Traffic-Related Air Pollution and Dementia Incidence in a Seattle-Based, Prospective Cohort Study

Magali N. Blanco

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

Doctor of Philosophy

University of Washington
2021

Reading Committee:
Lianne Sheppard, Chair
Timothy Larson,
Sara Adar

Program Authorized to Offer Degree:
Department of Environmental and Occupational Health Sciences

University of Washington

Abstract

Traffic-Related Air Pollution and Dementia Incidence

in a Seattle-Based, Prospective Cohort Study

Magali N. Blanco

Chair of the Supervisory Committee:

Professor Lianne Sheppard

Department of Environmental and Occupational Health Sciences

Department of Biostatistics

Dementia has been considered a major global public health priority (WHO, 2012). It is very common in older adults (Alzheimer's Association, 2015b; Seshadri & Wolf, 2007; Weuve et al., 2014) and characterized by the progressive and irreversible loss of memory and mental abilities (Alzheimer's Association, 2008). Those affected often experience other comorbidities, disability and early death (Alzheimer's Association, 2019; Arrighi et al., 2010). No cure currently exists for progressive dementias, and the associated healthcare costs exceed those of other age-related conditions (Alzheimer's Association, 2015a).

Recently, animal and human studies have begun reporting on air pollution neurotoxicity (J. L. Allen et al., 2017), including dementia (Carey et al., 2018; Chang et al., 2014; Chen,

Kwong, Copes, Hystad, et al., 2017; Chen, Kwong, Copes, Tu, et al., 2017; Kilian & Kitazawa, 2018; Oudin et al., 2016; Power et al., 2016). Traffic-related air pollutants (TRAP) such as nitrogen dioxide (NO_2), ultrafine particulates (UFP) and black carbon (BC) are important components of community air pollution that can vary substantially over space and time (Fujita et al., 2014; Liu et al., 2018). TRAP exposure has been shown to be associated with neurotoxicity (Block & Calderón-Garcidueñas, 2009; Calderón-Garcidueñas et al., 2002, 2003; Calderón-Garcidueñas, Mora-Tiscareño, et al., 2008; Calderón-Garcidueñas, Solt, et al., 2008) and pathologies such as Alzheimer's Disease (AD) in animals (Yan et al., 2015, 2016) as well as cognitive deficits, including late-life dementia, though the evidence has been stronger for some pollutants than others (Carey et al., 2018; Chang et al., 2014; Chen, Kwong, Copes, Hystad, et al., 2017; Chen, Kwong, Copes, Tu, et al., 2017; Kilian & Kitazawa, 2018; Oudin et al., 2016; Power et al., 2011, 2016; X. Yu et al., 2020). In particular, research indicates that UFPs may play an important role in the adverse health effects associated with particulate matter (Brown et al., 2001; Donaldson, 2001; N. Li et al., 2003; Lundborg et al., 2001; Manigrasso et al., 2017; Oberdorster et al., 1994; Seaton et al., 1995; Stone et al., 2000). Still, epidemiologic studies investigating dementia and long-term TRAP exposure are limited due to the absence of models that appropriately capture long-term human exposure to TRAP (Rivera et al., 2012). This study addresses this gap in the literature through three specific aims:

In Aim 1, we use fine-scale, long-term NO_2 exposure as well as road proximity to assess the association between TRAP and late-life all-cause and AD dementia incidence in a community-based prospective cohort study. This study was conducted using the Adult Changes in Thought (ACT) cohort, a well-characterized, Seattle-based, prospective cohort study of aging and the brain among elderly individuals (65+ years) that has been ongoing since 1994 (Kukull et

al., 2002; L. Wang et al., 2006). Participants were assigned long-term NO_2 exposure based on a spatiotemporal model that incorporates decades of local air quality monitoring data based on residential history. Our primary analyses indicated that for every additional 5 ppb increase in 10-year average NO_2 exposure, the hazard of all-cause and AD dementia is estimated to be 1% (HR: 1.01, 95% CI: 0.91, 1.11) and 2% (HR: 1.02, 95% CI: 0.91, 1.13) greater, respectively, after adjusting for important potential confounders. Sensitivity and secondary analyses investigating the impact of different exposure windows, model adjustments, exposure quality and more were in agreement, supporting the robustness of our results. These findings are in line with the literature and a recent meta-analysis indicating that there is no evidence of an association between NO_2 and dementia incidence.

In Aim 2, we leverage a highly innovative mobile monitoring campaign specifically designed to assess spatially-granular, long-term TRAP exposure for the ACT cohort (Blanco et al., 2019; Stanley, 2019) to characterize otherwise unavailable annual-average UFP and BC exposure. We calculate weighted UFP and BC averages from repeated short-term monitoring samples and use these to build universal kriging models with partial least squares regression to summarize hundreds of geographic covariate predictors. The hold-out model validation results indicated low model bias and high precision (RMSE: 933 pt UFP/cm^3, 58 ng BC/m^3; R^2: 0.87 for UFP, 0.85 for BC). Predicted annual average UFP and BC exposure for ACT cohort locations had a median (IQR) of 6,782 (1,788) pt/cm^3 and 525 (134) ng/m^3, respectively. Similar to past studies, predicted concentration were highest near the downtown, industrial and airport areas as well as along major highways. Sensitivity analyses taking different approaches for dealing with extreme observations, calculating annual averages and building models all resulted in very similar results, strengthening the robustness of these exposure models. These findings support

the use of these prediction models for future epidemiologic investigations of TRAP exposure in the ACT cohort.

Aim 3 extends the exposure surfaces developed in Aim 2 for 2019 back to 1995 in order to characterize otherwise unavailable, spatially granular, long-term BC and UFP exposure for the ACT cohort. We use time-varying values of emission indicators (highway emissions) and surrogates (population density and green space; hereafter referred to jointly as "indicators") known to be strongly associated with TRAP along with observations of air pollution trends over time to extrapolate model predictions back in time. We validate models against historical observations at air monitoring sites. Results from these models showed that annual average BC and UFP exposure estimates for the ACT cohort were generally higher and more variable for earlier years. Locations near Seattle and along major roadways saw the sharpest drops in BC levels, while locations near the Sea-Tac Airport saw the sharpest drops in UFP levels over time. Models captured overall spatial and temporal pollution trends, though they were conservative and underpredicted observed concentrations at AQS sites. These models provide an understanding of how these otherwise poorly characterized pollutants may have changed over time in the Puget Sound, an important gap in the field.

Until now, investigations of TRAP exposure have been largely limited to short-term human exposure and animal studies despite the growing body of evidence linking some TRAPs to brain health. In one of the first truly long-term epidemiologic studies of TRAP exposure, we found no evidence that elevated levels of long-term NO_2 exposure is associated with an increased risk of late-life dementia incidence. Furthermore, we are one of the first to build annual-average UFP and BC exposure models from a novel and extensive mobile monitoring campaign specifically designed to assess exposure in a long-standing, community-based, prospective

cohort study of aging and the brain. These models can be used to further advance the field and support epidemiologic investigations of dementia incidence and long-term TRAP exposure, including UFPs and pollutant mixtures.

Table of Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

Acronym/Abbreviation	Definition
AADT	Annual average daily traffic
ACT	Adult Changes in Thought
AD	Alzheimer's Disease
ANOVA	Analysis of variance
AP	Air pollution
APOE	Apolipoprotein E
AQS	Air Quality System (EPA)
BC	Black carbon
CASI	Cognitive Abilities Screening Instrument
CI	Confidence interval
CO	Carbon monoxide
CO_2	Carbon dioxide
CV	Cross validation
EC	Elemental carbon
EF	Emission factor
EPA	Environmental Protection Agency
Geocovariate	Geographic covariate
HR	Hazard ratio
IPW	Inverse probability weighting
IQR	Interquartile range
LUR	Land use regression
MESA Air	Multi-Ethnic Study of Atherosclerosis and Air Pollution
MOVES	MOtor Vehicle Emissions Simulator (EPA)
NDVI	Normalized Difference Vegetation Index
NO	Nitrogen monoxide
NO_2	Nitrogen dioxide
NOx	Oxides of nitrogen
PLS	Partial least squares
$PM_{2.5}$	Particulate Matter with aerodynamic diameter less than 2.5 μm
PSCAA	Puget Sound Clean Air Agency

R^2	Coefficient of determination
RMSE	Root mean square error
SD	Standard deviation
SES	Socioeconomic status
TOD	Time of day (e.g., TOD2 is for time of day binned into two categories)
TOW	Time of week (e.g., weekday vs weekend; TOW7 is for seven days of the week)
TOW7 TOD21	Time of week (7 days) and time of day (21 hours) adjusted annual averages
TRAP	Traffic-related air pollution
Trimmed mean	Mean calculated after trimming extreme observations (e.g., those below 5^{th} quantile and above the 95^{th} quantile)
UFP	Ultrafine particulate
UK	Universal kriging
UW	University of Washington
Windsorized mean	Mean calculated after setting extreme observations (e.g., below 5^{th} quantile or above the 95^{th} quantile) equal to a lower threshold (e.g., the 5^{th} or 95^{th} quantile)
WSDOE	Washington Department of Ecology
WSDOT	Washington Department of Transportation

Acknowledgements

I would like to acknowledge the many people who contributed to and supported this research.

Doctoral Supervisory Committee. First and foremost, I would like to thank my Committee members Dr. Lianne Sheppard, Dr. Edmund Seto, Dr. Timothy Larson, Dr. Sara Adar and Dr. Julian Marshall for supporting this work through their expertise in environmental exposure science, air pollution epidemiology and quantitative analyses. A special thank you to Lianne for being a role model of excellent science and a dedicated mentor throughout my development as a researcher. She supported and guided my work; gave me her invaluable and honest input; was engaged in my coursework; and continually encouraged my professional development.

The Sheppard Lab and field team. An additional thank you to the Sheppard Lab (Amanda Gassett, Gail Li, Roy Pardee, Rachel Shaffer, Chris Zuidema, Nancy Carmona and Annie Doubleday) for providing me with feedback and resources along the way. I am especially grateful to Amanda for her hands-on support on the survival analysis in Aim 1 and the mobile monitoring campaign, which was foundational for Aims 2-3. Furthermore, Aims 2-3 of this work would not have been possible without our dedicated mobile monitoring field team: Tim Gould, Jim Sullivan and Dave Hardie, who helped us implement this campaign and drove thousands of miles over the course of a year (in a hybrid) to collect air pollution data.

School of Public Health faculty and staff. There were countless other faculty and staff in the School of Public Health who took time out of their day to counsel me and improve this research, including: Brian High, John Yocum, Elena Austin, Marco Carone, my professors and many others.

Family and friends. Thank you to my family and friends who encouraged and showed interest in my work. A special thank you to my mom, for always putting my education first and for allowing me to be me. To my partner, Stacy, for having an open heart, being steady in my life, inspiring me to follow my dreams, and motivating me when things get tough.

Funding. Finally, this work would not be possible without the substantial generosity of various funding sources: the Adult Changes in Thought – Air Pollution (ACT-AP) Study (NIEHS, NIA, R01ES026187); BEBTEH: Biostatistics, Epidemiologic & Bioinformatic

Training in Environmental Health (NIEHS, T32ES015459); the Achievement Rewards for College Scientists (ARCS) Foundation Fellowship; the Latino Center for Health Fellowship; and the Pacific Northwest Section of American Industrial Hygiene Association (AIHA) Scholarship.

Dedication

To Stacy – my foundation and better half.

Introduction

Dementia

Dementia has been considered a major public health priority and has received significant global attention (WHO, 2012). Dementia is very common in older adults (Alzheimer's Association, 2015b; Seshadri & Wolf, 2007; Weuve et al., 2014) and characterized by the progressive and irreversible loss of memory and mental abilities that affect daily life (Alzheimer's Association, 2008). Those affected may experience more memory loss, difficulty finding words, confusion and disorientation, along with other comorbidities such as depression, anxiety, paranoia and hallucinations (Mayo Clinic, 2019). In addition to poor health, it is often associated with disability, oftentimes requiring individuals to live in a nursing home (Arrighi et al., 2010). Those affected are at higher risk of death, with people age 65 or older, for example, living an average of four to eight years after an Alzheimer's Disease (AD) diagnosis (Alzheimer's Association, 2019). No cure currently exists for progressive dementias, and the associated healthcare costs exceed those of other age-related conditions (Alzheimer's Association, 2015a). In 2019, dementias cost the United States about $290 billion in health care, long-term care, and hospice (Alzheimer's Association, 2019).

There are several risk factors for dementia (Alzheimer's Association, 2019, 2020). Age is the largest risk factor, with those over age 65 being at higher risk. The risk of AD, for example, doubles every five years after age 65 until almost one-third of individuals are affected by 85 years of age or older. Those with a family history of dementia and genetic carriers of the apolipoprotein E (APOE) ϵ4 allele are more likely to develop dementia. Heart health has been linked to brain health since it is responsible for pumping large quantities of blood to the brain through complex blood vessel networks. Conditions that damage this network (e.g., heart disease, diabetes, stroke, high blood pressure and high cholesterol) may thus increase the risk of cognitive decline. Healthy aging has been linked to brain health, including greater cognitive reserve and resilience (Alzheimer's Association, 2020). Having a poor diet, a lack of exercising, being overweight, consuming tobacco and excess alcohol may increase the risk of dementia. Gender, education and socioeconomic status (SES) have also been associated with a higher risk

of dementia incidence, with females, those with lower education and lower SES being at higher risk.

Dementia typically progresses slowly and can take decades to develop (Howell et al., 2017; Matthews et al., 2019; Weuve et al., 2018). It is often the result of multiple neurodegenerative disorders including AD, vascular contributions, frontotemporal disorders, Lewy body dementia, and others (Brenowitz et al., 2017; Jack et al., 2018). AD is involved in most dementia cases and has been most widely investigated. The Amyloid Cascade Hypothesis is the most prominent theory attempting to explain the pathogenesis of AD (Alzheimer's Queensland, 2014). An initiating event is hypothesized to occur while individuals are still cognitively unimpaired and involves the abnormal processing of ß-amyloid (Aß) peptide. This leads to the formation of Aß plaques in the brain and, after some period, neurodegeneration is initiated. This process is marked by increased cerebrospinal fluid (CSF) tau and cerebral atrophy (the loss of neurons). Years later, cognitive impairment is finally detected in individuals, first as mild cognitive impairment, followed by dementia. As such, the etiologically relevant exposure periods of interest may occur years or decades before the first signs of cognitive deficits are detected in individuals.

By 2030, the number of older Americans age 65 or older is projected to reach 74 million, or 20% of the total U.S. population (Alzheimer's Association, 2019). Given the tremendous burden that dementia places on individuals and the healthcare system, governments have become increasingly concerned about the likely increase in dementia cases (Wu et al., 2016).

Scientific recommendations currently highlight the importance of primary (reduce the risk and increase cognitive reserve), secondary (early detection) and tertiary (after dementia is present) interventions, with primary interventions being the most effective (Wu et al., 2016). Some work, for example, has indicated that the incidence of dementia in high-income countries may have declined over time, and that this could be explained by rising education levels along with improved controls for adverse cardiovascular health outcomes (Alzheimer's Association, 2019). This suggests that societal changes resulting from primary prevention efforts may effectively reduce the incidence of dementia.

Traffic-Related Air Pollution and its Effects on the Brain

Traffic-related air pollution (TRAP) is a complex mixture of gases and particulate matter (PM) that greatly contributes to overall air pollution (AP). Nitrogen oxides (Chen, Kwong, Copes, Tu, et al., 2017; Clifford et al., 2016; Karner et al., 2010; Oudin et al., 2016; Power et al., 2016; Ranft et al., 2009; Tzivian et al., 2014) (NOx, composed of nitrogen monoxide [NO] and nitrogen dioxide [NO_2]), ultrafine particulates (UFP or $PM_{0.1}$) and black carbon (BC) (Clifford et al., 2016; HEI, 2010; Power et al., 2016) have been identified as important TRAPs. NO is a gas that is emitted from tailpipe emissions that quickly reacts with oxygen in the atmosphere to form NO_2 (Sternberg, 2015). BC is a significant component of the PM that is emitted from gas and, in particular, diesel engines as a result of the incomplete combustion of fuel (US EPA, 2019c). Particles are considered UFPs if they are less than 100 nm (0.1 μm) in diameter. Many vehicle emissions are in this size range (Hatzopoulou et al., 2013; Weichenthal et al., 2015), with diesel engines emitting higher levels of UFPs than gasoline engines (US EPA, 2018b).

While extensive documentation has linked AP to adverse cardiovascular and pulmonary outcomes (J. L. Allen et al., 2017), studies have also begun to report on TRAP's neurotoxicity. Epidemiological investigations have linked elevated exposure levels of NOx, NO_2, BC, UFPs and other traffic pollutants (Babadjouni et al., 2017; Brunekreef & Holgate, 2002; Carey et al., 2018; Chang et al., 2014; Chen, Kwong, Copes, Hystad, et al., 2017; Chen, Kwong, Copes, Tu, et al., 2017; Clifford et al., 2016; de Prado Bert et al., 2018; Heusinkveld et al., 2016; Howard, 2009; Kilian & Kitazawa, 2018; Killin et al., 2016; Kulick et al., 2020; Oudin et al., 2016; Power et al., 2011, 2016; Seaton et al., 1995; US EPA, 2018b; Weichenthal et al., 2020; Yuchi et al., 2020), to an increased risk of late-life dementia incidence, worse cognitive performance and other neurological effects. Studies have reported hazard ratios (HR) for dementia incidence ranging from about 0.97-1.20 for a 5 ppb increment of NO_2 or NOx exposure in areas with exposure levels similar to the Puget Sound region (Carey et al., 2018; Chen, Kwong, Copes, Hystad, et al., 2017; Chen, Kwong, Copes, Tu, et al., 2017; Kilian & Kitazawa, 2018; Oudin et al., 2016; Power et al., 2016).

There are two proposed pathways for how particles reach the brain (Block & Calderón-Garcidueñas, 2009; Calderón-Garcidueñas, Solt, et al., 2008; González-Maciel et al., 2017; Kilian & Kitazawa, 2018). The first is an indirect pathway where small particles (e.g., BC and

UFPs) are inhaled deep into the lung, enter the blood stream and eventually target other organs including the brain. The second is a direct pathway where very small particles (e.g., UFPs) pass through the olfactory epithelium in the nose and are directly deposited into the brain. Mechanisms of particle-induced neurotoxicity include oxidative stress, metal homeostasis, protein homeostasis, neuroinflammation, ß-amyloid deposition, disruption of the blood brain barrier and activated microglia (Block & Calderón-Garcidueñas, 2009; Calderón-Garcidueñas et al., 2002, 2003; Calderón-Garcidueñas, Mora-Tiscareño, et al., 2008; Calderón-Garcidueñas, Solt, et al., 2008; Heusinkveld et al., 2016). Given their small size, UFPs may thus play an especially important role in the adverse health effects associated with PM (Manigrasso et al., 2017). Compared to larger particles, UFPs can additionally adsorb other toxic substances (e.g., polycyclic aromatic hydrocarbons), are less effectively phagocytized and cleared by alveolar macrophages, may induce more oxidative stress and mitochondrial damage, and can cause a greater inflammatory response (Brown et al., 2001; Donaldson, 2001; N. Li et al., 2003; Lundborg et al., 2001; Oberdorster et al., 1994; Seaton et al., 1995; Stone et al., 2000).

As for NOx, the NO_2 fraction has been the most strongly associated with adverse health effects (US EPA, 2016). While the exact neurological mechanisms are unclear, animal studies have shown that NO_2 inhalation negatively affects spatial learning and memory, increases amyloid β_{42} accumulation, promotes AD-like pathologies (Yan et al., 2016), and induces morphological and energy metabolism changes in brain mitochondria (including decreased ATP-production and increased reactive oxygen species) that may induce neurological disorders (Yan et al., 2015).

Limitations of Epidemiologic Investigations of Dementia

Epidemiologic investigations of long-term TRAP exposure and dementia incidence have been significantly limited by a lack of historical, spatially-granular TRAP data or exposure models (Power et al., 2016; Rivera et al., 2012). Traditionally, air pollution studies have largely relied on the EPA's AQS regional air monitors and/or study specific monitoring campaigns to collect air pollution data and aid in their assessment of human exposure. AQS monitors have generally collected pollutant concentrations for decades, but they are spatially sparse and measure few traffic-related pollutants that do not include UFPs (Özkaynak et al., 2013). On the

other hand, study-specific monitoring campaigns tend to be spatially denser and collect a wider array of pollutants, but they are generally short-term.

Because TRAP can vary substantially between microenvironments and over time (Enayati Ahangar et al., 2017; Fujita et al., 2014; Karner et al., 2010; Liu et al., 2018), studies that utilize traditional regional monitoring methods may misclassify exposure. This is particularly true for UFPs, BC and NOx because their concentrations can quickly decay with increasing distance from a source due to chemical reactions and dispersion (Karner et al., 2010; Patton, Perkins, et al., 2014; Riley et al., 2014). In particular, UFPs tend to be more unstable than other pollutants (Karner et al., 2010). After being produced, small particles routinely undergo chemical transformations, adsorb to surfaces, evaporate, coagulate (stick to each other) and undergo condensation, often times growing out of the ultrafine size range (Zhu et al., 2002). Locations roughly 300-1,000 m downwind of sources may thus have similar concentrations to background levels (US EPA, 2018b; Zhu et al., 2002). As for NO_2 and BC, recent work has indicated that concentrations may vary by a factor of 2-8 within neighborhoods and individual blocks (Apte et al., 2017).

Similarly, many long-term exposure studies have not fully accounted for temporal trends in TRAP, using limited sampling designs (e.g., business hours) and/or short-term exposure estimates (e.g., one year) as surrogates for long-term exposure, which may lead to exposure misclassification (Power et al., 2016). Sampling during limited time periods may produce biased long-term averages since samples may not capture the diurnal, seasonal, and longer-term TRAP exposure trends. Exposure, for example, can change in the short-term over the course of a day and between seasons, following meteorological cycles and vehicle activity. Exposures can also change in the long-term as a result of the progressively lower TRAP levels and due to new residential addresses. TRAP levels, for example, began to decrease more sharply after the signing of the 1990 Clean Air Act amendments and after 2007 when new heavy-duty diesel highway engine requirements were placed in order to reduce PM an NOx emissions (Khalek et al., 2015; Ruehl et al., 2015; US EPA, 2017b, 2018d, 2018a).

Furthermore, the use of broader diagnostic criteria, the increased likelihood of patients receiving medical attention and changes in education levels have likely changed dementia incidence and diagnoses over time (Wu et al., 2016). Most studies, however, have assessed dementia incidence based on medical records rather than regular, in-person visits, which may

introduce a substantial amount of outcome misclassification given that dementia tends to be underdiagnosed (National Institute of Aging, 2014). This may result in numerous false positive and false negatives (Power et al., 2016). Furthermore, while medical record studies can allow for larger sample sizes, these may be less practical and generalizable in areas without universal health care access such as the United States. Given that socioeconomic status is associated with both air pollution exposure and a dementia diagnosis through healthcare access, this may also result in biased estimates.

A 2016 review by Power et al. identified a number of studies that have investigated "long-term" air pollution exposure (1+ years) and dementia incidence or poor cognition (Power et al., 2016). They reviewed 18 published articles from around the globe published through 2015, most of which found adverse dementia-related outcomes in humans associated with air pollution. Specifically, there were three studies that had investigated long-term TRAP exposure and dementia incidence. These studies reported an increased risk in dementia or AD incidence associated with higher long-term NO_2, CO (Chang et al., 2014), ozone (C. R. Jung et al., 2015) and NOx (Oudin et al., 2016) exposure. The designs of these studies highlight common exposure assessment approaches in the field and stress the need for truly long-term, spatially granular exposure assessment in epidemiologic investigations of dementia. Chang et al. (2014), for example, assessed participant exposure for a duration of up to 12 years, but participants were assigned a fixed exposure based on their entire follow-up duration (i.e., the averaging period was longer for participants who were followed longer; and future exposure was used to characterize risk at earlier times). Moreover, exposure based on a sparse network of regulatory monitors. Jung et al. 2015 similarly assigned participants a fixed exposure, this time at baseline, despite the fact that participants were followed for up to 10 years. Exposure was based on three regulatory monitors within 25 km and a participant's postal code. Oudin et al. (2016) also assigned participants a fixed exposure, this time based on a participant's last year of follow-up. Only Oudin et al. (2016) assessed dementia incidence based on regular (every five years), in-person assessments of cognition rather than medical records.

Recently, Yu et al. (2020) published a meta-analysis of NO_2, NOx and dementia incidence based on a slightly larger set of longitudinal cohort studies (X. Yu et al., 2020). They reported a positive pooled relative ratio (RR) of 1.02 (95% CI: 0.99, 1.04) for a 5 ppb increase in NO_2 or NOx, though confidence intervals were consistent with both positive and negative

effects. Larger, (though not necessarily better designed studies in terms of exposure and outcome assessment) received a larger weight in the pooled RR. As such, this analysis was again limited to the designs of the existing studies in the field.

Mobile Monitoring Campaigns

In an effort to address the exposure assessment limitations of many of past studies, study-specific mobile monitoring campaigns have become more widely used in cities throughout the world (Apte et al., 2017; Hatzopoulou et al., 2017; Kerckhoffs et al., 2016; Patton, Perkins, et al., 2014; Van den Bossche et al., 2015; Xie et al., 2017). Typically, a vehicle is equipped with air monitors capable of measuring pollutants with high spatial and temporal resolution, and short-term sampling repeatedly occurs at predefined locations. By collecting short-term samples over extended periods of time, campaigns can capture the temporal variability of a location and calculate unbiased long-term averages without the need for continuous fixed site monitoring (Apte et al., 2017; Hatzopoulou et al., 2017).

Mobile monitoring campaigns have several advantages over traditional fixed site monitoring. For one, they can sample locations that may be more complex and representative of human exposure (e.g., near intersections). Moreover, mobile monitoring campaigns can be more efficient than fixed site monitoring since they can sample many locations within a relatively short period of time using a single monitoring platform. This has allowed campaigns to use high quality instrumentation to measure various pollutants simultaneously.

Campaigns have conducted short-term *stationary* monitoring where the platform collects samples while stopped, and/or short-term *mobile* monitoring where the platform collects samples while in motion. Most short-term stationary campaigns have collected 15-60 minutes of samples during 1-3 days per location, while short-term mobile campaigns have generally collected 1-300 seconds of measurements per road segment and used these to estimate long term averages (Apte et al., 2017; Saha, Li, et al., 2019). Campaigns have generally lasted anywhere from a few weeks to a few months. Campaigns generally calculate site means or medians as a measure of that site's long-term central tendency.

A few studies have investigated how many sampling locations and repeat samples should be collected to improve the robustness of long-term average prediction models (Hatzopoulou et

al., 2017; Messier et al., 2018). Messier et al. (2018) found that daytime annual averages in Oakland, CA could be accurately estimated with only a fraction of all roads within a domain (~30%) as long as these were representative of the area as a whole (Messier et al., 2018). They further concluded that as few as four driving days could accurately estimate the daytime annual average, with increasing drives resulting in diminishing returns in terms of model performance. Hatzopoulou et al. (2017) found that about 150-200 locations, each with about 10-12 visits was sufficient to produce robust LUR models for Montreal, Canada (Hatzopoulou et al., 2017). These studies suggest that robust annual average estimates can be calculated from mobile monitoring campaigns as long as: a) the sampling locations are representative of the study area, and b) sampling locations are visited roughly 4-12 times.

While mobile monitoring has shown promise in better characterizing the spatial variability of TRAP when compared to traditional monitoring methods, the existing campaigns may not be appropriate for epidemiologic investigation. For one, many campaigns have solely measured from the center of roadways while in motion where concentrations may be much higher than residential dwellings meters away (Alexeeff et al., 2018; Kerckhoffs et al., 2016). Some evidence has suggested that on-road samples are associated with residential exposure and adverse health outcomes, though this research is limited (Simon et al., 2017). Moreover, campaigns have sampled either high concentration areas or locations that are representative of a city's geography (e.g., highways, industrial and commercial areas), which are likely not representative of where the majority of people spend most of their time. Some work, for example, has indicated that pollution levels differ substantially among major road classes, with median BC, NO and NO_2 concentrations being roughly 2-5 times higher on highways than residential streets (Apte et al., 2017). Third, most studies have conveniently only conducted sampling during weekday business hours since a technician is required to operate the platform. These sampling methodologies may result in biased results since TRAP levels may be significantly different at night and on weekends due to changes in meteorology and traffic patterns (Batterman et al., 2015). UFP concentrations, for example, have been shown to be higher during morning and afternoon hours than midnight hours (Saha, Li, et al., 2019). Weekends also tend to have lower TRAP concentrations from reduced vehicle counts. Fourth, most campaigns have sampled over the course of a few weeks or months (Hoek et al., 2008; Xie

et al., 2017). This is likely to result in biased and imprecise long-term estimates, limiting the application of these campaigns for long-term exposure studies.

Land Use Regression for Exposure Assessment

Regardless of the method used to collect air pollution samples, epidemiologic studies in the field most often use land use regression (LUR), the use of geographic covariate (geocovariate) predictors such as land use and population density in linear regression, to assign participants a predicted exposure level (Hatzopoulou et al., 2017; Minet et al., 2017; Weichenthal, Ryswyk, et al., 2016). LUR models have been used to successfully predict the spatial variability of many TRAPs (including UFPs, BC and NO_2) in cities throughout the world (Bellander et al., 2001; Hystad et al., 2012; Kerckhoffs et al., 2016, 2017; Lane et al., 2016; Minet et al., 2018; Montagne et al., 2015; Patton et al., 2015; Patton, Collins, et al., 2014; Simon et al., 2018; Xie et al., 2017; C. H. Yu et al., 2016; Zwack et al., 2011). Advantages of LUR models include their ability to capture the variation of air pollutants at fine spatial scales, and the ability to easily build these given a set of geocovariates (Hystad et al., 2012; Jerrett et al., 2005; Lindhjem et al., 2012; Patton et al., 2017; Sharma et al., 2004).

Various studies have used pollutant emission estimates from models such as the US EPA's Motor Vehicle Emissions Simulator (MOVES) (Koupal et al., 2011; US EPA, 2018c) as predictors in LUR models. MOVES is regularly used by regulatory agencies to estimate roadway emission factors (EF, g/vehicle-mile) for specific years and counties based on important predictors of vehicle emissions including vehicle type and age, road type and speed, fuel composition, meteorological conditions and other factors (US EPA, 2017a; WSDOE, 2018). These emission factors are combined with the number of vehicles on roadways to calculate total emissions. Validation studies have shown good model accuracy (US EPA, 2012), suggesting that MOVES may appropriately capture long-term temporal patterns in vehicle emissions when other local monitoring data are not available.

There are four common approaches have been used to extrapolate exposure surfaces backwards or forwards in time in order to develop spatiotemporal models (Bellander et al., 2001; Hystad et al., 2012; S. Y. Kim et al., 2017; Levy et al., 2015; Meng et al., 2019; Mölter et al., 2010b; Montagne et al., 2015; R. Wang et al., 2013). These approaches have been used either on

their own or in combination. The first involves fitting a LUR model with a set of covariate predictors and afterwards using earlier or later values of a given covariate to extrapolate model predictions to a different time. An advantage to this approach is that it can account for changing spatial contrasts over time (e.g., population density may change to different degrees at different locations over time) such that different sites receive their own local temporal adjustment. This approach, however, assumes that the same covariate-outcome relationship holds true over time. For example, that for every additional 1,000 people within a given radius, TRAP levels increase by X amount. The second approach for extrapolating model predictions over time involves adjusting model predictions using a global temporal trend adjustment, such that an entire exposure surface is moved up or down. This approach assumes that the spatial contrasts of an original exposure surface remain constant over time. Some studies, for example, have applied a temporal trend to a current exposure knowing that TRAP concentrations have steadily dropped over time to better estimate what past exposure levels may have been. The third approach for extrapolating LUR predictions involves fitting separate models for an earlier and a later time, each with time-relevant independent variables (e.g., TRAP observations for a specific year) and possibly covariate predictors (e.g., that year's population density). Independent variables can be interpolated from the original sampling times, and new, year-specific models can be refit in order to obtain recalibrated coefficients. Like the first approach, an advantage of this method is that prediction sites can each receive their own trend adjustment rather than a global adjustment (approach two). Finally, the fourth approach for extrapolating model predictions involves using air pollution concentrations for the same site but a different time period as a covariate predictor. This approach takes advantage of serial correlation. A disadvantage of the latter two approaches is that air pollution observations are needed for each site from two distinct periods, limiting these approaches, for example, to studies using long-term regulatory monitoring.

Preliminary Work

Given the critical importance of cognition in an aging population and the need to improve long-term air pollution exposure assessment in epidemiologic studies, our team has made several contributions to the field. For one, we are collaborators on the Adult Changes in Thought (ACT) study, a long-standing, Seattle-based prospective cohort study of aging and the brain. Second, we

have developed a spatiotemporal model for predicting long-term air pollution levels with high spatial resolution at exact participant residences (Cohen et al., 2009; Keller et al., 2015; Lindström et al., 2014; Sampson et al., 2011; Szpiro et al., 2010). Finally, we recently completed an extensive mobile monitoring campaign specifically designed to characterize annual average TRAP concentrations for epidemiologic application, with a specific focus on the ACT cohort (Blanco et al., 2019). The remainder of this section further describes these efforts.

The Adult Changes in Thought Study

The ACT Study is Seattle-based prospective cohort study of aging and the brain. ACT enrolled its first cohort between 1994-1996 from Kaiser Permanente enrollees and since then has continued to add participants through expansion (2000-2002) and replacement (2004+) cohorts (Kukull et al., 2002; L. Wang et al., 2006). Over 5,000 participants have enrolled to date, with roughly 2,000 active participants at risk during any given year. ACT has used the same methodologies and diagnostic criteria since the inception of the first cohort to identify cases through in-person visits. Cognitively unimpaired, elderly individuals (65+ years) are enrolled and assessed every two years for the incidence of dementia. Participants are administered the Cognitive Abilities Screening Instrument (CASI), which combines common screening tests including the Mini-Mental State Examination (MMSSE) and the Hasegawa Dementia Rating Scale to quantitatively assess attention, concentration, orientation, short- and long-term memory, language abilities and judgement, among other functions (Dublin et al., 2011). Participants with CASI scores lower than 86/100 or those for whom there is clinical concern trigger a full dementia diagnosis evaluation that includes a physical and neurological examination with laboratory testing, imaging and a battery of neuropsychological tests (Teng et al., 2004). Standard clinical criteria based on the Diagnostic and Statistical Manual of Mental Disorders (DSM) – IV are used to diagnose dementia (American Psychiatric Association, 1994; McKhann et al., 1984). Those who do not meet the criteria for dementia or AD continue to be followed biennially as dementia-free patients. Those who meet the criteria for dementia or AD are considered incident cases and tested the following year to verify clinical diagnosis. The final case status reflects this final annual test (i.e., initial cases can be corrected to be non-cases again). Cases are given a *diagnosis* date associated with their last biennial visit and an *onset* date

halfway between their last two biennial visits. Cases are no longer followed biennially. ACT additionally collects other relevant health information from participants such as genetic susceptibilities and lifestyle factors. To date, ACT has enrolled over 5,500 individuals.

Spatiotemporal Models

Our team has developed various air pollution prediction models that have been extensively used for predicting long-term air pollution levels at participant residences, including the ACT cohort (Cohen et al., 2009; Keller et al., 2015; Lindström et al., 2014; Sampson et al., 2011; Szpiro et al., 2010). Models can use regional air quality data, measurements taken at participant homes, data from "snapshot" campaigns that measure local near-field TRAP gradients, and hundreds of geocovariates such as indicators of traffic, population density and land-use (Table 1) (MESA Air, 2019). We have recently developed a highly spatially-resolved, spatiotemporal NO_2 exposure prediction model for ACT cohort locations in the spatiotemporal modeling area (Figure 1).

Table 1. Geocovariates Available from the MESA Air Geodatabase

Covariate	**Source**	**Years Collected**	**Description**
Airports	National Emissions Inventory Database	2000	• meters to airport • meters to large airport
Coastline	TeleAtlas	2000	• meters to coast
Railroads, railyards	TeleAtlas	2000	• meters to railroad • meters to railyard
Roads	TeleAtlas	1990, 2000, 2010	• meters to A1, A2, A3 road • length of A1, A2, A3 roads within a buffer (50 m – 5 km)
Intersections	TeleAtlas	2010	• meters to <type> intersection • number <type> intersection within buffer (500 m – 3 km)

Population	US Census Bureau	1990, 2000, 2010	• population count within radius (500 m - 15 km)
Land use, commercial land use	MRL National Landcover Dataset	2006	•meters to commercial and services land use •land use <type> (e.g., commercial, residential, urban, cropland, mixed forest, streams, beaches) within radius (50 m – 15 km)
Ports	National Geospatial Intelligence Agency	2002	• meters to small, medium, large port
Emission sources	National Emissions Inventory Database	2002	• short and tall stack pollutant emissions (NOx, SO_2, $PM_{2.5}$, PM_{10}, CO) within radius (3-30 km)
Truck routes	Bureau of Transportation Statistics	2009	• meters to truck route • length of truck route within radius (50 m – 15 km)
Impervious surface	National Landcover Dataset	2006	• percent of an area covered with an impervious surface (e.g., pavement, concrete) within radius (50 m – 5 km)
Elevation	National Elevation Dataset		• elevation (m) • elevation within radius (1-5 km)
Normalized Difference Vegetation Index (NDVI)	University of Maryland	2006	• NDVI (percentile: 25, 50, 75; seasonal: summer, winter) within a radius (250 m – 10 km)

See MESA Air Data Organization and Operations Procedures document Section 1.11 (geocovariates) and Table 10 for more details (MESA Air, 2019).

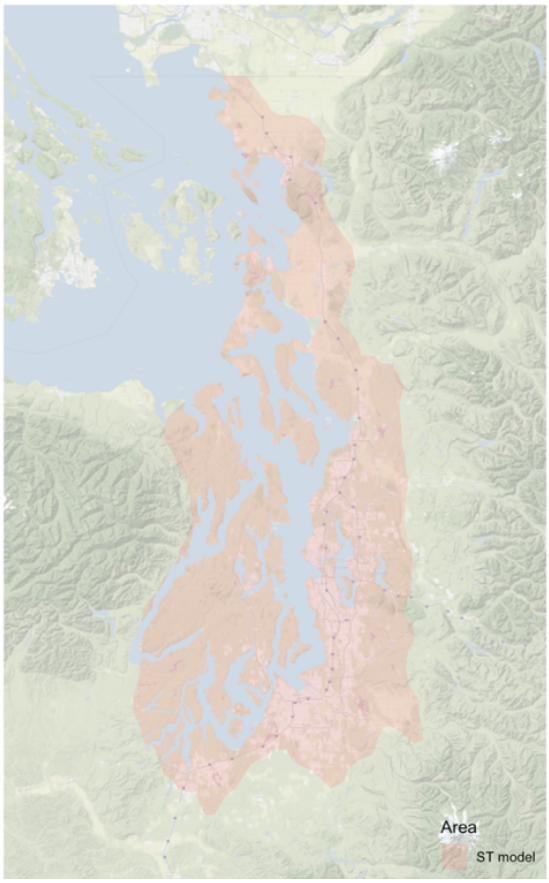


Figure 1. The spatiotemporal modeling area in the Puget Sound

Mobile Monitoring Campaign

Given that we do not have prediction models for many traffic-related pollutants, we designed an extensive mobile monitoring campaign to characterize the 2019 annual average TRAP exposure levels of the ACT cohort (Blanco et al., 2019). The campaign consisted of 309 stop locations off the side of the road, 5 of which were AQS sites (Figure 2). Stop sites in our monitoring area were visited an average of 27 times, each time for two minutes, between March 2019 – March 2020 (about 280 drive days) between the hours of 5 AM and 12 AM. Our goal was to sample each site during each season, time of week and time of day, though in the end not all sites were sampled during each time combination due to the logistical considerations of a mobile monitoring campaign. Results from a mobile monitoring simulation study using California AQS

site data indicated that both our planned and realized measurement approaches should have resulted in unbiased annual averages (Blanco et al., 2021).

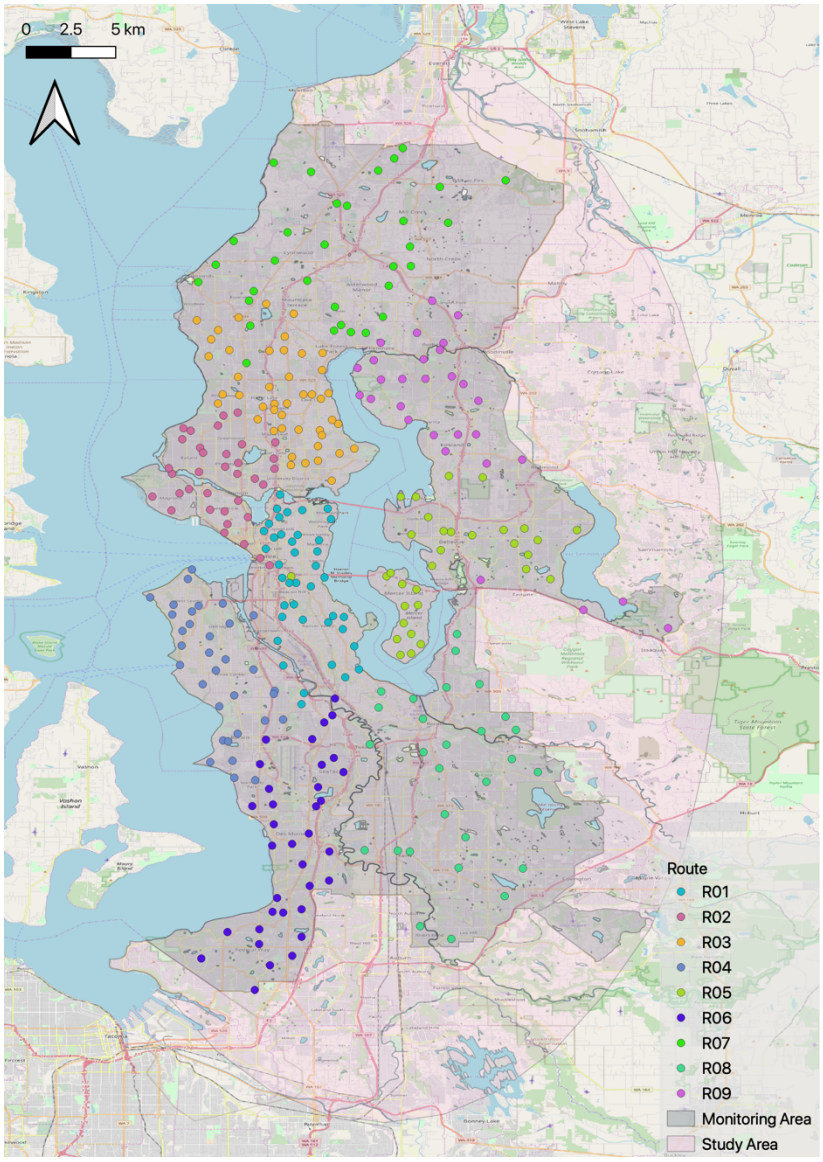


Figure 2. ACT-AP TRAP Mobile Monitoring stop locations (colored dots) within the monitoring (gray) and study area (pink).

Our mobile monitoring platform was equipped with high-quality instrumentation that measured various pollutants, including UFPs and BC, every 1-60 seconds (Table 1 - Table 2). UFPs were sampled using various instruments including PTRAK 8525 (20-1000 nm particles).

BC was collected from aethalometers (microAeth MA 200, Aethlabs) collecting measurements at an 880 nm wavelength (AethLabs, 2018; X. Wang et al., 2016).

Table 2. Air pollutants and other parameters measured using mobile monitoring. Bolded instruments indicate which were used in this work.

Parameter	Instrument	Manufacturer	Measurement Range	Limit of Quantification	Time Resolution
Particles[a]					
UFP					
10-420 nm (13-bin PSD)	NanoScan 3910	TSI	10^2-10^6 pt/cm^3	10 pt/ cm^3	60 sec
10-700 nm	DiSCmini	Testo	10^3-10^6 pt/cm^3	500-2,000 pt/cm^3 [b]	1 sec
20-1,000 nm	**PTRAK 8525**	**TSI**	**0 - 5 x 10^5 pt/cm^3**	**1 pt/cm^3**	**1 sec**
50-1,000 nm	PTRAK 8525, with diffusion screen	TSI	0 - 5 x 10^5 pt/cm^3	1 pt/cm^3	1 sec
BC	**microAeth MA200**	**AethLabs**	**0-1 mg BC/m^3**	**30 ng BC/m^3 [c]**	**10 sec**
Light scattering ($PM_{2.5}$)	M903	Radiance Research	0 - >1 km^{-1}	10^{-6} m^{-1}	10 sec
Gases					
NO_2	CAPS NO_2	Aerodyne Research, Inc.	0-2,000 ppb	2 ppbv	1 sec
CO_2	SenseAir CO2 K-30-FS sensor	CO2Meter.com	0-5,000 ppm (vol)	100 ppmv	1 sec

CO	CO Monitor T15N	Langan, Inc.	0-200 ppm	0.1 ppm	1 sec
Other					
Temperature	Onset UX100-011	HOBO	-4-158°F		1 sec
Relative humidity	Onset UX100-011	HOBO	0-95%		1 sec
Positioning & real-time tracking	DG-500	US GlobalSat	0-515 m/sec speed	2.5 m	1 sec

[a]PSD: particle size distribution; pt: particles

[b]estimate; detection limit is dependent on particle size

[c]for a 5 min time base, 150 ml/min flow rate

Study Aims

Our expertise in the field of long-term air pollution exposure assessment and late-life dementia positions us well to address important gaps in the field. In general, few studies have investigated long-term TRAP exposure and dementia. Among those that have, most have been significantly limited by a lack of sufficiently granular spatiotemporal TRAP data to appropriately characterize TRAP exposure. This body of work addresses these gaps twofold. First, we are one of the first to use fine-scale, truly long-term NO_2 exposure to assess the association between TRAP and dementia incidence. We hypothesized that elevated NO_2 exposure would be associated with a greater risk of dementia. Second, we characterize otherwise unavailable, fine-scale, long-term UFP and BC exposures for the ACT cohort by leveraging observations from our innovative mobile monitoring campaign and historical TRAP emissions data. The specific aims of this study are:

Aim 1: Investigate the relationship between individual-level, long-term TRAP exposure, using NO_2 as a surrogate, and the incidence of late-life dementia in the ACT cohort.

Aim 2: Characterize the fine-scale spatial variability of annual-average UFP and BC exposure levels for the ACT cohort by utilizing mobile monitoring observations.

Aim 3: Characterize the long-term temporal trend of UFP and BC exposure levels for the ACT cohort by utilizing historical TRAP emission indicators.

Both the incidence of dementia and population exposures to TRAP are predicted to increase in the near future (Alzheimer's Association, 2019; Sperling, 2018; Wortmann, 2013). Given that relatively few studies that have investigated the association between dementia and long-term TRAP exposure, this work is a significant contribution to the field.

Aim 1: Long-term NO_2 exposure and late-life dementia incidence

Methods

Baseline & Outcome Measures

We extracted participant baseline characteristics from ACT study records, including: birthday, enrollment and end of follow-up dates, APOEϵ*4* gene allele status, gender, race, educational attainment, smoking status, physical activity habits, ACT cohort (original, expansion or replacement), vascular health indicators (e.g., atherosclerosis, hypercholesterolemia, diabetes, hypertension, stroke history), heart disease summary (e.g., myocardial infarction), body mass index (BMI) and CASI score (adjusted for accuracy using Item Response Theory). Appendix A Figure 57 shows the directed acyclic graph (DAG) used to guide our models.

We collected the neighborhood-level median household income from 2,000 US Census tracts based on the longest-lived address at or prior to baseline.

All-cause dementia (our primary outcome) and AD dementia (our secondary outcome) outcomes were assigned based on the most recent follow-up visit as of September 30, 2018 based on standard DSM – IV criteria (dementia) and the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) (probable or possible AD) (American Psychiatric Association, 1994; McKhann et al., 1984). Participants were assigned follow-up time based on their most recent biennial visit. Cases were those who were diagnosed with all-cause dementia at or shorty after their most recent visit. Non-cases were those without a diagnosis of all-cause dementia by the most recent visit because they had left the study, died or had not yet had their next biennial visit. Individuals with only baseline observations were excluded from this analysis (795/5,546 = 14%).

Residential Histories & Exposure Estimates

We assigned participants space- and time-varying long-term NO_2 exposure based on our spatiotemporal air pollution prediction model. This model is similar to our previous work, which

combines land use regression (LUR) and geostatistical smoothing. The model was built using observations collected between 1995 and 2019 from ten regulatory monitoring sites (two sites were used to characterize the long-term trend), low-cost sensors at 112 community and ACT participant homes (2017-04-15 to 2019-05-25), 110 locations from three, seasonal snapshot campaigns (2018-04-28 to 2019-03-02) and over 200 other local research sites (1999-2019). The model used partial least squares (PLS) scores that summarize hundreds of geographic covariate predictors (e.g., road proximity, truck route length, normalized difference vegetation index [NDVI], population density within an X-meter radius, land use type, etc.) to three PLS components, one smoothed time-trend estimated from a singular value decomposition (SVD) of the regulatory monitoring data with eight degrees of freedom per year, and geostatistical smoothing for the spatially varying long-term mean and time trend. Summarizing the average of all measurements at each location, the model had a cross-validated R^2 and root mean square error (RMSE) of 0.68 and 3.91 ppb at the ten regulatory monitoring locations and 0.64 and 3.06 ppb at community and ACT participant homes. Predictions from 1995 were projected back in time to estimate long-term exposures starting prior to 1995. We assessed the impact of this extrapolation assumption in sensitivity analyses by restricting our ten-year exposure analyses to 2005 or later. Of note, these are preliminary predictions that will have minor modifications in the near future.

Person-years within the spatiotemporal modeling region (Figure 1) were included in our analysis. We calculated time-varying, ten-year average NO_2 exposure estimates from two-week spatiotemporal model predictions for individuals based on their residential histories. Residential gaps were imputed in two ways in order to assign individuals exposure estimates at any given time as long as they lived within the spatiotemporal modeling region. First, since information about the earliest residential histories varied for different people, a participant's first available address was extended back if we otherwise had insufficient residential time to estimate their long-term exposure. This was typically done for participants who enrolled early on in the study (~1994) since address histories were available starting in 1989 for most people (6/10 desired years), though some had histories dating back to the 1970s or earlier. Second, residential gaps after a first known address were imputed by assuming that individuals moved halfway through the gap.

We characterized the address imputation quality as being high if the imputation: a) was based on the first address on record and only extended backwards by a duration equal to or less

than the amount of time lived at that location, up to 10 years (e.g., if a person lived at their earliest recorded residence for five years and we imputed their address another five years back); or b) a residential gap after a first known address was less than two years. The address imputation quality was characterized as being low if these criteria were not met (e.g., an address gap greater than two years was imputed). Exposure estimates with high address imputation quality were analyzed in a sensitivity analysis.

Descriptive Analyses

In descriptive analyses, we summarized baseline participant characteristics overall as well as characterized NO_2 (and $PM_{2.5}$) exposure predictions by calendar time and participant age.

Survival Analysis

We used Cox proportional hazards regression models to investigate the association between time-varying, ten-year average NO_2 exposure and age of late-onset, all-cause dementia. We stratified the baseline hazard functions in these models by APOEϵ4 (1+ allele) since carriers may have a higher hazard of dementia incidence as they age, and adjusted for potential confounding by gender, baseline SES indicators (education, median Census tract household income, race) and calendar year. Gender and SES have been shown to be associated with both air pollution exposure and dementia incidence (L. P. Clark et al., 2014; Gouveia & Fletcher, 2000; Howell et al., 2017; Matthews et al., 2019; Weuve et al., 2018). Calendar time was included to adjust for changes that have occurred in both air pollution and dementia incidence over time. For example, individuals born earlier may have been exposed to higher TRAP concentrations over their lifetimes and they may have had lower education levels (associated with a higher dementia incidence). Each variable's reference category was selected based on the category with the largest sample size to improve model precision. Our model was:

$$h(age) = h_0(age|APOE)e^{\beta_1 X_{NO_2}(age) + \beta_2 I_{gender} + \beta_3 \boldsymbol{I}_{edu} + \beta_4 \boldsymbol{I}_{income} + \beta_5 I_{race} + \beta_6 \boldsymbol{I}_{year}} \quad (1.1)$$

where:

- $h(age)$ is the hazard of the outcome with age as the time axis.
- $h_0(age|APOE)$ is the baseline hazard with age as the time axis and stratified by APOE status (yes/no).
- $X_{NO_2}(age)$ is the time-varying NO_2 exposure on the age axis as a continuous variable. Ten-year individual exposure at any given age is assigned based a range of calendar year predictions from our spatiotemporal model. Long-term, rolling exposure averages are calculated every year.
- I_{gender} is a binary indicator variable for gender: female (reference), male
- $\boldsymbol{I}_{edu}$ is a vector of indicator variables for education at baseline: less than high school, high school or GED (reference), Bachelor's Degree, Master's Degree, Doctoral Degree, Other
- $\boldsymbol{I}_{income}$ is a vector of indicator variables for Census tract annual income at baseline: [$0 - $35k], ($35k - 50k], ($50k - $75k] (reference), > $75k
- I_{race} is a binary indicator variable for race: white (reference), not white
- $\boldsymbol{I}_{year}$ is a vector of indicator variables for calendar year, binned into 2-year categories except three years for 2016-2018 (2004 as the reference)

We accounted for delayed entry, right censoring, used the Efron approximation analysis method to handle any event ties (122 ties), and used robust standard errors in case the model was misspecified (Kleinbaum & Klein, 2012).

We used inverse probability weighting (IPW) to address possible selection bias since approximately 12% of participants had missing APOEϵ4 values. IPW assumes that observations are missing at random and works by upweighting observations that are more likely to have a missing covariate (APOEϵ4 in this case) and downweighting observations less likely to have a missing covariate. Since other covariates had small degrees of missingness (each between 0.1% and 2.1% of the time), we assigned missing covariate values based on each covariate's mean (for continuous variables, e.g., CASI) or mode (for categorical variables, e.g., race) prior to calculating model weights. We calculated participant weights using Lasso logistic regression to build a model for APOE availability based on the strongest predictors in our data and taking the inverse. Weights were stabilized to reduce their variability by multiplying them by the

probability of APOE availability given gender status. The final dataset included IPW weights for each person and recoding for other missing covariate values.

Sensitivity Analyses

We performed several sensitivity analyses to better understand the robustness of our results (Table 3). All sensitivity models used the same adjustment covariates as our primary model ("Model 2") and a 10-year exposure window unless otherwise specified. Separate models were run for all-cause and AD dementia.

We ran reduced and extended models to further adjust for potential confounders and mediators. A reduced model (Model 1) stratified by APOEe4 status and included time-varying NO_2 exposure but did not adjust for any additional covariates found in our primary model (Model 2). An extended model (Model 3a) built on our primary model and further adjusted for potential confounding by smoking status (never, former or current) and regular exercise activity (at least 15 minutes 3 times per week). A second extended model (Model 3b) built on Model 3a by further adjusting for ACT cohort (original, expansion or replacement). An extended and mediation model (Model 4) built on model 3a and further adjusted for potential mediating factors including baseline: vascular health indicators (past occurrence of stroke, transient ischemic attack or carotid endarterectomy), heart disease indicators (myocardial infarction, angina, coronary artery bypass grafting or angioplasty). Extended models (Models 3-4) used baseline covariate values (e.g., vascular health) since adjusting for these as time-varying covariates could dampen the observed NO_2 effect given that some of these may have also been along the causal pathway. These models thus assume that these covariates are constant over time. A co-pollutant model (Model 5) based on our primary model additionally adjusted for potential confounding by time-varying $PM_{2.5}$ exposure. A model stratified by all of the confounders in our primary model other than calendar time (which resulted in small strata combination bins of size one) to ensure that potential confounders were sufficiently controlled for (M6). Appendix A Figure 63 shows cumulative hazard curves for each covariate in this model. See Table 3 for additional details.

We looked at the impact of adjusting for time differently through various models. First, we ran a model using onset age (halfway between the last two study visits for cases and the last visit for non-cases) rather than diagnosis age (last study visit for both cases and non-cases) as the

event time, since this was similar what past ACT studies had done (Kukull et al., 2002). This may introduce some bias, however, since follow-up time for non-cases is about one year longer than cases. We also ran other models that adjusted for time differently by: 1) adjusting for 2- and 5-year birth cohort bins instead of 2-year calendar time, and 2) using calendar year as the time axis while with adjustment for 2-year birth cohort.

We ran restricted analyses that investigated the impact of exposure assessment quality. One model was restricted to 2005 and later when spatiotemporal model predictions were not extrapolated back based on the 1995 surface. Two models only included person-time that we believed had the best residential address quality and consequently good exposure predictions. The "highest" address quality analysis was restricted to person-time without any address imputation and with an exact address at least 80% of the time. The "high" address quality analysis was less restricted by also including person-time with a high-quality imputation (e.g., residential gaps less than 2 years). Finally, a model assessed 10-year NO_2 exposure as a fixed baseline rather than a time-varying covariate, as has been commonly done in the literature. This analysis was thus at the person, rather than person-year level.

We ran a model not using IPW to better understand how much our results could have been impacted by APOEϵ4 missingness.

Finally, we used NO_2 exposure predictions from models built by the Center for Air, Climate, & Energy Solutions (CACES) to verify the sensitivity our results given our exposure model (S.-Y. Kim et al., 2020). Annual CACES models were fit using data from all U.S. regulatory monitoring sites with available data between 1979-2015. The start date of the 1979 model was extended back to 1978-02-01 to match the time period of the $PM_{2.5}$ spatiotemporal model, while the end date was extended to 2019-04-02. Annual average universal kriging models were built using PLS to summarize roughly 350 geographic predictors (from the MESA Air geodatabase). We used predictions made at two location types within Washington State: a) the census blocks of cohort locations, and b) exact cohort locations. A small degree of missingness resulted at locations in blocks that had only recently been populated (only blocks with subjects as of ~2017 had predictions) and from the fact that the formaldehyde covariate was missing for some locations. Missingness from these models is further discussed in the results and discussion.

Table 3. Sensitivity Analyses

Sensitivity Analysis	Description	Rationale
Reduced and extended models[a]	Run the following models: M1 (reduced): Age (time axis), NO_2 (time-varying) + stratified(APOEϵ4) M2 (primary): M1 + gender + education + median household income + race + calendar time M3a (extended confounding adjustment 1): M2 + smoking + physical activity M3b (extended confounding adjustment 2): M3a + ACT cohort M4 (extended and mediation 1): M3a + vascular health + heart disease + BMI M5 (co-pollutants): M2 + $PM_{2.5}$ (time-varying) M6 (more stratification):	M1: Stratifies by APOEϵ4 since APOEϵ4 carriers may have different hazards of dementia incidence as they age. Does not adjust for potential confounding. M2: includes important potential confounders M3, M5 and M6: further adjust for potential confounding M3b: The Expansion cohort was originally intended to have greater demographic variability than the Original cohort, though the resulting cohorts were in the end similar. I thus do not expect to see a significant impact by ACT cohort. If differences had been observed, however, these could have been associated with both exposure and outcome. M4: adjusts non-direct effects of NO_2 on dementia onset (Ahlskog et al., 2011; HealthLine, 2016; Kivimäki et al., 2018; Lee et al., 2014)

	M1 + stratified(gender) + stratified(education) + stratified(median household income) + stratified(race) + calendar time	M5 adjusts for correlated pollutants; this is important for valid inferences M6 stratifies by all of the confounders in our primary model except for calendar time to ensure that potential confounders were sufficiently controlled for while ensuring that stratification bins have more than one person-year.
Dementia onset age	Uses age at dementia onset rather than age at dementia diagnosis A dementia *diagnosis* occurs during the last biennial visit while an *onset* date is assigned as being halfway between the last two biennial visits	This is the analysis followed by other ACT studies (Kukull et al., 2002), though it may introduce some bias since the follow-up time for controls is defined at their last study visit and thus is slightly longer by about a year
Calendar year time axis	Model used calendar year as the time axis instead of age and adjusted for 2-year birth cohort.	Verify that model results are similar since both adjustments should ensure that we are comparing people who are of similar age during the same time period.
Birth cohort adjustment	Model adjusted for 2- or 5-year birth cohort bins instead of 2-year calendar time	Verify that model results are similar since both adjustments should ensure that we are comparing people who are the same age during the same general time period
No exposure model extrapolation (2005+)	A 10-year exposure analysis that begins in 2005 since the spatiotemporal model begins in 1995	Investigate whether results change when the model is extrapolated back prior to 1995 under a fixed exposure surface assumption

"Highest" address quality	Analysis was restricted to person-time without any address imputation (e.g., no address gaps or a gap with the same address before and after) and with an exact address (vs. a geocoded block-level address) at least 80% of the time	These individuals may have the most accurate exposure estimates
"High" address quality	Analysis was restricted to person-time with a high quality imputation (e.g., residential gaps less than 2 years) and an exact address at least 80% of the time	These individuals may have more accurate exposure estimates than those not included in this analysis
Fixed exposure at study entry	Uses fixed exposure at study entry rather than time-varying exposure. This is a person- rather than a person-year -level analysis. Since we are adjusting for exposure year (calendar time), models will compare those entering the cohort at similar times.	Similar to studies that assign a fixed exposure at baseline.
No IPW	Does not reweigh observations to adjust for missingness in the APOE$\epsilon 4$ covariate	Investigate how not accounting for a large degree of missing data may result in biased estimates, as is the case if data are not missing completely at random
CACES exposure model (S.-Y. Kim et al., 2020)	Use NO_2 exposure predictions form the CACES model at the residence level (similar to our primary analysis) and at the Census block level.	Check whether our results are similar when using a different exposure model, and at a cruder spatial resolution

[a]All model covariates are baseline values unless otherwise stated (i.e., only NO_2 and $PM_{2.5}$ are time-varying)

Secondary Analyses

In secondary analyses, we investigated: a) different exposure time windows ranging from 1-20 years, including some lagged exposure windows, b) effect modification by APOE status, gender or race (white vs non-white), and c) road proximity exposure rather than predicted NO_2 exposure.

We used different exposure windows since the optimal exposure window for dementia progression is unknown, though it is hypothesized to be quite long (e.g., 10-20 years). Shorter exposure time windows reflected the common practice in the field of using short exposure windows as surrogates of long-term exposure.

A model based on our primary model further adjusted for effect modification by APOEϵ4 status since carriers may have a higher hazard of dementia for any given exposure level as they age. The purpose of this analysis was to determine whether the exposure risks are different for APOEϵ4 carriers and non-carriers. Similarly, we were interested in seeing if the risk of dementia was different for different genders or races.

We investigated time-varying major road proximity because we believed that this could be a better surrogate for overall TRAP, which is a complex mixture of both gases and particles, than NO_2 alone. Treating this as a time-varying analysis allowed us to incorporate multiple addresses (movers) rather than assuming that all individuals resided at the same baseline address the entire exposure period of interest. We censored road distances for each participant address so that the maximum was 500 m since any locations beyond this were expected to be similarly impacted by road emissions. We calculated a new major road proximity estimate for each participant-year based on that participant's residential history during the past ten years by calculating the geometric mean distance to an A1, A2 and A3 roads, as defined by Census Feature Class Codes (CFCC) (US Census Bureau, 2018). The purpose of calculating average ten-year distance in this manner was to incorporate address histories while reducing the influence of very large distances on the average. We categorized average road proximity over a ten-year period as being near a major road if it was on average ≤ 300 m from a primary highway with limited access (A1 road) or ≤ 50 m from a primary roads without limited access (A2 or A3 road), midway to a major road if it was ≥ 300 m from an A1 road and within 51-150 m from an A2 or A3, and away from a major road otherwise (≥ 300 m from an A1 and ≥ 150 m from an A2 or A3;

Table 4). We conducted a separate analysis investigating the impact on dementia of spending 10 years living "near" versus "far" from a major road. Others have also modeled major roadway proximity using similar approaches (Chen, Kwong, Copes, Tu, et al., 2017; Gan et al., 2010; E. H. Wilker et al., 2015; Elissa H Wilker et al., 2014; Yuchi et al., 2020).

Table 4. Definition of time lived near (A-D), midway (E) or far (F) from a major road (A1-A3) within the past 10 yrs, as detailed in the methods.

A2 or A3	**A1**	
	≤300 m	**>300 m**
<= 50 m	A	D
51-150 m	B	E
>150 m	C	F

Finally, we conducted some competing risk analyses to address whether loss to follow-up was associated with a higher risk of death prior to a dementia diagnosis. The Cox proportional hazards model assumes independent censoring, such that losses to follow up are independent of the outcome. However, this assumption may be violated if, for example, higher air pollution levels are associated with a higher risk of death prior to dementia. Air pollution in this instance may appear to be protective of dementia. This may be particularly concerning in an aging population where death is not a rare event (Waller et al., 2020).

Various approaches have been taken to address competing risks (Waller et al., 2020). We took a common approach, which includes building cause-specific Cox models for the potential competing risk to determine whether the exposure interest (e.g., TRAP) is associated with the competing risk of interest (e.g., death prior to dementia). Given the interval censoring design of this study, we considered cases to be those who had died within two years of their last biennial visit dementia-free. All other individuals (dementia cases, and non-dementia cases that did not die within two years of their last biennial visit) were non-cases. We used age as the time axis (as in our primary analysis) and age at the last biennial study visit as the follow-up time. Of note, the assigned age at death in this analysis was slightly lower than the official age at death to allow for

identical definitions of events and non-events, and since exposure assessment was only available for a participant up until their last study visit. This approach thus added a small degree of measurement error, though it is similar to that seen in the diagnosis age for cases given the interval censoring in this study (i.e., in the primary analysis, cases were given a diagnosis at last study visit rather than halfway between the last two visits). Models included the adjustment covariates described for the primary analysis (M2), a co-pollutant analysis (M5) and an analysis with additional stratification beyond APOEϵ4 (M6).

Results

Descriptive Analyses

The ACT cohort consisted of 5,546 people who were followed between 1994 and 2018. After dropping 795 individuals (14%) who only having baseline observations and 6 (0.1%) individuals who were either outside of the spatiotemporal modeling region or had missing address histories altogether, 4,745 (86%) people remained. The final dataset used in the primary analysis included 4,165 individuals since the APOEϵ4 was missing for about 12.2% of individuals. The following results are for the primary analysis unless otherwise specified.

Table 5 describes overall baseline participant characteristics for all individuals in the "follow-up" cohort (those with observations past baseline and a valid address within the spatiotemporal modeling region at least some of the time) and those included in the primary analysis. Participant characteristics were similar for both the follow-up and primary analysis cohort. Overall, those in the primary analysis enrolled in the study at a median (interquartile range [IQR]) age of 73 (9) and were followed for an average (standard deviation [SD]) of 10 (6) years. About 27% and 22% of the participants developed all-cause (1,141) or AD (924) dementia. More than half of participants were female (58%), most where white (90%). Most participants had a high school or bachelor's degree (62%) and lived in an area where the median 2,000 Census Tract annual income ranged from $35,000-$75,000. About half of participants had never smoked (48%), while the other half had been former smokers (47%). Most reported exercising regularly (at least 15 minutes three times a week, 73%). Most were overweight, 40% had hypertension and 17% had heart disease. Over one quarter (26%) were APOEϵ4 carriers.

Half (51%) of participants were enrolled in the original ACT cohort, while the rest entered during the Expansion or Replacement phases. The stabilized weights from IPW calculated for our Cox models had a mean (SD) of 1.0 (0.09) and ranged from 0.88-1.68.

Table 5. ACT cohort characteristics by all-cause dementia diagnosis for individuals in the entire cohort and those included in the primary analysis (complete variable analysis with IPW).

	Follow-Up Cohort[a]	**Cohort in Primary Analysis[b]**
Participants (n, %)	4745 (100%)	4165 (88%)
Person-years (n, %)	45898	41331
Entry age, years (median, IQR)	73 (9)	73 (9)
Follow-up years (mean, SD)	10 (5)	10 (6)
10yr Avg Exposure at Entry, ppb (mean, SD)	16 (3)	16 (3)
Dementia Cases (n, %)	1266 (27%)	1137 (27%)
Alzheimer's Disease Cases (n, %)	1020 (22%)	920 (22%)
Birth Cohort (n, %)		
1909 or earlier	177 (4%)	162 (4%)
1910-1914	351 (7%)	325 (8%)
1915-1919	703 (15%)	626 (15%)
1920-1924	964 (20%)	889 (21%)
1925-1929	893 (19%)	789 (19%)
1930-1934	490 (10%)	425 (10%)
1935 or later	1164 (25%)	947 (23%)
Female (n, %)	2781 (59%)	2417 (58%)
White Race (n, %)	4245 (90%)	3762 (90%)
Education (n, %)		
Less than High School	404 (9%)	345 (8%)
High School or GED	1822 (38%)	1622 (39%)
Bachelor's Degree	1086 (23%)	965 (23%)
Master's Degree	717 (15%)	613 (15%)
Doctorate Degree	281 (6%)	244 (6%)
Other	432 (9%)	374 (9%)
Census Tract Income (n, %)		
< $35k	437 (9%)	383 (9%)
$35-50k	1474 (31%)	1289 (31%)
$50-75k	2311 (49%)	2039 (49%)
> $75k	520 (11%)	452 (11%)
Smoker (n, %)		
Never	2316 (49%)	2017 (48%)
Former	2197 (46%)	1947 (47%)
Current	229 (5%)	199 (5%)
Regular exercise (n, %)	3417 (72%)	3018 (72%)
BMI (n, %)		
Underweight	41 (1%)	34 (1%)
Normal	1499 (32%)	1304 (31%)
Overweight	1988 (42%)	1763 (42%)
Obese	1214 (26%)	1062 (26%)

Vascular Health (n, %)		
Hypertension	1946 (41%)	1680 (40%)
Diabetes	508 (11%)	434 (10%)
Cardiovascular Disease	420 (9%)	360 (9%)
Heart Disease	786 (17%)	690 (17%)
APOE carrier (n, %)	1103 (26%)	1102 (26%)
Baseline CASI (mean, SD)	0.34 (0.7)	0.35 (0.7)
ACT Cohort (n, %)		
Original	2324 (49%)	2132 (51%)
Expansion	741 (16%)	651 (16%)
Replacement	1677 (35%)	1380 (33%)
IPW Weight, (mean, SD)	1.01 (0.09)	1.00 (0.09)

[a]this group includes individuals with follow-up times post baseline and with a valid address within the spatiotemporal modeling region at least some of the time.

[b]this group includes with APOE availability

On average, ten-year exposure windows had NO_2 predictions from the spatiotemporal model about 99% of the time (99% and 98% for the CACES model at block- and residential-level, respectively), exact addresses (the highest quality) about 97% of the time, and imputed addresses about 9% of the time. About 95% of ten-year exposure windows had either no address imputation or imputation of the highest quality. About 38% of participants lived in one location during their study enrollment.

Annual NO_2 exposure predictions for the entire study period had median (IQR) concentration 14 (5) ppb and ranged from 3-42 ppb (Figure 3, Appendix A Table 17). Concentrations varied within each year (based on participant location) and generally decreased over time. The average SD and IQR for 10-year NO_2 exposure predictions for any given two-year period (within calendar bin exposure variability) was 2.9 ppb and 3.1 ppb, respectively.

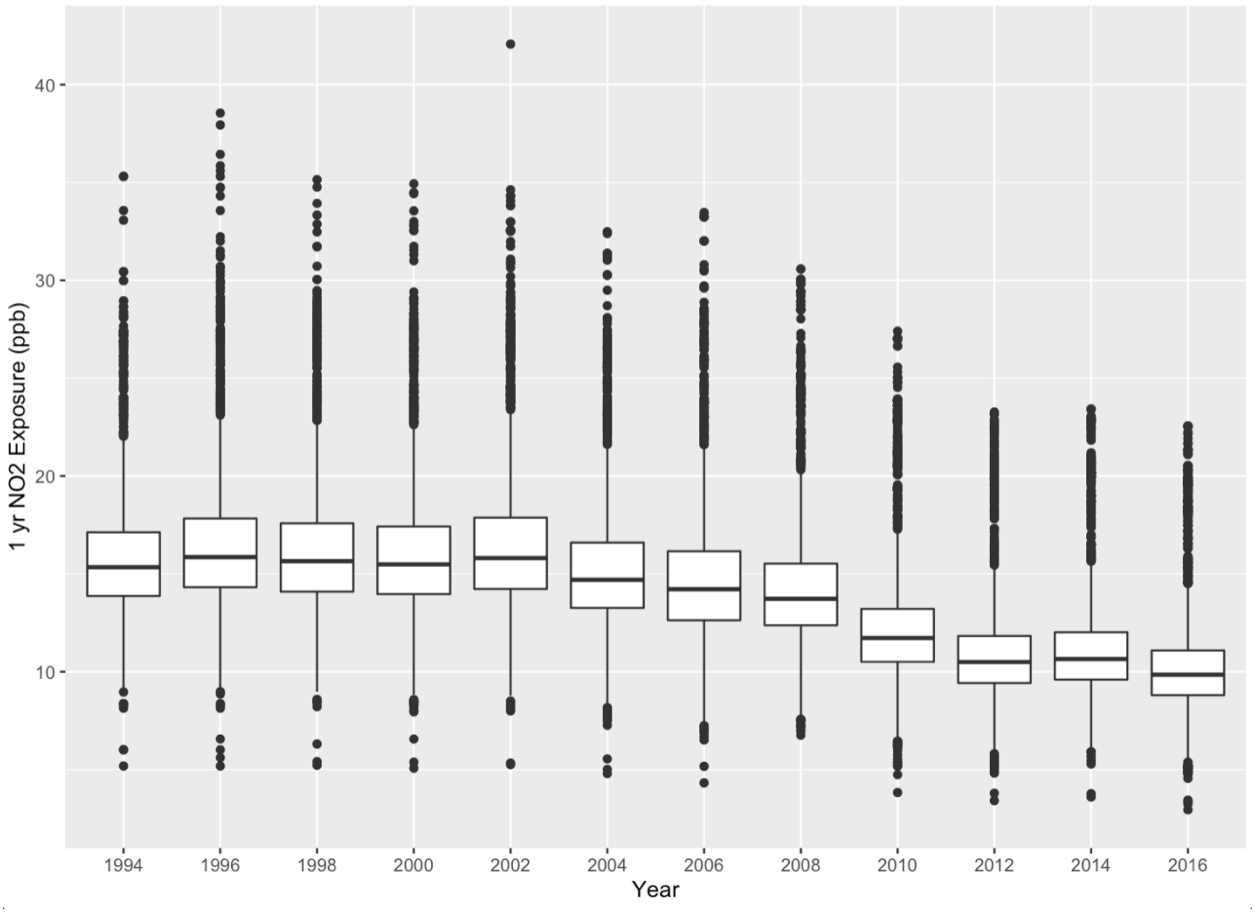


Figure 3. Annual NO_2 exposure predictions for person-years in the primary analysis

Starting in 2005, 1- and 10- year average exposure predictions from the spatiotemporal model were generally higher and more right skewed (had more extreme observations) than those from the CACES model (Appendix A Figure 58 - Figure 59). Prior to 2005, spatiotemporal model predictions were generally lower than the CACES model since we projected 1995 predictions back in time and earlier years with higher concentrations were missing from the model. The annual IRQs were similar for all the models.

Figure 4 shows ten-year NO_2 exposure by participant age and calendar year for person-years in the primary analysis (the spatiotemporal model). Each point represents a person-year, such that participants generally have multiple points. Within a calendar year, exposure typically remains the same for different ages. As participants age, they move to later calendar year bins where concentrations are typically lower starting around 2005 but remain similar for earlier

periods since the same 1995 exposure surface was used as a surrogate for earlier time periods. See Appendix A Figure 60 - Figure 61 for comparable $PM_{2.5}$ plots.

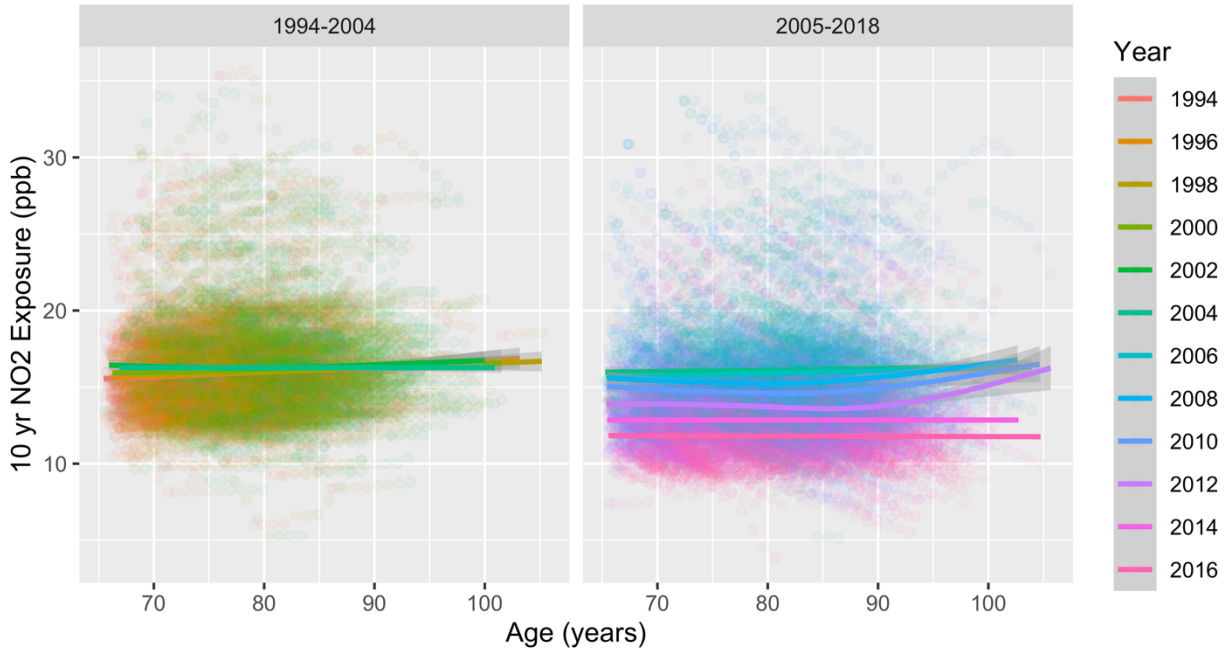


Figure 4. Ten-year average NO_2 exposure predictions by participant age and birth cohort

Survival Analyses

Our primary analyses indicated that for every additional 5 ppb increase in 10-year average NO_2 exposure, the hazard of all-cause and AD dementia is estimated to be 1% (HR: 1.01, 95% CI: 0.91, 1.11) and 2% (HR: 1.02, 95% CI: 0.91, 1.13) greater, respectively, after adjusting for APOEϵ4 status, gender, SES indicators (race, median neighborhood-level income and educational attainment) and calendar year (Figure 5, Appendix A Table 18). Both of these effects, however, were consistent with both positive and negative findings.

