrogen dioxide air pollution, *Environ. Sci. Technol.* **51**, 6957–6964 (2017).

[104] Ramboll, "User's guide Comprehensive Air Quality Model with Extensions version 7.20" (2022). `https://camx-wp.azurewebsites.net/Files/CAMxUsersGuide_v7.20.pdf`. Accessed Apr 10, 2023.

[105] K. W. Appel, *et al.*, *Air Pollution Modeling and its Application XXV 35* (Springer, 2018), pp. 69–73.

[106] G. A. Grell, *et al.*, Fully coupled "online" chemistry within the wrf model, *Atmos. Environ.* **39**, 6957–6975 (2005).

[107] I. Bey, *et al.*, Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, *J. Geophys. Res. Atmos.* **106**, 23073–23095 (2001).

[108] E. A. Gilmore, *et al.*, An inter-comparison of the social costs of air quality from reduced-complexity models, *Environ. Res. Lett.* **14**, 074016 (2019).

[109] C. W. Tessum, J. D. Hill, J. D. Marshall, InMAP: A model for air pollution interventions, *PloS ONE* **12**, e0176131 (2017).

[110] N. Z. Muller, R. Mendelsohn, The Air Pollution Emission Experiments and Policy Analysis Model (APEEP) Technical Appendix, *Yale University: New Haven, CT, USA* **1** (2006).

[111] N. Z. Muller, Boosting GDP growth by accounting for the environment, *Science* **345**, 873–874 (2014).

[112] B. Sergi, I. Azevedo, S. J. Davis, N. Z. Muller, Regional and county flows of particulate matter damage in the us, *Environ. Res. Lett.* **15**, 104073 (2020).

[113] J. Heo, P. J. Adams, EASIUR User's Guide Version 0.2, *Atmosphere* **148**, 112 (2015).

[114] L. R. Henneman, C. Choirat, C. Ivey, K. Cummiskey, C. M. Zigler, Characterizing population exposure to coal emissions sources in the United States using the HyADS model, *Atmos. Environ.* **203**, 271–280 (2019).

[115] N. Fann, K. R. Baker, C. M. Fulcher, Characterizing the $PM_{2.5}$-related health benefits of emission reductions for 17 industrial, area and mobile emission sectors across the US, *Environ. Int.* **49**, 141–151 (2012).

[116] M. Amann, *et al.*, Cost-effective control of air quality and greenhouse gases in Europe: Modeling and policy applications, *Environ. Model. Softw.* **26**, 1489–1501 (2011).

[117] Y. Xie, H. Dai, H. Dong, T. Hanaoka, T. Masui, Economic impacts from $PM_{2.5}$ pollution-related health effects in China: a provincial-level analysis, *Environ. Sci. Technol.* **50**, 4836–4843 (2016).

[118] P. Thunis, *et al.*, $PM_{2.5}$ source allocation in European cities: A SHERPA modelling study, *Atmos. Environ.* **187**, 93–106 (2018).

[119] F. Zhang, *et al.*, Estimation of abatement potentials and costs of air pollution emissions in China, *J. Environ. Manage.* **260**, 110069 (2020).

[120] M. Crippa, G. Janssens-Maenhout, D. Guizzardi, R. Van Dingenen, F. Dentener, Contribution and uncertainty of sectorial and regional emissions to regional and global $PM_{2.5}$ health impacts, *Atmos. Chem. Phys.* **19**, 5165–5186 (2019).

[121] A. L. Goodkind, C. W. Tessum, J. S. Coggins, J. D. Hill, J. D. Marshall, Fine-scale damage estimates of particulate matter air pollution reveal opportunities for location-specific mitigation of emissions, *Proc. Natl. Acad. Sci. U.S.A* **116**, 8775–8780 (2019).

[122] A. L. Goodkind, C. W. Tessum, J. Coggings, J. Hill, J. Marshall, "InMAP source-receptor matrix (ISRM) dataset [Dataset]" (2019). `https://zenodo.org/record/2589760`. Accessed Mar 15, 2023.

[123] J. Guo, *et al.*, Impact of various emission control schemes on air quality using WRF-Chem during APEC China 2014, *Atmos. Environ.* **140**, 311–319 (2016).

