Title
A breath of fresher air: improving methods for PM2.5 exposure assessment from Mongolia to California

Author
Hill, Lawson Andrew

Publication Date
2017

Peer reviewed|Thesis/dissertation

A breath of fresher air: improving methods for $PM_{2.5}$ exposure assessment from Mongolia to California

By

Lawson Andrew Hill

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Environmental Health Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in Charge:

Professor Kirk R Smith, Chair
Professor John Balmes
Professor Alan E Hubbard

Summer 2017

Abstract

A breath of fresh*er* air: Improving methods for $PM_{2.5}$ exposure assessment from California to Mongolia

by

Lawson Andrew Hill

Doctor of Philosophy in Environmental Health Sciences

University of California, Berkeley

Professor Kirk R Smith, Chair

Airborne particulate matter smaller than 2.5 μm in diameter ($PM_{2.5}$) is among the biggest determinants of disease world-wide. In 2013, exposure to $PM_{2.5}$ caused an estimated 7 million deaths and 189 million Disability-Adjusted Life Years (DALYs). About half of this calculated burden arises from the indoor residential use of solid fuels – like wood, coal, dung, and crop waste – for cooking, a practice common to about 2.8 billion people. The rest is attributable to ambient concentrations produced by combustion sources like power plants, heating utilities and appliances, and motor vehicles. Despite the ubiquity of the problem, many monitoring, research, and policy efforts employ proxies of total exposure like outdoor ambient concentrations. These proxies are inadequate for the quantification of actual exposures and can prove misleading when used to estimate health effects. More work is needed to progress the use of total exposure – which is often a complex function of a person's interaction with numerous environments with varying $PM_{2.5}$ concentrations – as a proper $PM_{2.5}$ risk metric.

The first research chapter of this dissertation, Chapter 2, estimates changes in total $PM_{2.5}$ exposure from indoor concentrations (including contributions from second hand smoke), outdoor concentrations, and time-activity in 2014 and in 2024 under alternative emissions policy pathways in one of the most polluted capital cities on Earth: Ulaanbaatar, Mongolia. Ulaanbaatar's air pollution crisis is seasonal; heavy use of residential coal heating during

its harsh and lengthy winters produces some of the worst air pollution in the world, but many cities are more polluted on an annual basis. With this in mind, seasonal exposure patterns are considered separately to produce estimates of annual exposures. These values are combined with projected background disease and population levels and some of the latest available exposure-response functions to project $PM_{2.5}$-related health impacts. Policy pathways are estimated for business as usual; moderate reductions in heating, power plant, and motor vehicle emissions; and major reductions in the same sectors. The analysis estimates a 2014 population-wide annual average exposure of 59 μg/m3, which increases to 60 $\mu g/m^3$ in 2024 under business as usual but falls to 32 $\mu g/m^3$ and 12 $\mu g/m^3$ under moderate and major emissions control policies, respectively. Annual $PM_{2.5}$-related deaths and DALYs are estimated at about 1,400 and 40,000, respectively, in 2014. Under business as usual, about 18,000 deaths and 530,000 DALYs are accrued through 2024. Exposure reductions resulting from the moderate control policy pathway avert an estimated 110,000 DALYs and 4,000 deaths from the business as usual pathway between 2014-2024. An estimated 240,000 DALYs and 8,000 deaths are averted under major reduction policies. In all, Chapter 2 highlights the need for aggressive action, especially related to residential heating and tobacco smoking, to avert a growing pollution crisis in Ulaanbaatar.

Chapter 3 presents some of the first personal $PM_{2.5}$ exposure measurements conducted in rural Lao women cooking primarily with wood. Measurements were taken during a stove intervention program in which traditional open fire and bucket stoves were ostensibly replaced with an ACE-1 fan stove. Average 48-hour concentrations before and after the intervention are reported at 123 $\mu g/m^3$ and 81 $\mu g/m^3$, respectively. Measurements of kitchen concentrations, ambient concentrations, and other environmental data are combined with an extensive set of survey responses to reliably model mean 48-hour average $PM_{2.5}$ exposures before and after the intervention using machine learning, ensemble, and cross-validation techniques (for the full model: $r^2 = 0.26$, predicted mean before intervention = 120 $\mu mg/m^3$, predicted mean after intervention = 88 $\mu g/m^3$).

Chapter 4 proposes the use of a household appliance, the smart smoke detector, as a tool for cost-effectively monitoring indoor $PM_{2.5}$ concentrations, which are often overlooked by regulatory monitoring networks and health effects research. A particularly popular smart smoke detector, the Nest Protect, is reverse engineered. Its onboard optical sensor is co-opted and characterized for the real-time measurement of $PM_{2.5}$ mass concentrations. Very good agreement is observed between processed Nest Protect signal and output from a co-located research grade monitor, the DustTrak II ($r^2 > 0.99$).

The final chapter, Chapter 5, reiterates the thread common among Chapters 3-5 – advancing $PM_{2.5}$ risk science through better estimation of total exposures – and discusses key areas for future research.

Für Elise.

Table of Contents

Preface

This doctoral dissertation was completed under the primary advising of Professor Kirk R. Smith, and with a great deal of support from friends, family, and colleagues.

Chapter 2 is a slightly modified version of a manuscript [1] which, at the time of this dissertation filing, had been submitted and was under consideration for peer-reviewed publication. While a truly collaborative project, I was the lead author responsible for primary composition of the manuscript, projecting demographics, applying the exposure model approach (which was developed collaboratively), modeling background disease rates and indoor pollution concentrations with input from co-authors and from data collected by co-authors and other colleagues, and calculating attributable disease burdens, among other contributions. The manuscript benefitted from the hard work of all co-authors and adapted text from a precursor report [2] to the Ministry of Environment and Green Development – which funded the study through the Clean Air Foundation of Ulaanbaatar. Portions of the precursor report were drafted by co-authors. For a period of time overlapping the beginning of this study, several co-authors were consultants to Social Impact, Inc. (SI) in conducting a *separate* study to evaluate the impacts of a stove subsidy program in Ulaanbaatar. Some of the openly accessible data collected as part of that study were used as inputs to the current study, but SI did not pay salaries as part of this study, have any role in study design, data collection, analysis, decision to publish, or preparation of the manuscript. Editing and advice from anonymous peer reviewers was received during the manuscript submission process. Permission to include this work in my dissertation was received from all co-authors and a representative of the Ministry of Environment and Green Development of Mongolia.

Chapter 3 evolved from a study funded in part by the World Bank, during which the data analyzed in the chapter were collected, and a subsequent collaborative report to the Ministry of Health and Inter-Ministerial Clean Stove Initiative of the Lao People's Democratic Republic.[3] I was responsible for drafting several sections of that report and led the analysis of gravimetric exposure and gravimetric kitchen air pollution concentration

---

[1] Submitted as: "Hill, LD, Edwards, R, Turner, JR, Damdinsuren, Y, Olkhanud, P, Odsuren, M, Guttikunda, S, Ochir, C, Smith, K. Health assessment of future $PM_{2.5}$ exposures from indoor, outdoor, and environmental tobacco smoke concentrations under alternative policy pathways in Ulaanbaatar, Mongolia. *In review. 2017.*" The title was revised during review to (at the time of dissertation filing): "Health assessment of future $PM_{2.5}$ exposures from indoor, outdoor, and second hand tobacco smoke concentrations under alternative policy pathways in Ulaanbaatar, Mongolia."

[2] Ochir, C, Smith, KR, Hill, LD, Olkhanud, P, Damdinsuren, Y, Odsuren, M, Edwards, R, Turner, JR. Air pollution and health in Ulaanbaatar. Final Project Report. Prepared for the Ministry of the Environment and Green Development, Ulaanbaatar, Mongolia. 2014.

[3] Hill LD, Pillarisetti A, Delapena S, Garland C, Jagoe K, Koetting P, Pelletreau A, Boatman MR, Pennise D, Smith KR. Air pollution and impact analysis of a pilot stove intervention: report to the Ministry of Health and Inter-Ministerial Clean Stove Initiative of the Lao People's Democratic Republic. 2015.

data. I was also involved in the collaborative processes of developing data collection tools, training the field teams, and managing data collection, among other aspects of the project. The work presented in Chapter 3 adds a great deal of analysis to the original report, including the implementation of super learning and machine learning to model personal exposures. Permission to use elements of that report in my thesis was received from relevant parties, including all co-authors. None of the individuals or groups who submitted the ministry report were responsible for choosing or disseminating the particular stove evaluated therein. None had any financial or other interest in the stove or any other competing stove at the time of the report, nor do I at the time of filing this dissertation.

Chapter 4 is not adapted from any previously "published" manuscripts, but did benefit from the assistance of colleagues as discussed in the Acknowledgments. Most notable are the contributions made by sensor expert Tracy Alan of EME Systems (Berkeley, CA) who guided all aspects of the reverse engineering work and mentored me through the signal co-opting process and Ajay Pillarisetti of UC Berkeley who provided invaluable input and feedback throughout the project. Chapters 1 and 5 are introduction and conclusion texts intended to tie together chapters 2-4.

Acknowledgements

I may not have followed my heart to UC Berkeley were it not for the funding that Dr. Kirk Smith, Dr. Michael Bates, Maria Hernandez, Norma Firestone, and others at the School of Public Health helped me procure: the James A. Buchanan Scholarship, Reshetko Fellowship, Preston Scholarship, and School of Public Health Power Top-Off Award. The School of Public Health Block Grant helped me to focus more on my studies, and less on financial worries. During my final two years at UC Berkeley, I was funded by the National Science Foundation Systems Approach to Green Energy Integrative Graduate Education and Research Traineeship Program (NSF SAGE IGERT) through the UC Berkeley Center for Green Chemistry. Under this program, Dr. Daniel Kammen, Dr. Marty Mulvihill, Dr. Tom McKone, and Tom McKeag provided invaluable mentorship.

There are too many friends and colleagues to mention here whose kindness, support, and gentle direction have carried me through my time at UC Berkeley. Most of all, my true love and my rock, Elise and our dog-child Jackson. Cynthia, Bruce, and Vicki– my loving and supportive parents – nurtured the curiosity and drive to achieve that led me to pursue higher education. I would be nothing without them. I am proud to call Alie and Jenna – my unending sources of cheer and support – my sisters and Ian, Anna, Grant, and Alissa my siblings-in-law (and adventure!). My parents-in-law, Tanya and Don, have been a most unexpected and delightful addition to my support network; the last stretches of writing this dissertation would have been immensely less satisfying without the den provided in their quiet, picturesque perch in the hills of Edmonds, WA.

In my five years at UC Berkeley, I have shared great comradery with and received sage advice from my office mates in University Hall: Nick Lam, Zoë Chafe, Manish Desai, Ajay Pillarisetti, and Amanda Northcross. The UC Berkeley Environmental Health Science Division's Norma Firestone, Terry Jackson, and the inimitable Maria Hernandez quelled or prevented innumerable administrative crises. Dr. Michael Bates was instrumental in fulfilling my goal to research air quality in the Himalayas, offering sage tutelage along the way. Ajay Pillarisetti provided feedback and assistance during the analysis phases of Chapters 2-4, aided with field efforts in Chapter 3, occasionally joined in on Chapter 4 device teardown sessions, and provided brainstorming support and literary advice throughout the smart smoke detector research endeavor. Ajay has never failed to lend insightful advice, editing help, or project input. I am also grateful to the large (and growing!) open source software community, which fostered my education in programming and hardware-monkeying, and facilitated aspects of my research and analysis.

Chapter 2 was a truly collaborative project. I thank the co-authors whose expertise, connections, and, well, gumption were instrumental in navigating the complexities of such a large scale and cooperative project: Dr. Rufus Edwards of UC Irvine, Dr. Jay Turner at Washington University in Saint Louis, Yuma Damdinsuren who was an independent consultant on the project, Dr. Purevdorj Olkhanud of the Mongolian National University

of Medical Sciences (MNUMS) during the project period, Munkhtuul Odsuren of MNUMS during the project period, Dr. Sarath Guttikunda of the Desert Research Institute, Dr. Chimedsuren Ochir of MNUMS during the project period, and Dr. Kirk Smith of UC Berkeley. I am grateful to the National Statistics Office of Mongolia, the Statistics Department of Ulaanbaatar, and the Health Development Center of the Ministry of Health and Sports which provided access to various demographics and health databases during Chapter 2 work. I am grateful to SI and the Millennium Challenge Corporation for their open access databases of household measurements conducted as part of the impact evaluation of the Energy and Environment Projects. I thank Paul Chung and Alan Hubbard of UC Berkeley and Nick Lam from the University of Illinois for their advice during the project. I acknowledge Boldkhuu Nanzad of the Ministry of Energy of Mongolia for advocating for the research involved in Chapter 2, and Dana Charron of Berkeley Air Monitoring Group and Maria Hernandez for their assistance with contractual matters in both Chapters 2 and 3. The final Chapter 2 analysis benefits from comments made by a number of participants at a workshop presenting preliminary results conducted as part of the impact evaluation of the Energy and Environment Projects.

The success of Chapter 3 was fostered by the generosity and backing of local leaders, Lao Government officials, and, of course, the kind study participants who allowed us into their homes, kitchens, and lives. I am immensely grateful for Samantha Delapena, Charity Garland, and David Pennise at Berkeley Air Monitoring Group, who handled gravimetric filter preparations, weighing, and related tasks (and hassles!) for the data collected in Chapter 3. The Berkeley Air Monitoring Group crew, including Kirstie Jagoe, also put a great deal of effort into the field preparation, team training, design, initial field work, and data quality assurance involved in the study from which Chapter 3 evolved. I owe a debt of gratitude to Philipp Koetting and Aurelie Pelletreau at LIRE along with Mette Rohr Boatman at Geo-Sys who assembled, managed, and led the field teams who collected the data used in Chapter 3: Keolamphan Bouthasing, Chanthee Khammavong, Laura Magni, Thomas Motmans (who served as interim field manager for a short period), Phetdavone Narmmaxay, Souksengdao Phetsaphangthong, Soraia Sadid, Lamduan Thammavong, and Khamlah Vongnakhone. I would like to thank the folks at Geo-Sys who worked with me to revisit Lao questionnaire audio recordings to address odd and missing values. In addition, I am thankful for the guidance and patience of Natsuko Toba and Rutu Dave at the World Bank, Ken Newcombe of C-Quest Capital, and Maria Hernandez, and the support of SNV Vientiane, the stove disseminator.

Chapter 4 would have come and gone as a pipe dream were it not for the assistance of the brilliant Tracy Allen of EME Systems, who advised the Nest Protect teardown and sensor access processes at his workshop in Berkeley, CA. A portion of Tracy's efforts, time using EME Systems equipment and workshop space, and the electronic components required to develop this project were funded through the NSF SAGE IGERT Innovation Fund. Ajay Pillarisetti edited this Innovation Fund proposal. Both Tracy and Ajay also generously provided feedback on a draft of this chapter. I would also like to thank Berkeley Air

Monitoring Group, which kindly lent some of the equipment used in or with the gravimetric sampling train and particle-removal train (during temperature response testing) during Chapter 4 laboratory validation work.

I owe great thanks to my graduate examination and dissertation committees, which provided invaluable feedback during all stages of this work (Drs. John Balmes, Ellen Eisen, Alan Hubbard, Mark Nicas, and Kirk Smith).

Finally, I would like to acknowledge in no uncertain terms my primary advisor and dissertation chair, Dr. Kirk Smith, whose wise advice and shared experiences have vastly improved the efficiency and impact of my work. Kirk has always found time to advise – and with great perspicacity – amidst his grueling schedule. I aspire to put into my life's work even a fraction of the verve and vigor that have gone into his.

Chapter 1

Introduction

1.1 Background

Air pollution is a leading cause of disease around the world (Forouzanfar et al. 2015). Of special concern is particulate matter smaller than 2.5 μm in diameter ($PM_{2.5}$), a common byproduct of combustion. Exposure to $PM_{2.5}$ has been linked to the development of lung cancer (LC), ischemic heart disease (IHD), stroke, chronic obstructive pulmonary disorder (COPD), and acute lower respiratory infection (ALRI) in children (Smith et al. 2014). From these diseases alone, researchers estimate that exposure to $PM_{2.5}$ is responsible for about 7 million deaths and nearly 189 million disability-adjusted life years (DALYs [4]) every year (Institute for Health Metrics & Evaluation 2017). The true impact may be higher. A growing body of evidence suggests that $PM_{2.5}$ exposures may increase the risk of dementia, Alzheimer's, and general cognitive impairment (Clifford et al. 2016; Oudin et al. 2015; Power et al. 2016); adverse birth outcomes (Brauer et al. 2008; Dadvand et al. 2013); and tuberculosis (Cohen and Mehta 2007; Pokhrel et al. 2010; Sumpter and Chandramohan 2013), among other health effects.

The health impacts of $PM_{2.5}$ are of special concern in regions where household cooking and heating are commonly performed with solid fuels – like wood, dung, and coal – and inefficient appliances or open fires. Use of solid fuels for cooking results in daily average household $PM_{2.5}$ concentrations (also called household air pollution, or HAP) that have been measured with consistency in kitchens at around 1,000 $\mu g/m^3$ (Balakrishnan et al. 2014), or 40 times the World Health Organization 24-hour health-based guideline (World Health Organization 2014), and personal $PM_{2.5}$ exposures (in women) that have been measured in the realm of 250 $\mu g/m^3$ (Balakrishnan et al. 2014). An estimated 2.8 billion people are thusly affected (Bonjour et al. 2013). Solid fuel cooking also impacts outdoor concentrations; cooking emissions produced an estimated 12.5% of ambient $PM_{2.5}$ globally

[4] DALYs are a time-based measure of disease burden summing years lost to premature death and lost years of healthy life.

in 2010 (Chafe et al. 2014). In all, about 40% of the estimated 7 million premature deaths and about 45% of the 189 million DALYs from exposure to $PM_{2.5}$ are attributed to cooking-related HAP exposures.

Solid fuel heating is less well studied, but is considered important to global $PM_{2.5}$ exposures (Chafe et al. 2015; Smith and Pillarisetti 2012). Residential solid fuel heating has been attributed to 13 – 21% of ambient $PM_{2.5}$ in Europe and about 10% in Canada and the US (Chafe et al. 2015), with higher values in winter and in certain regions (Balakrishnan et al. 2014; Davy et al. 2011; Naeher et al. 2007; Ward and Lange 2010). Outdoor and direct indoor (fugitive) emissions can also lead to increased HAP concentrations. In some regions solid fuel heating-related HAP concentrations may be on par with those of households cooking with solid fuels. This applies especially in low-income areas with harsh winters and a high prevalence of inefficient stoves, like China and Mongolia. While $PM_{2.5}$ exposures from heating-related HAP are not explicitly included in the previously mentioned $PM_{2.5}$ disease burden values – due in part to the difficulty of distinguishing cooking-related emissions from heating-related emissions – the effects of solid fuel heating on morbidity and mortality are likely substantial.

Health impacts from $PM_{2.5}$ exposures are not limited to hyper-polluted areas or high-HAP environments. High risks have been observed at concentrations common in parts of the world where the air is considered quite clean. In High Income countries, for example, ambient air pollution was indicated as the eighth leading cause of death in 2015 (Institute for Health Metrics & Evaluation 2017). In the United States, approximately 100,000 people every year die prematurely due to $PM_{2.5}$ exposures.

In spite of the scale of the problem, exposure measurement methods – or methods that allow for a comprehensive assessment of the "contact between receptors (such as people or ecosystems) and physical, chemical, or biological stressors" (National Research Council 2012) – are relatively underutilized or underdeveloped in the evaluation of $PM_{2.5}$ health risks. In the context of HAP, the second most important environmental risk factor for mortality (Institute for Health Metrics & Evaluation 2017), $PM_{2.5}$ exposures are commonly estimated from kitchen air pollution (KAP) concentrations or fuel-type indicators (e.g. wood, coal, natural gas). When assessing exposures from kitchen concentration measurements, an aggregate ratio of exposure concentration to KAP concentration is measured in a number of homes and then applied by simple multiplication to KAP concentrations either measured or estimated[5] in the remainder of the population of interest, e.g. (Smith et al. 2014). This ratio, called the kitchen exposure factor or KEF, cannot account for the variations in human-environment interaction that underpin actual personal exposure concentrations. This is supported by the considerable variability observed in measured KEF (Armendáriz-Arnez et al. 2008; Clark et al. 2010; Northcross et al. 2010), and suggests that a universal KEF may be a misleading proxy of exposure.

[5] Often from sophisticated models.

At the population level, $PM_{2.5}$ exposures are typically approximated from outdoor ambient concentrations. However, people spend upwards of 90% of their time indoors (Jenkins et al. 1992), where air pollution concentrations can differ considerably from those experienced outdoors (Brunekreef et al. 2005; Naeher et al. 2000). In areas were activities commonly produce indoor $PM_{2.5}$ emissions – like using cookstoves or smoking tobacco – outdoor measurements will tend to underestimate indoor concentrations. Where this is not the case, outdoor concentrations will overestimate indoor concentrations due to various particle-clearing processes that occur during infiltration. Adding to this error are monitoring gaps caused by governmental agency resource limitations and the high expenses associated with operating common research-grade monitors. The US EPA State and Local Ambient Monitoring Station (SLAMS) network, for example, has only about one population $PM_{2.5}$ exposure monitoring site per 350,000 people (United States Environmental Protection Agency 2016). The SLAMS network is one of the most-dense in existence, yet it still leaves large, exposure-relevant geographic regions with little-to-no reliable air quality data. Methods exist to estimate concentrations in-between monitors (e.g. land use regression), but are plagued by an inability to account for sources between monitoring sites (Jerrett et al. 2005a) and can introduce severe bias when applied in the estimation of health effects (Alexeeff et al. 2015). Satellite imaging is gaining popularity as a tool for addressing the low densities in ambient monitoring networks (van Donkelaar et al. 2014), yet remains limited by spatial resolution on the order of kilometers and a focus on outdoor concentrations only.

The potential for exposure misclassification in current monitoring frameworks may have consequences for those wishing to reduce the disease burden of $PM_{2.5}$. The efficiency and effectiveness of air quality interventions depend upon accurate estimates of population risks. A misestimation of exposures can either lead to an overestimation of risk, rendering policies and projects inefficient, or an underestimation of risk, leaving vulnerable populations unprotected. Moreover, evidence is mounting to suggest that the relationship between $PM_{2.5}$ exposure and risk is non-linear for several diseases (Burnett et al. 2014). This means that the change in risk per unit of $PM_{2.5}$ exposure is different at different levels of exposure. At certain concentration ranges, exposure misclassification can result in disproportionate overestimation or underestimation of risk. The effect varies by disease, but is particularly pronounced in ALRI in young children across a range of concentrations relevant to exposures in developed and developing nations alike. In some other diseases, the risk-response relationship flattens out at moderate-to-high exposures. Policy makers in regions with populations afflicted by exposures in this range must thereby aim to bring exposures down to truly clean levels – rather than making modest marginal reductions – in order to see significant risk benefits.

1.2 Layout and primary contributions to the literature

The work discussed in this dissertation was inspired by the extensive scale of the health impacts of $PM_{2.5}$ and the current inadequacies of the environmental health research and policy fields in assessing related exposures. Chapter 2 estimates population-wide $PM_{2.5}$ exposures and related health impacts in a highly polluted city with heavy residential solid fuel use. Exposures were constructed from micro-environmental concentrations (indoor vs. outdoor) and the amount of time spent by the population in each and accounted for effects from second hand smoke. Chapter 3 moves beyond exposure constructions to explore how population-wide exposures in the context of solid fuel use can be predicted from survey and environmental data using advanced machine learning and ensemble modeling techniques. Chapter 4 is motivated in part by the sparsity of indoor $PM_{2.5}$ concentration data available for analyses like those performed in Chapters 2 and 3, and demonstrates the potential for a popular and relatively inexpensive household product to reliably measure real-time $PM_{2.5}$ concentrations.

Chapter 2 provides one of the first calculations of city-wide $PM_{2.5}$ exposures in Ulaanbaatar, Mongolia – one of the most polluted capital cities in the world (The World Bank 2011). Some work has been done to characterize $PM_{2.5}$ concentrations in the city (Allen et al. 2013; Enkhbat et al. 2016; Greene et al. 2014a) and exposures of individuals (Nakao et al. 2016; Ulzii et al. 2015), but population-wide exposures integrated across seasons and environments have, to my knowledge, remained un-estimated. Chapter 2 provides Ulaanbaatar's policy makers with actionable information by calculating population exposure and disease burden under current emissions policies and two more-stringent alternative pathways between 2014-2024. Annual average city-wide exposures are calculated from estimates of indoor concentrations, outdoor concentrations, secondhand smoke impacts, and time-activity as well as projections of background disease burden and population. The presentation of methods for doing so may prove useful for investigators interested in performing similar work elsewhere. The analysis is also one of the first to apply the $PM_{2.5}$ exposure-response functions (Burnett et al. 2014) of the 2010 Comparative Risk Assessments of the Global Burden of Disease Study (Lim et al. 2012) in a forward projection of exposure-related disease burden.

Chapter 3 presents $PM_{2.5}$ exposure values measured in women cooking with solid fuel in Lao PDR, which are among the first to be published, and adds to the small pool of HAP data available for Southeast Asia (Huang et al. 2013; Morawska et al. 2011; Shimada and Matsuoka 2011). The application of cross-validated models using advanced machine learning and ensemble techniques in Chapter 3 shows that personal $PM_{2.5}$ exposures in a population cooking with solid fuels can be reliably predicted from survey data, environmental information, and HAP measurements along with personal measurements in just a subset of the group. Statistical procedures also reveal which types of variables may provide the most utility in such a model. Together, these analyses help to lay the groundwork of an exposure modeling framework that will reduce the resources required to

reliably estimate $PM_{2.5}$ exposures in homes were solid fuel cooking and heating are common.

Chapter 4 discusses the use of the smart smoke detector – an increasingly popular household product – to expand $PM_{2.5}$ monitoring networks and dramatically increase the quantity of indoor concentration data available for estimating exposures and informing health-related behavior. This concept is validated by reverse engineering a Nest Protect smart smoke detector, co-opting its onboard optical sensor, and characterizing that sensor's response to $PM_{2.5}$ alongside an industry standard $PM_{2.5}$ monitor, the DustTrak II. Chapter 5 concludes with a brief summary of the common thread in chapters 2-4, and a discussion of future work that might expand upon the analyses therein.

Chapter 2 [6]

Health assessment of future $PM_{2.5}$ exposures from indoor, outdoor, and second hand tobacco smoke concentrations under alternative policy pathways in Ulaanbaatar, Mongolia

2.1 Introduction

Air pollution is responsible for over 10% of global deaths annually (Forouzanfar et al. 2015). As discussed in Chapter 1, the greatest potential for air pollution related disease exists in regions in which households rely on solid fuels to meet various energy needs. Through this lens air pollution in Asia, where residential solid fuel use is common (Bonjour et al. 2013), has been heavily studied – e.g. (Balakrishnan et al. 2013; Baumgartner et al. 2011; Chafe et al. 2014; Smith 2000; Venkataraman et al. 2010; Zhang and Smith 2007) – but those efforts have focused primarily on charismatic nations like China and India. Only recently have researchers begun to direct their attention to the remote country of Mongolia, where an air pollution crisis is brewing.

Mongolia's capital city, Ulaanbaatar (UB), is home to nearly half of the nation's three million residents. Driven initially by political-economic changes incurred after the fall of the Soviet Union and hastened by periodic bouts of harsh weather and famine (Dore and Nagpal 2006), called dzud, UB is now experiencing rapid rural-to-urban migration and an increasing demand for coal. Between 1990 and 2014, the population of UB grew from 570,000 people (~27% of the Mongolian population) to nearly 1.4 million (~46% of the population) (Mongolian Statistical Information Service 2016). These factors, together, lead

[6] This chapter is a slightly modified version of a manuscript submitted for peer-review as: "Hill, LD, Edwards, R, Turner, JR, Damdinsuren, Y, Olkhanud, P, Odsuren, M, Guttikunda, S, Ochir, C, Smith, K. Health assessment of future $PM_{2.5}$ exposures from indoor, outdoor, and environmental tobacco smoke concentrations under alternative policy pathways in Ulaanbaatar, Mongolia. *In review. 2017.*" The title was revised during review to (at the time of dissertation filing) to: "Health assessment of future $PM_{2.5}$ exposures from indoor, outdoor, and second hand tobacco smoke concentrations under alternative policy pathways in Ulaanbaatar, Mongolia." This also applies to all figures, tables, and pictures therein. At the time of dissertation filing, a final response had not been received from reviewers.

to winter outdoor air pollution concentrations in UB that are among the worst in the world. About half of UB residents currently live in houses or gers—traditional yurt-like dwellings— heated by simple, coal-fired stoves. Recent measurements of fine particulate matter ($PM_{2.5}$) have shown citywide wintertime average concentrations as high as 250 $\mu g/m^3$ and annual average outdoor concentrations that are over seven times higher than the World Health Organization (WHO) health-based guidelines established to minimize morbidity and mortality risk and three times higher than Mongolian national standards (Allen et al. 2013; United Nations Children's Fund Mongolia 2016). Daily averages are often even higher; between January-February 2016, the United States Embassy in UB reported 20% of 24-hour average values greater than 400 $\mu g/m^3$ (OpenAQ 2016).

The high air pollution levels in UB arise from a combination of high anthropogenic emissions, geography, and meteorology (Allen et al. 2013; United Nations Children's Fund Mongolia 2016). Geographically, the city lies in a valley surrounded on three sides by mountain slopes onto which the peri-urban areas of the city have spread. Although these districts are referred to as the "ger areas," approximately 60% of the dwellings in these districts are one and two-story houses constructed by local residents. In winter months, UB is strongly affected by the mid-continental Siberian anticyclone, a semi-permanent high-pressure system that is characterized by stagnant air masses centered over northern Mongolia. The anticyclone forms because of intense cooling of the surface layers of air over the continent during this season, leading to a well-developed temperature inversion in the lower atmosphere. These two factors by themselves would not lead to severe air pollution without high pollutant emissions largely caused by the burning of raw coal for winter heating. Recent analyses estimate that about two thirds of the annual average $PM_{2.5}$ concentration in UB is from combustion sources with nominally equal contributions from coal-fired power plants and residential heating (Davy et al. 2011). However, this split seems unlikely given the power plants have tall stacks which release emissions above the very shallow wintertime inversion layers. In winter, up to 70% of $PM_{2.5}$ emissions in the ger areas are attributable to residential heating (World Bank Asia Sustainable and Alternative Energy Program 2009). Although UB's vehicle fleet is rapidly increasing (Allen et al. 2013), the contribution to ambient $PM_{2.5}$ concentrations is estimated to be only in the range of 5-10% (Davy et al. 2011) due to the large impact of coal emissions. Ultimately, once fine particulate matter is emitted into the valley, its dispersion is constrained by a very low atmospheric mixing height (Guttikunda et al. 2013).

The Mongolian government has undertaken efforts in this decade to reduce air pollution, including the subsidization of cleaner-burning coal stoves, elimination of many institutional heat only boilers (HOB), and the promotion of energy efficiency (Greene et al. 2014a). However, expansion of the peri-urban areas combined with the use of high-emission residential heating stoves continues to produce high $PM_{2.5}$ concentrations in UB, especially in winter. Current residents are facing exposures never before experienced by Mongolians, and urgent measures are needed to reduce consequent health impacts. A 2013 analysis (Allen et al. 2013) of outdoor concentrations in UB conservatively estimated that

$PM_{2.5}$ was responsible for 29% of cardiopulmonary deaths, 40% of lung cancer deaths, and nearly 10% of all-cause mortality. The World Bank (The World Bank 2011) places the greatest health insults in ger areas, where UB's poorest and most vulnerable populations reside, and project steady increases in medical and economic impacts into the foreseeable future without significant policy changes. Leaders agree with the need to eventually reduce emissions across all energy sectors, but the benefits of doing so quickly rather than more slowly, a choice with substantial differences in costs and strategy, have remained unclear. To assist Mongolia's policymakers at this critical juncture, we modeled future air pollution exposures and related health burdens in UB through 2024 under business as usual (BAU) and two alternative energy policy pathways: moderate emissions reductions across power, heating, and vehicle sectors (Pathway 1); and a major shift in these same sectors to clean fuels and technologies (Pathway 2). These pathways were based on technologies that have been successfully adopted by other countries, and represent progressively more aggressive adoption of clean fuels in order to inform on what measures may result in reduced health impacts across the population.

Our estimates improve on previous local and regional population-wide burden assessments (Allen et al. 2013; Forouzanfar et al. 2015; Lim et al. 2012; The World Bank 2011) by focusing on personal exposures and incorporating local measurements of indoor concentrations and stove emissions. Our analysis employs an ambient air quality model with high spatial and temporal resolution. In contrast, hybrid satellite methods, which have been used in recent regional-scale evaluations (Forouzanfar et al. 2015; Lim et al. 2012), have coarser spatial resolution and have been found particularly unreliable in UB because of poor resolution for winter nighttime (Brauer et al. 2016; van Donkelaar et al. 2014). The use of spatiotemporally-resolved models of ambient air pollution based on local emissions allows current policies for urban development in UB to be evaluated in relation to the health impacts of alternative policy pathways. The methods developed for this analysis are among the first to incorporate the exposure-response functions of the 2010 Comparative Risk Assessments of the Global Burden of Disease Study (GBD) in a *forward-looking* analysis that focuses not only or separately on outdoor or indoor pollution, but on *total exposure* of the population.

2.2 Methods

Personal $PM_{2.5}$ exposures and related disease burdens were modeled for UB residents through 2024 under BAU and two alternative policy pathways. Indoor and outdoor concentration estimates were combined with time-activity data, census information, demographics projections, and estimated smoking rates to provide population-weighted total exposure estimates. Figure 2.1 shows a summary flow chart of the exposure and disease burden calculation framework. Disease-specific estimates of health burden were produced using a modified version of the Household Air Pollution Intervention Tool (Pillarisetti et al. 2016) and projections of city-wide background disease. Data handling, mathematical modeling, and figure creation for indoor concentrations, demographic projections,

exposure modeling, and health burden estimation were performed in R (R Core Team 2016) and Microsoft Excel for Mac 2011 and 2016.

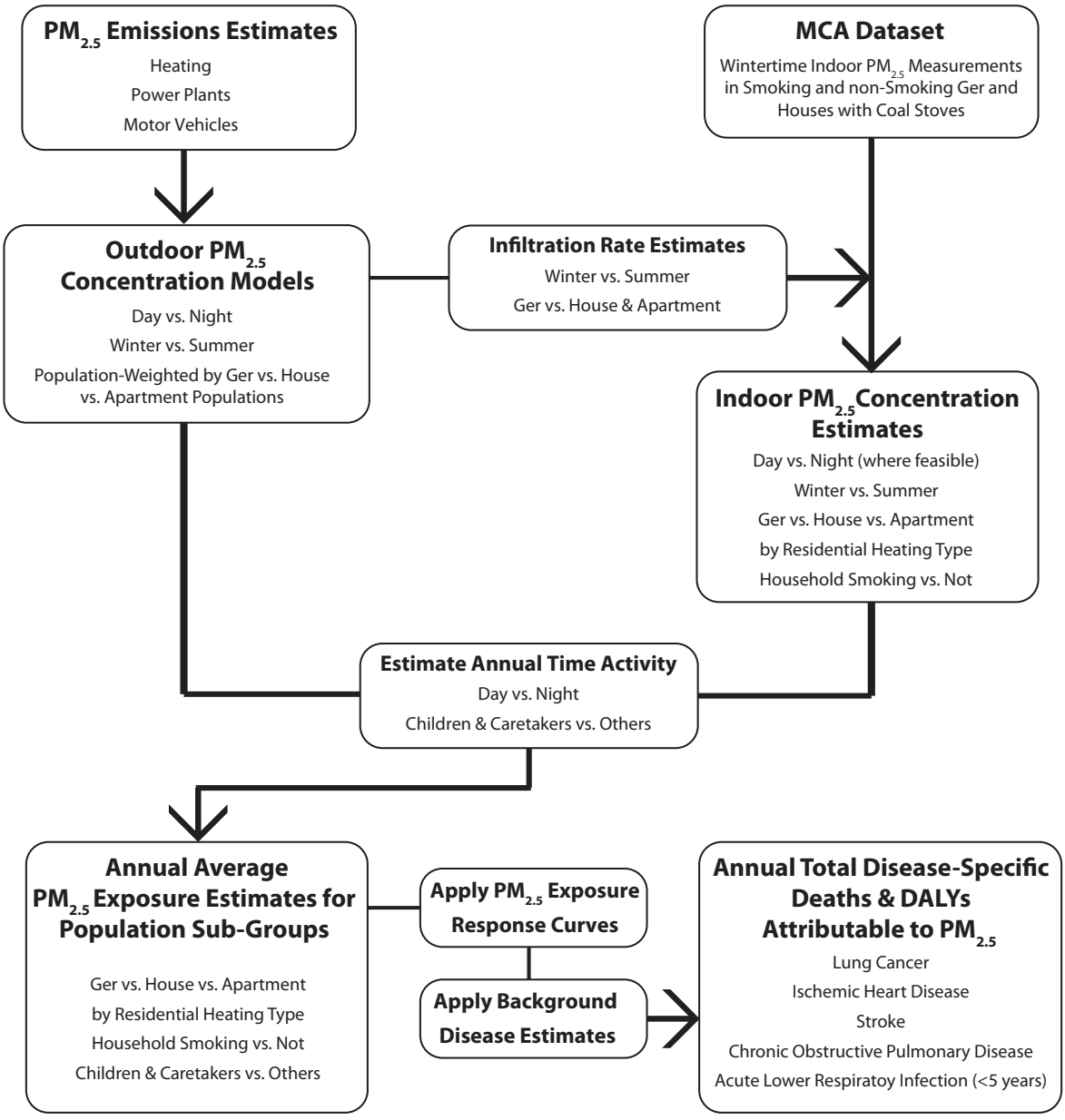

Figure 2.1. High-level flow chart of the general exposure and disease analysis approach. Annual average exposures were estimated for population sub-groups in 2014 and in 2024 under BAU and two alternative policy pathways from outdoor and indoor concentration models and time activity estimates using the approach summarized above. These exposures were applied to disease-specific exposure response curves to produce estimates of population attributable fraction (PAF) which were applied to background disease rates to quantify attributable disease burden. Detailed data descriptions and methods – including how interim year (2015 – 2023) disease burdens were calculated – are included in the manuscript and Appendix B.

