HEI
Health Effects Institute

NUMBER 216
JANUARY 2024

Walter A. Rosenblith New Investigator Award

RESEARCH REPORT

Scalable Multipollutant Exposure Assessment Using Routine Mobile Monitoring Platforms

Joshua S. Apte, Sarah E. Chambliss, Kyle P. Messier, Shahzad Gani, Adithi R. Upadhya, Meenakshi Kushwaha, and V. Sreekanth

INCLUDES A COMMENTARY BY THE INSTITUTE'S REVIEW COMMITTEE

healtheffects.org

Scalable Multipollutant Exposure Assessment Using Routine Mobile Monitoring Platforms

Joshua S. Apte, Sarah E. Chambliss, Kyle P. Messier, Shahzad Gani, Adithi R. Upadhya, Meenakshi Kushwaha, and V. Sreekanth

---

with a Commentary by the HEI Review Committee

Research Report 216
Health Effects Institute
Boston, Massachusetts

Trusted Science · Cleaner Air · Better Health

Publishing history: This document was posted at *www.healtheffects.org* in January 2024.

Citation for report:

Apte JS, Chambliss SE, Messier KP, Gani S, Upadhya AR, Kushwaha M, et al. 2024. Scalable Multipollutant Exposure Assessment Using Routine Mobile Monitoring Platforms. Research Report 216. Boston, MA:Health Effects Institute.

 Library of Congress Catalog Number for the HEI Report Series: WA 754 R432.

ISSN 2688-6855 (online)

CONTENTS

ABOUT HEI

The Health Effects Institute is a nonprofit corporation chartered in 1980 as an independent research organization to provide high-quality, impartial, and relevant science on the effects of air pollution on health. To accomplish its mission, the Institute

- Identifies the highest-priority areas for health effects research
- Competitively funds and oversees research projects
- Provides an intensive independent review of HEI-supported studies and related research
- Integrates HEI's research results with those of other institutions into broader evaluations
- Communicates the results of HEI's research and analyses to public and private decision-makers.

HEI typically receives balanced funding from the U.S. Environmental Protection Agency and the worldwide motor vehicle industry. Frequently, other public and private organizations in the United States and around the world also support major projects or research programs. HEI has funded more than 380 research projects in North America, Europe, Asia, and Latin America, the results of which have informed decisions regarding carbon monoxide, air toxics, nitrogen oxides, diesel exhaust, ozone, particulate matter, and other pollutants. These results have appeared in more than 260 comprehensive reports published by HEI, as well as in more than 2,500 articles in the peer-reviewed literature.

HEI's independent Board of Directors consists of leaders in science and policy who are committed to fostering the public–private partnership that is central to the organization. The Research Committee solicits input from HEI sponsors and other stakeholders and works with scientific staff to develop a Five-Year Strategic Plan, select research projects for funding, and oversee their conduct. The Review Committee, which has no role in selecting or overseeing studies, works with staff to evaluate and interpret the results of funded studies and related research.

All project results and accompanying comments by the Review Committee are widely disseminated through HEI's website (www.healtheffects.org), reports, newsletters and other publications, annual conferences, and presentations to legislative bodies and public agencies.

ABOUT THIS REPORT

Research Report 216, *Scalable Multipollutant Exposure Assessment Using Routine Mobile Monitoring Platforms*, presents a research project funded by the Health Effects Institute and conducted by Dr. Joshua S. Apte of the University of California, Berkeley, and his colleagues. This research was funded under HEI's Walter A. Rosenblith New Investigator Award Program, which provides support to promising scientists in the early stages of their careers. The report contains three main sections:

The **HEI Statement**, prepared by staff at HEI, is a brief, nontechnical summary of the study and its findings; it also briefly describes the Review Committee's comments on the study.

The **Investigators' Report**, prepared by Apte and colleagues, describes the scientific background, aims, methods, results, and conclusions of the study.

The **Commentary**, prepared by members of the Review Committee with the assistance of HEI staff, places the study in a broader scientific context, points out its strengths and limitations, and discusses the remaining uncertainties and implications of the study's findings for public health and future research.

This report has gone through HEI's rigorous review process. When an HEI-funded study is completed, the investigators submit a draft final report presenting the background and results of the study. This draft report is first examined by outside technical reviewers and a biostatistician. The report and the reviewers' comments are then evaluated by members of the Review Committee, an independent panel of distinguished scientists who are not involved in selecting or overseeing HEI studies. During the review process, the investigators have an opportunity to exchange comments with the Review Committee and, as necessary, to revise their report. The Commentary reflects the information provided in the final version of the report.

CONTRIBUTORS

RESEARCH COMMITTEE

David A. Savitz, Chair *Professor of Epidemiology, School of Public Health, and Professor of Obstetrics and Gynecology and Pediatrics, Alpert Medical School, Brown University*

David C. Dorman *Professor, Department of Molecular Biomedical Sciences, College of Veterinary Medicine, North Carolina State University*

Christina H. Fuller *Associate Professor, School of Environmental, Civil, Agricultural and Mechanical Engineering, University of Georgia College of Engineering*

Marianne Hatzopoulou *Professor, Civil and Mineral Engineering, University of Toronto, Canada, Research Chair in Transport Decarbonization and Air Quality*

Amy H. Herring *Sara & Charles Ayres Distinguished Professor of Statistical Science, Global Health, Biostatistics, and Bioinformatics, Duke University*

Heather A. Holmes *Associate Professor, Department of Chemical Engineering, University of Utah*

Neil Pearce *Professor of Epidemiology and Biostatistics, London School of Hygiene and Tropical Medicine, United Kingdom*

Evangelia (Evi) Samoli *Professor of Epidemiology and Medical Statistics, Department of Hygiene, Epidemiology and Medical Statistics, School of Medicine, National and Kapodistrian University of Athens, Greece*

Neeta Thakur *Associate Professor of Medicine, University of California, San Francisco*

Gregory Wellenius *Professor, Department of Environmental Health, Boston University School of Public Health and Director, BUSPH Center for Climate and Health*

REVIEW COMMITTEE

Melissa J. Perry, Chair *Dean, College of Public Health, George Mason University*

Sara D. Adar *Associate Professor and Associate Chair, Department of Epidemiology, University of Michigan School of Public Health*

Kiros T. Berhane *Cynthia and Robert Citron-Roslyn and Leslie Goldstein Professor and Chair, Department of Biostatistics, Mailman School of Public Health, Columbia University*

Ulrike Gehring *Associate Professor, Institute for Risk Assessment Sciences, Utrecht University*

Michael Jerrett *Professor, Department of Environmental Health Sciences, Fielding School of Public Health, University of California, Los Angeles*

Frank Kelly *Henry Battcock Chair in Community Health and Policy and Director of the Environmental Research Group, Imperial College London School of Public Health, United Kingdom*

Jana B. Milford *Professor Emerita, Department of Mechanical Engineering and Environmental Engineering Program, University of Colorado, Boulder*

Jennifer L. Peel *Professor of Epidemiology, Department of Environmental and Radiological Health Sciences, Colorado State University, and the Colorado School of Public Health*

Eric J. Tchetgen Tchetgen *University Professor and Professor of Biostatistics and Epidemiology, Perelman School of Medicine, and Professor of Statistics and Data Science, The Wharton School, University of Pennsylvania*

John Volckens *Professor, Department of Mechanical Engineering, Walter Scott Jr. College of Engineering, Colorado State University*

HEI PROJECT STAFF

Allison Patton *Staff Scientist (Study Oversight)*

Dan Crouse *Staff Scientist (Report Review)*

Kristin Eckles *Senior Editorial Manager*

Hope Green *Editorial Project Manager*

Mary Brennan *Consulting Editor*

HEI STATEMENT

Synopsis of Research Report 216

New Approaches to Air Pollution Exposure Assessment Using Mobile Monitoring

BACKGROUND

It is challenging to estimate exposures to outdoor air pollutants that vary highly over short distances and over short periods of time. Researchers are increasingly measuring air pollution using mobile monitoring by affixing monitoring devices to vehicles traveling systematically and repeatedly along road networks. Data collected this way can be used to produce maps of street-level exposure estimates. Questions remain, however, about the validity and use of these data in different locations and for different purposes. Dr. Joshua Apte and colleagues at the University of California, Berkeley, sought to improve mobile monitoring approaches and to test their suitability in a high-income country (United States) and a low- and middle-income country (India). Their study was funded through HEI's Request for Applications 16-1: Walter A. Rosenblith New Investigator Award.

APPROACH

Apte and colleagues considered the following overarching research questions: Does large-scale mobile monitoring produce useful results? What insights about traffic-related air pollution dynamics and patterns can be revealed by mobile monitoring? What are the potential limitations of mobile monitoring? The study builds on previous research by the investigators through which they collected a large amount of mobile monitoring data using Google Street View cars equipped with tools to measure several traffic-related pollutants, including black carbon, nitrogen oxides, and ultrafine particles.

First, they evaluated the extent to which observations from mobile monitoring collected during weekday work hours represented long-term observations of black carbon measured at fixed-site monitors in Oakland, California. For this analysis, they used two existing datasets, namely, over 300 hours of mobile measurements and data from about 100 fixed-site monitors that provided 100 days of continuous measurements to assess temporal and spatial variability.

Next, they explored and evaluated several approaches for mapping air quality in Oakland using those mobile monitoring data. Specifically, they compared an approach using repeated, full-coverage sampling (i.e., mobile monitoring on all roads, sampled many times) with alternative strategies that included using data from fewer roads or fewer sampling days, and supplementing the data with spatial prediction models. The investigators also evaluated the feasibility of mobile monitoring in a different setting by collecting over 400 hours of data over 19 months in the Malleshwaram neighborhood of Bangalore, India.

What This Study Adds

- This study evaluated the use of mobile monitoring for several air pollution mapping and exposure assessment applications.
- Apte and colleagues compared measurements collected through mobile monitoring with measurements collected at fixed-site locations and used the mobile monitoring data to develop maps of estimated potential exposure.
- They evaluated and compared such data and approaches in Oakland, California, and Bangalore, India.
- In both locations, they produced relatively reproducible maps of traffic-related air pollution with data from relatively few repeated drive passes.

KEY RESULTS

The investigators found that patterns of black carbon obtained using mobile monitoring in Oakland were very similar to the concentrations observed at the fixed monitoring sites. In addition, mobile measurements captured road-level variability and measurements along highways that were not available from the fixed-site monitors.

This Statement, prepared by the Health Effects Institute, summarizes a research project funded by HEI and conducted by Dr. Joshua S. Apte at the University of California, Berkeley, and colleagues. Research Report 216 contains both the detailed Investigators' Report and a Commentary on the study prepared by the Institute's Review Committee.

Next, Apte and colleagues produced maps of pollutant concentrations on sampled road segments using all available data (see Figure, left panel) and reduced datasets (middle and right panel). Visual inspection suggested that the various modeling approaches captured key features of the long-term concentrations of nitrogen oxide and black carbon. Maps developed using data from fewer roads or drive days resulted in negligible decreases in model predictions and performance even with substantial decreases in data requirements. The map produced with the full dataset, however, (i.e., many dozen drive passes on all roads, total drive time about 1,300 hours) contained localized pollution hotspots at intersections and locations with emissions sources that were not apparent in the other maps.

The campaign in Bangalore showed similarly that the highest concentrations were observed along highways, and the lowest concentrations were observed on smaller, residential streets. Concentrations of ultrafine particles were about four times higher, and those for black carbon about 100 times higher, in Bangalore than in Oakland. Despite differences in fleet composition, population density, and mean pollutant concentrations between the two locations, mobile monitoring produced relatively stable maps with data from about 10 drive days in both locations, with diminishing returns to precision with additional sampling beyond that.

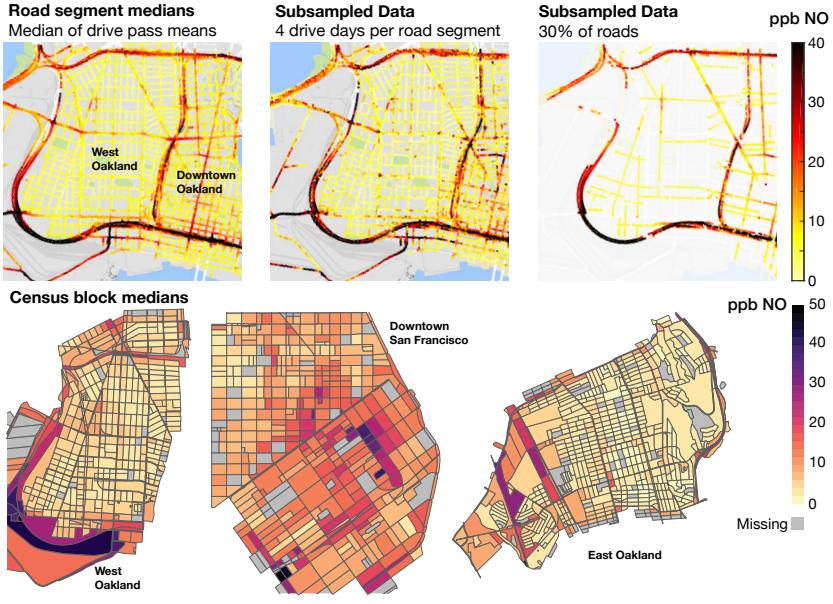
Results of this study also showed that some pollutants appear to be better suited for collection through mobile monitoring than others. Generally, pollutants with a high degree of spatial variation and a low degree of temporal variation were the best suited to this kind of approach.

Whether the results are generalizable to other pollutants or other locations (including to wider areas within California, Bangalore, or elsewhere) remains to be determined. The Committee also wondered about the suitability of mobile-measured air pollution data for use in epidemiological analyses or for regulatory purposes. For example, measurements collected in the middle of the road are likely different from those collected at roadsides or at other locations that might be closer to where people live. Additionally, the data used in this study were collected only during daytime hours on weekdays and do not reflect patterns during the times of day when people might be more likely to be at home (i.e., in the evenings, at night, and on weekends).

In summary, this study showed that mobile monitoring can be used to produce relatively reproducible maps of traffic-related air pollution with data from relatively few repeated drive passes, contributing interesting insights about collecting and working with mobile-measured air pollution data.

INTERPRETATION AND CONCLUSIONS

In its independent evaluation of the Investigators' Report, the HEI Review Committee commended the investigators for conducting one of the largest, most extensive studies examining the potential applications, strengths, and limitations of mobile monitoring. The rich datasets used by the investigators allowed them to explore and identify the relative trade-offs between intensive, repeated mobile monitoring and several alternative approaches. The study showed that mobile monitoring produced relatively reproducible maps for several traffic-related air pollutants with data from relatively few repeated drive passes, in two very different settings.



Statement Figure. Maps showing daytime median concentrations of nitrogen oxide in Oakland, CA, during 2015–2017, based on all available data (left), data from four randomly selected days (middle), and data from 30% of the arterial and residential roads (right). Source: Investigators' Report Figure 3.

INVESTIGATORS' REPORT

Scalable Multipollutant Exposure Assessment Using Routine Mobile Monitoring Platforms

Joshua S. Apte[1,2], Sarah E. Chambliss[2], Kyle P. Messier[2,3], Shahzad Gani[2], Adithi R. Upadhya[4], Meenakshi Kushwaha[4], and V. Sreekanth[5]

[1]Department of Civil & Environmental Engineering, University of California, Berkeley; [2]Department of Civil, Architectural & Environmental Engineering, University of Texas; [3]Environmental Defense Fund, Austin, Texas; [4]ILK Labs, Bangalore, Karnataka, India; [5]Centre for the Study of Science, Technology & Policy, Bangalore, Karnataka, India

ABSTRACT

Introduction The absence of spatially resolved air pollution measurements remains a major gap in health studies of air pollution, especially in disadvantaged communities in the United States and lower-income countries. Many urban air pollutants vary over short spatial scales, owing to unevenly distributed emissions sources, rapid dilution away from sources, and physicochemical transformations. Primary air pollutants from traffic have especially sharp spatial gradients, which lead to disparate effects on human health for populations who live near air pollution sources, with important consequences for environmental justice. Conventional fixed-site pollution monitoring methods lack the spatial resolution needed to characterize these heterogeneous human exposures and localized pollution hotspots. In this study, we assessed the potential for repeated mobile air quality measurements to provide a scalable approach to developing high-resolution pollution exposure estimates. We assessed the utility and validity of mobile monitoring as an exposure assessment technique, compared the insights from this measurement approach against other widely accepted methods, and investigated the potential for mobile monitoring to be scaled up in the United States and low- and middle-income countries.

Methods Our study had five key analysis modules (M1–M5). The core approach of the study revolved around repeated mobile monitoring to develop time-stable estimates of central-tendency air pollution exposures at high spatial resolution. All mobile monitoring campaigns in California were completed prior to beginning this study. In analysis M1, we conducted an intensive summerlong sampling campaign in West Oakland, California. In M2, we explored the dynamics of ultrafine particles (UFPs*) in the San Francisco Bay Area. In analysis M3, we scaled up our multipollutant mobile monitoring approach to 13 different neighborhoods with ~450,000 inhabitants to evaluate within- and between-neighborhood heterogeneity. In M4, we evaluated the coupling of mobile monitoring with land use regression models to estimate intraurban variation. Finally, in M5, we reproduced our mobile monitoring approach in a pilot study in Bangalore, India.

Results For M1, we found a moderate-to-high concordance in the time-averaged spatial patterns between mobile and fixed-site observations of black carbon (BC) in West Oakland. The dense fixed-site monitor network added substantial insight about spatial patterns and local hotspots. For M2, a seasonal divergence in the relationship between UFPs and other traffic-related air pollutants was evident from both approaches. In M3, we found distinct spatial distribution of exposures across the Bay Area for primary and secondary air pollutants. We found substantially unequal exposures by race and ethnicity, mostly driven by between-neighborhood concentration differences. In M4, we demonstrated that empirical modeling via land use regression could dramatically reduce the data requirements for building high-resolution air quality maps. In M5, we developed exposure maps of BC and UFPs in a Bangalore neighborhood and demonstrated that the measurement technique worked successfully.

Conclusions We demonstrated that mobile monitoring can produce insights about air pollution exposure that are externally validated against multiple other analysis approaches, while adding complementary information about spatial patterns and exposure heterogeneity and inequity that is not readily obtained with other methods.

This Investigators' Report is one part of Health Effects Institute Research Report 216, which also includes a Commentary by the Review Committee and an HEI Statement about the research project. Correspondence concerning the Investigators' Report may be addressed to Dr. Joshua S. Apte, Department of Civil and Environmental Engineering and School of Public Health, 661 Davis Hall, University of California, Berkeley, CA, 94720; email: apte@berkeley.edu. No potential conflict of interest was reported by the authors.

Although this document was produced with partial funding by the United States Environmental Protection Agency under Assistance Award CR–83998101 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and therefore may not necessarily reflect the views of the Agency, and no official endorsement by it should be inferred. The contents of this document also have not been reviewed by private party institutions, including those that support the Health Effects Institute; therefore, it may not reflect the views or policies of these parties, and no endorsement by them should be inferred.

* A list of abbreviations and other terms appears at the end of this volume.

INTRODUCTION

Air pollution affects billions of people worldwide (Apte et al. 2015; Burnett et al. 2018; Cohen et al. 2017). Ambient pollution measurements play a crucial role in both air pollution epidemiology and air quality management, yet the global scope of ground-based air pollution observations is limited (Apte et al. 2021; Carvalho 2016). For many regions in low- and middle-income countries (LMICs), especially in populous parts of Asia and Africa, robust air quality monitoring is largely absent (Apte et al. 2021; Martin et al. 2019). For example, India has fewer than 250 continuous air quality monitoring stations providing routine data for a population of 1.2 billion. Even in the United States, ground-based monitoring is sparse relative to the needs of exposure assessment, with a median of one to four ambient monitors per million urban inhabitants. Urban air pollution concentrations can vary sharply over short distances (<< 1 km) owing to unevenly distributed emissions sources, dilution, and physicochemical transformations (Karner et al. 2010; Marshall et al. 2008; Zhang et al. 2004). Accordingly, even where present, conventional fixed-site pollution monitoring methods lack the spatial resolution needed to characterize heterogeneous human exposures and localized pollution hotspots. This challenge might increase in the future as large regional sources are controlled while local or idiosyncratic emissions remain.

Some of the enduring mysteries of air pollution epidemiology relate to the health effects of multipollutant mixtures of traffic-related air pollution. A large body of evidence suggests that within-urban exposure gradients may have substantial human health impacts that are often not fully quantified (Crouse et al. 2015b; HEI 2010, 2022; Jerrett et al. 2005). Near-roadway populations are frequently observed to experience health effects in excess of what would be predicted by background levels (HEI 2010). Many primary pollutants, including UFPs, elemental carbon (EC), BC, nitrogen oxides (NO_x, including nitric oxide [NO] and nitrogen dioxide [NO_2]), and coarse particles, are sharply elevated above background levels in these environments, as are noise levels (Apte et al. 2011; Boogaard et al. 2010; HEI 2010; Karner et al. 2010). There is suggestive toxicological and epidemiological evidence to conclude that elements of these mixtures have important human health effects (HEI 2010, 2022). Nevertheless, at the population scale, the strongest and most consistent health effects are generally found for particulate matter with aerodynamic diameter ≤2.5 μm ($PM_{2.5}$) and ozone, which are predominantly secondary pollutants without sharp gradients. Here, a "chicken-and-egg" problem exists for air pollution exposure assessment and epidemiology. Pollution standards are informed by epidemiological evidence. Routine monitoring is expensive. Most monitoring therefore considers regulated pollutants with well-understood health effects. In turn, large epidemiological studies historically have emphasized pollutants for which routine monitoring data exist. Alternative scalable population exposure assessment techniques can therefore contribute to a richer understanding of the human health effects of the many regulated and unregulated air pollutants that vary over fine spatial scales.

Another compelling concern relates to environmental justice and the systematic racial and ethnic inequality in air pollution exposures and their associated health effects. For decades, environmental justice advocates have documented the unequal environmental burdens placed on racial and ethnic minorities in the United States. These disparities have frequently arisen as the result of inequitable or even explicitly racist planning decisions that concentrated traffic, industries, and other locally unwanted pollution sources in lower-income, less-White communities (Bullard 1993, 2020; Lane et al. 2022; Morello-Frosch et al. 2001; Rothstein 2017). Because air pollution levels vary spatially over length scales that are comparable to the spatial scales of racial segregation in U.S. cities, fine-scale intraurban pollution gradients can contribute a substantial fraction of the overall nationwide disparity in air pollution exposures for primary air pollutants (Clark et al. 2022; Jbaily et al. 2022; Lane et al. 2022; Liu et al. 2021). Even for pollutants with a strong secondary contribution, such as $PM_{2.5}$, recent modeling studies suggest that nearly all major source sectors in the U.S. economy create racially disparate air pollution exposures (Tessum et al. 2021). However, intraurban air pollution disparities by race and ethnicity tend to be especially large for primary pollutants, given that these pollutants have sharper spatial gradients (Clark et al. 2017, Clark et al. 2022; Demetillo et al. 2021; Lane et al. 2022; Liu et al. 2021). Thus, whereas a small number of centralized monitors in a city may be adequate to assess long-term trends in compliance with air quality standards, this conventional monitoring approach is not fully capable of characterizing systematic disparities in air pollution exposure.

Advances in air pollution exposure assessment approaches over the past 2 decades have helped address limitations of data coverage and spatial resolution, which are associated with central-site ambient monitoring. These methods include satellite remote sensing, chemical transport models, land use regression (LUR) models, low-cost sensor networks, and direct personal exposure measurements. Although each approach has distinct positive attributes, important limitations remain. First, satellite remote sensing instruments and chemical transport models are spatially coarse (>1–10 km resolution) and cannot characterize the fine-scale gradients (10–300 m) that drive population exposure to local emissions, such as traffic. Satellites are unable to measure some pollutants of key health concern (e.g., UFPs and EC). Dispersion models and chemical transport models are only as reliable as their underlying emissions inventories, and thus cannot reveal unexpected sources. LUR models are capable of estimating levels at high spatial resolution. However, these models provide limited temporal information, require local training datasets to be available or collected, and struggle to predict the tails of air pollution distributions, especially where idiosyncratic local sources exist.

Scalable new methods to measure how air pollution concentrations vary within cities and over time could therefore provide important insights applicable to epidemiology, atmospheric science, management, environmental justice, and public awareness. Two such measurement approaches that have become increasingly common over the past half-decade are dense low-cost sensor networks, and routine (i.e., repeated or high-frequency) mobile monitoring using fleet vehicles. These two measurement-based approaches to map pollution are potentially quite complementary (Chambliss et al. 2020). Dense networks of continuously operating fixed sensors can provide time-resolved (i.e., sub-hourly resolution) data, but at a finite number of observation locations. For example, the crowdsourced PurpleAir $PM_{2.5}$ network reports data every 2 minutes at thousands of unique locations across California alone. However, even such a dense network is usually spatially incomplete. In contrast, routine mobile monitoring can provide "wall-to-wall" coverage of an entire urban domain by repeatedly sampling air pollution on every city block (Apte et al. 2017), thus providing a very dense map of exposure estimates that are averaged over time from several repeated sampling runs. Some relative advantages of mobile monitoring include the ability to measure multiple pollutants at once (especially those pollutants for which robust low-cost sensors do not exist) and to produce high-resolution exposure estimates without the need to establish hundreds to thousands of dedicated fixed monitoring sites. Thus, this study aimed to test the proposition that routine mobile monitoring using fleet vehicles may be a highly scalable, efficient, and affordable approach for obtaining high-resolution air pollution exposure data.

The proposal that gave rise to this report was drafted in 2017 and was inspired by our pilot studies from 2015–2016 using Google's Street View mapping vehicles to develop very large mobile monitoring datasets of urban air pollution. Our pilot study with the Google Street View sampling approach in Oakland, California (Apte et al. 2017) revealed stable, fine-scale pollution patterns at 10^4–10^5 times greater spatial resolution than would be possible with conventional ambient monitors. To be clear, the use of vehicles for mobile air pollution sampling is not new — it dates back to the 1970s and likely to work by Haagen-Smit in the 1950s (Apte et al. 2011; Boogaard et al. 2010; Bukowiecki et al. 2002; Padró-Martínez et al. 2012; Westerdahl et al. 2005; Whitby et al. 1975). However, many earlier mobile monitoring studies had been constrained by limited statistical power and the absence of an efficient approach for systematic data analysis (Brantley et al. 2014). These classical mobile monitoring efforts often involved purpose-built mobile labs with finicky instruments and relied on specialized research staff (either professional scientists or graduate students) as drivers. As a result, data were generally collected during short, intensive campaigns, rather than on a routine basis. Accordingly, relatively few mobile monitoring datasets had sufficient repetition frequency (i.e., statistical power) to reveal consistent long-term spatial patterns or to evaluate changes in spatial patterns over time.

In contrast, the concept of routine monitoring emphasizes using fleet vehicles (e.g., Street View cars) that are equipped with robust instruments that require only infrequent attention by trained personnel. The vehicles are then operated on a routine regular basis by a professional driving staff within a fixed spatial domain (i.e., driving about their daily business in a city). In this scheme, it is possible to rapidly amass a large monitoring dataset on nearly every urban road at high-repeat frequency. Data reduction algorithms convert repeated 1-Hz samples into stable, precise time-averaged concentrations at high resolution (e.g., 30 m). Preliminary analysis suggested that one routinely operated vehicle with instrumentation costing $50,000–$100,000 could provide precise 30-m annual-average exposure estimates for ~250,000 people; at scale, fewer than 500 such vehicles could provide high-resolution exposure data for the ~110 million inhabitants of the 25 largest urban areas in the United States (Apte et al. 2017). Given the potential of this approach, this study therefore aimed to further evaluate the robustness and validity of this sampling approach using external datasets and, in parallel, explore approaches to further improving the efficiency and scalability of mobile monitoring. Quite fortuitously, our work did not occur in isolation: since the publication of our pilot study from Oakland in 2017, dozens of new mobile monitoring efforts using similar repeated sampling approaches have been conducted in cities around the world.

SPECIFIC AIMS

In this study, we explore the potential of repeated mobile monitoring to be used as a robust, routine, and scalable approach for characterizing spatial gradients of urban air pollution and their resulting implications for human exposure, air quality management, environmental justice, and other societal impacts. Although routine mobile monitoring has begun to gain increasing favor as a useful monitoring tool, this relatively new measurement approach has not been sufficiently evaluated in terms of its validity and capabilities. We therefore sought to address the following five inter-related aims:

Aim 1. Validate mobile monitoring as an exposure assessment technique via comparison against fixed observation networks. Because fixed-site observations are often treated as the reference standard for air pollution measurements, we sought to compare how the spatial patterns from mobile monitoring aligned with those derived from both available regulatory fixed-site data as well as specially built sensor networks.

Aim 2. Compare insights from mobile air pollution measurements against the insights that can be derived from other spatially resolved air pollution assessment techniques. Here, we sought to understand how the insights from mobile monitoring might complement and diverge from those that can be derived from other exposure assessment techniques, including regulatory observations, dense lower-cost sensor networks, and statistical exposure models.

Aim 3. Investigate the potential for scaling of mobile monitoring techniques through direct observation and through models. We sought to understand how mobile monitoring could be applied to increasing large study domains (not just neighborhoods, but full cities and regions) while minimizing the amount of sampling effort that would be required to accomplish this goal.

Aim 4. Develop a dense mobile monitoring air pollution dataset for an Indian city. Through a case study in an Indian city, Bangalore, we sought to investigate whether mobile monitoring might represent a viable path forward for adding air pollution data in lower-resource settings that currently lack robust air pollution monitoring infrastructure.

Aim 5. Evaluate the utility of mobile monitoring. In this overarching and cross-cutting aim, we sought to evaluate the utility of mobile monitoring for a range of exposure assessment, environmental justice, and air quality management applications by considering the following questions. Does mobile monitoring produce useful results? In what ways and for what exposure assessment applications is mobile monitoring effective? What policy and societally relevant insights are revealed by mobile monitoring? What complementary or additional insights can be revealed by mobile monitoring? What are the potential limitations of mobile monitoring for these applications?

METHODS AND STUDY DESIGN

The study design incorporated five inter-related analysis modules (M1–M5), each of which contributed to multiple study aims. Analyses M1–M4 focused on measurements collected in the San Francisco Bay Area, whereas analysis M5 assessed transferability of the mobile monitoring approach to an LMIC context (Bangalore, India). **Table 1** indicates how each analysis contributes to the five overarching study aims. **Table 2** provides further details on the methods and results. Each of the five analyses is summarized below.

Analysis M1: Intensive comparison of mobile and fixed-site monitoring in Oakland, California. We conducted an intensive experiment to evaluate the capabilities of mobile monitoring in the representation of time-stable spatial patterns by comparing repeated mobile air pollution measurements against a large set of continuous fixed-site measurements from a sampling campaign in West Oakland, California. For this analysis, we leveraged data that had been collected in 2017, prior to the start of this study. First, as part of the West Oakland "100 × 100" Study, Caubel et al. (2019) deployed approximately 100 custom-built low-cost aerosol BC detectors (ABCDs) that provided 100 days of continuous measurements at 97 near-road and 3 background fixed sites during the summer of 2017 and shared the resulting data with us. In parallel, two concurrently operated Google Street View cars were equipped as mobile laboratories that collected over 300 hours of in-motion BC measurements using a photoacoustic extinctiometer (PAX). We evaluated the degree to which the repeated mobile measurements were capable of representing time-stable (campaign average) concentrations measured at each of the fixed-site monitors. The complete results of this analysis were reported by Chambliss and colleagues (2020).