[124] J. Sciare, *et al.*, Comparison between simulated and observed chemical composition of fine aerosols in Paris (France) during springtime: contribution of regional versus continental emissions, *Atmos. Chem. Phys.* **10**, 11987–12004 (2010).

[125] D. Wu, J. C. H. Fung, T. Yao, A. K. H. Lau, A study of control policy in the Pearl River Delta region by using the particulate matter source apportionment method, *Atmos. Environ.* **76**, 147–161 (2013).

[126] M. Diamantopoulou, K. Skyllakou, S. N. Pandis, Estimation of the local and long-range contributions to particulate matter levels using continuous measurements in a single urban background site, *Atmos. Environ.* **134**, 1–9 (2016).

[127] M. Viana, *et al.*, Source apportionment of particulate matter in Europe: a review of methods and results, *J. Aerosol Sci.* **39**, 827–849 (2008).

[128] P. Thunis, On the validity of the incremental approach to estimate the impact of cities on air quality, *Atmos. Environ.* **173**, 210–222 (2018).

[129] X. Chang, *et al.*, Contributions of inter-city and regional transport to $PM_{2.5}$ concentrations in the Beijing-Tianjin-Hebei region and its implications on regional joint air pollution control, *Sci. Total Environ.* **660**, 1191–1200 (2019).

[130] D. Chen, *et al.*, Estimating the contribution of regional transport to $PM_{2.5}$ air pollution in a rural area on the North China Plain, *Sci. Total Environ.* **583**, 280–291 (2017).

[131] K. M. Wagstrom, S. N. Pandis, G. Yarwood, G. M. Wilson, R. E. Morris, Development and application of a computationally efficient particulate matter apportionment algorithm in a three-dimensional chemical transport model, *Atmos. Environ.* **42**, 5650–5659 (2008).

[132] K. Wagstrom, S. Pandis, Contribution of long range transport to local fine particulate matter concerns, *Atmos. Environ.* **45**, 2730–2735 (2011).

[133] G. Kiesewetter, *et al.*, Modelling NO_2 concentrations at the street level in the GAINS integrated assessment model: projections under current legislation, *Atmos. Chem. Phys.* **14**, 813–829 (2014).

[134] A. Clappier, C. A. Belis, D. Pernigotti, P. Thunis, Source apportionment and sensitivity analysis: two methodologies with two different purposes, *Geosci. Model Dev.* **10**, 4245–4256 (2017).

[135] P. Thunis, *et al.*, Source apportionment to support air quality planning: Strengths and weaknesses of existing approaches, *Environ. Int.* **130**, 104825 (2019).

[136] B. Zhao, *et al.*, A modeling study of the nonlinear response of fine particles to air pollutant emissions in the Beijing–Tianjin–Hebei region, *Atmos. Chem. Phys.* **17**, 12031–12050 (2017).

[137] P. Thunis, *et al.*, Overview of current regional and local scale air quality modelling practices: Assessment and planning tools in the EU, *Environ. Sci. Pol.* **65**, 13–21 (2016).

[138] J. S. Apte, *et al.*, High-resolution air pollution mapping with Google street view cars: exploiting big data, *Environ. Sci. Technol.* **51**, 6999–7008 (2017).

[139] A. F. Both, A. Balakrishnan, B. Joseph, J. D. Marshall, Spatiotemporal aspects of real-time $PM_{2.5}$: low-and middle-income neighborhoods in Bangalore, India, *Environ. Sci. Technol.* **45**, 5629–5636 (2011).

[140] M. Beekmann, *et al.*, In situ, satellite measurement and model evidence on the dominant regional contribution to fine particulate matter levels in the Paris megacity, *Atmos. Chem. Phys.* **15**, 9577–9591 (2015).

[141] J. Cyrys, M. Pitz, J. Heinrich, H.-E. Wichmann, A. Peters, Spatial and temporal variation of particle number concentration in Augsburg, Germany, *Sci. Total Environ.* **401**, 168–175 (2008).

[142] Á. Gómez-Losada, J. C. M. Pires, R. Pino-Mejías, Characterization of background air pollution exposure in urban environments using a metric based on Hidden Markov Models, *Atmos. Environ.* **127**, 255–261 (2016).

