



Advancing Data Science Methods for Environmental Health Policy Design and Evaluation

Citation

Considine, Ellen. 2025. Advancing Data Science Methods for Environmental Health Policy Design and Evaluation. Doctoral Dissertation, Harvard University Graduate School of Arts and Sciences.

Link

<https://dash.harvard.edu/handle/1/42719478>

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

<https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFIOWI2M2EwYTg>

Accessibility

<https://accessibility.huit.harvard.edu/digital-accessibility-policy>

Share Your Story

The Harvard community has made this article openly available.

Please share how this access benefits you. [Submit a story](#)

HARVARD
Kenneth C. Griffin



GRADUATE SCHOOL
OF ARTS AND SCIENCES

THESIS ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Department of Biostatistics

have examined a dissertation entitled

"Advancing Data Science Methods for Environmental Health Policy Design
and Evaluation"

presented by Ellen M. Considine

candidate for the degree of Doctor of Philosophy and hereby
certify that it is worthy of acceptance.

Rachel Nethery
Signature Rachel Nethery (Apr 30, 2025 09:12 EDT)

Typed name: Prof. Rachel Nethery

Francesca dominici
Signature _____

Typed name: Prof. Francesca Dominici

Danielle Braun
Signature Danielle Braun (Apr 30, 2025 20:01 EDT)

Typed name: Dr. Danielle Braun

Rajarshi Mukherjee
Signature Rajarshi Mukherjee (May 2, 2025 14:44 EDT)

Typed name: Prof. Rajarshi Mukherjee

Signature _____

Typed name:

Date: April 29, 2025

Advancing Data Science Methods for Environmental Health Policy Design and Evaluation

A DISSERTATION PRESENTED
BY
ELLEN M. CONSIDINE
TO
THE DEPARTMENT OF BIostatISTICS

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN THE SUBJECT OF
BIostatISTICS

HARVARD UNIVERSITY
CAMBRIDGE, MASSACHUSETTS
APRIL 2025

©2025 – ELLEN M. CONSIDINE
ALL RIGHTS RESERVED.

Advancing Data Science Methods for Environmental Health Policy Design and Evaluation

ABSTRACT

To date, the field of environmental health has been primarily focused on characterizing the public health impacts of environmental exposures such as air pollution and temperature. However, beyond studying impacts, there is an increasingly recognized need for designing data-driven policies or strategies to (a) reduce both the overall health burden and health disparities associated with current environmental factors and (b) adapt to emergent threats such as climate change. Meanwhile, diverse and growing data sources, paired with modern data science methods, hold the potential for expanding the types of environmental health science and policy questions that can be answered.

This dissertation employs a wide range of data sources and novel analytic techniques to break ground on the frontier of data-driven environmental health policy design, which often involves evaluating existing policies along the way. Meanwhile, an overarching theme of this work is the bridging of different scientific domains and methodological areas, such as decision science with environmental justice, artificial intelligence with climate & health, and causal inference with remote sensing of both human activity and environmental factors.

Chapter 1 proposes a Monte Carlo-based methodology to compare realistic strategies for measuring and reporting daily air quality information. Specifically, we investigate the usefulness of low-cost air quality sensors, which are often lauded for their potential to fill spatiotemporal gaps in air quality information – especially for underserved populations, making them a central tool in environmental justice efforts. However, because many of these sensors are purchased and deployed by concerned citizens with financial means, the extra measurements are skewed towards more privileged areas. Also,

low-cost sensors have lower accuracy compared to reference-grade monitors. To characterize these tradeoffs, we design and implement a simulation study based closely on real data to evaluate the accuracy and equity of information from individuals' nearest air quality instrument (sensor or reference monitor) under both real and hypothetical low-cost sensor deployment scenarios. By varying the number of sensors deployed, the amount of sensor measurement error (noise), and the relative placement of the sensors (e.g. at schools, near major roads, and in communities with environmental justice concerns), we are able to analyze and make recommendations for how a local or regional government or organization might deploy the most effective and equitable network of low-cost air quality sensors, given their budget constraints. Further, the demonstrated simulation methodology can be adapted for other environmental monitoring objectives.

Chapter 2 develops a framework with which reinforcement learning (RL) can be used to optimize the issuance of heat alerts. Heat alerts are a practical and low-cost intervention to mitigate the public health impacts of extreme heat. However, current practice for issuing heat alerts does not take advantage of modern data science methods to optimize local alert criteria. To fill this gap, we harness RL (a branch of artificial intelligence) to build a model for whether to issue a heat alert on a given day, accounting for sequential dependence – which in the heat alert setting is due to both alert fatigue and finite resources/ability (of individuals/communities) to take health-protective measures. To use RL, we have to overcome several major incompatibilities between standard RL methods and the heat alert setting, which extend to environmental health / climate & health more generally. First, the relatively small and easily confounded signal in heat-alert-health relationships challenges the ability of RL algorithms to identify relevant effects, much less to optimize heat alert issuance. Second, mainstream RL methods are not suitable for settings with significant spatial heterogeneity, which is a known feature of heat alert health impacts. To address these challenges, we use a combination of statistical modeling, cutting-edge RL techniques, and conceptually simple yet effective modifications such as restricting alerts to extremely hot days, to learn heat alert issuance policies that reduce the adverse health impacts

of extreme heat compared to the current U.S. National Weather Service policy. We also prioritize interpretable characterization of the RL results, offering intuitive insights about which counties across the nation stand to benefit the most from implementing heat alert-RL. A major contribution of this project is establishing a connection between the environmental health and artificial intelligence communities.

Chapter 3 combines spatiotemporal causal inference methods and remotely sensed data to quantify the air quality impacts of national-level plastic waste policies, via the mechanism of trash burning at open dump sites. Especially in low- and middle-income countries (LMICs) without adequate solid waste management infrastructure, trash burning is a huge environmental public health problem that is difficult to quantify on large scales due to its distributed, intermittent, and not-infrequently covert nature. This burden is compounded by many high-income countries exporting massive amounts of plastic waste, both with and without LMICs' consent. Using Indonesia as a case study, we develop a strategy to quantify the local air quality impacts of China's 2018 waste import ban (and subsequent diversion of waste to other Southeast Asian countries), using a collection of remotely sensed data products to overcome the lack of ground-level monitoring and to strengthen the causal argument. Methodologically, we combine two strains of causal inference, ultimately considering the proximity to ports (from which international plastic waste enters the country) as an induced continuous exposure at locations where open dumping has been detected, and using past years of data as controls, conditional on meteorologic variation. Additionally, we extend past work's uncertainty quantification strategy to account for residual spatial dependence. This project not only reveals a statistically significant increase in Indonesian air pollution attributable to imported plastic waste, but also lays groundwork for future environmental policy evaluations in data-scarce settings.

Contents

TITLE PAGE	i
COPYRIGHT	ii
ABSTRACT	iii
TABLE OF CONTENTS	vi
LIST OF FIGURES	ix
LIST OF TABLES	xii
ACKNOWLEDGEMENTS	xv
 I INVESTIGATING USE OF LOW-COST SENSORS TO INCREASE ACCURACY AND EQUITY OF REAL-TIME AIR QUALITY INFORMATION	 I
1.1 Introduction	1
1.2 Materials & Methods	5
1.3 Results	15
1.4 Discussion	22

2	OPTIMIZING HEAT ALERT ISSUANCE WITH REINFORCEMENT LEARNING	28
2.1	Introduction	28
2.2	Related Work	31
2.3	Problem Setup	33
2.4	BROACH: An RL Environment for Optimizing Heat Alert Issuance	35
2.5	Learning & Evaluation	41
2.6	Conclusion & Future Work Directions	50
3	A SPATIOTEMPORAL, QUASI-EXPERIMENTAL CAUSAL INFERENCE APPROACH TO CHARACTERIZE THE EFFECTS OF GLOBAL PLASTIC WASTE EXPORT AND BURNING ON AIR QUALITY, USING REMOTELY SENSED DATA	53
3.1	Introduction	53
3.2	Data & Definitions	57
3.3	Methods	63
3.4	Results	75
3.5	Discussion	79
	REFERENCES	83
	APPENDIX A APPENDIX TO CHAPTER 1	98
A.1	Supplemental Figures and Tables for All Analyses	98
A.2	Supplemental Figures and Tables for the Analysis Weighted by Population Density	102
A.3	Supplemental Figures and Tables for the Analysis Not Weighted by Population Den- sity	108
	APPENDIX B APPENDIX TO CHAPTER 2	116
B.1	Heat Alerts MDP Terminology	116

B.2	Details of the Data Set	116
B.3	Details of the Bayesian Rewards Model	118
B.4	Details of the RL Analysis	121
B.5	Summary of the Workflow Using BROACH	124
B.6	Approximate Confidence Interval Calculation	125
B.7	Discussion of Absolute Health Benefits	125
B.8	Extended Results from the RL Analysis	126
APPENDIX C APPENDIX TO CHAPTER 3		132
C.1	Estimation Details	132
C.2	Supplementary Figures and Tables	133
C.3	Simulations Exploring the Spatial Weighted Bootstrap	140

List of figures

1.1	Visual Abstract	4
1.2	Contextual Maps	7
1.3	AQ Information Reporting Diagram	9
1.4	Distance and MAE	19
1.5	Unhealthy-Healthy Misclassifications	21
2.1	Overview of the Heat Alerts RL Framework	30
2.2	Unobserved Mediation of the Effect of Heat Alerts on Hospitalizations by Height- ened Awareness	37
2.3	Map of the Counties Considered and Their Regional Climate Zone Classifications	37
2.4	Example of Heat Alert Policies in a Setting Where Using RL Would Be Beneficial .	45
2.5	Comparison Between Policies' Performance, Stratified by Climate Zone	49
3.1	Summary of PM _{2.5} Time Series at Indonesian Waste Dump Sites	60
3.2	Port Proximity Index at Indonesian Waste Dump Sites	61
3.3	Main Results: Increase in PM _{2.5} (Post China Ban) vs Port Proximity Index	77
3.4	Pre-trends Test	77
3.5	Number of Active Fires Over Time at Indonesian Waste Dump Sites	78

A.1	Spatial Distribution of Marginalized Groups	98
A.2	Annual Average PM _{2.5} for Each Population Subset	99
A.3	Simulated and Empirical Sensor Measurement Error Distributions	99
A.4	Distance to the Nearest AQ Instrument, Among Observations Misclassified by More Than One Level of the AQI	102
A.5	Underclassifications of the AQI	103
A.6	Overclassifications of the AQI	106
A.7	Distance and MAE, Unweighted by Population Density	109
A.8	Unhealthy-Healthy Misclassifications, Unweighted by Population Density	111
A.9	Distance to the Nearest AQ Instrument, Among Observations Misclassified by More Than One Level of the AQI, Unweighted by Population Density	112
A.10	Underclassifications of the AQI, Unweighted by Population Density	113
A.11	Overclassifications of the AQI, Unweighted by Population Density	114
B.1	Plots of the Quantile Heat Index, Baseline Hospitalization Rate, and Alert Effec- tiveness Estimated by the Rewards Model, Stratified by Climate Zone.	119
B.2	Convergence of the Rewards Model	122
B.3	Fitted Coefficients from the Rewards Model	123
B.4	Example of Heat Alert Policies Under a Small Alert Budget	126
B.5	Example of Heat Alert Policies in a Setting Where Using RL Would Not Be Beneficial	127
B.6	Distributions of Different Policies' Heat Alert Timing and Lengths of Alert Streaks	128
B.7	Post-hoc Contrastive Analysis Between Four Methods	128
B.8	Post-hoc Contrastive Analysis Between Four Methods, Using More Interpretable Features	130
B.9	Post-hoc Contrastive Analysis Between the Best RL and National Weather Service	131

C.1	Number of Active Fires in Indonesia Over Time	133
C.2	Detection Date of Waste Dump Sites in Indonesia	134
C.3	Map of Waste Dump Sites in Indonesia, Indicating the Number of Active Fires at Each Site	134
C.4	Map of Waste Dump Sites in Jakarta, Indicating the Number of Active Fires at Each Site	135
C.5	Illustration of the Effect of a Smaller Bandwidth in LLKR	135
C.6	Confounder-Naive Results	136
C.7	Illustrating the Results (Confidence Intervals) with Different Spatial Correlation Magnitudes	137
C.8	Pre-trends Test Sensitivity Analysis	138
C.9	Sensitivity Analysis Including Site Area (Footprint) as a Covariate	138
C.10	Number of Active Fires vs Waste Dump Site Area	139
C.11	Number of Active Fires vs Port Proximity at Waste Dump Sites	139
C.12	Simulation Study Results Examples (Different Seeds)	142
C.13	Simulation Study Bias	143
C.14	Simulation Study Mean Absolute Error (MAE)	144
C.15	Simulation Study Confidence Interval Coverage	145
C.16	Simulation Study Confidence Interval Width	146

List of Tables

1.1	Comparing Impacts of Different Sensor Accuracies	27
2.1	Summary of the State Space, Specifying Exogenous and Endogenous Components	38
2.2	RL Evaluation Results	46
A.1	Descriptions of Contextual Datasets	100
A.2	Hypothetical Numbers of LCS in Well-Known Counties	101
A.3	Summary Statistics of the Di et al. Estimates	101
A.4	Descriptive Statistics, Weighted by Population Density	104
A.5	Non-normalized Descriptive Statistics, Weighted by Population Density	105
A.6	Comparing Impacts of Different Sensor Accuracies for LCS at All Schools	107
A.7	Descriptive Statistics, Unweighted by Population Density	108
A.8	Comparing Impacts of Different Sensor Accuracies, Unweighted by Population Density	110
A.9	Comparing Impacts of Different Sensor Accuracies for LCS at All Schools, Un- weighted by Population Density	115
B.1	MDP/RL Terminology in the Context of Heat Alerts	117
B.2	Descriptive Statistics of Heat Index, Alerts, and Hospitalizations	118

B.3	Descriptive Statistics of the 30 counties Used in Policy Optimization	129
B.4	Comparison of Results Using 2014-2016 as the Evaluation Years (Sensitivity Analysis) as Opposed to 2007,2011,2015 (Main Analysis)	130

TO GEOFF AND MICHELLE CONSIDINE, MY WONDERFUL PARENTS.

Acknowledgments

For the past five years, it has been such a privilege to participate in the National Studies on Air Pollution and Health (NSAPH) research group under the guidance of Francesca Dominici, Danielle Braun, and Rachel Nethery. Francesca, your pioneering vision and leadership in societally-impactful data science is a constant source of awe and inspiration. Thank you for bringing together such a vibrant research community and including me in it. Danielle, over time it has only become more apparent to me how many balls you keep in the air – with grace – to maintain and build momentum and transparency in NSAPH. Thank you for your candid feedback and gems of wisdom about everything from mentoring to grant applications. Rachel, even during our very first conversation I was struck by the sense that you would be an excellent advisor. From the get-go, you provided me with an ideal balance of freedom and support, while also nudging me to improve both as a researcher and as a communicator. Envisioning my next career chapter, my lofty goal as a mentor will be to emulate you.

Other integral contributors to this dissertation are Mauricio Tec, Priyanka deSouza, Greg Welle-nius, and Rajarshi Mukherjee. Mauricio, I am so glad that our paths crossed when they did and that we persevered in overcoming the numerous roadblocks to use reinforcement learning methods for environmental health. I learned more from you than from many of my graduate level courses combined. Thank you for showing me the ways of the computer scientists, with patience and good humor. Priyanka, getting to know you during my first year at Harvard was so serendipitous. Your creative, clear-eyed approach to science and community engagement is deeply inspiring. Thank you for helping me stay grounded in why this work matters and encouraging me to fight academic tunnel vision. Greg, it was a pleasure collaborating with you during the middle of my PhD. Thank you for helping me think critically about algorithm-driven decision making for environmental health, as well as

how to communicate key findings for a broad audience. I came away from every discussion with you with a greater appreciation for the nuances of climate & health research, social impacts, and policy decisions. Rajarshi, thank you so much for serving on my dissertation committee, in addition to the many other hats you wear in the Biostatistics Department. Your insightful questions and unfailingly positive attitude made my committee meetings and evaluations both constructive and enjoyable.

The wide range of this dissertation was enabled by a Graduate Research Fellowship from the National Science Foundation (NSF), paired with an Environmental Statistics Training Grant from the National Institute of Environmental Health Sciences (NIEHS). A huge thank you to both NSF and NIEHS / the National Institutes of Health at large, for investing in me and more broadly in the future of science to promote planetary health (“the health of human civilization and the natural systems on which it depends”).

Additionally, I would like to thank Harvard University and the T.H. Chan School of Public Health (HSPH) for providing a suite of resources which have supported and inspired my dissertation work. The computation for all three of the projects herein was supplied by the Harvard FAS Division of Science Research Computing. My years as a PhD student were also greatly enriched by engagement with the Harvard Data Science Initiative, the HSPH-NIEHS Center, the HSPH Pedagogy Fellows, the Salata Institute for Climate & Sustainability, and the HSPH Center for Climate, Health, and the Global Environment (C-CHANGE), alongside countless seminars hosted by groups across the university.

My ability to access and make full use of these resources was developed during my undergraduate years at the University of Colorado (CU) Boulder – through courses, as a data science intern in Earth Lab, and through additional mentoring. I find it hard to put into words just how fortunate I feel to have started working with my research mentor Colleen Reid eight years ago. In the Department of Applied Math, I want to give special thanks to Anne Dougherty, Vanja Dukic, Will Kleiber, and Eric Vance for helping me find my path in applied statistics and pursue prestigious opportunities; my ap-

proach to the latter was also transformed by the expert guidance of Deborah Viles. Other communities at CU Boulder helped me to both expand my horizons and find my center. In particular, the Engineering Honors Program, Engineers Without Borders, and Norlin Scholars left lasting impacts. Beyond CU Boulder, I would also like to thank the Goldwater and Astronaut Scholarship Foundations, for commending and investing in my pursuit of a research career focused on social/environmental impact.

Starting graduate school during the Covid-19 pandemic, the camaraderie I found in my cohort and among other Biostatistics PhD students lifted my spirit and helped me push through challenging courses. Over the years, I have continued to appreciate this department's community, as well as other pockets of community I have discovered in the School of Public Health and Harvard at large.

Outside of these institutions, my journey would not have been possible without the support of my family, and would not have been pleasant without my friendships, many of which were sustained via phone and Zoom calls during these years. Because listing all the names of individuals close to my heart would be excessive in this setting, I will leave it at this. To my husband Nate, thank you for bringing a smile to my face every single day of the last 4.5 years, including the six semesters when we lived 2,000 miles apart. To my dad, thank you for raising me to think critically and for helping me learn how to strive with intentionality for things that will matter in the long run. To my mom, thank you for helping me understand myself as a whole person and for showing me how to develop and maintain the human connections that fuel and inform everything else. And to Nate, my parents, and my brother Wyatt: thank you for loving me in my entirety, for believing in my strength while also reminding me that I don't always have to be strong.

1

Investigating Use of Low-Cost Sensors to Increase Accuracy and Equity of Real-Time Air Quality Information*

1.1 INTRODUCTION

DECADES OF RESEARCH HAVE DOCUMENTED THE ADVERSE HEALTH IMPACTS OF BOTH SHORT- AND LONG-TERM EXPOSURE TO AIR POLLUTION. In this study, we focus on fine particulate matter (PM_{2.5}), an air pollutant that has been associated with various adverse health outcomes^{36,37,156}. In

*This work has been published. <https://pubs.acs.org/doi/abs/10.1021/acs.est.2c06626> The full citation is: Considine EM, Braun D, Kamareddine L, Nethery RC, deSouza P. Investigating Use of Low-Cost Sensors to Increase Accuracy and Equity of Real-Time Air Quality Information. *Environmental Science & Technology*. 2023 Jan 9;10.1021/acs.est.2c06626.

the US, although air quality (AQ) across the years has been improving overall, disparities between PM_{2.5} concentrations experienced by subpopulations persist^{78,24,141}. In addition, certain parts of the US are experiencing, or are likely to experience, higher air pollution (including PM_{2.5} exposure) from climate-change related events and processes such as exacerbated wildfires and dust storms¹¹⁶. To develop effective AQ management plans and address key concerns of equity, accurate and high resolution AQ monitoring data are needed.

The Federal Reference Method or Federal Equivalence Method (FRM or FEM) monitors deployed by the US Environmental Protection Agency (EPA) are considered the “gold standard” for measuring AQ. However, due to the prohibitively high capital (USD \$10,000+) and operating costs¹⁰³ of these instruments, they have been deployed in less than a third of all US counties⁸¹. Even in counties that have an EPA monitor, the often substantial within-county variability in AQ can render measurements unrepresentative of the pollution levels experienced by many residents of the county⁴⁶. As these monitors tend to be deployed in more populous locations^{59,163}, residents in rural areas tend to be farther away from monitors.

In recent years, low-cost AQ sensors (< USD \$2,500 as defined by the EPA Air Sensor Guidebook¹⁵⁴, but often much cheaper) have been gaining attention as a supplement to FRM/FEM monitoring, and many are deployed by private citizens in addition to public and private entities³². Key motivations of individuals (ascertained from online product reviews) include managing health impacts of wildfire smoke and other air pollution as well as detecting air pollution sources of concern³³. Networks of these sensors can help increase the spatial resolution and frequency of AQ measurements⁹⁷. However, the measurements from low-cost sensors (henceforth, LCS) have lower accuracy than EPA monitors, and can be further affected by environmental conditions such as relative humidity, temperature, and aerosol composition^{32,77,28,95}. Recent work has shown that algorithmic correction can reduce the error in LCS, but does not eliminate the problem^{25,176,99,68}. These studies have also highlighted the challenge of developing corrections that are transferable across measurements collected in

different locations and/or time periods. Despite these drawbacks, research has shown that LCS can be useful in identifying air pollution hotspots, expanding local awareness about AQ and health, and alerting the public of short-term changes in AQ, which may facilitate reduction of exposure to air pollution, e.g. by individuals choosing not to exercise outdoors on a high pollution day^{154,14,105,22}.

To our knowledge, the question of how incorporating measurements from LCS into real-time reporting systems affects the accuracy of people's AQ information remains unanswered. In addition to sensor measurement error, an important factor that must be considered is the spatial distribution (both spatial density and relative placement) of LCS. In this study, we assume that individuals view AQ data from the nearest instrument – EPA monitor or LCS – (e.g., as shown on a smartphone app) as their current AQ exposure. Despite the potential for LCS to increase and democratize access to AQ information, a 2021 study found that locations of PurpleAir (one of the most widely used LCS brands, costing around \$250 per sensor) in the US tend to be disproportionately located in neighborhoods that have higher incomes and higher percentages of white residents, compared both to the locations of EPA monitors and to the US overall³². This suggests that, relative to more privileged groups, residents of marginalized neighborhoods (who, as previously noted, tend to experience worse air quality) may have less access to information about their local AQ, a precursor to adaptive actions which could be taken to protect their health.

In summary, while collection and dissemination of LCS data have been increasing, there is a need for evaluating the impact of integrating LCS data into AQ reporting platforms. In this study, we investigate how altering the number of LCS deployed, amount and type of sensor measurement error, and relative placements of LCS affects the accuracy of daily AQ information available to individuals from their nearest AQ instrument (defined to be either an LCS or an EPA monitor). Our main objectives are to (a) increase the nuance with which various groups (scientists, policy makers, community organizations, etc.) can think about and discuss tradeoffs when it comes to measuring and reporting AQ to the public and (b) to suggest directions for future work to both the LCS instrument / data

science and environmental health / policy communities. We focus on $PM_{2.5}$ and base our LCS investigation on PurpleAir sensors, the data from which is used in a number of popular regional real-time AQ maps^{75,2}.

Our analysis consists first of simulating realistic LCS $PM_{2.5}$ measurements under numerous hypothetical LCS deployment and sensor measurement error scenarios in the state of California. Then, we compare the local AQ information available to individuals in each simulation scenario with that produced by (i) EPA monitors only, as well as (ii) the existing PurpleAir sensor network. We dedicate special attention to evaluating how each scenario impacts disparities in AQ information accuracy for marginalized groups, such as those living in communities with high rates of poverty or with high proportions of nonwhite or Hispanic residents. Our findings can be used to inform decisions about (a) where to place LCS to make real-time AQ reporting more accurate and equitable, (b) how many LCS to deploy, (c) whether existing sensor calibration approaches yield sufficient accuracy to justify use of LCS for real-time AQ reporting, and (d) what amount of error is “tolerable” for future LCS deployments.

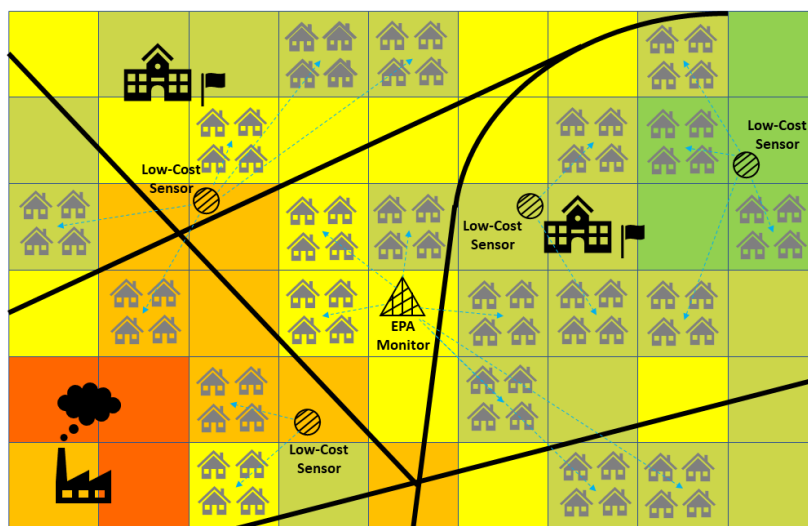


Figure 1.1: Visual abstract, previewing our approach.

In section 2 (Materials & Methods), we describe each step of the analysis and the sources of data used. Section 3 (Results) includes a comparison of the impact of different types and amounts of sensor measurement error for LCS at current PurpleAir locations, as well as a comparison across different LCS placement strategies and numbers of LCS deployed. Section 4 (Discussion) includes conclusions, limitations, and some ideas for future investigation.

1.2 MATERIALS & METHODS

1.2.1 STUDY SETTING AND OVERVIEW

To evaluate the potential impact of LCS measurements on localized AQ information accuracy, our study leverages real data on EPA monitor locations, PurpleAir LCS locations, and sociodemographic and geospatial features in California. Our choice to situate the study in California was primarily motivated by California’s widespread LCS uptake (California contains over half of the US’s PurpleAir sensors), in part prompted by concerns about increasing air pollution from wildfires^{18,58,62,94}.

To ensure that our simulations accounted for realistic spatial and temporal variability in $PM_{2.5}$, we assumed the “true” (error-free and comprehensive) daily ambient $PM_{2.5}$ concentrations were those obtained from an ensemble model predicting $PM_{2.5}$ exposures daily at 1 km x 1 km in 2016, created by Di et al. (2016 was the most recent year for which these predictions were available)³⁵. These estimates agreed well with ground-based reference measurements: the 10-fold cross-validated R^2 was 0.86 for the US overall and 0.80 for the Pacific coast states (including California). We also considered the current locations of EPA monitors ($n=154$ in California) to be fixed, and in our simulations, we set the $PM_{2.5}$ measurements from each of these monitors to be equal to the Di et al. $PM_{2.5}$ estimates at these locations (i.e., we assumed that there was no error in the measurements from these monitors).

Details of our simulation procedure are provided in the following subsections. Here is a brief overview to serve as a roadmap:

1. Using the placement strategy specified by the simulation scenario, select hypothetical locations of LCS
2. For each 1 km x 1 km grid in California, identify the grid centroid's nearest AQ instrument (among the real EPA monitor locations and hypothetical LCS locations)
3. For each day in 2016, simulate LCS PM_{2.5} measurements by adding simulated device measurement error to the "true" PM_{2.5} estimates from Di et al. at each hypothetical LCS's location
4. Evaluate the accuracy and equity of AQ information observed (based on the measurements from the nearest AQ instrument) across all 475,772 grids in California and 366 days in 2016

1.2.2 SELECTING LCS LOCATIONS

Table A.1 (in Appendix A) describes the data sets, data processing steps, and sampling methods used to select locations for LCS in each simulation.

To guide hypothetical LCS deployment strategies focused on environmentally and socially marginalized communities, we leveraged the CalEnviroScreen (CES) index, developed by the California Office of Environmental Health Hazard Assessment, which describes both environmental and socioeconomic-demographic marginalization at the Census tract level⁷, as well as its environmental component, henceforth referred to as the Pollution Score. The Pollution Score incorporates data on air pollution (ozone, PM_{2.5}, diesel PM emissions, toxic chemical releases from facilities) and traffic density; pesticides, groundwater threats, impaired water bodies, and drinking water contamination index; solid waste, hazardous waste, and cleanup sites. The Pearson correlation between the Pollution Scores and the PM_{2.5} estimates from Di et al. is 0.48. The socioeconomic-demographic disadvantage index used by CES incorporates data on asthma, low birth weight, cardiovascular disease, education, linguistic isolation, poverty, unemployment, and housing burden. The CES Score is a product of these environmental and socioeconomic-demographic indices.

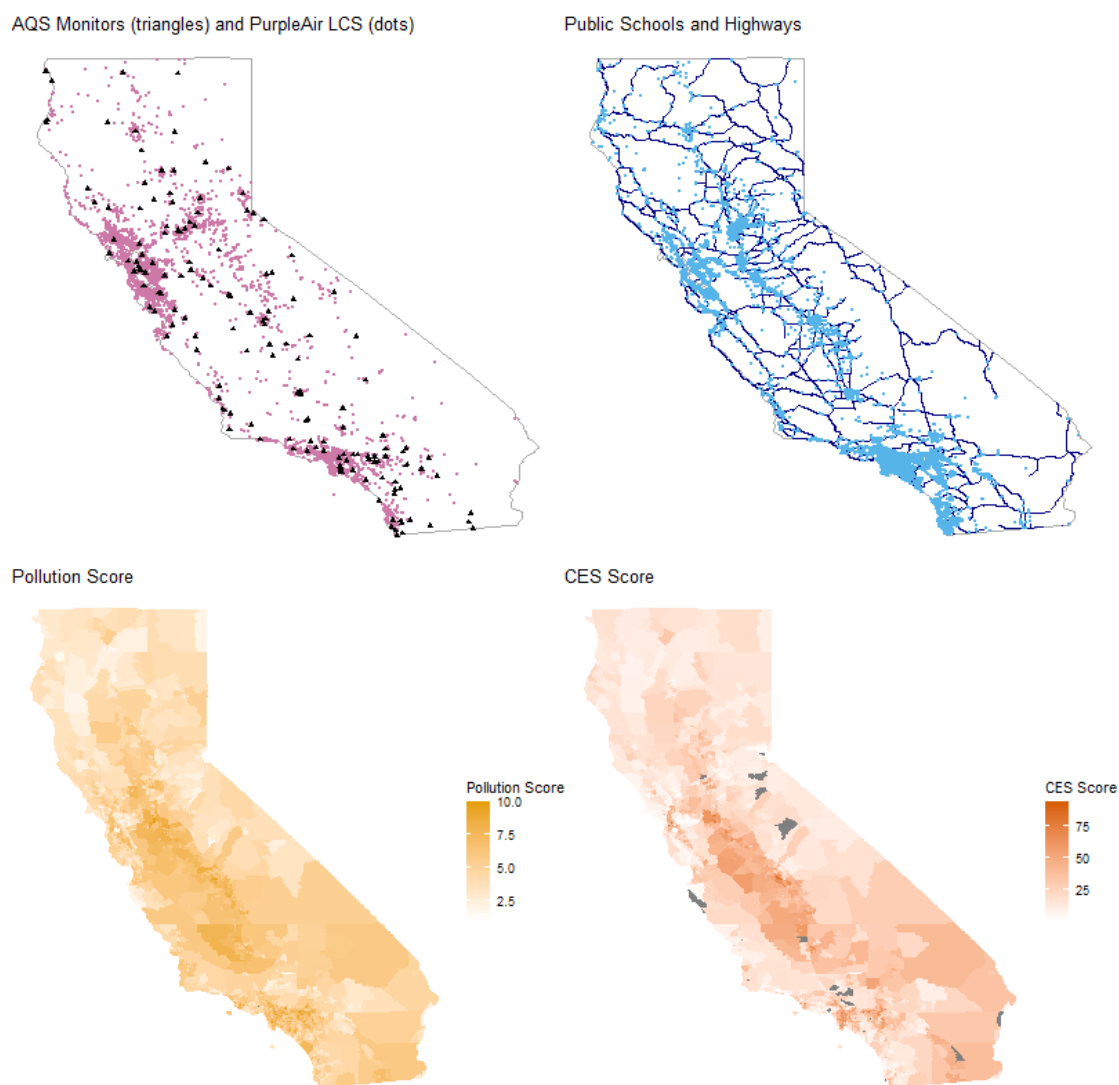


Figure 1.2: Visualizing data used to guide the hypothetical LCS placement strategies. Locations of EPA monitors and PurpleAir LCS (panel a), public schools and major roads in California (panel b), as well as maps of CES (panel c) and Pollution Score (panel d). Data sources (all public domain / open access): EPA AQS and PurpleAir (panel a), NCES and NHPN (panel b), California EPA's OEHHA (panels c and d).

In this study, we considered the following five hypothetical LCS placement strategies, illustrated in Figure 1.2: (a) at randomly selected real outdoor PurpleAir locations, (b) at randomly selected public schools, (c) at randomly selected locations favoring proximity to major roads, (d) at randomly

selected locations favoring high CES Score, and (e) at randomly selected locations favoring high Pollution Score. For the last three placement strategies, “favoring” refers to weighted random sampling (respectively using nearby road lengths, CES Score, and Pollution Score as weights) to determine placement locations. We compared each of these placement strategies across different numbers of sensors deployed (0, 50, 100, 250, 500, and 1000 LCS to show the trends). To provide context, average numbers of LCS assigned to the Los Angeles, Sacramento, Imperial counties (a large city, a medium-small city, and a well-known environmental justice focus area) under each placement strategy are provided in Table A.2.

1.2.3 OBSERVING AQ INFORMATION FROM THE NEAREST AQ INSTRUMENT

We assumed that all individuals in each 1 km x 1 km grid in California observed daily AQ measurements from the instrument (either EPA monitor or LCS based on simulated placement strategy) nearest to their grid centroid, as shown in Figure 1.3. In the True Air Pollution Exposure column (left), the background color in each grid cell represents the “true” air pollution that individuals experience (obtained from the Di et al. estimates). Note that these true exposures are most often not observed. The colors inside the triangles and circles represent the measurements from EPA monitors and LCS, respectively. LCS measurement error is represented in the bottom row, where the color inside the circle differs from the background color of the grid it’s in. In the Shown Air Pollution Exposure column (right), the background color in each grid is the air pollution measurement that individuals observe from their nearest AQ instrument. The differences between the AQ that individuals experience and the AQ that they are shown are indicated by the red and blue X’s in Figure 1.3. The red X’s indicate cells where AQ is over-classified, i.e., the AQ shown to residents is worse than the AQ truly experienced. The blue X’s indicate cells where AQ is under-classified, i.e., the AQ shown to residents is better than the AQ truly experienced.

Figure 1.3 illustrates three distinct sources of AQ reporting error to which we will refer through-

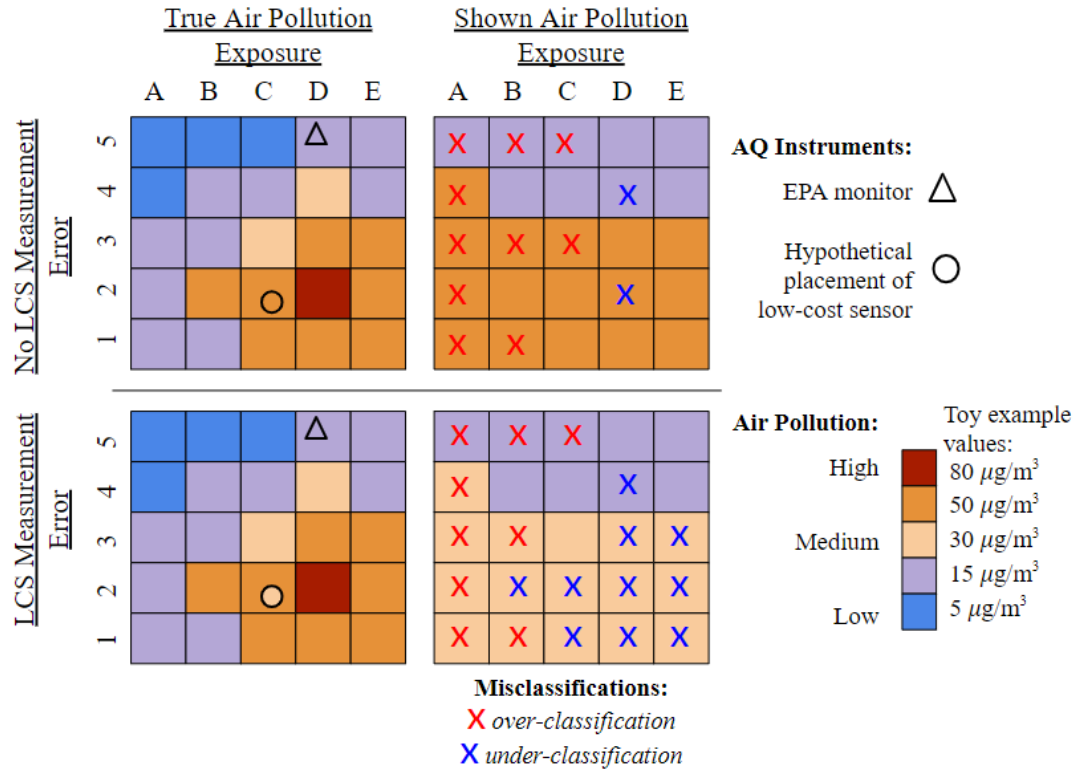


Figure 1.3: An illustration of the assumed process of AQ information reporting (from each grid's nearest AQ instrument), under two scenarios: without LCS measurement error (top row of panels) and with LCS measurement error (bottom row of panels). In the True Exposure column, the background color of each grid represents the (unobserved) air quality an individual in that grid experiences, whereas in the Shown Exposure column, the background color represents the air quality an individual in that grid observes from their nearest AQ instrument.

out the rest of this paper: (i) distance to the nearest AQ instrument, (ii) local variability in air quality, and (iii) sensor measurement error. As an example of distance-based errors, the distance between an individual in A₅ and the nearest AQ instrument is large, so an individual in A₅ is unlikely to be shown accurate measurements of their air pollution exposure (in Figure 1.3, they would be shown that their exposure is $15 \mu\text{g}/\text{m}^3$ instead of the true exposure, which is $5 \mu\text{g}/\text{m}^3$). As an example of local variability-based errors, while an individual in C₅ is close to an EPA monitor (D₅), local variability in AQ between C₅ and D₅ results in misclassification of C₅'s AQ (they would be shown that their exposure is $15 \mu\text{g}/\text{m}^3$, while their true exposure is $5 \mu\text{g}/\text{m}^3$). Even if a cell contains an LCS, sensor

measurement error may still result in reporting error, as in the case of an individual in C2: under the setting of device measurement error, they would be shown that their exposure is $30 \mu\text{g}/\text{m}^3$, instead of their true exposure, which is $50 \mu\text{g}/\text{m}^3$. These effects can also co-occur, as for an individual in D2: the nearest AQ instrument is in cell C2, which, in addition to having lower air pollution than D2, also suffers from LCS measurement error. In this study, we help disentangle these effects.

1.2.4 SIMULATING SENSOR MEASUREMENT ERROR

In each hypothetical (simulated) LCS deployment scenario, daily $\text{PM}_{2.5}$ “measurements” from LCS were generated by adding sensor measurement error to the Di et al. $\text{PM}_{2.5}$ estimates, in several different ways. First, we selected measurement error distributions informed by a tiered target for AQ instrument accuracy proposed at an EPA workshop¹⁶⁸. The proposal is that AQ measurements (i) for regulatory purposes require accuracy of $\pm 10\%$ of the true average $\text{PM}_{2.5}$ in that area, (ii) for mapping spatial gradients and monitoring microenvironments require accuracy of $\pm 25\%$, and (iii) for hotspot detection require accuracy of $\pm 50\%$. In our simulations, sensor measurement errors were generated both (a) differentially with respect to true $\text{PM}_{2.5}$ and (b) non-differentially (with error magnitude not varying across $\text{PM}_{2.5}$ levels). The former was motivated by empirical LCS observations, and accounts for the possibility of some spatial and temporal correlation in the sensor measurement errors due to some spatial and temporal smoothing induced by the Di et al. modeling approach. The latter assumes independence of the sensor measurement errors (post calibration of the LCS).

Second, we simulated LCS measurement errors in a manner enabling assessment of a nation-wide correction algorithm for PurpleAir sensors developed by EPA researchers¹⁰. Specifically, we sampled errors from the empirical distribution of residuals obtained by comparing measurements from EPA monitors in California to the corrected measurements from collocated PurpleAir sensors.

These procedures are detailed in the following subsections. Comparison of the characteristics and effects of all these different types and amounts of sensor measurement error is facilitated by Table 1.1

in the Results.

1.2.5 GENERATING ERRORS BASED ON PROPOSED PERFORMANCE TARGETS

As a benchmark, we first ran simulations using no sensor measurement error (as if the LCS were as accurate as the EPA monitors). Then, we used two levels of both non-differential and differential sensor measurement error. For a given accuracy level (either $\pm 10\%$ or $\pm 25\%$, as proposed in the EPA performance targets workshop¹⁶⁸), we simulated non-differential sensor measurement error by sampling from a normal distribution with mean zero and standard deviation of 5 (the unweighted 2016 average PM_{2.5} in California) multiplied by the accuracy level, yielding 0.5 and 1.25 respectively. In mathematical notation:

$$\{\text{LCS PM}_{2.5}\} = \{\text{Di et al. PM}_{2.5}\} + \varepsilon_n$$

where $\varepsilon_n \sim \text{Normal}(\text{mean} = 0, \text{sd} = \{0.1 \text{ or } 0.25\} * 5)$.

The differential sensor measurement error, on the other hand, was motivated by our observation of the positive correlation between the magnitude of the EPA correction residuals and the true PM_{2.5} (described in the following subsection). For the two accuracy levels, we simulated differential sensor measurement error by sampling from a normal distribution with mean zero and standard deviation of the true PM_{2.5} (for each grid and day) multiplied by the accuracy level. In mathematical notation:

$$\{\text{LCS PM}_{2.5}\} = \{\text{Di et al. PM}_{2.5}\} + \varepsilon_d$$

where $\varepsilon_d \sim \text{Normal}(\text{mean} = 0, \text{sd} = \{0.1 \text{ or } 0.25\} * \{\text{Di et al. PM}_{2.5}\})$.

1.2.6 SAMPLING FROM THE EMPIRICAL DISTRIBUTION OF PURPLEAIR ERRORS

To obtain an empirical distribution of measurement errors for our simulated LCS, we compared PurpleAir measurements and EPA monitor measurements for collocated EPA monitor and LCS pairs in California. Note that these data were not used directly in the simulations: the estimates from Di et al. were considered to be the true exposure. We obtained daily $PM_{2.5}$ averages for both the monitors and LCS from the year 2020 (as opposed to 2016, the year for which we had 1 km x 1 km estimates from Di et al. to use as “truth” in the simulations) because most of the PurpleAir sensors have been deployed since 2016. We used the purpleair Python package to obtain measurements of $PM_{2.5}$, temperature, and relative humidity from PurpleAir sensors (channel A on 1/11/22 and channel B on 1/23/22), and downloaded the EPA monitor measurements from the EPA Air Quality System annual summary files repository¹⁵⁵.