Analysis M2: Spatiotemporal analysis of ultrafine particle dynamics using mobile and fixed sensors in the San Francisco Bay Area. We evaluated how the spatiotemporal patterns of UFPs are correlated with other traffic-related air pollutants that are more routinely monitored, such as NO_x, BC, and carbon monoxide (CO). In the San Francisco Bay Area, concentrations of UFPs have been routinely monitored by the regulatory network since about 2011, providing a unique opportunity to compare insights about the association

Table 1. Relationship Between Aims 1–5 and Analyses M1–M5

	Aim 1 Validate Method	**Aim 2** Compare to Other Methods	**Aim 3** Scale Method	**Aim 4** Test in India	**Aim 5** Evaluate Utility
M1: Intensive comparison of mobile and fixed-site monitoring in Oakland, California	✓	✓			✓
M2: Spatiotemporal analysis of UFP dynamics using mobile and fixed sensors in the San Francisco Bay Area	✓	✓			✓
M3: Assessment of local- and regional-scale air pollution disparities in the San Francisco Bay Area using mobile monitoring		✓			✓
M4: Scaling hyperlocal air quality mapping through mobile monitoring and LUR		✓	✓		✓
M5: Mobile monitoring in Bangalore, India			✓	✓	✓

Table 2. Summary of Objectives, Approaches, and Results for Analyses M1–M5

	M1: Comparison of Mobile and Fixed-Site Monitoring in Oakland	**M2: Spatiotemporal Analysis of UFPs**
How Analysis Addresses Overall Research Aims	*Aim 1:* Validate mobile monitoring against fixed-site measurements *Aim 2*: Compare insights from mobile monitoring with those from fixed measurements *Aim 5:* Evaluate relative strengths of mobile and fixed-site sampling approaches	*Aim 1:* Evaluate whether mobile monitoring corroborates an observation from regulatory data *Aim 2:* Compare insights on UFP dynamics from mobile monitoring and regulatory monitoring *Aim 3:* Evaluate how mobile monitoring offers complementary information on seasonal patterns of UFP concentrations
Key Measurements	BC *Mobile measurements:* PAX on Google Street View cars *Fixed-site measurements:* Custom low-cost ABCD on buildings and utility poles	UFPs, NO, NO_2 + supplementary species *Mobile measurements:* NO, NO_2, UFPs + BC on Google Street View cars *Regulatory measurements:* NO, NO_2, UFPs + BC, CO at 4 fixed monitoring stations
Period of Measurements	May 19, 2017–Aug. 27, 2017	*Mobile measurements:* May 2015–Dec. 2017 *Regulatory measurements:* Full year, 2015
Geographic Coverage of Measurements	West Oakland, California (Figure 1) *Mobile measurements:* ~ 170 km of road network, 10 km^2 *Fixed-site measurements:* 100 fixed locations within the neighborhood	*Mobile measurements:* West Oakland and Downtown Oakland (Figure 1) *Regulatory measurements at 4 sites:* Sebastopol (rural), Livermore (suburban), Redwood City (urban), and Laney College (near road)
Populations Covered	West Oakland, CA (~28,000 people)	*Mobile measurements:* West Oakland and Downtown Oakland (~50,000 people) *Regulatory domain:* Entire San Francisco Bay Area (~7 million)
Statistical Analysis and Modeling Approaches	*Fixed-site:* Time-averaged concentrations at each site *Mobile data:* Time averages of repeated drive passes within a spatial buffer distance of each fixed site Assessment of the concordance between mobile and fixed-site averages using R^2, MAE, and other metrics of agreement	*Mobile measurements:* Seasonal weekday, daytime spatial patterns determined by computing medians of repeated drive-pass–mean concentrations along 30-m road segments *Regulatory measurements:* Seasonal diurnal profiles of hourly regulatory data
Key Results	Repeated mobile monitoring can reproduce time-averaged, fine-scale spatial patterns of BC with good fidelity, precision, and accuracy relative to a fixed-site sensor network	Data from mobile monitoring corroborates a surprising insight from regulatory data: Patterns of UFPs and NO_x are coupled in the winter months (indicative of a common primary traffic source), but sharply decoupled in the winter. UFPs in the Bay Area appear to be substantially driven by secondary formation during the summer months

Continued next page

Table 2, continued. Summary of Objectives, Approaches, and Results for Modules M1–M5

	M3: Assessment of local- and regional-scale air pollution disparities in the San Francisco Bay Area using mobile monitoring	**M4: Scaling hyperlocal air quality mapping through mobile monitoring and LUR**
How Analysis Addresses Overall Research Aims	*Aim 2*: Compare insights between mobile monitoring and LUR models for assessing population exposure and disparities for NO_2 *Aim 5*: Evaluate utility of mobile monitoring for assessing population exposure distributions and racial and ethnic exposure disparities at large scale	*Aim 3:* Explore whether and how statistical LUR-K models can make mobile monitoring more scalable by replacing labor-intensive measurements with statistical predictions trained on a more limited set of observations *Aim 5:* Evaluate the utility of dense "data-only" mobile monitoring approach that covers every city block vis-à-vis an alternative approach where mobile monitoring data are used only for training an LUR-K model
Key Measurements	NO, NO_2, BC, UFPs	NO, BC
Period of Measurements	May 2015–Dec. 2017	May 2015–May 2017
Geographic Coverage of Measurements	The 13 communities across the San Francisco Bay Area (93 km^2) that are mapped in Figure 1	West Oakland, Downtown Oakland, East Oakland (Figure 1) ~ 490 km of road network, 30 km^2
Populations Covered	~ 450,000 people; in this analysis data were explicitly aggregated to census-block geographies to permit assessment of the demographic factors and social disparities associated with air pollution gradients	~ 103,000 people in these three neighborhoods
Statistical Analysis and Modeling Approaches	Aggregation of repeated drive-by on-road measurements to estimate median long-term weekday, daytime median concentrations for surrounding U.S. Census blocks Computation of cumulative population-weighted exposure distributions for full population and by race and ethnicity Partitioning of total spatial variation in population exposure into within- and between-neighborhood components Assessment of relative racial and ethnic disparities at different moments of the cumulative exposure distribution	Computation of long-term weekday, daytime road segment median concentrations for repeated drive-pass mean concentrations at the 30-m road segment scale Development of LUR-K models to evaluate ability to make out-of-sample spatial predictions at unmonitored locations Monte Carlo simulations of spatial and temporal coverage in mobile mapping to assess the trade-off between the amount of data collected and fidelity of LUR-K models
Key Results	Repeated mobile monitoring can represent exposure heterogeneity across a large urban region Across the entire Bay Area region, within-neighborhood gradients account for a large (~30% for UFPs and NO_2) to dominant (>50% for BC and NO) fraction of the overall heterogeneity in the population-concentration distribution Mobile monitoring captures a much wider range of variation in the NO_2 exposure distribution than does a common nationwide NO_2 LUR model Substantial racial and ethnic disparities are driven mostly by intra-neighborhood segregation	The best-performing LUR-K models we developed are limited in their ability to capture full spatial heterogeneity we measured with data-only maps (max R^2 ~ 0.65) An advantage of LUR-K modeling is that there is very little penalty in model performance that arises from using a simulated mobile monitoring campaign with 10–50 times less data. It is possible to drive only a fraction of roads a few times and develop models that are nearly as good as the best models we trained. Data-only maps from repeated driving are superior to LUR-K models in terms of detecting idiosyncratic or unexpected spatial features and hotspots

Continued next page

Table 2, continued. Summary of Objectives, Approaches, and Results for Modules M1–M5

	M5: Mobile monitoring in Bangalore, India
How Analysis Addresses Overall Research Aims	*Aim 4:* Test mobile monitoring approach in an Indian city *Aim 5:* Evaluate the utility of mobile monitoring in the Indian context
Key Measurements	BC, UFPs, CO_2
Period of Measurements	July 2019–March 2020
Geographic Coverage of Measurements	Residential neighborhood in Bangalore (Malleshwaram) and supplemental transects in the Central Business District and between urban core and rural periphery
Populations Covered	~ 100,000 people live in the middle-income neighborhood of Malleshwaram
Statistical Analysis and Modeling Approaches	Computation of long-term weekday, daytime road segment median concentrations for repeated drive-pass mean concentrations at the 30-m road segment scale Monte Carlo simulations to assess the trade-offs between the number of repeated mobile monitoring visits and the fidelity of the resulting spatial concentration maps
Key Results	Mobile monitoring resolves time-stable spatial patterns with high fidelity in Malleshwaram and elsewhere in our domain Localized pollution gradients are sharp and reach very high concentrations in the near-road environment Observed a convergence to time-stable spatial patterns with fewer than 20 repeated mobile sampling runs over 1 year Some questions about the degree to which on-road concentrations are representative of population exposures away from roadways, especially given the persistent traffic congestion in parts of the Bangalore road network Slow traffic speeds in Bangalore present logistical challenges for mobile monitoring

ABCD = aerosol black carbon detector; BC = black carbon; LUR = land use regression; MAE = mean absolute error; PAX = photoacoustic extinctiometer; UFPs = ultrafine particles.

between UFPs and other traffic-related pollutants from both mobile and fixed-site perspectives. For this assessment, we integrated seasonally resolved maps from spatially intensive mobile monitoring in Oakland and time-resolved regulatory monitoring data of UFPs, NO_x, and other traffic-related air pollutants from multiple fixed sites across the San Francisco Bay Area, which include near-highway, urban, suburban, and rural sites. Taking a seasonal perspective, we examined the role that new particle formation plays in producing spatiotemporal patterns of UFPs that differ from other traffic-related pollutants that are often used as a proxy for UFPs. The complete results of this analysis were reported by Gani and colleagues (2021).

Analysis M3: Assessment of local- and regional-scale air pollution disparities in the San Francisco Bay Area using mobile monitoring. Disparities in air pollution exposure arise from variation at multiple spatial scales: along urban-to-rural gradients, between individual cities within a metropolitan region, within individual neighborhoods, and between city blocks. We systematically compared spatial variation in concentrations of NO, NO_2, BC, and UFPs at several scales, from hyperlocal (<100 m) to regional (>10 km), with a view to assessing consequences for outdoor air pollution experienced by residents of different races and ethnicities. To do so, we used the mobile monitoring dataset collected previously using Google Street View cars deployed in diverse communities across the San Francisco Bay Area. Overall, we collected full-coverage street-by-street monitoring in 13 distinct neighborhoods (93 km^2 and 450,000 residents) in four counties of the San Francisco Bay Area. We assessed how spatial variation at the within- and between-neighborhood levels affected racial and ethnic disparities and overall heterogeneity in population exposures. In addition, we compared our measurements with a widely used national empirical model of NO_2 to assess how insights from hyperlocal in-situ measurements differ from more conventional model-based assessments of exposure heterogeneity and disparity. The complete results of this analysis were reported by Chambliss and colleagues (2021).

Analysis M4: Scaling hyperlocal air quality mapping through mobile monitoring and land use regression. We explored and evaluated approaches to reduce data requirements for mapping a city's air quality using mobile monitors. To do so, we compared the increasingly common approach of repeated, full-coverage sampling with a set of hypothetical

alternative sampling strategies, whereby LUR models are developed using a more spatially and/or temporally limited set of monitoring data. To do so, we performed a set of data experiments on our extensive dataset of repeated air pollution sampling in Oakland, California. We considered two conceptually different approaches to reduce the sampling requirements for mapping urban air quality. First, we explored a "data-only" approach in which we attempted to minimize the number of repeated visits needed to reliably estimate concentrations for all roads. Second, we combined mobile measurements with an LUR-kriging (LUR-K) model to predict pollutant concentrations at unobserved locations; here, measurements from only a subset of roads and/or repeat visits are considered. For each set of simulated sampling scenarios, we evaluated trade-offs between sampling effort and the fidelity of the resulting exposure datasets. The complete results of this analysis were reported by Messier and colleagues (2018).

Analysis M5: Mobile monitoring in Bangalore, India. We sought to evaluate how repeated mobile sampling protocols developed in the United States could transfer to the distinct setting of mapping air pollution in dense cities in LMIC. Here, we considered the case study of Bangalore, India. In one of the few large-scale wall-to-wall monitoring exercises conducted in India, we constructed a mobile air quality laboratory and collected over 400 hours of on-road data over a period of 19 months in Bangalore. Over 22 repeat measurements, we covered diverse road segments ranging from highways to small streets, from peri-urban to business district to a residential neighborhood. We compared the insights we derived in Bangalore with those from the San Francisco Bay Area and other high-income settings.

DATA COLLECTION AND QUALITY ASSURANCE

Mobile Data Collection Instrumentation and Procedures: Bay Area Campaign (Analyses M1–M4)

To measure air pollution in the San Francisco Bay Area, two Google Street View cars were equipped with the Aclima mobile platform (Aclima, Inc., San Francisco, CA), which consists of fast-response air pollution instruments, an inlet system for particle- and gas-phase species, and a high-performance data acquisition and telemetry system. As noted above, these data were all collected prior to the start of this study. Full details of the sampling system have been described by Apte and colleagues (2017). The cars were equipped to measure the following species: NO_x, BC, and UFPs. The monitors employed were fast-response (1-Hz) laboratory-grade analyzers. NO was measured using chemiluminescence (Model CLD64, EcoPhysics AG, Switzerland). NO_2 was measured using a 450 nm cavity-attenuation phase-shift spectroscope (Model T500U, Teledyne Inc., San Diego, CA). BC was measured using a photoacoustic extinctiometer (PAX, Droplet Measurement Technologies, Boulder, CO) (Arnott et al. 1999). UFPs were measured using a water-based condensation particle counter with an effective minimum detection size of particle diameter > 2.5 nm (Model 3788, TSI Inc., Shoreview, MN). Each car was equipped with two independent global positioning system receivers with nominal ~1 m precision. The independent gas and aerosol inlet systems were designed to minimize self-sampling and particle-sampling losses. Extensive predeployment testing indicated that self-sampling was only observed to occur in rare circumstances when the car was reversed into its own exhaust plume after idling for a period at a fixed location during low-wind conditions. As described in detail by Apte and colleagues (2017), Aclima employed multiple calibration and quality assurance protocols, including frequent field calibration and/or zero-checks for all instruments, periodic manufacturer-based recalibration, cross-comparison of instrumentation between the two vehicles, and post hoc corrections to ensure synchronization and consistent response times among all sample streams.

The two Google Street View cars collected data used for this study between May 2015 and December 2017. A brief description of the sampling objectives and study design follows. Cars were based out of a garage that served as a calibration facility in San Francisco, California (marked in **Figure 1**) for ~ 90% of the study. For brief periods in May–July 2015 and July–August 2017, cars were parked overnight at alternative facilities in Mountain View and San Bruno, California. Drivers started and ended daily shifts at the garage, collecting data during daytime driving shifts that lasted 6–8 hours. Each day, our study team provided drivers a daily sampling assignment, typically a sequence of 1–5 km^2 polygons, within which the driver was tasked with driving every road at least once in an order of their choice. Our study design had two main emphases, which we illustrate in maps in Figure 1. First, we focused intensive sampling in a set of three socioeconomically diverse neighborhoods in Oakland, California (Figure 1b): West Oakland (~10 km^2); Downtown Oakland (5 km^2), and, to a somewhat lesser extent, East Oakland (15 km^2). With a focus on these three neighborhoods, we sought to develop an intentionally oversampled dataset, with many dozen repeated samples over more than 1300 hours of sampling over the 32 months of data collection. This especially intensive sampling program facilitated methodological investigations of mobile sampling study design and allowed for assessment of how spatial patterns of air quality differed among communities. Nearly all measurements were collected during daytime (8 a.m.–6 p.m.) hours on weekdays. In West Oakland, we focused on additional intensive measurements from May–August 2017 to coincide with our deployment of 100 fixed-site BC sensors. During this focused 3-month period, we collected ~ 304 hours of valid data during extended daylight hours (6 a.m.–8 p.m.) on 46 weekdays and 12 weekend days.

We complemented our intensive Oakland measurements with an additional ~1000 hours of measurements in 10 additional diverse neighborhoods (63 km^2, Figure 1a) distributed throughout the San Francisco Bay Area. These study areas, ranging in size from 2.4 to 15.7 km^2 land area, were selected to provide a range of land uses (e.g., industrial, commercial, dense residential, and light residential), atmospheric and cli-

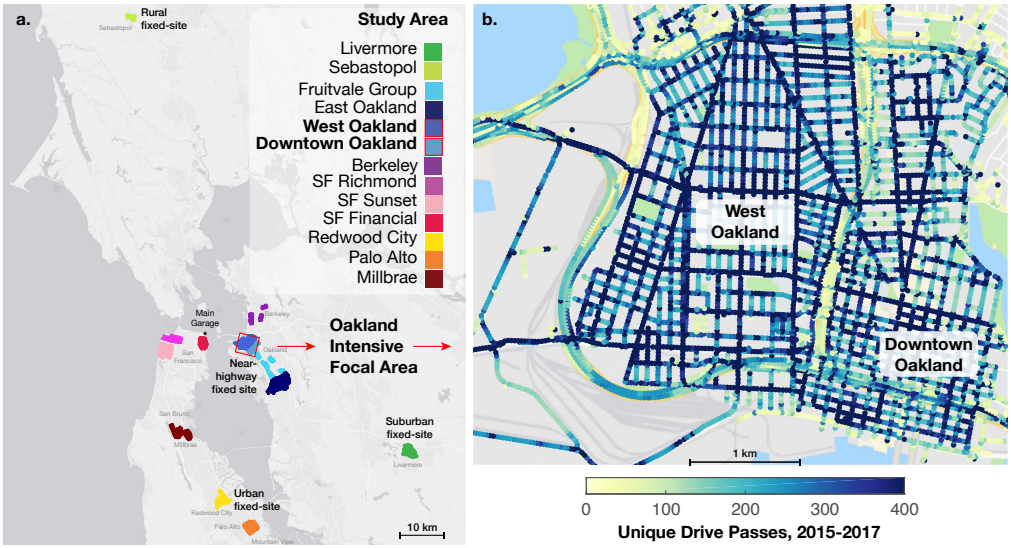


Figure 1. Maps of Bay Area study domain (93 km², ~500,000 individuals). (a) Individual neighborhood-sized study areas were distributed throughout the San Francisco Bay Area, spanning diverse conditions from the urban core of San Francisco, Oakland, and Berkeley to suburban and rural areas, such as Livermore and Palo Alto. Map notes the approximate locations of four multipollutant fixed monitoring sites operated by the Bay Area Air Quality Monitoring District (BAAQMD) that we used in the analysis M2. We report on the within- and between-neighborhood heterogeneity in pollutant concentrations across the overall study domain in analysis M3. **(b)** We focused especially intensive repeated monitoring within the inset Oakland Intensive Focal Area, comprising the West Oakland and Downtown Oakland neighborhoods. From 2015–2017, most 30-m road segments in this study design were sampled on 200–400 unique drive passes. (Panel **a** is adapted from Chambliss et al. 2021.) Map data for panel **b** © 2018 Google.)

mate conditions (upwind vs. downwind; marine-influenced vs. continental, share of open or green space, traffic density, demographic composition, and historical housing policy). Study areas were distributed within the counties of San Francisco, Alameda, San Mateo, and Santa Clara in the San Francisco Bay Area, with one background location in Sonoma County.

Our overall sampling program in Oakland (analyses M1, M2, and M4) and across the wider San Francisco Bay Area (analysis M3) was designed to minimize time-of-day or seasonal sampling biases. To do so, we undertook several design features. Study areas were repeatedly visited on a rotating schedule designed to assess long-term average concentrations indicative of typical weekday, daytime conditions. During a visit, the driver would follow a Google Street View–based driving protocol to visit every road segment within the neighborhood at least once, driving with the normal flow of traffic, with a typical speed of 25–35 km/hr on nonhighway road segments (Apte et al. 2017; Chambliss et al. 2021). For large study areas, a smaller subunit would be assigned for full coverage in a single day's driving, with full sampling occurring over multiple days. Visits to each area and subunit were distributed over different times of day and different seasons. When sampling a particular area, drivers were instructed to avoid following the same route each day. When multiple smaller subunits were sampled in a day, we randomized the sequence to avoid unintentional time-of-day biases. Across the full Bay Area domain, the median cumulative sampling time of each census block was 19 minutes, collected during a median of 47 unique visits over 20 days. Sampling coverage in Oakland neighborhoods, which were more intensively sampled, was approximately 2 to 3 times higher than the average neighborhood in the study. The statistical methods section, below, describes our evaluation of the temporal representativeness of our measurement coverage.

Mobile Data Collection Instrumentation and Procedures: Bangalore Campaign (Analysis M5)

Our measurement package in Bangalore consisted of instruments for measuring BC, UFPs, $PM_{2.5}$, carbon dioxide (CO_2), meteorological parameters, and a global positioning system (GPS). We measured BC via filter-based light absorption (Hansen et al. 1984) using a microAethalometer (model AE51, Aethlabs, San Francisco, CA). We corrected raw BC measurements for filter-loading and vibration artifacts following the method described by Apte and colleagues (2011). We measured UFPs using a battery-operated, isopropanol-based condensation particle counter (Model 3007, TSI, Inc., Shoreview, MN). This instrument measures particle number concentrations for particles >10 nm in diameter. To extend the concentration range of the instrument beyond the manufacturer's specified limit of 10^5 particles/cm^3, we used a custom-fabricated diluter that reduced concentrations 5.5-fold (Apte et al. 2011; Ban-Weiss et al. 2009). We tested the diluter regularly to ensure stable performance. We attempted to measure $PM_{2.5}$ using a DustTrak aerosol photometer (Model

8530, TSI Inc., Shoreview, MN). However, performance of the $PM_{2.5}$ measurements was not satisfactory, as this device appeared to be strongly influenced by Bangalore's particular combination of high relative humidity and combustion-dominated aerosol composition to produce implausible and evidently unreliable concentration estimates. We therefore did not report these measurements here (Kushwaha et al. 2022). We measured concentrations of CO_2 via nondispersive infrared spectroscopy (Model 840, LI-COR Biosciences, Lincoln, NE). Because CO_2 is a strong tracer of fossil fuel combustion, we used CO_2 concentrations as an indicator of the degree to which our measurements were influenced by the fresh exhaust of traffic emissions. To estimate the localized CO_2 increment, ΔCO_2, associated with local combustion emissions (as opposed to atmospheric background levels), we computed a daily time series with equation $\Delta CO_2 = CO_{2,measured} - CO_{2,min}$. Here $CO_{2,min}$ reflects the daily minimum in on-road CO_2 concentration encountered in a sampling run. For all instruments, we performed a factory calibration at the outset of the study, and all of the particle-phase instruments were zero-checked daily with a high-efficiency particulate air filter. Finally, we geolocated the position and speed of our mobile sampling vehicle with a Garmin GPSMap 64s, with a nominal precision of ~ 3 meters.

We integrated these four instruments into our mobile platform, a compressed natural gas-powered hatchback car (Maruti-Suzuki Celerio), which we selected as being an especially low-emissions vehicle model available locally. The instruments were mounted near the rear, passenger-side window, and oriented with short sampling lines to sample out the open window. To minimize the vibration that instruments could suffer due to poor road conditions, instruments were cushioned and strapped with bungee cords.

We conducted mobile monitoring of air pollution in four regions in Bangalore (Bengaluru), a large city of more than 12 million people in the southern state of Karnataka, India (**Figure 2**). The study regions included (1) an urban residential area in north Bangalore (Malleshwaram) (2) Bangalore's central business district, (3) a peri-urban area, and (4) peri-urban–urban transects. In this report, we emphasize our results from Malleshwaram for two reasons. First, this neighborhood was a completely sampled domain. Second, because of the design of the road network, Malleshwaram was the only area in which we were able to execute a block-by-block repeated sampling design comparable to our Bay Area measurements. The total road length covered in Malleshwaram was ~62 km (of ~150 km of total road length monitored), comprising highways (28%), arterial roads (24%), and residential roads (48%). Road classification was obtained from OpenStreetMap.org, the most widely used open-source global dataset on road networks. For the ~10% of study roads tagged as "unclassified" in OpenStreetMap, we used visual observations to assign a road type.

The mobile monitoring campaign ran from July 10, 2019, through March 12, 2020, a period of time representative of all

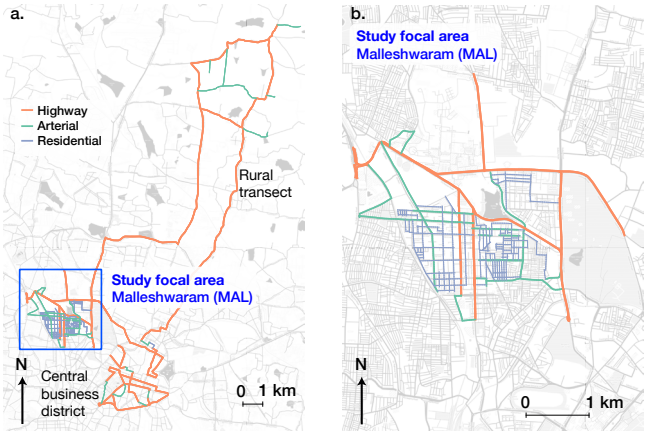


Figure 2. Bangalore study domains. (a) Maps of the overall study domain across Bangalore used for analysis M5, which is situated in roughly the northwestern quadrant of Bangalore. Within the middle-class residential and commercial neighborhood of Malleshwaram (MAL, highlighted in blue inset box), we conducted block-by-block mobile monitoring in a manner most analogous to our monitoring protocol in the San Francisco Bay Area. This study focal area is mapped in detail in (b) showing the mixture of a dense residential street grid bounded by highways and arterial roads.

seasons except the warmest part of summer (i.e., April – May). Because of the specific characteristics of our instruments in Bangalore, including limited battery life and endurance of the instrument's working fluid, we were generally limited to a 4-hour sampling period. Sampling was carried out on weekdays between 9 a.m. and 1 p.m. local time, so our maps best represent late-morning conditions on weekdays. These conditions capture two distinct regimes: Bangalore's peak rush hour and the mid-day pollution minima during periods of high atmospheric dispersion. We divided Malleshwaram into two subdomains, which were covered on different days. Our study design involved collecting approximately one weekly sample of the full Malleshwaram domain over two consecutive days, resulting in 44 days of data collection and 22 repeated drive days for each road segment.

Regulatory Air Pollutant Observations in the San Francisco Bay Area

To compare our insights from mobile air pollution monitoring with those from conventional fixed-site monitoring, we incorporated multiple years of quality-assured hourly fixed-site monitoring data from the Bay Area Air Quality Management District. Specific analytic applications of these data included an assessment of time-of-day sampling bias (analyses M2, M3, and M4), as well as a spatiotemporal assessment of UFP dynamics (analysis M2). Although each site differs in terms of the specific suite of pollutants monitored, instrumentation models were consistent from site to site. Particle number concentrations were measured using condensation particle counters (CPC, TSI, model 3783). NO_x, the sum of NO and NO_2, was measured using chemiluminescence analyzers

(Thermo Scientific, model 42i). BC was measured using aethalometers (Teledyne, model 633, equivalent to a Magee Scientific model AE33), and CO was measured using gas filter correlation CO analyzers (Thermo Scientific, model 48i).

In analysis M2, we evaluated the correlation between particle number concentrations (a strong proxy for UFP concentrations) and other traffic-related air pollutants. For this analysis, we selected four Bay Area Air Quality Management District fixed sites that were representative of a gradient in traffic influence: near-highway (Laney College), urban (Redwood City), suburban (Livermore), and rural (Sebastopol). Each site measured UFPs, CO, and NO_x; two sites additionally measured BC. To estimate annual averages, we used 2015 data, because that year had almost full coverage for these measured pollutants at all sites. For other analyses, including time-series correlations, we incorporated a full 4 to 6 years with available hourly data (typically 2011–2018) for these fixed sites.

Mobile and Fixed-Site Black Carbon Measurements During the 100 × 100 Study (Analysis M1)

In analysis M1, described in detail by Chambliss and colleagues (2020), we sought to compare mobile and fixed-site air pollution measurements in the West Oakland study domain during a 100-day period between May and August 2017. To do so, we used data from an existing dense fixed-site network ("100 × 100 BC Network") comprising 100 sites representative of residential, industrial, and high-traffic microenvironments at an average density of 6.7 sites per km^2, as described in detail by Caubel and colleagues (2019). The fixed-site network comprised 128 custom-built low-cost ABCDs, which were custom-built by Lawrence Berkeley National Laboratory (Caubel et al. 2019). In brief, the ABCD operates similarly to an aethalometer, which uses a filter-based light absorption technique to relate light attenuation on a filter to changes in BC mass loading (Hansen et al. 1984). This measurement approach for BC is distinct from the photoacoustic detection principle for BC used in the Google Street View cars in the San Francisco Bay Area measurements (M1–M4). Attenuation measurements were corrected for temperature, relative humidity, and loading artifacts before making a final determination of mass concentration. Post-correction data at a 1-hour averaging time showed a fleet average precision of 9.2% and accuracy of 24.6% evaluated relative to a commercial BC instrument (Aethalometer model AE33, Magee Scientific, Berkeley, CA). As configured here, the ABCDs measured at a maximum with averaging times of 2 seconds to 1 minute used in our analysis.

One or more low-cost ABCDs were installed at each fixed site, mounted at a height of 1.5 m on fences, porches, etc., at a median distance of 15 m from the nearest road. Of the 100 sites, 97 were located within 30 m of the road network covered by mobile monitoring, and 3 at upwind background sites along the San Francisco Bay. Network operation during the 100-day period (May 19 through August 27, 2017) was detailed in Caubel and colleagues (2019). Two mobile labs drove in West Oakland on 57 days during the same 100-day period, including 46 weekdays and 12 weekend days, for a total of 304 sampling hours. Mobile monitoring was limited to daytime hours, with most coverage from 8 a.m. to 6 p.m. Mobile labs repeatedly sampled air quality in a "blackout" pattern (Apte et al. 2017), covering all roads within subsections of West Oakland. Subsections of the West Oakland domain were driven on a rotating schedule to minimize temporal sampling bias.

The sampling design provided multiple ways in which the mobile lab could be located near the fixed-site measurements. The principal approach involved opportunistic "drive-by colocation": during normal on-road driving for the mobile lab, the mobile lab passed near the fixed-site instruments. Over the course of the campaign, this type of brief drive-by colocation occurred dozens of times for each monitor. In total, 88 hours of data were collected for which the mobile lab was within 150 m of a fixed-site instrument. For the median site in the fixed-site network, the mobile lab passed within 150 m of that site approximately 120 times over the course of the summerlong study.

In addition, the mobile lab was parked periodically near two fixed sites with ABCDs and commercial BC instruments (Aethalometer model AE33, Magee Scientific). These colocations provide an in situ comparison among the three detection methods. We collected a total of 3.7 hours of data during this type of intentional stationary colocation. As described in detail by Chambliss and colleagues (2020), we performed further experiments and analyses to evaluate differences in instrumental response and precision between the mobile and fixed-site measurement methods for BC. The three ABCDs were colocated with both PAXs for 183 hours in a semi-enclosed garage along the Embarcadero in San Francisco where routine mobile lab maintenance was performed. Major nearby BC sources included diesel vehicles and marine vessels, and concentrations of BC at a 1-minute averaging time ranged from <0.1–8 $\mu g/m^3$. Finally, we conducted manipulation experiments using filtered inlets for both the ABCDs and PAXs.

STATISTICAL METHODS AND DATA ANALYSIS

DATA REDUCTION PROCEDURES FOR ROAD-BASED DATA REPRESENTATION

In analyses M1, M2, M4, and M5, we used mobile air pollution measurements to estimate the time-stable average pollutant concentrations along roadways, which should be considered representative of the weekday, daytime conditions under which we undertook air pollution sampling. To develop these time-averaged estimates, we built on the data reduction scheme first introduced by Apte and colleagues (2017). The first step involved dividing the measurement domain into 30-m road segments. For the core Oakland

domain, this network had ~20,000 such road segments. For the Bangalore study areas, this network had ~5,000 such road segments. In the original scheme of Apte and colleagues (2017), all 1-Hz observations collected in each road segment were weighted equally in computing the mean concentration for a given road segment. Here, to ensure that each repeated drive through a given road segment (a drive pass), which had varying numbers of highly correlated 1-Hz measurements, was represented equally in our analysis, we updated our data reduction scheme as follows using a method described by Messier and colleagues (2018). First, we reduced the measurements for each drive pass through a 30-m road segment (typically ~3–10 seconds) into a single drive pass mean concentration. We then computed the median of repeated drive pass mean concentrations as our core metric for analysis. This approach has the effect of treating each drive pass as the unit of observation, which is conceptually superior to treating each individual time-resolved 1 Hz measurement as an independent unit of observation. Because this "median of drive pass means" approach incorporates information from numerous repeated drive passes, it is robust to anomalous or idiosyncratically polluted drive passes, even as it produces a concentration map (e.g., **Figure 3a**) that is highly correlated (R^2 > 0.9) with the data reduction approach used in Apte and colleagues (2017).

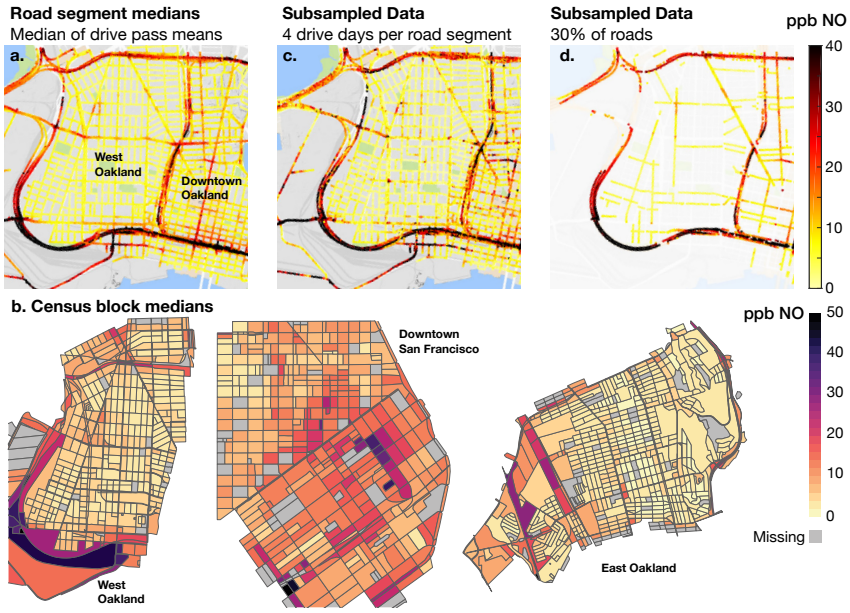


Figure 3. Examples of alternative data reduction schemes for estimating daytime median concentrations over dozens of visits. (a) Median of drive-pass mean concentrations, aggregated at the 30-m road segment, for NO concentrations measured between 2015–2017, mapped here for West Oakland and Downtown Oakland. **(b)** Census-block mean NO concentrations for West Oakland, Downtown San Francisco, and East Oakland, as used for analysis M3. **(c)** Example of a subsampled map for NO, created here using four randomly selected drive days per road segment. For our scaling analysis M4, we used such maps directly as a hypothetical "data-only" map. We also employed these subsampled maps as inputs to training an LUR-K model. With a small number of drive days, these subsampled maps present a noisier estimate of the long-term average spatial pattern as compared to the full dataset mapped in **(a)**. **(d)** Example of a road network map based on the full dataset but subsampled spatially to represent 30% of the arterial and residential road network. In our scaling analysis M4, we used similar maps as spatially restricted training datasets for developing LUR-K models. (Panels **a**, **c**, and **d** are adapted from Messier et al. 2018; panel b is adapted from Chambliss et al. 2021; and map data in panes **a**, **c**, and **d** © 2018 Google.)