Our sensitivity and secondary analyses for all-cause and AD dementia were in line with these results with some variations (Figure 5 - Figure 6, Appendix A Table 18 - Table 19). In particular, the $PM_{2.5}$ co-pollutant models (M5) produced slightly lower estimates, while the analysis restricting to 2005 and later resulted in wider confidence intervals. Analyses using the CACES exposure model produced wider confidence intervals. NO_2 interaction with APOE,

gender or race in both the all-cause dementia and AD models was uncertain (APOE interaction p-value: 0.42 for all-cause dementia, 0.27 for AD; gender interaction p-value: 0.06 for all-cause, 0.07 for AD; race interaction p-value: 0.75 for all-cause dementia, 0.63 for AD). The HR confidence interval comparing participants living midway versus far from a major road was very wide, likely due to few participants residing in the midway category.

Figure 5. NO_2 (5 ppb) hazard ratios for all-cause dementia and AD from primary and sensitivity analyses

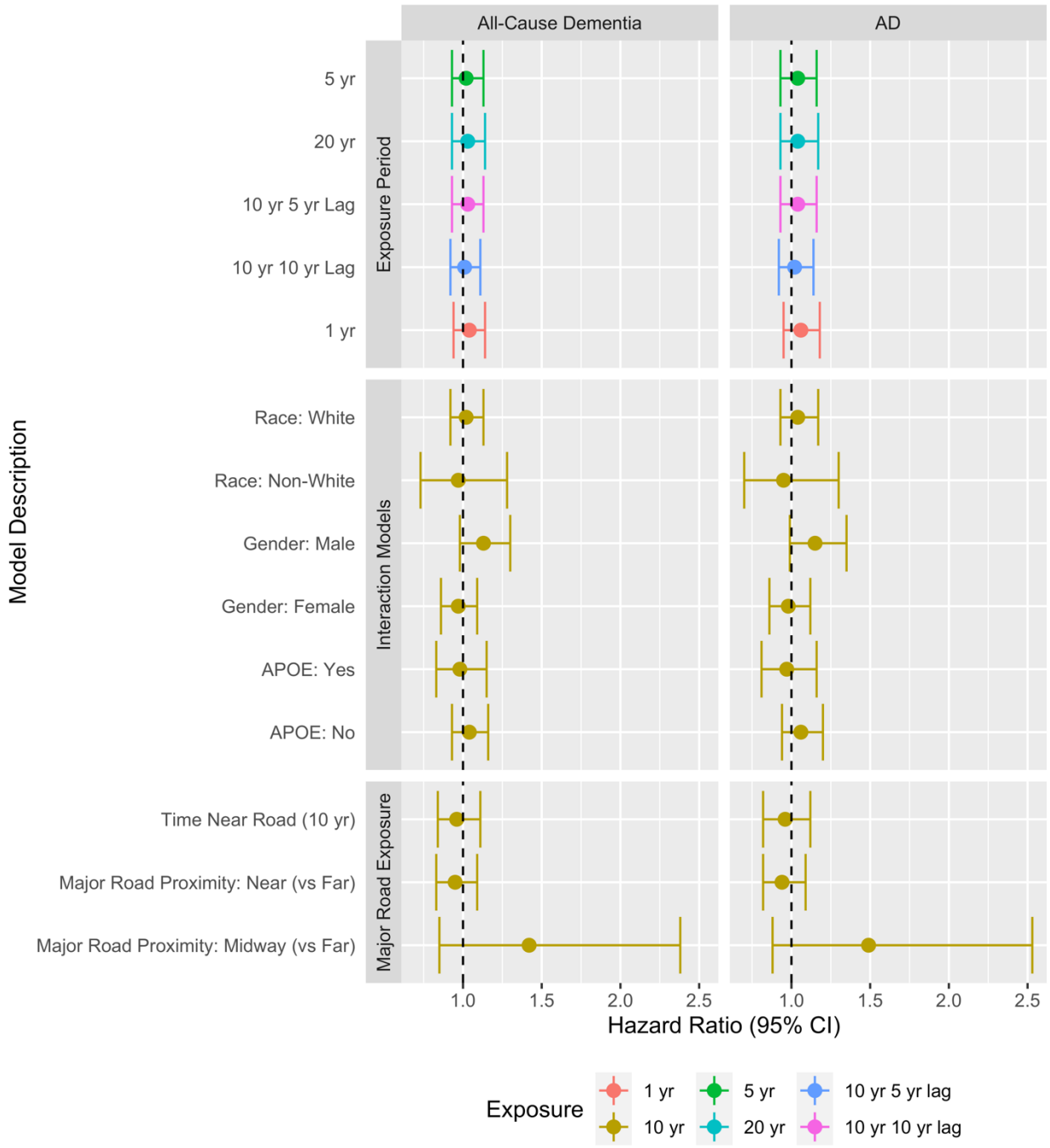


Figure 6. NO_2 (5 ppb) hazard ratios for all-cause dementia and AD from secondary analyses

Figure 7 and Appendix Table 20 show the results from our competing risks analyses looking at the association between 10-year NO_2 exposure and death before all-cause dementia or AD as a competing risk. As a reminder, cases were those who left the study without a dementia

diagnosis and died within two years after their last follow-up. NO_2 hazard ratios for death before all-cause and AD dementia were compatible with both positive and negative effects.

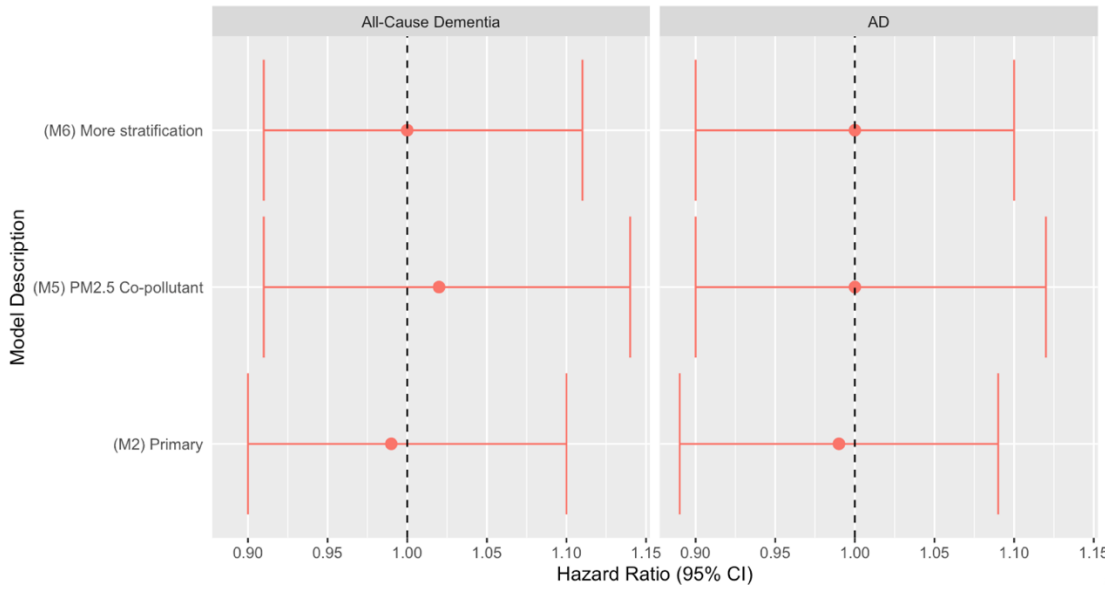


Figure 7. Competing risk analyses using cause-specific Cox proportional hazard models. Figure shows hazard ratios for death before dementia or AD for a 5 ppb increase in NO_2 exposure.

Discussion

Novelty

This study was one of the first to investigate the association between truly long-term TRAP exposure and late-life dementia incidence in the ACT cohort, a community-based, prospective cohort study. We used a spatiotemporal model that leveraged regulatory and supplemental monitoring data in the Puget Sound and excellent residential history records to assess ten-year, time-varying NO_2 exposure. This type of assessment is particularly relevant for dementia, which progresses slowly, and has been a notable limitation in the field (Power et al., 2016). Furthermore, we assessed outcome based on regular, in-person visits using consistent protocols over time. Most other studies have relied medical records, which have a limited ability to capture cases (National Institute of Aging, 2014; Power et al., 2016; Wu et al., 2016). Our extensive collection of participant covariates such as APOEϵ4 status, comorbidities, SES

indicators and lifestyle factors allowed us to control for potential confounding more thoroughly than what some other studies have been able to achieve. Finally, we conducted an extensive set of sensitivity and secondary analyses to test many of our assumptions related to confounding, proper temporal adjustment, exposure assessment quality and duration, interaction terms and more. Among these was an analysis of major road proximity exposure, which, to the best of our knowledge, is one of the first to investigate this as a time-varying exposure by incorporating participant residential histories. Furthermore, we conducted an analysis looking at death as a competing risk. Few other air pollution studies have reported conducting this analysis, despite death being a particularly relevant concern in studies of with older participants (Waller et al., 2020).

Cohort Characteristics

The demographic composition of the ACT cohort generally reflects the composition of the greater Seattle area, and is thus generalizable (Kukull et al., 2002). The cohort is comprised of both men and women, most of whom were former or never smokers, overweight and with some comorbidities. In accordance with the area, however, most are also white, well-educated, live in middle-income neighborhoods and report exercising regularly. About one quarter of participants developed late-life, dementia, with most of these cases being attributed to AD (81%: 920 AD/1,137 all-cause). Generally, females, those with lower education levels and APOEϵ4 carriers were more likely to develop dementia. These characteristics are in accordance with what others have found (Alzheimer's Association, 2019; Chêne et al., 2015). Most cases were in the original ACT cohort because this earlier study enrollment is associated with longer participant follow-up times. As expected, the risk of dementia incidence increased with age (Appendix A Figure 62).

Long-Term NO_2 Exposure

Almost all of the participant address histories were of exact quality (the highest quality; rather than at the zip code level, for example), and few had residential gaps, giving us confidence that the prediction locations were of high quality throughout our study period. This was made possible to us through various address data sources including Kaiser Permanente billing records,

which ensured that many participants had address histories dating back to 1989 with some having histories dating back to the 1970s or earlier. This is a particular strength of this study since our records indicated that most people moved at least once during the exposure periods of interest, with some moving as many as 17 times. Most other studies, on the other hand, have used fixed exposure at a single residence, typically at baseline.

As expected, 10-year NO_2 predictions varied across location within any given year and generally decreased over time starting in the early 2000s. One thing to note is that we extrapolated the 1995 exposure surface back in time in order to include participants enrolled in the ACT cohort prior to 2005. Concentrations trends were thus more constant prior to 2005. This approach assumes that the shape (rank order) of the exposure surface was the same prior to 1995. Our additional analyses for Aim 3 indicated that this assumption may be reasonable since, at least for of $PM_{2.5}$, exposure surfaces change but remain similar over time (Appendix C). Furthermore, this extrapolation should not have had a significant impact on our results since our models adjusted for time. Calendar time is an important potential confounder given the ACT cohort's long duration of over 25 years (during which many new participants have been enrolled), the generally decreasing trends in air pollution over time, and the potential for dementia incidence trends to have changed over time. This has been a less important consideration in other studies because typically a single (closed) cohort is enrolled at baseline and exposure is fixed to a single time period. Our calendar time-adjustment should thus make our open cohort more comparable to past studies.

Given our maximum predicted annual average concentration of around 39 ppb for earlier years (1994-2000) and 23 ppb for more recent years (2012-2018), our predicted exposure estimates can be generally considered low and are below the current annual average National Ambient Air Quality Standards (NAAQS) for NO_2 of 53 ppb (US EPA, 2018e). These NO_2 concentrations were similar to those reported by Beelen et al. (2010) for 36 European areas between 2008-2011 of 7-42 ppb (Rob Beelen et al., 2013). Similarly, Larkin et al. (2017) predicted global and North American annual average NO_2 concentrations from ground monitors between 2011-2013 of around 11 ppb (Larkin et al., 2017). These reported concentrations are slightly lower than what we observed, though this is likely a result of the low concentrations observed outside of large metropolitan areas.

Moreover, the average NO_2 concentrations used in this study were similar to other epidemiologic studies of TRAP and dementia incidence in Sweden (~9 ppb) (Oudin et al., 2016), London (~19 ppb) (Carey et al., 2018) and Canada (~15 and 31 ppb) (Chen, Kwong, Copes, Hystad, et al., 2017; Chen, Kwong, Copes, Tu, et al., 2017; Yuchi et al., 2020).

Primary, Sensitivity and Secondary Analyses

We did not find any evidence of an effect of long-term NO_2 exposure on late-life dementia incidence. Our sensitivity and secondary analyses were compatible with these results. The slightly lower hazard ratio point estimate from our extended and mediation (M4) model suggest that our primary model may have had a minor degree of confounding and/or that some of the estimated NO_2 effect may have come from cardiovascular health indicators. Similarly, $PM_{2.5}$ may have been responsible for part of the estimated NO_2 effect, as suggested by the slightly lower NO_2 hazard ratio in the co-pollutant model (M5).The fact that all of our alternative time adjustment ("time variant") models (i.e., adjusting for birth cohort instead of calendar year; using calendar year as the time axis; using onset age rather than diagnosis age) produced similar HRs to those of our primary analysis reassures us that adjusting for time differently, as has been done in the literature, would not have changed our results. Conservative analyses only looking at high exposure quality data showed similar results. This makes sense since most of the data were of high address quality. Additional sensitivity analyses suggest that extrapolating our exposure model prior to 1995 was not unreasonable. The wider confidence intervals for the 2005+ only models were likely a result of a smaller sample size. It makes sense that the model without IPW resulted in similar findings since the participants included in these analyses (with APOE availability) were similar to those in the follow-up cohort (which included APOE missingness), as indicated by the tight model weight estimates near 1.0. A possible explanation for the wider HR confidence intervals resulting from using the CACES exposure models could be due to the reduced variability of these models when compared to the spatiotemporal model. The CACES model relies solely on regulatory monitors, while the spatiotemporal model incorporates additional data from supplementary monitoring campaigns including participant homes and high-concentration areas such as near roads. This feature makes common exposure models such as the CACES less capable of detecting small effect sizes from quickly decaying traffic pollutants.

Given that residence- and Census block- level predictions were very similar, this suggests that the exact prediction location may not matter. Increased exposure misclassification may be more evident and problematic, however, in models with higher spatial resolutions or when exposure is assessed at cruder spatial resolutions such as Census Tracts rather than blocks.

While some work indicates that dementias progress slowly and may have long exposure periods of interest, we did not learn anything new from using shorter or longer exposure periods. We can't know, however, whether different exposure windows would have resulted in meaningfully different results had we seen a clearer effect in our primary analyses.

It was unclear from our data whether APOEϵ4 status, gender or race modified the association between NO_2 exposure and dementia incidence. It's possible that these variables do not act as effect modifiers in this population, though our results were unclear, and they could still be confounders or precision variables given that they are associated with dementia. The fact that males had a slightly higher HR than females despite generally being less likely to develop dementia suggest that changes in TRAP exposure could impact dementia incidence in males more so than females, although this effect size remains uncertain. Furthermore, these additional subgroup analyses should be interpreted with caution since most of these were post hoc.

Though the results of our major road proximity analyses were uncertain, they were in line with our primary analyses of NO_2 exposure. This suggests that it may be appropriate for studies aiming to investigate overall TRAP to use these more easily obtainable road proximity covariates rather than building a spatiotemporal model, which requires a substantial amount of work. Furthermore, road proximity serves as a more complete surrogate of overall TRAP exposure, something our primary and even $PM_{2.5}$ co-pollutant models fall short of. As a whole, these results indicate that it may not be overall TRAP or NO_2 specifically that puts individuals at an increased risk of developing dementia, though other pollutants cannot be ruled out. This is supported by the attenuated HR estimates in our $PM_{2.5}$ co-pollutant models and by a parallel survival analysis investigating $PM_{2.5}$ exposure and dementia incidence in the ACT cohort. This study found a positive HR for every additional 1 $\mu g/m^3$ increase in $PM_{2.5}$ of 1.16 (95% CI: 1.03, 1.31) (Shaffer et al., 2021). These findings are also in line with others in the field, including a recent meta-analysis of cohort studies investigating $PM_{2.5}$ exposure and dementia, which reported a pooled relative risk per 5 $\mu g/m^3$ of 1.08 (95% CI: 1.03, 1.13) (X. Yu et al., 2020). Furthermore, it's possible that late-life cognition may be more strongly impacted by other, less-

commonly studied traffic pollutants such as UFPs. There is some evidence that UFPs may impact cognitive function more so than other pollutants, though much of this work has so far been limited to short-term exposure and animal studies (US EPA, 2019e). Analyses of UFP exposure may be particularly relevant for the ACT cohort given its setting near the Seattle-Tacoma International (Sea-Tac) Airport, which has been linked to increased UFP levels, even when compared to other related pollutants like BC (Austin et al., 2019).

We did not find evidence in our data that death was a competing risk in the NO_2-dementia relationship (a violation of the Cox model assumption of independent censoring). NO_2 hazard ratios for death before all-cause and AD dementia were consistent with both positive and negative effects. As a reminder, the rationale for conducting this analysis was that if higher TRAP exposures were causing people to leave the study due to death before dementia, those that remained in the study would be more likely to have lower exposures, and they would also have the opportunity to develop dementia. In turn, higher TRAP exposures could appear to be protective of dementia.

Other Epidemiologic Studies

A recent meta-analysis of longitudinal cohort studies investigating NO_2, NOx and dementia incidence by Yu et al. (2020) reported similar findings consistent with both positive and negative effects. Authors reported a pooled relative ratio (RR) of 1.02 (95% CI: 0.99, 1.04) for a 5 ppb increase in NO_2 or NOx, as calculated from a random effect model that used the inverse variance and the number of participants to weight the study-specific log RRs (X. Yu et al., 2020). These results were initially surprising given that many individual studies have reported hazard ratios that are mostly consistent with positive effects ranging from 0.97-1.20 for a 5 ppb increase in NO_2 or NOx, even in lower-concentration areas comparable to the Puget Sound (Peters et al., 2019; X. Yu et al., 2020). This perceived positive effect reported by individual studies may be due to publication bias, which occurs when results that are positive and consistent with what is expected are more likely to be published (Song et al., 2013). The following paragraphs detail these studies, highlighting some of the differences in exposure and outcome assessment when compared to this study. Primarily, these involve using medical records to assess outcome rather than regular follow-up visits, and assessing exposure at baseline, based

on short-term measurements and/or crude spatial resolution models. Figure 8 and Appendix Table 21 further summarize and compare the findings of these studies. We converted mass concentration ($\mu g/m^3$) HRs to volume mixing ratio (ppbv) HRs using a 1.88 $\mu g/m^3$ per 1 ppb conversion factor, which assumes standard atmospheric conditions (1 ATM pressure, 25˚C) and a molecular weight for NO_2 (Martine et al., 2011; X. Yu et al., 2020).

Oudin et al. (2016) investigated the association between annual average NOx exposure at baseline and dementia incidence in the Betula study, a population-based cohort of 1,806 Swedish residents 55+ years of age (Oudin et al., 2016). The cohort started in 1988 and is assessed every 5 years. Additional participants were added during the second follow-up period (1993-1995). NO_x exposure was assigned based on residential address between 1993-1995 from a LUR model fit using 2009-2010 data (median ~ 9 ppb). The authors reported a HR (95% CI) for dementia incidence of 1.05 (0.98, 1.12) for every 10 $\mu g/m^3$ increase in NO_x.

Carey et al. (2018) investigated NO_2 exposure and dementia incidence from a retrospective cohort study using the primary care data of 130,978 adults ages 50-79 from 75 practices in the Greater London area (Carey et al., 2018). Subjects were followed between 2005-2013 for an average of 6.9 years, and NO_2 exposure was assigned at the postcode level based on a 2004 dispersion model. Authors reported an increased risk of dementia of 1.16 (1.05, 1.28) for every 7.5 $\mu g/m^3$ increase in NO_2 exposure.

Chen et al. (2017) looked at NO_2 exposure and dementia incidence from a population-based cohort study using the health care data of 2,165,268 Ontario adults ages 55-85 (Chen, Kwong, Copes, Tu, et al., 2017). Dementia-free subjects from this Ontario Population Health and Environment Cohort (ONPHEC) were enrolled April 1, 2001 and followed through 2012. As a sensitivity analysis, authors assessed NO_2 exposure at baseline from LUR-satellite models for the years 1998-2001. Authors reported an increased risk of dementia of 1.07 (1.06, 1.08) for every 11.3 ppb increase in NO_2. In a similar study using ONPHEC, Chen et al. (2017) once again looked at NO_2 exposure and dementia incidence, this time using the health care data of about 2,066,639 Ontario adults ages 55+ who were followed through 2013 (Chen, Kwong, Copes, Hystad, et al., 2017). Annual exposure predictions were assigned at the postal code level between 1994-2013. The final analysis used a time-varying NO_2 exposure defined as a 5-year moving average with a 2-year lag. Authors reported an increased risk of dementia of about 1.10 (1.08, 1.12) for every 14.2 ppb increase in NO_2.

Cerza et al. (2019) looked at air pollution exposure and primary or secondary hospitalizations for dementia in 21,548 older residents (ages 65-100) living in Rome, Italy. Those who completed a 2001 population census were enrolled and followed through either a dementia hospitalization, the year 2013, 100 years of age, death or migration, whichever came first. Exposure was assessed based on residential address at baseline (~2001) using NO_2 measurements from 40 sites during three two-week periods in 2010 (nine years later, near the end of the study period). Authors reported a decreased risk of dementia hospitalizations for every additional 10 $\mu g/m^3$ of NO_2, though the observed risk reduction was only about 3% (HR: 0.97, 95% CI: 0.96-0.99), after adjusting for sex, age, education, birthplace, marital status and census block SES.

Ilango et al. (2019) looked at the role of cardiovascular disease (CVD) in the relationship between air pollution and dementia among 34,391 older adult residents of Ontario, Canada. Participants were those who had completed Canadian Community Health Surveys between 1996-2003. Exposure was assessed at the 1 km^2 resolution from exposure models of annual average concentrations for any given three-year period prior to the completion of a health survey (e.g., ~1996). This was followed by a CVD (e.g., 1996-2001) and a dementia (e.g., 2001-2013) follow-up period. Authors reported an association between NO_2 (5 ppb) and dementia incidence that was compatible with mostly positive but some negative effects (HR: 1.10, 95% CI: 0.99, 1.19), with about 9% of the effect of NO_2 on dementia being mediated through CVD.

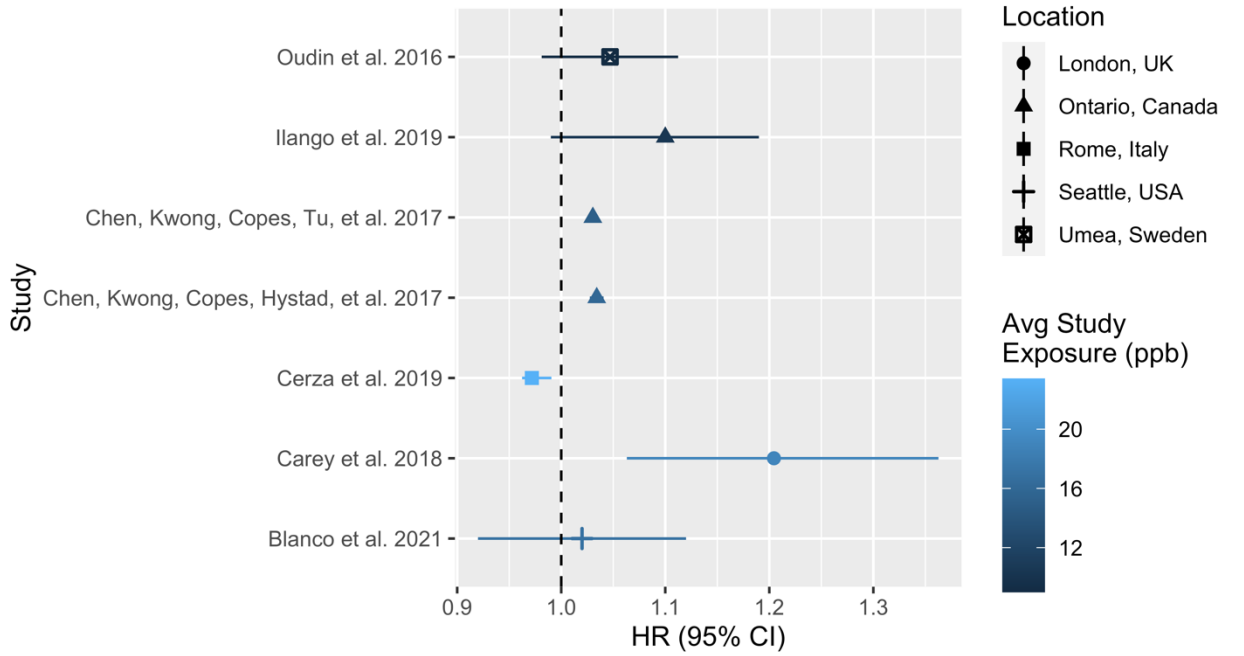


Figure 8. Estimated dementia HRs for a 5 ppb increase in NO_2 or NO_x from different studies, including this one. Some hazard ratios were converted to volume concentrations (ppb) from mass concentrations ($\mu g/m^3$) assuming a 1.88 $\mu g/m^3$ per 1 ppb conversion factor. Oudin et al.'s 2016 HR is for NOx. The HRs for all other studies are for NO_2.

Strengths of this Study

Looking at what point estimates fall within the range of our 95% confidence intervals, our results are consistent with other individual studies in the field other than Carey et al. (2018) and with a recent meta-analysis of NO_2, NOx and dementia incidence (X. Yu et al., 2020). Slight differences between our study and those of others may be attributable to different settings, populations and study design. Specifically, strengths of this study include the use of a large, long-standing (1994+), community-based, prospective cohort; the assessment of dementia incidence through regular, in-person follow-ups and the use of a standardized protocol applied consistently throughout the study period; and the assignment of long-term, time-varying exposure estimates largely based on exact residential addresses histories.

To date, many studies in this field have been limited to the use of more readily available but less reliable methods of exposure assessment. These include assessing exposure at one time point (e.g., baseline), over a short period of time (e.g., one year) and/or at a crude spatial resolution (e.g., residential zip code) (Chang et al., 2014; C. R. Jung et al., 2015; Loop et al., 2013; Oudin et al., 2016). A strength of this study is our use of historical NO_2 observations to assign long-term, time-varying exposure predictions at exact participant residences. We expect that this approach improves exposure assessment. Nonetheless, our secondary analyses using shorter and longer exposure periods did not change our results. This suggests that using short-term exposure as a surrogate for long-term exposure may be valid if the exposure surface remains similar over time.

Furthermore, this study used a prospective cohort and regular in-person visits to assess dementia incidence. A particular strength of the ACT cohort study is its use of similar follow-up methods and in-depth diagnostic procedures over time along with the collection of important potential confounder and precision variables (e.g., education, genetic susceptibilities). Many other studies have relied on medical records to assess dementia incidence and collect adjustment variables (Carey et al., 2018; Chang et al., 2014; Chen, Kwong, Copes, Tu, et al., 2017; Chen, Kwong, Copes, Villeneuve, et al., 2017; X. Yu et al., 2020). The use of medical records to assess dementia may introduce a substantial amount of outcome misclassification given that dementia tends to be underdiagnosed (National Institute of Aging, 2014), and may result in numerous false positive and false negatives (Power et al., 2016). Moreover, while medical record studies can allow for larger sample sizes, these may be less practical and generalizable in areas without universal health care access such as the United States. When time trends are also taken into account, it's also likely that dementia incidence and diagnoses have changed over time as a result of broader diagnostic criteria, increases in patient medical attention and changes in education levels (Wu et al., 2016). Since air pollution trends have also dropped over time, time can thus be an important confounder in medical record studies.

Limitations

There are some notable features of this study that should be recognized. First, we did not adjust for factors that impact personal exposure to TRAP other than residential outdoor levels

such as building characteristics (e.g., floor location, air exchange and infiltration rate), indoor sources of NO_2 (e.g., wood smoke, use of gas stoves), past occupational exposures or time activity patterns including time away from home (R. Allen et al., 2003; Jarvis et al., 2010; K. H. Jung et al., 2011; Klepeis et al., 2001; Park & Kwan, 2017; Vardoulakis et al., 2020; Zipprich et al., 2002). This approach is commonly taken in air pollution studies because: a) outdoor air pollution (unlike indoor air pollution) is regulated under the Clean Air Act (US EPA, 2019f), b) outdoor air pollution often behaves in predictable ways based on the surrounding environment (e.g., proximity to major roads), and c) characterizing outdoor air pollution infiltration and dispersion, indoor sources and individual behavior is not practical in epidemiologic cohort studies due to its complexity and resource requirements. Some work has shown, however, that indoor and outdoor NO_2 concentrations can be quite similar in residential buildings, offices and schools, though more variation exists in homes with major NO_2 sources such as gas stoves (Blondeau et al., 2005; Hu & Zhao, 2020; Salonen et al., 2019; Shrestha et al., 2019).

As noted above, since the spatiotemporal NO_2 exposure model began in 1995, we assumed that the spatial structure of the exposure surface had been the same during prior years and extrapolated this exposure surface back in time. Our additional analyses for Aim 3 of $PM_{2.5}$ (Appendix C) indicated that while the exposure surfaces likely do change over time, there is still a positive relationship between the more recent and earlier years (2018 vs 1995-2017; i.e., low- vs high- concentration sites are still generally distinguished). Some important features of the spatiotemporal model predictions used in this analysis include its large reliance on regulatory monitoring sites, particularly for earlier years when few sites and little to no study data were available. As such, most of the model's spatial variability comes from more recent years (1999+), while the temporal trend is established from two regulatory monitors in Seattle. These features of this model increase the potential for exposure misclassification for earlier years. Our model predictions may thus be impacted by both classical- and Berkson-like measurement error, meaning that predictions have the potential to be noisy and/or biased, which may in turn impact the health effect parameter estimates (Szpiro et al., 2011; Szpiro & Paciorek, 2013). This is typical, however, of many epidemiologic investigations where exposure is assessed based on prediction models.

An advantage to using the spatiotemporal exposure model over other models such as the CACES national model is that the spatiotemporal model leveraged more local information from

supplementary monitoring, and particularly measurements from participant homes, which is less common in the literature. Because they are different models, each with different features, however, this resulted in some missingness due to the spatial coverage differences between the models. While 10,985 locations (88% of the 11,610 cohort locations on record) had predictions from the spatiotemporal model and both CACES locations (block and exact), 172 (1%) only had spatiotemporal model predictions (e.g., locations were in unpopulated blocks), and 364 (3%) only had CACES predictions (locations were outside the spatiotemporal modeling region but within Washington State). Six hundred and twenty-five locations (5%) had no predictions from either model (e.g., they were outside of Washington State), while the rest had predictions from the spatiotemporal model and the CACES model at either the block- or residential- level. Nonetheless, long-term exposure predictions were generated for many of person-years in our analyses, provided that they were within a modeling region at least some of the time, though these were calculated from a fewer number of years. Coverage was still generally high, however, given that, on average, ten-year exposure windows had available NO_2 predictions 99% of the time from the spatiotemporal model and CACES model at the block-level, and 98% of the time from the CACES model at the residential-level.

Moreover, we additionally imputed some addresses when there were administrative gaps in residential address histories in order to estimate long-term exposure estimates. These were for a small fraction of all time-varying exposure estimates (~9%), however, and most imputations were of high quality, meaning that we had high confidence that the imputed addresses were accurate. Sensitivity analyses only including exposure estimates without address imputation did not change our results.

Our sensitivity analyses included baseline covariate values for all adjustment variables other than time-varying $PM_{2.5}$. This may have resulted in residual confounding when dealing with truly time-varying covariates (e.g., smoking status, stroke history), though it reduced the possibility of adjusting for covariates along the causal pathway. Importantly, this was not a concern for our primary analyses. The utilization of fixed covariates as surrogates for truly time-varying covariates has also been used in other TRAP exposure studies investigating dementia (Carey et al., 2018; Chang et al., 2014; Chen, Kwong, Copes, Hystad, et al., 2017; Chen, Kwong, Copes, Tu, et al., 2017; Oudin et al., 2016; Yuchi et al., 2020). Regarding median neighborhood-level household income, since this covariate was gathered from 2,000 Census data, this analysis

assumed that the neighborhood rank order did not change over time (e.g., the same neighborhoods have had the highest median neighborhood income levels over time).

Only a small fraction of the ACT cohort locations were "midway" to a major road (> 300 m from an A1 and within 51-150 m from anA2-3; 1%), limiting the use of the ACT cohort for investigating the impact of living midway to a major road when compared to living away from a major road (> 300 m from an A1 and > 150 m form an A2-3). This is highlighted by the large HR confidence interval for this category. Still, the "near-road" category (≤ 300 m of an A1 or ≤ 50 m from an A2-3), was associated with 27% of the total person-years in our data, and this was arguably the most scientifically interesting category. Furthermore, a minor discrepancy between this and our primary analyses using NO_2 exposure predictions from our spatiotemporal model is that road proximities were always available for a participant's entire exposure window of interest as long as residential address was available (or imputed). Our primary analyses using NO_2 concentration, on the other hand, were only available for addresses within the spatiotemporal modeling region. We thus we had more complete (fewer "missing") road exposure estimates. This is a minor discrepancy, however, given that most 10-year exposure windows had address histories inside the modeling region (99%).

A feature of this cohort is that individuals are older when they enroll in this study. The median entry age was 73 (IQR: 9) and ranged from 65-101. Since dementia becomes more prevalent after 65 years of age (WebMD, 2018), our study is best positioned to characterize the effects of long-term TRAP exposure on late-life dementia incidence. Moreover, though individuals are randomly invited to participate, this cohort may have some volunteer bias and possibly represent "healthy survivors" who were not affected by air pollution at earlier ages. If true, this could have dampened the NO_2-dementia relationship we observed. In addition, since ACT is an older population, losses to follow-up may not be random, but a result of individuals with very poor cognition no longer being able to complete follow-ups (Power et al., 2016). This is a minor concern, however, since ACT has achieved an impressive Completeness of Follow-up Index (the number of completed participant visits, divided by the number of possible visits) of 94.5% by giving participants the option of completing clinic, home or, as a last alternative, phone follow-ups (T. Clark et al., 2002).

Moreover, some cases may be missed (false negatives) since participants are screened for dementia using the CASI assessment tool and only receive a full clinical dementia diagnostic

evaluation if their scores are low enough a provider submits a referral. This study is also subject to interval censoring since follow-ups occur every two years. This follow-up frequency, however, is still higher than other cohort studies. Furthermore, since dementia progresses slowly, this should not have impacted our conclusions in a meaningful way.

Additional features of this analysis include the fact that there may be residual confounding due to variable categorization (e.g., income) and the use of self-reported variables (e.g., smoking status); there were a substantial number of participants with missing APOEϵ4 values (though we used IPW and these results were similar to sensitivity analyses without IPW); and our primary analyses did not differentiate between the effect of NO_2 and other correlated TRAPs (though sensitivity analyses of co-pollutant $PM_{2.5}$ models and road proximity resulted in similar conclusions). Finally, this study's sample size was smaller than most studies in the field, particularly medical record-based studies, which often include tens of thousands or millions of people (X. Yu et al., 2020). Combined with generally low NO_2 exposure levels in the region, an exposure model that began in 1995, and a limited within-calendar bin exposure variability (after adjusting for calendar time, there was a mean SD of 2.9 ppb), this could have contributed to our wider HR confidence intervals when compare to other studies. Still, the prospective cohort design of this study will have identified cases more accurately than medical-record studies. Additionally, many other studies in low concentration areas have used exposure models built from more recent, lower concentration years.

Conclusion

We did not find strong evidence of an effect of long-term NO_2 exposure on late-life dementia incidence. These results are generalizable since the demographic composition of the ACT cohort generally reflects the population of the greater Seattle area (Kukull et al., 2002). Furthermore, the design of this study allowed us to characterize individual, long-term exposure and assess dementia incidence through regular, in-person visits, a limitation of most studies in the field. The alignment of our sensitivity and secondary analyses with our primary analyses strengthens our findings. Future studies may want to investigate other pollutants and pollutant mixtures given that TRAP is a complex mixture gases and particles, and since other related pollutants such as $PM_{2.5}$ have been shown to be important determinants of dementia incidence.

Aim 2: Characterizing the spatial variability of ultrafine particulate matter and black carbon concentrations for the greater Seattle area from a mobile monitoring campaign

Overview

The aim of this study was to characterize 2019 annual average UFP and BC exposure levels for the ACT cohort using observations from an extensive mobile monitoring campaign specifically designed for epidemiologic application. We presented an approach for estimating annual average TRAP levels at stop locations from high-resolution, temporally unbalanced and non-normally distributed data collected from short-term, repeated measurements of TRAP. We build universal kriging models with partial least squares and use these models to characterize UFP and BC levels for the cohort and the greater Seattle area. Finally, we distinguish between the UFP and BC exposure surfaces by identifying land features that predict differences in these surfaces.

Methods

Instrumentation & Quality Assurance

We used total UFP count concentrations collected during our 2019 mobile monitoring campaign from unscreened PTRAK instruments (Table 2). These instruments collected particles ranging from 0.02-1 μm with a high temporal resolution (1 sec) and had more complete observations than the other UFP instruments on the platform, the DiSCminis (Testo). Unscreened PTRAKs were more inclusive of smaller particles, which may significantly contribute to number concentrations (Kukull et al., 2002), than screened PTRAKs (50-1000 nm sample range).

We used BC mass concentration estimates from aethalometers (microAeth MA 200, Aethlabs) used in the mobile monitoring campaign. Measurements were taken at 880 nm

wavelengths (which can be interpreted as BC) every 10 seconds (AethLabs, 2018; X. Wang et al., 2016).

Instruments were calibrated by the manufacturer before the start of the mobile monitoring campaign. Field technicians checked BC and UFP instruments for a zero or near zero response on five and six separate dates throughout the study, respectively (UFP & BC: 5/30/2019, 10/2/2019, 1022/2019, 11/21/2019, 1/13/2020; UFP: 3/13/2020).

Duplicate instruments were collocated in the field every few weeks and in the lab several times over the course of the year. This process of collocating particle instruments is common practice since there is no standard for field calibration of particle instruments.

Data Quality Control

All sampling days ("runs") with UFP readings below 300 pt/cm^3 were flagged and manually inspected. UFP readings were dropped if there was evidence of an instrument malfunction (N = 1,019 observations, 0.09% of the total data). This was often related to an alcohol wick issue and was verified by inspecting the field notes. P-TRAK concentrations at the instrument limit of 500k pt/cm^3 and aethalometer readings ≤ 0 ng/m^3 or > 99th percentile were visually inspected using time series plots to ensure that instrument readings did not deviate from the general trend observed for other particle instruments.

We compared median stop concentration estimates from collocated PTRAK (UFP) and aethalometer (BC) instruments. We used linear least squares regression to fit a best fit line to estimate the association between duplicate instruments. We further calculated the root mean square error (RMSE) and mean squared error- (MSE) based coefficient of determination (R^2) to quantify bias and precision between collocated instruments.

We linked instrument readings to specific stops based on GPS location. Readings were adjusted for a 10-second lag associated with the estimated time required for a volume of sampled air to travel from the manifold inlet on the roof of the car, into the manifold and into the sampling instrument.

Stop Concentration Estimates

We calculated the median UFP and BC concentration for every two-minute stop (i.e., using about 120 one-second PTRAK readings and 12 ten-second aethalometer readings) in the mobile monitoring campaign. We used the median to capture the central tendency of the data because the data were right skewed, and the median is known to be more robust to outliers than the mean. Other mobile monitoring campaigns have taken a similar approach (Apte et al., 2017). We evaluated using stop means in sensitivity analyses.

Annual Averages

We calculated annual average UFP and BC concentrations at each mobile monitoring stop using a 5% trimmed, weighted mean. First, the lowest and highest 5% of each site's stop concentrations (10% total) were first dropped. We assumed that these extreme observations were erroneous or uncharacteristic of the general underlying distribution (e.g., a short-term construction project or road maintenance), particularly since few samples were collected at each site (~28 stops per site were scheduled), and these were from brief, two-minute stops. Trimmed means have been used in past air pollution studies (Correia et al., 2013; Dominici et al., 2006; Samet et al., 2002; Van den Bossche et al., 2015) because they are robust to outliers and may capture diurnal pollutant variations more so than medians (Cleveland & Guarino, 1976).

Next, we used a regression-based approach to estimate season-, time-of-week- (weekend vs. weekday) and time-of-day (5 categories: 3-8 AM, 9-11 AM, 12-3 PM, 4-8 PM, 9 PM - 2 AM) averages. This allowed us to borrow information across sites, though it assumed that all sites had the same temporal patterns (e.g., higher concentrations at specific times of day). The rationale for using a regression-based approach rather than taking unweighted averages using only the observations collected at each site (site-specific averages) was that the mobile monitoring campaign was logistically unable to sample all sites at all times of the day within each season and time of week. We would have thus greatly reduced our site sample size if we had restricted our analysis to only included sites that had samples within each season, time-of-week and time-of-day combination. Finally, site-, season-, time-of-week- and time-of-day-specific regression concentration estimates were weighted appropriately to estimate annual

average concentrations for each site. We investigated whether this approach for estimating the site annual average was appropriate by conducting an analysis of variance (ANOVA) looking at the impact of site, season, day of the week and hour on the day on concentration variability.

Second, we compared the influence of using different approaches for estimating annual averages at sites using the mobile monitoring dataset itself since long-term UFP data were limited outside of this study. Specifically, we estimated annual averages using regression and site-specific approaches that weighted site-specific averages differently, for example by season alone or by season and time of week (Table 6). We ran ANOVAs looking at the effect of site and estimation method on the concentration variability.

Table 6. Weights used for estimating site annual average

Temporal Adjustment	**Weight Assigned**	**Rationale**	**Regression and Site-Only Approaches Used?**
Unweighted	NA	Take a simple average because the existing data may sufficiently capture the temporal variability at each site	Both Estimates are the same using either approach
Season	0.25 to each season	Heating seasons are associated with lower UFP levels than non-heating seasons	Both
Season + TOW[1]	1/4 * 5/7 for each season on weekdays 1/4 * 2/7 for each season on weekends	Weekdays generally have higher traffic counts	Both
Season + TOW[1] +	1/4 * 5/7 * 9/24 for each season on weekdays for	TRAP concentrations vary throughout the day as a result	Both

TOD[2] (2 categories)	between the hours of 9 AM and 5 PM 1/4 * 2/7 * 16/24 for each season on weekends for between the hours of 6 PM and 8 AM	of diurnal environmental conditions and traffic patterns	
Season + TOW[1] + TOD[2] (5 categories)	1/4 * 5/7 * 6/24 for each season on weekdays for the hours of 3-8 AM Etc.	More specific times of day categories	Both Only some sites are included in the site-only approach
Season + day of the week + hour of the day	1/4 * 1/7 * 1/24 for each season, day and hour combination. The first and last sampling hours were upweighted to account for no sampling in the middle of the night.	More specific times of day categories	Regression only No sites were sampled within each season, day of the week and hour of the day

[1]TOW: time of week (weekday vs weekend)
[2]TOD: time of day

We conducted several sensitivity analyses to understand the impact of these analytical decisions for dealing with extreme values and unbalanced data on our final site annual average estimates (Table 8). This included using 10% trimmed annual averages, five percent Windsorized means (setting observations below the 5th quantile and above the 95th quantile set to that respective quantile), using site-only, unweighted annual averages and adjusting annual averages for day and hour more finely (7 instead of 2 days; 21 instead of 5 times of the day) in sensitivity analyses.

Study Area

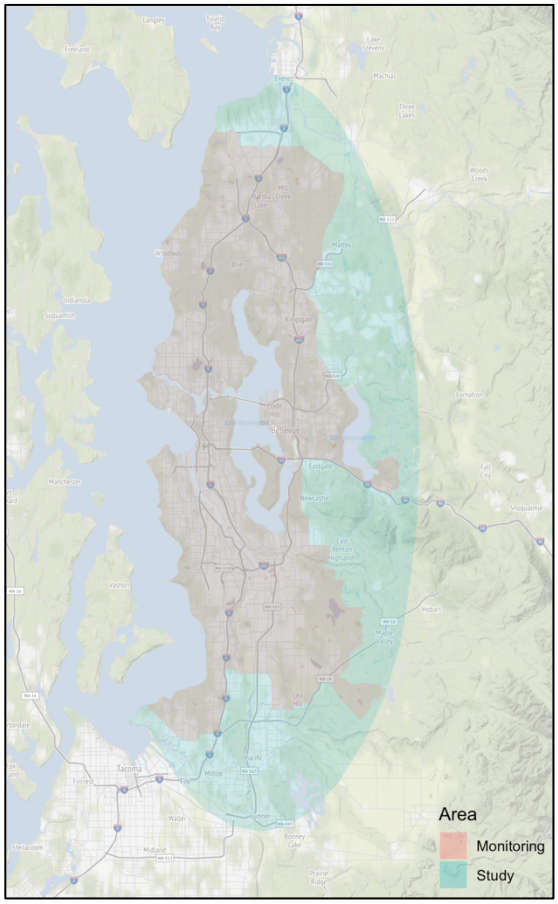


Figure 9. Map of the mobile monitoring and study areas.

We designated a study area that was slightly larger than the monitoring area (Figure 9). The monitoring area was where the mobile monitoring occurred and was roughly 1,100 land km^2. This area captured about 90% of the present and past ACT locations on record (Table 7). The study area was geographically similar to the monitoring area but was slightly larger at roughly 1,900 land km^2. It captured slightly more, about 92%, of all ACT locations on record. Locations beyond this study area were not included in this analysis because they were geographically different than the monitoring area and would have required too much extrapolation of the fitted model. Figure 71 - Figure 72 in Appendix B illustrates this by showing the distribution of geocovariates (summarized using PLS) for sites inside and outside of the monitoring area. Overall, the PLS score distribution is similar for mobile monitoring locations and ACT cohort locations within the study area, but different for ACT cohort locations outside of the study area. Since PLS scores are linear combinations of geographic covariates, this indicates that sites outside of the study area are overall geographically different than mobile monitoring stops.

Table 7. ACT cohort locations included in the mobile monitoring, and study areas relative to all U.S. ACT cohort locations on record.

Area	No. Locations	Proportion of Total
United States	26,225	1.00
Spatiotemporal Modeling	25,687	0.98
Study Area	24,150	0.92
Mobile Monitoring	23,606	0.90

Geocovariates

Geocovariates for this analysis were extracted from the MESA Air geodatabase (Table 1) for mobile monitoring stops and ACT cohort locations (MESA Air, 2019). There were 835 initial geocovariates, including population density estimates from the most recent, 2010 U.S. Census. Geocovariates were reduced and or adjusted before using partial least squares (PLS) regression following the approach described by Keller et al. (2015) (Appendix B Table 30) (Keller et al., 2015). First, variables were excluded if they lacked variability (less than 20% of the data were different from the most common value) since these were not likely to improve, and could even worsen, the model fit. Next, all proportion land use variables were excluded where the maximum proportion observed in the data was less than 10%. Low values for these variables indicated that these land use types made up a small fraction of the land relative to other land use variables and would not likely have a meaningful impact on observed pollutant concentrations. Variables were dropped if the distribution in mobile monitoring stops was very different than the distribution in the ACT cohort in an effort to reduce model extrapolation later on. This included variables where the standard deviation (SD) in the ACT cohort was more than five times the SD of the mobile monitoring data. Variables were dropped if too many outliers were observed in the data (>2% of the total data). Proximity variables (e.g., distance to closest major roadway) were set to a minimum distance of 10 m to protect against any possible ACT location geocoding issue since we did not believe that participant locations were closer than 10 m (e.g., the middle of a

roadway). These were log-transformed since air pollutants exponentially decay with increasing distance from a source.

Universal Kriging

We built universal kriging (UK) models using PLS to predict annual average UFP and BC concentrations. PLS is a technique that can take many features and summarize these into a smaller number of new features (James et al., 2013). New features are linear combinations of all of the original features, with variables that account for most of the variation in the response (i.e., the ones most strongly associated with the response) having heavier weights ("loadings"). PLS reduces the risk of overfitting the models while allowing for the use all of the available geocovariates. UK jointly runs regression (e.g., LUR) and geostatistical smoothing (kriging). The LUR piece takes geocovariate predictors (e.g., derived PLS features) while kriging attempts to model part of the structure that resides in the error term not otherwise captured by regression. Kriging operates under the premise that observations closer in space are more correlated than observations further away (unlike traditional linear regression which assumes that observations are all independent). As such, predictions for any particular location are influenced by both that location's geocovariates as well as correlated nearby observations. UK has been shown to perform better than LUR or ordinary kriging alone (Mercer et al., 2011; Young et al., 2016).

In this analysis, we used PLS regression to produce dimension-reduced geocovariate summary predictors of annual average UFP and BC (using the "plsr" function in the *pls* package, version 2.7.2) (Mevik, 2019). Pollutant concentrations and proximity geocovariates were modeled on the log-scale since these were log-normal. Next, we built UK models with PLS feature predictors and an exponential variogram to account for correlation among locations as a function of distance (using the "krige.conv" function in the *geoR* package, version 1.8.1), such that:

$$\log(\text{TRAP}) = \alpha + \sum_{m=1}^{M} \theta_{m,s} Z_{m,s} + \varepsilon \quad (2.1)$$

where:

- log(TRAP) is the log-transformed annual average TRAP pollutant concentration, either UFP (pt/cm^3) or BC (ng/m^3)
- $Z_{m,s}$ are dimension-reduced summaries of spatially-varying (s) geocovariates from the PLS procedure
- M is the number of PLS components selected through cross-validation (see results)
- α and $\theta_{m,s}$ are estimated coefficients modeled using UK, and
- ε is the residual term with mean zero and a geostatistical structure modeled as an exponential function with range ϕ_i, partial sill σ_i^2 and nugget γ_i^2 ($\epsilon_s \sim N(0, \sum(\phi_i, \sigma_i^2, \gamma_i^2))$)

For validation purposes, we split the data so that 10% was in a held-out validation set, and the remaining 90% was used for ten-fold cross-validation. We used the cross-validation set to select the number of PLS features (1-6 considered) and the kriging range parameter (5-50% of the maximum range considered) for the UK model using 10-fold cross-validation. A model was fit using the modeling parameters selected from the ten-fold cross-validation to these data (90%) and evaluated against the held-out validation set (10%). A final model was fit to all the data.

We plotted the PLS component loadings of the final model (i.e., from all of the mobile monitoring locations) for each pollutant to characterize the covariates and buffer sizes most strongly associated with pollutant variability.

We compared the distribution of PLS scores (linear combinations of geocovariates) between the monitoring locations and ACT locations within the study area. Different distributions between groups would be suggestive of extrapolation outside of the monitoring area.

We compared the UK model predictions (LUR and kriging combined) to the prediction from the LUR piece alone to evaluate the effect of kriging on the final UK predictions. If LUR alone was responsible for most of variability in the UK predictions (rather than the kriging), this would suggest that we could predict at locations outside of the monitoring area where we had geocovariate estimates but no mobile monitoring observations with which to krige the data as long as we did not extrapolate the geocovariate space.

We used the UK models to predict annual average UFP and BC concentrations at ACT locations within the study area. For display purposes, we used the UK models to make predictions on a grid to map exposure surfaces for the study area.

Furthermore, we characterized the geocovariate predictors of differences between the UFP and BC exposure surfaces. We did so by fitting least squares linear models to pollutant estimates (from mobile monitoring locations) and predictions (at cohort and grid locations), using UFP as the dependent variable and BC as the independent predictor of interest. We mapped the residuals to illustrate locations where UFP levels were lower or higher than would be expected based on BC levels. We also ran lasso regression models, selecting the tuning parameter lambda through cross-validation in order to identify geocovariate predictors of the residuals. The same geocovariates were used in this lasso regression as were used in our primary models (Appendix B Table 22). These were all standardized.

Sensitivity Analyses

We built similar BC and UFP prediction models in sensitivity analyses to assess the robustness of our models when dealing with extreme observations and unbalanced data differently (Table 8). Alternative models were built: 1) using site means (rather than medians); 2) calculating 10% (rather than 5%) trimmed annual means; 3) calculating Windsorized (rather than 5% trimmed) annual means; 4) calculating unweighted, site-specific annual average (rather than using regression to adjust estimates by season, time of week and time of day); 5) using site annual averages that were adjusted for sampling day and hour more finely (7 days of the week, 21 hours of the day); 6) modeling pollutants and proximity variables on the native scale (rather than on the log scale). Table 8 further details these analyses and the rationale.

Table 8. Sensitivity Analyses

Focus of Analysis	Analysis	Description	Rationale
Central tendency of raw, high-	Stop means	calculate two-minute stop averages rather than medians (n = ~28 means/site)	Means are more influenced by extreme values than medians, and extreme,

resolution data at two-minute stops			transient values may be characteristic of specific locations
Annual averaging of stop observations	10% trimmed mean	Drop each site's stop observations below the 10th quantile and above the 90th quantile (~ 6 observations/site = 28*(0.10*2))	Reduces the inclusion of outliers, though it decreases sample size
	Windsorized mean	Set each site's stop observations below the 5th percentile and above the 95th percentile equal to that respective quantile	Extreme observations may not be completely erroneous but in fact part of the true site variation, but we wish to reduce the potentially high influence of these extreme observations on the site mean
	Unweighted mean	Calculate means for each stop based on the observed data at that location alone	Analyses indicated that most of the pollutant variability occurs within sites, and that temporal factors influence levels to a much smaller degree
	TOW7 TOD21	Calculate means for each stop using a regression approach that more thoroughly adjusted for sampling day and hour of the day	Adjust more thoroughly for potential differences in concentrations at finer temporal time scales

UK model	Native scale modeling	Model pollutants and proximity variables on the native scale rather than on the log-scale	Model results may be more interpretable
	Predict within the monitoring area	Only include predictions for ACT locations within the study area	ACT locations outside of the monitoring area may have a slightly different geocovariate distribution than the monitoring sites

Results

Instrumentation & Raw Data

Two P-TRAKs (PMPT_93 and 94) and two aethalometers (BC_63 and 66) collected UFP and BC readings during the mobile monitoring campaign (Figure 10). P-TRAKs collected samples between 3/8/2019 and 3/17/2020. Aethalometers collected samples between 2/22/2019 and 3/17/2020.

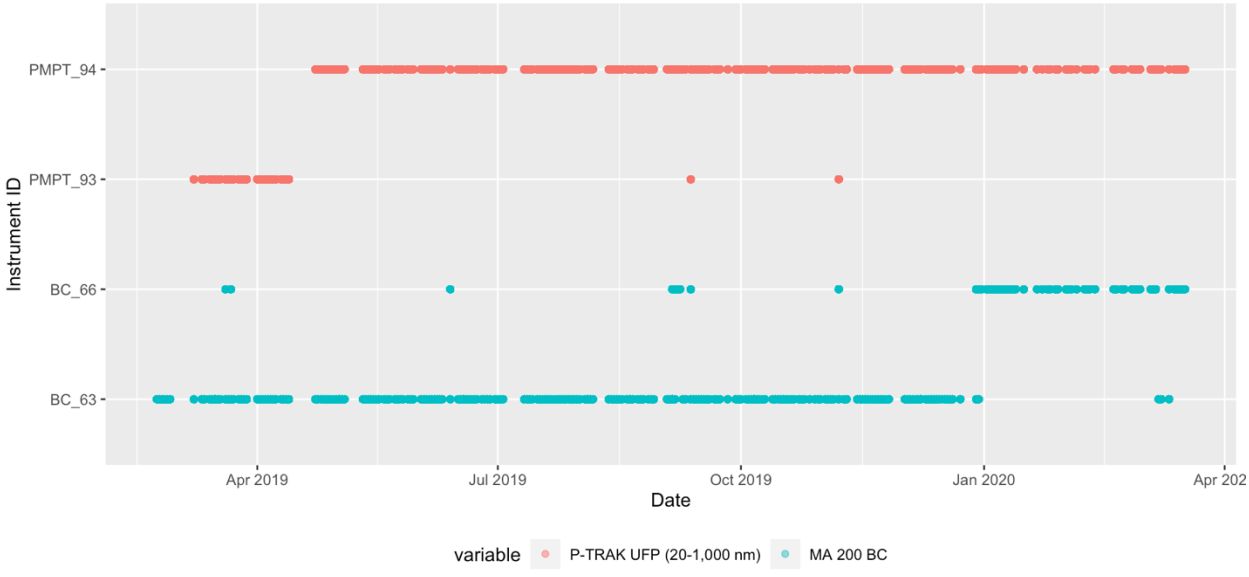


Figure 10. Instruments used over time in the mobile monitoring campaign

There were 1,177,944 raw 1-second UFP readings collected from P-TRAKs. These had a median (IQR) of 5,900 (5,570) pt/cm^3 and ranged from 0 to 500,000 pt/cm^3, respectively. There were 125,866 raw 10-second BC readings collected from aethalometers. These had a median (IQR) of 457 (697) ng/m^3 and ranged from -11,797 to 37,357 ng/m^3 (Appendix B Table 23).

Data Quality Control

Data quality checks with respect to zero, low and high concentration instrument checks as well as instrument collocations suggested good instrument performance (see Appendix B: Data Quality Control for details). Overall, P-TRAKs and aethalometers performed well, with aethalometer readings being noisier and more variable than P-TRAK readings. During zero checks, the reported instrument readings were generally centered around 0 (Appendix B Table 24). We observed unusually low PTRAK readings at stops on four occasions (1,019 one-second readings, 0.09% of total readings; Appendix B Figure 64). These readings were dropped since there were clear indications of instrument malfunctions. We observed very high UFP or BC readings on a couple of occasions (Appendix B Figure 65 - Figure 66). We kept these readings since they correlated well with other particle instruments which similarly reported elevated concentration levels. Collocated field and lab P-TRAKS were in good agreement with an overall RMSE and MSE-based R^2 of 575 pt/cm^3 and 0.98, respectively (Appendix B Figure 67). Collocated aethalometers were also in good agreement (RMSE: 152 ng/m^3; R^2: 0.91; Appendix B Figure 68).

Stop Concentrations

After trimming stop concentrations below the 5th and above the 95th percentile (about the lowest and highest measurement from each stop), calibrated UFP median stop concentrations (instruments were calibrated to the mean of collocated duplicate instruments) had a median (IQR) of 5,817 (4,493) pt/cm^3 and ranged from 1,050-51,935 pt/cm^3. BC stop concentrations had a median (IQR) of 411 (358) ng/m^3 and ranged from -31-3,606 ng/m^3. See Appendix B Figure 69 and Table 25 for additional details of the stop distributions before and after trimming.

Figure 11 shows the total stop UFP and BC samples that were collected at all sites, after trimming. All seasons and days of the week were visited a similar number of times. Sites were visited between the hours of 4 AM and midnight (though not all sites were visited during each hour), with most stop samples being collected during the morning and afternoon hours of 6-10 AM and 2-5 PM. On average (SD), each site had 24 (SD = 1) UFP stop samples and ranged from 21-27 stops. UFP samples were collected at all sites an average of 5-7 times per season, 17 times on weekdays and 7 times on weekends. Sample size results were similar for BC since instruments were generally collocated on the mobile monitoring platform.

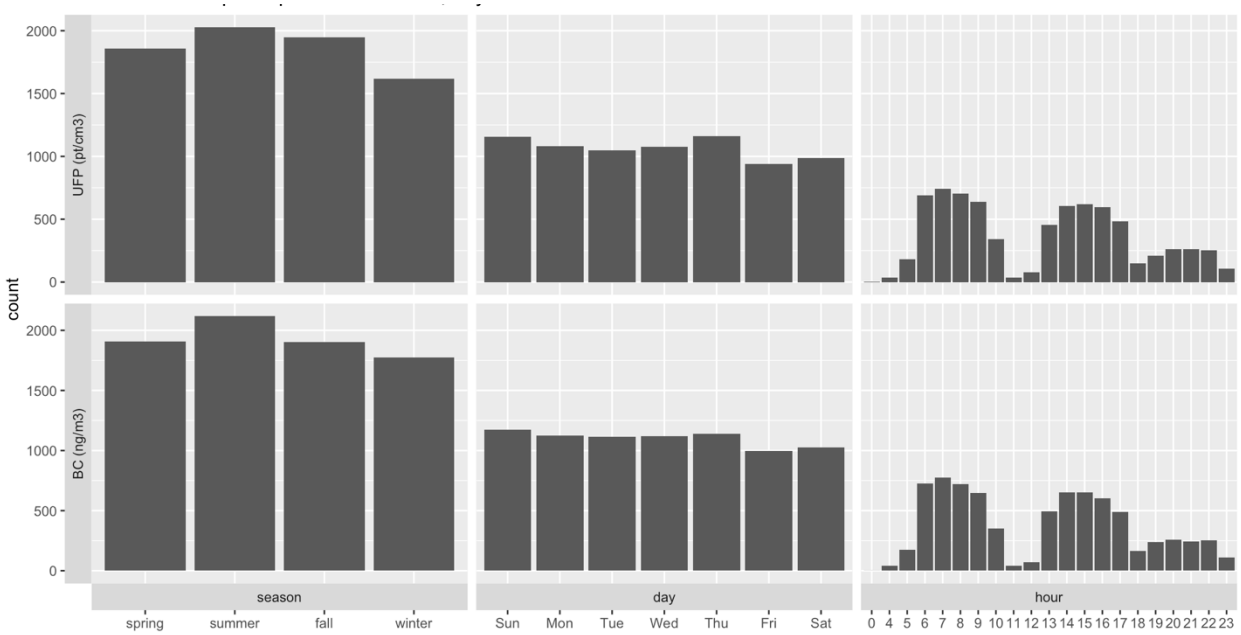


Figure 11. Total number of stop UFP and BC stop samples collected at all sites (~24 samples/site x 309 sites)

UFP stop concentrations were generally higher during the fall and winter and on weekdays. Concentrations peaked during late morning hours and evening hours (Figure 12). Results were similar for BC but with the morning concentration peak occurring a few hours earlier.

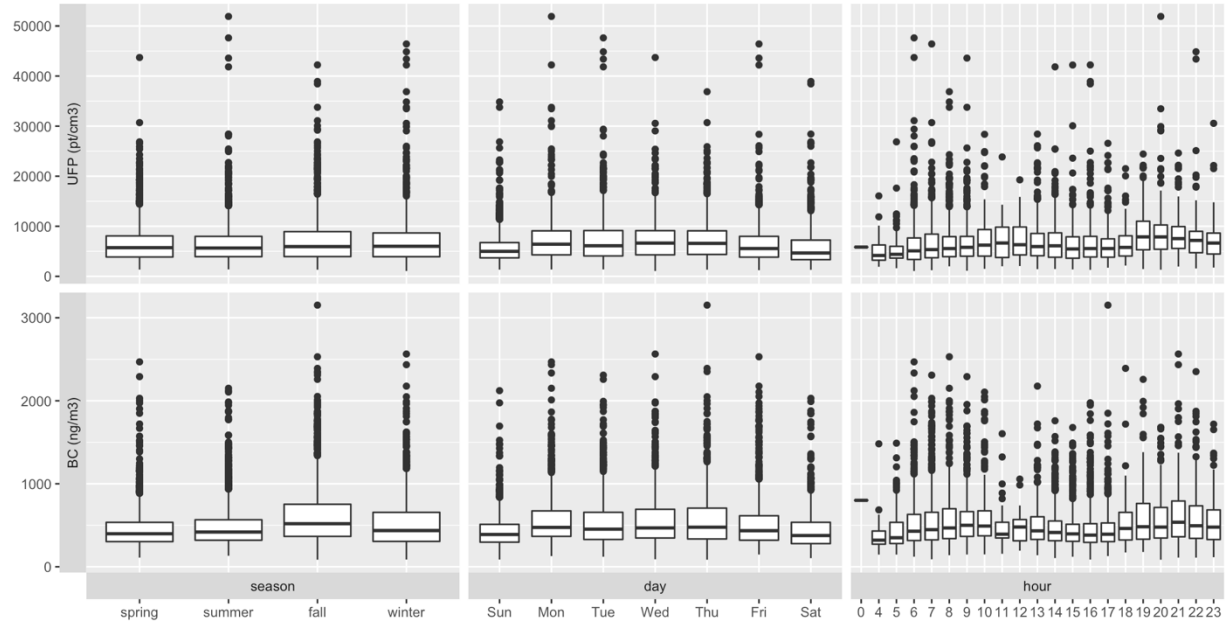


Figure 12. UFP and BC stop concentrations over time

Annual Averages

Season, time of week and time of day adjusted site annual average UFP concentrations at mobile monitoring sites had a median (IQR) of 6,511 (2,284) pt/cm^3 and ranged from 4,054-16,004 pt/cm^3. Sites near and south of downtown along I-5 generally had the highest annual average concentrations, while more distal sites generally had lower concentrations (Figure 13). Adjusted annual average BC concentrations at mobile monitoring sites had a median (IQR) of 506 (160) ng/m^3 and ranged from 306-1,384 ng/m^3. Elevated BC concentrations were also seen near and south of downtown along I-5.

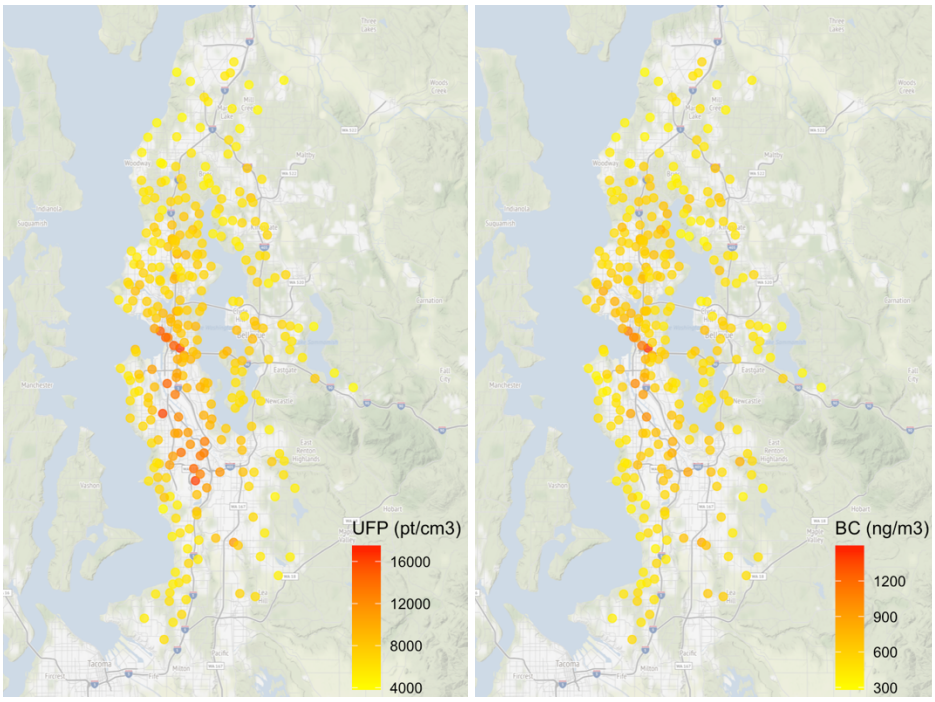


Figure 13. Map of annual average UFP (left) and BC (right) estimates at mobile monitoring sites

ANOVAs looking at the impact of site, season, day of the week and hour of the day on concentration variability indicated that most of the variability was seen within sites at any particular time (about 73% for UFPs and 77% for BC), followed by between sites (about 22% for UFPs and 12% for BC), and a minor fraction was seen across different seasons, days of the week and hours of the day (about 4% for UFPs and 12% for BC; Appendix B Table 26 - Table 27) .

Annual average estimates for each site were generally very similar when we looked at the 25 sites with available estimates for all of the different time-adjusted comparisons (i.e., all sites had the presented estimates available for UFP and/or BC, Appendix B Figure 70**Error! Reference source not found.**). ANOVA results looking at the effect of site and estimation method on concentration variability indicated that most of the variability was seen between sites (~99% for both UFPs and BC), while a tiny fraction (<1% for both UFPs and BC) was due to the estimation method (Appendix B Table 28 - Table 29). This indicated that using any of the examined averaging approaches would result in similar estimates.

Geocovariates

There were 183 geocovariates remaining after cleaning these (Appendix B Table 30). Geocovariates included proximity and buffer estimates for indicators of airports, coastlines, elevation, imperviousness, land use, NDVI, population density, ports, railroads and rail yards, roads (including truck routes) and water.

Universal Kriging

Validation

For the UFP UK model, three PLS components were selected through cross-validation (CV; Table 10). The variogram parameters for fitting the structured part of the residual term selected through CV were a plotting distance of 13,994 m (20% of the maximum distance between any two locations). The out-of-sample (validation set) RMSE and MSE-based R^2 for the UFP UK model were 933 pt/cm^3 and 0.87, respectively. Three PLS components and a plotting distance of 3,498 m (5% of the maximum distance between any two locations) was also selected for the BC model through CV. The out-of-sample RMSE and MSE-based R^2 for the BC UK model were 58 ng/m^3 and 0.85, respectively. The final UFP model had a partial sill of 0.010, a range of 4,160 m and a nugget of 0.011 (Appendix B Figure 73). The BC model had a partial sill of 0.014, a range of 960 m and a nugget of 0.004.

The following paragraphs detail the PLS regression and UK model fit. Figure 14 - Figure 15 show the geocovariate component loadings used in the PLS scores of the final UK models. The figure highlights which geocovariates and buffers had the largest weight (loading) in the PLS scores. In general, each covariate contributes differently to each component, as seen by changes in the loading magnitude and direction. Focusing on the first component (which characterizes most of the response in the outcome), NDVI seems to generally characterize most of the variability in UFP and BC (with smaller buffers having stronger associations), followed by distance to the nearest port and rail yard, imperviousness, land use, length of major roadways. Proximity to places such as coast lines and commercial areas as well as elevation have a smaller relative contribution. In general, UFP and BC geocovarite loadings are similar, with some of the largest differences seen in the second and third components for airport proximity, major road intersections, elevation, railroads, truck routes and waterways (Appendix B Figure 74).

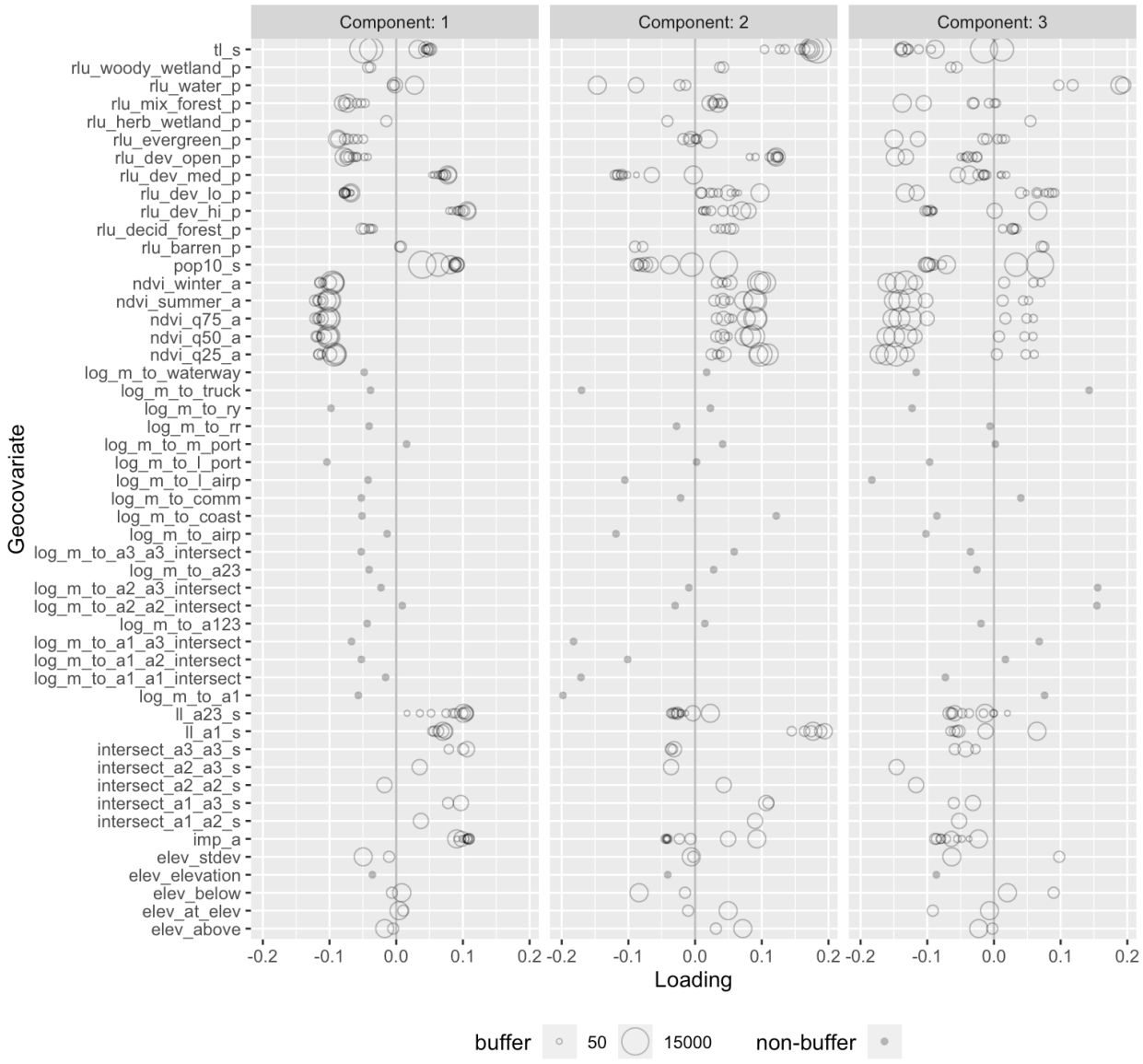


Figure 14. PLS geocovariate component loadings for the UFP model. See Appendix B Table 29 for geocovariate label descriptions.

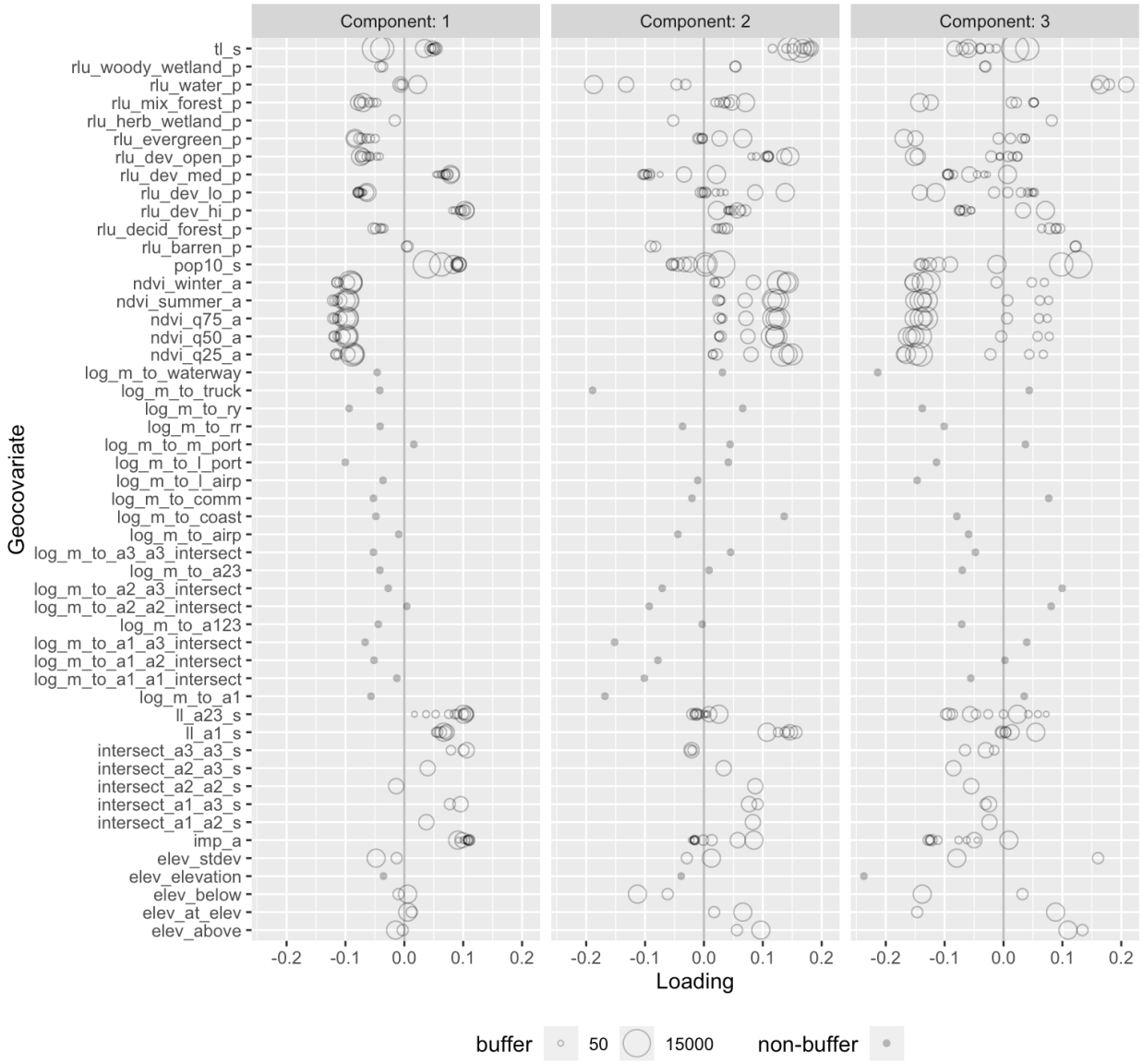


Figure 15. PLS geocovariate component loadings for the BC model. See Appendix B Table 29 for geocovariate label descriptions.