Following the 24-hour quality assurance methods of the EPA researchers who developed a national correction for PurpleAir data¹⁰, we identified all pairs of EPA monitors and outdoor PurpleAir sensors within 50 meters of each other. Then, also following their quality assurance methods, we only considered EPA monitor averages based on at least 18 hours of observation, and PurpleAir averages based on at least 90% of the 2-minute observations, where the difference in daily averages of the two PurpleAir channels was less than $5 \mu g/m^3$ or 61%. Then, we averaged the two PurpleAir channels. This procedure yielded 5,759 paired daily observations from 22 EPA monitor-PurpleAir locations. To make this distribution more realistic for 2016, we removed 48 daily observations where the EPA monitor $PM_{2.5}$ was greater than $112 \mu g/m^3$ (the maximum value in the 2016 Di et al. estimates), yielding $n = 5,711$.

The EPA correction equation developed for PurpleAir sensors in the US is

$$PM_{2.5} = 0.524 \times \{\text{PurpleAir } PM_{2.5}\} - 0.0862 \times \{\text{Relative Humidity}\} + 5.75$$

After applying this correction to the PurpleAir data, we compared the corrected values to the collocated EPA monitor measurements. The resulting RMSE (root-mean squared error) was $5.33 \mu\text{g}/\text{m}^3$ and R^2 was 0.81. We refer to the difference between these EPA monitor and corrected LCS measurements as the empirical residuals.

Figure A.3 shows that the magnitude of empirical residuals increased as the true $\text{PM}_{2.5}$ concentration (as measured by the EPA monitors) increased. Sensor measurement error increasing with $\text{PM}_{2.5}$ concentration (i.e., differential measurement error) has been observed by some other groups as well²⁸. Such a relationship was not observed for relative humidity nor ambient temperature. Thus, for the empirically based LCS measurement error simulation, we sampled from the distribution of residuals corresponding to the decile of true $\text{PM}_{2.5}$ (the Di et al. estimates) at each grid and day. A summary of the Di et al. estimates mapped into these deciles is given in Table A.3.

1.2.7 EVALUATING THE AQ INFORMATION

The final step of each simulation was to evaluate the error between the true AQ exposures (which are most often unobserved) and the exposures reported by the nearest AQ instrument, summarized across all the grids and days. We evaluated accuracy in AQ reporting using the mean absolute error (MAE) in reported $\text{PM}_{2.5}$ concentrations and the misclassification rates of the U.S. AQ index, or AQI¹¹.

The AQI classifies AQ into six levels, with different public health recommendations for each. Green = “Good”, Yellow = “Moderate”, Orange = “Unhealthy for Sensitive Groups”, Red = “Unhealthy”, Purple = “Very Unhealthy”, and Maroon = “Hazardous”. AQI is often reported as a combination of air pollutants, however, for this analysis we used the single-pollutant version for $\text{PM}_{2.5}$ ⁴⁷. We hypothesize that most people use these classifications to inform their activity rather than the exact concentrations of $\text{PM}_{2.5}$ (or any other air pollutant), so we calculated the percent of over- and under-classifications, the percent of misclassifications greater than one level (e.g. Orange → Green or Yellow → Red), and what we term Unhealthy-Healthy Misclassifications (UHM): the fraction of days that a

healthy (H) classification is shown, out of the days that are truly unhealthy (U). This last metric may be of the most concern for public health. For this dichotomous variable, we defined Green and Yellow to be healthy, and Orange through Maroon to be unhealthy.

When calculating these metrics, weighting each grid by its population density allows us to evaluate the accuracy of AQ information available to individuals in California. We also performed the calculations unweighted by population density, which represents averaging across the land area instead of averaging across individuals. However, as the population-weighted metrics are more relevant for public policy (e.g., for health impact assessments), we focus on these results in the main text; the unweighted results are provided in the SI.

For each combination of sensor placement strategy, number of sensors, and type/amount of sensor measurement error, we ran 100 replicates and averaged the metrics across them to account for random sampling variability. The results are robust to the number of replicates (i.e., 50 vs 100).

1.2.8 EQUITY ANALYSIS

In addition to the overall population metrics (averaging across all grids in California), we calculated the metrics for marginalized subsets of the population, to determine if certain sensor placement strategies resulted in more equitable or less equitable access to AQ information.

We obtained socio-demographic features from the 2016 American Community Survey³ using the R package `tidycensus`¹⁵⁹, at the finest spatial resolution for which they were available: Census block group (CBG) level for race/ethnicity and population density, and Census tract level for socioeconomic status. To merge these features with the AQ data, we performed an overlay of the CBG shapefile with the 1 km x 1 km grid centroids from the Di et al. estimates. Any block group or tract that did not contain a grid centroid was ignored. Although this procedure tends to exclude more CBGs with smaller land area and higher population density, our main analysis weighting by population density counteracts possible bias.

For this analysis, we used percentage of non-white individuals (all but non-Hispanic white) for the percent marginalized by race/ethnicity and the percent of people living under the poverty line for those marginalized by socioeconomic status. Our decision to use one minus the percent of non-Hispanic white people to represent disadvantage by race/ethnicity (elsewhere referred to as “% nonwhite” for verbal simplicity) was informed by a preliminary calculation showing that Hispanic white people on average experience higher pollution and socioeconomic disadvantage than the overall white population. We defined CBGs with high % nonwhite and high % poverty to be CBGs that fell into the top quintiles of these measures across the 1 km x 1 km grid centroids ($\geq 58.1\%$ and $\geq 23.5\%$ respectively). For the 0.3% of CBGs that were missing data on % nonwhite, we substituted in Census tract-level % nonwhite. Data for % poverty was only available at the tract level. Only for one tract with population > 0 (tract 6037920200 with population 5,000) were we unable to retrieve tract-level data, and thus we omitted that tract from the analysis.

In summary, for the equity analysis we calculated the AQ reporting error in CBGs with high % poverty and high % nonwhite residents. Maps of % poverty and % nonwhite residents in California are provided in Figure A.1.

1.2.9 SOFTWARE AVAILABILITY

Code used to download and process the datasets, run the analyses, and generate the figures and tables can be found at https://github.com/ElleenConsidine/LCS_placement_sims

1.3 RESULTS

1.3.1 DESCRIPTIVE STATISTICS

Summary statistics of average annual $PM_{2.5}$ (according to the 1km x 1km estimates by Di et al.), % poverty, CES Scores, % nonwhite, and population density, for population subgroups as well as loca-

tions (centroids of 1 km x 1 km grid cells) targeted in each of the LCS placement strategies, are shown in Table A.4. One key observation, consistent with the environmental justice literature^{78,24,141}, is that CBGs with high % poverty or high % nonwhite have higher annual average PM_{2.5} than the population overall. As shown in Figure A.2, these marginalized subgroups also experience far more days classified as unhealthy by the AQI (level orange and higher) than the overall population. One differentiator between the two marginalized subgroups (high % nonwhite and high % poverty) is that CBGs with high % nonwhite tend to have higher population density than CBGs with high % poverty.

Another observation from Table A.4, consistent with external findings³², is that among real EPA monitor locations, PurpleAir locations, and the hypothetical LCS placement strategies considered, PurpleAir locations have by far the lowest % poverty. By contrast, EPA monitor locations have higher % poverty than any of the LCS placement strategies considered. Among the LCS placement strategies considered, schools and locations favored by CES Score have the highest % poverty. LCS placements at schools also have the highest % nonwhite out of any of the EPA monitor, PurpleAir, or hypothetical LCS locations.

Finally, while schools and PurpleAir locations tend to be in CBGs with higher population density, locations chosen to favor proximity to roads, CES Score, and Pollution Score tend to have lower population density.

1.3.2 COMPARING DIFFERENT TYPES AND AMOUNTS OF SENSOR MEASUREMENT ERROR, ASSUMING LCS ARE PLACED AT CURRENT PURPLEAIR LOCATIONS

The average distance to the nearest EPA monitor is 10.11 km for the population overall, 8.94 km for CBGs with high % nonwhite residents, and 8.69 km for CBGs with high % poverty. When we include LCS at all current locations of outdoor PurpleAir sensors, the average distance to the nearest AQ instrument drops dramatically, to 2.41 km, 2.49 km, and 2.82 km, respectively. Note that these results, like all those in the main text, are weighted by population density.

Table 1.1 summarizes how the accuracy of daily AQ information changes when we compare the scenario where people only have access to EPA monitors, with the scenario where people have access to EPA monitors and LCS at current locations of outdoor PurpleAir sensors. Under the scenario with LCS at PurpleAir locations, we compare the different sensor measurement error types, as described in the Methods. The first column of Table 1.1 shows the amount of sensor measurement error under each measurement error type (calculated as the standard deviation of the mean-zero simulated errors). For all the differential sensor measurement error scenarios (10%, 25%, and empirical residual-based), marginalized subgroups on average experience higher sensor measurement error because their air pollution exposure is higher.

Next, we use several metrics to describe the accuracy of observed AQ information under each scenario: (a) absolute error (deviation from the true exposure value), captured using the mean absolute error (MAE) and 95th percentile (to illustrate the upper end of the error distribution in addition to the mean), (b) rate of misclassification (either over- or under-classification) of the AQI, and (c) rate of Unhealthy-Healthy misclassifications (UHM), which we define as the fraction of days with unhealthy AQI that are misreported as healthy AQI.

Table 1.1 shows that when the LCS have no sensor measurement error (i.e. they are as accurate as EPA monitors in our simulations), deploying them at all the real locations of PurpleAir sensors roughly halves the MAE and 95th percentile of error in daily reported air quality. These improvements are smaller for CBGs with high % poverty, likely because PurpleAir sensors tend to be situated in more socioeconomically privileged areas.

Counterintuitively, even in the absence of sensor measurement error, placement of LCS at current PurpleAir locations leads to increases in the rates of under-classification of the AQI and UHMs. These reductions in classification accuracy are likely due to local variability in AQ and the fact that PurpleAir locations have lower annual average $PM_{2.5}$ than the state overall (Table A.4). This issue is exacerbated by sensor measurement error.

With non-differential sensor measurement error, LCS (at all the current PurpleAir locations) with error magnitudes of $\pm 10\%$ and $\pm 25\%$ both generally improve on the no-LCS case except for CBGs with high % poverty, where the MAE increases slightly. The impact of sensor measurement error is likely exacerbated for these CBGs with high % poverty due to the socioeconomic bias of PurpleAir locations. By contrast, while 10% differential sensor measurement error improves on the no-LCS case in terms of absolute error, 25% differential error and empirical residual-based error worsen the real-time AQ reporting for all groups and by all metrics, except for some small reductions in over-classification of the AQI.

These results highlight the potential for (i) LCS to reduce both the distance to the nearest AQ instrument and the absolute error in daily $PM_{2.5}$ reporting, (ii) the accuracy of classification-based AQ information to diverge from the accuracy of concentration-based AQ information, (iii) different AQ information outcomes for different subsets of the population, which is related to LCS placement characteristics, and (iv) the dependence of these insights on the type and amount of sensor measurement error.

1.3.3 ADDITIONALLY COMPARING PLACEMENT STRATEGIES AND NUMBERS OF SENSORS DEPLOYED

We now discuss how different hypothetical LCS placement strategies and numbers of LCS deployed affect access to real-time AQ information under both non-differential and differential sensor measurement error scenarios.

Figure 1.4 shows the average distance to the nearest AQ instrument as well as the MAE resulting from simulations under each hypothetical LCS deployment scenario (i.e., each combination of placement strategy and number of LCS deployed) and for each sensor measurement error type. The vertical scales of each plot are different, to facilitate close inspection of the lines. The plot for 10% non-differential measurement error is not shown because it is very similar to panel b (the results with

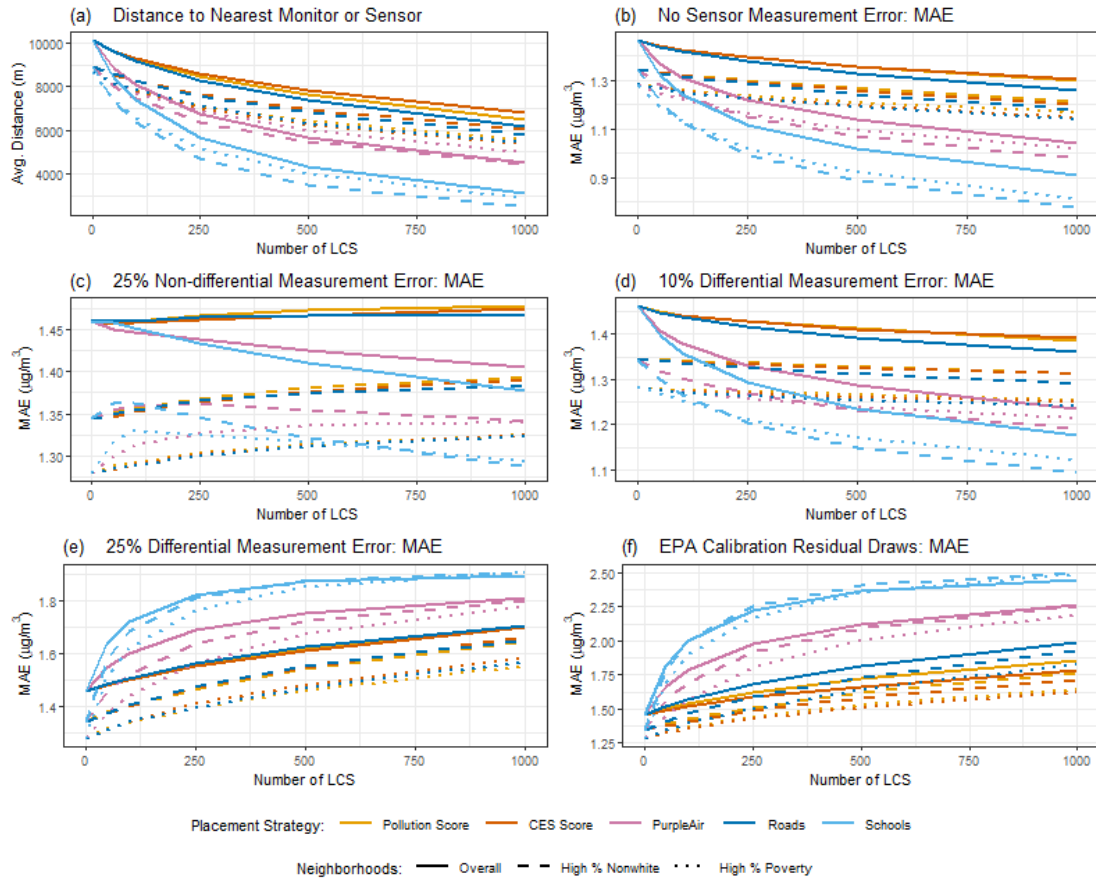


Figure 1.4: Distance to the nearest AQ instrument and mean absolute error (between what is reported vs. experienced) resulting from different numbers of LCS deployed, LCS placement strategies, and sensor measurement error types and amounts. All results were calculated using 366 days and averaged across 100 simulation replicates, weighted by population density. Panel a shows the distance to the nearest AQ instrument (monitor or LCS), panel b shows MAE when there is no LCS device measurement error, panel c shows MAE when LCS device measurement error is 25% non-differential, panel d shows MAE when LCS device measurement error is 10% differential, panel e shows MAE when LCS device measurement error is 25% differential, and panel f shows MAE when LCS device measurement error is sampled from the empirical distribution of residuals from PurpleAir LCS measurements (calibrated with the EPA equation) compared with collocated EPA monitor measurements.

no sensor measurement error).

The six panels in Figure 1.4 help distinguish between the three different components of error in real-time AQ reporting. First, we observe that with zero or low amounts of LCS measurement error, much of the average error in daily AQ reporting is due to individuals' distance to the nearest

AQ instrument, as illustrated by the similarity in line ordering and slopes between panels a, b, and d. Another observation, however, is that the impact of reduction in distance to the nearest AQ instrument can be confounded by local variability in AQ. For example, while LCS placements favoring CBGs with high Pollution Score result in greater reduction in distance to the nearest AQ instrument than placements favoring CBGs with high CES Score (panel a), the CES Score placement strategy results in lower MAE (panel b; the solid yellow line hides the solid orange line). This phenomenon is explained by local variability in AQ because the Pollution Score primarily highlights areas with large local sources of pollution, so measurements in those pollution “hotspots” may not be representative of air pollution even in nearby communities.

For low amounts of sensor measurement error (i.e. 10% non-differential and differential, as shown in panel d), reductions in daily AQ reporting error due to decreased distance to the nearest AQ instrument mitigate the impact of the LCS measurement error, improving MAEs across the board. When sensor measurement error is increased to 25% non-differential (panel c), deploying LCS only improves MAE under certain placement strategies: at schools and in current PurpleAir locations. This is largely because these placement strategies prioritize areas with higher population density.

With 25% differential and empirical residual-based sensor measurement errors (panels e and f), the impact of reduced distance to the nearest AQ instrument is overshadowed by the increased error in the sensor measurements, worsening AQ reporting across the board. Under these large amounts of sensor measurement error, placements favoring high CES Score result in the least error in AQ reporting. In nearly all cases, marginalized subgroups experience lower MAE than the population overall.

Figure 1.5 shows the rate of UHMs for all placement strategies, numbers of LCS deployed, and sensor measurement error types and amounts. One of the most noticeable patterns from Figure 1.5 is how the rate of UHMs increases when any LCS are introduced under all placement strategies and sensor measurement error types. This is especially true for LCS placements based on CES Score and Pollution Score (for the population overall and CBGs with high % nonwhite): the UHM rates stay ba-

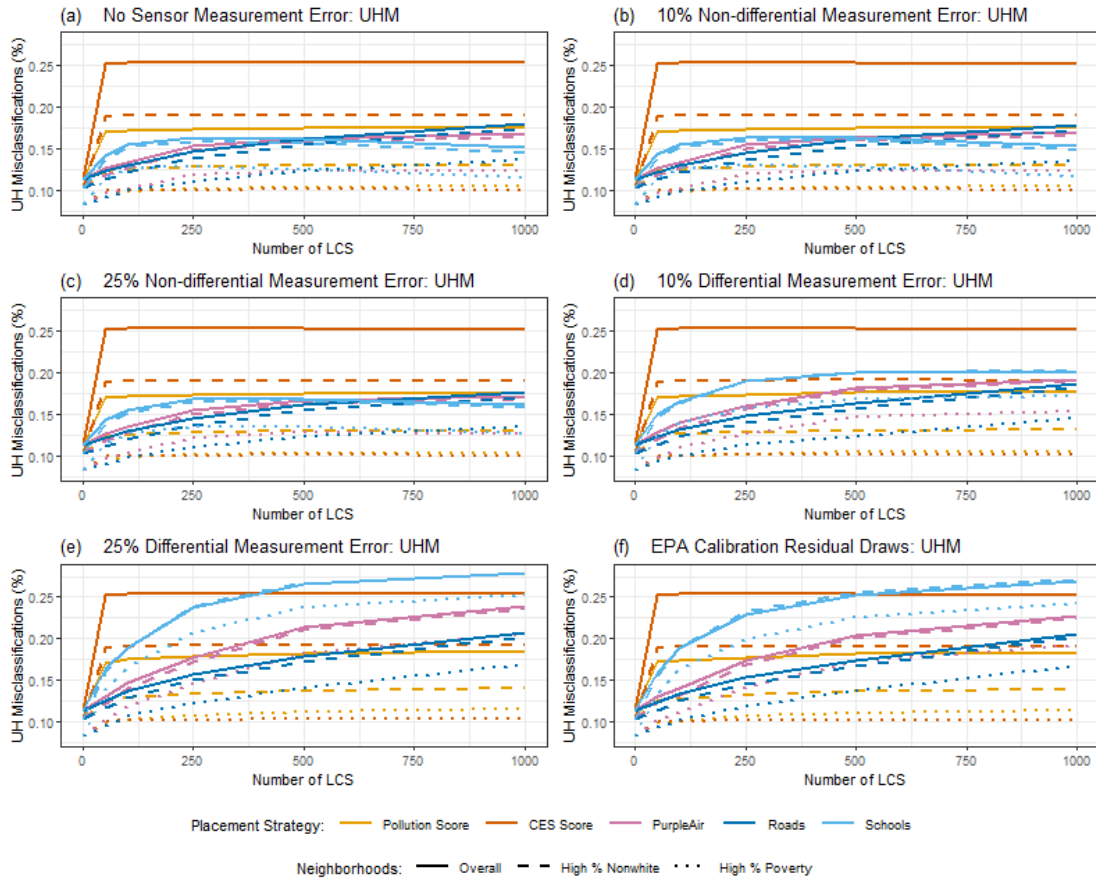


Figure 1.5: Rates of Unhealthy-Healthy Misclassifications (UHMs) resulting from different numbers of LCS deployed, LCS placement strategies, and sensor measurement error types and amounts. UH misclassification occurs when the AQ is unhealthy (Orange+) but is reported as healthy (Green or Yellow); the UHM rate is calculated by dividing the fraction of UH misclassifications by the total fraction of unhealthy days experienced by each group. All results were calculated using 366 days and averaged across 100 simulation replicates, weighted by population density. Panel a shows UHM rate when there is no LCS device measurement error, panel b shows UHM rate when LCS device measurement error is 10% non-differential, panel c shows UHM rate when LCS device measurement error is 25% non-differential, panel d shows UHM rate when LCS device measurement error is 10% differential, panel e shows UHM rate when LCS device measurement error is 25% differential, and panel f shows UHM rate when LCS device measurement error is sampled from the empirical distribution of residuals from PurpleAir LCS measurements (calibrated with the EPA equation) compared with collocated EPA monitor measurements.

sically constant despite the changing number of LCS. We posit that this is due to high local variability in AQ in Census tracts with high Pollution or CES Scores.

Another important observation is that although marginalized subgroups experience a higher num-

ber of unhealthy days in absolute terms, the fraction of those unhealthy days misclassified as healthy is lower than for the population overall. And, crucially, the CES and Pollution Score-based placements lead to the lowest UHM rates for CBGs with high % poverty. However, school locations are the only placement strategy resulting in decreasing UHM rates for large numbers of LCS deployed (when sensor measurement error is nondifferential, in panels a through c). A similar reversal is observed in panel c of Figure 1.4 for marginalized subgroups. To investigate whether the school placement strategy might produce the lowest rate of UHMs for many LCS deployed, we ran simulations with LCS at all the schools ($n=7,548$). The results are summarized in Table A.6. Notably, with 10% nondifferential sensor measurement error, this produces lower UHM rates than the EPA monitors alone. With 25% nondifferential sensor measurement error, the UHM rates are only slightly higher (e.g., 11.8% vs 11.4% for the population overall).

One last note on Figure 1.5 is that although generally increasing sensor measurement error increases the rate of UHMs, the empirical residual-based errors (panel f) result in slightly lower UHM rates than the 25% differential scenario (in panel e) for LCS at schools and PurpleAir locations.

Further insight can be gained from Figure A.4, which illustrates the contributions of distance to nearest instrument, local variability in AQ, and sensor measurement error to large misclassifications of the AQI (off by more than one class) for each of the LCS placement strategies.

1.4 DISCUSSION

In this study, we investigated the utility of including measurements from LCS in real-time AQ reporting using simulations based closely on real data. By comparing different types and amounts of sensor measurement error as well as different LCS placement strategies and numbers of LCS deployed, we were able to differentiate between the impacts of three components of error in daily AQ reporting: distance to the nearest AQ instrument, local variability in AQ, and sensor measurement error.

Our findings offer several key insights and suggestions. One of the most important is that *the value of using LCS for real-time AQ reporting depends strongly on the amount and type of sensor measurement error, and also on the metric to which people pay attention* (i.e., absolute concentration vs. AQI classification). Considering MAE, deploying LCS assuming 10% measurement error (either differential or non-differential) improves daily AQ reporting (compared to only using EPA monitors) across the board. However, among the placement strategies considered here, only LCS placements at schools and PurpleAir locations improve AQ reporting when sensor measurement error is assumed to be 25% and non-differential. Deploying LCS assuming 25% differential measurement error worsens daily AQ reporting across the board.

By contrast, introducing any number of LCS with any amount of sensor measurement error tends to increase the rate of UH misclassifications, which may be even more relevant than absolute PM_{2.5} concentrations for public health (whereas metrics like MAE may be more relevant for AQ-health science). This indicates that *the existing EPA monitor network in California is relatively good at reporting whether the AQ is healthy or unhealthy*. However, the UHM rate begins decreasing as more than 500 LCS are deployed at schools when sensor measurement error is assumed to be 10% or 25% and non-differential. If sensor measurement error is 10% nondifferential, then deploying LCS at all schools in California (n=7,548) produces a lower rate of UHMs than EPA monitors alone (it is about the same if sensor measurement error is 25% nondifferential). Of course, organizations deploying LCS will need to balance these considerations with their budgets for purchasing and maintaining LCS.

Accounting for both absolute concentration and AQI classification metrics, *it appears that placing LCS at schools results in the most accurate and equitable distribution of daily AQ information* when sensor measurement error is assumed to be less than 25% and nondifferential and more than 500 LCS are deployed. The latter is quite realistic given that in California in 2021, there were 4,343 1 km x 1 km grids with PurpleAir, not to mention other brands of LCS. From a health standpoint, children's relatively high vulnerability to air pollution⁵¹ further motivates the strategy of placing LCS at schools.

For our empirically-based simulation, the degree of error injected into our simulated LCS measurements was drawn from the empirical distribution of errors between collocated California EPA monitor and PurpleAir sensor data after correction using a national equation developed by EPA researchers¹⁰. This degree of measurement error is believed to most closely reflect error in the publicly available LCS measurements available through many programs and platforms today. Our simulations show that, under these conditions, using data from LCS in real-time AQ reporting would worsen the accuracy of information from people's nearest AQ instrument, both for the overall population and for the marginalized groups considered. This result suggests that *more region-specific LCS calibration procedures may be necessary for this application of LCS*, which aligns with the findings of several recent studies advocating for region-specific corrections^{176,21}.

When sensor measurement error is assumed to be high, LCS placement strategies that prioritize those burdened by environmental pollution and sociodemographic injustice result in the most equitable provision of AQ information. Also, across all levels of sensor measurement error and numbers of LCS deployed, selecting LCS placements based on CES Score results in the lowest rate of UHMs for CBGs with high % poverty. Lastly, our simulations revealed that AQ information accuracy under the CES and Pollution Score-based placements is often affected by high local variability in air quality, which makes sense because these locations tend to be near major sources of air pollution. This indicates that *while placing LCS in environmental justice hotspots may benefit those in the immediate community, integrating their data into wider AQ reporting platforms may lead to worsening of real-time AQ information for those outside the immediate community*.

Balancing policy priorities related to LCS deployment will not be easy. To balance the needs of people in less densely populated communities with those of people in more densely populated communities, our analysis unweighted by population density (results shown in Appendix A) suggests that *some strategic deployments near major roads (especially in less densely populated areas) might also be beneficial*.

These results can inform future investment in LCS networks for equitable AQ monitoring programs in the US, and the methods used can inform similar studies in other locales. While our findings are based on PurpleAir sensors for PM_{2.5}, this work informs accuracy targets and larger concerns about LCS placement and calibration across brands of PM_{2.5} sensors, and possibly for other air pollutants. However, it is important to note that this analysis has focused on the use of LCS for provision of real-time AQ information to the public. *These insights do not necessarily transfer to other applications of LCS.* For example, when used for research purposes, LCS have been shown to help capture neighborhood-scale PM_{2.5} when fused with satellite data⁷⁴ or incorporated into either a kriging model^{96,13,60} or a machine learning model with spatially-varying correction for the LCS¹⁴. That said, *the simulation methodology developed for this study could be adapted for many other research questions.*

1.4.1 LIMITATIONS AND FUTURE DIRECTIONS

Design-related limitations of our study are that we only used data from California and that we assumed individuals view the AQ information from the instrument nearest to their location of residence as their personal exposure. A technical limitation of this analysis is our use of the Di et al. daily 1 km x 1 km estimates for the “true” exposures. AQ may vary substantially over 24 hours and over a 1 km x 1 km square, and as several studies have observed, the PM_{2.5} patterns sensed by LCS are often different (i.e. affected more by local sources) than EPA monitors^{160,137}, which Di et al. used to train their model. We also did not consider any differences in LCS performance due to varying PM composition, which have been observed elsewhere¹⁶⁸.

Future work might consider more nuance in the LCS measurement error problem, such as accounting for sensor “drift”^{25,31,98} and varying particle composition / meteorological conditions, as has been explored by EPA researchers who have proposed a different correction method for wildfire smoke measured by PurpleAir sensors⁷⁹. This might be addressed by considering more spatial and temporal correlation in the LCS measurement errors. In terms of LCS placement, while we chose rel-

atively simple selection strategies to facilitate comprehension and comparison, future research could harness more sophisticated statistical and/or atmospheric modeling techniques^{13,82,38} to identify locations yielding key spatiotemporal information or to prioritize some LCS placements near EPA monitors for the purposes of sensor calibration. Finally, there is more work to be done investigating how access to real-time AQ information and/or alerts translates into public health and economic benefits, as several studies have begun exploring^{39,128,166,111,17}.

DATA AND MATERIALS AVAILABILITY

Code used to download and process the datasets, run the analyses, and generate the figures and tables can be found at https://github.com/EllenConsidine/LCS_placement_sims. All data used in this study are publicly available and our analytic (processed) dataset is on Harvard Dataverse, at <https://doi.org/10.7910/DVN/QR4N7V>. Descriptions of the data sources are in the Materials & Methods section as well as our GitHub README file.

Type/Amount of Sensor Measurement Error	Std. Dev. of Error ($\mu\text{g}/\text{m}^3$)	MAE ($\mu\text{g}/\text{m}^3$)	95th Pct. of Errors ($\mu\text{g}/\text{m}^3$)	Under-classified (%)	Over-classified (%)	UHM
Overall Population						
No LCS (only EPA monitors)	–	1.46	4.45	2.05	6.79	11.37
No Sensor Error	0	0.79	2.79	2.12	2.37	15.02
10% Non-differential	0.5	0.94	2.89	2.41	2.80	15.51
25% Non-differential	1.25	1.33	3.52	3.07	4.13	16.27
10% Differential	0.88	1.10	3.29	3.19	3.61	20.11
25% Differential	2.19	1.85	5.39	5.04	6.45	28.38
EPA Correction Residual	3.32	2.45	7.16	8.27	6.02	27.10
Decile Draws						
Population living in CBGs with high % nonwhite						
No LCS (only EPA monitors)	–	1.34	4.14	2.15	6.53	10.35
No Sensor Error	0	0.75	2.59	2.22	2.48	13.98
10% Non-differential	0.5	0.89	2.71	2.53	2.95	14.57
25% Non-differential	1.25	1.29	3.39	3.31	4.38	15.51
10% Differential	0.92	1.08	3.21	3.45	3.84	19.87
25% Differential	2.30	1.90	5.53	5.66	6.88	28.52
EPA Correction Residual	3.46	2.51	7.40	9.43	6.33	27.28
Decile Draws						
Population living in CBGs with high % poverty						
No LCS (only EPA monitors)	–	1.28	4.06	2.09	5.79	8.40
No Sensor Error	0	0.81	2.80	2.22	2.68	10.06
10% Non-differential	0.5	0.94	2.90	2.53	3.10	10.36
25% Non-differential	1.25	1.31	3.52	3.31	4.40	11.38
10% Differential	0.97	1.13	3.41	3.46	3.91	15.99
25% Differential	2.44	1.94	5.75	5.78	6.71	25.55
EPA Correction Residual	3.65	2.54	7.59	9.60	6.16	24.20
Decile Draws						

Table 1.1: Comparing Impacts of Different Sensor Accuracies. Results (weighted by population density) when there are no LCS vs. LCS at all real PurpleAir locations ($n = 4,343$), assuming different kinds and amounts of sensor measurement error (ME), averaged across 100 simulation replicates to account for randomness in the sensor measurement error generation. Unless otherwise specified, “errors” refer to the difference between the true exposure experienced at each grid centroid and the exposure reported from the nearest AQ instrument. AQI under-classification is when the true exposure class is greater than what someone is shown, and over-classification is when the true exposure class is less than what someone is shown. Rate of UH misclassification (UHM) is the fraction of days with unhealthy AQI (Orange+) that are misreported as healthy AQI (Green or Yellow).

2

Optimizing Heat Alert Issuance with Reinforcement Learning*

2.1 INTRODUCTION

EXTENSIVE EVIDENCE LINKS EXPOSURE TO EXTREME HEAT TO INCREASES IN MORBIDITY AND MORTALITY⁴⁰. Heat alerts are a practical and low-cost intervention to mitigate these effects⁴¹ by encouraging protective measures such as hydrating more, avoiding physical exertion outdoors, and opening cooling centers. However, studies investigating the effectiveness of heat alerts have observed

*This work has been published. <https://ojs.aaai.org/index.php/AAAI/article/view/35009> The full citation is: Considine EM, Nethery RC, Wellenius GA, Dominici F, and Tec M. (2025). Optimizing Heat Alert Issuance with Reinforcement Learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(27) [AI for Social Impact Track], 27922-27931. <https://doi.org/10.1609/aaai.v39i27.35009>

mixed results^{165,171}. Developing methods to optimize the issuance of heat alerts for public health is an open problem.

Various challenges stand in the way of solving this problem. First, issuing too many heat alerts may lead to alert fatigue¹⁰⁷ and health-protective resource depletion on both individual and community/institutional levels. Sequential decision-making (SDM) methods offer a promising yet unexplored avenue for tackling this issue. Second, despite representing a significant health threat on the population scale, local health impacts of heat—and therefore heat alerts—are small and easily confounded in observational data sets¹⁶⁴. Rare events and low signal (small effects) have been shown to challenge algorithmic decision-making^{52,126}. Third, mainstream SDM methods are not suitable for spatially heterogeneous settings in which a single policy is not equally effective in all regions or dynamically changing contexts¹¹⁷. However, attempting to identify independent policies for each location drastically reduces the amount of available historical data.

In this paper, we lay the foundation for addressing these challenges by introducing a framework for optimizing heat alert issuance using reinforcement learning (RL). RL allows learning SDM policies for determining when to issue heat alerts, aiming to minimize the population health risk as a negative reward signal. The ultimate vision for this vein of research is deploying data-driven policies to alert the public about extreme heat events (and, eventually, other environmental exposures such as extreme cold or wildfire smoke). In addition to our analysis breaking ground towards this goal, a variety of follow-up work is facilitated by the publication of our SDM environment in an open-source Python package, `weather2alert`[†], compatible with the Gymnasium framework for RL¹⁴⁵. This challenging, data-driven environment can be used as a benchmark. Further, our methodology to create this environment can be used as a blueprint for other RL simulators, particularly for problems with exogenous state space components and multiple locations or individuals.

The two main components of our heat alerts RL framework are visualized in Figure 2.1 and sum-

[†]<https://github.com/NSAPH-Projects/weather2alert>

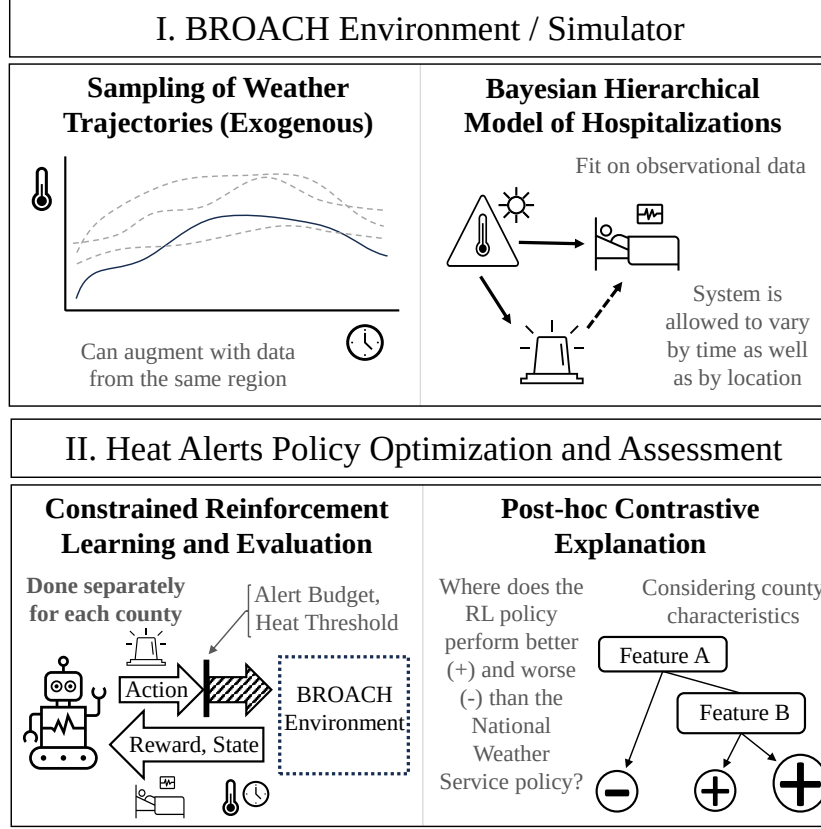


Figure 2.1: Overview of the heat alerts RL framework.

marized below. First, we create a realistically challenging SDM environment for heat alert issuance, structured as a Markov Decision Process (MDP) under a budget constraint. Key components of the dataset used to create this environment are heat alerts issued by the U.S. National Weather Service (NWS), hospitalization records from Medicare, ambient heat index, and other covariates informed by the epidemiological literature. To estimate the rewards under both observed (NWS) and counterfactual alert policies, we fit a statistical model of hospitalizations that allows for spatiotemporal heterogeneity, uncertainty quantification, and fast inference using variational Bayes¹⁴⁷. For the state-transition model, we exploit a factorization of the state space into an endogenous component with known transition dynamics and an exogenous component sourced from observed weather trajec-

ries. This use of real data allows us to make grounded inferences about the effectiveness of both observed and counterfactual heat alert policies in the U.S. We call this environment structure Bayesian Rewards Over Actual Climate History (BROACH).

Second, we use our environment to train RL models and then to evaluate the RL policies as well as the NWS and alternative baseline policies—all subject to the same alert budget. Our evaluation identifies issues during training that lead to poor performance of standard, widely-used RL algorithms. We address these issues by introducing conceptually simple modifications which enable learning of policies that reduce hospitalizations relative to the NWS policy. These modifications include restricting the RL model to issue alerts only on extremely hot days and optimizing policies for each location separately. Lastly, we perform contrastive explanation^{157,108} of the RL policies by comparing their attributes (e.g., how often does the policy issue alerts on consecutive days) within and across locations to those of the NWS and alternative baseline policies. Specifically, we use visualizations and Classification and Regression Trees (CART) to illustrate systematic differences between these policies and identify characteristics of locations where RL-based policies might offer the greatest improvements in public health.

2.2 RELATED WORK

HEAT ALERT OPTIMIZATION

Recent approaches to improving the issuance of heat alerts include (i) the development of a causal inference technique for stochastic interventions to infer whether increasing the probability of issuing a heat alert would be beneficial¹⁷¹ and (ii) a comparison of methods to identify optimal thresholds above which heat alerts should always be issued¹⁰⁰. Crucially, neither of these approaches addresses the complications of sequential dependence, i.e., the potential for alert fatigue and running out of resources to deploy precautionary measures. To our knowledge, SDM techniques have not yet been

explored in a climate and health setting.

RL WITH EXOGENOUS STATES

Several recent papers utilize a decomposition of the state space into exogenous vs. endogenous components, but none of these methods apply directly in our setting. Efroni et al.⁴³ provide theoretical analysis and guarantees for sample efficiency in tabular (finite state) environments with an unknown subset of exogenous variables. Sinclair et al.¹³³ introduce “hindsight learning” for situations with a known reward function and an exact solver under a fixed exogenous trajectory. Lee et al.⁸⁸ also use hindsight, but do so in the context of identifying latent states within a partially observable MDP. Contrary to Levine et al.⁹⁰ and Efroni et al.⁴⁴, we do not consider exogenous states as distractors since weather patterns are crucial for decision-making.

STATISTICAL MODELING FOR RL ENVIRONMENTS

Using a statistical model for the reward function in an RL environment is common practice in settings with empirical data. Bayesian models in particular have been used to identify the effect of intervention in health-related applications and dynamic treatment regimes^{93,139,174}. For spatiotemporal or otherwise heterogeneous environments, previous studies have also used a combination of local modeling and global modeling that takes into account the spatial or hierarchical structure to create simulated environments for RL^{169,92,1}.

CONSTRAINED LEARNING

The use of domain knowledge-based policy restrictions in RL has been found to speed up learning and to guide the learned policy in pragmatic directions¹⁰⁶. This supports our restriction of RL heat alerts to very hot days. More generally, incorporating “cost” or “safety” constraints is of interest for

many RL applications^{87,19}. A common approach in the constrained RL literature is to use Lagrange multipliers^{61,125}, however, these methods do not enforce a strict constraint/budget. To enable direct comparison with the observed NWS policy, we impose a strict alert budget, which can be viewed as a simplified variation of the indicator approach formulated by Xu et al.¹⁷². An alternative approach to handling both the alert budget and restriction of alerts to very hot days would be to model the extent of distributional overlap with the observed NWS policy, and penalize actions that have too little overlap¹⁷². However, our approach has the benefit of being relatively interpretable.

CONTRASTIVE POLICY EXPLANATIONS

There is a large literature on post-hoc explainability in RL⁷². Contrastive analysis has been used to explain differences between RL policies and more familiar / intuitive policies^{157,108}. We note that while CART has been used to try and mimic RL policies¹²¹, our use of it to analyze differences between policies has not been previously documented.

2.3 PROBLEM SETUP

2.3.1 RL PRELIMINARIES

RL methods typically operate within the framework of Markov Decision Processes or MDPs¹³⁵. A finite-horizon or episodic MDP with horizon (episode length) H is defined as a tuple $\mathcal{M} = \langle S, \mathcal{A}, R, P, d_0, \gamma \rangle$, where S is the set of states, $\mathcal{A}$ is the set of actions, $R : S \times \mathcal{A} \rightarrow \mathbb{R}$ is the expected reward function, $P : S \times \mathcal{A} \rightarrow \Delta(S)$ is the transition function[‡], $d_0 \in \Delta(S)$ is the initial state distribution, and $0 < \gamma \leq 1$ is the time discount factor[§]. A (non-stationary) policy is defined as a collection of decision rules $\pi = \{\pi_t : S \rightarrow \Delta(\mathcal{A})\}_{t=0}^{H-1}$, mapping from states to probability distributions over actions at

[‡] $\Delta(\cdot)$ is the set of probability distributions over the set “ $\cdot$ ”.

[§]Note that in our analysis, $\gamma = 1$ works best, so in practice there is no discounting.

each time step. The state value function is $V_t^\pi(s) := \mathbb{E}_\pi[\sum_{b=t}^{H-1} \gamma^{b-t} R(s_b, a_b) | s_t = s]$ and the state-action value function is $Q_t^\pi(s, a) := R(s, a) + \gamma \mathbb{E}_{s' \sim P(s, a)}[V_t^\pi(s')]$. The policy optimization goal is to identify π^* that maximizes the expected cumulative reward

$$J(\pi) := \mathbb{E}_{s_0 \sim d_0}[V_0^\pi(s_0)]. \quad (2.1)$$

Under regularity conditions¹²², the optimal policy is deterministic and the unique solution (up to ties) of the Bellman optimality equation

$$Q_t^*(s, a) = R(s, a) + \gamma \mathbb{E}_{s' \sim P(s, a)}[\max_{a'} Q_{t+1}^*(s', a')], \quad (2.2)$$

where Q_t^* denotes the state-action value function of π_t^* . The optimal action satisfies $\pi_t^*(s) = \arg \max_a Q_t^*(s, a)$.

RL algorithms search for the optimal policy of an MDP by interacting with an environment that generates rewards and transitions, where the RL algorithm does not need knowledge of the transition function¹³⁵.

2.3.2 ISSUING HEAT ALERTS AS A CONSTRAINED MDP

The heat alerts MDP components are summarized here. See also Table B.1 of Appendix B.1. Each MDP $\mathcal{M}_k$ corresponds to a geographic area $k \in \mathcal{K}$, and episodes are indexed as $j \in \mathcal{J}$. In this work, geographic areas are U.S. counties, and each episode spans the warm season / summer from May 1st to September 30th of a specific year ($H = 152$ days).