In the original scheme of Apte and colleagues (2017), a multiplicative time-of-day factor based on central-site monitoring was used to adjust for diurnal variation in ambient air quality. However, this adjustment factor had only a minor (±10%) effect on long-term average spatial patterns. Given that temporal adjustment factors can introduce their own biases — time-of-day patterns of air quality differ spatially within a neighborhood — for analyses M1, M2, M3, and M5, we employed a more parsimonious approach that simply omitted the time-of-day adjustment. Because we completed initial data processing for analysis M4 before determining that time-of-day-adjustment was not required, analysis M4 retained the multiplicative time-of-day adjustment factor approach described in Apte and colleagues (2017).

To ensure that temporal sampling biases did not unduly influence our measured spatial patterns, we undertook the following assessments. As described in detail by Chambliss and colleagues (2021), we used the complete time-series datasets from regulatory fixed sites to evaluate the space–time patterns of our final sampling datasets. For evaluation purposes, hourly reference site measurements were used to calculate two multiplicative adjustment factors: (1) diurnal adjustment, the ratio of the daytime (8 a.m.–6 p.m.) median concentration to that hour's measurement, and (2) annual adjustment, the ratio of the annual median of daytime weekday daily median concentrations to the daytime median of the sampling day. These adjustment factors were applied to the mobile monitoring time-series data, and then we mapped the spatial patterns of these resulting adjustment factors to evaluate temporal bias in the sampling of those six study areas. In principle, a perfectly balanced sampling campaign would result in no spatial patterns of these adjustment factors. A spatial signature that remains in these adjustment factors reflects a temporal sampling bias.

Using this temporal adjustment approach, we found minimal time-of-day sampling biases in most of our dataset. When

considered on average by road class (residential, arterial, or highway), we found a maximum of ±5% bias resulting from the diurnal adjustment factor in the Oakland dataset, which is very small compared to the spatial variability we observed. However, some road segments were more strongly affected. Highways that were used exclusively to access a particular neighborhood from our nighttime garage operations base tended to be oversampled in one direction in the morning and another direction in the afternoon. For example, Apte and colleagues (2017) noted relatively large temporal biases in the measurements along the San Francisco Bay Bridge approaches, given that this bridge links the garage with sampling sites in Oakland and Berkeley. Thus, caution should be used in interpreting concentrations along a small number of the highways in our dataset. For our broader Bay Area measurements (analysis M3), time-of-day biases in neighborhood-average concentrations were generally <15% for NO_2, BC, and UFPs, and ~30% for NO. We also assessed seasonal biases using the annual adjustment factor described above. These biases were moderately larger, <10% by road class, and generally within <25% for individual neighborhoods. In general, these biases are quite small relative to the very large concentration gradients we observed.

Finally, Chambliss and colleagues (2021) provided the details of a set of bootstrap resampling exercises we undertook to ascertain that our overall estimates of spatial variation in pollutant concentrations were not strongly influenced by the sampling error, especially relative to the high degree of spatial variation in pollutant concentrations. Bootstrap resampling generally indicated only mild sensitivity to the particular permutations of sampling days at individual locations, which suggests that our spatial patterns were robustly estimated.

Overall, we concluded that our sampling design was not strongly influenced by time-of-day biases, but moderately influenced by choice of the particular days and seasons we sampled in. Do these biases matter? If the goal is to precisely quantify the annual average conidtions level of air pollution at a location relative to a standard, a ±10%–25% bias may be meaningful. However, if the goal is to characterize patterns of air pollution within and between neighborhoods, these biases at particular locations are quite small relative to the very large concentration gradients (factors of 2 to 8 times) we observed. Thus, for the purposes of assessing spatial patterns of air pollution, we assessed that our sampling design was robust to time-of-day and seasonal biases.

DATA REDUCTION PROCEDURES FOR REPRESENTING AIR POLLUTION BY CENSUS GEOGRAPHIES (ANALYSIS M3)

In analysis M3, we developed estimates of time-averaged air pollution concentrations for census block geographies for 13 communities around the San Francisco Bay Area (Figure 1a). Communities ranged in size between 95 and 930 census blocks (median: 447 blocks), with a total of 6,362 blocks sampled. Census blocks are the smallest aggregation unit used by the U.S. Census Bureau, with geographies that correspond roughly to city blocks in the urban cores. For our study domain, the mean census block had a land area of ~14,000 m^2 (roughly equivalent to a 120 m × 120 m square), with a mean population of 70 people.

As described by Chambliss and colleagues (2021), we calculated concentrations for each census block as the median of surrounding roads, typically located within 50–100 m from the block center point. The geographic assignment of on-road measurements to census blocks involved a two-step process. First, for each 30-m road segment surrounding every census block, we computed the median of drive-pass mean concentrations (Messier et al. 2018), as described above. Then, we calculated census block concentrations as the median of concentrations at every adjacent or intersecting 30-m road segment, using a 10-m buffer to capture road segments a small distance from the census block edge. Blocks varied in size and shape but were virtually always surrounded by roads, with a median perimeter of 447 meters. Accordingly, our census block estimates integrated measurements from 15–20 road segments but still revealed substantial fine-scale concentration variation (**Figure 3b**). In some cases — near highways and strong point sources — pollution gradients may vary over finer spatial scales than those captured by census block spatial units. However, the integration of multiple road segments provided an increase in the total number of visits and total sampling time per spatial unit, which reduced sampling error and measurement uncertainty. Although on-road measurements were not a perfect approximation of concentrations throughout a census block, our comparisons in analysis M1 of mobile and fixed-site observations in West Oakland showed no evidence of bias in on-road concentrations due to increased proximity to on-road emissions.

COMPARISON OF MOBILE AND FIXED-SITE AIR POLLUTION MEASUREMENTS (ANALYSIS M1)

Assessment of Instrumental Precision and Uncertainty

Here, we discuss key analytical considerations for the comparison of our data between mobile (PAX) and fixed-site (ABCD) BC measurement approaches in the Summer 2017 100 × 100 Study in analysis M1. In comparing measurements from two different detection methods, we assumed that both methods would respond equivalently to BC particles of varying source or age under all relevant environmental conditions. This assumption was reasonable, because (1) the ABCD measurements included adjustments for humidity effects and a filter loading artifact, (2) the garage colocation measurements showed a strong linear correlation between the two analyzers, and (3) previous evaluations validated the relative instrumental response of photoacoustic and filter-based BC measurements under laboratory and field conditions (Arnott et al. 2003; Tasoglou et al. 2018).

This combination of manipulation and colocation experiments provided multiple important insights about the compa-

rability of measurement techniques used in the mobile (PAX) and fixed-site (ABCD) observations. Here, we summarize from the detailed discussion provided by Chambliss and colleagues (2020). First, for both measurement techniques, the inherent instrumental noise at the finest temporal resolution (PAX, 1 second; ABCD, 2 seconds) often substantially exceeded the ambient concentrations of BC that we sought to measure. Thus, an instantaneous comparison of BC measurements between colocated mobile and fixed samplers was not generally possible. However, second, as expected, instrumental precision improved dramatically with increasing measurement averaging times for both the fixed ABCDs and mobile PAXs. We quantified noise as the standard deviation around zero (σ_0) of measurements made with filtered air. Indicative estimates of noise σ_0 for the mobile PAXs were 0.59, 0.31, and 0.16 μg/m^3 for 1-second, 10-second, and 1-minute averaging times. Considering the ABCDs, indicative estimates of σ_0 were 0.14 and 0.03 μg/m^3 for 1- and 20-minute integration periods, respectively, and ~0.001 μg/m^3 for a 24-hour integration period. For each road segment and sampling site, we estimated instrumental precision and a limit of detection (LOD), where the instrumental precision is defined as $\pm 2 \times \sigma_0$, and LOD = $3 \times \sigma_0$. Finally, despite the effect of instrumental noise, colocated PAX and ABCD measurements revealed highly comparable and unbiased measurements: for 20-minute averaged data, pairwise comparisons between the measurement methods resulted in an R^2 = 0.85 to 0.91 with low (<10–15%) systematic bias between the two methods.

A key overarching insight from our investigations of instrumental noise and precision is that comparisons between mobile and fixed-site measurements benefit substantially from time averaging. Crucially, because our repeated mobile measurements over the course of the summerlong 100 × 100 Study resulted in many dozen repeated drive passes near each fixed-site monitor, mobile estimates of long-term average BC concentrations near each fixed-site monitor were often quite precise, even while instantaneous samples were not. For a 20-minute averaging time, a typical time-integration period for the repeated mobile visits to a fixed site over our measurement campaign, our estimate of noise σ_0 for the PAX was only 0.08 μg/m^3, with a corresponding instrumental precision of ±0.16 μg/m^3 and LOD of 0.24 μg/m^3. While the cleanest road segments in our sampling domain had BC levels below this limit of detection, time-averaged concentrations at typical locations were substantially above this LOD: ~0.5 μg/m^3 on nonhighway road segments, and 1–2 μg/m^3 at pollution hotspots.

Data Aggregation Techniques

Next, we iteratively developed a data aggregation technique, described in this section, to address the issue that our mobile and fixed-site observations did not overlap perfectly in space. Whereas the mobile measurements were collected on roadways during in-motion sampling, the fixed-site monitors were located at a median distance of 15 m from the nearest road—typically on fences, lamp posts, front porches, and streetside building faces at a height of 1.5 meters. As described by Chambliss and colleagues (2020), we constructed radial spatial buffers of fixed distance from each fixed-site monitor to link these two distinct datasets. To process mobile data for this comparison, mobile lab GPS coordinates were used to estimate the instantaneous distance of the mobile lab from each fixed site. The series of time-resolved (1 second) mobile measurements made within a given buffer length from a fixed site make up a single unique sample visit, for which we calculated the mean of mobile measurements and the mean of contemporaneous fixed-site measurements. For this assessment, we included only measurements made between 9 a.m. and 5 p.m. We parametrically repeated this pairing between mobile and fixed-site monitors for buffer distances ranging from 30 to 150 meters. For larger radial buffers, our estimated measurement precision benefitted from the inclusion of a larger number of 1-Hz data points, thereby increasing the number of data points that exceed their respective limit of detection. In contrast to the reduction in instrumental noise at higher radial buffer lengths, however, is the spatial mismatch error that arises from including mobile measurements that were collected from microenvironments that may differ from those where the fixed-site sampler is located. By parametrically varying the buffer radius, we were therefore able to gain insight into the relative influence of these two sources of error on the comparability between mobile and fixed-site measurements. For our core analyses, we selected a buffer radius of 95 m, which appeared to best balance between these two sources of error. At a 95-m buffer length, the 97 fixed sites included for analysis received a median (10th–90th percentile range) of 73 (27–142) unique drive-by sampling visits over the full campaign, with a corresponding total of 29 (10–58) minutes of total in-motion sampling at that site.

ADDITIONAL DATASETS FOR U.S. CENSUS GEOGRAPHIES (ANALYSIS M3)

We compared our census block NO_2 concentration estimates with block-face-average concentrations of NO_2 for 2015 from a nationwide exposure model. This integrated-empirical-geographic (IEG) model was produced by the Center for Air, Climate and Energy Solutions (CACES) and is estimated in a manner conceptually similar to that of an LUR model. Nationwide, the IEG model predicted NO_2 concentrations with a cross-validation R^2 of ~0.85 and a normalized root-mean-square error (NRMSE) of ~20%.

To develop population-based estimates of exposure concentration distributions, we obtained U.S. Census Bureau block-level population data via the IPUMS National Historical Geographic Information System. We used data for the year 2010, the most recent year for which block-level data were available . Using the racial and ethnic designations provided by the U.S. Census Bureau, we categorized populations identifying as Latino and/or Hispanic in one group ("Hispanic"), and then categorized non-Hispanic populations by race: Asian, Black, White, and "Other" (including those of

Native American, Pacific Islander, multiracial, or other racial identity). The racial composition of our study population is broadly representative of the Bay Area as a whole, although it includes more neighborhoods with a high proportion of Black residents (Chambliss et al. 2021). In 2010, approximately 450,000 people lived in the census blocks that constituted our 13 mobile-monitoring sampling areas.

SCALING ANALYSIS: LAND USE REGRESSION AND DATA SUBSAMPLING (ANALYSIS M4)

In analysis M4, we investigated approaches to reducing the field data collection intensity required for producing high-resolution mobile maps, with a view to increasing the scalability of mobile monitoring. To do so, we used three intensively sampled domains in our San Francisco Bay Area study area (West Oakland, Downtown, and East Oakland), which had such a high repeated-measurement frequency that they permitted us to conduct a series of structured Monte Carlo subsampling investigations. As reported in detail by Messier and colleagues (2018), we developed models for two distinct pollutants, NO and BC. Overall, we collected approximately 3.5 million (NO) and 3.7 million (BC) 1-Hz observations in a 30-km^2 domain with ~19,000 road segments, with each road segment sampled a minimum (median) of 10 (41) times between May 2015 and May 2017. Because the overarching insights from the analyses for the two pollutants were very similar, we emphasize the results for NO here. We considered two broad classes of approaches to reducing the measurement frequency requirements for developing a high-fidelity estimate of spatial patterns.

- **Data Only**. Following Apte and colleagues (2017), this model-free approach maps concentrations solely based on repeated observations while attempting to minimize the number of repeated visits to each road. For this scaling approach, all roads must be sampled, but there is the possibility of substantially reducing the number of repeated samples at each location, at the cost of reducing the precision and accuracy of the resulting estimated concentrations surface. To implement this approach, starting with the full 2 years of observations, we developed a subsampled dataset with N driving days at each 30-m road segment from the full 2 years of observations. We estimated the long-term concentrations at each 30-m road segment as the median of drive pass means for this subsample (see **Figure 3c**).
- **LUR-Kriging Mode**. For this alternative approach, we employed our mobile air pollution measurements as training data for a statistical air quality model that combined LUR and LUR-K. By using geographic predictor variables (land use) as model inputs, it is possible to make model predictions at unobserved locations within the sampling domain. Accordingly, the LUR-K modeling approach allowed us to reduce sampling repetition and explore the consequences of using only a subset of all roads in the measurement domain as training domain. To implement this approach, we trained LUR-K models using a subset of the full 2-year dataset. We considered three alternative approaches to subsampling data to train LUR-K models, described briefly here and in detail in Messier and colleagues (2018). First, we considered a "drive day" sampling scheme: mobile monitors collect N days of data for all 30-m road segments in the domain, and then an LUR-K prediction for all road segments was trained on this temporal subsample of measurements. Second, we considered a "road coverage" sampling scheme, where all 2 years of data for only a portion of the roads in the sampling domain were included for training an LUR-K model. Third, we consider "joint" scenarios in which LUR-K predictions were developed based on a subsampled dataset where a limited number of repeated observations were collected on a limited number of roads.

To summarize, we simulated a total of four different approaches for reducing data requirements for mobile sampling: (1) data-only mapping based on a reduced subset of drive days, and LUR-K modeling based on (2) a reduced subset of drive days, (3) a reduced subset of road coverage, and (4) joint scenarios where road coverage and drive days were simultaneously reduced. For scenarios (1) and (2), we developed 16 scenarios where we randomly selected without replacement N = [1, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 45] days with valid measurements within the Oakland sampling domain from our full set of 2 years of repeated observations, preserving at least 95% of all road segments in the domain to ensure our domain did not change substantially from one subsample to the next. For approaches (3) and (4), we developed nine scenarios where we subsampled the road network to develop a map with varying levels of spatial coverage between 10% and 90% of the full set of arterial and residential roads in the domain. To ensure spatial contiguity, we sampled the dataset by street names. We did not subsample the small number of highways in the spatial domain.

For each model scenario, we conducted 100 Monte Carlo draws, and for each draw either developed a full data-only map or trained an LUR-Kriging model, as appropriate. In each case, we evaluated the fidelity of the resulting concentration field against the full years of monitoring data across the entire domain, with the R^2 and NRMSE as our evaluation metrics.

A summary of the overall LUR-K model development approach is provided here, with further details available in Messier and colleagues (2018). Models were fitted to predict the observed distribution of 30-m median-of-drive-pass mean pollutant concentrations, with the goal of capturing the high spatial resolution heterogeneity present in the full dataset at baseline. After examining the datasets for normality, we used log-transformed data for NO and untransformed data for BC. Overall model performance was assessed using untransformed data. LUR-K models were selected following a similar approach developed for the European Study of Cohorts for Air Pollution Effects (ESCAPE) studies (Raaschou-Nielsen

et al. 2013), wherein an ordinary least squares LUR was fit using a modified stepwise procedure. Variables were added based on an increase in model R^2; variables were required to be statistically significant to enter the model; variables were constrained to a priori assumption of physical interpretations (i.e., sources are expected to increase pollution therefore their coefficients are positive); and variance inflation was maintained below 3.

We used a candidate set of 121 geographic predictor variables, which included binary road classifications, binary local truck routes, local zoning classifications, normalized difference vegetative index (NDVI), percent landcover, road length, population density, and continuous point source variables (such as National Priority Listing sites, airports, and ports). Continuous variables had a distance hyperparameter, such as exponential decay distance (Messier et al. 2012) or buffer size, with a minimum buffer size of 50 meters. See Messier and colleagues (2018) for the full details of the LUR candidate predictor variables. In brief, we constructed 121 candidate input variables for each road segment as follows:

- Binary road classification using the OpenStreetMap dataset, with indicator variables for highways, arterial roads, and residential streets
- The total road lengths for highways, major arterials, residential roads, and total roads within a given distance of each road segment were calculated from the OpenStreetMap data in 50, 100, 250, 500, 1000, and 2500 meter buffers
- Binary classification of roads that are designated truck routes and roads from which trucks are restricted, based on City of Oakland data
- Binary zoning variables representing City of Oakland zoning for residential, commercial, and industrial land uses
- The average NDVI within buffer radius lengths of 50, 100, 250, 500, 1000, and 2500 meters
- Variables for the average percentage coverage of the following satellite-based National Land Cover Database land cover types: Open, Developed Low, Developed Medium, Developed High, Evergreen Forest, Deciduous Forest, Mixed Forest, and Impervious Surface. Buffer sizes calculated include 50, 100, 250, 500, 1000, and 2500 meters
- The mean elevation within circular buffers (50, 100, 250, 500, 1000, and 2500 m) was calculated in Google Earth Engine using a 10-m resolution digital elevation model.
- Proxy variables for potentially contributing point sources, including ports, airports, National Priority Listing sites, and Toxic Release Inventory sites, computed based on either (1) a minimum-inverse-distance metric or (2) a sum-of-exponentially decaying contributions model (Messier et al., 2012, 2018)

We developed a modified *K*-fold cross-validation scheme with spatial clustering to evaluate our model performance. In conventional practice, *K*-fold cross-validation involves randomly assigning observational data into *K* distinct folds, which are iteratively used for either model training or evaluation. This conventional approach has two key conceptual limitations for our application. The first was physical realism. Because a key goal was to explore how a spatially restricted dataset of vehicle-borne air pollutant observations can be used to make predictions at unobserved locations, our training datasets needed to represent physically realistic driving patterns, rather than a randomly selected set of disconnected road segments sprinkled throughout a city (in other words: our cars drive, they don't teleport). The second consideration was spatial autocorrelation. Because air pollution data are spatially autocorrelated, testing our models on spatially random cross-validation data would have overestimated our predictive ability, because the models would be informed by near-neighbor information.

To address these two conceptual issues, we used a genetic algorithm to define $K = 10$ contiguous, similarly sized spatially clustered cross-validation groups to minimize the spatial autocorrelation of near-neighbors in the cross-validation. Owing to the high spatial density of mobile monitoring samples, this cluster approach to cross-validation reduces the effect of the extremely close neighbors and more rigorously approximates out-of-sample prediction performance. In 10-fold cross-validation, the subsampled road segments selected for model training were divided into $K = 10$ spatially clustered folds. We then cycled through the 10 possible permutations of $K–1 = 9$ folds, each time training an LUR-K model on 9 of the 10 folds, while reserving data from the tenth fold for independent model evaluation.

For each set of $K = 10$ folds, we apply the fitted LUR-K model to make predictions in the single held-out spatial cluster. In conventional LUR modeling practice, model predictions would be compared against the data withheld from the training dataset for each of the *K* folds. In contrast, our core analyses of LUR-K model performance compared model predictions for each road segment within the held-out cluster with the long-term median-of-drive-pass-mean concentrations at those locations. Whereas the former analysis approach provides information on how well the model reproduces the training dataset, the latter approach summarizes how well the model predicts long-term average concentrations.

Based on the modified-stepwise model selection procedure described above, we fit models for log-NO and BC. For log-NO, geographic information system (GIS) covariates selected in 5 or more of the 10 folds included road-type indicators (e.g., highway roads and residential roads), the local truck route indicator, NDVI within a 50-m buffer, and elevation. For BC, GIS covariates selected in 5 or more of the 10 folds included the highway road type indicator, the local truck route indicator, the sum of exponentially decaying contributions from U.S. EPA toxic release sites within 5 km, and the sum of expo-

nentially decaying contributions from port land uses within 5 km. Although our modeling approach is not necessarily intended to produce models that are physically interpretable, we note that the set of selected variables and the signs of their coefficients generally comport with known sources and dynamics of BC and NO.

RESULTS AND INTERPRETATION

COMPARISON OF ESTIMATES FROM MOBILE AND FIXED-SENSOR NETWORKS (ANALYSIS M1)

Spatial Patterns

Figure 4 shows the distribution of daytime (9 a.m.–5 p.m.) median concentrations estimated by mobile monitoring and by the dense network of 97 ABCD sensor sites during the May–August 2017 period. In Figure 4a and b, we contrast the spatial patterns revealed by the fixed-site network with the more spatially complete patterns from mobile monitoring on every road segment. The mobile monitoring map shows the same general spatial patterns as fixed-site daytime medians. Measurements at the road segment level also reveal localized patterns not detected by the fixed-site network, with examples marked 1–4 on the map in Figure 4b. Mobile monitoring provides measurements on highways where placement of fixed-site monitors may be infeasible. Mobile coverage near example 1 shows the increase in concentration on elevated sections of Interstates 880 and 580 compared to the adjacent road network, as well as concentration reduction with distance from highways. Industrial activity near example 2, including a cement plant and metals recycling facility, is reflected in elevated concentrations at nearby fixed sites, while mobile monitoring also captures several additional highly localized hotspots. Road-segment medians also show hotspots corre-

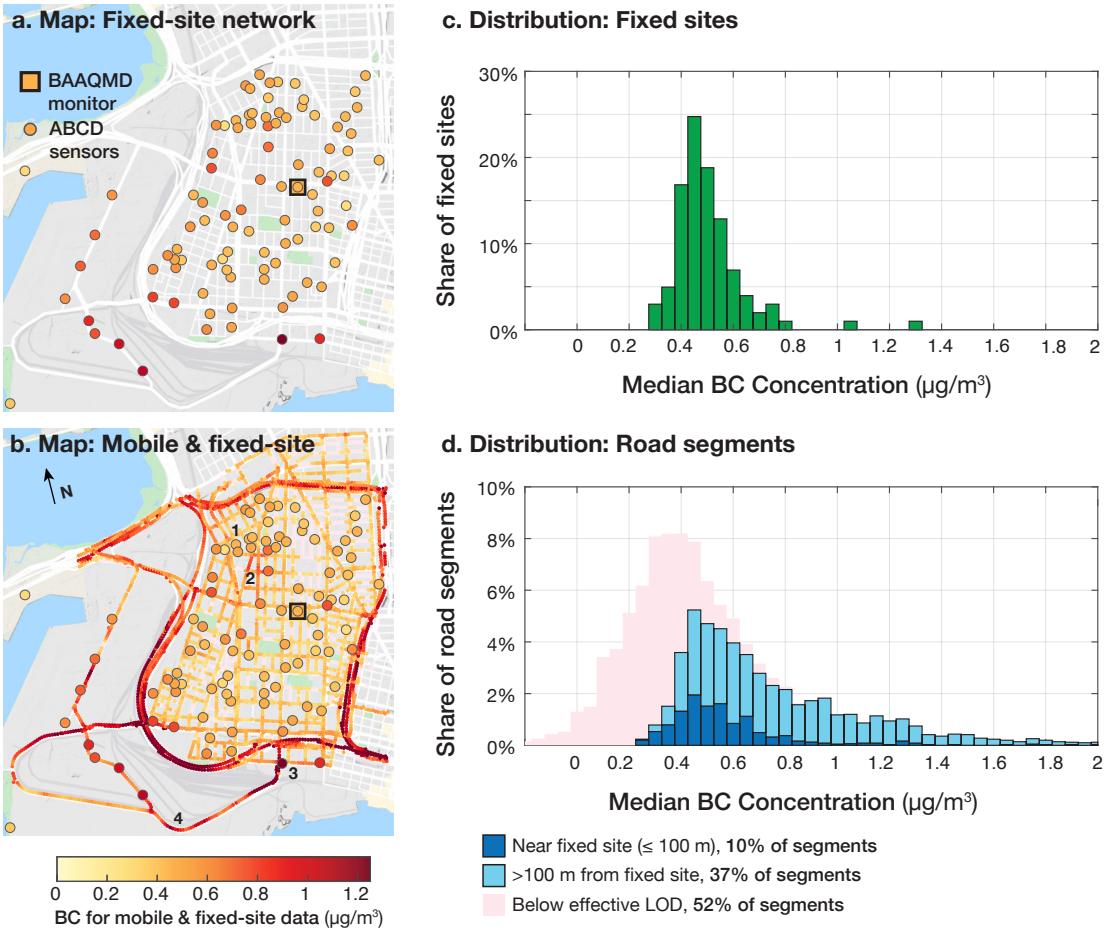


Figure 4. Contribution of mobile monitoring to additional spatial coverage beyond the fixed-site network, as used in analysis M1. (a) Median daytime fixed-site BC concentrations for the sensor network. (b) Median daytime road segment BC concentrations overlaid with the fixed-site sensor network, illustrating the additional spatial coverage provided by mobile monitoring. Numbers indicate four sites called out in text. (c) Histograms of median daytime fixed-site concentrations. (d) Median mobile monitoring concentrations by road segment within the study domain. Mobile monitoring data in (d) are divided into data collected within 100 m of a fixed site, which constitutes 22% of spatial coverage over the effective LOD, and data collected beyond 100 m from a fixed site, constituting 78% of spatial coverage.

sponding to specific routes such as the intersection segment at example 3, which acts as a funnel for truck traffic to the Port of Oakland. Concentration peaks along roads like the designated truck route around the Port of Oakland south of example 4 may reflect persistent small-scale differences in patterns of traffic congestion. Thus, mobile monitoring adds local context to the more precise, time-resolved measurements at fixed sites.

The map visualization in Figure 4b emphasizes how mobile monitoring fills in gaps even for an unusually dense fixed-site sensor network (median pairwise distance among nearest-neighbor fixed sites is ~160 m). As further illustration of the potential value of mobile monitoring, 78% of mobile monitoring data was collected at a distance greater than 100 m from any fixed site (Figure 4d). The upper tail (top 5%) of the mobile distribution (Figure 4d) shows highway road segments, many of which exceed 1.25 μg/m^3. Fixed-site hotspots (>0.80 μg/m^3) appear as isolated peaks of 1–2 monitors in Figure 4c, matched in Figure 4d by a small share of near-site mobile monitoring data and a large share of additional mobile monitoring data collected in interstitial areas. Nonetheless, the overall median among fixed sites (0.48 μg/m^3) closely matches the median concentration of nonhighway road segments (0.44 μg/m^3). This similarity suggests that nonhighway data collected on-road are broadly representative of near-road concentrations despite closer proximity to tailpipe emissions.

Assessment of Correspondence Between Mobile and Fixed-Site Median Concentrations

We assessed the fidelity with which long-term average daytime fixed site concentrations could be represented by drive-by mobile sampling. For each fixed site, we computed the median daytime (9 a.m.–5 p.m.) concentration measured by the continuously operating ABCD samplers. The corresponding mobile-derived estimate for this long-term average concentration was computed as the (temporal) median of all visit-level (spatial) means of the time-resolved mobile measurements collected within the spatial buffer radius. Thus, the temporally sparse but repeated mobile monitoring visits to the area immediately surrounding each fixed monitoring site are used to estimate the true time-integrated median concentration of each temporally continuous fixed-site dataset. In essence, this is a spatial evaluation, as we are assessing the ability of temporally sparse mobile measurements to reproduce the spatial pattern of concentrations that are measured by fixed-site monitoring. Across the network of n sites with valid data (i.e., n = 97 sites for a buffer length of 95 m), we computed two comparison metrics between these pairwise estimates: the ordinary Pearson R^2 coefficient of determination and the mean absolute error (MAE, expressed as μg/m^3).

Figure 5 presents our assessment of spatial correlation at multiple spatial scales. At a buffer distance of up to 95 m, mobile monitoring reproduces fixed-site daytime median concentrations with an MAE within the bounds of the mobile instrument's precision limits (MAE = 0.11 μg/m^3, as compared to 95% precision ± 0.15 μg/m^3). The majority of the fixed-site concentration points are clustered within the range of 0.4 to 0.7 μg/m^3 (Figure 5a), typical of residential and commercial area concentration. Approximately 20% of points occur at concentrations greater than 0.7 μg/m^3, indicative of high traffic or industrial activity (cf. Figure 4). Fixed-site and mobile in-motion medians are reasonably well correlated (R^2 = 0.51), despite method differences and temporal sparsity.

At buffer length scales longer and shorter than 95 m, two competing influences are at play. Because the measurement sample size — and thus instrumental precision — declines substantially at smaller radial buffer lengths, the MAE for the mobile-to-fixed site comparison increases sharply (Figure 5e), and the R^2 for this spatial comparison declines to values in the range of ~0.35–0.45 (Figure 5d). In contrast, at buffer radii longer than 95 m, although sample sizes further increase (Figure 5b), there are diminishing returns to the instrumental precision (Figure 5c). However, at longer spatial distances, we believe that the spatial mismatch errors between mobile and fixed-site observations for some sites may become sufficiently important that the R^2 of the spatial comparison declines again toward ~0.35. Thus, for this particular choice of instrument (PAX), pollutant (BC), and study setting (West Oakland during cleaner summer conditions), there appears to be a local optimum of the spatial scale at ~95 m for comparing mobile and fixed-site measurements.

Next, we explored the degree to which the specific timing of a finite number of incompletely randomized mobile drive passes might lead to a misestimation of the central-tendency concentrations measured from continuous observations at each fixed site. To quantify this dimension of sampling error, we compared the true daytime median concentrations at each ABCD fixed site with the median of subsampled concentrations measured by the ABCD sensors when they were being passed by the mobile samplers. These subsampled ABCD concentrations reproduced the true daytime median concentrations better than our estimates from the mobile monitoring, with minimal bias: a R^2 of 0.74 and MAE of 0.09 μg/m$_3$. It is worth noting here that this extremely temporally sparse subsample of ABCD measurements, representing ~70 point-in-time measurements per site, successfully approximated the median of continuous fixed-site measurements. Here, a key inference is that a moderate number of very temporally sparse mobile measurements, if measured with sufficient analytical precision, can reproduce the long-term average of continuous measurements. In the particular case of our 100 × 100 experiment, instrumental precision became the binding constraint, especially given the very low summer-daytime BC concentrations (~0.5 μg/m^3) and the rather high LOD for our mobile BC instrument (0.22 μg/m^3). It is reasonable to expect that the comparison of mobile to fixed site might have been even stronger (1) under more polluted conditions (winter in Oakland or a more polluted city) or (2) for measuring pollutants for which the 1-Hz instrumental precision was less of a measurement issue.