[143] P. Lenschow, *et al.*, Some ideas about the sources of PM_{10}, *Atmos. Environ.* **35**, S23–S33 (2001).

[144] X. Querol, *et al.*, Speciation and origin of PM_{10} and $PM_{2.5}$ in selected European cities, *Atmos. Environ.* **38**, 6547–6555 (2004).

[145] S. Squizzato, *et al.*, A procedure to assess local and long-range transport contributions to $PM_{2.5}$ and secondary inorganic aerosol, *J. Aerosol Sci.* **46**, 64–76 (2012).

[146] J. G. Watson, J. C. Chow, Estimating middle-, neighborhood-, and urban-scale contributions to elemental carbon in Mexico City with a rapid response aethalometer, *J. Air Waste Manag. Assoc.* **51**, 1522–1528 (2001).

[147] J. Antonelli, J. Schwartz, I. Kloog, B. A. Coull, Spatial multiresolution analysis of the effect of $PM_{2.5}$ on birth weights, *Ann. Appl. Stat.* **11**, 792 (2017).

[148] R. Beelen, *et al.*, Mapping of background air pollution at a fine spatial scale across the European Union, *Sci. Total Environ.* **407**, 1852–1867 (2009).

[149] P. Kumar, *et al.*, Ultrafine particles in cities, *Environ. Int.* **66**, 1–10 (2014).

[150] M. Lin, J. Hang, Y. Li, Z. Luo, M. Sandberg, Quantitative ventilation assessments of idealized urban canopy layers with various urban layouts and the same building packing density, *Build. Environ.* **79**, 152–167 (2014).

[151] S. Manson, J. Schroeder, D. Van Riper, S. Ruggles, IPUMS National Historical Geographic Information System: Version 14.0 [Database]. Minneapolis, MN: IPUMS. 2019 (2019).

[152] R. A. Engel, M. E. Marlier, D. P. Lettenmaier, On the causes of the summer 2015 Eastern Washington wildfires, *Environ. Res. Commun.* **1**, 011009 (2019).

[153] M. Eeftens, *et al.*, Spatial and temporal variability of ultrafine particles, NO_2, $PM_{2.5}$, $PM_{2.5}$ absorbance, PM_{10} and PMcoarse in Swiss study areas, *Atmos. Environ.* **111**, 60–70 (2015).

[154] C. N. Hewitt, Spatial variations in nitrogen dioxide concentrations in an urban area, *Atmos. Environ. Part B - Urban Atmos.* **25**, 429–434 (1991).

[155] L. Wang, Z. Liu, Y. Sun, D. Ji, Y. Wang, Long-range transport and regional sources of $PM_{2.5}$ in Beijing based on long-term observations from 2005 to 2010, *Atmos. Res.* **157**, 37–48 (2015).

[156] M. L. Bell, K. Ebisu, Environmental inequality in exposures to airborne particulate matter components in the United States, *Environ. Health Perspect.* **120**, 1699–1704 (2012).

[157] M. L. Miranda, S. E. Edwards, M. H. Keating, C. J. Paul, Making the environmental justice grade: the relative burden of air pollution exposure in the United States, *Int. J. Environ. Res. Public Health* **8**, 1755–1771 (2011).

[158] R. D. Bullard, G. S. Johnson, Environmental justice: Grassroots activism and its impact on public policy decision making, *J. Soc. Issues* **56**, 555–578 (2009).

[159] W. Liang, M. Yang, Urbanization, economic growth and environmental pollution: Evidence from China, *Sustain. Comput.: Inform. Syst.* **21**, 1–9 (2019).

[160] K. H. Zhang, S. Shunfeng, Rural–urban migration and urbanization in China: Evidence from time-series and cross-section analyses, *China Econ. Rev.* **14**, 386–400 (2003).

[161] S. Li, H. Sato, T. Sicular, *Rising inequality in China: Challenges to a harmonious society* (Cambridge University Press, 2013).

[162] Y. Xie, X. Zhou, Income inequality in today's China, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 6928–6933 (2014).

[163] K. W. Chan, L. Zhang, The hukou system and rural-urban migration in China: Processes and changes, *China Q.* **160**, 818–855 (1999).