The action a_t on day t is either 1 (issue a heat alert) or 0 (do not issue a heat alert). To mitigate alert fatigue, we adopt a strict action budget-constrained approach, optimizing the policy subject to $\sum_{t=0}^{H-1} a_t \leq b$.

The reward $r_t = R(s_t, a_t)$ is the expected rate of heat-related hospitalizations at time t , transformed

such that fewer hospitalizations correspond to a greater reward. Specifically, if $\rho_t(a_t) := \rho(s_t, a_t)$ represents the *per capita* rate of heat-related hospitalizations on day t if action $a_t \in \{0, 1\}$ is taken at state s_t , then $r_t \propto -\rho_t(a_t) + \text{const.}$

The state vector s_t contains the factors underlying the effect of heat on hospitalizations and the effectiveness of heat alerts at reducing hospitalizations. It can be decomposed as $s_t = (\xi_t, x_t)$, where ξ_t is the exogenous component and x_t is the endogenous component. The exogenous component is defined as the portion of the state space that is not influenced by the agent’s actions a_t , such as the weather. The transition function of the endogenous component (heat alert history) is known and deterministic. Hence, the full state transition function admits a factorization

$$\begin{aligned} P(s_{t+1}|s_t, a_t) &= P((\xi_{t+1}, x_{t+1}) | (\xi_t, x_t), a_t) \\ &= P_\xi(\xi_{t+1}|\xi_t)P_x(x_{t+1}|x_t, a_t) \end{aligned} \tag{2.3}$$

2.4 BROACH: AN RL ENVIRONMENT FOR OPTIMIZING HEAT ALERT ISSUANCE

In creating an interactive heat alert environment, our design objectives are (i) allow use of general-purpose RL algorithms, (ii) be grounded in real data, and (iii) handle the low signal and heterogeneity of health impacts of heat and heat alerts across space and time. To meet these objectives, we introduce a methodology termed Bayesian Rewards Over Actual Climate History (BROACH). BROACH can be generalized to other climate-related events and interventions.

2.4.1 DATA SOURCES

Here we provide an overview of the data used in the RL environment; Appendix B.2 contains more details.

HEAT ALERTS AND HEAT INDEX

We use the dataset of Weinberger et al.¹⁶⁴ with daily, county-level records (2006-2016, May-September) of heat alerts issued by the NWS, as well as heat index, which is a measure of the combined effect of temperature and humidity. To understand the observed heat alert data, note that while the decision to issue an alert is based on temperature thresholds, it is also strongly affected by the discretion of the local office⁶⁶. Analysis by Hondula et al.⁷³ suggests that spatial variability in the current NWS/local office approach is not well aligned with the health risk from heat.

HEAT-RELATED HOSPITALIZATIONS

We merged the heat alert and weather data with daily, county-level hospitalization counts from Medicare. We included cause-specific hospitalizations previously found to be associated with extreme heat in the Medicare population: septicemia, peripheral vascular disease, urinary tract infections, and diabetes mellitus with complications¹⁵. We excluded heat stroke and fluid and electrolyte disorders (which were also found to be associated with extreme heat in Medicare) because Weinberger et al.¹⁶⁴ observed a positive association between heat alerts and hospitalizations with these causes. They hypothesized it was due to increased awareness of heat-related symptoms and seeking medical care. Figure 2.2 illustrates this scenario as an unobserved mediation problem¹²⁰, which affects the identification of the preventive effect of heat alerts. We refer to the remaining causes, which we pool together into one outcome variable, as not-obviously heat-related (NOHR) hospitalizations.

COUNTIES

We consider the 761 U.S. counties with a population greater than 65,000 to avoid locations with very few hospitalizations and to focus on the most populous areas. Figure 2.3 shows the counties considered, which account for approximately 75% of the population and 25% of the number of counties.

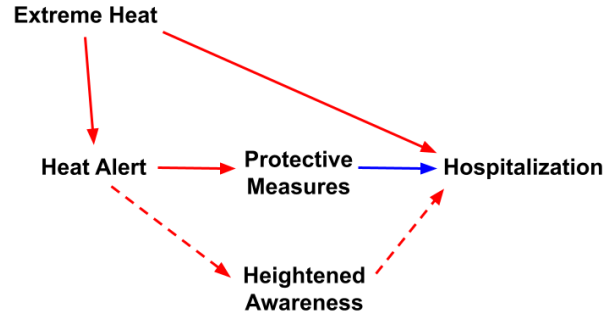


Figure 2.2: Unobserved mediation of the effect of heat alerts on hospitalizations by heightened awareness of heat-related symptoms and seeking medical care.

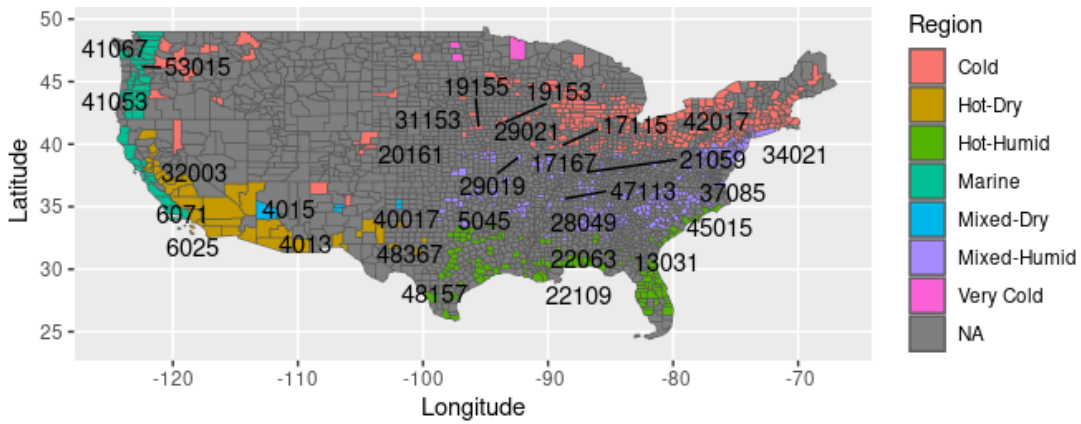


Figure 2.3: Map of the counties considered and their regional climate zone classifications. All the colored counties are used in the Bayesian rewards model and RL environment; the 30 counties with annotated FIPS codes are used in the RL experiments.

COUNTY-LEVEL CHARACTERISTICS

Spatial heterogeneity in the health effects of heat and heat alerts can occur due to people in different regions being affected differently by the same absolute temperatures due to geographical self-selection and climate adaptation, differential health susceptibility and/or individual agency associated with socioeconomic status and population density, and variable response to heat alerts due to local policy and political ideology^{112,175,48,29}. To characterize this spatial heterogeneity, we compiled a set of county-level covariates: population density and median household income¹⁵¹, regional classifications of cli-

mate zones¹⁵³, broadband (internet) usage data¹⁰², presidential election returns¹⁰⁴, and fine particulate matter (PM_{2.5})⁶³. For the latter, there is strong evidence of adverse synergistic health impacts of air pollution and heat⁵.

2.4.2 THE STATE SPACE

In addition to alert fatigue, time-varying factors that can modify the effect of heat on hospitalizations and the effectiveness of heat alerts include relative excess heat, for example, compared to recent days⁶⁹, and the timing of heat waves within the warm season⁴.

Table 2.1 summarizes the variables used in the state space of the MDP. Recall that the state space can be factorized as $s_t = (\xi_t, x_t)$, where ξ_t is the exogenous component and x_t is the endogenous component. The exogenous component contains the quantile value of the day’s observed heat index (QHI)—within each county across warm seasons for all the years—which allows us to account for long-term climate adaptation. It also contains time-varying factors which can modify the heat-health relationship^{69,4}: day of summer (which is equivalent to t), weekend status, and excess heat compared to the last three days. The endogenous component contains the number of heat alerts issued in the last 14 days, an indicator of whether a heat alert was issued yesterday, and the remaining alert budget for that episode.

Component	Variable
exogenous (ξ_t)	quantile of heat index (QHI), excess heat, day of summer, weekend indicator
endogenous (x_t)	Number of heat alerts issued in the last 14 days, indicator of whether a heat alert was issued yesterday, and the remaining alert budget

Table 2.1: The state space, decomposed into exogenous and endogenous components.

2.4.3 THE REWARD FUNCTION: A HIERARCHICAL MODEL OF HOSPITALIZATIONS

To learn the expected hospitalization rate, and thereby the transformed reward r_t , we train a Bayesian hierarchical model with the desiderata of (a) facilitating obtainment of confidence intervals when evaluating RL policies and (b) allowing for spatial heterogeneity in the health effects of heat and heat alerts. Let $\rho_t^{(k,j)}(a)$ be the expected per capita NOHR hospitalization rate at time t in county-summer (k, j) when taking action a . We propose the scaled reward function $r_t^{(k,j)}(a) := C_1 - C_2 \rho_t^{(k,j)}(a)$ where C_1 and C_2 are positive scaling constants. (Note that choices of C_1 and C_2 do not change the optimal policy, but can help stabilize learning.) We choose $C_1 = 1$ and C_2 as the reciprocal of the observed hospitalization rate through the entire summer.

Let $y_t^{(k,j)}(a)$ denote the counterfactual outcome representing the number of NOHR hospitalizations for action $a \in \{0, 1\}$, and $n^{(k,j)}$ denote the total population susceptible to hospitalization in county k in episode j . We introduce two functions—to be learned—governing the counterfactual hospitalization rate in county k as a function of $s_t^{(k,j)}$. First, $\lambda_k: S \rightarrow \mathbb{R}^+$ denotes the baseline rate of hospitalizations when no alert is issued. Second, $\tau_k: S \rightarrow (0, 1)$ expresses the effectiveness of issuing an alert as a multiplicative reduction factor. The top-level hospitalization model is

$$\begin{aligned} y_t^{(k,j)}(a) &\sim \text{Poisson}(n^{(k,j)} \rho_t^{(k,j)}(a)), \\ \rho_t^{(k,j)}(a) &:= \lambda_k(s_t^{(k,j)})(1 - a \cdot \tau_k(s_t^{(k,j)})), \end{aligned} \tag{2.4}$$

for all $(k, j) \in \mathcal{K} \times \mathcal{J}$ and $t \in \{0, \dots, H-1\}$. The Poisson distribution/loss function is the natural choice for count data and recommended for health outcomes data due to its correspondence to the Cox survival model⁸.

To estimate the functions λ_k and τ_k under limited data and low health signal, we incorporate a data-driven random effects prior (using the county-level features) and inject domain knowledge on the signs of certain coefficients. We train this model using variational inference, using the Python

package Pyro⁶⁵. The details of this approach, as well as model diagnostics, are in Appendix B.3.

Note that the fitted coefficients in the alert effectiveness component indicate the presence of alert fatigue. To utilize this rewards model within the BROACH environment, we draw from the variational posterior distribution at the start of each episode, and then use those coefficients to calculate $r_t^{(k,j)}$ given $(s_t^{(k,j)}, a_t^{(k,j)})$ for $t = 0, \dots, H - 1$.

2.4.4 THE ALERT BUDGET

Each episode in our MDP starts with a fixed alert budget b . This can be specified as a constant, sampled from a range of values, or set to the observed (NWS) value. As the agent steps through the episode, once $\sum_{b=0}^t a_b = b$, it is prohibited from issuing any more alerts: it can only obtain the reward $r_b^{(k,j)}(0)$ for days $b = t + 1, \dots, H - 1$.

2.4.5 THE TRANSITION FUNCTION: OBSERVING ACTUAL WEATHER/CLIMATE HISTORY

We leverage the decomposition equation 2.3 to use real data for the exogenous trajectories, and introduce a data augmentation scheme in the context of single-county RL.

SAMPLING OF EXOGENOUS TRAJECTORIES

Recall that the endogenous aspect of our state space, heat alert history, is deterministically updating, so does not need to be modeled probabilistically. The key observation in the heat alert setting is that the more complicated aspect of the environment, the weather/climate, also does not need to be modeled because it is completely exogenous to heat alert decision-making. In this setting, developing a model for the weather is not only unnecessary, but would unavoidably introduce error. Instead, we sample real observed weather trajectories.

REGIONAL DATA AUGMENTATION

The issue with this approach is that there are only 11 years of exogenous trajectories (2006-2016) available per county during which we also have heat alert data—which is needed when we use observed alert budgets (paired with observed weather) for comparison with the NWS. To mitigate the potential for RL overfitting to these 11 years, we propose data augmentation by sampling exogenous trajectories from other counties in the same regional climate zone[¶]¹⁵³ during RL training and validation / tuning.

2.5 LEARNING & EVALUATION

We conduct experiments to (a) assess the ability of standard RL algorithms to learn effective heat alert policies in the BROACH environment, (b) test various modifications to these algorithms, and (c) compare the resulting policies to the NWS policy and other intuitive baselines. We then investigate heterogeneous performance of the RL policies to provide domain-relevant insights. Code for all steps is available in our GitHub repository^{||}.

2.5.1 POLICY CONSTRAINT: VERY HOT DAYS

Foreshadowing our findings, standard RL algorithms struggle to learn to conserve their alert budgets for later in the summer, resulting in poor performance. A major modification we implement to encourage more effective behavior is restricting the issuance of heat alerts to days above a QHI threshold, optimized separately for each county. To optimize the QHI threshold, we test the sequence of values between 0.5 and 0.9 (ensuring overlap with the NWS policy), by every 0.05, and select the value that yields the best return on our validation set (specified in the following section). Names of models including this adaptation have the suffix “.QHI”.

[¶]Due to the large size/geographic range of the “Cold” zone, we grouped its counties into eastern and western subsets.

^{||}https://github.com/NSAPH-Projects/heat-alerts_RL

2.5.2 EXPERIMENTAL SETUP

For all training and evaluations, to directly compare counterfactual alert policies with the NWS policy, we fix b in each county-summer to the observed number of heat alerts.

RL BASELINES

We investigate using four common RL algorithms: Deep Q-learning (DQN), Quantile Regression Deep Q-learning (QRDQN), Trust Region Policy Optimization (TRPO), and Advantage Actor-Critic (A2C)—additional information on these algorithms is in Appendix B.4.2. We use standard implementations of these methods available in the `Stable-Baselines3` Python library¹²³. To ground our analysis and discussion, note that DQN and QRDQN are off-policy methods which learn deterministic policies by directly solving for the optimal value function using the Bellman optimality condition in equation 2.2; exploration is induced only during training (e.g. epsilon-greedy). Whereas, TRPO and A2C are on-policy methods which learn stochastic policies by direct optimization of the expected return in equation 2.1, using refinements of the policy gradient theorem¹³⁶; exploration is inherent.

NWS AND SIMPLE ALTERNATIVE BASELINES

To evaluate each of the RL policies, we compare them to the observed NWS policy as well as several counterfactual baselines: most simplistically, randomly selecting b days on which to issue alerts (RANDOM) and selecting the b days with the highest QHI that summer (TOPK)—the latter is an oracle policy (not implementable in the real world) because to use it, we would have to know the whole summer’s daily QHI in advance. We also implement the general guidelines that NWS recommends in the absence of local criteria for issuing heat alerts (BASIC.NWS): alert if the heat index is $\geq 100^\circ\text{F}$ in northern states and $\geq 105^\circ\text{F}$ in southern states⁶⁶. Lastly, we implement a policy of always issuing alerts on days above a per-county optimized QHI threshold until the budget runs out (AA.QHI).

TRAINING DATA

The counties for which NWS issued few heat alerts had high-variance estimates of heat alert effectiveness in the rewards model. Therefore, for the RL experiments, we selected a subset of 30 counties in which at least 75 heat alerts were issued during 2006-2016; spread across the five major climate regions, prioritizing those with higher variance in estimated alert effectiveness across days. See map in Figure 2.3 and Appendix B.4.1 for more details.

EVALUATION METRIC

The main metric we use for policy comparison is the average cumulative reward per episode (“average return”). To estimate this metric fairly, we hold three years of data (2007, 2011, and 2015) out of RL training, referred to as the evaluation years^{**}. For the purposes of RL hyperparameter tuning (details in Appendix B.4.3) and selection of the optimal QHI threshold for each county, we consider the regionally augmented weather trajectories and associated heat alert budgets from the evaluation years *excluding the county of interest* to be the validation data set. The final evaluation results (reported in tables and figures) are obtained by drawing 1,000 sets of coefficients from the Bayesian rewards model posterior and calculating the return under each policy with weather and budgets *only from the county of interest* during the evaluation years. These validation/tuning and evaluation procedures are described algorithmically in Appendix B.5. After calculating the average returns for the 30 counties under each competing policy, we compare each policy’s returns with the associated returns under the NWS policy using a Wilcoxon-Mann-Whitney test. This nonparametric test allows us to compare the distribution of differences in returns, accounting for the fact that the counties are highly heterogeneous, so the differences between the policies are not normally distributed.

^{**} Appendix B.4.4 shows the results of a sensitivity analysis where the models are trained on 2006-2013 and evaluated on 2014-2016.

SENSITIVITY ANALYSIS

In the main analysis, we evaluate each county-specific implementation of DQN, QRDQN, TRPO, and A2C by allowing each algorithm to deploy the type of policy that it optimizes during training: deterministic for DQN and QRDQN and stochastic for TRPO and A2C. As a sensitivity analysis, we also deterministically evaluate the TRPO and A2C policies, selecting whichever action has a higher probability under the policy function. The names of these models have the prefix “DET”. In another sensitivity analysis, we investigate whether the RL models can learn to anticipate subseasonal variation in the warm season by augmenting the RL state space with information about the future. Specifically, we test whether inclusion of the change in heat index over each of the next 10 days, as well as the 50th-100th (by every 10) percentiles of QHI for the remainder of the summer, improves RL performance. In real-time deployments of RL models, this kind of future information would not be known but could be sourced from weather forecasts or climate model projections. Of course, such forecasts / model projections are not perfect, but we start with perfect future information as a proof of concept. The names of models that include future information have the suffix “.F”.

2.5.3 RESULTS AND DISCUSSION

Figure 2.4 illustrates how the different policies look in practice, specifically in a scenario where using RL is beneficial. Figures B.4 and B.5 show scenarios where the alert budget is small and where using RL is not beneficial compared to the alternatives.

Table 2.2 provides the main results of our RL experiments.

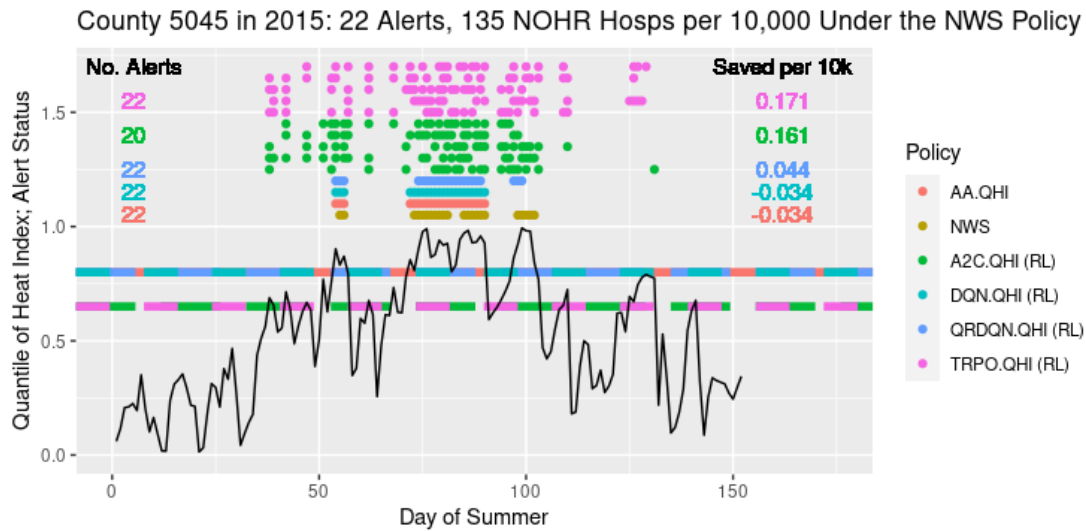


Figure 2.4: An example of observed and counterfactual heat alert policies for a single summer (2015 is the most recent year in our evaluation set), with estimates of the number of NOHR hospitalizations saved (compared to NWS) per 10,000 Medicare enrollees under each policy. The dashed lines indicate the optimized QHI threshold of the policy in the same color. The multiple horizontal lines of pink and green dots indicate five different samples from TRPO.QHI and A2C.QHI respectively. For these two policies, the number of NOHR hospitalizations saved is the average of all their evaluations over 2015.

QHI RESTRICTION IS NECESSARY FOR THE STANDARD RL ALGORITHMS TO PERFORM WELL

Several counterfactual policies perform worse than NWS (as indicated by negative median differences in returns): RANDOM, BASIC.NWS, and the RL models without the QHI restriction. By contrast, the RL models with the QHI restriction perform significantly better than NWS, as indicated by their significant positive median difference in returns. A rough estimate if A2C.QHI were implemented as-is across all counties in the U.S. (using the Medicare population from 2011, the midpoint of our study period) is that we could see a reduction of 222 NOHR hospitalizations per year (approximate 95% CI = (-491, 1131); details of this CI calculation are in Appendix B.6). Discussion of this absolute benefit to public health is in Appendix B.7.

Policy	Median Diff.	WMW	P-value	NOHR Hosps Saved vs NWS (vs Zero) Per Summer	
				Median per 10,000 Medicare	Approximate Total
*TOPK	0.022	406	0.0002	0.023 (1.96)	113 (9,603)
RANDOM	-0.015	177	0.88	-0.013 (1.98)	-64 (9,697)
BASIC.NWS	-0.286	30	1.0	-0.283 (1.45)	-1,388 (7,110)
AA.QHI	0.03	348	0.0025	0.029 (1.99)	143 (9,747)
DQN	-0.123	43	0.99995	-0.102 (1.83)	-502 (8,989)
QRDQN	-0.117	51	0.99991	-0.098 (1.85)	-480 (9,046)
TRPO	-0.065	97	0.99742	-0.063 (1.9)	-309 (9,321)
A2C	-0.063	100	0.99689	-0.062 (1.9)	-301 (9,317)
DQN.QHI	0.03	370	0.00242	0.029 (1.99)	142 (9,746)
QRDQN.QHI	0.035	344	0.01121	0.031 (1.99)	153 (9,767)
TRPO.QHI	0.038	338	0.0154	0.038 (2.02)	185 (9,897)
A2C.QHI	0.042	344	0.01121	0.045 (2.02)	222 (9,902)
DQN.F	-0.417	40	0.99996	-0.401 (0.49)	-1,966 (2,408)
QRDQN.F	-0.42	48	0.99993	-0.42 (0.47)	-2,059 (2,322)
TRPO.F	-0.062	103	0.99625	-0.061 (1.9)	-297 (9,322)
A2C.F	-0.063	101	0.99669	-0.062 (1.9)	-302 (9,304)
DQN.QHI.F	-1.895	1	1	-1.901 (0)	-9,317 (0)
QRDQN.QHI.F	0.014	238	0.45904	0.012 (1.61)	61 (7,876)
TRPO.QHI.F	0.046	345	0.01062	0.048 (2.02)	236 (9,886)
A2C.QHI.F	0.036	343	0.01183	0.038 (2.02)	187 (9,915)
DET.TRPO.QHI	-0.662	57	0.99985	-0.579 (0)	-2,835 (2)
DET.A2C.QHI	-0.886	90	0.99837	-0.794 (0.16)	-3,888 (807)
DET.TRPO.QHI.F	0.032	271	0.21723	0.034 (1.83)	168 (8,948)
DET.A2C.QHI.F	0.02	287	0.13335	0.021 (2.01)	104 (9,836)

Table 2.2: Comparison between the average return of each counterfactual policy and that of the NWS policy on the evaluation years, summarized across counties (e.g. “Median Diff.” is the median difference in average return). The last two columns also provide a comparison to the counterfactual with zero heat alerts issued. WMW is the Wilcoxon-Mann-Whitney statistic (higher is better); its associated p-value is also included. The first policy, marked by *, requires oracle knowledge of the future weather. The approximate total number of hospitalizations is based on 49 million Medicare enrollees in 2011²³.

THE FUTILITY OF INCLUDING FUTURE INFORMATION

Across both plain and QHI-restricted models, TRPO was the only one that benefited from including future information in the state space. While A2C was minimally impacted, QRDQN and DQN were made noticeably worse. The latter may be due to the future information dramatically increasing the size of the state space/neural network parameters, which our hyperparameter tuning did not fully counteract. Note that for the TRPO models, the benefit observed with the use of future information would likely be smaller in practice due to forecast / prediction error—the simulation of which is beyond the scope of this paper. Therefore, we focus on the RL models without future information in the remainder of this text.

THE NEED FOR STOCHASTIC POLICIES

Interpreting policies from the TRPO and A2C models deterministically fails in the heat alerts setting under our rewards model. Additional discussion of this result is in Appendix B.8.2. This finding raises the question: *Are stochastic policies palatable from a domain perspective?* On the one hand, it is less immediately satisfying that some aspect of heat alerts issuance would be left to chance. On the other hand, in the context of human-in-the-loop decision making—which would likely be the case for public organizations such as NWS¹³⁴—an algorithm reporting probabilities is more informative than reporting only a binary action. In any case, if in the future a heat alerts RL model was running continuously online, it would likely need to utilize exploration (incorporating some randomness into its actions) to update itself over time.

COMPARING POLICIES THAT PERFORM SIGNIFICANTLY BETTER THAN NWS

We note that the performance and behavior of DQN.QHI is practically identical to AA.QHI—not just in its numerical summary, but also in its policy behavior. QRDQN.QHI performs a bit better than

DQN.QHI, but not as well as the stochastic policies. Between those, A2C.QHI performs better than TRPO.QHI, though their policies tend to display similar characteristics. Therefore, we focus on A2C.QHI as the best RL policy in the post-hoc analysis. The non-RL counterfactual policies that perform well are that of always issuing an alert on days above an optimized QHI threshold (AA.QHI) and of issuing alerts on the b hottest days of the summer (TOPK). We note that the ordering of the Wilcoxon-Mann-Whitney statistics for TOPK, AA.QHI, and A2C.QHI is opposite the ordering of these policies' median difference in returns (compared to NWS). This is due to the longer tails of the A2C.QHI differences, as illustrated in Figure 2.5. For interpreting these results, recall that TOPK is an oracle policy. However, AA.QHI is an implementable alternative to A2C.QHI, so we must consider it more seriously.

2.5.4 POST-HOC CONTRASTIVE EXPLANATION

To characterize differences between the best RL policies, the NWS policy, and alternative baselines, we consider both stationary features included upstream in the analysis (e.g., climate region and median household income) as well as characteristics of both the NWS and counterfactual alert policies, namely the distributions of the days of summer on which alerts are issued and the lengths of alert streaks (sequences of repeated alerts). We start with these alert characteristics, descriptive histograms of which are in Figure B.6. Both A2C.QHI and AA.QHI tend to issue heat alerts earlier in the summer than NWS. A2C.QHI also tends to issue shorter streaks of alerts (i.e., fewer consecutive days of alerts), which makes sense given its ability to learn sequential dependence encoded by the rewards model.

Figure 2.5 illustrates the varying performance of TOPK, AA.QHI, and A2C.QHI relative to the NWS policy, across climate regions and counties. Note that large heterogeneity in the health effects of heat alert intervention across counties is consistent with previous findings¹⁷¹. A striking visual pattern is the increasing vertical spread from left to right: TOPK performs the least heterogeneously across counties, followed by AA.QHI, and finally A2C.QHI. Similarly, the humid climate regions display more heterogeneity across counties than the others, across all the policies. This is partially due to those

regions being larger and thus overrepresented among our 30 counties (see Table B.3), but that does not fully explain the discrepancy.

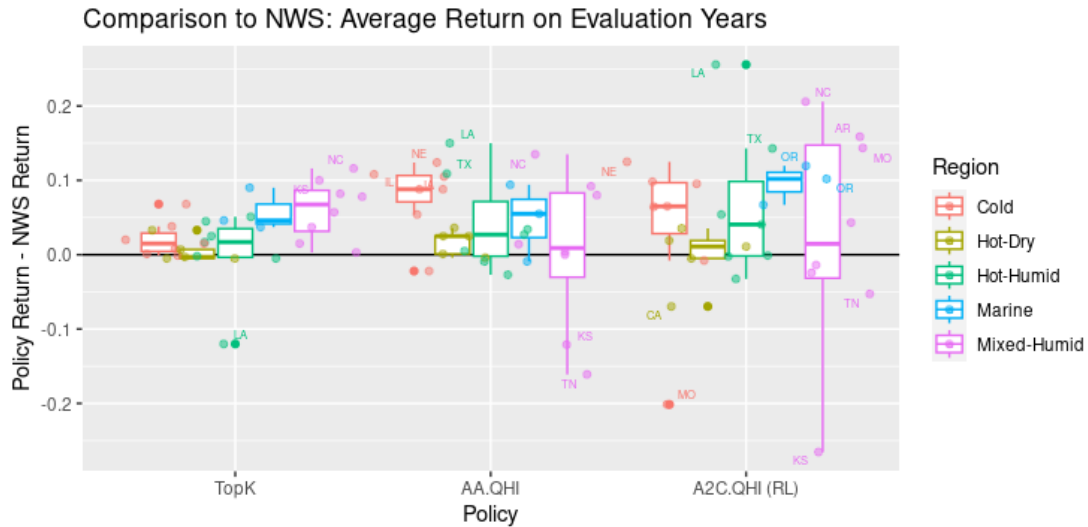


Figure 2.5: Different policies' average return relative to NWS across different climate regions, over the evaluation years. Outlier points are labeled with their state abbreviation. Note that TOPK is an oracle policy.

To structure our investigation of why some counties see smaller (or even no) benefits from the application of the heat alerts RL, we fit CART on both the numeric difference in average return (compared to NWS) and a polytomous indicator of which policy performed the best. To prevent CART overfitting to our 30 counties and to facilitate interpretation, we ensure that the minimum number of counties in each leaf (terminal) node is greater than or equal to five for the regression and three for the polytomous classification.

RL PERFORMS BEST IN COUNTIES WITH LARGER HEAT ALERT-HEALTH SIGNAL AND LONGER HEAT WAVES

A classification tree distinguishing A2C.QHI, AA.QHI, NWS, and TRPO.QHI is shown in Figure B.7. The most predictive feature of the counterfactual policies’ performance relative to NWS is the size of

the alert budget. The fact that the other policies find it harder to improve on NWS with more alerts may be because in such settings it is less critical to discriminate days when heat alerts will be most effective, given that our rewards model assumes more alerts never hurt in an absolute sense. The second split in the tree indicates that the RL performs best when there is greater variation in alert effectiveness across days, in other words, when there is more signal in the heat alert-health relationship that can be leveraged by the RL. The third split highlights how it is better to use RL in counties that tend to experience more prolonged heat waves, which cause AA.QHI to issue more alerts in a row, increasing the likelihood of alert fatigue. Running the same polytomous classification on a set of more conventional covariates (Figure B.8), we see a nearly-identical tree except with median household income in place of alert effectiveness and humidity in place of prolonged heat waves. Both of these associations make sense from a domain science perspective.

ADDITIONALLY, RL PERFORMS BETTER THAN NWS EARLIER IN THE SUMMER

Conducting regression on the difference in average returns between A2C.QHI and NWS using CART (shown in Figure B.9) generated complementary insights. The most predictive feature is the median day of summer on which the RL issued heat alerts: A2C.QHI performs better in counties for which it identified that it is optimal to issue alerts earlier in the summer. Whereas, in counties for which it is optimal to issue alerts later in the summer, there was less room for improvement by the RL because the NWS already tends to issue alerts later in the summer.

2.6 CONCLUSION & FUTURE WORK DIRECTIONS

This work lays the foundation of SDM for climate & health, by (1) formulating and building an SDM pipeline to optimize heat alert issuance, via integrating traditional statistical methods with cutting-edge RL techniques, and (2) offering insights into scenarios where RL improves vs. fails to improve

on the observed policy or simpler alternatives.

Ultimately, we found that it was necessary to restrict alerts to days with higher QHI for standard RL algorithms (DQN, QRDQN, TRPO and A2C) to be effective. Our post-hoc analysis enabled investigation of where RL performed better than the NWS policy and the simpler alternative AA.QHI. Future work might explore the use of methods which ensure a new policy is never worse than the existing policy, such as safe policy learning or model predictive control. Alternatively, we could develop a preliminary predictive model to select counties that are likely to benefit from RL, using intuitions similar to those in our CART analysis.

We anticipate several common questions about our approach. The first is how to move beyond a fixed alert budget, important in a changing system (under climate change, the number of extreme heat events is expected to increase). From a practical standpoint, real monetary budgets from stakeholders could be used, for instance “how many times can we afford to open our cooling center(s) this summer?” Future work could also explore modeling alert fatigue as part of the RL environment, for instance by incorporating tools from behavioral science. RL algorithms that are intentionally robust to distribution shift also merit investigation in a climate & health setting.

A second question is whether offline (batch) RL methods, which use only observed data in place of a simulator, could be used to circumvent dependence on the specification of a rewards model. Several major challenges stand in the way of an offline approach, highlighted in our heat alerts example. First, there is a tension between trying to improve the observed policy and controlling distribution shift⁹¹: for instance, we see that NWS often issues alerts in streaks, making it hard for offline RL to learn and optimize the impact of an individual alert. Second, assessing the performance of offline RL using off-policy evaluation methods is made difficult by long episodes, large potential for distribution shift (especially in the absence of a modification such as restricting alerts to very hot days), and high degree of autocorrelation in the reward¹⁴⁸.

We close with a broader view of climate & health decision making. Ideally, an optimal heat alert

system would account for different types of health impacts experienced by different demographic groups—future work could explore multi-objective RL. Similarly, it is important to recognize that the issuance of heat alerts is only a first step in reducing the public health impacts of extreme heat⁴⁸: there is much more work to be done analyzing and expanding local actions in response to heat alerts. If such on-the-ground interventions are able to increase the effectiveness of heat alerts, then this is likely to increase the ability of SDM methods such as RL to identify these effects and help us continue improving our strategies in the future.

3

A Spatiotemporal, Quasi-experimental Causal Inference Approach to Characterize the Effects of Global Plastic Waste Export and Burning on Air Quality, Using Remotely Sensed Data

3.1 INTRODUCTION

THE OPEN BURNING OF WASTE, INCLUDING PLASTIC WASTE, IS AN URGENT GLOBAL HEALTH ISSUE ¹¹⁸. Open burning is defined as burning “in open fires without managing for the emission of

byproducts, such as gases and ash, into the ambient air or soil”¹¹⁹. An estimated 40-65% of total municipal solid waste is open-burned in low- and middle-income countries (LMICs), largely as a result of two billion people around the world receiving no municipal solid waste collection¹¹⁸. As will be highlighted in this paper, LMICs also receive large amounts of waste from high-income countries. Regardless of the source of the waste, waste burning threatens public health due to the air pollutants emitted, such as fine particulate matter^{167,9}, abbreviated PM_{2.5}. There is a large body of scientific evidence linking air pollution, and specifically PM_{2.5}, to health consequences ranging from respiratory and cardiovascular disease, to cancer, to reproductive and neurological disorders, to mortality¹¹⁰.

A decade ago, Wiedinmyer et al.¹⁶⁷ presented the first comprehensive, global estimates of air pollution (and greenhouse gas) emissions due to open waste burning, using an emissions factor approach based on guidelines from the IPCC (Intergovernmental Panel on Climate Change). This approach relies on rough country-level approximations for the amount of waste generated and open-burned, which were derived as a deterministic function of national waste generation rate and waste collection efficiency estimates. The estimated amount of open waste burned was then converted into waste burning-attributable emissions estimates through multiplying by a literature-informed emissions factor. This and other inventories have subsequently been paired with chemical transport modeling to estimate the impact of open burning on regional air quality and attributable mortality⁵⁶. While these studies provide evidence that the air pollution contributions of waste burning may be substantial, their methods require strong assumptions backed up by very limited empirical data, and adaptation of these methods to conduct policy impact evaluations, which are crucial for future policy making, would necessitate further strong assumptions about the impact of a policy on the various inputs.

In this study, we focus on plastic waste, whose global generation more than doubled from 2000-2019¹¹⁴ and whose transport / trade from higher- to lower-income countries is a major environmental justice issue. In contrast to earlier studies estimating emissions from burning all kinds of waste, recent studies have focused on more robust estimation of plastic waste generation in specific regions⁹ and

detecting molecular tracers from plastic burning⁷⁶. However, these methods do not easily scale to large regions and/or time periods.

To generate robust empirical evidence about the impacts of plastic waste burning on air quality at scale, here we leverage a case study featuring a large plastic waste quasi-experiment. On January 1, 2018, China banned the import of plastic waste^{*}. Previously, between 1988 and 2016, they received 45% of global plastic waste exports¹⁶. The 38 OECD (Organisation for Economic Co-operation and Development) countries, most of which are classified as high-income, contributed 64% to all exports in this time frame, with the United States, Japan, and Germany topping the list¹⁶. Unsurprisingly, in the wake of China's ban there was a marked increase in plastic waste exports to other countries in the East Asia & Pacific region, which were already receiving 30% of global exports¹⁶. In this paper, we focus on Indonesia, which in 2018 became a net importer of plastic waste⁵⁵. As of 2020, it was estimated that 48% of Indonesia's plastic waste is openly burned⁵⁵.

In this paper, we seek to quantify the increase in air pollution at open dump sites in Indonesia during 2018-2019 due to (the burning of) increased plastic waste in the wake of China's 2018 ban. However, data scarcity presents significant obstacles to such an analysis. Indonesia, like many LMICs, has very limited ground-level ("in situ") air quality monitoring. The country's first reliable air quality monitors were deployed in Jakarta during our study period¹⁷³. For context: as of 2020, 51% of the world's governments did not produce air pollution data^{115†}. In recent years, remotely sensed data have been widely used to fill gaps in in-situ monitoring¹³². In addition to the lack of air quality data, to our knowledge there is no detailed record of Indonesia's plastic waste imports along the lines of what Unfried & Wang¹⁴⁹ used to investigate domestic air quality impacts of China's ban – and in any case, use of an official record might obscure the contribution of informal and/or covert transportation of waste, which is suggested by analysis of the Indonesian waste sector¹⁶¹. Therefore, we develop an

^{*}China also implemented a second ban in January 2019, tightening its initial policy. However, Unfried & Wang¹⁴⁹ found that the impact of this second ban was minimal.

[†]Note that OpenAQ has reported updates to this statistic in 2022 and 2024.

analytic strategy relying on data products derived from remote sensing, including $\text{PM}_{2.5}$ estimates, the identification of locations where waste has been openly dumped, and proximity to ports from which plastic waste can enter Indonesia (data sources are detailed in Section 3.2). To our knowledge, this is the first analysis providing empirical estimates of how a policy-driven change in plastic waste quantity – and associated changes in burning – impacts air quality, in a setting without in-situ air quality monitoring or waste quantity data.

Motivated by this case study, we also propose a new, multiply-robust causal inference approach to estimation and inference for quasi-experimental study designs with a “universal” but dose-dependent intervention. A growing body of work uses quasi-experimental study designs to estimate the air quality impact of a policy or intervention, with a recent surge focused on the Covid-19 lockdown – which resembles our setting in causing a nationwide shift in emissions-relevant behaviors^{67,20,34}. Policies that are adopted “universally” across a country/region of interest present a challenge for many conventional quasi-experimental analytic methods, which typically assume the existence of a control group. To address this, Heffernan et al.⁶⁷ developed a framework for machine learning-based comparative interrupted time series (CITS), in which U.S. city-specific models were used to estimate the difference in average daily air pollution before and after the lockdown dates, using past years of data as controls. A closely related approach was taken by Chen et al.²⁰, who used synthetic controls in place of CITS. However, these methods that model each location’s effect separately do not satisfy our needs, because we are interested in characterizing the relationship between changes in air quality and each location’s “dose” of exposure to the treatment / intervention, which in our case is port proximity as defined in Section 3.2.3. That is, we wish to estimate a causal exposure-response curve.

Our proposed method merges ideas from this literature on quasi-experiments with a “universal” intervention and ideas from the burgeoning literature on multiply-robust causal exposure-response curve estimation. Specifically, we build closely upon recent work by Hettinger et al.⁷¹, who introduced a multiply-robust estimator for causal exposure-response curves within the difference-in-differences

(DiD) framework. Derived from the efficient influence function or EIF (see Section 3.3.2), this estimator accounts for confounding both between treated and control units and across levels of exposure to the treatment / intervention, which can be thought of as a dosage. We propose a related EIF-based multiply-robust estimator for causal exposure-response curves in quasi-experimental designs with a universal intervention. We also propose an uncertainty quantification approach that accounts for both spatial and temporal correlation. We apply this method to remotely sensed data for the Indonesia plastic waste case study to estimate the change in air pollution at open dump sites due to increased plastic waste in the wake of China’s 2018 ban, as a function of a dump site’s port proximity.

3.2 DATA & DEFINITIONS

We compile a monthly analytic dataset spanning 2012-2019, not including more recent years to avoid confounding by the Covid-19 pandemic. This dataset will be made available on Harvard Dataverse post-acceptance of this paper.

3.2.1 SPATIAL UNITS OF ANALYSIS: WASTE DUMP SITES

Waste dump sites serve as our spatial units of analysis in this study. While plastic makes up only 12% of global municipal solid waste, the total aggregation of waste is a key proxy for aggregations of plastic waste⁸⁵. Indonesia does not provide official data on dump site locations in the country; thus, we source these data from remote sensing products. Global Plastic Watch (GPW) is a data platform containing satellite imagery-derived classifications of open dump sites, also known as waste aggregations. Kruse et al.⁸⁵ used machine learning to do this classification, and evaluated their performance in Indonesia. The GPW dataset has been used by high-profile groups such as ClimateTrace to estimate emissions from the waste sector¹²⁴.

We use all dump sites in Indonesia detected by GPW prior to 2020 for which PM_{2.5} data were

available in the Asia region PM_{2.5} data files from¹⁵⁸, described below. This results in 356 sites included in our analyses, out of 357 total detected by GPW in Indonesia. Site footprints are provided monthly starting in 2015, but there are frequent gaps and GPW notes that the average across time is more reliable than any given time point. Averaged across time (through the end of 2019), the median area of open dump sites in Indonesia is 2,317m² (IQR=[881m², 5,982m²]), but this distribution is highly right-skewed, with a maximum value of 288,424m² (0.29km²).

WHY NON-DUMP SITES SHOULD NOT BE USED AS “CONTROL” LOCATIONS

To motivate our methodological approach (in which we use past years as controls for future years at the GPW dump sites), briefly consider what trying to identify separate “control” locations would look like.

Let us walk through several of the causal identifying assumptions made in⁷¹, which are standard in DiD-type analyses. **Positivity** requires that controls have a non-zero probability of being treated. In our setting, this might be reasonable if we selected non-dump locations with similar characteristics to dump sites, including their proximity to major roads and population density. The **Stable Unit Treatment Value Assumption (SUTVA)** requires that (i) there is no interference between units and (ii) there are no hidden variations in treatment¹²⁷. In our setting, spillover of air pollution becomes a concern if controls are spatially proximate to dump sites, which would be made more likely if we prioritized non-dump locations with similar landcover characteristics. The second aspect of SUTVA might be violated in our case study if we accidentally selected some “control” locations where waste is dumped and burned, but dumping there went undetected in the satellite imagery. **Conditional Parallel Trends** requires that the pre-post difference in the treated group would be the same (conditional on covariates) as that in the control group, in the absence of treatment. Here, consider that Indonesia’s domestic plastic waste generation is increasing by about 5% per year⁵⁵. Since we know that a substantial fraction of this waste is taken to dump sites and burned, we would expect the long-term

trend in air quality over dump sites to be different than that over non-dump sites, regardless of waste being imported.

Taken together, these considerations illustrate why attempting to use non-dump sites as controls would likely lead to violations of causal identifying assumptions and biased estimates. Thus, we use only waste dump sites in our analysis and assume that they are universally treated by the China plastic waste ban starting in 2018.