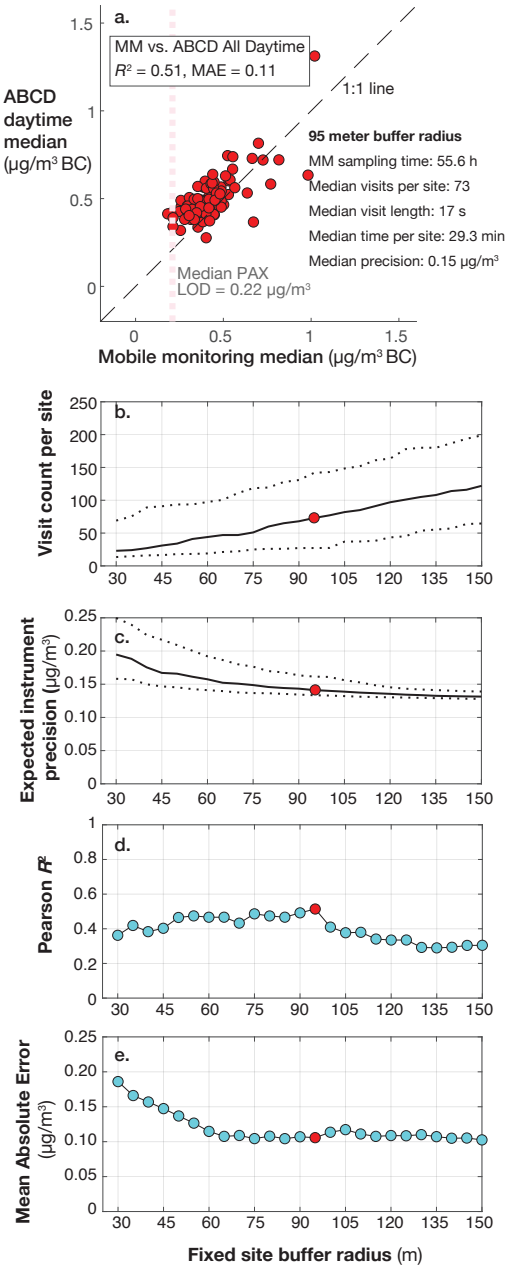


Figure 5. Comparison between mobile and fixed-sites observations during the 100 × 100 Study (analysis M1) in West Oakland. (a) Using a fixed-site radius of 95 m, pairwise correlation between median BC concentrations measured by the photoacoustic spectrometer (PAX) instrument during mobile monitoring visits (MM) and ABCD fixed-sensor measurements made during all daytime hours throughout the campaign (ABCD daytime median). Given nearly continuous monitoring, the LOD for ABCD daytime medians is <<0.01 µg/m³, while the median PAX LOD is 0.22 µg/m³ for this radius. **(b)** Visit count increases with an increasingly large buffer radius used to aggregate mobile measurements around each fixed site, thereby improving (c) the expected instrumental precision (2σ) of the mobile monitoring PAX instrument. **(d)** and **(e)** illustrate the Pearson R^2 and mean absolute error (MAE, µg/m³), respectively, for a pairwise comparison between the fixed-site ABCD sensors and the aggregated mobile measurements within a variable buffer radius given by the abscissa. We interpret the local maximum Pearson R^2 at 95 m (see **d**) and the accompanying minimum in MAE near this distance as reflecting the interplay between two factors. Whereas the instrumental precision of mobile measurements improves with the greater aggregation afforded by a larger buffer radius (see **c**), this increasing buffer radius imposes a trade-off in terms of increased spatial mismatch error between the location of the fixed sensor and increasingly distant on-road mobile measurements. (Adapted from Chambliss et al. 2020.)

SPATIOTEMPORAL ASSESSMENT OF THE RELATIONSHIP BETWEEN UFPS AND OTHER TRAFFIC-RELATED AIR POLLUTANTS (ANALYSIS M2)

In analysis M2, we undertook a detailed investigation of the relationship between UFPs and other traffic-related air pollutants in the Bay Area, with a view to evaluating this relationship from the perspectives of fixed-site and mobile monitoring. The detailed results of this investigation are presented by Gani and colleagues (2021).

Figure 6a presents diurnal profiles for UFPs and NO_x for the year 2015 stratified by season and weekday/weekend, for four regulatory fixed-sites with continuous UFP monitoring. For both UFPs and NO_x, annual average traffic-related air pollutant concentrations follow a consistent and strong gradient by site. At the near-highway site (Laney College, located ~10 m from the I-880 highway in Downtown Oakland), annual average concentrations were 29,000/cm³ UFPs and 34.7 ppb NO_x. At the urban site (Redwood City), annual average concentrations were 11,900/cm³ UFPs and 18.8 ppb NO_x. At the suburban site (Livermore), annual average concentrations were annual average concentrations were 10,100/cm³ UFPs and 17.4 ppb NO_x. Finally, at the rural site (Sebastopol), annual average concentrations were considerably lower: 3500/cm³ UFPs and 8.4 ppb NO_x.

By stratifying the diurnal profiles of UFPs and NO_x at each site by winter/summer and weekday/weekend (Figure 6a), we gained insight into coupling and divergence between the air pollutant dynamics of UFPs as compared with NO_x. During winter conditions, we observed a tight coupling of the diurnal (hour-of-day) concentration profiles for UFPs and NO_x. For the urban, rural, and suburban sites, both UFPs and NO_x show the double-peaked diurnal profile that is commonly observed for primary air pollutants. This profile arises because of the competing influence of traffic emissions (which generally peak at morning and evening commute times) and the effect of dilution into the atmospheric boundary layer, which is strongest from midday into the afternoon. Thus, the morning and evening peaks arise from traffic emissions that are diluted into a shallow atmospheric boundary layer, while a midday trough emerges because of the strong effect of daytime dilution. During winter, the overall shape of the diurnal profiles for both pollutants is similar on weekdays and weekends, but with lower concentrations on weekends, as might be expected from the lower level of traffic emissions on weekends. For the

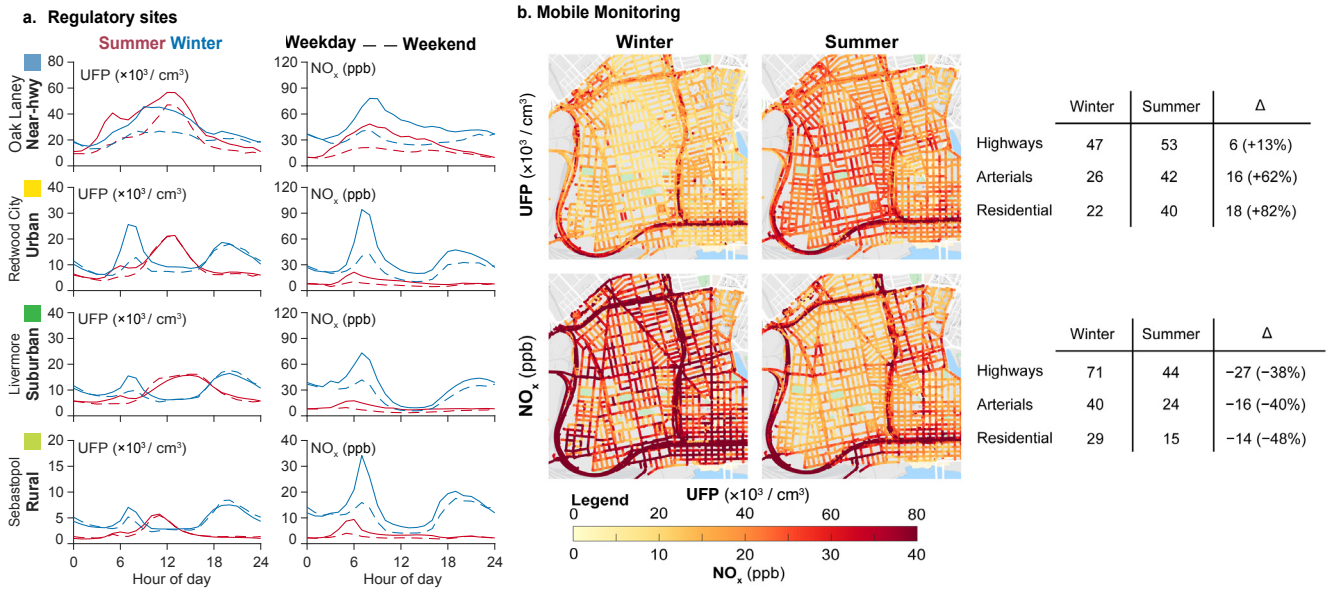


Figure 6. Spatiotemporal dynamics of UFPs in the Bay Area. (a) Concentrations of UFP particle number count (UFP PN) and NO_x at the four monitoring sites indicated in Figure 1. Diurnal cycles (hour-of-day) mean concentrations are presented for summer and winter months on weekdays and weekends. During winter months, both UFPs and NO_x follow similar diurnal cycles that are characteristic for primary combustion–derived pollutants, with weekend decline in concentrations that are in line with the expected major traffic source for both pollutants. During summer months, the UFP diurnal cycles decouple sharply from NO_x, and have a daytime peak that is highly suggestive of secondary new particle formation. **(b)** Summer and winter maps of daytime UFP PN and NO_x concentrations from mobile monitoring in the core Oakland Intensive Focal Area. Whereas the NO_x concentrations show the expected seasonal decrease in concentrations from winter to summer, UFP concentrations increase from winter to summer months, with the greatest relative increases on residential road segments. An overall consequence is much lower spatial variability in UFP exposure during summer months than for other primary pollutants. (Adapted from Gani et al. 2021.)

near-highway site, which is strongly influenced by sustained truck traffic throughout the day on I-880, wintertime concentrations of UFPs and NO_x do not experience a mid-day trough, but the two pollutants have similar diurnal profiles. Finally, Gani and colleagues (2021) reported on additional observations of other primary traffic-related air pollutants, including BC and CO. Wherever these pollutants are measured, they show similar wintertime diurnal profiles to NO_x and UFPs.

In summertime, we found that the diurnal profiles of UFPs and NO_x decouple substantially. In particular, UFP profiles show a strong divergence from the archetypal diurnal profiles that one would expect for a conserved pollutant, whereas NO_x retains that diurnal profile. Considering the hourly time-series correlation between UFPs and NO_x during winter days, winter nights, summer days, and summer nights, the Pearson correlation coefficient was lowest at each site during summer days (Gani et al. 2021). Summertime diurnal profiles for NO_x show lower average concentrations than wintertime, as would be expected by the substantial seasonal increase in dilution during warmer months. In addition, the second daily peak disappears from the diurnal profile, given that the evening decrease in mixing height happens well after the afternoon commute period during the summer months. Concentrations of NO_x are lower on weekends, as would be expected for a traffic-related pollutant. Other primary pollutants, such as BC and CO, share this diurnal profile at sites with available data (Gani et al. 2021). In contrast to these primary pollutants, the principal summertime peak concentration in UFPs during summer months happens between 11 a.m. and 3 p.m., a time of day when other pollutants are close to their daily minima. Additionally, daytime peak UFP concentrations show minimal difference between weekdays and weekends during the summer, thus producing *higher* weekend peak concentrations of UFPs in summer than in winter at all monitoring sites.

In **Figure 6b**, we map the average spatial patterns of daytime UFPs and NO_x for Oakland by road segment on the basis of routine mobile monitoring with Google Street View cars. Whereas on-road concentrations in NO_x decrease from winter to summertime, consistent with the seasonal increase in atmospheric ventilation, the on-road concentrations of UFPs increased substantially in summertime. For both UFPs and NO_x, we observed the smallest relative seasonal changes — but in opposite directions — for highway road segments, with the largest relative seasonal changes seen on residential streets. Because residential streets are less strongly affected by highly localized sources of air pollution (i.e., traffic on that street) than are arterials and highways, residential streets are where we would expect to see the largest impact of regional-scale processes influenced by seasonality. Daytime concentrations of NO_x declined by 48% from winter to summer on residential

streets but increased by 82% from winter to summer for UFPs. As further evidence of the seasonally shifting relationship between NO_x and UFPs, the average on-road ratio of UFP:NO_x shows little time-of-day variation during winter months, but during summer shows a strong three- to fourfold increase from the morning into the afternoon hours (Gani et al. 2021). This increase in the UFP:NO_x ratio during summer afternoons is most prominent on residential streets, which are the least influenced by traffic, and least prominent on highways.

In combination, our data from both fixed sites and mobile monitoring presents strong circumstantial evidence of a major nontraffic source of UFPs, especially during summer daytime hours. Unlike other traffic-related air pollutants, UFP concentrations can be strongly affected by nucleation, a process of new particle formation from atmospheric vapors. These nucleation events have been widely observed in urban, regional, and background environments spanning a range of conditions from pristine to polluted (Boy and Kulmala 2002; Costabile et al. 2009; Gani et al. 2020; Kulmala et al. 2004; O'Dowd et al. 2010; Sellegri et al. 2010; Vakkari et al. 2011). A growing body of evidence shows elevated UFP concentrations during periods with increased solar radiation (Betha et al. 2013; Brines et al. 2015; Hudda et al. 2010; Salma et al. 2011; Shen et al. 2011). In a comparative study of multiple cities with a Mediterranean climate (Rome, Barcelona, Madrid, and Los Angeles), Brines and colleagues reported that while traffic was the dominant contributor to UFP concentrations, under sunny conditions new particle formation could lead to UFPs becoming decoupled from other traffic-related air pollutants.

To summarize, we found daytime peaks in UFP concentrations at multiple sites during the warmer months that were not observed for other primary traffic-related pollutants. To provide context, consider that we measured a twofold increase of UFP concentrations during mid-day hours relative to the morning rush hours, while in contrast we found a twofold decrease in NO_x and other traffic-related air pollutant concentrations during the same period. This is strong evidence that new particle formation can complement traffic as a major source of ambient UFP exposure. In approximate terms, this finding implies that for the half of the year when new particle formation is common in the San Francisco Bay Area, approximately half or more of the UFP concentrations might be attributed to new particle formation during the peak hours for this photochemical process. Because the spatiotemporal variation in NO_x concentrations differs from UFP concentrations, using NO_x (or other traffic-related air pollutants) as a proxy for UFPs could result in inaccuracies in estimating UFP exposure.

CHARACTERIZING THE HETEROGENEITY OF AIR POLLUTION EXPOSURES WITHIN AND AMONG NEIGHBORHOOD BY RACE AND ETHNICITY ACROSS THE SAN FRANCISCO BAY AREA (ANALYSIS M3)

Description of Multipollutant Exposure Gradients

In analysis M3, we analyzed the spatial distribution of population exposure based on the census-block level estimates of BC, NO, NO_2, and UFPs in 13 different study areas sampled across the Bay Area. Here, we refer to the population-weighted distribution of concentrations as an estimate of "exposure;" we acknowledge that individual exposure also depends on many other factors (e.g., diurnal activity, indoor infiltration and dynamics, and physiology).

Over these 13 different study areas, the population-weighted mean (range of study area means) concentrations were 0.31 μg/m^3 (0.18–0.60) for BC, 4.6 ppb (0.9–10.6) for NO, 8.2 ppb (3.3–13.1) for NO_2, and 19,100/cm^3 (6,900–33,700) for UFPs.

In discussing spatial variation, we refer to gradients among neighboring blocks (~100 m) as "hyperlocal," variation within each study area (~1 km) as "local," and variation among study areas (~10 km) as "regional." Among the four pollutants, NO showed the highest-magnitude hyperlocal peaks, with a typical ratio of 10× between a peak and local median (Chambliss et al. 2021). BC, NO_2, and UFPs (peak ratios 3.1×, 2.7×, and 2.6×, respectively) exhibited shallower hyperlocal gradients and more diffuse peaks. Within many individual study areas, the correlation between block-level concentrations of individual pollutants was quite variable, with a rather low correlation between UFPs and other pollutants (interquartile range Pearson's $r \sim 0.4–0.7$), but high correlations between NO and NO_2 ($r \sim 0.8–0.9$). These differences demonstrate the importance of measuring multiple pollutants. Furthermore, these patterns likely differ from those of other important pollutants like fine particulate matter ($PM_{2.5}$) and air toxics, both in location and degree of local and regional variation.

The exposure variation that we were able to quantify here reflects the interaction between multiscale gradients of air pollutant concentrations (regional, local, and hyperlocal) and the spatial distribution of census-block populations. **Figure 7** shows the full distribution of exposure levels within each study area. Comparing median exposures between the most- and least-polluted study areas, concentrations varied by factors of 4, 5, 6, and 28 for BC, NO_2, UFPs, and NO, respectively, while within-neighborhood interdecile (10th–90th percentile) ranges showed variation up to a factor of 4 for BC, NO_2, and UFPs, and a factor of 19 for NO. Generally, neighborhoods with higher BC and NO medians also displayed a wider range of exposures, while NO_2 and UFP ranges remained more consistent across neighborhoods. This difference in the spatial patterns between the two exclusively primary pollutants (BC and NO) and the two pollutants with substantial secondary formation (NO_2 and UFPs) is consistent with prior mobile monitoring studies (e.g., Apte et al. 2017) and expectations about air pollutant dynamics. Chambliss and colleagues (2021) described an analysis where we partitioned variability into local and regional components by decomposing the sum-of-squared deviation from the mean (SSD) of each resident versus the study area mean and of all study areas versus the grand mean. This analysis found that local gradients contributed the majority of exposure variation for primary pollutants NO and BC (52% and 63% of SSD,

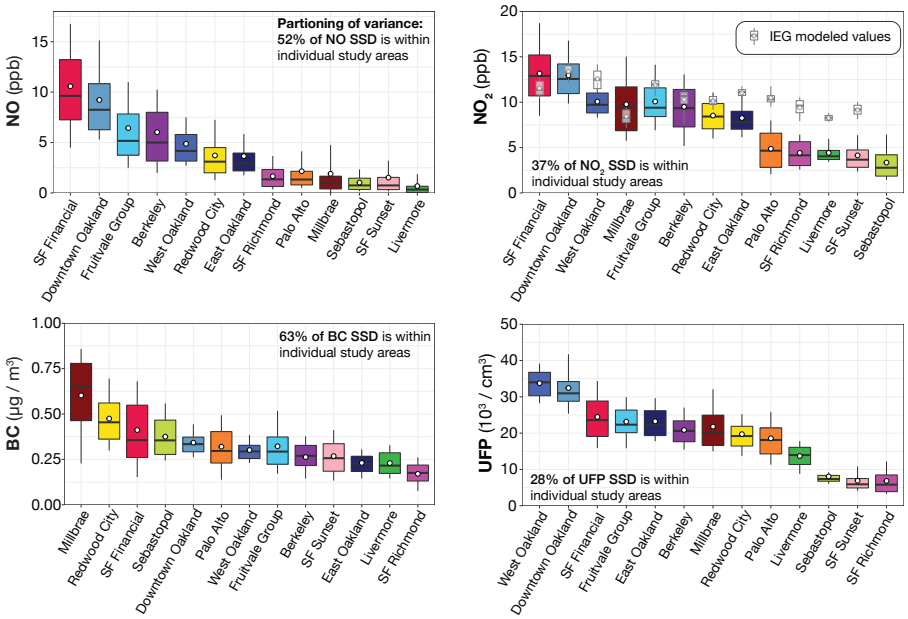


Figure 7. Exposure distributions for the 13 different San Francisco Bay Area neighborhood study areas mapped in Figure 1 (analysis M3). Distributions reflect population-weighted exposures based on daytime weekday census-block concentrations. Areas are shown in order of descending median concentration. Whisker ends represent 10th and 90th percentiles, box boundaries represent upper and lower quartiles, the center bar marks the median, and circle represents the mean. Gray box plots in the NO_2 panel represent modeled exposure estimates from the CACES national-scale integrated empirical geographic (IEG) regression model. Using a partitioning-of-variance technique (sum of squared deviations, SSD), we decomposed the overall heterogeneity in the population-weighted concentration distributions into within- and between-study area components. For the two pollutants at left (NO and BC; both dominated by primary emissions), more than half of the heterogeneity in population exposures occurs within individual neighborhoods. In contrast, for the two pollutants at right (NO_2 and UFPs; substantial contribution from secondary chemistry), between-neighborhood differences account for more of the SSD than within-neighborhood heterogeneity. (Adapted from Chambliss et al. 2021.)

respectively), but the minority for NO_2 and UFPs (37% and 28%, respectively). A subset of study areas accounted for a disproportionate share of local variation: for example, the San Francisco Financial District and East Oakland (24% of study population) accounted for roughly 50% of local exposure variation for NO and BC and 40% for UFPs and NO_2. These study areas represent denser urban settings with a greater mix of land uses. In general, we found slightly lower spatial variation in population-weighted exposure distributions within study areas, as compared with the variation in measured block-average concentrations. This discrepancy arises because populations tend not to be concentrated in the census blocks with the most extreme concentrations within our study area.

We compared our estimates of exposure gradients for NO_2 with those from the IEG model of the CACES for the year 2015 (Kim et al. 2020). Compared to the mobile monitoring data, the IEG model predicted higher median and mean exposure (respectively, 2.8 ppb [36%] and 2.5 ppb [30%] higher). However, our mobile monitoring observations show a substantially greater range of exposures, both within and between neighborhoods, as compared to the IEG model predictions (overlaid in gray in Figure 7). Whereas our mobile NO_2 observations show a ratio of 4.6 between the neighborhood median exposures among the highest and lowest study areas, the IEG model shows a ratio of only 1.6. Likewise, across the full Bay Area domain, we found a population-weighted interquartile range (IQR) of 6.1 ppb with mobile monitoring, as compared with a population-weighted IQR of 2.2 ppb for the IEG model. This result suggests that the national IEG model may miss some localized influences and may underestimate total population disparity and, by extension, the potential range of health risks. In the future, spatially resolved models for other pollutants, such as NO, BC, or UFPs, may enable further comparison between empirical model predictions and mobile monitoring observations.

Assessment of Exposure Disparities by Race and Ethnicity Across the Bay Area

To illustrate how mobile monitoring data can provide useful new insights that are not possible with other existing methods (Specific Aim 5), we undertook an environmental justice analysis of our full Bay Area measurements. We assessed how the spatial distribution of air pollutant concentrations across the Bay Area leads to disparate air pollution exposures by race

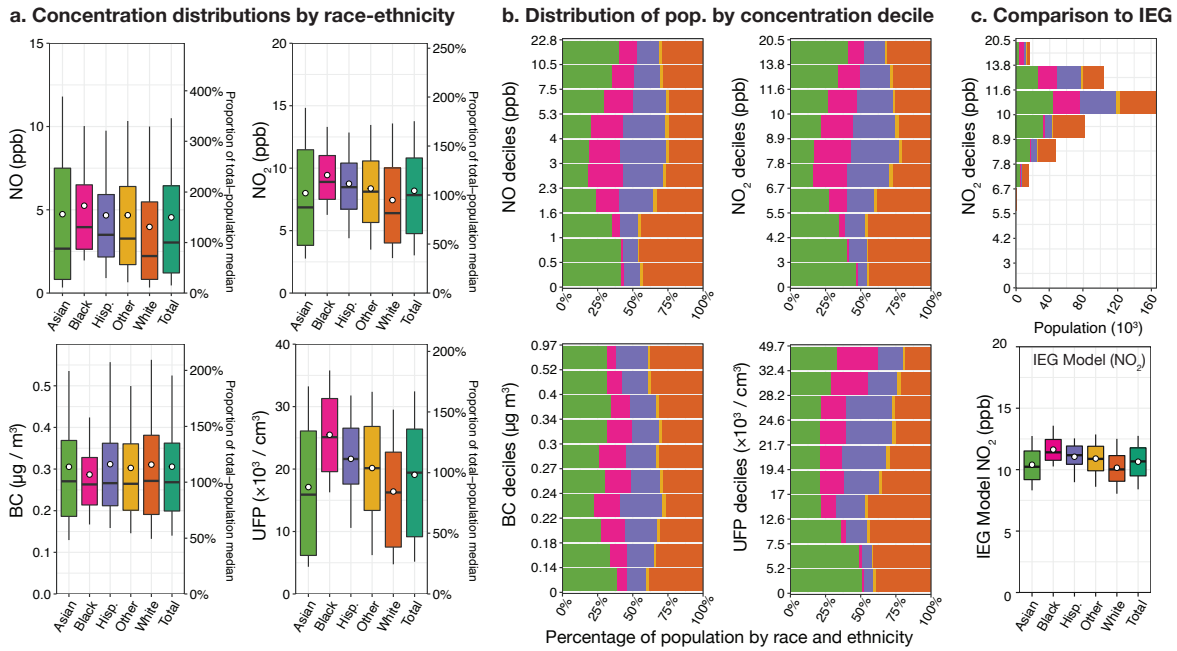


Figure 8. Variation in total exposure distributions from analysis M3 by racial or ethnic group, indicated by five distinct colors. The white, Asian, Black, and Other race groups only include those identifying as non-Hispanic. The distributions shown in the box-and-whisker plots **(a)** include the median (central bar), mean (white circle), upper and lower quartiles (box boundaries) and upper and lower 90th and 10th percentiles (whiskers). For NO, NO_2, and UFPs, population-weighted mean exposure concentrations are lowest for the White population and highest for the Black population. The division of the total-population exposure distribution into concentration deciles **(b)** shows the division of the population within the decile by race and ethnicity, with decile boundary concentrations indicated on the y axis and the racial and ethnic color key provided by the box-and-whisker plots. In general, the White population is most strongly represented in the lowest concentration deciles. Also evident in **(a)** and **(b)** are the broad and bimodal exposure distributions for the Asian population. Column **(c)** presents a comparison analysis for the CACES IEG exposure model, which predicts higher but less variable exposures for every racial and ethnic group (cf. Figure 7). Although both mobile monitoring measurements **(b)** and the IEG exposure model **(c)** predict roughly similar rank-ordering of mean NO_2 exposures by race and ethnicity, the IEG model substantially underestimates the within-group heterogeneity in exposures. (Adapted from Chambliss et al. 2021.)

and ethnicity. **Figure 8a** presents estimates of exposure distributions for five major population groups in our study area: White (33%), Asian (31%), Hispanic (21%), Black (14%), and other (2%). On average, the White population is exposed to lower NO, NO_2, and UFPs than other groups, with a median exposure 16% to 27% below the total population median, while medians for the Black and Hispanic populations are higher by 8% to 30% depending on pollutant (Figure 8a). The spatial detail provided by our method reveals nuances in disparity patterns beyond differences in medians. **Figure 8b** illustrates the weighting of each racial or ethnic group by the exposure deciles of the total population. Overall, the White population is strongly overrepresented in the lowest deciles of the concentration distributions. The Asian population is overrepresented at the extremes, with the high end driven by the communities in Downtown Oakland and the San Francisco Financial District, and the low end driven by less-polluted coastal locations. The Black and Hispanic populations are strongly underrepresented at the low end and concentrated toward the higher deciles, giving rise to higher average exposures for those groups. Apart from distinctly higher ranges of NO_2 and UFP exposure among Black and Hispanic populations, the range of exposures *within* racial and ethnic groups tends to be large compared with the range among groups. This finding holds especially for the Asian population (Figure 8a), which is bimodally distributed (Figure 8b) between some of the cleanest (coastal) and most polluted (downtown) areas.

Figure 9 illustrates how populations of each of the four largest racial and ethnic groups are distributed with respect to pollutant levels by neighborhood. Two key insights emerge from this visualization. First, for each pollutant, the shape and magnitude of the population–concentration distributions differ substantially among racial and ethnic groups. Second, Figure 9 highlights the role of regional demographic patterns in shaping the distinct distributions of exposure among racial and ethnic groups. For example, our Oakland study areas comprise most of the Hispanic and Black populations in our domain, whereas Oakland's study areas contribute only a small fraction of the White and Asian populations. These aggregate exposure profiles reveal an overall pattern of racial and ethnic disparities: higher concentration ranges in predomi-

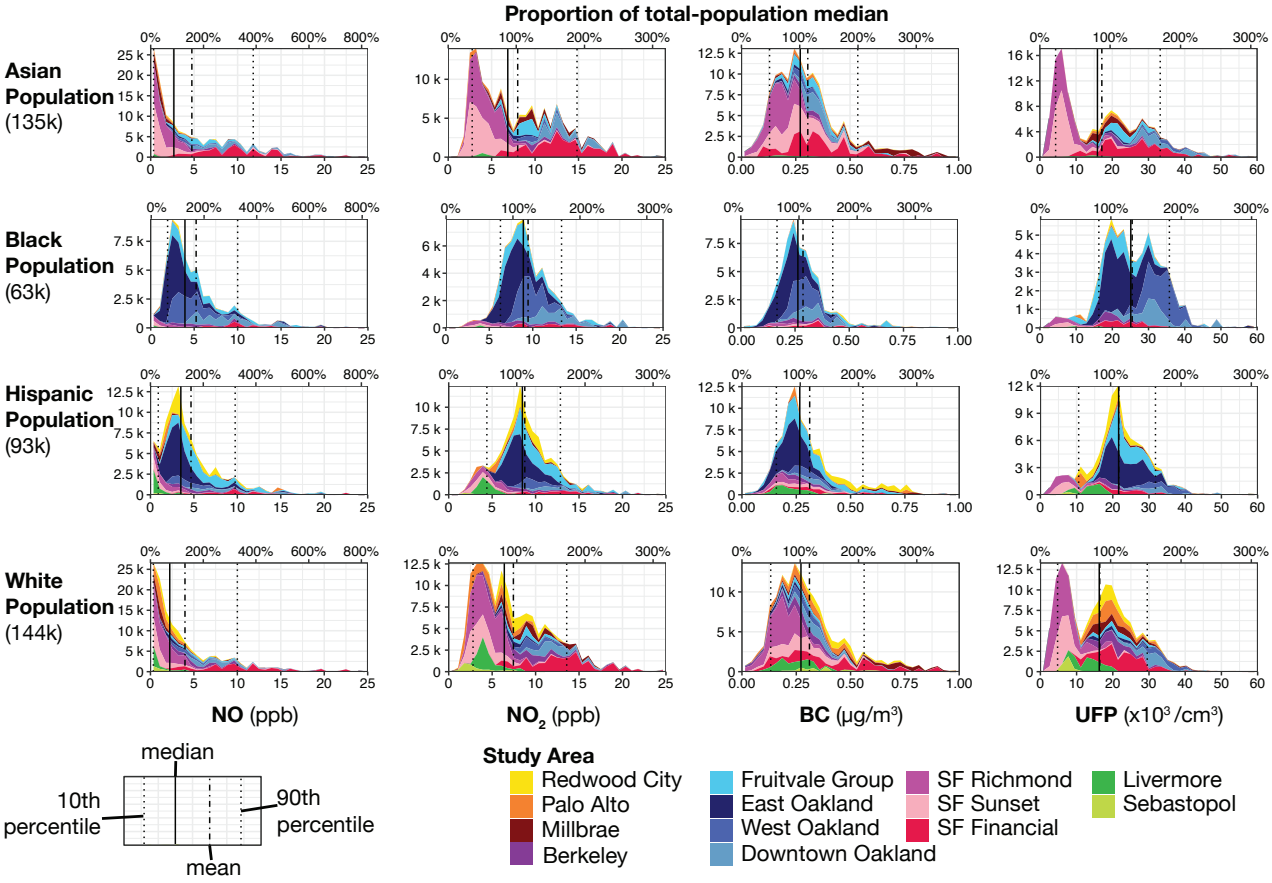


Figure 9. Population distribution. Distribution of exposure stratified by racial and ethnic groups (not shown is the "Other" racial category, population: 10,000). Height on the *y* axis indicates the population at a given concentration level summed over all study areas. Because the racial and ethnic composition differs sharply from neighborhood to neighborhood — a consequence of historical segregation patterns — concentration distributions in some study areas contribute much more to specific race and ethnicities (e.g., East Oakland for the Hispanic and Latino population and West Oakland for the Black population). Vertical lines show the indicated statistics (mean, median, 10th and 90th percentiles) for each race and ethnicity. (Adapted from Chambliss et al. 2021.)

nantly Black and Hispanic neighborhoods result in higher mean exposure for those groups. Notably, many of the neighborhoods with the highest average pollution exposures in our measurement dataset were subjected to overtly racially discriminatory housing policies, such as redlining (Lane et al. 2022).

Within our study domain, the national IEG model reproduced our observation of the lowest mean exposure for the White population and highest for the Black population, with moderately higher exposure for the Hispanic population (**Figure 8c**). However, the IEG model distributed the study populations into much narrower bands of concentrations as compared to our observations. Thus, it generally does not predict the same magnitude of disparity between mean concentrations among racial and ethnic groups and does not show a disproportionate share of people of Asian descent in the highest exposure categories. Modeled disparities may therefore miss an important dimension of racial and ethnic exposure disparity.

ASSESSING THE POTENTIAL OF STATISTICAL MODELS TO REDUCE SAMPLING EFFORT AND INCREASE SCALABILITY OF MOBILE MONITORING (ANALYSIS M4)

We examined strategies to efficiently develop air quality maps from mobile monitoring data, either via a "data-only" scheme that averages repeated measurements or via LUR-K models trained on repeated measurements. The overarching results of our analysis are that robust LUR-K models can be effectively developed even with very sparse mobile monitoring data, but that the data-only approach outperforms LUR-K in precision (R^2) after a small number of drive days.

Figure 10 illustrates representative results and residuals for these two approaches (left column: maps of daytime NO, right column: residuals of estimated NO in comparison to LUR long-term measurement (Figure 3a). Visual inspection suggests that each approach recreates some key features of

the long-term observed concentrations. NO concentrations are elevated strongly on highways (and modestly on arterials) relative to residential streets. Elevated NO levels in Downtown Oakland are evident in each of the maps. However, the full dataset (Figure 3a) contains numerous localized pollution hotspots at road intersections, industries, and other emissions sources, only some of which are reproduced in the Figure 10 maps. Similar patterns emerged for the BC maps, which are reported on in further detail by Messier and colleagues (2018).

Figure 10a depicts an example of a data-only map for NO concentrations in our Oakland domain (excluding East Oakland), which incorporates 4 days of sampling at each location (85 total hours of measurement, approximately 6% of the full measurement dataset). Even with only 4 days of measurement data per road segment, the key spatial patterns of NO are evident (median $R^2 = 0.65$; median NRMSE = 1.18), including pollution hotspots near some industries and intersections. However, there is evident "noise" — errors that appear to be approximately randomly distributed in space — that arises because of the limited number of samples at each location.