Baseline and the pathways

Variations in the emissions trajectories of heating, power, and traffic sectors were considered in relation to policy pathways from baseline that followed business as usual or one of two alternative policy approaches: moderate restrictions in addition to those in place at baseline (Pathway 1) and additional aggressive restrictions (Pathway 2). The variations examined were most detailed for household heating, which has been identified as the single largest contributor to outdoor air pollution in UB (Guttikunda et al. 2013). Household heating types in Ulaanbaatar include stoves fired with raw coal; stoves fired with semi-coke fuels (semi-coke stoves), which are relevant only to houses and gers; coal-fired low pressure boilers (LPB) used to heat radiator systems in houses; and heating sources that produce no indoor emissions at the point-of-use (Clean Indoor Use, or CIU heat), such as HOB or centrally-distributed steam that is produced during combined electricity and heat generation, which can be deployed in all home types. Table 2.1 and Table 2.2 show a summary of baseline and 2024 under BAU and the two alternative energy policy pathways considered.

Table 2.1 . Summary of the assumptions made for emissions sources, by category [1]

	Household Heating	Power Plants	Vehicles
Baseline (2014)	20,000 LPB-heated houses; semi-coke coal in gers & houses in Bayangol district; apartments heated with CIU heat including HOB units; MCA stoves in all other gers & houses	Four CHP: CHP-2, CHP-3 (two units), and CHP-4	Nearly 100% growth over values from 2010 – the most recent inventory at the time of analysis
BAU (2024)	All homes, except LPB and clean-heat homes, transitioned to MCA stoves	Addition of CHP-5, which meets U.S. New Source Performance Standards	2.5%/year growth from 2014 and addition of Euro III standards
Pathway 1 (2024)	Transition of half of non-LPB houses to clean heat; replacement of remaining MCA stoves with "Future Tech" raw coal stove; 50% of HOB decommissioned, others retrofitted with controls	Addition of CHP-5; high-efficiency retrofits of CHP-2, CHP-3, and CHP-4.	2.5%/year growth from 2014 and addition of Euro V standards
Pathway 2 (2024)	Transition of all ger and houses to clean heat; all HOB decommissioned	Addition of CHP-5; CHP-3,-4 retrofitted; CHP-2 replaced by renewables and/or imports	Increased mass-transit ridership; improved traffic flow; Euro VI standards; among others

1. LPB = low pressure boiler, CIU = heating method with clean indoor use, HOB = heat only boiler, MCA = improved coal stove distributed by the Mongolian government and the U.S. Millennium Challenge Account, CHP = combined heat and power plant.

Table 2.2. Assignment of heating type of all UB households in BAU and each alternative pathway.

	Proportion of Gers	*Proportion of Houses*	*Proportion of Apartments*
2014 – Baseline			
MCA Stove [1]	92.9%	75.2%	0%
Stove w/Semi-Coke Fuel	7.1%	6.0%	0%
Low Pressure Boiler	0%	18.8%	0%
Heat w/Clean Indoor Use	0%	0%	100%
2024 – BAU			
MCA Stove	100%	82.9%	0%
Low Pressure Boiler	0%	17.1%	0%
Heat w/Clean Indoor Use	0%	0%	100%
2024 – Pathway 1			
Future Tech Stove	100%	41.4%	0%
Low Pressure Boiler	0%	17.1%	0%
Heat w/Clean Indoor Use	0%	41.4%	100%
2024 – Pathway 2			
Clean	100%	100%	100%

1. Improved coal stove distributed by the Mongolian government and United States Millennium Challenge Account (MCA).

Recent efforts by the Mongolian Government, the U.S. Millennium Challenge Corporation and Millennium Challenge Account (MCA) (Greene et al. 2014a), and the World Bank have aimed to replace the majority of traditional raw-coal stoves with coal stoves that substantially reduce outdoor emissions, hereafter "MCA stoves." For this reason, complete penetration of MCA stoves was assumed in gers and houses not using LPB or semi-coke coal. At the time of the study, gers and houses in the Bayangol district were not targeted by MCA stove dissemination plans due to a concurrent semi-coke coal intervention underway in the district. Homes in this district were assumed to rely on traditional stove technology but with cleaner semi-coke coal. Based on 2012 census data 7% of all ger and 6% of all houses in the city were assigned to the Bayangol district at baseline (Statistics Department of Ulaanbaatar 2013). We assumed 20,000 LPB households with one LPB per home, based on information provided to us by government officials at the Clean Air Fund. This was consistent with data showing 14,186 LPB households in 2010 (National Statistics Office of Mongolia 2012). All apartment households were assigned CIU heating at the household, provided either by steam heat or heat only boiler (The World Bank 2011). UB currently employs four combined heat and power (CHP) plants in the generation of heat and electricity— CHP-2, CHP-3 (two units), and CHP-4. Baseline assumed this setup. Motor vehicle data for UB were sparse, and so BAU and the alternative pathways were based on plausible assumptions about trends in total fleet emissions. Values were scaled from a 2010 inventory (Japan International Cooperation Agency 2013) of vehicle exhaust emissions from travel on major and minor roads— this excluded emissions related to brake wear, tire wear, or re-suspended dust. Baseline $PM_{2.5}$ emissions were estimated as 1.7 times the 2010

inventory which was the growth in the number of registered vehicles between 2010 and 2013. This is likely an overestimate because of the partial offset from a vehicle travel day ban program which requires most vehicles not be used one day per week based on license plate number.

Business as Usual assumed no major changes from trends underway at the time of the study (mid-2013) by 2024. This included a transition to MCA stoves of all homes not employing clean heat or LPB at baseline. Because HOB and LPB are outdated technologies, no net increases in the number of HOBs or LPB-heated homes were assumed. BAU retained the power plant emissions in 2024 and included a new 820 MW power plant (CHP-5) to be located 15km east of the UB Central Business District, as supported by recent government plans to develop a 450 MW plant and expand it to 820 MW shortly thereafter (HJI Group Corporation and MonEnergy Consult Co. Ltd. 2011; Kohn 2013). CHP-5 was assumed to meet the U.S. New Source Performance Standards for electric utility power plants. Vehicle emissions assumed a growth of 1.3 times the 2014 inventory based on an emissions growth rate of 2.5%/year over the 10-year period and Euro III emissions standards – a higher growth rate seemed unreasonable given the existing transportation network infrastructure.

Pathway 1, or moderate emissions reductions, assumed all changes in BAU as well as some moderate improvements. All 20,000 LPB-heated homes from baseline remained as such under Pathway 1. Remaining non-LPB houses were assumed equally split between clean heating and an even cleaner hypothetical coal stove, called the 'Future Tech' stove, which improved the emissions performance of the MCA stoves by the same percentage as the MCA stoves improved upon the traditional stoves and improved indoor concentrations by 20% compared to those in MCA stove homes. Half of all HOB units from baseline were assumed decommissioned by 2024 under Pathway 1, and the other half were assumed retrofitted with cyclone control technologies. All other households were assumed to rely on clean heat from other sources. In addition to the power plant assumptions under BAU, Pathway 1 assumed high efficiency control devices, such as electrostatic precipitators, installed on units CHP-2, CHP-3, and CHP-4. This is a significant upgrade to the existing CHP infrastructure, which includes wet scrubbers or electrostatic precipitators, depending on the facility. For vehicles, the BAU rate of growth was assumed for Pathway 1, but with the implementation of Euro V emissions standards.

Pathway 2, or transition to cleaner fuels and technologies, assumed feasible but ambitious rates of change in all sectors by 2024. Solid fuel combustion was assumed eliminated in households. CHP-3 and CHP-4 were assigned high efficiency control technologies. CHP-2 was decommissioned by 2024 and replaced with renewables and/or imports (i.e. sources with negligible impacts on UB air quality). A 50% reduction in traffic emissions over Pathway 1 was assumed, opportunities for which include but are not limited to higher adoption rates for mass transit use, transportation network enhancements to improve traffic flow, and adoption of Euro VI standards, which include an additional 50% reduction in PM emission rates from heavy duty diesel vehicles compared to Euro V standards.

Throughout BAU and the alternative pathways smoking prevalence among households (not individuals) was maintained at 45% of households. This figure was based

on a series of surveys (Greene et al. 2014b, 2014c) of gers and houses in UB conducted by the Millennium Challenge Corporation during the 2012-2013 winter and corroborated by recent studies of national smoking rates (Demaio et al. 2014; World Health Organization 2010).

Population and household numbers

Demographic conditions were estimated for 2006 through 2024. The methods and sources used are described in detail in Appendix B. Briefly, city-wide population data all residents and, specifically, those < 5 years old were estimated from historical data and National Statistics Office medium growth projections (National Statistics Office of Mongolia 2012). Detailed information on the spatial distribution of population and household number by household type was obtained from 2012 city census data (Statistics Department of Ulaanbaatar 2013). Estimates were made for the number, type, and general location of households in UB through 2024 using historical demographic data and official city projections. Household types most relevant to the UB context were identified as gers, single-family houses, and multi-family apartments. Gers in the peri-urban area of UB are circular traditional dwellings with multiple felt layers and a waterproof outer shell covering a wooden lattice frame (Figure 2.2). Houses in the peri-urban area are often locally constructed wood, cement, or brick structures and, while they vary considerably in layout and construction, are generally in the style of traditional western houses in which one extended family resides. Apartments are identified as buildings within which two or more families are living and, in most cases, are large complexes that house dozens of families.

The 2014 population of UB was estimated at 1,355,176 residents distributed among 86,246 gers, 106,353 houses, and 179,718 apartments. Projections indicated that by 2024 the population would grow nearly 40% and experience a substantial shift to apartment dwelling with over 65% of the population living in multi-family buildings (Table 2.3). It should be noted that while our stove-number projection technique employed an underlying assumption that each household typically relied on a single stove, some homes may employ more than one stove (e.g. a 'home' may consist of two ger, each heated by their own stove) (Greene et al. 2014a). In stove-heated homes, outdoor models accounted for this by inflating stove estimates and, thus, stove emissions upwards by 20% (Greene et al. 2014a), while the exposure model was applied using a single primary stove type for each household.

Figure 2.2. Left, a traditional ger dwelling. Right, houses typical of the peri-urban regions of UB. (Credit: L. Drew Hill)

Table 2.3. Estimates of age-specific city-wide population and household number by home type, and estimated household size

Year	Population [1]		Number of Households [2]			Pop. Per Household [2]
-	Total Pop.	Pop. 0-4 Years	Ger	House	Apartment	
2014	1,355,176	148,219	86,246	106,353	179,718	3.64
2015	1,407,196	155,551	88,547	109,191	197,539	3.56
2016	1,459,516	158,438	90,684	111,826	216,586	3.48
2017	1,511,836	161,325	92,616	114,209	236,854	3.41
2018	1,564,157	164,212	94,323	116,313	258,369	3.34
2019	1,616,477	167,099	95,781	118,112	281,152	3.27
2020	1,668,797	169,986	96,967	119,574	305,219	3.20
2021	1,715,748	168,427	96,997	119,611	330,782	3.13
2022	1,762,700	166,869	96,667	119,204	357,645	3.07
2023	1,809,651	165,310	95,954	118,324	385,792	3.02
2024	1,856,603	163,752	94,834	116,943	415,195	2.96

1. Interpolated from five-year "medium growth" (version 1b) projections identified in the 2010 Population and Housing Census of Mongolia Report (National Statistics Office of Mongolia 2012).
2. Estimated using the techniques and sources described in Appendix B.

Outdoor ambient air quality modeling

Air quality modeling was conducted to estimate outdoor $PM_{2.5}$ mass concentrations. The modeling methodology followed that used by Social Impact for an impact evaluation of the Energy-Efficient Stove Subsidy Program of the Millennium Challenge Mongolia Energy and Environment Project (MCA impact evaluation), detailed in the full report (Greene et al. 2014a). Emission categories were expanded to include LPB, HOB, motor vehicles, and Combined Heat & Power plants in addition to residential heating stoves. Sources not

included in the model were heating stoves in kiosks, industrial emissions including kilns, re-suspended road dust, and windblown dust.

Emissions data

Table 2.4 summarizes total annual emissions from the major sources considered at baseline and in 2024 under BAU the two alternative development pathways.

Table 2.4. Estimated annual $PM_{2.5}$ emissions from major sources (tons/yr)

Pathway	Vehicles	Power Plants	Heat Only Boilers	Household Stoves & LPB
2014 Baseline	384	11,500	1,300	1,700
2024 BAU	500	12,000	1,300	1,900
2024 Pathway 1	96	1,900	390	640
2024 Pathway 2	48	1,830	0	0

Residential heating stove emissions were assumed zero during the summer period, April through September. An MCA stove emissions profile was taken from values reported in the MCA impact evaluation as weighted by the sales-based prevalence of three variations of the MCA stove (Ulzii, Khas, and Dul) detailed in the same publication (Greene et al. 2014a). Data on the emissions profiles of low-pressure boilers and semi-coke coal stoves in UB were unavailable, and so they were conservatively assigned the emissions profiles of MCA stoves. "Future Tech" stove outdoor emissions were assigned by applying to the MCA emissions profile the same ~ 60% reduction estimated for the transition from traditional coal stoves to MCA stoves during the Social Impact evaluation (Greene et al. 2014a). HOB emissions were informed by a PM_{10} HOB emissions inventory prepared by the Japanese International Cooperation Agency (JICA) (Japan International Cooperation Agency 2013). All HOB stack emissions were assumed to be in the $PM_{2.5}$ size range. These inventories were used with no modifications for the baseline (2014) and BAU (2024) $PM_{2.5}$ emission inventories. Pathway 1 assumed that in 2024, an overall 70% reduction in HOB emissions would be reached, which was consistent with decommissioning 50% of HOB and adopting high efficiency cyclones as a control strategy on all others (Japan International Cooperation Agency 2013). Pathway 2 assumed all HOBs would be decommissioned by 2024.

Baseline power plant emissions were taken from a recent JICA (Japan International Cooperation Agency 2013) PM_{10} emission inventory for CHP-2, CHP-3, and CHP-4. All stack emissions were assumed to be in the $PM_{2.5}$ size range. CHP-5 was assumed to meet the U.S. New Source Performance Standards for electric utility power plants, which is 0.015 lb PM/MMBtu (United States Environmental Protection Agency 2006), and have an electricity generation rate of 1870 kWh/ton coal and a coal heat content of 19.53 MMBtu/ton. Assuming the plant would operate continuously throughout the year, the

estimated PM emissions were 511 tons/year, all of which were assumed to be $PM_{2.5}$. In Pathway 1, the PM capture rate for the high efficiency control devices installed on units CHP-2, CHP-3, and CHP-4 was conservatively assumed at 98%. The 98% capture rate was applied to an assumed uncontrolled emission factor of 16.6 kg PM/ton coal. The renewable and imported generation capacity in Pathway 2 assumed no emissions-related impacts on UB.

A simple scaling approach as previously described was used for motor vehicle emissions pathways, and did not account for changes in the fleet composition over time as insufficient details for the JICA 2010 inventory were available to make more sophisticated projections. All exhaust emissions were assumed to be in the $PM_{2.5}$ size range. For diesel vehicles the Euro V PM emission standards are 80% -93% lower than the Euro III standards depending on vehicle class. There were no Euro standards for PM emissions from gasoline-fueled vehicles and thus 90% overall reduction would not be realized between BAU and Pathway 1. However, gasoline vehicle Total Hydrocarbon standards are 50% lower for Euro V compared to Euro III. This could result in some PM reductions for the cold wintertime conditions, which favor semi-volatile gaseous compounds entering the particle phase. Overall, Pathway 1 employed a 75% reduction in vehicle emissions in 2024 compared to those at baseline. Pathway 2 assumed a simple 50% reduction in traffic emissions over Pathway 1, as previously discussed.

Model techniques

Existing power plants were modeled as point sources using available geographic location and stack properties data (Japan International Cooperation Agency 2013). Residential heating stoves, HOB, and motor vehicle emissions were modeled as area sources. The greater UB region was discretized into 6,298 grid cells, each with dimension 1 km × 1 km. Emissions were allocated to these grid cells and the center of each grid cell was used as a receptor site for which modeled $PM_{2.5}$ concentrations were generated. Modeling was conducted at hourly resolution using the ISCST3 Gaussian dispersion model (United States Environmental Protection Agency 1995a, 1995b) with pre- and post-processing using GIS (Environmental Systems Resource Institute 2012). Two satellite districts – Baganuur and Bagakhangai– were excluded from modeling. These districts are remote, low population zones that are not contiguous with UB's other districts. No HOB or LPB emission sources were assigned to these districts; all previously described HOB and LPB sources were included in the modeled districts only. Seasonal-average $PM_{2.5}$ concentrations for these districts, which together account for less than 3% of all households, are assumed to be at the 10^{th} percentile (decile) of population-weighted $PM_{2.5}$ concentration distributions for the remaining districts that were modeled.

Estimation of residential heating stove emissions and allocation of these emissions in space and time followed approaches developed for the Social Impact MCA impact evaluation (Greene et al. 2014a). The baseline pathway treated all stoves as MCA stoves, which were

increased proportionally to adjust for population growth to 2024. Stoves emissions were modeled as area sources corresponding to the 1 km × 1 km grids.

Motor vehicle and HOB emissions were spatially allocated using emissions fields from 2010 with a resolution of 0.01°× 0.01° (Guttikunda et al. 2013) that were re-projected in GIS by contouring the data and then calculating area weighted means for the 1 km × 1 km grids used for the modeling. HOB emissions were temporally allocated using weights employed in a 2013 examination of particulate pollution in UB (Guttikunda et al. 2013). Motor vehicle emissions were held constant for each season and were allocated to hour of day using a typical urban profile with morning and afternoon rush hour peaks.

Year 2012 population by dwelling type (ger, house, and apartment) at the level of Khoroo, or Mongolian administrative sub-division similar to a sub-district, was also allocated to the 1 km × 1 km grids using area weighted sums. Projected changes in the peri-urban population between 2012 and the baseline and 2024 BAU and alternative pathway years were distributed across grid cells in proportion to the 2012 peri-urban population for both gers and houses. Projected changes in the population residing in apartments were allocated in proportion to the total population in each grid.

Air quality modeling was conducted at hourly resolution using meteorology data from April 2012 through March 2013 provided to us by the National Agency for Meteorology, Hydrology, and Environmental Monitoring of Mongolia – data available to other users upon written request to the Environmental Monitoring Department at what is now the National Agency for Meteorology and Environmental Monitoring of Mongolia. Un-modeled emission sources were assumed to have a spatially and temporally constant contribution of 10 $\mu g/m^3$ across the city and over the ten-year assessment period. The model underestimated outdoor $PM_{2.5}$ measurements conducted during the 2012-2013 winter heating season and these measurement data (Greene et al. 2014b, 2014a, 2014c) were used to calibrate the model. The hourly modeled concentration fields were used to construct daytime (8:00-18:00) and nighttime 18:00-8:00) average concentrations for summer (April through September) and winter (October through March). These gridded concentration estimates were combined with the gridded population data to estimate citywide population-weighted outdoor $PM_{2.5}$ concentrations by home type (ger, house, apartment). Further detail about the modeling and calibration are provided in Appendix B.

Indoor air quality estimates

The vast majority of gers and houses in peri-urban areas heat with raw coal lit by small amounts of wood in small chimney stoves, while apartment households almost exclusively employ CIU heat that creates no indoor emissions. These differences combined with variations in outdoor particle infiltration between building types likely result in substantially different indoor concentrations between gers, houses, and apartments. Indoor concentrations of $PM_{2.5}$ were thus estimated by home type, household heating

source, presence of second hand smoke, and season. Estimates were made for 2014 (baseline) and 2024 under BAU and the two alternative policy pathways.

Indoor air concentrations in homes with heating stoves were estimated by applying linear modeling techniques to data collected during the 2012-2013 winter season discussed in (Greene et al. 2014a) and accessible freely online (Greene et al. 2014b, 2014c). This model was designed to account for the impacts of stove type, home type, and household smoking status on household $PM_{2.5}$ concentrations while controlling for when the samples were taken. Methods and results are described in more detail in Appendix B. Wintertime indoor concentrations in homes with "Future Tech" stoves (Pathway 1) were assumed 20% lower than those found in MCA stove homes based on the assumption that such a stove would be designed to reduce indoor fugitive emissions. Wintertime indoor concentrations in homes with LPB and semi-coke stoves were assigned those of homes with MCA stoves, in accordance with the conservative emissions assignments for these stove types as discussed above.

Indoor concentrations in homes that employ CIU heating sources like district heating, HOB, electric heat, or gas-based heat were estimated differently from those with heating stoves. Such concentrations were assumed to be governed primarily by SHS and by the penetration of outdoor $PM_{2.5}$ into the indoor environment. Infiltration efficiencies were estimated at 64% in the summer and 53% in the winter for houses and apartments, and 100% in the summer and 70% in the winter for gers based on blower door tests and relevant literature detailed in Appendix B. Indoor concentrations in households with clean heating were estimated by linearly applying infiltration efficiencies to home-type specific population-weighted outdoor ambient concentrations.

Smoking rates in Mongolia are among the highest in the world (Demaio et al. 2014; World Health Organization 2010). As previously discussed, SHS was assumed present in 45 percent of households. Recent nation-wide bans on public indoor smoking (World Health Organization Western Pacific Region 2012) suggest indoor SHS may only make considerable contributions to exposure in personal, private indoor environments. Indoor concentration estimates thereby conservatively assume SHS occurred only indoors at home and thus contributions from SHS were applied only to nighttime indoor concentration estimates.

For simplicity and due to limited information on Mongolian workplace environments, the concentration profiles of the indoor environments in which the population spends their time away from home were assumed the same as those of their home indoor environments.

Time activity

Time activity information was informed by a recent survey of UB households (Greene et al. 2014b, 2014c) as well as an understanding of the local job market and commuting patterns

(Chilkhaasuren and Baasankhuu 2010). It is expected that small children spend more time indoors than the rest of the population, and so time activity was calculated separately for children (< 5 years old) and non-children (> 5 years old). Because the social pension fund provides a homecare allowance to mothers in UB with children less than 2 years old (United Nations Children's Fund Mongolia 2016) and because day care and nursery are not financially accessible to many households, it is assumed that a subset of the non-child population is charged with taking care of the children, and so every child < 5 years old was assigned one non-child as a "caretaker". Time spent by residents in each microenvironment likely differs by socio-demographics and season, but because no data on this breakdown were available, we conservatively assumed that the average day for "non-children" would be spent 75% indoors and 25% outdoors, the average day for children and caretakers would be spent approximately 100% indoors, and the average nighttime for all residents would be spent approximately 100% indoors.

Exposure estimation

The UB population was divided into sub-groups based on the major exposure-related features of the indoor model and time activity estimates: home type, heating type, presence of SHS at home, and age. More specifically, sub-groups were made for children (< 5 years old), caretakers (≥ 5 years, assumed 1 per child), and all others (≥ 5 years) in smoking and non-smoking households representing each of the following home-heating combinations: gers with MCA stoves or semi-coke coal stoves, future tech stoves, or clean heat; houses with MCA stoves or semi-coke coal stoves or low pressure boilers, future tech stoves, or clean heat; and apartments with clean heat. For exposure estimation, children, caretakers, and non-children were distributed evenly to each household, and exposures were not distinguished by gender. Baganuur and Bagakhangai—the two districts for which ambient air quality estimates were handled outside of the outdoor models—were assigned the same distribution of population sub-groups as the overall population, with the exception that none of the LPB homes were included in these excluded districts as previously discussed. Household proportions in Baganuur and Bagakhangai were identified in the 2012 city census (Statistics Department of Ulaanbaatar 2013) as comprising 3.0%, 2.1%, and 3.4% of UB's total apartment, house, and ger households, respectively. These proportions were assumed constant through 2024. Population sub-groups totaled 18 at baseline and in 2024 under BAU and Pathway 2. Pathway 1 included 30 sub-groups due to the presence of additional heating types.

Average annual $PM_{2.5}$ exposure concentrations for each population sub-group "i" were estimated at baseline and in 2024 under each pathway "j" (BAU, Pathway 1, or Pathway 2) by averaging seasonal exposure values ($_S$; winter as April – September, summer as October – March) calculated from indoor ($_{in}$) and outdoor ($_{out}$) concentrations I at night ($_N$; 18:00-8:00) and during the day ($_D$; 8:00-18:00) as weighted by the fraction of time (t) during a typical 24-hour period spent in each environment during the specified time period (Equation 1.1 1).

$$\underset{Average\ Exposure_{i,j}}{Annual} = \sum_{s=\ summer}^{s=\ winter} \left(\frac{1}{2} \cdot \left(\begin{matrix} \left(t_{in,D,i} \cdot C_{in,D,i}\right) + \left(t_{out,D,i} \cdot C_{out,D,i}\right) + \\ \left(t_{in,N,i} \cdot C_{in,N,i}\right) \end{matrix} \right) \right)_S \quad (1.1)$$

Citywide population-weighted average exposures were calculated by aggregating the exposure concentrations of each sub-group from i = 1 to i = "n", where n is the total number of sub-groups in each pathway-year j, as weighted by their representative fraction (λ) of the total population (Equation 1.2 2).

$$\underset{Average\ Exposure}{Pop.Weighted\ Annual}_{j} = \sum_{i=1}^{n} \left(\lambda_{i,j} \cdot \underset{Average\ Exposure_{i,j}}{Annual} \right) \quad (1.2)$$

Estimating health effects

Burden attributable to $PM_{2.5}$ exposures was calculated for lung cancer, ischemic heart disease (IHD), stroke, and chronic obstructive pulmonary disorder (COPD) in all UB residents as well as acute lower respiratory tract infection (ALRI) in children (ages 0-4 years) for 2014 – 2024 using a version of the Household Air Pollution Intervention Tool (HAPIT) (Pillarisetti et al. 2016) that was modified to accommodate UB-specific background data and projections through 2024. Disease-specific background mortality for the capital city was projected through 2024 using historical data for 2006-2012 provided by the Health Development Center of the Ministry of Health and Sports in conjunction with the Mongolian National University of Medical Sciences. Deaths were obtained for 2006-2012 that matched the ICD-10 codes used in the 2010 GBD (Lozano et al. 2012) to estimate illness from $PM_{2.5}$. Mortality projections were adjusted to better agree with national estimates produced by the Institute for Health Metrics and Evaluation (Lim et al. 2012). The methods and results are described further in Appendix B.

Average annual exposures were used to calculate disease-specific relative risks (RR) of mortality due to $PM_{2.5}$ exposure in each population sub-group. Mean, lower bound, and upper bound RR were taken from the integrated exposure-response functions produced by Burnett et al (Burnett et al. 2014) and applied using a counterfactual exposure of 12.0 μg/m^3 (i.e. RR = 1 at 12.0 μg/m^3). This counterfactual represents the US Environmental Protection Agency's annual for $PM_{2.5}$, which is the strictest national $PM_{2.5}$ standard in the world. Mean, lower, and upper estimates of disease-specific RR for each sub-group "i" as well as the proportion of the population that each sub-group represents, "P", were then applied to Equation 1.3 3 (Lim et al. 2012) to produce a population-wide estimate of the fraction of background mortality from each disease, "k", attributable to air pollution exposure (population attributable fraction, or PAF).

$$PAF_{j,k} = \frac{\sum_{i=1}^{n} P_{i,j} \cdot (RR_{i,j,k} - 1)}{\sum_{i=1}^{n} P_{i,j} \cdot (RR_{i,j,k} - 1) + 1} \quad (1.3)$$

Disease-specific PAF estimates for 2015-2023 were linearly interpolated from baseline and 2024 PAF values under BAU and each alternative pathway. Finally, PAF values were applied to estimated background mortality estimates to produce disease-specific estimates of total attributable death in each year. Morbidity was calculated in the form of disability-adjusted life years (DALYs). DALYs are widely used to take into account both the age distribution of premature mortality and the severity of non-fatal diseases. Disease-specific DALY estimates were calculated using the national disease-specific Death: DALY ratio produced during the 2010 GBD (Lim et al. 2012). The ratios were assumed constant throughout the projection period. The modified HAPIT used in this analysis did not include considerations for exposure cessation lag (Cameron and Ostro 2004; Doll et al. 2004); the estimated health impacts of exposure were assumed to be immediately incurred for simplicity of interpretation. Health burden was calculated not only for BAU and each alternative pathway ("accrued") but also in terms of benefit of each alternative pathway over BAU exposure levels ("averted").

2.3 Results

Outdoor ambient air quality models

Figure 2.3 shows modeled wintertime (October-March) average outdoor $PM_{2.5}$ concentrations for BAU and the alternative pathways. Pathway 2 is excluded because the highest average concentration attributed to the modeled sources was ~2 $\mu g/m^3$ (12 $\mu g/m^3$ when accounting for non-modeled sources). For the panels shown in Figure 2.3, the model predicts large variations in $PM_{2.5}$ mass concentrations across UB with highest concentrations in the ger areas where residential stoves and HOBs have the largest impact. Given these large spatial variations, outdoor concentration levels between baseline, 2024 BAU, 2024 Pathway 1, and 2024 Pathway 2 are compared using population-weighted measures. Table 2.4 presents the population-weighted mean outdoor $PM_{2.5}$ for modeled districts. Pathway 1 reduces 2024 wintertime population mean concentration by 65% compared to the 2024 BAU pathway, but the wintertime mean concentration value of 55 $\mu g/m^3$ is still quite high. Figure 2.4 shows box plots for the population-weighted distribution of wintertime outdoor concentrations. For BAU and each of the alternative policy pathways, 10% of the population resides in areas with $PM_{2.5}$ outdoor concentrations ~50% higher than the mean pathway-specific outdoor concentration reported in Table 2.5.

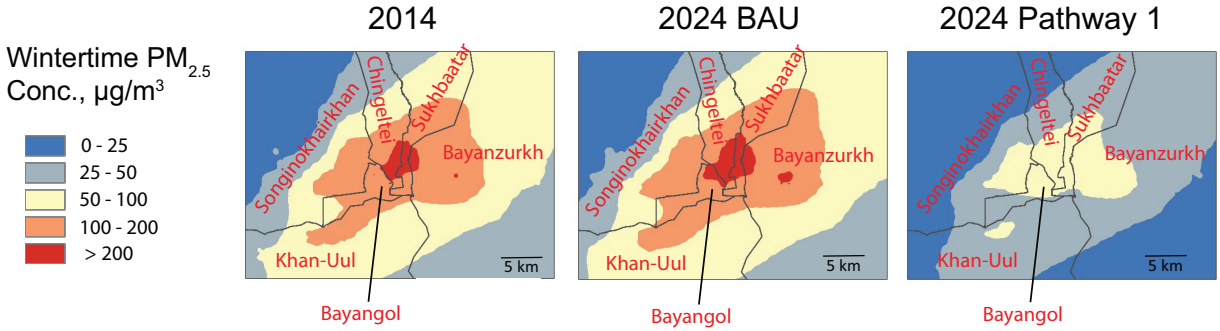

Figure 2.3. Winter average outdoor $PM_{2.5}$ concentrations for baseline and 2024 under BAU and Pathway 1. Figure created by Dr. Jay Turner, who performed the outdoor modeling. Khoroo (district) names modified by LAH.

Table 2.5. Modeled population weighted mean $PM_{2.5}$ outdoor concentrations ($\mu g/m^3$) by season and year

Pathway	Summer	Winter			
	Total Pop.	Total Pop.	Ger Pop.	House Pop.	Apt. Pop.
2014	16	141	140	148	137
2024 BAU	19	156	154	163	154
2024 Pathway 1	11	55	55	58	55
2024 Pathway 2	11	11	11	11	12

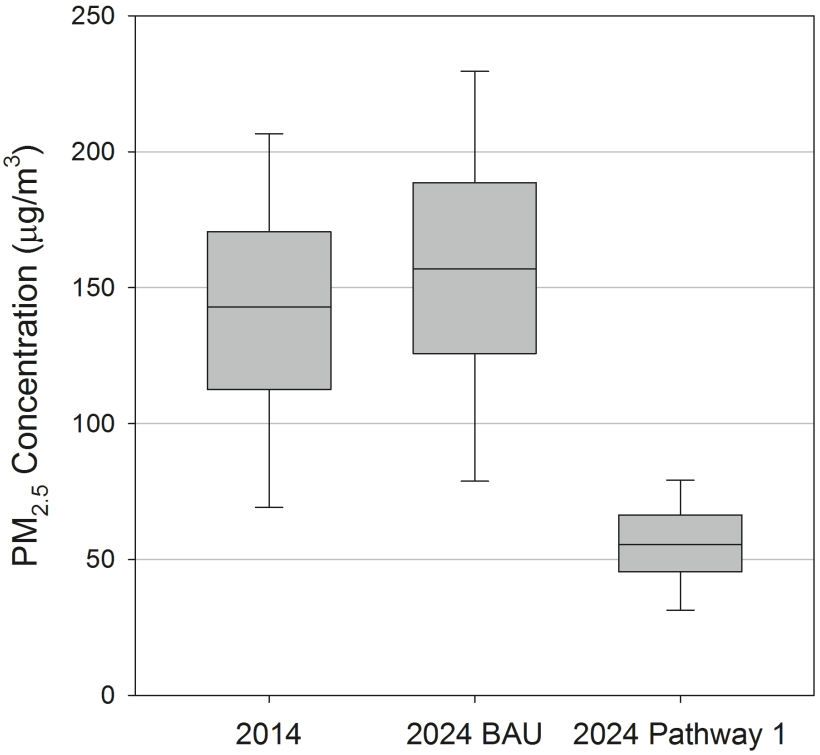

Figure 2.4. Population weighted wintertime outdoor $PM_{2.5}$ concentrations. Whiskers represent 10th and 90th percentile concentrations. Figure created by Dr. Jay Turner, who performed the outdoor modeling.

Indoor air quality estimates

Average wintertime indoor concentrations for non-smoking homes heating with MCA stoves, LPB (houses only), and stoves using semi-coke coal were modeled at 107.0 µg/m^3 for gers and 118.3 µg/m^3 for houses. Wintertime indoor concentrations for homes with Future Tech heating stoves were assigned at 20% lower than homes with MCA stoves. Population-weighted indoor wintertime concentrations for non-smoking homes with clean heating at baseline and in 2024 are shown in Table 2.6. Estimates of population-weighted summertime indoor concentrations in non-smoking homes at baseline and in 2024 are shown in Table 2.7. Based on the results of the indoor concentration model, a contribution from SHS of 18.1 µg/m^3 in gers and 20.0 µg/m^3 in houses and apartments was applied to the nighttime indoor concentration estimates of smoking households.