3.2.2 OUTCOME: FINE PARTICULATE MATTER (PM_{2.5}) CONCENTRATIONS

Our outcome of interest is PM_{2.5} concentrations at dump sites. While in-situ measurements are not available across Indonesia, van Donkelaar et al.¹⁵⁸ estimated global monthly average PM_{2.5} at 0.01° × 0.01° (approximately 1km) resolution using a combination of remotely sensed aerosol optical depth (AOD), chemical transport modeling, and in-situ measurements from around the world. Briefly, they relate AOD to surface PM_{2.5} concentrations using geophysical relationships observed between in-situ PM_{2.5} measurements and AOD simulated from a chemical transport model. They then apply geographically weighted regression to predict (and later adjust for) the residual bias with the in-situ monitors. At monitor sites in Asia from 2015-2019, they report cross-validated R^2 values ranging from 0.59–0.86. Unsurprisingly, estimates over regions with fewer in-situ monitors have higher uncertainty; however, this data is widely used for exposure assessment and epidemiology studies.

We extract these monthly PM_{2.5} values for 2012-2019 at the GPW sites, for each site taking the value of the grid cell in which the site centroid lies. The average PM_{2.5} concentration at the dump sites in 2012-2017 (prior to the China ban) was 25.7 $\mu\text{g}/\text{m}^3$. The time series of concentrations averaged across sites (Figure 3.1) shows a strong seasonal pattern across our study period.

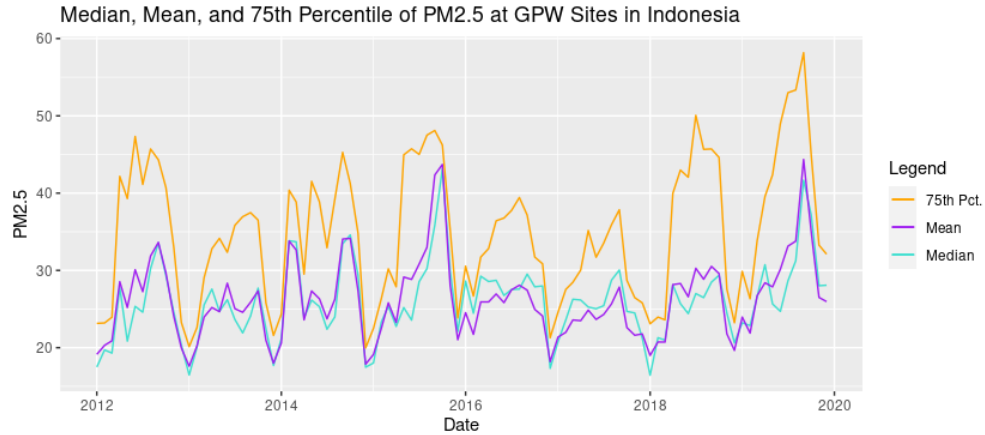


Figure 3.1: Median, mean, and 75th percentile of $PM_{2.5}$ across GPW sites in Indonesia, 2012-2019.

3.2.3 EXPOSURE: PORT PROXIMITY

Because the policy intervention is on waste imports and in Indonesia any such imports should come through ports, presumably dump sites closer to port activity should be more heavily affected by waste imports. Thus, we consider each dump site's proximity to port activity to be a proxy measure of treatment dose (or "exposure"). Following other recent work assessing the domestic impacts of China's plastic waste ban Unfried & Wang¹⁴⁹, we could use distance to the nearest port as a proxy. However, that would not take into account the amount of shipping activity in each location. Therefore, we used data from shipping signals remotely sensed via the Automated Identification System, specifically a derived measure of cargo ship loitering from the Global Maritime Traffic Density Service⁵⁴. For our continuous exposure metric, we calculated an inverse-distance-weighted sum of the 2018 loitering data around every GPW site. Due to the strongly right-skewed distribution of this port proximity index, we use a quantile version throughout the analysis, as shown in Figure 3.2. This transformed variable is henceforth referred to simply as the "port proximity".

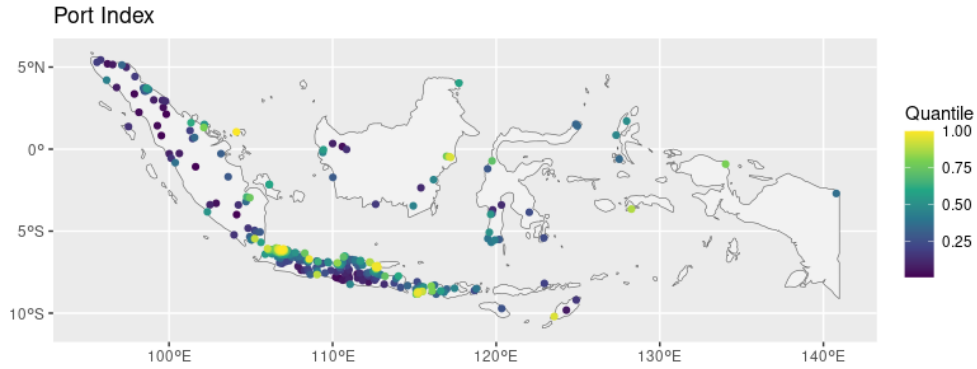


Figure 3.2: Quantile of the port proximity index for each GPW site in Indonesia.

3.2.4 COVARIATES: METEOROLOGY, POPULATION DENSITY, AND FIRE LOCATIONS

As in previous analyses of pre/post-intervention air quality data, the primary covariates in our analyses are meteorological factors^{67,20,34}. By chance, meteorological conditions may differ in the pre- and post-intervention periods, also influencing air quality, and thus failure to adjust for them may bias our estimates of the treatment effects. We obtained ERA5 monthly-aggregated data for 2012-2019²⁶ via Google Earth Engine⁵⁷. This climate reanalysis product, which integrates remotely-sensed and in-situ observations with physics-based models, is available at a spatial resolution of $0.25^\circ \times 0.25^\circ$ (approximately 28km). We take the value at each dump site to be that of the ERA5 grid cell the dump site centroid lies within. We use average air temperature at 2m height, dewpoint temperature at 2m height (indicating humidity), total precipitation, surface pressure, and the u- and v-components of wind speed as covariates in our model.

Population density is another potential confounder in our case study. The GPW API provides an estimate of the population living within 1, 5, and 10 kilometers of each dump site, sourced from WorldPop¹³⁸. Because we observed similar patterns across the sites between these variables, we chose to adjust for population living within 1km for simplicity (it is straightforward to think of this as pop-

ulation density).

Lastly, we obtained the locations of active fires from VIIRS, the Visible Infrared Imaging Radiometer Suite¹⁰⁹. The availability of this remotely sensed product (375 m resolution daily) starting in 2012 influenced our selection of the study period start. A time series of the total number of active fires across Indonesia is shown in Figure C.1. After obtaining the maximum footprint of each dump site as detected by GPW, we performed a spatial overlay of the active fire points. To account for imprecise detection of site boundaries as well as shape fluctuation over time, we used a 1000 m buffer around each site for the overlay. Throughout our analysis, we consider non-waste dump fires to be those that occurred more than 1000 m away from any GPW site. We use the monthly count of fires occurring outside of open dump sites in the province as a covariate in our main analysis. The purpose of this is to adjust for potential changes over time in wildfire activity and agricultural burning practices, which could also contribute to changes in PM_{2.5} during the study period.

To assess plausibility that changes in waste burning are contributing to any changes in air quality post-intervention, we conduct a post-hoc exploratory analysis of the temporal trend in the number of VIIRS active fire points at open dump sites in Indonesia during the study period. Increases in active fire points at open dump sites from 2018 onwards would provide informal evidence that changes in air quality may be (at least partially) due to increases in waste burning.

The VIIRS fire overlay also justifies our analysis of all dump site locations from 2012-2019, regardless of when GPW first detected them (prior to 2020). As shown in Figure C.2, there were many locations at which major fires occurred prior to their date of first detection as a dump site by GPW. Having observed fires at a site before it was classified as a dump site by GPW increases the likelihood that the site was being used as a dump site prior to the GPW designation.

3.3 METHODS

Analytic code and detailed information on data processing will be made available on GitHub post-acceptance of this paper.

3.3.1 SETUP AND ASSUMPTIONS

Let $\mathcal{O}$ be the observed data, containing: Y the outcome ($\text{PM}_{2.5}$), D the dose of exposure (port proximity), and $\mathbf{X}$ the covariates. The latter includes the population living within 1km of the site, monthly averaged meteorology (temperature, humidity, precipitation, pressure, and wind speed), and the number of active fires detected in the province (that month) which did not overlap with any GPW site. To help capture spatial heterogeneity, we also include fixed effects for site ID and province in our model; the spatial boundaries for the latter are shown in Figure C.3. Lastly, to account for longer-term trends, we include time $t = 1, \dots, 96$ to represent the months between 2012 and 2019, inclusive.

Furthermore, let B (for “ban”) be a pre- vs. post-intervention indicator. $B = 0$ indicates dates before policy implementation, and $B = 1$ indicates dates after (in our case, 1/1/2018 onwards). Because all units (dump sites) in the analysis are treated post-intervention, B is also a binary indicator of treatment status. We use B rather than the more commonly used A to illustrate this departure from DiD.

Each observation is indexed by its location i ($N = 356$ dump sites) and time point t . When $B_{it} = 0$, we say $D_{it} = \emptyset$, meaning D is undefined – while each location i has a port proximity prior to China’s ban, this feature has not yet been activated to represent the dose of treatment. We define the potential outcomes as $Y_{it}(B_{it} = b, D_{it} = \delta)$, i.e., the outcome that would have been observed for a given unit under intervention status b and dose δ . This definition of the potential outcomes requires that the future cannot affect the past (arrow of time, call this A1) and SUTVA (A2), discussed in section 3.2.1. Mathematically, SUTVA requires that $Y_{it}(\mathbf{B}_t, \mathbf{D}_t) = Y_{it}(B_{it}, D_{it})$, where $\mathbf{B}_t$ and $\mathbf{D}_t$ denote

the population-level vectors of intervention and dosage.

Our aim is to estimate a causal exposure-response curve quantifying the causal effect on Indonesian air quality of China’s 2018 plastic waste import ban at different levels of port proximity, δ . In the DiD setting, Hettinger et al.⁷¹ terms this the “Average Dose Effect on the Treated”, or ADT. While noting again that our formulation is different, we still see fit to call our effect of interest the ADT:

$$ADT(\delta) = \Psi(\delta) = E[Y(1, \delta) - Y(0, \emptyset) | B = 1]$$

In our case study, this estimand describes the average difference in air quality post-China ban at open dump sites in Indonesia, if all dump sites were exposed to port proximity dose δ . To identify the $ADT(\delta)$, we require

$$(A_3) \text{ Consistency: } Y_{it}(b, \delta) = Y_{it} \text{ when } (B_{it}, D_{it}) = (b, \delta)$$

(A4) Positivity: there exists $\varepsilon > 0$ such that

$$\begin{aligned} \text{(i)} \quad & \varepsilon \leq \pi_B(\mathbf{x}) < 1 \quad \forall \mathbf{x} \in \mathbb{X}, \quad \pi_B(\mathbf{x}) = P(B = 1 | \mathbf{x}) \\ \text{(ii)} \quad & \varepsilon \leq \pi_D(\delta | \mathbf{x}, B = 1) \quad \forall \mathbf{x} \in \mathbb{X} | B = 1, \delta \in \mathbb{D}, \quad \pi_D(\delta | \mathbf{x}, B = 1) = p(D = \delta | B = 1, \mathbf{X} = \mathbf{x}) \end{aligned}$$

In contrast to Hettinger’s Conditional Counterfactual Parallel Trends Between Treated and Control, we require

$$(A_5) \text{ Conditional Ignorability Between Time Periods: } E[Y(0, \emptyset) | B = 1, \mathbf{X}] = E[Y(0, \emptyset) | B = 0, \mathbf{X}]$$

This means that there is *no residual time-varying confounding*. Assembling an adequate set of covariates to justify this assumption relies on domain knowledge. Its plausibility can be strengthened by adjusting for seasonal and/or long-term trends, e.g., through inclusion of a time variable in the adjustment set. This assumption is needed due to the absence of control locations in our application.

Lastly, in contrast to Hettinger’s Conditional Counterfactual Parallel Trends Among Treated Between Doses, we require

(A6) Conditional Ignorability Among Treated Between Doses:

$$E[Y(1, \delta)|B = 1, D = \delta, \mathbf{X}] = E[Y(1, \delta)|B = 1, \mathbf{X}] \quad \forall \delta \in \mathbb{D}$$

This means that there is *no residual spatial confounding*, in other words we would expect the post-policy air quality to be the same at all dump sites with covariates $\mathbf{X}$, if they were assigned dose δ (regardless of their observed port proximity). The plausibility of this assumption may be strengthened by adjusting for region-specific fixed effects (in our case, province).

Assuming these six conditions, our $ADT(\delta)$ is identifiable with observed data:

$$\Psi(\delta) = E[Y(1, \delta) - Y(0, \emptyset)|B = 1] \tag{3.1}$$

$$= E[E[Y(1, \delta) - Y(0, \emptyset)|B = 1, \mathbf{X}]|B = 1] \tag{3.2}$$

$$= E[E[Y(1, \delta)|B = 1, D = \delta, \mathbf{X}] - E[Y(0, \emptyset)|B = 0, \mathbf{X}]|B = 1] \tag{3.3}$$

$$= E[E[Y|B = 1, D = \delta, \mathbf{X}] - E[Y|B = 0, \mathbf{X}]|B = 1] \tag{3.4}$$

Where line (1) is by definition, (2) by iterated expectations, (3) by A6 on the lefthand side and A5 on the righthand side, and (4) by A1-3.

3.3.2 MOTIVATIONS FOR EFFICIENT INFLUENCE FUNCTION-BASED ESTIMATION

A conventional approach to estimate an exposure-response curve in a quasi-experimental setting would be to use a regression model like this one:

$$E[Y_{it}] = \beta_0 + \beta \mathbf{X}_{it} + \tau B_t + f_\varphi(D_i)B_t$$

where $f_\varphi(D)$ is a flexible (nonparametric) function, with parameters φ . There are several drawbacks of this approach. To begin, even if we allow a flexible relationship between D and Y , we are not allowing for flexible relationships between $\mathbf{X}$ and Y , nor interactions between the components of $\mathbf{X}$. Importantly, we are not addressing confounding between the dose D and trends in the outcome – for instance, if the overall (or seasonal) trend in air pollution is affected by population density, which is static in our dataset and which we know to be correlated with D (see Section 3.4.2). To address this, we could make a much more complicated linear regression (e.g., by using splines and more interaction terms), or use a nonparametric machine learning algorithm to model Y given $D, \mathbf{X}, B$. However, this outcome model might still be misspecified. This motivates the search for a doubly-robust estimator, which incorporates both a model for the outcome and a model for the propensity score (the probability of receiving treatment), which explicitly addresses confounding, and will produce results that are consistent (reliable in a probabilistic sense) if at least one of these models is correctly specified.

We turn to estimators based on the the Efficient Influence Function (EIF). Without going into the technical details (interested readers should see Kennedy⁸³ and Schuler & van der Laan¹³⁰), the EIF for a specific estimand can be used to construct an estimator that is doubly robust and attains the nonparametric efficiency bound (a nonparametric analog of the Cramer-Rao bound), at least asymptotically. In practice, one has to derive the EIF only for a novel estimand, and in many cases this can build closely on past work. Kennedy et al.⁸⁴ provided a doubly-robust estimation method for causal exposure-response curves in observational data settings, which Hettinger et al.⁷¹ then extended to a

DiD framework, including conditions for “multiple robustness” which will be described in the following section. We further extend this multiply-robust estimator to a setting without control locations, which as described in Section 3.1 is another common type of quasi-experiment. This work contributes to the nascent application of EIF-based estimation techniques in quasi-experimental settings.

3.3.3 ESTIMATION

The major difference between our estimator and that of Hettinger et al.⁷¹ is that instead of modeling the pre-post difference in outcomes, we are just modeling the outcome Y because we are using pre-intervention years rather than separate control locations. This does not substantively change the steps of the proofs in Hettinger et al.⁷¹ to (1) derive the EIF for the ADT[‡] and (2) show that the resulting estimator is multiply robust. At a high level, after obtaining the EIF, we construct the estimator via the method of estimating equations, which involves setting the EIF equal to zero and solving for Ψ ¹³⁰. Given our goal of estimating $\Psi(\delta)$, the relevant parts of this estimator are:

$$\begin{aligned}\xi(D, \mathbf{X}, B, Y; \mu_1, \pi_D) &= B * \left[\frac{f(D|B=1)}{\pi_D(D|\mathbf{X}, B=1)} (Y - \mu_1(D, \mathbf{X})) + m(D|B=1) \right] \\ \tau(\mathbf{X}, B, Y; \mu_0, \pi_B) &= \frac{1}{P(B=1)} \left[\frac{(1-B) * \pi_B(\mathbf{X})}{(1 - \pi_B(\mathbf{X}))} (Y - \mu_0(\mathbf{X})) + B * \mu_0(\mathbf{X}) \right]\end{aligned}$$

[‡]Technically, they derive the EIF for the integral of $ADT(\delta)$ to satisfy the pathwise differentiability requirement.

where

$$\begin{aligned}
\mu_0(\mathbf{X}) &= E[Y|B = 0, \mathbf{X}] \\
\mu_1(D, \mathbf{X}) &= E[Y|B = 1, D, \mathbf{X}] \\
m(D|B = 1) &= \int_{\mathbb{X}} \mu_1(D, \mathbf{x}) dP(\mathbf{x}|B = 1) = E[Y|B = 1, D] \\
f(D|B = 1) &= \int_{\mathbb{X}} \pi_D(D|B = 1, \mathbf{x}) dP(\mathbf{x}|B = 1)
\end{aligned}$$

In words: μ_0 and μ_1 are the outcome regressions in the pre- and post-intervention periods, respectively, and $m(D|B = 1)$ is the marginalized outcome regression in the post-intervention (treated) period, averaged across the covariates so that it only depends on the dose of exposure D . π_B is the propensity score for being intervened upon (in our case, being in the treated period), π_D is the generalized propensity score (GPS) for receiving dose D of treatment, given being in the treated period, and $f(D|B = 1)$ is the marginalized GPS.

Overall, this estimator is unusual in that it has separate functions ξ and τ to estimate the outcome in the presence and absence of treatment – in our case meaning the years after intervention and before intervention respectively. The form of both functions, including both an outcome model and a propensity score-based weight, foreshadows the multiple robustness conditions of this estimator: we need at least one of *each pair* (μ_1, π_D) and (μ_0, π_B) to be correctly specified for the results to be consistent.

To estimate $\Psi(\delta)$, we follow this procedure:

1. Estimate nuisance functions $\mu_1, \mu_0, \pi_B, \pi_D$
2. Calculate $\hat{\xi}(D, \mathbf{X}, B, Y; \hat{\mu}_1, \hat{\pi}_D)$ and $\hat{\tau}(\mathbf{X}, B, Y; \hat{\mu}_0, \hat{\pi}_B)$ on the empirical data using the estimated nuisance functions

3. Regress (smooth) $\hat{\xi}$ on D to obtain $\hat{\Psi}_D(\partial)$, and take the empirical mean of $\hat{\tau}$ to obtain $\hat{\Psi}_0$
4. $\hat{\Psi}(\partial) = \hat{\Psi}_D(\partial) - \hat{\Psi}_0$

The estimation of the nuisance functions can be carried out using any flexible modeling technique, e.g., tree ensembles or SuperLearner. Local linear kernel regression (LLKR) is employed for the smooth regression in step 3. The methods and implementation details used in our case study analysis are described in Section 3.3.5.

3.3.4 UNCERTAINTY QUANTIFICATION

For uncertainty quantification, Hettinger et al.⁷¹ derive a sandwich variance estimator, but mention various limitations of that approach, including that it relies on a normal approximation, that incorporating uncertainty in nuisance function estimation is challenging, and that extensions would be required to account for correlated data. To avoid these complications, here we use a novel variant of the nonparametric bootstrap to obtain pointwise confidence intervals for the ADT. This approach avoids making asymptotic assumptions in our finite sample setting and is designed to respect the complex spatial and temporal correlation structures in our data. Rather than discrete bootstrapping, where observations are either included in a sample or left out entirely, we use weights drawn from an Exponential distribution to allow for small but nonzero inclusion, which helps to improve the observed support of D in the samples⁷¹. (Note that these weights are scaled to reflect the observed sample sizes in the pre- and post-intervention time periods.) Sampling one weight per location (dump site) rather than per individual observation accounts for temporal correlation across repeated observations. An important extension we propose, motivated by our case study, is to adapt the weights to account for spatial correlation in addition to temporal.

SPATIAL WEIGHTED BOOTSTRAP

Because our set of covariates likely does not explain all spatial correlation in the data, we need to account for residual spatial dependence in our uncertainty quantification. For instance, note that there are some clusters of dump sites near Jakarta (shown in Figure C.4) which we would be hard-pressed to call independent of one another, even conditional on our covariate set. Therefore, in our weighted bootstrap procedure we would like a site's weight to be larger (or smaller) if other sites nearby have larger (or smaller) weights. We achieve this through *spatially-correlated sampling of the bootstrap weights*.

To build intuition, imagine the following procedure to generate the weights, given a spatial correlation matrix Σ :

1. Sample from a Multivariate Normal distribution with mean zero and variance-covariance (in this case, correlation) matrix Σ .
2. Apply the inverse probability transform to obtain samples with a different (target) marginal distribution, such as the Exponential.

In practice, we can estimate Σ by modeling the empirical spatial variogram of the residuals (see the next section for more details on this estimation and explanation of which model's residuals are used). This approach of generating bootstrap weights from an estimated spatial correlation function is substantiated by Kurisu et al.⁸⁶, who show the mathematical validity of what they term the spatially dependent wild bootstrap (SDWB). While the SDWB paper assumes a Gaussian random field, they note that this assumption can be relaxed at the expense of additional technical complexity. The issue, described in depth by Tsoukalas et al.¹⁴⁶, is that the magnitude of correlation is not preserved during the mapping from Gaussian random variables to those from another distribution using the inverse probability transform. To end up with the desired spatial correlation in our Exponential weights,

we must first estimate an “equivalent correlation” Σ^* to feed into the Multivariate Normal sampling procedure. Namely, Σ^* should be specified such that, when the Multivariate Normal samples are inverse probability transformed to generate Exponential weights, the resulting weights have the desired spatial correlation (Σ). The estimation of Σ^* can be achieved using numerical integration, which is facilitated by the `anySim` package in R¹⁴⁶.

The complete proposed spatial weighted bootstrap procedure is detailed below.

1. Estimate the empirical variogram of the model residuals and use it to estimate the desired spatial correlation matrix for the weights, Σ (see next section).
2. Estimate the equivalent correlation matrix Σ^* for the desired marginal distribution of the weights (in our case, Exponential).
3. Sample weights $\tilde{w}_i$ (one per site) from a Multivariate Normal distribution with mean zero and variance-covariance matrix Σ^* .
4. Apply the inverse probability transform to $\tilde{w}_i$ to obtain weights w_i with a marginal Exponential distribution and correlation matrix Σ .
5. Use w_i as weights for each observation from location i in the dataset ($t = 1, \dots, 96$), and carry these weights through nuisance function estimation and regression on D to obtain an estimate of the ADT.
6. Repeat steps 3-5 to generate the bootstrap estimates of the ADT.

Point-wise confidence intervals for $ADT(\delta)$ can be formed by extracting the appropriate percentiles (in our case, 2.5 and 97.5) of the estimates across the bootstrap samples.

ESTIMATING THE RESIDUAL SPATIAL CORRELATION FUNCTION

In conventional outcome-model-only settings, residual spatial correlation refers to spatial autocorrelation in the residuals obtained from subtracting the model-predicted outcome values $\hat{Y}$ from the observed values Y . In our setting, the precise definition of residual spatial correlation is less evident due to the multiple stages of modeling that go into our estimation: both outcome and propensity score models are plugged into the EIF-derived equations ξ and τ , then $\hat{\xi}$ is regressed on D .

For convenience, let $\hat{\xi}_{it}$ denote $\hat{\xi}(D_{it}, X_{it}, B_{it}, Y_{it})$. Here, we focus on spatial correlation in the residuals $\hat{\xi}_{it} - \hat{\Psi}_D(D_{it})$ for all observations in the treated period $B = 1$. (Note that the counterfactual outcome $\hat{Y}_0$ is constant over space.) Intuitively, we address residual spatial correlation from the final stage of modeling because that is the stage in which our effect of interest is estimated, and it represents the “last opportunity” to explain spatial correlation in the data via modeling. From a more technical perspective, Lee et al.⁸⁹ explain that for doubly-robust methods utilizing LLKR, the model-based uncertainty arising from the second-stage estimation (the LLKR) is greater than that of the first-stage estimation (the nuisance models). Additionally, note that in settings like ours where the exposure of interest is spatially explicit, calculating residuals based on the final effect curve (which has been marginalized over all the other covariates) induces a different spatial structure than if we were to analyze residuals $Y_{it} - \hat{\xi}_{it}$.

After calculating the residuals $\hat{\xi}_{it} - \hat{\Psi}_D(D_{it})$, we estimate the empirical variogram using the robust estimator introduced by Cressie & Hawkins²⁷, which is less susceptible to influential outliers than the basic method-of-moments approach. We then fit a covariance function, $Cov(m; L, \sigma^2)$, to the empirical variogram, where m is the distance between each pair of sites, L is a bandwidth or “range” parameter, and σ^2 is the magnitude of the variance. The corresponding correlation function $Corr(m; L)$ is obtained by removing the constant σ^2 .

SIMULATIONS TO INVESTIGATE THE SPATIAL WEIGHTED BOOTSTRAP

A pseudo-simulation study was conducted to evaluate the properties of the proposed Spatial Weighted Bootstrap confidence intervals. In Appendix C.3, we provide full details on the data generating process, the calculation of performance metrics, and the results of this simulation study. Crucially, these simulations illustrate that our Spatial Weighted Bootstrap improves coverage of the confidence intervals relative to a Non-spatial Weighted Bootstrap, across a variety of scenarios.

3.3.5 CASE STUDY DATA APPLICATION

In our case study analysis, we use the Ranger algorithm in the `caret` package in R to estimate the nuisance functions, using separate models for $\mu_1, \mu_0, \pi_B, \pi_D$. Details of these models' hyperparameter tuning and estimation of π_D (the GPS) and $f(D|B = 1)$ are in Appendix C.1. However, note that while we referred to a common set of covariates ($\mathbf{X}$) throughout the previous sections, the set of covariates used in the two propensity score models is actually slightly smaller than that described in Section 3.3.1. The time (month) variable is not included in the model for π_B , because the set of covariates in a flexible machine learning model must not be able to perfectly predict the intervention⁶⁷ – in our case, time steps before and after 1/1/2018. Similarly, the dump site ID is not included in the model for π_D , because the port proximity value is unique to each site. The time variable is also not included in the model for π_D because our port proximity values are static over time.

To regress $\hat{\xi}$ on D (step 3 in the estimation procedure from Section 3.3.3), we use the `locpol` package in R to perform LLKR with a Gaussian kernel. This package utilizes leave-one-out cross-validation (LOO CV) to select the bandwidth parameter for the kernel regression. Therefore, naively applying the method to our dataset, which contains repeated observations (over time) at each unique value of D (one per dump site), selects an inappropriately small bandwidth. The resulting confidence intervals indicate high instability, as shown in Figure C.5. To address this issue, we instead apply the LOO CV

to each time step separately (across the 356 dump sites), and then calculate the average bandwidth across the time steps – recall that for the LLKR stage, we are just considering the post-treatment time period $B = 1$. This average bandwidth value is then used in the LLKR model fit on all time points simultaneously.

For the spatial weighted bootstrap, we compared several standard spatial covariance functions (Exponential, Linear, Spherical, and Matérn), ultimately selecting the Matérn as the one with the best RMSE when fit to the empirical variogram of the residuals from the LLKR stage. Therefore, the entries of Σ for our spatially-correlated weights have the form

$$\text{Corr}(m; L, \nu) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{m}{L}\right)^\nu K_\nu\left(\frac{m}{L}\right)$$

where Γ is the gamma function, K_ν is the modified Bessel function of the second kind, and ν is an additional smoothness parameter which is optimized alongside L .

The empirical spatial range was $\hat{L} = 26.8\text{km}$, and the optimal smoothness parameter was $\hat{\nu} = 0.35$. To gain intuition for the spatially weighted bootstrap, we also ran an analysis ignoring the possibility of spatial correlation and another using $L = 10\text{km}$ (with the same smoothness $\nu = 0.35$), to span the closely-clustered and numerous-fire sites near Jakarta (shown in Figure C.4). The upper lefthand plot in Figure C.7 illustrates the resulting magnitudes of correlation between sites at different distances, when using 10km and 26.8km as the spatial range parameter. Confidence intervals for the ADT (with or without spatial correlation in the weights) are calculated using 100 bootstrap samples.

For illustrative purposes, we compare the results from our multiply-robust EIF-based method to (i) a plain outcome regression, which is the same as that used inside the multiply-robust estimator, and (ii) a flexible but confounder-naïve approach. The latter estimates $E[Y|D = \delta, B = 1] - E[Y|B = 0]$, where the first expectation uses the same flexible regression on D as the main method (LLKR) and the second expectation is a sample mean. We also re-ran the analysis using 2017 and, separately, 2016 &

2017 as the “treated” period, to check for pre-trends / policy anticipation (whether the effect began before 2018). Conducting such tests is standard practice for quasi-experiments.

Lastly, while dump site size is not directly accounted for in our main analyses, we conducted a sensitivity analysis to assess how this factor might influence our results, as we thought it could be an important predictor of air pollution heterogeneity across open dump sites. As mentioned in Section 3.2, Global Plastic Watch estimates the area (footprint) of each dump site. In reality, the dump sites are continuously fluctuating in size, but the measurements at any given time point are less reliable than their average over time; there are also frequent gaps in the remotely sensed detection time series. Therefore, we took the area at each site to be the average across all measurements.

We did not adjust for GPW site size as a covariate in the main analysis because the area of the dump sites as measured after 1/1/18 is technically a post-treatment variable which can be affected by treatment, referred to as a “descendant” of treatment (note that only 68% of the dump sites have at least one area estimate prior to 2018, so we would have to exclude many sites, potentially biasing our causal effect curve, in order to include only pre-2018 area as a covariate). Descendants of treatment should not be conditioned on in causal inference analyses⁷⁰. We ran a sensitivity analysis to investigate the impact of including site size (averaged across 2012-2019) as a covariate. For simplicity, this sensitivity analysis uses the same nuisance function hyperparameters and spatial range (for the bootstrap weights) as the main analysis.

3.4 RESULTS

3.4.1 STATISTICAL ANALYSIS RESULTS

Figure 3.3 shows the estimated ADT curve and 95% bootstrap confidence interval (CI) from the main analysis. Positive ADT estimates provide evidence of post-China ban increases in $PM_{2.5}$ concentrations over GPW sites, compared to those expected under business-as-usual. Thus, our results indicate

that at open dump sites in Indonesia, the monthly $PM_{2.5}$ in 2018-2019 was higher than expected under business-as-usual (except for dump sites nowhere near port activity), with the effect ranging from $-0.17-1.72 \mu g/m^3$ (95% CI = (0.84, 2.66) for the maximum value) depending on the port proximity. As expected, we see smaller effects at locations with lower port proximity – the effect is not statistically significantly different from zero until $\delta = 0.2$ (at which $\Psi(\delta) \geq 0.76$), above which it remains significant. The decrease of the effect curve between medium and high levels of port proximity ($\delta \in [0.7, 1]$) could reflect smaller increases in dumping/burning in densely populated areas very close to ports, where there is likely more government oversight and less open space. Integrating across all values of the estimated effect curve yields an average increase at dump sites of $1.14 \mu g/m^3$ attributable to China's ban on plastic imports, corresponding to a 4.4% increase in overall $PM_{2.5}$ in 2018-2019 compared to 2012-2017. Note that if we only consider dump sites above the 20th quantile of port proximity (where the effect is statistically significant), the average increase is $1.38 \mu g/m^3$, corresponding to a 5.1% increase at those sites.

The results from our proposed multiply-robust method are similar to those produced using an outcome regression alone (Figure 3.3). But variations in both the curve (point estimate) and the confidence interval indicate that it is still important to guard against misspecification of the outcome model by incorporating propensity scores. By contrast, the confounder-naïve results look very different (Figure C.6): the mean effect ranges from -6.4 to $7.2 \mu g/m^3$. This divergence from our main result is unsurprising, but demonstrative.

The results of using different values as the range parameter for the bootstrap weights' spatial correlation are shown in Figure C.7. A range of 10km produces confidence intervals 74% the width, on average, of those in the main analysis (with a range of 26.8km). Ignoring spatial dependence entirely in the weighted bootstrap narrows the confidence intervals by more than a factor of two (on average, they are 45% the width of those in the main analysis). This demonstrates that accounting for residual spatial correlation is essential to avoid under-estimation of uncertainty in this application.

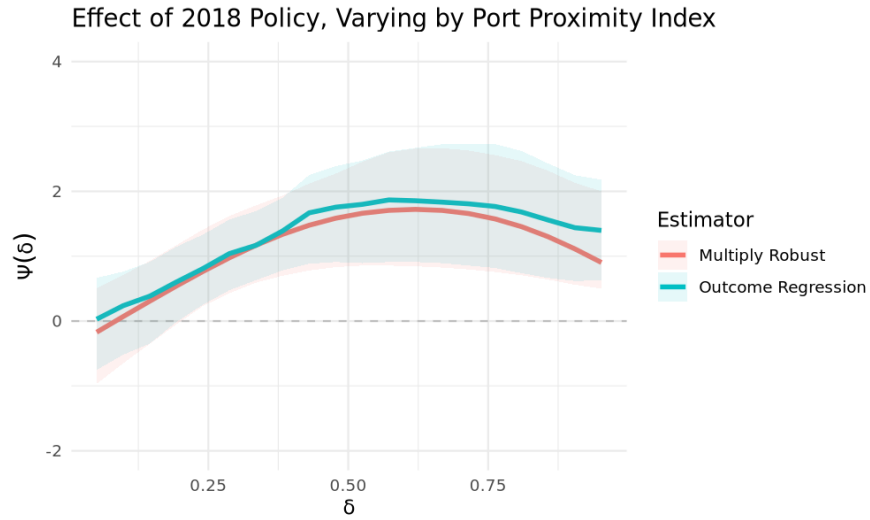


Figure 3.3: Average dose-effect on the treated (ADT) estimates and bootstrap 95% confidence intervals for our case study, estimated using both our proposed multiply robust method and a conventional outcome regression. ADT estimates (y-axis) quantify the average change in $PM_{2.5}$ ($\mu g/m^3$) post-China ban, compared to concentrations expected under business-as-usual, at open dump sites in Indonesia for a given port proximity (x-axis). The pointwise confidence intervals are generated using our spatial weighted bootstrap with Matérn correlation range=26.8km (as observed in the empirical residuals).

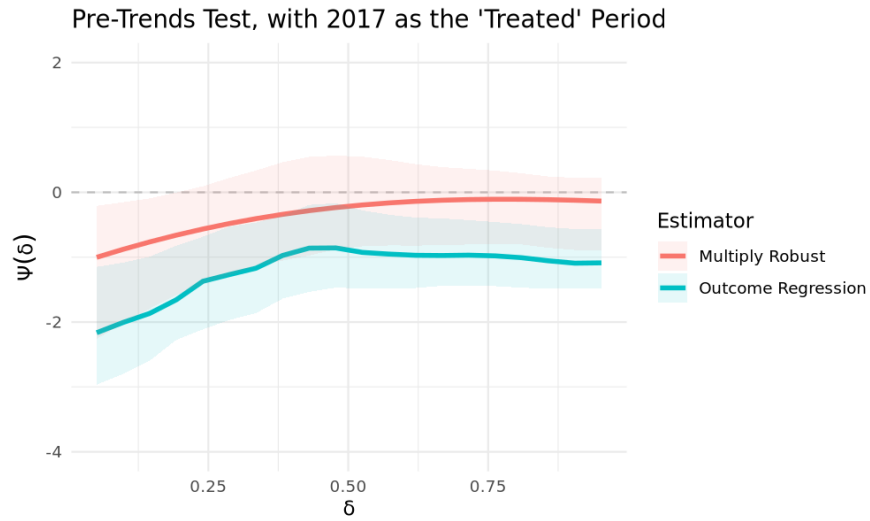


Figure 3.4: Pre-trends test using 2017 as the “treated” period; sampling bootstrap weights with spatial correlation (range=26.8km).

As shown in Figure 3.4, we did not observe a significant increase in $PM_{2.5}$ when using 2017 as the “treated” period (to check for pre-trends / policy anticipation) – in fact, the mean effect is negative, indicating lower $PM_{2.5}$ in 2017 compared to preceding years. Using 2016 & 2017 as the “treated” period (Figure C.8) yields the same conclusion.

Including GPW site area as a covariate in our sensitivity analysis did not substantially change the results. The multiply-robust curve in Figure C.9 is slightly less peaked (its maximum value is 1.45 rather than 1.72), but the confidence interval widths are practically identical.

3.4.2 POST-HOC EXPLORATORY DATA ANALYSIS RESULTS

To complement our main analysis, we investigated the occurrence of active fires at the GPW dump sites throughout our study period. Figure 3.5 shows the time series (aggregated by month) of VIIRS-detected fires overlapping with the dump sites. We observe a large increase in the number of fires in 2018-2019.

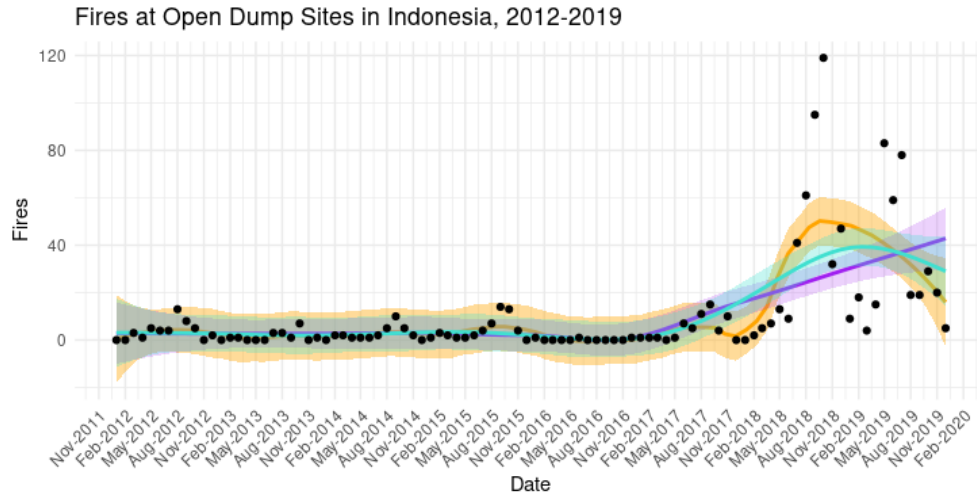


Figure 3.5: The number of fires detected by VIIRS overlapping with GPW sites (within a 100m buffer) across Indonesia. The colored trend lines are estimated using several different smoothing techniques: the orange using LOESS with span=0.25, the purple using LOESS with span=0.75, and the turquoise using a generalized additive model.

The distribution of the number of fires detected by VIIRS at each dump site is highly skewed, as shown in Figures C.3 and C.4 (maps of Indonesia overall and the Jakarta region respectively). This is somewhat explained by the area of the dump sites, but not entirely (Figure C.10). In keeping with our main analysis, the sites with the most fires are not those with the highest port proximity (Figure C.11) – there is a spike around $\delta = 0.7$, due to the cluster of large dump sites to the west of Jakarta, which partially motivated our spatial weighted bootstrap approach. We also note that the correlation between the port proximity and (square root-transformed) population living within 1km of a site is 0.6. While this collinearity might be an issue for estimating our effect of interest in a linear model, it is less of a concern for a doubly-robust nonparametric estimator in which we marginalize over the covariates.

3.5 DISCUSSION

In this case study, we observed a 4.4% increase in $PM_{2.5}$ at dump sites in Indonesia in 2018-2019 compared to 2012-2017. This effect is on a similar scale to the results of Unfried & Wang¹⁴⁹, who assessed the domestic impacts of China’s plastic waste import ban and estimated an average reduction of $3.7 \mu g/m^3$, corresponding to 8.5% reduction of overall $PM_{2.5}$ concentrations in China as a result of their 2018 ban. It makes sense that our observed effect would be smaller, given that Indonesia was only one of the countries to which waste that would otherwise have gone to China was diverted. Unfried & Wang¹⁴⁹ similarly did not observe an anticipation effect in mid-2017, when China announced its upcoming ban to the World Trade Organization.

Methodologically, we demonstrate the viability of pairing spatiotemporal causal inference methods with remotely sensed data products to quantify the air quality impacts of large-scale plastic waste policies via the mechanism of waste burning. On the data side, we harmonize areal classifications of open dump sites, cargo ship loitering densities, meteorologic estimates, active fire points, and high-

resolution air pollution estimates. For the statistical analysis, we combine a multiply-robust estimator for causal exposure-response curves within the difference-in-differences framework with the technique of using pre-intervention years as controls for post-intervention years. Additionally, we illustrate the importance of accounting for spatial dependence in the uncertainty quantification stage, and propose a spatial weighted bootstrap approach for doing so. Given that there are numerous documented quasi-experiments created by plastic waste policies globally⁸⁰, which have largely occurred in countries lacking in-situ air quality monitoring data, we anticipate that the analytic framework developed here can be much more widely applied.

This study has a few limitations which merit remarks. First, while we are focused on the air quality impact of China's 2018 ban and resultant increase in plastic waste in Indonesia, the observed difference between 2018-2019 and 2012-2017 might be affected by another policy that was implemented around the same time (2017-2018): Indonesia's National Action Plan on Marine Plastic Debris. The goal of this policy was to reduce the amount of plastic waste ending up in the ocean – with limited emphasis on reducing the amount of plastic waste being generated⁶; by 2021 Indonesia achieved an estimated 15% reduction of marine plastic debris from 2017 levels¹⁶¹. Ethnographic studies from around the world have found that campaigns to raise awareness about plastic pollution often contribute to increased open burning, as communities seek to cut down on litter and deal with waste that has been collected¹¹⁸. Hence, it is possible that these two policies were interacting to generate the observed increase in PM_{2.5} at dump sites in Indonesia. Our analysis showing no pre-trends effect when using 2017 as the “treated” year guards against this possibility somewhat, however the effect of the National Action Plan could lag or increase after 2017.

Second, while our interest is primarily on air pollution generated by the burning of waste, it is possible that some of the observed increase in PM_{2.5} at open dump sites is attributable to the *transportation* of this waste, for instance from diesel trucks making more trips to dump sites and, for sites close to the coast, from cargo ships. Our post-hoc analysis illustrating the uptick in active fires overlapping

with dump sites offers some evidence that at least part of the observed effect could plausibly be due to increases in waste burning. Unfortunately, speciated $PM_{2.5}$ estimates, which might enable the disentangling of emissions from waste burning and transportation, tend to be available only at coarse spatiotemporal resolutions globally and rely more heavily on models⁴⁹, which is likely to frustrate the application of quasi-experimental statistical methods. For smaller scale studies, Islam et al.⁷⁶ demonstrated the viability of measuring a molecular tracer of plastic burning with in-situ monitors. Alternatively, Scott et al.¹³¹ applied machine learning-based image classification to identify smoke plumes in high-resolution (3m x 3m) satellite imagery over the largest municipal landfill in the Maldives. However, in addition to being computationally intensive, this approach is not directly translatable into air pollution concentrations, which are easier to link to the rich literature on health impacts, and hence tend to be more relevant for policy makers.

Third, while remotely sensed data products enable application of quasi-experimental methods in data-scarce settings, we must be conscious of their potential for measurement error / misclassification. In this case study, we believe that the GPW dump site classifications are the data product in which errors would be most consequential for our analysis. Because all the GPW sites detected in Indonesia were manually validated⁸⁵, there should be no false positive classifications. However, false negative classifications are possible and could affect our analysis in two ways. On the one hand, because we include the number of active fires in each province *not overlapping with dump sites* as a covariate in $\mathbf{X}$, any fires mistakenly classified as non-dump fires will reduce our estimated policy effect. On the other hand, if there is spillover of $PM_{2.5}$ from false negative dump sites to nearby GPW-detected dump sites, then this could increase our estimated *per-site* policy effect.