The PLS score distributions were overall similar for stops in the mobile monitoring region and ACT cohort locations in the study area for both the UFP and BC model (Appendix B Figure 71 - Figure 72). Some locations outside of the monitoring, but still within the monitoring area (N ~ 300) had slightly lower scores for PLS components one and two.

UK LUR Predictions vs Kriging

The mean, LUR piece of the UFP and BC UK models explained 89% and 84%, respectively, of the variability in the UK model predictions (LUR + kriging) at ACT cohort locations within the study area. The remaining variability in UK model predictions can be attributed to kriging (Figure 16). Figure 17 illustrates the degree of kriging at cohort locations in the study area for both the UFP and BC models. Locations with higher degrees of kriging (where paired points deviate most from the one-to-one line) indicate places where the observed concentrations from nearby mobile monitoring locations more heavily influenced that location's prediction (e.g., if there was an important geocovariate predictor missing from the LUR piece of the model, such as being northeast [downwind] of the airport). Higher degrees of kriging in UFP predictions were seen near the industrial district (the northernmost hotspot on the map) and at locations generally northeast of the airport (the southernmost hotspot on the map). The BC model had elevated degrees of kriging throughout the study area.

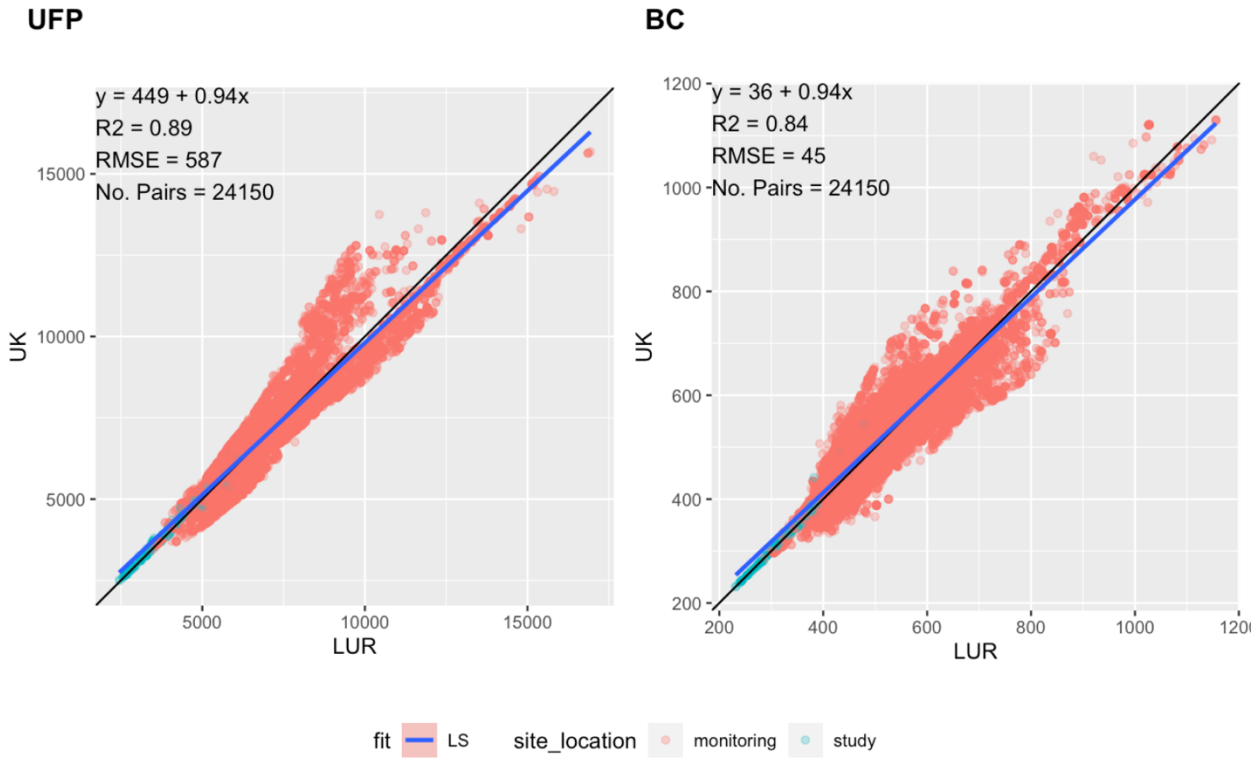


Figure 16. Decomposition of annual average predictions from the UK model: LUR vs. LUR + kriging at ACT cohort locations within the study area for UFP (pt/cm^3; left) and BC (ng/m^3; right)

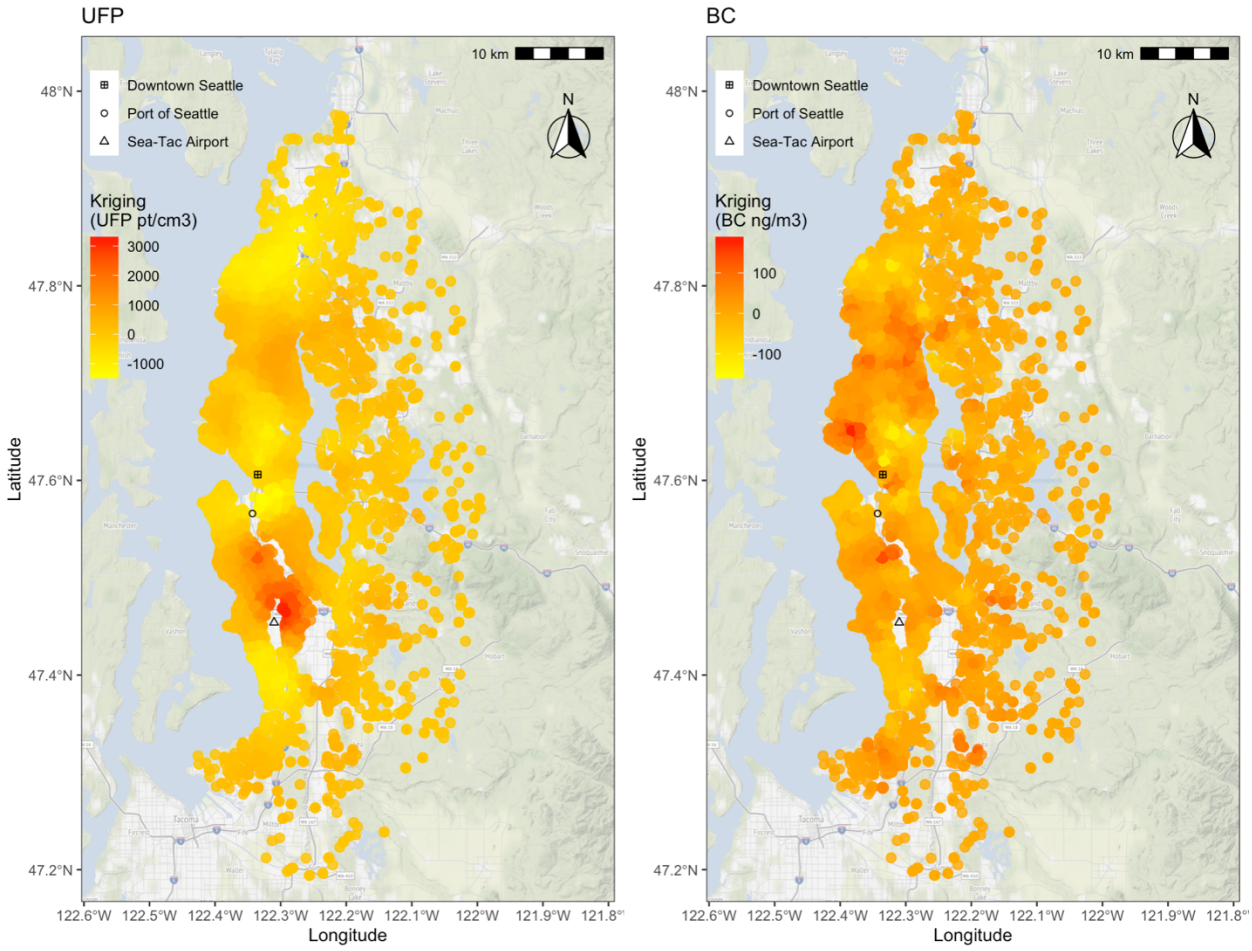


Figure 17. Kriging contributions in UK model predictions of annual average UFP (left) and BC (right) for ACT locations within the study area.

Universal Kriging Model Predictions

Annual average UFP exposure predictions for ACT cohort locations within the study area had a median (IQR) of 6,782 (1,788) pt/cm^3 and ranged from 2,491-15,682 pt/cm^3 (Table 9). BC exposure predictions had a median (IQR) of 525 (134) ng/m^3 and ranged from 232-1,130 ng/m^3.

Table 9. UFP and BC UK model predictions for ACT locations

Pollutant	N	Min	Q05	Mean	SD	Median	IQR	Q95	Max
UFP (pt/cm^3)	24,150	2,491	5,017	7,101	1,745	6,782	1,788	10,803	15,682
BC (ng/m^3)	24,150	232	389	538	115	525	134	727	1,130

In general, annual average UFP and BC predictions were highest near the downtown, industrial and airport areas as well as along major highways, with concentrations decreasing at locations away from these features (Figure 18).

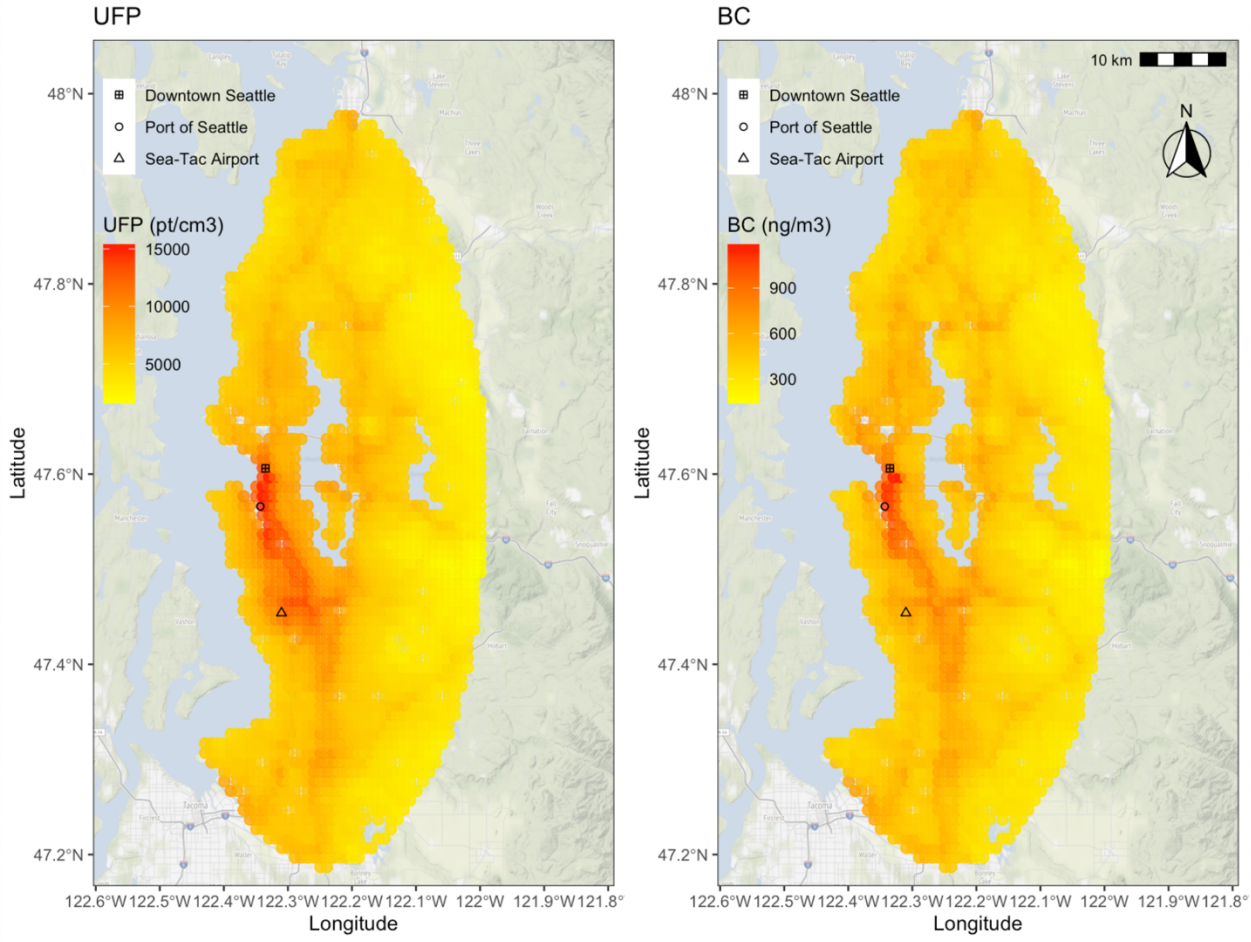


Figure 18. Annual average UFP (left) and BC (right) model predictions for the study area

Figure 19 shows linear regression fits for predicting UFP from BC levels. Panels show regression fits when estimates from mobile monitoring stops are used and regression fits when Cohort and Grid Predictions are used (three datasets). Plots show generally moderate to strong linear relationships (regression-based R^2: 0.65-0.90) between BC and UFP. Large positive residuals (concentration > 95th quantile) are generally near the Sea-Tac Airport, while large negative residuals (concentration < 5th quantile) are generally in the northern part of Seattle and along highways (Figure 20, Appendix B Table 31). Distance to large airport (log_m_to_l_airp) was consistently selected through lasso regression as an important predictor of large residuals in all three datasets (Figure 21). Distance to rail yard (log_m_to_ry) and distance to A1-A3 intersects (log_m_to_a1_a3_intersect) were also selected for the estimates and cohort predictions datasets, while distance to large port (log_m_to_l_port) was selected for the grid predictions dataset, among other variables.

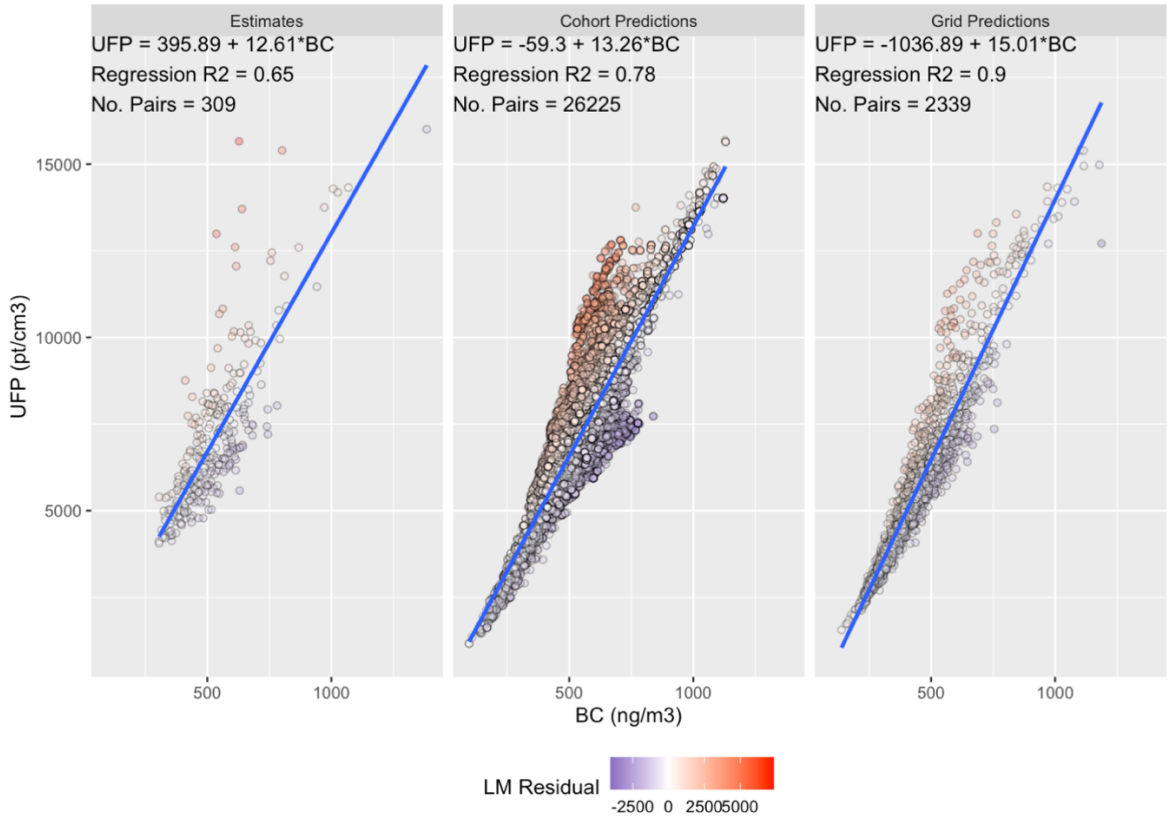


Figure 19. Comparison of UFP and BC estimates (from mobile monitoring observations) and model predictions (at ACT cohort grid locations) with linear regression fits. Pairs are colored by the linear regression residual.

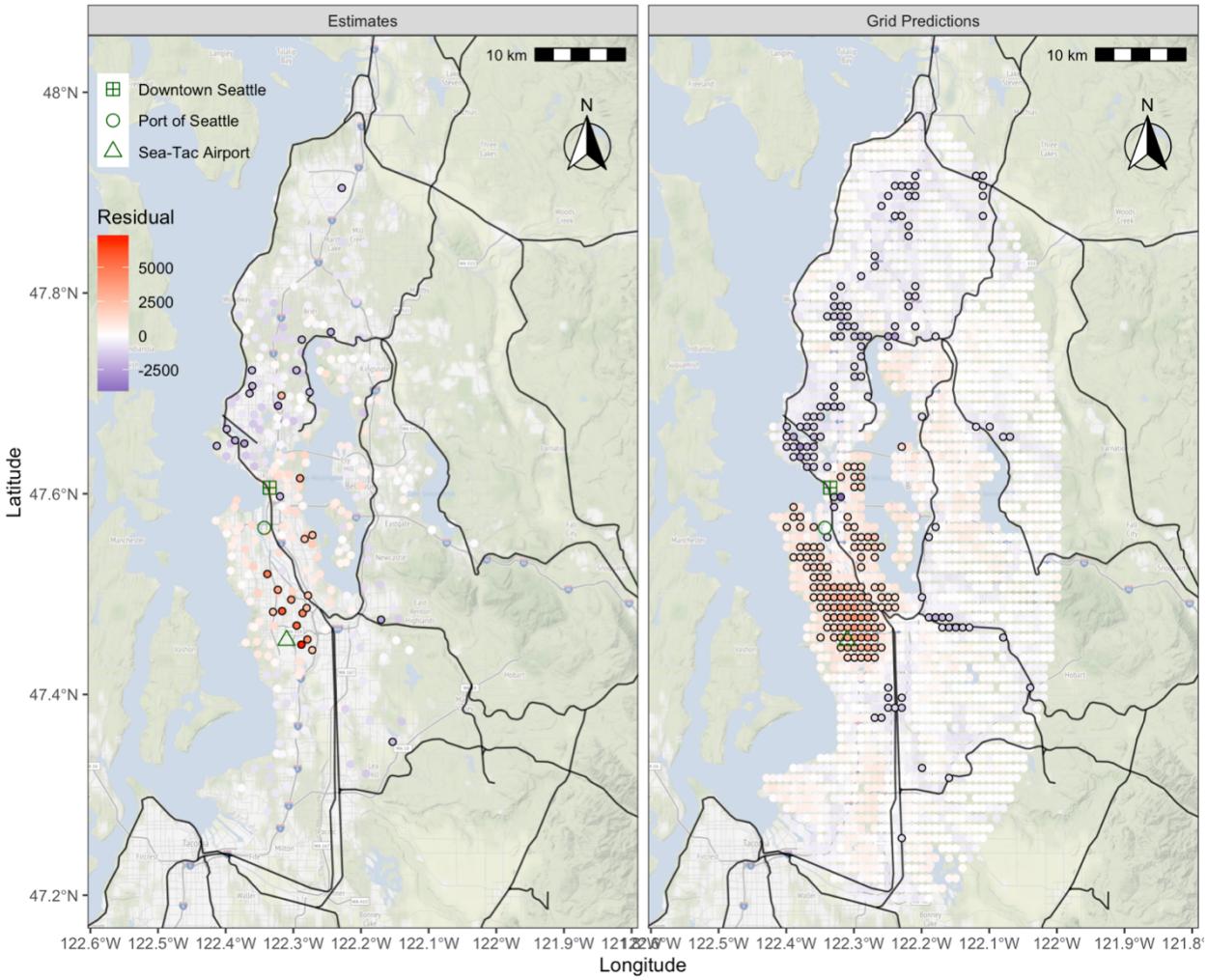


Figure 20. UFP residuals for UFP-BC linear regressions using estimates (from mobile monitoring observations) and model predictions (at grid locations). Black circles highlight the locations with high residuals in each dataset (< 5th quantile or > 95th quantile; see Appendix B Table 31 for values); black lines are railroads.

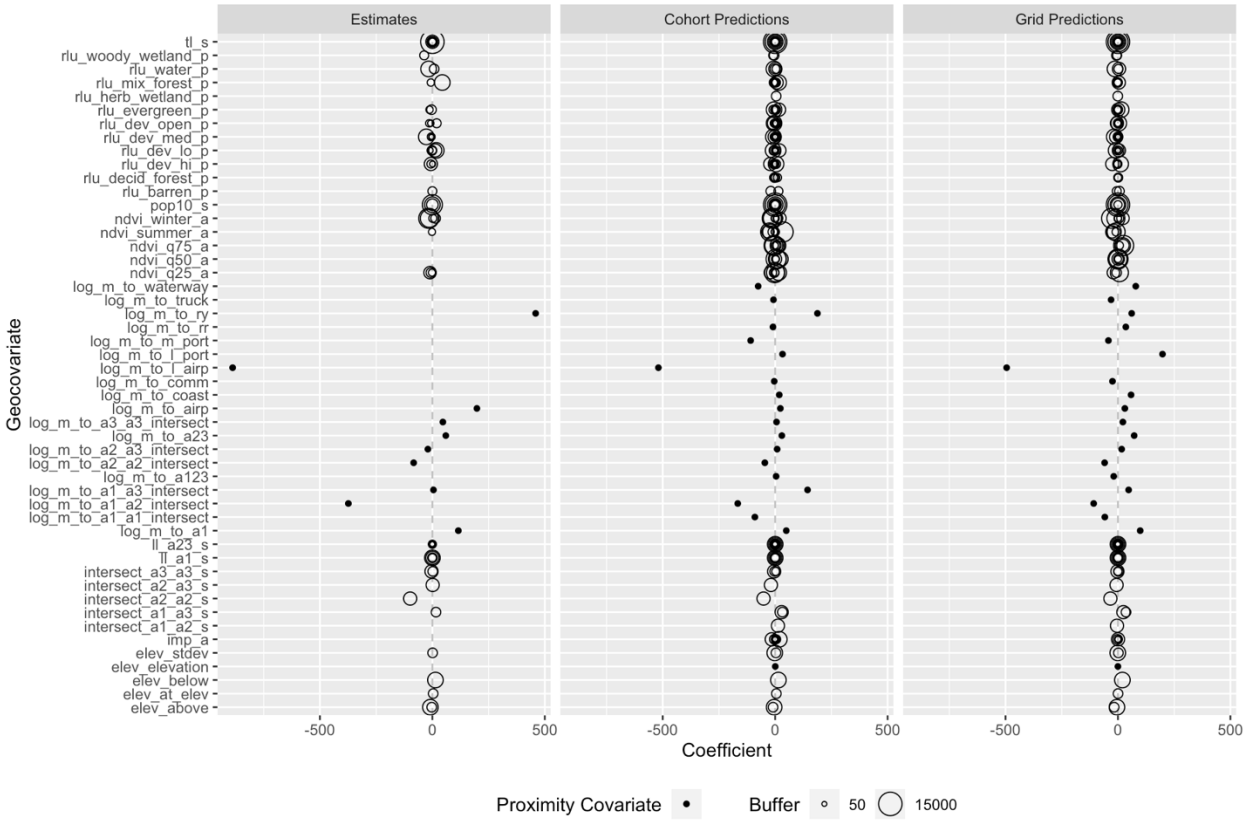


Figure 21. Lasso regression coefficient estimates for UFP residuals from UFP-BC linear regressions using estimates (from mobile monitoring observations) and model predictions (at ACT cohort grid locations). Covariates are all standardized.

Sensitivity Analyses

Annual average UFP and BC estimates at mobile monitoring stops were similar for primary and sensitivity analyses that consider different ways of estimating stop means. UFP estimates were slightly more similar across sensitivity analyses than BC estimates (Figure 22).

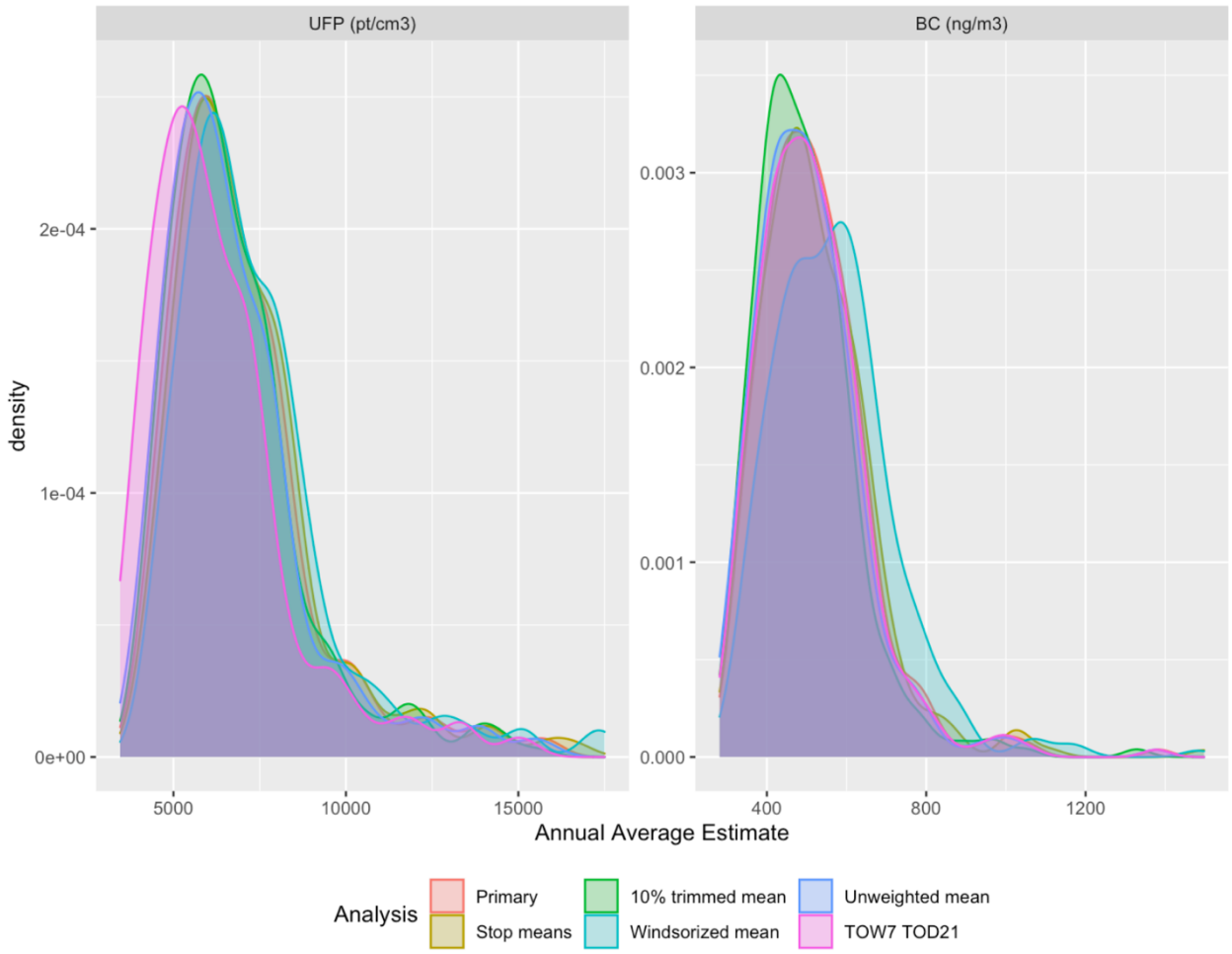


Figure 22. Distribution of annual average UFP and BC site concentration estimates for primary and sensitivity analyses. The left-most UFP density curve is for the TOW7 DOD21 analysis, while the right-most BC curve is for the Windsorized mean analysis.

The cross-validation (CV) and hold-out validation results were similar for sensitivity and primary analyses (Table 10). The number of PLS components selected through CV for UFP and BC sensitivity models ranged from 2-3 components. Hold-out RMSE and MSE-based R^2 values for the UFP models ranged from 850-1,239 pt/cm^3 and 0.84-0.89, respectively. Hold-out RMSE and MSE-based R^2 values for the BC models ranged from 57-70 ng/m^3 and 0.80-0.83, respectively.

Table 10. UK modeling parameters chosen through cross-validation for sensitivity and primary analyses

Analysis	PLS Comp	Variogram Max Dist Fraction	Variogram Dist (m)	CV[a] RMSE	CV [a] R^2	Hold Out RMSE	Hold Out R^2
UFP (pt/cm3)							
Primary	3	0.20	13994	984	0.77	933	0.87
Stop means	2	0.10	6785	1131	0.72	857	0.89
10% trimmed mean	3	0.30	20991	932	0.78	850	0.88
Windsorized mean	2	0.10	6997	1113	0.76	1239	0.84
Unweighted mean	2	0.10	6997	959	0.78	851	0.89
TOW7 TOD21[b]	2	0.10	6997	954	0.79	858	0.89
Native scale	2	0.20	13813	977	0.78	946	0.87
BC (ng/m3)							
Primary	3	0.05	3498	72	0.71	58	0.85
Stop means	3	0.05	3498	84	0.65	66	0.81
10% trimmed mean	3	0.10	6997	66	0.73	57	0.83
Windsorized mean	2	0.20	13994	87	0.68	70	0.82
Unweighted mean	2	0.05	3498	72	0.71	63	0.81
TOW7 TOD21[b]	2	0.05	3498	73	0.70	64	0.82
Native scale	3	0.05	3498	80	0.65	66	0.80

[a]Cross-validated; [b]Time of week (7 days) and time of day (21 hours) adjusted annual averages

UFP and BC predictions from primary and sensitivity analyses were all similar (Figure 23). On average, predictions for both pollutants from the stop means, Windsorized and native scale modeling analyses were slightly higher, while predictions from the 10% trimmed mean and unweighted annual average analyses were slightly lower. Maps of annual average UFP and BC prediction differences show where predictions from sensitivity analyses differed from the primary analyses (Figure 24 - Figure 25). For UFPs, the stop means analysis generally shows higher predictions throughout the study area, while the 10% trimmed mean largely does not change the predictions other than lowering them near the Sea-Tac Airport. The Windsorized analysis increases predictions at most places, with a special emphasis near the airport. The unweighted mean and native scale analysis change predictions in both the positive and negative direction but at different places on the map. The time-of-week (7 days) and time of day (21 hours) analysis reduces predictions throughout most of the study area other between Seattle and the airport near I-5. Overall, most analyses seem to consistently impact estimates near the airport where the highest UFP levels were observed. We also see changes in the BC predictions from each sensitivity analysis, though the spatial patterns are different. The Sea-Tac Airport, for example, does not stand out as much. Furthermore, the prediction contrasts seem more dramatic, possibly due to the measurement instrument's (aethalometer) inherent variability in its readings.

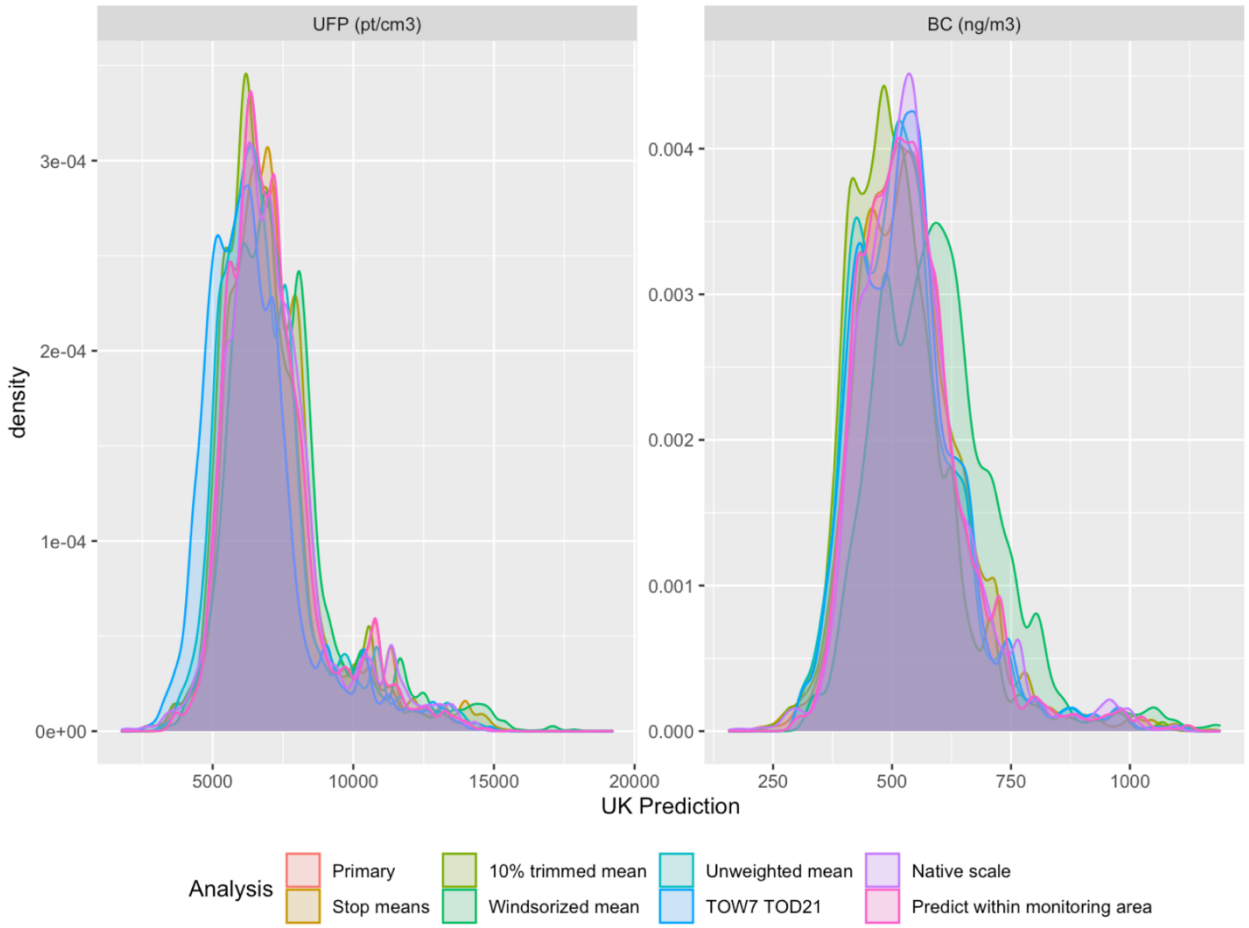


Figure 23. Distribution of UFP and BC predictions from primary and sensitivity analyses for ACT cohort locations in the study area. The left-most UFP density curve is for the TOW7 DOD21 analysis, while the right-most BC curve is for the Windsorized mean analysis.

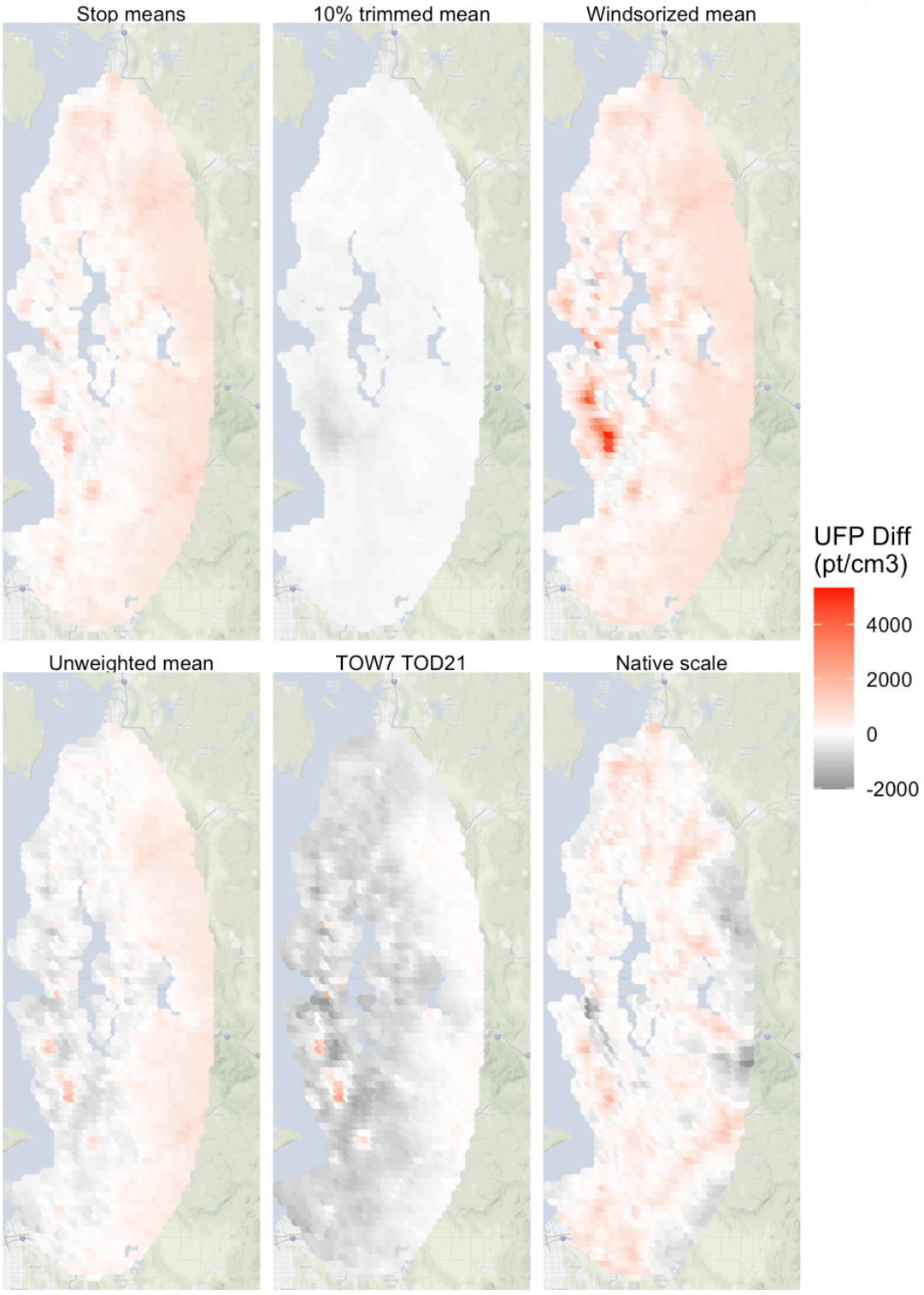


Figure 24. Annual average UFP prediction differences between sensitivity analyses and the primary analysis. A positive value indicates a higher predicted concentration from the sensitivity analysis compared to the primary analysis. Note that the windoszired map did not trim extreme observations but instead set these equal to a less extreme threshold (5^{th} or 95^{th} quantile).

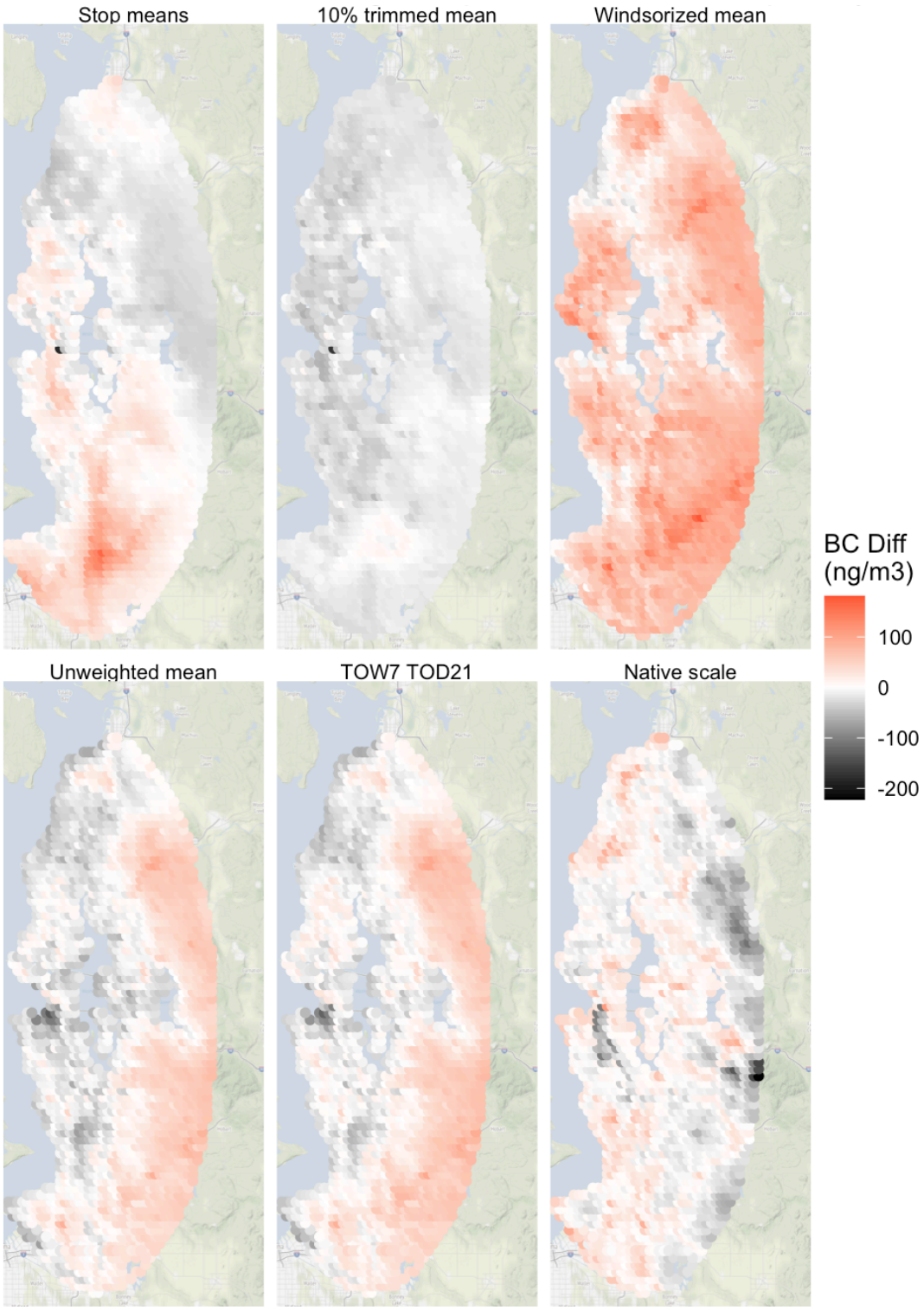


Figure 25. Annual average BC prediction differences between sensitivity analyses and the primary analysis. A positive value indicates a higher predicted concentration from the sensitivity analysis compared to the primary analysis.

Discussion

Novelty

This study was one of the first to characterize the long-term UFP and BC exposure surface with high spatial resolution for the greater Seattle area. This was made possible through our innovative mobile monitoring campaign, which we specifically developed for the assessment of long-term TRAP exposure in epidemiologic applications. This campaign is one of the largest mobile monitoring campaigns conducted when considered from many perspectives: the spatial extent covered (~1,100 km^2), sampling density (309 stops along 9 unique routes), sampling frequency (~5 AM – 11 PM, 7 days a week) and duration (~6 hours of driving per day, 275 days over one-year period covering all seasons). We were also able to differentiate between the UFP and BC exposure surfaces, highlighting how these differ most near major sources including large airports, rail yards and large road intersections. These findings suggest that using UFP surrogates such as BC could result in increased exposure misclassification near major pollutant sources and highlights the need for the additional monitoring of UFPs.

Given that the use of mobile monitoring campaigns for exposure assessments is relatively new, this study presented an approach for estimating annual averages from short-term, repeated measurements of TRAP, a research area that has received limited attention in the literature. Specifically, we presented an approach for reducing large amounts of high-resolution, temporally unbalanced and non-normally distributed data into long-term central tendencies that can afterwards be used in modeling. ANOVA models and descriptive analyses indicate that annual average estimates were minimally impacted when sites were visited during limited times of the day.

Universal Kriging Model

PLS Component Loadings

This study was one of first to use UK with PLS to build UFP and BC prediction models. The approach allowed us to build pollutant exposure models using a larger number of covariates and associated buffers than what other studies have used to characterize TRAP concentrations. The variability in component loadings for different covariate buffer sizes particularly highlighted the importance of buffer sizes. Typically, one or a few covariate buffer sizes were most strongly associated with UFP and BC concentrations.

Covariates with strong positive associations with UFP and BC were indicators of traffic (e.g., distance to and length of roadways, number of intersections), population density, developed land use, ports and railyards. Covariates with strong negative associations with UFP and BC were indicators of vegetation (e.g., forest, NDVI). This is in line with much of the literature in this field (Abhijith et al., 2017; Apte et al., 2017; Hatzopoulou et al., 2017; Hoek et al., 2008; Kerckhoffs et al., 2016; Lane et al., 2016; Patton et al., 2015; Patton, Collins, et al., 2014; Saha, Li, et al., 2019; Weichenthal, Ryswyk, et al., 2016; Xie et al., 2017). Airport proximity was a strong predictor of large differences between the UFP and BC exposure surface, with closer proximities being more likely to have relatively higher UFP levels. A recently published study investigating the UFP profiles associated with aircraft and roadway traffic also found differences in the UFP and BC exposure surfaces, particularly near major sources (Austin et al., 2021).

Smaller buffers for NDVI, population density and imperviousness tended to be more strongly associated with both UFP and BC concentrations. This is in line with our understanding that these pollutants are strongly impacted by nearby sources but quickly disperse or react, causing them to have small spatial scales. A synthesis of field studies by Karnet et al. (2010), for example, found that UFP and elemental carbon (associated with BC) concentrations dropped down to background levels within 200-400 m of roadways (Karner et al., 2010). Other studies have reported similar findings, with buffers within a few hundred meters of a target location often being most associated with UFPs and BC (Jones et al., 2020; Saha, Li, et al., 2019; Weichenthal, Ryswyk, et al., 2016).

Interestingly, water land use covariates were weakly associated with both UFP and BC concentrations. While proximity to bodies of water can sometimes be an indicator of a lack of emission sources, this is not necessarily true for Seattle since its urban and industrial areas are both located near bodies of water.

Distribution of PLS Scores

PLS score distributions were similar between the mobile monitoring stops and ACT cohort locations in the study area. Still, there was some indication of slightly different PLS scores for some of the ACT locations within the study area but outside of the monitoring area. This suggests that predictions for ACT cohort locations in the study area are mostly within the geocovariate modeling space, though some of the more distal locations (particularly the easternmost locations) may be on the edge. Thus, model predictions at these more distal areas may not be as accurate, and this could have important implications for epidemiologic investigations that use these predictions.

Residual Model Parameters

The ranges used in the UK residual models for UFP and BC were fairly short (~1-2 km, with BC having slightly higher ranges). This is in line with the understanding that UFPs and BC have small spatial scales, especially when compared to other pollutants such as $PM_{2.5}$ (Karner et al., 2010; Mercer et al., 2011; Sampson et al., 2013; Young et al., 2016). Sampson et al. (2013), for example, reported variogram ranges from 35-2,944 km for PLS-UK $PM_{2.5}$ models build from geocovariates and regional AQS sites throughout the U.S. (Sampson et al., 2013). Young et al. (2016) similarly reported variogram ranges for PLS-UK $PM_{2.5}$ models built form a combination of regional AQS site monitoring data and satellite images ranging from roughly 100-600 km (Young et al., 2016). On the other hand, Mercer et al. (2011) reported slightly shorter variogram ranges of about 6 km – 15 km for UK NOx models built from about 150 sites in Los Angeles designed to capture both local concentrations and very near-field TRAP gradients (Mercer et al., 2011). NOx has been shown to generally decay slower than UFPs but faster than $PM_{2.5}$ (Karner et al., 2010).

Model Decomposition

The LUR, not the kriging piece of the UK models predicted most of the variability in the predictions (79%). This suggests that model predictions rely most heavily on geocovariates rather than the additional spatial correlation from nearby observations from the mobile monitoring campaign. Given that we have geocovarites for all ACT locations in the study area and that the distribution of these are very similar to those of the mobile monitoring stops, the application of these models for all ACT locations in the study area should be informative. Furthermore, the areas that were most heavily kriged by the model were generally those with elevated predicted pollutant concentration. For the UFP model, the areas near the Port of Seattle, Industrial District and the Sea-Tac Airport were the most heavily kriged. It was less clear whether kriging was systematic in the BC model, though the areas near the airport, along railroads and along busy truck routes tended to be have higher degrees of kriging. Since these features have existed for decades, this supports the extrapolation of the 2019 spatial surface kriging back to the mid 1990s (Aim 3). It also suggests that developing a covariate to better characterize this area (e.g., distance to airport landing strip, being downwind of the airport, aviation traffic, industrial facilities, etc.) could allow the LUR piece of the model to better capture changes in pollutant levels near these places.

Model Fit

Models performed very well in out-of-sample validation. The UFP and BC model predicted 87% and 85% of the variability in the observed concentrations, respectively. These performances are higher than what many short-term stationary and mobile monitoring campaigns have achieved (UFP: R^2 ~ 13-72%, BC: R^2 ~ 12-86%) and closer to that of many fixed site continuous sampling campaigns (UFP: R^2 ~ 67-89%, BC: R^2 ~ 39-82%; Figure 26, Table 11, Table 12) [UFP: (Abernethy et al., 2013; Rob Beelen et al., 2013; Cattani et al., 2017; Farrell et al., 2016; Hankey & Marshall, 2015b; Hoek et al., 2011; Kerckhoffs et al., 2016; Minet et al., 2018; Montagne et al., 2015; Patton et al., 2015; Ragettli et al., 2014; Rivera et al., 2012; Sabaliauskas et al., 2015; Saha, Li, et al., 2019; Saraswat et al., 2013; Simon et al., 2018; van

Nunen et al., 2017; Weichenthal, Ryswyk, et al., 2016; Weichenthal, Van Ryswyk, et al., 2016; Wolf et al., 2017; C. H. Yu et al., 2016); BC: (R Beelen et al., 2007; Brauer et al., 2003; Carr et al., 2002; Dodson et al., 2009; Hankey & Marshall, 2015b; Henderson et al., 2007; Hochadel et al., 2006; Kerckhoffs et al., 2016; Timothy Larson et al., 2007, 2009; Montagne et al., 2015; Morgenstern et al., 2007; Ryan et al., 2007; Saraswat et al., 2013; Su et al., 2013; C. H. Yu et al., 2016)]. This strong model performance held when we restricted our comparison to other studies that have also modeled site averages from multiple site visits rather than using individual site visits in their models and evaluations (see "multi-site avg modeled" and "analysis uses site mean" in Figure 26 – Figure 28 and Table 11 – Table 12).

There are several features of our study design that could have impacted our strong model performances. For UFPs, Saha et al. (2019) reported that short-term stationary (collecting short-term samples while stopped, as opposed to while moving or traditional fixed long-term sampling) studies like ours have generally sampled between 60-644 sites, sampled each site between 15 minutes and 3 hours, and collected between 1-5 repeat samples per site. Similarly, BC studies like this one have generally sampled 26-161 sites, sampled each site about 30 minutes, and collected about 2-3 samples per site. Campaigns with more sites counts have generally collected fewer repeat samples per site. Compared to earlier studies, we sampled more sites than most fixed and short-term stationary studies (309 sites). This dense monitoring network covered a larger geographic area and likely allowed us to capture hotspots that may have otherwise been missed by more sparse monitoring networks. Additionally, we visited each site for shorter periods of time (2 minutes) and collected more repeat samples per site than what most studies have done (~24; UFP: Figure 27, BC: Figure 28). And while our resulting total site sampling durations (~48 minutes) were similar to other short-term stationary studies, we sampled year-around during each day of the week and most times of the day, which most studies have not done. This design may have better captured the temporal variability of TRAP despite collecting short-term samples, thus improving the accuracy and precision of our long-term estimates. The collection of stationary (vs mobile) samples may have additionally helped. Moreover, long-term averaging reduces variability in the data and, assuming this variability is not predictable (i.e., noise), could have resulted in more precise (higher R^2) estimates than if we had modeled concentrations for smaller time scales (e.g., 1-second readings or stop medians).

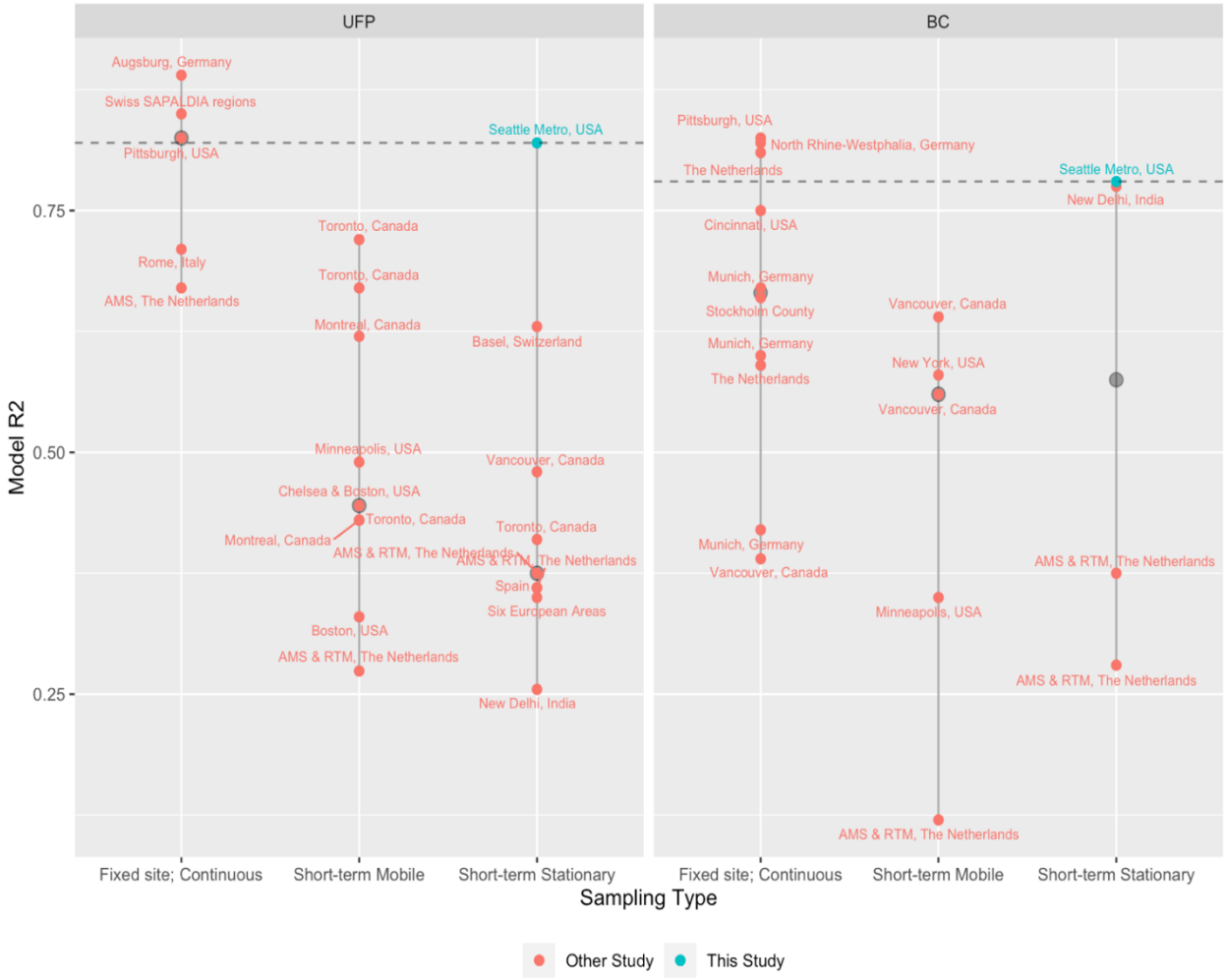


Figure 26. Model R^2 estimates from other UFP and BC studies. Gray dot and vertical line are the median and range. Horizontal dashed line is the R^2 for this study. Plots show the average R^2 from a study if multiple models were presented without a clear primary model. See Table 11 - Table 12 for details.

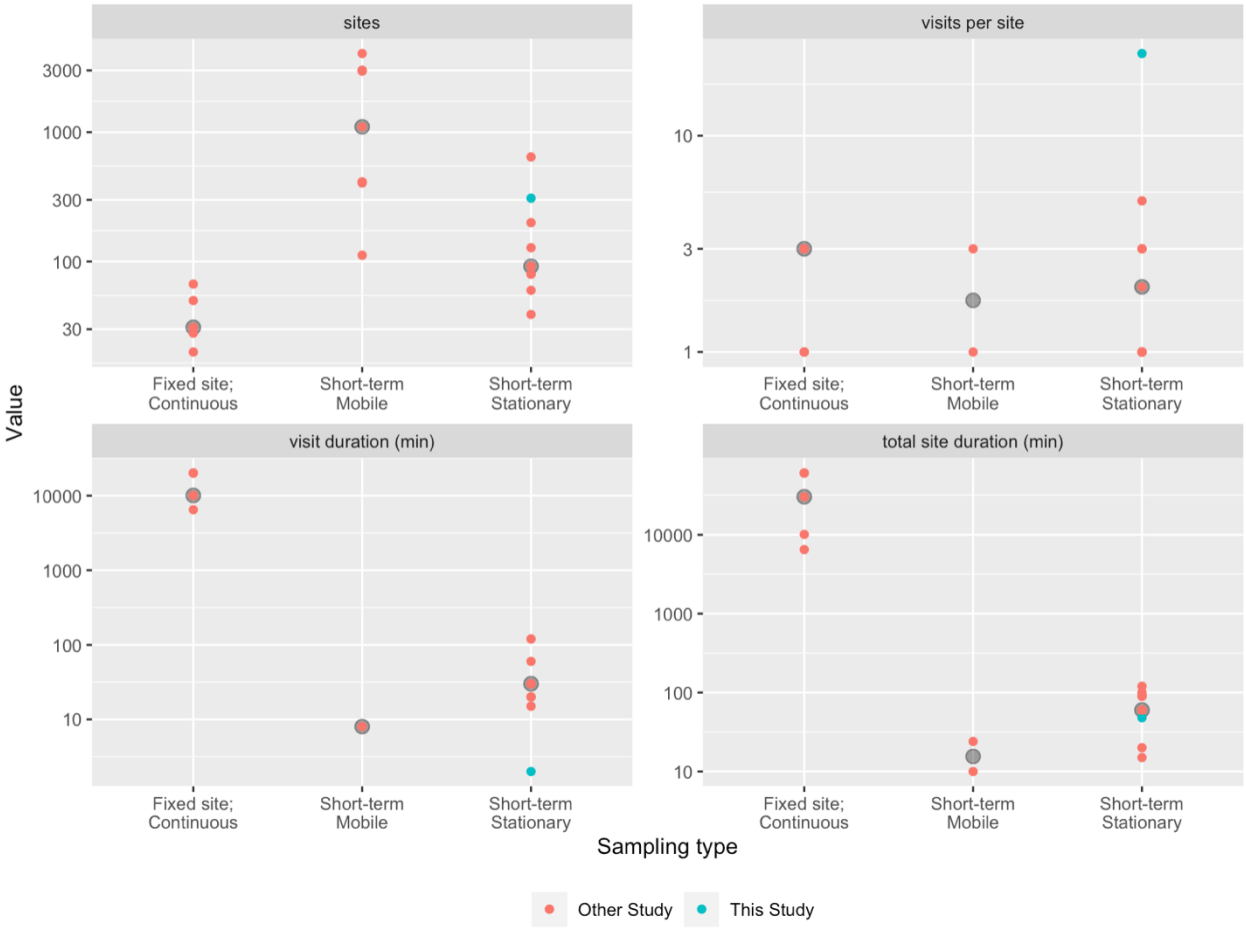


Figure 27. Sampling approaches across other UFP studies (see Table 11 for details). Gray dot is the median. Note that little data were available for short-term mobile studies regarding visit duration (N=1), total site duration (N=2) or visits per site (N=2). The single study under short-term mobile visit duration of ~ 8 min was conducted with pedestrians (Sabaliauskas et al. 2015).

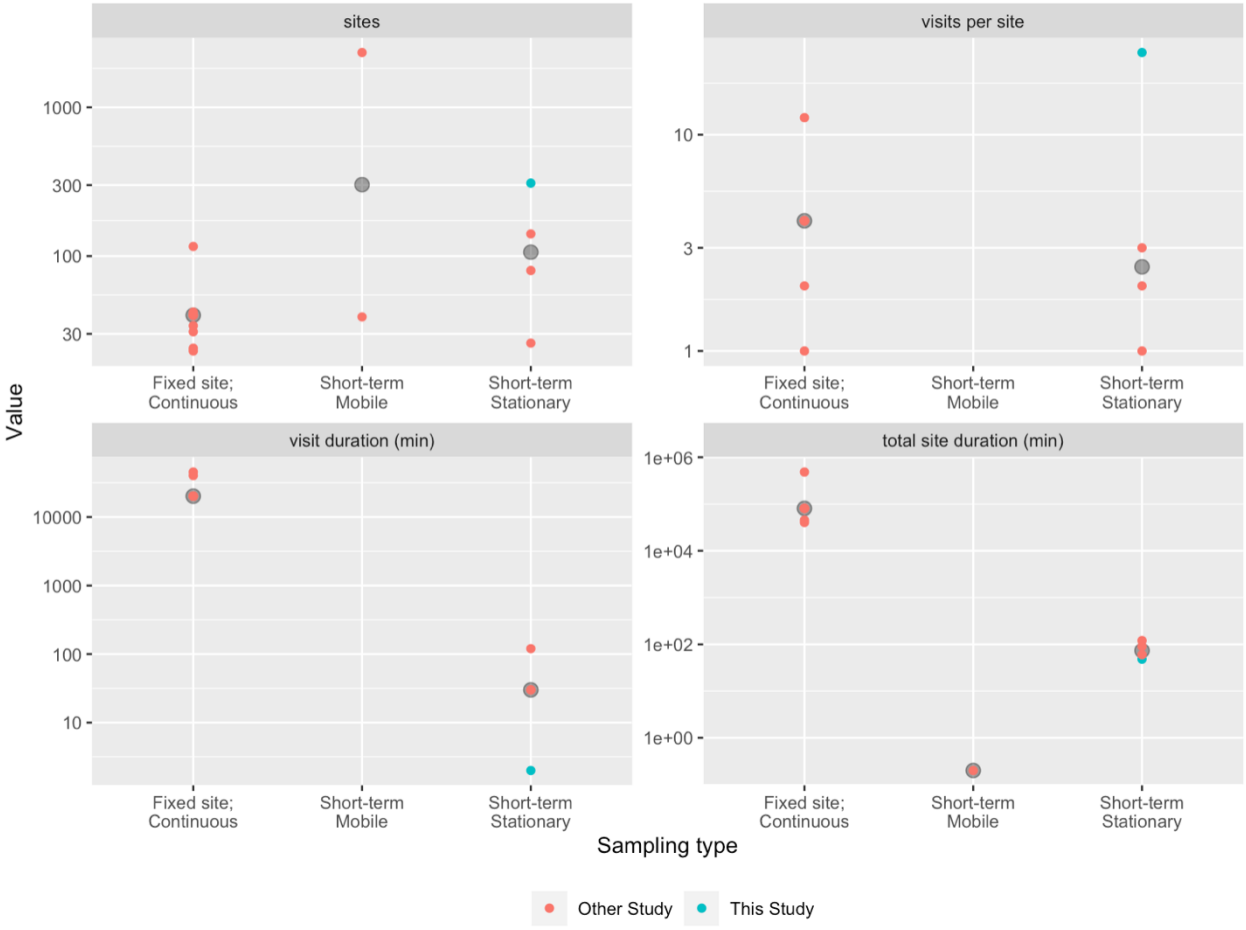


Figure 28. Sampling approaches across other BC studies (see Table 12 for details). Gray dot is the median.

Table 11. Model R^2 estimates from other LUR UFP studies. Modified from Saha, Li, et al., 2019 Table S6.

Study	Location[a]	Sampling	Analysis uses site mean	Model R2[b]
Fixed site; Continuous				
Beelen et al. (2007)	The Netherlands	continuous sampling/site; 23 regional air quality sites	yes	0.59
Brauer et al. (2003)	The Netherlands	2 wk/site; 4 repeats; 40 sites	yes	0.81
Brauer et al. (2003)	Munich, Germany	2 wk/site; 4 repeats; 40 sites	yes	0.67

Brauer et al. (2003)	Stockholm County	2 wk/site; 4 repeats; 42 sites	yes	0.66
Carr et al. (2002)	Munich, Germany	4 wk/site; 12 repeats; 34 sites	yes	0.40-0.80
Henderson et al. (2007)	Vancouver, Canada	2 wk/site; 2 repeats; 116 sites	yes	0.39
Hochadel et al. (2006)	North Rhine-Westphalia	2 wk/site; 4 repeats; 40 sites	yes	0.82
Morgenstern et al. (2007)	Munich, Germany	2 wk/site; 4 repeats; 40 sites	yes	0.42
Ryan et al. (2007)	Cincinnati, Ohio, USA	24 sites	no	0.75
Saha, Li, et al. (2019)	Pittsburgh, PA, USA	3-6 wk/site; 31 sites	yes	0.80-085
Short-term mobile				
Hankey & Marshall et al. (2015)	Minneapolis, MN	Bicycle; 85 hr over 42 runs; 1,101 aggregation loc	no	0.28-0.42
Kerckhoffs et al. (2016)	AMS & RTM, The Netherlands	car; ~12s/road segments; 2964 total segments (2336 used in model)	yes	0.12
Larson et al. (2007)	Vancouver, Canada	car	yes	0.64
Larson et al. (2009)	Vancouver, Canada	car; 39 sites	no	0.56
Su et al. (2013)	New York, USA	car; 10 nights	no	0.57
Short-term stationary				
Saraswat et al. (2013)	Delhi, India	1-3 hrs/site; no repeat, 26 sites	no	0.69-0.86
Blanco et al. (2021)	Seattle Metro, WA, USA	2 min/site x 24 repeats; 309 sites	yes	0.77-0.87
Kerckhoffs et al. (2016)	AMS & RTM, The Netherlands	30 min/site; 2 repeats; 161 sites (141 used in model)	yes	0.28

Montagne et al. (2015)	AMS & RTM, The Netherlands	30 min/site; 3 repeats; 80-81 sites/city	yes	0.35-0.40

[a]AMS = Amsterdam; MMa = Maastricht; RTM = Rotterdam; UT = Utrecht

[b] R^2 ranges are presented if multiple estimates were presented with no clear primary model. Most studies do not clarify what R^2 they are presenting (in-sample vs out-of sample or cross-validated; regression- vs MSE- based). We present a range for ours, with the lower value being the cross-validated MSE-based R^2 and the higher value being the out-of-sample MSE-based R^2.

Table 12. Model R^2 estimates from other LUR BC studies

Study	Location[a]	Sampling	Analysis uses site mean	Model R2[b]
Fixed site; Continuous				
Beelen et al. (2007)	The Netherlands	continuous sampling/site; 23 regional air quality sites	yes	0.59
Brauer et al. (2003)	The Netherlands	2 wk/site; 4 repeats; 40 sites	yes	0.81
Brauer et al. (2003)	Munich, Germany	2 wk/site; 4 repeats; 40 sites	yes	0.67
Brauer et al. (2003)	Stockholm County	2 wk/site; 4 repeats; 42 sites	yes	0.66
Carr et al. (2002)	Munich, Germany	4 wk/site; 12 repeats; 34 sites	yes	0.40-0.80
Henderson et al. (2007)	Vancouver, Canada	2 wk/site; 2 repeats; 116 sites	yes	0.39
Hochadel et al. (2006)	North Rhine-Westphalia, Germany	2 wk/site; 4 repeats; 40 sites	yes	0.82
Morgenstern et al. (2007)	Munich, Germany	2 wk/site; 4 repeats; 40 sites	yes	0.42
Ryan et al. (2007)	Cincinnati, USA	24 sites	no	0.75

Short-term mobile				
Hankey & Marshall et al. (2015)	Minneapolis, USA	Bicycle; 85 hr over 42 runs; 1,101 aggregation loc	no	0.28-0.42
Kerckhoffs et al. (2016)	AMS & RTM, The Netherlands	car; ~12s/road segments; 2964 total segments (2336 used in model)	yes	0.12
Larson et al. (2007)	Vancouver, Canada	car	yes	0.64
Larson et al. (2009)	Vancouver, Canada	car; 39 sites	no	0.56
Su et al. (2013)	New York, USA	car; 10 nights	no	0.57
Short-term stationary				
Saraswat et al. (2013)	New Delhi, India	1-3 hrs/site; no repeat, 26 sites	no	0.69-0.86
Blanco et al. (2021)	Seattle Metro, USA	2 min/site x 24 repeats; 309 sites	yes	0.71-0.85
Kerckhoffs et al. (2016)	AMS & RTM, The Netherlands	30 min/site; 2 repeats; 161 sites (141 used in model)	yes	0.28
Montagne et al. (2015)	AMS & RTM, The Netherlands	30 min/site; 3 repeats; 80-81 sites/city	yes	0.35-0.40

[a]AMS = Amsterdam; MMa = Maastricht; RTM = Rotterdam; UT = Utrecht

[b] R^2 ranges are presented if multiple estimates were presented with no clear primary model. Most studies do not clarify what R^2 they are presenting (in-sample vs out-of sample [cross-validated or hold-out]; regression- vs MSE- based). We present a range for ours, both of which are out-ofo-sample. The lower value is the cross-validated MSE-based R^2 while the higher value is the hold-out MSE-based R^2.

In terms of modeling, we used PLS regression to summarize hundreds of geographic covariate-buffer combinations. Most other studies have started with fewer covariates and used covariate selection methods (e.g., forward stepwise regression) that eliminate "non-informative" covariates rather than downweighing them like PLS. This may have allowed us to explain more of the TRAP variability, particularly in such a large, geographically diverse study area.

Moreover, our dense monitoring network means that randomly selected test set locations for our model validations were likely near training set locations and thus similar. Since this monitoring network was specifically designed to be representative of the ACT cohort, this indicates that models should perform similarly well when predicting at ACT cohort locations within the study area. Notably, we performed hold-out validation in addition to cross-validation and calculated an MSE-based R^2. This has not always been applied in mobile monitoring LUR studies and can often result in lower model performance than in-sample and regression-based R^2 approaches (Montagne et al., 2015). Interestingly, we saw slightly better performances in hold-out rather than cross-validation even after using different hold-out validation sets to evaluate the potential variability of our model performance estimates (data not shown). Our higher out-of-sample model performance may have to do with having a small number of test sites, all of which were near the center of our study area where our monitoring network was dense. Ten-fold cross-validation, on the other hand, allowed most sites (90%) to be in a test set, including those in areas where our monitoring network was sparser (i.e., near the edges).

Exposure Estimates for the ACT Cohort

Average UFP exposure predictions for ACT cohort locations within the study area were around 7,000 pt/cm^3 and ranged from about 2,400-15,000 pt/cm^3. Average BC predictions were around 500 ng/m^3 and ranged from about 200-1,100 ng/m^3. The estimated six-fold difference between the lowest and highest predictions suggests that there is a large degree of variation in UFP and BC exposure for cohort members residing in the greater Seattle area.

Our study results were generally lower and less variable than what other studies have observed in cities around the world investigating UFPs and BC (Figure 29) (UFP: (Abernethy et al., 2013; Farrell et al., 2016; Hankey & Marshall, 2015a; Kerckhoffs et al., 2016, 2017; Minet et al., 2018; Montagne et al., 2015; Patton et al., 2015; Ragettli et al., 2014; Rivera et al., 2012; Saha, Zimmerman, et al., 2019; Saraswat et al., 2013; Simon et al., 2018; van Nunen et al., 2017; Weichenthal, Ryswyk, et al., 2016; Weichenthal, Van Ryswyk, et al., 2016; C. H. Yu et al., 2016); BC: (Apte et al., 2017; R Beelen et al., 2007; Brauer et al., 2003; Carr et al., 2002; Dodson et al., 2009; Hankey & Marshall, 2015b; Henderson et al., 2007; Hochadel et al., 2006; Kerckhoffs et al., 2016, 2017; Timothy Larson et al., 2007, 2009; Minet et al., 2018; Montagne

et al., 2015; Morgenstern et al., 2007; Ryan et al., 2007; Saraswat et al., 2013; Su et al., 2013; C. H. Yu et al., 2016)). When compared to other long-term UFP exposure measurements in the U.S., our results were similar to suburban and rural locations (Presto, 2020).

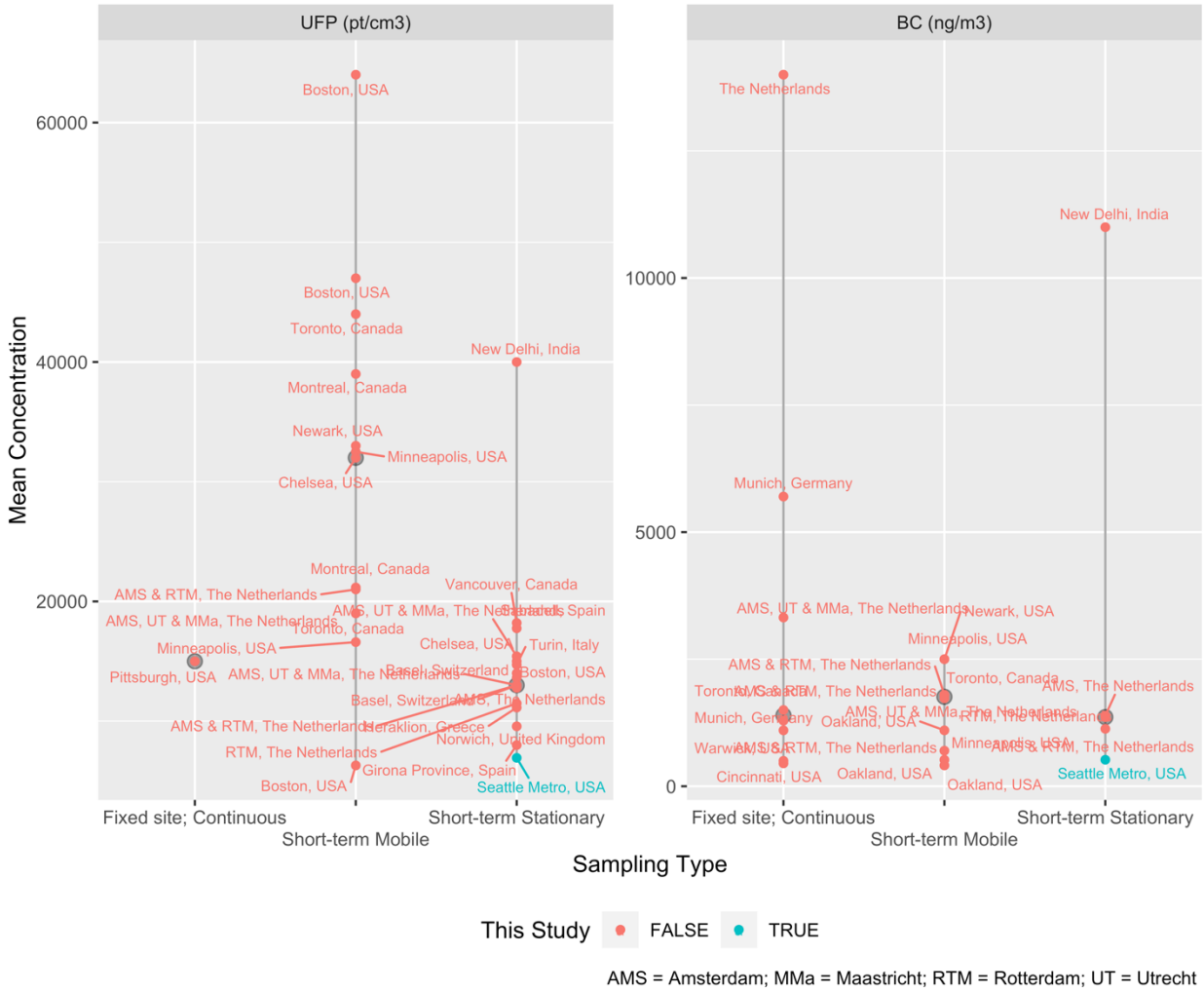


Figure 29. Mean UFP and BC levels reported in other studies. Mean UFP and BC for this study (Seattle Metro, WA) is for ACT cohort locations.

There are some possible reasons for our lower TRAP concentration estimates when compared to other studies. For one, we sampled a wider geographic area over a wider range of times than what many studies have done. Most studies, for example, have focused their sampling efforts near major sources (e.g., highways, truck routes, urban areas, airports) and during known high concentration times (e.g., during rush hour times, on weekdays). Our annual averaging time

was also larger than some of the previous studies, and this may have reduced both concentration estimates and their variability. Furthermore, Seattle generally has cleaner air than other major cities where some of these studies have been conducted.