Table 2.6. Estimated average wintertime indoor $PM_{2.5}$ concentrations for non-smoking homes with CIU heat, by home type and year

	2014	2024 BAU	2024 Pathway 1	2024 Pathway 2
Ger	98	108	39	8
House	79	86	31	6
Apartment	73	82	29	6

Table 2.7. Estimated average summertime indoor $PM_{2.5}$ concentrations for all non-smoking homes, by home type and year

	2014	2024 BAU	2024 Pathway 1	2024 Pathway 2
Ger	15	17	11	11
House	10	11	7	7
Apartment	11	13	7	7

Exposure

Table 2.8 shows population-weighted average exposure estimates of 59 $\mu g/m^3$ in 2014, and 60 $\mu g/m^3$, 32 $\mu g/m^3$, and 12 $\mu g/m^3$ in 2024 under BAU, Pathway 1, and Pathway 2, respectively, with the greatest exposures consistently affecting ger-dwelling residents. A continuation of current policy trends (BAU) slightly increased population exposures by 2024. In contrast, the modest control measures of Pathway 1 reduced exposures by 45% compared to 2014 levels. The shift to clean technologies in Pathway 2 reduced population exposures by 80%. With the exception of Pathway 2, wintertime exposures in gers and houses dominated city-wide average exposures. Summertime concentrations varied only modestly across BAU and the alternative pathways. Figure 2.5 shows the relative contributions from exposures experienced indoors and outdoors. Exposures incurred indoors accounted for most of the annual averages with a large portion of this resulting from SHS, especially in Pathways 1 and 2. Substantial differences in average annual exposures were seen between home types, with those in houses and gers receiving the highest exposures, and those in apartments on average about 30% lower.

Table 2.8. Estimated population-weighted average exposures by home type, year, and season ($\mu g/m^3$)

	Annual				Winter				Summer			
	2014	2024 BAU	2024 Path 1	2024 Path 2	2014	2024 BAU	2024 Path 1	2024 Path 2	2014	2024 BAU	2024 Path 1	2024 Path 2
Ger	66	68	52	14	113	114	87	13	20	22	16	15
House	70	71	44	12	125	126	76	12	15	17	13	12
Apartment	50	56	25	12	83	93	36	12	17	19	13	13
All Population	**59**	**60**	**32**	**12**	**102**	**102**	**51**	**12**	**17**	**19**	**13**	**13**

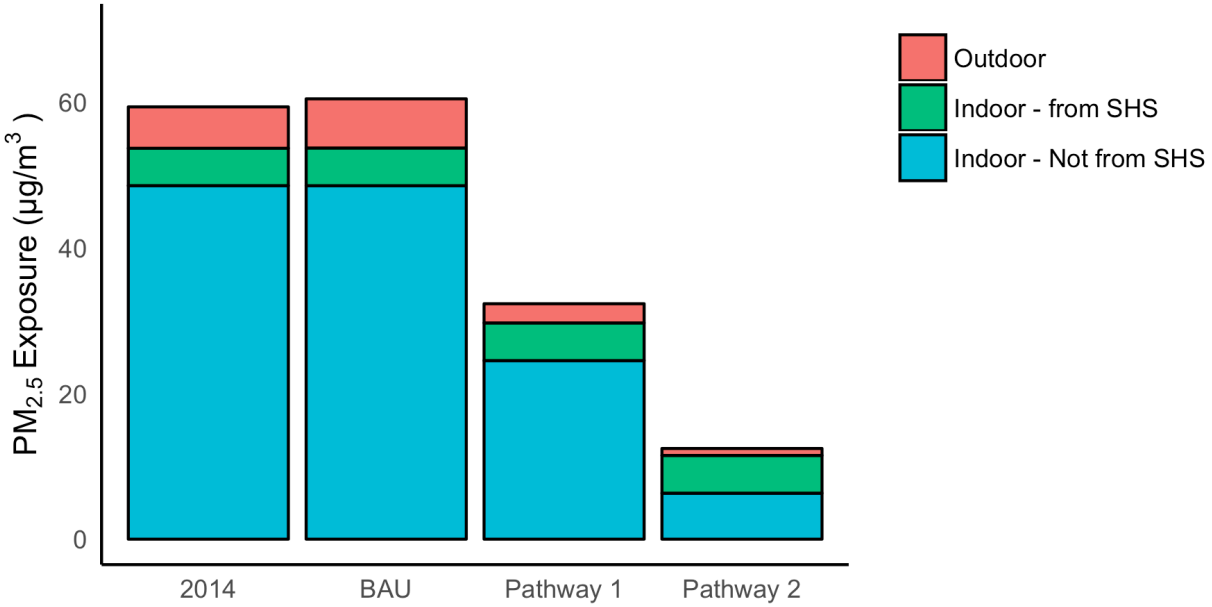

Figure 2.5. Exposures in 2014 and 2024 under BAU and alternative policy pathways, by environment. Indoor exposures are stratified by SHS and non-SHS environments. The difference between indoor and outdoor contribution to total exposure is primarily from the disproportionately high fraction of time spent indoors.

Health impacts

We estimated that 33% (lower: 23%, upper: 42%) of all ALRI deaths in children, 19% (lower: 9%, upper: 28%) of all COPD deaths, 27% (lower: 19%, upper: 42%) of all IHD deaths, 24% (lower: 8%, upper: 34%) of all lung cancer deaths, and 42% (lower: 14%, upper: 54%) of all stroke deaths could be attributed to $PM_{2.5}$ exposure in UB in 2014 – a total of 1,400 attributable deaths (lower: 710, upper: 1,900) and 40,000 attributable DALYs (lower: 22,000, upper: 55,000) (Table 2.9). Deaths and DALYs attributable to $PM_{2.5}$ at baseline were dominated by cardiovascular disease (Figure 2.6). This pattern was consistent throughout

the analysis period. Child ALRI comprised about 25% of attributable morbidity when totaled over the entire 2014-2024 period under each pathway. Over the period 2014-2024, an estimated 18,000 attributable deaths (lower: 9,300, upper: 25,000) and 530,000 attributable DALYs (lower: 290,000, upper: 720,000) were accrued under a BAU pathway. Figure 2.7 shows that these values were reduced by about 20% under Pathway 1, and about 45% under Pathway 2. Nearly all of the deaths averted by Pathways 1 and 2 resulted from IHD and stroke (Figure 2.8). Averted DALY's in both alternative pathways were dominated by child ALRI, IHD, and stroke.

Table 2.9. Excess deaths and DALYs attributable to $PM_{2.5}$, BAU, Pathway 1, & Pathway 2 (per 1000 capita values rounded to two significant digits)

	Accrued, 2014-2024 (per avg. 1000 Capita)	Lower, Upper Bounds	Incurred in 2014 (per 1000 Capita)	Lower, Upper Bounds	Incurred in Final Year of BAU or Pathway, 2024 (per 1000 Capita)	Lower, Upper Bounds
Deaths						
BAU	18,000 (11)	9,300 -25,000			1,800 (0.99)	980 - 2,600
Pathway 1	14,000 (9.0)	7,500 -20,000	1,400 (1.0)	710 - 1,900	1,200 (0.63)	630 - 1,700
Pathway 2	9,800 (6.4)	5,200 -14,000			310 (0.16)	180 -450
DALYs						
BAU	530,000 (330)	290,000 -720,000			55,000 (30)	31,000 - 77,000
Pathway 1	420,000 (260)	230,000 -590,000	40,000 (30)	22,000 - 55,000	34,000 (18)	18,000 - 49,000
Pathway 2	290,000 (190)	160,000 -400,000			8,100 (4.4)	4,400 - 12,000

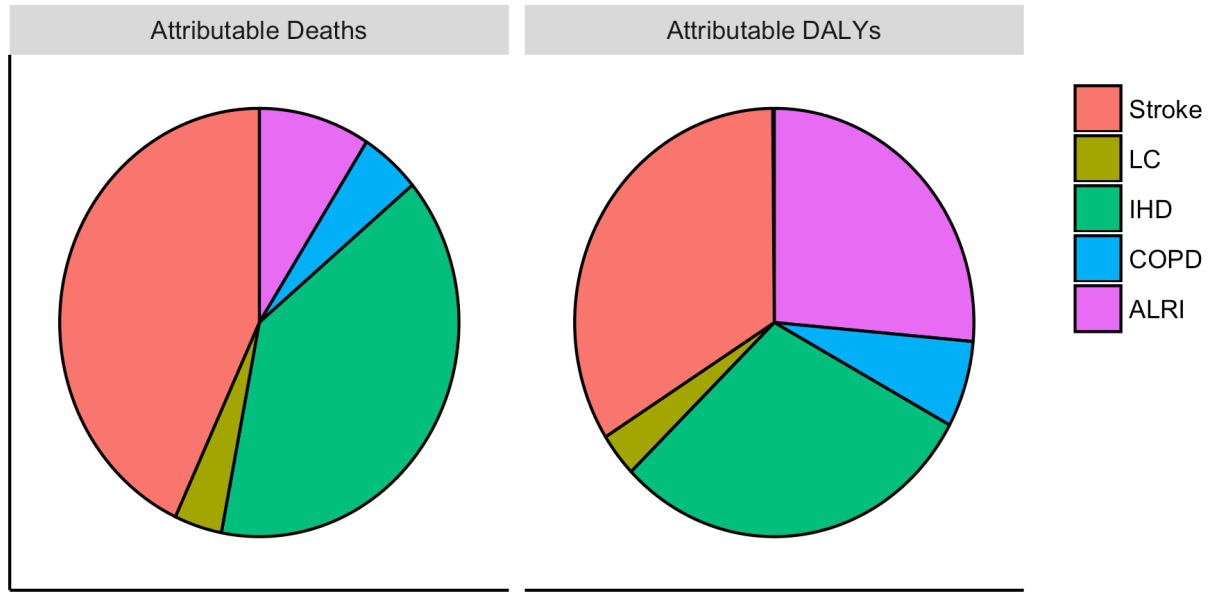

Figure 2.6. Distribution by disease of burden attributable to air pollution in UB at baseline. Note the higher importance for ALRI in the DALY distribution because it affects young children.

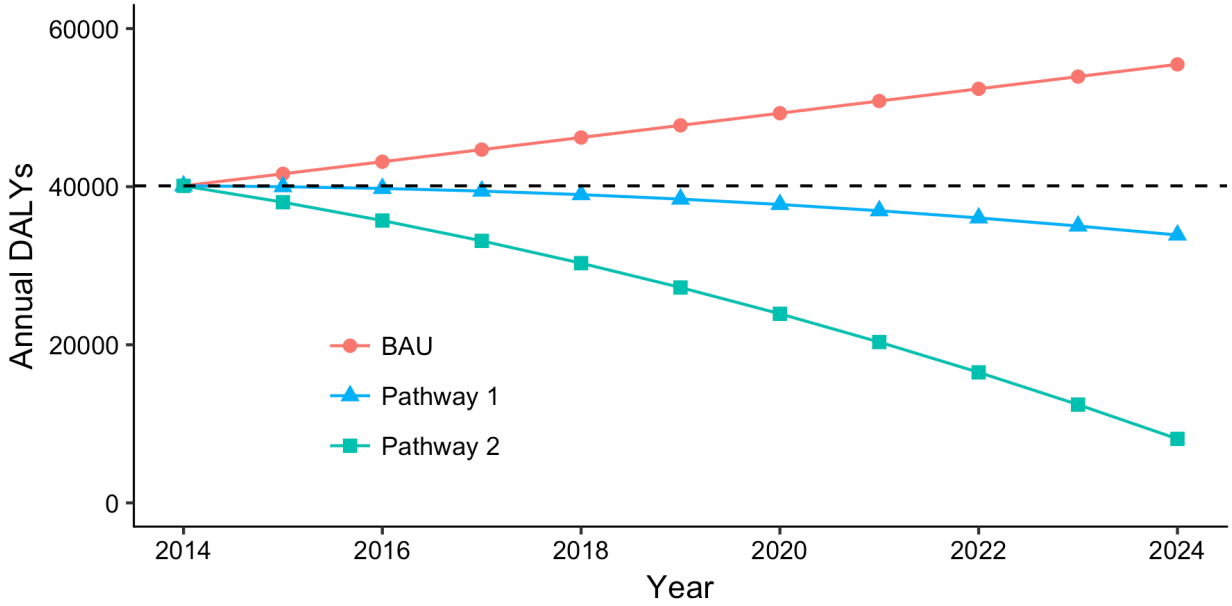

Figure 2.7: Estimated disease burden of $PM_{2.5}$ over the assessment period for BAU and each alternative pathway. Burden of measure is annual DALYs, or DALYs/year. Baseline value (2014) is marked as a dashed line. Pathway 1 averts about 110,000 total DALYs from BAU policies. The stronger reduction measures of Pathway 2 avert about 240,000 total DALYs.

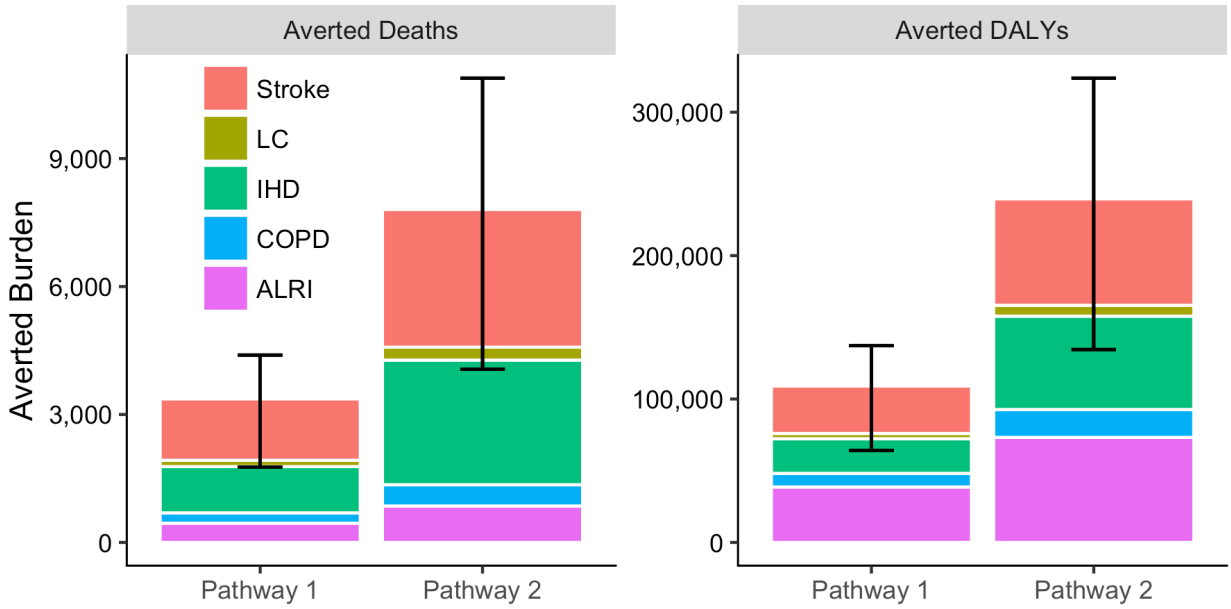

Figure 2.8. Estimated burden averted from BAU by measures taken in Pathways 1 and 2 (2014-2024). Pathway 2 would save more lives than Pathway 1 by more than a factor of 2. Lower and upper bounds on total values are shown as whiskers. Note the greater importance of child ALRI in averted DALYs as compared to averted deaths.

The prolonged reduction period resulted in about 9,800 (lower: 5,200, upper: 14,000) unavoidable deaths and 290,000 (lower: 160,000, upper: 400,000) unavoidable DALYs under the most rigorous reduction pathway between 2014-2024 (Figure 2.8), but annual reductions were substantial. In 2024, an estimated 1,800 (lower: 980, upper: 2,600) deaths and 55,000 (lower: 31,000, upper: 77,000) DALYs were still incurred under BAU. This was reduced by about 35% under Pathway 1 and about 85% under Pathway 2 (Table 2.9). Ulaanbaatar's rapidly increasing population was accounted for by an examination of annual per capita burden (Table 2.9 and Figure 2.9). Changes in annual per-capita burden between baseline and 2024 under all pathways were similar to those estimated for total burden.

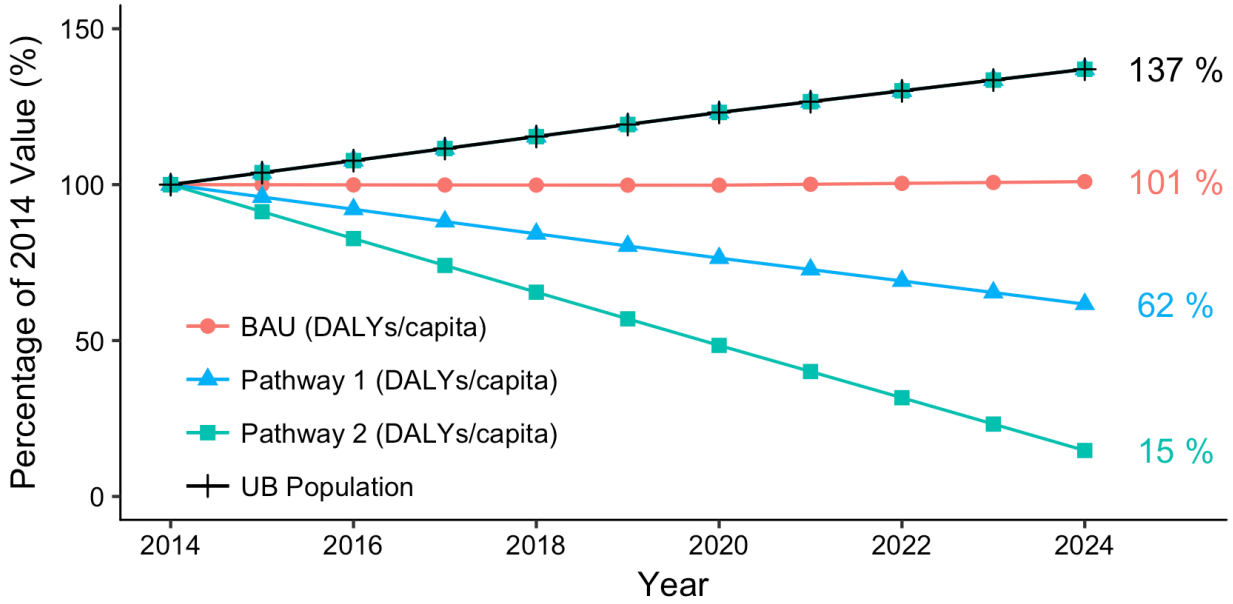

Figure 2.9. Relative projected health impacts per capita for BAU and alternative pathways (2014-2024). Relative projected urban population is also shown. Note that 2014 values are set at 100%. Pathway 2 would reduce impacts to near-counterfactual levels by 2024.

2.4 Discussion

Modeled average annual exposures in Ulaanbaatar (estimated at 59 μg/m^3 in 2014) remained high, despite a wide range of pollution reduction measures recently enacted by the Mongolian government, including ambient air quality standards (United Nations Children's Fund 2016), energy efficiency programs (Greene et al. 2014a), anti-smoking laws (World Health Organization Western Pacific Region 2012), and improved stove subsidies (Greene et al. 2014a). Our models indicated that this trend was driven primarily by high wintertime indoor concentrations influenced substantially by the infiltration of outdoor pollution. In gers and houses heating with LPB (houses only), MCA stoves, and stoves using semi-coke coal, modeled wintertime indoor concentrations were more than ten times higher than the WHO $PM_{2.5}$ annual Air Quality Guideline (10 μg/m^3) (World Health Organization 2014), a recommendation considered necessary to be fully health protective. For regions with very high concentrations, a more reasonable context for indoor concentration comparisons may be the WHO Air Quality Guideline interim targets, which are designed as achievable incremental policy goals. Even still, modeled indoor concentrations in gers and houses heating with LPB (houses only), MCA stoves, and stoves using semi-coke coal are more than three times higher than the highest interim target (35 μg/m3).

Our estimates indicated that 33% (lower: 23%, upper: 42%) of all deaths from ALRI in children and 19% (lower: 9%, upper: 28%), 27% (lower: 19%, upper: 42%), 24% (lower: 8%,

upper: 34%), and 42% (lower: 14%, upper: 54%) of all deaths from COPD, IHD, lung cancer, and stroke, respectively, could be attributed to $PM_{2.5}$ exposures in 2014. $PM_{2.5}$ related mortality at baseline and as accrued under all pathways was driven by cardiovascular disease, while attributable morbidity was more evenly distributed between IHD, Stroke, and ALRI in children (< 5 years). These estimates and trends were consistent with global and national estimates from the GBD (Forouzanfar et al. 2015), and identified children as a local population of particular vulnerability. They were also within the range of citywide findings from Allen et al. (Allen et al. 2013), which were based on earlier exposure-response functions, different disease groupings, a lower counterfactual of 7.5 $\mu g/m^3$, and outdoor concentrations alone.

A business as usual approach to energy policies in UB will have little impact on citywide $PM_{2.5}$ exposures by 2024 yet may result in a substantial increase in total health burden because of large increases in projected urban population. A package of policies targeting reductions in both indoor and outdoor emissions from household coal stoves alongside aggressive improvements in power and traffic sectors could reduce annual average population-weighted $PM_{2.5}$ exposures by nearly 80% and annual per-capita attributable health burden by about 85% by 2024. A package of more moderate emissions control policies, including cleaner-burning coal stoves and modest improvements to the city's power plants and vehicle fleets, may reduce $PM_{2.5}$ exposures by 45% but would have less effect on health burden due in part to the non-linearity of the relationship between $PM_{2.5}$ exposure and risk for many diseases (Burnett et al. 2014). When energy-related emissions are ultimately reduced, environmental tobacco will play an important role in local disease burden if not aggressively targeted by regulators.

Our investigation builds upon a small but growing body of air quality research in Mongolia (Allen et al. 2013; Amarsaikhan et al. 2014; Davy et al. 2011; Enkhbat et al. 2016; Guttikunda et al. 2013; Nakao et al. 2016; The World Bank 2011; Ulzii et al. 2015; Warburton et al. 2013) and is the first to both examine and predict population-wide $PM_{2.5}$ exposures in UB that are integrated across environments and account for contributions from SHS. We are aware of two studies that have directly measured personal $PM_{2.5}$ exposures in individuals in UB, both of which are difficult to interpret within the context of our analysis and so were not explored in depth. One of these studies (Ulzii et al. 2015) reported primarily on peak concentrations and is thus difficult to compare with our estimates of longer term averages. The other study (Nakao et al. 2016) examined personal $PM_{2.5}$ exposures between 9:00 and 17:00 in hospital patients in March and July. However, that study reported only on "indoor" and "outdoor" personal exposure levels, the definitions of which are unclear, did not disclose a sample size, and did not distinguish exposure statistics by relevant demographics like heating type or smoking status.

Most inferences about the population health impacts of $PM_{2.5}$ in UB and greater Mongolia have relied on outdoor concentrations modeled from emissions, chemical transport estimates, and/or measurements taken from a small number of outdoor, fixed-site

monitors (Allen et al. 2013; Davy et al. 2011; Guttikunda et al. 2013) or outdoor concentrations measured using hybrid satellite techniques (Brauer et al. 2016). But satellite estimates have been found particularly unreliable in UB because of poor resolution in winter and at night, when concentrations are highest in the region (Jerrett et al. 2005b; Pinto et al. 2004), and evidence suggests that intra-urban outdoor air quality is highly heterogeneous and that city-wide outdoor estimates rooted in data from a small number of sites may lead to considerable error when applied to health impact evaluation (Alexeeff et al. 2015). Our approach included outdoor modeling techniques, but tempered issues of exposure misclassification by incorporating indoor concentrations weighted by locally relevant and age-specific time activity estimates.

A 2011 study by the World Bank (The World Bank 2011) projected air-pollution related health effects in UB using only outdoor air pollution as an indicator of exposure. The current analysis improves upon the methods used in that study by implementing more-advanced exposure estimation methods, a newer exposure-response technique, more-nuanced population and background disease projections, local household stove emissions data, detailed indoor air pollution measurement data, a baseline scenario inclusive of recent MCA efforts to distribute improved coal heating stoves, and an analysis inclusive of all UB districts. Our analysis also benefits from the emissions and survey data collected during the 2012-2013 MCA stove implementation project (Greene et al. 2014b, 2014a, 2014c). Nevertheless, the general trends identified in the World Bank report are in agreement with ours: heavy reductions in $PM_{2.5}$ emissions, particularly those from household stoves, are needed to make appreciable impacts on the current pollution-related health burden in UB. In contrast, the high attributable disease rates remaining in our baseline, which included 100% replacement of traditional raw coal stoves with improved MCA stoves, and Pathway 1, which assumed even greater stove improvements and a switch by a large fraction of homes to CIU heating, suggests World Bank estimates that improved stoves can produce exposure reductions commensurate with a full transition to electric heating are unrealistic.

Although few measurements exist with which to compare our values, a study (Enkhbat et al. 2016) of indoor concentrations during the 2015 winter season in UB found geometric mean concentrations in apartments (52.8 $\mu g/m^3$, 95% CI: 39-297 $\mu g/m^3$) that were generally in agreement with our 2014 non-smoking clean apartment estimate (73 $\mu g/m^3$). We expect our indoor estimates of homes heated with stoves to be more robust, as they were derived from the largest database of measured indoor concentrations of homes and gers in Ulaanbaatar, to date. Recently published measurements of indoor concentrations in gers using traditional stoves during the 2015 winter season (127.8 $\mu g/m^3$, 95%CI: 86-190 $\mu g/m^3$) are consistent with the value produced by our model for gers using traditional stoves (113.3 $\mu g/m^3$) (Enkhbat et al. 2016). We are unaware of published measurements in cleanly heated houses for comparison.

There are several limitations to our study. Concerning the outdoor air quality modeling, Gaussian dispersion models are overly simplistic to capture all of the transport, dispersion,

and terrain characteristics for UB wintertime conditions. Air quality modeling errors from the use of a Gaussian dispersion model are lumped together with emission inventory errors when calibrating the model to air quality observations. It is not clear how these errors propagate through to 2024. Un-modeled emission sources were assumed to have an impact of 10 $\mu g/m^3$ and this simplification influences the exposure estimates, especially for Pathway 2 where modeled emission source contributions are low.

Indoor concentrations in 'clean' heating homes, which were calculated by applying infiltration factors to outdoor concentrations, may have propagated any error incurred by the outdoor model. These methods, which employed climate-relevant but non-local infiltration rates, could further be improved by future work characterizing local building infiltration rates. In addition, the linear model used to estimate wintertime indoor concentrations in stove-heated homes, as described in Appendix B, proved a poor predictor of individual indoor concentrations with particular weakness predicting very high and low values (adj-r^2 = 0.13). However, evidence suggests that stove type, SHS, and various structural characteristics associated with home type play important roles in shaping indoor concentration profiles in households using solid fuels (Balakrishnan et al. 2013; Chowdhury et al. 2013; Li et al. 2016). In order to place measurement-based constraints on these relationships at the population level, the full model was kept, despite the low r^2. More complex cross-validated modeling was attempted with these covariates using the SuperLearner machine learning package in R (van der Laan et al. 2007), but did not improve the fit. Moreover, the assumption that residential indoor $PM_{2.5}$ concentrations reflect indoor concentrations in general is overly simplistic. Future research should elucidate $PM_{2.5}$ concentrations throughout workplace, recreational, and other indoor environments in order to inform a more-nuanced population-wide indoor exposure model. This is especially important if a portion of the population works in high-exposure settings like coal mines.

The use of linear models to project background disease rates through 2024 may not reflect future trends in areas like rural to urban migration, economic development, regulatory shifts, and healthcare improvements which may have non-linear impacts on disease-influencing factors (e.g. introduction of pneumococcal conjugate vaccines). This general limitation is highlighted by the weak fit of the linear background disease model to historical data for several diseases, as demonstrated in Appendix B. Background disease projections may also have been substantially underestimated because of latent disease at the time of estimation. Because air pollution is a relatively new problem in UB (only gaining traction in the 1990s) and because of the rapid, recent influx of people from cleaner rural parts of Mongolia, the full impact of diseases that require several decades after exposure to develop, like lung cancer, were likely not fully represented in the background disease data used in our projections, biasing our burden estimates downward (Cameron and Ostro 2004; National Institutes of Health 2007). On the other hand, our decision to not consider exposure cessation lag in calculation of averted deaths and DALYs may have accelerated the accrual of estimated health benefits and resulted in modestly inflated estimates for the

time periods given, except for ALRI which is an acute outcome. The effect of cessation lag on the final results may have been attenuated by the large fraction of burden caused by ALRI, especially DALYs.

Our BAU and alternative policy pathways were limited by gaps in the literature, too. For example, while recent government anti-smoking campaigns (World Health Organization Western Pacific Region 2012) suggest that smoking rates are likely to change, it is unclear how this will impact household SHS prevalence and related indoor concentrations. Implementation of a steady rate allowed the importance of addressing this issue to be clearly identified in the models. Information on the number of building structures and, thus, heating sources per home was also inconsistent. City census data seem to be reported in terms of primary residence type (Statistics Department of Ulaanbaatar 2013), while satellite imagery and field experience indicated that, in peri-urban areas, a considerable number of residences include multiple ger or house structures. This may have contributed to uncertainty in our exposure estimates which were, in part, based on the assumption that the indoor environment of each resident was dictated by a single combination of structure, heat source, and SHS presence.

Attributable burden estimates in UB may, in general, be underestimated. Evidence suggests that cold-air exposures may increase sensitivity to risk factors for cardiovascular diseases (Luo et al. 2014). Ulaanbaatar's temperatures are typically much colder than the regions that inform Burnett et al.'s exposure-response models (Burnett et al. 2014), suggesting a risk misclassification biasing our burden estimates downward. Differential bias in the calculation of averted burden in our projections may have resulted from winter exposures dominating the annual average most in BAU, less in Pathway 1, and least in Pathway 2, thus leading to further underestimation of averted burden. It is also not clear at present whether pneumonia incidence is related to winter time exposures, or annual average exposures, although hospital records of pneumonia incidence indicate the majority of the burden is in winter months. The distinction becomes important for non-linear dose response curves where the wintertime exposures are on a much flatter section of the curve.

2.5 Conclusion

The analyses performed in this chapter demonstrate that $PM_{2.5}$ exposures in Ulaanbaatar has reached a critical level, and immediate measures must be taken to reduce its health impacts on the city's growing population. Current exposures are projected to produce unprecedented levels of respiratory illness, especially in children, and cardiovascular disease. Using some of the latest available exposure-response techniques and novel data on local emissions and indoor concentrations, this analysis is the most holistic view of population-wide air quality exposures in Ulaanbaatar to date. The results highlight the need for aggressive actions, including the elimination of residential coal burning and the reduction of current smoking rates, if the health burden of air pollution is to be reduced. Our conclusions support recent findings that $PM_{2.5}$ emissions, especially from household

heating, contribute substantially to mortality and morbidity from cardiovascular and pulmonary disease in the city. In addition, without efforts to moderate indoor concentrations, the full benefits of pollution reductions in UB will not be realized.

Chapter 3

Machine-learned modeling of $PM_{2.5}$ exposures from household air pollution, ambient concentrations, and comprehensive surveys in homes cooking with solid fuels in rural Lao PDR

3.1 Introduction

Approximately 2.8 billion people meet most of their cooking needs with solid fuels like wood or coal (Bonjour et al. 2013). Chapter 1 points out that the resulting $PM_{2.5}$ emissions, called household air pollution (HAP), produce an estimated four million deaths annually from cardiovascular and respiratory diseases (Lim et al. 2012). Burden assessments like this one are calculated similarly to those produced in Chapter 2: by applying estimates of population annual average exposures to air pollution risk-response relationships elucidated in epidemiological studies. Unfortunately, compared to the extent of the problem, relatively few exposure data exist for HAP globally. Endeavors to estimate HAP-related disease burden thereby commonly employ fuel-type indicators or modeled or measured estimates of 24-hour average indoor $PM_{2.5}$ concentrations as proxies of total exposure (Balakrishnan et al. 2013; Smith et al. 2014).

Personal exposures are, however, more than a function of fuel type, and even the most sophisticated indoor concentration modeling procedures are incapable of assessing the human-environment interactions that affect actual exposures (Steinle et al. 2013). In the specific context of HAP, exposures can be a factor of heterogeneity of pollution within the home, including vertical stratification (Johnson et al. 2011; Kandpal et al. 1995) and changes in dilution relative to distance from the stove or windows; the amount of time spent outside of the household; and $PM_{2.5}$ sources unaccounted for in HAP measurements or indicators like traffic, pollution at a neighbor's home during a visit, or high outdoor ambient values. Errors in exposure estimation caused by relying on such proxies may lead to significant bias in the estimation of related health burdens, especially at points along the exposure-response relationship where risk may be highly non-linear, i.e. be changing rapidly with exposure (Smith et al. 2014).

The general paucity of HAP exposure data stems from the difficulty, invasiveness, and resource-intensive nature of current exposure measurement techniques. Gold standard methods for measuring exposure require outfitting individuals with delicate consumable media and expensive pumps that are bulky, heavy, and noisy. Recently developed equipment like the MicroPEM (RTI International, Research Triangle Park, North Carolina, USA) and AirChek pump series (SKC Inc., Eighty Four, Pennsylvania, USA) have, in part, addressed concerns of participant discomfort with lighter and quieter pumps, but such devices are often expensive and produce delicate samples not well-suited to the rugged environments that characterize solid fuel using regions and transport to and from those regions, which can be quite remote. [7]

The need for a better means of accurately assigning air pollution exposures in the solid fuel cooking context is now more important than ever. A market is coalescing around the sale of Disability-Adjusted Life Years (DALYs) averted (aDALYs) from HAP intervention projects much like the market that has in recent years formed around the offset of greenhouse gases (Anenberg et al. 2017; Smith et al. 2015). To be effective, such a market must have at its disposal a tool for estimating personal exposures that is both accurate and practical. Evidence suggests that variability in both HAP concentrations and personal exposures may be explained in part by household and behavioral indicators like fuel type, stove type, kitchen structure, and cooking duration (Balakrishnan et al. 2002, 2013). It stands to reason that the relationship *between* indoor HAP concentrations and personal exposures may also be partially explained by such indicators. To date, this hypothesis has not been thoroughly explored.

Chapter 2 improved upon the estimation of annual average $PM_{2.5}$ exposures in a specific population through the applications of modeled HAP concentrations, outdoor ambient models, and time activity estimates. This chapter, Chapter 3, pivots to a broader investigation of the utility of a variety of indoor HAP data and easier-to-assess household and behavioral indicators to predict 24-hour average $PM_{2.5}$ exposures – an oft used estimate of annual average exposures – in a more traditional rural HAP context. Machine learning – a type of artificial intelligence – and super learning – the production of a single "super" learner by combining a set of candidate learners, like random forests or neural networks, as weighted by their predictive utility (van der Laan et al. 2007) – are applied. Particular attention is paid to survey indicators similar to or drawn from the Demographic and Health Survey (DHS). The DHS is an internationally administered survey designed to collect

[7] Technologies in this field are rapidly advancing. A number of smaller, lighter, and quieter gold standard devices have been produced, but not yet thoroughly field tested. For example, the Enhanced Children's Monitor (Chartier 2015) – which operates on the same principles as the MicroPEM but reduces the size by half and overall weight by nearly two thirds – and the Ultrasonic Personal Aerosol Sampler (Volckens et al. 2017) – which developers claim operates in virtual silence and can be produced from approximately $150 worth of components – are now being validated for use in a major multi-center randomized control trial. While making great improvements to participant comfort and, in some cases, total expense, such devices may still suffer from the issues of sample durability identified above.

accurate and representative data on demographics and health around the world (ICF International 2011). DHS questionnaires are structured, standardized, and frequently administered in over 90 countries. Insight into the power of various DHS indicators to predict personal exposures in this context may prove useful for future disease burden research, aDALY measurement schemes, and, in general, cookstove monitoring and evaluation programs.

Data collection was carried out during a World Bank stove intervention program in the People's Democratic Republic of Lao (Lao PDR), where about 95% of households cook with solid fuels on traditional appliances (Bonjour et al. 2013). This work also provides one of the first estimates of exposures among rural Lao women cooking with solid fuel, and adds to the small number of studies on HAP in the region (Huang et al. 2013; Mengersen et al. 2011; Morawska et al. 2011), where HAP is considered the number three cause of ill-health (Institute for Health Metrics & Evaluation 2017) and the cause of an estimated 12% of all deaths and 7% of DALYs annually.