A more general limitation of our approach is that we cannot estimate the effects of distributed waste burning, i.e., at the household level or at communal dump sites that were too small to be detected by GPW. Future research, and likely investment in ground-level air quality monitors (and/or lower cost sensors), will be needed to estimate these distributed burning emissions, which may ultimately pose

greater risks to public health due to higher frequency of burning, greater proximity to people, and lower dispersive dilution of air pollution at the ground level¹¹⁸ compared to taller plumes of smoke resulting from larger fires at larger dump sites.

Despite its limitations, this study provides strong evidence that the export of plastic waste to Indonesia, and subsequent open burning of import-related waste in Indonesia, has resulted in higher concentrations of PM_{2.5} – at least near large open dump sites. This motivates future policies to (a) reduce the amount of plastic waste being exported from the Global North to the Global South and (b) reduce the amount of waste being openly burned. Both will require substantial investments in waste management infrastructure and/or dramatic shifts in the amount of plastic waste being generated. With respect to waste burning, it has been observed that simply prohibiting open burning is seldom effective¹¹⁸, though burning at major sites may be easier to monitor and restrict¹³¹.

On the global stage, negotiations by the United Nations are currently underway to develop a Global Plastic Treaty – with an international legally binding instrument on plastic pollution¹⁵⁰ – under the Basel Convention on the Control of Transboundary Movements of Hazardous Wastes and Their Disposal. The effectiveness of this and future policies, at least for improving air quality but potentially also for other kinds of environmental outcomes, can in turn be evaluated using methods similar to those presented in this paper.

References

- [1] Agarwal, A., Alomar, A., Alumootil, V., Shah, D., Shen, D., Xu, Z., & Yang, C. (2021). PerSim: Data-Efficient Offline Reinforcement Learning with Heterogeneous Agents via Personalized Simulators. In *Advances in Neural Information Processing Systems*, volume 34 (pp. 18564–18576).: Curran Associates, Inc.
- [2] AirNow (2022). Airnow fire and smoke map. accessed 2022-11-03.
- [3] American Community Survey (2022). Data Profiles. The United States Census Bureau. <https://www.census.gov/programs-surveys/acs/data/eo-data/>. Accessed: 2022-01-27.
- [4] Anderson, G. B. & Bell, M. L. (2011). Heat Waves in the United States: Mortality Risk during Heat Waves and Effect Modification by Heat Wave Characteristics in 43 U.S. Communities. *Environmental Health Perspectives*, 119(2), 210–218.
- [5] Anenberg, S. C., Haines, S., Wang, E., Nassikas, N., & Kinney, P. L. (2020). Synergistic health effects of air pollution, temperature, and pollen exposure: a systematic review of epidemiological evidence. *Environmental Health*, 19, 130.
- [6] Arifin, Z., Falahudin, D., Saito, H., Mintarsih, T. H., Hafizt, M., & Suteja, Y. (2023). Indonesian policy and researches toward 70 *Marine Policy*, 155, 105692.
- [7] August, L. (2021). Calenviroscreen 3.0. <https://oehha.ca.gov/calenviroscreen/report/calenviroscreen-30>. Accessed: 2021-11-06.
- [8] Austin, P. C. (2017). A tutorial on multilevel survival analysis: methods, models and applications. *International Statistical Review*, 85(2), 185–203.
- [9] Bardales Cruz, M., Saikawa, E., Hengstermann, M., Ramirez, A., McCracken, J. P., & Thompson, L. M. (2023). Plastic waste generation and emissions from the domestic open burning of plastic waste in Guatemala. *Environmental Science: Atmospheres*, 3(1), 156–167.
- [10] Barkjohn, K. K., Gantt, B., & Clements, A. L. (2021). Development and application of a united states-wide correction for pm2.5 data collected with the purpleair sensor. *Atmospheric Meas. Tech.*, 14(6), 4617–4637.
- [11] Basics, A. (2021). Airnow.gov. <https://www.airnow.gov/aqi/aqi-basics>. Accessed: 2021-11-06.

- [12] Bell, A. & Jones, K. (2015). Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data. *Political Science Research and Methods*, 3(1), 133–153.
- [13] Bi, J., Carmona, N., Blanco, M. N., Gassett, A. J., Seto, E., Szpiro, A. A., Larson, T. V., Sampson, P. D., Kaufman, J. D., & Sheppard, L. (2022). Publicly available low-cost sensor measurements for pm_{2.5} exposure modeling: Guidance for monitor deployment and data selection. *Environ. Int.*, 158, 106897.
- [14] Bi, J., Stowell, J., Seto, E. Y. W., English, P. B., Al-Hamdan, M. Z., Kinney, P. L., Freedman, F. R., & Liu, Y. (2020). Contribution of low-cost sensor measurements to the prediction of pm_{2.5} levels: A case study in imperial county, california, usa. *Environ. Res.*, 180, 108810.
- [15] Bobb, J. F., Obermeyer, Z., Wang, Y., & Dominici, F. (2014). Cause-Specific Risk of Hospital Admission Related to Extreme Heat in Older Adults. *JAMA*, 312(24), 2659–2667.
- [16] Brooks, A. L., Wang, S., & Jambeck, J. R. (2018). The Chinese import ban and its impact on global plastic waste trade. *Science Advances*, 4(6), eaato131.
- [17] Buonocore, J. J., Robinson, L. A., Hammitt, J. K., & O’Keeffe, L. (2021). Estimating the potential health benefits of air quality warnings. *Risk Analysis*, 41(4), 645–660.
- [18] Burke, M., Driscoll, A., Heft-Neal, S., Xue, J., Burney, J., & Wara, M. (2021). The changing risk and burden of wildfire in the united states. *Proc. Natl. Acad. Sci.*, 118(2), e2011048118.
- [19] Carrara, N., Leurent, E., Laroche, R., Urvoy, T., Maillard, O.-A., & Pietquin, O. (2019). Budgeted reinforcement learning in continuous state space. *Advances in Neural Information Processing Systems*, 32.
- [20] Chen, K. L., Henneman, L. R. F., & Nethery, R. C. (2021). Differential impacts of COVID-19 lockdowns on PM_{2.5} across the United States. *Environmental Advances*, 6, 100122.
- [21] Chu, H.-J., Ali, M. Z., & He, Y.-C. (2020). Spatial calibration and pm_{2.5} mapping of low-cost air quality sensors. *Sci. Rep.*, 10(1), 22079.
- [22] Clarity (2022). Download - guide to maximizing your air quality budget in a post-covid world. accessed 2022-01-27.
- [23] CMS (2011). *2011 CMS Statistics*. Technical report, U.S. Department of Health and Human Services.
- [24] Colmer, J., Hardman, I., Shimshack, J., & Voorheis, J. (2020). Disparities in pm 2.5 air pollution in the united states. *Science*, 369(6503), 575–578.
- [25] Considine, E. M., Reid, C. E., Ogletree, M. R., & Dye, T. (2021). Improving accuracy of air pollution exposure measurements: Statistical correction of a municipal low-cost airborne particulate matter sensor network. *Environ. Pollut.*, 268, 115833.

- [26] Copernicus Climate Change Service (2017). ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). Accessed: 2024-09-06.
- [27] Cressie, N. & Hawkins, D. M. (1980). Robust estimation of the variogram: I. *Journal of the international Association for Mathematical Geology*, 12, 115–125.
- [28] Crilley, L. R., Singh, A., Kramer, L. J., Shaw, M. D., Alam, M. S., Apte, J. S., Bloss, W. J., Hildebrandt Ruiz, L., Fu, P., Fu, W., Gani, S., Gatari, M., Ilyinskaya, E., Lewis, A. C., Ng'ang'a, D., Sun, Y., Whitty, R. C. W., Yue, S., Young, S., & Pope, F. D. (2020). Effect of aerosol composition on the performance of low-cost optical particle counter correction factors. *Atmospheric Meas. Tech.*, 13(3), 1181–1193.
- [29] Cutler, M. J., Marlon, J. R., Howe, P. D., & Leiserowitz, A. (2018). The Influence of Political Ideology and Socioeconomic Vulnerability on Perceived Health Risks of Heat Waves in the Context of Climate Change. *Weather, Climate, and Society*, 10(4), 731–746.
- [30] Dahl, K., Licker, R., Abatzoglou, J. T., & Delet-Barreto, J. (2019). Increased frequency of and population exposure to extreme heat index days in the United States during the 21st century. *Environmental Research Communications*, 1(7), 075002.
- [31] Delaine, F., Lebental, B., & Rivano, H. (2019). In situ calibration algorithms for environmental sensor networks: A review. *IEEE Sens. J.*, 19(15), 5968–5978.
- [32] deSouza, P. & Kinney, P. L. (2021). On the distribution of low-cost pm2.5 sensors in the us: Demographic and air quality associations. *J. Expo. Sci. Environ. Epidemiol.*, 31(3), 514–524.
- [33] deSouza, P. N. (2022). Key concerns and drivers of low-cost air quality sensor use. *Sustainability*, 14(1), 584.
- [34] Dey, T., Tyagi, P., Sabath, M. B., Kamareddine, L., Henneman, L., Braun, D., & Dominici, F. (2021). Counterfactual time series analysis of short-term change in air pollution following the covid-19 state of emergency in the united states. *Scientific Reports*, 11(1), 23517.
- [35] Di, Q. (2019). An ensemble-based model of pm2.5 concentration across the contiguous united states with high spatiotemporal resolution. *Environ. Int.*, 13.
- [36] Di, Q., Dai, L., Wang, Y., Zanobetti, A., Choirat, C., Schwartz, J. D., & Dominici, F. (2017a). Association of short-term exposure to air pollution with mortality in older adults. *JAMA*, 318(24), 2446–2456.
- [37] Di, Q., Wang, Y., Zanobetti, A., Wang, Y., Koutrakis, P., Choirat, C., Dominici, F., & Schwartz, J. D. (2017b). Air pollution and mortality in the medicare population. *N. Engl. J. Med.*, 376(26), 2513–2522.

- [38] Diggle, P. J. & Giorgi, E. (2019). Preferential sampling of exposure levels. In *Handbook of Environmental Statistics* (pp. 477–490). CRC Press.
- [39] D’Antoni, D., Auyeung, V., Walton, H., Fuller, G. W., Grieve, A., & Weinman, J. (2019). The effect of evidence and theory-based health advice accompanying smartphone air quality alerts on adherence to preventative recommendations during poor air quality days: A randomised controlled trial. *Environmental International*, 124, 216–235.
- [40] Ebi, K. L., Capon, A., Berry, P., Broderick, C., de Dear, R., Havenith, G., Honda, Y., Kovats, R. S., Ma, W., Malik, A., Morris, N. B., Nybo, L., Seneviratne, S. I., Vanos, J., & Jay, O. (2021). Hot weather and heat extremes: health risks. *The Lancet*, 398(10301), 698–708.
- [41] Ebi, K. L., Teisberg, T. J., Kalkstein, L. S., Robinson, L., & Weiher, R. F. (2004). Heat Watch/Warning Systems Save Lives: Estimated Costs and Benefits for Philadelphia 1995–98. *Bulletin of the American Meteorological Society*, 85(8), 1067–1074.
- [42] Economic Research Service, U. D. o. A. (2022). What is rural? <https://www.ers.usda.gov/topics/rural-economy-population/rural-classifications/what-is-rural.aspx#:~:>. Accessed: 2022-01-03.
- [43] Efroni, Y., Foster, D. J., Misra, D., Krishnamurthy, A., & Langford, J. (2022). Sample-efficient reinforcement learning in the presence of exogenous information. In *Conference on Learning Theory* (pp. 5062–5127).
- [44] Efroni, Y., Misra, D., Krishnamurthy, A., Agarwal, A., & Langford, J. (2024). Provable rl with exogenous distractors via multistep inverse dynamics. *ICML Workshop on Reinforcement Learning Theory*.
- [45] Eimer, T., Benjamins, C., & Lindauer, M. (2022). Hyperparameters in Contextual RL are Highly Situational. arXiv:2212.10876 [cs].
- [46] English, P. B., Olmedo, L., Bejarano, E., Lugo, H., Murillo, E., Seto, E., Wong, M., King, G., Wilkie, A., Meltzer, D., Carvlin, G., Jerrett, M., & Northcross, A. (2017). The imperial county community air monitoring network: A model for community-based environmental monitoring for public health action. *Environ. Health Perspect.*, 125(7), 074501.
- [47] Equation, T. A. (2021). Air quality and aqi info. airnow discussion forum. <https://forum.airnowtech.org/t/the-aqi-equation/169>. Accessed: 2021-11-06.
- [48] Errett, N. A., Hartwell, C., Randazza, J. M., Nori-Sarma, A., Weinberger, K. R., Spangler, K. R., Sun, Y., Adams, Q. H., Wellenius, G. A., & Hess, J. J. (2023). Survey of extreme heat public health preparedness plans and response activities in the most populous jurisdictions in the United States. *BMC public health*, 23(1), 811.

- [49] European Commission, Joint Research Centre (2023). EDGAR - Emissions Database for Global Atmospheric Research: v8.1. Accessed: 2024-10-23.
- [50] for Education Statistics, N. C. (2021). School locations & geoassignments 2020-2021, education demographic and geographic estimates. <https://nces.ed.gov/programs/edge/geographic/schoollocations>. Accessed: 2021-11-06.
- [51] for Environment, E. C. & Health, B. O. (2005). *Effect of Air Pollution on Children's Health and Development: A Review of the Evidence*. World Health Organization, Europe.
- [52] Frank, J., Mannor, S., & Precup, D. (2008). Reinforcement learning in the presence of rare events. In *Proceedings of the 25th international conference on Machine learning* (pp. 336–343).
- [53] Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by browne and draper). *Bayesian Analysis*, 1(3), 515–534.
- [54] Global Maritime Traffic Density Service (2021). GMTDS Data. <https://www.globalmaritimetraffic.org>. A service of MapLarge. Accessed ship loitering data on 2024-09-05.
- [55] Global Plastic Action Partnership (2020). *Radically Reducing Plastic Pollution in Indonesia: A Multistakeholder Action Plan National Plastic Action Partnership*. Technical report, World Economic Forum. Accessed: 2024-10-08.
- [56] Gordon, J. N. D., Bilsback, K. R., Fiddler, M. N., Pokhrel, R. P., Fischer, E. V., Pierce, J. R., & Bililign, S. (2023). The Effects of Trash, Residential Biofuel, and Open Biomass Burning Emissions on Local and Transported PM_{2.5} and Its Attributed Mortality in Africa. *GeoHealth*, 7(2), e2022GH000673. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1029/2022GH000673>.
- [57] Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18–27. Big Remotely Sensed Data: tools, applications and experiences.
- [58] Goss, M., Swain, D. L., Abatzoglou, J. T., Sarhadi, A., Kolden, C. A., Williams, A. P., & Diffenbaugh, N. S. (2020). Climate change is increasing the likelihood of extreme autumn wildfire conditions across california. *Environ. Res. Lett.*, 15(9), 094016.
- [59] Grainger, C. & Schreiber, A. (2019). Discrimination in ambient air pollution monitoring? *AEA Pap. Proc.*, 109, 277–282.
- [60] Gressent, A., Malherbe, L., Colette, A., Rollin, H., & Scimia, R. (2020). Data fusion for air quality mapping using low-cost sensor observations: Feasibility and added-value. *Environ. Int.*, 143, 105965.

- [61] Guin, S. & Bhatnagar, S. (2023). A policy gradient approach for finite horizon constrained markov decision processes. In *2023 62nd IEEE Conference on Decision and Control (CDC)* (pp. 3353–3359).: IEEE.
- [62] Gupta, P., Doraiswamy, P., Levy, R., Pikelnaya, O., Maibach, J., Feenstra, B., Polidori, A., Kiros, F., & Mills, K. C. (2018). Impact of california fires on local and regional air quality: The role of a low-cost sensor network and satellite observations. *GeoHealth*, 2(6), 172–181.
- [63] Hammer, M. S., van Donkelaar, A., Li, C., Lyapustin, A., Sayer, A. M., Hsu, N. C., Levy, R. C., Garay, M. J., Kalashnikova, O. V., Kahn, R. A., Brauer, M., Apte, J. S., Henze, D. K., Zhang, L., Zhang, Q., Ford, B., Pierce, J. R., & Martin, R. V. (2020). Global Estimates and Long-Term Trends of Fine Particulate Matter Concentrations (1998–2018). *Environmental Science & Technology*, 54(13), 7879–7890. Publisher: American Chemical Society.
- [64] Han, D., Mulyana, B., Stankovic, V., & Cheng, S. (2023). A Survey on Deep Reinforcement Learning Algorithms for Robotic Manipulation. *Sensors (Basel, Switzerland)*, 23(7), 3762.
- [65] Harpole, A., Zingale, M., Hawke, I., & Chegini, T. (2019). pyro: a framework for hydrodynamics explorations and prototyping. *Journal of Open Source Software*, 4(34), 1265.
- [66] Hawkins, M. D., Brown, V., & Ferrell, J. (2017). Assessment of NOAA National Weather Service Methods to Warn for Extreme Heat Events. *Weather, Climate, and Society*, 9(1), 5–13.
- [67] Heffernan, C., Koehler, K., Zamora, M. L., Buehler, C., Gentner, D. R., Peng, R. D., & Datta, A. (2024). A causal machine-learning framework for studying policy impact on air pollution: A case-study in COVID-19 lockdowns. *American Journal of Epidemiology*, (pp. kwae171). Epub ahead of print.
- [68] Heffernan, C., Peng, R., Gentner, D. R., Koehler, K., & Datta, A. (2022). Gaussian process filtering for calibration of low-cost air-pollution sensor network data. Submitted 2022-03-28, ArXiv220314775, accessed 2022-05-06.
- [69] Heo, S., Bell, M. L., & Lee, J.-T. (2019). Comparison of health risks by heat wave definition: Applicability of wet-bulb globe temperature for heat wave criteria. *Environmental Research*, 168, 158–170.
- [70] Hernán, M. A. & Robins, J. M. (2020). *Causal Inference: What If*. Boca Raton: Chapman & Hall/CRC.
- [71] Hettinger, G., Lee, Y., & Mitra, N. (2025). Multiply robust difference-in-differences estimation of causal effect curves for continuous exposures. *Biometrics*, 81(1), ujafo15.
- [72] Heuillet, A., Couthouis, F., & Díaz-Rodríguez, N. (2021). Explainability in deep reinforcement learning. *Knowledge-Based Systems*, 214, 106685.

- [73] Hondula, D. M., Meltzer, S., Balling, R. C., & Iñiguez, P. (2022). Spatial Analysis of United States National Weather Service Excessive Heat Warnings and Heat Advisories. *Bulletin of the American Meteorological Society*, 103(9), E2017–E2031. Publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society.
- [74] Huang, K., Bi, J., Meng, X., Geng, G., Lyapustin, A., Lane, K. J., Gu, D., Kinney, P. L., & Liu, Y. (2019). Estimating daily pm_{2.5} concentrations in new york city at the neighborhood-scale: Implications for integrating non-regulatory measurements. *Sci. Total Environ.*, 697, 134094.
- [75] IQAir (2022). Live animated air quality map. accessed 2022-11-03.
- [76] Islam, M. R., Welker, J., Salam, A., & Stone, E. A. (2022). Plastic Burning Impacts on Atmospheric Fine Particulate Matter at Urban and Rural Sites in the USA and Bangladesh. *ACS Environmental Au*, 2(5), 409–417.
- [77] Jayaratne, R., Liu, X., Ahn, K.-H., Asumadu-Sakyi, A., Fisher, G., Gao, J., Mabon, A., Maza-heri, M., Mullins, B., Nyakcrillu, M., Ristovski, Z., Scorgie, Y., Thai, P., Dunbabin, M., & Morawska, L. (2020). Low-cost pm_{2.5} sensors: An assessment of their suitability for various applications. *Aerosol Air Qual. Res.*
- [78] Jbaily, A., Zhou, X., Liu, J., Lee, T.-H., Kamareddine, L., Verguet, S., & Dominici, F. (2022). Air pollution exposure disparities across us population and income groups. *Nature*, 601(7892), 228–233.
- [79] Johnson, K., Holder, A., Frederick, S., & Clements, A. (2020). Purpleair pm_{2.5} u.s. correction and performance during smoke events. https://cfpub.epa.gov/si/si_public_record_Report.cfm?dirEntryId=349513&Lab=CEMM. Accessed: 2021-11-06.
- [80] Karasik, R., Virdin, J., & Wilson, J. (2023). Plastics Policy Inventory. <https://nicholasinstitute.duke.edu/plastics-policy-inventory/>. Accessed: 2024-10-30.
- [81] Keller, J. P. & Peng, R. D. (2019). Error in estimating area-level air pollution exposures for epidemiology. *Environmetrics*, 30(8).
- [82] Kelp, M., Lin, S., Kutz, J. N., & Mickley, L. J. (2022). A new approach for determining optimal placement of pm_{2.5} air quality sensors: Case study for the contiguous united states. *Environmental Research Letters*, 17(3), 034034.
- [83] Kennedy, E. H. (2022). Semiparametric Doubly Robust Targeted Double Machine Learning: A Review. *arXiv preprint arXiv:2203.06469*.
- [84] Kennedy, E. H., Ma, Z., McHugh, M. D., & Small, D. S. (2017). Nonparametric methods for doubly robust estimation of continuous treatment effects. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(4), 1229–1245.

- [85] Kruse, C., Boyda, E., Chen, S., Karra, K., Bou-Nahra, T., Hammer, D., Mathis, J., Maddalene, T., Jambeck, J., & Laurier, F. (2023). Satellite Monitoring of Terrestrial Plastic Waste. *PLOS ONE*, 18(1), e0278997. arXiv:2204.01485 [cs, eess].
- [86] Kurisu, D., Kato, K., & Shao, X. (2023). Gaussian Approximation and Spatially Dependent Wild Bootstrap for High-Dimensional Spatial Data. *Journal of the American Statistical Association*, (pp. 1–13).
- [87] Laber, E. B., Wu, F., Munera, C., Lipkovich, I., Colucci, S., & Ripa, S. (2018). Identifying optimal dosage regimes under safety constraints: An application to long term opioid treatment of chronic pain. *Statistics in Medicine*, 37(9), 1407–1418. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/sim.7566>.
- [88] Lee, J. N., Agarwal, A., Dann, C., & Zhang, T. (2023). Learning in POMDPs is Sample-Efficient with Hindsight Observability. arXiv:2301.13857 [cs, stat].
- [89] Lee, S., Okui, R., & Whang, Y.-J. (2017). Doubly robust uniform confidence band for the conditional average treatment effect function. *Journal of Applied Econometrics*, 32(7), 1207–1225.
- [90] Levine, A., Stone, P., & Zhang, A. (2024). Multistep inverse is not all you need. *Workshop on Reinforcement Learning Beyond Rewards at the Reinforcement Learning Conference*.
- [91] Levine, S., Kumar, A., Tucker, G., & Fu, J. (2020). Offline Reinforcement Learning: Tutorial, Review, and Perspectives on Open Problems. arXiv:2005.01643 [cs, stat]. arXiv: 2005.01643.
- [92] Li, Y., Zheng, Y., & Yang, Q. (2018). Dynamic Bike Reposition: A Spatio-Temporal Reinforcement Learning Approach. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 1724–1733). London United Kingdom: ACM.
- [93] Liao, P., Greenewald, K., Klasnja, P., & Murphy, S. (2020). Personalized HeartSteps: A Reinforcement Learning Algorithm for Optimizing Physical Activity. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(1), 18:1–18:22.
- [94] Liu, J. C., Wilson, A., Mickley, L. J., Dominici, F., Ebisu, K., Wang, Y., Sulprizio, M. P., Peng, R. D., Yue, X., Son, J.-Y., Anderson, G. B., & Bell, M. L. (2017). Wildfire-specific fine particulate matter and risk of hospital admissions in urban and rural counties. *Epidemiology*, 28(1), 77–85.
- [95] Liu, X., Jayaratne, R., Thai, P., Kuhn, T., Zing, I., Christensen, B., Lamont, R., Dunbabin, M., Zhu, S., Gao, J., Wainwright, D., Neale, D., Kan, R., Kirkwood, J., & Morawska, L. (2020). Low-cost sensors as an alternative for long-term air quality monitoring. *Environ. Res.*, 185, 109438.

- [96] Lu, T., Bechle, M. J., Wan, Y., Presto, A. A., & Hankey, S. (2022). Using crowd-sourced low-cost sensors in a land use regression of pm_{2.5} in 6 us cities. *Air Qual. Atmosphere Health*, 15(4), 667–678.
- [97] Lu, Y., Giuliano, G., & Habre, R. (2021). Estimating hourly pm_{2.5} concentrations at the neighborhood scale using a low-cost air sensor network: A los angeles case study. *Environ. Res.*, 195, 110653.
- [98] Malings, C., Tanzer, R., Hauryliuk, A., Kumar, S. P. N., Zimmerman, N., Kara, L. B., Presto, A. A., & Subramanian, R. (2019). Development of a general calibration model and long-term performance evaluation of low-cost sensors for air pollutant gas monitoring. *Atmospheric Measurement Techniques*, 12(2), 903–920.
- [99] Malings, C., Tanzer, R., Hauryliuk, A., Saha, P. K., Robinson, A. L., Presto, A. A., & Subramanian, R. (2020). Fine particle mass monitoring with low-cost sensors: Corrections and long-term performance evaluation. *Aerosol Sci. Technol.*, 54(2), 160–174.
- [100] Masselot, P., Chebana, F., Campagna, C., Lavigne, E., Ouarda, T. B., & Gosselin, P. (2021). Machine Learning Approaches to Identify Thresholds in a Heat-Health Warning System Context. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 184(4), 1326–1346.
- [101] Metzger, K. B., Ito, K., & Matte, T. D. (2010). Summer Heat and Mortality in New York City: How Hot Is Too Hot? *Environmental Health Perspectives*, 118(1), 80–86.
- [102] Microsoft AI for Good Research Lab (2021). U.S. Broadband Usage Percentages Dataset.
- [103] Miller, D. (2021). National park service. accessed 2022-01-03.
- [104] MIT Election Data and Science Lab (2018). County Presidential Election Returns 2000-2020.
- [105] Morawska, L., Thai, P. K., Liu, X., Asumadu-Sakyi, A., Ayoko, G., Bartonova, A., Bedini, A., Chai, F., Christensen, B., Dunbabin, M., Gao, J., Hagler, G. S. W., Jayaratne, R., Kumar, P., Lau, A. K. H., Louie, P. K. K., Mazaheri, M., Ning, Z., Motta, N., Mullins, B., Rahman, M. M., Ristovski, Z., Shafiei, M., Tjondronegoro, D., Westerdahl, D., & Williams, R. (2018). Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: How far have they gone? *Environ. Int.*, 116, 286–299.
- [106] Mu, T., Theocharous, G., Arbour, D., & Brunskill, E. (2021). Constraint Sampling Reinforcement Learning: Incorporating Expertise For Faster Learning. arXiv:2112.15221 [cs].
- [107] Nahum-Shani, I., Smith, S. N., Spring, B. J., Collins, L. M., Witkiewitz, K., Tewari, A., & Murphy, S. A. (2017). Just-in-Time Adaptive Interventions (JITAs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support. *Annals of Behavioral Medicine: A Publication of the Society of Behavioral Medicine*, 52(6), 446–462.

- [108] Narayanan, S., Lage, I., & Doshi-Velez, F. (2022). (When) Are Contrastive Explanations of Reinforcement Learning Helpful? arXiv:2211.07719 [cs].
- [109] NASA FIRMS (2020). Viirs Active Fire (375m S-NPP, 2012-2019). https://firms.modaps.eosdis.nasa.gov/active_fire/. Accessed: 2024-06-18.
- [110] National Institute of Environmental Health Sciences (2024). Air pollution and your health. Accessed: October 22, 2024.
- [111] Neidell, M. & Kinney, P. L. (2010). Estimates of the association between ozone and asthma hospitalizations that account for behavioral responses to air quality information. *Environmental Science & Policy*, 13(2), 97–103.
- [112] Ng, C. F. S., Ueda, K., Takeuchi, A., Nitta, H., Konishi, S., Bagrowicz, R., Watanabe, C., & Takami, A. (2014). Sociogeographic Variation in the Effects of Heat and Cold on Daily Mortality in Japan. *Journal of Epidemiology*, 24(1), 15–24.
- [113] Nobre, W. S., Schmidt, A. M., & Pereira, J. B. (2021). On the effects of spatial confounding in hierarchical models. *International Statistical Review*, 89(2), 302–322.
- [114] OECD (2022). Plastic Pollution is Growing Relentlessly as Waste Management and Recycling Fall Short. <https://www.oecd.org/en/about/news/press-releases/2022/02/plastic-pollution-is-growing-relentlessly-as-waste-management-and-recycling-fall-short.html>. Accessed: 2024-10-30.
- [115] OpenAQ Team (2020). *Open Air Quality Data: The Global State of Play*. Technical report, OpenAQ. Accessed: 2025-04-08.
- [116] Orru, H., Ebi, K. L., & Forsberg, B. (2017). The interplay of climate change and air pollution on health. *Curr. Environ. Health Rep.*, 4(4), 504–513.
- [117] Padakandla, S. (2021). A survey of reinforcement learning algorithms for dynamically varying environments. *ACM Computing Surveys (CSUR)*, 54(6), 1–25.
- [118] Pathak, G., Nichter, M., Hardon, A., & Moyer, E. (2024). The Open Burning of Plastic Wastes is an Urgent Global Health Issue. *Annals of Global Health*, 90(1), 3.
- [119] Pathak, G., Nichter, M., Hardon, A., Moyer, E., Latkar, A., Simbaya, J., Pakasi, D., Taqueban, E., & Love, J. (2023). Plastic pollution and the open burning of plastic wastes. *Global Environmental Change*, 80, 102648.
- [120] Pearl, J. (2012). The mediation formula: A guide to the assessment of causal pathways in nonlinear models. *Causality: Statistical perspectives and applications*, (pp. 151–179).
- [121] Puiutta, E. & Veith, E. M. (2020). Explainable Reinforcement Learning: A Survey. arXiv:2005.06247 [cs, stat].

- [122] Puterman, M. L. (2014). *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons.
- [123] Raffin, A., Hill, A., Gleave, A., Kanervisto, A., Ernestus, M., & Dormann, N. (2021). Stable-baselines3: Reliable reinforcement learning implementations. *Journal of Machine Learning Research*, 22(268), 1–8.
- [124] Raniga, K., Davitt, A., Lewis, C., Sridhar, L., Gans, L., & McCormick, G. (2023). Waste Sector: Estimating CH₄ Emissions from Solid Waste Disposal Sites. Emissions modeling by WattTime.
- [125] Ray, A., Achiam, J., & Amodei, D. (2019). *Benchmarking Safe Exploration in Deep Reinforcement Learning*. Technical report, OpenAI.
- [126] Romoff, J., Henderson, P., Piche, A., Francois-Lavet, V., & Pineau, J. (2018). Reward estimation for variance reduction in deep reinforcement learning. In *Conference on Robot Learning* (pp. 674–699): PMLR.
- [127] Rubin, D. B. (1980). Randomization analysis of experimental data: The fisher randomization test comment. *Journal of the American Statistical Association*, 75(371), 591–593.
- [128] Saberian, S., Heyes, A., & Rivers, N. (2017). Alerts Work! Air Quality Warnings and Cycling. *Resources, Energy and Economics*, 49, 165–185.
- [129] Sardegna, C. (2021). Github. https://github.com/ReagentX/purple_air_api. Accessed: 2021-04-16.
- [130] Schuler, A. & van der Laan, M. (2023). *Introduction to Modern Causal Inference*. Available at <https://alejandroschuler.github.io/mci>.
- [131] Scott, S. R., Hailemariam, P. E., Bhave, P. V., Bergin, M. H., & Carlson, D. E. (2023). Identifying Waste Burning Plumes Using High-Resolution Satellite Imagery and Machine Learning: A Case Study in the Maldives. *Environmental Science & Technology Letters*, 10(8), 642–648. Publisher: American Chemical Society.
- [132] Shin, M., Kang, Y., Park, S., Im, J., Yoo, C., & Quackenbush, L. J. (2020). Estimating ground-level particulate matter concentrations using satellite-based data: a review. *GIScience & Remote Sensing*, 57(2), 174–189.
- [133] Sinclair, S. R., Vieira Frujeri, F., Cheng, C.-A., Marshall, L., Barbalho, H. D. O., Li, J., Neville, J., Menache, I., & Swaminathan, A. (2023). Hindsight learning for MDPs with exogenous inputs. In A. Krause, E. Brunskill, K. Cho, B. Engelhardt, S. Sabato, & J. Scarlett (Eds.), *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research* (pp. 31877–31914): PMLR.

- [134] Stuart, N. A., Hartfield, G., Schultz, D. M., Wilson, K., West, G., Hoffman, R., Lackmann, G., Brooks, H., Roebber, P., Bals-Elsholz, T., Obermeier, H., Judt, F., Market, P., Nietfeld, D., Telfeyan, B., DePodwin, D., Fries, J., Abrams, E., & Shields, J. (2022). The Evolving Role of Humans in Weather Prediction and Communication. *Bulletin of the American Meteorological Society*, 103(8), E1720–E1746. Publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society.
- [135] Sutton, R. S. & Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. The MIT Press, second edition.
- [136] Sutton, R. S., McAllester, D., Singh, S., & Mansour, Y. (1999). Policy gradient methods for reinforcement learning with function approximation. *Advances in neural information processing systems*, 12.
- [137] Tanzer, R., Malings, C., Hauryliuk, A., Subramanian, R., & Presto, A. A. (2019). Demonstration of a low-cost multi-pollutant network to quantify intra-urban spatial variations in air pollutant source impacts and to evaluate environmental justice. *Int. J. Environ. Res. Public Health*, 16(14), 2523.
- [138] Tatem, A. J. (2017). WorldPop, open data for spatial demography. *Scientific Data*, 4(1), 170004.
- [139] Tec, M., Duan, Y., & Müller, P. (2023). A comparative tutorial of bayesian sequential design and reinforcement learning. *The American Statistician*, 77(2), 223–233.
- [140] Tec, M., Trisovic, A., Audirac, M., Woodward, S., Hu, J. K., Khoshnevis, N., & Dominici, F. (2024). Space: The spatial confounding environment. In *International Conference in Learning Representations (ICLR)*.
- [141] Tessum, C. W., Apte, J. S., Goodkind, A. L., Muller, N. Z., Mullins, K. A., Paoella, D. A., Polasky, S., Springer, N. P., Thakrar, S. K., Marshall, J. D., & Hill, J. D. (2019). Inequity in consumption of goods and services adds to racial–ethnic disparities in air pollution exposure. *Proc. Natl. Acad. Sci.*, 116(13), 6001–6006.
- [142] The Boards of Trustees of the Federal Hospital Insurance and Federal Supplementary Medical Insurance Trust Funds (2023). *2023 Annual Report*. Technical report, Centers for Medicare & Medicaid Services.
- [143] The Commonwealth Fund (2021). Medicare Advantage vs. Traditional Medicare: How Do Beneficiaries’ Characteristics and Experiences Differ? <https://www.commonwealthfund.org/publications/issue-briefs/2021/oct/medicare-advantage-vs-traditional-medicare-beneficiaries-differ>. Accessed: 2024-01-01.

- [144] Tomczak, M., Swaroop, S., & Turner, R. (2020). Efficient Low Rank Gaussian Variational Inference for Neural Networks. In *Advances in Neural Information Processing Systems*, volume 33 (pp. 4610–4622).: Curran Associates, Inc.
- [145] Towers, M., Terry, J. K., Kwiatkowski, A., Balis, J. U., Cola, G. d., Deleu, T., Goulão, M., Kallinteris, A., KG, A., Krimmel, M., Perez-Vicente, R., Pierré, A., Schulhoff, S., Tai, J. J., Shen, A. T. J., & Younis, O. G. (2023). Gymnasium.
- [146] Tsoukalas, I., Kossieris, P., & Makropoulos, C. (2020). Simulation of Non-Gaussian Correlated Random Variables, Stochastic Processes and Random Fields: Introducing the anySim R-Package for Environmental Applications and Beyond. *Water*, 12(6).
- [147] Tzikas, D. G., Likas, A. C., & Galatsanos, N. P. (2008). The variational approximation for bayesian inference. *IEEE Signal Processing Magazine*, 25(6), 131–146.
- [148] Uehara, M., Shi, C., & Kallus, N. (2022). A Review of Off-Policy Evaluation in Reinforcement Learning. arXiv:2212.06355 [cs, math, stat].
- [149] Unfried, K. & Wang, F. (2022). Importing Air Pollution? Evidence from China’s Plastic Waste Imports.
- [150] United Nations Environment Programme (UNEP) (2024). Addressing Global Plastic Pollution: A Timeline Leading up to the Intergovernmental Negotiating Committee. Accessed: 2024-10-23.
- [151] U.S. Census Bureau (2014). 2009-2013 American Community Survey 5-year County-level Estimates of Population and Median Household Income.
- [152] U.S. Department of Transportation, F. H. A. (2021). National highway planning network. <https://www.fhwa.dot.gov/planning/processes/tools/nhpn/index.cfm>. Accessed: 2021-11-06.
- [153] U.S. Energy Information Administration (2020). Climate Zones - DOE Building America Program.
- [154] U.S. Environmental Protection Agency (2014). How to use air sensors: Air sensor guidebook. Accessed 2021-11-14.
- [155] U.S. Environmental Protection Agency (2021). Airdata website file download page. https://aqs.epa.gov/aqsweb/airdata/download_files.html. Accessed: 2021-11-14.
- [156] U.S. Environmental Protection Agency, Air Quality Planning Unit (2022). How Does PM Affect Human Health? accessed 2022-05-06.

- [157] van der Waa, J., van Diggelen, J., Bosch, K. v. d., & Neerincx, M. (2018). Contrastive Explanations for Reinforcement Learning in terms of Expected Consequences. *arXiv:1807.08706* [cs, stat].
- [158] van Donkelaar, A., Hammer, M. S., Bindle, L., Brauer, M., Brook, J. R., Garay, M. J., Hsu, N. C., Kalashnikova, O. V., Kahn, R. A., Lee, C., Levy, R. C., Lyapustin, A., Sayer, A. M., & Martin, R. V. (2021). Monthly Global Estimates of Fine Particulate Matter and Their Uncertainty. *Environmental Science & Technology*, 55(22), 15287–15300. Publisher: American Chemical Society.
- [159] Walker, K. (2022). Tidycensus. <https://cran.r-project.org/web/packages/tidycensus/index.html>.
- [160] Wang, W.-C. V., Lung, S.-C. C., & Liu, C.-H. (2020). Application of machine learning for the in-field correction of a pm2.5 low-cost sensor network. *Sensors*, 20(17), 5002.
- [161] Wang, Y. & Karasik, R. (2022). *Plastic Pollution Policy Country Profile: Indonesia*. Technical Report NI PB 22-05, Duke University, Durham, NC.
- [162] Wang, Y., Miller, A. C., & Blei, D. M. (2019). Comment: Variational autoencoders as empirical bayes. *Statistical Science*, 34(2), 229–233.
- [163] Watson, J. G., Chow, J. C., DuBois, D., Green, M., & Frank, N. (1997). *Guidance for the Network Design and Optimum Site Exposure for PM_{2.5} and PM₁₀*. EPA-454/R-99/022 PB-99-157513/XAB, Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC (United States); Nevada Univ. System, Desert Research Inst., Reno, NV (United States); National Oceanic and Atmospheric Administration, Las Vegas, NV (United States). Accessed 2022-01-03.
- [164] Weinberger, K. R., Wu, X., Sun, S., Spangler, K. R., Nori-Sarma, A., Schwartz, J., Requia, W., Sabath, B. M., Braun, D., Zanobetti, A., Dominici, F., & Wellenius, G. A. (2021). Heat warnings, mortality, and hospital admissions among older adults in the United States. *Environment International*, 157, 106834.
- [165] Weinberger, K. R., Zanobetti, A., Schwartz, J., & Wellenius, G. A. (2018). Effectiveness of National Weather Service heat alerts in preventing mortality in 20 US cities. *Environment International*, 116, 30–38.
- [166] Wen, X.-J., Balluz, L., & Mokdad, A. (2009). Association between media alerts of air quality index and change of outdoor activity among adult asthma in six states, brfss, 2005. *Journal of Community Health*, 34(1), 40–46.
- [167] Wiedinmyer, C., Yokelson, R. J., & Gullett, B. K. (2014). Global Emissions of Trace Gases, Particulate Matter, and Hazardous Air Pollutants from Open Burning of Domestic Waste. *Environmental Science & Technology*, 48(16), 9523–9530. Publisher: American Chemical Society.

- [168] Williams, R., Duvall, R., Kilaru, V., Hagler, G., Hassinger, L., Benedict, K., Rice, J., Kaufman, A., Judge, R., Pierce, G., Allen, G., Bergin, M., Cohen, R. C., Fransioli, P., Gerboles, M., Habre, R., Hannigan, M., Jack, D., Louie, P., Martin, N. A., Penza, M., Polidori, A., Subramanian, R., Ray, K., Schauer, J., Seto, E., Thurston, G., Turner, J., Wexler, A. S., & Ning, Z. (2019). Deliberating performance targets workshop: Potential paths for emerging pm_{2.5} and o₃ air sensor progress. *Atmospheric Environ. X*, 2, 100031.
- [169] Wu, Q., Chen, X., Zhou, Z., Chen, L., & Zhang, J. (2021). Deep Reinforcement Learning With Spatio-Temporal Traffic Forecasting for Data-Driven Base Station Sleep Control. *IEEE/ACM Transactions on Networking*, 29(2), 935–948.
- [170] Wu, X., Mealli, F., Kioumourtzoglou, M.-A., Dominici, F., & Braun, D. (2024). Matching on generalized propensity scores with continuous exposures. *Journal of the American Statistical Association*, 119(545), 757–772. PMID: 38524247.
- [171] Wu, X., Weinberger, K. R., Wellenius, G. A., Dominici, F., & Braun, D. (2023). Assessing the causal effects of a stochastic intervention in time series data: are heat alerts effective in preventing deaths and hospitalizations? *Biostatistics*.
- [172] Xu, H., Zhan, X., & Zhu, X. (2022). Constraints Penalized Q-learning for Safe Offline Reinforcement Learning. arXiv:2107.09003 [cs].
- [173] Yosephine, L. (2016). U.S. embassy, Jakarta cooperate to monitor air quality. *The Jakarta Post*.
- [174] Zajonc, T. (2012). Bayesian Inference for Dynamic Treatment Regimes: Mobility, Equity, and Efficiency in Student Tracking. *Journal of the American Statistical Association*, 107(497), 80–92.
- [175] Zanobetti, A., O'Neill, M. S., Gronlund, C. J., & Schwartz, J. D. (2013). Susceptibility to Mortality in Weather Extremes: Effect Modification by Personal and Small Area Characteristics In a Multi-City Case-Only Analysis. *Epidemiology (Cambridge, Mass.)*, 24(6), 809–819.
- [176] Zusman, M., Schumacher, C. S., Gassett, A. J., Spalt, E. W., Austin, E., Larson, T. V., Carvlin, G., Seto, E., Kaufman, J. D., & Sheppard, L. (2020). Calibration of low-cost particulate matter sensors: Model development for a multi-city epidemiological study. *Environ. Int.*, 134, 105329.