In Figure 10, panels b–e illustrate four alternative approaches to training an LUR-K model to predict concentrations at every 30-m road segment. Figure 10b represents a scenario in which the LUR-K model incorporates 2 full years of measurement data for training. It might seem illogical to develop an LUR-K model based on what is already a high-fidelity data-only map, but this model represents a conceptually useful starting point because it is trained on a dataset that represents a best-case scenario in terms of data quality. Accordingly, this model achieves the best performance ($R^2 = 0.60$, NRMSE = 1.05) of all LUR-K models in our analysis. The predictions capture regional and local variability, but often fail to correctly predict fine-scale hotspots. GIS covariates selected in each of the 10-folds included road-type indicators (highway roads, residential roads, etc.) local truck route indicator, greenness (NDVI) within a 50-m buffer, distance to the port, and elevation.

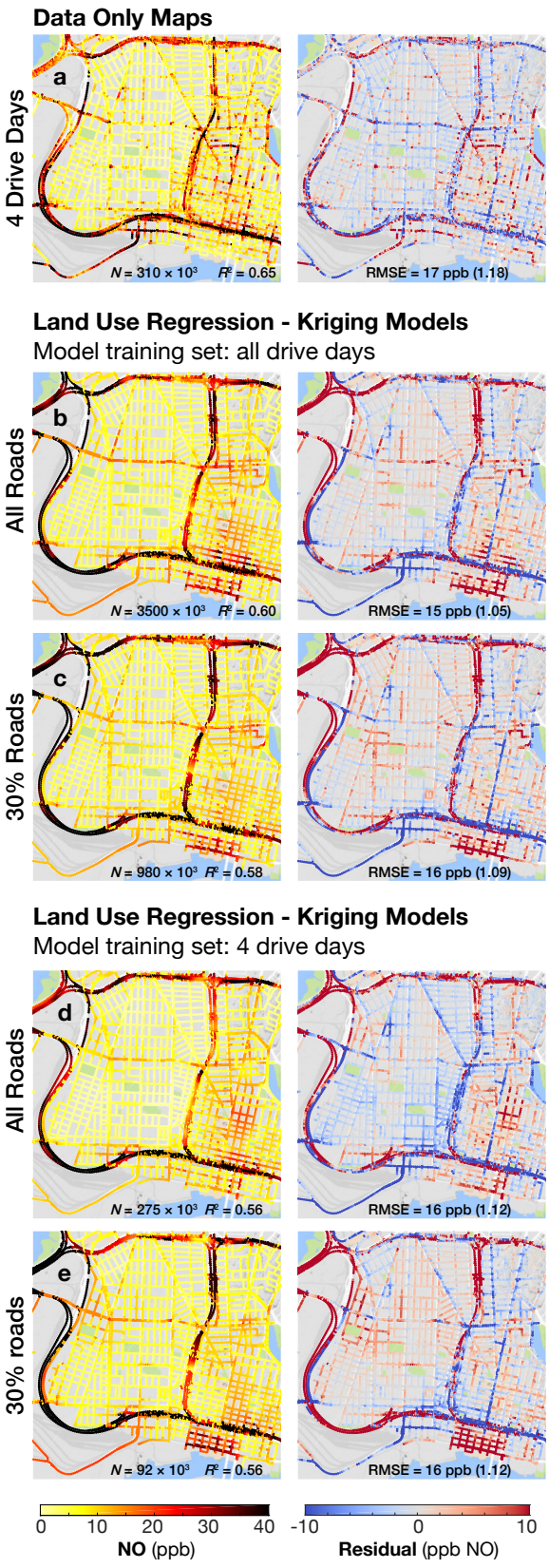


Figure 10. Examples of air quality maps constructed using sampling and/or modeling approaches. (a) Data-only map drawn from Monte Carlo subsample with 4 days at each road segment. Residuals are computed as the difference between subsampled 4-day map and the long-term concentrations shown in Figure 3a. **(b)** 10-fold cross-validation LUR-K prediction (and residual) surface trained on all the road segments and the entire 2-year dataset. **(c)** 10-fold cross-validation LUR-K predicting (and residual) surface trained the full Oakland dataset and 30% subsample of road segments. Note the similarity in predictions between **(b)** and **(c)**. **(d)** 10-fold cross-validation LUR-K predictions trained on a 4-day subsample for the full Oakland domain of road segments. **(e)** 10-fold cross-validation LUR-K predictions trained on a 4-day subsample and 30% of the road segments. The number of training data are given as *N*, and span a range from 92,000 independent points (~25 hour of sampling to cover 30% of roads in domain four times each) up to the size of the full Oakland dataset. For all panels, the R^2 is based on the log-transformed NO data and the RMSE is calculated in untransformed space. Normalized RMSE values are provided in parentheses. (Map data © 2018 Google; Figure adapted from Messier et al. 2018.)

One potentially attractive feature of combining mobile monitoring with LUR-K models is that effective models may be developed using a very limited set of training data. In Figure 10, panels c–e illustrate that the LUR model performance remains essentially similar even when the amount of model training data is substantially restricted. In Figure 10c, the training dataset is restricted to a subset of roads accounting for all highways and a random set of 30% of the nonhighway road network (20% of the full dataset hours), resulting in only a negligible change in model predictions and performance (median R^2 = 0.58, median NRMSE = 1.09). In Figure 10d, the training dataset is restricted to only 4 days of observation, resulting in a different model with a slight decrement in performance (median R^2 = 0.56, median NRMSE = 1.12), but with a large drop in training data requirements (6% of full dataset; ~80 hours). Figure 10e illustrates an example of a model trained on a highly restricted dataset (30% road coverage, 4 days of observation) with a dramatic reduction in data requirements (2% of full dataset; ~25 hours) and accompanied by a slight reduction in model performance (median R^2 = 0.56; median NRMSE = 1.12).

Figure 11 presents a more systematic evaluation of these results for the four approaches to reduce data requirements: (1) data-only mapping with reduced sampling frequency, (2) LUR-K models with reduced sampling frequency, (3) reduced road coverage, and (4) a combination of both lower sampling frequency and lower road coverage Figure 11a presents an evaluation of our results by the number of repeated drive days. Considering both R^2 and NRMSE, for fewer than 5 drive days over the full sampling domain, the LUR-K modeling approach tends to outperform a data-only map. This result arises because a predictive model overcomes some of the high degree of instability that occurs at the road segment level when the number of repeated samples is very small. However, with an increasing number of drive days, the performance of the LUR-K model saturates very quickly at an R^2 of 0.55–0.6 and an NRMSE of ~1–1.2. With only a small number of drive days (typically ~4–6 days), the data-only mapping approach outperforms the best LUR-K models (R^2 > 0.7), with dramatic increases in performance with increasing numbers of drive days before saturating with R^2 > 0.9 above about 15 drive days. Crucially, the data-only mapping approach results in a more spatially random set of errors as compared to the LUR-K model (compare for example Figure 10a vs. Figure 10b).

One potentially important advantage of the LUR-K modeling approach is that models can successfully be trained for a full domain while collecting only a small number of road segments, whereas by definition a data-only map requires measurements on every road segment. Figure 11b presents a key insight: the performance of an LUR-K model is quite insensitive to the percentage of road segments sampled. Even with only 20% to 30% of the surface streets in a domain sampled, the LUR-K model performance approaches the performance of the best LUR-K models we could develop, with R^2 ~ 0.55–0.6 and NRMSE between 1.05 and 1.2. Importantly, this reduced spatial coverage can be combined with reduced sampling frequency. While sampling this subset of roads just once produced a sharp decline in LUR model performance (R^2 ~ 0.4–0.5), training models on just four repeated samples produced models only mildly inferior to models using far more repeated samples. Thus, data requirements can be relaxed in two dimensions at once — both in terms of the

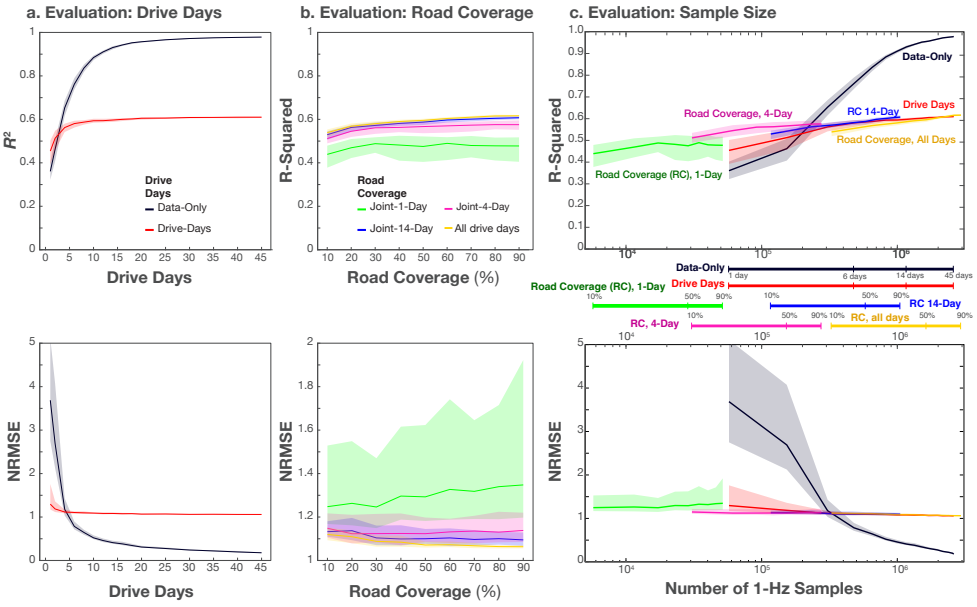


Figure 11. Evaluation of performance from scaling analysis. Performance evaluation for subsampled maps in terms of R^2 (upper row) and NRMSE (lower row). Column **(a)** presents subsampling schemes based on number of drive days per road segment. The black trace indicates the data-only mapping scheme, and the red trace indicates an LUR-K model trained with an equal number of measurements days on all road segments in the domain. Shaded area and solid lines represent, respectively, the interquartile range and median model performance for 100 Monte Carlo permutations of our full dataset. Column **(b)** presents an evaluation of LUR-K models trained on a specified fraction of road segments on the domain for four alternative levels of repetition frequency on each road segment. "Joint" models refer to cases with jointly reduced road coverage and drive days per road segment. Column **(c)** rescales results from **(a)** and **(b)** on a common abscissa representing the number of 1-Hz samples used to develop the exposure estimates. The line segments in the middle correspond in color to lines in the figures. The line segments provide context to the number of drive days or percentage of road segments for the corresponding number of 1-Hz samples. The *x* axis is on the log-scale. (Adapted from Messier et al. 2018.)

number of repeated samples and in terms of the number of roads sampled for training — while training an LUR-K model.

Figure 11c integrates these insights by presenting our evaluation statistics relative to the overall number of 1-Hz data samples required for producing a given result. The overall number of measurements used for our mapping simulations varies by more than a factor of 100 between our full dataset and our most restricted exercise. Overall, data-only maps substantially outperform the best LUR-K models we developed. Our LUR-K models approach their upper-bound performance quickly and then show little value from increased sampling. Over two orders of magnitude of sample size (~10–1000 hours of sampling), our LUR-K models consistently had R^2 ~ 0.5–0.6. As is evident in Figure 11b and c, reducing road coverage is a particularly effective approach to reducing the data requirements for training an LUR-K model. Thus, the value of this empirical modeling approach may be principally in that it can develop moderately good exposure predictions based on minimal sampling data, even if it never achieves the full potential of the much more measurement-intensive data-only mapping approach. The overall selection of approach (data-only mapping vs. predictive model) would need to consider the relative costs of potentially laborious ongoing data collection versus the availability of the specialized skillset and analytical time required to develop models.

In this analysis, we did not have access to detailed information on the types of industrial activities present at specific addresses, nor did we have information on smaller-scale sites and/or unpermitted commercial and industrial sites. In addition, we did not include publicly available information on businesses, restaurants, or other points of interest that are increasingly used in LUR model development (Lu et al. 2019, 2021). However, our measurement dataset is clearly influenced by localized pollution hotspots. Although some of these hotspots are sites that generate their own pollution, many also include locations where larger commercial vehicles congregate, such as factories and warehouses. Future work might also usefully consider whether further improvements in predictive model performance could be achieved using a more extensive set of spatial covariates, including data from ground-based and aerial/satellite imagery (Ganji et al. 2020; Qi and Hankey 2021; Qi et al. 2022; Weichenthal et al. 2019) and from scrape-able datasets of points of interest, such as the Google points-of-interest databases (Lu et al. 2019, Lu et al. 2021). Finally, although we did not account for meteorology, we speculate that accounting for the direction of the prevailing wind in Oakland might have offered an additional improvement in our model performance.

APPLICATION OF MOBILE MONITORING IN BANGALORE, INDIA (ANALYSIS M5)

The spatial mean (median) for the road segments in our core study domain in Malleshwaram was ~26 μg/m³ (15 μg/m³) for BC, ~81,000/cm³ (62,000/cm³) for UFPs, and 49 ppm (42 ppm) for ΔCO_2. Given the timing of our sampling, these estimates should be taken to represent typical morning-time concentrations on nonsummer weekdays. Mean on-road concentrations across the full study area were 47, 22, and 10 μg/m³ for BC; 116,000, 65,000, and 42,000/cm³ for UFPs; and 69, 44, 34 ppm for ΔCO_2, respectively, for highways, arterial roads, and residential roads. **Figure 12** presents maps for Malleshwaram of the spatial patterns of the median of drive-pass mean concentrations for these three pollutants. A clear structure in the spatial patterns of the pollutants emerged, with a strong rank ordering in concentration by road type (highways > arterials > residential streets), and rather similar spatial patterns for all three pollutants. As in the San Francisco Bay Area, localized multipollutant hotspots were evident in multiple locations throughout Malleshwaram, especially in congested traffic areas. Near these hotspots, our estimates of time-stable

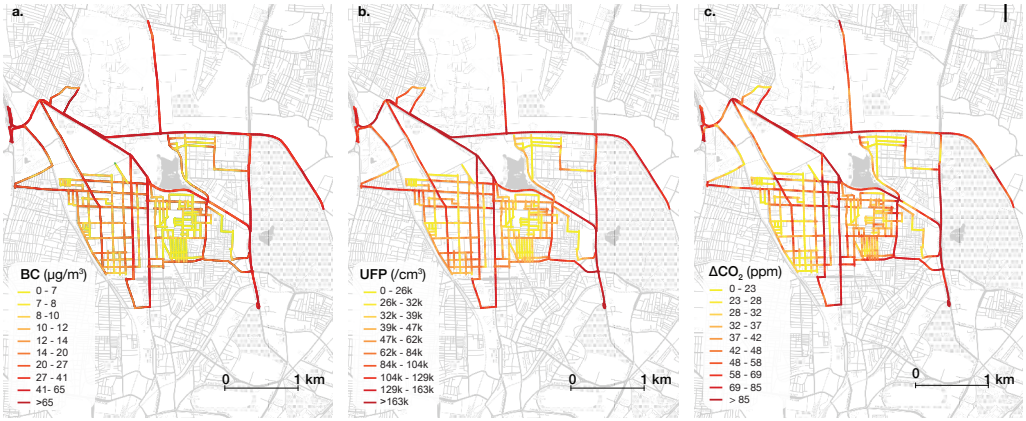


Figure 12. Maps of median of drive-pass mean concentrations of (a) BC, (b) UFPs, and (c) ΔCO2 for analysis M5 in Bangalore, India. Concentrations represent the median weekday 9 a.m.–1 p.m. concentration for each 30-m road segment over 22 repeated drive passes. Color scales are based on deciles of the road segment concentration distribution for each pollutant. Note the relatively high degree of concordance in spatial patterns among the three pollutants, with the lowest concentrations generally observed on residential streets, and the highest concentrations at congested highway junctions (compare with the domain maps in Figure 2).

median concentrations showed rather sharp spatial variation, with concentrations varying by a factor of about two- to threefold over distances of ~ 100 meters. These sharp spatial contrasts were also evident when comparing concentrations on quieter residential lanes with those on neighboring arterials one block away. This pattern was especially evident on the eastern side of Malleshwaram, where residential lanes are less used for pass-through vehicle traffic.

Our measurements of on-road concentrations are consistent with — or somewhat lower than — other reports of on-road air pollutant concentrations in India. For example, roughly a decade earlier and using very similar instrumentation sampling from an auto-rickshaw-based mobile laboratory, Apte and colleagues reported concentrations of 42 μg/m^3 BC and 280,000 UFPs/cm^3 in Delhi's traffic conditions (Apte et al. 2011). Our measured on-road increments of CO_2 were quite typical for urban roadway conditions (e.g., Westerdahl et al. 2005). However, our BC and UFP concentration measurements were dramatically higher than what we measured with a similar study design in the San Francisco Bay Area. Most spectacularly, on-road concentrations of BC were approximately 100 times higher in our Malleshwaram domain than in the Bay Area. This result likely arises principally from the very high share of poorly controlled diesel engines operating in the Bangalore vehicle fleet. Estimates of UFP concentrations in Malleshwaram were approximately fourfold higher than in the San Francisco Bay Area. Here, we caution that the choice of condensation particle counters used for measuring UFPs in Bangalore likely resulted in an undercount of the total number concentration, given the large number of combustion particles in the 2.5–10 nm size range that were detectable by our instrumentation in California but not India. In comparison with earlier results from Delhi, the considerably lower UFP concentrations in Bangalore may result in part from a sharp difference in the characteristics of the vehicle fleet: Delhi has a considerably higher prevalence of vehicles powered by compressed natural gas, which tend to have lower particle mass emissions but much higher particle number emissions (Apte et al. 2011; Hallquist et al. 2013).

Following the methods developed by Apte and colleagues (2017), we assessed the stability of our on-road concentration estimates in Malleshwaram in two different ways. First, we sought to understand the degree to which the overall heterogeneity in road segment median concentrations emerged from systematic spatial variability versus from stochastic variation in the time-resolved measurements that give rise to these road segment medians. To assess this question, we computed the intra-class correlation (ICC) metric on our concentration datasets grouped by 30-m road segments. The ICC metric varies from 0 to 1, and higher values of ICC indicate that the overall variance in a dataset is attributable to systematic differences between groups, rather than random heterogeneity within groups. For our application, an ICC of 0.75–1 would indicate large and systematic spatial differences in concentrations, with comparatively smaller contributions to heterogeneity from the stochastic variation. Here, we found ICC values of 0.81–0.92 for the three pollutants we measured in Malleshwaram, indicating that our estimated long-term spatial patterns were robust to the stochastic variations in concentration among repeated drive passes. (For comparison, Apte and colleagues [2017] reported an ICC of 0.8–0.95 for the pollutants measured in Oakland.)

Second, following the method from our analysis M4, we conducted Monte Carlo resampling of our data-only maps of BC and UFPs to assess how much repeated sampling would be needed to converge to the spatial patterns we measured with our full dataset of 22 repeated drives. **Figure 13** presents the results of this analysis, which are analogous to the results presented for the Oakland data-only maps in Figure 11. The gain in R^2 with the inclusion of each additional ride data increased rapidly until about 7 sampling days for both BC and UFPs and slowly thereafter. At ~10 sampling days, R^2 was 0.9. Similarly, NRMSE curves showed that the error rapidly decreased with the inclusion of each additional sampling day, with NRMSE < 20% after ~10 days for UFPs and ~15 days for BC. Despite the considerably different setting (Bangalore vs. Oakland) and the dramatically higher pollutant concentrations in India, these subsampling results suggest that mobile monitoring produces stable maps after about 10 drive days, with diminishing returns to precision from additional sampling beyond this level of repeated sampling. These conclusions about data-only maps are thus in line with the conclusions of Apte and colleagues (2017) (~10–20 drive days were usually sufficient) and from our analysis M4 presented above (diminishing returns were reached after 10–15 drive days; see Figure 12a).

DISCUSSION AND CONCLUSIONS

Our study had the overarching aims of assessing and validating the suitability of routine mobile monitoring for large-scale multipollutant air pollution exposure assessment,

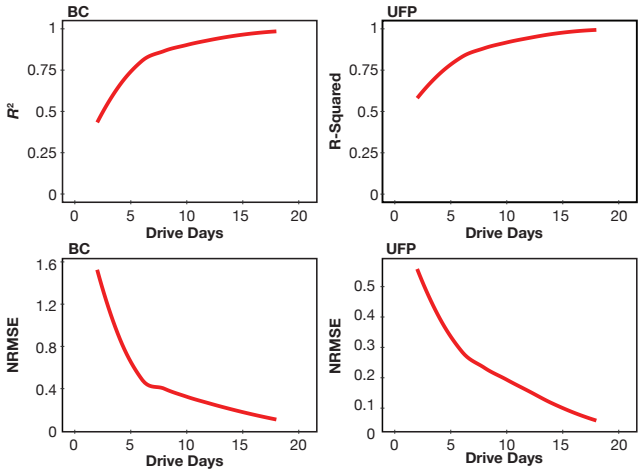


Figure 13. Monte Carlo subsampling analysis for the Malleshwaram neighborhood in Bangalore.

and to then apply this technique at scale. Specifically, we sought to (1) validate whether routine mobile monitoring could reproduce observed patterns of air pollution measured at fixed sites, (2) compare the insights derived from mobile monitoring with those from other measurement and modeling methods, (3) explore how statistical modeling could be paired with mobile sampling to further improve the efficiency and scalability of this exposure assessment technique, and (4) pilot the usage of mobile monitoring to fill data gaps in a lower-resource, low-data setting (Bangalore, India). Finally, (5) we sought an integrative assessment of cross-cutting questions related to the utility and efficacy of mobile monitoring vis-à-vis applications of mobile monitoring to atmospheric science, exposure assessment, environmental justice, air quality management, and other societally useful ends.

Table 2 summarizes analyses M1–M5 of this study, and how they relate to our specific aims. In analysis M1, we conducted an intensive summerlong sampling campaign in West Oakland, CA, where two Google Street View cars repeatedly mapped block-by-block air quality while driving around an exceptionally dense grid of 100 stationary BC monitors. In analysis M2, we explored the dynamics of UFPs in the San Francisco Bay Area and examined converging lines of evidence from both mobile and fixed-site monitoring about the association between UFPs and other traffic-related pollutants. In analysis M3, we scaled up our multipollutant mobile monitoring approach to 13 different neighborhoods with nearly 500,000 inhabitants, evaluated how the within- and between-neighborhood heterogeneity in concentrations affected population exposure and environmental disparities, and compared our insights with those from a widely used empirical exposure model. In analysis M4, we evaluated the advantages and trade-offs for coupling mobile monitoring with statistical LUR models to estimate intraurban variation in air pollution in a data-efficient manner. Finally, in analysis M5, we reproduced our mobile monitoring approach in a pilot study in Bangalore, India.

VALIDATION OF MOBILE MONITORING DATASETS

In analyses M1 and M2, we compared the insights we derived from mobile sampling with those from fixed-site sampling. Our detailed methods comparison during the 100 × 100 Study (see Chambliss et al. 2020 for details) (analysis M1) revealed that mobile monitoring was capable of capturing much — but not all — of the high-resolution spatial variation in air pollution that was observable by a dense fixed-site network for measuring BC. Our analysis revealed mobile monitoring could capture the overall concentration gradients with moderate-to-high reliability from the cleanest to most polluted locations in our West Oakland sampling domain during some of the lowest-concentration conditions in the year (daytime during the summer). There were many advantages for considering this mobile-to-fixed site comparison for BC measurement, as opposed to other pollutants. BC is a primary conserved pollutant that provides us with sharp spatial gradients to compare our methods. While a robust low-cost sensor exists for BC (Caubel et al. 2018), NO and NO_2 remain difficult to reliably measure with a low-cost sensor, and UFPs cannot yet be measured with one. However, the relatively high signal-to-noise ratio for our mobile measurements of BC meant that our mobile measurements had high associated measurement uncertainty. Our analysis also revealed this imprecision was the key factor in degrading our comparison between mobile and fixed-site observations. Indeed, had we conducted this validation study during a different time period (e.g., during nonrainy winter days) or in a more polluted setting, such as Bangalore, it is likely that we would have found an even higher degree of concordance between mobile and fixed-site observations.

Analysis M1 also revealed several complementary aspects of mobile and fixed-site monitoring used in combination. As shown in Figure 4, our mobile monitoring provided a large amount of spatial coverage (and associated representativeness) that was missing from the fixed-site monitoring network. This finding is especially notable given that the 100 × 100 Study monitoring network was likely the densest neighborhood-scale BC monitoring effort undertaken to date (Caubel et al. 2019). This analysis also highlights the utility of a fixed-site network in providing distributed real-time observations, which are especially valuable in the context of noisy mobile BC measurements that have especially poor precision when considered on a time-resolved basis (Chambliss et al. 2020).

It is instructive to note here that we used fixed-site monitoring here in a limited capability, whereby the measurement datasets were used independently. Nonetheless, there is substantial potential to integrate mobile and fixed-site data into a hybrid monitoring product that combines the distinct advantages of each sampling paradigm (i.e., dense spatial coverage and continuous temporal coverage, respectively). However, the data analysis methods for fusing air quality datasets that are, respectively, temporally and spatially sparse are still not highly developed and remain an important area for further research. In new work not supported by HEI, we are developing new methods for spatiotemporal modeling that can fuse mobile and fixed-site sensor data. The data from our 2017 100 × 100 Study provided a unique opportunity to demonstrate and evaluate these new methods.

In analysis M2, we found strong alignment between mobile and fixed-site observations in studying the seasonal variable association between UFPs and NO_x. From both mobile and fixed-site data, we found a tight association in the spatial patterns and diurnal cycles of UFPs and NO_x during winter months, with strong evidence that traffic is a major source during winter conditions. In the summer months, mobile and fixed-site data again are in concurrence, but the result diverges from the winter data. Here, we found that daytime UFP concentrations in the Bay Area appear to be strongly influenced by secondary new particle formation events, with little association between UFPs and other traffic-related

pollutants, such as NO_x. A consequence is that the intraurban spatial gradients of UFPs are strongly attenuated during summer daytime conditions. This result illustrates how routine mobile monitoring can reveal facets of exposure patterns that are not well characterized by short-term studies. The result also further demonstrates the value in future monitoring efforts of combining detailed mobile mapping with a small number of fixed-site monitors that can provide time-resolved data. In the absence of the routine fixed-site UFP monitoring data in the Bay Area, which is quite rare for regulatory agencies to undertake in the United States, we likely would have not been able to explain this unique feature of the spatial patterns in our mobile dataset.

A related question — albeit one that we did not investigate directly in this project — is whether the spatially resolved concentration fields that we mapped in this project would in fact constitute a valid exposure measure that would provide utility for epidemiological studies. The preliminary evidence on this topic is positive. To date, our San Francisco Bay Area concentration datasets from this study have been used as exposure measures for four epidemiological studies. Using electronic health records for a population of ~41,900 adults living in a ~25-km^2 area of Oakland, our collaborators found clear and statistically significant associations between road-segment level estimates of NO, NO_2, and BC and adverse cardiovascular events among the older population (Alexeeff et al. 2018). Second, in a study of ~8,800 births in the San Francisco Bay Area, our collaborators found an elevated risk of preterm birth for children born to Black and Latina mothers (Riddell et al. 2021). Third, for electronic health records of pregnant women (N = 1,095) living in Downtown and West Oakland, our collaborators identified statistically significant associations between pollutant exposure (especially NO_2 and UFPs) and preeclampsia (Goin et al. 2021). Finally, for ~25,700 older subjects living in the Oakland domain, our collaborators found statistically significant associations between BC and (especially) NO_2 exposures and multiple measures of healthcare expenditures captured by electronic health records (Alexeeff et al. 2022). These studies illustrate how hyperlocalized exposure measures can enable health studies even on very small and spatially localized populations and support the inference that our on-road concentration measures may provide a useful measure of air pollution exposures. Recently, several other Canadian and European studies have utilized mobile monitoring data as an exposure assessment strategy for assessing within-urban and even within-country spatial variations in air pollution exposure (Bouma et al. 2023; Weichenthal et al. 2020).

Of course, there are limitations in the use of mobile monitoring data for epidemiological exposure estimates. First, there are concerns of temporal representativeness. Our mobile monitoring campaigns were designed to represent daytime spatial patterns on weekdays, which are not necessarily representative of conditions at night or on weekends, and therefore not completely representative of annual average conditions. Although we found in analysis M1 that BC concentrations measured on-road had good correspondence with measurements at building façades in Oakland, for populations who live far from roads, as is more common in rural and suburban areas, on-road measurements may not be especially useful in estimating exposures. More broadly, it may be infeasible to provide mobile monitoring for the very large populations used for some large epidemiological cohort studies of air pollution (e.g., ACS, Pope et al. 2002; Medicare, Di et al. 2017; and CanCHEC, Crouse et al. 2015a). Finally, people don't spend their full lives at home, so the benefits of more spatially precise exposure assessments from mobile monitoring may be offset by an additional misclassification error that arises from not accounting for population mobility. In sum, although mobile monitoring may enable new types of epidemiological studies by capturing sharp spatial gradients over small spatial areas, the potential of mobile monitoring data for epidemiological studies is not fully resolved with this study and is likely context-dependent.

COMPARISON OF INSIGHTS FROM MOBILE MONITORING WITH OTHER MEASUREMENT APPROACHES

Whereas routine mobile monitoring excels at providing spatially *intensive* exposure measurements (i.e., block-by-block coverage), the mobile monitoring approach does not automatically guarantee spatially *extensive* measurements. For extensive air pollution mapping, approaches such as satellite remote sensing and LUR modeling can provide large-scale exposure estimates. National and global exposure model datasets using remote sensing and/or LUR are increasingly common and often publicly available. In analysis M3, we investigated how the widely used CACES IEG exposure models (Kim et al. 2020) performed in estimating NO_2 at the scale of census blocks. Here, we compared our insights from the full-scale mobile monitoring of NO_2 across 13-neighborhood Bay Area domain of ~500,000 people (Figure 1a) with the CACES IEG predictions. Whereas the census-block predictions of the CACES NO_2 model performed quite well in reproducing the rank-ordering of the neighborhood-median NO_2 concentrations that we measured by mobile monitoring (Figure 7), the model missed nearly all of the within-neighborhood exposure heterogeneity that we measured. This finding suggests that traditional national-scale LUR models may struggle to predict local-scale heterogeneity within individual communities.

We found that the national-scale CACES IEG model performed adequately in estimating the scale of average racial and ethnic group NO_2 disparities that we measured across our study domain (Figure 8). This somewhat surprising result arises because between-neighborhood segregation is a stronger driver of the systemic inequality in air pollution than is the hyperlocal within-neighborhood heterogeneity (Figure 9). To put this differently: if one lives in a racially segregated U.S. city, the demographics of one's own city block are likely to be quite similar to the demographics of the next few city blocks in either direction, but demographics might be quite different

for a neighborhood a few miles away. In contrast, our data show how air pollution levels can vary sharply over length scales of even just a few city blocks. This hyperlocal variation therefore matters considerably more for heterogeneity in the overall population exposure than it does for estimating racial and ethnic exposure disparities (Chambliss et al. 2021). Other recent work has affirmed this conclusion using other lines of evidence. Clark and colleagues illustrated using the CACES IEG models that estimates of racial and ethnic group exposure inequalities in NO_2 and $PM_{2.5}$ are highly scale-dependent at coarser scales of aggregation (e.g., counties and states), but show little sensitivity at the finest spatial scales (e.g., U.S. census blocks, block groups, and tracts) (Clark et al. 2022). Demetillo and colleagues have demonstrated that comparatively coarse satellite remote sensing estimates (length scales of several km) of NO_2 are capable of resolving substantial racial and ethnic group exposure inequalities, even though NO_2 concentrations themselves vary over considerably finer scales (Demetillo et al. 2020, Demetillo et al. 2021).

ASSESSMENT OF MOBILE MONITORING IN BANGALORE, INDIA

Our measurements in Bangalore, India (analysis M4) demonstrated that mobile monitoring is a viable technique for estimating fine-scale concentration gradients in an LMIC city. Our mapping exercise revealed fine-scale patterns of spatial heterogeneity in air pollutant concentrations within our study neighborhood of Malleshwaram that were highly reminiscent of what we measured in the San Francisco Bay Area, albeit at dramatically higher concentrations. The high coherence in spatial patterns among the three pollutants we measured, as well as good metrics of measurement stability (e.g., ICC by road segment), provide confidence that the observed spatial patterns are likely to reflect true conditions, rather than measurement artifacts. Another key aspect of the success of this part of the study likely relates to our process. Our study team in Bangalore was led by co-authors Kushwaha, Upadhya, and Sreekanth. Although highly experienced in air pollution measurement and spatial statistics, the field team was conducting a mobile monitoring study for the first time as part of this report. Working together as a team, we adapted the data collection and analysis protocols from Oakland to the Indian context and undertook substantial learning-by-doing to develop a workable study protocol for Bangalore. While the difficulty of this endeavor should not be understated, the fact that we successfully developed these results in Bangalore provides evidence that this mobile monitoring approach can be successfully adapted to LMIC contexts.