3.2 Methods

The data analyzed in this chapter were collected during a cookstove evaluation program in the rural Xonboury District of the Savannakhet Province of Lao People's Democratic Republic (PDR). Briefly, the larger program sought to evaluate the displacement of traditional tripod and bucket type cookstoves with the African Clean Energy (ACE)-1 battery-powered[8] blower stove in three villages: Houaymouange, Vangkhonkham, and Aho. These villages will hereafter be identified as villages A, B, and C, in no particular order, to reduce the possibility of identifying participants. Twenty-four (24) households from each village were enrolled for a total of 72 households, a size chosen to detect a minimum 40% difference between pre- and post-intervention kitchen air pollution (KAP) concentrations – a more specific variant of HAP – assuming a paired coefficient of variation of approximately 0.80 (Edwards et al. 2007). Selection criteria included wood as the primary household cooking fuel and that the main cook be 18 years of age or older and not pregnant. Recruitment heavily preferred that the main cook not smoke, but this was not a hard selection criterion due to limited availability of eligible households within the study area. While not a hard-and-fast selection criterion, homes without kerosene lighting were also given preference.[9] Each household was encouraged to use their new stove during the post-intervention study period, but no criticism or sanction was imposed if they did not comply.

Main cooks from 12 of the 24 households in each village recruited into the larger intervention study were enrolled in personal exposure measurement activities (36 total) on a convenience basis. This number was chosen to optimize available resources. A more-rigorous preference for non-smoking than the larger intervention program was afforded by

[8] The ACE-1 comes with a small solar panel to recharge the internal battery of the stove.
[9] At baseline, no study households reported kerosene lighting.

the smaller sample size; no smokers were enrolled in personal exposure measurement activities. This chapter reports on the 36 households and participants enrolled in personal exposure measurement activities, hereafter simply called "study households" or "participants," and on the HAP and personal exposure activities conducted as part of the academic work of this thesis, hereafter identified as the "study," "work," etc. Concurrent environmental information like meteorology and ambient $PM_{2.5}$ concentrations are also considered.

The sampling scheme generally took the form of a before-and-after study, relative to the cookstove intervention. Each study household participated in a baseline survey, a round of air pollution monitoring and follow-up surveys prior to receiving the ACE-1 stove (hereafter referred to as the "Before" period), and a round of air pollution monitoring and follow-up surveys several weeks after receiving the ACE-1 stove ("After"). This time between stove dissemination and measurement allowed cooks to gain some familiarity with the new stove. Dissemination occurred over the course of a single week in January 2015. Distributors made follow-up visits to villages and households to address technical problems and usage queries through the following 2 weeks.

Each round of air pollution sampling occurred in 3-week segments, with one week in each village. The first round spanned December 2014 to January 2015, and the second round, January to February 2015. Both sampling periods generally entailed 4 days of continuous KAP monitoring (not included in this analysis); a post-monitoring survey to provide energy-use and exposure related information during the 4-day round of KAP measurement (not included in this analysis); 2 days of simultaneous gravimetric measurement of KAP and personal $PM_{2.5}$ exposure nested within that 4-day continuous KAP monitoring period; a separate post-monitoring survey to provide data specific to each 2-day set of nested gravimetric measurements; and ambient air pollution and meteorology measurements at a central cite in each village. This detailed sampling scheme is summarized in Figure 3.1. Stove use monitoring system (SUMS) data, accelerometer data from personal exposure setups, and continuous personal $PM_{2.5}$ concentration measurements were also collected, but are not discussed in this chapter. Baseline questionnaire, post-monitoring questionnaire, stove use monitoring, and KAP and exposure sampling protocols used to collect the data analyzed in this chapter were approved by the University of California, Berkeley Committee for Protection of Human Subjects, protocol number 2014-06-6457.

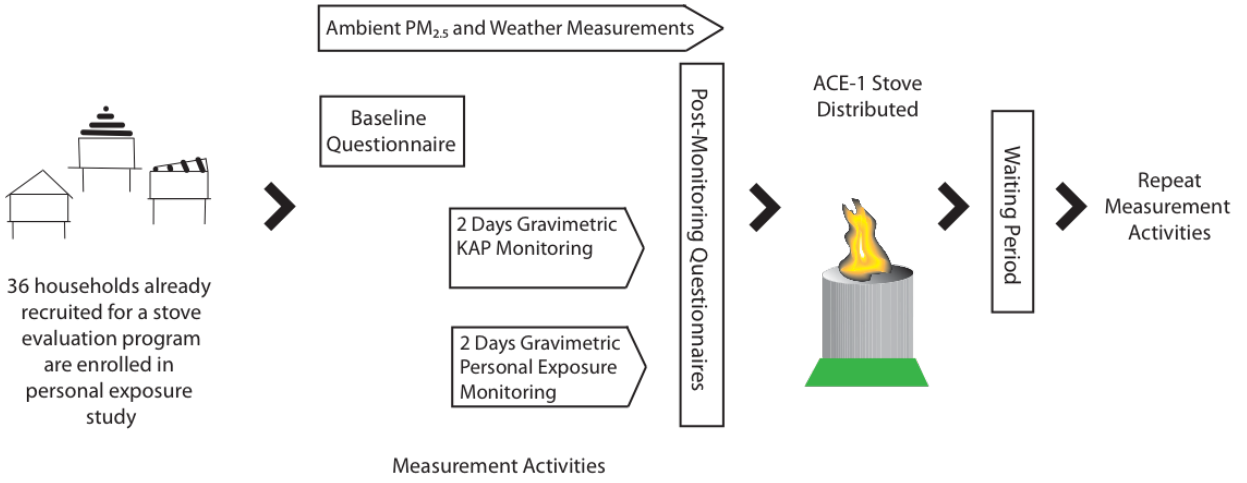

Figure 3.1. Key aspects of the sampling scheme.

Air pollution measurements

Gravimetric air pollution measurements of KAP and personal $PM_{2.5}$ exposures were taken both before and after stove dissemination. Sampling was performed using Triplex cyclones (Mesa Labs, Butler, NJ, USA) with 2.5µm size cut at 1.5 LPM, AirCheck XR5000 pumps (SKC, Inc., Eighty Four, PA, USA) set to approximately 1.5 LPM. Cyclones were fit with 37mm Teflon filters with support ring and 2µm pores (Pall Corporation, Port Washington, NY, USA). Gravimetric KAP and personal $PM_{2.5}$ exposure samples were collected for two 24-hour periods (48 hours total). Completed sampling periods were consecutive in all but one household where the two 24-hour samples were separated by 2 days due to issues with the ACE-1 stove that were discovered after the first day of After sampling.

KAP monitors were placed on the kitchen wall approximately 1.5m from the floor and 1m from the edge of the main cooking stove, a standardized location meant to represent the approximate breathing zone of a woman standing near the stove. Where possible, monitors were placed 1m away from major ventilation sources such as windows, eaves, and doors. Field teams installed push-pins during the Before sampling period to assist in the duplication of monitor placement in the After sampling period, though sometimes push pins fell out or were moved by the household. It should also be noted that the distances between monitoring equipment and the stove and between monitoring equipment and sources of ventilation were measured from the continuous KAP monitors previously mentioned (but not used in this analysis). Gravimetric sample inlets were placed directly adjacent to the continuous monitor, which may have produced a several centimeter difference in true distance from the primary stove and true minimum distance from ventilation features between households. It is also important to note that while monitor placement was intended at 1m from the edge of the primary cooking stove, both the traditional and improved stoves are relatively portable and may have been moved during

the sample period. A small number of duplicate KAP filter measurements were taken for quality assurance and quality control. A significant discrepancy was not observed between duplicate KAP filters.

Measurements of personal exposure and gravimetric KAP concentrations were taken concurrently. A custom vest garment was designed to hold personal exposure monitoring equipment with the gravimetric sampling inlet approximately in the breathing zone while limiting discomfort (Figure 3.2). Participants were instructed to place the garment next to their bed at night and in a nearby location while bathing.

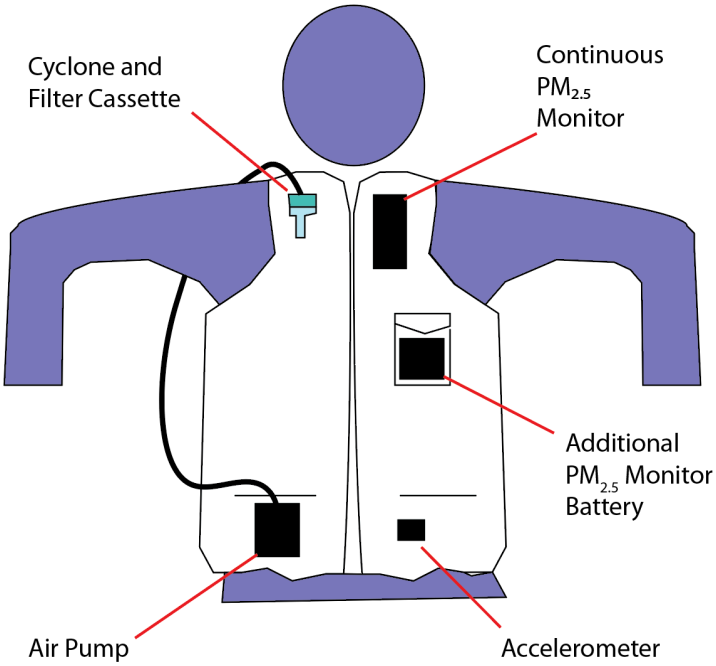

Figure 3.2. An example of the sampling garments used.

The default method of calculating sample runtime for gravimetric samples was to subtract sample start times from sample end times as marked by field technicians on sample forms. As a redundancy check, runtimes were calculated from the pumps' internal timers. These timers tend to overestimate actual run time, because they can also include the several minutes that the field team operates the pump (without considerable filter contamination) while adjusting and measuring its flow rate. These measures were compared against each other and any major outliers (> 15 minutes of difference) were examined and runtimes manually discerned – field notes or other oddities sometimes also compelled manual analysis. Flow rates through gravimetric sampling trains were measured at the start and end of each gravimetric sample. Air flow through each filter during sampling was calculated by averaging pre- and post-sample pump flow rates and multiplying by sample runtime.

Standardized sampling forms were used throughout KAP and personal exposure field work to minimize errors during the process of linking data media (e.g. gravimetric filters) with corresponding household identifications (ID), sample periods, and notes. To reduce transcription errors, the original paper-based KAP and exposure sampling forms used in the field were entered into an electronic database two separate times. Discrepancies between duplicate entries were reconciled with further review of the sampling forms.

Ambient 24-hour $PM_{2.5}$ measurements were collected using a MiniVol $PM_{2.5}$ Sampler (Air Metrics, Springfield, OR, USA). Forty-seven (47) mm Teflon filters with built-in support ring and 2µm pores (Pall Corporation, Port Washington, NY, USA) were used for sampling with a flow rate of approximately 5 L/min. The MiniVol was placed in a central location in each village where it would be safe and would not be disturbed, such as on the roof of a home, in a tree, or on a pole. Placement rotated with KAP and exposure sampling. For example, while KAP sampling occurred in Village A, the MiniVol was placed at a central site in Village A.

Measurements of outdoor temperature (°C), relative humidity (%), barometric pressure (mb), wind speed (km/hour), wind direction (45° increments), and precipitation (mm) were planned for Monday through Saturday of every sampling week using a Vantage Vue 6250 wireless weather station with Integrated Sensor Suite 6357 and a 6351 console (Davis Instruments, Hayward, CA, USA) co-located with the MiniVol in each village. Meteorology readings were observed and manually entered into a field form once in the mid-morning ("morning") and again in the later morning or early afternoon ("afternoon").

Mass deposition on gravimetric samples was determined using the weights of filters before and after sampling. Field blanks were used to correct for filter contamination or mass loss not specific to sampling-based loading processes. Both 37mm and 47mm field blanks were collected. Notably, 37mm field blanks from the Before and After periods were treated slightly differently[10], though a significant difference was not seen between the mean Before lab blank mass deposition and the mean After lab blank deposition. Weighing and mass calculations were performed by colleagues at Berkeley Air Monitoring Group (Berkeley, CA, USA). During the study, an equipment malfunction in the weighing facility was discovered that affected and invalidated many pre-sample filter weight measurements. Using over 250 sample filters unaffected by the malfunction, Berkeley Air Monitoring Group was able to develop a method of recovering pre-sample filter weights (Garland et al. *in preparation*). Briefly, post-sample filters were sonicated in a solvent bath to remove PM mass loaded onto the filter during sampling. Filters were allowed to dry, and then reweighed. The filter's pre-

[10] In general, 37 mm field blanks were transported to the field, prepared in a KAP cyclone setup and placed on the wall as normal sample filters, allowed to sit inactive for some period of time, and then taken down, capped, transported, and stored as a normal sample filter. The majority of 37mm blanks from the Before period remained in the cyclone on the wall for a full ~ 24-hour sampling period before being taken down, while a small number of Before and all of the After 37mm blanks remained in the cyclone for only about 5 minutes before being removed, capped, placed in a secure location for the remainder of the ~ 24-hour sampling period, and then transported.

sample weight was taken as this post-sonication mass. Filter tearing during the sonication process did not significantly affect pre-sample weight estimates, and so filters torn during sonication were not excluded. Final mass deposition from sampling was calculated as the difference between the post-sample (pre-sonication) filter mass and post-sonication filter mass, as adjusted for the average post-sonication field blank mass from each sampling period. The method has been shown to recover pre-sample filter weights with high reliability ($r^2 > 0.99$), and is currently in preparation for peer review. For consistency, all KAP and personal gravimetric samples were treated with this method. Filters from ambient PM measurement were not affected by this equipment malfunction, and thus not treated for sonication. Ambient measurement filters were corrected for three 47mm field blanks collected in the Before sampling period and two 47mm field blanks collected in the After sampling period.

Questionnaires

Three separate questionnaires were administered during the personal exposure study: a baseline questionnaire, a post-KAP monitoring questionnaire, and a post-personal exposure monitoring questionnaire. Questionnaires were drafted in English and translated into Lao with input from colleagues with experience in the field, rural Lao context, or both.[11] An English version of each questionnaire is available online (Hill LD et al. 2015). Questionnaires were administered in Lao by trained surveyors familiar with local customs using Microsoft Excel-based survey tools and portable computers. Administration was recorded (audio). Questionnaires were reviewed for missing and suspect answers and manually adjusted during two separate rounds of quality assurance (QA) —one shortly after sampling concluded in 2015, and another in late 2016 and early 2017 prior to publication of this chapter. Audio recordings were consulted when possible; recordings were available for the 2015 QA period but, by the 2016/2017 period, recordings of all post-monitoring questionnaires and five of the baseline questionnaires had been destroyed.

The baseline questionnaire was administered to households in December 2014, and was designed to gain information on home characteristics, demographics, cooking behavior, time activity, locally relevant exposure-related activities, and energy use at various times of the year and with a primary focus on the season during which the questionnaire was administered. The baseline questionnaire also included a series of questions as a double-check for compliance with exclusion criteria. Most baseline questionnaires were administered in the first week of December, not long after administration of consent.

[11] Multiple questions— especially on the topics of demographics, fuel use, kitchen location, and smoking— were drafted from or inspired by the Lao Social Indicator Survey (Ministry of Health and Lao Statistics Bureau of the Ministry of Planning and Investment 2012), the general Demographic and Health Surveys (ICF International 2011), and the National Health and Nutrition Examination Survey (Centers for Disease Control and Prevention 2013). Content was also drawn from the pool of survey tools developed over the years by project collaborators and affiliates of the Household Energy, Climate, and Health research group at the University of California, Berkeley.

About 10% of baseline questionnaires were administered the fourth week of December, either during or just before the first household air pollution sample.

Post-monitoring questionnaires were administered during equipment takedown in each sample period. Post-monitoring questionnaires included questions similar to those on the baseline questionnaire regarding cooking behavior, time activity, energy use, and locally relevant exposure-related behavior. The primary difference between the post-monitoring and baseline questionnaires was the time period of interest; while the baseline questionnaire asked about various times of the year, the post-KAP monitoring questionnaire asked specifically about the several days of continuous KAP measurement, and the post-personal exposure monitoring questionnaire asked only about the 2 days of simultaneous gravimetric exposure and KAP measurement. Questions posed in the post-KAP and post-personal exposure monitoring questionnaires were nearly identical.

Only a subset of response variables from the post-personal monitoring questionnaire and the baseline questionnaire was used in analysis. Variables from all post-monitoring questions were included, because they pertain directly to the air pollution measurement periods of interest. As the baseline and post-monitoring questionnaires overlap heavily, baseline questionnaire variables considered for analysis were limited to topics covered exclusively in the baseline questionnaire. Baseline questionnaire topics considered for analysis included household characteristics and demographics, architectural characteristics of each respondent's primary house and kitchen, respondent age and smoking status, attitudes toward cooking smoke and its impacts on health, wood fuel preparation, electricity access, and electric appliance use. Baseline questionnaire variables were combined by household with both the Before and After post-monitoring datasets.

A primary goal of the analysis was to understand the utility of DHS-type survey questions in the prediction of personal $PM_{2.5}$ exposure concentrations in the wood-fuel cooking context. In Lao PDR, DHS questions are administered in the form of the Lao Social Indicator Survey (LSIS) (Ministry of Health and Lao Statistics Bureau of the Ministry of Planning and Investment 2012). LSIS questions for which reasonable overlap with the baseline or post-monitoring questionnaires existed or could be produced during data processing covered the topics of household size by age, education level of various household members, drinking water source, sanitation, ethnic identity, architectural characteristics of the main house, primary fuel type and cooking location, household asset and financial status, and land and livestock ownership.[12]

Several existing variables were modified and new variables created by combining individual responses or questions. Of particular note is a cooking exposure activity score created from post-monitoring questionnaire responses about cooking behavior. This score was produced

[12] LSIS questions HL6, HH11, ED3, ED4, WS1, WS8, WS9, HC1C, HC3, HC4, HC5, HC6, HC7, HC8, HC9, HC10, HC12, HC14, and HC15.

separately for cooking performed at home in the morning, afternoon, and evening. A total score was created by combining the sum of home-specific scores with a score calculated for all cooking performed away from home. This score was calculated by multiplying the amount of self-reported time spent cooking, parsed in minutes; by the amount of that cooking time spent in the kitchen as reported on a scale of 1 (very little) to 5 (nearly all of it); by the amount of that time cooking in a kitchen spent within arm's length of the stove, reported on the same 1 to 5 scale. A set of exposure scores was also created for grilling at home and total grilling (at home and away).

Final dataset

The final dataset was selected to optimize the external validity, or transportability, of this analysis to balance predictive power with a broader interpretation of model results. The analysis set included all responses from the post-personal monitoring questionnaires and the previously described responses selected from the baseline questionnaires, with some modifications from the aforementioned variable creation and grouping. The full set included both individual 24-hour KAP concentration and personal $PM_{2.5}$ exposure measurements, 48-hour average KAP concentration and personal $PM_{2.5}$ exposure estimates, and average ambient $PM_{2.5}$ concentration estimates for each village by sampling period. Variables were also developed from sample form data to account for statistical effects from: sample period, distance of equipment from primary stove and floor, and whether a ventilation source was noted as < 1m from the equipment on which each 24-hour sample was begun. While not transportable, variables for household ID and village were retained as a means of improving statistical analyses, as described in more detail below. Meteorology data were also retained. It should be noted that many variables were found to have perfect correlation with at least one other variable, which is not unexpected for such a large database. The application of a common variable screening method did not substantially alter SuperLearner model [13] performance and computational burden was not high, so, for simplicity, variables were not actively culled for high correlation in any of the datasets.

KAP and personal $PM_{2.5}$ exposure samples with a start or end flow rate of less than 1.4 LPM or greater than 1.6 LPM were excluded. Filter samples with runtimes greater than 28 hours or less than 20 hours were discarded in order to avoid samples unrepresentative of a full day activity cycle. Sampling forms were manually examined to find and account for errors or issues affecting the samples; when non-conformity[14] was discovered or considerable error expected, such samples were omitted from analysis. In two samples where the primary KAP filter was discarded for failing to meet the aforementioned standards, it was possible to reassign a duplicate KAP filter as the primary KAP sample.

[13] Variable screening was explored on the Full dataset predicting 48-hour average exposures.

[14] For example, one household had only a single day of gravimetric exposure measurement due to equipment problems.

Missing or invalid observations in independent questionnaire and KAP concentration variables were imputed using k-nearest neighbors (k = 20) on the full dataset,[15] with the exception of outcome variables, household ID, variables considered personally identifying, and meteorology data. Prediction of exposure values as stratified by sampling period would later be used as a model performance metric, and so variables for sampling period and stove type, which was a near-perfect proxy for sampling period, were also excluded from the dataset during imputation. Missing meteorology data were imputed separately using k-nearest neighbors (k = 5) on the entire meteorology data subset, which also included date, village, period, and whether the missing data were from a "morning" or "afternoon" weather observation. A binary variable to indicate imputation of an observation was created for each imputed variable. Rows with missing 24-hour exposure estimates were dropped after imputation. Variables of a categorical class were converted into indicator variables with j-1 levels, where "j" is the number of unique responses recorded in each categorical variable. The least frequent response for each categorical variable was selected as the variable's reference level.

Statistical analysis

Analyses were performed in the R statistical program, version 3.3.2 (R Core Team 2016). Data were analyzed using the SuperLearner (Polley et al. 2016) and randomForest (Liaw and Wiener 2002) packages. All SuperLearner processing was performed with the following learners: random forest, an ensemble of random forest and bagging that uses conditional inference trees for base learning ("cForest"), extreme gradient boosting, neural networks with a single hidden layer, support vector machines, and 10-fold cross-validated (CV) generalized linear modeling with regularization.[16] Most households contributed two sets of data to the models (Before and After measurements); values from the same household were kept together (by fold) during cross-validation. Variable importance for the prediction of 48-hour exposure was assessed using the randomForest package. Variables for household ID and village were not included in the variable importance analysis. Model performance was evaluated using the coefficient of determination (r^2) of observed exposure values regressed on predicted exposure values, and by comparing the similarity of observed and predicted exposure values as stratified by sampling period (Before vs. After). The latter was chosen to reflect real-world use in the context of monitoring and evaluating an exposure intervention program.

The kitchen exposure factor, or the ratio of a person's exposure to their KAP concentration, is often used along with KAP concentration data to quantify personal $PM_{2.5}$ exposure in the cookstove context. KEF can be estimated from the literature to produce exposure estimates from nothing more than measured KAP values, allowing investigators to skip costly and

[15] The full dataset after excluding rows with incomplete personal exposure data.

[16] Learners as identified by SuperLearner: SL.randomForest, SL.cforest, SL.xgboost, SL.nnet, SL.svm, SL.glmnet. All were used with default settings.

burdensome personal exposure measurement campaigns. KEF can also be estimated onsite from concurrent exposure and KAP measurements in a subset of the study population; this involves some exposure measurement, but researchers are spared the burden of outfitting the entire population with monitors. The predictive value of KEF was explored in this dataset using 10-fold CV: a mean 48-hour KEF was calculated on approximately 90% of sample data and multiplied across individual 48-hour KAP concentration measurements on the remaining ~ 10% of data to produce estimates of exposure for those participants. This was performed ten times with each iteration leaving out a unique ~ 10% of the data until cross validated exposure predictions were calculated for all participants. This procedure was performed on all KAP and exposure data and, again, on the dataset as stratified by sampling period (5 folds per stratum). Before and After samples from the same household were kept together during CV fold creation.

The ability of the data to predict 48-hour average $PM_{2.5}$ exposure was explored using 10-fold CV SuperLearner methods on three primary variable sets. The first training set included all original and imputed independent variables with the exception of village, as it is largely non-transportable, and sampling period, as it is a direct aspect of the pre-determined model performance metric ("Full" set). Stove type, as a near-perfect proxy for sampling period, was also excluded. For data quality assurance, an iteration of this model was also produced using the un-imputed dataset.[17] The second training set excluded all KAP and meteorology variables from the Full set ("Full Without KAP" set). The third set included only KAP, meteorology, and ambient concentration data from the Full set ("KAP Without Surveys" set). Exposure outcomes were not transformed for modeling.

The Full set was also explored with CV SuperLearner using each of the first day and second day 24-hour average exposure measurements as the outcomes of interest. When predicting individual 24-hour average outcomes, only gravimetric and meteorology data from the day of measurement were included. For example, when predicting the first 24-hour average exposure, the second day KAP concentration, related meta data, and meteorology data were excluded from analysis.

A group of four training datasets that focused on LSIS-type questionnaire data was also explored with CV SuperLearner in order to better understand the predictive power of DHS indicators and the adjuvant power of other measurements. These included a set of only the LSIS-type questionnaire variables ("LSIS Only"); a set of LSIS-type, ambient $PM_{2.5}$ concentration, and all meteorology variables ("LSIS and Outdoor"); a set of LSIS-type variables and self-reported wood fuel use (kg) from the post-monitoring questionnaires ("LSIS and Wood Use"); a set of LSIS-type variables and self-reported heating variables from

[17] The un-imputed "Full" dataset consisted of only 37 samples (n_{Before} = 13, n_{After} = 24) and 278 covariates, compared to the 60 samples (n_{Before} = 27, n_{After} = 33) and 305 covariates included in the post-imputation dataset. Differences in covariates arose from the changes to data processing outputs caused by data missingness. In general, the types of covariates included in the un-imputed and post-imputation datasets were similar.

post-monitoring questionnaires ("LSIS and Heating"); and a set of LSIS-type variables and the previously described exposure score ("LSIS and Exposure Scores").

The distributions of observed and predicted exposure values did not always match, and so exposure values were left untransformed during t-testing. For example, the distribution of observed After exposures (48-hour average) was normal, while After exposures predicted using the Full Without KAP dataset were distributed log-normally. Paired student's t-tests (n_{Before} = 54, n_{After} = 66) were used to compare observed and predicted exposure distributions as stratified by sampling period. Both paired and unpaired student's t-tests were used to assess the difference in observed exposures between the Before and After sample periods. The same was performed for KAP measurements. The non-parametric, unpaired Mann-Whitney Wilcoxon rank sum test was used to assess the difference in average outdoor $PM_{2.5}$ concentration between the Before and After sampling periods. Unpaired student's t-tests were used to assess the difference in various continuous meteorological variables between the Before and After sample periods. When testing for significant differences in survey indicators between Before and After periods, an unpaired student's t-test was used for continuous variables and a Fisher's Exact test was used for categorical and binary variables. All significance tests were two-sided with an α of 0.05.

3.3 Results

Air pollution and meteorological conditions

A total of 6 variables related to KAP measurement were available for use. Viable 24-hour KAP measurements from the first day of sampling (Day 1) were collected in 34 households in the Before period and 35 households in the After period. Viable 24-hour KAP measurements from the second day of sampling (Day 2) were collected in 32 households in the Before period and 35 households in the After period. Overall, combined 48-hour average KAP concentration estimates were calculated in 31 households in the Before period and 34 households in the After period. Values for the 15 missing or invalid KAP estimates were imputed. Indicator variables for imputation of Day 1, Day 2, and the 48-hour average concentrations were created, bringing the total number of KAP variables available for analysis to 9.

Viable 24-hour personal $PM_{2.5}$ exposure concentrations for Day 1 were collected for 34 participants in the Before period and 34 participants in the After period. Viable 24-hour personal exposure concentrations for Day 2 were collected for 28 participants in the Before period and 34 participants in the After period. Overall, combined 48-hour personal exposure estimates were calculated for 27 participants in the Before period and 33 participants in the After period. No personal concentration estimates were imputed, because it is the outcome of interest.

Missing personal $PM_{2.5}$ exposure concentration values determined the final sample size for analysis: 33 After samples and 27 Before samples from 34 unique participants (60 total samples; 19 in village A, 22 in village B, and 19 in village C). Summary pollution concentration statistics for the 60 complete samples are shown in Table 3.1. Concentration distributions are demonstrated in Figure 3.3. Unpaired t-tests on the untransformed data from all 60 samples showed that the 48-hour average personal $PM_{2.5}$ exposure concentration in the Before period was significantly different from that of the After periods ($p < 0.001$). The same was true of 48-hour average KAP concentrations ($p < 0.001$). Day 1 average concentrations were marginally different from paired Day 2 averages ($p = 0.07$), driven by a significant difference ($p < 0.01$) in the Before period. Average KAP concentrations were not significantly different between Days 1 and 2. Paired t-tests of Before and After 48-hour samples in the subset of participants with both Before and After measurements ($n = 52$) proved significant for both exposure and KAP ($p < 0.001$). KEF varied significantly between the Before and After period when tested using the paired and unpaired methods described above ($p < 0.01$).

Table 3.1. Kitchen Air Pollution and Personal $PM_{2.5}$ Exposure Concentrations, by Period and Sample Day

		Kitchen Air Pollution Concentration				Personal $PM_{2.5}$ Exposure Concentration				KEF	
	n	mean ($\mu g/m^3$)	+/- 95% CI b	GM	GSD	mean ($\mu g/m^3$)	+/- 95% CI b	GM	GSD	Mean	SD
Before	*27*										
Day 1		499 a	182	376	2.1	139 a	28	124	1.6	0.45 a	0.43
Day 2		470 a	163	350	2.1	107 a	19	98	1.5	0.39 a	0.36
48 Hour Avg.		462 a	144	370	2.0	123 a	22	113	1.5	0.42 a	0.42
After	*33*										
Day 1		131 a	30	113	1.7	78	12	71	1.6	0.70	0.32
Day 2		116	15	109	1.5	83	14	69	2.2	0.80 a	0.50
48 Hour Avg.		124 a	20	114	1.5	81	11	75	1.5	0.72	0.29

Note: GM = Geometric Mean (calculated as $\exp(\text{mean}(\log(X_1, X_2, \ldots, X_n)))$), GSD = Geometric Standard Deviation (calculated as $\exp(\text{sd}(\log(X_1, X_2, \ldots, X_n)))$), KEF = Kitchen Exposure Factor

a. Values comprising this arithmetic mean are distributed log-normally ($p < 0.05$)

b. For non-transformed data

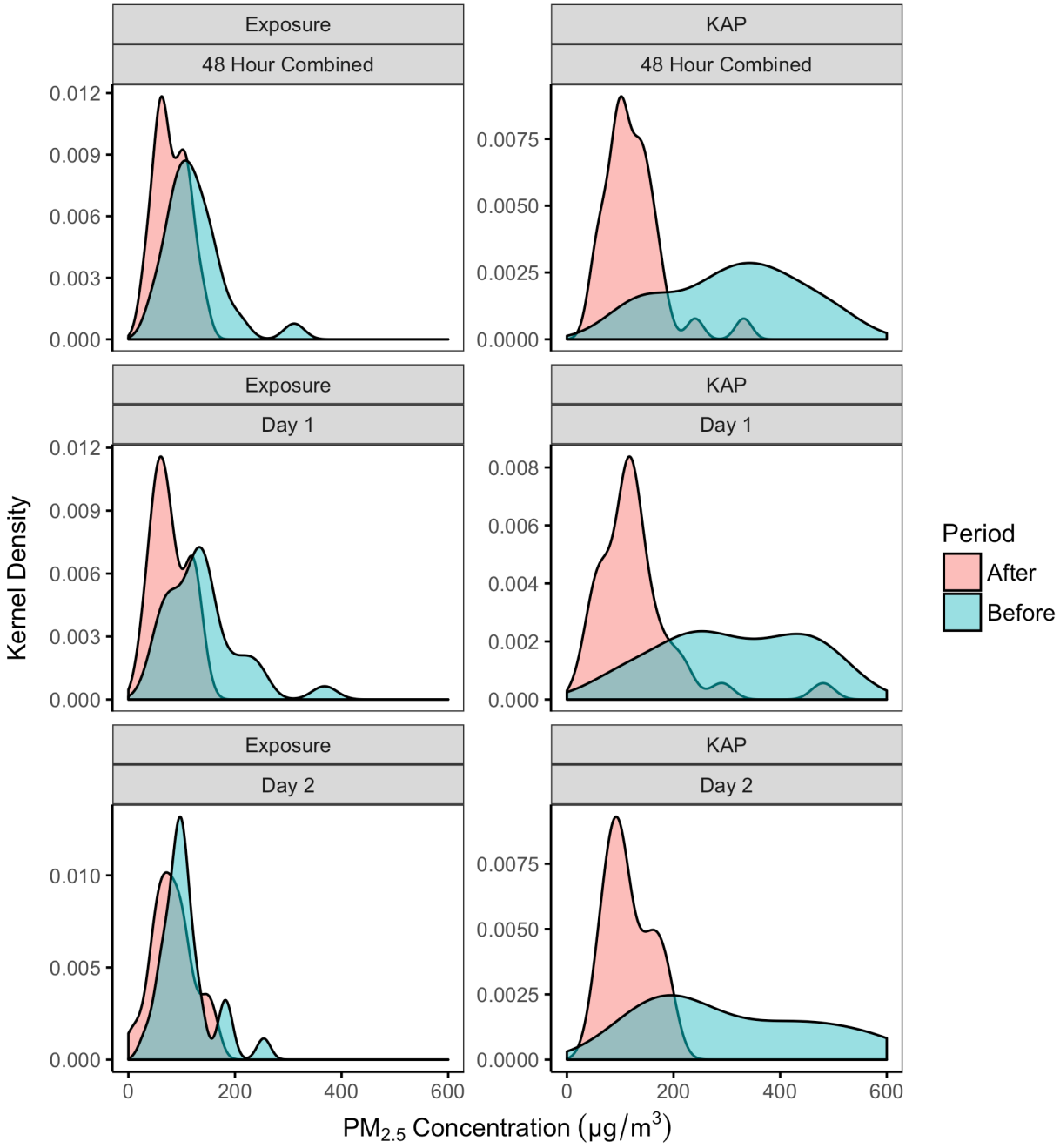

Figure 3.3. A kernel density plot showing distributions of personal $PM_{2.5}$ exposure (left) and Kitchen $PM_{2.5}$ Air Pollution (KAP; right), by sampling period (blue and red) and sampling time (from top to bottom). Concentrations along the x-axis are limited to 600 µg/m^3 to allow easier interpretation of all graphs in the same figure – excluded were 6 points from the 48 Hour Combined KAP, 6 points from the Day 1 KAP, and 5 points from the Day 2 KAP plots.

Budget constraints and initial difficulties with equipment resulted in too few viable samples to produce ambient $PM_{2.5}$ concentration estimates for each day of measurement. Thus, individual 24-hour measurements were aggregated to produce a single estimate of average ambient $PM_{2.5}$ concentration for each village during each sampling period (Table 3.2). Each sample was assigned an ambient concentration based on the village in which the participant lived and the sampling period. A significant difference in average ambient concentrations between sampling periods was not observed ($p = 0.74$).

Table 3.2. Summary Statistics for Outdoor Ambient $PM_{2.5}$ Concentrations, by Sampling Period and Village

	n	Mean ($\mu g/m^3$)	Max ($\mu g/m^3$)	Min ($\mu g/m^3$)	SD ($\mu g/m^3$)
Before	7	52	73	26	16
Village A	2	49	50	47	-
Village B	3	57	66	43	12
Village C	2	49	73	26	-
After	14	53	147	15	34
Village A	4	57	82	15	30
Village B	5	70	147	38	44
Village C	5	34	57	20	15

The difference between Before and After concentrations was not statistically significant ($p = 0.74$)

Six (6) meteorology variables were available for analysis. All meteorology observations were missing for 2 of the 36 intended morning weather measurements and 2 of the 36 intended afternoon weather measurements (about 6% of all meteorology data). Two (2) additional morning and 2 additional afternoon measurements of wind direction were also missing (about 1% of all meteorology data). The median observation times of un-imputed morning and afternoon measurements were 9:06 AM and 12:05 PM, respectively. This did not differ considerably by sample period. All missing observations were imputed using the full meteorology dataset, and variables for date, time of day (morning vs. afternoon), village, and sampling period. New variables were created to indicate imputation. Key meteorology data are summarized post-imputation in Table 3.3. Wind direction was then split into binary indicators for northeast (NE), east (E), southeast (SE), southwest (SW), west (W), and northwest (NW) with north (N) as a reference level. New meteorology variables were created to distinguish each reading by “morning” and “afternoon”. These variables were assigned to individual 24-hour gravimetric measurements by gravimetric measurement start date. After this merge, 21 variables were left without variation; uniform variables are of no use to modeling, and so they were dropped, leaving a total of 47 meteorology variables for consideration during analysis.