Of note, while UFPs are defined as particles less than 100 nm (0.1 μm) in diameter, the P-TRAKs used in this analysis measured particles 20-1,000 nm in diameter. Since most particles by count are below 100 nm (Ramachandran, 2005; US EPA, 2019e), our instruments' upper size range should minimally impact our concentrations, while the lower size range (20 nm rather than 0 nm) may have led to slightly underestimated total UFP count concentrations. Of the studies that we reviewed, none strictly defined UFPs as particles less than 100 nm (Appendix B Figure 75). Studies often use different particle size ranges with some counting particles up to 3,000 nm in diameter. Our analyses indicate, however, that these discrepancies are not solely responsible for the large concentration differences reported. Still, given that there is no gold standard for the collection of UFPs, the EPA has noted that this discrepancy in the particle ranges collected in different "UFP" studies makes it difficult to truly compare results across studies, and has advocated for a more uniform definition (US EPA, 2019e).

Mapping UFPs and BC for the Greater Seattle Area

The results of this analysis highlight TRAP hotspots and concentration gradients that may have otherwise been missed by traditional stationary monitoring or mobile monitoring designs not intended for long-term, spatially granular, human exposure assessment. Models predicted elevated pollutant levels near highways, downtown Seattle, the Port of Seattle, Seattle's International and Industrial districts, the Sea-Tac Airport and other locations. Other studies have also found elevated UFP and BC concentrations associated with vehicles (Apte et al., 2017; Karner et al., 2010; T Larson et al., 2017; Riley et al., 2014), aircraft activity (Austin et al., 2019; Dodson et al., 2009; Hatzopoulou et al., 2017; Hudda et al., 2014; Riley, Schaal, et al., 2016), ports and marine activity (Kozawa et al., 2009; Lack & Corbett, 2012), and industry (Apte et al., 2017; Hatzopoulou et al., 2017; Riffault et al., 2015). When we consider that these areas house and employ a large number of individuals, public health interventions and policies to reduce TRAP exposure levels at these locations could make a significant impact. From an environmental justice perspective, individuals with a lower socioeconomic status could

particularly benefit from these efforts given that they are more likely to reside in areas with higher levels of air pollution and be impacted by adverse health outcomes (Chi et al., 2016; L. P. Clark et al., 2017; Hajat et al., 2013, 2015).

Moreover, our findings suggest that there likely added value of monitoring for UFPs despite these generally being strongly correlated with BC, a more frequently monitored pollutant, particularly for epidemiologic cohorts with locations near large UFP and BC emission sources such as airports and rail yards. At these locations, the relative concentrations of each of these pollutants may differ, making it difficult to predict one from the other. If BC models were to be used as a surrogate for UFP exposure in epidemiologic studies such as the ACT cohort, for example, this could result in a larger degree of exposure misclassification for participants living near these locations.

Sensitivity analyses

The predicted annual average UFP and BC concentrations were similar across sensitivity analyses, suggesting that our findings are robust and unlikely to be an artifact of our modeling decisions.

Overall, UFP and BC predictions were slightly higher for the Windsorized, stop means and native scale modeling sensitivity analyses than the primary analyses since higher stop concentrations were included in these stop annual average estimates. Locations near the Sea-Tac Airport generally saw the largest UFP prediction change when taking different approaches, more so than BC. This reflects the particularly high UFP levels that were observed in this area when compared to the rest of our study area and in relation to BC. It also suggests that approaches that allow for high concentrations to be more influential (e.g., Windsorizing) could better characterize the variability of the UFP exposure surface (e.g., due to flight path patterns). If this exposure assessment approach was applied for epidemiologic investigations, this could result in a greater degree of exposure misclassification for higher UFP concentration areas, which could possibly impact epidemiologic inferences. Both alternative methods of UFP exposure assessment and the impacts on epidemiologic inferences should be studied in the future.

On the other hand, UFP and BC predictions were generally slightly lower from the 10% trimmed mean sensitivity analyses than the primary analyses because fewer high concentrations

were included in these annual averages. UFP predictions were slightly lower when we used regression adjusted annual average site estimates that more thoroughly adjusted for day of the week and time of day to build UK models. This may have been a result of upweighting a few low-concentration hours that had few samples (e.g., 11 AM) more heavily than in our primary analysis. The weights used in this analysis are thus most representative of the sites that were sampled during more times of the day, though this appeared to have a small effect, particularly on BC predictions.

Maps of prediction differences relative to our primary analyses indicated, however, that some methods are more likely to result in both over- and under-estimation of site concentrations, depending on their location. If we consider predictions from the primary analysis to be closest to the true annual average, this suggests that some of these methods could result in differential misclassification if applied to epidemiologic investigations. Native scale modeling predictions of UFP, for example, tended to be higher throughout most of our modeling region except the easternmost areas where predictions tended to be lower than our primary model. We saw similar over- and under-estimation patterns for BC predictions from models that used stop means (vs medians), unweighted means and finer temporal adjustment methods to estimate the annual averages used to build our models. Future work is needed to better understand whether these differences could have a meaningful impact on epidemiologic inferences.

Limitations

This work had several potential limitations. First, we did not include difficult to obtain, built environment covariates that can influence TRAP concentrations such as street width or building height (Hoek et al., 2008; Weissert et al., 2018). Our prediction models, however, did have other geocovariates that could serve as surrogates (e.g., population density, road networks).

Second, we did not use pollutant dispersion modeling to characterize decreasing pollutant concentrations with increasing source distance. A review by Hoek et al. (2008), however, concluded that LUR can often perform better than or equivalent to other methods such as dispersion modeling alone, which can make too many unrealistic assumptions (Hoek et al., 2008). In this analysis we included dispersion indicators in the LUR piece of our UK models, including proximity and buffered covariates (e.g., distance from roadways, roadway length

within a buffer). We then allowed our models to weigh the covariates most associated with TRAP concentrations most heavily (e.g., closer distances and smaller buffers). Other studies have achieved a similar result by using, for example, variable selection regression methods such as forward stepwise selection (Hoek et al., 2008).

Third, we collected some of our geocovariates years earlier, and as a consequence, some of them may have slightly changed over time. We expected that these changes were small, however, and assumed that geocovariates had maintained the same rank order over time (i.e., the same neighborhoods had the highest values of a given covariate).

Future Directions

Future analyses using these data may want to develop prediction models that incorporate the mobile (non-stationary) monitoring data. The mobile data collected immediately before and after taking a stationary sample could be used to increase the sampling "duration" and possibly stabilize site estimates. Moreover, road segments with repeat visits could be used to increase the number of "sites" used to build prediction models.

If this study were to be repeated, a more extensive monitoring campaign could ensure that sampling occurred at all sites during all times of the day, something that was infeasible during our campaign. Our work investigating mobile monitoring design, however, indicates that the reduced sampling hours at some locations likely had a minimal impact on the accuracy of the annual average estimates at those locations (Blanco et al., 2019). Pairing this campaign with continuous, fixed-site sampling of UFPs could further support model validation efforts and be especially important since regulatory, continuous sampling of this pollutant does not currently exist. Notably, however, having only a few continuous sampling sites could still mischaracterize the temporal trends at some sites, and would be costly to maintain with the existing sensor technologies.

Future models would benefit from the inclusion of additional built environment covariates such as distance to plane landing paths, apartment height and bus routes. High density areas such as downtown Seattle could benefit the most.

The particle size distribution and source of UFPs is a scientifically interesting question that should be considered. Different sources have been shown to emit UFPs of specific size

ranges and chemical composition (Austin et al., 2019; Riley, Gould, et al., 2016; Ye et al., 2020), which may be associated with varying adverse health effects. UFPs near the Sea-Tac Airport, for example, have been shown to predominantly be in the smaller size ranges (Austin et al., 2019).

Conclusion

In this study, we described how increasingly popular mobile monitoring campaigns can be used to specifically assess long-term exposure for epidemiologic investigations. This study was one of the first to use short-term stationary measurements from mobile monitoring to estimate long-term averages at locations representative of a study population. We believe that these estimates may be more accurate and precise than more common, drive-by only campaigns that average mobile observations over large areas (e.g., street blocks) that may or may not be geographically similar to the study population of interest. Differences between these monitoring designs may be particularly important for pollutants that quickly decay with increasing distance from their sources, such as UFPs and BC. Next steps include understanding the impact of this improved exposure assessment on epidemiologic inferences.

We were also able to characterize annual average UFP and BC concentrations with high spatial precision for the greater Seattle area using our dense mobile monitoring network. The long-term spatial variability of these pollutants has historically been poorly understood, particularly for UFPs. Our models identified the downtown, International and Industrial districts and airport areas as having elevated annual average UFP and BC concentrations. Many of these areas are residential and business areas where people may be exposed for long periods of time.

Our models performed very well in out-of-sample validation, and our sensitivity analyses indicated that our results were robust. These findings strengthen our results and support the use of these predictions for future epidemiologic investigations of long-term TRAP exposure, particularly for the ACT cohort.

Aim 3: Characterizing the long-term temporal trend of ultrafine particulates and black carbon concentrations in the greater Seattle Area

Overview

Aim 3 builds on our Aim 2 work, which focused on predicting 2019 annual-average UFP and BC exposure levels for the ACT cohort based on a mobile monitoring campaign specifically designed for epidemiologic application (MESA Air, 2019). We use previously documented methods from the literature (Bellander et al., 2001; Hystad et al., 2012; S. Y. Kim et al., 2017; Levy et al., 2015; Meng et al., 2019; Mölter et al., 2010b; Montagne et al., 2015; R. Wang et al., 2013) to extrapolate the UFP and BC exposure surfaces back in time in order to predict long-term exposure for the ACT cohort.

First, we incorporate spatially- and temporally- varying TRAP emission indicators into our UK models, including highway vehicle emissions, population density and Normalized Difference Vegetation Index (NDVI). Note that population density and NDVI were used in Aim 2, but these were fixed for one time period rather than being time-varying. Highway emissions are important sources of TRAP, while population density and NDVI act as surrogates for vehicle emissions in areas where these data are otherwise unavailable (e.g., suburban areas away from major roadways). We use these models to predict past TRAP concentrations using surrogates for historical vehicle emissions (e.g., population density in 2,000).

Second, we apply a temporal trend adjustment factor developed from historical elemental carbon (EC) observations to further account for temporal trends in TRAP not otherwise captured by our time-varying covariates, for example neighborhood traffic emissions.

While long-term UFP exposure is our primary scientific question of interest, validating this model is challenging due to the extremely limited historical UFP data. This work thus focusses on BC, a correlated TRAP for which more historical AQS site data are available to validate our models. We use the findings from this work to afterwards support our work on UFPs.

Methods

Geocovariates

We collected new space- and time-varying vehicle emission indicator geocovariates for our UK models based on the following data requirements. First, geocovariates needed to vary spatially at fine enough resolutions (e.g., at the block level rather than the county level). As in Aim 2, this spatial contrast was used to build 2019 TRAP prediction models. Second, geocovariates needed to be available sufficiently back in time, starting around 1995, and thereafter collected regularly through the present. This allowed us to use historical covariate values in our models to extrapolate the 2019 exposure surface back in time. Finally, geocovariates needed to have similar collection approaches over time. We did not want to use data that had been collected differently over time since this could result in varying model prediction performances and exposure assessments of varying quality over time, which would have been especially problematic for our longitudinal exposure assessment. For example, if vehicle counts were collected at more or different locations over time. The vehicle emission indicators that met our data requirements included major roadway EC emissions, population density and NDVI (Table 13).

Table 13. New Covariates to be Collected for Aim 3

Covariate	**Temporal Resolution**	**Spatial Resolution**	**Rationale**
Distance-weighted AADT per 100 m highway road link	~5-year averages between 1995-2018	100 m road links	Used to calculate emissions Calculated from AADT highway point counters for 1990-2018[a]

Vehicle elemental carbon emission factors: g/vehicle-mile[b]	~5-year averages between 1995-2019	County-level	Used to calculate emissions Accounts for features that have historical impacted vehicle emissions (e.g., vehicle age distribution, fleet distribution, fuel composition, temperature, relative humidity, road speeds)
Total grams of EC emissions within eight buffers (50, 100, 150, 300, 400, 500, 750, 1000 m radii)	~5-year estimates between 1995-2019; linear interpolation/extrapolation for other years	100 m road links	Total emissions mass has greatly changed over time
Population density[c] within eight buffers (0.5, 1, 2, 2.5, 3, 5, 10, 15 km radii)	1990, 2000, 2010; linear interpolation for other years; using 2011-2018 ACS survey results for interpolation post 2010	US Census Block Group	Positively associated with TRAP
NDVI[d] within 35 35 buffer-quantile-season combinations (5 quantile seasons: Q25 all seasons, Q50 all seasons, Q75 all seasons, Q50 summer, Q50	1990-1993, 2010, 2019; linear interpolation for other years	12 m	"Greenness" indicator Negatively associated with

winter; each for 7 buffers: 0.25, 0.50, 1, 2.5, 5, 7.5, 10 km radii)			TRAP and land development

Sources: [a]WSDOT; [b]US EPA; [c]US Census Bureau, decennial censuses and annual ACS surveys; [d]Landsat 5 (1990-1993, 2010) and Landsat 8 (2019)

Emissions

To estimate BC emissions (grams) on major roads, we collected annual average daily traffic (AADT) count information (vehicle-miles) from the Washington Department of Transportation (WSDOT) and estimated elemental carbon (EC) vehicle emission factors (grams/vehicle-mile) from the US EPA's MOtor Vehicle Emission Simulator (MOVES) system (US EPA, 2018c; WSDOT, 2019).

AADT estimates at highway point locations were for total AADT on all lanes, including high-occupancy vehicle (HOV) lanes, from 1995-2018. We calculated about 5-year average AADT estimates (e.g., 1995-1999, 2015-2018) with the available data for each location since some locations had missing data years. We only included locations for which we had averages during every five-year period, starting in 1995 (i.e., dropped locations without at least one reading every 5 years). For simplicity, we excluded AADT point estimates along ramps. Ramp AADT estimates were generally lower and had a different density distribution than AADT estimates along main highways. In addition, their inclusion would have entailed appropriately adding and subtracting these from highway counts based on their relatively down- or up-stream location.

Next, we used ArcMap's (version 10.5.1) Network Analysis tools (Esri, 2019) and RStudio (version 1.2.5033, using R version 3.6.2) (RStudio., 2019) to estimate distance weighted AADT along every 100 m of the highway road network based on these AADT estimates at point locations (Figure 30). The two nearest AADT point estimates (based on highway driving distance) were used to calculate a weighted average for each 100 m road segment. For quality assurance purposes, we compared these estimates to estimates from the WADOT, which were available for 2005-2018. To smooth AADT estimates, we grouped 100 m road links into mutually exclusive groups of 5 based on location and assigned each road link its group median.

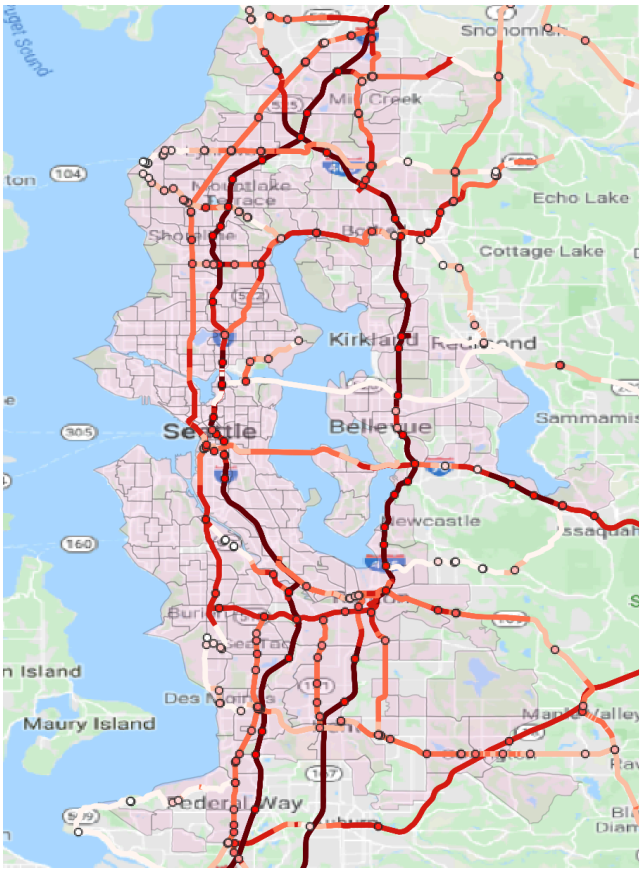


Figure 30. Example of average AADT estimates (2015-2018) on 100 m highway road links based on the two nearest counters, prior to smoothing. Dots = highway vehicle point counters; lines = 100 m road links; pink area = mobile monitoring area (for display purposes – AADT was calculated for the entire study area of interest). Darker colors indicate higher AADT estimates.

We used MOVES (version 2014b) to estimate EC emission factors (g/vehicle-mile) for King County (where the majority of our study area was situated) from 1990-2019. BC emission factors were unavailable, though these have been shown to be highly correlated with EC since they have similar sources but are measured using different methods (Janssen et al., 2017; Salako et al., 2012). Emission factors were estimated for restricted roadways (e.g., included highways but not smaller, neighborhood roads) from gasoline and diesel fuels. Other fuels were not included since gasoline and diesel fuels compromise the majority of fuels used, and data on other, more recent fuels may have less accurate and less consistent information over time. We used the default county settings for other parameters, as has been done by past work (Snyder et al., 2014), since not all of the inputs were available dating back to about 1995. Estimated

emission factors for 1991-1998 were not available through MOVES, and so they were calculated from a linear interpolation of 1990 and 1999 emission factors. We estimated about five-year average emission factors from 1995-2019 and converted these from g/vehicle-mile to g/vehicle-meter for consistency with our five-year average AADT estimates.

We calculated EC emissions (g) by multiplying 100 m AADT estimates (vehicle-meters) by MOVES emission factors (g/vehicle-meters).

We used PostgresSQL (PSQL, version 12.1) and Python (version 2.7.13) to estimate total emission about every five years between 1995-2019 within various buffer sizes from 50 m to 1k m radii around each point location of interest, including mobile monitoring locations (for model building), ACT cohort locations (in order to generate predictions), and a grid (for visualizing and mapping our results).

Given the large number of covariate-buffer combinations for each location and year, we used PLS regression to estimate a single linear combination of our covariates, using the 2019 mobile monitoring observations of BC as the outcome and emissions variables for 2019 as the predictors used to build the model, such that:

$$\hat{Z}_i^m = \sum_{j=1}^{p} \hat{\phi}_j^m X_{i,j} \quad (3.1)$$

where:

- $\hat{Z}$ are the scores estimated from PLS regression, which uses the 2019 mobile monitoring observations of BC as the outcome, or independent variable
- m is the PLS component number, where m=1 for space-time-varying geocovariates (or m = {1,2,3} for purely space-varying covariates, as described below)
- i is an index defined as s, t for space-time-varying covariates (or defined as s for purely space-varying covariates, as described below)
- p are our original emissions geocovariates (or population density, NDVI or purely space-varying covariates, as described below)
- X are the standardized (mean centered and divided by the standard deviation) geocovariate values
- $\hat{\phi}$ are the estimated loadings from PLS regression

We used this model to generate a single linear combination of our geocovariates (scores) from the first PLS component for the mobile monitoring locations (for model building later on), ACT cohort locations (prediction locations) and a grid (for visualization/mapping) for every year with available emissions estimates (about every five years). We used linear interpolation to estimate PLS scores for other years.

Population Density

Our data team collected decennial population density from the US Census Bureau for 1990, 2000 and 2010 at the block group level (US Census Bureau, 2019a). We estimated population density within various buffers ranging from 500 m to 10 km around each location of interest following the standard operating procedures used by MESA Air (MESA Air, 2019).

Following the same PLS regression approach described above for emissions, we fit a model to the 2019 mobile monitoring observations based on the 2010 population density buffered covariates and used this model to estimate PLS scores for all years with available data (1990, 2000, 2010). Given the trend in population growth in the area (US Census Bureau, 2019b), we linearly interpolated population density PLS scores for 1991-1999 using a 1990 and 2000 linear fit, and for 2001-2009 using a 2000 and 2010 linear fit. Since the 2020 Decennial US Census was unavailable, in order to estimate population density after 2010, we used the annual US Census ACS survey to estimate the average annual population rate of change between 2010 and 2018 (~1.8% increase) for the Seattle-Tacoma-Bellevue metropolitan area (US Census Bureau, 2020). This annual rate of change was used to adjust 2010 US Census PLS score estimates in order to calculate population density PLS scores between 2011 and 2019. Unlike earlier interpolations using decennial Census data at the individual level, this approach of interpolation was at the annual level and thus the same for all locations.

NDVI

Following the methods described by MESA Air, we gathered satellite images of NDVI collected throughout 1990-1993, 2010 and 2019 from the United States Geological Service (USGS, via Google Earth Engine) (Google, 2019; MESA Air, 2019). The images came from

Landsat 5 (1990-1993, 2010) and Landsat 8 (2019). Years 1990-1993 were grouped since each of these years was incomplete due to cloud coverage. We used PSQL and Python to calculate annual NDVI quantiles within various buffers ranging from 250 m to 10 km radii for each location of interest.

Again, following a similar PLS regression approach as described above for emission estimates, we fit a model to the 2019 mobile monitoring observations based on 2019 NDVI covariates and used this model to estimate PLS scores for all years with available data (1990-1993, 2000 and 2019). We linearly interpolated NDVI PLS scores for other years using available data from the surrounding years. Specifically, we interpolated 1994-2009 using a 1990-1993 and 2010 fit, and 2011-2018 using a 2010 and 2019 fit.

Space Covariates

We followed a similar PLS regression approach as described above to generate three linear combinations of the remainder of our purely space-varying covariates (e.g., airport and road proximity; Table 1), fitting a model to the 2019 mobile monitoring observations based on all of our purely space-varying covariates. The first three PLS components were chosen since this was the number selected via cross-validation for our purely spatial model in Aim 2.

Temporal Trend Adjustment

In addition to gathering new spatiotemporally-varying geocovariates, we characterized the temporal trend of TRAP for the period of interest: 1995-2019. The rationale was that our spatiotemporally-varying geocovariates may not fully capture the historically large reductions in TRAP, and that this could improve our UK model predictions. Specifically, we gathered historical observations of EC at the Beacon Hill AQS site, which was the only site with long-term TRAP data within our study area (Figure 31). This site is part of the EPA's Chemical Speciation Network and has collected daily EC averages about every three days since 1996 using an EC $PM_{2.5}$ LC TOR method (US EPA, 2019b).

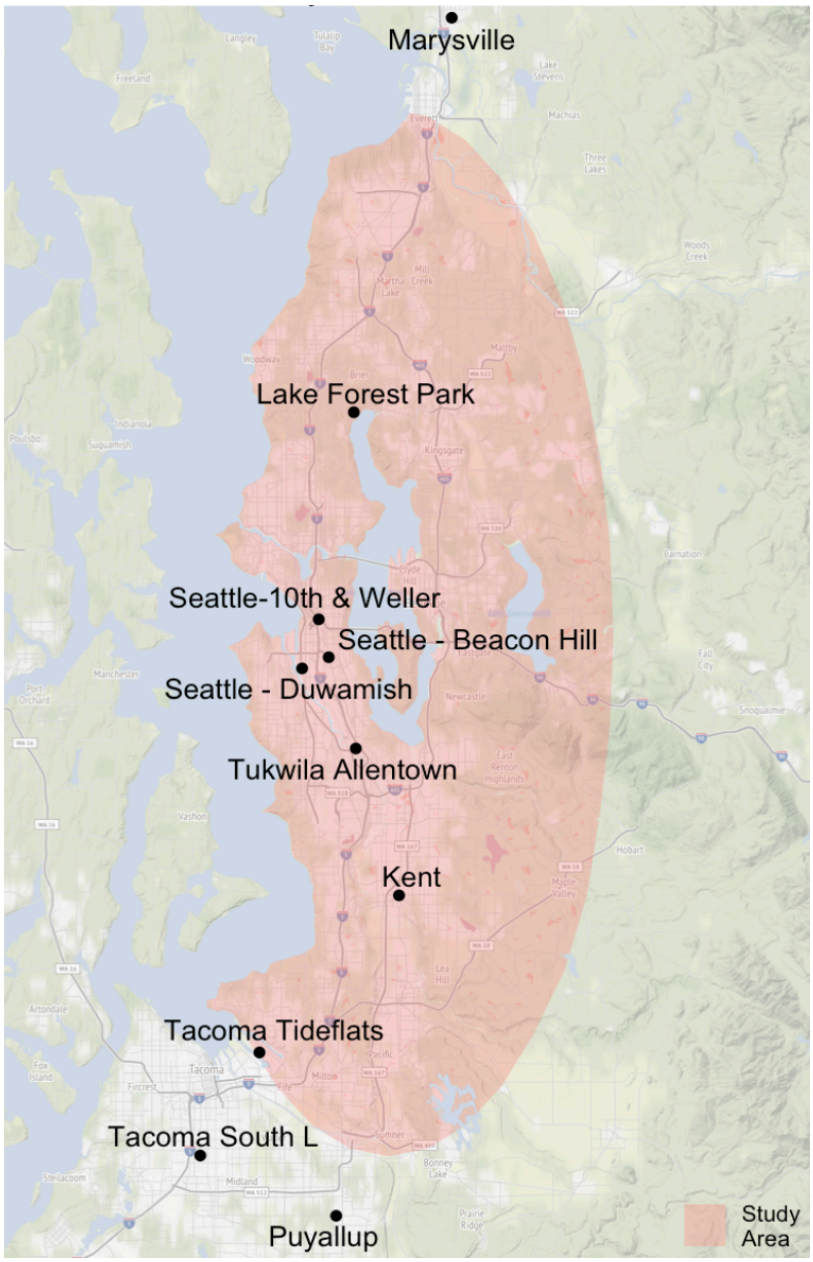


Figure 31. AQS sites within and near our study area with historical BC and/or EC measurements.

We followed data completeness criteria similar to those developed by MESA Air to address missingness and ensure that the estimated annual averages from these data were unbiased (MESA Air, 2019). Specifically, each year needed to have: a) at least 82/122 (67%) sampling days, b) a maximum of 45 days between measurements to ensure adequate coverage throughout the year, and c) a similar number of samples collected during heating (October - March) and

non-heating (April - September) months. We calculated annual averages with the years that met these criteria.

To smooth annual averages and address missing years, we fit a natural cubic spline (3 degrees of freedom) model ($\hat{f}(t)$) to these data with knots at 2005 and 2012 and boundary knots at 1997 and 2018 (the first and last years with available data). We used this model to predict annual average EC concentrations (EC_t), extrapolating it to 1995 and 2019:

$$\widehat{EC}_t = \hat{f}(t). \quad (3.2)$$

Next, we estimated trend adjustment factors ($\hat{\alpha}_t^{trend}$), defined as the ratio of the predicted EC concentration for any given study year t ($\widehat{EC}_t$) and the predicted EC concentration for 2019, the reference year ($\widehat{EC}_{t=2019}$; Figure 82):

$$\hat{\alpha}_t^{trend} = \widehat{EC}_t / \widehat{EC}_{t=2019} \quad (3.3)$$

UK Model with Partial Least Squares and a Trend Adjustment

We built a UK model for BC with geocovariate PLS scores and adjusted these predictions using a long-term temporal trend. Mobile monitoring locations were used to build the model after excluding AQS collocation sites (n = 304 remaining sites) since these were later used to validate the model.

Specifically, we hypothesized that we could predict TRAP concentrations at a particular location and time given that location's geocovariates (some of which could vary over time) and a temporal trend that accounted for generally decreasing TRAP concentrations over time, with some prediction error due to some unknowns.

Our assumed underlying model was:

$$Conc_{s,t} = exp\left(\theta_0 + \sum_{f=1}^{3} \theta^f Z_s^f + \theta^{pop} Z_{s,t}^{pop} + \theta^{NDVI} Z_{s,t}^{NDVI} + \theta^{EC} Z_{s,t}^{EC} + \epsilon_s\right)\alpha_t^{trend} + \epsilon_t \quad (3.4)$$

where:

- $Conc_{s,t}$ denotes the annual-average BC (ng/m^3) concentration for a particular location s within our study area and year t, where $t = 1995 - 2019$.
- Z are PLS scores (dimension-reduced, linear combinations of geocovariate predictors) for locations s. They are fixed over time (Z_s) or are for a specific year t ($Z_{s,t}$). Specifically:
  - f denotes the purely space-varying geocovariates that are fixed (f) in time
  - pop denotes population density geocovariates for 8 buffers ranging from 500 m – 15 km
  - $NDVI$ denotes NDVI geocovariates for 35 buffer-quantile-season combinations (5 quantile seasons: Q25 all seasons, Q50 all seasons, Q75 all seasons, Q50 summer, Q50 winter; each for 7 buffers ranging from 250 m – 10 km)
  - EC denotes elemental carbon emissions geocovariates for 8 buffers ranging from 50 m – 1 km
- θ are the model coefficients
- α_t^{trend} is a temporal trend describing decreasing TRAP concentrations over time not otherwise captured by the time-varying PLS scores
- ϵ_s is the residual term that varies over space with mean zero and a geostatistical structure assumed to be an exponential function with range ϕ_s, partial sill σ_s^2 and nugget γ_s^2 ($\epsilon_s \sim N(0, \sum(\phi_s, \sigma_s^2, \gamma_s^2))$).
- ϵ_t is the residual term for a given year

Given our data from the mobile monitoring campaign, we hypothesized the underlying model for the 2019 exposure surface ($t = 2019$) to be the following:

$$Ln\left(Conc_{s,t=2019}\right) = \theta_0 + \sum_{f=1}^{3} \theta^f Z_s^f + \theta^{pop} Z_{s,t=2019}^{pop} + \theta^{NDVI} Z_{s,t=2019}^{NDVI} + \theta^{EC} Z_{s,t=2019}^{EC} + \epsilon_s \quad (3.5)$$

where:

- $Ln(Conc_{s,t=2019})$ is the log-transformed annual-average BC (ng/m^3) concentration for 2019

- α_t^{trend} and ϵ_t are omitted in this model since the model is defined for a single, reference year

We estimated the modeling parameters for this UK model. Model predictions are given by both regression coefficients and kriging of the residuals:

$$Ln(\widehat{Conc_{s,t=2019}}) = \hat{\theta}_0 + \sum_{f=1}^{3} \hat{\theta}^f \hat{Z}_s^f + \hat{\theta}^{pop} \hat{Z}_{s,t=2019}^{pop} + \hat{\theta}^{NDVI} \hat{Z}_{s,t=2019}^{NDVI} + \hat{\theta}^{EC} \hat{Z}_{s,t=2019}^{EC} + \hat{\epsilon}_s \quad (3.6)$$

where:

- $(Ln(\widehat{Conc_{s,t=2019}}))$ is the predicted log-transformed annual-average BC (ng/m^3) concentration for 2019
- $\hat{Z}$ scores are estimated from PLS regression used to predict $Ln(\widehat{Conc_{s,t=2019}})$ from geocovariate predictors (see Equation 3.1 for details)
- $\hat{\theta}$ are the estimated model coefficients derived from measured TRAP concentrations and PLS scores from our 2019 mobile monitoring campaign
- $\hat{\epsilon}_s$ is the kriging contribution to the overall model prediction. It is a linear combination of the model residuals that is are weighted based on distance by the estimated covariance matrices for the monitored and unmonitored locations, and between monitored locations, such that residuals at monitoring locations closer to prediction sites are given a higher weight. These parameters are estimated for $t = 2019$, the mobile monitoring campaign year

We estimated TRAP concentrations from 1995-2019 at ACT cohort locations within our study area using a two-step approach. First, the fitted model was used to estimate Ln TRAP concentrations at new locations and for specific times based on space- and space-time- varying geocovariates. Of note, an important feature of this fitted model is its reliance on the PLS

regression loadings and the estimated kriging contribution from our 2019 mobile monitoring campaign. Second, we exponentiated and temporally adjusted these values:

$$\widehat{Conc}_{s,t} = exp\left(Ln(\widehat{Conc}_{s,t=2019})\right)\hat{\alpha}_t^{trend} \quad (3.7)$$

where:

- $(\widehat{Conc}_{s,t})$ is the predicted annual-average BC (ng/m^3) concentration for a given location within our study area and study period
- $\hat{\alpha}_t^{trend}$ is the predicted temporal trend

We used this model to predict BC exposure for ACT cohort locations over time from this model. For visualization purposes, we plotted predictions on a grid.

Validation

Validation Dataset

We validated model predictions against historical BC observations from AQS sites within and near our study area (Figure 31). Sites represented a mixture of suburban, residential, commercial, urban center and industrial areas (Appendix C Figure 76).

Daily average BC readings from aethalometer readings at an 880 nm wavelength were collected using the PSCAA's Air Graphing Tool (PSCAA, 2020). To account for discrepancies in the concentration sensitivities of the older (Magee Aethalomter Model AE33) and newer (Magee Scientific TAPI M633) version of the same instrument, we followed the PSCAA's guidance and increased older readings by 32%.

Similar to our temporal trend adjustment approach, we followed data completeness criteria to address missingness and ensure that the estimated annual averages from these data were unbiased. Specifically, each site-year combination needed to have: a) at least 244/365 sampling days (67% of the maximum available data), and b) a maximum of 45 days between measurements to ensure adequate coverage throughout the year. In addition, we dropped 10th &

Weller's 2015 average since it appeared to greatly deviate from that site's trend. This year contained more zero readings than other years and was also excluded from the EPA's data files. Moreover, the Duwamish Valley site was also excluded since its temporal trend appeared to be different than the other AQS sites. This site has historically had different air pollution sources that are unique to the area, including a substantial amount of heavy-duty diesel trucks, idling traffic, boat emissions and heavy train traffic. We calculated annual averages for the site-years that met these prerequisites.

Model Validation

We evaluated the performance of our UK model against these historical observations of BC from AQS sites starting in 2003. We calculated the out-of-sample RMSE and MSE-based R^2 overall (all years and all AQS sites), by year, and by AQS site. The purpose of these stratified analyses was to ensure that: 1) model predictions were not increasingly biased back in time, and 2) the model bias was not different for different locations. This was important to check because varying model performances for different times or places could later result in poor predictions for some years and/or locations for potentially unknown reasons.

Secondary Analysis of UFPs

Following a similar methodology as that described for BC, we built long-term UFP exposure prediction models as a secondary analysis.

We used EC emissions as a surrogate for UFP emissions since MOVES does not produce emission factors for UFP, and these are highly correlated (Hagler et al., 2009; Kerckhoffs et al., 2016; Westerdahl et al., 2005).

We were able to validate this model to a small degree using UFP observations from a measurement study conducted by Kim et al. (2004) at Beacon Hill during 2001 (E. Kim et al., 2004). In this study, hourly UFP volume concentrations ($\mu m^3/cm^3$ air) were collected for various bin sizes during most hours and months of the year. To make these comparable to count concentrations from the P-TRAKs used in our mobile monitoring study, we converted each volume concentrations to a range of particle number concentrations (PNC; pt/cm^3) by dividing

volume concentration by the minimum, median and maximum volume of a particle in that bin (i.e., that bin's minimum, median and maximum particle diameter (d) was converted to a spherical volume (V; μm^3/pt) using $V = \frac{4}{3}\pi\left(\frac{d}{2}\right)^3$. We aggregated several bin concentrations to calculate the total PNC for particles 0.02-1.0 μm in diameter. Next, we calculated a weighted annual average using these PNCs. Since no observations were collected during the months of July or December or the midnight hours, we upweighted the neighboring months (June and August, November and January) and hours (11 PM and 1 AM) appropriately.

Sensitivity Analyses

We conducted a sensitivity analysis to assess the robustness of our primary (BC) and secondary (UFP) long-term exposure models using NOx emissions rather than EC emissions in our prediction models. NOx is a reliable measure of tailpipe emissions (mostly comprised of NO) and has a similar concentration decay pattern as UFP and BC (Karner et al., 2010; Saha, Zimmerman, et al., 2019).

Results

Emissions Geocovariate

The average AADT at highway point locations slightly increased over time, going from about 43,303 vehicles from 1995-1999 to 49,441 vehicles from 2015-2018 (Appendix C Figure 77 and Table 32). Five-year average (SD), distance-weighted AADT estimates on 100 m highway road links based on these point locations also generally increased over time, going from an average of 48,107 vehicles between 1995-1999 to 57,074 vehicles between 2015-2018 (Appendix C Figure 78 and Table 33). This pattern was generally consistent at the road segment level, where AADT estimates generally (but not always) increased over time at any particular road segment (Appendix C Figure 79). These distance weighted AADT estimates on 100 m road links were similar to those available from the WSDOT starting in 2005 for road segments of varying lengths (Figure 32), though they slightly underestimated high AADT locations. This

precision and accuracy of our estimates when compared to those available from WSDOT was consistent over time (R_2 = 0.45-0.60, RMSE = 35,967-42,772 AADT). AADT estimates were very similar before and after smoothing for all time periods (R^2 = 0.98, best fit line slope = 0.99; Appendix C Figure 80).

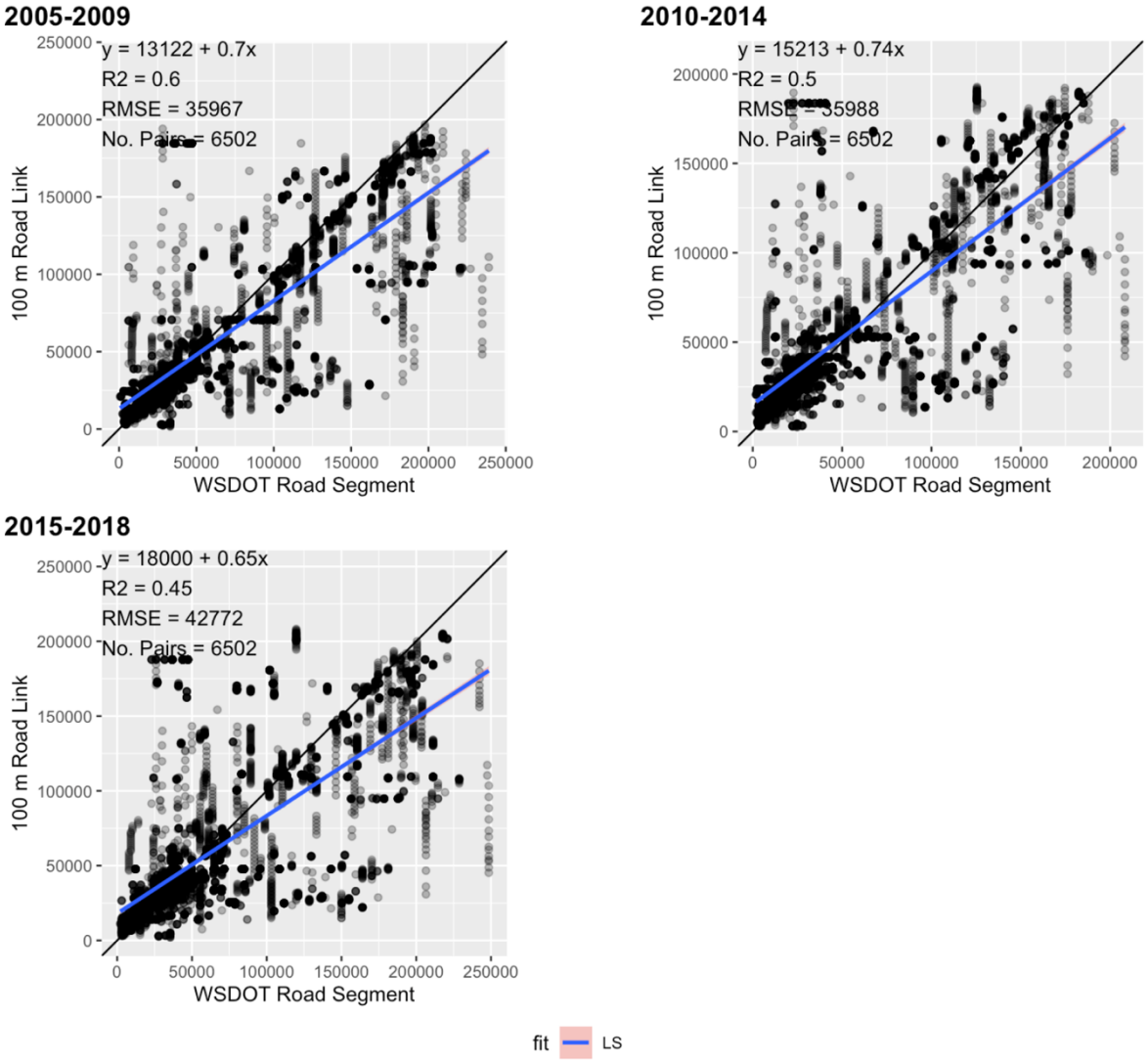


Figure 32. Comparison of the AADT estimates used in this study from 100 m road links and WSDOT-defined road segments. Our 100 m road links are are weighted averages of the two nearest AADT counters. WSDOT AADT estimates on road segments are only available 2005+. LS = least squares fit.

MOVES emission factors for EC ranged from 7e-5 g/meter to 1e-5 g/meter, with lower emission factors over time (Figure 33). Interpolated values from 1991-1999 appeared to be in

line with the temporal trend observed. The five-year average emission factors had a similar pattern to the one-year average emission factors.

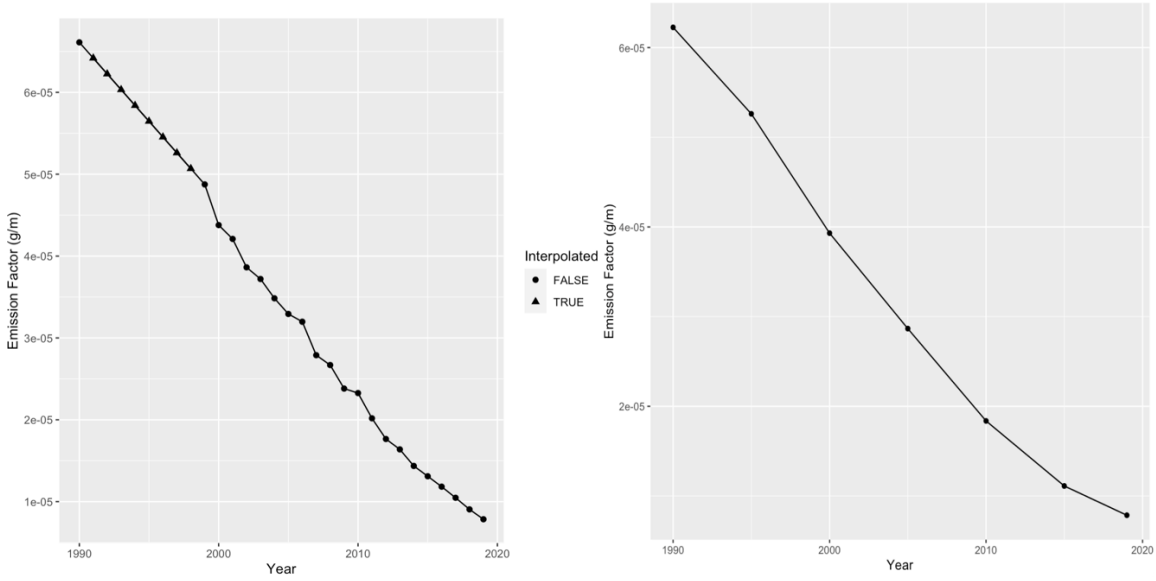


Figure 33. One- (left) and five- (right) year average EC emission factors from MOVES.

Total emission estimates based on AADT and emission factors ranged from about 0-1,025 g per 100 m road link (Figure 34 - Figure 35, Appendix C Table 34). Emissions generally decreased and were less variable for later time periods.

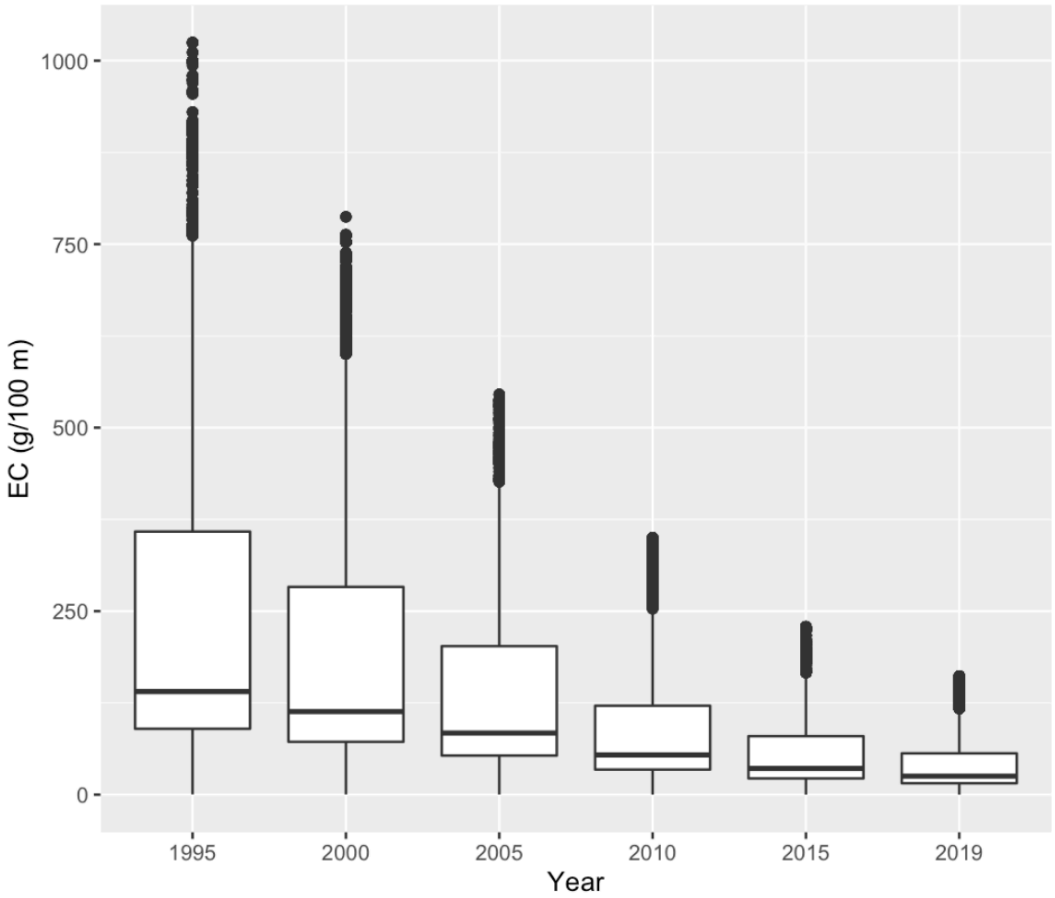


Figure 34. EC emissions over time on restricted roads near our study area

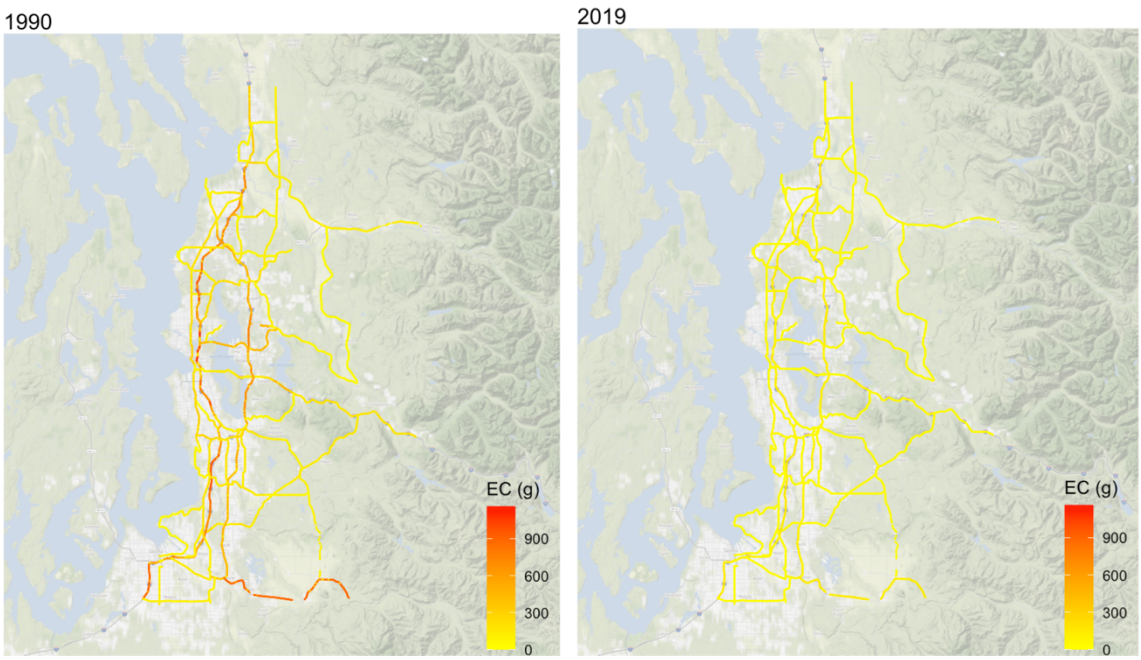


Figure 35. EC emissions on 100 m road links for two example years: 1990 (left) and 2015 (right)

Geocovariate PLS Scores

Figure 36 shows the distribution of PLS scores for individual time-varying covariate models (EC, NDVI and population density [POP]) at mobile monitoring locations (used to build each PLS model) and cohort locations (where predictions were made) for years with available covariate data before the interpolation of scores for other years. The PLS score distributions of EC emissions, population density and NDVI at the monitoring locations at the most recent times were generally similar to the PLS score distributions at cohort locations, and earlier times tended to have higher EC emissions PLS scores, lower population density PLS scores and higher NDVI PLS scores. Earlier years additionally had more variability in EC emissions since there were more very high concentration areas. Note that the final 2019 BC models were built using 2019 EC emissions PLS scores, 2019 NDVI PLS scores, and interpolated 2019 population density PLS scores (not the 2010 scores shown in Figure 36 prior to interpolation) at mobile monitoring locations.

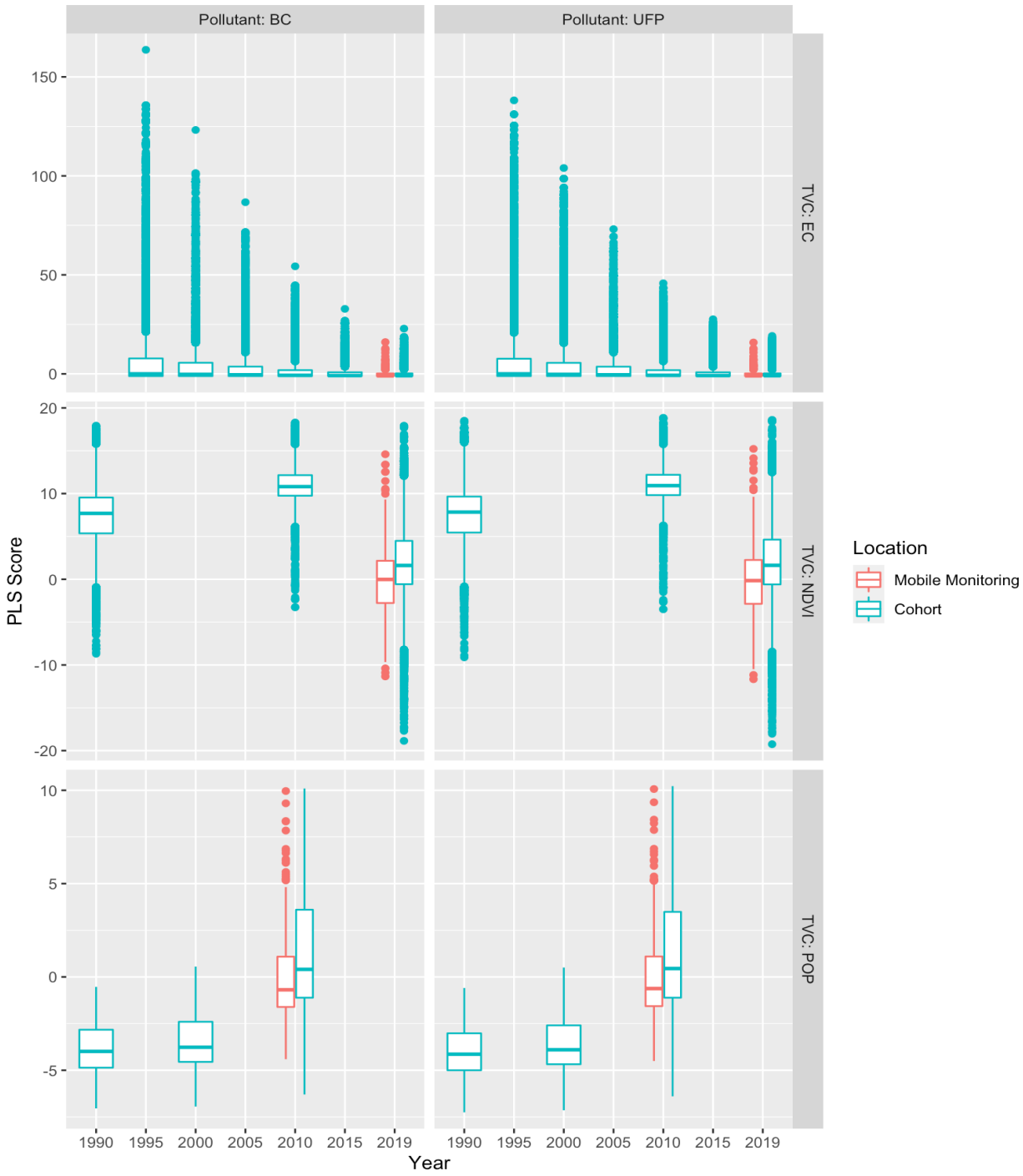


Figure 36. Distribution of PLS scores for individual time-varying covariate (TVC) PLS models (EC, NDVI, population density [pop]) at mobile monitoring locations (used to build the PLS models) and cohort locations (where predictions were made; based on available data years and before interpolation of scores for other years)

Figure 37 show the distribution of PLS scores for cohort locations where predictions were made for both years when data were available to predict PLS scores and for years when PLS scores were linearly interpolated (e.g., between decennial censuses). Overall, EC emissions and NDVI decrease over time, while population density (POP) increases over time. Note that these plots do not include mobile monitoring locations since predictions were not made at these locations back in time. Mobile monitoring locations, however, would also see PLS score changes over time as a result of the interpolation process.

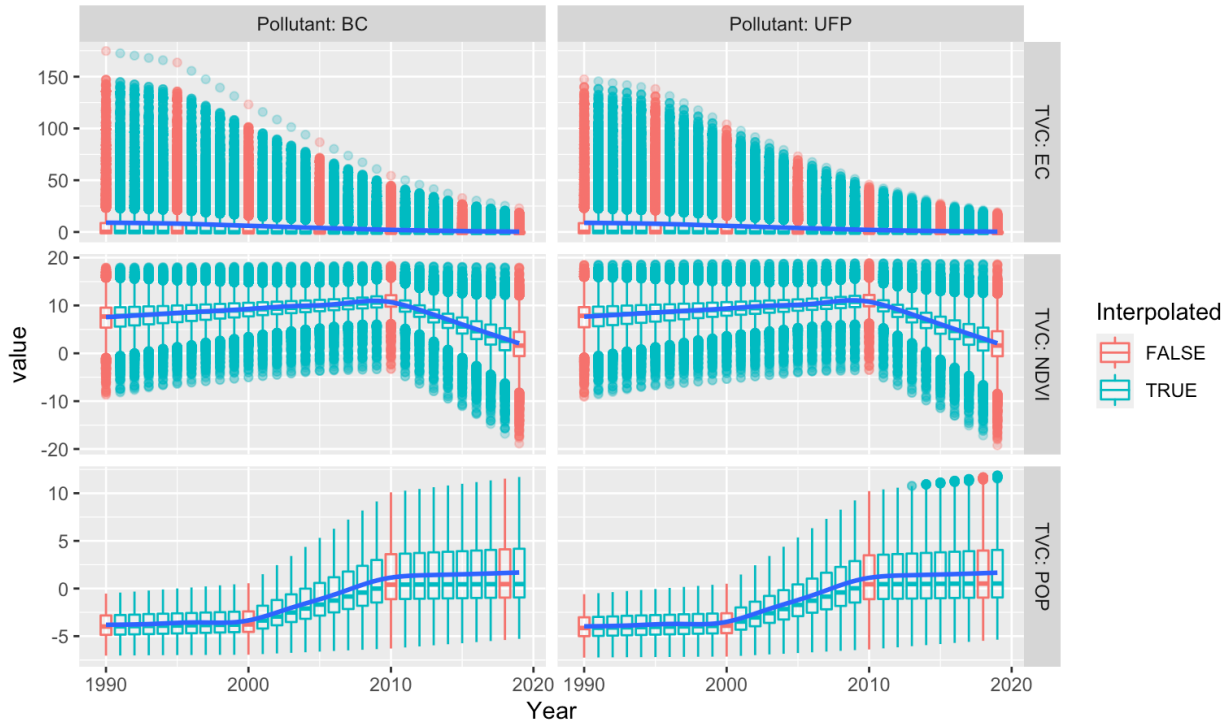


Figure 37. Distribution of PLS scores from each time-varying covariate (TVC) model at cohort locations for the BC and UFP models. Scores are for years with covariate data (no interpolation) and for years that were linearly interpolated. Note, 2018 population density scores are not labeled as being "interpolated" since ACS Survey data between 2010 and 2018 used to inflate 2010 scores in an informative way.

The rank order of the population density PLS scores appeared to remain constant (Figure 38 - Figure 39). The center of Seattle, for example, had the highest population density PLS scores, while the eastern part of the study area had lower PLS scores over time. As for EC emissions, locations that saw the sharpest decreases were those near major highways (Figure 40 -

Figure 41). NDVI had the smallest change near the Seattle-Tacoma, Bellevue and Edmunds area and largest change in the eastern part of the study are (Figure 42 - Figure 43).

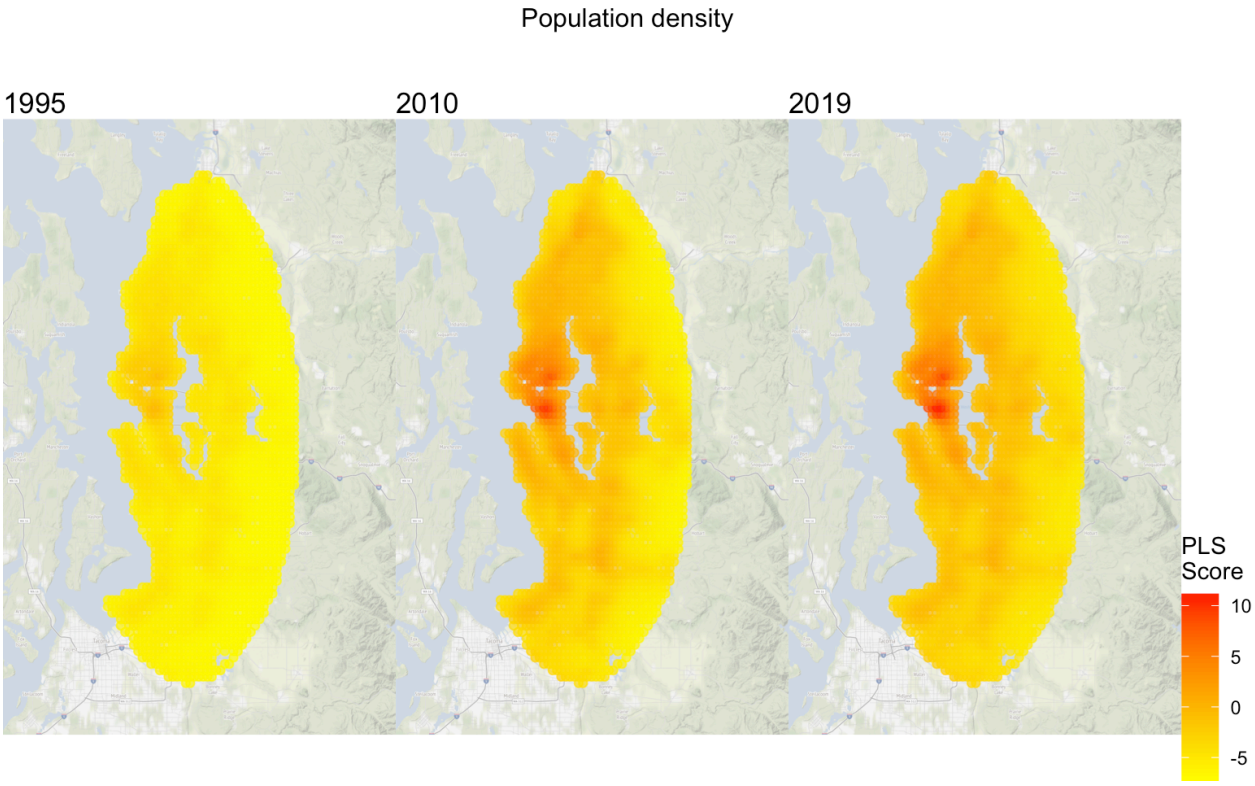


Figure 38. Map of population density PLS scores for various time periods

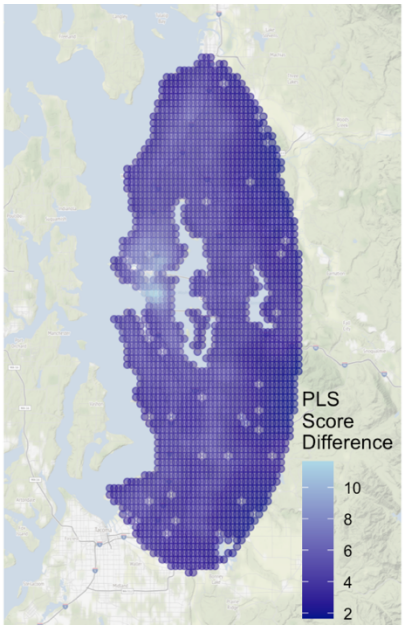


Figure 39. Change in population density PLS score between 1995 and 2019

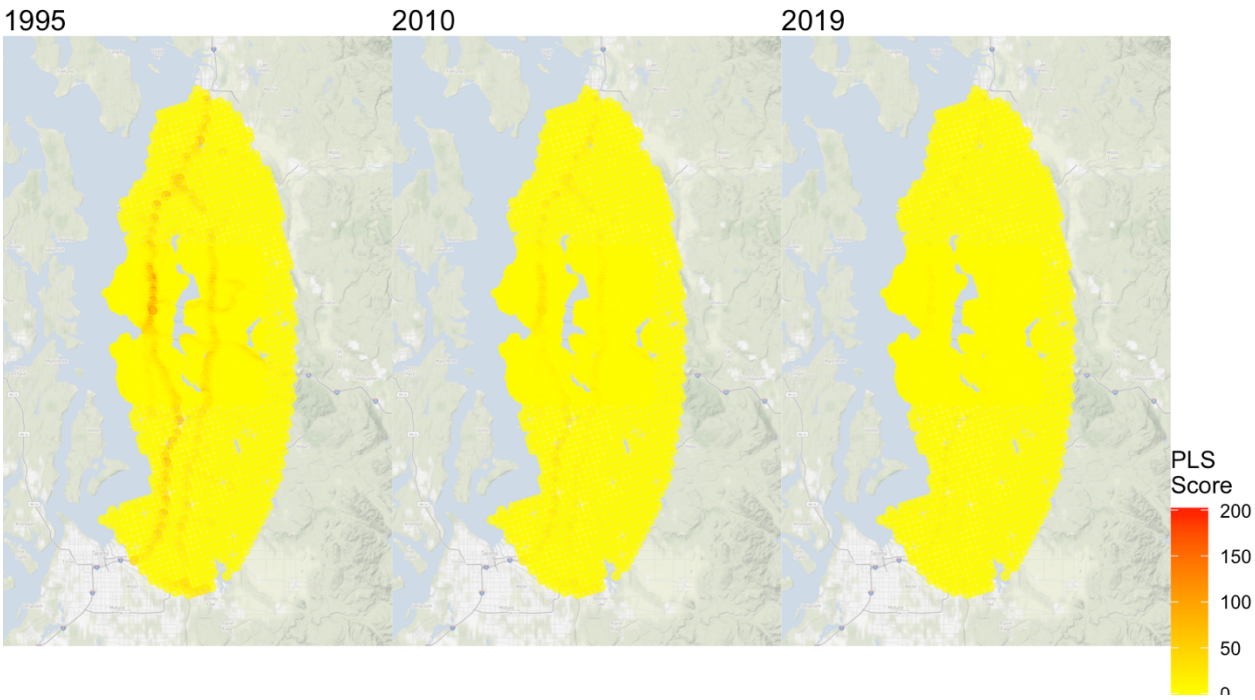


Figure 40. Map of EC emissions PLS scores for various time periods

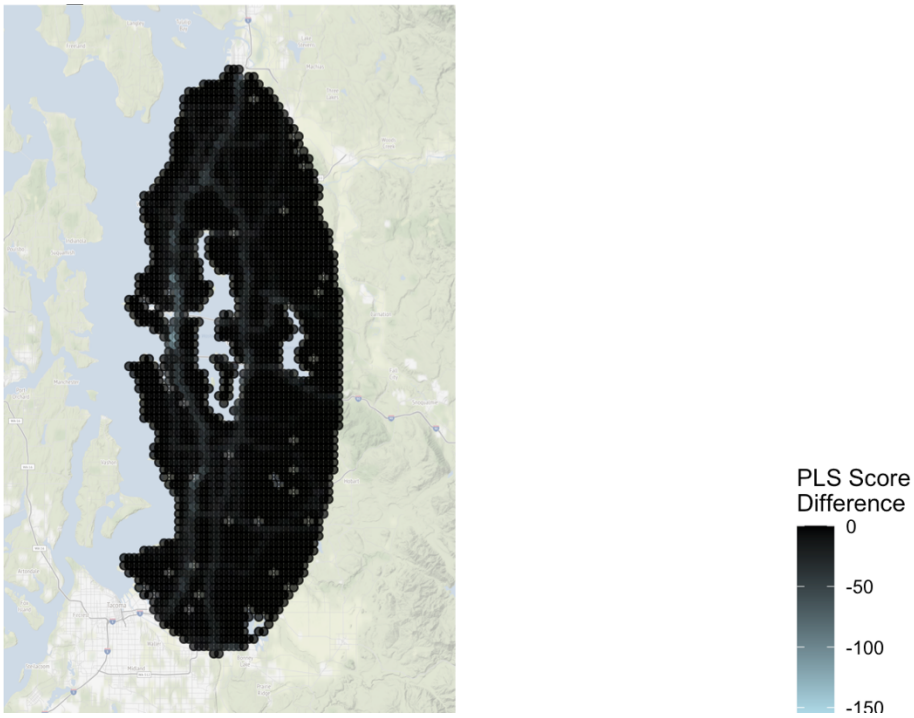


Figure 41. Change in EC PLS score between 1995 and 2019

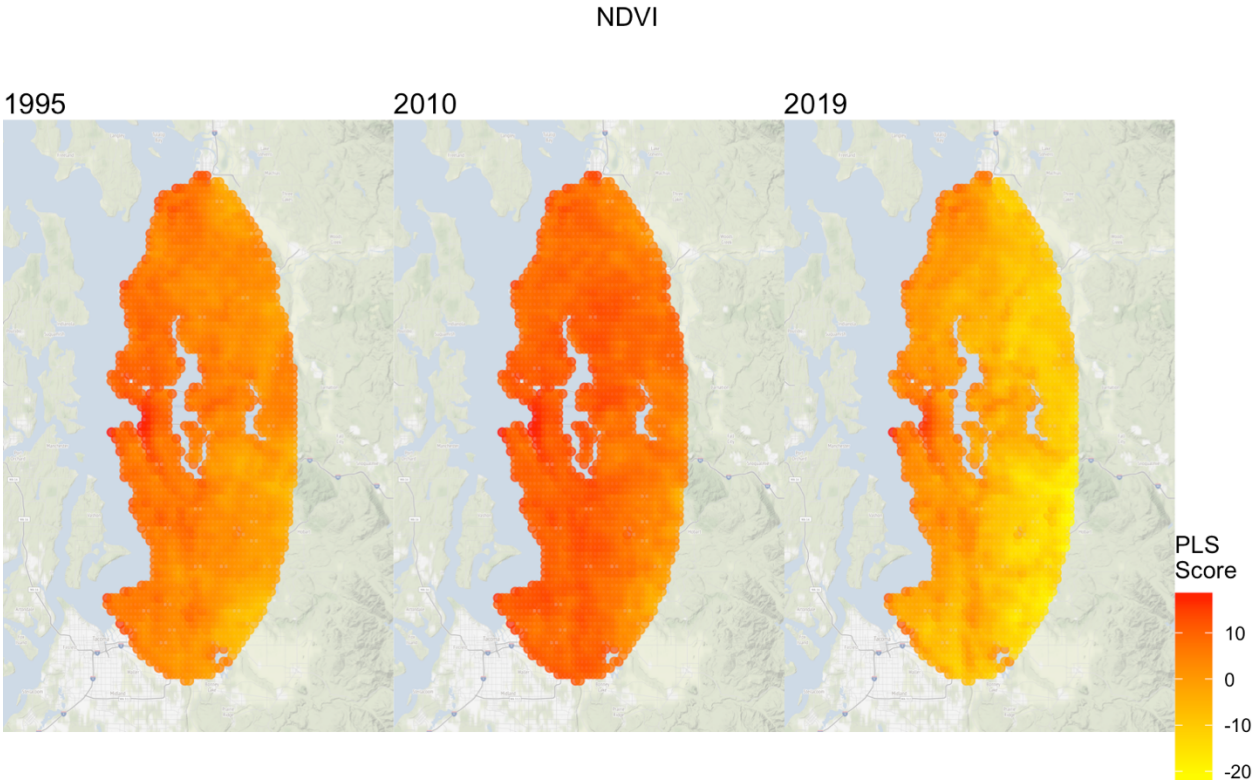


Figure 42. Map of NDVI PLS scores for various time periods

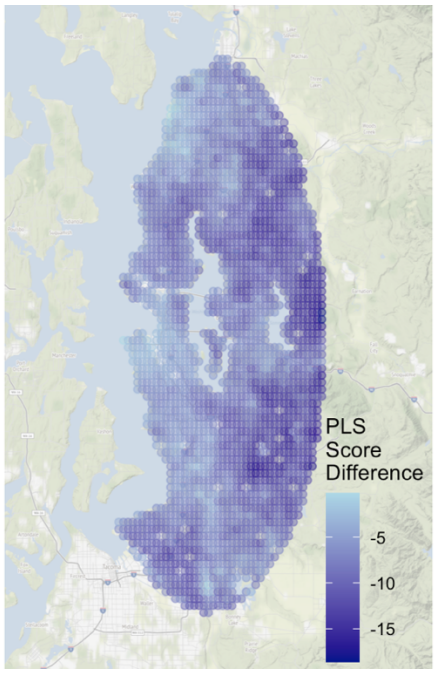


Figure 43. Change in NDVI PLS score between 1995 and 2019

Table 14 shows the change in variability (SD scale) for each pollutant and TVC between 1995 and 2019. Both the UFP and BC pollutant models resulted in similar changes over time. The variability of EC in 2019 dropped to 15% of its value in 1995, the variability of population density increased by 160%, and the variability of NDVI increased by 46% of its value.

Table 14. Change in variability of time-varying covariate PLS scores between 1995 and 2019.

Pollutant Model	TVC	SD 1995	SD 2019	SD Ratio
BC	EC	19.47	2.86	0.15
BC	NDVI	3.05	4.46	1.46
BC	POP	1.33	3.46	2.60
UFP	EC	19.17	2.82	0.15
UFP	NDVI	3.12	4.57	1.46
UFP	POP	1.33	3.43	2.59

Temporal Trend Adjustment

EC observations at Beacon Hill were available between 1996-2018, though four of these years did not meet our inclusion criteria and were dropped (1996, 2000, 2001, 2006; Appendix C Figure 81). The natural cubic spline used to smooth annual averages generally fit the data well and was extrapolated to 1995-1996 and 2019 when no observations were available (Figure 44). The resulting ratio adjustment factors using the spline's predicted EC concentration for 2019 as the reference ranged from >1.0 to ~ 2.3, with increasingly higher adjustment factors for earlier years (Appendix C Figure 82).

Figure 44. Annual average EC at Beacon Hill: Observed and predicted concentrations from a natural cubic spline with knots at 2005 and 2012 and boundary knots at 1997 and 2018.

UK Model Predictions

Predicted BC and UFP exposure concentrations for the ACT cohort were generally higher and more variable during earlier years (Figure 45). BC predictions ranged from 213-5,225 ng/m^3 between 1995 and 2019 (Appendix C Table 35). The median (IQR) BC concentration, dropped from 1,269 (409) ng/m^3 in 1995 to 524 (141) ng/m^3 in 2019. Map predictions generally showed higher BC concentrations near the center of Seattle and along major highways. This spatial contrast faded over time, though predicted concentrations remained relatively higher near major highways (Figure 46).

UFP predictions ranged from 2,245-55,571 pt/cm^3 between 1995 and 2019 (Appendix C Table 36). The median (IQR) UFP concentration, for example dropped from 16,073 (4,939) pt/cm^3 in 1995 to 6,753 (1,822) pt/cm^3 in 2019. Map predictions indicated that UFP

concentrations were generally higher between the center of Seattle and the Sea-Tac Airport, as well as along major highways. Some of this spatial contrast faded over time as concentrations decreased across the entire area, though areas near the center of Seattle and the Sea-Tac Airport remained elevated (Figure 47).