Some key limitations from our Bangalore mobile monitoring experience should also be emphasized here. First, the relatively low traffic speeds in Bangalore (typically ~10-15 km/hr, slower during rush hour) limit the amount of data that can be collected in a single sampling session, which erodes the efficiency of the mobile monitoring approach. For our study, the practical impact of these low traffic speeds was amplified by the limited battery life (~ 4 to 5 hours) of our portable instruments. Second, our data collection approach required study personnel to join the vehicle's driver for each sampling run to aid in wayfinding and ensure adequate instrument performance. This consideration meant that our data collection ended up being especially labor-intensive and physically taxing. These logistical considerations should not be overlooked in designing future studies, but some aspects could be resolved with a more refined mobile laboratory and instrumentation package. Third, given that very dense traffic can be quite common in some parts of Bangalore, the interpretation of our data is somewhat difficult. Our measurements are representative of on-road concentrations. Our conclusions from Oakland in analysis M1 imply that on-road measurements can succeed in representing exposure-relevant concentrations at the front façade of buildings and homes. However, it is not necessarily clear that our Bangalore measurements in especially congested locales would necessarily be representative of exposure concentrations outside of the immediate roadway environment, especially when our car was trapped in dense gridlock. This concern might be assuaged by the rather moderate on-road ΔCO_2 increments we measured (95th percentile of road segment averages = 85 ppm ΔCO_2). Future work could investigate this possible concern by applying a paired mobile and fixed-site study design like that from analysis M1 in an Indian context. Indeed, it is reasonable to expect that some of the measurement precision constraints we experienced in analysis M1 might be less of an issue in India, given the higher signal-to-noise ratio under polluted conditions and the longer averaging times afforded by slower traffic.

SCALING MOBILE MONITORING: TO MODEL OR TO MEASURE?

In analysis M5, we compared the relative strengths of data-only mobile monitoring (i.e., a model-free approach of repeated sampling) and spatiotemporally restricted observations to train an LUR model. In general, the latter approach is more common in the exposure assessment community, and LUR models based on mobile monitoring are increasingly common in major health studies (e.g., Brauer et al. 2008; Kerckhoffs et al. 2022; Weichenthal et al. 2020; Wu et al. 2021). Our assessment in analysis M5 suggests that there are some clear trade-offs to this hybrid approach that fuses mobile measurements with LUR-K models. On the benefits side, we demonstrated that relatively little repeated sampling (~4–10 days) is required to build stable or parsimonious LUR models and that a modest (10%–30%) stratified sample of the urban road network was sufficient to adequately train our LUR-K models in Oakland (Messier et al. 2018). Thus, there are very large potential gains in data-collection efficiency in this hybrid modeling approach compared with extensive repeated sampling of every road segment. If one were to scale up such a modeling approach, it's conceivable that a rotating sampling protocol inspired by this approach in a finite and modest number of large and small cities might be sufficient to develop a multipollutant mobile-monitoring prediction

surface for an entire country's urban population, which might provide substantial benefit for future health studies. On the other hand, we learned some significant disadvantages of the LUR-K modeling approach relative to the model-free data-only mapping approach. One relative benefit of data-only mapping is that the data analysis approach itself is far simpler: one does not need to undertake the time-consuming extra step of developing model covariates, training a model, and then rigorously evaluating that model. Another key benefit of data-only mapping is that the prediction fidelity of a data-only map can dramatically exceed that of an LUR-K model (Figure 11) and the residuals for a data-only map contain relatively little spatial bias structure, which is not so for LUR-K models (compare Figure 10a with the rest of Figure 10).

It is worthwhile to contemplate whether the data-only repeated mobile mapping approach we demonstrated here is truly scalable. In Apte and colleagues (2017), we postulated that ~500 air mapping vehicles would be required to generate an annual air quality map of the 25-largest urban areas in the United States, accounting for 50% of the total U.S. urban population. Here, we comment on what it would take to accomplish mobile monitoring at such a large scale. First, we believe that considerably fewer vehicles may be necessary for such a campaign, in part because there may not be a strong need to map air pollution in each city annually. For example, a once-in-3-years mapping exercise for each city would allow for ongoing assessment of the long-term evolution of air pollution sources and patterns in each urban area and only require 100 air mapping vehicles with some optimization of sampling intensity. Second, we estimate required capital costs of $150,000 per vehicle at scale to equip each vehicle with fast-response monitors for BC, NO, NO_2, CO_2, and UFPs. Although lower-cost sensors are available, the very careful attention to calibration and performance that is required may not yet be feasible at such a large scale. If the capital costs of a vehicle are amortized over a nominal 7.5-year lifespan, the capital cost per vehicle-year is ~$20,000/yr. Operational costs for a full-time driver ($40/hr salary and benefits, 2000 hr/yr) add another $80,000/yr; fueling and maintenance costs for the vehicle and instruments add another $50,000/yr. Thus, the total cost of a vehicle system's usage may come closer to $150,000/yr.

In our experience, the routine mapping techniques we developed here could be readily transferred from a research setting to a more operational setting, perhaps provided by private vendors. Beyond the capital and operational costs we discussed above, additional costs would include providing routine calibration quality assurance and quality control (1 technician per 10 cars), developing daily drive plans (1 dispatcher per 10 cars), routine analysis (1 analyst per 10 cars), and central management (1 per 10–20 cars). At average salary and benefit costs of ~$200,000/yr per staff member, staffing and analysis costs might add $75,000/yr per vehicle. Rounding up, the entire enterprise of routine mobile monitoring might cost $250,000/yr per vehicle. With a fleet of 100 mapping vehicles providing once-in-a-3-year coverage, this would equate to $25 million in annual expenditures, or approximately $1 million per year for each of the 25 largest urban areas in the United States. These costs might be quite reasonable. For example, consider that the annual monitoring and analysis budgets for the Bay Area Air Quality Management District (San Francisco Bay Area) and the South Coast Air Quality Management District (Los Angeles Area) are ~$10 million/yr and ~$30 million/yr, respectively. When considering the power of such data for identifying emissions hotspots, characterizing exposure distributions and environmental inequalities, and enabling accountability studies, these expenditures might be quite worthwhile.

It is also instructive to note that this sampling approach is labor intensive. Many of the costs associated with this rough budgetary sketch do not arise directly from the capital costs of the vehicle and instruments, but rather from those of the drivers, technicians, analysts, and management. (This same feature is true of the budgets of conventional air monitoring networks). Thus, the costs of routine mobile monitoring might scale to lower levels in LMIC settings where labor costs are also generally lower. However, this enterprise would require unique skills and expertise that are often in short supply in LMIC settings. The prospects for scaling this approach more generally in a setting like India are perhaps somewhat less favorable than in the United States, especially given the highly constrained public budgets for environmental protection.

OTHER LESSONS LEARNED

It is worthwhile to convey a few final lessons learned during the execution of this study. First, as has been noted by others (e.g., Brantley et al. 2014), subtle aspects of study design are exceptionally important in designing a mobile monitoring data collection scheme that can robustly estimate time-averaged pollution patterns. In general, it is far preferable to design a measurement campaign to have spatiotemporally balanced measurements from the outset, rather than to attempt to "nudge" a measurement distribution that was not collected in a temporally representative manner. One reason why it is advantageous to combine mobile and fixed-site multipollutant monitoring is that continuous time series from multiple fixed sites can aid in the assessment of the temporal representativeness of mobile monitoring data (Chambliss et al. 2021).

There is some irony that much of this study relied on Google Street View cars. In our study, these vehicles were operated under the direction of our research team, permitting a carefully balanced and repeated study design. However, more generally, these vehicles are tasked to revisit the same locations to collect imagery only very infrequently, typically every few years, which would pose challenges for interpreting very temporally sparse air pollution data. Other vehicle fleets, such as urban taxis, which tend to have more random and frequent drive patterns, might be better suited to the task of the routine mobile monitoring. However, there may also be value in simply having vehicle fleets that are dedicated to the

task of routine mobile monitoring, thereby permitting more carefully designed sampling strategies.

A second observation is that not all pollutants are equally suited to mobile monitoring. In general, pollutants with a high degree of spatial variation and a low degree of temporal variation appear to be the best suited to the type of routine mobile-monitoring approach we employed here. In the environments we considered in this study, it appears that urban primary pollutants fit this description. However, other important pollutants, such as $PM_{2.5}$, do not match this description well. Efforts to develop representative maps of $PM_{2.5}$ via mobile monitoring often encounter two difficulties. First, the spatial variation in $PM_{2.5}$ is generally small: whereas the time-stable patterns of many primary pollutants may vary by a factor of 2–8× across even a small neighborhood, the spatial variation in $PM_{2.5}$ across an entire urban domain might be only 1.3–2×. This is because, for most urban areas, the regional background aerosol is the dominant source of fine particle mass (Jimenez et al. 2009). Related to this point, a second challenge for mapping $PM_{2.5}$ via mobile monitoring is that episodic variation in regional background concentrations affects $PM_{2.5}$ much more strongly than for more localized primary pollutants. Accordingly, with a small number of repeated samples, the temporal sampling bias for $PM_{2.5}$ can be quite large. We speculate that these two factors explain why comparatively few well-designed mobile monitoring studies have reported robust results for $PM_{2.5}$. Thankfully, there are other modern exposure assessment techniques, such as low-cost sensors and remote sensing, that can provide high-quality, spatially resolved estimates of urban $PM_{2.5}$ concentrations.

IMPLICATIONS OF FINDINGS

The past decade has witnessed a renaissance in tool development for spatially resolved exposure assessment. These techniques include model-based or model-informed approaches, including chemical transport models, reduced-complexity mechanistic models, and empirical models — all of which increasingly leverage aspects of machine learning. Observational techniques that have progressed rapidly include low-cost sensors, wearable exposure monitors, chemically resolved real-time instrumentation (e.g., mass spectrometry for aerosol and gas-phase species), and satellite remote sensing. Increasingly, there is crossover and inspiration occurring within this multitude of rapidly developing techniques. Moreover, there is a growing appreciation that different research questions require different observational strategies.

Routine mobile monitoring, too, has experienced a rapid expansion in both method development and practical applications. At the time this study's proposal was submitted for review at HEI in early 2017, few studies had been published exploring the idea of routine mobile monitoring. Our pilot study using Google Street View cars for air pollution mapping in Oakland was published in mid-2017 and has now been cited over 450 times. Since then, a few dozen scientific studies using these cars for mobile air pollution monitoring have been published, and scientist-led campaigns have been conducted in places as diverse as London, Copenhagen, Amsterdam, Dublin, Austin, Houston, Los Angeles, Salt Lake City, and Denver. Of course, repeated mobile monitoring need not be conducted using Google Street View cars as a platform, as we demonstrated with our own analyses in Bangalore. Using a wide array of platforms and instrumentation, from trash trucks with low-cost sensors (deSouza et al. 2020), to telemetry-equipped taxis (Yu et al. 2022), to mobile laboratories equipped with aerosol mass spectrometers (Gu et al. 2018; Shah et al. 2018), repeated mobile monitoring is now finding widespread application in air quality studies. As one metric of the expansion of the scientific literature on mobile monitoring, we conducted a Web of Science search for the keywords "air pollution" and "mobile monitoring." Since 1979, 209 peer-reviewed articles with this specific keyword combination have been indexed — 113 were published between 2018 and 2022, as compared to 60 papers published over the previous 5-year period. Similarly, of the 4,871 papers (Web of Science) citing this body of literature, 3,406 have been published since 2018. Routine mobile monitoring is also increasingly common outside of the academic sphere. Using a suite of lower-cost sensors developed by Aclima, Google Street View cars are collecting mobile monitoring data during ordinary driving for imagery in dozens of cities on multiple continents. Aclima has developed its own mobile platforms and has attracted substantial government funding — often from environmental justice–inspired monitoring initiatives — to develop air pollution datasets for cities in California, New York, and elsewhere. As of the writing of this report (early 2023), the U.S. EPA has selected nine community-led mobile monitoring efforts for $4M in funding nationwide, and the State of California is preparing a $30M Statewide Mobile Monitoring initiative (CA: *https://ww2.arb.ca.gov/statewide-mobile-monitoring-initiative*, NY: *https://www.governor.ny.gov/news/governor-hochul-announces-launch-first-statewide-mobile-air-monitoring-initiative*).

One key lesson from this emerging body of work is that routine mobile monitoring is quite useful for capturing time-stable patterns of pollution, especially for pollutants with sharp spatial gradients. Because many localized pollution hotspots are indicative of proximate pollution sources, mobile monitoring can provide useful screening-level identification of potentially unknown sources that may be off the "radar screen" of scientists and government agencies. In contrast, mobile monitoring approaches likely are less useful, at least on their own, for identifying air pollution sources that are highly episodic or transient. Although our study provides examples of how mobile and fixed-site monitoring data can be integrated and compared, future analysis efforts could go much further to integrate mobile and fixed-site observation data using mechanistic or statistical modeling approaches. Again, the specific approach likely needs to be dictated by the analytical goals, which may be quite diverse, ranging from spatiotemporal exposure prediction to environmental justice studies to constraining source emission rates.

There remains a great societal need for continued air pollution observations. In the United States efforts to mitigate climate change are accelerating, with a strong emphasis on transitioning major distributed sources of fossil carbon emissions (such as vehicles and households) to electric-powered substitutes. This transition could sharply redraw the spatial geography of urban air pollutant emissions and their impacts and higher-resolution air pollution observation techniques, including mobile monitoring, will be important to quantify their impact. At the same time, major environmental policy efforts in the United States are increasingly focused on reducing exposure disparities for historically disadvantaged groups. Policy efforts for the clean energy transition and environmental justice are increasingly intertwined, as is evident, for example, in the recently passed Inflation Reduction Act of 2022 (Levy 2022). Notably, this act earmarked more than $100 million dollars for enhanced air pollution monitoring, with a key focus on environmental justice communities. In LMIC, the context is different: monitoring networks are not yet fully established and can be designed from the ground up to take advantage of more recent developments in measurement and modeling (Brauer et al. 2019; Gani et al. 2022); at the same time, efforts to control pollution are also accelerating. Thus, the lessons learned from this study and many other studies of hyperlocal air pollution variation in urban areas will be applicable to devising monitoring strategies, accountability studies, and epidemiological analyses that can assess the real-world impact of these ambitious efforts to protect human health and well-being.

ACKNOWLEDGMENTS

In addition to the HEI funding that enabled this work to be possible, other funders contributed resources to the research presented herein. We particularly wish to acknowledge the support of Environmental Defense Fund and the Center for Air, Climate, and Energy Solutions (CACES), which was supported under Assistance Agreement No. R835873 awarded by the U.S. Environmental Protection Agency (U.S. EPA). This work has not been formally reviewed by U.S. EPA. The views expressed in this document are solely those of authors and do not necessarily reflect those of the Agency. The U.S. EPA does not endorse any products or commercial services mentioned in this publication. In addition, Google provided in-kind support for the mobile monitoring in California.

This work benefitted greatly from the mentorship of Prof. Michael Brauer (University of British Columbia) and Prof. Adam Szpiro (University of Washington). Many others contributed to the development of the ideas and datasets reported in the core journal articles that we published under HEI support. Here, we wish to acknowledge Jonathan Gingrich, Carlos Pinon, Brian LaFranchi, Jai Asundi, Pratyush Agrawal, Crystal Upperman, Melissa Lunden, Allen Robinson, Julian Marshall, Chelsea Preble, Julien Caubel, Troy Cados, Ramon Alvarez, Thomas Kirchstetter, Jonathan Choi, Steven Hamburg, Christopher Portier, Ananya Roy, and Roel Vermeulen. We also gratefully acknowledge Maria Harris, Elena Craft, Cassandra Ely, Fern Uennatornwaranggoon, Karin Tuxen-Bettman, Alexander Cooper, Rebecca Moore, David Herzl, Arjun Raman, Mille Chu Baird, Margaret Gordon, Brian Beveridge, Phil Martien, David Holstius, Rivkah Gardner-Frolick, Mark Campmier, and the Aclima and Google Street View operations teams.

DATA AVAILABILITY

Data from analyses M1–M4 are publicly available. Time-averaged maps of air pollution at the 30-m road segment aggregation level are available for the years 2015–2016 in the online supplemental information (SI) of Apte and colleagues (2017). Time-averaged maps of concentrations for the full 3-year campaign, aggregated to U.S. Census block geographies, are available in the supplementary materials of Chambliss and colleagues (2021). The underlying spatiotemporally resolved mobile monitoring data from the campaign are available on request from Google via the following link: *https://docs.google.com/forms/d/e/1FAIpQLSf_4GIkK1tmVMFRSxz42KgvOM3Z3NGeOFFje_FS8FBbz1vTig/viewform.*

Time-resolved fixed-site BC data from the 100 × 100 campaign of ABCD sensors in Oakland are available from Environmental Defense Fund via the following link: *https://www.edf.org/airqualitymaps/oakland/request-time-series-air-pollution-data.*

REFERENCES

Alexeeff SE, Roy A, Shan J, Liu X, Messier K, Apte JS, et al. 2018. High-resolution mapping of traffic related air pollution with Google Street View cars and incidence of cardiovascular events within neighborhoods in Oakland, CA. Environ Health 17:38; *https://doi.org/10.1186/s12940-018-0382-1*.

Alexeeff SE, Roy A, Shan J, Ray GT, Quesenberry CQ, Apte J, et al. 2022. Association between traffic related air pollution exposure and direct health care costs in northern California. Atmos Environ 287:119271; *https://doi.org/10.1016/j.atmosenv.2022.119271*.

Apte JS, Kirchstetter TW, Reich AH, Deshpande SJ, Kaushik G, Chel A, et al. 2011. Concentrations of fine, ultrafine, and black carbon particles in auto-rickshaws in New Delhi, India. Atmos Environ 45:4470–4480; *https://doi.org/10.1016/j.atmosenv.2011.05.028*.

Apte JS, Marshall JD, Brauer M, Cohen AJ. 2015. Addressing global mortality from ambient $PM_{2.5}$. Environ Sci Technol 49:8057–8066; *https://doi.org/10.1021/acs.est.5b01236*.

Apte JS, Messier KP, Gani S, Brauer M, Kirchstetter TW, Lunden MM, et al. 2017. High-resolution air pollution mapping with Google Street View cars: Exploiting big data. Environ Sci Technol 51:6999–7008; *https://doi.org/10.1021/acs.est.7b00891*.

Apte J, Seraj S, Chambliss S, Hammer M, Southerland V, Anenberg S, et al. 2021. Air inequality: Global divergence in urban fine particulate matter trends. ChemRxiv 26 May; https://doi.org/10.26434/chemrxiv.14671908.v1.

Arnott WP, Moosmüller H, Rogers FC, Jin T, Bruch R. 1999. Photoacoustic spectrometer for measuring light absorption by aerosol: Instrument description. Atmos Environ 33:2845–2852; *https://doi.org/10.1016/S1352-2310(98)00361-6*.

Arnott WP, Moosmüller H, Sheridan PJ, Ogren JA, Raspet R, Slaton WV, et al. 2003. Photoacoustic and filter-based ambient aerosol light absorption measurements: Instrument comparisons and the role of relative humidity. J Geophys Res Atmos 108(D1); *https://doi.org/10.1029/2002JD002165*.

Ban-Weiss GA, Lunden MM, Kirchstetter TW, Harley RA. 2009. Measurement of black carbon and particle number emission factors from individual heavy-duty trucks. Environ Sci Technol 43:1419–1424; *https://doi.org/10.1021/es8021039*.

Betha R, Spracklen DV, Balasubramanian R. 2013. Observations of new aerosol particle formation in a tropical urban atmosphere. Atmos Environ 71:340–351; *https://doi.org/10.1016/j.atmosenv.2013.01.049*.

Boogaard H, Kos GPA, Weijers EP, Janssen NAH, Fischer PH, van der Zee SC, et al. 2010. Contrast in air pollution components between major streets and background locations: Particulate matter mass, black carbon, elemental composition, nitrogen oxide and ultrafine particle number. Atmos Environ 45:650–658; *https://doi.org/10.1016/j.atmosenv.2010.10.033*.

Bouma F, Janssen NAH, Wesseling J, van Ratingen S, Strak M, Kerckhoffs J, et al. 2023. Long-term exposure to ultrafine particles and natural and cause-specific mortality. Environ Int 175:107960; *https://doi.org/10.1016/j.envint.2023.107960*.

Boy M, Kulmala M. 2002. Nucleation events in the continental boundary layer: Influence of physical and meteorological parameters. Atmos Chem Phys 2:1–16; *https://doi.org/10.5194/acp-2-1-2002*.

Brantley HL, Hagler GSW, Kimbrough ES, Williams RW, Mukerjee S, Neas LM. 2014. Mobile air monitoring data-processing strategies and effects on spatial air pollution trends. Atmos Meas Tech 7:2169–2183; *https://doi.org/10.5194/amt-7-2169-2014*.

Brauer M, Guttikunda SK, Nishad KA, Dey S, Tripathi SN, Weagle C, et al. 2019. Examination of monitoring approaches for ambient air pollution: A case study for India. Atmos Environ 216:116940; *https://doi.org/10.1016/j.atmosenv.2019.116940*.

Brauer M, Lencar C, Tamburic L, Koehoorn M, Demers P, Karr C. 2008. A cohort study of traffic-related air pollution impacts on birth outcomes. Environ Health Perspect 116:680–686; *https://doi.org/10.1289/ehp.10952*.

Brines M, Dall'Osto M, Beddows DCS, Harrison RM, Gómez-Moreno F, Núñez L, et al. 2015. Traffic and nucleation events as main sources of ultrafine particles in high-insolation developed world cities. Atmos Chem Phys 15:5929–5945; *https://doi.org/10.5194/acp-15-5929-2015*.

Bukowiecki N, Dommen J, Prevot ASH, Richter R, Weingartner E, Baltensperger U. 2002. A mobile pollutant measurement laboratory-measuring gas phase and aerosol ambient concentrations with high spatial and temporal resolution. Atmos Environ 36:5569–5579; *https://doi.org/10.1016/S1352-2310(02)00694-5*.

Bullard RD. 1993. The legacy of American apartheid and environmental racism. Available: *https://scholarship.law.stjohns.edu/cgi/viewcontent.cgi?article=1460&context=jcred* [accessed 21 October 2023].

Bullard RD. 2020. Dumping in Dixie: Race, Class, and Environmental Quality, Third Edition. Oxfordshire UK:Routledge.

Burnett R, Chen H, Szyszkowicz M, Fann N, Hubbell B, Pope CA 3rd, et al. 2018. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. Proc Natl Acad Sci USA; 115:9592–9597; *https://doi.org/10.1073/pnas.1803222115*.

Carvalho H. 2016. The air we breathe: Differentials in global air quality monitoring. Lancet Respir Med 4:603–605; *https://doi.org/10.1016/s2213-2600(16)30180-1*.

Caubel JJ, Cados TE, Kirchstetter TW. 2018. A new black carbon sensor for dense air quality monitoring networks. Sensors 18:738; *https://doi.org/10.3390/s18030738*.

Caubel JJ, Cados TE, Preble CV, Kirchstetter TW. 2019. A distributed network of 100 black carbon sensors for 100 days of air quality monitoring in West Oakland, California. Environ Sci Technol 53:7564–7573; *https://doi.org/10.1021/acs.est.9b00282*.

Chambliss SE, Pinon CPR, Messier KP, LaFranchi B, Upperman CR, Lunden MM, et al. 2021. Local- and regional-scale racial and ethnic disparities in air pollution determined by long-term mobile monitoring. Proc Natl Acad Sci USA; 118:e2109249118; *https://doi.org/10.1073/pnas.2109249118*.

Chambliss SE, Preble CV, Caubel JJ, Cados T, Messier KP, Alvarez RA, et al. 2020. Comparison of mobile and fixed-site black carbon measurements for high-resolution urban pollution mapping. Environ Sci Technol 54:7848–7857; *https://doi.org/10.1021/acs.est.0c01409*.

Clark LP, Harris MH, Apte JS, Marshall JD. 2022. National and intraurban air pollution exposure disparity estimates in the United States: Impact of data-aggregation spatial scale. Environ Sci Technol Lett 9:786–791; *https://doi.org/10.1021/acs.estlett.2c00403*.

Clark LP, Millet DB, Marshall JD. 2017. 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; *https://doi.org/10.1289/ehp959*.

Cohen AJ, Brauer M, Burnett R, Anderson HR, Frostad J, Estep K, et al. 2017. 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. The Lancet 389:1907–1918; *https://doi.org/10.1016%2FS0140-6736(17)30505-6*.

Costabile F, Birmili W, Klose S, Tuch T, Wehner B, Wiedensohler A, et al. 2009. Spatio-temporal variability and principal components of the particle number size distribution in an urban atmosphere. Atmos Chem Phys 9:3163–3195; *https://doi.org/10.5194/acp-9-3163-2009*.

Crouse DL, Peters PA, Hystad P, Brook JR, van Donkelaar A, Martin RV, et al. 2015a. Ambient $PM_{2.5}$, O_3, and NO_2 exposures and associations with mortality over 16 years of follow-up in the Canadian census health and environment cohort (CANCHEC). Environ Health Perspect 123:1180–1186; *https://doi.org/10.1289/ehp.1409276*.

Crouse DL, Peters PA, Villeneuve PJ, Proux M-O, Shin HH, Goldberg MS, et al. 2015b. Within- and between-city contrasts in nitrogen dioxide and mortality in 10 canadian cities: A subset of the Canadian census health and environment cohort (CANCHEC). J Expo Sci Environ Epidemiol 25:482–489; *https://doi.org/10.1038/jes.2014.89*.

Demetillo MAG, Harkins C, McDonald BC, Chodrow PS, Sun K, Pusede SE. 2021. Space-based observational constraints on NO_2 air pollution inequality from diesel traffic in major us cities. Geophy Res Lett 48:e2021GL094333; *https://doi.org/10.1029/2021GL094333*.

Demetillo MAG, Navarro A, Knowles KK, Fields KP, Geddes JA, Nowlan CR, et al. 2020. Observing nitrogen dioxide air pollution inequality using high-spatial-resolution remote sensing measurements in Houston, Texas. Environ Sci Technol 54:9882–9895; *https://doi.org/10.1021/acs.est.0c01864.*

deSouza P, Anjomshoaa A, Duarte F, Kahn R, Kumar P, Ratti C. 2020. Air quality monitoring using mobile low-cost sensors mounted on trash-trucks: Methods development and lessons learned. Sustain Cities Soc 60:102239; *https://doi.org/10.1016/j.scs.2020.102239*.

Di Q, Wang Y, Zanobetti A, Wang Y, Koutrakis P, Choirat C, et al. 2017. Air pollution and mortality in the medicare population. N Engl J Med 376:2513–2522; *https://doi.org/10.1056/nejmoa1702747*.

Gani S, Bhandari S, Patel K, Seraj S, Soni P, Arub Z, et al. 2020. Particle number concentrations and size distribution in a polluted megacity: The Delhi aerosol supersite study. Atmos Chem Phys 20:8533–8549; *https://doi.org/10.5194/acp-20-8533-2020*.

Gani S, Chambliss SE, Messier KP, Lunden MM, Apte JS. 2021. Spatiotemporal profiles of ultrafine particles differ from other traffic-related air pollutants: Lessons from long-term measurements at fixed sites and mobile monitoring. Environ Sci Atmos 1:558–568; *https://doi.org/10.1039/D1EA00058F*.

Gani S, Pant P, Sarkar S, Sharma N, Dey S, Guttikunda SK, et al. 2022. Systematizing the approach to air quality measurement and analysis in low and middle income countries. Environ Res Lett 17:021004; *https://doi.org/10.1088/1748-9326/ac4a9e*.

Ganji A, Minet L, Weichenthal S, Hatzopoulou M. 2020. Predicting traffic-related air pollution using feature extraction from built environment images. Environ Sci Technol 54:10688–10699; *https://doi.org/10.1021/acs.est.0c00412*.

Goin DE, Sudat S, Riddell C, Morello-Frosch R, Apte JS, Glymour MM, et al. 2021. Hyperlocalized measures of air pollution and preeclampsia in Oakland, California. Environ Sci Technol 55:14710-14719; *https://doi.org/10.1021/acs.est.1c02151*.

Gu P, Li HZ, Ye Q, Robinson ES, Apte JS, Robinson AL, et al. 2018. Intracity variability of particulate matter exposure is driven by carbonaceous sources and correlated with land-use variables. Environ Sci Technol 52:11545–11554; *https://doi.org/10.1021/acs.est.8b03833*.

Hallquist ÅM, Jerksjö M, Fallgren H, Westerlund J, Sjödin Å. 2013. Particle and gaseous emissions from individual diesel and CNG buses. Atmos Chem Phys 13:5337–5350; *https://doi.org/10.5194/acp-13-5337-2013*.

Hansen ADA, Rosen H, Novakov T. 1984. The aethalometer — An instrument for the real-time measurement of optical absorption by aerosol particles. Sci Total Environ 36:191–196; *https://doi.org/10.1016/0048-9697(84)90265-1*.

Health Effects Institute (HEI). 2010. Traffic-Related Air Pollution: A Critical Review of The Literature on Emissions, Exposure, and Health Effects. Special Report 17. Boston, MA:Health Effects Institute.

Health Effects Institute (HEI). 2022. Systematic Review and Meta-Analysis of Selected Health Effects of Long-Term Exposure to Traffic-Related Air Pollution. Special Report 23. Boston, MA:Health Effects Institute.

Hudda N, Cheung K, Moore KF, Sioutas C. 2010. Inter-community variability in total particle number concentrations in the eastern Los Angeles air basin. Atmos Chem Phys 10:11385–11399; *https://doi.org/10.5194/acp-10-11385-2010*.

Jbaily A, Zhou X, Liu J, Lee T-H, Kamareddine L, Verguet S, et al. 2022. Air pollution exposure disparities across U.S. population and income groups. Nature 601:228–233; *https://doi.org/10.1038/s41586-021-04190-y*.

Jerrett M, Burnett RT, Ma R, Pope CA 3rd, Krewski D, Newbold KB, et al. 2005. Spatial analysis of air pollution and mortality in Los Angeles. Epidemiology 16:727–736; *https://doi.org/10.1097/01.ede.0000181630.15826.7d*.

Jimenez JL, Canagaratna MR, Donahue NM, Prevot ASH, Zhang Q, Kroll JH, et al. 2009. Evolution of organic aerosols in the atmosphere. Science 326:1525–1529; *https://doi.org/10.1126/science.1180353*.

Karner AA, Eisinger DS, Niemeier DA. 2010. Near-roadway air quality: Synthesizing the findings from real-world data. Environ Sci Technol 44:5334–5344; *https://doi.org/10.1021/es100008x*.

Kerckhoffs J, Khan J, Hoek G, Yuan Z, Ellermann T, Hertel O, et al. 2022. Mixed-effects modeling framework for Amsterdam and Copenhagen for outdoor NO_2 concentrations using measurements sampled with Google Street View cars. Environ Sci Technol 56:7174–7184; *https://doi.org/10.1021/acs.est.1c05806*.

Kim S-Y, Bechle M, Hankey S, Sheppard L, Szpiro AA, Marshall JD. 2020. Concentrations of criteria pollutants in the contiguous U.S., 1979–2015: Role of prediction model parsimony in integrated empirical geographic regression. PLOS ONE 15:e0228535; *https://doi.org/10.1371/journal.pone.0228535*.

Kulmala M, Vehkamäki H, Petäjä T, Dal Maso M, Lauri A, Kerminen VM, et al. 2004. Formation and growth rates of ultrafine atmospheric particles: A review of observations. J Aerosol Sci 35:143–176; *https://doi.org/10.1016/j.jaerosci.2003.10.003*.

Kushwaha M, Sreekanth V, Upadhya AR, Agrawal P, Apte JS, Marshall JD. 2022. Bias in $PM_{2.5}$ measurements using collocated reference-grade and optical instruments. Environ Monit Assess 194:610; *https://doi.org/10.1007/s10661-022-10293-4*.

Lane HM, Morello-Frosch R, Marshall JD, Apte JS. 2022. Historical redlining is associated with present-day air pollution disparities in U.S. cities. Environ Sci Technology Lett 9:345–350; *https://doi.org/10.1021%2Facs.estlett.1c01012*.

Levy JI. 2022. 2022 Inflation Reduction Act: Climate Investments are Public Health Investments. Am J Public Health 112:1525–1525; *https://doi.org/10.2105/ajph.2022.307089*.

Liu J, Clark Lara P, Bechle Matthew J, Hajat A, Kim S-Y, Robinson AL, et al. 2021. Disparities in air pollution exposure in the United States by race/ethnicity and income, 1990–2010. Environ Health Perspect 129:127005; *https://doi.org/10.1289/ehp8584*.

Lu T, Lansing J, Zhang W, Bechle MJ, Hankey S. 2019. Land use regression models for 60 volatile organic compounds: Comparing Google Point of Interest (POI) and city permit data. Sci Total Environ 677:131–141; *https://doi.org/10.1016/j.scitotenv.2019.04.285*.

Lu T, Marshall JD, Zhang W, Hystad P, Kim S-Y, Bechle MJ, et al. 2021. National empirical models of air pollution using microscale measures of the urban environment. Environ Sci Technol 55:15519–15530; *https://doi.org/10.1021/acs.est.1c04047*.

Marshall JD, Nethery E, Brauer M. 2008. Within-urban variability in ambient air pollution: Comparison of estimation methods. Atmos Environ 42:1359–1369; *https://doi.org/10.1016/j.atmosenv.2007.08.012*.

Martin RV, Brauer M, van Donkelaar A, Shaddick G, Narain U, Dey S. 2019. No one knows which city has the highest concentration of fine particulate matter. Atmos Environ X 3:100040; *https://doi.org/10.1016/j.aeaoa.2019.100040*.

Messier KP, Akita Y, Serre ML. 2012. Integrating address geocoding, land use regression, and spatiotemporal geostatistical estimation for groundwater tetrachloroethylene. Environ Sci Technol 46:2772–2780; *https://doi.org/10.1021/es203152a*.