Table 3.3. Summary Statistics for Key Meteorology Data, by Sampling Period and Village

	Mean Barometric Pressure (mb) a	**Mean Temperature (°C)** b	**Mean RH (%)**	**Mean Wind Speed (km/h)** b	**Most Observed Wind Direction** c	**Mean Afternoon Precipitation (mm)**
Before	1017.1	24.3	56	6.1	SW	0.3
Village A	1018.7	23.3	53	8.5	SE	0.0
Village B	1017.3	23.3	54	6.5	NW	0.8
Village C	1015.2	26.3	61	3.2	SW	0.0
After	1015.9	25.9	54	3.9	NW	0.2
Village A	1015.4	27.8	53	4.2	SE	0.0
Village B	1015.0	24.5	57	3.2	NW	0.7
Village C	1017.3	25.3	51	4.3	NW	0.0

Note: with the exception of mean afternoon precipitation, "mean" values are taken as the average among all morning and afternoon measurements.

a. Difference between Before and After concentrations marginally significant using unpaired student's t-test ($p < 0.10$)

b. Difference between Before and After concentrations statistically significant using un-paired student's t-test ($p < 0.05$)

c. Difference between Before and After values not tested for significance

Questionnaires

A total of 246 variables from the baseline and post-monitoring questionnaires were available for analysis (not including variables for household ID and village). Eight (8) observations across 8 post-monitoring questionnaire variables were missing from the full dataset (<< 1 % of all included observations). Missing values were imputed, and an indicator variable was created for each imputation. Cooking exposure scores were created, adding 6 new variables to the set. Questionnaire data were truncated to match the 27 participants with complete exposure data from the Before period and the 33 participants with complete exposure data from the After period. In this data subset, 28 post-monitoring questionnaire variables and 44 baseline questionnaire variables had only 1 unique response, and so were dropped. Of the remaining questionnaire variables, 57 were categorical. These categorical variables were converted into a total of 120 indicator variables. After variables for sampling period and stove type were excluded, the total number of questionnaire variables for inclusion in the Full analysis set was 248. A summary of key questionnaire indicators is shown in Table 3.4.

Table 3.4. Summary of Selected Questionnaire Responses

	All Study Data Mean (SD or %)		**Dataset as Analyzed** a Mean (SD or %)		
	Before	*After*	*Before*	*After*	p-value
n	*36*	*36*	*27*	*33*	-
Age of Cook (years; reported during Baseline Questionnaire)	35 (11)		34 (11)	35 (11)	0.95
Female Head of Household	4 (11%)		4 (15%)	4 (12%)	1.00
Self Reported Spending (KIP/month)	316,000 (710,000)		201,000 (343,000)	314,000 (745,000)	0.44
Electricity Access	36 (100%)		27 (100%)	33 (100%)	-
Primary Biomass Stove is ACE 1	0	36 (100%) b	0	33 (100%) b	<0.01*
Primary Biomass Stove is Open Fire	30 (83%)	0 b	24 (89%)	0 b	<0.01*
Reported Wood Use for Cooking (kg)	10.5 (5.1)	2.4 (1.2)	10.2 (4.5)	2.4 (1.3)	<0.01*
Kitchen Size (m3)	40 (33)		41 (32)	42 (34)	0.91
Cooking Location at Home					
In House (Separate Room)	1 (3%)	0	1 (4%)	0	0.44
In House (Elsewhere)	30 (83%)	33 (92%)	22 (81%)	30 (91%)	
In Separate Building	5 (14%)	3 (8%)	4 (15%)	3 (9%)	
Time Spent Cooking per Day at Home (minutes)	543 (211)	364 (146)	542 (223)	368 (147)	<0.01*
Total Cooking Exposure Score	6,896 (3,441)	4,356 (3,498)	7,423 (3,643)	4,353 (3,516)	<0.01*
Time Spent Grilling per Day at Home (minutes)	86 (113)	29 (59)	87 (123)	31 (61)	0.04*
Total Grilling Exposure Score	3,744 (6,537)	1,414 (5,375)	4,284 (7,315)	1,542 (5,603)	0.12
Smoking Occurred in House	28 (78%)	27 (75%)	23 (85%)	25 (76%)	0.52
Smoking Occurred in Kitchen	5 (14%)	3 (8%)	4 (15%)	3 (9%)	0.69
Ever Used a Heat Source During Sampling	23 (64%)	10 (28%)	20 (74%)	9 (27%)	<0.01*
Time Activity During 48 Hour Sampling (hours)					
Kitchen	8.9 (3.3)	7.4 (3.2)	9.3 (3.5)	7.5 (3.2)	0.04*
Inside Home, but Not In Kitchen	21.6 (5.2)	27.7 (5.2)	22.4 (5.4)	27.8 (5.4)	<0.01*
Inside, at a Job	0.1 (0.5)	0	0.1 (0.6)	0	0.33
Inside, Elsewhere	2.4 (3.1)	1.1 (2.0)	2.0 (3.1)	0.9 (1.4)	0.08
Outdoors, at a Job Site	6.6 (5.2)	6.2 (6.8)	6.5 (5.2)	6.7 (6.8)	0.88
Outdoors, Elsewhere	8.4 (4.3)	5.6 (5.7)	7.7 (4.1)	5.1 (5.3)	0.04*

a. Twelve samples were dropped prior to analysis for a lack of outcome data. This is described in the text below. Some metrics in the full dataset may be reported as constant between Before and After sections in the full dataset, but not in the dataset as analyzed. This is due to the fact that in the dataset as analyzed, some household samples were retained in one sampling period but not the other. Tests for significant differences between Before and After responses were performed only for the dataset as analyzed.

b. Field notes indicate that at least one participant's ACE-1 stove was non-operational during one of the two days of After period sampling.

* Significant difference.

Prediction models

Summary statistics for models produced using all 48-hour exposure datasets are shown in Table 3.5. Paired t-tests compare observed values with those predicted by the models—this is a relevant statistic for the cross-validated context if we consider the prediction models constructed for each training sample as fixed (Dudoit and van der Laan 2005). Table 3.6 summarizes relative predictive performance of the super learner components for the 48-hour average exposure model run on the Full dataset. Figure 3.4 shows observed exposure values plotted against values predicted using CV SuperLearner with each of the three primary datasets and against values predicted using the KEF method with stratification. Among the three primary datasets, the model produced using the Full Without KAP dataset estimated Before and After exposures with the least reliability. Little difference was observed between the predictive power of the Full and Full Without Surveys datasets, which both produced models that accurately predicted Before and After concentrations. Both models demonstrated strength in predicting individual exposure values with $r^2 \sim 0.25$. The model produced using the Full un-imputed dataset performed similarly well to the model produced on the true Full dataset.

Table 3.5. 48-Hour Model Results - Observed vs. Predicted

		Before Sampling Period			After Sampling Period		
Model	**r2 (Observed vs. Predicted)**	**Mean**	**SD**	**p-value (against Observed)**	**Mean**	**SD**	**p-value (against Observed)**
Observed		123.2	54.5		80.8	30.5	
Predicted							
Full	0.26	120.1	17.9	0.76	88.1	16.4	0.15
Full – Un-imputed Dataset	0.49	117.6	10.8	0.46	86.8	15.3	0.19
Full without KAP	0.01	105.5	11.6	0.12	91.7	6.9	0.07
Full Without Surveys	0.27	119.0	23.4	0.68	87.1	22.7	0.21
KEF Only	0.03	240.8	206.1	0.02	67.4	58.2	0.00
KEF Only - Stratified	0.02	199.5	170.5	0.04	89.0	40.3	0.28
LSIS Only	0.00	97.5	6.0	0.02	97.7	6.5	0.01
LSIS and Outdoor	0.31	119.7	26.0	0.72	85.8	16.4	0.30
LSIS and Wood Use	0.05	102.7	14.5	0.06	91.9	9.1	0.06
LSIS and Heating	0.04	108.1	17.7	0.16	91.7	10.7	0.09
LSIS and Exposure Scores	0.01	103.6	8.8	0.08	98.2	10.9	0.00

Table 3.6. Candidate and Super Learner Performance

Algorithm	Average RMSE (µg/m³) *
Super Learner	40.1
Discrete SL	40.0
SL.randomForest_All	40.0
SL.nnet_All	48.1
SL.glmnet_All	44.6
SL.xgboost_All	50.1
SL.svm_All	47.7
SL.cforest_All	43.2

* Square root of the mean squared wrror output produced by the CV SuperLearner model run on Full Data with 48 Hour Exposures

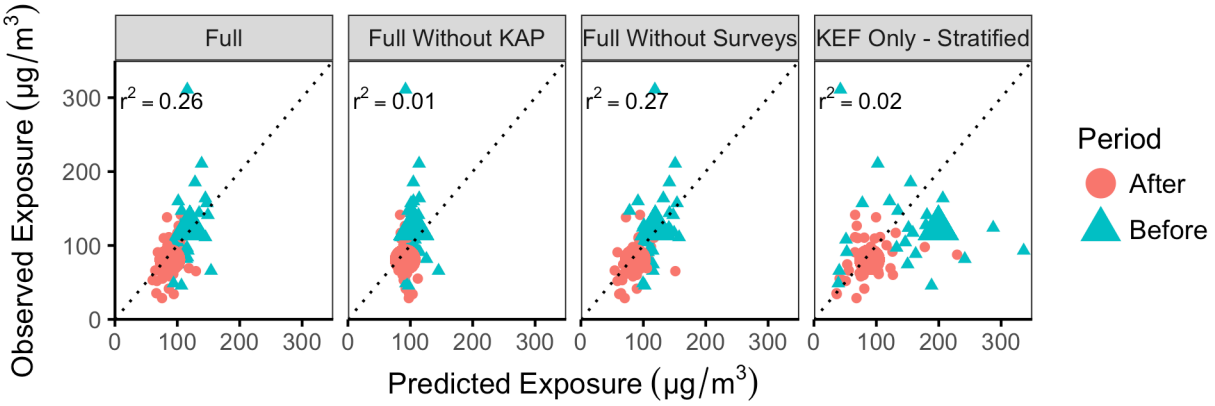

Figure 3.4. Observed 48-hour average exposures plotted against predicted values for the three primary models produced with CV SuperLearner and the better-performing of the two traditional KEF-based models – KEF with stratification. Individual concentrations are depicted by smaller points, and mean concentrations, by larger points. All data are stratified by sampling period. The coefficient of determination (r^2) for the regression of observed values on predicted values is shown in the top left corner of each panel. X and Y axes were limited to 350 µg/m³ to improve interpretability of the figures; these bounds were exceeded by 4 predicted values (max X = 861.3 µg/m³) in the KEF Only – Stratified panel.

KEF proved a poor predictor of exposures when applied traditionally, especially in the Before sampling period. Stratification by sampling period improved results slightly, with most improvement seen in the After period predictions. 48-hour exposures were also modeled from 48-hour average KAP concentrations with SuperLearner [18] as a

[18] Excluding learners reliant on forests and regularization (i.e. SL.randomForest, SL.cforest, and SL.glmnet), which did not function with only than one predictor variable.

complementary analysis of what a KEF-type model would look like with the benefits of super learning and machine learning. This improved performance slightly (r^2 = 0.09), producing mean Before and After exposures that were closer to measured values (Before: 85.7 $\mu g/m^3$, p = 0.19; After: 108.3 $\mu g/m^3$, p = 0.36).

The model produced on LSIS Only data did not successfully delineate mean Before and After concentrations. Addition of ambient $PM_{2.5}$ and meteorology data dramatically improved the LSIS dataset's predictive power, and produced the most-predictive model of all. The addition of heating variables, wood use variables, and exposure score variables all improved performance slightly.

Variable importance analyses uncovered several useful data groupings. They are demonstrated in Table 3.7, which shows the 10 most important variables in the Full, Full Without KAP, and one of the LSIS datasets – LSIS and Heating. Relative importance is demonstrated in terms of percent increase in mean squared error (MSE) caused by removing each variable during Random Forest selection. Meteorology and KAP data made the greatest contributions to the Full data model. The sets of 10 most important variables were very similar between the Full and Full Without Surveys models. The model produced on the Full set without meteorology or KAP data was less robust, and produced relatively poor estimates of sample period mean exposures. Generally, LSIS-type variables related to household size, ethnicity, and cooking location made the greatest contributions to LSIS model performance. Among all models, a select few variables produced a considerably larger impact than the others. For example, relative importance in the Full dataset drops considerably after Day 1 24-hour KAP concentration, the fourth most important variable.

Table 3.7. 10 Most Importance Variables for Prediction for Select Datasets

Full Dataset		Full Without KAP		LSIS and Heating	
Variable	**% MSE Increase**	**Variable**	**% MSE Increase**	**Variable**	**% MSE Increase**
Morning Wind Speed - Day 1	407	Combined Cooking Exposure Score	147	Indoor Heating - Yes	319
Morning Wind Speed - Day 2	257	Wood Fuel Use (kg)	52	Ethnicity - Lao	87
Morning %RH - Day 1	187	Home Cooking Exposure Score - Morning	48	Any Heating	31
24-Hour KAP Concentration - Day 1	150	Indoor Heating Type - None (a)	45	Total Household Size	29
48-Hour KAP Concentration	73	Ambient PM2.5 Concentration	34	Number of Household Members aged 5-14 Years	19
Combined Cooking Exposure Score	41	Indoor Heating - Yes (a)	34	Ethnicity - Katang	14
Afternoon Pressure - Day 1	32	Number of Pigs Owned	28	Kitchen Type - Cooks In House (Elsewhere)	13
Afternoon Wind Speed - Day 1	31	Home Cooking Exposure Score - Afternoon	25	Number of Pigs Owned	8
Morning Wind Direction, Blowing West - Day 2	29	Combined Grilling Exposure Score	24	Highest Education in Home - Lower Secondary	8
Afternoon Temperature - Day 2	29	Owns Radio	14	Kitchen Type - Cooks In Separate Building	6

a. Perfectly correlated with each other

Observed Day 1 and Day 2 24-hour average exposure values plotted against the values predicted using CV SuperLearner on the Full dataset are shown in Figure 3.5. Models created with either Day 1 or Day 2 measurements accurately predicted sample period means ($p > 0.15$). The model trained on Day 1 measurements produced a much stronger fit ($r^2 = 0.32$) than the model trained on Day 2 measurements ($r^2 = 0.02$) or any of the models produced using 48-hour average exposures as the outcome of interest (highest 48-hour average $r^2 = 0.25$).

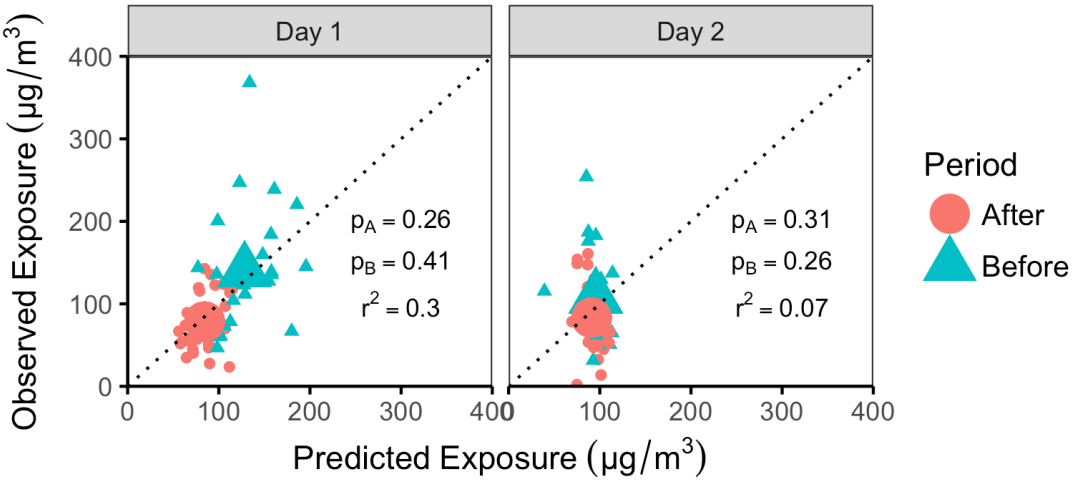

Figure 3.5. Observed 24-hour average exposures plotted against predicted values produced with CV SuperLearner and the Full dataset. Individual concentrations are depicted by smaller points, and mean concentrations, by larger points. All data are stratified by sampling period. The coefficient of determination (r^2) for the regression of observed values on predicted values, the p-value for a paired student's t-test for significance between observed and predicted Before values (p_B), and the same for After values (p_A) are shown in the bottom right corner of each panel.

3.4 Discussion

KAP and exposure concentrations in the study group

The personal $PM_{2.5}$ exposure measurements taken in this study are, to my knowledge, the first in solid-fuel-using cooks in Lao PDR. Values were within the range of those experienced by cooks using solid fuels, globally. A recent review of the literature (Balakrishnan et al. 2014) estimated the mean of 24-hour average exposures for women cooking with solid fuels around the world at 267 µg/m³ (SD: 297 µg/m³). Another study modeled piecemeal estimates of exposure concentrations in Lao PDR in cooking & eating, heating, and illumination related indoor micro-environments (Shimada and Matsuoka 2011). Together, the combination of the values estimated for those environments suggests a daily average exposure concentration of about 266.3 µg/m³ in cooking & eating micro-environment, 151.9 µg/m³ in a heating micro-environment, and 0.9 µg/m³ in an illumination context. A distinction was not made between rural and urban residents. The annual average exposure value for women cooking with solid fuels assumed in the household air pollution comparative risk assessment of the Global Burden of Disease 2010 project (GBD 2010) (Smith et al. 2014) was 337 µg/m³ (95% CI: 238, 479).[19] These values are between 120 - 170% greater than the mean 48-hour exposure measurement in the present analysis of Lao

[19] The GB 2010 HAP exposure value was based on an exposure model developed for India.

women using traditional wood cookstoves (123 μg/m^3, 95% CI: 101, 145) and about 230 - 320% greater than that of ACE-1 stove users (81 μg/m^3, 95% CI: 70, 92).

The lower nature of exposures experienced in Lao women may be attributed to the hybrid indoor-outdoor nature of cooking in the region. Dwellings and kitchens in Lao homes were observed to be highly ventilated relative to other areas in which the research team has worked. Rooms often have very large eaves, windows, and walls that do not reach the roof. A similar ACE-1 intervention in nearby Cambodia (Berkeley Air Monitoring Group 2015) conducted after the Lao study reported mean 48-hour average cook exposures that were even lower: 66 μg/m^3 (95% CI: 52, 80) in while using traditional biomass stoves and 47 μg/m^3 (95% CI: 35, 59) while using the ACE-1 stove.

The KAP measurements presented in this chapter are among the first in Southeast Asia. The Cambodia ACE-1 project reported mean 48-hour average KAP concentrations of 183 μg/m^3 (95% CI: 123, 243) in households cooking with traditional biomass stoves and 111 μg/m^3 (95% CI: 73, 149) after a switch to ACE-1 biomass stoves. A 2011 campaign observed 12-hour average concentrations of particulate matter smaller than 10μm in diameter (PM_{10}) in other Lao households (Viantiane and Bolikhamxay provinces) where traditional biomass cooking stoves were used (Morawska et al. 2011). Their measured PM_{10} concentrations averaged approximately 1200 μg/m^3, or nearly three times the mean 48-hour average $PM_{2.5}$ concentrations observed during the Before period in Savannakhet (462 μg/m^3, 95% CI: 318, 606). Twenty-four (24) hour mean kitchen concentrations have been measured at 972 μg/m^3 (SD: 876 μg/m^3) on average in the literature for homes cooking with solid fuels around the world (Balakrishnan et al. 2014).

Kitchen $PM_{2.5}$ concentrations measured in Lao were higher than those measured in Cambodia. This may be explained by seasonality. The Cambodia fieldwork was completed between July – August when precipitation is typically greater in the region than during the Lao study period (Central Intelligence Agency 2017). Increased rainfall may result in more atmospheric particle clearing. Meteorological conditions were not discussed in the Cambodia report (Berkeley Air Monitoring Group 2015). Differences in household characteristics between the two sites may also have played a role. It is possible that cultural and household characteristics, including food type and traditional stove type, are systematically different between the Cambodia study site and Savannakhet in a way that appreciably affects indoor $PM_{2.5}$ concentrations. Such a theory is supported by the 2011 Lao study, which found significant differences in household PM_{10} concentrations by various household characteristics and by district.

PM_{10} concentrations measured in the 2011 study (12-hour means in Bolikhamxay and Vientiane provinces of 1183 μg/m^3 and 1275 μg/m^3) were considerably higher than the $PM_{2.5}$ concentrations presented in this chapter. As discussed above, it is possible that systematic differences in household characteristics or cultural practices existed between the study locations in a way that affected concentrations. The 2011 PM_{10} study also only measured 12-

hours at a time, incorporating all or most major cooking events and excluding nighttime concentrations, which likely produced an overestimation of the daily average. But the difference between PM_{10} and $PM_{2.5}$ should not be ignored. By definition, the mass concentration of PM_{10} can never be lower than that of $PM_{2.5}$ in the same parcel of air; $PM_{2.5}$ is, by definition, a subset of PM_{10}. For example, a study of outdoor air quality in Bhaktapur, Nepal (Pokhrel et al. 2015) – a region where much cooking with biomass fuels occurs – measured the ratio of $PM_{2.5}/PM_{10}$ at 0.55. Application of this ratio to the PM_{10} measurements would produce an expected $PM_{2.5}$ concentration of roughly 660 $\mu g/m^3$, which is closer to the 48-hour mean measured in Savannakhet kitchens.

Lessons from the ACE-1 intervention data

KAP concentrations and personal exposures fell significantly between the Before and After periods ($p < 0.001$). By itself and from a health risk standpoint, this would suggest the ACE-1 stove intervention was a success. However, this analysis did not hypothesize the causal link between ACE-1 use and exposure reduction. Causal inference methods identify and explain the factors that directly alter an outcome (e.g. an X unit increase in variable A causes a Y unit increase in outcome B), but they do not necessarily produce models with high predictive power (Shmueli 2010). Alternatively, predictive methods are intended to estimate new or future values of an outcome from a set of covariates with the greatest possible accuracy and precision, but do not necessarily elucidate direct relationships between an outcome and covariates. A third method type, descriptive statistics, identifies associations observed between covariates and an outcome (e.g. patients diagnosed with concussion are on average Z% wealthier than the general population [20]). Those relationships may be neither causal (e.g. wealth did not cause the concussion; taking hard hits as a well-paid professional football player did) nor helpful in predicting the outcome. The analysis presented in this chapter focuses on predictive and descriptive methods, and so the discussion will not explore causal links in depth. However, analyses performed on several questionnaire and environmental indicators during exploration of predictive models pose compelling questions for future work.

Important weather indicators changed between the Before and After periods. Mean temperature was lower during the Before period ($p < 0.05$) while mean wind speed was greater ($p < 0.05$). Such changes may decrease physical comfort and encourage increased heating, thus increasing smoke exposures among participants. In fact, the use of heating during sampling was significantly higher ($p < 0.01$) during the Before period at 74% of participants, compared to 27% in the After period.[21] The effect may be to amplify the observed Before vs. After decrease in exposures. Future work should explore study designs that allow for delineation of the exposure effects of seasonality. Both meteorological and cultural events should be considered – e.g. harvest periods, changes in ingredients and thus

[20] This example and allusions to it throughout the text are merely hypothetical, and may not be true.
[21] The most commonly reported source of heat was from a fire or cooking stove.

cooking types, religious holidays, and celebrations with pyrotechnic or major cooking components.

The amount of self-reported fuel used and time spent in the kitchen dropped significantly between the Before and After periods ($p < 0.0001$ and $p < 0.05$, respectively). Time spent cooking at home and, specifically, grilling at home also fell significantly ($p < 0.01$ and $p < 0.05$, respectively). These reductions may have been initiated by an improved cooking efficiency afforded by the new stove, and, thus, may be the link that proves ACE-1-based reductions in exposures. However, such relationships should be carefully examined in studies wishing to demonstrate causality. The exposure scores developed during this study may prove useful for understanding how new stoves impact exposure-related cooking behavior, but suffer from the same causal inference issues as fuel use, cooking time, and time spent in the kitchen, and so should be used with caution.

Error in the data may have been incurred during questionnaire administration. While reviewing audio recordings, it became clear that some interviewers inferred specific information like time activity patterns or occupation from more general participant responses such as a broad description of daily life. Interviewers were trained in survey administration procedures prior to the study, and questionnaire phrasing was carefully considered for contextual relevance. However, not all queries could be made perfectly appropriate— like questions about time spent, when locals do not have much experience applying Western notions of minutes and hours— and subtle differences may have existed in how each interviewer perceived more complex or less-intuitive questions and responses. The translation process between English and Lao may have magnified the issue. Future studies should expend sufficient effort to ensure all questions are clear and understood to mean the same thing among all interviewers. In addition, such error could be reduced with more timely quality assurance measures such as audio review at the end of each study day and frequent check-ins with interviewers.

The presence or absence of the household head during the interview may also have resulted in erroneous or biased responses. Absence of the household head is expected to affect responses related to issues of which the cook may not have a firm understanding – like finances – while their presence may influence the cook to provide answers with less candor or to agree with input from the household head on issues of which the head has less understanding. The household head was reported present during about two thirds of baseline questionnaires, which did not differ significantly by sample period. This was not tracked during post-monitoring questionnaires. One might avoid this issue by interviewing both the household head and cook separately when such bias is expected, though this may not always be culturally or ethically appropriate. In addition, recall bias may have affected various questionnaire responses, especially those about behavior during the multiple days of monitoring. More frequent survey administration or, preferably, implementation of non-survey tools like diaries or a device like the Time Activity Monitoring System (Allen-Piccolo

et al. 2009) – an ultrasound-based device for measuring time spent in various micro-environments – may reduce recall bias in future work.

Predicting $PM_{2.5}$ exposures with traditional and new methods

Models rooted in machine learning and super learning produced reliable estimates of mean exposures in both the Before and After sample periods. The best models calculated individual 48-hour $PM_{2.5}$ exposures with reasonable accuracy (r^2: 0.26 - 0.31), and predicted period-specific mean exposure values that were within 10% of and not significantly different from measured values. No other cross-validated prediction modeling could be found for HAP *exposures*, but at least one notable instance does appear in the HAP *concentration* literature (Balakrishnan et al. 2013): an application of 24-fold cross-validation to a model of 24-hour KAP concentrations produced from measurements of a small number of household-level, DHS-type variables in 617 households from 24 villages in India. Measured 24-hour KAP concentrations were predicted with a reliability considered modest for the estimation of individual exposures ($r^2 = 0.31$) but acceptable for prediction at the population level. Their model or a variation thereof was ultimately employed in the GBD 2010 to estimate global HAP concentrations.[22]

Under cross validation, the machine learning and super learning methods applied to larger covariate sets strongly outperformed the more conventional method of linear KEF-based exposure assessment. KEF-based predictions overestimated mean Before exposure by 60 - 95%, and produced estimates of the After period mean that were only marginally similar to the measured value. This is supported by the wide variation observed in individual and period-aggregated KEF values. Application of super learning and machine learning did improve the predictive performance of KAP (a mathematical variant of KEF when used as the sole predictor), but did not produce results on par with the best larger datasets. Researchers and practitioners should approach the predictive use of KEF with caution.

Poor generalizability of the bivariate relationship between indoor area concentration and exposure is supported by work performed elsewhere. A study of homes cooking with wood in Guatemala showed 48-hour average $PM_{2.5}$ exposure estimates for mothers that were typically 70% lower than 48-hour average KAP measurements in homes cooking with open fires but only 35% lower in homes using chimney stoves (Northcross et al. 2010). Despite average exposure values that were *lower* than average kitchen concentrations, 33% of measured exposure concentrations were *greater* than corresponding KAP levels. Women cooking with solid fuels over traditional open fires in Mexico (Armendáriz-Arnez et al. 2008) were shown to have mean 24-hour average $PM_{2.5}$ exposures over 75% lower than mean 48-hour average KAP concentrations. This relationship changed after the introduction of a Patasari chimney stove. In women cooking solely with the Patasrai, this

[22] Predicted HAP concentrations were later converted to exposure estimates using a set of exposure factor ratios.

difference dropped to about 70%; in women cooking with a Patsari who also maintained an open fire outside, it dropped to about 55%; and in women cooking with a Patsari who also maintained an open fire in the kitchen it dropped to about 40%. A study of cookstove users in rural China (Baumgartner et al. 2011) demonstrated a significant relationship between individual 24-hour kitchen concentrations and exposures in adults, but found no such connection in children and did not test predictive performance from a cross-validated standpoint.

The relative performance of the models examined in this analysis provides insight into which types of covariates may be of most predictive utility. Models trained on outdoor environmental data and KAP-related data were most accurate. Variable impact analysis identified morning wind speed, morning relative humidity, and KAP concentrations as producing the greatest impacts on model error. However, a direct causal link between these variables and exposures should not be automatically assumed.

For example, it is unclear what direct impact wind speed might have on exposures. To better understand this relationship, a simple linear model of 48-hour average exposure was produced for each sampling period. Each model included only two covariates: Day 1 morning wind speed and Day 2 morning wind speed. Holding Day 2 morning wind speed constant, a significant ($p < 0.01$) positive relationship between Day 1 morning wind speed and 48 hour average exposure was observed during the After sampling period. No other non-intercept covariates were significant within the two models. While not an overwhelming amount of evidence, the outcome of this process makes for a good thought exercise. The significant, positive relationship does not support the most intuitive potential cause: wind-related HAP dilution. Perhaps this particular relationship is mediated instead by physical comfort—as Day 1 morning wind speeds increase, comfort decreases, and Day 1 heating-related activities become more prominent. This seems counter-intuitive, because survey queries about heating were included in several models and their impact ratings did not exceed that of Day 1 morning wind speed. However, wind speeds could, theoretically, be a better marker of heating-related increases in exposures than a direct survey query if the language of that query does not address more nuanced behaviors, such as standing closer to otherwise-normal cooking fires, or if the survey tool is prone to error.

The relationship could be even more complex. For example, suppose that, all other factors constant, exposure-relevant heating behavior is considerably different when a person receives much of their heating from their cooking appliance and primarily cooks with the ostensibly more-efficient and more-contained ACE-1 stove than with an open fire. Because wind speeds were different ($p < 0.05$) between the Before period – when no homes cooked with an ACE-1 stove – and the After period – when all homes reportedly cooked with the ACE-1 stove – some of the influence of the wind speed variable may actually come from a stove-specific heating effect. Causal analyses could help to elucidate such complexities.

Questionnaire data demonstrated some predictive power, but were outperformed by other information. This suggests that accurate prediction of group mean exposures in the context of rural solid-fuel cooking might not require lengthy surveys. However, there is reason to believe that certain survey-related queries would be of use when predicting exposures in populations with household characteristics and behavior patterns that are more heterogeneous than those of the Lao study population. Previous studies have been able to explain a good deal of statistical variation in 24-hour and 8-hour average HAP exposures with data ascertained or ascertainable by questionnaire (Baumgartner et al. 2011; Clark et al. 2010), but few, if any, have investigated cross-validated prediction. Exploration of predictive modeling in larger, more variable populations is needed to better understand this phenomenon.

LSIS data produced the least well-performing super learning models. Poor performance of the LSIS data model in this specific study population may stem from the fact that the LSIS dataset does not include variables with much differentiation between seasons or cooking appliance. This theory was explored by training models on LSIS data combined with wood fuel usage (kg), heating, exposure scores, and outdoor environmental information, which all varied to some extent by sample period and stove type. Each new variable or variable set improved LSIS model performance considerably. Globally administered DHS surveys like LSIS may be improved by adding a small number of questions about stove type, fuel usage, and heating behaviors. It is also likely that DHS indicators provide more predictive benefit in regions in which household characteristics, like kitchen and stove type, are more variable or seasonal effects, like heating, are less variable. In support of the usefulness of DHS indicators, the GBD 2010 model was able to reliably predict KAP concentrations entirely from DHS-like data (Balakrishnan et al. 2013).

The single best candidate learner on the Full dataset was random forest, with an average root mean squared error (RMSE) of 40.0 $\mu g/m^3$, followed by cForest at 43.2 $\mu g/m^3$. Generalized linear modeling with regularization performed about as well as cForest with an RMSE of 44.6 $\mu g/m^3$. Researchers interested in HAP exposure prediction should explore decision trees and regression with regularization, and may benefit from the use of a super learner, which, in this analysis, did not improve upon the best individual candidate learner (RMSE: 40.1 $\mu g/m^3$). The neural network package did not outperform most other candidates (RMSE: 48.1 $\mu g/m^3$). However, this package used only a single hidden layer. Future analyses should investigate the utility of more complex neural networks.

3.5 Conclusion

$PM_{2.5}$ exposures and KAP concentrations in rural Lao PDR were within the range of those expected among women cooking with solid fuels worldwide. Both exposures and KAP were lower after transition from an open fire to the ACE-1 stove. Causal factors were not assessed, but descriptive and predictive analyses did highlight the importance of understanding the effects of multiple factors, like meteorology, in any such monitoring and evaluation project.

This analysis also demonstrated that exposures can be accurately predicted at the population level using a diverse set of indicators and advanced statistical techniques.

Future investigators may wish to consider the following when designing studies to model exposure, especially within the context of a Before and After paradigm:

- KAP and KEF alone should not be relied upon as unbiased predictors of exposure.
- Care should be taken to reduce the impact of exposure-related non-intervention covariates. For example, seasonality. Measured outdoor temperature – a covariate with a potential impact on exposure due to its relation with heating behavior, among other things– was significantly different between the Before and After periods in this study. This confounds the observable relationship between the ACE-1 intervention and exposures. In some instances, confounders may be measured and adjusted for or employed as predictors.
- Covariates with high "importance" in HAP exposure prediction models are not necessarily causal factors for the associated changes in exposure. Instead, they may be good proxies of one or more unmeasured causal factors.
- Ample effort should be expended to ensure questionnaires are culturally appropriate and that participants are able to answer candidly. Structured, timely, and repeated quality control measures – like reviewing participant responses with interviewers and discussing unanticipated points of confusion at the end of each sample day or week – should be used during periods of questionnaire administration.

The prediction methods explored in this chapter have several strengths relative to current models, like the GBD 2010:

- Exposure, the metric of interest to health investigators, was directly predicted.
- Machine learning allowed for the exploration of hidden, more complex, and certain non-linear relationships precluded by the common linear and log-linear regression approaches.
- Super learning leveraged ensemble methods to improve predictive power.
- The exploration of datasets with dozens to hundreds of covariates allowed for a broad investigation of best-case prediction power, and demonstrated the relatively poor performance of relying solely on traditional KEF or even more-advanced KAP-only proxies.
- Despite the inclusion of more covariates than samples, cross-validation methods limited error from over fitting.

The limitations of this work highlight areas where researchers might advance the field substantially:

- A sample size of 60 exposure measurements from a single district in Lao PDR is unlikely to provide the heterogeneity required to produce inferences or models with wide external validity.

- Causal links were not actively explored, and so little insight was provided into factors that may be targeted to reduce exposures. This can be addressed by field work explicitly structured for explanatory modeling or the implementation of methods designed to extract causal inference from observational data (Shmueli 2010).
- Collection of the number of survey indicators and environmental measurements used in this chapter can be resource intensive. Approaches focused on more targeted variable sets employing, where appropriate, lessons learned in this chapter may produce models that are more manageable in the context of resource limitations and less onerous for researchers and participants alike.
- The impact of compliance on prediction power was not analyzed. Future work should assess compliance, perhaps through the use of accelerometers on monitors. This will also help investigators better understand how their measured exposures relate to true exposures.
- Comprehensive longitudinal measurements of both predictors and exposures were not collected in each sampling period. Expanded collection and more in-depth analysis of repeated or long-term measurements would better capture between and within-individual variability. This has been shown to considerably improve the reliability of predictions of HAP-related exposures (McCracken et al. 2009).

The monitoring devices used can have a major impact on cost, comfort, and sample size.

- Gravimetric monitors are increasingly becoming smaller, quieter, longer-lasting, and more integrated. This may have the effect of increasing compliance and reducing error from human handling. Academics and practitioners should stay informed as devices like the Enhanced Children's MicroPEM (Chartier 2015) and Ultrasonic Personal Aerosol Sampler (Volckens et al. 2017) receive field validation.
- In general, filter media are ill-suited for the rugged environments that typify high-KAP settings. Filter media are also expensive—about $10 per sample after analysis. Use of data from media-free continuous monitors may reduce cost, complexity, and human-related error.[23] Media aside, gravimetric monitors are also more expensive than many continuous monitors now on the market.[24]
- Continuous monitors provide the added benefit of finer temporal resolution, which can increase data and allow more-acute exposure events to be modeled. They also require calibration, typically via co-location with a gold-standard monitor. However, a small number of gold standard samples can be used to calibrate many continuous monitors.
- The use of continuous monitors stands to have two major and opposite impacts on precision: increased sample sizes allowed by lower costs, greater temporal

[23] If field teams are not trained properly, sample loss from human error may also be high.