[164] J. Chen, C. Zhou, S. Wang, J. Hu, Identifying the socioeconomic determinants of population exposure to particulate matter ($PM_{2.5}$) in China using geographically weighted regression modeling, *Environ. Pollut.* **241**, 494–503 (2018).

[165] T. Li, *et al.*, Mining of the association rules between industrialization level and air quality to inform high-quality development in China, *J. Environ. Manage.* **246**, 564–574 (2019).

[166] X. Liu, M. Wang, How polycentric is urban China and why? A case study of 318 cities, *Landsc. Urban Plan.* **151**, 10–20 (2016).

[167] F. Wu, J. Xu, A. G.-O. Yeh, *Urban development in post-reform China: state, market, and space* (Routledge, 2006).

[168] H. Chen, B. Jia, S. Lau, Sustainable urban form for Chinese compact cities: Challenges of a rapid urbanized economy, *Habitat Int.* **32**, 28–40 (2008).

[169] S. Zheng, M. E. Kahn, Land and residential property markets in a booming economy: New evidence from Beijing, *J. Urban Econ.* **63**, 743–757 (2008).

[170] W. Wu, Migrant housing in urban China: choices and constraints, *Urban Aff. Rev.* **38**, 90–119 (2002).

[171] L. Xiao, Q. Qiu, L. Gao, Chinese housing reform and social sustainability: Evidence from post-reform home ownership, *Sustainability* **8**, 1053 (2016).

[172] Y. Zhao, Y. Hu, J. P. Smith, J. Strauss, G. Yang, Cohort profile: the China health and retirement longitudinal study (CHARLS), *Int. J. Epidemiol.* **43**, 61–68 (2014).

[173] C. Liu, *et al.*, Associations between ambient fine particulate air pollution and hypertension: a nationwide cross-sectional study in China, *Sci. Total Environ.* **584**, 869–874 (2017).

[174] National Bureau of Statistics of China, (NBS), "communiqué of the seventh national population census (no. 2)" (NBS, 2021). `http://www.stats.gov.cn/english/PressRelease/202105/t20210510_1817187.html`. Accessed Feb 13, 2022.

[175] China Health and Retirement Longitudinal Study, (CHARLS), "2013 charls wave2 (baseline) [data file and code book]" (2015). `https://charls.charlsdata.com/pages/Data/2013-charls-wave2/en.html`. Accessed Nov 4, 2017.

[176] F. Afridi, S. X. Li, Y. Ren, Social identity and inequality: The impact of China's hukou system, *J. Public Econ.* **123**, 17–29 (2015).

[177] Z. Liu, Institution and inequality: the hukou system in China, *J. Comp. Econ.* **33**, 133–157 (2005).

[178] J. Whalley, S. Zhang, A numerical simulation analysis of (hukou) labour mobility restrictions in China, *J. Dev. Econ.* **83**, 392–410 (2007).

[179] X. Wu, D. J. Treiman, Inequality and equality under Chinese socialism: The hukou system and intergenerational occupational mobility, *Am. J. Sociol.* **113**, 415–445 (2007).

[180] Y. Zhao, Labor migration and earnings differences: The case of rural China, *Econ. Dev. Cult. Change* **47**, 767–782 (1999).

[181] China Health and Retirement Longitudinal Study, (CHARLS), "2011 charls wave1 (baseline) [data file and code book]" (2013). https://charls.charlsdata.com/pages/Data/2011-charls-wave1/en.html. Accessed Nov 4, 2017.

[182] S. Kolenikov, G. Angeles, *et al.*, The use of discrete data in PCA: theory, simulations, and applications to socioeconomic indices, *Chapel Hill: Carolina Population Center, University of North Carolina* **20**, 1–59 (2004).

[183] S. Vyas, L. Kumaranayake, Constructing socio-economic status indices: how to use principal components analysis, *Health Policy Plan.* **21**, 459–468 (2006).

[184] N. Zhao, Y. Liu, G. Cao, E. L. Samson, J. Zhang, Forecasting China's GDP at the pixel level using nighttime lights time series and population images, *GIsci. Remote Sens.* **54**, 407–425 (2017).