Messier KP, Chambliss SE, Gani S, Alvarez R, Brauer M, Choi JJ, et al. 2018. Mapping air pollution with Google Street View cars: Efficient approaches with mobile monitoring and land use regression. Environ Sci Technol 52:12563–12572; *https://doi.org/10.1021/acs.est.8b03395*.

Morello-Frosch R, Pastor M, Sadd J. 2001. Environmental justice and southern California's "riskscape": The distribution of air toxics exposures and health risks among diverse communities. Urban Aff Rev 36:551–578; *https://doi.org/10.1177/10780870122184993*.

O'Dowd C, Monahan C, Dall'Osto M. 2010. On the occurrence of open ocean particle production and growth events. Geophy Res Lett 37; *https://doi.org/10.1029/2010GL044679*.

Padró-Martínez LT, Patton AP, Trull JB, Zamore W, Brugge D, Durant JL. 2012. Mobile monitoring of particle number concentration and other traffic-related air pollutants in a near-highway neighborhood over the course of a year. Atmos Environ 61:253–264; *https://doi.org/10.1016%2Fj.atmosenv.2012.06.088*.

Pope CA 3rd, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K, et al. 2002. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. JAMA 287:1132–1141; *https://doi.org/10.1001/jama.287.9.1132*.

Qi M, Dixit K, Marshall JD, Zhang W, Hankey S. 2022. National land use regression model for NO_2 using Street View imagery and satellite observations. Environ Sci Technol 56:13499–13509; *https://doi.org/10.1021/acs.est.2c03581*.

Qi M, Hankey S. 2021. Using Street View imagery to predict street-level particulate air pollution. Environ Sci Technol 55:2695–2704; *https://doi.org/10.1021/acs.est.0c05572*.

Raaschou-Nielsen O, Andersen ZJ, Beelen R, Samoli E, Stafoggia M, Weinmayr G, et al. 2013. Air pollution and lung cancer incidence in 17 European cohorts: Prospective analyses from the European study of cohorts for air pollution effects (ESCAPE). Lancet Oncol 14:813–822; *https://doi.org/10.1016/s1470-2045(13)70279-1*.

Riddell CA, Goin DE, Morello-Frosch R, Apte JS, Glymour MM, Torres JM, et al. 2021. Hyper-localized measures of air pollution and risk of preterm birth in Oakland and San Jose, California. Int J Epidemiol 50:1875–1885; *https://doi.org/10.1093/ije/dyab097*.

Rothstein R. 2017. The Color of Law: A Forgotten History of How Our Government Segregated America. New York:Liveright Publishing.

Salma I, Borsós T, Weidinger T, Aalto P, Hussein T, Dal Maso M, et al. 2011. Production, growth and properties of ultrafine atmospheric aerosol particles in an urban environment. Atmos Chem Phys 11:1339–1353; *https://doi.org/10.5194/acp-11-1339-2011*.

Sellegri K, Laj P, Venzac H, Boulon J, Picard D, Villani P, et al. 2010. Seasonal variations of aerosol size distributions based on long-term measurements at the high altitude Himalayan site of Nepal climate observatory-pyramid (5079 m), Nepal. Atmos Chem Phys 10:10679–10690; *https://doi.org/10.5194/acp-10-10679-2010*.

Shah RU, Robinson ES, Gu P, Robinson AL, Apte JS, Presto AA. 2018. High-spatial-resolution mapping and source apportionment of aerosol composition in Oakland, California, using mobile aerosol mass spectrometry. Atmos Chem Phys 18:16325–16344; *https://doi.org/10.5194/acp-18-16325-2018*.

Shen XJ, Sun JY, Zhang YM, Wehner B, Nowak A, Tuch T, et al. 2011. First long-term study of particle number size distributions and new particle formation events of regional aerosol in the north China plain. Atmos Chem Phys 11:1565–1580; *https://doi.org/10.5194/acp-11-1565-2011*.

Tasoglou A, Subramanian R, Pandis SN. 2018. An inter-comparison of black-carbon-related instruments in a laboratory study of biomass burning aerosol. Aerosol Sci Technol 52:1320–1331; *https://doi.org/10.1080/02786826.2018.1515473*.

Tessum CW, Paolella DA, Chambliss SE, Apte JS, Hill JD, Marshall JD. 2021. $PM_{2.5}$ polluters disproportionately and systemically affect people of color in the United States. Sci Adv 7:eabf4491; *https://doi.org/10.1126/sciadv.abf4491*.

Vakkari V, Laakso H, Kulmala M, Laaksonen A, Mabaso D, Molefe M, et al. 2011. New particle formation events in semi-clean South African savannah. Atmos Chem Phys 11:3333–3346; *https://doi.org/10.5194/acp-11-3333-2011*.

Weichenthal S, Hatzopoulou M, Brauer M. 2019. A picture tells a thousand…exposures: Opportunities and challenges of deep learning image analyses in exposure science and environmental epidemiology. Environ Int 122:3–10; *https://doi.org/10.1016/j.envint.2018.11.042*.

Weichenthal S, Olaniyan T, Christidis T, Lavigne E, Hatzopoulou M, Van Ryswyk K, et al. 2020. Within-city spatial variations in ambient ultrafine particle concentrations and incident brain tumors in adults. Epidemiology 31:177–183; *https://doi.org/10.1097/ede.0000000000001137*.

Westerdahl D, Fruin S, Sax T, Fine PM, Sioutas C. 2005. Mobile platform measurements of ultrafine particles and associated pollutant concentrations on freeways and residential streets in Los Angeles. Atmos Environ 39:3597–3610; *https://doi.org/10.1016/j.atmosenv.2005.02.034*.

Whitby KT, Clark WE, Marple VA, Sverdrup GM, Sem GJ, Willeke K, et al. 1975. Characterization of California aerosols—i. Size distributions of freeway aerosol. Atmos Environ 9:463–482; *https://doi.org/10.1016/0004-6981(75)90107-9*.

Wu AH, Fruin S, Larson TV, Tseng C-C, Wu J, Yang J, et al. 2021. Association between airport-related ultrafine particles and risk of malignant brain cancer: A multiethnic cohort study. Cancer Res 81:4360–4369; *https://doi.org/10.1158/0008-5472.can-21-1138*.

Yu YT, Xiang S, Zhang T, You Y, Si S, Zhang S, et al. 2022. Evaluation of city-scale disparities in $PM_{2.5}$ exposure using hyper-localized taxi-based mobile monitoring. Environ Sci Technol 56:13584–13594; *https://doi.org/10.1021/acs.est.2c02354*.

Zhang KM, Wexler AS, Zhu YF, Hinds WC, Sioutas C. 2004. Evolution of particle number distribution near roadways. Part ii: The 'road-to-ambient' process. Atmos Environ 38:6655–6665; *https://doi.org/10.1016/j.atmosenv.2004.06.044*.

ABOUT THE AUTHORS

Joshua S. Apte is an Associate Professor of Environmental Engineering and Environmental Health Sciences at the University of California, Berkeley. Previously, he was an Assistant Professor at the University of Texas (2015–2020) and the ITRI-Rosenfeld Postdoctoral Fellow at the Lawrence Berkeley National Laboratory (2013–2014). Apte received his Ph.D. in Energy and Resources from the University of California, Berkeley, in 2013. Apte's research focuses on methods for air pollution exposure assessment using large-scale measurement and modeling datasets.

Sarah E. Chambliss received her Ph.D. in Environmental Engineering from the University of Texas at Austin in 2020 and is currently a postdoctoral fellow at the University of Texas in the Department of Statistics and Data Science. Chambliss's current research focuses on statistical methods for air pollution exposure assessment.

Kyle Messier is a tenure-track Stadtman Investigator at the National Institute of Environmental Health Sciences, where he directs the Spatiotemporal Health Analytics Group. Messier was previously an EDF/Kravis postdoctoral fellow at the University of Texas from 2015–2018 and earned his Ph.D. in Environmental Science and Engineering from the University of North Carolina, Chapel Hill, in 2015. Messier's research focuses on spatiotemporal and geostatistical methods for environmental exposure assessment.

Shahzad Gani is Assistant Professor in the Centre for Atmospheric Sciences at the Indian Institute of Technology, Delhi. Gani received his Ph.D. in Environmental Engineering from the University of Texas at Austin in 2019 and was then a postdoctoral researcher at the Institute for Atmospheric and Earth System Research (INAR) at the University of Helsinki. Gani's research focuses on the atmospheric dynamics of particulate matter, especially in the formation and fate of ultrafine particles.

Adithi R. Upadhya is a Data Scientist at ILK Labs, Bangalore, and holds an MS in Geoinformatics from Bharati Vidhyapeeth Institute of Environment Education and Research, India. Upadhya's research focuses on open-source data analysis tools for environmental exposure datasets.

Meenakshi Kushwaha is the Co-founder and Director of Research at ILK Labs, Bangalore, and holds an MPH in Environmental and Occupational Health from the University of Washington. Kushwaha's research interests lie at the intersection of environmental health, equity, and development in low- and middle-income countries.

Sreekanth Vakacherla is a Senior Scientific Advisor at Environmental Defense Fund. Previously, Dr. Sreekanth was Senior Research Scientist at the Centre for the Study of Science, Technology & Policy (CSTEP), Bangalore, India, and earned his Ph.D. in Atmospheric Aerosols from Andhra University in 2008. Dr. Sreekanth's research focuses on the physicochemical characterization of particulate matter and its application to exposure assessment and atmospheric processes.

OTHER PUBLICATIONS RESULTING FROM THIS RESEARCH

Analysis M1 (mobile vs. fixed site comparison): Chambliss SE, Preble CV, Caubel JJ, Cados T, Messier KP, Alvarez RA, et al. 2020. Comparison of mobile and fixed-site black carbon measurements for high-resolution urban pollution mapping. Environ Sci Technol 54:7848–7857; *doi:10.1021/acs.est.0c01409*.

Analysis M2 (spatiotemporal aspects of UFPs vs. NO_x from mobile and fixed-site perspective): Gani S, Chambliss SE, Messier KP, Lunden MM, Apte JS. 2021. Spatiotemporal profiles of ultrafine particles differ from other traffic-related air pollutants: Lessons from long-term measurements at fixed sites and mobile monitoring. Environ Sci Atmos 1:558–568; *doi.10.1039/D1EA00058F*.

Analysis M3 (within vs. between neighborhood comparison in San Francisco Bay Area): Chambliss SE, Pinon CPR, Messier KP, LaFranchi B, Upperman CR, Lunden MM, et al. 2021. Local- and regional-scale racial and ethnic disparities in air pollution determined by long-term mobile monitoring. Proc Natl Acad Sci USA 118:e2109249118; *doi:10.1073/pnas.2109249118*.

Analysis M4 (scaling with LUR-K models): Messier KP, Chambliss SE, Gani S, Alvarez R, Brauer M, Choi JJ, et al. 2018. Mapping air pollution with Google Street View cars: Efficient approaches with mobile monitoring and land use regression. Environ Sci Technol 52:12563–12572; *doi:10.1021/acs.est.8b03395*.

Research Report 216, *Scalable Multipollutant Exposure Assessment Using Routine Mobile Monitoring Platforms*, J.S. Apte et al.

INTRODUCTION

Accurately estimating people's exposure to various pollutants is essential for evaluating and understanding the health effects associated with the pollutants. Accurate estimates of exposure are also essential for identifying disparities in exposure so that policies can be developed to reduce such disparities if they exist. It is challenging, however, to estimate exposures to outdoor air pollutants that vary highly in space and time. Most air pollution datasets tend to have adequate resolution and accuracy either over space or time, but not both. For example, researchers typically conduct targeted, short-term sampling campaigns used to develop land use regression (LUR*) models or acquire data from fixed-site monitoring networks or chemical transport models with hourly output, but typically resources are not available to obtain both. Fixed-site networks — even those in North America and Western Europe — still have relatively limited spatial coverage in many areas, particularly in suburban and rural locations, and insufficient density to capture small-scale (within-city) variations of pollution.

In recent years, researchers have increasingly used routine mobile monitoring by affixing monitoring devices to vehicles and making measurements while systematically and repeatedly traveling a road network. Such mobile monitoring can provide a very dense map of street-level exposure estimates across a given urban area (Apte et al. 2017; Klompmaker et al. 2015; Messier et al. 2018; Patton et al. 2015; Weichenthal et al. 2016). Although the use of mobile monitoring for mapping local concentrations of traffic-related air pollution is becoming more common, many questions remain. For example, how do on-road measurements compare to data from fixed sites, can the method be scaled up to larger areas, and in which contexts is the approach appropriate and feasible? Also, how much data need to be collected (in terms of spatial coverage and repeated samples) to develop satisfactory, robust maps of long-term patterns of air pollution concentrations?

To investigate and develop further the utility of mobile monitoring, Dr. Joshua Apte of the University of Texas at Austin, submitted an application to HEI titled "Scalable Multipollutant Exposure Assessment using Routine Mobile Monitoring Platforms" in response to HEI's Request for Applications 16-1: Walter A. Rosenblith New Investigator Award. This award was established to provide support for an outstanding new investigator at the assistant professor level to conduct research in the area of air pollution and health; it is unrestricted with respect to the topic of research. Dr. Apte proposed to assess the utility of mobile monitoring data collected previously by fleet vehicles (i.e., Google Street View cars) equipped with instruments to routinely monitor air pollution. His application focused on the utility of the data and the scalability of approaches, and it proposed several related analyses based in two cities: Oakland, California, USA, and Bangalore, India.

HEI's Research Committee recommended funding Dr. Apte's application because it thought that the work proposed was novel and could affect how air pollution health research is done in the future. They appreciated his proposed use of an existing large-scale mobile monitoring dataset along with new measurements to be collected in India that would allow him to evaluate approaches in two very different settings. They also liked the focus on traffic-related air pollutants, especially ultrafine (<0.1 µm) particles (UFPs) for which fixed-site monitoring data are sparse. Additionally, they thought the large amount of data that he would analyze and collect had the potential to contribute significantly to exposure assessment for future epidemiological studies. The study started in 2018 and continued when Dr. Apte moved to the University of California, Berkeley.

This Commentary provides the HEI Review Committee's independent evaluation of the study. It is intended to aid the sponsors of HEI and the public by highlighting both the strengths and limitations of the study and by placing the results presented in the Investigators' Report into a broader scientific and regulatory context.

SCIENTIFIC AND REGULATORY BACKGROUND

Patterns of air pollution around traffic sources are characterized by high spatial and temporal variability related to meteorological conditions, varying emission rates, and other

Dr. Joshua S. Apte's 3-year study, "Scalable Multipollutant Exposure Assessment Using Routine Mobile Monitoring Platforms," began in January 2018. Total expenditures were $426,752. The draft Investigators' Report from Apte and colleagues was received for review in October 2022. A revised report, received in May 2023, was accepted for publication in June 2023. During the review process, the HEI Review Committee and the investigators had the opportunity to exchange comments and clarify issues in both the Investigators' Report and the Review Committee's Commentary.

* A list of abbreviations and other terms appears at the end of this volume.

factors (HEI 2022; Park and Kwan 2017; Zhou and Levy 2007). UFPs, compared to some other air pollutants, have especially high spatial and temporal variability. UFPs originate from anthropogenic sources — primarily industrial emissions and combustion of fossil fuels for transportation, energy production, and heating — and from such natural sources as forest fires and marine aerosols, such as sea salt (Moreno-Ríos et al. 2022). They can also form in the atmosphere when combustion processes emit hot, supersaturated vapors that undergo nucleation and condensation while being cooled to ambient temperatures and through chemical reactions in the atmosphere (Sioutas et al. 2005). Their dispersion, transport, and duration of suspension in the atmosphere are affected by environmental and meteorological conditions, including topography, local wind direction and speed, temperature variations, and precipitation, among other factors.

Some of the major challenges in conducting epidemiological studies of air pollution exposure and health include the difficulty of assigning exposures to study participants accurately and quantifying the influence of exposure measurement error on estimated health risks. Those issues are especially challenging for some components of particulate matter (e.g., UFPs) and gaseous outdoor air pollutants, such as nitrogen dioxide (NO_2) and ozone that vary highly in space and time (HEI Review Panel on Ultrafine Particles 2013).

In the past, many studies relied on data from a few fixed-site monitors to assign exposure to study participants, partly because those were the only data available. To improve exposure assessment resources, researchers have deployed additional fixed-site monitors in specific areas (e.g., busy streets). That approach is particularly needed for measuring UFPs for which fixed-site monitoring networks are lacking. Moreover, in many locations in low- and middle-income countries (LMICs), there are few to no permanent fixed-site regulatory air pollution monitors; thus, creative approaches are needed. More recently, researchers have started to use satellite data to cover regions where no monitors exist and mobile monitoring platforms with real-time instrumentation to measure highly resolved spatial trends in air pollution concentrations (e.g., Apte et al. 2017; Minet et al. 2018; Patton et al. 2014; Riley et al. 2014).

Mobile monitoring strategies can involve on-road mobile measurements made while driving predefined strategic routes, or repeated short-term measurements made while in a parked vehicle at many locations. Data collected through mobile monitoring have been used to develop LUR models and other air pollution maps (Klompmaker et al. 2015; Messier et al. 2018; Patton et al. 2015; Weichenthal et al. 2016). Air pollution maps estimated from such monitoring are being increasingly applied in epidemiological studies (e.g., Alexeeff et al. 2018; Corlin et al. 2018). As noted above, however, questions remain about the scalability of mobile monitoring approaches and their applications in different contexts. The current study was designed to improve on these approaches and to test their applicability in a high-income country and an LMIC.

SUMMARY OF APPROACH AND METHODS

STUDY OBJECTIVES

Dr. Apte and colleagues sought to evaluate and assess the utility of mobile monitoring for a range of air pollution exposure assessment applications. The study builds on previous research by the investigators during which they collected a large amount of mobile monitoring data using Google Street View cars equipped with tools to measure nitric oxide (NO), NO_2, black carbon (BC), UFPs, and fine particulate matter <2.5 $\mu g/m^3$ in diameter ($PM_{2.5}$) in Oakland, California.

For this study, they specified the following overarching questions: Does large-scale mobile monitoring produce useful results? In what ways and for what exposure assessment applications is mobile monitoring effective? What complementary or additional insights can be revealed by mobile monitoring? What are the potential limitations of mobile monitoring? To address these overarching questions, the investigators proposed the following aims:

1. Validate intensive mobile monitoring as an exposure assessment technique via comparison with observations from a network of fixed-site monitors.
2. Compare insights from mobile air pollution measurement campaigns with those derived from other approaches and data sources, including observations from regulatory networks, dense low-cost sensor networks, and statistical exposure models.
3. Investigate the potential for scaling of mobile monitoring techniques through both direct observation and modeling, to better understand how mobile monitoring could be applied to larger study domains while minimizing the amount of monitoring effort required.
4. Investigate whether mobile monitoring might be a viable option for collecting air pollution data in a low-resource setting that currently lacks robust air pollution monitoring infrastructure.
5. Probe the rich multipollutant dataset with data mining techniques to understand how sources influence population exposures.

Aims 1 through 3 were addressed by working with data collected previously from fixed-site stations and mobile monitoring campaigns for BC, NO, nitrogen oxides (NO_x), NO_2, and UFPs in Oakland, California. Aim 4 was addressed by conducting a new mobile monitoring campaign for BC, UFPs, $PM_{2.5}$, and carbon dioxide (CO_2) in Bangalore, India. Aim 5 was eventually dropped due to time constraints. The investigators organized their study into five interrelated analysis modules (M1–M5) that each contributed to multiple study aims. They are described below and summarized in the **Commentary Table** with key features and findings.

METHODS

Analysis M1: Intensive comparison of mobile and fixed-site monitoring of black carbon in Oakland, California

The purpose of analysis module M1 was to evaluate the capabilities of mobile monitoring for representing long-term spatial patterns of black carbon by comparing repeated mobile air pollution measurements with data from a large set of continuous fixed-site monitors. For this analysis, the investigators leveraged mobile-monitoring data that they had collected and described previously (see Apte et al. 2017 and sidebar) along with data from a dense network of low-cost, fixed-site BC monitors custom-built and deployed by colleagues at the University of California, Berkeley (Caubel et al. 2019).

The BC monitors deployed by Caubel and colleagues were installed at 100 sites in residential, industrial, and high-traffic microenvironments at an average density of 6.7 sites per km^2 in West Oakland. The instruments were mounted at a height of 1.5 m on fences, porches, or street poles at a median distance of 15 m from the nearest road. Of these 100 sites, 97 were located within 30 m of the road network covered by the mobile monitoring described in the sidebar, and three were located at upwind background sites along the San Francisco Bay. This network was in operation during a 100-day period between May and August 2017. Apte and colleagues computed the median daytime concentration at each site. They then calculated the ordinary Pearson R^2 coefficient of determination between the median concentration of BC of all drive pass means within 95 meters of the 97 custom-built BC detectors with valid data. They chose a distance of 95 meters because the precision of the fixed-site detectors to estimate on-road concentrations decreased notably at distances greater than 95 meters. In total, the mobile monitoring vehicles sampled roads within 95 meters of these fixed-site detectors for nearly 56 hours, with a median of 73 drives past each site. Each visit of a mobile monitoring vehicle to a fixed site lasted about 17 seconds for a median total time of 29.3 minutes at each site.

Analysis M2: Spatiotemporal analysis of traffic-related air pollution dynamics using mobile and fixed sensors in the San Francisco Bay Area

The purpose of this analysis module was to evaluate how the spatiotemporal patterns of UFPs compared with other traffic-related air pollutants that are monitored routinely. For this module, the investigators made use of the mobile monitoring data collected in 10 neighborhoods across the San Francisco Bay area, as described in the sidebar. For this analysis, the investigators compared particle number concentrations (as their proxy for UFPs) obtained through the mobile monitoring with concentrations of NO_x obtained at four regulatory fixed-site monitoring stations operated by the Bay Area Air Quality Management District. Specifically, they used hourly data from 2011 to 2018 from regulatory sites representative of a gradient in traffic influence, namely, near-highway, urban, suburban, and rural.

Analysis M3: Assessment of local- and regional-scale air pollution disparities in the San Francisco Bay Area using mobile monitoring

This analysis was not part of the original application and study plan but was included in the investigators' final report to present the totality of analyses that the investigators conducted with mobile monitoring datasets. The purpose of this analysis was to describe how variability in concentrations of air pollution affected estimates of population exposure and environmental disparities in the San Francisco Bay Area. This analysis module also made use of the mobile monitoring datasets described earlier. Here, the investigators estimated long-term pollution concentrations of BC, NO, NO_2, and UFPs for 6,362 census blocks in 13 communities around the San Francisco Bay Area. The communities ranged in size from 95 to 930 census blocks (median: 447 blocks). The mean census block had an area of about 14,000 m^2 (equivalent to 120 meters × 120 meters) with a mean population of 70 people. The investigators estimated pollution concentrations for each block as the median of observations from roads within about 100 meters of the block center point.

They used U.S. Census Bureau block-level population data for the year 2010, the most recent year for which block-level data were available, to describe the populations in the 13 communities. Specifically, they used the racial and ethnic designations provided by the U.S. census to summarize proportions of populations described as Latino or Hispanic in one group ("Hispanic") and then categorized non-Hispanic populations by race: Asian, Black, White, and "Other," including those of Native American, Pacific Islander, multiracial, or other racial identity. In 2010, about 450,000 people lived in these areas.

The investigators used the pollution and population datasets together to describe distributions of the various pollutants within each community and to describe the exposure distributions according to the racial and ethnic compositions of the population.

Analysis M4: Scaling air quality mapping of NO and BC through mobile monitoring and land use regression in Oakland, California

The purpose of analysis module M4 was to evaluate the advantages and trade-offs of coupling mobile monitoring with LUR and Kriging approaches to estimate intraurban variation in air pollution in a data-efficient manner. This analysis module made use of the mobile monitoring datasets described earlier. Here, Apte and colleagues investigated approaches to reduce the intensity of field data collection required for producing high-resolution pollution maps of NO and BC from mobile monitoring data. For this analysis, they focused on West Oakland, Downtown Oakland, and East Oakland. They considered two broad approaches to data reduction for developing reliable estimates of spatial patterns, namely a "data only" approach and a "land use regression-Kriging model (LUR-K)" approach.

Commentary Table. Key Details of the Five Analysis Modules[a]

Analysis Module	Research Aims Addressed	Pollutants Examined	Period of Measurement	Geographic Location	Key Findings
Analyses That Focus on Comparing Mobile Monitoring with Fixed-Site Monitoring Data					
M1: Intensive comparison of mobile and fixed-site monitoring in Oakland	Validate intensive real-time mobile monitoring as an exposure assessment technique via comparison with fixed observation networks. Compare insights from mobile air pollution measurement campaigns with those derived from other approaches and data sources.	BC	May 2017 – August 2017	West Oakland	Repeated mobile monitoring can reproduce time-averaged, fine-scale spatial patterns of BC with good fidelity, precision, and accuracy relative to a fixed-site sensor network.
M2: Spatiotemporal analysis of traffic- related air pollution dynamics using mobile and fixed sensors in the San Francisco Bay Area	Validate intensive real-time mobile monitoring as an exposure assessment technique via comparison with fixed observation networks. Compare insights from mobile air pollution measurement campaigns with those derived from other approaches and data sources.	BC, CO, NO_x, UFPs	*Mobile measurements*: May 2015 – December 2017 *Regulatory measurements*: Full year, 2015	Mobile measurements: West Oakland and Downtown Oakland Fixed sites: Sebastopol, Livermore, Redwood City, and Laney College	Data from mobile monitoring corroborates a surprising insight from regulatory data: patterns of UFPs and NO_x are coupled in the winter months (indicative of a common primary traffic source), but sharply decoupled in the summer. UFPs in the Bay Area appear to be substantially driven by secondary formation during the summer months.
Analyses that focus on uses and applications of mobile monitoring data					
M3[b]: Assessment of local- and regional-scale air pollution disparities in the San Francisco Bay Area using mobile monitoring	Validate intensive real-time mobile monitoring as an exposure assessment technique.	BC, NO, NO_2, UFPs	May 2015 – December 2017	13 communities across the San Francisco Bay Area	Repeated mobile monitoring can capture exposure heterogeneity across a large urban region. Across the entire Bay Area region, within-neighborhood gradients account for a large to dominant fraction of the overall heterogeneity in the population-concentration distribution. Substantial racial/ethnic disparities are driven mostly by intra-neighborhood segregation.
M4: Scaling air quality mapping of NO and BC through mobile monitoring and spatial modeling in Oakland	Investigate the potential for scaling of mobile monitoring techniques through both direct observation and modeling.	BC, NO	May 2015 – May 2017	West Oakland, Downtown Oakland, East Oakland	With LUR-K modeling, it is possible to drive only a fraction of roads a few times and develop models that are nearly as good as the best models they developed. Data-only maps from repeated driving are superior to LUR-K models in terms of detecting idiosyncratic or unexpected spatial features and hotspots.
M5: Mobile monitoring in Bangalore, India	Investigate whether mobile monitoring might be a viable option for collecting air pollution data in a low-resource setting.	BC, CO_2, UFPs	July 2019 – March 2020	Residential neighborhood in Bangalore (Malleshwaram) and supplemental transects in surrounding areas	Mobile monitoring produced time-stable spatial patterns in Malleshwaram and elsewhere in the study domain. Observed a convergence to time-stable spatial patterns with fewer than 20 repeated mobile monitoring runs over 1 year. Slow traffic speeds in Bangalore present logistical challenges for mobile monitoring.

[a] Source: Investigators' Report Table 2

[b] As described below, this analysis was not part of the original study plan.

SIDEBAR

Prior to applying to this RFA, Apte and colleagues had already collected a large amount of mobile-monitoring data in the San Francisco Bay Area. Briefly, the investigators had equipped two Google Street View cars with instruments for measuring BC, NO_x, and particle number concentrations (a strong proxy for UFPs). Drivers of the vehicles conducted 6–8-hour long shifts between 8 a.m. and 6 p.m. between May 2015 and December 2017. They were assigned 1–5-km^2 areas to cover each day within which they were asked to drive each road in that area at least once, in any order. They conducted intensive monitoring in West Oakland, Downtown Oakland, and East Oakland (totaling over 1,300 hours of monitoring) and added an additional 300 hours in West Oakland alone. They also sampled 1,000 hours in 10 other neighborhoods in the greater San Francisco Bay Area to cover locations with various land uses (e.g., industrial, commercial, dense residential, and light residential), atmospheric and climate conditions, share of open or green space, traffic density, and demographic composition.

The investigators used the air pollution measurements to estimate long-term, average concentrations of the pollutants along roadways that represented the weekday, daytime conditions of the period sampled in these locations. For this task, they divided the measurement domains into 30-meter road segments (equivalent to about 3–10 seconds of observation). For the core Oakland domain, this network included about 20,000 such segments. First, they calculated the mean of all measurements in each 30-meter road segment for each individual drive pass (i.e., the mean of all observations taken during that single 3–10-second period of a drive pass). Then, they computed the median of all repeated drive pass mean concentrations to use as their core metric for analysis. These datasets were used in the various analysis modules described in the investigators' report.

For the data-only approach, they mapped concentrations of pollutants based exclusively on data from the mobile observations, with no support from spatial modeling techniques. Here, they attempted to minimize the number of repeated visits to each road at the cost of reducing the precision and accuracy of the resulting estimated concentrations.

For the LUR-K approach, they applied their mobile-measured observations in a statistical model that combined LUR and Kriging. Briefly, LUR is a spatial modeling technique that uses observations of pollutant concentrations at given locations as the dependent variable and data describing such characteristics as road density and land use as the independent variables, in a multivariate regression model to estimate pollutant concentrations at unsampled locations. Kriging, on the other hand, is a method of spatial interpolation whereby values are predicted at unsampled locations based on measurements taken at nearby locations. As such, for the LUR-K approach, pollution concentrations can be estimated at unsampled locations and mobile observations are not needed from every road in the study domain.

The investigators simulated several variations of approaches to reducing data requirements for mobile sampling:

- Data-only mapping based on mobile monitoring data from a reduced subset of drive days (i.e., sampling on all highway and nonhighway roads, but only 4 days of sampling on each segment).
- Data-only mapping based on mobile monitoring data from a reduced subset of roads sampled (i.e., sampling on all highways and on a random selection of 30% of the nonhighway roads, including all days of sampling).
- LUR-K modeling based on mobile monitoring data from the reduced subset of drive days.
- LUR-K modeling based on mobile monitoring data from the reduced subset of roads sampled.
- Joint scenario with LUR-K modeling where drive days and roads sampled were reduced simultaneously.

Ultimately, they used visual inspection and analyzed model residuals, coefficients of determination (R^2), and normalized root mean square errors (NRMSEs) to compare and evaluate the various approaches.

Analysis M5: Mobile monitoring in Bangalore, India

The purpose of analysis M5 was to investigate their mobile monitoring approach in a low-resource setting. This analysis was set in Bangalore, India, which is located in the southern state of Karnataka, and has a population greater than 12 million people. For this analysis module, Apte and colleagues combined instruments for measuring BC, UFPs, $PM_{2.5}$, CO_2, meteorological parameters, and GPS into a mobile monitoring platform mounted in a compressed-natural gas-powered hatchback car. They used CO_2 concentrations as an indicator of the degree to which their measurements were influenced by the fresh exhaust of traffic emissions.

The investigators conducted mobile monitoring in four regions, including streets in urban residential areas (Malleshwaram), the central business district, and in peri-urban areas. Drivers conducted shifts of about 4 hours long between 9 a.m. and 1 p.m. between July 2019 and March 2020, which covered all seasons except the hottest summer months. As such, results generally represent late morning conditions on weekdays.

Similar to the analysis process described in analysis M1, the investigators used the mobile air measurements to estimate long-term, average pollutant concentrations representative of the period sampled. As was done in Oakland,

they divided the measurement domains into 30-meter-long road segments, which in this case was about 5,000 segments. Again, they computed the median of the repeated drive pass mean concentrations to use as their core metric for analysis.

All modules described above were conducted at various times between May 2015 and March 2020. The key findings from the analyses are presented below.

SUMMARY OF KEY RESULTS

ANALYSIS M1: INTENSIVE COMPARISON OF MOBILE AND FIXED-SITE MONITORING OF BLACK CARBON IN OAKLAND, CALIFORNIA

The investigators found that the spatial patterns of BC produced with their mobile monitoring data were similar to the daytime medians calculated with observations from the 97 fixed-site detectors. The correlation (R^2) between the measurements at the fixed sites and the mobile measurements sampled within 95 meters was 0.51. The correlations varied but were approximately 0.5 for measurements within distances of 50–90 meters and were in the range of 0.4 to 0.3 for measurements within distances of 100 to 150 meters (Investigators' Report [IR], Figure 5d). Although their results were influenced somewhat by the choice of days and seasons in which they sampled pollutants, they ultimately concluded that their mobile monitoring design was sufficiently robust for the purpose of characterizing spatial patterns of air pollution.

Overall, the median concentration of BC measured along all nonhighway road segments within 95 meters (i.e., 0.44 μg/m^3) matched closely the median concentration among the fixed sites of 0.48 μg/m^3, suggesting that the data collected on-road were broadly representative of the near-road concentrations based on data from fixed sites. The mobile measurements had the advantage of detecting road-level variability not available from the fixed-site monitors, as well as estimates on highways, where placement of fixed-site monitors would likely be infeasible.