[24] This may soon change. Developers claim to be able to produce the Ultrasonic Personal Aerosol Sampler for $150 in components. It is unclear how much it will ultimately cost.

resolution, and, possibly, decreased sample loss will help improve precision, while the indirect nature of continuous measurement principles (i.e. that which requires calibration) and signal noise present in many of the lower-cost devices will reduce it. These factors should be accounted for in sample size calculations.

Chapter 4

Bridging a dumb gap: smart smoke detectors as a tool for consumer and regulatory $PM_{2.5}$ monitoring

4.1 Introduction

Chapter 1 describes a vast health threat from $PM_{2.5}$. Despite the magnitude and urgency of this threat, current resources available for monitoring actual human $PM_{2.5}$ exposures are limited. As discussed in Chapter 3, reliable gold standard $PM_{2.5}$ monitors are expensive – often priced in the thousands to tens of thousands of dollars (Lewis and Edwards 2016). High costs present a barrier for scientists and regulators wishing to assess exposures across any considerable spatial and temporal scale, and prevent access to the public. Large monitoring networks and related data repositories exist, but the most common and accessible of these are regulatory networks comprised of expensive specialized devices that require frequent maintenance by skilled professionals and provide data at a rate of one measurement per day or hour (Sheppard et al. 2012). While reliable at the point of measurement, sites tend to be sparsely distributed, limiting spatial resolution. These limitations hinder efforts to assess human exposures, as $PM_{2.5}$ concentrations are highly heterogeneous in both space and time (Bell et al. 2011; Chow et al. 2002; Kloog et al. 2011; Levy et al. 2014). This is especially true in developing nation settings where monitoring networks are often poorly developed or non-existent.

As the name implies, regulatory networks are managed by government agencies rather than consumers or grass roots organizations, and so have been designed to respond to the needs of central government (Sheppard et al. 2012). They focus on outdoor environments despite estimates that people spend upwards of 90% of their time indoors (Jenkins et al. 1992), and employ criteria for monitor placement that can give preference based on proximity to sources rather than only the representativeness of population exposures. The US EPA State and Local Ambient Monitoring Station network is a good example. Among the world's most comprehensive, the SLAMS network operates with approximately 2,100 outdoor sites across the country (United States Environmental Protection Agency 2016). About 1,000 of these

sites measure $PM_{2.5}$ and just over 800 do so for the explicit purpose of population exposure— or about one site per every 350,000 Americans. Reliance upon networks like this engenders exposure classification error that hinders risk science and mitigation efforts (Jerrett et al. 2005b; Lioy and Smith 2013; Sheppard et al. 2012).

Methods have been developed to geospatially resolve concentrations between central monitors, including land use regression and kriging. These methods can reduce error, but have considerable flaws. Kriging-based interpolation fails to account for sources in between monitors like commercial zones or highways. LUR is more accurate at assessing relative concentration differences between sites, but generally provides low temporal resolution (Kloog et al. 2011). Overall, kriging and LUR methods have been shown to produce bias when used to estimate exposures for the calculation of health effects, most notably in acute outcomes (Alexeeff et al. 2015; Hoek et al. 2008). Estimates can be further honed using remote-sensing satellite techniques to resolve complex surface patterns in air pollution and provide insight in regions that lack monitoring networks (van Donkelaar et al. 2014). This is particularly useful in lower resource settings that lack robust monitoring networks. However, the spatial resolution of current satellite tools is functionally limited to 1km x 1km, which can overlook exposure-relevant hyperlocal variations in outdoor concentrations like those near roadways, construction, and industry (Jerrett et al. 2005a).

Perhaps most importantly, the aforementioned methods are limited to outdoor ambient assessments, which may be unrepresentative or differentially representative of indoor concentrations depending upon a number of local factors (e.g., building characteristics, heating and ventilation culture, $PM_{2.5}$ size distribution, etc.) (Hänninen et al. 2011; Meier et al. 2015). This is particularly relevant in regions with significant indoor sources (Sagar et al. 2016) and, conversely, in regions with high outdoor concentrations and low building infiltration rates (Zhou et al. 2016). For example, recent literature indicates that the government in Ulaanbaatar operates a network of four fixed-site tapered oscillating microbalance $PM_{2.5}$ monitors (Allen et al. 2013).[25] It does not appear that any routine population-wide indoor monitoring is performed. Yet, as shown in Chapter 2, the heavy use of coal heating in winter produces outdoor concentrations that may substantially differ from indoor concentrations. The 2014 average wintertime indoor concentration estimated (in Chapter 2) in non-smoking apartments with clean (at the point of use) indoor heating sources illustrates this point, calculated at about half the value of population-weighted average outdoor wintertime concentrations.[26]

No routine $PM_{2.5}$ monitoring appears to occur in Lao PDR (Asian Development Bank and Clean Air Initiative for Asian Cities Center 2006; Morawska et al. 2011)—which itself is

[25] Though the situation is changing rapidly.

[26] This is a function of the assumption that indoor air quality in apartments using heating sources that are clean at the point of use are moderated by infiltration of outdoor pollution at a ratio of 0.53 (outdoors to indoors) in winter. This value was informed by relevant literature, and is discussed in more detail in Chapter 2 and Appendix B.

illustrative of the disproportionate measurement need in developing nations. Were an outdoor-oriented network like those in the US or Ulaanbaatar to be constructed in the region in which the measurements in Chapter 3 were conducted, indoor concentrations would likely be underestimated by about an order of magnitude, and exposures, by about 50 - 100% or more.[27]

The inadequacies of current large scale $PM_{2.5}$ monitoring can be tied to the expense and complexity of available air quality monitors. Proof of this can be found in Appendix D of Code of Federal Regulations 40 Part 58 (US National Archives and Records Administration 2016), the federal code that governs regulatory air quality monitoring networks in the US—which are among the most comprehensive in the world. This code specifically acknowledges that the "optimum size of a particular network involves trade-offs among data needs and available resources." The high costs of reliable $PM_{2.5}$ monitors also affect the ability of the public to individually control and act upon air quality data. Citizens wishing to understand air pollution information must rely on the transparency, publicity, and timeliness of monitoring agencies like the US EPA, which has no specific reporting timelines (US National Archives and Records Administration 2016) and suffers from the flaws discussed above.

Low-cost sensors designed specifically to monitor and communicate both outdoor and indoor air quality are becoming increasingly popular. However, availability to consumers is limited to various crowd-funding campaigns, internet promotions, and regional grass-roots monitoring campaigns, (e.g., Air Visual 2016; Purple Air 2016; Speck 2017). Major penetration into neighborhoods, homes, or work places has not yet been shown. Perhaps more importantly, low-cost sensors run a wide gamut of quality with limited laboratory or field validation (Lewis and Edwards 2016). Hobbyists and researchers have, increasingly, been assembling and testing their own (Austin et al. 2015; Holstius et al. 2014; Kelly et al. 2017), but this requires a great deal of skill and knowledge not accessible to most.

<u>The case for smart smoke detector based $PM_{2.5}$ monitoring</u>

Overall, current $PM_{2.5}$ monitoring networks are inadequate for the provision of actionable, spatially and temporally resolved exposure assessments, especially in developing countries. An emerging category of home electronics, smart smoke detectors (SSD), may offer a partial solution to the inadequacies of current $PM_{2.5}$ monitoring networks. "Smart" devices use processing and network capabilities to send and receive data between each other and the cloud. SSD simply perform the function of a standard smoke detector with added "smart" capabilities. Standard optical smoke detectors use an optical sensor to measure the amount of light scattered when a beam is sent through a chamber of air. When increased scattering levels indicate a high level of smoke, an alarm sounds. While smoke detectors employ this mechanism in a binary system, the scattering itself varies with the amount of

[27] Based on Chapter 3 ambient, KAP, and exposure measurements.

$PM_{2.5}$ of a given suspended aerosol (Hinds 1999). Past researchers have capitalized on this relationship to modify commercially available smoke detectors to cheaply and reliably measure $PM_{2.5}$ concentrations (Edwards et al. 2006; Litton et al. 2004). Modified smoke detectors have since been used in a wide range of published studies around the world (Balakrishnan et al. 2013; Chowdhury et al. 2013; Northcross et al. 2012). To date and to my knowledge little, if any, [28] work has been published characterizing the stock onboard sensors of *smart* smoke detectors to do the same. SSD stand to offer the same sensing functionality *alongside* processing and internet capabilities without the need to make costly, tedious hardware modifications to the commercially available unit. Onboard processors could facilitate data cleaning and calibration while internet access could allow data to be uploaded to offsite databases, remotely calibrated, and shared. Internet connectivity could potentially be used to upload PM monitoring abilities to any SSD already in place, as well. Importantly, SSD exist primarily *indoors* – where people spend most of their time.

This paper discusses an innovative concept for bringing affordable, reliable $PM_{2.5}$ measurement to the public while filling gaps in government-operated monitoring networks: smart smoke detectors as smart $PM_{2.5}$ sensors.

A brief primer on smoke detector prevalence

The primary purpose of smoke detectors – detecting fires – has already proved highly beneficial in the public health context. A study of fatal fires in the US found that smoke detectors produced a 61% reduction in related mortality risk (Marshall et al. 1998). Indeed, because of this and their now-low costs, smoke detectors are nearly ubiquitous in many nations. For example, in a recent survey of US homes (Ballesteros and Kresnow 2007) about 95% reported having at least one smoke detector – about 93% of those reported at least one operational smoke detector per floor. The prevalence of SSD is much lower, but still considerable in the context of the size and density of typical $PM_{2.5}$ monitoring networks. According to a recent analysis, about 6% of Americans (20 million people) own a smart smoke/carbon monoxide detector (The Harris Poll 2015a) and 27% of US internet users are interested in purchasing one (The Harris Poll 2015b). If SSD can be remotely programmed to reliably measure $PM_{2.5}$, a large number of potential monitors exist in-situ. It should be noted that the distribution of smart smoke detectors may be clustered around areas where wealthier, early technology adopters reside. While not necessarily representative of an entire population, SSD coverage on the order of millions of people is a great advantage.

[28] Some smart smoke detector companies, like Birdi (Birdi 2016), reportedly have or are experimenting with particle monitoring, but it is unclear whether any employ the original smoke detector sensor itself. At least one Nest Labs, Inc. patent (Mittleman et al. 2014) even suggests the potential use of an *additional* light source as a particle counter or indicator of general air quality. $PM_{2.5}$ characterization literature of these devices is difficult to come by.

Smoke detector prevalence is lower in developing nations, particularly where major indoor sources like cookstoves, which could cause daily false alarms, are common. Smart smoke detectors, which are a relatively new technology and substantially more expensive than standard smoke detectors, are likely to be even less common. Still, the low price point of smart smoke detectors may justify construction of a network of SSD in these regions, depending on electricity and internet access. The combination of fire risk reduction and air pollution monitoring may even allow such a network to pay for itself in countries with high annual fire and pollution related losses. A study conducted in the US in the 1990's (Haddix et al. 2001) found that every $1 USD spent on a smoke detector giveaway program provided over $28 USD in health-based savings. The smoke detectors used in that study cost (~ $50) approximately half of the off-the-shelf price of popular smart smoke detectors today, and so the ratio of cost to fire-risk-benefits of a similar smart smoke detector intervention may be higher. However, added air $PM_{2.5}$ monitoring capabilities may provide financial value in the form of reductions in other risks.

Laboratory characterization and proof of concept

As previously discussed, common smoke detectors modified to log real-time data from their internal sensors have been shown to track strongly with research-grade continuous particulate monitors in laboratory and field tests ($r^2 > 0.99$) (Chowdhury et al. 2007; Edwards et al. 2006; Litton et al. 2004). The addition of "smart" abilities to air quality monitors has been shown advantageous in the processing and transmission of real-time data (Al-Ali et al. 2010; Ding and Song 2016; Postolache et al. 2009; Ray 2016; Soldo et al. 2012). This Chapter explores the intersection of these concepts through the laboratory examination of a well-known smart smoke detector: the Nest Protect 2nd Gen Smoke + Carbon Monoxide Alarm (Part No. S3000BWES, Nest Labs, Palo Alto, CA, USA). This device can be purchased for about $100. Proof of concept for an SSD-based $PM_{2.5}$ monitor is sought by reverse engineering and characterizing the device's key hardware. The ability of the onboard smoke detection sensor to provide reliable measurements of $PM_{2.5}$ across a range of concentrations relevant to both developed and developing nation contexts is explored through laboratory analysis of the sensor's response to dynamic fluctuations in $PM_{2.5}$ concentrations alongside an industry standard monitor.

4.2 Methods

Reverse engineering the smart smoke detector

This section describes selected aspects and results of the Nest Protect reverse engineering process. No guidance in this process was received from the manufacturer, and units were purchased off the shelf. Speakers were disconnected to avoid alarm sounds. Plastic housing was removed to reveal the optical light scatter chamber, circuitry, and various electronic components (Figure 4.1). Components were indexed and researched via manufacturer

specification sheets where possible. The two light emitting diodes (LED) of the optical chamber (infrared [IR] and blue; Figure 4.1) were connected to a LeCroy DDA-125 oscilloscope (Teledyne LeCroy, Inc., Chestnut Ridge, NY, USA), and LED firing rates and voltages were characterized. Component leads were explored with the oscilloscope until those adequate for receiving output from the single photodiode sensor of the optical chamber, which intercepts light from both the IR and blue LEDs, were identified.

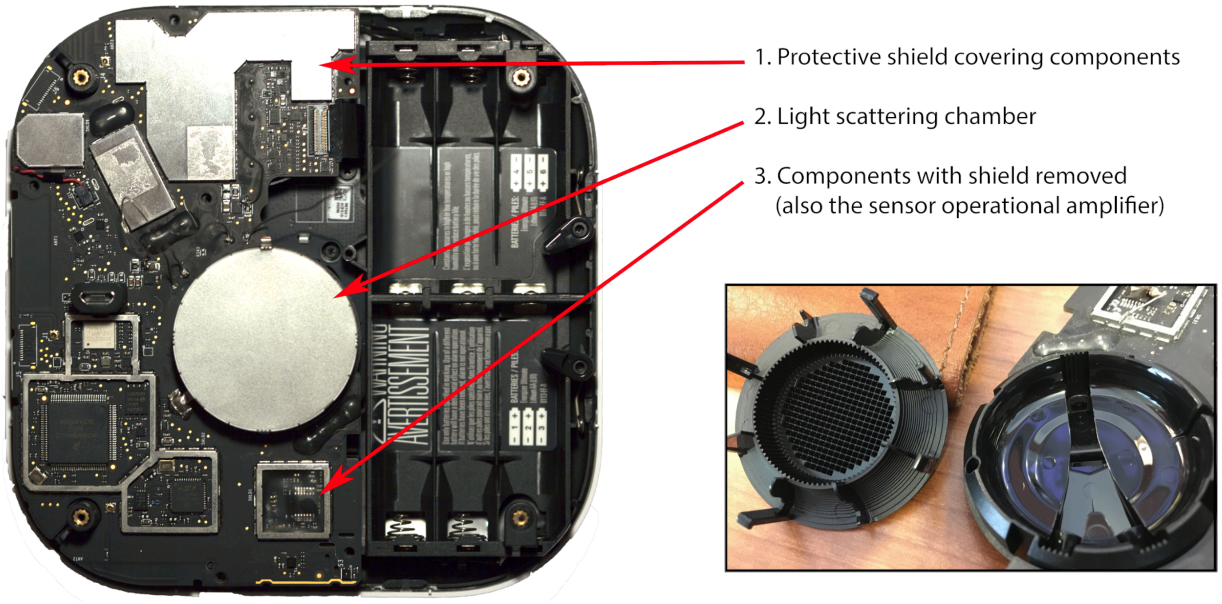

Figure 4.1. (Left) Nest protect without back-plate. Protective shields have been removed from the lower part of the device, revealing several onboard components. (Bottom Right) Light scattering chamber (right) with lid (left) removed. Scattering angles of both LEDs were measured at 45 degrees forward.

Thresholds of 1.3V and 3.1V were measured for the IR and blue LED, respectively. Each LED was positioned to allow forward scattering at about 45 degrees into the photodiode. Exact wavelengths of the LEDs were not measured. LED are controlled by unique pins on the a Freescale MKL16Z128VLH4 microcontroller (Freescale Semiconductor, Inc., Austin, TX, USA), which is a 48 MHz ARM Cortex-M0+ core microcontroller with 128 KB flash memory and 16 KB static random-access memory (SRAM) (Freescale Semiconductor 2014).

Raw output from the optical sensor is passed through an operational amplifier (STMicroelectronics, Geneva, Switzerland, part no. TSV634A) which amplifies the signal and adjusts it for a sloping baseline before passing it on to another component. The ADC for converting optical sensor output into digital signal for processing and use was not identified. However, at least two onboard chips, the MKL16Z128VLH4 and MK24FN1M0VLL12 (Freescale), have at least one 16-bit successive approximation register ADC (Freescale Semiconductor 2014; NXP Semiconductors 2016), and so one of these chips may serve this purpose.

LEDs fire at a rate of 0.1 Hz for the IR LED and approximately 0.005 Hz for the blue LED. When in an alarm state (i.e. high smoke concentrations), firing rate increases to 0.5 Hz for both LEDs. The IR LED fires at 250 mA for approximately 300 µs (Figure 4.2). The blue LED fires at 68 mA between 5ms to 6ms after the beginning of the IR LED pulse, and lasts approximately 300 µs. Firing rate was the only observed difference in LED function between alarm and non-alarm states.

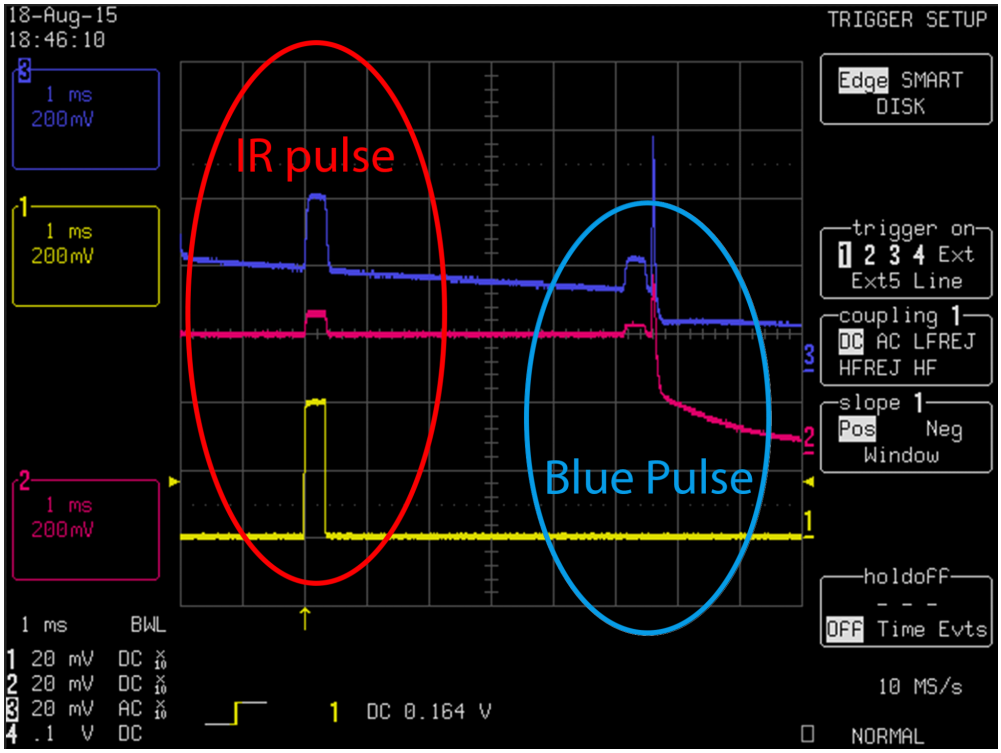

Figure 4.2. Oscilloscope trace. The yellow line (bottom) represents the signal of the IR LED firing lead. Magenta (middle) represents optical sensor output as received from Pin 7 (baseline-slope corrected) of the operational amplifier. Periwinkle (top) represents unadjusted optical sensor output. X-axis ticks = 1 ms. Y-axis ticks = 200 mV.

Multiple microprocessors and memory chips were identified. Of particular note is the Freescale MK24FN1M0VLL12, a 32-bit 120 MHz ARM cortex M4 microcontroller with 1 MB program flash memory and 256 KB of embedded SRAM (NXP Semiconductors 2016). In addition, a Micron N25Q 25Q128A 11E40 memory chip (Micron Technology, Inc., Boise, ID, USA) that appeared to be connected with the MK24FN1M0Vll12 microcontroller provides 128 Mb of 108 MHz serial NOR flash memory (Micron Technologies 2014, 2015). Nest product literatures confirms temperature and humidity sensing capabilities (Nest Labs 2015b, 2015a).

Multiple communication microchips and antennae were confirmed. A Dialog DA14580 (Dialog Semiconductor PLC, Reading, UK) and adjacent antenna provides low-power Bluetooth V4.2 (Dialog Semiconductor 2016) for mid-range communication. An Ember EM3581 ZigBee system on a chip (Silicon Laboratories, Inc., Sunnyvale, CA, USA) provides a 2.4 GHz IEEE 802.15.4-2003 transceiver & lower MAC for low-speed area communication and mesh networking capability. Standard Wi-Fi capabilities are confirmed in product literature (Nest Labs 2015b).

Co-opting the smart smoke detector optical light scatter chamber

This analysis explored the $PM_{2.5}$ concentration response of the Nest Protect. Factory set LED pulses were not modified. No application program interface (API) exists for accessing raw optical sensor data. Instead, sensor data were retrieved by connecting relevant leads from the IR LED, blue LED, and the photo-receptor operational amplifier to a custom circuit board interfaced with a Raspberry Pi (Raspberry Pi 3 model B, Raspberry Pi Foundation, Cambridge, UK running Ubuntu 16.04.1) through its general-purpose input/output (GPIO) pins. Three such "Nest Devices" were developed for testing. Specifically, the LED leads were connected in such a way that a brief digital "high" signal would be passed to one Raspberry Pi GPIO pin whenever the IR LED fired, and another GPIO pin when the blue LED fired. The photodiode sensor output lead was buffered with a 1x gain and connected by shielded wiring to a 13-bit differential analog-to-digital converter (Microchip MCP3304, Microchip Technology Inc., Chandler, Arizona, USA) to convert the raw sensor output voltage into a digital signal readable by the Raspberry Pi. A 13-bit ADC was chosen to imitate the minimum expected resolution of the Nest Protect's onboard components as inferred during teardown. The ADC differential input was connected via the wire shielding and post-signal amplifier to a metal ground on the Nest Protect. This allowed for removal of electrical interference experienced within this system (from sensor, through amplifier, to ADC). Rather than expending the time and effort to connect to the onboard temperature and humidity sensor, an SHT15 temperature and relative humidity (RH) sensor (version 3, Sensirion AG, Staefa, Switzerland) was affixed to each Nest Device and interfaced with the Raspberry Pi.

A custom program was written in C (Bell Labs, Murray Hill, NJ, USA) to read and log optical sensor data. Specifically, five rapid, successive, differential readings at 12-bit resolution were taken by the ADC from the optical sensor output when either of the LEDs fired. Readings were averaged, labeled according to LED type and datetime, and logged. Timing of the programmatic samples in relation to the signal was confirmed by oscilloscope trace using a Tektronix THS730A (Tektronix, Inc., Wilsonville, OR, USA). Five-point averages of the digital ADC output were translated back into voltage using Equation 4.1.

$$Voltage = \frac{Average\ ADC\ Value}{Max\ ADC\ Value} \cdot V_{ref} \quad (4.1)$$

V_{ref} is the reference voltage of 3.3V, and the Max ADC Value is calculated as $2^{(\text{ADC resolution in bits})} - 1$, which in this case is 2^{12} -1 or 4095 bits. Under normal conditions, the C program checked for an IR and blue LED fire every 9 and 180 seconds, respectively. If an LED fire was registered, the ADC would take a reading and the sample would be logged. As discovered during teardown, LED firing rates differ between alarm and non-alarm states. The transition between the non-alarm and alarm states appeared to be controlled by the photodiode signal level, with higher signals eliciting an alarm state. The thresholds at which this transition occurred were roughly approximated experimentally for the three devices using wood-stick match smoke – about 0.48 V for IR signals and about 0.40 V for blue signals. Changes in LED firing rates were accounted for programmatically by mimicking the Nest Protect's internal threshold system: if an ADC sample taken under standard program conditions exceeded the approximated thresholds, the C program would begin checking for fires from each LED at 1 Hz until values again dropped below the threshold. To ensure minimal interference from Linux operating system processes – which can supersede and thus lag the timing of user programs– the C program was run with maximum system priority.

A program was written in Python (version 2.7.12, Python Software Foundation, Delaware, USA) to read and log datetime, temperature and RH readings from the SHT15 sensor, and temperature readings from the Raspberry Pi onboard system temperature sensor at a rate of 0.1 Hz. Upon program startup, this 0.1 Hz cycle was set to be triggered by the first subsequent IR LED firing in order to approximately sync environmental condition readings with optical sensor readings. To ensure minimal timing interference from Linux operating system processes but also avoid interfering with the more time sensitive optical sensor program, the Python program was run with the second highest system priority.

Laboratory validation

The response characteristics of optical sensor systems can vary with changes in temperature (Edwards et al. 2006). Nest Devices were tested for the influence of temperature on LED and optical sensor operation. A temperature test was performed by placing devices inside a sealed 38 L (liter) bin at effectively zero-levels of $PM_{2.5}$, placing that bin inside of an incubator, and varying incubator temperatures between 0 and 48 °C (measured with the aforementioned SHT15 sensors) over the course of about 330 minutes.[29] $PM_{2.5}$ was removed from the bin by pump and HEPA filter at 5 L/minute for about 35 minutes. Negligibly low $PM_{2.5}$ concentrations were confirmed by connecting the bin outlet to a DustTrak for 10 minutes after the initial zeroing process (non-gravimetrically adjusted mean: 0.000 $\mu g/m_3$).

[29] Characterization of a sensor's temperature-response by placing it in a near-particle-free environment under varying temperatures is a method commonly applied in the Household Energy, Climate, and Health (HECH) research group at the University of California, Berkeley—e.g. Edwards et al. 2006 – and Berkeley Air Monitoring Group.

Nest Device sensitivities to variations in $PM_{2.5}$ concentration were tested in comparison to two DustTraks (DustTrak II, model 8530, TSI, Inc., Shoreview, MN, USA). A 55-gallon stainless steel barrel was used as a concentration chamber. [30] The chamber was outfitted with a sealable lid with mixing fan. Nest Devices were placed inside of the chamber, while DustTraks were connected by tubing to individual sampling outlets (Figure 4.3). Particulate matter was produced in a fume hood using small wood pieces and a combustion source, and pulled into the chamber. Generally, testing involved bringing the $PM_{2.5}$ concentration of the chamber from 0 $\mu g/m^3$ to a predetermined level, sealing the chamber, allowing the concentration to settle for between 30-45 minutes, venting the chamber for approximately 30 minutes with HEPA filtered inlet air, and then repeating with a new $PM_{2.5}$ concentration. Testing target concentrations included approximately 10 mg/m^3, 1.5 mg/m^3, 0.75 mg/m3, 0.075 $mg/^3$, and 0.015 mg/m^3. Particulate production primarily consisted of wooden doll-house shingles [31] combusted with an electric soldering iron, but did briefly incorporate a kerosene lighter and small wooden sticks during the highest concentration test. An hour of < 0.001 $\mu g/m^3$ readings was taken after the last test concentration.

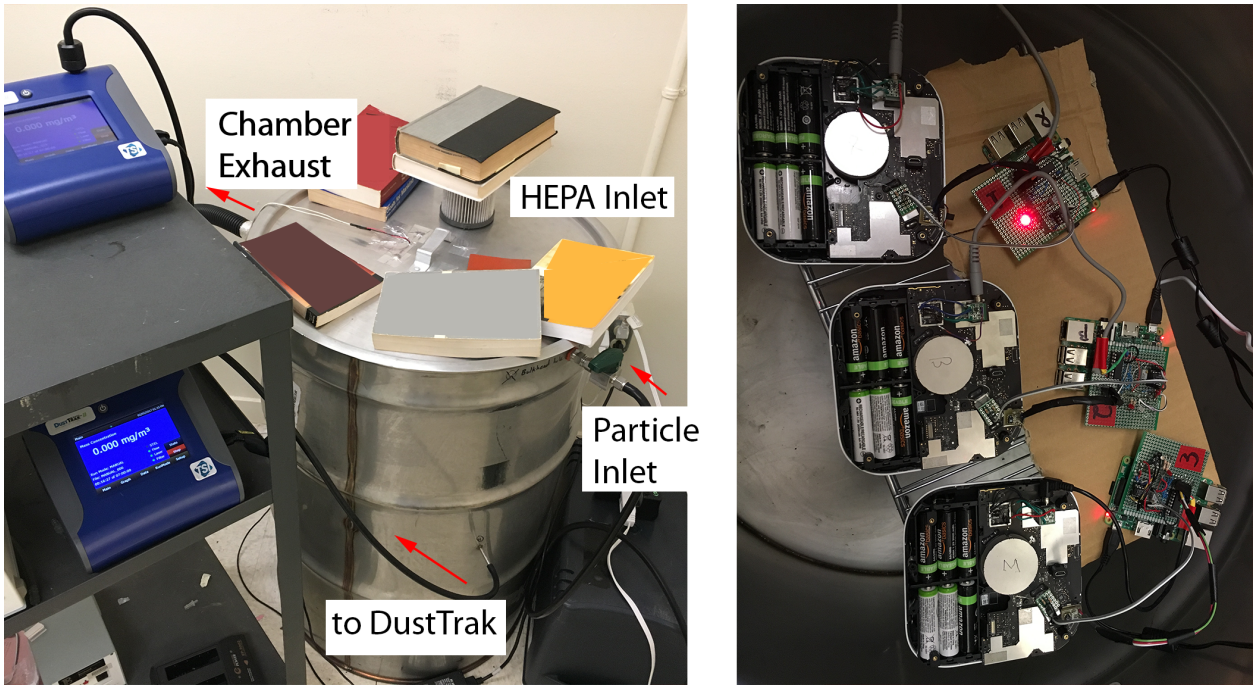

Figure 4.3. Calibration chamber used for $PM_{2.5}$ testing (left), and placement of Nest Devices as tested (right). Books are used as ballast to ensure a tight seal on the lid.

[30] Variations of the chamber test method are commonly applied in the HECH research group – e.g. Chowdhury et al. 2007 – and Berkeley Air Monitoring Group.
[31] Chosen for their uniformity.

For testing Nest Device $PM_{2.5}$ sensitivities, DustTraks were set at a Log Interval of 1 second. Each DustTrak was loaded with an inline 37 mm Teflon filter with support ring (Pall Corporation, Port Washington NY, USA), zero-calibrated according to manufacturer recommendations, flow calibrated to 3.0 LPM (liters per minute; 0.050 liters per second [LPS]) using a Bios Defender 510 volumetric flow meter (Mesa Labs, Bulter, NJ, USA), and then fitted with a DustTrak 2.5 µm-cut impactor. Filters were weighed on a Mettler Toledo XP2U balance (Mettler Toledo, Columbus, OH, USA) with 0.0001 mg precision. Filters were conditioned for at least 48 hours in a temperature and RH controlled room prior to weighing, and stored at -20 °C between sampling and post-sample weighing. Notably, the chamber mixing fan was not turned until about 2 minutes into testing, and DustTraks were not fitted with impactors for approximately the first four minutes of testing, or about 1% of total test time. These first minutes of sampling were at near-zero PM concentrations, and so are not expected to affect overall device response metrics or the size distribution of sampled aerosols. Due to the late addition of impactors, pre-sample zero and flow calibrations were performed without them. Post-sample flow testing with impactors attached produced a flow of 3.0 LPM on each device and so no major impact on flow rate is suspected.

Signal calibration

The relationship between the $PM_{2.5}$ mass concentrations in the air and the amount of light scattered in the optical chamber is linear for a given aerosol (Hinds 1999). However, this relationship depends upon characteristics of the aerosol of interest – like distributions of size, shape, and refractive index. For this reason, sensing methods require calibration against a gold standard (Wallace et al. 2011). Individual calibration factors for each DustTrak were produced by comparison with in-line gravimetric measurements using Equation 4.2.

$$k_i = \frac{\sum_{t_{0,i}}^{t_{end,i}} PM_{t,i}}{\left(\frac{M_i}{\sum_{t_{0,i}}^{t_{end,i}} \left(F_i \cdot 0.001\ m^3/_L\right)}\right)} \tag{4.2}$$

Where k_i is the gravimetric calibration factor for DustTrak "i", $t_{o,i}$ is the beginning of the sampling period for the sample taken with DustTrak "i", $t_{end,\,i}$ is the end of the sampling period for DustTrak "i", $PM_{t,i}$ is the continuous DustTrak reading (mg/m^3) at time "t" for DustTrak "i", "M_i" is the mass (mg) accrued on the inline filter of DustTrak "i", and "F_i" is the flow rate (LPS) for DustTrak "i" which is assumed constant over the measurement period at 0.050 LPS.

Nest Device voltage readings were cleaned, temperature-adjusted, and calibrated to the mass-adjusted DustTrak values. A small number of low-value outliers were discovered in raw Nest Device output. They may have been due to slight lags in the highly time-sensitive

ADC-read program caused by other top-priority operating system processes occurring simultaneously. This could have caused the ADC to read partially or wholly from a signal period unrelated to LED firings. A conservative decision rule for excluding outliers was set at a difference between a reading and its previous reading of ≥ 5 times the preceding 6-point rolling standard deviation, calculated as the standard deviation of the 6 previous readings.

Outlier-adjusted voltages from Nest Devices during temperature tests were linearly regressed against temperature readings to produce device and LED specific temperature calibrations. Where significant, these calibrations were applied to outlier-adjusted $PM_{2.5}$ test output prior to analysis. Zero-degree LOESS was used to smooth temperature-adjusted LED signals as a function of time. Spans of 0.0025 and 0.01 were used for IR LED LOESS and blue LED LOESS, respectively. $PM_{2.5}$ concentration responses for Nest Devices were obtained from a linear model produced by regressing average gravimetrically adjusted DustTrak readings against modeling smoothed voltage output.

Noise in optical scattering measurements (voltage) can be introduced by Rayleigh scattering by gas molecules and instrument electronics (Friedlander 2000). In the case of the Nest Devices, noise is likely governed primarily by electronics. For a given aerosol, voltage noise can be converted to mass concentration-equivalent noise by applying the aforementioned calibration techniques to the signal prior to calculating noise. Signal noise was calculated during the post-$PM_{2.5}$ testing zero period and defined as one standard deviation (SD) in the LOESS-smoothed voltage. This definition (1 SD) was chosen for its use in previous low-cost sensor analyses (Litton et al. 2004). Mass concentration-equivalent noise was calculated as one standard deviation in the LOESS-smoothed, gravimetrically corrected signal during the same period. The lower limit of detection (LOD) was calculated as 3 times the noise during the same zero period, a method used in other low-cost sensor analyses (Austin et al. 2015; Wang et al. 2015) and similar to generally popular methods (Specker 1968; Wallace et al. 2011).

4.3 Results

Laboratory validation: temperature test

Strong and significant temperature responses were observed for the IR LED response in all Nest Devices (Table 4.1, Figure 4.4). Only one Nest Device, “Epsilon”, exhibited a significant blue LED response ($p < 0.05$).