[185] F. R. Stevens, A. E. Gaughan, C. Linard, A. J. Tatem, Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data, *PloS ONE* **10**, e0107042 (2015).

[186] N. J. Horton, S. R. Lipsitz, Multiple imputation in practice: comparison of software packages for regression models with missing variables, *Am. Stat.* **55**, 244–254 (2001).

[187] Y. Long, Redefining Chinese city system with emerging new data, *Appl. Geogr.* **75**, 36–48 (2016).

[188] K. Wu, Y. Long, Q. Mao, X. Liu, Featured graphic. mushrooming jiedaos, growing cities: an alternative perspective on urbanizing China, *Environ. Plan. A* **47**, 1–2 (2015).

[189] Y. Long, Y. Shen, X. Jin, Mapping block-level urban areas for all Chinese cities, *Ann. Assoc. Am. Geogr.* **106**, 96–113 (2016).

[190] S. Bontemps, *et al.*, GLOBCOVER 2009 products description and validation report, *ESA Bull* **136**, 10013 (2011).

[191] Beijing City Lab, "2data 22: Urban areas of China in 2012 (by various methods)" (2015). http://www.beijingcitylab.com. Accessed Oct 8, 2020.

[192] D. Norbäck, *et al.*, Asthma and rhinitis among chinese children—indoor and outdoor air pollution and indicators of socioeconomic status (ses), *Environ. Int.* **115**, 1–8 (2018).

[193] X. Zhao, W. Zhou, L. Han, D. Locke, Spatiotemporal variation in $PM_{2.5}$ concentrations and their relationship with socioeconomic factors in China's major cities, *Environ. Int.* **133**, 105145 (2019).

[194] Y. Wang, *et al.*, Spatial decomposition analysis of NO_2 and $PM_{2.5}$ air pollution in the united states, *Atmos. Environ.* **241**, 117470 (2020).

[195] J. G. Su, *et al.*, An index for assessing demographic inequalities in cumulative environmental hazards with application to Los Angeles, California, *Environ. Sci. Technol.* **43**, 7626–7634 (2009-15).

[196] S. Wang, S. Gao, S. Li, K. Feng, Strategizing the relation between urbanization and air pollution: Empirical evidence from global countries, *J. Clean. Prod.* **243**, 118615 (2020).

[197] S. Lerner, *Sacrifice zones: The front lines of chemical exposure in the United States* (Cambridge (MA): The MIT Press, 2010).

[198] V. Castán Broto, M. Sanzana Calvet, Sacrifice zones and the construction of urban energy landscapes in concepción, Chile, *J. Political Ecol.* **27**, 279–299 (2020).

[199] K. Valenzuela-Fuentes, E. Alarcón-Barrueto, R. Torres-Salinas, From resistance to creation: Socio-environmental activism in Chile's "sacrifice zones", *Sustainability* **13**, 3481 (2021).

[200] A. A. Baccarelli, *et al.*, Air pollution exposure and lung function in highly exposed subjects in Beijing, China: a repeated-measure study, *Part. Fibre. Toxicol.* **11**, 1–10 (2014).

[201] J. Baumgartner, *et al.*, Patterns and predictors of personal exposure to indoor air pollution from biomass combustion among women and children in rural China, *Indoor Air* **21**, 479–488 (2011).

[202] X. Du, Q. Kong, W. Ge, S. Zhang, L. Fu, Characterization of personal exposure concentration of fine particles for adults and children exposed to high ambient concentrations in Beijing, China, *J. Environ. Sci.* **22**, 1757–1764 (2010).

[203] W. Du, X. Li, Y. Chen, G. Shen, Household air pollution and personal exposure to air pollutants in rural China–a review, *Environ. Pollut.* **237**, 625–638 (2018).

[204] C. Sun, M. E. Kahn, S. Zheng, Self-protection investment exacerbates air pollution exposure inequality in urban China, *Ecol. Econ.* **131**, 468–474 (2017).

[205] S. A. Venners, *et al.*, Indoor air pollution and respiratory health in urban and rural China, *Int. J. Occup. Environ. Health* **7**, 173–181 (2001).

[206] X. Xu, *et al.*, Occupational and environmental risk factors for asthma in rural communities in China, *Int. J. Occup. Environ. Health* **2**, 172–176 (1996).