ANALYSIS M2: SPATIOTEMPORAL ANALYSIS OF TRAFFIC-RELATED AIR POLLUTION DYNAMICS USING MOBILE AND FIXED-SITE MONITORS IN THE SAN FRANCISCO BAY AREA

The investigators compared diurnal profiles of UFPs and NO_x stratified by season and weekday or weekend at the four regulatory fixed-site locations. During winter conditions, they observed generally consistent diurnal (hour-of-day) patterns for UFPs and NO_x (IR, Figure 6a). The summertime diurnal patterns for each pollutant, however, notably differed; observations for NO_x were generally lower than those for UFPs. For example, there were daytime peaks in UFP concentrations at multiple sites during the warmer months that were not observed with NO_x. Observations of NO_x were also notably lower in summer than in winter and lowest on weekend days.

Overall, the investigators concluded that daytime UFP concentrations in this area, especially during summer, appeared to be influenced strongly by nontraffic sources of UFPs, including secondary new particle formation events. Given the differences in spatiotemporal patterns of NO_x concentrations compared to those of UFPs, they suggested that using NO_x (or other traffic-related air pollutants) as a proxy for UFPs could result in inaccuracies in estimating UFP exposure.

ANALYSIS M3: ASSESSMENT OF LOCAL- AND REGIONAL-SCALE AIR POLLUTION DISPARITIES IN THE SAN FRANCISCO BAY AREA USING MOBILE MONITORING

The population-weighted means of the measured pollutants among the 13 communities were: 0.31 μg/m^3 for BC, 4.6 ppb for NO, 8.2 ppb for NO_2, and 19,100 cm^3 for UFPs. Generally, correlations between block-level concentrations of the individual pollutants were variable, and they observed the lowest correlations between UFPs and the other pollutants (interquartile range Pearson's r ranged from 0.4 to 0.7).

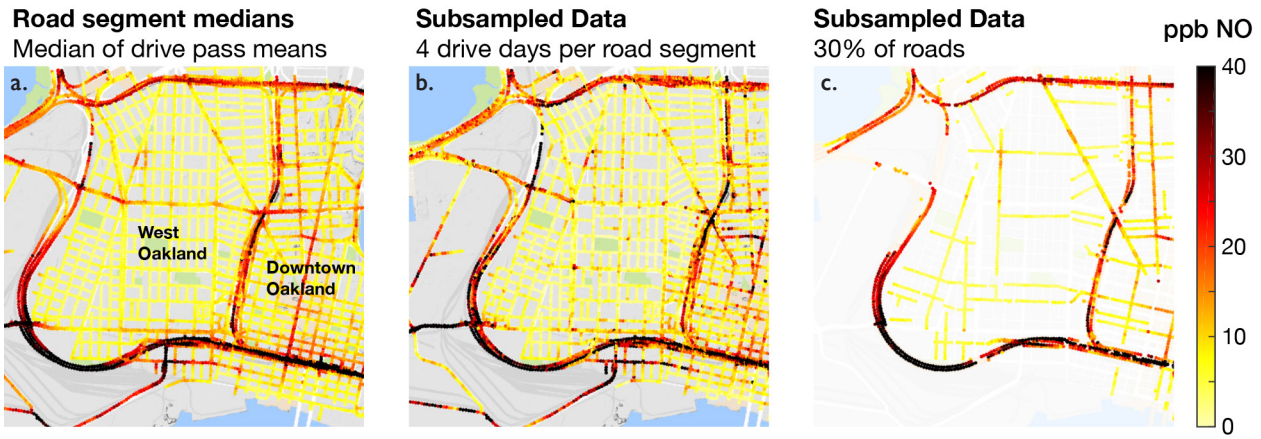
In this study area, based on data from the 2010 U.S. Census, 33% of the population was Non-Hispanic White, 31% was Asian, 21% was Hispanic, and 14% was Black. The investigators found that Non-Hispanic White populations were exposed to lower concentrations of NO, NO_2, and UFPs than other groups, with median exposures 16% to 27% below the total population median, while Black and Hispanic populations were exposed to concentrations 8% to 30% higher than the total population medians (IR, Figure 8a).

This analysis found that differences in population exposures to NO and BC were driven mostly by variability in concentrations within individual neighborhoods (i.e., very local-scale variability; within 1 km), whereas differences in exposures to NO_2 and UFPs across the area were driven principally by differences in larger-scale, neighborhood-level mean concentrations.

ANALYSIS M4: SCALING AIR QUALITY MAPPING OF NO AND BC THROUGH MOBILE MONITORING AND LAND USE REGRESSION IN OAKLAND, CALIFORNIA

Apte and colleagues produced maps of pollutant concentrations on sampled road segments using the various approaches described earlier. Visual inspection suggested that each approach had generally good face validity and captured key features of the long-term concentrations of NO and BC. For example, in all cases, concentrations appeared lowest on residential streets, and highest on highways and in the downtown area of Oakland. **Commentary Figure 1** presents maps of NO patterns created with the data-only approach using all available data (left panel) and reduced datasets (middle and right panel).

The map produced using the data-only approach with the full dataset (i.e., many dozen drive passes on all roads, with a total drive time of about 1,300 hours) contained many



Commentary Figure 1. Maps showing results of data reduction schemes for estimating daytime median concentrations of NO in Oakland, California, during 2015–2017 using a data-only approach. (**a**) Median of drive-pass mean concentrations using all available data (all roads, all drive passes). (**b**) Four randomly selected drive days per road segment (all roads, fewer drive passes). (**c**) All drive days but subsampled to represent 30% of the arterial and residential roads (fewer roads, all drive passes). Source: IR Figure 3.

localized pollution hotspots at intersections and locations with industries or other emissions sources that were not apparent in the maps created with the reduced datasets. The data-only map produced with a dataset restricted to only four drive days of observation, but coverage of all streets (i.e., 6% of the full dataset; about 80 hours in total; middle panel of Commentary Figure 1), resulted in only a slight decrease in performance, but with a substantial drop in mobile-monitoring data requirements. The panel on the right of Commentary Figure 1 shows the estimated NO concentrations based on all available days of observation but limited to only 30% of the arterial and residential roads.

The LUR-K approaches developed using either a sampling of all roads, but from a reduced subset of drive days, or a subset of roads, but sampled many times, both resulted in only negligible decreases in model predictions and performance.

Finally, the LUR-K model based on a highly restricted dataset (i.e., 30% road coverage and only four days of observation) also reflected only a moderate reduction in model performance despite the substantial reduction in data requirements. More details can be found in IR Figures 3 and 10 for maps created using the various alternative approaches.

Ultimately, the overarching conclusion from this analysis was that viable LUR-K models could be developed even with little mobile monitoring data. Although the data-only approach outperformed the LUR-K in precision with only a modest number of repeated samples (i.e., <10 repeated days), this was at the cost of having to sample every road for which exposure measurements are desired.

ANALYSIS M5: MOBILE MONITORING IN BANGALORE, INDIA

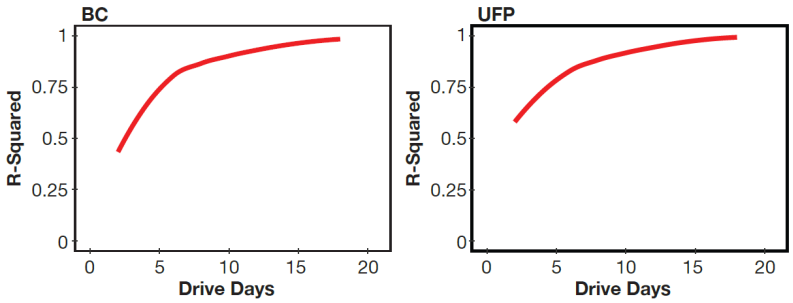
Due to various logistical issues, the work in India was not as extensive as originally planned, and so the investigators focused on the results from Malleshwaram, a large, urban neighborhood of Bangalore. This area was the only one for which they were able to conduct complete block-by-block repeated monitoring comparable to that of their San Francisco Bay Area campaign. Their study design involved collecting one weekly sample of the entire Malleshwaram area over two consecutive days, resulting in 44 days of data collection and 22 repeated drive days for each road segment. In total, they sampled about 150 km of roads across Bangalore, about 62 km of which were in Malleshwaram.

The spatial means (and medians) representing morning-time concentrations on the nonsummer weekdays for the road segments in the Malleshwaram study domain were about 26 $\mu g/m^3$ (15 $\mu g/m^3$) for BC and about 81,000 cm^3 (62,000 cm^3) for UFPs. Similar to the maps for the San Francisco Bay Area, the maps produced here again had strong face validity with the highest observations observed along highways, lower observations on major arterial roads, and the lowest observations on smaller, residential streets (with similar patterns for all three pollutants). The observed concentrations of BC and UFPs in Malleshwaram were both much higher than those observed in the San Francisco Bay Area, with the observations for UFPs about four times higher in Malleshwaram than in the Bay Area and those for BC about 100 times higher. The investigators suggested that this finding is likely due to the high proportion of older diesel engines operating in India.

As was done in analysis M4, the investigators examined how many repeated samples would be needed to capture the spatial patterns observed with the full dataset of 22 repeated drives. Here, they observed that including information from each additional drive pass increased rapidly until about 7 sampling days, and then only minimally thereafter (**Commentary Figure 2**).

As such, despite the differences in terms of fleet composition, population density, and mean pollutant concentrations

between the two settings of Malleshwaram and the San Francisco Bay Area, the reduced sampling results in both locations suggested that mobile monitoring produced relatively stable maps after about 10 drive days, with diminishing returns to precision with additional sampling beyond that. The finding from this analysis module, therefore, is consistent with those presented earlier for analysis M4 and from previous work by these investigators (Apte et al. 2017).



Commentary Figure 2. Subsampling analysis for the Malleshwaram neighborhood in Bangalore. Source: IR Figure 13.

HEI REVIEW COMMITTEE'S EVALUATION

STUDY DESIGN, DATASETS, AND ANALYTICAL APPROACHES

In its independent evaluation of the study, the Review Committee noted that at the time of funding, in June 2017, this study proposed the largest, most extensive campaign to examine the potential applications, strengths, and limitations of mobile monitoring. Overall, the HEI Review Committee was impressed with the extent to which the investigators described, compared, and analyzed the data.

The Committee noted several strengths of the study design. For example, Apte and colleagues compiled a large amount of data from several sources, including mobile monitoring data in several locations (in two countries) and data from several fixed-site networks. In addition, the richness of the data allowed the investigators to explore many issues, including the comparability of long-term observations from fixed-site monitors with observations collected through mobile monitoring and the utility of mobile monitoring data for describing spatial gradients in pollution. Their application of these measurements to estimate potential population-level exposures was also appreciated as an enhancement to our understanding of environmental inequities within the population. The datasets also allowed the investigators to evaluate the feasibility of applying these approaches in different settings. The wide spatial and temporal extent of data used here also allowed the investigators to conduct simulation studies to evaluate various logistical and study design considerations that can affect the potential benefits and costs associated with mobile monitoring. Another strength of the study is the examination of the performance of air quality models that integrate mobile monitoring data into LUR-K modeling.

The rich datasets used by the investigators also allowed them to explore and identify the relative trade-offs between intensive repeated sampling and several alternative approaches to data reduction, including reducing the number of roads sampled and the number of repeated passes on given roads. The Committee agreed with the investigators that in both the San Francisco Bay Area and in Malleshwaram, mobile monitoring produced relatively reproducible maps for several traffic-related pollutants with data from relatively few repeated drive passes.

The Committee also noted some limitations to the approach. For example, one issue with the mobile monitoring is that the drivers drove some routes and areas always in the same order and at the same time of day. This pattern of data collection makes it difficult to disentangle whether the pollution concentrations in a given location are indeed representative of the daytime typical average conditions, or if the concentrations for that location in fact represent temporal trends much higher than average levels occurring during rush hour or lower than average values during a low-traffic time of day.

The Committee also wondered whether the results are generalizable to other pollutants, longer periods, or to other locations (including to wider areas within the San Francisco Bay and Bangalore areas, as well as to other locations in the United States or elsewhere). For example, all the comparisons between mobile and fixed-site measurements from analysis M1 pertain to only one pollutant (BC), one study area (West Oakland), and cover a relatively short period (May-August 2017). Similar analyses of other pollutants would be useful in the future. It is also difficult to know the extent to which the observed correspondences between UFPs and NO_x described in analysis M2 would apply to other locations with different geographies, mixes of vehicles, or kinds of point sources of air pollution. Finally, the monitored area in Bangalore (i.e., Malleshwaram) comprised only a few square kilometers so might not accurately capture variations that might have been observed elsewhere in the very large city or in the surrounding regions.

It is important to acknowledge that the limitations above, along with a few other issues, might affect the suitability of mobile measured air pollution data for use in epidemiological analyses when used as the only data source. For example, we would expect on-road measurements to be different from those observed at fixed-site stations because they are collected in the middle of the road rather than at roadsides or other locations that might be closer to where people live. This is in contrast to measurements from fixed-site monitors, and even satellite-based measurements, that can be collected in a vari-

ety of locations, including away from busy roads. The mobile monitoring was also performed only during daytime hours on weekdays and does not reflect concentrations during the times of day when people might be more likely to be at home (i.e., in the evenings, at night, and on weekends). Moreover, most cohort studies have information on the residential addresses of individuals for the purpose of estimating air pollution exposures. Given the intensiveness of mobile monitoring, there will often be a mismatch between the period captured by the mobile measurements and the window of exposure most relevant for epidemiological purposes, especially if the focus is on the health impacts of long-term exposures.

Nonetheless, these measurements did provide additional spatial resolution that might not be captured by the limited fixed-site monitoring network or area-based satellite measurements. Additionally, mobile measurements might be especially useful for estimating exposures for commuters, especially cyclists and pedestrians. Overall, the Committee agrees with the investigators that there are further opportunities to explore these kinds of rich datasets, especially for combining the mobile measured data with fixed-site data to develop exposure models for use in epidemiological analyses and for identifying disparities in population exposures.

FINDINGS AND INTERPRETATION

Generally, the Committee found that the report presented a comprehensive and thoughtful discussion of the findings from the numerous research modules and analyses. Results from this study answered important questions and contributed interesting insights about collecting and working with mobile-measured air pollution data.

The descriptive analyses of BC, NO_x, and UFPs provide valuable new insights about their spatial and temporal patterns, and particularly, how they compare with those of other traffic-related pollutants in different contexts. For example, the investigators were able to identify that patterns of UFPs and NO_x shared similar spatial and hourly patterns during winter months in the San Francisco Bay Area, a result indicating a common primary traffic source during this season. However, during summer months the patterns were dissimilar, with the suggestion that summer concentrations of UFPs in this area were more strongly influenced by new secondary particle formation rather than primary emissions. The Committee agreed with this conclusion and felt that these data highlight nicely the value of combining detailed mobile mapping with at least a few fixed-site monitors that can provide long-term data.

The Committee also agreed with the investigators that some pollutants appear to be better suited for mobile monitoring than others. Generally, pollutants with a high degree of spatial variation and a low degree of temporal variation, such as NO_x, should be among the best suited to this kind of approach. In contrast, $PM_{2.5}$ is likely less suited for this approach because it tends to have relatively low spatial variability within an urban area. Similar conclusions could be made about the kinds of locations that would benefit most from mobile monitoring. Specifically, locations with greater heterogeneity in local sources will benefit from the richer spatial information of a mobile monitoring campaign.

The Committee thought this report highlighted what we can learn about spatial patterns of traffic-related air pollution and population exposures when mobile monitoring data are leveraged. Importantly, mobile monitoring can provide measurements directly on highways where fixed-site monitoring is infeasible. This has value for better capture of emissions from the vehicle fleet and for reflecting exposures to drivers on the road. Apte and colleagues also demonstrated clearly that mobile monitoring is able to detect localized pollution hot spots, such as at specific intersections and along designated truck routes, which would not be captured by measurements from fixed-site stations alone.

The investigators estimated potential population exposures by averaging observations collected on surrounding streets to the centers of city blocks. Here, they found that across the San Francisco Bay Area, Non-Hispanic White populations were exposed to lower concentrations of NO, NO_2, and UFPs than other groups, and Black and Hispanic populations were exposed to higher-than-average concentrations of those pollutants. The Committee saw the value in using these data for characterizing environmental disparities and was generally satisfied with this approach though they acknowledge the potential challenge of disentangling differences in concentration due to time and space as discussed above.

An especially important aspect of this study was a detailed analysis to determine how much mobile monitoring data are needed to get relatively accurate maps of long-term patterns of traffic-related air pollution along roadways. The Committee noted that the investigators demonstrated that adequate pollution maps were produced by models supported by LUR-K approaches that used relatively limited data from mobile monitoring. Importantly, this study showed that sampling on every road is not needed for the model output to be effective. The investigators also showed that maps produced with only mobile monitoring data (i.e., without support from the spatial modeling approaches) outperformed the LUR-K in precision with only a modest number of repeated samples (i.e., fewer than 10 repeated days), but at the cost of having to sample every road. Researchers using these methods for epidemiological studies will need to evaluate the extent to which the added cost of mobile monitoring yields sufficient improvements to exposure modeling and prediction.

A novel aim of this study was the investigators' efforts to implement mobile monitoring in a low-resource setting, namely Bangalore, India, with traffic patterns and pollution concentrations that are very different from those in the San Francisco Bay Area. The investigators demonstrated that with sufficient funding and expertise, mobile monitoring was a viable technique for estimating fine-scale concentrations of

traffic-related air pollution in that area. They noted that key challenges of conducting mobile monitoring in this setting included the low traffic speeds (typically 10-15 km/h), which limited the area that could be covered in a given sampling session, and that the instruments used required study personnel to accompany the drivers at all times to ensure the instruments were operating properly. Both of those issues limited the efficiency of the process, as compared to that undertaken in Oakland. The Review Committee perceived the work in India as a feasibility study given the small sampling area that was ultimately sampled. Therefore, although the Review Committee commends the investigators for undertaking this analysis, they note that more work is needed to know if this is a feasible approach in other LMICs, and perhaps in India more broadly.

Another aim of this study was to investigate the potential for scaling mobile monitoring techniques to larger study domains (i.e., not just neighborhoods, but across entire cities and regions). The analyses with LUR-K modeling demonstrated how the mobile monitoring data could be leveraged for creating spatial models to cover areas where not all roads are sampled. The leveraging of measurements collected previously using Google Street View cars was a unique opportunity that the investigators benefited from in their study. However, a potential limiting factor for scaling or replicating these analyses is that Google Street View cars are not available on demand to other researchers or in other locations. Other fleet vehicles that regularly drive around cities, such as taxis or delivery trucks, are an alternate possibility but might be less suitable options for this purpose because they are driven less systematically through communities and researchers would have no control over the routes covered.

It is worth noting that mobile monitoring, in addition to being time-consuming and laborious, can be costly, especially in areas that do not have sufficient resources dedicated to air quality monitoring and research. The investigators estimated a cost of about $1 million per year (which would include vehicles, equipment, and salaries for drivers and analysts) to conduct mobile monitoring equivalent to what was done in Oakland in a large urban area in the United States. The investigators noted that costs to do this might be lower in LMICs settings where labor costs are generally lower, but personnel with the required training and expertise might not be readily available. Ultimately, these estimated costs are much higher than what might be expected for establishing or expanding a network of low-cost, fixed-site monitors to capture more detailed data on pollutant concentrations for epidemiological or regulatory purposes. A related question, therefore, is whether mobile monitoring is really needed in some locations, such as in LMICs, or would time and resources be better spent in building the basic air quality monitoring infrastructure first? Certainly, the answer will depend on the pollutant, location, and question of interest.

CONCLUSIONS

In this pioneering study, Apte and colleagues conducted very thorough analyses of the various strengths, limitations, and potential uses of mobile monitored air pollution data. They showed that mobile monitoring data (which provide dense spatial coverage) coupled with observations from fixed-site stations (which provide long-term temporal coverage) and spatial modeling approaches can produce robust maps of spatiotemporal patterns of traffic-related pollution that can capture highly localized hotspots of pollution. On their own, however, data from mobile monitoring can have important limitations and therefore careful consideration is needed before using them in exposure assessment or epidemiological analyses.

ACKNOWLEDGMENTS

The HEI Review Committee thanks the ad hoc reviewers for their help in evaluating the scientific merit of the Investigators' Report. The Committee is also grateful to Allison Patton for her oversight of the study, to Dan Crouse for assistance with review of the report and in preparing its Commentary, to Mary Brennan for editing of this Report and its Commentary, and to Kristin Eckles and Hope Green for their roles in preparing this Research Report for publication.

REFERENCES

Alexeeff SE, Roy A, Shan J, Liu X, Messier K, Apte JS, et al. 2018. High-resolution mapping of traffic related air pollution with Google Street View cars and incidence of cardiovascular events within neighborhoods in Oakland, CA. Environ Health 17:38; doi:10.1186/s12940-018-0382-1.

Apte JS, Messier KP, Gani S, Brauer M, Kirchstetter TW, Lunden MM, et al. 2017. High-resolution air pollution mapping with Google Street View cars: Exploiting big data. Environ Sci Technol 51:6999–7008; doi:10.1021/acs.est.7b00891.

Caubel JJ, Cados TE, Preble CV, Kirchstetter TW. 2019. A distributed network of 100 black carbon sensors for 100 days of air quality monitoring in West Oakland, California. Environ Sci Technol 53:7564–7573; doi:10.1021/acs.est.9b00282.

Corlin L, Woodin M, Hart JE, Simon MC, Gute DM, Stowell J, et al. 2018. Longitudinal associations of long-term exposure to ultrafine particles with blood pressure and systemic inflammation in Puerto Rican adults. Environ Health 17:33; doi:10.1186/s12940-018-0379-9.

HEI. 2022. Special Report 23. Systematic Review and Meta-analysis of Selected Health Effects of Long-Term Exposure to Traffic-Related Air Pollution. Boston, MA: Health Effects Institute.

HEI Review Panel on Ultrafine Particles. 2013. HEI Perspectives 3. Understanding the Health Effects of Ambient Ultrafine Particles. Boston, MA:Health Effects Institute.

Klompmaker JO, Montagne DR, Meliefste K, Hoek G, Brunekreef B. 2015. Spatial variation of ultrafine particles and black carbon in two cities: Results from a short-term measurement campaign. Sci Total Environ 508:266–275; doi:10.1016/j.scitotenv.2014.11.088.

Messier KP, Chambliss SE, Gani S, Alvarez R, Brauer M, Choi JJ, et al. 2018. Mapping air pollution with Google Street View cars: Efficient approaches with mobile monitoring and land use regression. Environ Sci Technol; doi:10.1021/acs.est.8b03395.

Minet L, Liu R, Valois MF, Xu J, Weichenthal S, Hatzopoulou M. 2018. Development and comparison of air pollution exposure surfaces derived from on-road mobile monitoring and short-term stationary sidewalk measurements. Environ Sci Technol 52:3512–3519; doi:10.1021/acs.est.7b05059.

Moreno-Ríos AL, Tejeda-Benítez LP, Bustillo-Lecompte CF. 2022. Sources, characteristics, toxicity, and control of ultrafine particles: An overview. Geosci Front 13:101147; doi:10.1016/j.gsf.2021.101147.

Park YM, Kwan M-P. 2017. Individual exposure estimates may be erroneous when spatiotemporal variability of air pollution and human mobility are ignored. Health Place 43:85–94; doi:10.1016/j.healthplace.2016.10.002.

Patton AP, Perkins J, Zamore W, Levy JI, Brugge D, Durant JL. 2014. Spatial and temporal differences in traffic-related air pollution in three urban neighborhoods near an interstate highway. Atmos Environ 99:309–321; doi:10.1016/j.atmosenv.2014.09.072.

Patton AP, Zamore W, Naumova EN, Levy JI, Brugge D, Durant JL. 2015. Transferability and generalizability of regression models of ultrafine particles in urban neighborhoods in the Boston area. Environ Sci Technol 49:6051–6060; doi:10.1021/es5061676.

Riley EA, Banks L, Fintzi J, Gould TR, Hartin K, Schaal L, et al. 2014. Multi-pollutant mobile platform measurements of air pollutants adjacent to a major roadway. Atmos Environ 98:492–499; doi:10.1016/j.atmosenv.2014.09.018.

Sioutas C, Delfino RJ, Singh M. 2005. Exposure assessment for atmospheric ultrafine particles (UFPs) and implications in epidemiologic research. Environ Health Perspect 113: 947–955.

Weichenthal S, Ryswyk KV, Goldstein A, Bagg S, Shekkarizfard M, Hatzopoulou M. 2016. A land use regression model for ambient ultrafine particles in Montreal, Canada: A comparison of linear regression and a machine learning approach. Environ Res 146:65–72; doi:10.1016/j.envres.2015.12.016.

Zhou Y, Levy JI. 2007. Factors influencing the spatial extent of mobile source air pollution impacts: A meta-analysis. BMC Pub Health 7:89; doi:1471-2458-7-89 [pii] 10.1186/1471-2458-7-89.

ABBREVIATIONS AND OTHER TERMS

ABCD	aerosol black carbon detector
BC	black carbon
CACES	Center for Air, Climate and Energy Solutions
CO	carbon monoxide
CO_2	carbon dioxide
GPS	global positioning system
ICC	intra-class correlation
IEG	integrated empirical geographic
LMIC	low-middle income country
LOD	limit of detection
LUR	land use regression
LUR-K	land use regression-kriging
MAE	mean average error
NO	nitric oxide
NO_2	nitrogen dioxide
NO_x	nitrogen oxides
NRMSE	normalized root-mean-square error
PAX	photoacoustic extinctiometer
PM	particulate matter
$PM_{2.5}$	fine PM, particulate matter with aerodynamic diameter ≤2.5 μm
ppb	parts per billion
RMSE	root-mean-square error
SSD	sum-of-square deviation
UFPs	ultrafine particles

RELATED HEI PUBLICATIONS

Number	Title	Principal Investigator	Date
Research Reports			
209	Associations of Air Pollution on the Brain in Children: A Brain Imaging Study	M. Guxens	2022
207	Characterizing Determinants of Near-Road Ambient Air Quality for an Urban Intersection and a Freeway Site	H.C. Frey	2022
202	Enhancing Models and Measurements of Traffic-Related Air Pollutants for Health Studies Using Dispersion Modeling and Bayesian Data Fusion	S. Batterman	2020
183	Development of Statistical Methods for Multipollutant Research		
	Part 1. Statistical Learning Methods for the Effects of Multiple Air Pollution Constituents	B.A. Coull	2015
	Part 2. Development of Enhanced Statistical Methods for Assessing Health Effects Associated with an Unknown Number of Major Sources of Multiple Air Pollutants	E.S. Park	2015
	Part 3. Modeling of Multipollutant Profiles and Spatially Varying Health Effects with Applications to Indicators of Adverse Birth Outcomes	J. Molitor	2016
Special Report			
23	Systematic Review and Meta-analysis of Selected Health Effects of Long-Term Exposure to Traffic-Related Air Pollution	HEI	2022

PDFs are available for free downloading at *www.healtheffects.org/publications*.

HEI BOARD, COMMITTEES, and STAFF

Board of Directors

Richard A. Meserve, Chair *Senior of Counsel, Covington & Burling LLP; President Emeritus, Carnegie Institution for Science; former Chair, U.S. Nuclear Regulatory Commission*

Jo Ivey Boufford *Clinical Professor of Global Health and Pediatrics, New York University*

Homer A. Boushey *Emeritus Professor of Medicine, University of California, San Francisco*

Jared L. Cohon *President Emeritus and Professor, Civil and Environmental Engineering and Engineering and Public Policy, Carnegie Mellon University*

Stephen Corman *President, Corman Enterprises*

Martha J. Crawford *Operating Partner, Macquarie Asset Management*

Ana V. Diez Roux *Dana and David Dornsife Dean and Distinguished University Professor of Epidemiology, Dornsife School of Public Health, Drexel University; Director, Drexel Urban Health Collaborative*

Michael J. Klag *Dean Emeritus and Second Century Distinguished Professor, Johns Hopkins Bloomberg School of Public Health*

Alan I. Leshner *CEO Emeritus, American Association for the Advancement of Science*

Catherine L. Ross *Regents' Professor Emerita, City and Regional Planning and Civil and Environmental Engineering, Georgia Institute of Technology; Chairman of the Board of Directors of the Auto Club Group, American Automobile Association*

Martha E. Rudolph *Environmental Attorney, Former Director of Environmental Programs, Colorado Department of Public Health and Environment*

Karen C. Seto *Frederick Hixon Professor of Geography and Urbanization Science, Yale School of the Environment*

Research Committee

David A. Savitz, Chair *Professor of Epidemiology, School of Public Health, and Professor of Obstetrics and Gynecology and Pediatrics, Alpert Medical School, Brown University*

David C. Dorman *Professor, Department of Molecular Biomedical Sciences, College of Veterinary Medicine, North Carolina State University*

Christina H. Fuller *Associate Professor, School of Environmental, Civil, Agricultural and Mechanical Engineering, University of Georgia College of Engineering*

Marianne Hatzopoulou *Professor, Civil and Mineral Engineering, University of Toronto, Canada, Research Chair in Transport Decarbonization and Air Quality*

Amy H. Herring *Sara & Charles Ayres Distinguished Professor of Statistical Science, Global Health, Biostatistics, and Bioinformatics, Duke University*

Heather A. Holmes *Associate Professor, Department of Chemical Engineering, University of Utah*

Neil Pearce *Professor of Epidemiology and Biostatistics, London School of Hygiene and Tropical Medicine, United Kingdom*

Evangelia (Evi) Samoli *Professor of Epidemiology and Medical Statistics, Department of Hygiene, Epidemiology and Medical Statistics, School of Medicine, National and Kapodistrian University of Athens, Greece*

Neeta Thakur *Associate Professor of Medicine, University of California, San Francisco*

Gregory Wellenius *Professor, Department of Environmental Health, Boston University School of Public Health and Director, BUSPH Center for Climate and Health*

Continues next page

HEI BOARD, COMMITTEES, and STAFF

Review Committee

Melissa J. Perry, Chair *Dean, College of Public Health, George Mason University*

Sara D. Adar *Associate Professor and Associate Chair, Department of Epidemiology, University of Michigan School of Public Health*

Kiros T. Berhane *Cynthia and Robert Citrone-Roslyn and Leslie Goldstein Professor and Chair, Department of Biostatistics, Mailman School of Public Health, Columbia University*

Ulrike Gehring *Associate Professor, Institute for Risk Assessment Sciences, Utrecht University*

Michael Jerrett *Professor, Department of Environmental Health Sciences, Fielding School of Public Health, University of California, Los Angeles*

Frank Kelly *Henry Battcock Chair in Community Health and Policy and Director of the Environmental Research Group, Imperial College London School of Public Health, United Kingdom*

Jana B. Milford *Professor Emerita, Department of Mechanical Engineering and Environmental Engineering Program, University of Colorado, Boulder*

Jennifer L. Peel *Professor of Epidemiology, Department of Environmental and Radiological Health Sciences, Colorado State University, and the Colorado School of Public Health*

Eric J. Tchetgen Tchetgen *University Professor and Professor of Biostatistics and Epidemiology, Perelman School of Medicine, and Professor of Statistics and Data Science, The Wharton School, University of Pennsylvania*

John Volckens *Professor, Department of Mechanical Engineering, Walter Scott Jr. College of Engineering, Colorado State University*

Officers and Staff

Elena Craft *President*

Robert M. O'Keefe *Vice President*

Ellen K. Mantus *Director of Science*

Donna J. Vorhees *HEI Energy CEO and Vice President*

Thomas J. Champoux *Director of Science Communications*

Jacqueline C. Rutledge *Director of Finance and Administration*

Jason Desmond *Deputy Director of Finance and Administration*

Emily Alden *Corporate Secretary*

Daniel S. Greenbaum *President Emeritus*

Annemoon M. van Erp *Deputy Director of Science Emeritus*

Amy Andreini *Science Communications Specialist*

Ayusha Ariana *Research Assistant*

Hanna Boogaard *Consulting Principal Scientist*

Aaron J. Cohen *Consulting Principal Scientist*

Dan Crouse *Senior Scientist*

Philip J. DeMarco *Compliance Manager*

Cloelle Danforth *Senior Scientist*

Kristin C. Eckles *Senior Editorial Manager*

Continues next page

HEI BOARD, COMMITTEES, and STAFF

(Staff, continued)

Elise G. Elliott *Staff Scientist*
Hope Green *Editorial Project Manager*
Lissa McBurney *Senior Science Administrator*
Janet McGovern *Executive Assistant*
Samantha Miller *Research Assistant*
Victor Nthusi *Consulting Research Fellow*
Pallavi Pant *Head of Global Health*
Allison P. Patton *Senior Scientist*
Quoc Pham *Science Administrative Assistant*
Yasmin Romitti *Staff Scientist*
Anna Rosofsky *Senior Scientist and Environmental Justice Program Lead*
Abinaya Sekar *Consulting Research Fellow*
Robert Shavers *Operations Manager*
Eva Tanner *Staff Scientist*
Alexis Vaskas *Digital Communications Manager*
Ada Wright *Research Assistant*

RESEARCH REPORT

NUMBER 216
JANUARY 2024

HEI

Health Effects Institute

75 Federal Street
Suite 1400
Boston, Massachusetts 02110, USA
+1-617-488-2300

healtheffects.org