Table 4.1. Regression of Outlier-Adjusted Raw Nest Device Response against Temperature During the Temperature Calibration Procedure

Device	LED Type	Adj-R2	Intercept (V)	Slope (V / °C) [1]	Model p-value
Alpha	IR	0.932	0.467	-5.7 E-04	< 2.2E-16
Beta	IR	0.947	0.483	-7.1 E-04	< 2.2E-16
Epsilon	IR	0.932	0.444	-4.3 E-04	< 2.2E-16
Alpha	Blue	0.001	0.367	7.8 E-06	0.3057
Beta	Blue	-0.003	0.373	6.2 E-06	0.4093
Epsilon	Blue	0.457	0.370	5.9 E-05	3.50 E-14

1. Reported to 2 figures, but applied during temperature calibration with 4 figures.

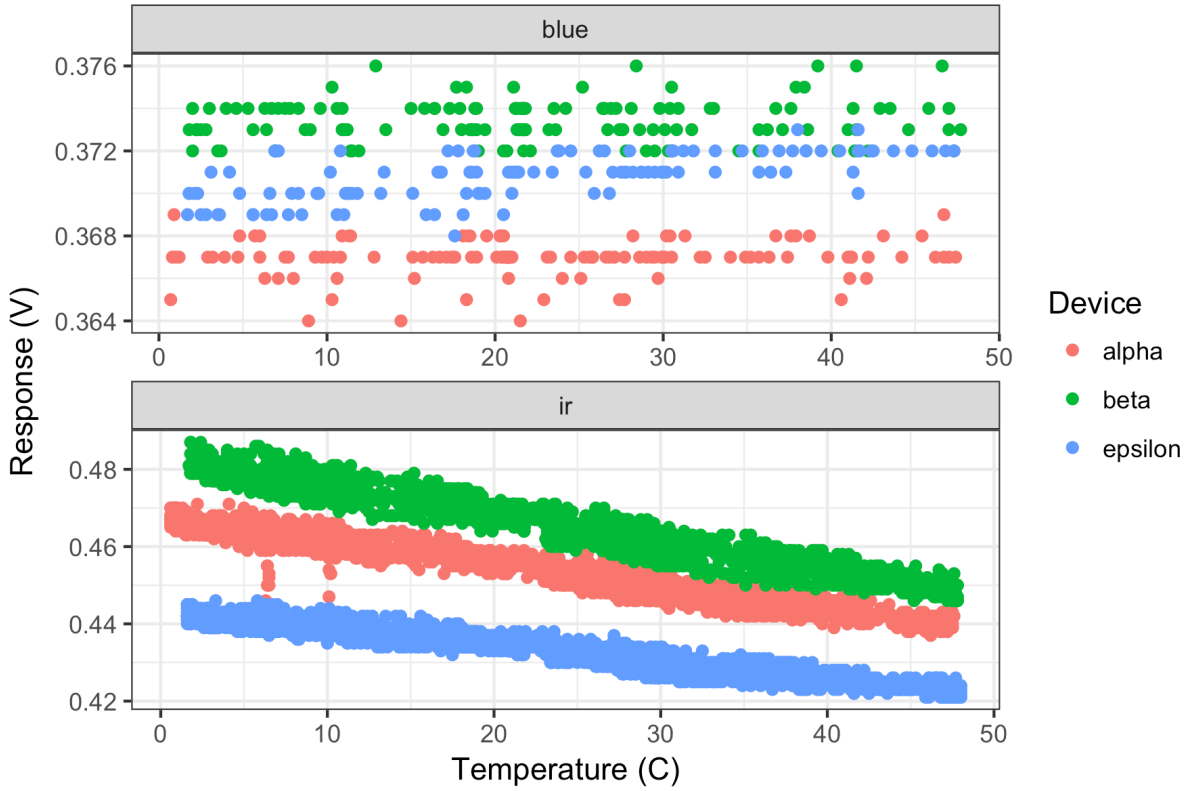

Figure 4.4. Measured temperature responses for each Nest Device, by LED type.

Laboratory validation: $PM_{2.5}$ testing

Comparison of gravimetric measurements against integrated DustTrak output produced calibration coefficients of 4.90 and 4.30 for each of the two monitors (i.e. the two DustTraks estimated concentrations at about 490% and 430% of actual levels). Regression of averaged, gravimetrically adjusted DustTrak output against LOESS-smoothed Nest Device output demonstrated correlations > 0.99 for all devices and both LED types (Table 4.2). Real-time gravimetrically adjusted DustTrak and LOESS-smoothed, temperature & $PM_{2.5}$-adjusted

Nest Device output are shown for the $PM_{2.5}$ testing in Figures 4.5 and 4.6. Noise and LOD values are provided in Table 4.3. Chamber temperatures ranged from 20 – 23 °C. RH was maintained between 53 - 60%.

Table 4.2. Regression of DustTrak II Measurements (Two-Device Average) against LOESS-Smoothed Nest Device Output

Device	LED Type	Adj-R2	Intercept (V)	Slope (V per mg/m^3)	Model p-value
Alpha	IR	0.997	-0.467	0.021	< 2.2E-16
Beta	IR	0.997	-0.482	0.016	< 2.2E-16
Epsilon	IR	0.998	0.442	0.024	< 2.2E-16
Alpha	Blue	0.991	0.372	0.013	< 2.2E-16
Beta	Blue	0.997	0.378	0.020	< 2.2E-16
Epsilon	Blue	0.992	0.372	0.016	< 2.2E-16

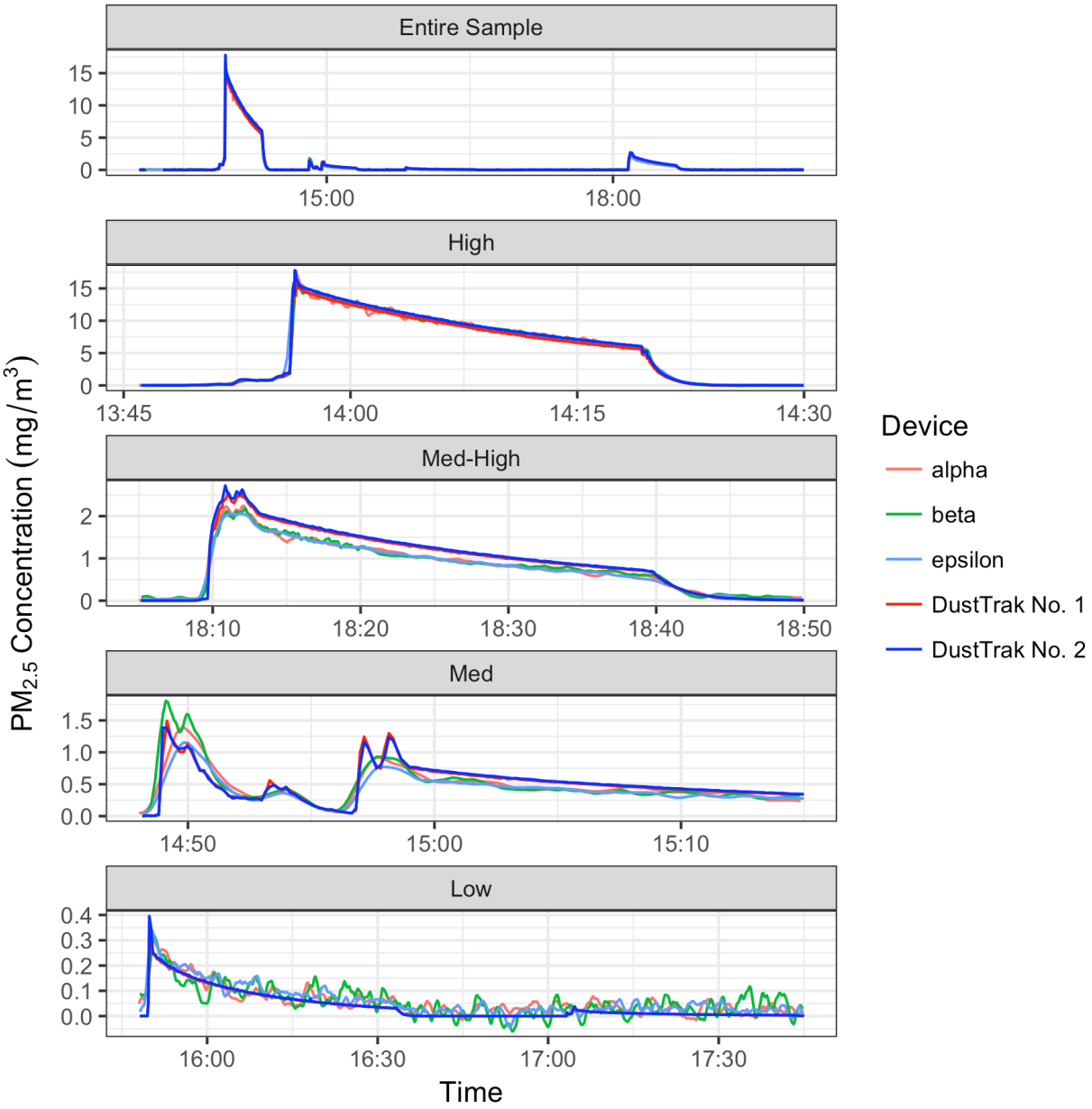

Figure 4.5. Overlaid Nest Protect IR (LOESS-smoothed, $PM_{2.5}$-adjusted) and DustTrak Output, by concentration range, during $PM_{2.5}$ testing. High corresponds to a concentration test range between about 15 – 5 mg/m^3. Med-High corresponds to a test range of about 2.5 – 1.5 mg/m^3. Med corresponds to a test range of about 1.5 – 0.5 mg/m^3. Low corresponds to a test range of about 0.4 – 0 mg/m^3

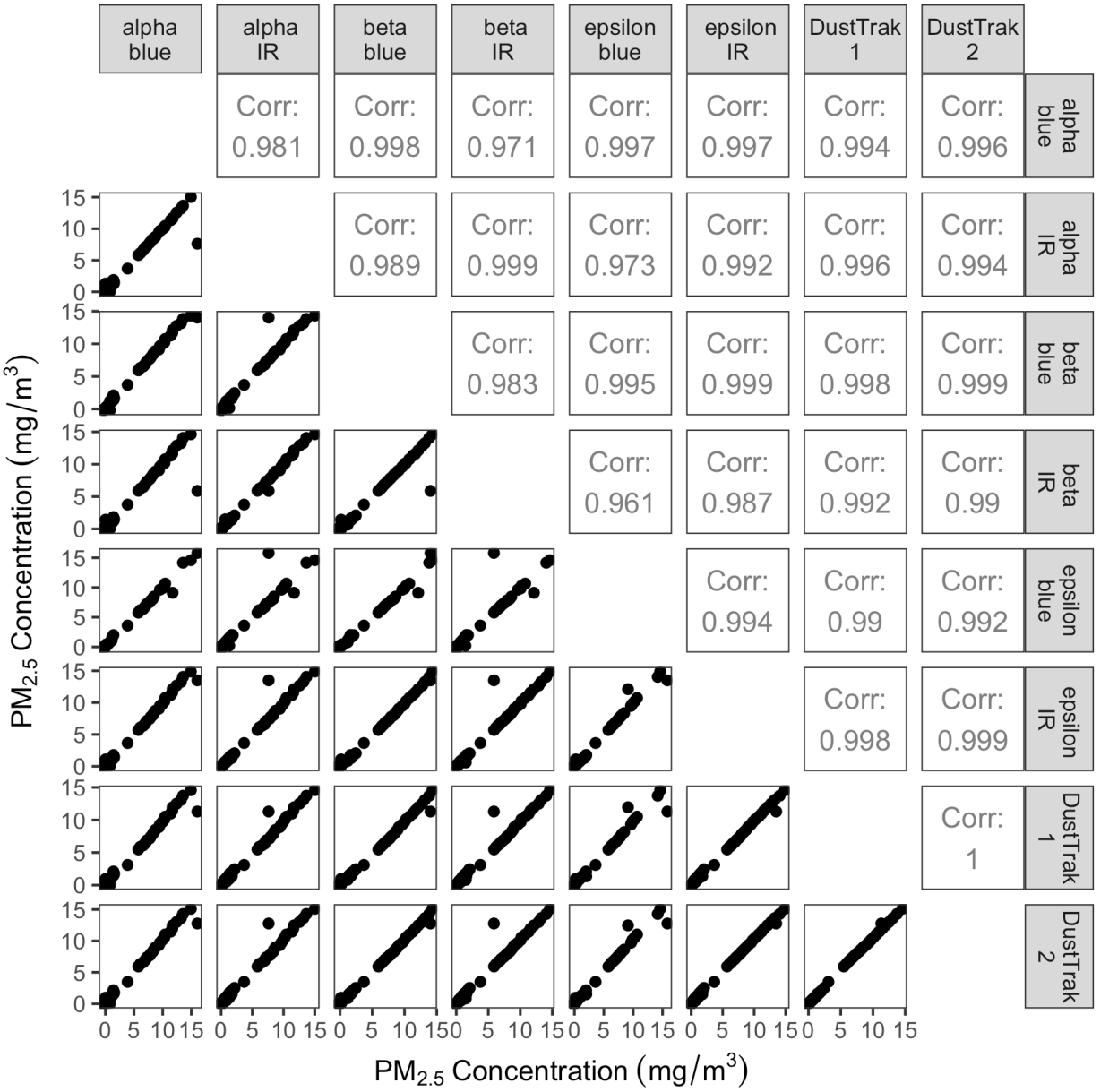

Figure 4.6. Pairwise correlations between gravimetrically adjusted DustTrak and LOESS-smoothed, gravimetrically adjusted Nest Device output during $PM_{2.5}$ testing. Output is averaged by minute to allow comparison between Nest Devices, because individual samples did not exactly overlap at a 1-second temporal resolution.

Table 4.3. Noise and Limit of Detection Values, by Device and LED Type, Determined During Zero Period

Device	LED	Noise (+/- V)	Noise (+/- mg/m3)	LOD (mg/m3)
alpha	IR	4.8 E-04	0.024	0.071
alpha	Blue	1.7 E-04	0.013	0.040
beta	IR	4.9 E-04	0.030	0.091
beta	Blue	1.4 E-04	0.007	0.022
epsilon	IR	5.2 E-04	0.021	0.064
epsilon	Blue	3.9 E-04	0.024	0.071

4.4 Discussion

$PM_{2.5}$ exposure is an important determinant of disease. Recent estimates indicate that 87 % of the global population experiences annual average $PM_{2.5}$ concentrations above the World Health Organization's guideline (0.010 mg/m^3) (Brauer et al. 2016). In 2015, an estimated 7.2 million people died from exposures to ambient and solid-fuel related $PM_{2.5}$ (Institute for Health Metrics & Evaluation 2017), representing over 10% of all annual deaths. Yet $PM_{2.5}$ monitoring networks are inadequate for the provision of actionable, spatially and temporally resolved assessments, especially in developing countries. An increasingly popular and affordable household product, the smart smoke detector, may provide a means of buttressing such networks. This chapter explores key principles required for such a network to function with a laboratory analysis of a specific smart smoke detector: the Nest Protect

Others have examined the benefits of smart air quality monitoring networks. The use of wireless networks – like cellular and mesh networks— to manage outdoor air quality sensors and collect and broadcast real-time data to a wide variety of communities through web interfaces like Google Maps has been explored in several projects, e.g. Al-Ali et al. 2010 and Murty et al. 2008. Methods have been presented for remotely collecting and transmitting outdoor and indoor pollutant data over Wi-Fi in a way that allows for resource-intensive calibration procedures to be performed at a central processing hub (Postolache et al. 2009). More recent efforts have successfully demonstrated that the Internet of Things (IoT) can be used to manage, transmit, store, analyze, and visualize air quality data remotely and in real-time via the cloud (Ray 2016). These studies have all focused on professional or custom-designed low-cost air pollution sensors, but it is probable that the same functions could be achieved with the stock sensors aboard SSD, which hold the unique advantage of an established and growing presence in residential and commercial buildings.

This chapter provides tangible evidence of the plausibility of this concept. Teardown confirmed the presence of hardware capable of reading, calibrating, storing, and communicating $PM_{2.5}$ concentrations. Each Nest Protect comes equipped with at least two microchips capable of converting raw analog sensor data into digital signal interpretable by digital microprocessors. Multiple onboard microchips meet specifications required to process raw signal and apply corrective algorithms. Onboard flash storage may allow data logging, and wireless radios indicate the potential for real-time communication of concentrations or raw voltages and remote software modification.

Among the most interesting components of the Nest Protect is the dual-LED optical chamber, with light sources in both the IR and blue wavelength ranges. The optical chambers of typical smoke detectors and low-cost PM sensors are limited to a single LED, often in the IR spectrum around 0.900 μm (Litton 2002). IR scatters efficiently for moderate to high concentrations and relatively large aerosols, but becomes decreasingly efficient as aerosol diameters drop below the wavelength of the emitted IR light (Friedlander 2000). For particles << 1 μm, blue light is scattered preferentially to red—in fact, this is why the sky is blue! As a result, low-cost IR-based PM sensors are relatively insensitive to low concentrations of particles smaller than about 0.5 μm (Litton et al. 2004). This is of particular relevance to $PM_{2.5}$ sensing. Vehicle exhaust aerosols have been measured with mass median diameter of 0.15 μm and volumetric distributions between 0.01 and 0.1 μm (Flagan and Seinfeld 2012); emissions from household combustion of cooking fuels like firewood, coal, LPG, dung cake, and kerosene have been measured with geometric mean diameters between 0.048 and 0.152 μm (Tiwari et al. 2014); and a study of urban aerosols in a major German city found that the majority of ambient $PM_{2.5}$ fell between 0.03 – 0.3 μm (Birmili et al. 2010) with similar trends observed in other Western cities (Van Dingenen et al. 2004). The inclusion of an additional, blue LED (about 0.475 μm in wavelength) in the Nest Protect optical chamber may provide an increased sensitivity at low concentrations, especially of smaller particles, compared to other low-cost sensors. This should be examined further.

In addition to better mass concentration measurements, the dual-LED architecture of the Nest Protect optical chamber could allow for the determination of particle size and surface area characteristics — capabilities currently unattainable in low-cost sensors. An examination of low-cost sensors demonstrated that simultaneous output from two sensing mechanisms with different characteristic relationships between output and aerosol size distribution can be evaluated against each other to calculate average particle size, total particle surface area, and specific surface area (Litton et al. 2004). This analysis looked specifically at a device with both an IR optical scattering chamber and an ionization chamber (note: recent regulations and restrictions on radioactive materials, among other issues, have made ionization chambers prohibitively difficult to work with for the purposes of low-cost sensing). While that analysis looked at two signals produced using distinctly different mechanisms (light scattering vs. ion absorption), it is possible that the same theory could apply to a dual-LED method like the one employed by the Nest Protect,

provided the relationship between particle size and mass concentration response differs between the two wavelengths. Moreover, their study suggests that concurrently measured output from two optical scatter signals of different calibration characteristics might allow (when used in comparison to each other) improved resolution of the optical scattering sensitivities of both signals. Further research should examine the potential for such improved mass concentration estimation and the calculation of average particle size and surface area properties in dual-LED devices.

Laboratory testing confirmed that Nest Protect optical sensors can be calibrated and used to produce highly reliable estimates of $PM_{2.5}$ mass concentration in relation to each other and to a more-expensive industry standard device ($r^2 \geq 0.99$) at up to 15 mg/m^3. This relationship is in or above the range of laboratory analyses of low-cost PM sensors. A recent comparison of the Shinyei PPD42NS (PPD), Samyoung DSM501A (DSM), and Sharp GP2Y1010AU0F (GPY) with a TSI SidePak AM510 produced linear correlations between r^2 = 0.88 and r^2 = 0.99 when tested in the presence of 0 to 1 mg/m^3 of sub-micron combustion aerosols (Wang et al. 2015). An examination of the PPD alongside a TSI Aerosol Particle Sizer produced correlations around r^2 = 0.99 at 0.001 to 0.050 mg/m^3 of polystyrene particles between 1 - 3 µm in diameter (Austin et al. 2015). Analyses of the UC Berkeley Particle and Temperature Sensor (UCB PATS) – a modified standard smoke detector that employed both optical and ionization sensing chambers— showed correlations with DustTrak and gravimetric measurements typically of $r^2 > 0.99$ when tested in the presence of oleic acid and wood smoke aerosols at concentration ranges similar to that of the present examination (Chowdhury et al. 2007; Edwards et al. 2006).

The Nest Device setup as tested incurred considerable noise, more so for the IR LED type than the blue LED type. The importance of this noise cannot be overstated, and the entirety of the Discussion should be considered with it in mind. As a result, LOD for wood smoke ranged from 0.022 to 0.091 mg/m^3. This is within the range of similar smoke-detector based devices tested on similar aerosols and designed to operate across a wide concentration range. In particular, the UCB-PATS has demonstrated a limit of detection between 0.030 to 0.050 mg/m^3 for wood smoke (Chowdhury et al. 2007). This range is reasonable for the measurement of high-concentration settings like kitchens of biomass users in developing countries or highly polluted cities like Ulaanbaatar, but is not adequate for the measurement of ambient concentrations in rural regions or US cities where concentrations are often around the US EPA annual average primary standard of 0.012 mg/m^3 (12 µg/m^3). The LOD of recently developed low-cost sensors specifically tailored to ambient air quality monitoring are lower. The PPD has demonstrated an LOD as low as 0.001 mg/m^3 (Austin et al. 2015; Holstius et al. 2014). In another test (Wang et al. 2015), the LOD of the PPD, DSM, and GP2Y were estimated at 0.006, 0.011, and 0.026 mg/m^3, respectively. A custom low-cost sensor, the Plantower PMS 3003, currently deployed in communities throughout the US has shown LOD between 0.001 and 0.011 mg/m^3 (Kelly et al. 2017).

One reason for the decreased LOD of other low-cost sensors is that many have been modified to provide higher sensitivity at lower concentrations, often by modifying the brightness of the chamber light source (e.g. LED), focusing or collecting optics, or adding air flow. However, this process increases scattering at both low and high concentrations, and thus comes at the cost of a lower maximum detection limit. Austin et al. (2015), for example, saw this limit in the PPD at around 0.9 mg/m^3, 1.2 mg/m^3, and 2.7 mg/m^3 for 1 μm, 2 μm, and 3 μm polystyrene particles, respectively. The upper detection limit of the UCB-PATS optical chamber, which is more similar to the Nest Device setup, has been reported at about 25 mg/m^3 (Berkeley Air Monitoring Group 2011). The present experiments did not approach sensor saturation for the Nest Devices even as concentrations reached 15 mg/m^3. In addition, the measured LED pulse voltages (0.4 V) were approximately 10 times lower than the Nest Protect power supply (4.5V). It is possible that simple software modifications (e.g. increased LED pulse voltages) could reduce noise and improve LOD. Future studies should experiment with such modifications and explore the true peak capacities of the chamber LEDs and photo sensor.

4.5 Conclusions

This analysis demonstrates the efficacy of using a popular SSD to produce reliable estimates of $PM_{2.5}$ mass concentration, and finds that current SSD might be made to do so with minor hardware modifications or software updates alone. However, there is considerable work to be done before this can be achieved. This study required co-opting and processing raw analog data with an off-board component system (the Raspberry Pi setup), because public access to raw sensor data via API was not supported. More work is needed to understand how the component systems as they exist in commercially available SSD can be used together or modified to effectively measure $PM_{2.5}$ without non-OEM (original equipment manufacturer) hardware add-ons and without voiding the smoke detection functionality that current owners rely on. In order to facilitate this, manufacturers should make raw sensor data accessible to the public,[32] collaborate with the research community, and perform (or publish any past) in-house analyses (if existing).

Deriving reliable and actionable $PM_{2.5}$ data from any light scattering sensor currently requires calibration against aerosol characteristics like morphology and refractive index. Without such corrective action, SSD output is virtually useless for the purposes of pollution inference. This study calibrated Nest Device output for laboratory testing by co-location with professional monitors, an option infeasible for large networks or individual consumers. More work is needed to develop frameworks for calibrating SSD against common and locally relevant aerosol mixtures, and, especially, to produce methods that remove human interaction from the calibration process. Moreover, producers of

[32] Public availability of all data (not just a user's own data) will, of course, produce privacy concerns which should be explored.

professional and low-cost optical $PM_{2.5}$ sensors typically distribute units as calibrated to one aerosol type. For example, the DustTrak II is shipped with calibration against Arizona road dust (TSI Inc 2013), and only cigarette smoke particle count calibration data are provided by the PPD manufacturer's datasheet (Gao et al. 2015). This system is inadequate for real world situations in which aerosol compositions are heterogeneous and frequently changing; measurements taken from monitors calibrated in such a way are prone to error and bias. SSD-based monitors and the optical sensor field, in general, would benefit from the development of calibration frameworks that update continuously based on patterns in local aerosol sources like traffic, construction, and industry and, when sensors are utilized indoors, building characteristics and resident behavior (e.g. smoking and cooking). Particularly technology-savvy manufacturers like Nest – an Alphabet, Inc. subsidiary – could apply their expertise and corporate connections in cloud communication, big data manipulation, and mathematics to develop and provide such a service in tandem with their SSD.

In this manner, applications could be developed to intercept, calibrate, and store SSD data as well as communicate them to consumers and related data repositories. Data repositories could provide critical evidence for epidemiology, exposure assessment, and related regulation. Government bodies, like the EPA and National Institute of Environmental Health Sciences, whose missions support enhanced environmental and health research should encourage manufacturers to support such efforts, providing funding and research support as necessary. Investigators should also explore the effects of traditional smoke detector placement – often on or near the ceiling (National Fire Protection Association 2016; Nest Labs 2017) and throughout multiple rooms with varying sources– on the relationship between calibrated SSD readings and the exposures actually experienced by a building's inhabitants. This is discussed in more detail in Chapter 5.

Grassroots and community monitoring groups interested in understanding hyper-local exposures or nearby sources should implore SSD companies to explore these possibilities and support such efforts by offering test locations, focus groups, and other resources. Manufacturers may, themselves, find the prospect lucrative for use in pollution-conscious regions or as a medical adjuvant for vulnerable populations. Future work should explore the market potential for SSD-based $PM_{2.5}$ monitors.

SSD are poised to revolutionize the way we measure and respond to $PM_{2.5}$. This efficacy study is among the first steps of many toward an innovative means of increasing the public accessibility and utility of $PM_{2.5}$ data, improving monitor coverage, and re-orienting the focus of large-scale networks from the outdoors to more exposure-relevant environments.

Chapter 5

Conclusions

5.1 Toward health-oriented pollution metrics

The previous chapters cover a range of topics in $PM_{2.5}$. Each is connected by the common theme of advancing the use of personal exposure as a metric of $PM_{2.5}$-related risk. Historically, regulators and advocacy groups have tended to look at air pollution through an environmental lens (Sagar et al. 2016). Monitoring networks, regulations, epidemiological analyses, and interventions have focused on environmental concentrations (e.g. outdoor ambient concentrations at US EPA SLAMS or indoor HAP concentrations) rather than actual exposures. Exposure, however, is considerably more closely related to biological dose than environmental indicators or other simplistic proxies and, as such, is considerably more relevant to health outcomes. While non-exposure methods have proved important first steps in understanding general trends and exposure-related determinants of disease, more comprehensive techniques are both needed and increasingly being made more feasible through advances in data availability, computing power, computer science, statistics, and low-cost sensors.

The work presented in Chapter 2 employs a combination of methods to estimate changes in population-wide $PM_{2.5}$ exposures and calculate related disease burden under several real-world policy scenarios in one of the world's most polluted capital cities. A more comprehensive understanding of the local population-wide health impacts of $PM_{2.5}$ pollution is achieved by combining micro-environmental concentrations with estimates of time activity and information on second hand tobacco smoke. Preliminary results of this analysis have already received attention from Mongolian media and high-ranking politicians.

Chapter 3 analyses leverage advanced statistical techniques and some of the first personal $PM_{2.5}$ exposure measurements in rural Lao to demonstrate that population exposures within the context of solid fuel use can be reliably predicted from environmental and survey indicators. Such indicators are often collected but used as simplistic or roughly adjusted

measures of $PM_{2.5}$-related risk. Chapter 3 demonstrates that, given the proper inputs, cross-validated machine learned ensemble models may be able to account for the complex interactions between various easier-to-collect (than actual exposure measurements) indicators that more directly govern population-wide $PM_{2.5}$ exposures. This may facilitate a shift in the focus of relevant policies, interventions, and project evaluations away from simplistic applications of such indicators toward explicit exposure estimates without much added effort or resource outlay.

Chapter 4 proposes the use of a popular internet-connected consumer device, the smart smoke detector, for measuring $PM_{2.5}$ concentrations indoors in real-time. A network of $PM_{2.5}$-measurement-enabled smart smoke detectors could be leveraged to help fill the gaps in existing environmental monitoring networks, provide consumers with actionable hyper-local pollution information, and cost-effectively aid in solving both current and unexpected $PM_{2.5}$ issues (e.g. unknown hotspots or sources or unanticipated events like wildfires and certain industrial releases (Kumar et al. 2015)) – a priority for the field of exposure science (Lioy and Smith 2013). In general, widespread implementation of $PM_{2.5}$ measurement features in smart smoke detectors would improve the availability of measurements in a micro-environment in which people spend a majority of their time (and is, thereby, imperative to the reconstruction of personal exposures) but that is poorly covered by monitoring networks and scientific studies.

5.2 Research needs and future work

This section identifies key areas for improvement of the work presented and opportunities for future research.

Improve Ulaanbaatar $PM_{2.5}$ exposure models with more data, less-simplistic assumptions

The exposure assessment produced in Chapter 2 provides Ulaanbaatar's policy makers with a more holistic understanding of local $PM_{2.5}$-related risk by combining time-weighted estimates of outdoor and indoor concentrations over a set of alternative emissions policy pathways. However, very little actual $PM_{2.5}$ exposure data have been collected in the city; more primary personal measurements are needed across a wide range of locally relevant environments. A campaign measuring winter and summer time *exposures* in children, adults, and elderly residents living in gers, houses, and apartments using solid fuels, gas, and centralized steam would go a long way toward validating the Chapter 2 results and informing future population-wide risk analyses.

Chapter 2 also includes relatively simplistic assumptions about $PM_{2.5}$ concentrations in non-residential environments. Researchers should explore the detailed impacts of non-residential conditions on exposures in the UB population, especially those of occupational environments which may include construction sites, tanneries, factories, mines, and heavy

traffic. State and local governmental organizations should facilitate exposure assessment efforts in the region by funding extensive occupational monitoring campaigns (for $PM_{2.5}$ and other more-toxic compounds expected in such work environments) and expanding outdoor air quality monitoring networks. Such efforts must be accompanied by data transparency and availability schemes that allow researchers to download or otherwise receive current and historical data in a form suitable for analysis (e.g. spreadsheet or comma-separated value file).[33] Investigators interested in improving on the UB population exposure estimates presented in Chapter 2 may wish to explore the impacts of updated and non-static tobacco smoking rates (if projecting future impacts); include more sophisticated modeling of background disease rates (if projecting future impacts); and gather infiltration rate information on a wide variety of local building types.

Bring prediction modeling to the public: an online resource for $PM_{2.5}$ exposure prediction

Chapter 3 demonstrates that total exposure estimates can be modeled from other, often easier-to-collect, sources of data. The models produced in Chapter 3 are not highly transportable to other settings, but the underlying framework may be. If used properly, the machine learning ensemble prediction framework has the potential to alleviate the financial, administrative, and participation-related burdens of $PM_{2.5}$ exposure research. Yet, the applied use of machine learning and ensemble packages like those used in Chapter 3 typically requires some understanding of software coding and data science. This remains a major hurdle to the widespread tailoring of the proposed prediction framework to individual projects, but cloud computing may allow a practical workaround.

Shiny (Chang et al. 2017) is a software package for R (R Core Team 2016) used to develop interactive online applications. Shiny could, in theory, be employed to design and host a user-friendly web application that automatically trains a machine-learned ensemble model from an inputted training dataset (a dataset with all variables of interest collected on a subset of the study population with exposure measurements). The application would output training statistics – like root mean squared error and the r^2 of observed vs. predicted values – and produce estimates of population-wide exposures from a larger prediction dataset (a dataset of predictor variables from either the entire study population or a subset representative thereof, but without exposure measurement data). To maximize the range of possible users, the framework should also handle various data pre-processing

[33] Recent developments in UB air quality data availability include the National Agency of Meteorology and Environmental Monitoring "Air Quality" webpage (National Agency of Meteorology and Environmental Monitoring 2017) that presents air quality information from monitors located throughout the city, a related mobile application (iOS or Android), and a data aggregation website called OpenAQ (OpenAQ 2017). The first two resources appear to be targeted towards the general public, while the latter is designed to provide downloadable datasets to more-data-savvy individuals. OpenAQ indexes data from the National Agency of Meteorology and Environmental Monitoring website (National Agency of Meteorology and Environmental Monitoring 2017) and from a website maintained by the US Diplomatic Post (AirNow Department of State 2017).

procedures. For example, missing values for independent variables should be automatically imputed and factor variables should be converted to indicator variables (perhaps with user input).

Although certain parameters would be required – like participant ID and, in the training set, time-standardized exposure measurements – the web-based exposure prediction framework would allow the user to collect and input as few or as many variables as they wish to collect. Based in part on the random forest analysis in Chapter 3, several parameters would be recommended, including HAP measurements, meteorological indicators, fuel and stove types, and survey responses regarding locally relevant cooking behaviors. These recommendations would evolve over time as the web application received and processed more data from a wider range of contexts.

Processing times are not expected to be unreasonably long. For example, the training process for the Full model from Chapter 3 (60 observations across about 300 total variables) performed on a laptop [34] lasted on the order of dozens of seconds; shiny currently allows application hosting on servers like Amazon Web Services that can be configured with substantially more computing resources. Application of this framework to particularly large datasets may increase processing time, but, because the bulk of the work is performed in the cloud, the user does not need to be engaged throughout the entire procedure and so user fatigue should still be minimal. This (processing performed in the cloud) also relaxes the user burden of maintaining high-powered computers for otherwise resource-intensive modeling— the user need only have enough computing and web power to access the web application, upload their datasets, and download the outputs.

The proposed framework draws inspiration in part from two existing shiny-based web tools: the highly successful Household Air Pollution Intervention Tool (HAPIT) (Pillarisetti et al. 2016; Pillarisetti and Smith 2017) and a drag-drop-and-click implementation of SuperLearner (Polley et al. 2016) produced by Professor Alan Hubbard at the University of California, Berkeley (Hubbard 2016). HAPIT is supported by the Global Alliance for Clean Cookstoves and is designed for comparing the relative health benefits of residential-fuel-based $PM_{2.5}$ exposure-reduction projects. The Berkeley SuperLearner web application allows users to upload a dataset in comma-separated value format, enter (text) or select (drop-down menu) independent and dependent variables, define the number of cross-validation splits, and, at the click of a button, conduct an automated cross-validated SuperLearner procedure. Missing values (non-outcome) are imputed and variables with non-numeric types are handled according to SuperLearner and general statistics conventions with only minor input required from the user. Summary statistics and a graph are produced.

[34] Apple MacBook Pro, 15-inch, late 2016, quad-core 2.6 GHz Intel Core i7, 16 GB 2133 MHz LPDDR3

Expand data inputs to improve the transferability of predictive $PM_{2.5}$ exposure models

With expanded data inputs, it may be possible to develop a set of models – not just a framework – that is generalizable across all manner of settings. The use of machine learning to predict large-scale outdoor $PM_{2.5}$ concentrations has been demonstrated with success (Beckerman et al. 2013; Gupta and Christopher 2009; Hu et al. 2017; Lary et al. 2014; Li et al. 2017; Liu et al. 2016; Niu et al. 2017; Reid et al. 2015, 2016; Zou et al. 2015), however a search of PubMed, [35] which only returned manuscripts focusing on environmental concentrations, suggests that not much work has been done using machine learning to predict large-scale total $PM_{2.5}$ *exposures*. More research is needed on this topic.

Future exploration of a generalizable machine learning model for predicting total $PM_{2.5}$ exposure should start at the level of population-wide exposures— the results of the Chapter 3 analyses indicate that low-bias models may be easier to achieve than high-precision models. To adequately demonstrate the external validity of this approach, such work should attempt to include as much data from as heterogeneous of a study population as possible. Regions with high diversity in meteorological conditions, cultural characteristics, and emissions sources like China and India are ideal for a proof-of-concept approach.

Exposure prediction models would benefit from including the types of data used in effective machine-learned environmental $PM_{2.5}$ concentration prediction, which include meteorological variables, satellite imagery (aerosol optical depth), and land use characteristics. Investigators should explore the utility of more individual-level parameters, like time activity and location – perhaps derived from mobile devices (de Nazelle et al. 2013) – or mobile phone imagery and thermometer data, which have been shown useful for predicting outdoor concentrations (Liu et al. 2016). DHS-type survey indicators did not prove to be highly useful predictors of $PM_{2.5}$ exposure in the Chapter 3 analyses, but it is likely that this is a context-specific phenomenon. Researchers should examine the predictive utility of indicators from common questionnaires like censuses or the DHS. In more technologically advanced regions, smart appliance information like electricity usage (e.g. as an indicator of air conditioning use or time spent in a specific indoor location) may add power to prediction models. As suggested in Chapter 4, smart home devices may eventually allow measurement of indoor pollution data, which surely would provide considerable utility to an exposure prediction model.