[207] Y. Zhao, S. Wang, K. Aunan, H. Martin Seip, J. Hao, Air pollution and lung cancer risks in China—a meta-analysis, *Sci. Total Environ.* **366**, 500–513 (2006).

[208] L. Zhang, D. Enarson, G. He, B. Li, M. Chan-Yeung, Occupational and environmental risk factors for respiratory symptoms in rural Beijing, China, *Eur. Respir. J.* **20**, 1525–1531 (2002).

[209] J. D. Marshall, W. J. Riley, T. E. McKone, W. W. Nazaroff, Intake fraction of primary pollutants: motor vehicle emissions in the South Coast Air Basin, *Atmos. Environ.* **37**, 3455–3468 (2003).

[210] J. D. Marshall, *et al.*, Inhalation intake of ambient air pollution in California's South Coast Air Basin, *Atmos. Environ.* **40**, 4381–4392 (2006).

[211] P. Pant, G. Habib, J. D. Marshall, R. E. Peltier, $PM_{2.5}$ exposure in highly polluted cities: A case study from New Delhi, India, *Environ. Res.* **156**, 167–174 (2017).

[212] C. Milà, *et al.*, When, where, and what? Characterizing personal $PM_{2.5}$ exposure in periurban india by integrating GPS, wearable camera, and ambient and personal monitoring data, *Environ. Sci. Technol.* **52**, 13481–13490 (2018).

[213] K. H. Chan, *et al.*, Trans-generational changes and rural-urban inequality in household fuel use and cookstove ventilation in China: a multi-region study of 0.5 million adults, *Int. J. Hyg. Environ. Health* **220**, 1370–1381 (2017).

[214] X. Duan, *et al.*, Household fuel use for cooking and heating in China: results from the first Chinese Environmental Exposure-Related Human Activity Patterns Survey (CEERHAPS), *Appl. Energy* **136**, 692–703 (2014).

[215] G. Geng, *et al.*, Drivers of $PM_{2.5}$ air pollution deaths in China 2002–2017, *Nat. Geosci.* **14**, 645–650 (2021).

[216] N. D. Rao, G. Kiesewetter, J. Min, S. Pachauri, F. Wagner, Household contributions to and impacts from air pollution in India, *Nat. Sustain.* **4**, 859–867 (2021).

[217] Z. Zhang, G. Zhang, B. Su, The spatial impacts of air pollution and socio-economic status on public health: Empirical evidence from China, *Socioecon. Plann. Sci.* **83**, 101167 (2022).

[218] U.S. Environmental Protection Agency, (EPA), "Benefits and costs of the clean air act 1990–2020, the second prospective study" (EPA, 2011). `https://www.epa.gov/clean-air-act-overview/benefits-and-costs-clean-air-act-1990-2020-second-prospective-study`. Accessed Jan 15, 2022.

[219] G. Torres, Introduction: Understanding environmental racism, *U. Colo. L. Rev.* **63**, 839 (1992).

[220] C. Dhillon, M. G. Young, Environmental racism and First Nations: A call for socially just public policy development, *Int. J. Humanit. Soc. Sci.* **1**, 25–39 (2010).

[221] M. J. Brady, S. Monani, Wind power! marketing renewable energy on tribal lands and the struggle for just sustainability, *Local Environ.* **17**, 147–166 (2012).

[222] Y. Wang, Y. Wang, H. Xu, Y. Zhao, J. D. Marshall, Ambient air pollution and socioeconomic status in China, *Environ. Health Perspect.* **130**, 067001 (2022).

[223] U.S. Environmental Protection Agency, (EPA), "Air pollutant emissions trends data" (EPA, 2021). `https://www.epa.gov/air-emissions-inventories/air-pollutant-emissions-trends-data`. Accessed Jan 15, 2022.

[224] B. A. Lee, *et al.*, Beyond the census tract: Patterns and determinants of racial segregation at multiple geographic scales, *Am. Sociol. Rev.* **73**, 766–791 (2008).

[225] S. F. Reardon, *et al.*, The geographic scale of metropolitan racial segregation, *Demography* **45**, 489–514 (2008).