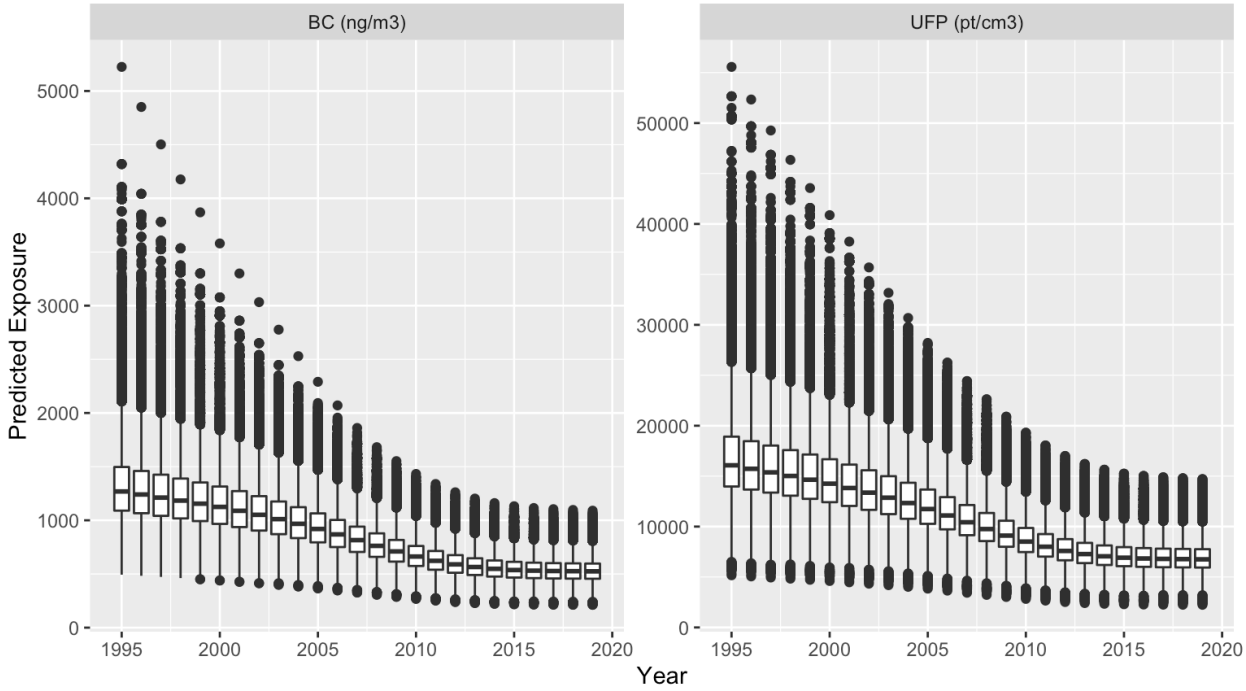


Figure 45. Predicted historical UFP and BC exposure for the cohort

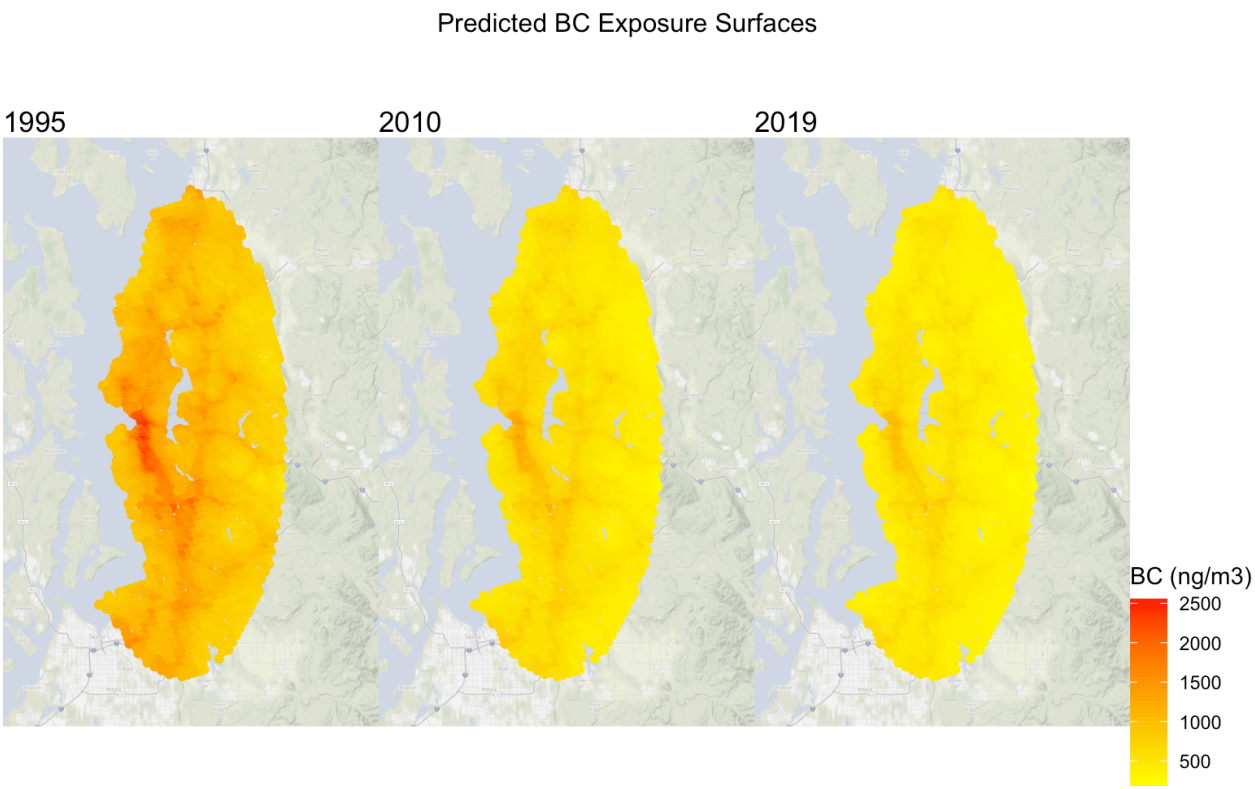


Figure 46. Predicted BC exposure surfaces over time

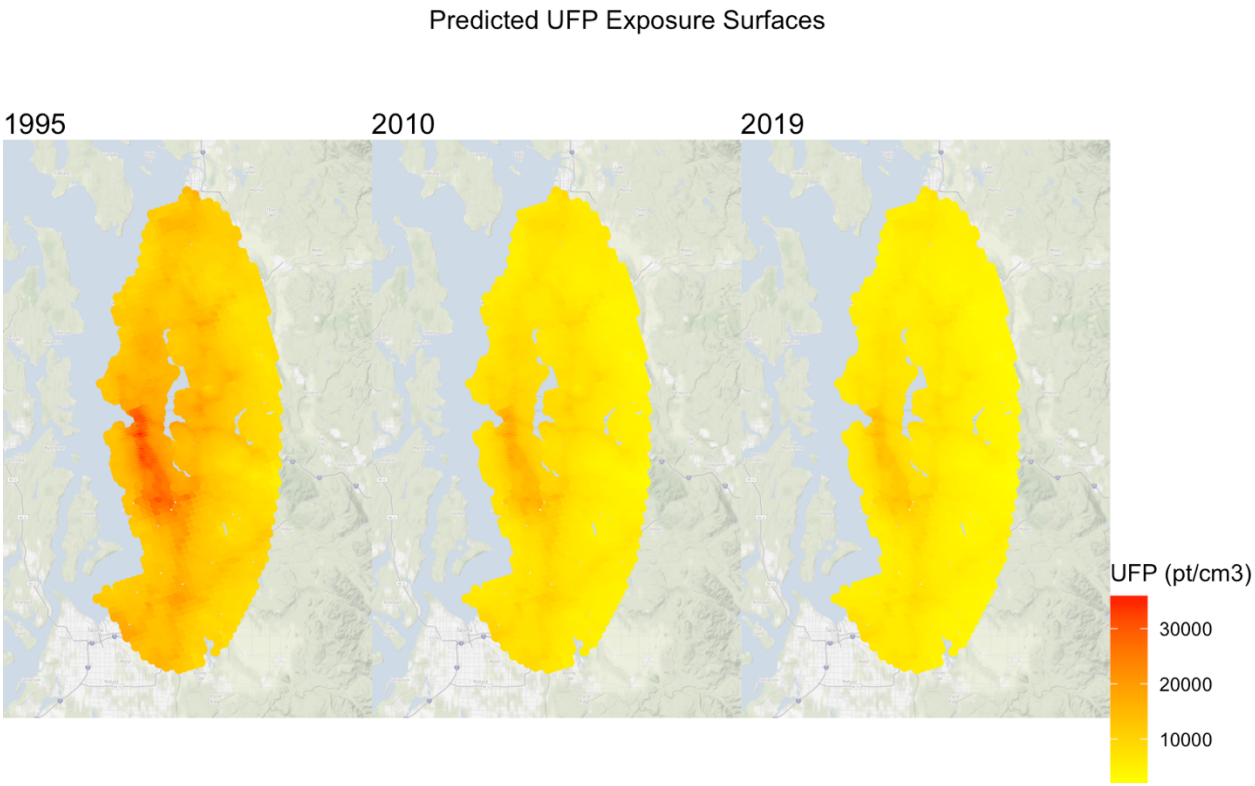


Figure 47. Predicted UFP exposure surfaces over time

Maps looking at how the predicted BC and UFP exposure surfaces changed between 1995 and 2019 indicated that the drop in concentration varied across the study area (Figure 48). UFPs, in particular, saw the largest change between the center of Seattle and the Sea-Tac Airport, while BC also saw a strong change along major highways.

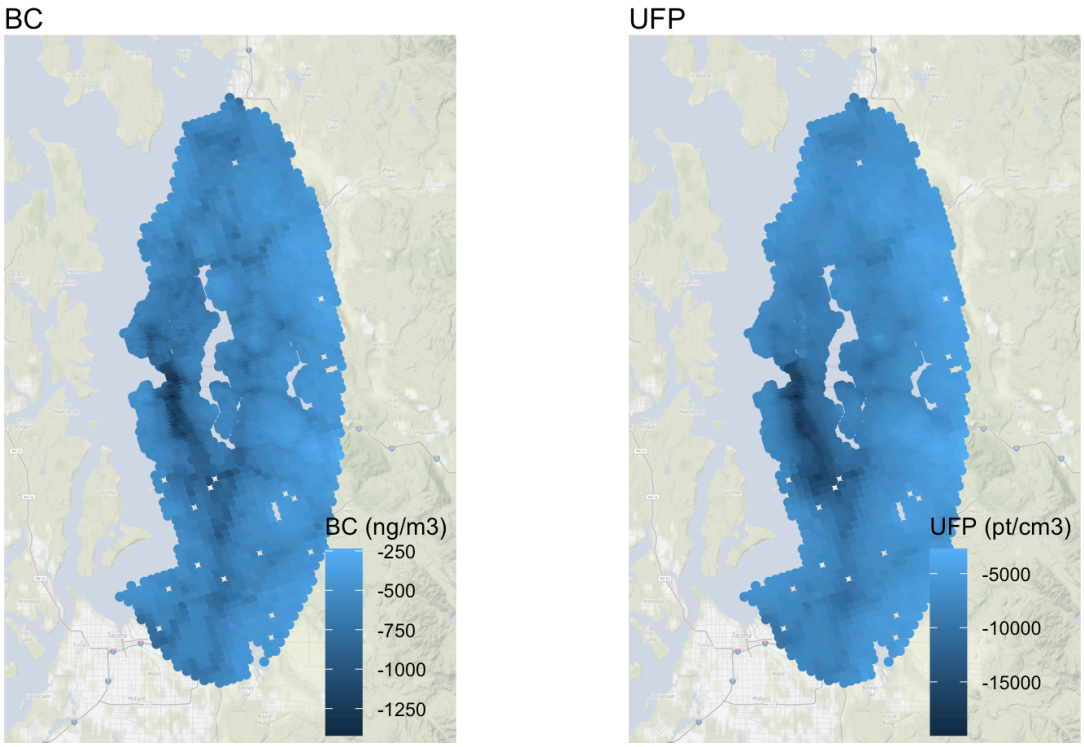


Figure 48. Reduction in the predicted BC and UFP exposure surfaces between 1995-2019

A comparison of model predictions for ACT cohort locations with and without a temporal adjustment showed that the trend adjustment impacted predictions substantially more than changing values of TVCs over time (Figure 49). Note that the annual predictions shown are for all cohort locations on record and are not time-varying, as would be conducted in a health analysis.

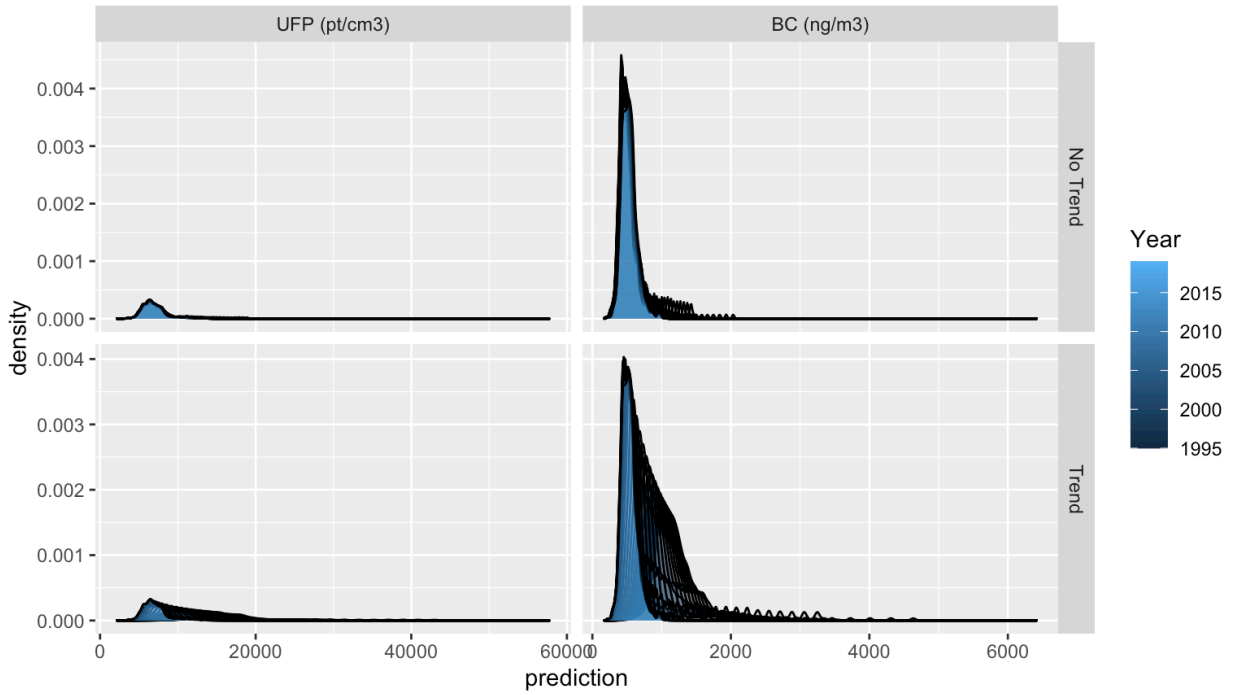


Figure 49. Model predictions at ACT cohort locations before and after applying a temporal trend adjustment. No trend plots are those where different values of time-varying covariates were used for any given year prior to using temporal trend adjustment factors for the final predictions. There is one density curve per year. Note that the annual predictions shown are for all cohort locations on record and are not time-varying, as would be conducted in a health analysis.

Validation

Validation Dataset

Sixty-five of the original 98 site-years met our inclusion criteria and were included in the final BC validation dataset (66%). Observations came from 9 different sites that had collected BC measurements between 2003-2019 (Figure 50). BC concentrations during this time had a median (IQR) of 1,109 (476) ng/m^3 and ranged from 630-1,833 ng/m^3. Each year had between 1-5 sites with annual averages. Median annual concentrations ranged from 783-1,554 ng/m^3, with typically higher concentrations during earlier years. Each site had between 2 and 12 years of data. Median site concentrations ranged from 836-1,365 ng/m^3.

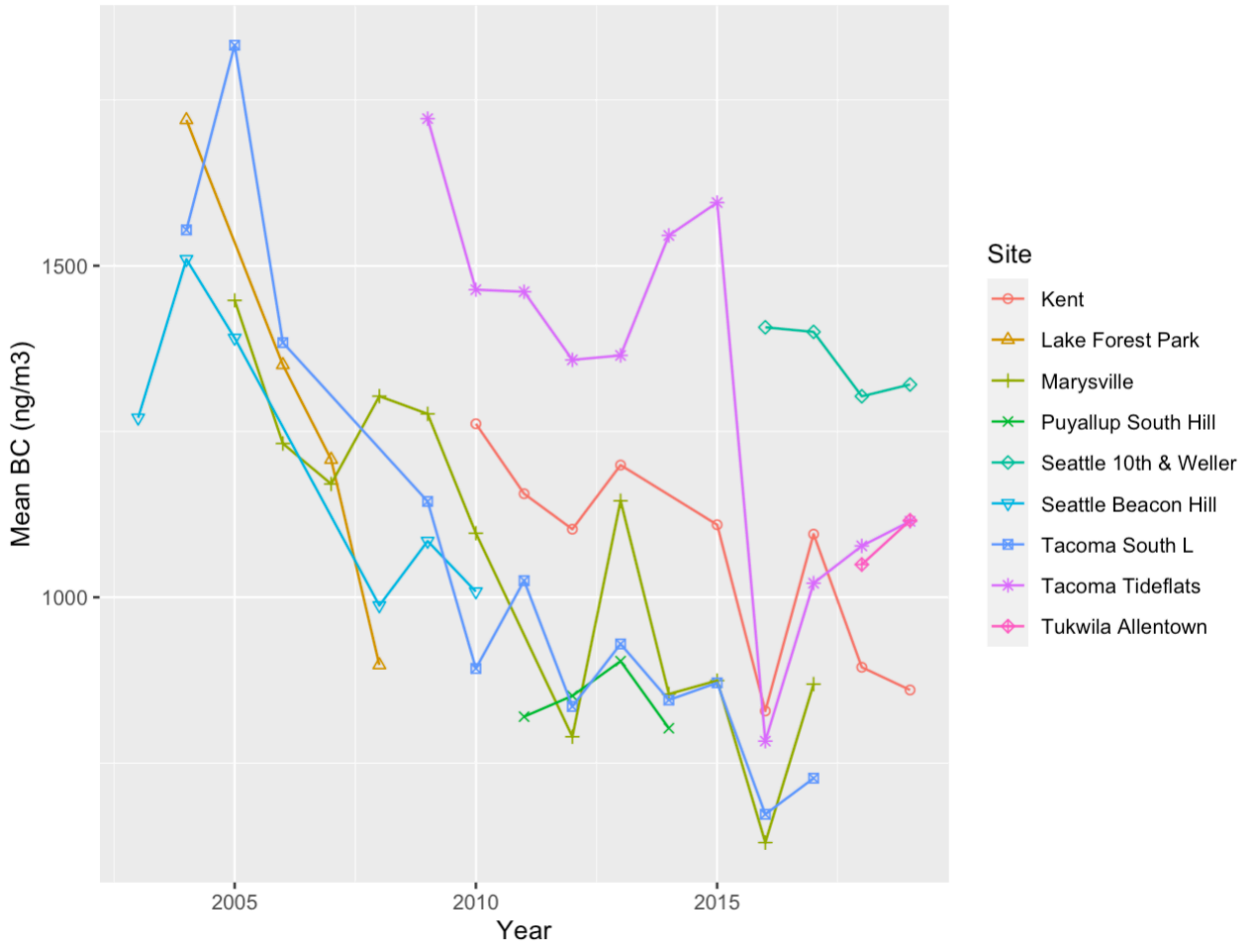


Figure 50. Historical annual average BC concentrations at AQS sites

The validation dataset for UFP consisted of a range of particle concentrations estimated from the volume concentrations presented by Kim et al. (2004) during their 2001 Beacon Hill study (E. Kim et al., 2004). The estimated median UFP concentration was 92,119 pt/cm^3 and ranged from 69,451-131,404 pt/cm^3.

Model Validation

The BC model captured the overall concentration trend at all of the available AQS site-years in our out-of-sample validation set, though the model underestimated observed concentrations (Figure 51). Model predictions were on average 61% of the value of the observed

concentrations (39% lower) and ranged from 40-86%. The model's out-of-sample RMSE was 502 ng/m^3, while the out-of-sample MSE-based R^2 (R^2_{MSE})and regression-based R^2 (R^2_{Reg}) were 0 and 0.50, respectively. The site-specific RMSE estimates ranged from 313-634 ng/m^3, with slightly better model performances at Beacon Hill and Kent than Tacoma Tideflats, though this difference was small (Figure 52, Table 15). Year-specific RMSE estimates ranged from 189-688 ng/m^3 (Figure 53, Appendix B Table 16). There was no clear temporal pattern in the model's performance. The R^2_{MSE} was 0 for all site-specific and most year-specific estimates since most predictions were not the same as the estimates. The R^2_{Reg} was higher and more variable for all estimates, presumably because this approach reports whether variables are linearly associated, regardless of whether estimates and predictions are the same.

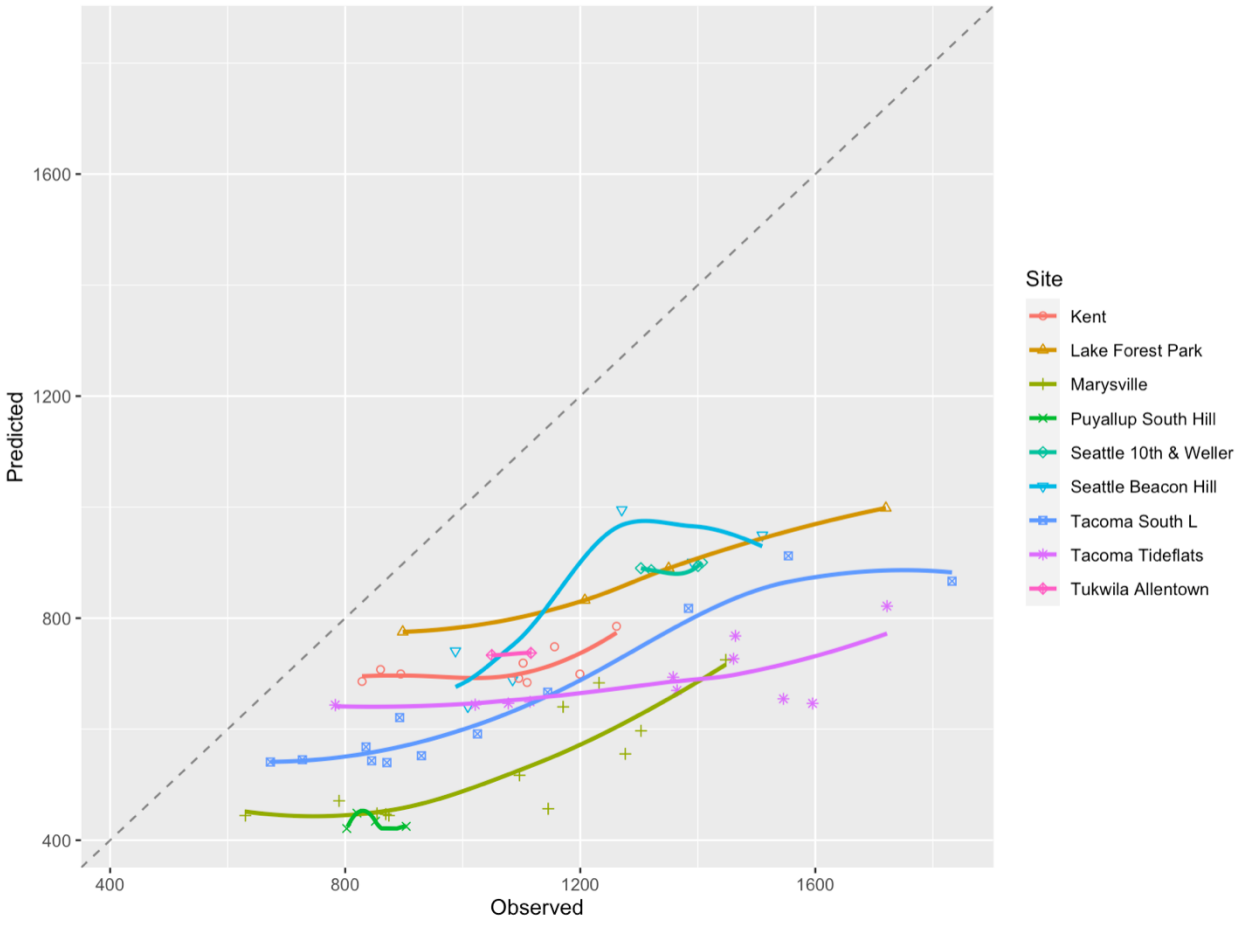


Figure 51. Observed vs predicted BC concentration (ng/m^3) at validation AQS sites. Diagonal dashed line is the 1-1 line.

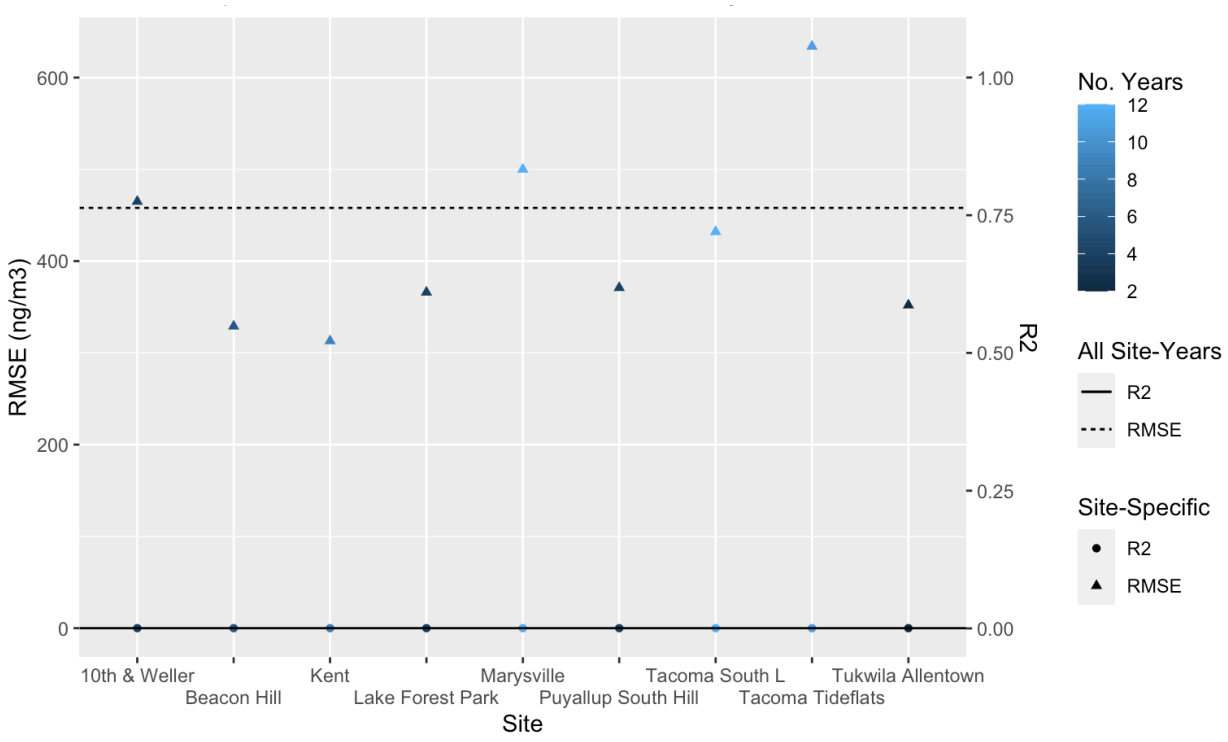


Figure 52. Validation results by site, including MSE-based R^2

Table 15. Validation results by site. N = number of years with BC observations.

Site	N	Years	RMSE (ng/m^3)	R^2_{MSE}	R^{2Reg}
Kent	9	2010, 2011, 2012, 2013, 2015, 2016, 2017, 2018, 2019	313	0	0.32
Lake Forest Park	4	2004, 2006, 2007, 2008	366	0	0.98
Marysville	12	2005, 2006, 2007, 2008, 2009, 2010, 2012, 2013, 2014, 2015, 2016, 2017	500	0	0.68
Puyallup South Hill	4	2011, 2012, 2013, 2014	371	0	0.05
Seattle 10th & Weller	4	2016, 2017, 2018, 2019	465	0	0.72
Seattle Beacon Hill	6	2003, 2004, 2005, 2008, 2009, 2010	329	0	0.71
Tacoma South L	12	2004, 2005, 2006, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017	432	0	0.89
Tacoma Tideflats	11	2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019	634	0	0.38
Tukwila Allentown	2	2018, 2019	352	0	1.00

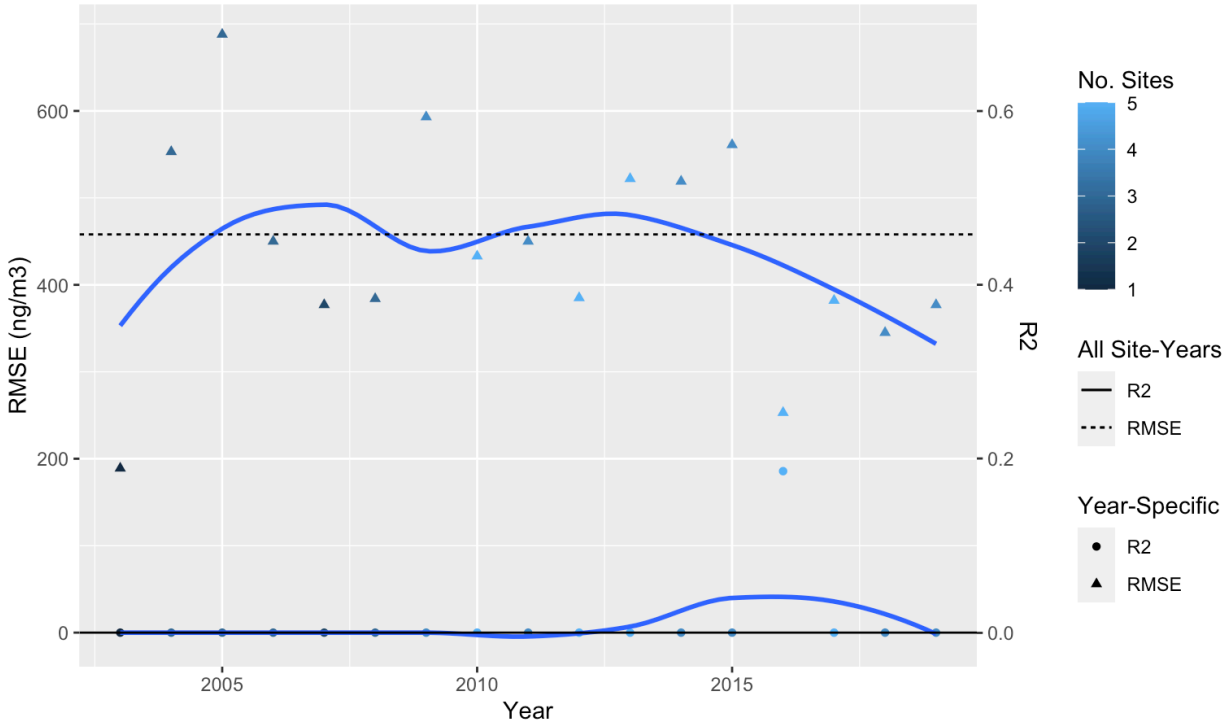


Figure 53. Validation results by year, including MSE-based R^2

Table 16. Validation results by year. N = number of sites.

Year	N	Sites	RMSE (ng/m^3)	R^2_{MSE}	R^2_{Reg}
2003	1	Seattle Beacon Hill	189	0.00	
2004	3	Lake Forest Park, Seattle Beacon Hill, Tacoma South L	553	0.00	0.64
2005	3	Marysville, Seattle Beacon Hill, Tacoma South L	688	0.00	0.05
2006	3	Lake Forest Park, Marysville, Tacoma South L	450	0.00	0.71
2007	2	Lake Forest Park, Marysville	377	0.00	1.00
2008	3	Lake Forest Park, Marysville, Seattle Beacon Hill	384	0.00	1.00
2009	4	Marysville, Seattle Beacon Hill, Tacoma South L, Tacoma Tideflats	593	0.00	0.44
2010	5	Kent, Marysville, Seattle Beacon Hill, Tacoma South L, Tacoma Tideflats	433	0.00	0.47
2011	4	Kent, Puyallup South Hill, Tacoma South L, Tacoma Tideflats	450	0.00	0.73
2012	5	Kent, Marysville, Puyallup South Hill, Tacoma South L, Tacoma Tideflats	385	0.00	0.68
2013	5	Kent, Marysville, Puyallup South Hill, Tacoma South L, Tacoma Tideflats	522	0.00	0.48
2014	4	Marysville, Puyallup South Hill, Tacoma South L, Tacoma Tideflats	519	0.00	0.78
2015	4	Kent, Marysville, Tacoma South L, Tacoma Tideflats	561	0.00	0.44
2016	5	Kent, Marysville, Seattle 10th & Weller, Tacoma South L, Tacoma Tideflats	253	0.19	0.90
2017	5	Kent, Marysville, Seattle 10th & Weller, Tacoma South L, Tacoma Tideflats	382	0.00	0.84
2018	4	Kent, Seattle 10th & Weller, Tacoma Tideflats, Tukwila Allentown	345	0.00	0.61
2019	4	Kent, Seattle 10th & Weller, Tacoma Tideflats, Tukwila Allentown	377	0.00	0.46

The UFP model underpredicted the lower range of the observed UFP concentration at Beacon Hill in 2001 by 51,313 pt/cm^3 (equivalent to the RMSE) (Figure 54). This was 74% lower than the lower range of the observed concentration.

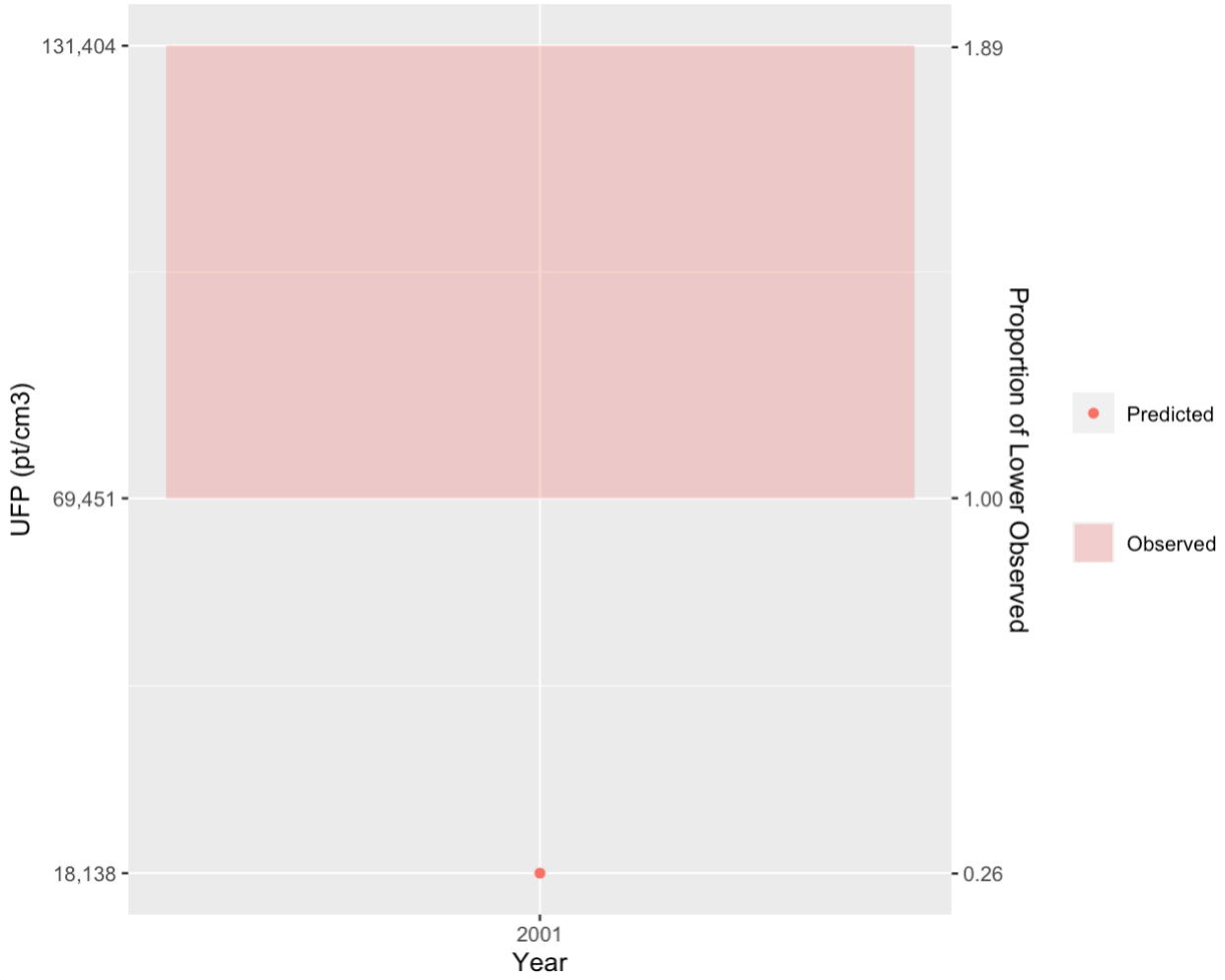


Figure 54. Observed (Kim et al., 2004) UFP concentration range and predicted UFP concentration for 2001 at Beacon Hill

Sensitivity Analyses

The resulting NOx emission factors gathered from MOVES and total emission estimates on 100 m road links had very similar temporal and spatial patterns as EC (Pearson correlation, R

= 0.998; Figure 55; Appendix C Figure 83 - Figure 85). As with EC, earlier times and locations near major highways generally had higher NOx emissions.

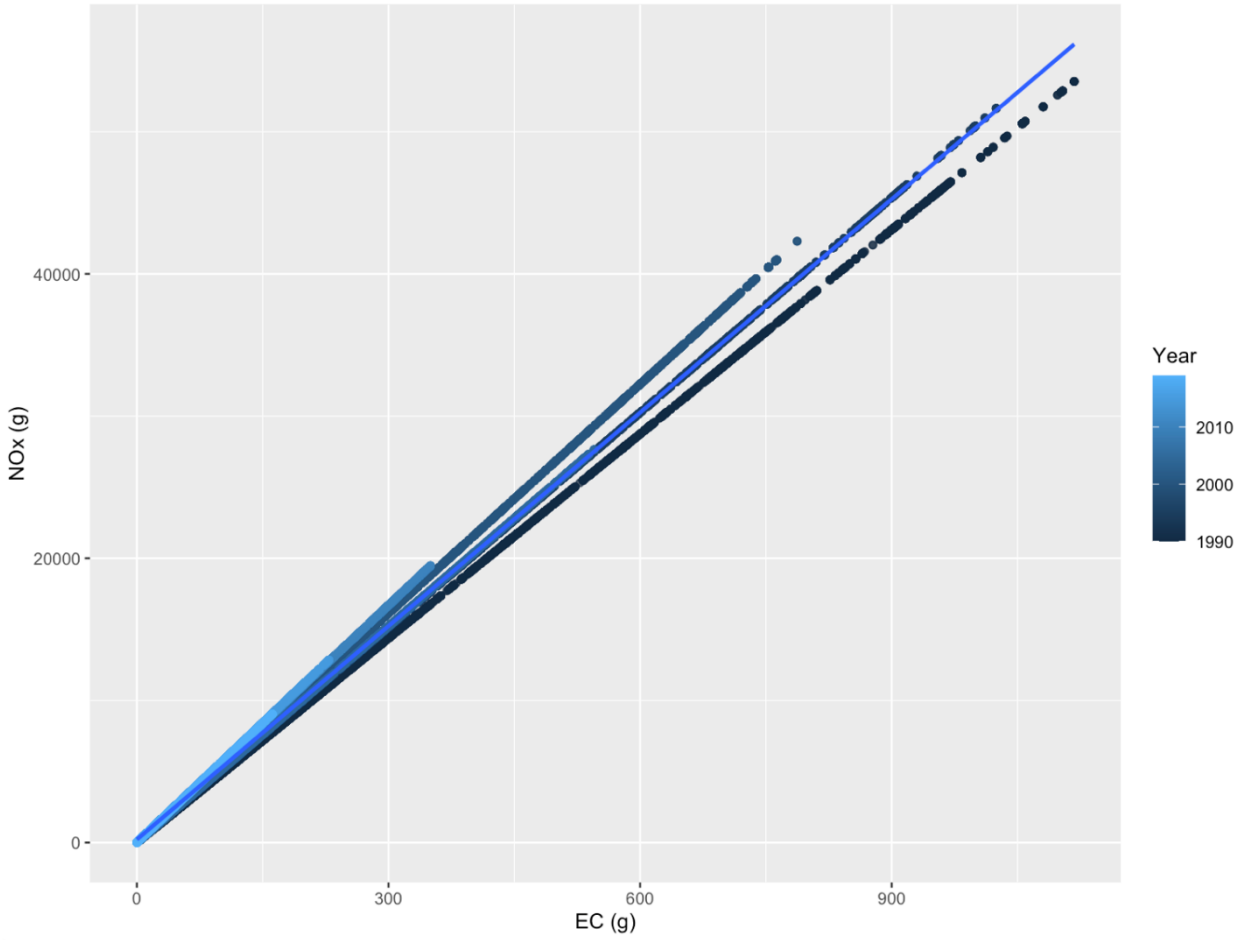


Figure 55. Comparison of final NOx and EC emissions on 100 m road links

As with EC PLS scores, NOx PLS scores from the BC and UFP models for ACT cohort locations were larger and more variable during earlier time periods, with locations near major highways having particularly high values during earlier times (Appendix C Figure 86 – Figure 88). NOx PLS scores generally decreased over time, with the largest decreases observed along major highways (e.g., I-5).

BC and UFP predictions from this NOx predictor model were almost identical to predictions from the primary model (BC: R^2~1.0, RMSE =13 ng/m^3; UFP: R^2 ~1.0, RMSE = 142 pt/cm^3; Figure 56). The NOx predictor BC model's overall out-of-sample RMSE and R^2 were the same as the primary model at 502 ng/m^3 and 0, respectively. The model's site- and year-specific

RMSE and R^2 were almost identical. The NOx predictor UFP model's overall out-of-sample RMSE and R^2 were also similar to those of the primary model at 51,737 pt/cm^3 and 0, respectively.

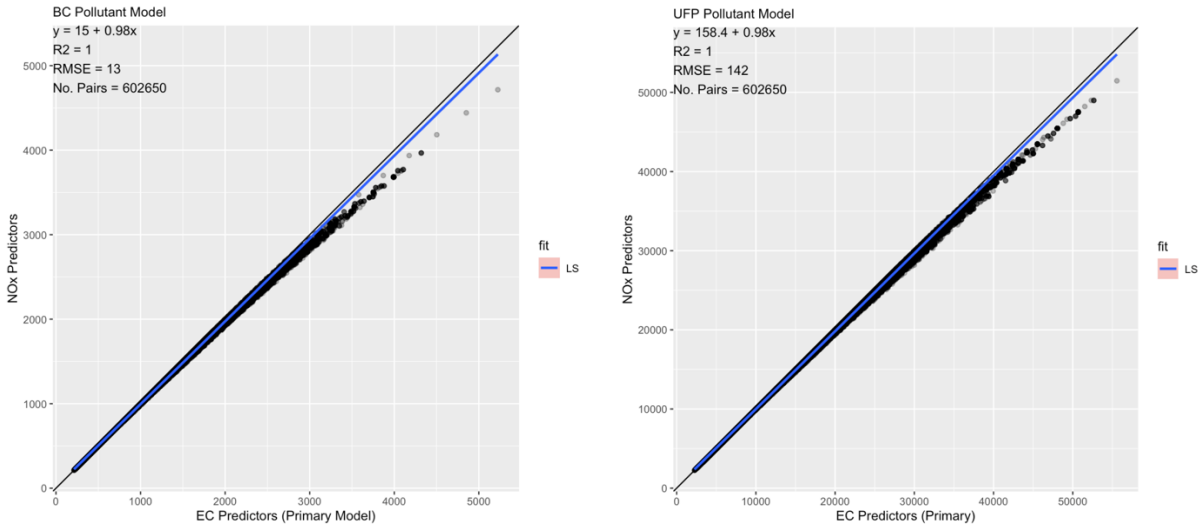


Figure 56. Comparison of BC (ng/m^3; left) and UFP (pt/cm^3; right) predictions at cohort locations using EC (primary) and NOx emission predictors. Prediction comparisons are for all study years (1995-2019).

Discussion

Novelty

In this work, we are the first to develop long-term UFP and BC exposure surfaces for the ACT cohort. A strength of this study is that, to the best of our knowledge, we are the first to extrapolate an exposure surface derived from a mobile monitoring campaign back in time. Most others have relied on dispersion models, satellite images, regulatory monitoring, and/or short-term fixed sampling with few repeat visits per site (Bellander et al., 2001; Hystad et al., 2012; S. Y. Kim et al., 2017; Levy et al., 2015; Meng et al., 2019; Mölter et al., 2010b; Montagne et al., 2015; R. Wang et al., 2013). As such, we were able to extrapolate a larger and likely more accurate exposure surface back in time when compared to other studies. Furthermore, most studies have focused extending NO_2, NOx and $PM_{2.5}$ surfaces back in time. Only one study has extrapolated UFP and BC exposure surfaces back in time, though this was done from a short-term monitoring campaign with only 3 repeat visits per site, and predictions were only

extrapolated back ten years (Montagne et al., 2015). Furthermore, in this study we document an approach for capturing decreasing air pollution trends over time despite only having a single regulatory monitoring site within our study area with long-term observations and the presence of missing data. Our extensive use of covariate-buffer predictors and interpolation of these at the annual level resulted in fine spatial adjustments at the site level. Despite having a limited validation dataset, we know that TRAP levels have changed over time, and at some places more than others. Our approach acknowledges this and produces exposure surfaces that we believe will yield improved long-term exposure assessment for epidemiologic cohorts when compared to more common methods involving short-term exposure surrogates. As a whole, we thus describe an approach using time-varying covariates (TVCs) and temporal adjustment factors under circumstances of limited exposure data in order to extrapolate otherwise unavailable exposure surface back in time. This approach builds on past work in the field and specifically applies it to an extensive mobile monitoring campaign designed for epidemiologic application.

Geocovariates

Using space- and time-varying highway vehicle emission covariates (TVCs), population density and NDVI allowed us to account for spatiotemporal changes in these important TRAP predictors. Major roadway emissions are important contributors to overall TRAP concentrations, while population density and NDVI are both strongly association with TRAP (Cattani et al., 2017; Hatzopoulou et al., 2017; Kerckhoffs et al., 2016; Messier et al., 2018; Minet et al., 2017; Montagne et al., 2015; Saha, Li, et al., 2019; R. Wang et al., 2013; Weichenthal, Ryswyk, et al., 2016; Xie et al., 2017; Yuchi et al., 2020). Population density and NDVI were especially useful at locations where vehicle emissions were otherwise unavailable, for example, away from major roads and in rural areas.

Highway Emissions

Given that TRAP emissions on road sections were unavailable prior to 2005, our approach of using point counter data, which was available starting in 1990, to interpolate distance-weighted AADT on road sections allowed us to use this covariate in our models. The moderate to good agreement between our AADT estimates on 100 m road links and WSDOT

road segment estimates starting in 2005 indicated that this approach was reasonable. A similar, but more complex interpolation approach is used by WSDOT. Though the smoothing of AADT estimates based on other nearby road links did not change most estimates, we incorporated this step since a few road links did see large changes. This difference may have been a consequence of our GIS approach for identifying nearby counters. The final smoothed road link emission estimates generally decreased and were less spatially variable over time. Emissions near major sources (i.e., highways) had the sharpest drops.

A few other studies have also used vehicle emissions within buffers in LUR models and found significant associations with TRAPs, including UFPs and BC (Hatzopoulou et al., 2017; Minet et al., 2017; Weichenthal, Ryswyk, et al., 2016). Many past studies, however, have used traffic counts in LUR models as a simpler emissions surrogate (Rob Beelen et al., 2013; Henderson et al., 2007; Kerckhoffs et al., 2016; Montagne et al., 2015; Saha, Li, et al., 2019; Xie et al., 2017). Given our focus on long-term exposure, however, we felt it necessary to also incorporate emission factors in order to better characterize changing emissions over time, which have been impacted by both generally cleaner running cars and an increase in vehicle miles driven.

PLS

We were able to more efficiently interpolate geocovariate values for years with unavailable data and more finely adjust model predictions over time by using PLS to summarize each geocovariate's many buffer combinations into a single linear combination.

Over time, we generally saw decreasing emission and NDVI PLS scores and increasing population density PLS scores, illustrating the temporal changes in these covariates. PLS scores at mobile monitoring locations (from 2019 emissions, 2019 NDVI and 2010 population density used to build each TVC model) were similar but slightly less variable than the PLS scores at cohort locations, indicating spatial compatibility (Figure 36). Over time, however, the distribution of cohort location PLS scores generally increased or decreased, moving away from the mobile monitoring distributions in 2019 (used to fit the models), suggesting that model predictions were being extrapolated outside of the geocovariate space used to fit the models. This was especially true for EC emissions, which were much higher at some locations during earlier

times than the highest values we observed during 2019. This approach of adjusting model predictions thus assumes that the same relationship held between each TVC predictor and TRAP since no other monitoring observations were available during earlier time periods. The spatiotemporal model used to generate predictions in Aim 1 is also heavily reliant on the spatial contrasts and temporal trends observed in more recent years, though not exclusively so as in this work, due to the scarcity of historical data (Keller et al., 2015).

Maps of TVC PLS scores further illustrated how temporal changes in these covariates were more pronounced in some areas than others. Between 1995 and 2019, for example, population density PLS scores saw the greatest increases near the Seattle area, while emissions PLS scores saw the greatest decreases along major highways. These trends highlight the importance of incorporating TVCs into our models, particularly when long-term exposure assessment is of interest, rather than solely relying on a temporal trend adjustment where all locations are adjusted to the same degree.

Trend

We estimated a temporal trend from Beacon Hill observations of EC between 1997-2018. A strength of this approach was the availability long-term observations that were collected using the same protocols over time from an AQS site located in the center of our study area.

In order to estimate temporal trend adjustment factors from long-term observations, past studies have used both ratio ($X_i/X_{reference}$) and additive ($X_i - X_{reference}$) approaches (Meng et al., 2019; Mölter et al., 2010a; R. Wang et al., 2013). We decided to use a ratio rather than an additive approach since our trend was based on EC, which: a) could had a different concentration magnitude than BC, and b) was based on different units than UFP (mass vs. number concentration).

The resulting EC trend was similar to the BC trend observed at AQS sites in our validation set, suggesting that this approach was appropriate. This trend reflected two emission-reducing policies that are believed to have drastically lowered particle levels in the Puget Sound (US EPA, 2019d). The first policy was enacted in 2007 and required new diesel trucks (important contributors of particle pollution) to use particle filters (Transport Policy, 2018; US EPA, 2001). The second policy was phased in between 2007-2014 and required lowered sulfur

levels in ship fuel (T Larson, 2017). Using a natural cubic spline to smooth EC observations and reduce random, year-to-year variability allowed us to capture this steep decrease in concentrations between 2007-2014 as well as the overall long-term trend. We were also able to extrapolate this trend to 1995 and 2019 where we otherwise had no observations.

As mentioned above, adjusting the entire exposure surface based on an annual temporal trend adjustment factor assumes that TRAP levels changed at all locations in a similar fashion (e.g., TRAP decreases by X% at all sites), though after the inclusion of TVCs in the universal kriging models, which themselves did change the predictions to different degrees over time at different locations. This could have resulted in some locations having more exposure misclassification than others, particularly those that are different from Beacon Hill, an urban background site. We may expect, for example, that very urban locations such as those near downtown areas and I-5 may have truly seen steeper decreases in TRAP over time. The inclusion of TVCs, which could capture these spatial changes over time, was meant to address this gap.

Still, while a temporal TRAP trend may be important for long-term exposure assessment, it may be less essential for epidemiologic investigations where calendar time is adjusted for. Our NOx-Dementia survival analysis presented in Aim 1 is an example of this since adjusting for calendar time while using age as the time axis meant that we were comparing individuals who were the same age during similar calendar years. Again, this highlights the importance of using TVCs in our exposure models, even if their impact on model predictions was relatively small compared to the trend adjustment.

In additional analyses, we characterized the contribution of the TVC, space covariates and trend adjustment factors (Appendix C: Additional Analyses of Space, Time-Varying Covariates and Trend Adjustment Factors). The following discussion pertains to UFPs, though BC findings were very similar. Figure 94 shows that UFP concentrations were more strongly correlated with the PLS space covariates (space1-3) than with PLS TVCs (pop2010, ec2019, ndvi2019). In a simplified linear regression model of UFP with standardized space and TVC predictors, the space covariates had larger coefficient estimates that contributed significantly to the model, while the space covariates had smaller coefficient estimates, and only the NDVI TVC contributed significantly to the model (Table 38). In additional linear regression models of UFP with space and/or TVC predictors, we again saw that models with space covariates alone explained a moderate degree of the variation in UFP (R^2=0.72), models with TVCs alone

explained a smaller degree of the variation (R^2=0.23), and models with both space and TVCs explained a similar amount of variation as the space-only models (R^2=0.73; Table 39). To better understand the variability in UFP explained by the space covariates in the presence of TVCs, we fit a model to the residuals of the TVC only model using space covariate predictors. The space covariates explained 51% of the variability of these residuals, suggesting that the TVCs explain roughly 45% of the variability that space covariates do (0.23/0.51 = 0.45).

We compared how the exposure surface changed over time before adding a temporal trend (changes in predictions over time were due to changing TVCs) and after adding a temporal trend (our primary model; Appendix C Figure 96). The model without a temporal trend predicted increases in UFP concentrations for earlier times, particularly at higher concentration sites. The model with a temporal trend predicted similar trends, though the increases were larger in scale and also impacted low-concentration sites. Maps of predicted UFP concentrations before and after adding temporal trend adjustment factors illustrate these patterns (Appendix C Figure 96).

These results suggest that while the space covariates were more strongly correlated to UFP and BC than the TVCs and thus weighted heavier in the models, over time TVCs still contributed to changes in the exposure surfaces. These changes were more evident for higher concentration sites which experienced the largest changes over time as a result of changing emissions, population density and NDVI, though the resulting changes were more subtle than those from the primary models which also include trend adjustment factors. This suggests that health analyses that adjust for time could still benefit from exposure surfaces derived from TVCs since these may reduce exposure misclassification.

UK Model

The final models predicted a varying degree of BC and UFP exposure levels for the ACT cohort over time. The distribution of predicted concentrations generally had the same shape as the 2019 predictions (the reference year for the temporal trend), but with higher, more variable concentrations during earlier times. In this study, the median and IQR of the final predicted exposure levels for the ACT cohort were cut to less than half from 1995 to 2019. These results are in line with changes in national and regional emissions as well as overall air quality since

1990 (US EPA, 2019a). These have shown, for example, roughly a 50-60% drop in ambient NO_2 and 78% drop in ambient CO levels.

Maps looking at changes in the predicted BC and UFP levels over time were similar, with the largest drops occurring near the center of Seattle, along I-5 near downtown and the Sea-Tac Airport, and along other major highways. There were some predicted differences, however, with larger drops in BC levels occurring near major highways than UFP, and larger drops in UFP levels occurring along I-5 near the airport than BC. Changes in BC levels may be a result of cleaner running diesel trucks along major highways (i.e., lower emission factors from large contributors) (Bureau of Transportation Statistics, 2020; US EPA, 2018c). Drops in UFP levels may be due to cleaner aircraft technologies and fuels, improved airport ground traffic flow, and additional modifications in the built environment meant to reduce ground-level aviation emissions (Aviation Benefits Beyond Borders, 2020).

While some studies have justified the use of the same exposure surface for long-term epidemiologic studies by arguing that the spatial concentration contrasts are conserved over time (Cesaroni et al., 2012; Eeftens et al., 2011), this may not be true for very long-term exposure studies, particularly those that do not adjust for time. In the case of dementia incidence in the ACT cohort, for example, the exposure period of interest may be over 25 years, over which the spatiotemporal changes in TRAP may be substantial. This may be especially true for places near major TRAP sources such as highways and major airports.

Validation

BC validation AQS sites represented a mixture of residential, suburban, urban center, commercial and industrial locations, indicating that these sites were generally diverse and representative of the ACT cohort.

BC

The BC model generally captured the observed decreasing TRAP concentration patterns over time and predicted elevated BC levels at the appropriate sites. The model, however, underpredicted observed AQS site concentrations. The fact that the temporal trend adjustment

factors changed model predictions more substantially than the time-varying covariates alone indicate that our models relied heavily on temporal trend adjustment factors from Beacon Hill observations of EC. A temporal trend derived from the actual pollutants of interest (i.e., BC and UFP) and from higher concentration locations that observed steeper decreases in air pollution over time could improve the model predictions. We did not conduct this analysis because historical data were limited, and this would have further reduced our BC validation set.

Moreover, the BC model performed slightly better at some locations than others, though these differences may not have been meaningful. We saw a higher model performance, for example, at Beacon Hill. Since EC from Beacon Hill was used to establish a temporal trend, and it is correlated with BC, this likely explains the slightly better model performances at this site. On the other hand, the model performed slightly worse at 10th & Weller. This difference could be due to different monitoring years (10th & Weller monitored during later years than Beacon Hill), or the different locations where the 10th & Weller and Beacon Hill AQS sites are located (urban center vs residential/suburban area).

Furthermore, it was difficult to assess whether the model's performance changed over time. This was because each year only had 1-3 site observations, and the sites included in each year varied from year-to-year. The variability in site- and year-specific R^2_{Reg} estimates ranging from poor to strong agreement may be partly due to this variability in the sites and years included in each estimate. Since we fit the model using more recent BC observations and used these to predict BC levels during earlier time periods, this resulted in the model being extrapolated outside of the modeled geocovariate space for the TVCs. In addition, there were no available observations prior to 2003, limiting our ability to assess the performance of our model between 1995-2002. To adequately assess how our model performed over time, we would have liked to have had observations from many AQS sites over the entire time period of interest.

Other possible reasons for the discrepancy between the observed and predicted model concentrations could have been related to unusual weather events during particular years and/or different sampling protocols including the equipment type and exact sampling location.

UFP

Like the BC model, the UFP model also underpredicted the UFP concentration at Beacon Hill during 2001. Our evaluation of this model was limited to this single site-year, however, making it difficult to thoroughly assess this model.

There are a few possible explanations for the model's underprediction of UFP concentration. First, this could be related our conversion of volume concentrations to number concentrations since this was comparable with what was measured in the mobile monitoring campaign. We used the minimum, median and maximum particle diameter of each bin during the conversion process in order to estimate a concentration range since particle diameter is an important factor of this conversion. There were other conversion factors, however, that could have affected this estimate (e.g., the standard atmospheric conditions assumption). Moreover, different sampling protocols between our mobile monitoring campaign and the 2001 study could have contributed the model's performance. Kim et al. (2004), for example, collected UFPs 4 m above ground, while our mobile monitoring campaign collected samples ~1.5 m above ground (vehicle rooftop). Both campaigns additionally used different instrumentation. While we used TSI P-TRAK 8525 particle counter, they used a differential mobility particle sizer (DMPS) consisting of a TSI Electrostatic Classifier 3080 with a model 3081 long column differential mobility analyzer and a TSI model 3010 condensation nuclei counter.

Sensitivity Analyses

BC and UFP predictions from the NOx predictor models were very similar to predictions from our primary models. Since EC and NOx are highly correlated, it makes sense that their predicted spatial and temporal patterns were very similar. Using only one of these pollutants in future work would likely suffice.

We chose not to conduct other sensitivity analyses since these would have been difficult to interpret given our limited validation data and would thus not have answered our scientific question of interest. For example, we considered using historical BC concentrations from AQS sites near our study area to build alternative temporal trend adjustment factors. And while we believed that this approach could have better characterized historical BC and UFP trends when

compared to EC at Beacon Hill alone, this would have left us with no out-of-sample observations to validate our models. Using a subset of the AQS sites for this analysis would have made it harder to see the spatial or temporal patterns that may have occurred over time.

Limitations

Time-Varying Geocovariates

The available model covariates that met all of our data inclusion criteria were limited. Most covariates did not vary sufficiently over space, were unavailable going back to the mid 1990's and/or had been collected differently over time. Still, the variables that we did use have been previously identified as strong predictors of TRAP (Abhijith et al., 2017; Apte et al., 2017; Hatzopoulou et al., 2017; Hoek et al., 2008; Kerckhoffs et al., 2016; Lane et al., 2016; Patton et al., 2015; Patton, Collins, et al., 2014; Saha, Li, et al., 2019; Weichenthal, Ryswyk, et al., 2016; Xie et al., 2017). Studies with shorter, more recent exposure periods of interest will have an expanded set of available covariates, for example public transportation routes and stops, Washington and National Emissions Inventories, more complete road networks, urban imperviousness, etc.

The lack of available data dating back to the mid 1990s meant that we used default MOVES EC emission factors for King County in order to estimate total BC and UFP vehicle emissions. These pollutants have all been shown to be strongly associated with each other (Janssen et al., 2017; Kerckhoffs et al., 2016; Salako et al., 2012), though here we assumed that this relationship was constant over time. This may be a strong assumption given that technological advances have reduced different vehicle emissions to varying degrees over time (US EPA, 2019a). Notably, the national- and county- scale default input tables for King County were almost identical (Appendix C Figure 93) for two example years chosen: 1999 and 2009. There was only one discrepancies for 1999 where the national-scale actually had slightly more complete vehicle miles traveled (VMT) by vehicle type information than the county-scale. Upon closer inspection of default input tables when running MOVES at the national-scale but for different counties, it also appears as though input tables vary, reassuring us that emission factors are to some degree customized for a given county even at the national scale. The advantage of

calculating emission factors at the national- rather than the county- scale, however, was that national-level inputs were used when default county-level input tables were otherwise unavailable. Still, it would have been ideal to run MOVES at the county-scale with actual input data for each modeling year rather than relying on default county inputs. In practice, however, it is less likely that investigators will have access to all of these inputs, particularly for large study areas spanning long time periods. During our initial research, for example, we learned that King County itself (WADOE, WADOT, etc.) did not have archives of MOVES emission factors or all of the required inputs required to run the model at the county-scale dating back to the 1990s. WADOE had emission factors for 2011, 2014 and 2017, and they had extrapolated these to estimate emissions between the years of 2005-2020. While it is possible to further extrapolate these emissions back to the 1990s, this approach may be unjustified given that emissions and air pollution trends saw large changes during the 1990s.

Furthermore, we acknowledge that the crude spatial resolution of MOVES limits the use of this covariate in spatial models. Emission factors are available at the county level, while AADT is only available on highways and freeways dating back to the 1990s. Historical emissions on small neighborhood streets are thus unavailable, though these may contribute to total TRAP levels to a relatively smaller degree than major roads since they generally have irregular traffic volumes, making them difficult to model. For simplicity, we additionally did not include arterials in this analysis. Arterials may contribute significantly to local TRAP levels, but methods for estimating road AADT from these are more complex. Our approach in this analysis was to use MOVES emission factors in conjunction with other higher spatial resolution covariates, including population density and NDVI. In this study, we have thus shown an approachable method for incorporating time-varying emissions into analyses as an improvement over simpler traffic covariates (e.g., roadway proximity) which are more likely to result in increasing levels of exposure measurement error over time. While imperfect, emission factors generally captured the significant decreases in vehicle emissions over time, which would have otherwise been unaccounted for.

To further improve on our approach, studies could extrapolate the earliest available data, given that they had a good understanding of how it has changed over time. For example, the distribution of vehicle class could be used to calculate class-specific AADT estimates, and these could be combined with class-specific MOVES emission factors to estimate more accurate

annual emission factors. This could result in a more accurate time-varying covariate if the vehicle class distribution in King County differed from the default values available. In turn, this could increase the influence of this covariate in the fitted, 2019 exposure models. An alternative approach could be to model more recent years (~ mid 2,000s+) when more county-specific inputs become available for MOVES (e.g., vehicle age and mix). We chose not to do so in this study because this would have cut our exposure period of interest by a significant amount.

As for population density, we used the metro-level population growth trends from annual ACS Surveys between 2010 and 2019 to inflate the 2010 decennial Census Bureau population density estimates since the 2020 estimates were not yet available. This approach assumed that population density increased by the same degree at each location (~2% per year). This approach missed some of the variability that was truly observed at different locations, though it captured the overall trend.

Furthermore, many of the monthly images for our study area used calculate NDVI: a) were stitched together from different days, and b) had cloud coverage. While most of our study area was within a single stitched area, cloud coverage generally shifted the NDVI estimate for any given pixel closer to 0 (water areas are otherwise closer to -1, and other land areas are variable but can be closer to 1). It was thus unclear whether the apparent decrease in NDVI from 1990 to 2010 and increase from 2010 to 2019 was true or reflective of an anomaly that occurred during one of the years. Heavier than normal cloud coverage, for example, could have artificially lowered the NDVI. Moreover, NDVI estimates came from satellites with different technologies (Landsat 5 and Landsat 8). This likely resulted in slight discrepancies across years, as has been seen for more recent NDVI imagery from Landsat 7 and Landsat 8 (P. Li et al., 2014; Xu & Guo, 2014). Still, our results showed that the inclusion of time-varying values for NDVI, population density and emissions had a minimal impact on the model predictions back in time.

A potential solution for the existing NDVI dataset would be for future analyses to download additional NDVI images and use a smooth line to recalculate NDVI. This may be particularly important if inconsistencies appear to exist between years. A more sound solution, however, is for future studies to ensure that NDVI images are radiometrically corrected, the process of calibrating pixel values and/or correcting for errors (Humboldt State University, 2019). This is particularly important when land changes are being assessed over time (i.e., multiple images are being used). The appropriate, Surface Reflectance Higher-Level data (Level

2 Data) are available for various Landsats dating back to Landsat 4 and can be ordered on demand at no charge (U.S. Geology Survey, 2021). If studies are only interested in looking back to around the year 2000 and do not require as fine of a spatial resolution as Landsat (~250 m vs 10-30 m), they may choose to use MODIS instead of Landsat images (NASA, 2020). These are more temporally resolved (i.e., have an orbit frequency of 1 vs 16 days) and may thus also better adjust for cloud coverage, and can have quality flags for each pixel. Finally, studies may opt to use other geocovariate indicators of land use such as enhanced vegetation index (EVI), which is similar to NDVI but has additional corrections that make it more sensitive to dense vegetation (USGS, 2021).

Modeling

A feature of this analysis is that we extrapolated the 2019 exposure surface back in time using TVCs and temporal trend adjustment factors. As such, our models rely on spatial rather than temporal contrasts of TRAP and assume that the same covariate-TRAP relationships hold over time, including times when the geocovariate values are outside of the modeling ranges (R. Wang et al., 2013). For example, some population density values in 1995 were lower than the lowest observed values in 2019. This relationship assumption may be less accurate as you go further back in time. For example, while population density was generally lower during earlier years, residential driving and public transportation patterns may have also differed, thus impacting the relationship between population density and TRAP. Still, TVCs seemed to impact model predictions to a small degree, especially when compared to the impact of the trend adjustment factors, suggesting that this assumption had a minimal impact on our predictions.

To further investigate the rank order stability of a pollutant exposure surface over time, we compared $PM_{2.5}$ spatiotemporal model predictions (the pollutant with the richest dataset) at ACT cohort locations in 2018 (the most recent year with complete predictions) to earlier years, starting with the year 1995 (Appendix C Figure 89 - Figure 92 and Table 37). The mean (SD) Pearson correlation coefficient (R) was 0.81 (0.12) and ranged from 0.43-0.96. We observed a spread of the predicted concentrations for any given year relative to the prediction for 2018 (Appendix C Figure 91) and lower R values for some years (Appendix C Table 37), suggesting that the concentration rank order (i.e., the exposure surface) did change over time, and for some

years more than others. This indicates that our use of historical trend adjustment factors may be a strong assumption, particularly since TVCs had a small impact on changing the exposure surface over time. Still, the trend adjustment does capture the strong, positive association we saw between $PM_{2.5}$ concentrations in 2018 and earlier years. Moreover, past work has shown that back-extrapolation of pollutant exposure surfaces can effectively characterize past exposure levels (S. Y. Kim et al., 2017; Meng et al., 2019; R. Wang et al., 2013). Of note, this analysis was limited given the existing data, and results may vary for pollutants with different sources and spatial decay patterns such as NO_2 or other TRAPs.

Validation

In terms of model validation, we had a limited set of AQS site observations with which to validate our model predictions. As such, our model results should be interpreted with caution. For BC, we had observations dating back to 2003 but were unable to validate model predictions prior to this when extrapolation may have been an increasing concern. Our UFP validation dataset was even more limited. We only had one year's worth of data for one site, there was some missingness in the data, and we had to make various assumptions in order to estimate particle number concentrations from the volume concentrations originally presented by Kim et al. (2004). We addressed some of these concerns by calculating a weighted annual average and estimating a range of possible particle number concentrations.

Furthermore, differences in data collection methods (e.g., instrumentation, exact collection location) between our 2019 mobile monitoring campaign and AQS sites over time may have additionally impacted our validation results.

Even with limitations, however, validating our model against AQS site observations informed us that our predictions were within a generally expected range, though slightly lower than the observations.

Future Directions

This work positions us to build similar exposure models of other traffic-related pollutants and conduct epidemiologic investigations of the association between long-term TRAP exposure and dementia in the long-standing ACT cohort.

Future work may want to repeat our mobile monitoring campaign in the future and use these additional data to improve model performances (e.g., by capturing site-specific temporal trends). Models that adjust model predictions using observations from other years, for example, may have reduced bias since pollutant concentrations are correlated across time periods (Correia et al., 2013; R. Wang et al., 2013). Alternatively, others have fit separate models to different monitoring campaigns to "recalibrate" coefficient estimates and allow for the relationship between TVCs and TRAP to vary over time, a limitation of this study. In general, more work on long-term, spatially granular UFP exposure assessment is necessary. Data in this field are incredible sparse despite an increasing body of evidence linking elevated exposure levels to adverse short- and long-term health outcomes including cognitive function (US EPA, 2018b) (Table 8-31, 8-38). The spatiotemporal differences between UFP and BC levels predicted by our models additionally stresses the need for more extensive monitoring of UFPs, all of which could help validate our models more robustly.

Conclusion

In this work, we developed long-term BC and UFP exposure prediction models for the long-standing ACT cohort. We extrapolated the exposure surfaces from a one-year mobile monitoring campaign by incorporating into our models emission indicator covariates that were available dating back to the early 1990s and varied over fine spatial scales along with temporal trend adjustment factors.

Annual average BC and UFP exposure estimates for the ACT cohort varied for any given year and were generally higher and more variable for earlier years. Locations near Seattle and along major roadways saw the sharpest drops in BC levels, while locations near the Sea-Tac

Airport saw the sharpest drops in UFP levels over time. Sensitivity analyses using NOx emissions resulted in similar exposure estimates.

While our models underpredicted historical TRAP concentrations at AQS sites and do not replace extensive, long-term exposure studies, they do provide an understanding of how these otherwise poorly characterized pollutants may have changed over time at a fine spatial resolution in the Puget Sound. These long-term BC and UFP exposure estimates for the ACT cohort will support future investigations of the role of long-term TRAP exposure on dementia incidence. This field that has been previously largely limited to short-term exposure, which may increase exposure misclassification and be inadequate for long-term health studies such as dementia (Power et al., 2016; Rivera et al., 2012).

Conclusions & Future Directions

Conclusions

Dementia has been considered a major global public health priority (WHO, 2012). And while recent animal and human studies have begun reporting on air pollution's neurotoxicity, including its association with dementia, many of these studies are extremely limited in their exposure assessment (J. L. Allen et al., 2017; Carey et al., 2018; Chang et al., 2014; Chen, Kwong, Copes, Hystad, et al., 2017; Chen, Kwong, Copes, Tu, et al., 2017; Kilian & Kitazawa, 2018; Oudin et al., 2016; Power et al., 2016). This body of work addressed some of the important limitations in the field by taking into account TRAP's spatial and temporal variability despite the limited monitoring data available. The ACT cohort, a community-based, prospective cohort study on aging and the brain, was the basis for this work.

First, we are one of the first to investigate time-varying, long-term NO_2 exposure and late-onset dementia incidence using predictions from our spatiotemporal model. The model leveraged decades of regional monitoring data and local data from smaller exposure assessment campaigns at locations representative of participant homes. In line with the current literature, we saw no strong evidence of an association between long-term NO_2 exposure on late-life dementia.

Second, our extensive 2019 TRAP mobile monitoring campaign allowed us to develop well-performing, spatially rich models of annual average UFP and BC concentrations. Using time-varying traffic-indicator covariates and historical trends in elemental carbon levels in the region, we extrapolated these models back to 1995 and characterized these otherwise poorly understood long-term exposure surfaces. To the best of our knowledge, these are one of the first long-term, spatially rich UFP and BC exposure prediction models specifically designed for epidemiologic application. The development of these models will position us to conduct future investigations of long-term TRAP exposure and dementia incidence in the ACT cohort.

Future Directions

Given the projected increase in dementia cases and higher vehicle miles driven in the near future (Alzheimer's Association, 2019; Sperling, 2018; Wortmann, 2013), exposure assessment methods that consider both the spatial and temporal aspects of TRAP will continue to be essential. While this work has already advanced our understanding of long-term traffic pollution exposure and its role in late-life dementia, we are well-positioned to continue to make strides in this increasingly appreciated but understudied field.

Refinement of this Work

A valuable next step in this work would be to refine the models presented in Aims 2-3 of this work, for example by incorporating new covariates that characterize the spatiotemporal trends in air pollution. Our work showed that locations northeast (downwind) of the airport had higher observed concentrations than would have been predicted by the existing geocovariates in the land use regression (LUR) piece of our universal kriging models. This suggests that the development of new covariates that better characterize being within the downwind impact area, proximity to the airplane landing strip and overall aviation traffic could improve our prediction models without having them rely on existing nearby observations to make these adjustments.

Further work could be done to understand the smaller scale temporal pattern differences between UFP, BC and other pollutants. For example, do the annual average exposure surfaces for UFP, BC and other pollutants differ as a result of large differences across pollutant levels during rush hours or nighttime?

As we noted in Aim 3, the NDVI geocovariate should also be improved, with an emphasis on gathering radiometrically corrected values that are comparable over time. The inclusion of new built environment covariates could be a valuable asset, as these are largely missing from our models other than as general land use and type covariates (e.g., industrial, green space). High-rise buildings, restaurants; processing, manufacturing and industrial facilities, for example, could characterize important sources and dispersion factors not otherwise captured by our existing covariates. Urban locations where these footprints are more likely to be greater would most benefit from these improvements. Finally, exploring other available data through

servers such as Google Earth Engine could enrich our access to time-varying covariates important for the characterization of long-term air pollution trends.

Comparable Exposure Models

It would also be valuable to put our results into context by developing models comparable to those of other mobile monitoring exposure assessment studies. We could develop models using subsets of the data, for example, just the mobile data; just data from regulatory monitoring collocation sites; just data from urban, higher population density locations, etc. to reflect the sampling choices of other campaigns. We understand more about the importance of collecting repeated, temporally-balanced samples from any given site in order to estimate unbiased long-term averages from our other work looking at NOx measurements from regulatory monitors in California (Blanco et al., 2021). This work suggests that shorter-term, two-season rush hours or business hours designs (which are common in the literature) are more likely to result in biased annual average estimates than year-around balanced designs similar to our 2019 campaign. Finally, it would be informative to build exposure models using simpler land use regression-only approaches with a fewer number of modeled geocovariates, as is common in the literature, and see how our model performances differ. These analyses should additionally report the in-sample regression-based R^2, the default output in most statistical software, which is likely what most other studies have reported. This estimate often performs better than the out-of-sample and MSE-based R^2 estimates that we reported, which are more scientifically interesting since they are more reflective of a model's accuracy and ability to predict and at locations it has never seen before (locations not used to build the model).

Additional Related Questions

There are additional related questions that could be answered with much of our existing data related to the ACT cohort and exposure assessment. A comparison of the ACT cohort characteristics against those of the greater Seattle area and older US population as a whole would be valuable. The demographic composition of the cohort was reported for the initial cohort, but as far as we know has not been formally published for the most recent cohort (Kukull et al.,

2002). Characterizing the air pollution exposure levels of this cohort relative to other populations would also be valuable.

Furthermore, the EPA has noted that a major limitation of integrating and assessing the results from different UFP studies has to do with the different size fractions that are collected by different studies (US EPA, 2019e). While UFPs are defined as particles under 100 nm in diameter, studies report total UFP count based on instruments with different particle size ranges, many of which measure particles larger than 100 nm and do not go below 10 nm. Our 2019 mobile monitoring campaign gives us a unique opportunity to explore this question. This campaign collected UFPs using different instruments with various particle size ranges over an extended area and period of time (Table 2). A first order question is to investigate whether the predicted exposure surfaces resulting from different instruments differ. Since most particles by count are in the smaller size ranges, instruments that measure larger particle sizes may still produce similar total counts. Instruments with lower measurement ranges, however, may miss very high particle concentration levels near major sources. A second question is to investigate where instruments disagree since different sources may produce different particle sizes within the UFP size range. Some studies, for example, have found that higher levels of very small UFPs are found near airports, suggesting that the UFP size ranges may differ over space (Austin et al., 2021). This is of scientific interest since smaller particles can target larger areas of the body and have been associated with greater neurotoxicity (Kilian & Kitazawa, 2018).

There is an opportunity to for us to characterize traffic pollutant decay curves with the multi-pollutant data collected form our mobile monitoring campaign. A notable study in this field published over a decade ago accomplished this by synthesizing the findings of different studies which were conducted using various protocols in different cities and time periods (Karner et al., 2010). Our mobile monitoring data positions us to investigate this question in more consistent manner. Because all of our data were collected near the center or edge of the road, however, this analysis would be limited. Still, there is the potential to characterize the immediate on-road decay curves along many different types of roads and land use settings, which is less well documented. Data from regulatory monitoring sites where mobile monitoring collocations were conducted could additionally be incorporated to characterize the concentrations of some pollutants away from the roadway and at distances more comparable to participant homes. Furthermore, data collected from small, low-traffic roads adjacent to large roads could

potentially be treated as “off-road” data. This would further shed light on whether measurements collected in the middle and edge of the roadway as well as near participant homes are similar enough or at least maintain the same rank order. If true, this could make it unnecessary for mobile monitoring campaigns to collect stationary measurements away from the center of the road, which is reflective of the approach that most campaigns have already taken since it is more efficient.

A Future Mobile Monitoring Campaign

An important limitation of Aim 3 where we were interested in characterizing long-term traffic pollution trends was the limited temporal information that was available. Repeating our campaign in the near future could better inform us about the region’s spatiotemporal trends in traffic pollution. We would want to follow the same sampling protocols and complete the same sampling routes in order the capture the within site variability over time and characterize the site features responsible for different degrees of variation. For example, do traffic pollution levels change more drastically over time at sites near major roads compared to sites away from roads? If so, is this change meaningful and should this be considered in long-term exposure models? Given sufficient funds, it would be valuable to conduct continuous sampling, especially of UFPs for which we otherwise have very limited information on, to better characterize diurnal, weekly and seasonal concentration patterns of individual pollutants and pollutant mixtures. Fixed site sampling locations should be representative of a variety of locations including urban and rural, industrial and residential, etc. This could further help us understand whether a single regulatory monitoring site could adequately characterize temporal changes across all sites and be useful enough to make temporal adjustments to site samples, like others have done for some pollutants (Hankey & Marshall, 2015b).

We would also want to expand sampling to include participant locations along sidewalks, front door entry ways and inside homes. Personal UFP samplers (e.g., Testo’s DiSCmini) could also be used to collect personal air samples. This would help us understand how well ambient air concentrations of various pollutants, including UFPs, reflect personal exposure. We would want to include different types of dwellings (e.g., single family, apartment, high-rise buildings) and collect home residence information to characterize lifestyle and behavioral patterns associated

with air pollution exposure. For example, major indoor pollutant sources (e.g., wood/gas stoves), home ventilation, whether filtration systems are used, etc.

While we have a good understanding that older populations spend the majority of their time indoors, it would be useful to confirm this by collecting time activity pattern information from a subset of participants. A growing area of concern, for example, is wildfire smoke exposure, which can expose large populations to elevated air pollution levels for extended periods of time. Given that air quality during wildfire seasons is expected to worsen in the coming decades, it would be valuable to understand how older, vulnerable populations like the ACT cohort may change their activity patterns during these times and as a result, be impacted differently. We believe, for example, that individuals are likely to reduce their outdoor activities and use indoor air filtration systems. This will certainly reduce personal exposure, but also the accuracy of our ambient air exposure predictions models which do not yet incorporate these increasingly important factors.

Taking all of this into account, gathering additional information through a follow-up campaign could be incredibly valuable as it could inform us about both the fine-scale spatial and long-term temporal patterns of traffic pollution, both of which are limited in the literature.

Epidemiologic Application

Finally, an important step is to use the existing models from Aims 2 and 3 of this work as well as those resulting from the proposed work above to better understand the role of traffic pollution on dementia incidence in the ACT cohort. Few epidemiologic studies, for example, have investigated short- or long-term UFP exposure. Our work positions us to advance this understanding especially well since we have one of the largest annual average exposure surfaces for UFPs, and the means to develop similar exposure surfaces for other traffic pollutants, including pollutant mixtures, an increasing area of interest. This work would be especially valuable in understanding whether the health inferences that we draw are meaningfully different when different exposure assessment approaches and models are applied. This is particularly relevant since the design of our 2019 mobile monitoring campaign was more resource and analytically intensive than many of the existing mobile monitoring campaigns conducted to date. Having the ability to draw similar conclusions from much simpler campaigns and analytical

approaches, for example, could facilitate future efforts focused on characterizing other poorly monitored pollutants.