Appendix to Chapter I

A.1 SUPPLEMENTAL FIGURES AND TABLES FOR ALL ANALYSES

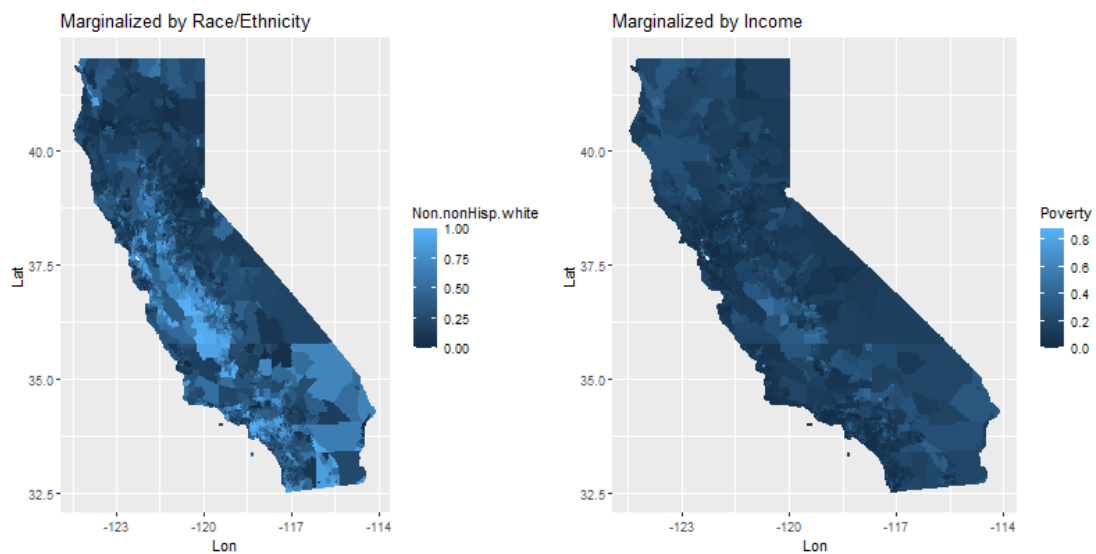


Figure A.1: Distributions of marginalized groups considered in the analyses: percent of individuals identifying as non-white or Hispanic, and percent of households with income below the poverty line. Data source: 2016 ACS³.

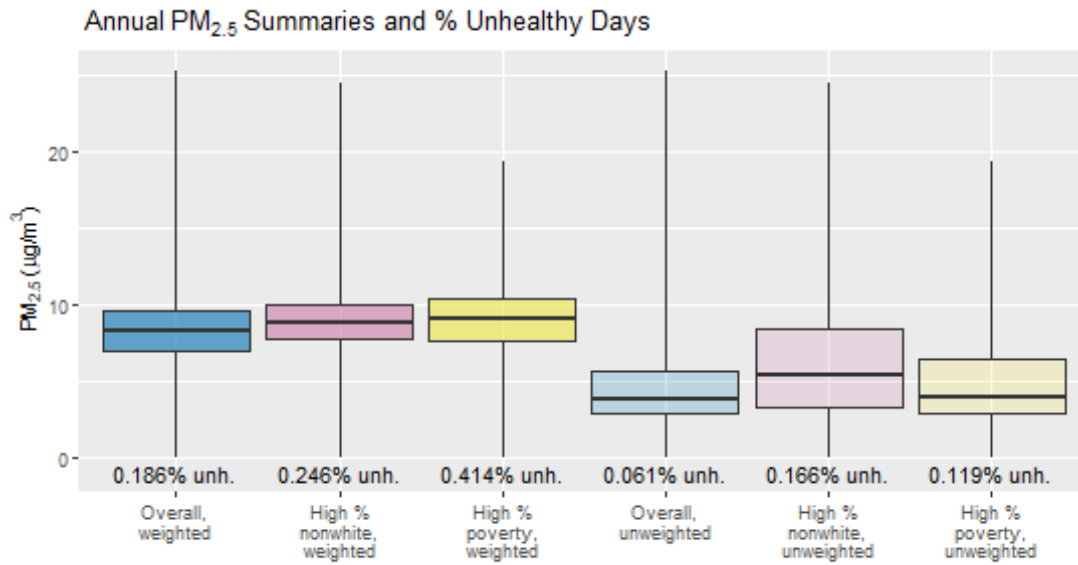


Figure A.2: Annual average PM_{2.5} ($\mu\text{g}/\text{m}^3$) summaries for population subsets, weighted and unweighted by population density; labeled with % Unhealthy (AQI classification = Orange and higher) days.

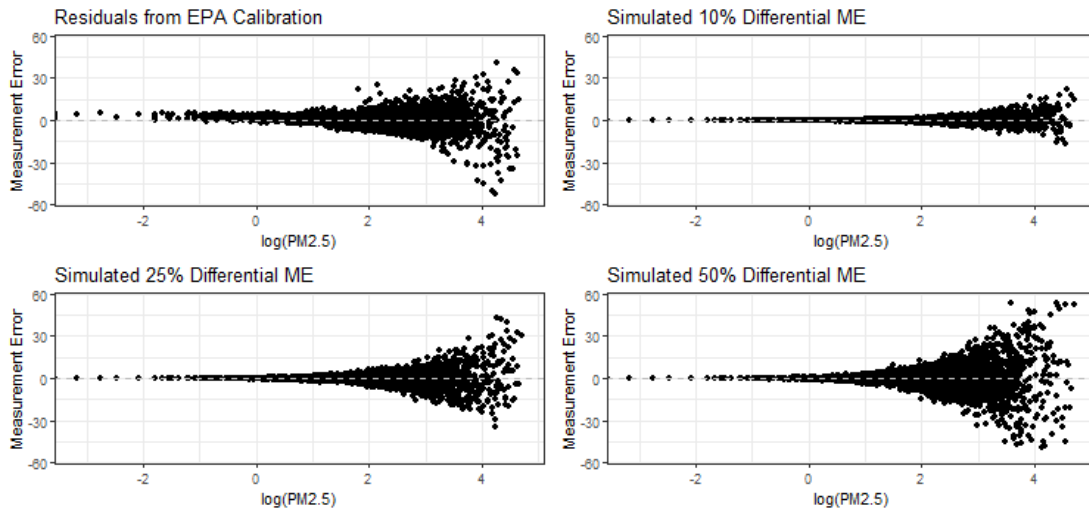


Figure A.3: Simulated and empirical sensor measurement error. Top left plot shows the empirical distribution of residuals in $\mu\text{g}/\text{m}^3$ (compared to AQS monitor measurements within 50m) from PurpleAir observations corrected using the EPA equation. The other three plots show the distributions resulting from our method of simulating sensor measurement error differentially with respect to true PM_{2.5} concentration. Note that 50% accuracy was not included in the full analysis simulations.

Placement strategy	Data source	Data processing	LCS location selection method (within each simulation)	Additional notes
Current PurpleAir Locations	The purpleair Python package ¹²⁹	Assigned each outdoor PurpleAir sensor to its nearest 1 km x 1 km grid centroid	Randomly selected grids with PurpleAir	4,343 grids with PurpleAir LCS
Public School Locations	The National Center for Education Statistics ⁵⁰	Assigned each school to its nearest 1 km x 1 km grid centroid	Randomly selected grids with schools	7,548 grids with schools
Near-Road	The National Highway Planning Network shapefile ¹⁵²	Summed the lengths of major roads within circular buffers with 500 m radii around each grid centroid	Randomly sampled from all grids, where the sampling probability for each grid was the sum of road lengths	4.5% of the grids had major roads; among these, the average length was 2km
CES Score	Cal Enviro Screen 3.0 ⁷	Assigned all grids the CES Score of the census tract their centroid falls within	Randomly sampled from all grids, where the sampling probability for each grid was the CES Score	<i>Additional notes are in the main text</i>
Pollution Score	Cal Enviro Screen 3.0 ⁷	Assigned all grids the Pollution Score of the census tract their centroid falls within	Randomly sampled from all grids, where the sampling probability for each grid was the Pollution Score	<i>Additional notes are in the main text</i>

Table A.1: Describing the data sets, data processing steps, and sampling methods used to select hypothetical locations for LCS in each simulation.

County	Population size (2016)	Avg. CES score	Purple Air	Schools	Near road locations	CES score based	Pollution score based
Sacramento	1,495,611	23	26	40	10	7	7
Imperial	186,019	38	1	7	26	46	34
Los Angeles	10,180,169	23	107	204	30	25	26

Table A.2: Average number of LCS (out of 1,000 deployments) placed in the counties of Sacramento, Imperial, and Los Angeles according to each placement strategy considered in this study.

Decile	Minimum	Q1	Median	Mean	Q3	Maximum
1	0.000	0.684	1.265	1.262	1.863	2.455
2	2.455	2.757	3.074	3.083	3.404	3.750
3	3.750	4.055	4.375	4.389	4.716	5.083
4	5.083	5.385	5.711	5.731	6.066	6.458
5	6.458	6.780	7.136	7.167	7.538	8.000
6	8.000	8.365	8.778	8.818	9.251	9.796
7	9.796	10.225	10.705	10.747	11.246	11.875
8	11.88	12.45	13.15	13.25	13.99	15.08
9	15.08	15.93	17.10	17.60	18.91	22.65
10	22.65	24.24	26.57	29.49	31.02	112.18

Table A.3: Summary statistics of the Di et al. estimates in each decile of AQS PM_{2.5} from our correction dataset, in $\mu\text{g}/\text{m}^3$. Matching on these deciles was used to draw empirical residuals from the EPA correction applied to collocated PurpleAir and AQS data.

A.2 SUPPLEMENTAL FIGURES AND TABLES FOR THE ANALYSIS WEIGHTED BY POPULATION DENSITY

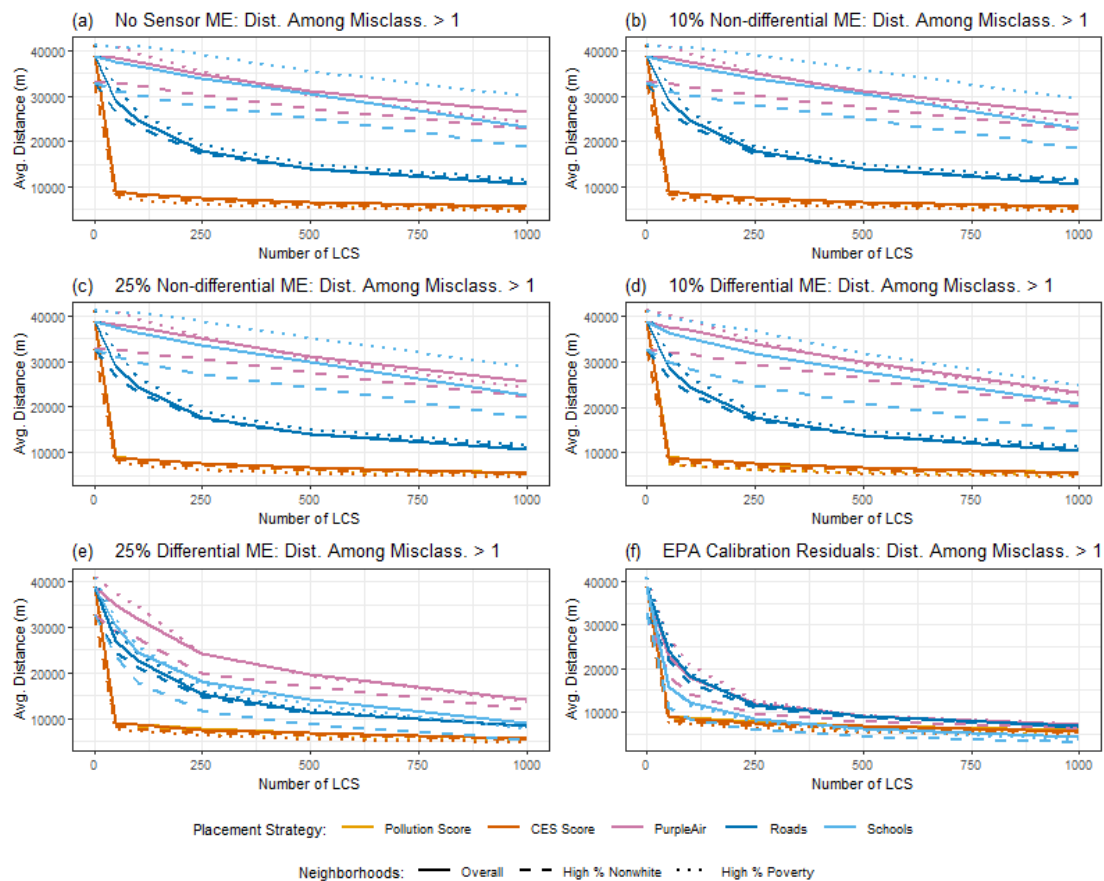


Figure A.4: Distance to the nearest AQ instrument, among observations misclassified by more than one level of the AQI (e.g. Red to Yellow, or Green to Orange). All results were calculated using 366 days and averaged across 100 simulation replicates, weighted by population density.

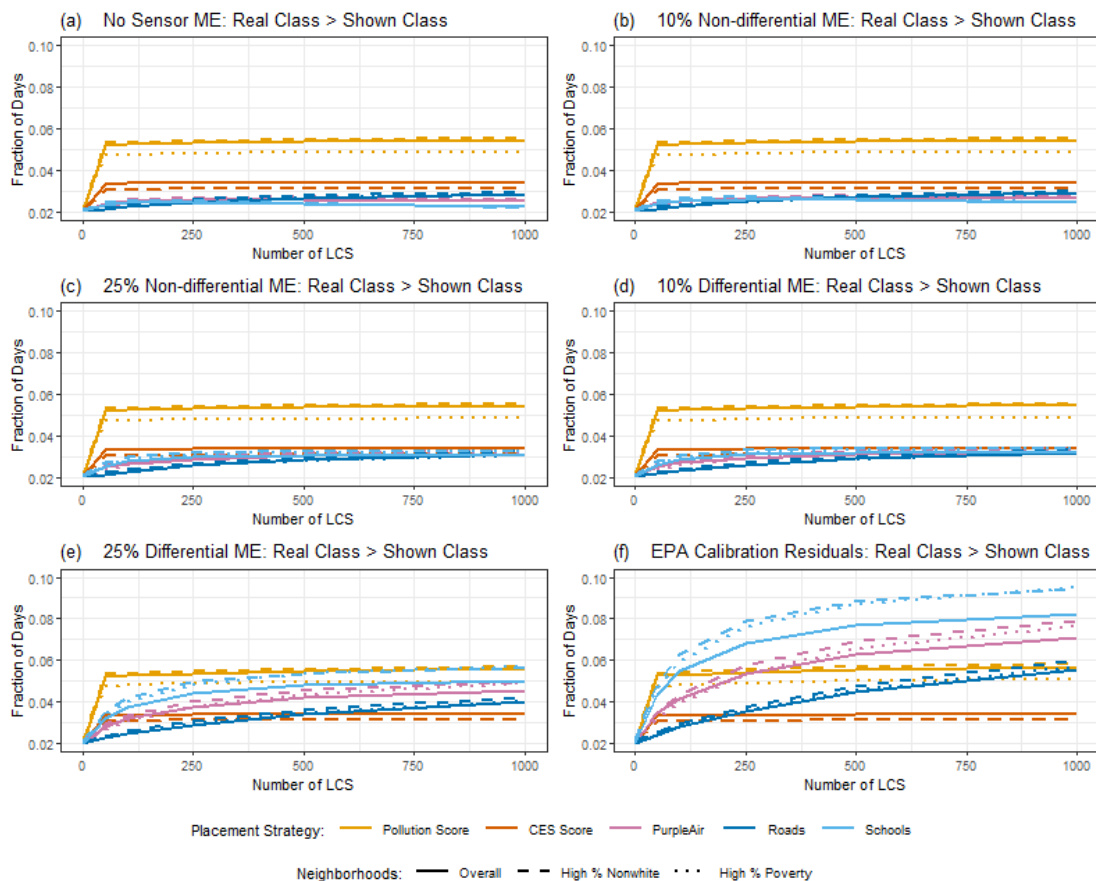


Figure A.5: Underclassifications of the AQI. The average fraction of days on which the real AQI class (what was truly experienced) was higher than the shown AQI class (reported from the nearest monitor or sensor), resulting from different numbers of LCS deployed, LCS placement strategies, and sensor measurement error types and amounts. All results were calculated using 366 days and averaged across 100 simulation replicates, weighted by population density.

Locations	Annual Avg. PM _{2.5} ($\mu\text{g}/\text{m}^3$)	% Poverty	CES Score	% Non- white or Hispanic	Population Density (un- weighted)
CA overall	0 (1)	0 (1)	0 (1)	0 (1)	0 (1)
EPA monitor sites	0.34 (1.01)	0.67 (1)	0.62 (1.06)	0.22 (0.96)	2.54 (3.99)
PurpleAir sites	-0.24 (0.83)	-0.33 (0.83)	-0.46 (0.9)	-0.26 (0.93)	2.51 (4.11)
School sites	0.23 (0.89)	0.17 (1)	0.23 (1.02)	0.33 (0.93)	3.96 (4.28)
Favoring by nearby roads (not wtd. by pop. density)	-1.37 (1.23)	0.08 (0.75)	-0.29 (0.81)	-0.93 (0.96)	0.3 (1.75)
Favoring by roads, where pop. density > 500	-0.28 (1.06)	-0.08 (1)	-0.16 (0.97)	-0.3 (1)	2.87 (3.67)
Favoring by Pollution Score (not wtd. by pop. density)	-1.7 (1.26)	0.17 (0.67)	-0.1 (0.79)	-0.81 (0.93)	0.02 (1.07)
Favoring by Pollution Score, where pop. density > 500	0 (1.03)	-0.08 (1)	0.03 (1.04)	-0.15 (0.96)	2.81 (3.48)
Favoring by CES Score (not wtd. by pop. density)	-1.67 (1.31)	0.25 (0.75)	0.11 (0.8)	-0.67 (0.96)	0.03 (1.16)
Favoring by CES Score, where pop. density > 500	0.12 (1.06)	0.17 (1.08)	0.41 (1.04)	0.11 (0.93)	3.13 (3.79)
High % nonwhite	0.31 (0.9)	0.33 (1)	0.49 (0.93)	0.78 (0.44)	0.29 (1.83)
High % poverty	0.37 (1.09)	1.5 (0.75)	0.92 (0.9)	0.7 (0.78)	0.02 (1.26)

Table A.4: Basic descriptive statistics, weighted by population density (unless otherwise specified) and normalized, to provide a sense of scale. Non-normalized results are in Table A.7. All cells are of the form “Mean (Standard Deviation)”. Note: wtd. = weighted. The rows that specify when population density greater than 500 / sq. mile is to account for both the weighting to select the LCS locations (i.e., by proximity to roads, Pollution Score, or CES Score) and the weighting by population density in our results. Population density less than 500 / sq. mile is considered rural by the USDA⁴².

Locations	Annual Avg. PM _{2.5} ($\mu\text{g}/\text{m}^3$)	% Poverty	CES Score	% Non- white or Hispanic	Population Density (un- weighted)
CA overall	8.23 (1.98)	0.16 (0.12)	28.11 (15.93)	0.61 (0.27)	249.81 (1,526.04)
EPA monitor sites	8.91 (2)	0.24 (0.12)	37.98 (16.95)	0.67 (0.26)	4,119.08 (6,090.22)
PurpleAir sites	7.76 (1.65)	0.12 (0.1)	20.84 (14.33)	0.54 (0.25)	4,075.79 (6,267.25)
School sites	8.69 (1.77)	0.18 (0.12)	31.74 (16.28)	0.70 (0.25)	6,296.3 (6,531.5)
Favoring by nearby roads (not wtd. by pop. density)	5.52 (2.43)	0.17 (0.09)	23.42 (12.96)	0.36 (0.26)	701.74 (2,670.86)
Favoring by nearby roads, where pop. density > 500	7.68 (2.1)	0.15 (0.12)	25.63 (15.5)	0.53 (0.27)	4,627.07 (5,604.17)
Favoring by Pollution Score (not wtd. by pop. density)	4.87 (2.49)	0.18 (0.08)	26.52 (12.51)	0.39 (0.25)	275.26 (1,633.18)
Favoring by Pollution Score, where pop. density > 500	8.23 (2.04)	0.15 (0.12)	28.51 (16.57)	0.57 (0.26)	4,535.05 (5,309.65)
Favoring by CES Score (not wtd. by pop. density)	4.93 (2.6)	0.19 (0.09)	29.88 (12.71)	0.43 (0.26)	296.21 (1,774.24)
Favoring by CES Score, where pop. density > 500	8.46 (2.09)	0.18 (0.13)	34.68 (16.54)	0.64 (0.25)	5,027.12 (5,779.83)
High % nonwhite	8.84 (1.78)	0.20 (0.12)	35.92 (14.87)	0.82 (0.12)	696.8 (2,789.15)
High % poverty	8.96 (2.16)	0.34 (0.09)	42.81 (14.28)	0.80 (0.21)	277.93 (1,922)

Table A.5: Non-normalized version of previous table. Basic descriptive statistics, weighted by population density (unless otherwise specified). All cells are of the form “Mean (Standard Deviation)”. Note: wtd. = weighted. The rows that specify when population density greater than 500 / sq. mile is to account for both the weighting to select the LCS locations (i.e. by proximity to roads, Pollution Score, or CES Score) and the weighting by population density in our results. Population density less than 500 / sq. mile is considered rural.

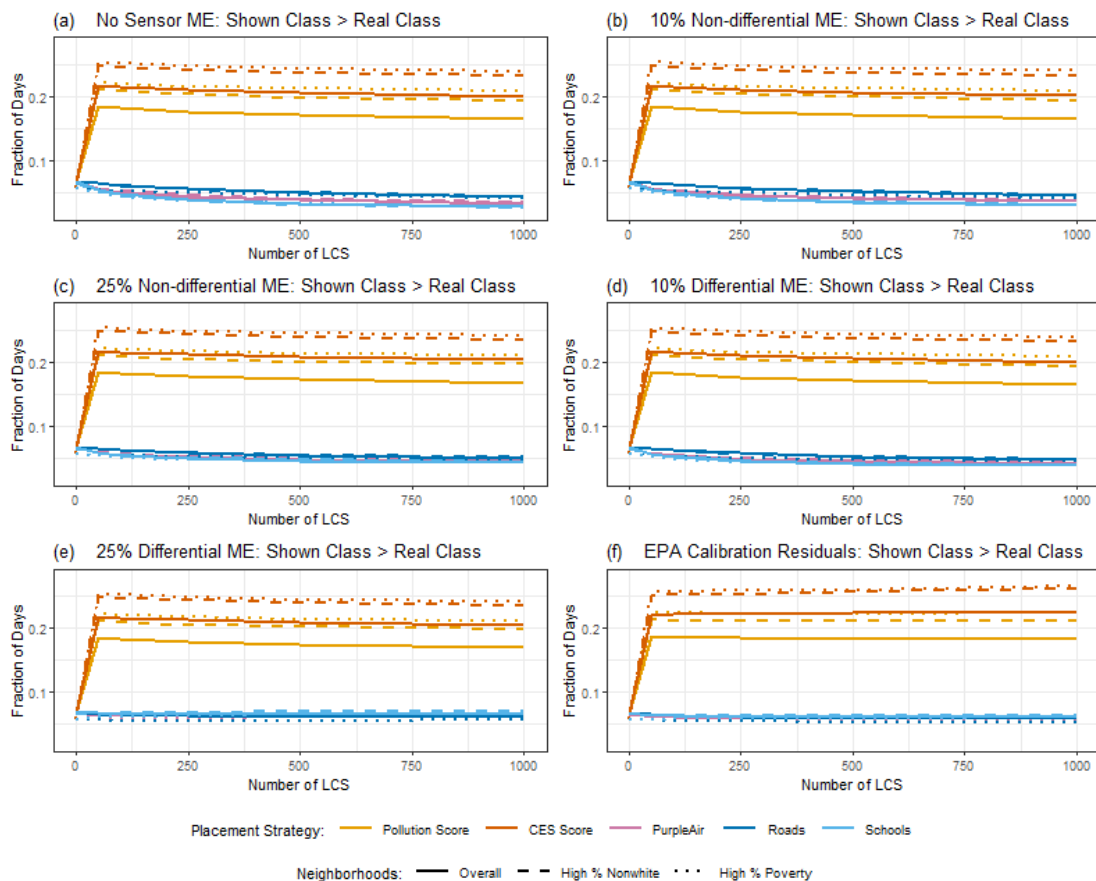


Figure A.6: Overclassifications of the AQI. The average fraction of days on which the shown AQI class (reported from the nearest monitor or sensor) was higher than the real AQI class (what was truly experienced), resulting from different numbers of LCS deployed, LCS placement strategies, and sensor measurement error types and amounts. All results were calculated using 366 days and averaged across 100 simulation replicates, weighted by population density.

Type/Amount of Sensor Measurement Error	Std. Dev. of Error ($\mu\text{g}/\text{m}^3$)	MAE ($\mu\text{g}/\text{m}^3$)	95th Pct. of Errors ($\mu\text{g}/\text{m}^3$)	Under-classified (%)	Over-classified (%)	UHM
Overall Population						
No LCS (only EPA monitors)	—	1.46	4.45	2.05	6.79	11.37
No Sensor Error	0	0.53	2.30	1.32	1.48	8.40
10% Non-differential	0.5	0.76	2.42	1.88	2.18	9.74
25% Non-differential	1.25	1.24	3.20	2.79	3.80	11.82
10% Differential	0.99	0.97	2.92	2.90	3.23	18.41
25% Differential	2.46	1.84	5.31	5.10	6.48	29.22
EPA Correction Residual	3.70	2.51	7.22	8.72	6.07	28.14
Decile Draws						
Population living in CBGs with high % nonwhite						
No LCS (only EPA monitors)	—	1.34	4.14	2.15	6.53	10.35
No Sensor Error	0	0.39	1.81	1.11	1.25	7.22
10% Non-differential	0.5	0.65	1.97	1.81	2.11	8.71
25% Non-differential	1.25	1.16	2.94	2.93	3.95	11.09
10% Differential	1.03	0.91	2.69	3.06	3.34	18.32
25% Differential	2.56	1.88	5.42	5.75	6.92	29.44
EPA Correction Residual	3.83	2.59	7.49	10.03	6.39	28.50
Decile Draws						
Population living in CBGs with high % poverty						
No LCS (only EPA monitors)	—	1.28	4.06	2.09	5.79	8.40
No Sensor Error	0	0.42	1.97	1.14	1.22	5.15
10% Non-differential	0.5	0.68	2.11	1.87	2.08	6.55
25% Non-differential	1.25	1.18	3.01	3.03	3.85	8.75
10% Differential	1.08	0.95	2.83	3.17	3.29	15.88
25% Differential	2.71	1.93	5.60	6.00	6.72	27.52
EPA Correction Residual	4.03	2.64	7.66	10.42	6.22	26.47
Decile Draws						

Table A.6: Comparing impacts of different sensor accuracies for LCS at all schools (similar to Table 1.1 in the main text), weighted by population density. Results when there are no LCS vs. LCS at all school locations ($n = 7,548$), assuming different kinds and amounts of sensor measurement error. Results are the average of 100 simulation replicates to account for randomness in the sensor measurement error generation. Unless otherwise specified, "errors" refer to the difference between the true exposure experienced at each grid centroid and the exposure reported from the nearest AQ instrument. AQI under-classification is when the true exposure class is greater than what someone is shown, and over-classification is when the true exposure class is less than what someone is shown. Rate of UH misclassification (UHM) is the fraction of days with unhealthy AQI (Orange+) that are misreported as healthy AQI (Green or Yellow).

A.3 SUPPLEMENTAL FIGURES AND TABLES FOR THE ANALYSIS NOT WEIGHTED BY POPULATION DENSITY

Locations	Annual Avg. PM _{2.5} ($\mu\text{g}/\text{m}^3$)	% Poverty	CES Score	% Non- white or Hispanic	Population Density (un- weighted)
CA overall	4.55 (2.32)	0.17 (0.08)	23.92 (11.95)	0.35 (0.24)	249.81 (1,526.04)
EPA monitor sites	7.43 (2.58)	0.20 (0.13)	29.23 (15.8)	0.53 (0.26)	4,119.08 (6,090.22)
PurpleAir sites	6.63 (2.05)	0.11 (0.09)	16.47 (12.32)	0.38 (0.24)	4,075.79 (6,267.25)
School sites	8.04 (2.08)	0.16 (0.11)	27.88 (16.04)	0.58 (0.27)	6,296.3 (6,531.5)
Favoring by nearby roads	5.52 (2.43)	0.17 (0.09)	23.42 (12.96)	0.36 (0.26)	701.74 (2,670.86)
Favoring by Pollution Score	4.87 (2.49)	0.18 (0.08)	26.52 (12.51)	0.39 (0.25)	275.26 (1,633.18)
Favoring by CES Score	4.93 (2.6)	0.19 (0.09)	29.88 (12.71)	0.43 (0.26)	296.21 (1,774.24)
High % nonwhite	5.77 (2.93)	0.24 (0.09)	38.31 (10.84)	0.74 (0.12)	696.8 (2,789.15)
High % poverty	4.85 (2.66)	0.30 (0.06)	31.6 (13.5)	0.52 (0.26)	277.93 (1,922)

Table A.7: Basic descriptive statistics, unweighted by population density (counterpart of Table A.5). All cells are of the form "Mean (Standard Deviation)". Note: wtd. = weighted.

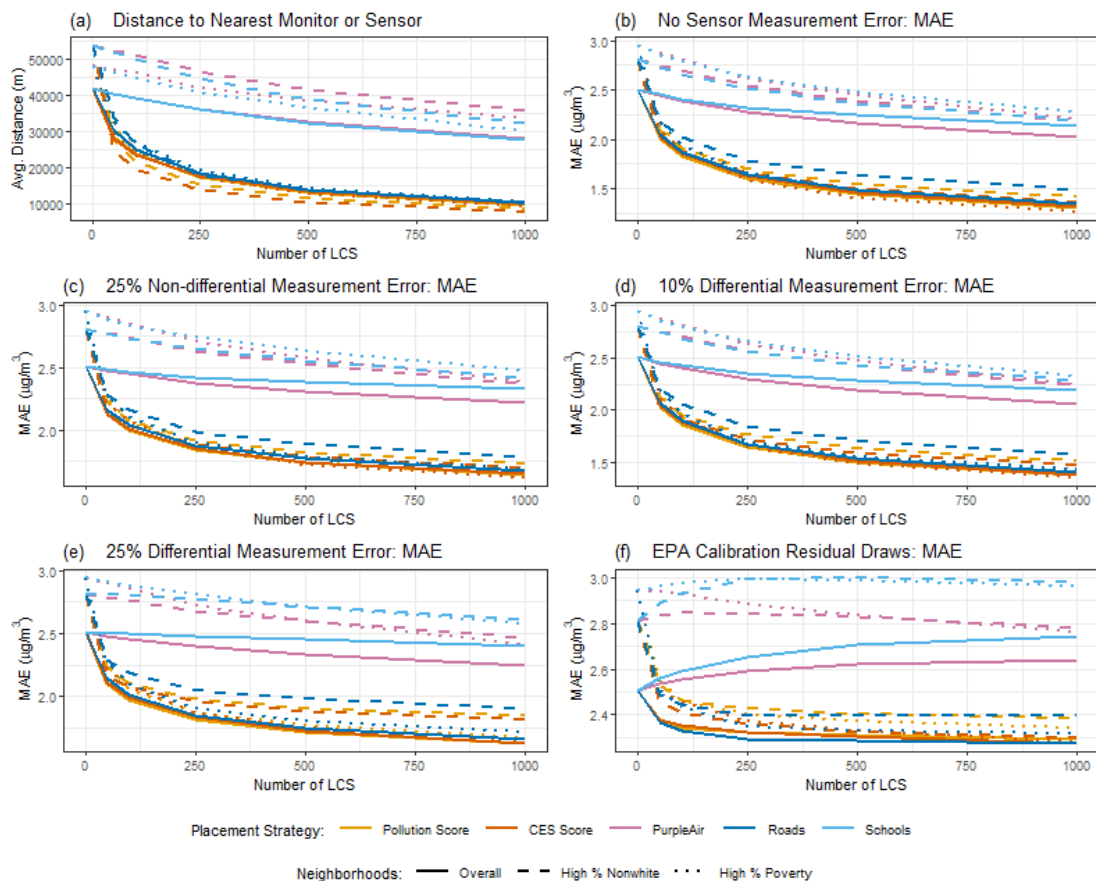


Figure A.7: Distance and MAE, unweighted by population density. Distance to the nearest monitor or sensor and mean absolute error (between what is reported vs. experienced) resulting from different numbers of LCS deployed, LCS placement strategies, and sensor measurement error types and amounts. All results were calculated using 366 days and averaged across 100 simulation replicates, unweighted by population density.

Type/Amount of Sensor Measurement Error	Std. Dev. of Error ($\mu\text{g}/\text{m}^3$)	MAE ($\mu\text{g}/\text{m}^3$)	95th Pct. of Errors ($\mu\text{g}/\text{m}^3$)	Under-classified (%)	Over-classified (%)	UHM
Overall Population						
No LCS (only EPA monitors)	—	2.51	8.08	1.17	7.09	17.67
No Sensor Error	0	1.79	5.87	1.36	3.59	25.95
10% Non-differential	0.5	1.85	5.92	1.39	3.66	25.50
25% Non-differential	1.25	2.08	6.19	1.47	4.09	25.42
10% Differential	0.78	1.86	5.99	1.51	3.84	26.88
25% Differential	1.95	2.12	6.68	1.85	4.86	32.45
EPA Correction Residual	3.06	2.68	7.34	2.62	4.61	31.19
Decile Draws						
Population living in CBGs with high % nonwhite						
No LCS (only EPA monitors)	—	2.81	9.43	1.58	10.56	9.86
No Sensor Error	0	1.79	5.60	2.11	3.76	21.41
10% Non-differential	0.5	1.84	5.65	2.19	3.86	20.69
25% Non-differential	1.25	2.07	5.95	2.37	4.43	20.86
10% Differential	0.91	1.88	5.75	2.44	4.14	22.84
25% Differential	2.27	2.23	6.65	3.12	5.51	29.00
EPA Correction Residual	3.43	2.72	7.56	4.50	5.26	27.94
Decile Draws						
Population living in CBGs with high % poverty						
No LCS (only EPA monitors)	—	2.95	9.40	1.11	9.63	9.89
No Sensor Error	0	1.83	5.88	1.58	3.11	16.82
10% Non-differential	0.5	1.89	5.93	1.63	3.17	15.73
25% Non-differential	1.25	2.11	6.21	1.73	3.64	15.97
10% Differential	0.88	1.90	6.00	1.78	3.39	18.04
25% Differential	2.20	2.18	6.70	2.20	4.56	24.47
EPA Correction Residual	3.38	2.70	7.38	3.11	4.42	23.74
Decile Draws						

Table A.8: Comparing impacts of different sensor accuracies, unweighted by population density (counterpart of Table 1.1 in the main text). Results when there are no LCS vs. LCS at all real PurpleAir locations ($n = 4,343$), assuming different kinds and amounts of sensor measurement error. Results are the average of 100 simulation replicates to account for randomness in the sensor measurement error generation. Unless otherwise specified, “errors” refer to the difference between the true exposure experienced at each grid centroid and the exposure reported from the nearest AQ instrument. AQI under-classification is when the true exposure class is greater than what someone is shown, and over-classification is when the true exposure class is less than what someone is shown. Rate of UH misclassification (UHM) is the fraction of days with unhealthy AQI (Orange+) that are misreported as healthy AQI (Green or Yellow).

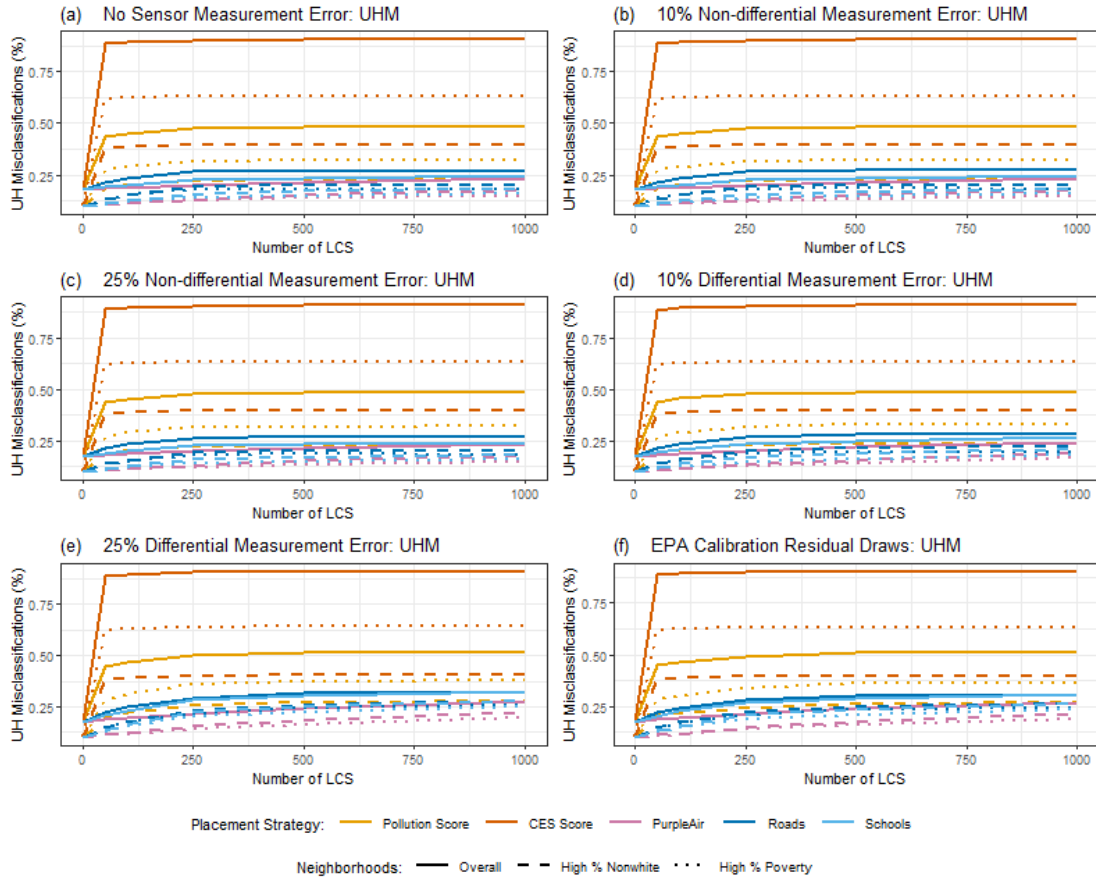


Figure A.8: Unhealthy-Healthy Misclassifications, unweighted by population density. Rates of UHMs resulting from different numbers of LCS deployed, LCS placement strategies, and sensor measurement error types and amounts. UH misclassification occurs when the air quality is unhealthy (Orange+) but is reported as healthy (Green or Yellow); the UHM rate is calculated by dividing the fraction of UH misclassifications by the total fraction of unhealthy days experienced by each group. All results were calculated using 366 days and averaged across 100 simulation replicates, unweighted by population density.

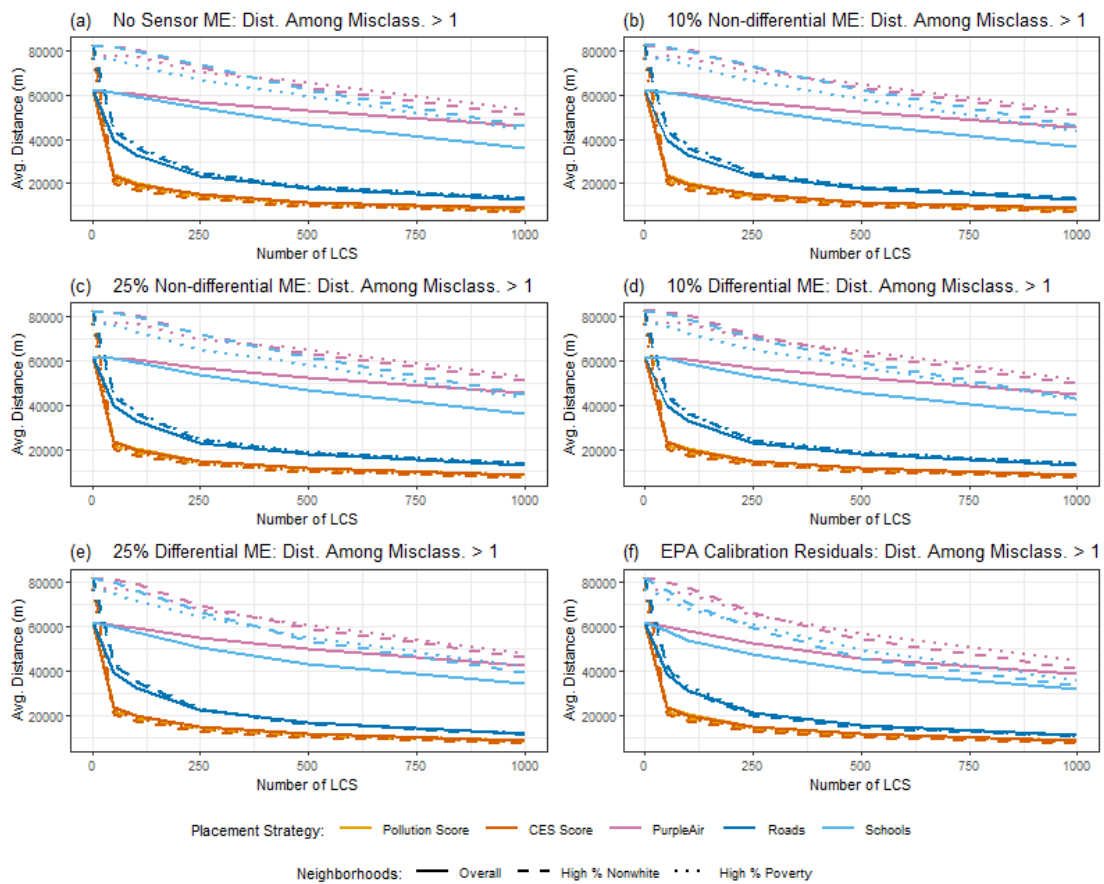


Figure A.9: Distance to the nearest monitor or sensor, among observations misclassified by more than one level of the AQI (e.g. Red to Yellow, or Green to Orange). All results were calculated using 366 days and averaged across 100 simulation replicates, unweighted by population density.

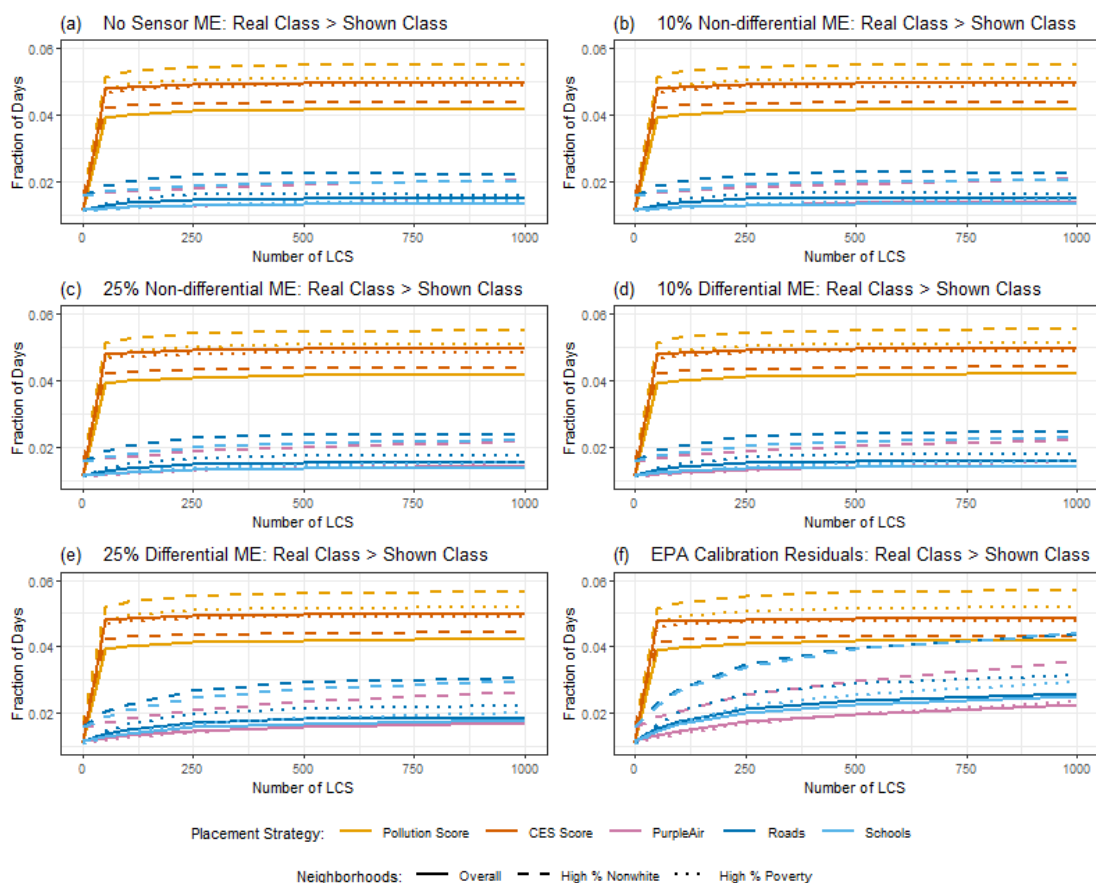


Figure A.10: Underclassifications of the AQI, unweighted by population density. The average fraction of days on which the real AQI class (what was truly experienced) was higher than the shown AQI class (reported from the nearest monitor or sensor), resulting from different numbers of LCS deployed, LCS placement strategies, and sensor measurement error types and amounts. All results were calculated using 366 days and averaged across 100 simulation replicates, unweighted by population density.