[226] M. A. G. Demetillo, *et al.*, Space-based observational constraints on NO_2 air pollution inequality from diesel traffic in major US cities, *Geophys. Res. Lett.* **48**, e2021GL094333 (2021).

[227] G. Adamkiewicz, *et al.*, Moving environmental justice indoors: Understanding structural influences on residential exposure patterns in low-income communities, *Am. J. Public Health* **101**, S238–S245 (2011).

[228] A. Rosofsky, J. I. Levy, M. S. Breen, A. Zanobetti, M. P. Fabian, The impact of air exchange rate on ambient air pollution exposure and inequalities across all residential parcels in Massachusetts, *J. Expo. Sci. Environ. Epidemiol.* **29**, 520–530 (2019).

[229] Y. M. Park, M. P. Kwan, Individual exposure estimates may be erroneous when spatiotemporal variability of air pollution and human mobility are ignored, *Health Place* **43**, 85–94 (2017).

[230] V. A. Southerland, *et al.*, Assessing the distribution of air pollution health risks within cities: a neighborhood-scale analysis leveraging high-resolution data sets in the bay area, California, *Environ. Health Perspect.* **129**, 037006 (2021).

[231] M. H. Forouzanfar, *et al.*, Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015, *Lancet* **388**, 1659–1724 (2016).

[232] Executive Office of the President, "Tackling the Climate Crisis at Home and Abroad" (2021). `https://www.federalregister.gov/documents/2021/02/01/2021-02177/tackling-the-climate-crisis-at-home-and-abroad`. Accessed Jan 15, 2022.

[233] U.S. Environmental Protection Agency, (EPA), "FY 2022-2026 EPA Strategic Plan Draft" (EPA, 2021). `www.epa.gov/system/files/documents/2021-10/fy-2022-2026-epa-draft-strategic-plan.pdf`. Accessed Jan 15, 2022.

[234] Office of Washington Governor, "Inslee signs climate change legislative package" (2021). https://www.governor.wa.gov/news-media/inslee-signs-climate-change-legislative-package. Accessed Jan 15, 2022.

[235] C. Garcia, "Assembly Bill No. 617: Nonvehicular air pollution: criteria air pollutants and toxic air contaminants" (California Office of Legislative Counsel, 2021). https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201720180AB617. Accessed Jan 15, 2022.

[236] State of New Jersey, "NJ environmental justice law and rules" (2020). https://www.nj.gov/dep/ej/policy.html. Accessed Jul 13, 2022.

[237] National Environmental Justice Advisory Council, (NEJAC), "20 Year Retrospective Report (1994 - 2014)" (NEJAC, 2016). https://www.epa.gov/sites/default/files/2019-10/documents/nejac_20_year_retrospective_report.pdf. Accessed Jan 15, 2022.

[238] W. H. Hansell, E. Hollander, D. John, "Putting community first: A promising approach to federal collaboration for environmental improvement" (National Academy of Public Administration, 2009; https://napawash.org/uploads/Academy_Studies/05-09.pdf).

[239] D. N. Pellow, R. J. Brulle, *Power, Justice, and the Environment: A Critical Appraisal of the Environmental Justice Movement* (The MIT Press, Cambridge, MA, 2005).

[240] L. Cushing, *et al.*, Carbon trading, co-pollutants, and environmental equity: Evidence from California's cap-and-trade program (2011–2015), *PLoS Med.* **15**, e1002604 (2018).

[241] C. M. Anderson, K. A. Kissel, C. B. Field, K. J. Mach, Climate change mitigation, air pollution, and environmental justice in California, *Environ. Sci. Technol.* **52**, 10829–10838 (2018).

[242] D. Hernandez-Cortes, K. C. Meng, Do environmental markets cause environmental injustice? Evidence from California's carbon market (2020). NBER Working Paper No. 27205.

[243] M. A. Bravo, *et al.*, Where is air quality improving, and who benefits? A study of $PM_{2.5}$ and ozone over 15 years, *Am. J. Epidemiol.* **191**, 1258–1269 (2022).

[244] Y. Wang, "Crosswalk between Source Classification Codes (SCCs) and 14 sectors of economy [Dataset]" (2023). https://zenodo.org/record/7739032. Accessed Mar 15, 2023.