Appendices

Appendix A: Aim 1

Methods

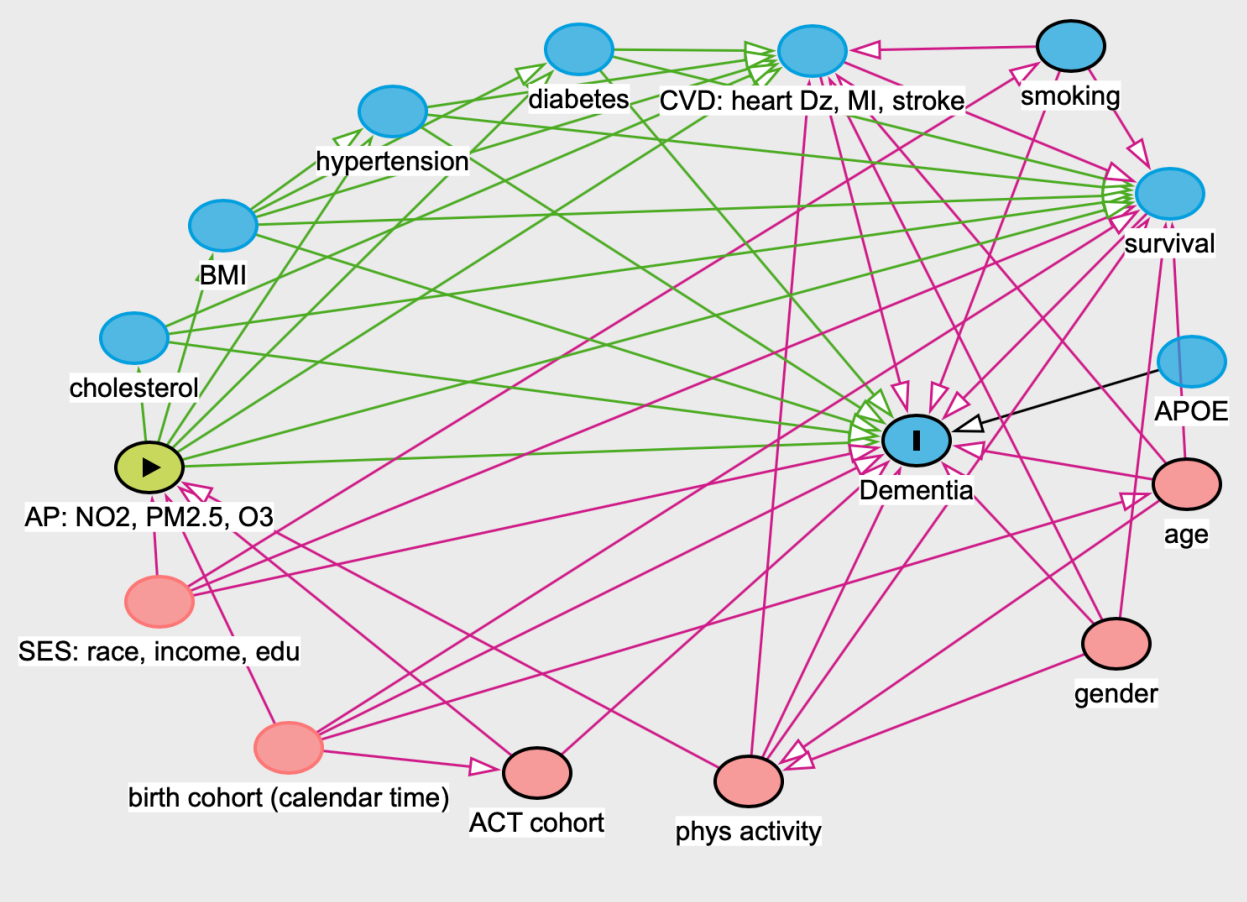


Figure 57. Directed acyclic graph (DAG) of the association between TRAP and dementia incidence with potential confounders and precision variables.

Table 17. Annual average NO2 (ppb) exposure predictions for person-years (N) in the primary analysis

Years	N	Min	Q05	Mean	SD	Median	IQR	Q95	Max
All	40,957	3	9	14	4	14	5	21	42
[1994,2000]	12,134	5	12	16	3	16	3	22	39
(2000,2006]	10,762	5	12	16	3	15	4	23	42
(2006,2012]	9,999	4	9	13	3	12	3	19	31
(2012,2018]	8,062	3	8	11	3	10	2	15	23

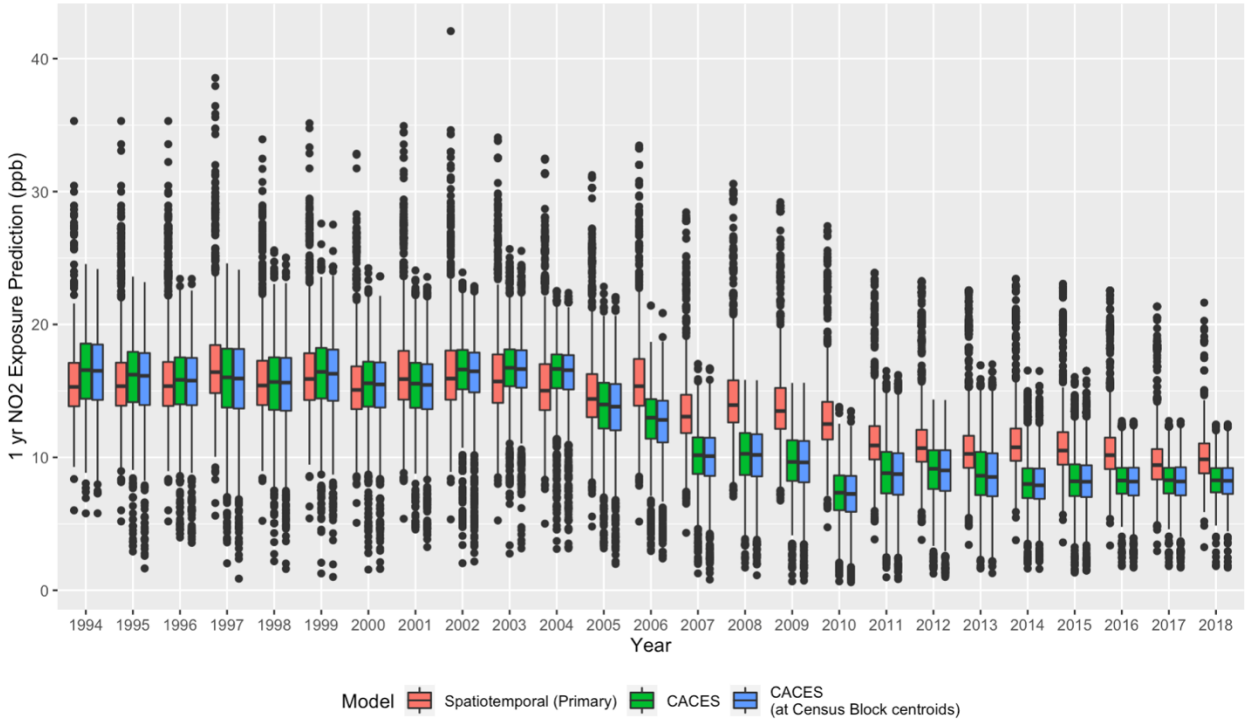


Figure 58. Comparison of 1-year NO_2 exposure predictions for the ACT cohort from different models.

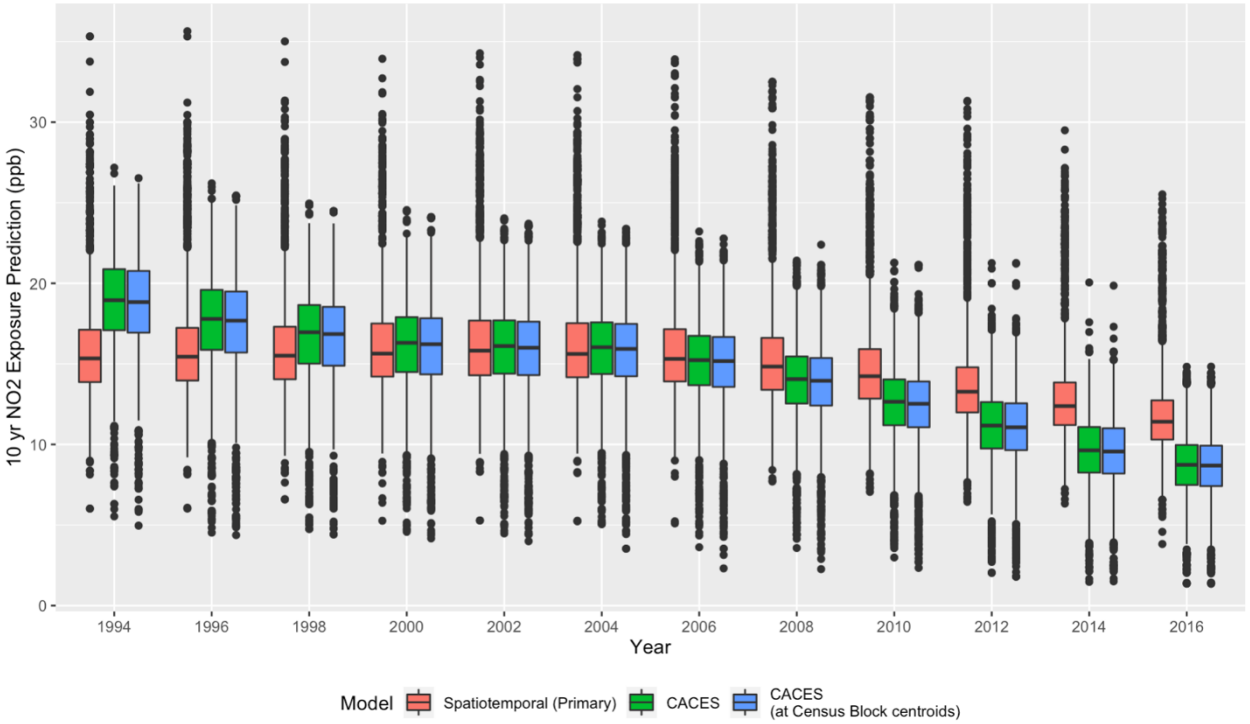


Figure 59. Comparison of 10-year exposure predictions for the ACT cohort from different models.

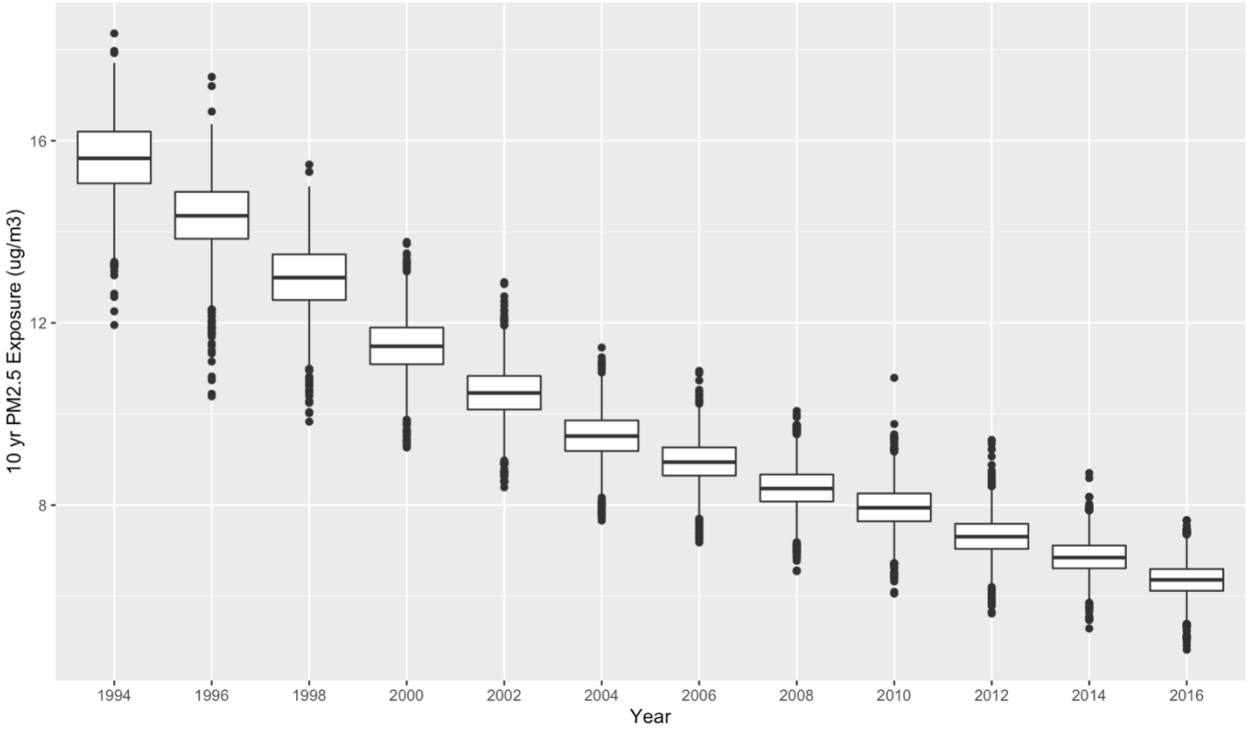


Figure 60. $PM_{2.5}$ exposure predictions for the ACT cohort over time

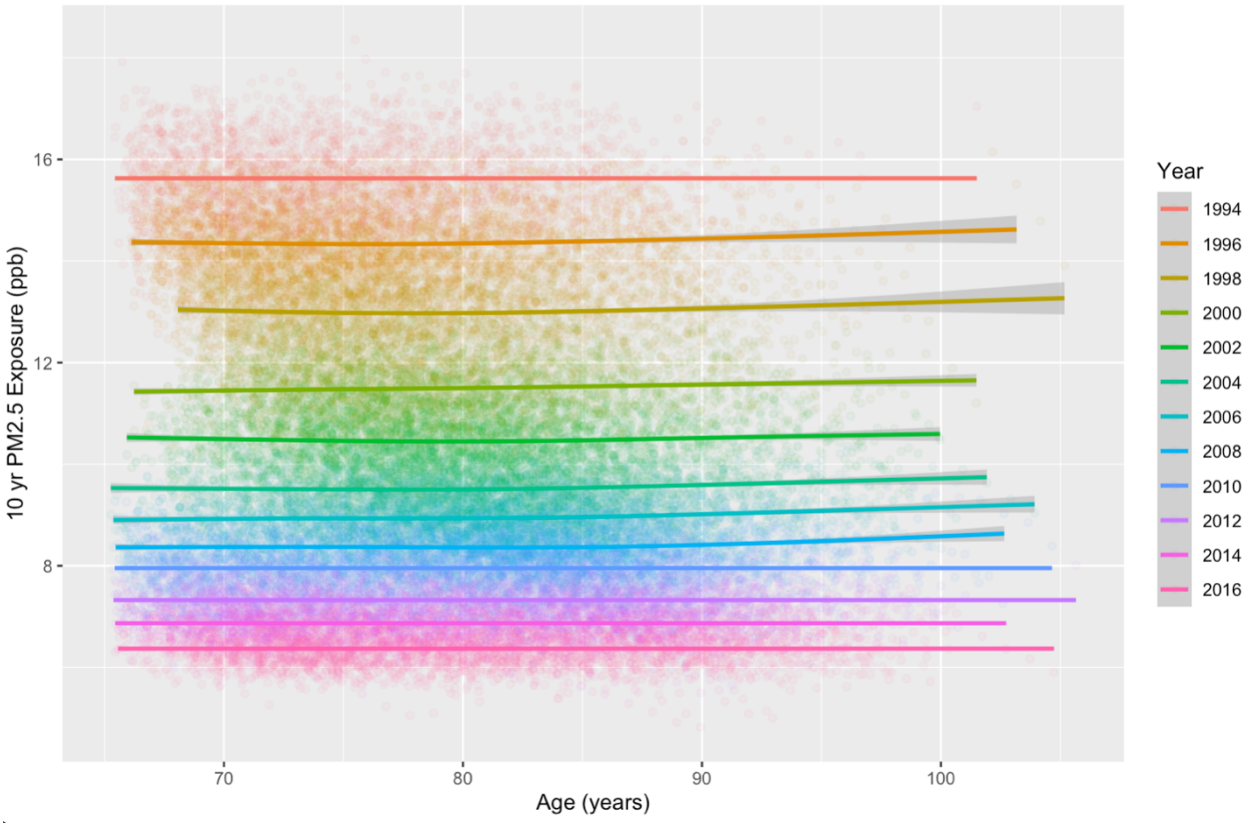


Figure 61. $PM_{2.5}$ exposure predictions for the ACT cohort by participant age and calendar year bin

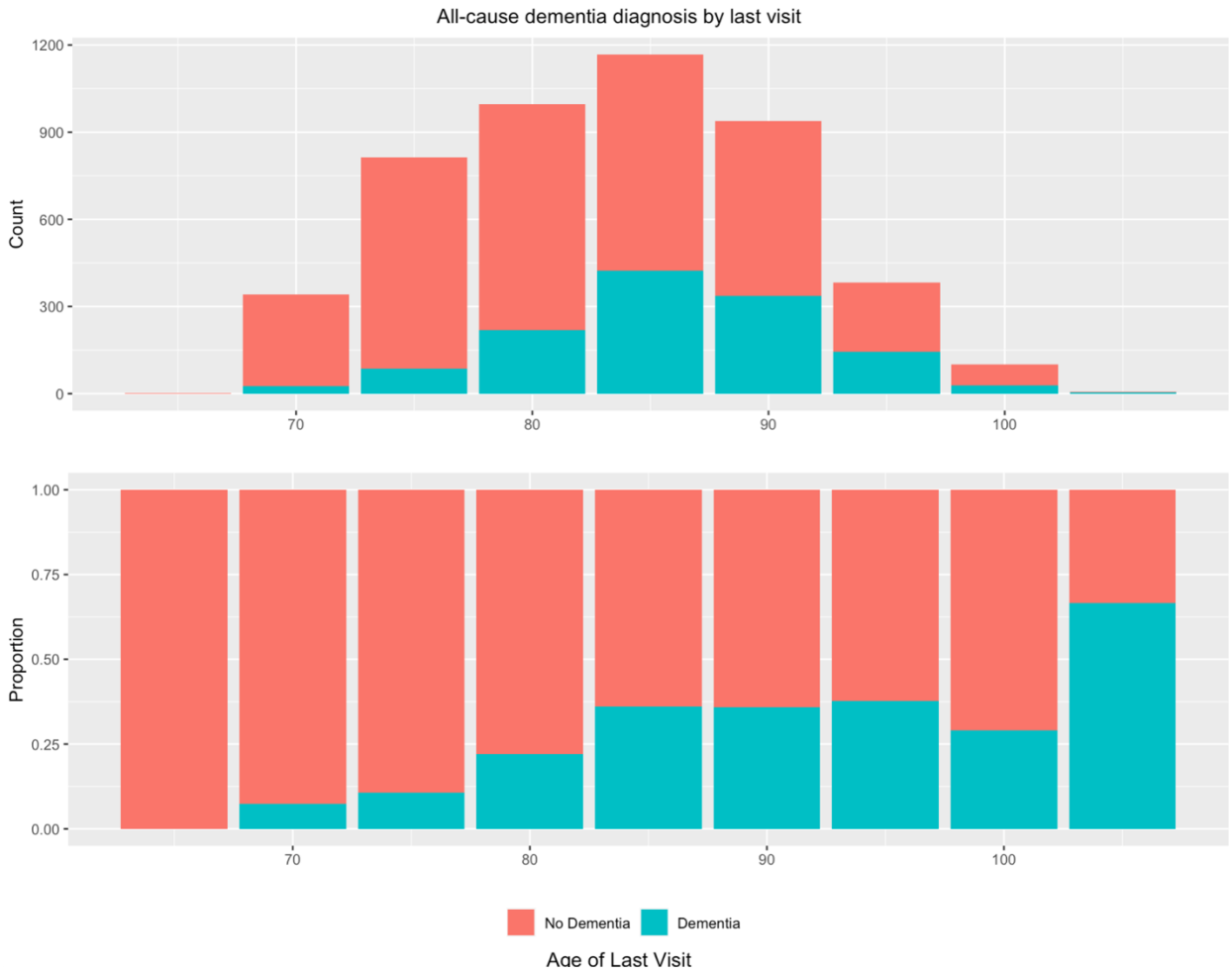
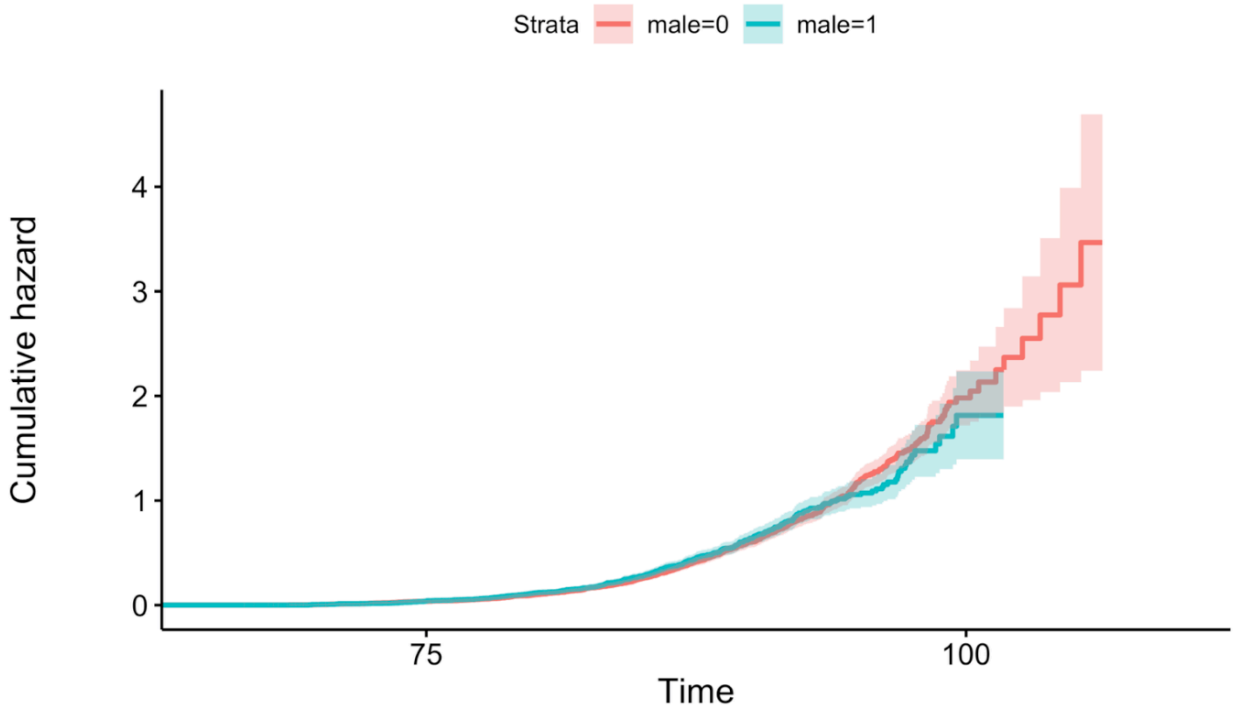
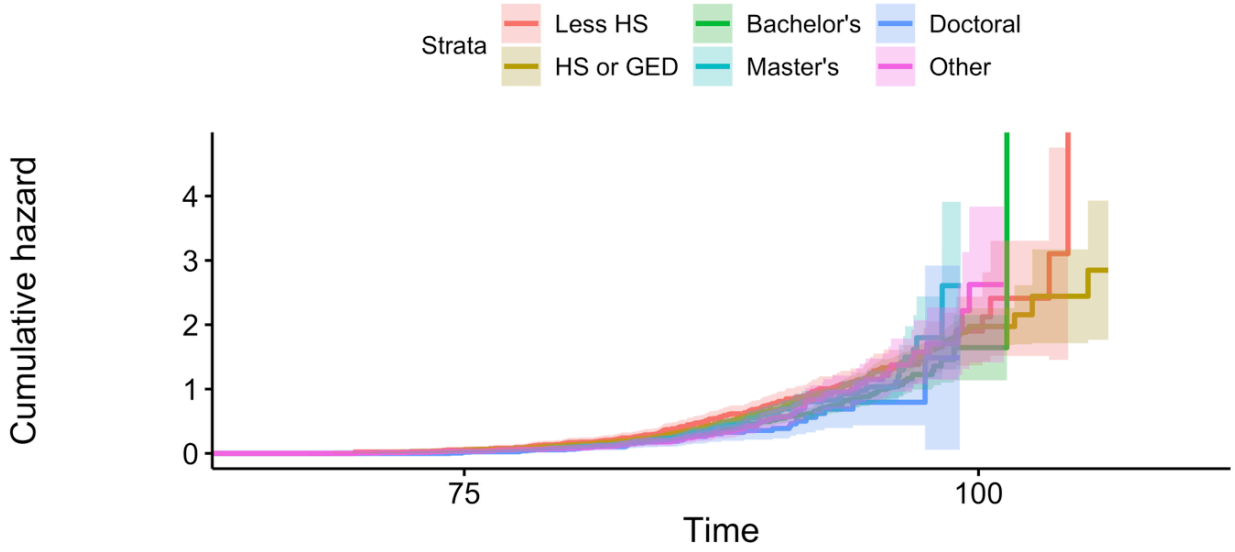


Figure 62. All-cause dementia diagnosis by age at last visit (non-cases are lost to follow-up, while cases are no longer followed)

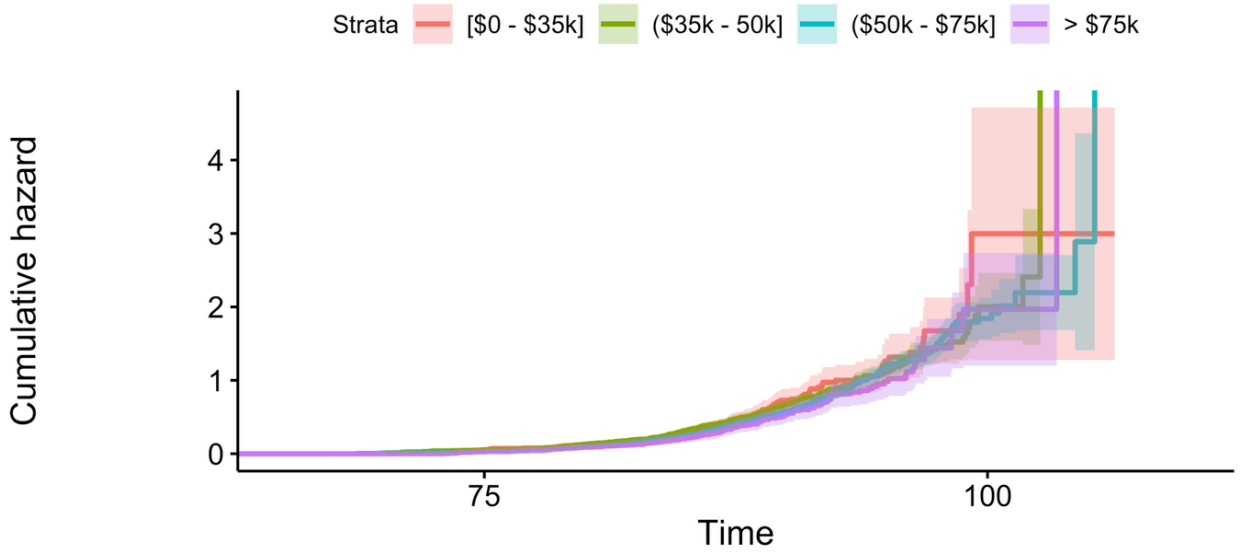
a) gender



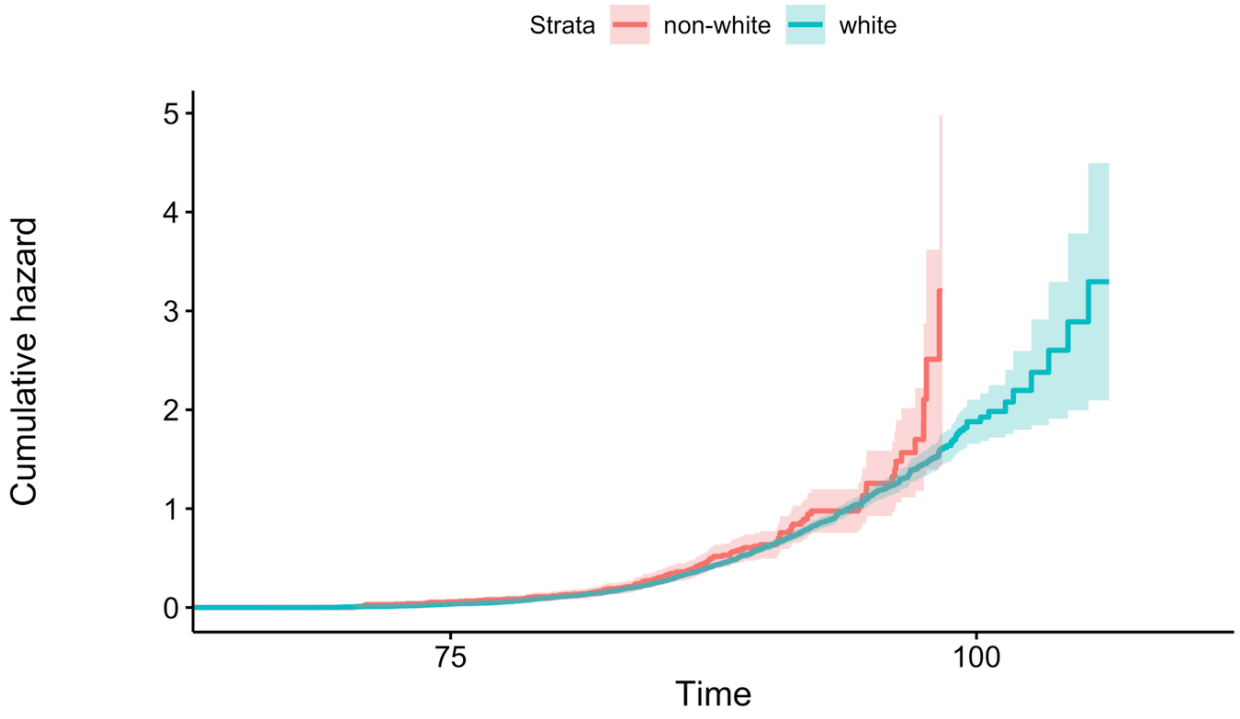
b) education at baseline



c) Census tract annual income at baseline



d) race



e) calendar year bin

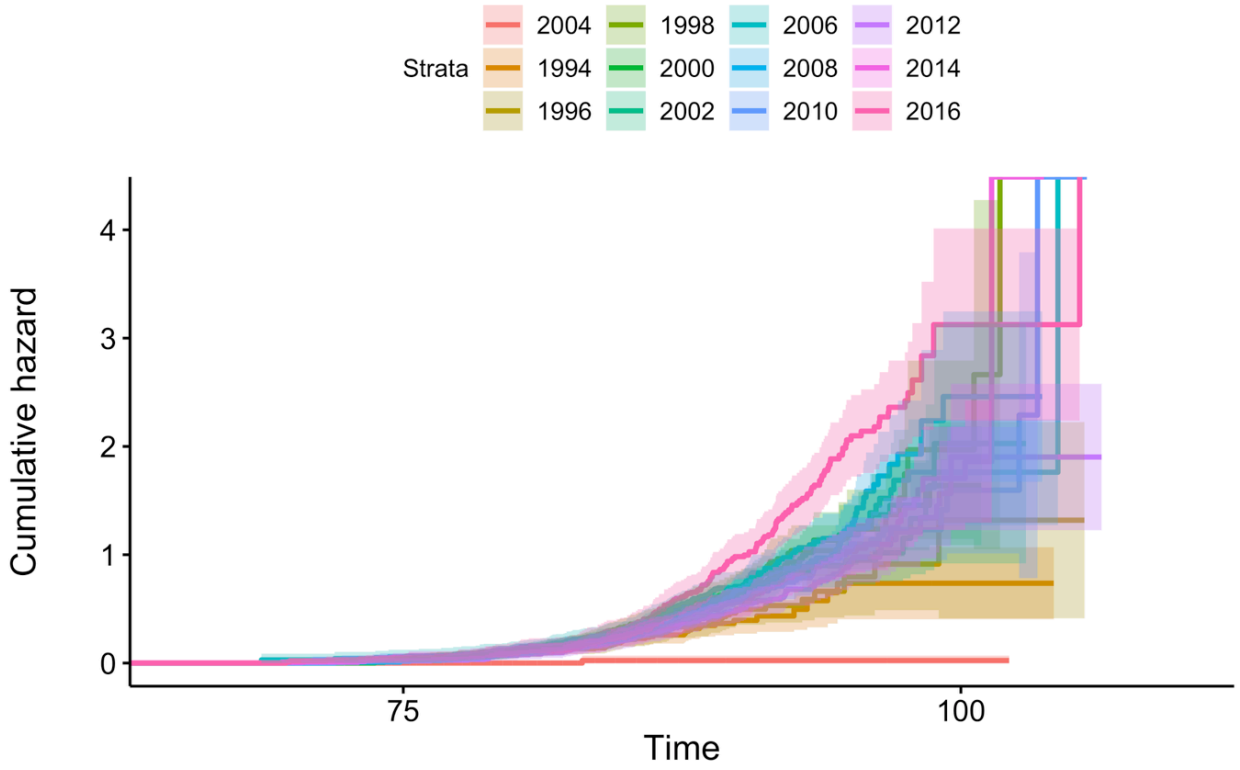


Figure 63. Plots of the cumulative hazard curves with confidence intervals of all-cause dementia for the primary model's confounder subcategories: a) gender, b) education at baseline, c) Census tract annual income at baseline, d) race, and e) 2-year calendar bin. Nonproportional hazard lines in the confounder subcategories indicate that the Cox model proportional hazards assumption may be violated.

Table 18. NO_2 (5 ppb) hazard ratios for all-cause dementia and AD from primary and sensitivity analyses

Outcome	Model Covariates	Type	Description	HR (95% CI)	N	Events
All-Cause Dementia	2	Primary	(M2) Primary Analysis	1.01 (0.91, 1.11)	41331	1136
All-Cause Dementia	1	Confounding & Mediation	(M1) Reduced	0.98 (0.9, 1.07)	41331	1136

All-Cause Dementia	3a	Confounding & Mediation	(M3a) Extended	1.01 (0.91, 1.11)	41331	1136
All-Cause Dementia	3b	Confounding & Mediation	(M3b) Extended, ACT Cohort	1.02 (0.92, 1.12)	41331	1136
All-Cause Dementia	4	Confounding & Mediation	(M4) Extended & Mediation	1.00 (0.91, 1.11)	41331	1136
All-Cause Dementia	5	Confounding & Mediation	(M5) PM2.5 Adjusted	0.94 (0.84, 1.06)	41324	1136
All-Cause Dementia	6	Confounding & Mediation	(M6) More Confounder Stratification	1.00 (0.90, 1.11)	41331	1136
All-Cause Dementia	2	Time Variants	Using Onset (vs Diagnosis) Age	0.99 (0.90, 1.09)	40007	1132
All-Cause Dementia	2	Time Variants	Cal Time Axis w/ 2 Yr Birth Cohort Adj	1.01 (0.91, 1.11)	41331	1136
All-Cause Dementia	2	Time Variants	5 Yr Birth Cohort (vs 2 Yr Cal Time)	1.01 (0.92, 1.11)	41331	1136
All-Cause Dementia	2	Time Variants	2 Yr Birth Cohort (vs 2 Yr Cal Time)	1.02 (0.93, 1.12)	41331	1136
All-Cause Dementia	2	Exposure Quality	Highest Quality Address Hx	0.99 (0.89, 1.11)	28276	941

All-Cause Dementia	2	Exposure Quality	High Quality Address Hx	1.00 (0.90, 1.11)	37895	1053
All-Cause Dementia	2	Exposure Quality	No model extrapolation (2005+)	1.04 (0.92, 1.18)	21672	733
All-Cause Dementia	2	Other	No IPW	1.01 (0.91, 1.11)	41331	1136
AD	2	Primary	(M2) Primary Analysis	1.02 (0.91, 1.13)	41331	920
AD	1	Confounding & Mediation	(M1) Reduced	0.95 (0.87, 1.05)	41331	920
AD	3a	Confounding & Mediation	(M3a) Extended	1.02 (0.91, 1.13)	41331	920
AD	3b	Confounding & Mediation	(M3b) Extended, ACT Cohort	1.02 (0.92, 1.14)	41331	920
AD	4	Confounding & Mediation	(M4) Extended & Mediation	1.02 (0.91, 1.13)	41331	920
AD	5	Confounding & Mediation	(M5) PM2.5 Adjusted	0.99 (0.87, 1.13)	41324	920
AD	6	Confounding & Mediation	(M6) More Confounder Stratification	1.03 (0.92, 1.15)	41331	920
AD	2	Time Variants	Using Onset (vs Diagnosis) Age	1.01 (0.90, 1.12)	40007	916

AD	2	Time Variants	Cal Time Axis w/ 2 Yr Birth Cohort Adj	1.02 (0.91, 1.14)	41331	920
AD	2	Time Variants	5 Yr Birth Cohort (vs 2 Yr Cal Time)	1.02 (0.92, 1.14)	41331	920
AD	2	Time Variants	2 Yr Birth Cohort (vs 2 Yr Cal Time)	1.04 (0.94, 1.16)	41331	920
AD	2	Exposure Quality	Highest Quality Address Hx	0.99 (0.88, 1.12)	28276	777
AD	2	Exposure Quality	High Quality Address Hx	1.00 (0.89, 1.12)	37895	853
AD	2	Exposure Quality	No model extrapolation (2005+)	1.04 (0.92, 1.19)	21672	635
AD	2	Other	No IPW	1.02 (0.92, 1.14)	41331	920

Table 19. NO_2 (5 ppb) hazard ratios for all-cause dementia and AD from secondary analyses

Outcome	Type	Description	HR (95% CI)	N	Events
All-Cause Dementia	Exposure Quality	Fixed Exposure at Entry	0.99 (0.90, 1.09)	4163	1137
All-Cause Dementia	Exposure Period	1 yr Exposure	1.03 (0.94, 1.13)	40957	1113

All-Cause Dementia	Exposure Period	5 yr Exposure	1.01 (0.92, 1.11)	41211	1126
All-Cause Dementia	Exposure Period	20 yr Exposure	1.01 (0.92, 1.12)	41371	1138
All-Cause Dementia	Exposure Period	10 yr 5 yr Lag Exposure	1.01 (0.92, 1.11)	41224	1135
All-Cause Dementia	Exposure Period	10 yr 10 yr Lag Exposure	1.00 (0.91, 1.10)	41320	1138
All-Cause Dementia	Major Road Exposure	Major Road Proximity: Near (vs Far)	0.94 (0.83, 1.07)	41376	1138
All-Cause Dementia	Major Road Exposure	Major Road Proximity: Midway (vs Far)	1.42 (0.85, 2.38)	41376	1138
All-Cause Dementia	Major Road Exposure	Time Near Road (10 yr)	0.96 (0.83, 1.10)	41376	1138
All-Cause Dementia	APOE Interaction	Interaction, APOE: No	1.03 (0.92, 1.14)	41331	1136
All-Cause Dementia	APOE Interaction	Interaction, APOE: Yes	0.96 (0.82, 1.12)	41331	1136
All-Cause Dementia	Race Interaction	Interaction, Race: White	1.01 (0.92, 1.12)	41331	1136
All-Cause Dementia	Race Interaction	Interaction, Race: Non-White	0.96 (0.73, 1.28)	41331	1136
All-Cause Dementia	Gender Interaction	Interaction, Gender: Male	1.11 (0.97, 1.28)	41331	1136

All-Cause Dementia	Gender Interaction	Interaction, Gender: Female	0.95 (0.85, 1.07)	41331	1136
All-Cause Dementia	CACES	Residence Level Exposure	1.08 (0.94, 1.23)	41037	1123
All-Cause Dementia	CACES	Block Level Exposure	1.05 (0.93, 1.19)	41162	1131
AD	Exposure Quality	Fixed Exposure at Entry	0.99 (0.88, 1.10)	4163	920
AD	Exposure Period	1 yr Exposure	1.04 (0.94, 1.16)	40957	903
AD	Exposure Period	5 yr Exposure	1.02 (0.92, 1.14)	41211	913
AD	Exposure Period	20 yr Exposure	1.02 (0.92, 1.14)	41371	921
AD	Exposure Period	10 yr 5 yr Lag Exposure	1.02 (0.92, 1.13)	41224	921
AD	Exposure Period	10 yr 10 yr Lag Exposure	1.01 (0.91, 1.12)	41320	921
AD	Major Road Exposure	Major Road Proximity: Near (vs Far)	0.93 (0.81, 1.08)	41376	921
AD	Major Road Exposure	Major Road Proximity: Midway (vs Far)	1.49 (0.88, 2.53)	41376	921
AD	Major Road Exposure	Time Near Road (10 yr)	0.95 (0.81, 1.11)	41376	921
AD	APOE Interaction	Interaction, APOE: No	1.05 (0.93, 1.18)	41331	920
AD	APOE Interaction	Interaction, APOE: Yes	0.94 (0.79, 1.12)	41331	920
AD	Race Interaction	Interaction, Race: White	1.02 (0.92, 1.15)	41331	920

AD	Race Interaction	Interaction, Race: Non-White	0.95 (0.70, 1.29)	41331	920
AD	Gender Interaction	Interaction, Gender: Male	1.14 (0.97, 1.33)	41331	920
AD	Gender Interaction	Interaction, Gender: Female	0.96 (0.85, 1.10)	41331	920
AD	CACES	Residence Level Exposure	1.00 (0.87, 1.16)	41037	909
AD	CACES	Block Level Exposure	0.99 (0.86, 1.13)	41162	915

Table 20. Competing risk analyses using cause-specific Cox proportional hazard models. Table shows hazard ratios for death before dementia or AD for a 5 ppb increase in 10-year average NO_2 exposure. The terminating event of interest (death before dementia) was censored at last study visit.

Outcome	Description	HR (95% CI)	N	Events
All-Cause Dementia	(M2) Primary	0.98 (0.89, 1.09)	41331	1086
All-Cause Dementia	(M5) PM2.5 Co-pollutant	1.00 (0.90, 1.12)	41324	1086
All-Cause Dementia	(M6) More stratification	0.99 (0.90, 1.10)	41331	1086
AD	(M2) Primary	0.98 (0.88, 1.08)	41331	1160
AD	(M5) PM2.5 Co-pollutant	0.98 (0.88, 1.10)	41324	1160
AD	(M6) More stratification	0.98 (0.89, 1.08)	41331	1160

Table 21. Estimated dementia HRs for a 5 ppb increase in NO_2 or NO_x from different studies, including this one. Some hazard ratios were converted to volume concentrations (ppb) from mass concentrations ($\mu g/m^3$) assuming a 1.88 $\mu g/m^3$ per 1 ppb conversion factor. Oudin et al.'s 2016 HR is for NOx. The HRs for all other studies are for NO_2.

Author Year	Location	N	Avg Exposure (ppb)[1]	Study HR	Study Units	HR for 5 ppb NO_2
Blanco et al. 2020	Seattle, USA	4,150	17.0	1.01 (0.91, 1.11)	5 ppb	1.01 (0.91, 1.11)
Carey et al. 2018	London, UK	130,978	19.0	1.16 (1.05, 1.28)	7.5 ug/m^3	1.20 (1.06, 1.36)
Cerza et al. 2019	Rome, Italy	350,844	23.4	0.97 (0.96, 0.99)	10 ug/m^3	0.97 (0.96, 0.99)
Chen, Kwong, Copes, Hystad, et al. 2017	Ontario, Canada	2,100,000	16.0	1.10 (1.08, 1.12)	14.2 ppb	1.03 (1.03, 1.04)
Chen, Kwong, Copes, Tu, et al. 2017	Ontario, Canada	243,611	15.0	1.07 (1.06, 1.08)	11.3 ppb	1.03 (1.03, 1.03)
Ilango et al. 2019	Ontario, Canada	34,391	10.4	1.10 (0.99, 1.19)	5 ppb	1.10 (0.99, 1.19)
Oudin et al. 2016	Umea, Sweden	1,806	9.0	1.05 (0.98, 1.12)	10 ug/m^3	1.05 (0.98, 1.11)

[1]Or median exposure if the average was not available; in ppb

Appendix B: Aim 2

Methods

Table 22. Geocovariate predictors available for modeling

No.	Kind	Covariate Label	Available Buffers	Description
1	airports	log_m_to_airp		log meters to closest airport
2	airports	log_m_to_l_airp		log meters to closest large airport
3	coast	log_m_to_coast		log meters to closest coastline
4	commercial and services	log_m_to_comm		log meters to closest commercial and services area
5	elevation	elev_above	1000, 5000	number of points (out of 24) more than 20 m and 50 m uphill of a location for a 1000 m and 5000 m buffer, respectively
6	elevation	elev_at_elev	1000, 5000	number of points (out of 24) within 20 m and 50 m of the location' elevation for a 1000 m and 5000 m buffer, respectively
7	elevation	elev_below	1000, 5000	number of points (out of 24) more than 20 m and 50 m downhill of a location for a 1000 m and 5000 m buffer, respectively
8	elevation	elev_elevation		elevation above sea level in meters
9	imperviousness	imp_a	50, 100, 150, 300, 400, 500, 750, 1000, 3000, 5000	average imperviousness
10	land use	rlu_barren_p	1000	proportion of barren land
11	land use	rlu_decid_forest_p	300, 400, 500, 750, 1000	proportion of deciduous forest

12	land use	rlu_dev_hi_p	100, 150, 300, 400, 500, 750, 1000, 3000, 5000	proportion of highly developed land (e.g., commercial and services; industrial; transportation, communication and utilities)
13	land use	rlu_dev_lo_p	50, 100, 150, 300, 400, 500, 750, 1000, 3000, 5000	proportion of low developed land (e.g., residential)
14	land use	rlu_dev_med_p	50, 100, 150, 300, 400, 500, 750, 1000, 3000, 5000	proportion of medium developed land (e.g., residential)
15	land use	rlu_dev_open_p	50, 100, 150, 300, 400, 500, 750, 1000, 3000, 5000	proportion of developed open land
16	land use	rlu_evergreen_p	150, 300, 400, 500, 750, 1000, 3000, 5000	proportion of evergreen forest
17	land use	rlu_herb_wetland_p	750, 1000	proportion of herb (nonforested) wetland
18	land use	rlu_mix_forest_p	150, 300, 400, 500, 750, 1000, 3000, 5000	proportion of mixed forest
19	land use	rlu_woody_wetland_p	750, 1000	proportion of woody wetland
20	NDVI	ndvi_q25_a	250, 500, 1000, 2500, 5000, 7500, 10000	NDVI (25th quantile)
21	NDVI	ndvi_q50_a	250, 500, 1000, 2500, 5000, 7500, 10000	NDVI (50th quantile)

22	NDVI	ndvi_q75_a	250, 500, 1000, 2500, 5000, 7500, 10000	NDVI (75th quantile)
23	NDVI	ndvi_summer_a	250, 500, 1000, 2500, 5000, 7500, 10000	average summer time NDVI
24	NDVI	ndvi_winter_a	250, 500, 1000, 2500, 5000, 7500, 10000	average winter time NDVI
25	population	pop_s	500, 1000, 1500, 2000, 2500, 3000, 5000, 10000, 15000	2000 population density
26	port	log_m_to_l_port		log meters to closest large port
27	port	log_m_to_m_port		log meters to closest medium port
28	railroads, rail yards	log_m_to_rr		log meters to closest railroad
29	railroads, rail yards	log_m_to_ry		log meters to closest rail yard
30	roads	intersect_a1_a3_s	1000, 3000	number of a1-a3 road intersections
31	roads	intersect_a2_a2_s	3000	number of a2-a2 road intersections
32	roads	intersect_a2_a3_s	3000	number of a2-a3 road intersections
33	roads	intersect_a3_a3_s	500, 1000, 3000	number of a3-a3 road intersections
34	roads	ll_a1_s	500, 750, 1000, 1500, 3000, 5000	length of a1 roads
35	roads	ll_a23_s	50, 100, 150, 300, 400, 500, 750, 1000, 1500, 3000, 5000	length of a2 and a3 roads
36	roads	log_m_to_a1		log meters to closest a1 road
37	roads	log_m_to_a1_a1_intersect		log meters to closest a1-a1 road intersection
38	roads	log_m_to_a1_a2_intersect		log_m_to_a1_a2_intersect

39	roads	log_m_to_a1_a3_intersect		log meters to closest a1-a3 road intersection
40	roads	log_m_to_a123		log meters to closest a1, a2 or a3 road
41	roads	log_m_to_a2_a2_intersect		log meters to closest a2-a2 road intersection
42	roads	log_m_to_a2_a3_intersect		log meters to closest a2-a3 road intersection
43	roads	log_m_to_a23		log meters to closest a2 or a3 road
44	roads	log_m_to_a3_a3_intersect		log meters to closest a3-a3 road intersection
45	truck routes	log_m_to_truck		log meters to closest truck route
46	truck routes	tl_s	300, 400, 500, 750, 1000, 1500, 3000, 5000, 10000, 15000	length of truck routes
47	water	rlu_water_p	500, 750, 1000, 3000, 5000	proportion of water

Results

Raw Readings

Table 23. Distribution of all, raw UFP and BC data collected.

Pollutant	N	Min	Q05	Mean	SD	Median	IQR	Q95	Max
UFP (pt/cm^3)	1,177,944	0	1,840	7,698	9,223	5,900	5,570	18,600	5e+05
BC (ng/m^3)	125,866	-11,797	-284	582	832	457	697	1,836	37,357

Data Quality Control

P-TRAKs and aethalometers recorded very low concentrations during zero checks throughout the study (UFP & BC: 5/30/2019, 10/2/2019, 1022/2019, 11/21/2019, 1/13/2020; UFP: 3/13/2020; Appendix B Table 24). Aethalometer readings were noisier and varied more than P-TRAK readings, but concentrations were generally centered around 0 ng/m^3. Note that a few aethalometer readings were very low due to the instability of the instrument at measuring short-term samples.

Table 24. P-TRAK and aethalometer readings during zero checks. The number of samples corresponds to 1- and 10-second readings for P-TRAKs and aethalometers, respectively.

Instrument	No. Days	No. Samples	Min	Q05	Mean	SD	Median	IQR	Q95	Max
P-TRAK (pt/cm^3)	6	2,738	0	0	38	150	15	25	187	3,380
Aethalometer (ng/m^3)	5	313	-1,162	-621	-1	391	0	479	629	1,304

We observed unusually low PTRAK readings at stops on four occasions (1,019 one-second readings, 0.09% of total readings; Appendix B Figure 64)**Error! Reference source not found.**. These readings were dropped since there were clear indications of instrument malfunctions such that readings precipitously dropped, remained flat and quickly increased again or remained flat. Two of these occasions (3/22/2019 and 7/13/2019) had field notes verifying that the instruments had had low alcohol wick issues.

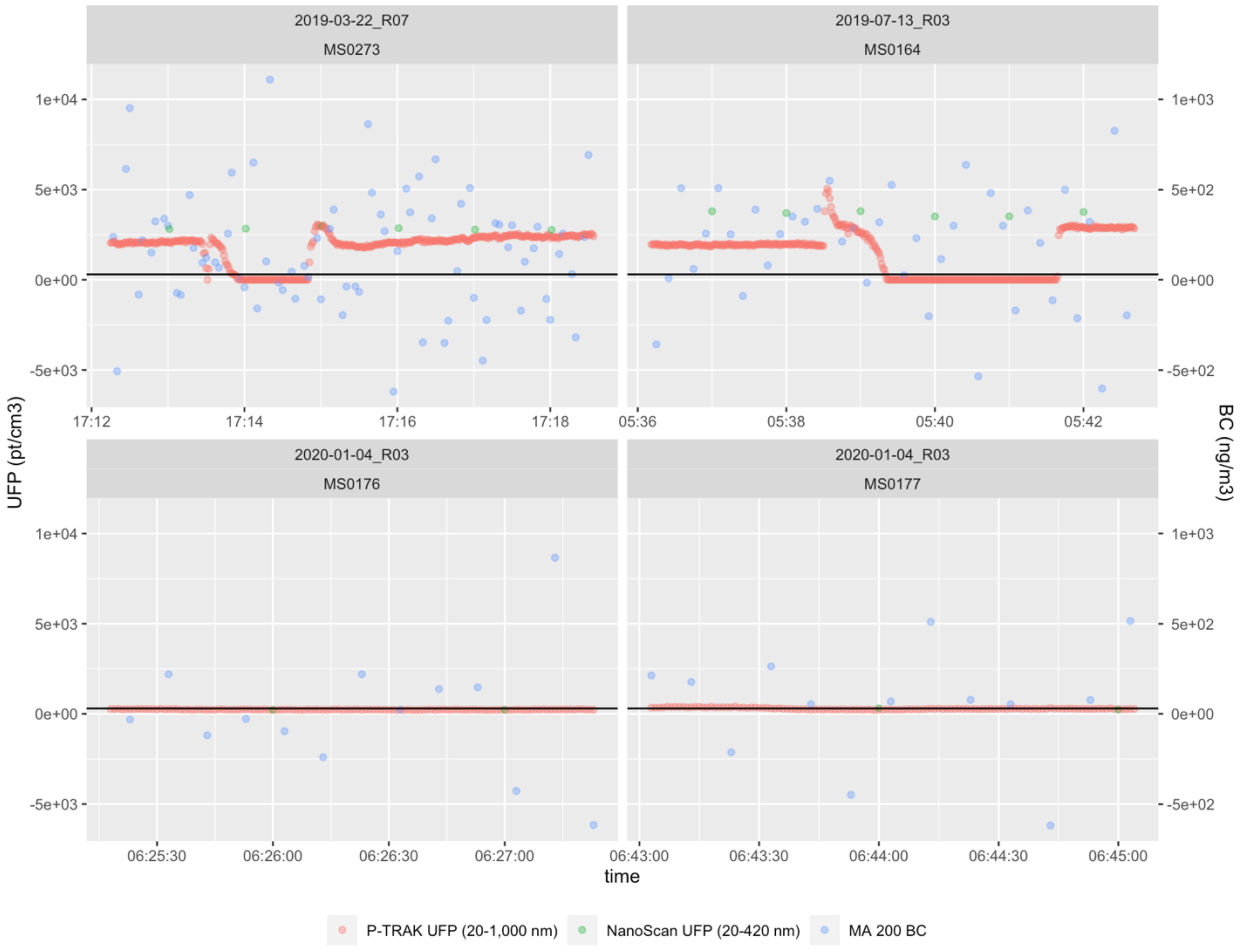


Figure 64. Time series plots for stops with low P-TRAK readings (<300 pt/cm^3) that were dropped. Note that the relevant time periods varied in duration and thus so do the x-axes.

We observed high P-TRAK readings at the instrument threshold of 500,000 pt/cm^3 at stops on two occasions (Appendix B Figure 65). These readings were kept since we also observed elevated UFP readings from other, NanoScan instruments during these times.

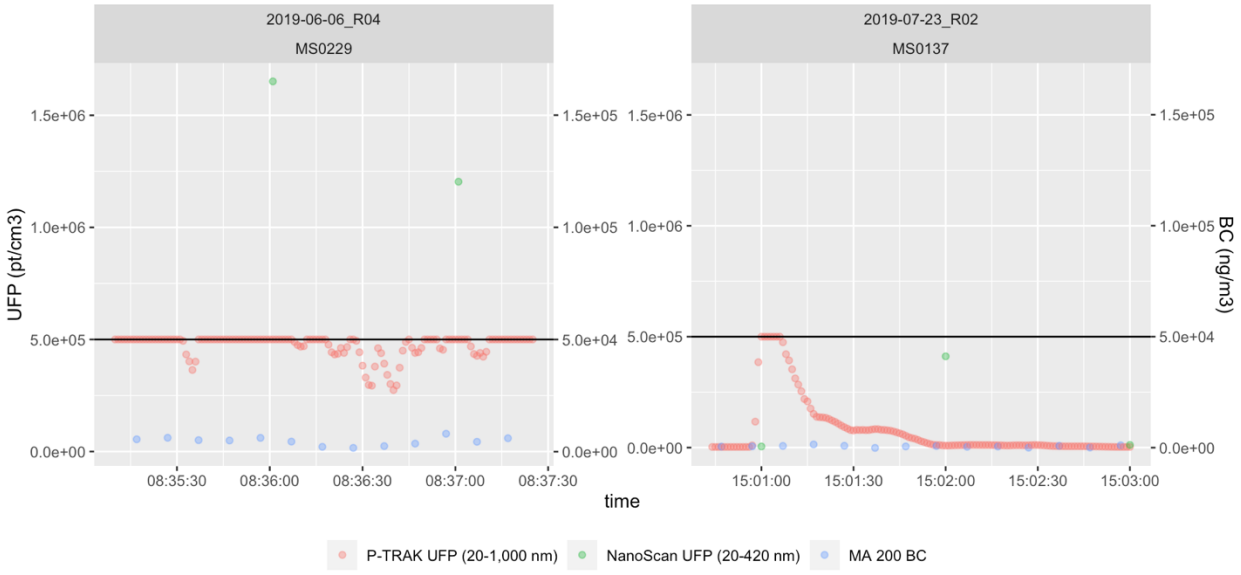


Figure 65. Time series plots of runs with P-TRAK (UFP) readings at the instrument threshold of 500,000 pt/cm^3

No BC readings reached the instrument threshold of 1 mg/m^3. We observed high BC readings above the 99th percentile of the data (~3,200 ng/m^3) on one occasion (Appendix B Figure 66). These readings were kept since we also observed elevated correlated UFP readings from P-TRAK instruments during these times.

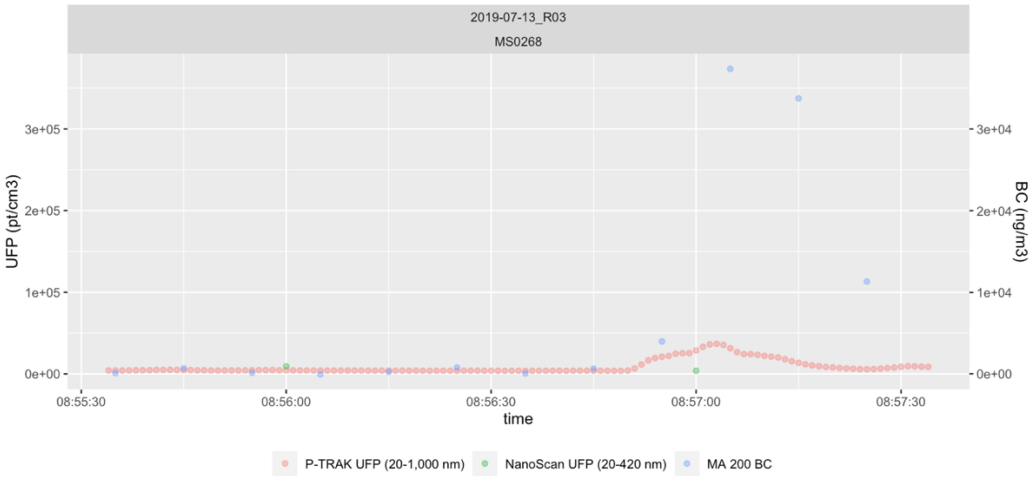


Figure 66. Time series plots of runs with aethalometer (BC) readings above the 99th percentile (~3,200 ng/m^3)

Concentrations from collocated P-TRAKS in the field (median stop concentrations, ~ 2 minutes) and lab (median 2-minute concentrations) were in good agreement (Appendix B Figure 67). The estimated RMSE and MSE-based R^2 parameters were 575 pt/cm^3 and 0.98, respectively.

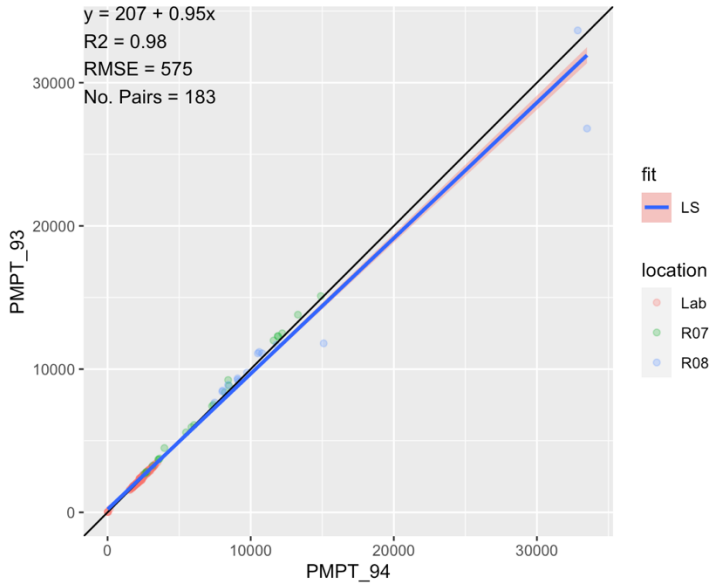


Figure 67. Comparison of collocated primary (PMPT_94) and secondary (PMPT_93) P-TRAK instruments (pt/cm^3).

Field (route: R01-08) and lab (Lab, Rm 114) collocated aethalometer instruments were in good agreement with an RMSE of 152 ng/m^3 and an R^2 of 0.91 (Appendix B Figure 68).

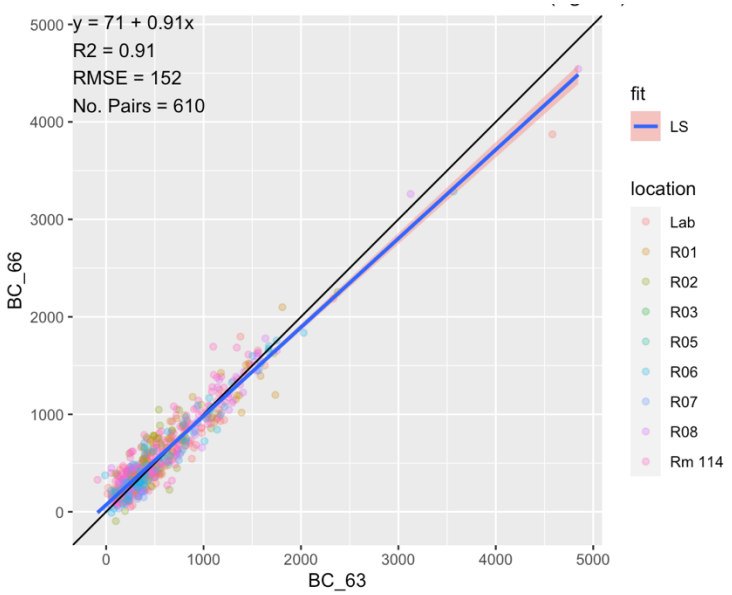


Figure 68. Comparison of collocated primary (BC_63) and secondary (BC_66) aethalometer instruments (ng/m3).

Stop Concentrations & Annual Averages

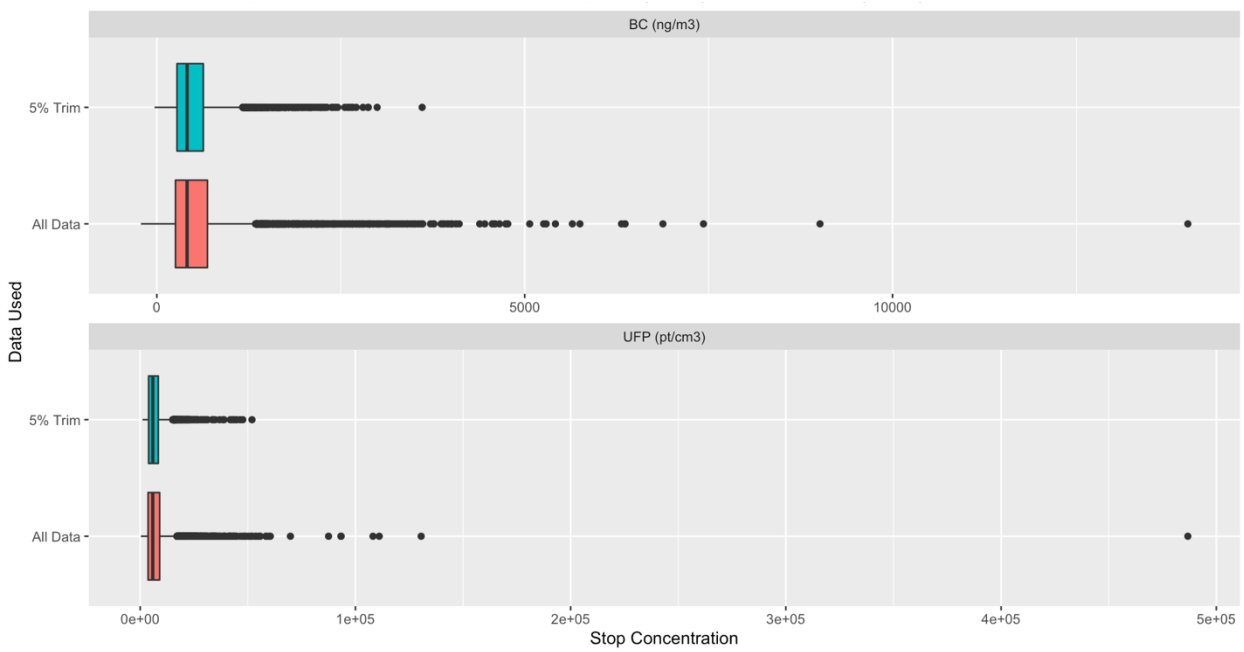


Figure 69. Distribution of UFP and BC concentrations using all stops and 5% trimmed stops

Table 25. Distribution of UFP and BC concentrations using all stops and 5% trimmed stops

Pollutant	Description	N	Min	Q05	Mean	SD	Median	IQR	Q95	Max
BC (ng/m3)										
	All data	8,931	-216	92	581	602	411	435	1,687	14,014
	5% trimmed	7,696	-31	152	507	352	411	358	1,196	3,606
UFP (pt/cm3)										
	All data	8,686	330	1,895	7,442	8,186	5,820	5,378	17,819	486,786
	5% trimmed	7,452	1,050	2,508	6,775	4,186	5,817	4,493	14,023	51,935

Table 26. ANOVA looking at the effects of site, season, day of the week and hour of the day on the variation of UFP stop median concentrations

	DF	SS	MS	VC	%Total
total	2329			17637316	100.0
season	3	411114550	137038184	63944	0.4
day	6	3221808823	536968137	491806	2.8
hour	20	1858518973	92925949	214196	1.2
site_id	308	33032865063	107249562	3931766	22.3
error	7110	91972141565	12935604	12935604	73.3

Table 27. ANOVA looking at the effects of site, season, day of the week and hour of the day on the variation of BC stop median concentrations

	DF	SS	MS	VC	%Total
total	638			126311.7	100.0
season	3	42870317	14290105.6	7369.3	5.8
day	6	32007049	5334508.1	4760.1	3.8
hour	20	19020340	951017.0	2339.1	1.9
site_id	308	141756312	460247.8	14643.8	11.6
error	7358	715192958	97199.4	97199.4	77.0

Figure 70. Annual average site concentration estimates by method (regression vs site-only data approach) and weight (s: 4 seasons; tow[#]: time of week splint into # times; tod[#]: time of day split into # hours; uw: unweighted) for sites sampled each season-, time-of-week and time of day. Only presenting sites with available estimates for all methods and weights presented (i.e., 25 sites had estimates for 2 different methods and 4 different weights).

Table 28. Variance components analysis (ANOVA) looking at the effects of site and estimation approach (method and weight) on the variation of UFP annual average concentrations

	DF	SS	MS	VC	%Total
total	314			4448060	100.0
method_weight	9	96750449	10750050	37572	0.8
site_id	308	12049662523	39122281	4399684	98.9
error	2430	26254644	10804	10804	0.2

Table 29. Variance components analysis (ANOVA) looking at the effects of site and estimation approach (method and weight) on the variation of BC annual average concentrations

	DF	SS	MS	VC	%Total
total	310.9			18288	100.0
method_weight	9	98084	10898	32	0.2
site_id	308	50074082	162578	18203	99.5
error	2442	130591	54	54	0.3

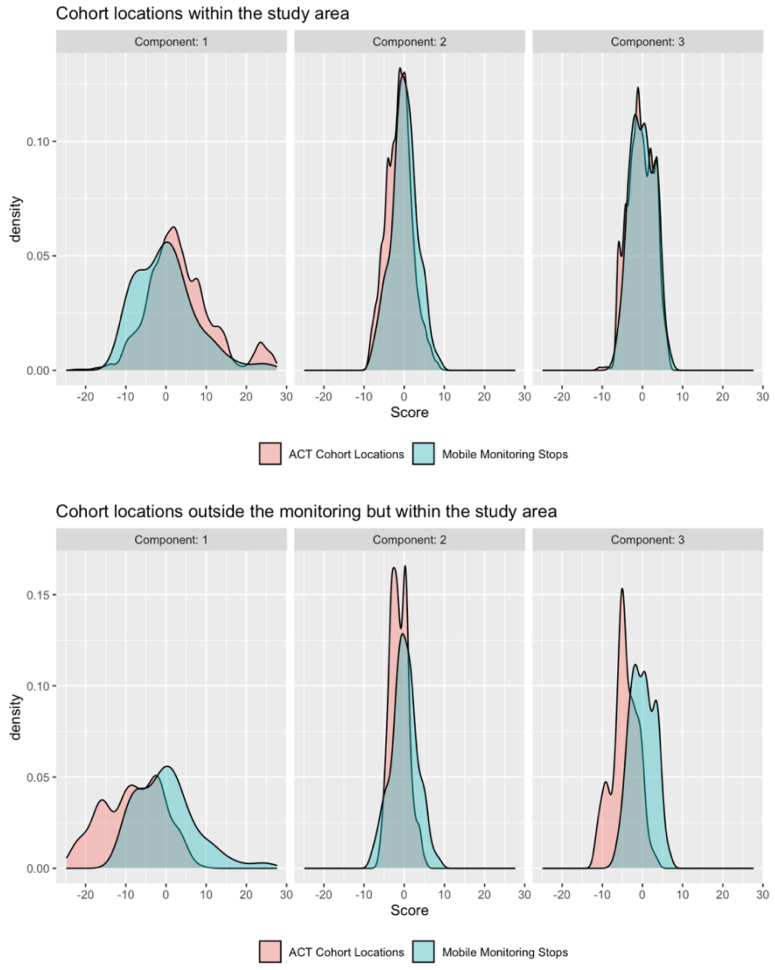
Table 30. Variables dropped prior to running PLS

Reason	N	Covariates
Common value (exclude variables with values less than 20% of being different from most common values)	110	ll_a1_s00050, ll_a1_s00100, ll_a1_s00150, ll_a1_s00300, ll_a1_s00400, rlu_water_p00050, rlu_ice_p00050, rlu_dev_open_p00050, rlu_dev_hi_p00050, rlu_barren_p00050, rlu_decid_forest_p00050, rlu_evergreen_p00050, rlu_mix_forest_p00050, rlu_shrub_p00050, rlu_grass_p00050, rlu_pasture_p00050, rlu_crop_p00050, rlu_woody_wetland_p00050, rlu_herb_wetland_p00050, rlu_water_p00100, rlu_ice_p00100, rlu_barren_p00100, rlu_decid_forest_p00100, rlu_evergreen_p00100, rlu_mix_forest_p00100, rlu_shrub_p00100, rlu_grass_p00100, rlu_pasture_p00100, rlu_crop_p00100, rlu_woody_wetland_p00100, rlu_herb_wetland_p00100, rlu_water_p00150, rlu_ice_p00150, rlu_barren_p00150, rlu_decid_forest_p00150, rlu_evergreen_p00150, rlu_mix_forest_p00150, rlu_shrub_p00150, rlu_grass_p00150, rlu_pasture_p00150, rlu_crop_p00150, rlu_woody_wetland_p00150, rlu_herb_wetland_p00150, rlu_water_p00300, rlu_ice_p00300, rlu_barren_p00300, rlu_shrub_p00300, rlu_grass_p00300, rlu_pasture_p00300, rlu_crop_p00300, rlu_woody_wetland_p00300, rlu_herb_wetland_p00300, rlu_water_p00400, rlu_ice_p00400, rlu_barren_p00400, rlu_shrub_p00400, rlu_grass_p00400, rlu_pasture_p00400, rlu_crop_p00400, rlu_woody_wetland_p00400, rlu_herb_wetland_p00400, rlu_water_p00500, rlu_ice_p00500, rlu_barren_p00500,

Reason	N	Covariates
		rlu_shrub_p00500, rlu_grass_p00500, rlu_pasture_p00500, rlu_crop_p00500, rlu_woody_wetland_p00500, rlu_herb_wetland_p00500, rlu_ice_p00750, rlu_shrub_p00750, rlu_grass_p00750, rlu_pasture_p00750, rlu_crop_p00750, rlu_herb_wetland_p00750, rlu_ice_p01000, rlu_shrub_p01000, rlu_grass_p01000, rlu_pasture_p01000, rlu_crop_p01000, rlu_ice_p03000, rlu_shrub_p03000, rlu_grass_p03000, rlu_pasture_p03000, rlu_crop_p03000, rlu_ice_p05000, rlu_shrub_p05000, rlu_grass_p05000, rlu_pasture_p05000, rlu_crop_p05000, tl_s00050, tl_s00100, tl_s00150, intersect_a1_a1_s00500, intersect_a1_a2_s00500, intersect_a1_a3_s00500, intersect_a2_a2_s00500, intersect_a2_a3_s00500, intersect_a1_a1_s01000, intersect_a1_a2_s01000, intersect_a2_a2_s01000, intersect_a2_a3_s01000, intersect_a1_a1_s03000, log_em_NOx_s03000, log_em_SO2_s03000, log_em_PM25_s03000, log_em_CO_s03000, log_em_PM10_s03000, log_m_to_s_port
Low land use (< 10%)	44	rlu_water_p00050, rlu_ice_p00050, rlu_pasture_p00050, rlu_crop_p00050, rlu_water_p00100, rlu_ice_p00100, rlu_pasture_p00100, rlu_crop_p00100, rlu_ice_p00150, rlu_pasture_p00150, rlu_crop_p00150, rlu_ice_p00300, rlu_pasture_p00300, rlu_crop_p00300, rlu_ice_p00400, rlu_pasture_p00400, rlu_crop_p00400, rlu_ice_p00500, rlu_pasture_p00500, rlu_crop_p00500, rlu_ice_p00750, rlu_shrub_p00750, rlu_grass_p00750, rlu_pasture_p00750, rlu_crop_p00750, rlu_ice_p01000, rlu_shrub_p01000, rlu_grass_p01000, rlu_pasture_p01000, rlu_crop_p01000, rlu_ice_p03000, rlu_barren_p03000, rlu_decid_forest_p03000, rlu_shrub_p03000, rlu_grass_p03000, rlu_woody_wetland_p03000, rlu_herb_wetland_p03000, rlu_ice_p05000, rlu_barren_p05000,

Reason	N	Covariates
		rlu_decid_forest_p05000, rlu_shrub_p05000, rlu_grass_p05000, rlu_woody_wetland_p05000, rlu_herb_wetland_p05000
Prediction sites more variable than monitoring sites (cohort SD > 5 times monitoring data SD)	22	rlu_water_p00050, rlu_pasture_p00050, rlu_crop_p00050, rlu_pasture_p00100, rlu_crop_p00100, rlu_pasture_p00150, rlu_crop_p00150, rlu_pasture_p00300, rlu_crop_p00300, rlu_pasture_p00400, rlu_crop_p00400, rlu_pasture_p00500, rlu_crop_p00500, rlu_pasture_p00750, rlu_crop_p00750, rlu_shrub_p01000, rlu_pasture_p01000, rlu_crop_p01000, rlu_shrub_p03000, rlu_shrub_p05000, rlu_grass_p05000, rlu_crop_p05000

Universal Kriging Models with Partial Least Squares



cohort locations outside the monitoring area but within the spatiotemporal modeling region

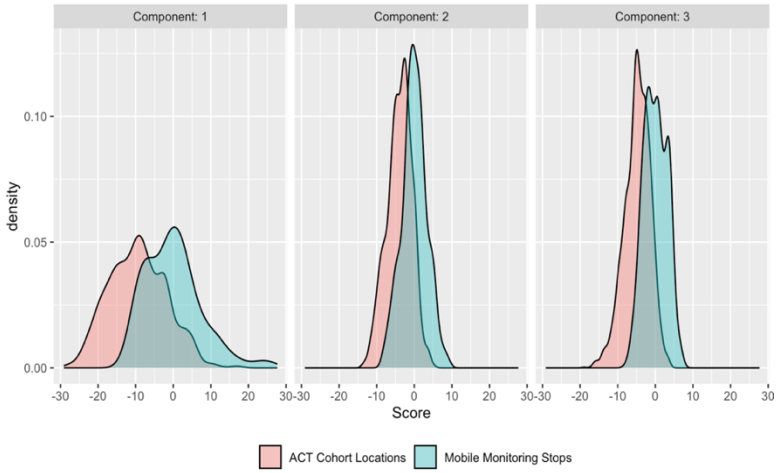
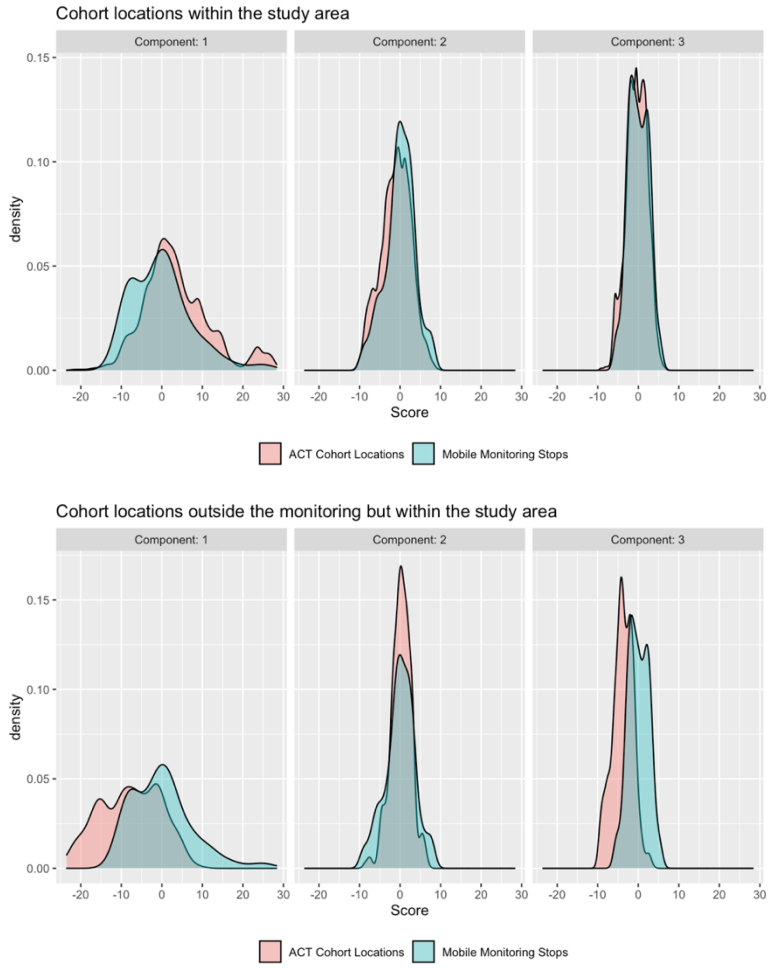


Figure 71. Distribution of UFP PLS component scores for mobile monitoring stops and ACT cohort locations within various areas



cohort locations outside the monitoring area but within the spatiotemporal modeling region

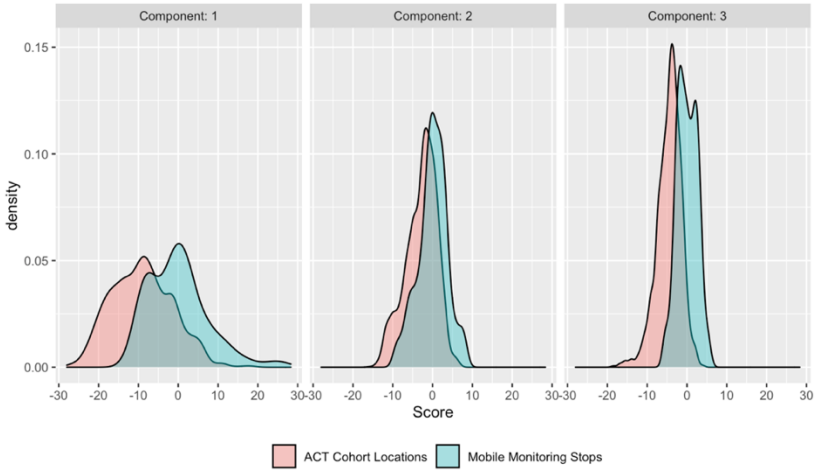


Figure 72. Distribution of BC PLS component scores for mobile monitoring stops and ACT cohort locations within various areas

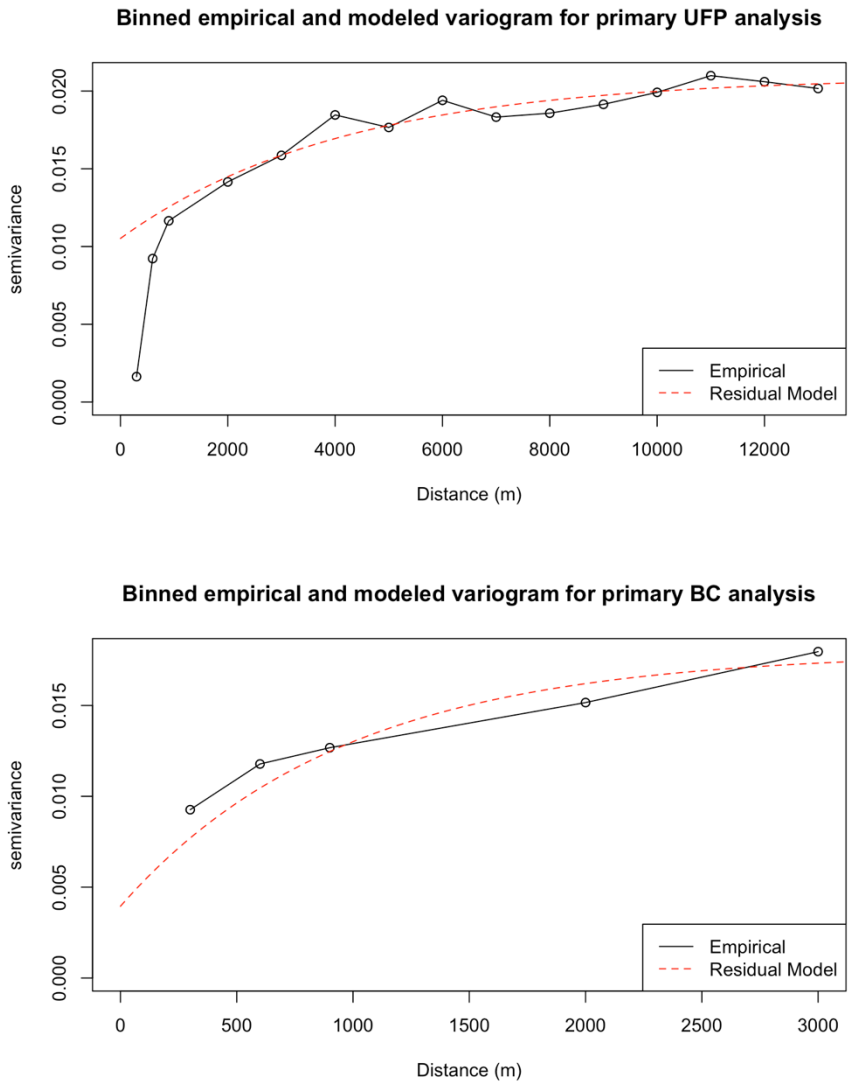


Figure 73. Binned empirical and modeled variogram for the UFP (top) and BC (bottom) model

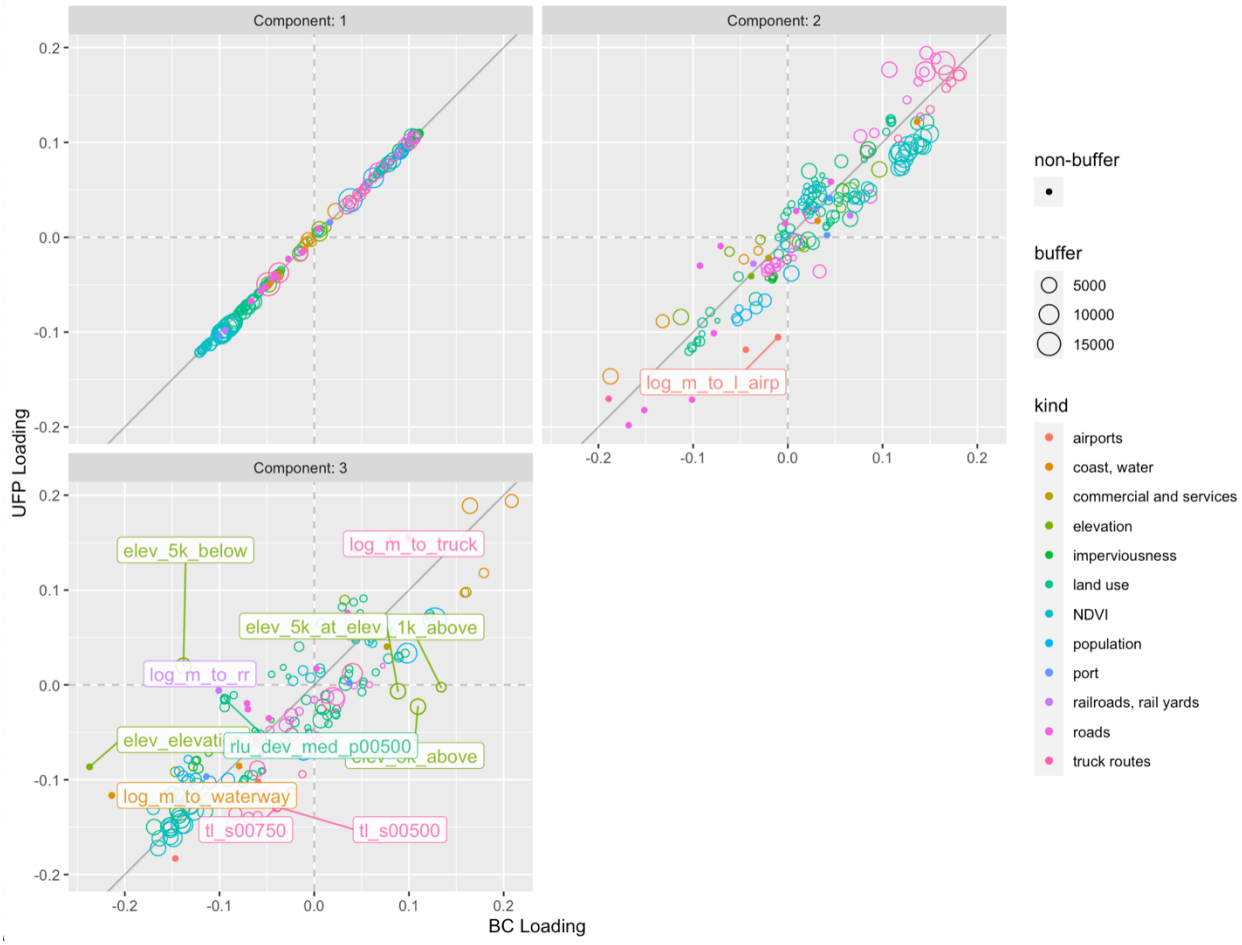


Figure 74. Comparison of UFP and BC PLS model loadings. Labeled pairs are those with the largest UFP-BC loading difference, as determined by the lowest and highest quantile from all three components (diff < 0.1 quantile or diff > 0.99 quantile).