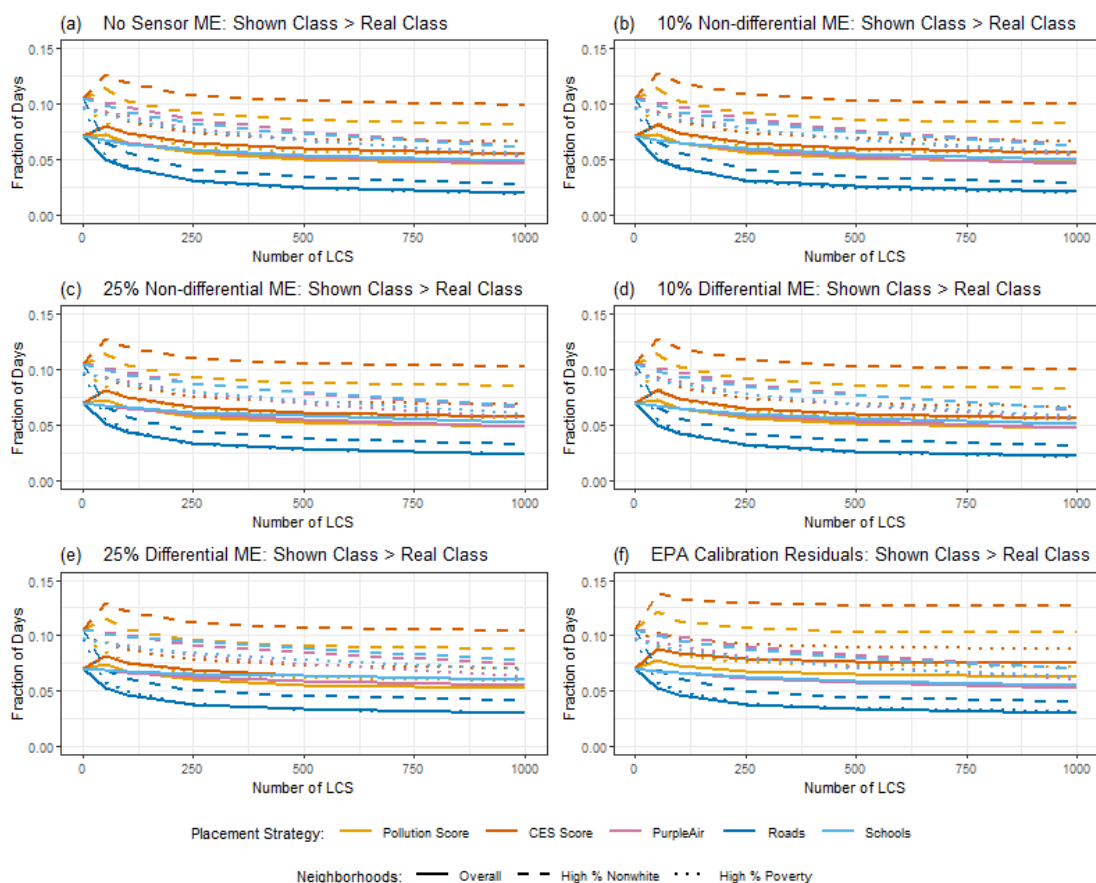


Figure A.11: Overclassifications of the AQI, unweighted by population density. The average fraction of days on which the shown AQI class (reported from the nearest monitor or sensor) was higher than the real AQI class (what was truly experienced), resulting from different numbers of LCS deployed, LCS placement strategies, and sensor measurement error types and amounts. All results were calculated using 366 days and averaged across 100 simulation replicates, unweighted by population density.

Type/Amount of Sensor Measurement Error	Std. Dev. of Error ($\mu\text{g}/\text{m}^3$)	MAE ($\mu\text{g}/\text{m}^3$)	95th Pct. of Errors ($\mu\text{g}/\text{m}^3$)	Under-classified (%)	Over-classified (%)	UHM
Overall Population						
No LCS (only EPA monitors)	—	2.51	8.08	1.17	7.09	17.67
No Sensor Error	0	1.78	5.74	1.35	3.40	21.91
10% Non-differential	0.5	1.84	5.80	1.38	3.50	21.91
25% Non-differential	1.25	2.08	6.10	1.48	3.96	22.16
10% Differential	0.93	1.86	5.87	1.53	3.70	25.05
25% Differential	2.33	2.14	6.63	1.91	4.83	32.76
EPA Correction Residual	3.54	2.71	7.34	2.78	4.55	31.36
Decile Draws						
Population living in CBGs with high % nonwhite						
No LCS (only EPA monitors)	—	2.81	9.43	1.58	10.56	9.86
No Sensor Error	0	1.72	5.60	1.95	3.65	15.16
10% Non-differential	0.5	1.80	5.66	2.04	3.84	15.57
25% Non-differential	1.25	2.06	5.98	2.29	4.53	16.10
10% Differential	1.01	1.86	5.77	2.38	4.18	19.82
25% Differential	2.53	2.28	6.79	3.24	5.83	29.46
EPA Correction Residual	3.79	2.81	7.76	4.98	5.37	27.90
Decile Draws						
Population living in CBGs with high % poverty						
No LCS (only EPA monitors)	—	2.95	9.40	1.11	9.63	9.89
No Sensor Error	0	1.75	5.79	1.46	3.20	12.63
10% Non-differential	0.5	1.82	5.84	1.51	3.33	12.97
25% Non-differential	1.25	2.07	6.13	1.65	3.87	13.42
10% Differential	1.03	1.84	5.93	1.71	3.57	16.97
25% Differential	2.58	2.17	6.75	2.25	4.91	27.11
EPA Correction Residual	3.86	2.76	7.50	3.39	4.63	25.37
Decile Draws						

Table A.9: Comparing impacts of different sensor accuracies for LCS at all schools, unweighted by population density (counterpart of Table A.6). Results when there are no LCS vs. LCS at all school locations ($n = 7,548$), assuming different kinds and amounts of sensor measurement error. Results are the average of 100 simulation replicates to account for randomness in the sensor measurement error generation. Unless otherwise specified, “errors” refer to the difference between the true exposure experienced at each grid centroid and the exposure reported from the nearest AQ instrument. AQI under-classification is when the true exposure class is greater than what someone is shown, and over-classification is when the true exposure class is less than what someone is shown. Rate of UH misclassification (UHM) is the fraction of days with unhealthy AQI (Orange+) that are misreported as healthy AQI (Green or Yellow).

B

Appendix to Chapter 2

B.1 HEAT ALERTS MDP TERMINOLOGY

Table B.1 summarizes the heat alerts MDP components.

B.2 DETAILS OF THE DATA SET

Evidence from the climate & health and environmental epidemiology literature guides our selection of relevant variables as well as feature engineering.

B.2.1 HEAT INDEX AND HEAT ALERTS

We follow Weinberger et al.¹⁶⁴ and Wu et al.¹⁷¹ in using population-weighted maximum heat index to describe heat exposure in each county. They also combined NWS heat “warnings” and “advisories” into the single category of “alerts”. In some locations, where the forecast zones in which the NWS issues alerts are smaller than county boundaries, the alert history was taken from whichever forecast zone contained the largest proportion of the 2010 county population.

B.2.2 MEDICARE HOSPITALIZATIONS

Medicare is available for people aged 65 years and older. The hospitalization claims data we use in this study is from Medicare Part A, fee-for-service. A 2021 report¹⁴³ found that fee-for-service enrollees do not significantly differ from managed care enrollees (the rest of the Medicare population).

RL Term	Notation	Meaning in the Heat Alerts Application
episode	$t \in \{0, \dots, H-1\}$	the $H = 152$ days of the warm season/summer in a specific year and county; H is the horizon
state	$s_t = (\xi_t, x_t)$	a vector of time-varying factors influencing the effect of heat on hospitalizations and the effectiveness of an alert; ξ_t is the exogenous component (e.g., heat index) and x_t is the endogenous component (e.g., history of heat alerts)
action	a_t	whether a heat alert is issued (1) or not (0) at time t
reward	$r_t = R(s_t, a_t)$	daily rate of heat-related hospitalizations at time t ; a negative sign transform is applied so that fewer hospitalizations corresponds to a greater reward
agent	—	the algorithmic decision-maker that implements a policy to determine when to issue heat alerts
policy	$\pi(a_t s_t)$	the heat alert issuance rule; the NWS policy is the rule that was used to issue heat alerts in the U.S. from 2006-2016
cumulative reward / return	$\sum_{b=t}^{H-1} r_b$	sum of remaining reward (transformed hospitalizations) at time t ; unless otherwise specified, “cumulative” implies $t = 0$
intervention / alert budget	$\sum_{t=0}^{H-1} a_t \leq b$	the maximum number of heat alerts that can be issued in a given county-summer; in our experiments, the observed number of NWS alerts

Table B.1: MDP/RL terminology in the context of the heat alerts environment.

B.2.3 DESCRIPTIVE STATISTICS OF HEAT, ALERTS, AND HOSPITALIZATIONS

Table B.2 provides descriptive statistics of the heat alerts dataset, both for the 30 counties in our RL analysis and all the counties with population $\geq 65,000$ (used to create the BROACH environment). The two sets of statistics are similar aside from the number of heat alerts issued because the counties in the RL analysis were selected specifically for having issued higher numbers of alerts. Other observations are the long tail of the number of NWS alerts per county-summer, that some heat alerts were issued on days that were not very hot in an absolute or relative (quantile) sense, that there were some county-summers with few “hot” days (90th percentile by county*) and others with many, and that the distribution of day-of-summer for NWS heat alerts and 90th percentile QHI by county are quite similar. A few other relevant statistics not shown in the table are that (across all the counties) there were a total of 40,387 heat alerts issued between 2006 and 2016, that 17.2% of NWS heat alerts were issued on days with heat index values that were *not* in their county’s top 90th percentile, and that 74.0% of days in the top 90th QHI by county did not have NWS heat alerts.

*Here, the 90th percentile was chosen to provide some intuition for the distribution of heat index and heat alerts. For context, Hondula et al. (2022) found that 95th percentile daily maximum heat index was the climatological indicator most strongly correlated with NWS heat alert frequency.

Variable	Min.	Q1	Median	Mean	Q3	Max.
*NOHR hosps per 1,000	5.4 (3.1)	12.5 (12.7)	14.8 (15.2)	15 (15.3)	17.3 (17.8)	25.7 (34.5)
*No. of NWS alerts	0 (0)	4 (0)	8 (3)	9.6 (4.8)	13 (7)	43 (57)
DMHI on alert days (°F)	70.5 (68.3)	100.3 (99.7)	103.5 (103.2)	102.5 (102.6)	106.2 (105.9)	121.1 (144)
Q-DMHI on alert days	18.9 (15.6)	91.6 (92.6)	95.8 (96.9)	93.4 (94.2)	98.2 (98.9)	100 (100)
*No. of 90th pct. HI days	0 (0)	9 (8)	14 (14)	15.4 (15.4)	21 (21)	46 (61)
DOS of NWS alerts	16 (2)	68 (69)	84 (84)	83.4 (83.5)	99 (99)	142 (150)
DOS of 90th pct. HI days	15 (1)	72 (69)	88 (85)	86.9 (85.2)	102 (102)	145 (153)

Table B.2: Summary statistics for our 30 counties of interest (and all the counties) of Medicare NOHR hospitalizations, NWS heat alerts, heat index, and day of summer for several features of interest. DMHI = daily maximum heat index, Q-DMHI = county quantile of DMHI (referred to simply as QHI in the main text), DOS = day of summer (out of 152). Variables given per county-summer are marked with *.

B.3 DETAILS OF THE BAYESIAN REWARDS MODEL

Recall our formulation for the expected per capita rate of NOHR hospitalizations at time t in county-summer (k, j) when taking action a :

$$\rho_t^{(kj)}(a) = \lambda_k(s_t^{(kj)})(1 - a \cdot \tau_k(s_t^{(kj)}))$$

THE BASELINE AND EFFECTIVENESS FUNCTIONS

Note that λ_k and τ_k are unique to each location k , allowing for spatial heterogeneity and thus reducing the possibility of unobserved spatial confounding^{12,113,140}. However, this design introduces the challenge of estimating these functions under limited data for each county (11 summers between 2006 and 2016). This challenge is amplified by the low signal in the health effects of heat and heat alerts¹⁶⁴. To address this, we adopt a (generalized) linear model for λ_k and τ_k paired with (i) borrowing statistical strength across counties using a data-driven random effects prior and (ii) injecting domain knowledge

on the signs of certain coefficients. The functions λ_k and τ_k are specified as

$$\begin{aligned}\lambda(s_t^{(k,j)}) &:= \exp\left(\beta_k^\top s_t^{(k,j)}\right), \\ \tau(s_t^{(k,j)}) &:= \text{sigmoid}\left(\delta_k^\top s_t^{(k,j)}\right),\end{aligned}\tag{B.1}$$

where β_k and δ_k are the model coefficients for each county $k \in \mathcal{K}$. While this specification is a linear combination of state variables, we allow for non-linear effects by using a piecewise linear basis expansion on the quantile of heat index (QHI) and a spline basis on the day of summer. The piecewise linear expansion on QHI in λ (the baseline function) has knots at the 25th and 75th percentiles. In τ (the alert effectiveness function), we include only a linear term for QHI—again, conditional on day of summer and excess QHI (compared to the average of the last 3 days, a time window informed by the literature on heat waves and health¹⁰¹)—to both stabilize effect identification and facilitate interpretation. The spline on day of summer (with 3 degrees of freedom) is included in both λ and τ .

Figure B.1 illustrates the interaction of λ_k and τ_k to capture the total effect of heat alerts on NOHR hospitalizations. For instance, we see that in early summer it is possible to have high heat alert effectiveness even when the heat index is not as high as it is later in the summer, possibly because people are less well-adapted to heat and/or more responsive to heat alerts early on.

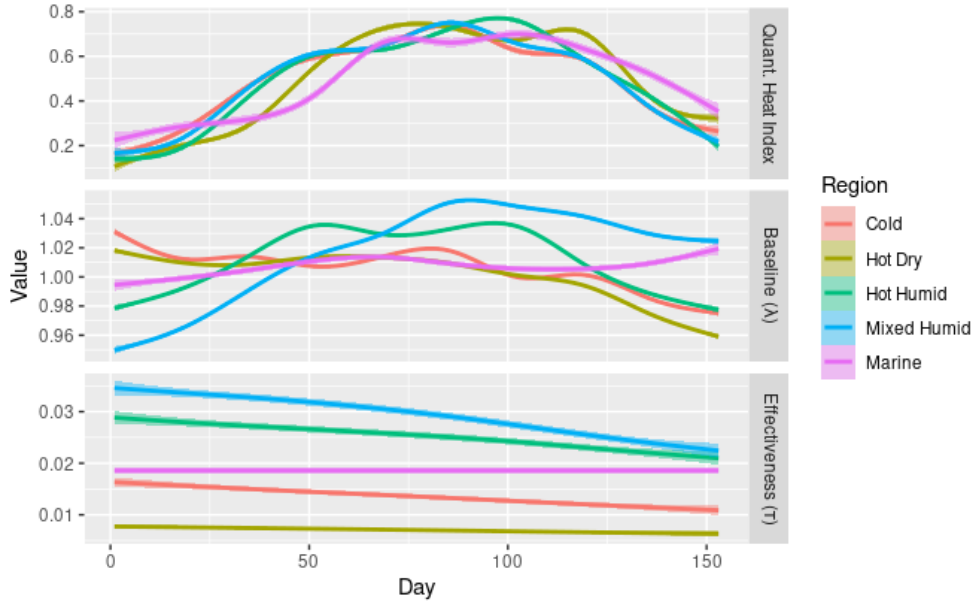


Figure B.1: Smoothed trajectories of observed quantile of heat index, modeled baseline NOHR hospitalization rate, and modeled alert effectiveness (assuming no past alerts) across days of summer; colored by climate region. Note that λ and τ are normalized as they are in the RL reward function.

DATA-DRIVEN PRIOR

We propose a data-driven prior distribution that allows borrowing statistical strength from other locations while preserving the possibility of spatial heterogeneity. Let w_k denote the vector of county-level features that are not included in the state vector but are predictive of the effect of heat and heat alerts on hospitalizations (see “County-level characteristics” in 2.4.1), and let $\gamma_k = [\beta_k; \delta_k] = (\gamma_k^\ell)_{\ell=1}^L$ denote the parameters corresponding to county k . Let $N(\cdot)$ denote a normal distribution. Then, the prior distribution of each γ_k^ℓ is set independently as

$$p_\theta(\gamma_k^\ell | \sigma_\ell, w_k) = \begin{cases} N(\gamma_k^\ell; f_\theta(w_k)_\ell, \sigma_\ell^2) & \text{no domain knowledge,} \\ \text{LogN}(\gamma_k^\ell; \exp(f_\theta(w_k)_\ell), \sigma_\ell^2) & \text{if } \gamma_k^\ell \in (0, \infty), \\ \text{NegLogN}(\gamma_k^\ell; -\exp(f_\theta(w_k)_\ell), \sigma_\ell^2) & \text{if } \gamma_k^\ell \in (-\infty, 0) \end{cases} \quad (\text{B.2})$$

where f_θ is a feed-forward neural network (with weights θ and L outputs) specifying the prior mean as a function of the spatial features, and σ_ℓ specifies the prior scale. We use a standard hyper-prior $p(\sigma_\ell) = \text{HalfCauchy}(\sigma_\ell)$ ⁵³.

The domain knowledge constraints we utilize are: (i) past heat alerts cannot increase the baseline hospitalization rate—in other words, alert fatigue can be identified only in τ —and (ii) higher QHI cannot decrease the effectiveness of heat alerts—note that this is conditional on day of summer.

The weights θ are learned from the data to characterize the prior distribution that best describes the latent space according to the data. As remarked in Wang et al.¹⁶², this procedure can be seen as a form of Empirical Bayes. According to this viewpoint, we can interpret the prior mean function f_θ in equation B.2 as an encoder that maps the county-level features w_k to a data-driven probability distribution over a latent space. The coefficients γ_k are latent variables in this space. The predictive model $\rho_t^{(k,j)}(a)$ in equation 2.4 acts as a decoder.

TRAINING

Traditional Bayesian inference methods using Markov Chain Monte Carlo are computationally prohibitive for our model due to the high dimensionality of parameters and the number of observations for each county. To address this issue, we use variational inference, which introduces an approximation of the true posterior distribution of the model parameters. We summarize the approach.

Let $(\gamma, \sigma) \sim q_\psi$ denote an approximation of the true posterior distribution where ψ are the learnable parameters of the approximating distribution q_ψ . Next, let $o_t^{(k,j)} = (s_t^{(k,j)}, a_t^{(k,j)}, y_t^{(k,j)}, n^{(k,j)})$ denote a data point in the training set and $\mathcal{L}(o_t^{(k,j)}; \gamma_k)$ be the Poisson log-likelihood corresponding to this data point (see equation 2.4). Variational inference attempts to maximize the likelihood of the observed data while penalizing the divergence between the variational posterior q_ψ and the prior $p_\theta(\gamma, \sigma | w) = \prod_{k,l} p_\theta(\gamma_k^\ell | \sigma_\ell, w_k) \prod_l p(\sigma_\ell)$. The optimization objective is to maximize the evidence

lower bound (ELBO) given by

$$\text{ELBO}(\psi, \theta) = \mathbb{E}_{q_\psi} \left[\sum_{k,j,t} \mathcal{L}(o_t^{(k,j)}; \gamma_k) \right] - D_{\text{KL}}(q_\psi(\gamma, \sigma) \parallel p_\theta(\gamma, \sigma|w)), \quad (\text{B.3})$$

where D_{KL} denotes the Kullback-Leibler divergence. To ensure the variational posterior is flexible enough to approximate the true posterior, we use a low-rank multivariate normal distribution¹⁴⁴. This approach uses a bijective transform of the parameters to an unconstrained space and then applies a low-rank factorization to the covariance matrix of the multivariate normal distribution. Thereby, it allows for efficient computation while approximating the true correlation structure of the joint posterior distribution. The data-driven prior parameters θ are learned jointly with the variational parameters ψ using stochastic optimization. Specifically, we train one neural network for β (the coefficients of the baseline function) and another for δ (the coefficients of the alert effectiveness function). Both networks have one hidden layer with 32 units—this modest architecture avoids overfitting the relatively small number of spatial features in our dataset.

INSPECTION OF THE TRAINED REWARDS MODEL

The rewards model clearly converged in its loss across epochs (shown in Figure B.2) and achieved higher predictive accuracy than standard machine learning algorithms (i.e., random forest and neural network). Predicting the rate of NOHR hospitalizations in the Medicare population in each county, this model has $R^2 = 0.102$ —using the calculation SSR/SST for comparability with earlier machine learning models (such as that we used when checking the signal in the mortality data). Predicting the absolute number of NOHR hospitalizations, it has $R^2 = 0.884$ —much higher because the value is mostly driven by county population size. The model also displayed very good coverage when we ran it on synthetic data (1,000 samples from the posterior predictive) using known coefficients: average coverage across parameters for a 90% CI was 0.897.

Figure B.3 shows one posterior sample across the 761 counties on which the model was fit. Note that the spread across counties is larger than the spread within counties, so the IQR bars are no larger for multiple samples. Overall, we see that the coefficients on past alerts in today’s alert effectiveness (τ) are both negative, *empirical evidence of alert fatigue*. In the baseline hospitalization rate (λ), we see a strong protective effect by weekend.

B.4 DETAILS OF THE RL ANALYSIS

B.4.1 SELECTION OF 30 COUNTIES

Out of the 170 counties with 75 or more heat alerts issued between 2066 and 2016, we selected 30 using the following criteria: for each climate region[†], select the county with the highest variance in

[†]Excluding “Mixed Dry” and “Very Cold” because there were very few counties in these regions with sufficient population size

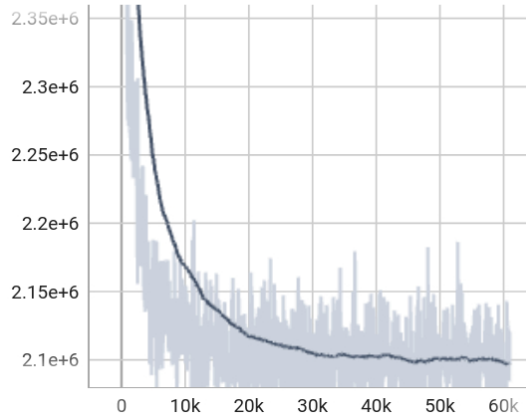


Figure B.2: Convergence of the ELBO in the Bayesian rewards model during training (80 epochs).

estimated heat alert effectiveness (indicating higher signal as identified by our rewards model). To get up to 30, repeat this selection, but first prioritize selecting a county from a U.S. state not already in the set, and if we must repeat a state, do not select a county that is directly adjacent to one already in the set. If there are no more eligible counties in a region, select from a larger region. Characteristics of these 30 counties are provided in Table B.3.

B.4.2 RL ALGORITHM DESCRIPTIONS

DQN is standard Deep Q-learning. QRDQN differs from DQN in that rather than estimating the average return from taking action a at state s , it estimates the distribution over the returns using quantile regression. Compared to other common stochastic policy gradient algorithms, TRPO is distinguished by its constraint on the size of the optimization step, so that the new policy lies within a “trust region” from the previous policy, in which local approximations of the policy function are still accurate⁶⁴. A2C improves on the basic setup of actor-critic by updating its policy model using the advantage function, which indicates how much better an action is than the alternative(s), rather than just the state-action value of that one action; this helps increase model stability⁶⁴.

B.4.3 HYPERPARAMETER TUNING

RL algorithms are known to be sensitive to their hyperparameters, especially in empirical settings⁴⁵. To address this concern, we tuned hyperparameters per county, using the average return metric. First, we did a broad sweep on one county from each of the five climate regions using TRPO.QHI, experimenting with values of learning rate $\in \{0.01, 0.001, 0.0001\}$, discount factor $\in \{1.0, 0.999, 0.99\}$, neural network architecture (number of hidden layers $\in \{1, 2, 3\}$ and number of hidden units $\in$

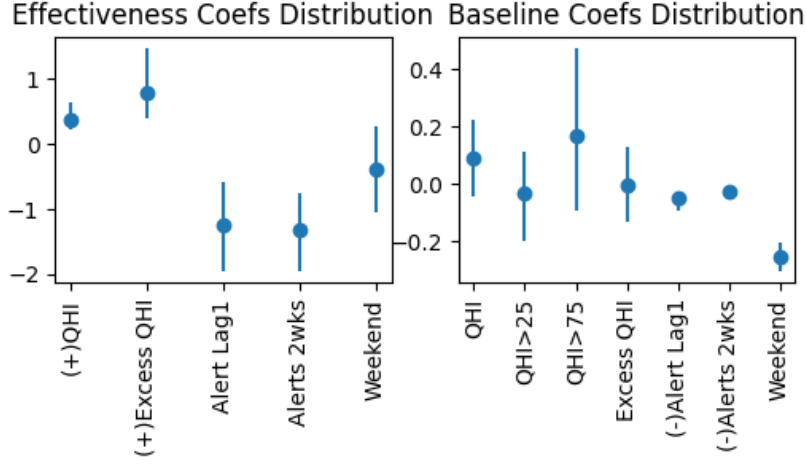


Figure B.3: A sample of posterior coefficients from the rewards model (in τ and λ respectively), displaying IQR and median. Bias and coefficients for the spline on day of summer are not shown, for ease of interpretability. (+)/(-) indicates a coefficient constrained to be positive/negative.

{16, 32, 64}), and size of the replay buffer $\in \{1024, 2048, 4096\}$; paired with possible QHI thresholds $\in \{0.55, 0.7, 0.85\}$ and both including and not including future information. After determining that learning rate=0.001 and discount factor=1.0 performed the best across the board, we tuned the remaining parameters for each county and each RL model (all combinations of TRPO, A2C, QRDQN and DQN with and without QHI restrictions and future information). To make the computation time more reasonable, we did a grid search on two values each for the number of hidden layers (2 or 3), number of hidden units (16 or 32), and size of the replay buffer (1,500 or 3,000). For QRDQN specifically, we also tested different values for the number of quantiles, which is closely tied to computation time for this algorithm. Using 10 quantiles rather than 20 ended up being not only faster but also better, potentially due to the latter overfitting.

B.4.4 SENSITIVITY ANALYSIS USING 2014-2016 AS THE EVALUATION YEARS

For real-world deployment of a heat alert-RL system, an ideal setup would be to train on years $\{1, \dots, x\}$ and evaluate on year $x + 1$. However, in this analysis we held out $\{2007, 2011, 2015\}$ as the evaluation set to provide a better representation of historical performance. Note that while days within each episode are consecutive, the 7-month gap between episodes lessens concern about correlated episodes.

To provide intuition for how sequential training-testing would look, we re-ran the best RL model (A2C.QHI) using 2006-2013 to train and 2014-2016 to evaluate. To reduce the computational burden, we used the optimized hyperparameters from the main analysis for each county. Table B.4 shows the aggregated results as well as results stratified by year. A few notable results are that (i) the ordering of the original results ($2007 < 2015 < 2011$) is the same as the ordering of median QHI, (ii) the 2015

results are similar in each evaluation scheme, and (iii) the 2016 result is the worst. However, this last result is overly bleak because in an online deployment we would not be evaluating 3 years ahead—the model would be updated at least once a year.

B.5 SUMMARY OF THE WORKFLOW USING BROACH

This section is meant to illustrate a toy example of our workflow. Imagine seven counties located in two climate regions, as shown in the figure below.

Climate Region A				Climate Region B		
1	2	3	4	5	6	7

Let the training years $TY = \{2006, 2008, 2009, 2010, 2012, 2013, 2014, 2016\}$ and evaluation years $EY = \{2007, 2011, 2015\}$. First, train the rewards model on counties 1-7 and years $TY \cup EY$ to obtain posterior distributions for $r_1(s, a), \dots, r_7(s, a)$. Then, for each county, follow the steps below. Consider county 1 for illustration. Our procedure is as follows:

```

1: if Training RL (with a QHI restriction) then
2:   for  $h$  in  $\{0.5, 0.55, \dots, 0.9\}$  do
3:     for  $i$  in 1 to  $N_T$  episodes do
4:       Sample  $r_{1,i}(s, a)$  from the posterior for  $r_1(s, a)$ 
5:       Sample  $\xi_i \sim TY \times \{\xi_1, \xi_2, \xi_3, \xi_4\}$ 
6:       Update  $\pi_{1,b}(a|s)$  s.t.  $a = o$  if  $QHI < h$ , using  $r_{1,i}$  and  $\xi_i$ 
7:     end for
8:     Select  $\pi_{1,b}^*$  (across  $i$ ) with the largest average return using  $\xi_i \sim EY \times \{\xi_2, \xi_3, \xi_4\}$ 
9:   end for
10: else if Tuning RL or AA.QHI then
11:   for  $h$  in  $\{0.5, 0.55, \dots, 0.9\}$  do
12:     for  $i$  in 1 to  $N_E$  episodes do
13:       Sample  $r_{1,i}(s, a)$  from the posterior for  $r_1(s, a)$ 
14:       Sample  $\xi_i \sim EY \times \{\xi_2, \xi_3, \xi_4\}$ 
15:       Calculate cumulative reward using  $\pi_{1,b}^*$  (or AA.QHI with  $h$ ),  $r_{1,i}$  and  $\xi_i$ 
16:     end for
17:   end for
18:   Select  $h^*$  with the largest return
19: else if Final evaluation for any policy (but using notation for optimized RL) then
20:   for  $i$  in 1 to  $N_E$  episodes do
21:     Sample  $r_{1,i}(s, a)$  from the posterior for  $r_1(s, a)$ 

```

```

22:         Sample  $\xi_i \sim EY \times \{\xi_1\}$ 
23:         Calculate cumulative reward using  $\pi_{1,h}^*$ ,  $r_{1,i}$  and  $\xi_i$ 
24:     end for
25: end if

```

B.6 APPROXIMATE CONFIDENCE INTERVAL CALCULATION

To generate an approximate confidence interval for the absolute number of NOHR hospitalizations saved by implementing our best RL model, we use the estimates from the final evaluation for A2C.QHI and NWS. However, rather than first calculating average returns per county (across the 1,000 evaluation episodes) and then taking the median across the 30 counties, we first take the median return (across the 30 counties) from each of the 1,000 evaluation episodes and then calculate the 2.5% and 97.5% quantiles of the resulting distribution of medians.

B.7 DISCUSSION OF ABSOLUTE HEALTH BENEFITS

Do we estimate substantive health benefits from implementing the improved heat alerts policies detailed in the main text? For general context, the median rate of NOHR hospitalizations per 10,000 Medicare enrollees per summer during our study period was 148 (Table B.2), and we estimated that deploying our best RL model *relative to issuing no heat alerts* would save 2 of these 148, or 1.4%, in the 30 counties in our RL analysis. Noting that the overall relative risk of a Medicare enrollee experiencing one of the NOHR hospitalization types on a heat wave day is about 1.10^{15} , this is substantial evidence in support of continuing to issue heat alerts to protect public health.

Compared to the observed NWS policy, A2C.QHI applied to the 30 counties in our analysis would save approximately 0.042 NOHR hospitalizations per 10,000 Medicare enrollees per summer, a reduction of only 0.03%. Multiplying this by 49 million, the size of the Medicare population in 2011 (the midpoint of our study period), yields about 222 NOHR hospitalizations saved per summer. However, this modest value should be further contextualized by several points.

Firstly, heat alerts by themselves (and in our case, the re-arrangement of the same number of heat alerts) are a very cost-effective intervention, so even small reductions in health harms are promising.

Secondly, if we take into account the large heterogeneity in the benefit of RL across counties with numbers from the CART regression (Figure B.9), assuming that we figure out how to implement a safe policy such that counties which would not benefit are unaffected, this number increases to 262 $((49m/10k) * (0.074 * 0.17 + 0.053 * 0.33 + 0.18 * 0.13))$ NOHR hospitalizations saved per summer. We obtain a nearly identical result from running a CART analysis using only variables which are available for all 761 counties that were included in the rewards model, and applying the results of that model across the whole set of counties. These are rough approximations given that they assume the proportion of counties in which RL alert policies were found to improve health outcomes in our sample is generalizable to the rest of the US. It is plausible that our sample of 30 counties is unrepresentative of

the U.S. as they tend to have higher variability in alert effectiveness (increasing potential improvements from RL) and also higher alert budgets.

Thirdly, these metrics reflect the climate and Medicare population between 2006 and 2016. We confidently anticipate that both the frequency of extreme heat events under climate change and the size of the Medicare population will continue increasing³⁰. For instance, while there were 49 million Medicare enrollees in 2011, by 2023 this number has increased to over 65 million, and by 2050 it is projected to be over 85 million¹⁴².

Fourthly, the numbers reported in this study should not be interpreted as the overall benefit to public health from improving heat alerts issuance strategies. The health data used in our study only reflect inpatient hospitalizations from NOHR causes billed to Medicare Part A, or hospital fee-for-service. Only about 60% of Medicare enrollees are in the fee-for-service population[‡], so hospitalizations of the other 40% do not appear in our dataset.

B.8 EXTENDED RESULTS FROM THE RL ANALYSIS

B.8.1 POLICY EXAMPLES

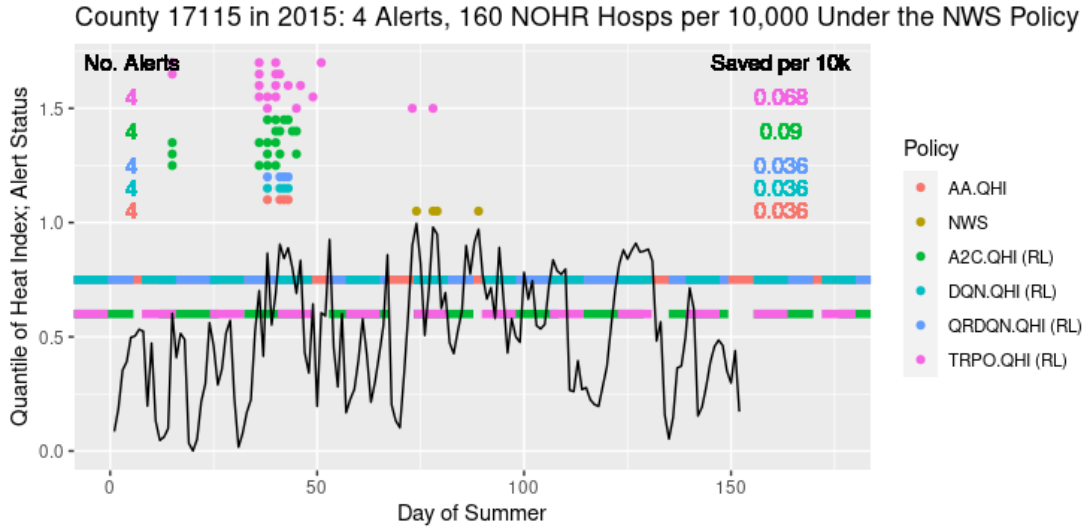


Figure B.4: Example of a setting with a relatively small heat alert budget.

[‡]https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Beneficiary-Snapshot/Downloads/Bene_Snapshot.pdf

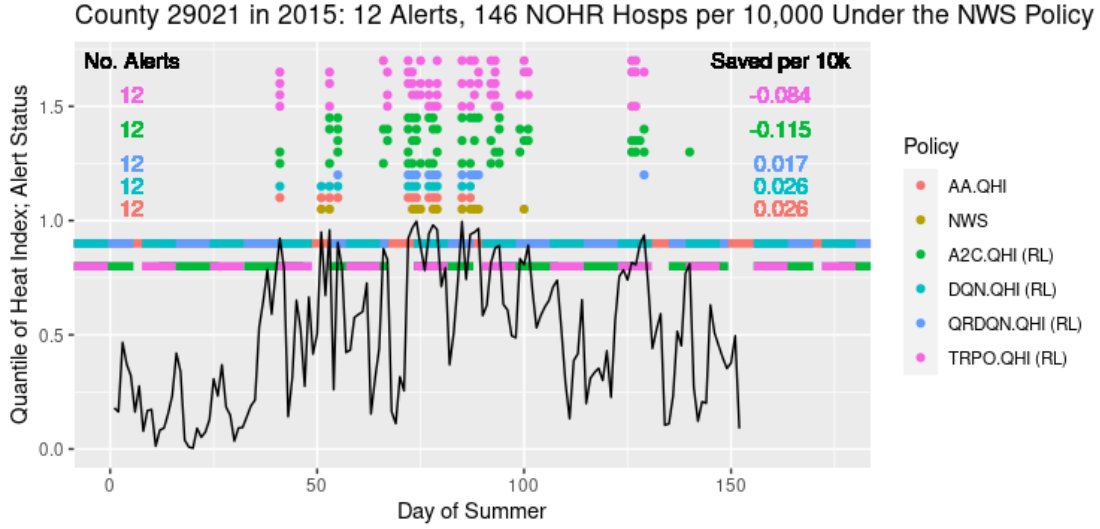


Figure B.5: Example of a setting where implementing RL would not be beneficial relative to nws or aa.qhi.

B.8.2 ADDITIONAL DISCUSSION OF STOCHASTIC VS DETERMINISTIC POLICIES

Interpreting policies from the TRPO and A2C models deterministically fails because many of the policy functions' estimated probabilities for issuing an alert are below 0.5, so always taking the action with the larger probability dramatically reduces the number of alerts issued—in some counties to zero. Note that policies which do not use all the alerts in their budget are penalized by the specification of our rewards model, resulting in `DET.TRPO.QHI` and `DET.A2C.QHI` no longer outperforming NWS with statistical significance across the 30 counties. Inclusion of future information mitigates this issue, but the results are not nearly as good as their stochastic counterparts.

Our big picture intuition for this result is that we are in a heavily data-constrained setting, which frustrates RL's ability to completely identify alert fatigue. The stochastic models address this by sampling alert days, inducing less alert streaks. Though a bit unsatisfying, we believe this is an acceptable compromise given that, as mentioned in the main text, an online heat alert-RL system would likely utilize stochastic exploration anyway, to update itself over time.

B.8.3 ADDITIONAL FIGURES FROM THE POST-HOC CONTRASTIVE ANALYSIS

Figure B.6 provides histograms of different policies' temporal characteristics, and figures B.7, B.8, and B.9 illustrate the CART analysis.

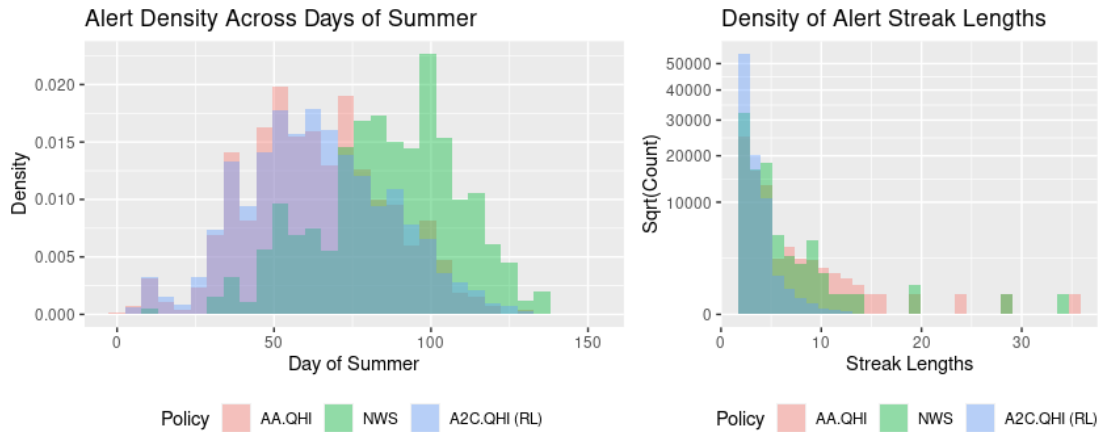


Figure B.6: Distributions of heat alert characteristics under different policies, over the evaluation years.

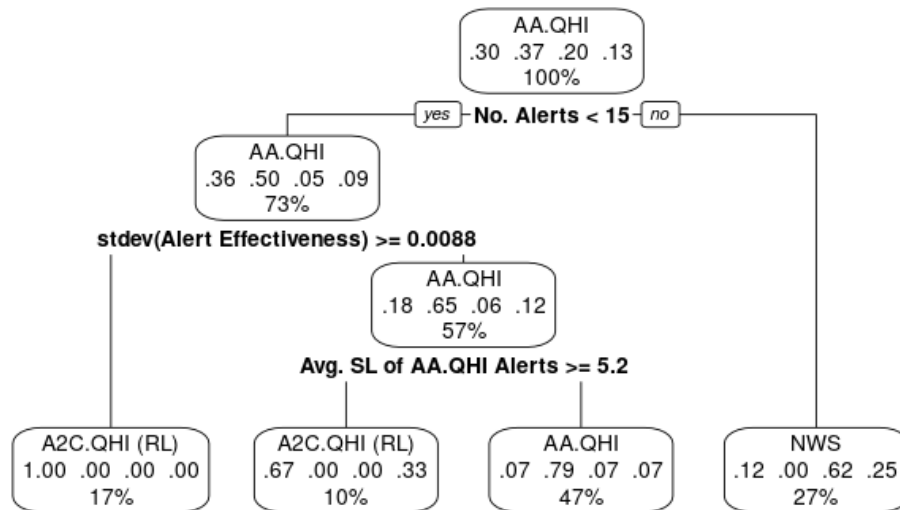


Figure B.7: Polytomous classification tree comparing A2C.QHI, AA.QHI, NWS, and TRPO.QHI across the 30 counties. Within each node, the policy at the top is that associated with the highest probability of performing best. The four probabilities beneath correspond to the four policies in alphabetical order, as written above. (Note that the two RL policies tend to be more similar to one another, as do the AA.QHI and NWS policies, so the distribution of probabilities is important to consider in addition to the overall classification of each node.) The percentage at the bottom is the fraction of counties represented by that node. “No. Alerts” is the average number of heat alerts per warm season during the evaluation years, “Alert Effectiveness” is the τ estimated by our rewards model, and SL is the streak length.