[245] Council on Environmental Quality (CEQ), "Climate and Economic Justice Screening Tool, version 1.0" (2022). https://screeningtool.geoplatform.gov. Accessed Dec 27, 2022.

[246] C. A. Pope III, N. Coleman, Z. A. Pond, R. T. Burnett, Fine particulate air pollution and human mortality: 25+ years of cohort studies, *Environ. Res.* **183**, 108924 (2020).

[247] K. D. Kochanek, S. L. Murphy, J. Xu, B. Tejada-Vera, Deaths: final data for 2014, *Natl. Vital Stat. Rep.* pp. 1–122 (2016).

[248] B. M. Gentry, A. L. Robinson, P. J. Adams, EASIUR-HR: A model to evaluate exposure inequality caused by ground-level sources of primary fine particulate matter, *Environ. Sci. Technol.* (2023).

[249] L. R. Henneman, M. M. Rasel, C. Choirat, S. C. Anenberg, C. Zigler, Inequitable exposures to US coal power plant–related $PM_{2.5}$: 22 years and counting, *Environ. Health Perspect.* **131**, 037005 (2023).

[250] J. S. Lefler, *et al.*, Air pollution and mortality in a large, representative us cohort: multiple-pollutant analyses, and spatial and temporal decompositions, *Environ. Health* **18**, 1–11 (2019).

[251] J. Liu, J. D. Marshall, Spatial decomposition of air pollution concentrations highlights historical causes for current exposure disparities in the United States, *Environ. Sci. Technol. Lett.* (2023).

[252] P. Knobel, *et al.*, Socioeconomic and racial disparities in source-apportioned $PM_{2.5}$ levels across urban areas in the contiguous US, 2010, *Atmos. Environ.* p. 119753 (2023).

[253] M. Burke, *et al.*, The changing risk and burden of wildfire in the United States, *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2011048118 (2021).

[254] R. Aguilera, T. Corringham, A. Gershunov, T. Benmarhnia, Wildfire smoke impacts respiratory health more than fine particles from other sources: Observational evidence from Southern California, *Nat. Commun.* **12**, 1493 (2021).

[255] J. D. Stowell, *et al.*, Associations of wildfire smoke $PM_{2.5}$ exposure with cardiorespiratory events in Colorado 2011–2014, *Environ. Int.* **133**, 105151 (2019).

[256] L. J. Cushing, S. Li, B. B. Steiger, J. A. Casey, Historical red-lining is associated with fossil fuel power plant siting and present-day inequalities in air pollutant emissions, *Nat. Energy* **8**, 52–61 (2023).

[257] X. Qiu, *et al.*, Deriving high-resolution emission inventory of open biomass burning in China based on satellite observations, *Environ. Sci. Technol.* **50**, 11779–11786 (2016).

[258] Y. Gao, *et al.*, Unveiling the spatial and sectoral characteristics of a high-resolution emission inventory of CO_2 and air pollutants in China, *Sci. Total Environ.* **847**, 157623 (2022).

[259] R. Wu, *et al.*, Reduced-complexity air quality intervention modeling over China: the development of InMAPv1.6.1-China and a comparison with CMAQv5. 2, *Geosci. Model Dev.* **14**, 7621–7638 (2021).

[260] L. P. Clark, *et al.*, Developing a low-cost passive method for long-term average levels of light-absorbing carbon air pollution in polluted indoor environments, *Sensors* **20**, 3417 (2020).

[261] L. Morawska, *et al.*, Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: How far have they gone?, *Environ. Int.* **116**, 286–299 (2018).

[262] M. Nyarku, *et al.*, Mobile phones as monitors of personal exposure to air pollution: Is this the future?, *PLoS One* **13**, e0193150 (2018).

[263] M. Shupler, *et al.*, Household and personal air pollution exposure measurements from 120 communities in eight countries: results from the PURE-AIR study, *Lancet Planet. Health* **4**, e451–e462 (2020).

[264] Y. M. Park, *et al.*, Personal exposure monitoring using GPS-enabled portable air pollution sensors: A strategy to promote citizen awareness and behavioral changes regarding indoor and outdoor air pollution, *J. Expo. Sci. Environ. Epidemiol.* pp. 1–11 (2022).