Table 31. UFP residual quantiles used to define a "large" residual from UFP-BC best fit lines.

Dataset	Q05	Q95
Estimates	-1582	2055
Cohort Predictions	-1162	1504
Grid Predictions	-858	1127

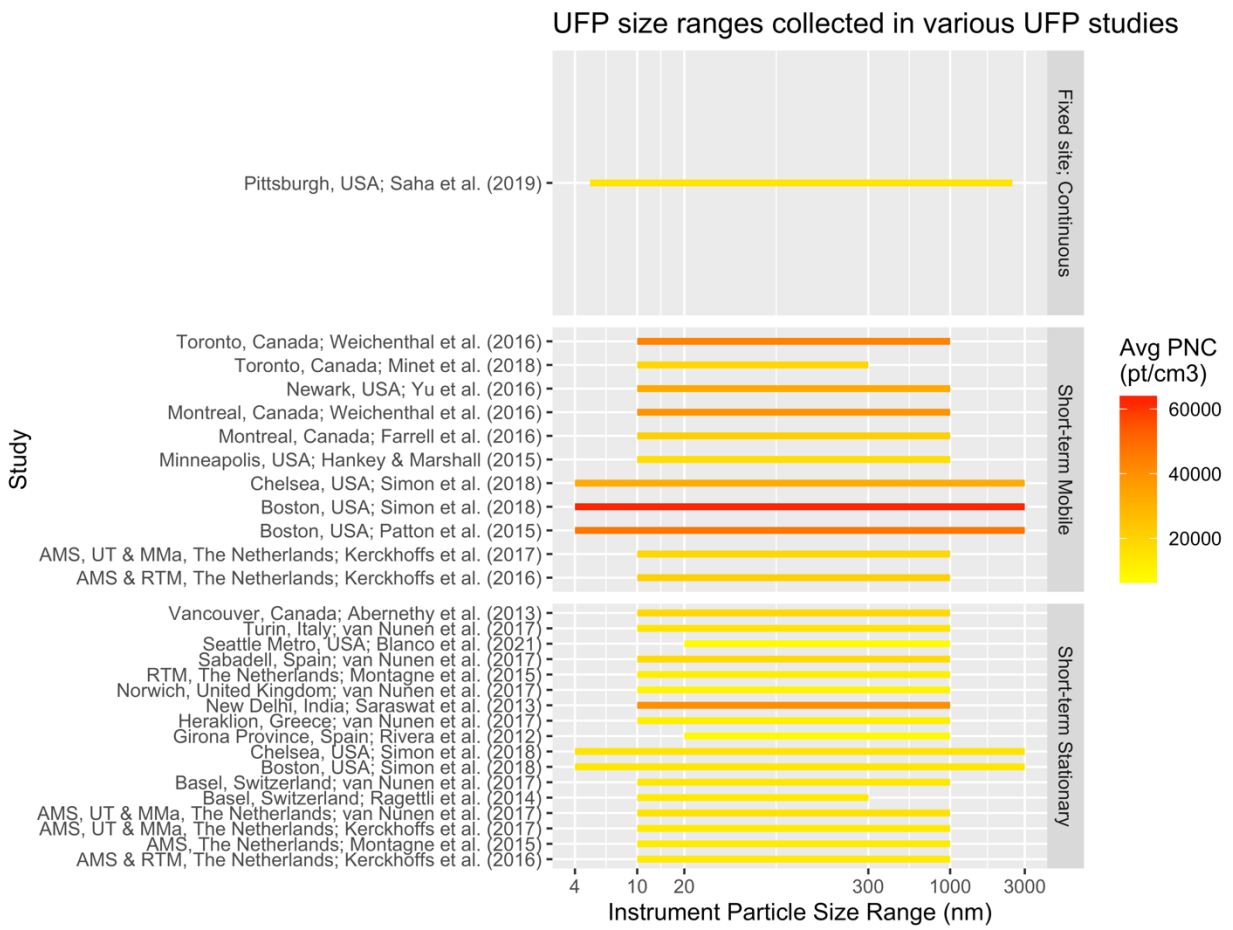


Figure 75. UFP size ranges and average concentrations reported in various studies, stratified by sampling mode. The x-axis is on the log scale.

Appendix C: Aim 3

Methods

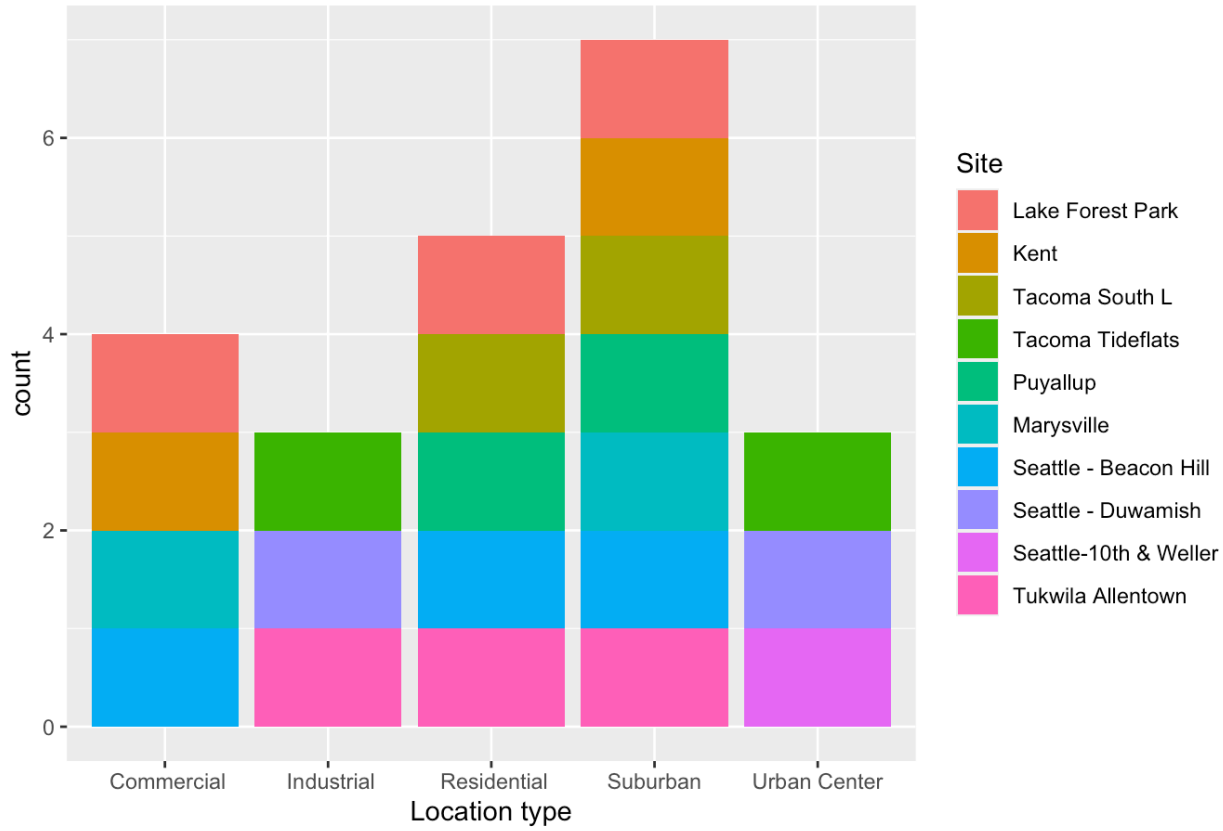


Figure 76. AQS sites near the study area by location type

Results

Primary

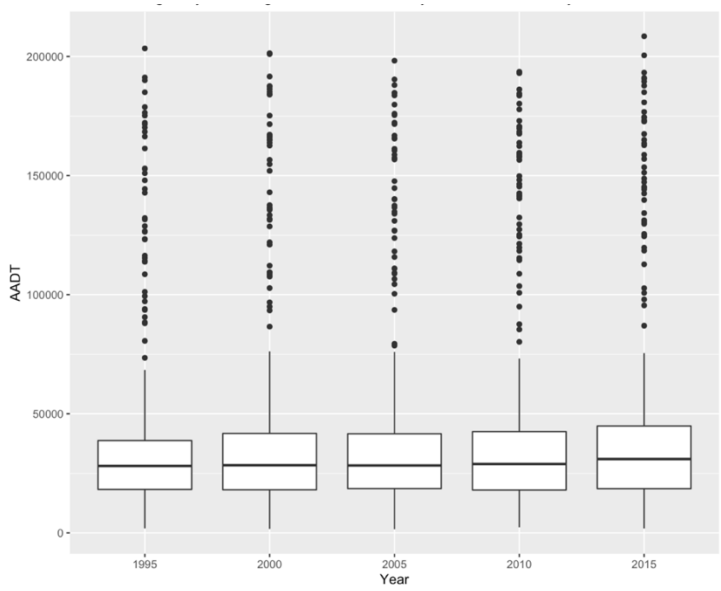


Figure 77. Five-year average AADT at highway point locations.

Table 32. Five-year average AADT at highway point locations. N = number of point locations with counters near the study area.

Year	N	Min	Q05	Mean	SD	Median	IQR	Q95	Max
Overall	1,270	1,580	6,054	46,318	48,262	29,000	24,462	167,020	208,500
1995	254	1,900	6,419	43,303	45,040	28,100	20,500	155,940	203,400
2000	254	1,700	6,320	45,964	48,568	28,400	23,650	166,233	201,400
2005	254	1,580	6,216	46,275	48,361	28,300	23,000	166,118	198,250
2010	254	2,340	5,745	46,608	48,636	28,900	24,500	165,130	193,600
2015	254	1,875	5,989	49,441	50,715	31,000	26,312	172,925	208,500

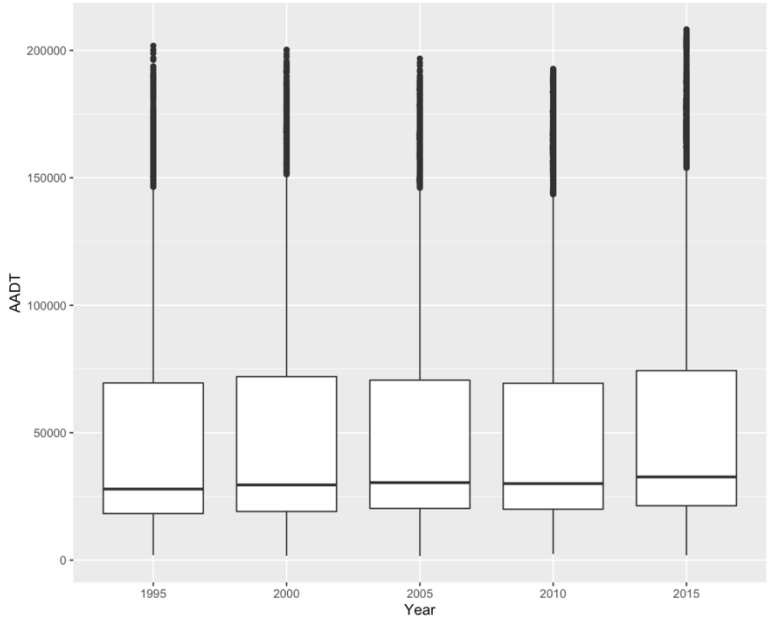


Figure 78. Five-year average AADT on 100 m highway road links between 1995-2018.

Table 33 . Five-year average AADT on 100 m highway road links between 1995-2018. N = the number of road links.

Year	N	Min	Q05	Mean	SD	Median	IQR	Q95	Max
Overall	51,560	1,604	8,767	52,955	50,416	30,404	51,210	168,016	208,246
1995	10,312	1,935	6,540	48,107	46,327	27,888	51,246	150,337	201,799
2000	10,312	1,725	9,494	52,608	50,711	29,512	52,903	168,016	200,268
2005	10,312	1,604	9,688	53,336	50,344	30,418	50,302	165,317	196,765
2010	10,312	2,396	9,100	53,651	50,717	30,028	49,412	165,026	192,758
2015	10,312	1,900	9,035	57,074	53,331	32,648	53,008	174,902	208,246

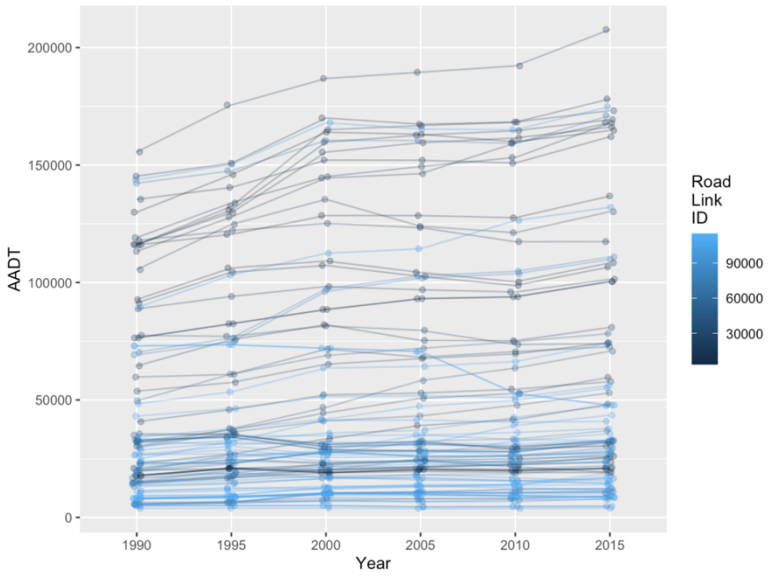


Figure 79. AADT on 100 m highway road links near the study area. Plot shows 100 randomly selected road links.

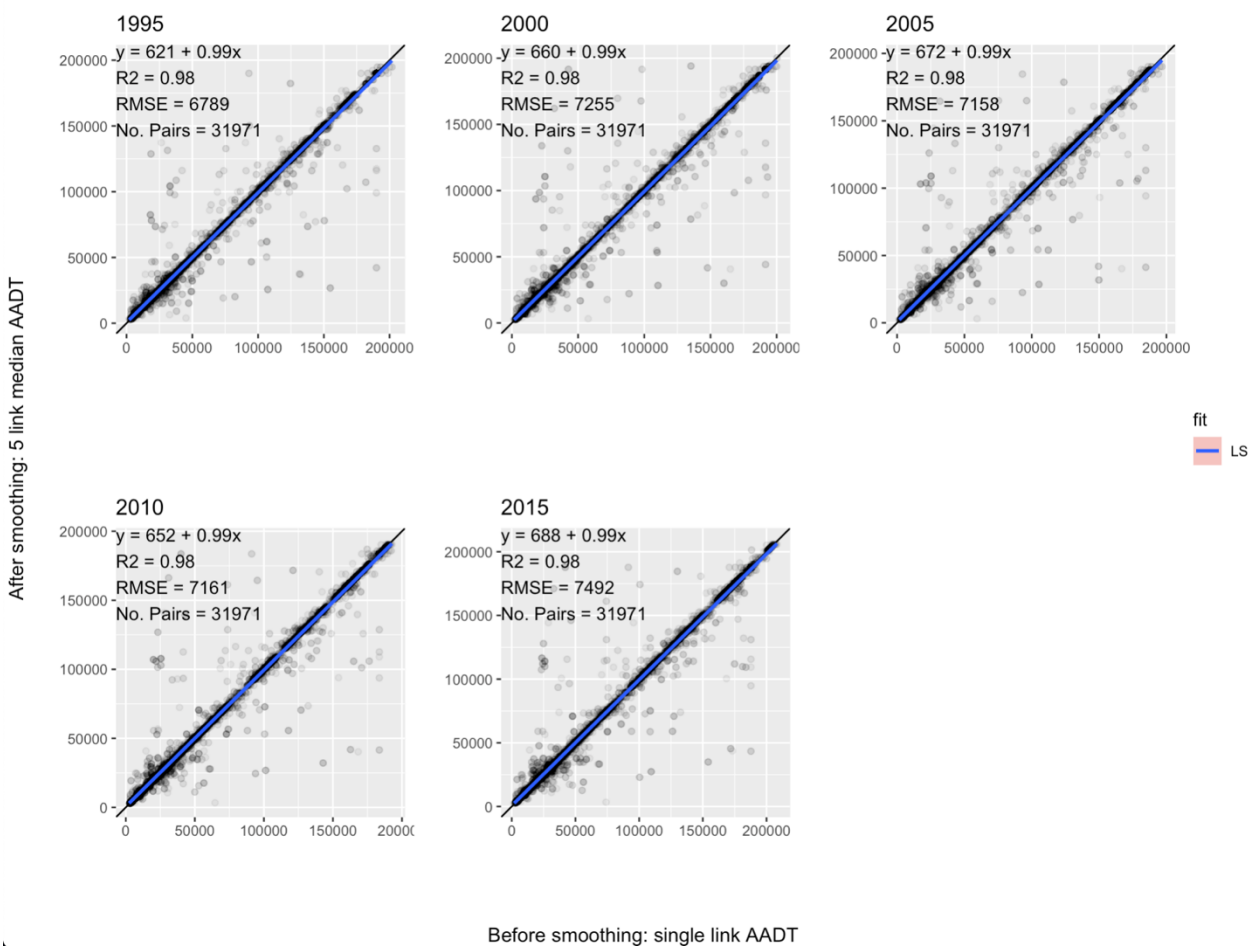


Figure 80. AADT on 100 m road links before and after smoothing. Smoothing was based other nearby road links (N=5). Points furthest away from the 1-1 line indicate the locations where smoothing had the largest impact.

Table 34. EC emissions (g/100 m) over time on restricted roads near our study area

Year	N	Min	Q05	Mean	SD	Median	IQR	Q95	Max
Overall	191,826	0	10	133	166	73	122	521	1,025
1995	31,971	0	31	246	244	140	269	791	1,025
2000	31,971	0	27	201	198	113	211	660	787
2005	31,971	0	22	148	143	84	149	474	545

2010	31,971	0	14	96	93	54	87	303	350
2015	31,971	0	9	61	59	36	58	194	229
2019	31,971	0	7	43	42	25	41	137	162

Figure 81. Number of daily EC samples available at Beacon Hill. Lighter-colored years are those that did not meet the data completeness criteria.

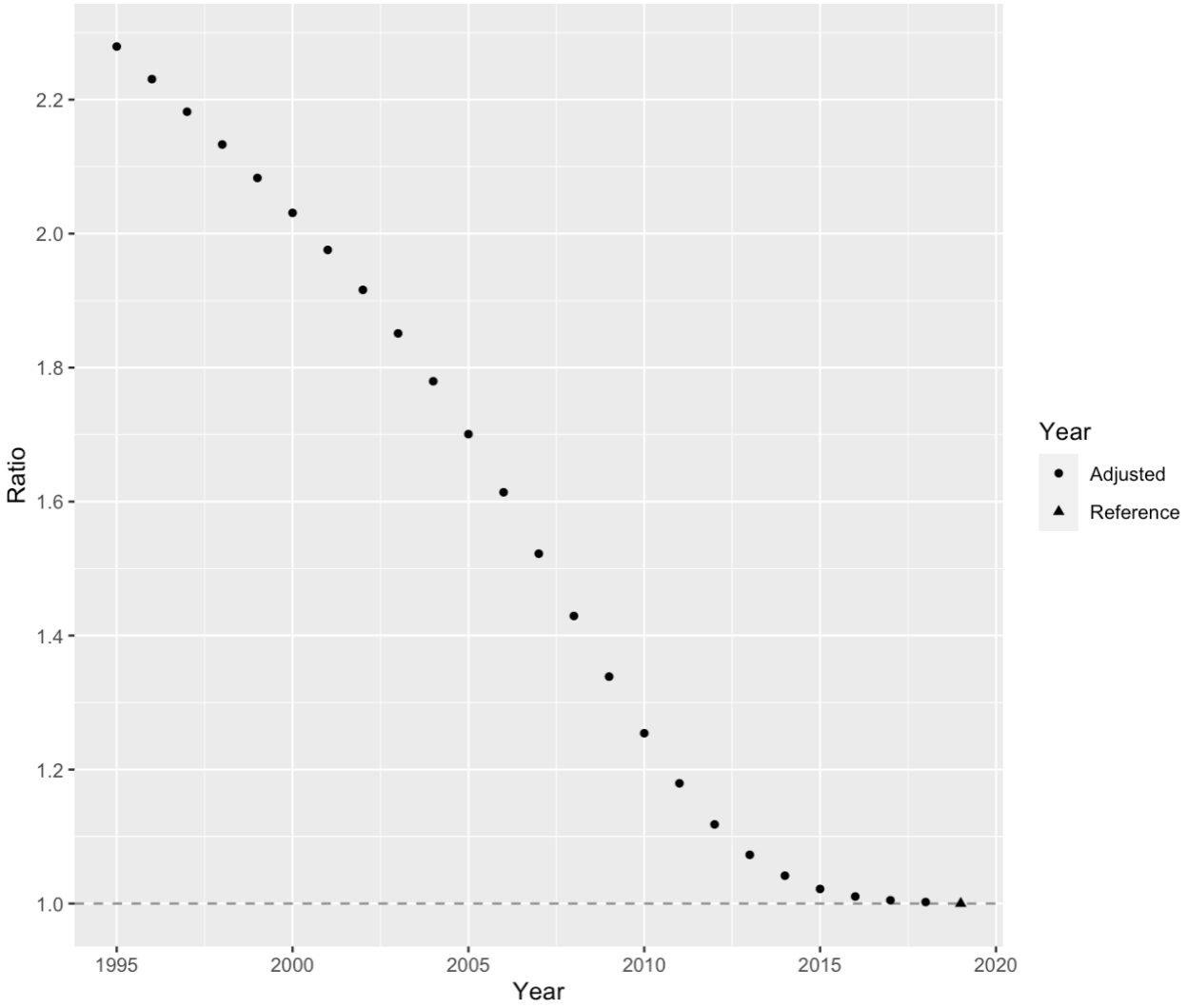


Figure 82. Annual adjustment factors based on EC readings at Beacon Hill, using 2019 as the reference year.

Table 35. Predicted historical BC exposure for the cohort

Year	N	Min	Q05	Mean	SD	Median	IQR	Q95	Max
Overall	602,650	213	440	883	393	805	500	1,574	5,225
1995	24,106	494	906	1,372	460	1,269	407	2,318	5,225
2000	24,106	440	806	1,195	349	1,125	348	1,929	3,580

2005	24,106	366	666	962	245	919	270	1,472	2,291
2010	24,106	269	484	684	155	663	186	1,009	1,432
2015	24,106	218	394	551	117	537	146	761	1,132
2019	24,106	213	384	536	111	524	141	726	1,090

Table 36. Predicted historical UFP exposure for the cohort

Year	N	Min	Q05	Mean	SD	Median	IQR	Q95	Max
Overall	602,650	2,245	5,683	11,509	5,247	10,560	6,445	20,765	55,571
1995	24,106	5,170	11,694	17,668	6,187	16,073	4,939	30,590	55,571
2000	24,106	4,605	10,359	15,443	4,889	14,254	4,254	25,837	40,874
2005	24,106	3,846	8,586	12,560	3,602	11,741	3,397	20,709	28,209
2010	24,106	2,829	6,273	9,019	2,384	8,511	2,392	14,410	19,304
2015	24,106	2,299	5,101	7,276	1,843	6,914	1,888	11,479	15,242
2019	24,106	2,245	4,987	7,086	1,760	6,753	1,822	11,095	14,712

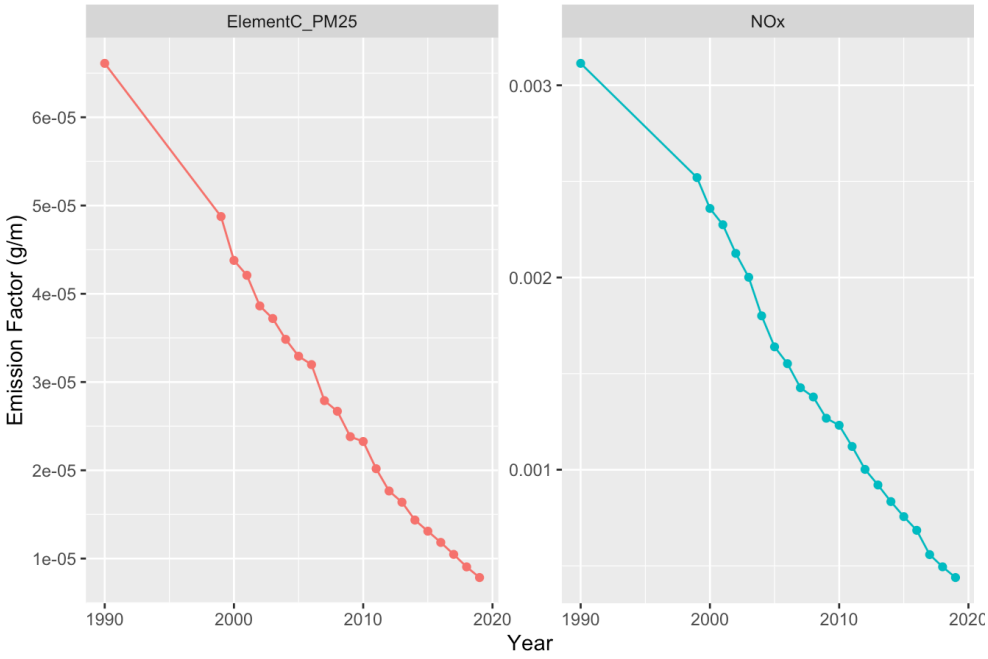


Figure 83. MOVES emission factors over time for EC (ElementC_PM25) and NOx.

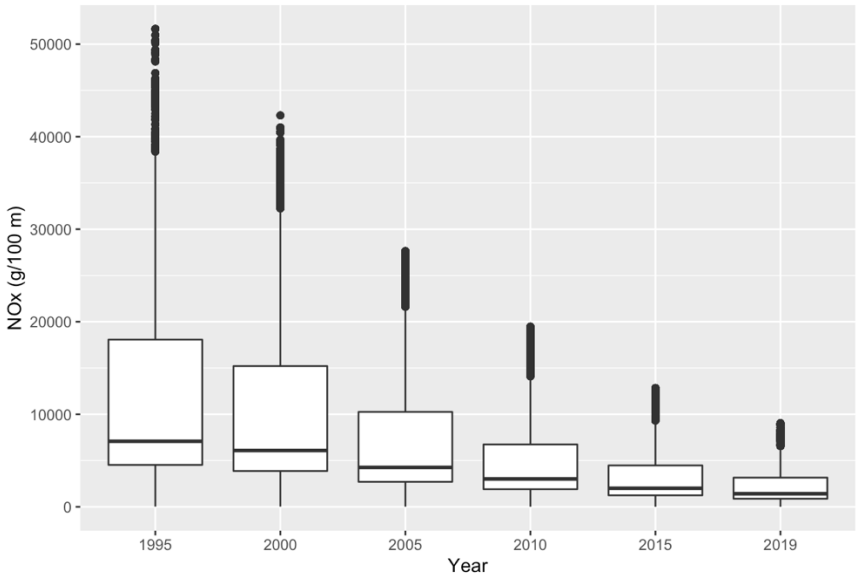


Figure 84. NOx emissions on 100 m road links over time, after smoothing.

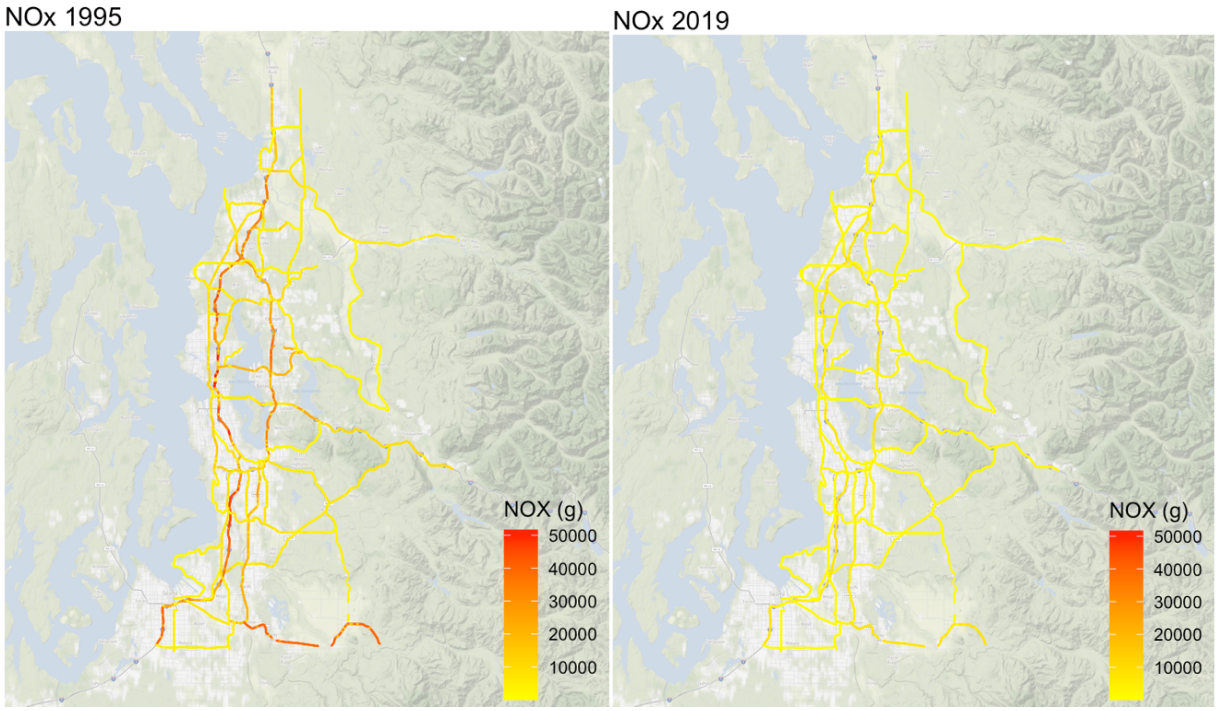


Figure 85. Map of NOx emissions on 100 m road links at the beginning (1995) and end (2019) of the study period

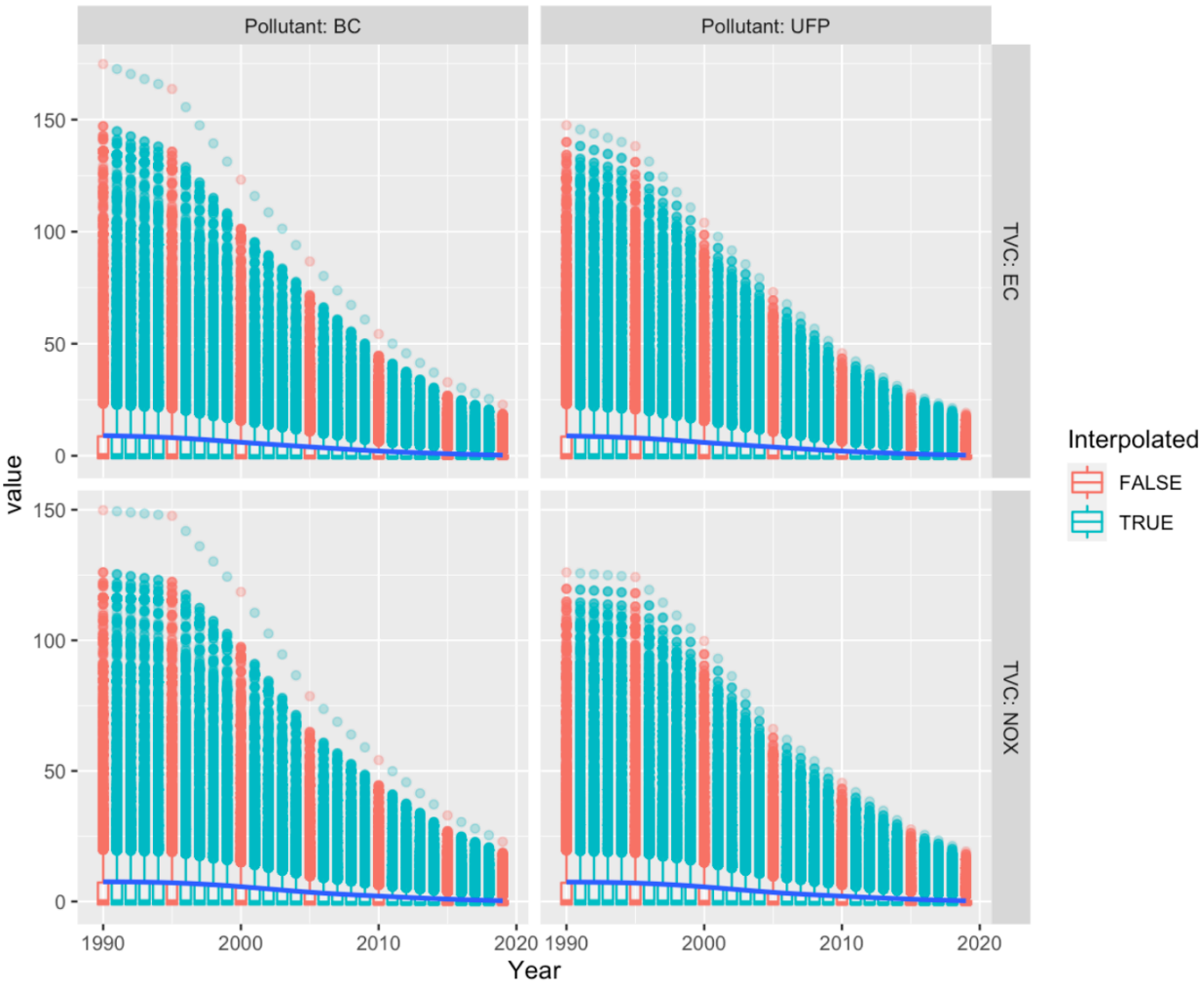


Figure 86. Distribution of observed and interpolated NOx and EC time-varying covariate (TVC) PLS scores at cohort locations for the BC and UFP models

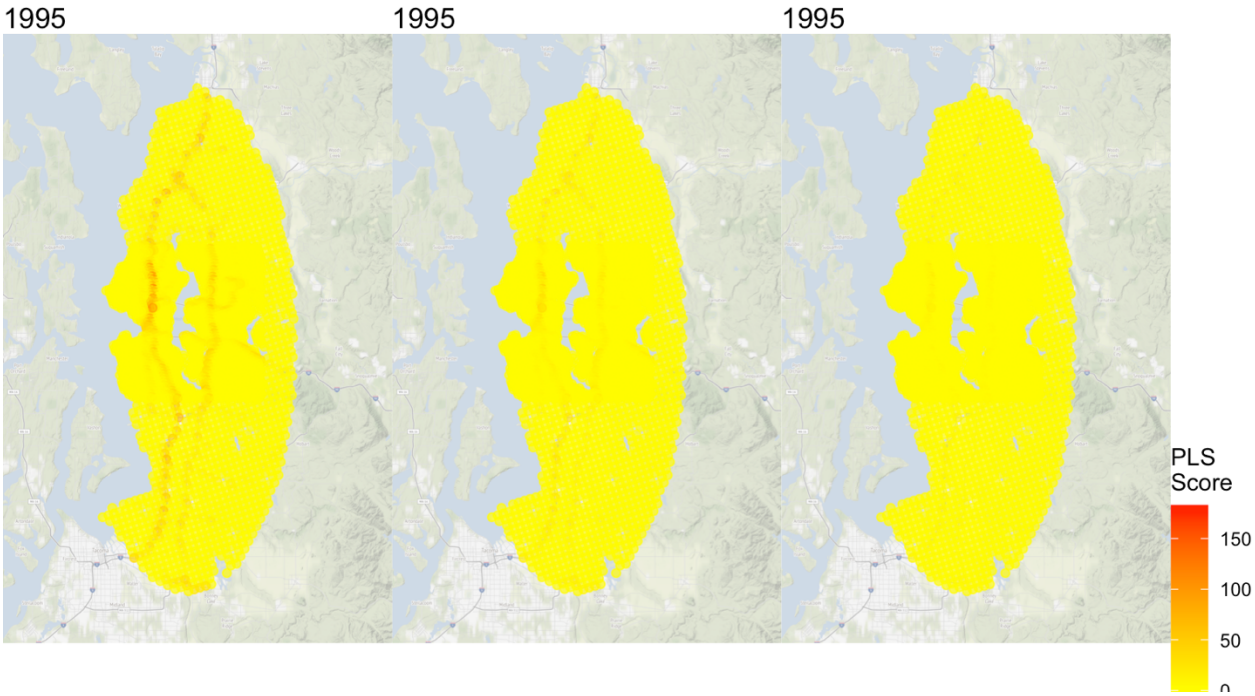


Figure 87. NOx PLS scores from the BC model for 1995, 2010 and 2019.

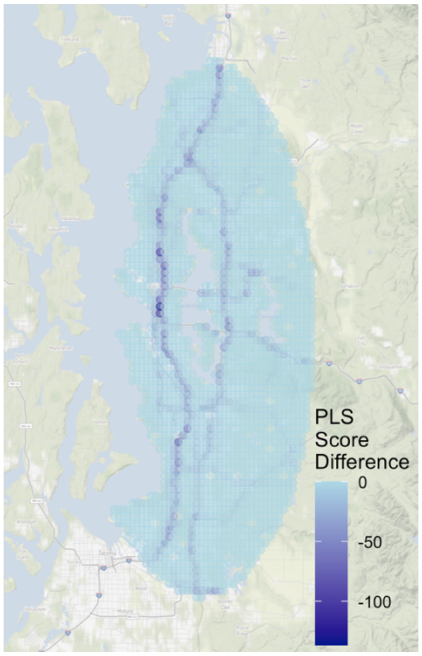


Figure 88. Change in NOx PLS scores from 1995-2019, based on the BC model

Discussion

Additional Analyses of the $PM_{2.5}$ Concentration Rank Order Stability Over Time

The purpose of this analysis was to characterize how the $PM_{2.5}$ spatial structure has changed over time. The rationale was that this could help us understand how other pollutants with more limited historical information (e.g., NO_2, UFP) may have also changed over time. Analyses use spatiotemporal $PM_{2.5}$ model predictions for ACT cohort locations at the two-week level averaged over the year. Predictions are at cohort locations since these are the locations we care about. The annual average for 2018 is used as the reference since it is the most recent year for which we had a full year's worth of predictions.

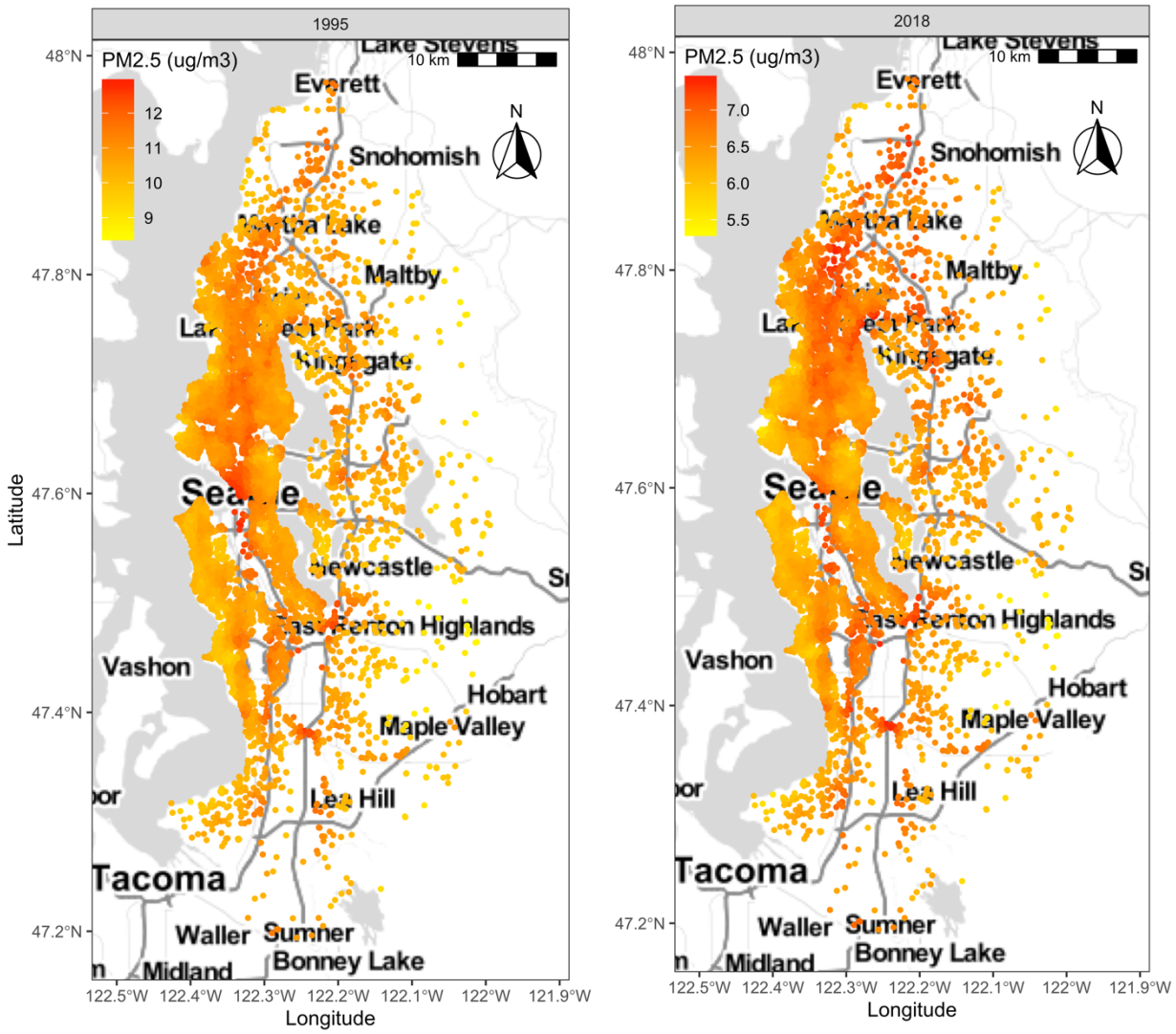


Figure 89. Spatiotemporal $PM_{2.5}$ model predictions at ACT cohort locations in 1995 (left) and 2018 (right)

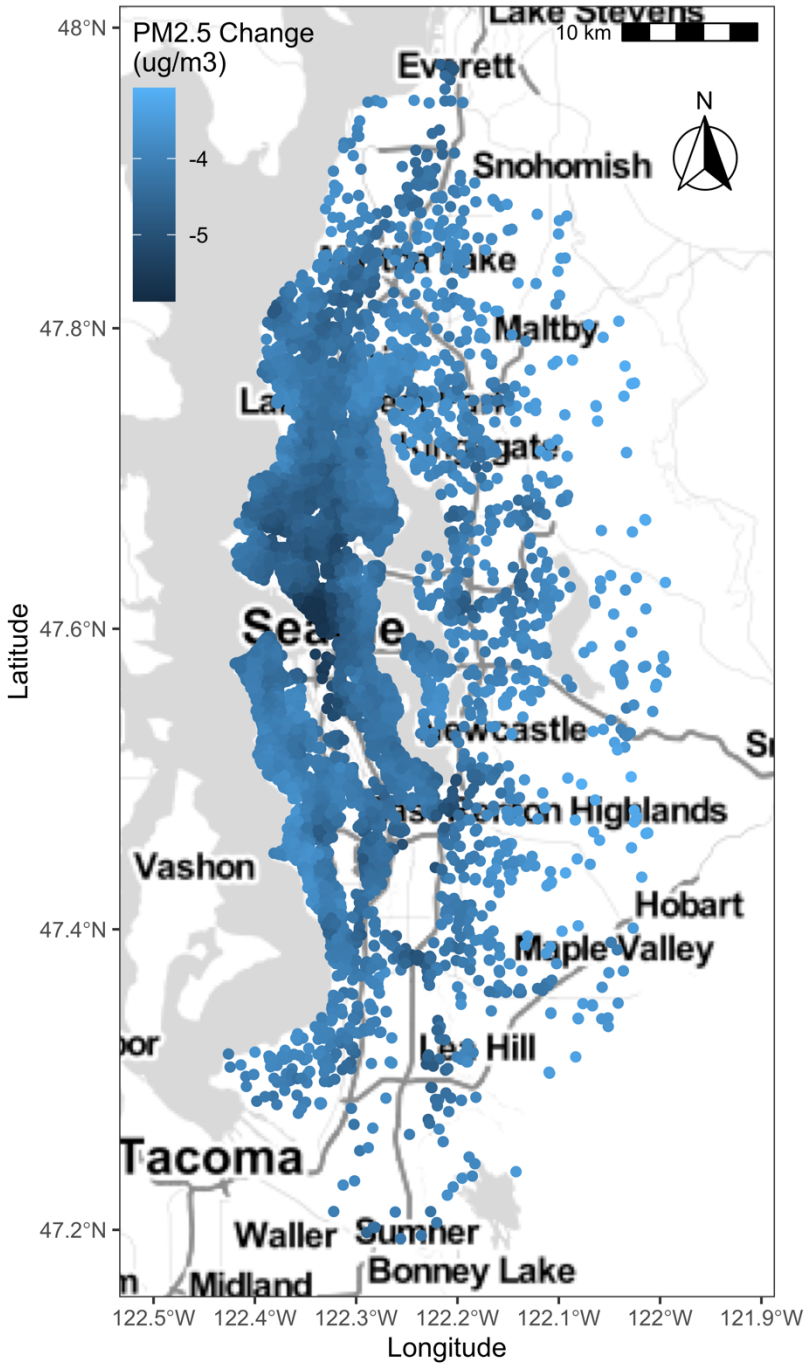


Figure 90. Reduction in spatiotemporal $PM_{2.5}$ Model Predictions at ACT cohort locations from 1995 to 2018

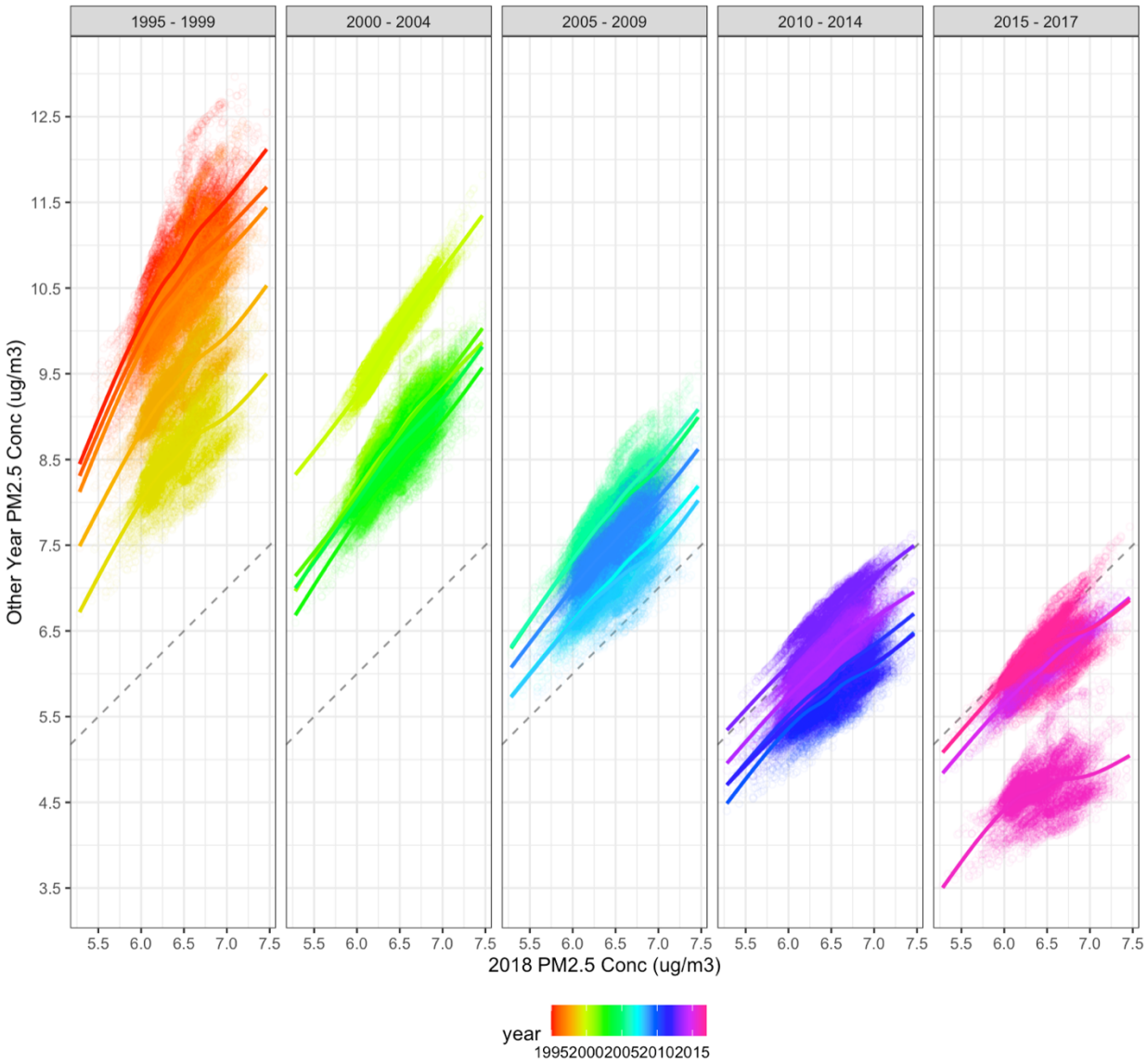


Figure 91. Comparison of spatiotemporal $PM_{2.5}$ model predictions at ACT cohort locations for 2018 and other years. Plot includes loess lines. Plots show that there was some change in the rank order of predictions (e.g., a concentration in 2018 was associated with varying concentration levels during earlier years).

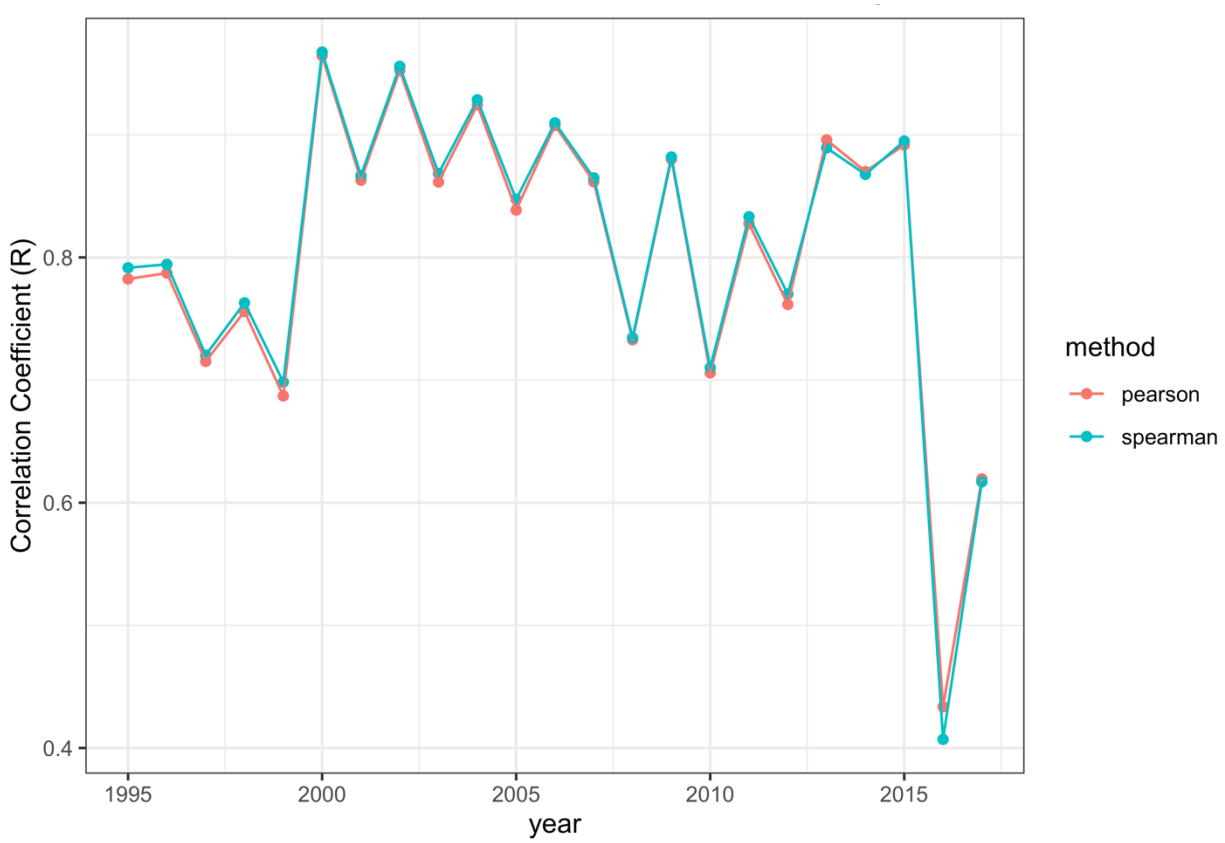


Figure 92. Correlation coefficients for spatiotemporal model predictions at ACT cohort locations for 2018 versus other years

Table 37. Distribution of correlation coefficients for spatiotemporal model predictions at ACT cohort locations for 2018 and other years

Method	Min	Mean	SD	Max
Pearson	0.43	0.81	0.12	0.96
Spearman	0.41	0.81	0.12	0.97

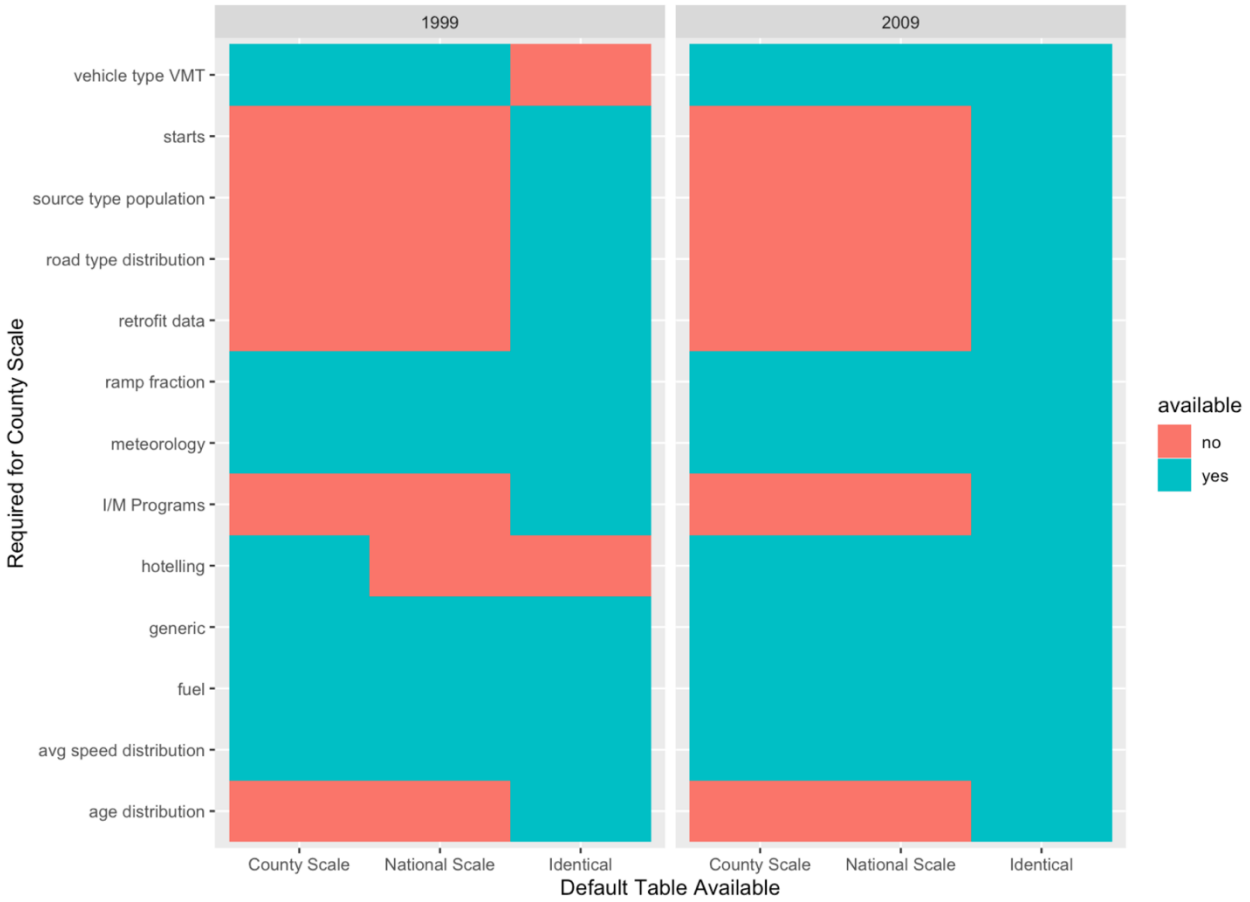


Figure 93. Required inputs for running MOVES at the county-scale and whether default inputs are available at the county- and national- scale for two example years (1999, 2009). The identical column shows whether the default county- and national- scale tables are the same (e.g., both are missing or both area available and with the same information). Note that hoteling hours were not used to estimate emission factors in this analysis. For 1999, the national-scale actually had slightly more complete vehicle type vehicle miles traveled (VMT) information than the county-scale.

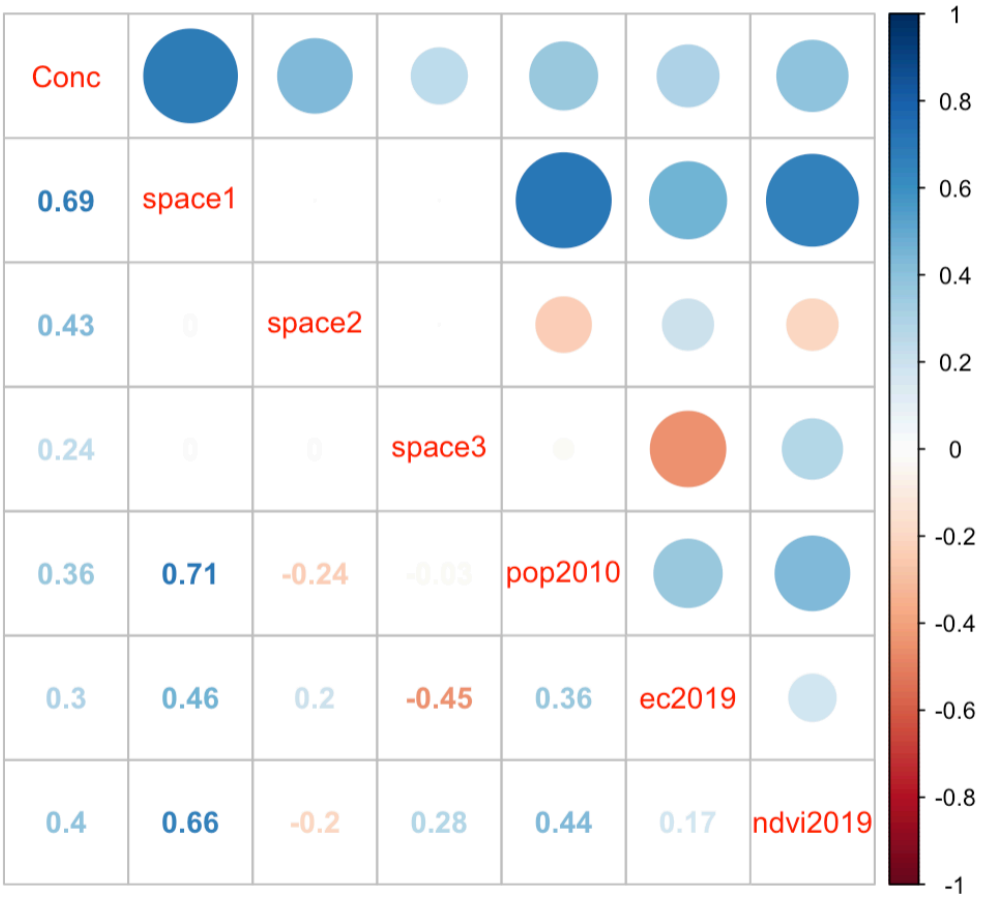


Figure 94. Pearson correlations (R) between UFP concentration and model covariates (PLS component scores).

Table 38. Linear regression UFP model with model covariates. This is a simplified model over the UK model to illustrate coefficient estimates and p-values. Using mobile monitoring stop observations to fit models. All covariates were standardized (yi-mean(yi)/sd(yi). Similar findings for BC model.

term	estimate	std.error	statistic	p.value	
(Intercept)	8.800	0.008	1118.322	0.000	*
space1	0.206	0.015	13.623	0.000	*
space2	0.103	0.009	11.064	0.000	*
space3	0.072	0.010	7.226	0.000	*
pop2010	-0.015	0.012	-1.210	0.227	
ec2019	0.003	0.011	0.285	0.776	
ndvi2019	-0.027	0.012	-2.222	0.027	*

Table 39. The variability explained an increasing number of PLS component predictors in least squares linear regression for UFP concentration.

Predictors in Model	R^2
space	**0.72**
pop	0.13
ec	0.09
ndvi	0.16
all TVCs	**0.23**
space + pop	0.72
space + ec	0.72
space + ndvi	0.73
space + pop + ec	0.72
space + all TVCs	**0.73**

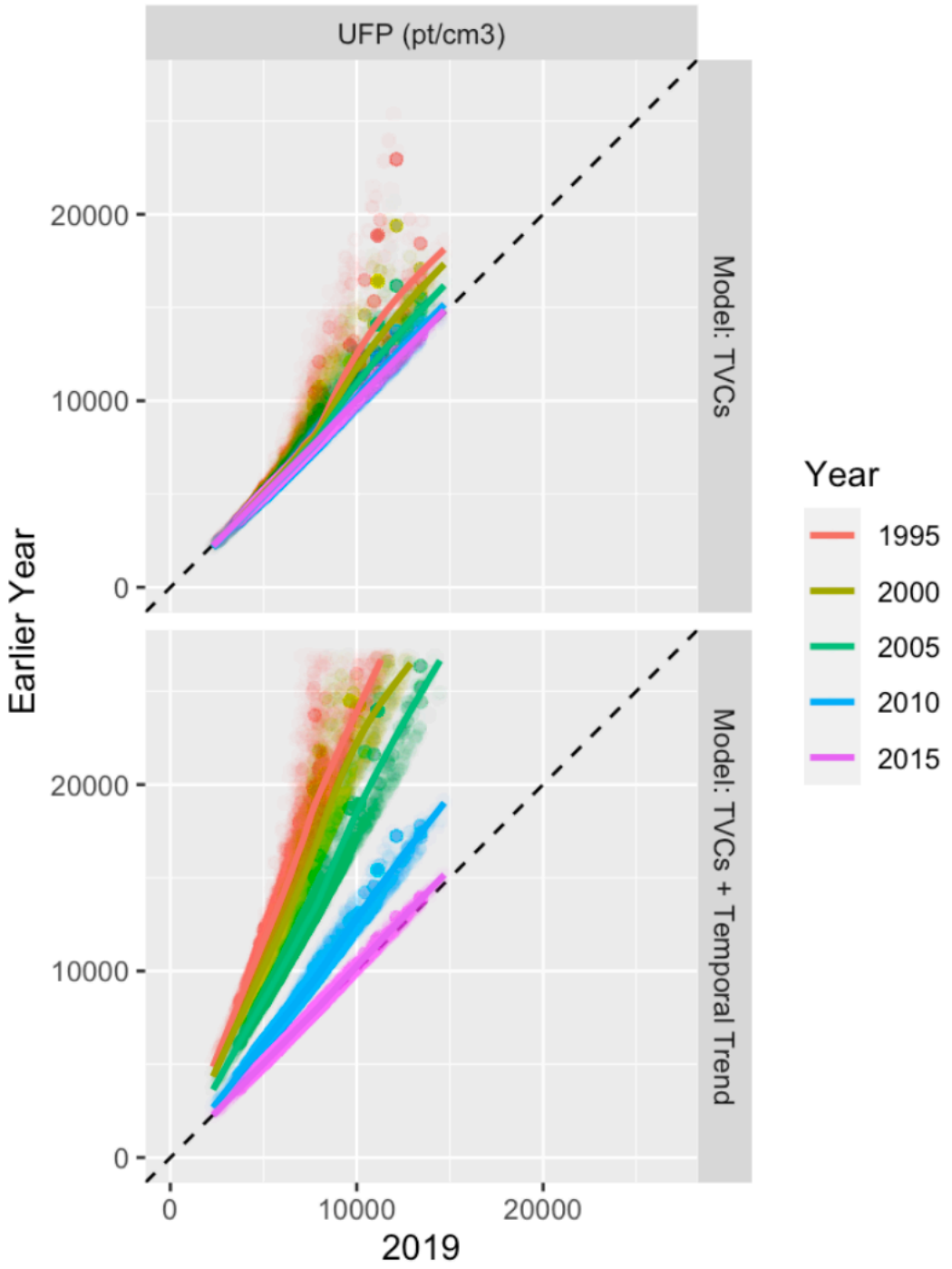


Figure 95. Comparison of model predictions for 2019 vs a few example years from the primary analysis (models include TVCs and temporal trend factors) and an analysis without trend factors.

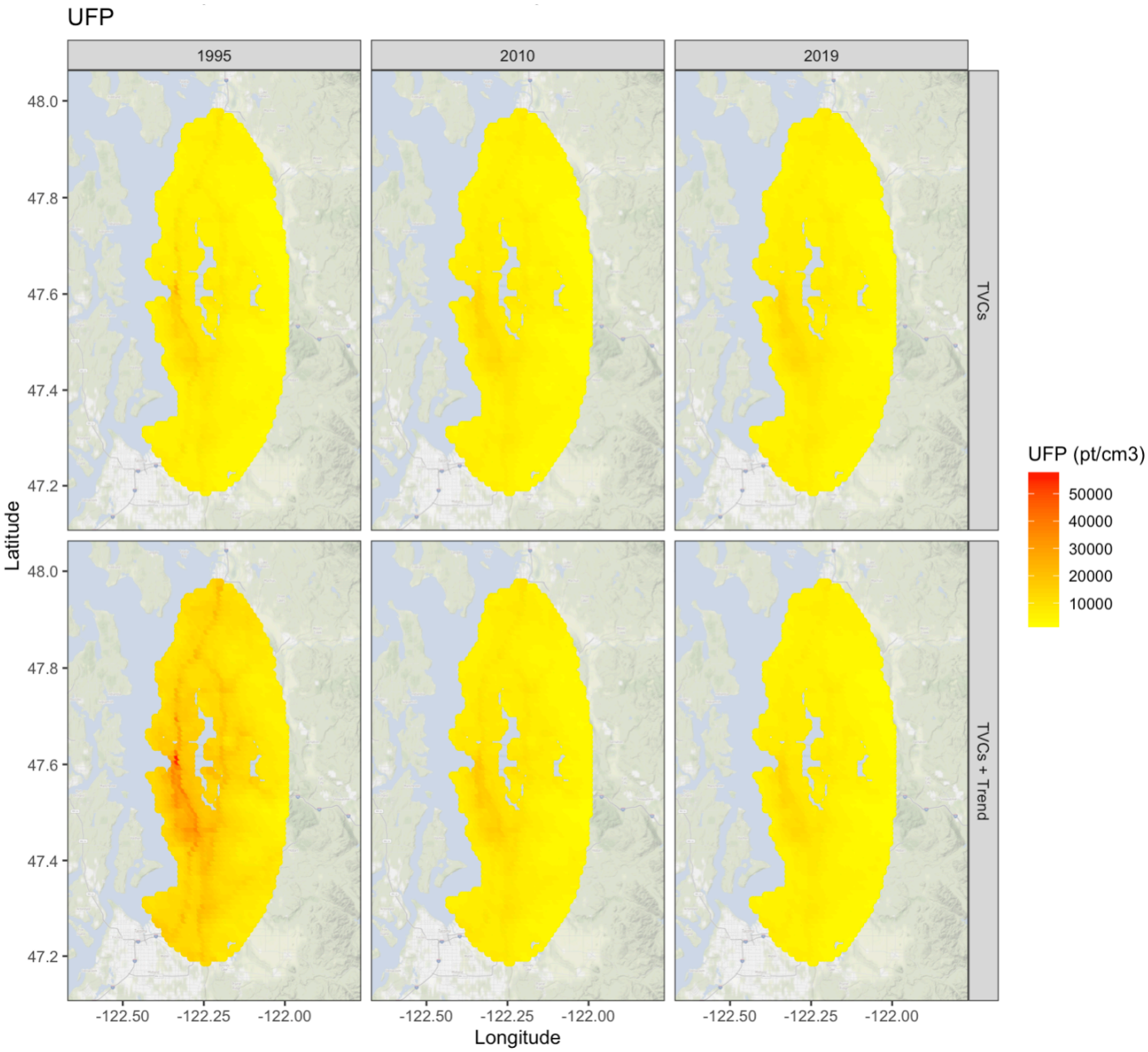


Figure 96. UFP predictions from the primary model (TVCs + Trend) vs a model without trend (TVCs).