County Fips (State name abbreviation)	Region	No. Alerts (Eval Years)	Population Density (/mi ²)	Median HH Income (USD)	Democratic Voters (%)	Broadband Usage (%)	Average PM _{2.5} ($\mu\text{g}/\text{m}^3$)
5045 (AR)	MxHd	61	179.00	50314.00	34.00	37.50	9.08
20161 (KS)	MxHd	47	120.00	43962.00	43.50	52.20	6.35
21059 (KY)	MxHd	24	212.00	46555.00	38.70	53.30	10.22
29019 (MO)	MxHd	50	242.00	48627.00	51.60	46.60	7.85
37085 (NC)	MxHd	31	200.00	44625.00	39.30	46.70	9.65
40017 (OK)	MxHd	63	133.00	63629.00	22.80	70.10	7.53
47113 (TN)	MxHd	58	176.00	41617.00	44.30	79.40	9.29
42017 (PA)	MxHd	25	1036.00	76555.00	51.00	85.10	9.14
41053 (OR)	Mrn	31	102.00	52808.00	45.60	79.20	3.24
41067 (OR)	Mrn	31	745.00	64180.00	57.70	82.70	4.28
53015 (WA)	Mrn	23	90.00	47596.00	49.40	53.60	3.59
13031 (GA)	HtHd	43	106.00	35840.00	38.90	57.50	9.01
22109 (LA)	HtHd	34	91.00	49960.00	27.90	55.50	8.26
28049 (MS)	HtHd	57	283.00	37626.00	69.90	36.50	8.50
45015 (SC)	HtHd	48	168.00	52427.00	41.20	66.20	8.72
48157 (TX)	HtHd	42	707.00	85297.00	47.80	83.70	8.64
48367 (TX)	HtHd	41	131.00	64515.00	18.30	52.20	7.84
22063 (LA)	HtHd	42	201.00	56811.00	13.70	47.60	8.87
4015 (AZ)	HtDr	29	15.00	39200.00	28.70	63.60	4.59
6025 (CA)	HtDr	40	42.00	41807.00	64.00	64.40	7.22
32003 (NV)	HtDr	29	251.00	52873.00	55.90	78.10	5.21
4013 (AZ)	HtDr	72	423.00	53596.00	44.00	77.40	6.38
6071 (CA)	HtDr	29	103.00	54090.00	51.60	77.10	6.38
17115 (IL)	Cold	22	190.00	46559.00	45.70	63.60	8.87
17167 (IL)	Cold	25	228.00	55449.00	46.00	66.50	8.75
19153 (IA)	Cold	24	764.00	59018.00	54.90	57.90	8.25
19155 (IA)	Cold	27	98.00	51304.00	44.30	42.00	7.89
34021 (NJ)	Cold	29	1639.00	73480.00	67.30	81.60	9.31
29021 (MO)	Cold	44	219.00	44363.00	43.70	48.80	7.97
31153 (NE)	Cold	27	681.00	69965.00	37.60	72.60	7.92

Table B.3: Descriptive statistics of the 30 counties used in policy optimization. Climate region abbreviations = {MxHD: Mixed-Humid, Mrn: Marine, HtHd: Hot-Humid, HtDr: Hot-Dry, Cold: Cold}. HH = household. PM_{2.5} is an annual average.

Evaluation Period	QHI Median	Median Diff. (RL-NWS)	WMW p-value
Main Analysis:			
{2007, 2011, 2015}		0.042	0.011
2007	0.51	0.006	0.013
2011	0.555	0.016	0.027
2015	0.519	0.01	0.010
Sensitivity Analysis:			
{2014, 2015, 2016}		0.007	0.121
2014	0.46	0.002	0.012
2015	0.519	0.009	0.022
2016	0.6	-0.005	0.777

Table B.4: Comparison of results using 2014-2016 as the evaluation years (sensitivity analysis) as opposed to 2007,2011,2015 (main analysis).

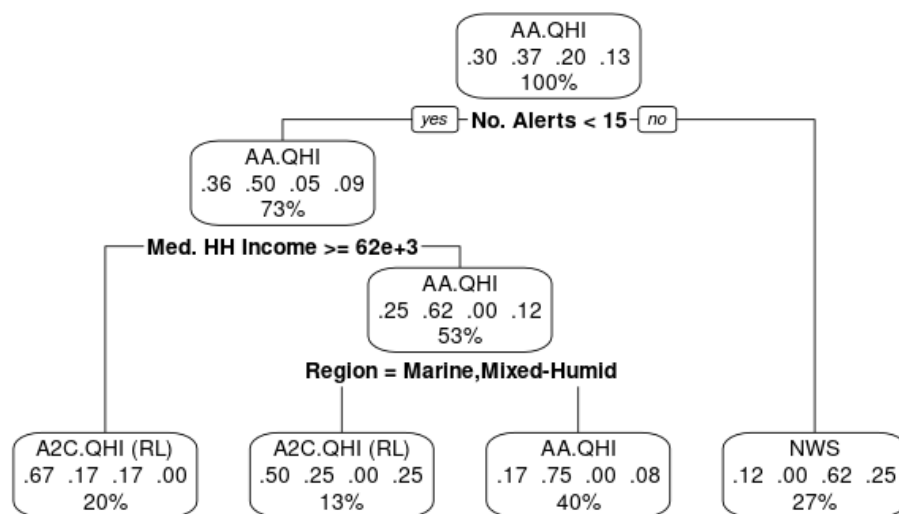


Figure B.8: Polytomous classification tree comparing A2C.QHI, AA.QHI, NWS, and TRPO.QHI across the 30 counties—using more conventional features than Figure B.7. Within each node, the policy at the top is that associated with the highest probability of performing best. The four probabilities beneath correspond to the four policies in alphabetical order, as written above. (Note that the two RL policies tend to be more similar to one another, as do the AA.QHI and NWS policies, so the distribution of probabilities is important to consider in addition to the overall classification of each node.) The percentage at the bottom is the fraction of counties represented by that node. “No. Alerts” is the average number of heat alerts per warm season during the evaluation years.

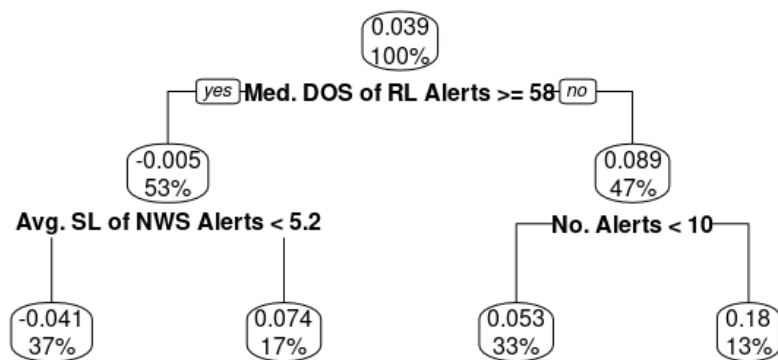


Figure B.9: Regression tree on the differences in average return between A2C.QHI and NWS. Within each node, the number at the top is the prediction (average) difference, with positive indicating that the RL is better and negative indicating that the NWS is better. The percentage at the bottom is the fraction of counties represented by that node. DOS is the day of summer, SL is the streak length, and “No. Alerts” is the average number of heat alerts per warm season during the evaluation years.



Appendix to Chapter 3

C.1 ESTIMATION DETAILS

C.1.1 NUISANCE FUNCTION HYPERPARAMETERS

For each of the nuisance functions $(\mu_1, \mu_0, \pi_B, \pi_D)$, we use the Ranger algorithm (from the R caret package) with `num.trees=500`. For the regressions (all but π_B , which is a binary classification), we use `min.node.size=5` and select the minimum value of `mtry` (the number of variables to possibly split at in each node) beyond which the R^2 and RMSE do not substantially improve (a difference within 0.01 for R^2 and 0.1 for RMSE). For π_B , we limit `mtry` such that the empirical probabilities do not get too large (>0.99) or too small (<0.001).

To prevent overfitting, we use cross-validation with $k=2$ folds to select `mtry`. For μ_0 , we sample observations into the folds by year block (2012 & 2013, 2014 & 2015, 2016 & 2017) to emulate this model's extrapolation to the $B = 1$ years. For π_D , we sample observations by GPW site to account for D being a static variable in our dataset, even though $\mathbf{X}$ contains some time-varying covariates. For μ_1 and π_B , we randomly sample observations from $\mathcal{O}_{B=1}$ and $\mathcal{O}$ respectively.

C.1.2 GENERALIZED PROPENSITY SCORE (GPS) ESTIMATION

Following the GPS estimation method utilized by Wu et al.¹⁷⁰, we first fit a model to D given $\mathbf{X}$, within $B = 1$. Note that while in our dataset, D is static (derived from the 2018 GMTDS data), early experimentation revealed that the time-varying covariates are meaningfully predictive of D . Our intuition for this is that the spatial variation in the *average* values of the meteorological covariates, such as humidity, is informative of how close one is to the coast, which is in turn associated with port proximity.

After fitting the initial model for D , we fit a model to the residuals $D_i - \hat{D}_i$ given $\mathbf{X}$, within $B = 1$. (For efficiency, we use the same hyperparameters for this residual model as for the initial model of D .) Then, we apply a kernel density estimator to $(D_i - \hat{D}_i)/\hat{\varepsilon}_i$ to obtain π_D . For $f(D|B = 1) = \int_{\mathbb{X}} \pi_D(D|B = 1, \mathbf{x})dP(\mathbf{x}|B = 1)$, we interpolate over the observed $\mathbf{X}$ for specified values of D .

C.2 SUPPLEMENTARY FIGURES AND TABLES

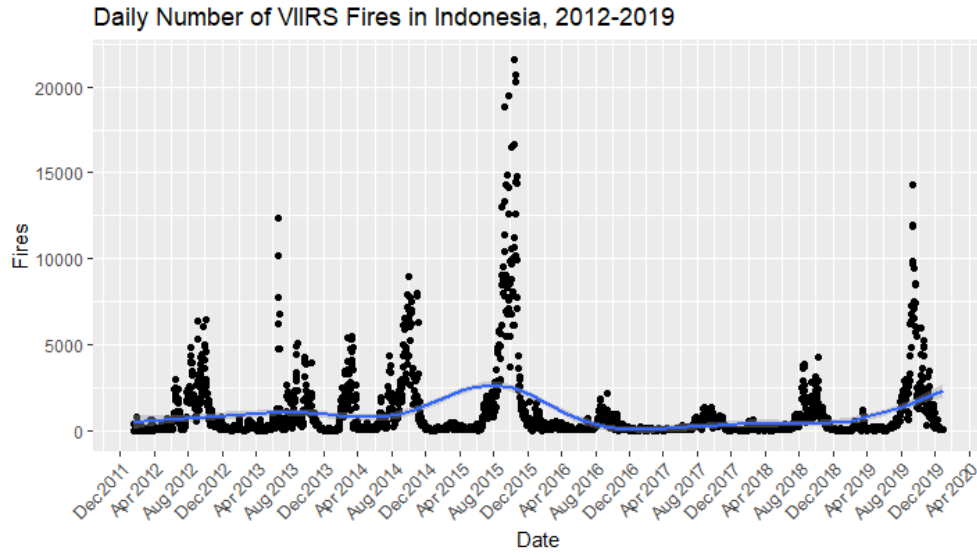


Figure C.1: Sum of daily VIIRS fires detected anywhere in Indonesia (not only at waste dump sites), with smoothed trend shown.

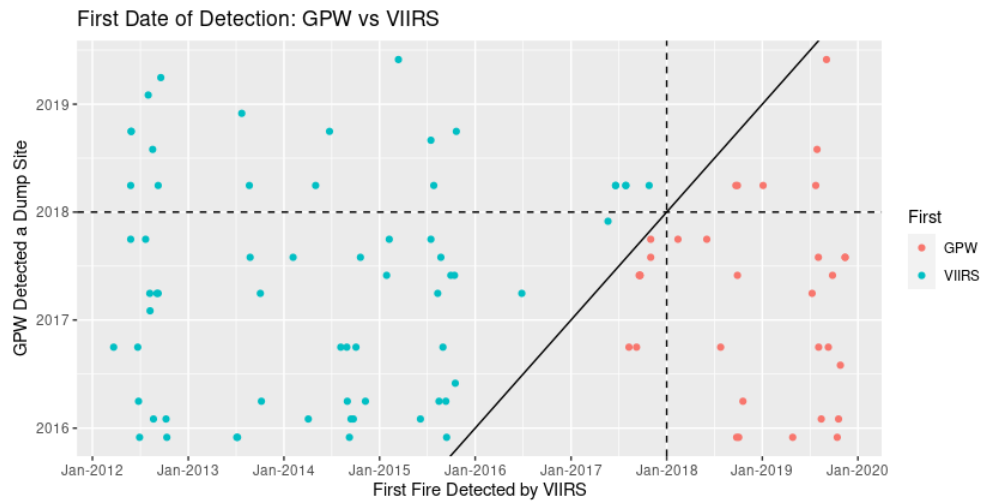


Figure C.2: Detection date of each dump site (at which active fires occurred) based on GPW's algorithm and VIIRS.

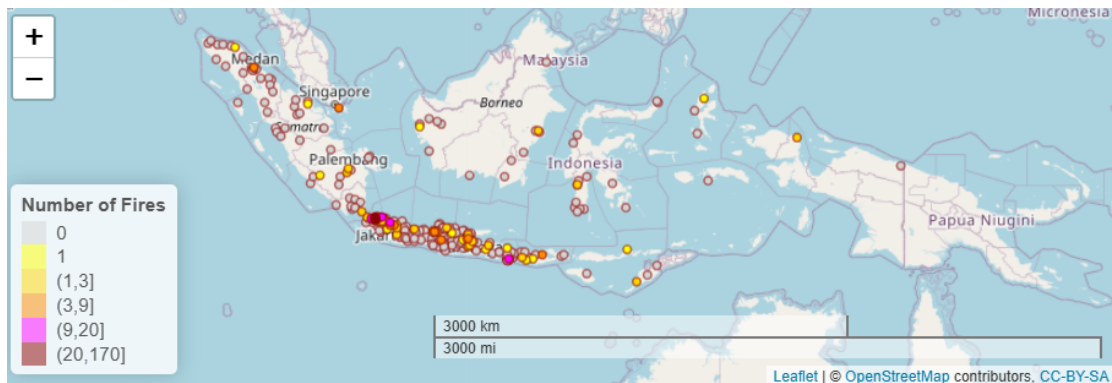


Figure C.3: Map of Indonesia showing the number of fires detected by VIIRS overlapping with each dump site (detected by Global Plastic Watch) in 2012-2019 and the boundaries of the provinces (in gray).

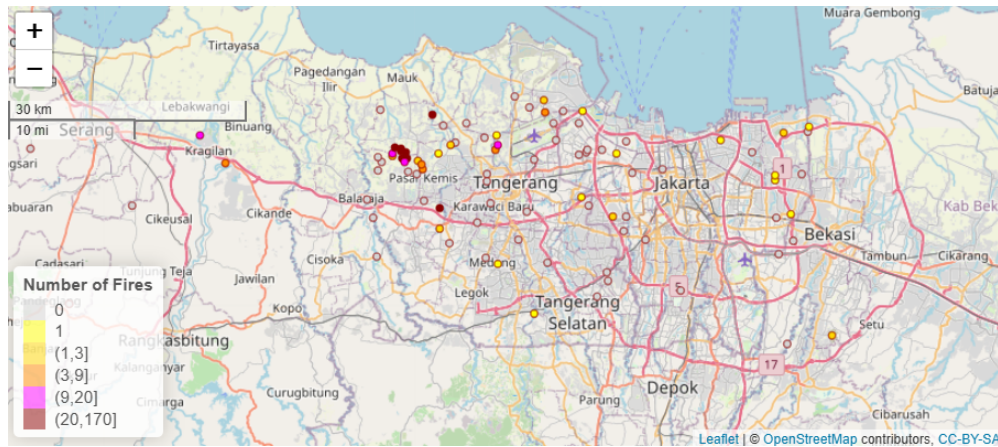


Figure C.4: Zoomed-in map of Jakarta showing the number of fires detected by VIIRS overlapping with each GPW site in 2012-2019.

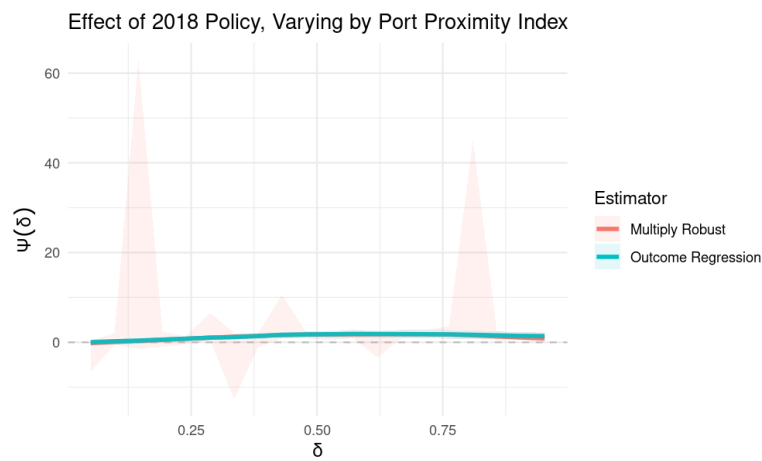


Figure C.5: Main analysis run with LLKR bandwidth selected on all data (with repeated observations at each dump site).

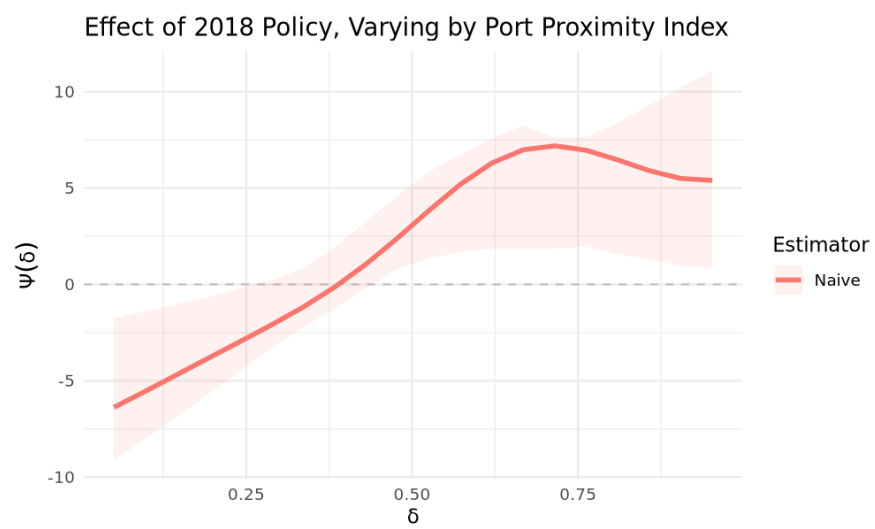


Figure C.6: The confounder-naive results. Sampling bootstrap weights with spatial correlation (range=26.8km).

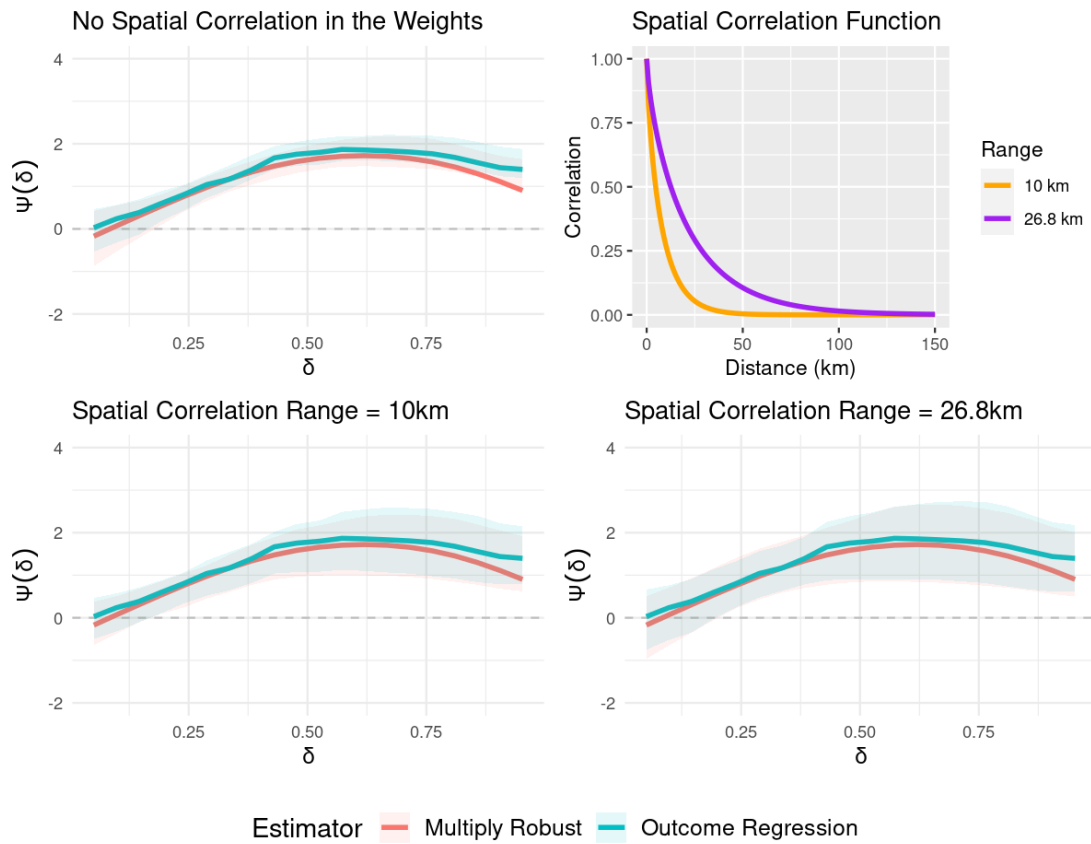


Figure C.7: Comparing the results when the bootstrap weights have spatial autocorrelation range 0, 20, and 26.8 km respectively. The top righthand plot shows the Matérn correlation decay versus distance, using these different values for the range parameter.

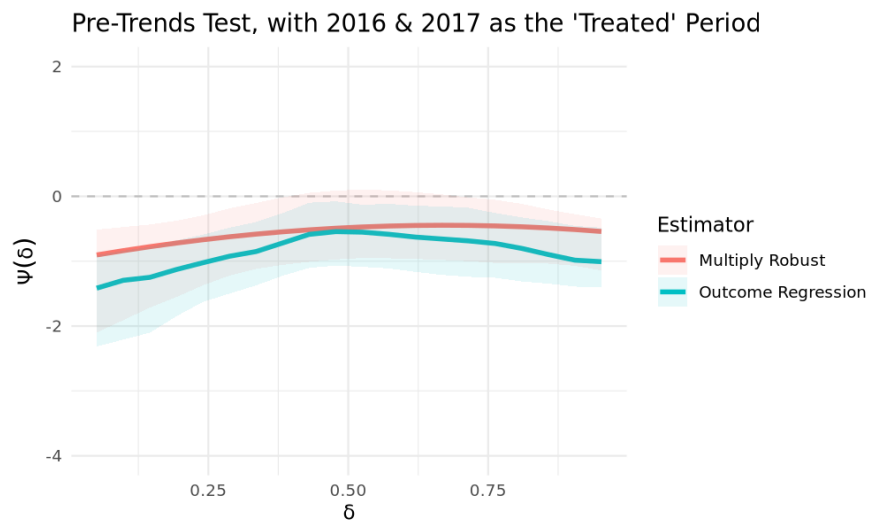


Figure C.8: Pre-trends test for 2016-2017. Sampling bootstrap weights with spatial correlation (range=26.8km).

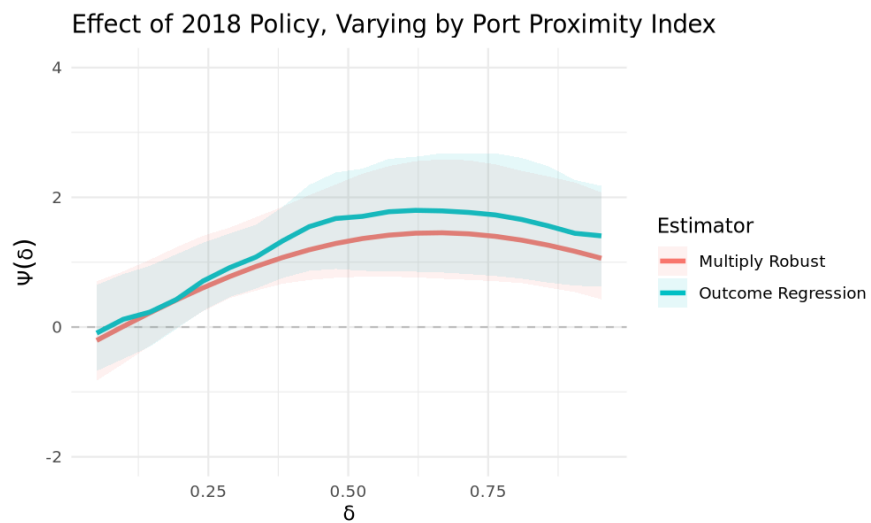


Figure C.9: Sensitivity analysis including GPW site area as a covariate. Sampling bootstrap weights with spatial correlation (range=26.8km).

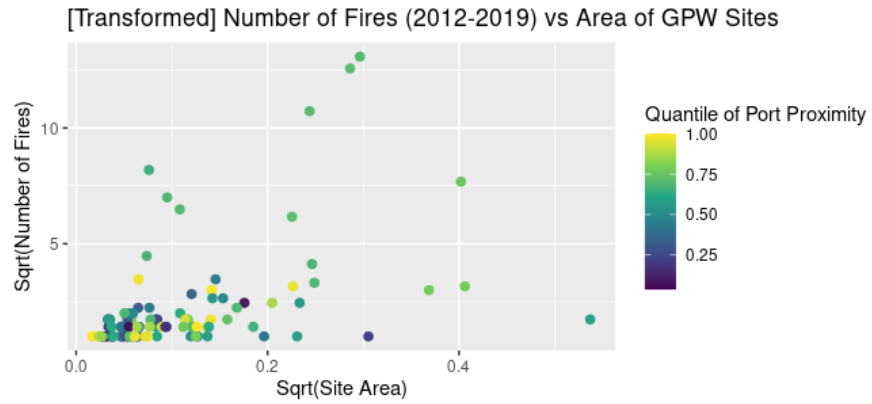


Figure C.10: Square root of number of fires detected by VIIRS vs square root of GPW site area, colored by the quantile of port proximity.

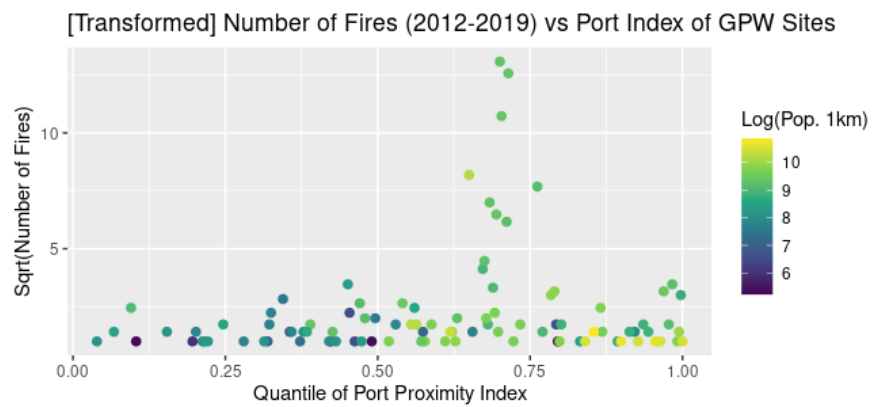


Figure C.11: Square root of number of fires detected by VIIRS vs quantile of port proximity, colored by the log of population living within 1km of each GPW site.

C.3 SIMULATIONS EXPLORING THE SPATIAL WEIGHTED BOOTSTRAP

In this appendix, our primary aim is to evaluate the characteristics of the confidence intervals generated by our Spatial Weighted Bootstrap procedure, relative to those generated by a Non-spatial Weighted Bootstrap. Here, calculation of the ADT curve’s bias and Mean Absolute Error (MAE) is an intermediate step to better understand the confidence intervals, as Hettinger et al.⁷¹ present extensive investigation of a very similar ADT estimator.

To do this evaluation, we develop simulations which utilize the real (observed) exposure D and covariates $\mathbf{X}$ to simulate the outcome Y .

C.3.1 SIMULATING THE OUTCOME

First, we compute the mean function for the outcome data generating process using the real exposure and covariate data and setting “true” parameter values governing the relationship between these variables and the outcome in our simulated data. Our choice of these parameter values was guided by (a) using domain knowledge to select the sign of each coefficient, (b) producing an outcome distribution similar to that in the real data, and (c) producing a plausible ADT curve, similar to that observed in our main analysis.

The mean function for each location i and time t is given by:

$$\mu_{it} = B_t L_1(\mathbf{X}_{it}, D_{it}) + (1 - B_t) L_0(\mathbf{X}_{it})$$

where

$$L_0(\mathbf{X}_{it}) = -9.5 + 0.075(\text{Temp}_{it}) + 0.075(\text{Pop}_i) + 0.0001(\text{Temp}_{it})(\text{Pop}_i) + 0.15(\text{Time}_t) + 0.001(\text{Time}_t)(\text{Pop}_i) + 1(\text{Other_fires_province}_{it}) - 18(\text{Precip}_{it})$$

and

$$L_1(\mathbf{X}_{it}, D_{it}) = -12.5 + 0.075(\text{Temp}_{it}) + 0.075(\text{Pop}_i) + 0.0007(\text{Temp}_{it})(\text{Pop}_i) + 0.15(\text{Time}_t) + 0.001(\text{Time}_t)(\text{Pop}_i) + 1.5(\text{Other_fires_province}_{it}) - 20(\text{Precip}_{it}) + 2(D_{it}) \left[-3(D_{it}) + 0.4(\text{Pop}_i) + 0.2(\text{Time}_t) \right]$$

where “Temp” is temperature, “Pop” is population [density], “Time” is the fifth-root of the count of days since 01/01/2012, “Other_fires_province” is the fourth-root of the number of fires in the province not overlapping with any GPW site, and “Precip” is precipitation. For both L_0 and L_1 , we truncated any negative values ($\sim 0.06\%$) to zero post hoc.

We calculate the “true” ADT for each exposure value δ as

$$\Psi(\delta) = \frac{1}{96N} \sum_{i=1}^N \sum_{t=1}^{96} \left[L_1(\mathbf{X}_{it}, \delta) - L_0(\mathbf{X}_{it}) \right]$$

This true ADT curve is shown in Figure C.12.

We simulate both spatially-structured and independent error in the outcome, as follows:

$$\tilde{Y}_{it} = \mu_{it} + \sigma_{\text{sp}} \varepsilon_{\text{sp}} + \sigma_{\text{ind}} \varepsilon_{\text{ind}}$$

where $\varepsilon_{\text{sp}} \sim MVN(0, \Sigma)$, where the correlation function to generate the entries of Σ is $\text{Corr}(m; L) = \exp(-m/L)$ [Exponential form] and $\varepsilon_{\text{ind}} \sim N(0, 1)$. The parameters L (the true spatial range), σ_{sp} , and σ_{ind} are constants, which we specify to obtain four distinct simulation scenarios:

1. Scenario 1: $L = 100$ km, $\sigma_{\text{sp}} = 2$, $\sigma_{\text{ind}} = 0.5$ (the “base case”)
2. Scenario 2: $L = 500$ km, $\sigma_{\text{sp}} = 2$, $\sigma_{\text{ind}} = 0.5$ (increased spatial range)
3. Scenario 3: $L = 100$ km, $\sigma_{\text{sp}} = 3$, $\sigma_{\text{ind}} = 0.5$ (increased spatial sigma)
4. Scenario 4: $L = 100$ km, $\sigma_{\text{sp}} = 2$, $\sigma_{\text{ind}} = 1.5$ (increased noise)

These four scenarios allow us to investigate the influence of different spatial error structures on the accuracy of our estimation and inference procedures. For each simulation scenario, we generate 100 simulated datasets (using 100 different seeds for random number generation), each having the same true parameter values and (real) exposure and covariates but different synthetic outcome variables.

C.3.2 ESTIMATING THE EFFECT CURVE AND CONFIDENCE INTERVAL

When implementing our ADT estimation procedure on each simulated dataset, we correctly specify the outcome models’ mean function but estimate the propensity score models π_B and π_D using Ranger (the same algorithm as in our main analysis).

When applying the Spatial Weighted Bootstrap, we fit both Exponential and Power covariance functions (the latter has correlation function $\text{Corr}(m; a, z) = 1 - zm^a$) to the empirical variogram of the residuals from the LLKR stage and use whichever model has the smaller RMSE to generate the spatial weights. As noted in Section 3.3.4, the spatial structure of the LLKR residuals is different than the spatial structure of the outcome (here, $\tilde{Y}$), so traditional notions of correctly vs incorrectly specifying the spatial correlation model do not directly apply here.

C.3.3 EVALUATING THE PERFORMANCE OF THE SPATIAL AND NON-SPATIAL WEIGHTED BOOTSTRAPS

We compare our ADT curve estimates from each simulated dataset to the true ADT curve by computing the pointwise bias and MAE, which are then summarized by averaging across all 100 simulation



Figure C.12: Examples of the ADT estimate from our multiply robust estimation procedure applied to (four instances of) simulated data from our base case scenario, compared to the true ADT curve.

replicates within each scenario. We then calculate the pointwise coverage and width of the 95% confidence intervals computed using the Spatial Weighted Bootstrap, and compare the results with those from a Non-spatial Weighted Bootstrap.

C.3.4 SIMULATION RESULTS AND DISCUSSION

The bias and MAE results are shown in Figures C.13 and C.14, respectively. The confidence interval coverage and width results are shown in Figures C.15 and C.16, respectively.

First, we observe that the magnitude of the bias ranges between 0.1 and 0.5, while the true ADT ranges between 0.1 and 2.6, with an average value of 1.9. From the MAE plot, we see that the estimator is less stable on the tails (resulting in larger errors), which is to be expected for methods like LLKR. Our intuition is that the asymmetry of the true ADT curve contributes to the bias being higher at higher values of port proximity.

Note that in addition to the finite sample bias of LLKR, which results in imprecise centering of the ADT's confidence intervals⁷¹, here we also do not know the correct specification of the propensity score models π_B and π_D , instead estimating them from the data. Foreshadowing our coverage results, the simulations in Hettinger et al.⁷¹ indicate that misspecification of the propensity score models has

a greater impact on coverage than misspecification of the outcome models.

In Figure C.15, we see that neither the Spatial nor Non-spatial Weighted Bootstrap achieves 95% coverage at any value of δ (port proximity). However, the Spatial Weighted Bootstrap has consistently higher coverage than the Non-spatial method. This makes sense, as the Spatial method generates wider confidence intervals (as shown in Figure C.16). Mirroring the results of the MAE plot, we see that the confidence interval coverage tends to be highest in the lower-middle range of δ (0.25-0.5) for both the Spatial and Non-spatial Weighted Bootstrap.

Now, we compare the differences in results across our four simulation scenarios. Note that both Scenario 2 (increasing L) and Scenario 3 (increasing σ_{sp}) generate greater similarity in the value of outcomes that are spatially proximate, through different mechanisms, relative to Scenario 1. Whereas, Scenario 4 simply increases the random noise in the outcomes.

In Figures C.15 and C.16, we see that inducing greater spatial signal (Scenarios 2 and 3) increases the width of the Spatial Weighted Bootstrap's confidence intervals, resulting in higher coverage. By contrast, the Non-spatial approach responds to increasing L by narrowing its confidence intervals (resulting in lower coverage) and responds to increasing σ_{sp} by widening its confidence intervals (resulting in higher coverage). In Scenario 4, the Spatial approach actually narrows its confidence intervals, likely because increasing σ_{ind} makes it harder to detect the spatially structured error, resulting in under-estimation of the spatial correlation.

In Figure C.14, the only parameter configuration that stands out is Scenario 3 (increased σ_{sp}), which results in higher MAE (lower precision in the point estimate) across values of δ . In Figure C.13, none of the configurations have consistently different bias.

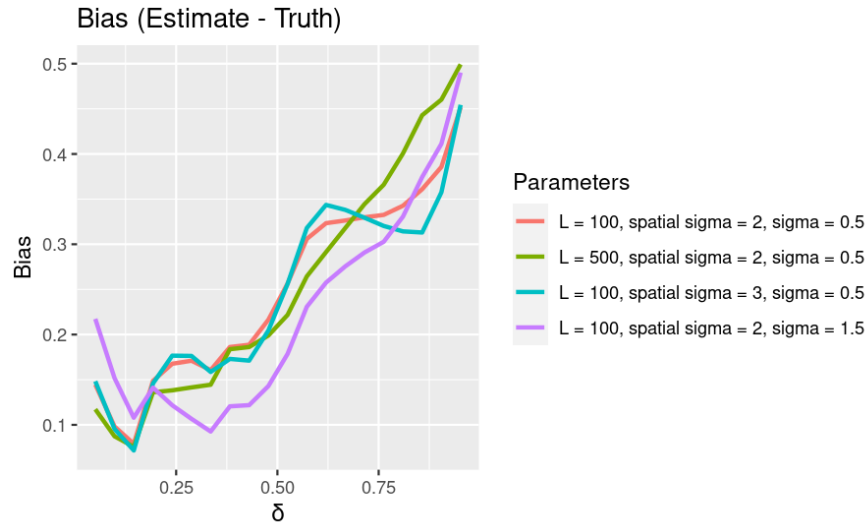


Figure C.13: Bias of the estimated ADT compared to the true ADT in our simulations.

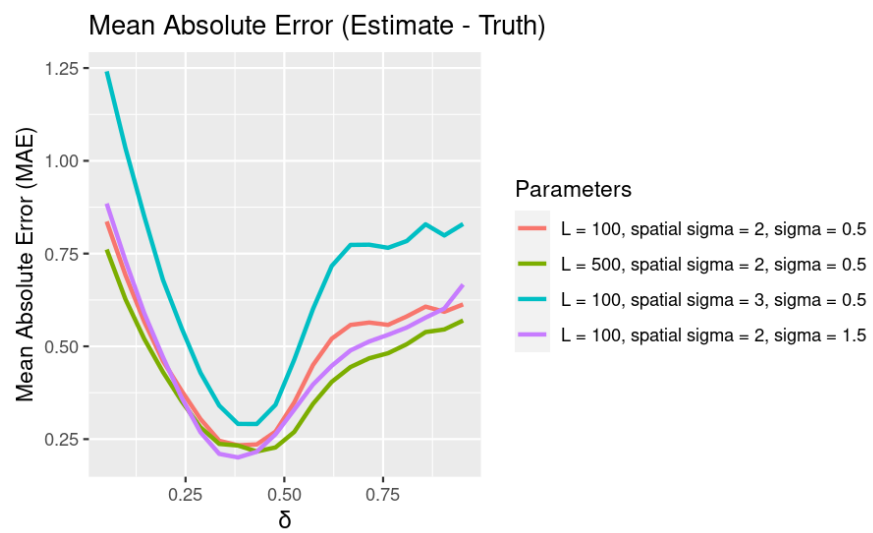


Figure C.14: Mean Absolute Error (MAE) of the estimated ADT compared to the true ADT in our simulations.

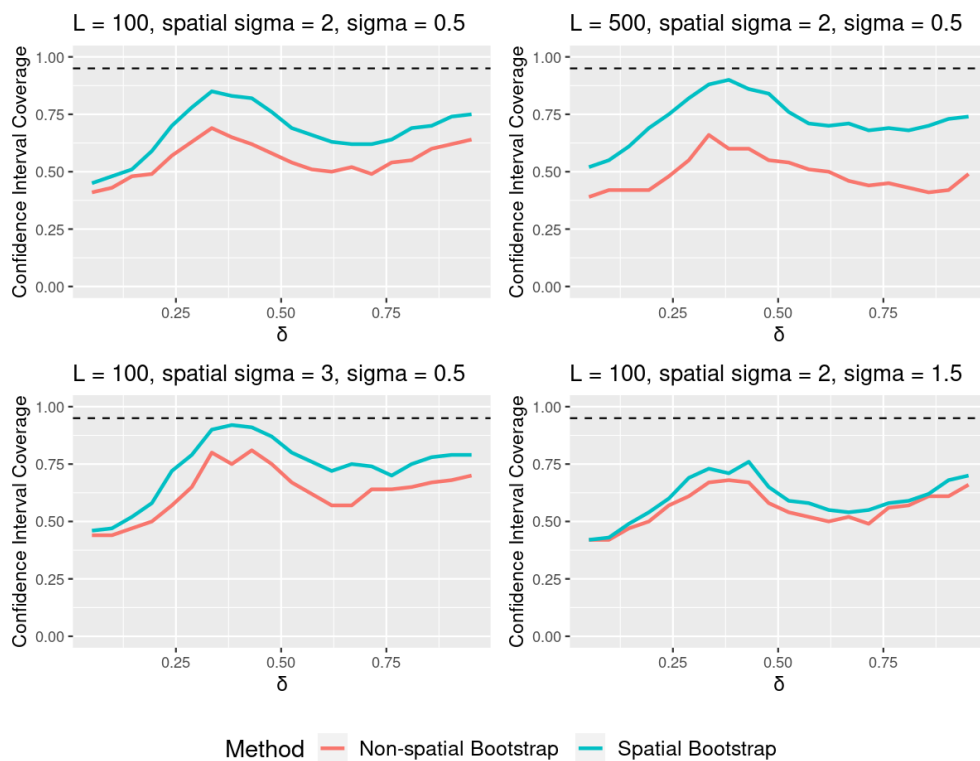


Figure C.15: Coverage of the 95% confidence intervals in our simulations. The dashed line indicates the desired 95% coverage.

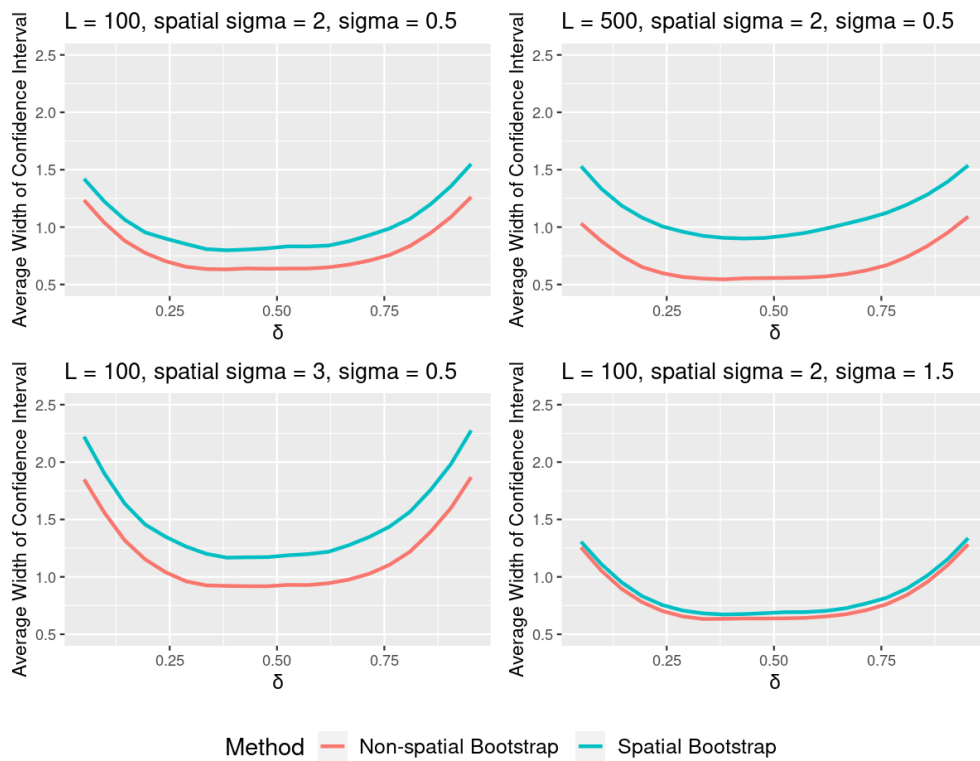


Figure C.16: Width of the confidence intervals in our simulations.

THIS THESIS WAS TYPESET using L^AT_EX, originally developed by Leslie Lamport and based on Donald Knuth's T_EX. The body text is set in 11 point Egenolff-Berner Garamond, a revival of Claude Garamont's humanist typeface. A template that can be used to format a PhD thesis with this look and feel has been released under the permissive MIT (X11) license, and can be found online at github.com/suchow/Dissertate or from its author, Jordan Suchow, at suchow@post.harvard.edu.