References

Abernethy, R. C., Allen, R. W., McKendry, I. G., & Brauer, M. (2013). A land use regression model for ultrafine particles in Vancouver, Canada. *Environmental Science & Technology*, *47*(10), 5217–5225.

Abhijith, K. V, Kumar, P., Gallagher, J., McNabola, A., Baldauf, R., Pilla, F., Broderick, B., Di Sabatino, S., & Pulvirenti, B. (2017). Air pollution abatement performances of green infrastructure in open road and built-up street canyon environments – A review. *Atmospheric Environment*, *162*, 71–86. https://doi.org/https://doi.org/10.1016/j.atmosenv.2017.05.014

AethLabs. (2018). *microAeth MA Series MA200, MA300, MA350 Operating Manual*. https://aethlabs.com/sites/all/content/microaeth/maX/MA200 MA300 MA350 Operating Manual Rev 01 Oct 2017.pdf

Ahlskog, J. E., Geda, Y. E., Graff-Radford, N. R., & Petersen, R. C. (2011). Physical exercise as a preventive or disease-modifying treatment of dementia and brain aging. *Mayo Clinic Proceedings*, *86*(9), 876–884. https://doi.org/10.4065/mcp.2011.0252

Alexeeff, S. E., Roy, A., Shan, J., Liu, X., Messier, K., Apte, J. S., Portier, C., Sidney, S., & Eeden, S. K. Van Den. (2018). High-resolution mapping of traffic related air pollution with Google street view cars and incidence of cardiovascular events within neighborhoods in Oakland, CA. *Environmental Health*, *17*(38), 1–13.

Allen, J. L., Klocke, C., Morris-Schaffer, K., Conrad, K., Sobolewski, M., & Cory-Slechta, D. A. (2017). Cognitive Effects of Air Pollution Exposures and Potential Mechanistic Underpinnings. In *Current environmental health reports*. https://doi.org/10.1007/s40572-017-0134-3

Allen, R., Larson, T., Sheppard, L., Wallace, L., & Liu, L. J. S. (2003). Use of real-time light scattering data to estimate the contribution of infiltrated and indoor-generated particles to indoor air. *Environmental Science and Technology*. https://doi.org/10.1021/es021007e

Alzheimer's Association. (2008). *Alzheimer's & Dementia*. https://www.alz.org/dementia/types-of-dementia.asp

Alzheimer's Association. (2015a). 2015 Alzheimer's disease facts and figures. *Alzheimer's and Dementia*, *11*(3), 332–384. https://doi.org/10.1016/j.jalz.2015.02.003

Alzheimer's Association. (2015b, November). *Alzheimer's Disease: How the disease progresses*. Mayo Clinic. http://www.alz.org/alzheimers_disease_stages_of_alzheimers.asp

Alzheimer's Association. (2019). *2019 Alzheimer's Disease Facts and Figures*. https://www.alz.org/media/documents/alzheimers-facts-and-figures-2019-r.pdf

Alzheimer's Association. (2020). *Causes and Risk Factors*.

Alzheimer's Queensland. (2014, December). *Types of Dementia*. AAQ Factsheet. https://doi.org/10.1590/S0102-695X2011005000114

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders : DSM-IV.* (4th ed.). Washington, DC: American Psychiatric Association.

Apte, J. S., Messier, K. P., Gani, S., Brauer, M., Kirchstetter, T. W., Lunden, M. M., Marshall, J. D., Portier, C. J., Vermeulen, R. C. H., & Hamburg, S. P. (2017). High-Resolution Air Pollution Mapping with Google Street View Cars: Exploiting Big Data. *Environmental Science and*

Technology, *51*(12), 6999–7008. https://doi.org/10.1021/acs.est.7b00891
Arrighi, H. M., Neumann, P. J., Lieberburg, I. M., & Townsend, R. J. (2010). Lethality of Alzheimer disease and its impact on nursing home placement. *Alzheimer Disease & Associated Disorders*, *24*(1), 90–95.
Austin, E., Xiang, J., Gould, T. R., Shirai, J. H., Yun, S., Yost, M. G., Larson, T. V, & Seto, E. (2021). Distinct Ultrafine Particle Profiles Associated with Aircraft and Roadway Traffic. *Environmental Science & Technology*, *55*(5), 2847–2858. https://doi.org/10.1021/acs.est.0c05933
Austin, E., Xiang, J., Gould, T., Shirai, J., Yun, S., Yost, M. G., Larson, T., & Seto, E. (2019). *Mobile ObserVations of Ultrafine Particles: The MOV-UP study report*. https://deohs.washington.edu/mov-up
Aviation Benefits Beyond Borders. (2020). *Improving air quality*. https://aviationbenefits.org/environmental-efficiency/improving-air-quality/
Babadjouni, R. M., Hodis, D. M., Radwanski, R., Durazo, R., Patel, A., Liu, Q., & Mack, W. J. (2017). Clinical effects of air pollution on the central nervous system; a review. *Journal of Clinical Neuroscience: Official Journal of the Neurosurgical Society of Australasia*, *43*, 16–24. https://doi.org/10.1016/j.jocn.2017.04.028
Batterman, S., Cook, R., & Justin, T. (2015). Temporal variation of traffic on highways and the development of accurate temporal allocation factors for air pollution analyses. *Atmospheric Environment*, *107*, 351–363. https://doi.org/10.1016/j.atmosenv.2015.02.047
Beelen, R, Hoek, G., Fischer, P., van den Brandt, P. A., & Brunekreef, B. (2007). Estimated long-term outdoor air pollution concentrations in a cohort study. *Atmospheric Environment (1994)*, *41*(26), 1343–1358.
Beelen, Rob, Hoek, G., Vienneau, D., Eeftens, M., Dimakopoulou, K., Pedeli, X., Tsai, M.-Y., Künzli, N., Schikowski, T., Marcon, A., Eriksen, K. T., Raaschou-Nielsen, O., Stephanou, E., Patelarou, E., Lanki, T., Yli-Tuomi, T., Declercq, C., Falq, G., Stempfelet, M., … de Hoogh, K. (2013). Development of NO2 and NOx land use regression models for estimating air pollution exposure in 36 study areas in Europe – The ESCAPE project. *Atmospheric Environment (1994)*, *72*, 10–23. https://doi.org/10.1016/j.atmosenv.2013.02.037
Bellander, T., Berglind, N., Gustavsson, P., Jonson, T., Nyberg, F., Pershagen, G., & Järup, L. (2001). Using geographic information systems to assess individual historical exposure to air pollution from traffic and house heating in stockholm. *Environmental Health Perspectives*, *109*(6), 633–639. https://doi.org/10.1289/ehp.01109633
Blanco, M., Doubleday, A., Austin, E., Marshall, J. D., Seto, E., Larson, T., & Sheppard, L. (2021). Design and evaluation of mobile monitoring campaigns for exposure assessment in epidemiologic cohorts. *Manuscript In Preparation*.
Blanco, M., Gassett, A., Seto, E., Larson, T., & Sheppard, L. (2019). Design and Implementation of a Mobile Monitoring Campaign for Epidemiology. *Poster Presented at the Health Effects Institute (HEI) Annual Conference*. https://www.healtheffects.org/sites/default/files/AnnualConferenceProgramBook2019_1.pdf
Block, M. L., & Calderón-Garcidueñas, L. (2009). Air pollution: mechanisms of neuroinflammation and CNS disease. *Trends in Neurosciences*, *32*(9), 506–516. https://doi.org/10.1016/j.tins.2009.05.009

Blondeau, P., Iordache, V., Poupard, O., Genin, D., & Allard, F. (2005). Relationship between outdoor and indoor air quality in eight French schools. *Indoor Air*, *15*(1), 2–12. https://doi.org/https://doi.org/10.1111/j.1600-0668.2004.00263.x

Brauer, M., Hoek, G., Vliet, P. van, Meliefste, K., Fischer, P., Gehring, U., Heinrich, J., Cyrys, J., Bellander, T., Lewne, M., & Brunekreef, B. (2003). Estimating Long-Term Average Particulate Air Pollution Concentrations: Application of Traffic Indicators and Geographic Information Systems. *Epidemiology (Cambridge, Mass.)*, *14*(2), 228–239. https://doi.org/10.1097/01.EDE.0000041910.49046.9B

Brenowitz, W. D., Keene, C. D., Hawes, S. E., Hubbard, R. A., Longstreth, W. T., Woltjer, R. L., Crane, P. K., Larson, E. B., & Kukull, W. A. (2017). Alzheimer's disease neuropathologic change, Lewy body disease, and vascular brain injury in clinic- and community-based samples. *Neurobiology of Aging*. https://doi.org/10.1016/j.neurobiolaging.2017.01.017

Brown, D. M., Wilson, M. R., MacNee, W., Stone, V., & Donaldson, K. (2001). Size-dependent proinflammatory effects of ultrafine polystyrene particles: A role for surface area and oxidative stress in the enhanced activity of ultrafines. *Toxicology and Applied Pharmacology*, *175*(3), 191–199. https://doi.org/10.1006/taap.2001.9240

Brunekreef, B., & Holgate, S. T. (2002). Air pollution and health. In *Lancet*. https://doi.org/10.1016/S0140-6736(02)11274-8

Bureau of Transportation Statistics. (2020). *Estimated U.S. Average Vehicle Emissions Rates per Vehicle by Vehicle Type Using Gasoline and Diesel*. US Department of Transportation. https://www.bts.gov/content/estimated-national-average-vehicle-emissions-rates-vehicle-vehicle-type-using-gasoline-and

Calderón-Garcidueñas, L., Azzarelli, B., Acuna, H., Garcia, R., Gambling, T. M., Osnaya, N., Monroy, S., Tizapantzi, M. D. R., Carson, J. L., Villarreal-Calderon, A., & Rewcastle, B. (2002). Air pollution and brain damage. *Toxicologic Pathology*, *30*(3), 373–389. https://doi.org/10.1080/01926230252929954

Calderón-Garcidueñas, L., Maronpot, R. R., Torres-Jardon, R., Henríquez-Roldán, C., Schoonhoven, R., Acuña-Ayala, H., Villarreal-Calderón, A., Nakamura, J., Fernando, R., Reed, W., Azzarelli, B., & Swenberg, J. A. (2003). DNA damage in nasal and brain tissues of canines exposed to air pollutants is associated with evidence of chronic brain inflammation and neurodegeneration. *Toxicologic Pathology*, *31*(5), 524–538. https://doi.org/10.1080/01926230390226645

Calderón-Garcidueñas, L., Mora-Tiscareño, A., Ontiveros, E., Gómez-Garza, G., Barragán-Mejía, G., Broadway, J., Chapman, S., Valencia-Salazar, G., Jewells, V., Maronpot, R. R., Henríquez-Roldán, C., Pérez-Guillé, B., Torres-Jardón, R., Herrit, L., Brooks, D., Osnaya-Brizuela, N., Monroy, M. E., González-Maciel, A., Reynoso-Robles, R., … Engle, R. W. (2008). Air pollution, cognitive deficits and brain abnormalities: A pilot study with children and dogs. *Brain and Cognition*, *68*(2), 117–127. https://doi.org/10.1016/j.bandc.2008.04.008

Calderón-Garcidueñas, L., Solt, A. C., Henríquez-Roldán, C., Torres-Jardón, R., Nuse, B., Herritt, L., Villarreal-Calderón, R., Osnaya, N., Stone, I., García, R., Brooks, D. M., González-Maciel, A., Reynoso-Robles, R., Delgado-Chávez, R., & Reed, W. (2008). Long-term air pollution exposure is associated with neuroinflammation, an altered innate immune response, disruption of the blood-brain barrier, ultrafine particulate deposition, and accumulation of amyloid β-42 and α-synuclein in children and young adult. *Toxicologic Pathology*, *36*(2),

289–310. https://doi.org/10.1177/0192623307313011
Carey, I. M., Anderson, H. R., Atkinson, R. W., Beevers, S. D., Cook, D. G., Strachan, D. P., Dajnak, D., Gulliver, J., & Kelly, F. J. (2018). Are noise and air pollution related to the incidence of dementia? A cohort study in London, England. *BMJ Open*, *8*(9), e022404. https://doi.org/10.1136/bmjopen-2018-022404
Carr, D., von Ehrenstein, O., Weiland, S., Wagner, C., Wellie, O., Nicolai, T., & von Mutius, E. (2002). Modeling annual benzene, toluene, NO2, and soot concentrations on the basis of road traffic characteristics. *Environmental Research*, *90*(2), 111–118. https://doi.org/10.1006/enrs.2002.4393
Cattani, G., Gaeta, A., Di Menno di Bucchianico, A., De Santis, A., Gaddi, R., Cusano, M., Ancona, C., Badaloni, C., Forastiere, F., Gariazzo, C., Sozzi, R., Inglessis, M., Silibello, C., Salvatori, E., Manes, F., & Cesaroni, G. (2017). Development of land-use regression models for exposure assessment to ultrafine particles in Rome, Italy. *Atmospheric Environment*, *156*, 52–60. https://doi.org/https://doi.org/10.1016/j.atmosenv.2017.02.028
Cesaroni, G., Porta, D., Badaloni, C., Stafoggia, M., Eeftens, M., Meliefste, K., & Forastiere, F. (2012). Nitrogen dioxide levels estimated from land use regression models several years apart and association with mortality in a large cohort study. *Environmental Health*, *11*(1), 48. https://doi.org/10.1186/1476-069X-11-48
Chang, K. H., Chang, M. Y., Muo, C. H., Wu, T. N., Chen, C. Y., & Kao, C. H. (2014). Increased risk of dementia in patients exposed to nitrogen dioxide and carbon monoxide: A population-based retrospective cohort study. *PLoS ONE*, *9*(8). https://doi.org/10.1371/journal.pone.0103078
Chen, H., Kwong, J. C., Copes, R., Hystad, P., van Donkelaar, A., Tu, K., Brook, J. R., Goldberg, M. S., Martin, R. V., Murray, B. J., Wilton, A. S., Kopp, A., & Burnett, R. T. (2017). Exposure to ambient air pollution and the incidence of dementia: A population-based cohort study. *Environment International*, *108*, 271–277. https://doi.org/10.1016/j.envint.2017.08.020
Chen, H., Kwong, J. C., Copes, R., Tu, K., Villeneuve, P. J., van Donkelaar, A., Hystad, P., Martin, R. V., Murray, B. J., Jessiman, B., Wilton, A. S., Kopp, A., & Burnett, R. T. (2017). Living near major roads and the incidence of dementia, Parkinson's disease, and multiple sclerosis: a population-based cohort study. *The Lancet*, *389*(10070), 718–726. https://doi.org/10.1016/S0140-6736(16)32399-6
Chen, H., Kwong, J. C., Copes, R., Villeneuve, P. J., Goldberg, M. S., Ally, S. L., Weichenthal, S., van Donkelaar, A., Jerrett, M., Martin, R. V., Brook, J. R., Kopp, A., & Burnett, R. T. (2017). Cohort Profile: The ONtario Population Health and Environment Cohort (ONPHEC). *International Journal of Epidemiology*, *46*(2), dyw030. https://doi.org/10.1093/ije/dyw030
Chêne, G., Beiser, A., Au, R., Preis, S. R., Wolf, P. A., Dufouil, C., & Seshadri, S. (2015). Gender and incidence of dementia in the Framingham Heart Study from mid-adult life. *Alzheimer's & Dementia*, *11*(3), 310–320. https://doi.org/https://doi.org/10.1016/j.jalz.2013.10.005
Chi, G. C., Hajat, A., Bird, C. E., Cullen, M. R., Griffin, B. A., Miller, K. A., Shih, R. A., Stefanick, M. L., Vedal, S., & Whitsel, E. A. (2016). Individual and neighborhood socioeconomic status and the association between air pollution and cardiovascular disease. *Environmental Health Perspectives*, *124*(12), 1840–1847.
Clark, L. P., Millet, D. B., & Marshall, J. D. (2014). National patterns in environmental injustice and inequality: Outdoor NO2 air pollution in the United States. *PLoS ONE*.

https://doi.org/10.1371/journal.pone.0094431
Clark, L. P., Millet, D. B., & Marshall, J. D. (2017). Changes in transportation-related air pollution exposures by race-ethnicity and socioeconomic status: Outdoor nitrogen dioxide in the United States in 2000 and 2010. *Environmental Health Perspectives*, *125*(9), 97012.
Clark, T., Altman, D. G., & De Stavola, B. L. (2002). Quantification of the completeness of follow-up. *Lancet*, *359*(9314), 1309–1310. https://doi.org/10.1016/S0140-6736(02)08272-7
Cleveland, W. S., & Guarino, R. (1976). Some Robust Statistical Procedures and Their Application to Air Pollution Data. *Technometrics*, *18*(4), 401–409. https://doi.org/10.2307/1268655
Clifford, A., Lang, L., Chen, R., Anstey, K. J., & Seaton, A. (2016). Exposure to air pollution and cognitive functioning across the life course - A systematic literature review. In *Environmental Research* (Vol. 147). https://doi.org/10.1016/j.envres.2016.01.018
Cohen, M. A., Adar, S. D., Allen, R. W., Avol, E., Curl, C. L., Gould, T., Hardie, D., Ho, A., Kinney, P., Larson, T. V., Sampson, P., Sheppard, L., Stukovsky, K. D., Swan, S. S., Liu, L. J. S., & Kaufman, J. D. (2009). Approach to estimating participant pollutant exposures in the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). *Environmental Science and Technology*. https://doi.org/10.1021/es8030837
Correia, A. W., Pope 3rd, C. A., Dockery, D. W., Wang, Y., Ezzati, M., & Dominici, F. (2013). Effect of air pollution control on life expectancy in the United States: an analysis of 545 U.S. counties for the period from 2000 to 2007. *Epidemiology (Cambridge, Mass.)*, *24*(1), 23–31. https://doi.org/10.1097/EDE.0b013e3182770237
de Prado Bert, P., Mercader, E. M. H., Pujol, J., Sunyer, J., & Mortamais, M. (2018). The Effects of Air Pollution on the Brain: a Review of Studies Interfacing Environmental Epidemiology and Neuroimaging. *Current Environmental Health Reports*, *5*(3), 351–364. https://doi.org/10.1007/s40572-018-0209-9
Dodson, R. E., Houseman, E. A., Morin, B., & Levy, J. I. (2009). An analysis of continuous black carbon concentrations in proximity to an airport and major roadways. *Atmospheric Environment*, *43*(24), 3764–3773.
Dominici, F., Peng, R. D., Bell, M. L., Pham, L., McDermott, A., Zeger, S. L., & Samet, J. M. (2006). Fine Particulate Air Pollution and Hospital Admission for Cardiovascular and Respiratory Diseases. *JAMA*, *295*(10), 1127–1134. https://doi.org/10.1001/jama.295.10.1127
Donaldson, K. (2001). Ultrafine particles. *Occupational and Environmental Medicine*, *58*(3), 211–216. https://doi.org/10.1136/oem.58.3.211
Dublin, S., Anderson, M. L., Haneuse, S. J., Heckbert, S. R., Crane, P. K., Breitner, J. C. S., McCormick, W., Bowen, J. D., Teri, L., McCurry, S. M., & Larson, E. B. (2011). Atrial Fibrillation and Risk of Dementia: A Prospective Cohort Study. *Journal of the American Geriatrics Society*, *59*(8), 1369–1375. https://doi.org/10.1111/j.1532-5415.2011.03508.x
Eeftens, M., Beelen, R., Fischer, P., Brunekreef, B., Meliefste, K., & Hoek, G. (2011). Stability of measured and modelled spatial contrasts in NO₂ over time. *Occupational and Environmental Medicine*, *68*(10), 765 LP – 770. https://doi.org/10.1136/oem.2010.061135
Enayati Ahangar, F., Heist, D., Perry, S., & Venkatram, A. (2017). Reduction of air pollution levels downwind of a road with an upwind noise barrier. *Atmospheric Environment*, *155*, 1–10. https://doi.org/10.1016/j.atmosenv.2017.02.001

Esri. (2019). *What is the ArcGIS Network Analyst extension?—ArcGIS Pro | ArcGIS Desktop*. https://pro.arcgis.com/en/pro-app/help/analysis/networks/what-is-network-analyst-.htm

Farrell, W., Weichenthal, S., Goldberg, M., Valois, M.-F., Shekarrizfard, M., & Hatzopoulou, M. (2016). Near roadway air pollution across a spatially extensive road and cycling network. *Environmental Pollution*, *212*, 498–507. https://doi.org/https://doi.org/10.1016/j.envpol.2016.02.041

Fujita, E. M., Campbell, D. E., Arnott, W. P., Johnson, T., & Ollison, W. (2014). Concentrations of mobile source air pollutants in urban microenvironments. *Journal of the Air and Waste Management Association*, *64*(7), 743–758. https://doi.org/10.1080/10962247.2013.872708

Gan, W. Q., Tamburic, L., Davies, H. W., Demers, P. A., Koehoorn, M., & Brauer, M. (2010). Changes in Residential Proximity to Road Traffic and the Risk of Death From Coronary Heart Disease. *Epidemiology*, *21*(5). https://journals.lww.com/epidem/Fulltext/2010/09000/Changes_in_Residential_Proximity_to_Road_Traffic.18.aspx

González-Maciel, A., Reynoso-Robles, R., Torres-Jardón, R., Mukherjee, P. S., & Calderón-Garcidueñas, L. (2017). Combustion-Derived Nanoparticles in Key Brain Target Cells and Organelles in Young Urbanites: Culprit Hidden in Plain Sight in Alzheimer's Disease Development. *Journal of Alzheimer's Disease*, *59*(1), 189–208. https://doi.org/10.3233/JAD-170012

Google. (2019). *Earth Engine Data Catalog*. https://developers.google.com/earth-engine/datasets/

Gouveia, N., & Fletcher, T. (2000). Time series analysis of air pollution and mortality: effects by cause, age and socioeconomic status. *Journal of Epidemiology and Community Health*, *54*(10), 750 LP – 755. https://doi.org/10.1136/jech.54.10.750

Hagler, G. S. W., Baldauf, R. W., Thoma, E. D., Long, T. R., Snow, R. F., Kinsey, J. S., Oudejans, L., & Gullett, B. K. (2009). Ultrafine particles near a major roadway in Raleigh, North Carolina: Downwind attenuation and correlation with traffic-related pollutants. *Atmospheric Environment*, *43*(6), 1229–1234. https://doi.org/10.1016/j.atmosenv.2008.11.024

Hajat, A., Diez-Roux, A. V, Adar, S. D., Auchincloss, A. H., Lovasi, G. S., O'Neill, M. S., Sheppard, L., & Kaufman, J. D. (2013). Air pollution and individual and neighborhood socioeconomic status: evidence from the Multi-Ethnic Study of Atherosclerosis (MESA). *Environmental Health Perspectives*, *121*(11–12), 1325–1333.

Hajat, A., Hsia, C., & O'Neill, M. S. (2015). Socioeconomic Disparities and Air Pollution Exposure: a Global Review. *Current Environmental Health Reports*, *2*(4), 440–450. https://doi.org/10.1007/s40572-015-0069-5

Hankey, S., & Marshall, J. D. (2015a). On-bicycle exposure to particulate air pollution: Particle number, black carbon, PM2.5, and particle size. *Atmospheric Environment*, *122*, 65–73. https://doi.org/10.1016/j.atmosenv.2015.09.025

Hankey, S., & Marshall, J. D. (2015b). Land Use Regression Models of On-Road Particulate Air Pollution (Particle Number, Black Carbon, PM2.5, Particle Size) Using Mobile Monitoring. *Environmental Science and Technology*, *49*(15), 9194–9202. https://doi.org/10.1021/acs.est.5b01209

Hatzopoulou, M., Valois, M. F., Levy, I., Mihele, C., Lu, G., Bagg, S., Minet, L., & Brook, J. (2017).

Robustness of Land-Use Regression Models Developed from Mobile Air Pollutant Measurements. *Environmental Science and Technology*, *51*(7), 3938–3947. https://doi.org/10.1021/acs.est.7b00366

Hatzopoulou, M., Weichenthal, S., Dugum, H., Pickett, G., Miranda-Moreno, L., Kulka, R., Andersen, R., & Goldberg, M. (2013). The impact of traffic volume, composition, and road geometry on personal air pollution exposures among cyclists in Montreal, Canada. *Journal of Exposure Science & Environmental Epidemiology*, *23*(1), 46–51. https://doi.org/10.1038/jes.2012.85

HealthLine. (2016). *Dementia Risk Factors: Cholesterol, Diabetes, and More*. https://www.healthline.com/health/dementia-risk-factors#medical-risk-factors

HEI. (2010, January). *Traffic-related air pollution: a critical review of the literature on emissions, exposure, and health effects*. Health Effects Institute. https://doi.org/10.1016/j.jcis.2012.08.055

Henderson, S. B., Beckerman, B., Jerrett, M., & Brauer, M. (2007). Application of Land Use Regression to Estimate Long-Term Concentrations of Traffic-Related Nitrogen Oxides and Fine Particulate Matter. *Environmental Science & Technology*, *41*(7), 2422–2428. https://doi.org/10.1021/es0606780

Heusinkveld, H. J., Wahle, T., Campbell, A., Westerink, R. H. S., Tran, L., Johnston, H., Stone, V., Cassee, F. R., Schins, R. P. F., H.J., H., T., W., A., C., R.H.S., W., L., T., H., J., V., S., F.R., C., & R.P.F., S. (2016). Neurodegenerative and neurological disorders by small inhaled particles. *NeuroToxicology*, *56*. https://doi.org/10.1016/j.neuro.2016.07.007

Hochadel, M., Heinrich, J., Gehring, U., Morgenstern, V., Kuhlbusch, T., Link, E., Wichmann, H.-E., & Krämer, U. (2006). Predicting long-term average concentrations of traffic-related air pollutants using GIS-based information. *Atmospheric Environment (1994)*, *40*(3), 542–553. https://doi.org/10.1016/j.atmosenv.2005.09.067

Hoek, G., Beelen, R., de Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P., & Briggs, D. (2008). A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric Environment*, *42*(33), 7561–7578. https://doi.org/10.1016/j.atmosenv.2008.05.057

Hoek, G., Beelen, R., Kos, G., Dijkema, M., Zee, S. C. van der, Fischer, P. H., & Brunekreef, B. (2011). Land Use Regression Model for Ultrafine Particles in Amsterdam. *Environmental Science & Technology*, *45*(2), 622–628. https://doi.org/10.1021/es1023042

Howard, V. (2009). Statement of Evidence: Particulate Emissions and Health (An Bord Plenala, on Proposed Ringaskiddy Waste-to-Energy Facility). In *Durham Environmental Watch*. http://www.durhamenvironmentwatch.org/Incinerator Health/CVHRingaskiddyEvidenceFinal1.pdf

Howell, J. C., Watts, K. D., Parker, M. W., Wu, J., Kollhoff, A., Wingo, T. S., Dorbin, C. D., Qiu, D., & Hu, W. T. (2017). Race modifies the relationship between cognition and Alzheimer's disease cerebrospinal fluid biomarkers. *Alzheimer's Research and Therapy*, *9*(1), 1–10. https://doi.org/10.1186/s13195-017-0315-1

Hu, Y., & Zhao, B. (2020). Relationship between indoor and outdoor NO2: A review. *Building and Environment*, *180*, 106909. https://doi.org/https://doi.org/10.1016/j.buildenv.2020.106909

Hudda, N., Gould, T., Hartin, K., Larson, T. V., & Fruin, S. A. (2014). Emissions from an

international airport increase particle number concentrations 4-fold at 10 km downwind. *Environmental Science and Technology*, *48*(12), 6628–6635. https://doi.org/10.1021/es5001566

Humboldt State University. (2019). *Radiometric Calibration and Corrections. Introduction to Remote Sensing*. Geospatial Online.

Hystad, P., Demers, P. A., Johnson, K. C., Brook, J., van Donkelaar, A., Lamsal, L., Martin, R., & Brauer, M. (2012). Spatiotemporal air pollution exposure assessment for a Canadian population-based lung cancer case-control study. *Environmental Health*, *11*(1), 22. https://doi.org/10.1186/1476-069X-11-22

Jack, C. R., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., Holtzman, D. M., Jagust, W., Jessen, F., Karlawish, J., Liu, E., Molinuevo, J. L., Montine, T., Phelps, C., Rankin, K. P., Rowe, C. C., Scheltens, P., Siemers, E., Snyder, H. M., … Silverberg, N. (2018). NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's and Dementia*, *14*(4), 535–562. https://doi.org/10.1016/j.jalz.2018.02.018

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to Statistical Learning. In *Springer*. Springer. https://doi.org/10.1007/978-1-4614-7138-7

Janssen, N. A. H., Gerlofs-Nijland, M. E., Lanki, T., Salonen, R. O., Cassee, F., Hoek, G., Fischer, P., Brunekreef, B., & Krzyzanowski, M. (2017). *Health effects of black carbon (2012)*. World Health Organization.

Jarvis, D. J., Adamkiewicz, G., Heroux, M.-E., Rapp, R., & Kelly, F. J. (2010). Nitrogen dioxide. In *WHO Guidelines for Indoor Air Quality: Selected Pollutants.* World Health Organization. https://www.ncbi.nlm.nih.gov/books/NBK138707/

Jerrett, M., Arain, A., Kanaroglou, P., Beckerman, B., Potoglou, D., Sahsuvaroglu, T., Morrison, J., & Giovis, C. (2005). A review and evaluation of intraurban air pollution exposure models. *Journal of Exposure Analysis and Environmental Epidemiology*, *15*(2), 185–204. https://doi.org/10.1038/sj.jea.7500388

Jones, R. R., Hoek, G., Fisher, J. A., Hasheminassab, S., Wang, D., Ward, M. H., Sioutas, C., Vermeulen, R., & Silverman, D. T. (2020). Land use regression models for ultrafine particles, fine particles, and black carbon in Southern California. *Science of The Total Environment*, *699*, 134234. https://doi.org/https://doi.org/10.1016/j.scitotenv.2019.134234

Jung, C. R., Lin, Y. T., & Hwang, B. F. (2015). Ozone, particulate matter, and newly diagnosed Alzheimer's disease: A population-based cohort study in Taiwan. *Journal of Alzheimer's Disease*, *44*(2), 573–584. https://doi.org/10.3233/JAD-140855

Jung, K. H., Bernabé, K., Moors, K., Yan, B., Chillrud, S. N., Whyatt, R., Camann, D., Kinney, P. L., Perera, F. P., & Miller, R. L. (2011). Effects of floor level and building type on residential levels of outdoor and indoor polycyclic aromatic hydrocarbons, black carbon, and particulate matter in New York City. *Atmosphere*, *2*(2), 96–109. https://doi.org/10.3390/atmos2020096

Karner, A. A., Eisinger, D. S., & Niemeier, D. A. (2010). Near-roadway air quality: Synthesizing the findings from real-world data. *Environmental Science and Technology*, *44*(14), 5334–5344. https://doi.org/10.1021/es100008x

Keller, J. P., Olives, C., Kim, S. Y., Sheppard, L., Sampson, P. D., Szpiro, A. A., Oron, A. P., Lindström, J., Vedal, S., & Kaufman, J. D. (2015). A unified spatiotemporal modeling

approach for predicting concentrations of multiple air pollutants in the multi-ethnic study of atherosclerosis and air pollution. *Environmental Health Perspectives*, *123*(4), 301–309. https://doi.org/10.1289/ehp.1408145

Kerckhoffs, J., Hoek, G., Messier, K. P., Brunekreef, B., Meliefste, K., Klompmaker, J. O., & Vermeulen, R. (2016). Comparison of ultrafine particle and black carbon concentration predictions from a mobile and short-term stationary land-use regression model. *Environmental Science and Technology*, *50*(23), 12894–12902. https://doi.org/10.1021/acs.est.6b03476

Kerckhoffs, J., Hoek, G., Vlaanderen, J., van Nunen, E., Messier, K., Brunekreef, B., Gulliver, J., & Vermeulen, R. (2017). Robustness of intra urban land-use regression models for ultrafine particles and black carbon based on mobile monitoring. *Environmental Research*, *159*(May), 500–508. https://doi.org/10.1016/j.envres.2017.08.040

Khalek, I. A., Blanks, M. G., Merritt, P. M., & Zielinska, B. (2015). Regulated and unregulated emissions from modern 2010 emissions-compliant heavy-duty on-highway diesel engines. *Journal of the Air & Waste Management Association*, *65*(8), 987–1001. https://doi.org/10.1080/10962247.2015.1051606

Kilian, J., & Kitazawa, M. (2018). The emerging risk of exposure to air pollution on cognitive decline and Alzheimer ' s disease e Evidence from epidemiological and animal studies. *Biomedical Journal*, *41*(3), 141–162. https://doi.org/10.1016/j.bj.2018.06.001

Killin, L. O. J., Starr, J. M., Shiue, I. J., & Russ, T. C. (2016). Environmental risk factors for dementia: a systematic review. *BMC Geriatrics*, *16*(1), 1–28. https://doi.org/10.1186/s12877-016-0342-y

Kim, E., Hopke, P. K., Larson, T. V., Maykut, N. N., & Lewtas, J. (2004). Factor analysis of Seattle fine particles. *Aerosol Science and Technology*, *38*(7), 724–738. https://doi.org/10.1080/02786820490490119

Kim, S.-Y., Bechle, M., Hankey, S., Sheppard, L., Szpiro, A. A., & Marshall, J. D. (2020). Concentrations of criteria pollutants in the contiguous US, 1979–2015: Role of prediction model parsimony in integrated empirical geographic regression. *PloS One*, *15*(2), e0228535.

Kim, S. Y., Olives, C., Sheppard, L., Sampson, P. D., Larson, T. V., Keller, J. P., & Kaufman, J. D. (2017). Historical prediction modeling approach for estimating long-term concentrations of PM2.5 in cohort studies before the 1999 implementation of widespread monitoring. *Environmental Health Perspectives*. https://doi.org/10.1289/EHP131

Kivimäki, M., Luukkonen, R., Batty, G. D., Ferrie, J. E., Pentti, J., Nyberg, S. T., Shipley, M. J., Alfredsson, L., Fransson, E. I., Goldberg, M., Knutsson, A., Koskenvuo, M., Kuosma, E., Nordin, M., Suominen, S. B., Theorell, T., Vuoksimaa, E., Westerholm, P., Westerlund, H., … Jokela, M. (2018). Body mass index and risk of dementia: Analysis of individual-level data from 1.3 million individuals. *Alzheimer's and Dementia*, *14*(5), 601–609. https://doi.org/10.1016/j.jalz.2017.09.016

Kleinbaum, D. G., & Klein, M. (2012). *Survival Analysis : A Self-Learning Text, Third Edition*. Springer Science+Business Media, LLC.

Klepeis, N. E., Nelson, W. C., Ott, W. R., Robinson, J. P., Tsang, A. M., Switzer, P., Behar, J. V., Hern, S. C., & Engelmann, W. H. (2001). The National Human Activity Pattern Survey (NHAPS): A resource for assessing exposure to environmental pollutants. *Journal of*

Exposure Analysis and Environmental Epidemiology, *11*, 231–252. https://doi.org/10.1038/sj.jea.7500165

Koupal, J., Beardsley, M., Brzezinski, D., Warila, J., & Faler, W. (2011). *U.S. EPA's MOVES2010 vehicle emission model: overview and considerations for international application*.

Kozawa, K. H., Fruin, S. A., & Winer, A. M. (2009). Near-road air pollution impacts of goods movement in communities adjacent to the Ports of Los Angeles and Long Beach. *Atmospheric Environment*, *43*(18), 2960–2970.

Kukull, W. A., Higdon, R., Bowen, J. D., McCormick, W. C., Teri, L., Schellenberg, G. D., Van Belle, G., Jolley, L., & Larson, E. B. (2002). Dementia and Alzheimer disease incidence: A prospective cohort study. *Archives of Neurology*, *59*(11), 1737–1746. https://doi.org/10.1001/archneur.59.11.1737

Kulick, E. R., Elkind, M. S. V., Boehme, A. K., Joyce, N. R., Schupf, N., Kaufman, J. D., Mayeux, R., Manly, J. J., & Wellenius, G. A. (2020). Long-term exposure to ambient air pollution, APOE-ε4 status, and cognitive decline in a cohort of older adults in northern Manhattan. *Environment International*, *136*(October 2019), 105440. https://doi.org/10.1016/j.envint.2019.105440

Lack, D. A., & Corbett, J. J. (2012). Black carbon from ships: a review of the effects of ship speed, fuel quality and exhaust gas scrubbing. *Atmospheric Chemistry & Physics*, *12*(9).

Lane, K. J., Levy, J. I., Scammell, M. K., Peters, J. L., Patton, A. P., Reisner, E., Lowe, L., Zamore, W., Durant, J. L., & Brugge, D. (2016). Association of modeled long-term personal exposure to ultrafine particles with inflammatory and coagulation biomarkers. *Environment International*, *92–93*, 173–182. https://doi.org/https://doi.org/10.1016/j.envint.2016.03.013

Larkin, A., Geddes, J. A., Martin, R. V, Xiao, Q., Liu, Y., Marshall, J. D., Brauer, M., & Hystad, P. (2017). Global land use regression model for nitrogen dioxide air pollution. *Environmental Science & Technology*, *51*(12), 6957–6964.

Larson, T. (2017). *Fine Particles in Seattle's Air: What, When, Where, Why and So What? Steve and Sylvia Burges Lecture.* https://www.ce.washington.edu/news/lecture/burges

Larson, T, Gould, T., Riley, E. A., Austin, E., Fintzi, J., Sheppard, L., Yost, M., & Simpson, C. (2017). Ambient air quality measurements from a continuously moving mobile platform: Estimation of area-wide, fuel-based, mobile source emission factors using absolute principal component scores. *Atmospheric Environment*, *152*, 201–211. https://doi.org/10.1016/j.atmosenv.2016.12.037

Larson, Timothy, Henderson, S. B., & Brauer, M. (2009). Mobile monitoring of particle light absorption coefficient in an urban area as a basis for land use regression. *Environmental Science and Technology*, *43*(13), 4672–4678. https://doi.org/10.1021/es803068e

Larson, Timothy, Su, J., Baribeau, A.-M., Buzzelli, M., Setton, E., & Brauer, M. (2007). A spatial model of urban winter woodsmoke concentrations. *Environmental Science & Technology*, *41*(7), 2429–2436.

Lee, B. J., Kim, B., & Lee, K. (2014). Air pollution exposure and cardiovascular disease. *Toxicological Research*, *30*(2), 71–75. https://doi.org/10.5487/TR.2014.30.2.071

Levy, I., Levin, N., Yuval, Y., Schwartz, J. D., & Kark, J. D. (2015). Back-extrapolating a land use regression model for estimating past exposures to traffic-related air pollution. *Environmental Science and Technology*. https://doi.org/10.1021/es505707e

Li, N., Sioutas, C., Cho, A., Schmitz, D., Misra, C., Sempf, J., Wang, M., Oberley, T., Froines, J., & Nel, A. (2003). Ultrafine particulate pollutants induce oxidative stress and mitochondrial damage. *Environmental Health Perspectives*, *111*(4), 455–460. https://doi.org/10.1289/ehp.6000

Li, P., Jiang, L., & Feng, Z. (2014). Cross-comparison of vegetation indices derived from landsat-7 enhanced thematic mapper plus (ETM+) and landsat-8 operational land imager (OLI) sensors. *Remote Sensing*, *6*. https://doi.org/10.3390/rs6010310

Lindhjem, C. E., Pollack, A. K., DenBleyker, A., & Shaw, S. L. (2012). Effects of improved spatial and temporal modeling of on-road vehicle emissions. *Journal of the Air and Waste Management Association*, *62*(4), 471–484. https://doi.org/10.1080/10962247.2012.658955

Lindström, J., Szpiro, A. A., Sampson, P. D., Oron, A. P., Richards, M., Larson, T. V., & Sheppard, L. (2014). A flexible spatio-temporal model for air pollution with spatial and spatio-temporal covariates. *Environmental and Ecological Statistics*, *21*(3), 411–433. https://doi.org/10.1007/s10651-013-0261-4

Liu, S., Hua, S., Wang, K., Qiu, P., Liu, H., Wu, B., Shao, P., Liu, X., Wu, Y., Xue, Y., Hao, Y., & Tian, H. (2018). Spatial-temporal variation characteristics of air pollution in Henan of China: Localized emission inventory, WRF/Chem simulations and potential source contribution analysis. *Science of the Total Environment*, *624*, 396–406. https://doi.org/10.1016/j.scitotenv.2017.12.102

Loop, M. S., Kent, S. T., Al-Hamdan, M. Z., Crosson, W. L., Estes, S. M., Estes, M. G., Quattrochi, D. A., Hemmings, S. N., Wadley, V. G., & McClure, L. A. (2013). Fine Particulate Matter and Incident Cognitive Impairment in the REasons for Geographic and Racial Differences in Stroke (REGARDS) Cohort. *PLoS ONE*, *8*(9), 1–9. https://doi.org/10.1371/journal.pone.0075001

Lundborg, M., Johard, U., Låstbom, L., Gerde, P., & Camner, P. (2001). Human Alveolar Macrophage Phagocytic Function is Impaired by Aggregates of Ultrafine Carbon Particles. *Environmental Research*, *86*(3), 244–253. https://doi.org/10.1006/enrs.2001.4269

Manigrasso, M., Natale, C., Vitali, M., Protano, C., & Avino, P. (2017). Pedestrians in Traffic Environments: Ultrafine Particle Respiratory Doses. *International Journal of Environmental Research and Public Health*, *14*(3), 288. https://doi.org/10.3390/ijerph14030288

Martine, V., David, M., Sandra, M., Payam, D., Anna, S., Judith, R., & Mark, N. (2011). Ambient Air Pollution and Risk of Congenital Anomalies: A Systematic Review and Meta-analysis. *Environmental Health Perspectives*, *119*(5), 598–606. https://doi.org/10.1289/ehp.1002946

Matthews, K. A., Xu, W., Gaglioti, A. H., Holt, J. B., Croft, J. B., Mack, D., & McGuire, L. C. (2019). Racial and ethnic estimates of Alzheimer's disease and related dementias in the United States (2015–2060) in adults aged ≥65 years. *Alzheimer's and Dementia*, *15*(1), 17–24. https://doi.org/10.1016/j.jalz.2018.06.3063

Mayo Clinic. (2019). *Dementia*.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease. *Neurology*, *34*(7), 939 LP – 939. https://doi.org/10.1212/WNL.34.7.939

Meng, J., Li, C., Martin, R. V., van Donkelaar, A., Hystad, P., & Brauer, M. (2019). Estimated

Long-Term (1981–2016) Concentrations of Ambient Fine Particulate Matter across North America from Chemical Transport Modeling, Satellite Remote Sensing, and Ground-Based Measurements. *Environmental Science & Technology*. https://doi.org/10.1021/acs.est.8b06875

Mercer, L. D., Szpiro, A. A., Sheppard, L., Lindström, J., Adar, S. D., Allen, R. W., Avol, E. L., Oron, A. P., Larson, T., Liu, L. J. S., & Kaufman, J. D. (2011). Comparing universal kriging and land-use regression for predicting concentrations of gaseous oxides of nitrogen (NOx) for the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). *Atmospheric Environment*. https://doi.org/10.1016/j.atmosenv.2011.05.043

MESA Air. (2019). *Data Organization and Operating Procedures (DOOP) for the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air) and Associated Studies*. MESA Air. https://www.uwchscc.org/MESAAP/Documents/MESAAirDOOP.pdf

Messier, K. P., Chambliss, S. E., Gani, S., Alvarez, R., Brauer, M., Choi, J. J., Hamburg, S. P., Kerckhoffs, J., Lafranchi, B., Lunden, M. M., Marshall, J. D., Portier, C. J., Roy, A., Szpiro, A. A., Vermeulen, R. C. H., & Apte, J. S. (2018). Mapping Air Pollution with Google Street View Cars: Efficient Approaches with Mobile Monitoring and Land Use Regression. *Environmental Science and Technology*, *52*(21), 12563–12572. https://doi.org/10.1021/acs.est.8b03395

Mevik. (2019). *Introduction to the pls Package*. https://cran.r-project.org/web/packages/pls/vignettes/pls-manual.pdf

Minet, L., Gehr, R., & Hatzopoulou, M. (2017). Capturing the sensitivity of land-use regression models to short-term mobile monitoring campaigns using air pollution micro-sensors. *Environmental Pollution*, *230*, 280–290. https://doi.org/10.1016/j.envpol.2017.06.071

Minet, L., Liu, R., Valois, M. F., Xu, J., Weichenthal, S., & Hatzopoulou, M. (2018). Development and Comparison of Air Pollution Exposure Surfaces Derived from On-Road Mobile Monitoring and Short-Term Stationary Sidewalk Measurements. *Environmental Science and Technology*, *52*(6), 3512–3519. https://doi.org/10.1021/acs.est.7b05059

Mölter, A., Lindley, S., de Vocht, F., Simpson, A., & Agius, R. (2010a). Modelling air pollution for epidemiologic research—Part I: A novel approach combining land use regression and air dispersion. *Science of the Total Environment*, *408*(23), 5862–5869.

Mölter, A., Lindley, S., de Vocht, F., Simpson, A., & Agius, R. (2010b). Modelling air pollution for epidemiologic research – Part II: Predicting temporal variation through land use regression. *Science of The Total Environment*, *409*(1), 211–217. https://doi.org/https://doi.org/10.1016/j.scitotenv.2010.10.005

Montagne, D. R., Hoek, G., Klompmaker, J. O., Wang, M., Meliefste, K., & Brunekreef, B. (2015). Land Use Regression Models for Ultrafine Particles and Black Carbon Based on Short-Term Monitoring Predict Past Spatial Variation. *Environmental Science and Technology*, *49*(14), 8712–8720. https://doi.org/10.1021/es505791g

Morgenstern, V., Zutavern, A., Cyrys, J., Brockow, I., Gehring, U., Koletzko, S., Bauer, C. P., Reinhardt, D., Wichmann, H.-E., & Heinrich, J. (2007). Respiratory health and individual estimated exposure to traffic-related air pollutants in a cohort of young children. *Occupational and Environmental Medicine*, *64*(1), 8–16. https://doi.org/10.1136/oem.2006.028241

NASA. (2020). *Vegetation Indices. MODIS Land*. National Aeronautics and Space

Adminnistration. https://modis-land.gsfc.nasa.gov/vi.html
National Institute of Aging. (2014, May). *Number of Alzheimer's deaths found to be underreported*. https://www.nia.nih.gov/news/number-alzheimers-deaths-found-be-underreported
Oberdorster, G., Ferin, J., & Lehnert, B. E. (1994). Correlation between particle size, in vivo particle persistence, and lung injury. *Environmental Health Perspectives*.
Oudin, A., Forsberg, B., Adolfsson, A. N., Lind, N., Modig, L., Nordin, M., Nordin, S., Adolfsson, R., & Nilsson, L. G. (2016). Traffic-related air pollution and dementia incidence in Northern Sweden: A longitudinal study. *Environmental Health Perspectives*, *124*(3), 306–312. https://doi.org/10.1289/ehp.1408322
Özkaynak, H., Baxter, L. K., Dionisio, K. L., & Burke, J. (2013). Air pollution exposure prediction approaches used in air pollution epidemiology studies. *Journal of Exposure Science and Environmental Epidemiology*, *23*(6), 566–572. https://doi.org/10.1038/jes.2013.15
Park, Y. M., & Kwan, M.-P. (2017). Individual exposure estimates may be erroneous when spatiotemporal variability of air pollution and human mobility are ignored. *Health & Place*, *43*, 85–94. https://doi.org/https://doi.org/10.1016/j.healthplace.2016.10.002
Patton, A. P., Collins, C., Naumova, E. N., Zamore, W., Brugge, D., & Durant, J. L. (2014). An hourly regression model for ultrafine particles in a near-highway urban area. *Environmental Science & Technology*, *48*(6), 3272–3280. https://doi.org/10.1021/es404838k
Patton, A. P., Milando, C., Durant, J. L., & Kumar, P. (2017). Assessing the Suitability of Multiple Dispersion and Land Use Regression Models for Urban Traffic-Related Ultrafine Particles. *Environmental Science and Technology*, *51*(1), 384–392. https://doi.org/10.1021/acs.est.6b04633
Patton, A. P., Perkins, J., Zamore, W., Levy, J. I., Brugge, D., & Durant, J. L. (2014). Spatial and temporal differences in traffic-related air pollution in three urban neighborhoods near an interstate highway. *Atmospheric Environment*. https://doi.org/10.1016/j.atmosenv.2014.09.072
Patton, A. P., Zamore, W., Naumova, E. N., Levy, J. I., Brugge, D., & Durant, J. L. (2015). Transferability and generalizability of regression models of ultrafine particles in urban neighborhoods in the boston area. *Environmental Science and Technology*, *49*(10), 6051–6060. https://doi.org/10.1021/es5061676
Peters, R., Ee, N., Peters, J., Booth, A., Mudway, I., & Anstey, K. J. (2019). Air Pollution and Dementia: A Systematic Review. *Journal of Alzheimer's Disease*, *70*, S145–S163. https://doi.org/10.3233/JAD-180631
Power, M. C., Adar, S. D., Yanosky, J. D., & Weuve, J. (2016). Exposure to air pollution as a potential contributor to cognitive function, cognitive decline, brain imaging, and dementia: A systematic review of epidemiologic research. *NeuroToxicology*, *56*, 235–253. https://doi.org/10.1016/j.neuro.2016.06.004
Power, M. C., Weisskopf, M. G., Alexeeff, S. E., Coull, B. A., Avron, S., & Schwartz, J. (2011). Traffic-related air pollution and cognitive function in a cohort of older men. *Environmental Health Perspectives*, *119*(5), 682–687. https://doi.org/10.1289/ehp.1002767
Presto, A. (2020). *Trends in Inter- and Intraurban Ultrafine Particle Trends in Inter Levels in the U.S.: Exposure Implications. Health Effects Institute Annual Conference*.

https://www.healtheffects.org/sites/default/files/presto-ultrafine-particle-hei-2020.pdf
PSCAA. (2020). *Air Quality Data. Puget Sound Clean Air Agency (PSCAA).* https://pscleanair.gov/154/Air-Quality-Data
Ragettli, M. S., Ducret-Stich, R. E., Foraster, M., Morelli, X., Aguilera, I., Basagaña, X., Corradi, E., Ineichen, A., Tsai, M. Y., Probst-Hensch, N., Rivera, M., Slama, R., Künzli, N., & Phuleria, H. C. (2014). Spatio-temporal variation of urban ultrafine particle number concentrations. *Atmospheric Environment*, *96*, 275–283. https://doi.org/10.1016/j.atmosenv.2014.07.049
Ramachandran, G. (2005). Occupational Exposure Assessment for Air Contaminants. In *Occupational Exposure Assessment for Air Contaminants*. https://doi.org/10.1201/9781420032154
Ranft, U., Schikowski, T., Sugiri, D., Krutmann, J., & Krämer, U. (2009). Long-term exposure to traffic-related particulate matter impairs cognitive function in the elderly. *Environmental Research*, *109*(8), 1004–1011. https://doi.org/10.1016/j.envres.2009.08.003
Riffault, V., Arndt, J., Marris, H., Mbengue, S., Setyan, A., Alleman, L. Y., Deboudt, K., Flament, P., Augustin, P., Delbarre, H., & Wenger, J. (2015). Fine and Ultrafine Particles in the Vicinity of Industrial Activities: A Review. *Critical Reviews in Environmental Science and Technology*, *45*(21), 2305–2356. https://doi.org/10.1080/10643389.2015.1025636
Riley, E. A., Banks, L., Fintzi, J., Gould, T. R., Hartin, K., Schaal, L. N., Davey, M., Sheppard, L., Larson, T., Yost, M. G., & Simpson, C. D. (2014). Multi-pollutant mobile platform measurements of air pollutants adjacent to a major roadway. *Atmospheric Environment*, *98*, 492–499. https://doi.org/10.1016/j.atmosenv.2014.09.018
Riley, E. A., Gould, T., Hartin, K., Fruin, S. A., Simpson, C. D., Yost, M. G., & Larson, T. (2016). Ultrafine particle size as a tracer for aircraft turbine emissions. *Atmospheric Environment*, *139*, 20–29.
Riley, E. A., Schaal, L. N., Sasakura, M., Crampton, R., Gould, T. R., Hartin, K., Sheppard, L., Larson, T., Simpson, C. D., & Yost, M. G. (2016). Correlations between short-term mobile monitoring and long-term passive sampler measurements of traffic-related air pollution. *Atmospheric Environment*, *132*. https://doi.org/10.1016/j.atmosenv.2016.03.001
Rivera, M., Basagaña, X., Aguilera, I., Agis, D., Bouso, L., Foraster, M., Medina-Ramón, M., Pey, J., Künzli, N., & Hoek, G. (2012). Spatial distribution of ultrafine particles in urban settings: A land use regression model. *Atmospheric Environment*, *54*, 657–666. https://doi.org/10.1016/j.atmosenv.2012.01.058
RStudio. (2019). *RStudio*. https://rstudio.com/
Ruehl, C., Herner, J. D., Yoon, S., Collins, J. F., Misra, C., Na, K., Robertson, W. H., Biswas, S., Chang, M.-C. O., & Ayala, A. (2015). Similarities and differences between “traditional” and “clean” diesel PM. *Emission Control Science and Technology*, *1*(1), 17–23.
Ryan, P. H., LeMasters, G. K., Biswas, P., Levin, L., Hu, S., Lindsey, M., Bernstein, D. I., Lockey, J., Villareal, M., Hershey, G. K. K., & Grinshpun, S. A. (2007). A Comparison of Proximity and Land Use Regression Traffic Exposure Models and Wheezing in Infants. *Environmental Health Perspectives*, *115*(2), 278–284. https://doi.org/10.1289/ehp.9480
Sabaliauskas, K., Jeong, C. H., Yao, X., Reali, C., Sun, T., & Evans, G. J. (2015). Development of a land-use regression model for ultrafine particles in Toronto, Canada. *Atmospheric Environment*, *110*, 84–92. https://doi.org/10.1016/j.atmosenv.2015.02.018
Saha, P. K., Li, H. Z., Apte, J. S., Robinson, A. L., & Presto, A. A. (2019). Urban Ultrafine Particle

Exposure Assessment with Land-Use Regression: Influence of Sampling Strategy [Research-article]. *Environmental Science and Technology*, *53*, 7326–7336. https://doi.org/10.1021/acs.est.9b02086

Saha, P. K., Zimmerman, N., Malings, C., Hauryliuk, A., Li, Z., Snell, L., Subramanian, R., Lipsky, E., Apte, J. S., Robinson, A. L., & Presto, A. A. (2019). Quantifying high-resolution spatial variations and local source impacts of urban ultrafine particle concentrations. *Science of The Total Environment*, *655*, 473–481. https://doi.org/https://doi.org/10.1016/j.scitotenv.2018.11.197

Salako, G. O., Hopke, P. K., Cohen, D. D., Begum, B. A., Biswas, S. K., Pandit, G. G., Lodoysamba, S., Wimolwattanapun, W., Bunprapob, S., Chung, Y.-S., Rahman, S. A., Hamzah, M. S., Davy, P., Markwitz, A., & Shagjjamba, D. (2012). Exploring the Variation between EC and BC in a Variety of Locations. *Aerosol and Air Quality Research*, *12*(1), 1–7. https://doi.org/10.4209/aaqr.2011.09.0150

Salonen, H., Salthammer, T., & Morawska, L. (2019). Human exposure to NO2 in school and office indoor environments. *Environment International*, *130*, 104887. https://doi.org/https://doi.org/10.1016/j.envint.2019.05.081

Samet, J. M., Dominici, F., Curriero, F. C., Coursac, I., & Zeger, S. L. (2002). Fine Particulate Air Pollution and Mortality in 20 U.S. Cities, 1987–1994. *New England Journal of Medicine*, *343*(24), 1742–1749. https://doi.org/10.1056/nejm200012143432401

Sampson, P. D., Richards, M., Szpiro, A. A., Bergen, S., Sheppard, L., Larson, T. V., & Kaufman, J. D. (2013). A regionalized national universal kriging model using Partial Least Squares regression for estimating annual PM2.5 concentrations in epidemiology. *Atmospheric Environment*. https://doi.org/10.1016/j.atmosenv.2013.04.015

Sampson, P. D., Szpiro, A. A., Sheppard, L., Lindström, J., & Kaufman, J. D. (2011). Pragmatic estimation of a spatio-temporal air quality model with irregular monitoring data. *Atmospheric Environment*, *45*(36), 6593–6606. https://doi.org/10.1016/j.atmosenv.2011.04.073

Saraswat, A., Apte, J. S., Kandlikar, M., Brauer, M., Henderson, S. B., & Marshall, J. D. (2013). Spatiotemporal land use regression models of fine, ultrafine, and black carbon particulate matter in New Delhi, India. *Environmental Science and Technology*, *47*(22), 12903–12911. https://doi.org/10.1021/es401489h

Seaton, A., Godden, D., MacNee, W., & Donaldson, K. (1995). Particulate air pollution and acute health effects. *The Lancet*, *345*(8943), 176–178. https://doi.org/10.1016/S0140-6736(95)90173-6

Seshadri, S., & Wolf, P. A. (2007). Lifetime risk of stroke and dementia: current concepts, and estimates from the Framingham Study. *Lancet Neurology*, *6*(12), 1106–1114. https://doi.org/10.1016/S1474-4422(07)70291-0

Shaffer, R. M., Blanco, M., Li, G., Adar, S. D., Carone, M., Szpiro, A., Kaufman, J. D., Larson, T., Larson, E. B., Crane, P. K., & Sheppard, L. (2021). Fine Particulate Matter and Dementia Incidence in the Adult Changes in Thought Study. *Submitted to Environmental Health Perspectives*.

Sharma, N., Chaudhry, K., & Rao, C. C. (2004). Vehicular pollution prediction modelling: a review of highway dispersion models. *Transport Reviews*, *24*(4), 409–435. https://doi.org/10.1080/0144164042000196071

Shrestha, P. M., Humphrey, J. L., Carlton, E. J., Adgate, J. L., Barton, K. E., Root, E. D., & Miller, S. L. (2019). Impact of Outdoor Air Pollution on Indoor Air Quality in Low-Income Homes during Wildfire Seasons. *International Journal of Environmental Research and Public Health*, *16*(19), 3535. https://doi.org/10.3390/ijerph16193535

Simon, M. C., Hudda, N., Naumova, E. N., Levy, J. I., Brugge, D., & Durant, J. L. (2017). Comparisons of Traffic-Related Ultrafine Particle Number Concentrations Measured in Two Urban Areas by Central, Residential, and Mobile Monitoring. *Atmospheric Environment (Oxford, England : 1994)*, *169*, 113–127. https://doi.org/10.1016/j.atmosenv.2017.09.003

Simon, M. C., Patton, A. P., Naumova, E. N., Levy, J. I., Kumar, P., Brugge, D., & Durant, J. L. (2018). Combining Measurements from Mobile Monitoring and a Reference Site to Develop Models of Ambient Ultrafine Particle Number Concentration at Residences. *Environmental Science and Technology*, *52*(12), 6985–6995. https://doi.org/10.1021/acs.est.8b00292

Snyder, M., Arunachalam, S., Isakov, V., Talgo, K., Naess, B., Valencia, A., Omary, M., Davis, N., Cook, R., & Hanna, A. (2014). Creating locally-resolved mobile-source emissions inputs for air quality modeling in support of an exposure study in Detroit, Michigan, USA. *International Journal of Environmental Research and Public Health*, *11*(12), 12739–12766. https://doi.org/10.3390/ijerph111212739

Song, F., Hooper, L., & Loke, Y. K. (2013). Publication bias: what is it? How do we measure it? How do we avoid it? *Open Access Journal of Clinical Trials*, *5*, 71–81.

Sperling, D. (2018). *Three Revolutions: Steering Automated, Shared, and Electric Vehicles to a Better Future*. Island Press. https://books.google.com/books?id=f0NEDwAAQBAJ

Stanley, S. (2019). *Traffic pollution and dementia. Environmental & Occupational Health Sciences*. University of Washington. https://deohs.washington.edu/hsm-blog/traffic-pollution-and-dementia

Sternberg, S. P. K. (2015). *Air Pollution: Engineering, Science, and Policy*. College Publishing. https://books.google.com/books?id=9Lj5vQAACAAJ

Stone, V., Tuinman, M., Vamvakopoulos, J. E., Shaw, J., Brown, D., Petterson, S., Faux, S. P., Borm, P., MacNee, W., Michaelangeli, F., & Donaldson, K. (2000). Increased calcium influx in a monocytic cell line on exposure to ultrafine carbon black. *European Respiratory Journal*, *15*(2), 297–303. http://erj.ersjournals.com/content/15/2/297

Su, J. G., Allen, G., Miller, P. J., & Brauer, M. (2013). Spatial modeling of residential woodsmoke across a non-urban upstate New York region. *Air Quality, Atmosphere & Health*, *6*(1), 85–94. https://doi.org/10.1007/s11869-011-0148-1

Szpiro, A. A., & Paciorek, C. J. (2013). Measurement error in two-stage analyses, with application to air pollution epidemiology. *Environmetrics*. https://doi.org/10.1002/env.2233

Szpiro, A. A., Paciorek, C. J., & Sheppard, L. (2011). Does more accurate exposure prediction necessarily improve health effect estimates? *Epidemiology*, *22*(5), 680–685. https://doi.org/10.1097/EDE.0b013e3182254cc6

Szpiro, A. A., Sampson, P. D., Sheppard, L., Lumley, T., Adar, S. D., & Kaufman, J. D. (2010). Predicting intra-urban variation in air pollution concentrations with complex spatio-temporal dependencies. *Environmetrics*, *21*(6), 606–631. https://doi.org/10.1002/env.1014

Teng, E. L., Hasegawa, K., Homma, A., Imai, Y., Larson, E., Graves, A., Sugimoto, K., Yamaguchi, T., Sasaki, H., Chiu, D., & White, L. R. (2004). The Cognitive Abilities Screening Instrument (CASI): A Practical Test for Cross-Cultural Epidemiological Studies of Dementia. *International Psychogeriatrics*, *6*(1), 45–58. https://doi.org/10.1017/s1041610294001602

Transport Policy. (2018). *US: Heavy-Duty: Emissions*.

Tzivian, L., Winkler, A., Dlugaj, M., Schikowski, T., Vossoughi, M., Fuks, K., Weinmayr, G., & Hoffmann, B. (2014). Effect of long-term outdoor air pollution and noise on cognitive and psychological functions in adults. *International Journal of Hygiene and Environmental Health*. https://doi.org/10.1016/j.ijheh.2014.08.002

U.S. Geology Survey. (2021). *Landsat Normalized Difference Vegetation Index*. Landsat Surface Reflectance-Derived Spectral Indices. https://www.usgs.gov/core-science-systems/nli/landsat/landsat-normalized-difference-vegetation-index?qt-science_support_page_related_con=0#qt-science_support_page_related_con

US Census Bureau. (2018). *Census Feature Class Codes (CFCC)*. https://www.maris.state.ms.us/pdf/cfcccodes.pdf

US Census Bureau. (2019a). *Population*. https://www.census.gov/topics/population.html

US Census Bureau. (2019b). *Quick Facts: King County, Washington; US*. https://www.census.gov/quickfacts/fact/table/kingcountywashington,US/PST045218

US Census Bureau. (2020). *American Community Survey (ACS)*. https://www.census.gov/programs-surveys/acs

US EPA. (2001). *Control of Air Pollution From New Motor Vehicles: Heavy-Duty Engine and Vehicle Standards and Highway Diesel Fuel Sulfur Control Requirements*. US Environmental Protection Agency; United States Environmental Protection Agency (US EPA). https://www.epa.gov/regulations-emissions-vehicles-and-engines/final-rule-control-air-pollution-new-motor-vehicles-heavy

US EPA. (2012). *MOVES Validation Efforts to Date*. US Environmental Protection Agency. https://www.epa.gov/sites/production/files/2016-06/documents/03-moves-validation-faca.pdf

US EPA. (2016). Integrated Science Assessment for Oxides of Nitrogen – Health Criteria. *United States Environmental Protection Agency*, *January*. https://www.epa.gov/isa/integrated-science-assessment-isa-nitrogen-dioxide-health-criteria

US EPA. (2017a). *National Emissions inventory*. United States Environmental Protection Agency.

US EPA. (2017b). *The Clean Air Act - Highlights of the 1990 Amendments*. US Environmental Protection Agency. https://www.epa.gov/clean-air-act-overview/clean-air-act-highlights-1990-amendments

US EPA. (2018a). *History of Reducing Air Pollution from Transportation in the United States*. US Environmental Protection Agency. https://www.epa.gov/transportation-air-pollution-and-climate-change/accomplishments-and-success-air-pollution-transportation

US EPA. (2018b). *Integrated Science Assessment for Particulate Matter*. US Environmental Protection Agency. https://www.epa.gov/isa/integrated-science-assessment-isa-particulate-matter

US EPA. (2018c). *Latest Version of MOtor Vehicle Emission Simulator (MOVES)*. US Environmental Protection Agency. https://www.epa.gov/moves/latest-version-motor-vehicle-emission-simulator-moves

US EPA. (2018d). *Our Nation's Air 2018. US Environmental Protection Agency*. US Environmental Protection Agency. https://gispub.epa.gov/air/trendsreport/2018/#welcome
US EPA. (2018e). *Primary National Ambient Air Quality Standards (NAAQS) for Nitrogen Dioxide*. US Environmental Protection Agency. https://www.epa.gov/no2-pollution/primary-national-ambient-air-quality-standards-naaqs-nitrogen-dioxide
US EPA. (2019a). Air Quality - National Summary. *US Environmental Protection Agency*. https://www.epa.gov/air-trends/air-quality-national-summary
US EPA. (2019b). *AirData Pre-Generated Data Files*. US Environmental Protection Agency. https://aqs.epa.gov/aqsweb/airdata/download_files.html
US EPA. (2019c). *Black Carbon Research*. US Environmental Protection Agency. https://www.epa.gov/air-research/black-carbon-research
US EPA. (2019d). *Diesel Fuel Standards and Rulemakings*. US Environmental Protection Agency. https://www.epa.gov/diesel-fuel-standards/diesel-fuel-standards-and-rulemakings
US EPA. (2019e). Integrated science assessment (ISA) for particulate matter (final report, Dec 2019). *US Environmental Protection Agency*. https://cfpub.epa.gov/ncea/isa/recordisplay.cfm?deid=347534
US EPA. (2019f). Regulatory Information by Topic: Air. *US Environmental Protection Agency*. https://www.epa.gov/regulatory-information-topic/regulatory-information-topic-air
USGS. (2021). *Landsat Enhanced Vegetation Index*. U.S. Geology Survey. https://www.usgs.gov/core-science-systems/nli/landsat/landsat-enhanced-vegetation-index?qt-science_support_page_related_con=0#qt-science_support_page_related_con
Van den Bossche, J., Peters, J., Verwaeren, J., Botteldooren, D., Theunis, J., & De Baets, B. (2015). Mobile monitoring for mapping spatial variation in urban air quality: Development and validation of a methodology based on an extensive dataset. *Atmospheric Environment*. https://doi.org/10.1016/j.atmosenv.2015.01.017
van Nunen, E., Vermeulen, R., Tsai, M.-Y., Probst-Hensch, N., Ineichen, A., Davey, M., Imboden, M., Ducret-Stich, R., Naccarati, A., & Raffaele, D. (2017). Land use regression models for ultrafine particles in six European areas. *Environmental Science & Technology*, *51*(6), 3336–3345.
Vardoulakis, S., Giagloglou, E., Steinle, S., Davis, A., Sleeuwenhoek, A., Galea, K. S., Dixon, K., & Crawford, J. O. (2020). Indoor Exposure to Selected Air Pollutants in the Home Environment: A Systematic Review. *International Journal of Environmental Research and Public Health*, *17*(23), 8972.
Waller, M., Mishra, G. D., & Dobson, A. J. (2020). A comparison of cause-specific and competing risk models to assess risk factors for dementia. *Epidemiologic Methods*, *1*.
Wang, L., Larson, E. B., Bowen, J. D., & Van Belle, G. (2006). Performance-based physical function and future dementia in older people. *Archives of Internal Medicine*, *166*(10), 1115–1120. https://doi.org/10.1001/archinte.166.10.1115
Wang, R., Henderson, S. B., Sbihi, H., Allen, R. W., & Brauer, M. (2013). Temporal stability of land use regression models for traffic-related air pollution. *Atmospheric Environment*, *64*, 312–319. https://doi.org/https://doi.org/10.1016/j.atmosenv.2012.09.056
Wang, X., Heald, C. L., Sedlacek, A. J., de Sá, S. S., Martin, S. T., Alexander, M. L., Watson, T. B., Aiken, A. C., Springston, S. R., & Artaxo, P. (2016). *Deriving brown carbon from multiwavelength absorption measurements: method and application to AERONET and*

Aethalometer observations.
WebMD. (2018). *Types of Alzheimer's: Early-Onset, Late-Onset, and Familial*. https://www.webmd.com/alzheimers/guide/alzheimers-types
Weichenthal, S., Olaniyan, T., Christidis, T., Lavigne, E., Hatzopoulou, M., Van Ryswyk, K., Tjepkema, M., & Burnett, R. (2020). Within-city Spatial Variations in Ambient Ultrafine Particle Concentrations and Incident Brain Tumors in Adults. *Epidemiology*, *31*(2). https://journals.lww.com/epidem/Fulltext/2020/03000/Within_city_Spatial_Variations_in_Ambient.4.aspx
Weichenthal, S., Ryswyk, K. Van, Goldstein, A., Bagg, S., Shekkarizfard, M., & Hatzopoulou, M. (2016). A land use regression model for ambient ultrafine particles in Montreal, Canada: A comparison of linear regression and a machine learning approach. *Environmental Research*, *146*, 65–72. https://doi.org/10.1016/j.envres.2015.12.016
Weichenthal, S., Van Ryswyk, K., Goldstein, A., Shekarrizfard, M., & Hatzopoulou, M. (2016). Characterizing the spatial distribution of ambient ultrafine particles in Toronto, Canada: A land use regression model. *Environmental Pollution*. https://doi.org/10.1016/j.envpol.2015.04.011
Weichenthal, S., Van Ryswyk, K., Kulka, R., Sun, L., Wallace, L., & Joseph, L. (2015). In-vehicle exposures to particulate air pollution in Canadian metropolitan areas: the urban transportation exposure study. *Environmental Science & Technology*, *49*(1), 597–605.
Weissert, L. F., Salmond, J. A., Miskell, G., Alavi-Shoshtari, M., & Williams, D. E. (2018). Development of a microscale land use regression model for predicting NO 2 concentrations at a heavy trafficked suburban area in Auckland, NZ. *Science of the Total Environment*, *619–620*, 112–119. https://doi.org/10.1016/j.scitotenv.2017.11.028
Westerdahl, D., Fruin, S., Sax, T., Fine, P. M., & Sioutas, C. (2005). Mobile platform measurements of ultrafine particles and associated pollutant concentrations on freeways and residential streets in Los Angeles. *Atmospheric Environment*, *39*, 3597–3610. https://doi.org/10.1016/j.atmosenv.2005.02.034
Weuve, J., Barnes, L. L., Mendes de Leon, C. F., Rajan, K. B., Beck, T., Aggarwal, N. T., Hebert, L. E., Bennett, D. A., Wilson, R. S., & Evans, D. A. (2018). Cognitive Aging in Black and White Americans. *Epidemiology*, *29*(1), 151–159. https://doi.org/10.1097/ede.0000000000000747
Weuve, J., Hebert, L. E., Scherr, P. A., & Evans, D. A. (2014). Deaths in the United States among persons with Alzheimer's disease (2010-2050). *Alzheimer's and Dementia*, *10*(2), e40-46. https://doi.org/10.1016/j.jalz.2014.01.004
WHO. (2012). *Dementia: A Public Health Priority*. World Health Organization. https://www.who.int/mental_health/publications/dementia_report_2012/en/
Wilker, E. H., Preis, S. R., Beiser, A. S., Wolf, P. A., Au, R., Kloog, I., Li, W., Schwartz, J., Koutrakis, P., DeCarli, C., Seshadri, S., & Mittleman, M. A. (2015). Long-Term Exposure to Fine Particulate Matter, Residential Proximity to Major Roads and Measures of Brain Structure. *Stroke*, *46*(5), 1161–1166. https://doi.org/10.1161/STROKEAHA.114.008348
Wilker, Elissa H, Ljungman, P. L., Rice, M. B., Kloog, I., Schwartz, J., Gold, D. R., Koutrakis, P., Vita, J. A., Mitchell, G. F., Vasan, R. S., Benjamin, E. J., Hamburg, N. M., & Mittleman, M. A. (2014). Relation of Long-Term Exposure to Air Pollution to Brachial Artery Flow-Mediated Dilation and Reactive Hyperemia. *The American Journal of Cardiology*, *113*(12), 2057–

2063. https://doi.org/10.1016/j.amjcard.2014.03.048
Wolf, K., Cyrys, J., Harciníková, T., Gu, J., Kusch, T., Hampel, R., Schneider, A., & Peters, A. (2017). Land use regression modeling of ultrafine particles, ozone, nitrogen oxides and markers of particulate matter pollution in Augsburg, Germany. *Science of the Total Environment*, *579*, 1531–1540. https://doi.org/10.1016/j.scitotenv.2016.11.160
Wortmann, M. (2013). Importance of national plans for Alzheimer's disease and dementia. *Alzheimer's Research and Therapy*, *5*(5), 40. https://doi.org/10.1186/alzrt205
WSDOE. (2018). *Washington State 2014 Comprehensive Emissions Inventory*. Washington State Department of Ecology. https://ecology.wa.gov/DOE/files/0d/0dfbc0d0-8485-4620-981b-d6636e1157ee.pdf
WSDOT. (2019). *Historic Traffic Volumes. Traffic Data GeoPortal. Washington State Department of Transportation*. Washington State Department of Transportation. https://www.wsdot.wa.gov/mapsdata/tools/trafficplanningtrends.htm
Wu, Y. T., Fratiglioni, L., Matthews, F. E., Lobo, A., Breteler, M. M. B., Skoog, I., & Brayne, C. (2016). Dementia in western Europe: Epidemiological evidence and implications for policy making. *The Lancet Neurology*, *15*(1), 116–124. https://doi.org/10.1016/S1474-4422(15)00092-7
Xie, X., Semanjski, I., Gautama, S., Tsiligianni, E., Deligiannis, N., Rajan, T. R., Pasveer, F., & Philips, W. (2017). A Review of Urban Air Pollution Monitoring and Exposure Assessment Methods. In *ISPRS International Journal of Geo-Information* (Vol. 6, Issue 12). https://doi.org/10.3390/ijgi6120389
Xu, D., & Guo, X. (2014). Compare NDVI Extracted from Landsat 8 Imagery with that from Landsat 7 Imagery. *American Journal of Remote Sensing*, *2*(2). https://doi.org/10.11648/j.ajrs.20140202.11
Yan, W., Ji, X., Shi, J., Li, G., & Sang, N. (2015). Acute nitrogen dioxide inhalation induces mitochondrial dysfunction in rat brain. *Environmental Research*. https://doi.org/10.1016/j.envres.2015.02.022
Yan, W., Yun, Y., Ku, T., Li, G., & Sang, N. (2016). NO2 inhalation promotes Alzheimer's disease-like progression: Cyclooxygenase-2-derived prostaglandin E2modulation and monoacylglycerol lipase inhibition-targeted medication. *Scientific Reports*, *6*(February), 1–17. https://doi.org/10.1038/srep22429
Ye, Q., Li, H. Z., Gu, P., Robinson, E. S., Apte, J. S., Sullivan, R. C., Robinson, A. L., Donahue, N. M., & Presto, A. A. (2020). Moving beyond Fine Particle Mass: High-Spatial Resolution Exposure to Source-Resolved Atmospheric Particle Number and Chemical Mixing State. *Environmental Health Perspectives*, *128*(1), 17009.
Young, M. T., Bechle, M. J., Sampson, P. D., Szpiro, A. A., Marshall, J. D., Sheppard, L., & Kaufman, J. D. (2016). Satellite-Based NO2 and Model Validation in a National Prediction Model Based on Universal Kriging and Land-Use Regression. *Environmental Science and Technology*. https://doi.org/10.1021/acs.est.5b05099
Yu, C. H., Fan, Z., Lioy, P. J., Baptista, A., Greenberg, M., & Laumbach, R. J. (2016). A novel mobile monitoring approach to characterize spatial and temporal variation in traffic-related air pollutants in an urban community. *Atmospheric Environment*, *141*, 161–173. https://doi.org/10.1016/j.atmosenv.2016.06.044
Yu, X., Zheng, L., Jiang, W., & Zhang, D. (2020). Exposure to air pollution and cognitive

impairment risk: a meta-analysis of longitudinal cohort studies with dose-response analysis. *Journal of Global Health*, *10*(1).

Yuchi, W., Sbihi, H., Davies, H., Tamburic, L., & Brauer, M. (2020). Road proximity, air pollution, noise, green space and neurologic disease incidence: A population-based cohort study. *Environmental Health: A Global Access Science Source*, *19*(1), 1–15. https://doi.org/10.1186/s12940-020-0565-4

Zhu, Y., Hinds, W. C., Kim, S., & Sioutas, C. (2002). Concentration and size distribution of ultrafine particles near a major highway. *Journal of the Air and Waste Management Association*. https://doi.org/10.1080/10473289.2002.10470842

Zipprich, J. L., Harris, S. A., Fox, J. C., & Borzelleca, J. F. (2002). An analysis of factors that influence personal exposure to nitrogen oxides in residents of Richmond, Virginia. *Journal of Exposure Analysis and Environmental Epidemiology*, *12*(4), 273–285. https://doi.org/10.1038/sj.jea.7500226

Zwack, L. M., Paciorek, C. J., Spengler, J. D., & Levy, J. I. (2011). Characterizing local traffic contributions to particulate air pollution in street canyons using mobile monitoring techniques. *Atmospheric Environment*, *45*(15), 2507–2514. https://doi.org/https://doi.org/10.1016/j.atmosenv.2011.02.035