Validate and calibrate smart smoke detectors to monitor exposure-relevant $PM_{2.5}$ in-field

The reverse engineering and laboratory characterization of the Nest Protect in Chapter 4 demonstrate that at least one popular smart smoke detector can be used to produce high-quality real-time estimates of $PM_{2.5}$. In theory, a network of smart smoke detectors could be used to sense indoor $PM_{2.5}$ concentrations at scale, improving current systems for

[35] Search terms: “machine learning PM2.5” and “machine learning PM2.5 exposure”.

estimating human exposures at both the population and individual levels. Field trials are needed to validate this theory.

Of particular concern is sensor calibration at scale. Optical sensors like the one in the Nest Protect output a voltage signal that must be translated into $PM_{2.5}$ mass concentration. As discussed in Chapter 4, the relationship between voltage and mass concentration depends on many factors, including the distributions of the shape, size, and refractive index of the aerosols being measured. Calibration factors or algorithms estimated in a laboratory can be used to translate voltage to mass concentrations in the field, but because environmental aerosol compositions are heterogeneous across time and space static laboratory-derived adjustments are unreliable.

An ideal calibration procedure for use at scale would respond to hyperlocal changes in aerosol composition in real-time. Such a framework should be explored, leveraging recent advancements in cloud computing – like publicly available traffic counts and readings from nearby regulatory monitors – and the intercommunication of smart smoke detectors within and between buildings. Researchers may also find ways to employ emerging mobile source technologies– like the recent addition of nitric oxide (NO), nitrogen dioxide (NO_2), and black carbon monitors to a small number of Google Street View cars (Apte et al. 2017) – or publicly available data therefrom in reducing error and bias in this framework.

A major issue related to the in-situ conversion of a smoke detector to (or its dual use as) an air quality monitor is placement and how it might affect observed concentrations compared to concentrations actually experienced by people within that environment. Smoke detector placement is relatively standardized, which could allow for good comparability of readings between devices in separate buildings. However, convention and, in some cases, regulation dictate that smoke detectors be placed on or close to the ceiling – e.g. (National Fire Protection Association 2016; Nest Labs 2017) – which is likely to put the sensor outside the typical breathing zone – commonly estimated at a height of around 1.5 m – e.g. (Balakrishnan et al. 2013; Parlar and Greim 2009). This is important because indoor air pollution concentrations have been shown to stratify vertically, especially when a major indoor combustion source is present (Johnson et al. 2011; Kandpal et al. 1995). Stratification may thereby result in misestimation of exposure-relevant concentrations by converted smoke detector readings. Such impacts should be examined across a wide range of building types and indoor emissions profiles. Pollution mixing within a building is also affected by its ventilation characteristics (Sherman and Walker 2010). Effects on the relationship between smoke detector readings and exposure-relevant concentrations from ventilation features like windows and HVAC in the context of various weather and climate patterns across building types and $PM_{2.5}$ source profiles should also be explored.

The US Fire Administration recommends that in residential settings smoke detectors be placed in every bedroom, outside sleeping areas, and on each level of the home (US Fire Adminsitration 2017). Investigators should examine the representativeness of

concentrations at monitor placement sites in each of these areas in relation to room-specific point sources (e.g. stove in a kitchen) and behaviors (e.g. lying down in a bedroom) as they pertain to concentrations experienced by dwellers.

References

Air Visual. 2016. Node. Available: https://airvisual.com/node [accessed 15 March 2017].

AirNow Department of State. 2017. Ulaanbaatar - PM2.5. Available: http://stateair.net/dos/#Mongolia$Ulaanbaatar [accessed 19 June 2017].

Al-Ali AR, Zualkernan I, Aloul F. 2010. A mobile GPRS-sensors array for air pollution monitoring. IEEE Sens. J. 10:1666–1671; doi:10.1109/JSEN.2010.2045890.

Alexeeff SE, Schwartz J, Kloog I, Chudnovsky A, Koutrakis P, Coull BA. 2015. Consequences of kriging and land use regression for PM2.5 predictions in epidemiologic analyses: insights into spatial variability using high-resolution satellite data. J. Expo. Sci. Environ. Epidemiol. 25:138–144; doi:10.1038/jes.2014.40.

Allen-Piccolo G, Rogers J V., Edwards R, Clark MC, Allen TT, Ruiz-Mercado I, et al. 2009. An ultrasound personal locator for time-activity assessment. Int. J. Occup. Environ. Health 15:122–132; doi:10.1179/oeh.2009.15.2.122.

Allen RW, Gombojav E, Barkhasragchaa B, Byambaa T, Lkhasuren O, Amram O, et al. 2013. An assessment of air pollution and its attributable mortality in Ulaanbaatar, Mongolia. Air Qual. Atmos. Heal. 6:137–150; doi:10.1007/s11869-011-0154-3.

Amarsaikhan D, Battsengel V, Nergui B, Ganzorig M, Bolor G. 2014. A study on air pollution in Ulaanbaatar city, Mongolia. J. Geosci. Environ. Prot. 2:123–128; doi:10.4236/gep.2014.22017.

Anenberg S, Kinney P, Newcombe K, Talyan V, Goyal A, Hewlett O. 2017. Gold Standard methodology to estimate and verify averted Mortality and disability adjusted life years (ADALYs) from cleaner household air. 1–52.

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

Armendáriz-Arnez C, Edwards RD, Johnson M, Zuk M, Rojas L, Jiménez RD, et al. 2008. Reduction in personal exposures to particulate matter and carbon monoxide as a result of the installation of a Patsari improved cook stove in Michoacan Mexico. Indoor Air 18:93–105; doi:10.1111/j.1600-0668.2007.00509.x.

Asian Development Bank, Clean Air Initiative for Asian Cities Center. 2006. Country synthesis report on urban air quality management: Lao PDR, discussion draft, December.

Austin E, Novosselov I, Seto E, Yost MG. 2015. Laboratory Evaluation of the Shinyei

PPD42NS Low-Cost Particulate Matter Sensor. J. Shamaned. PLoS One 10:e0137789; doi:10.1371/journal.pone.0137789.

Balakrishnan K, Ghosh S, Ganguli B, Sambandam S, Bruce N, Barnes DF, et al. 2013. State and national household concentrations of PM2.5 from solid cookfuel use: results from measurements and modeling in India for estimation of the global burden of disease. Environ. Heal. 12:77; doi:10.1186/1476-069X-12-77.

Balakrishnan K, Mehta S, Ghosh S, Johnson M, Brauer M, Zhang J, et al. 2014. Review 5: population levels of household air pollution and exposures. In *WHO Indoor Air Quality Guidelines: Household Fuel Combustion*, World Health Organization, Geneva, Switzerland.

Balakrishnan K, Parikh J, Sankar S, Padmavathi R, Srividya K, Venugopal V, et al. 2002. Daily average exposures to respirable particulate matter from combustion of biomass fuels in rural households of southern india. Environ. Health Perspect. 110:1069–1075; doi:10.1289/ehp.021101069.

Ballesteros MF, Kresnow M-J. 2007. Prevalence of residential smoke alarms and fire escape plans in the US: results from the Second Injury Control and Risk Survey (ICARIS-2). Public Health Rep. 122: 224–231.

Baumgartner J, Schauer JJ, Ezzati M, Lu L, Cheng C, Patz J, et al. 2011. Patterns and predictors of personal exposure to indoor air pollution from biomass combustion among women and children in rural China. Indoor Air 21:479–488; doi:10.1111/j.1600-0668.2011.00730.x.

Beckerman BS, Jerrett M, Serre M, Martin R V., Lee S-J, van Donkelaar A, et al. 2013. A hybrid approach to estimating national scale spatiotemporal variability of PM 2.5 in the contiguous United States. Environ. Sci. Technol. 11:130617085617008; doi:10.1021/es400039u.

Bell ML, Ebisu K, Peng RD. 2011. Community-level spatial heterogeneity of chemical constituent levels of fine particulates and implications for epidemiological research. J. Expo. Sci. Environ. Epidemiol. 21:372–384; doi:10.1038/jes.2010.24.

Berkeley Air Monitoring Group. 2015. Quantifying the health impacts of ACE-1 biomass and biogas stoves in Cambodia Final Report.

Berkeley Air Monitoring Group. 2011. UCB Particle and Temperature Sensor (UCB-PATS) Specifications. 1.

Birdi. 2016. Tech. Available: http://birdihome.com/tech.html [accessed 6 July 2017].

Birmili W, Heinke K, Pitz M, Matschullat J, Wiedensohler A, Cyrys J, et al. 2010. Particle number size distributions in urban air before and after volatilisation. Atmos. Chem. Phys. 10:4643–4660; doi:10.5194/acp-10-4643-2010.

Bonjour S, Adair-Rohani H, Wolf J, Bruce NG, Mehta S, Prüss-Ustün A, et al. 2013. Solid fuel use for household cooking: country and regional estimates for 1980-2010. Environ. Health Perspect. 121:784–790; doi:10.1289/ehp.1205987.

Brauer M, Freedman G, Frostad J, van Donkelaar A, Martin R V, Dentener F, et al. 2016. Ambient air pollution exposure estimation for the global burden of disease 2013. Environ. Sci. Technol. 50:79–88; doi:10.1021/acs.est.5b03709.

Brauer M, Lencar C, Tamburic L, Koehoorn M, Demers P, Karr C. 2008. A cohort study of

traffic-related air pollution impacts on birth outcomes. Environ. Health Perspect. 116:680–686; doi:10.1289/ehp.10952.

Brunekreef B, Janssen NAH, de Hartog JJ, Oldenwening M, Meliefste K, Hoek G, et al. 2005. Personal, indoor, and outdoor exposures to PM2.5 and its components for groups of cardiovascular patients in Amsterdam and Helsinki. Res. Rep. Health. Eff. Inst. 1-70–9.

Burnett RT, Pope CA, Ezzati M, Olives C, Lim S, Mehta S, et al. 2014. An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. Environ. Health Perspect. 122:397–403; doi:10.1289/ehp.1307049.

Cameron T, Ostro B. 2004. Advisory Council on Clean Air Compliance Analysis response to agency request on cessation lag - letter to the United States Environmental Protection Agency.

Centers for Disease Control and Prevention. 2013. National Health and Nutrition Examination survey questionnaire 2011-2012.

Central Intelligence Agency. 2017. Field listing :: climate. World Fact B. Available: https://www.cia.gov/library/publications/the-world-factbook/fields/2059.html#cb [accessed 5 July 2017].

Chafe ZA, Brauer M, Klimont Z, Dingenen R Van, Mehta S, Rao S, et al. 2014. Household cooking with solid fuels contributes to ambient PM2.5 air pollution and the burden of disease. Environ. Health Perspect.

Chafe Z, Brauer M, Heroux M, Klimont Z, Lanki T, Salonen R, et al. 2015. Residential heating with wood and coal: health impacts and policy options in Europe and North America.

Chang W, Cheng J, Allaire J, Xie Y, McPherson J. 2017. shiny: web application framework for R.

Chartier R. 2015. Pilot testing and evaluation of an enhanced children's micropem, abstract number 3043 [abstract]. 27th Conf. Int. Soc. Environ. Epidemiol. Addressing Environ. Heal. Inequalities.

Chilkhaasuren B, Baasankhuu B. 2010. Population and economic activities of Ulaanbaatar.

Chow JC, Engelbrecht JP, Watson JG, Wilson WE, Frank NH, Zhu T. 2002. Designing monitoring networks to represent outdoor human exposure. Chemosphere 49:961–978; doi:10.1016/S0045-6535(02)00239-4.

Chowdhury Z, Campanella L, Gray C, Al Masud A, Marter-Kenyon J, Pennise D, et al. 2013. Measurement and modeling of indoor air pollution in rural households with multiple stove interventions in Yunnan, China. Atmos. Environ. 67:161–169; doi:10.1016/j.atmosenv.2012.10.041.

Chowdhury Z, Edwards RD, Johnson M, Naumoff Shields K, Allen T, Canuz E, et al. 2007. An inexpensive light-scattering particle monitor: field validation. J. Environ. Monit. 9:1099–106; doi:10.1039/b709329m.

Clark ML, Reynolds SJ, Burch JB, Conway S, Bachand AM, Peel JL. 2010. Indoor air pollution, cookstove quality, and housing characteristics in two Honduran communities. Environ. Res. 110:12–18; doi:10.1016/j.envres.2009.10.008.

Clifford A, Lang L, Chen R, Anstey KJ, Seaton A. 2016. Exposure to air pollution and

cognitive functioning across the life course – a systematic literature review. Environ. Res. 147:383–398; doi:10.1016/j.envres.2016.01.018.

Cohen A, Mehta S. 2007. Pollution and tuberculosis: outdoor sources. PLoS Med. 4:e142; doi:10.1371/journal.pmed.0040142.

Dadvand P, Parker J, Bell ML, Bonzini M, Brauer M, Darrow LA, et al. 2013. Maternal exposure to particulate air pollution and term birth weight: a multi-country evaluation of effect and heterogeneity. Environ. Health Perspect. 121:367–373; doi:10.1289/ehp.1205575.

Davy PK, Gunchin G, Markwitz A, Trompetter WJ, Barry BJ, Shagjjamba D, et al. 2011. Air particulate matter pollution in Ulaanbaatar, Mongolia: determination of composition, source contributions and source locations. Atmos. Pollut. Res. 2:126–137; doi:10.5094/APR.2011.017.

de Nazelle A, Seto E, Donaire-Gonzalez D, Mendez M, Matamala J, Nieuwenhuijsen MJ, et al. 2013. Improving estimates of air pollution exposure through ubiquitous sensing technologies. Environ. Pollut. 176:92–99; doi:10.1016/j.envpol.2012.12.032.

Demaio AR, Nehme J, Otgontuya D, Meyrowitsch D, Enkhtuya P. 2014. Tobacco smoking in Mongolia: findings of a national knowledge, attitudes and practices study. BMC Public Health 14:213; doi:10.1186/1471-2458-14-213.

Dialog Semiconductor. 2016. DA14580, document number CFR0011-120-01, revision 3.4.

Ding F, Song A. 2016. Development and coverage evaluation of ZigBee-based wireless network applications. J. Sensors 2016:1–9; doi:10.1155/2016/2943974.

Doll R, Peto R, Boreham J, Sutherland I. 2004. Mortality in relation to smoking: 50 years' observations on male British doctors. BMJ 328:1519; doi:10.1136/bmj.38142.554479.AE.

Dore G, Nagpal T. 2006. Urban transition in Mongolia: pursuing sustainability in a unique environment. Environ. Sci. Policy Sustain. Dev. 48:10–24; doi:10.3200/ENVT.48.6.10-24.

Dudoit S, van der Laan MJ. 2005. Asymptotics of cross-validated risk estimation in estimator selection and performance assessment. Stat. Methodol. 2:131–154; doi:10.1016/j.stamet.2005.02.003.

Edwards R, Hubbard A, Khalakdina A, Pennise D, Smith KR. 2007. Design considerations for field studies of changes in indoor air pollution due to improved stoves. Energy Sustain. Dev. 11:71–81; doi:10.1016/S0973-0826(08)60401-9.

Edwards R, Smith KR, Kirby B, Allen T, Litton CD, Hering S. 2006. An inexpensive dual-chamber particle monitor: laboratory characterization. J. Air Waste Manag. Assoc. 56: 789–99.

Enkhbat U, Rule A, Resnick C, Ochir C, Olkhanud P, Williams D. 2016. Exposure to PM2.5 and blood lead level in two populations in Ulaanbaatar, Mongolia. Int. J. Environ. Res. Public Health 13:214; doi:10.3390/ijerph13020214.

Environmental Systems Resource Institute. 2012. ArcGIS desktop: release 10.1. Redlands CA.

Flagan RC, Seinfeld JH. 2012. *Fundamentals of air pollution engineering*. Dover Publications Inc, Mineola, NY.

Forouzanfar MH, Alexander L, Anderson HR, Bachman VF, Biryukov S, Brauer M, et al. 2015. Global, regional, and national comparative risk assessment of 79 behavioural,

environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 6736:1990–2013; doi:10.1016/S0140-6736(15)00128-2.

Freescale Semiconductor. 2014. Kinetis KL16 sub-family, document number KL16P64M48SF5, rev 5.

Friedlander SK. 2000. *Smoke, dust, and haze: fundamentals of aerosol dynamics*. Second Edi. Oxford University Press, New York, NY.

Gao M, Cao J, Seto E. 2015. A distributed network of low-cost continuous reading sensors to measure spatiotemporal variations of PM2.5 in Xi'an, China. Environ. Pollut. 199:56–65; doi:10.1016/j.envpol.2015.01.013.

Garland C, Delapena S, Pennise D. 2017. An alternative technique for determining gravimetric particle mass deposition on PTFE filter substrate: the particle extraction method [In Preparation].

Greene L, Turner J, Edwards R, Cutler N, Duthie M, Rostapshova O. 2014a. Impact evaluation results of the MCA Mongolia Energy and Environment Project Energy-Efficient Stove Subsidy Program.

Greene L, Turner J, Edwards R, Cutler N, Duthie M, Rostapshova O. 2014b. Mongolia - energy and environment project, stove subsidies component: get microdata. Impact Eval. results MCA Mong. energy Environ. Proj. energy-efficient stove Subsid. Progr. Available: https://data.mcc.gov/evaluations/index.php/catalog/133/get_microdata [accessed 9 August 2016].

Greene L, Turner J, Edwards R, Cutler N, Duthie M, Rostapshova O. 2014c. Social Impact 2012-2013 household survey data in support of the impact evaluation of the MCA Mongolia Energy and Environment Project Energy-Efficient Stove Subsidy Program.

Gupta P, Christopher SA. 2009. Particulate matter air quality assessment using integrated surface, satellite, and meteorological products: 2. A neural network approach. J. Geophys. Res. 114:D20205; doi:10.1029/2008JD011497.

Guttikunda SK, Lodoysamba S, Bulgansaikhan B, Dashdondog B. 2013. Particulate pollution in Ulaanbaatar, Mongolia. Air Qual. Atmos. Heal. 6:589–601; doi:10.1007/s11869-013-0198-7.

Haddix A, Mallonee S, Waxweiler R, Douglas M. 2001. Cost effectiveness analysis of a smoke alarm giveaway program in Oklahoma City, Oklahoma. Inj. Prev. 7: 276–281.

Hänninen O, Hoek G, Mallone S, Chellini E, Katsouyanni K, Gariazzo C, et al. 2011. Seasonal patterns of outdoor PM infiltration into indoor environments: review and meta-analysis of available studies from different climatological zones in Europe. Air Qual. Atmos. Heal. 4:221–233; doi:10.1007/s11869-010-0076-5.

Hänninen OO, Lebret E, Ilacqua V, Katsouyanni K, Künzli N, Srám RJ, et al. 2004. Infiltration of ambient PM2.5 and levels of indoor generated non-ETS PM2.5 in residences of four European cities. Atmos. Environ. 38:6411–6423; doi:10.1016/j.atmosenv.2004.07.015.

Hill LD, Pillarisetti A, Delapena S, Garland C, Jagoe K, Koetting P, et al. 2015. Appendix D. Forms used in field work. Air Pollut. Impact Anal. a Pilot Stove Interv. Rep. to Minist. Heal. Inter-Ministerial Clean Stove Initiat. Lao People's Democr. Repub. Available:

http://ehsdiv.sph.berkeley.edu/krsmith/publications/2015/Lao_Appendices_all_Jul_20_15.pdf [accessed 5 May 2017].

Hinds WC. 1999. *Aerosol Technology: Properties, Behavior, and Measurement of Airborne Particles, 2nd Edition*. 2nd ed. John Wiley & Sons, Inc., New York.

HJI Group Corporation, MonEnergy Consult Co. Ltd. 2011. Mongolia: Ulaanbaatar low carbon energy supply project using a public-private partnership model.

Hoek G, Beelen R, de Hoogh K, Vienneau D, Gulliver J, Fischer P, et al. 2008. A review of land-use regression models to assess spatial variation of outdoor air pollution. Atmos. Environ. 42:7561–7578; doi:10.1016/j.atmosenv.2008.05.057.

Holstius DM, Pillarisetti A, Smith KR, Seto E. 2014. Field calibrations of a low-cost aerosol sensor at a regulatory monitoring site in California. Atmos. Meas. Tech. 7:1121–1131; doi:10.5194/amt-7-1121-2014.

Hu X, Belle JH, Meng X, Wildani A, Waller LA, Strickland MJ, et al. 2017. Estimating PM 2.5 concentrations in the conterminous United States using the random forest approach. Environ. Sci. Technol. 51:6936–6944; doi:10.1021/acs.est.7b01210.

Huang K, Fu JS, Hsu NC, Gao Y, Dong X, Tsay S-C, et al. 2013. Impact assessment of biomass burning on air quality in Southeast and East Asia during BASE-ASIA. Atmos. Environ. 78:291–302; doi:10.1016/j.atmosenv.2012.03.048.

Hubbard AE. 2016. Cross-validated AUC using SL. Available: https://ahubb40.shinyapps.io/CVSLAUC/ [accessed 22 June 2017].

ICF International. 2011. Demographic and Health Surveys methodology - questionnaires: household, woman's, and man's.

Institute for Health Metrics & Evaluation. 2017. GBD compare. Available: http://vizhub.healthdata.org/gbd-compare/ [accessed 2 June 2017].

Jantunen MJ, Katsouyanni K, Lebret E, Maroni M, Saarela K, Zmirou D. 1998. Final report: air pollution exposures in European cities: the EXPOLIS study. 202.

Japan International Cooperation Agency. 2013. Capacity development project for air pollution control in Ulaanbaatar city Mongolia, final report.

Jenkins PL, Phillips TJ, Mulberg EJ, Hui SP. 1992. Activity patterns of Californians: use of and proximity to indoor pollutant sources. Atmos. Environ. Part A. Gen. Top. 26:2141–2148; doi:10.1016/0960-1686(92)90402-7.

Jerrett M, Arain A, Kanaroglou P, Beckerman B, Potoglou D, Sahsuvaroglu T, et al. 2005a. A review and evaluation of intraurban air pollution exposure models. J. Expo. Anal. Environ. Epidemiol. 15:185–204; doi:10.1038/sj.jea.7500388.

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

Johnson M, Pennise D, Lam N, Brant S, Charron D, Gray C, et al. 2011. Modeling indoor air pollution concentrations from stove emissions using a Monte Carlo single-box model. Atmos. Environ. 45: 3237–3243.

Kandpal JB, Maheshwari RC, Kandpal TC. 1995. Indoor air pollution from domestic cookstoves using coal, kerosene and LPG. Energy Convers. Manag. 36:1067–1072; doi:10.1016/0196-8904(94)00087-G.

Kelly KE, Whitaker J, Petty A, Widmer C, Dybwad A, Sleeth D, et al. 2017. Ambient and laboratory evaluation of a low-cost particulate matter sensor. Environ. Pollut. 221:491–500; doi:10.1016/j.envpol.2016.12.039.

Kloog I, Koutrakis P, Coull BA, Lee HJ, Schwartz J. 2011. Assessing temporally and spatially resolved PM2.5 exposures for epidemiological studies using satellite aerosol optical depth measurements. Atmos. Environ. 45:6267–6275; doi:10.1016/j.atmosenv.2011.08.066.

Kohn M. 2013. Mongolia selects GDF Suez-led group to build 450MW power plant. Bloomberg Business, August 26.

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

Lary DJ, Faruque FS, Malakar N, Moore A, Roscoe B, Adams ZL, et al. 2014. Estimating the global abundance of ground level presence of particulate matter (PM2.5). Geospat. Health 8:611; doi:10.4081/gh.2014.292.

Levy I, Mihele C, Lu G, Narayan J, Hilker N, Brook JR. 2014. Elucidating multipollutant exposure across a complex metropolitan area by systematic deployment of a mobile laboratory. Atmos. Chem. Phys. 14:7173–7193; doi:10.5194/acp-14-7173-2014.

Lewis A, Edwards P. 2016. Validate personal air-pollution sensors. Nature 535:29–31; doi:10.1038/535029a.

Li L, Zhang J, Qiu W, Wang J, Fang Y. 2017. An ensemble spatiotemporal model for predicting PM2.5 concentrations. Int. J. Environ. Res. Public Health 14:549; doi:10.3390/ijerph14050549.

Li T, Cao S, Fan D, Zhang Y, Wang B, Zhao X, et al. 2016. Household concentrations and personal exposure of PM2.5 among urban residents using different cooking fuels. Sci. Total Environ. 548–549:6–12; doi:10.1016/j.scitotenv.2016.01.038.

Liaw A, Wiener M. 2002. Classification and regression by randomForest. R News 2: 18–22.

Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. 2012. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 380:2224–60; doi:10.1016/S0140-6736(12)61766-8.

Lioy PJ, Smith KR. 2013. A discussion of exposure science in the 21st century: a vision and a strategy. Environ. Health Perspect. 23:1–1; doi:10.1289/ehp.1206170.

Litton CD. 2002. Studies of the measurement of respirable coal dusts and diesel particulate matter. Meas. Sci. Technol. 13:365–374; doi:10.1088/0957-0233/13/3/319.

Litton CD, Smith KR, Edwards R, Allen T. 2004. Combined optical and ionization measurement techniques for inexpensive characterization of micrometer and submicrometer aerosols. Aerosol Sci. Technol. 38:1054–1062; doi:10.1080/027868290883333.

Liu C, Tsow F, Zou Y, Tao N. 2016. Particle pollution estimation based on image analysis. H. Liued. PLoS One 11:e0145955; doi:10.1371/journal.pone.0145955.

Long CM, Suh HH, Catalano PJ, Koutrakis P. 2001. Using time- and size-resolved particulate

data to quantify indoor penetration and deposition behavior. Environ. Sci. Technol. 35:2089–2099; doi:10.1021/es001477d.

Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, et al. 2012. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 380:2095–128; doi:10.1016/S0140-6736(12)61728-0.

Luo B, Zhang S, Ma S, Zhou J, Wang B. 2014. Effects of different cold-air exposure intensities on the risk of cardiovascular disease in healthy and hypertensive rats. Int. J. Biometeorol. 58:185–194; doi:10.1007/s00484-013-0641-3.

Marshall SW, Runyan CW, Bangdiwala SI, Linzer MA, Sacks JJ, Butts JD. 1998. Fatal residential fires: who dies and who survives ? J. Am. Med. Assoc. 279:1633–1637; doi:279.20.1633.

McCracken JP, Schwartz J, Bruce N, Mittleman M, Ryan LM, Smith KR. 2009. Combining individual- and group-level exposure information. Epidemiology 20:127–136; doi:10.1097/EDE.0b013e31818ef327.

Meier R, Eeftens M, Phuleria HC, Ineichen A, Corradi E, Davey M, et al. 2015. Differences in indoor versus outdoor concentrations of ultrafine particles, PM2.5, PM absorbance and NO2 in Swiss homes. J. Expo. Sci. Environ. Epidemiol. 25:499–505; doi:10.1038/jes.2015.3.

Mengersen K, Morawska L, Wang H, Murphy N, Tayphasavanh F, Darasavong K, et al. 2011. Association between indoor air pollution measurements and respiratory health in women and children in Lao PDR. Indoor Air 21:25–35; doi:10.1111/j.1600-0668.2010.00679.x.

Micron Technologies. 2014. Micron serial NOR flash memory, document number 09005aef845665fe, revision S.

Micron Technologies. 2015. Technical note: serial flash memory device marking for the M 25P, M 25PE, M 45PE, M 25PX, and N25Q product families, document number 09005aef84fb7996, revision G.; doi:10.1080/02564602.1995.11416507.

Ministry of Health, Lao Statistics Bureau of the Ministry of Planning and Investment. 2012. Lao PDR social indicator survey (MICS-DHS) 2011-2012.

Mittleman A, Goldenson A, Dong W, Fadell T, Rogers M, Matsuoka Y, et al. 2014. Detector unit and sensing chamber therefor.

Mongolian Statistical Information Service. 2016. Mid-year total population, by regions, aimags and the capital (Mid-year). Stat. Database by Sect. Available: http://1212.mn/ [accessed 19 October 2016].

Morawska L, Mengersen K, Wang H, Tayphasavanh F, Darasavong K, Holmes NS. 2011. Pollutant concentrations within households in Lao PDR and association with housing characteristics and occupants' activities. Indoor Air 45:882–9; doi:10.1021/es102294v.

Murty RN, Mainland G, Rose I, Chowdhury AR, Gosain A, Bers J, et al. 2008. CitySense: an urban-scale wireless sensor network and testbed. Proc. 8th IEEE Conf. Technol. Homel. Secur.; doi:10.1109/THS.2008.4534518.

Naeher L, Brauer M, Lipsett M, Zelikoff J, Simpson C, Koenig JQ, et al. 2007. Woodsmoke health effects: a review. Inhal. Toxicol. 19.

Naeher LP, Smith KR, Leaderer BP, Mage D, Grajeda R. 2000. Indoor and outdoor PM2.5 and CO in high- and low-density Guatemalan villages. J. Expo. Anal. Environ. Epidemiol. 10:544–551; doi:10.1038/sj.jea.7500113.

Nakao M, Yamauchi K, Ishihara Y, Solongo B, Ichinnorov D. 2016. Effects of air pollution and seasonality on the respiratory symptoms and health-related quality of life (HR-QoL) of outpatients with chronic respiratory disease in Ulaanbaatar: pilot study for the comparison of the cold and warm seasons. Springerplus 5:1817; doi:10.1186/s40064-016-3481-x.

National Agency of Meteorology and Environmental Monitoring. 2017. Air quality. Available: http://www.agar.mn [accessed 19 June 2017].

National Fire Protection Association. 2016. Installing and maintaining smoke alarms. Available: http://www.nfpa.org/public-education/by-topic/smoke-alarms/installing-and-maintaining-smoke-alarms [accessed 22 June 2017].

National Institutes of Health. 2007. Biological sciences curriculum study. NIH Curric. Suppl. Ser. Available: http://www.ncbi.nlm.nih.gov/books/NBK20362/ [accessed 7 June 2017].

National Research Council. 2012. *Exposure science in the 21st century: a vision and a strategy committee on human and environmental exposure science in the 21st century*. National Research Council, Washington, DC.

National Statistics Office of Mongolia. 2003. Mongolian statistical yearbook 2002.

National Statistics Office of Mongolia. 2007. Mongolian statistical yearbook 2006.

National Statistics Office of Mongolia. 2011. Mongolian statistical yearbook 2010.

National Statistics Office of Mongolia. 2012. The 2010 population and housing census of Mongolia.

Nest Labs. 2015a. Nest split-spectrum white paper.

Nest Labs. 2015b. The new Nest Protect, SPEC SHEET.

Nest Labs. 2017. Where in the room should I install my Nest Protect? Available: https://nest.com/support/article/Where-in-the-room-should-I-install-my-Nest-Protect [accessed 22 June 2017].

Niu M, Gan K, Sun S, Li F. 2017. Application of decomposition-ensemble learning paradigm with phase space reconstruction for day-ahead PM 2.5 concentration forecasting. J. Environ. Manage. 196:110–118; doi:10.1016/j.jenvman.2017.02.071.

Northcross A, Chowdhury Z, McCracken J, Canuz E, Smith KR. 2010. Estimating personal PM2.5 exposures using CO measurements in Guatemalan households cooking with wood fuel. J. Environ. Monit. 12:873–8; doi:10.1039/b916068j.

Northcross AL, Katharine Hammond S, Canuz E, Smith KR. 2012. Dioxin inhalation doses from wood combustion in indoor cookfires. Atmos. Environ. 49:415–418; doi:10.1016/j.atmosenv.2011.11.054.

NXP Semiconductors. 2016. Kinetis K24F sub-family data sheet, document number K24P144M120SF5, revision 7.

OpenAQ. 2017. Mongolia: Ulaanbaatar. Available: https://openaq.org/#/countries/MN [accessed 19 June 2017].

OpenAQ. 2016. US Diplomatic Post: Ulaanbaatar. OpenAQ Data Mong. Available:

www.openaq.org [accessed 20 July 2016].
Oudin A, Forsberg B, Adolfsson AN, Lind N, Modig L, Nordin M, et al. 2015. Traffic-related air pollution and dementia incidence in northern Sweden: a longitudinal study. Environ. Health Perspect. 124:306–312; doi:10.1289/ehp.1408322.
Parlar H, Greim H, eds. 2009. Part III: air monitoring methods. In *The MAK-Collection for Occupational Health and Safety*, Vol. 11 of, p. 26, Deutsche Forschungsgemeinschaft, Weinheim.
Pillarisetti A, Mehta S, Smith KR. 2016. HAPIT, the Household Air Pollution Intervention Tool, to evaluate the health benefits and cost-effectiveness of clean cooking interventions. In *Broken Pumps and Promises* (E.A. Thomased.), pp. 147–169, Springer International Publishing, Cham.
Pillarisetti A, Smith KR. 2017. HAPIT 2.0. Available: https://hapit.shinyapps.io/HAPIT/ [accessed 20 June 2017].
Pinto JP, Lefohn AS, Shadwick DS. 2004. Spatial variability of PM 2.5 in urban areas in the United States. J. Air Waste Manage. Assoc. 54:440–449; doi:10.1080/10473289.2004.10470919.
Pokhrel AK, Bates MN, Acharya J, Valentiner-Branth P, Chandyo RK, Shrestha PS, et al. 2015. PM 2.5 in household kitchens of Bhaktapur, Nepal, using four different cooking fuels. Atmos. Environ. 113:159–168; doi:10.1016/j.atmosenv.2015.04.060.
Pokhrel AK, Bates MN, Verma SC, Joshi HS, Sreeramareddy CT, Smith KR. 2010. Tuberculosis and indoor biomass and kerosene use in Nepal: a case-control study. Environ. Health Perspect. 118:558–64; doi:10.1289/ehp.0901032.
Polley EC, LeDell E, Kennedy C, van der Laan MJ. 2016. SuperLearner: super learner prediction.
Postolache O, Pereira JMD, Girao PMBS. 2009. Smart sensors network for air quality monitoring applications. IEEE Trans. Instrum. Meas. 58:3253–3262; doi:10.1109/TIM.2009.2022372.
Power MC, Adar SD, Yanosky JD, 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; doi:10.1016/j.neuro.2016.06.004.
Purple Air. 2016. An air quality monitoring network built on a new generation of laser particle counters. Available: http://www.purpleair.org/welcome-to-purple-air [accessed 15 March 2017].
R Core Team. 2016. R: A language and environment for statistical computing.
Ray PP. 2016. Internet of Things Cloud Based Smart Monitoring of Air Borne PM2 . 5 Density Level.
Reid CE, Jerrett M, Petersen ML, Pfister GG, Morefield PE, Tager IB, et al. 2015. Spatiotemporal prediction of fine particulate matter during the 2008 northern California wildfires using machine learning. Environ. Sci. Technol. 49:3887–3896; doi:10.1021/es505846r.
Reid CE, Jerrett M, Tager IB, Petersen ML, Mann JK, Balmes JR. 2016. Differential respiratory health effects from the 2008 northern California wildfires: a spatiotemporal

approach. Environ. Res. 150:227–235; doi:10.1016/j.envres.2016.06.012.
Sagar A, Balakrishnan K, Guttikunda S, Roychowdhury A, Smith KR. 2016. India leads the way: a health-centered strategy for air pollution. Environ. Health Perspect. 124:116–117; doi:10.1289/EHP90.
Schmidt M, Lipson H. 2013. Eureqa (Version 0.99.6 beta).
Sheppard L, Burnett RT, Szpiro AA, Kim SY, Jerrett M, Pope CA, et al. 2012. Confounding and exposure measurement error in air pollution epidemiology. Air Qual. Atmos. Heal. 5:203–216; doi:10.1007/s11869-011-0140-9.
Sherman M, Walker I. 2010. Impacts of mixing on acceptable indoor air quality in homes. HVAC&R Res. 16:315–329; doi:10.1080/10789669.2010.10390907.
Sherman MH. 1987. Estimation of infiltration from leakage and climate indicators. Energy Build. 10:81–86; doi:10.1016/0378-7788(87)90008-9.
Shimada Y, Matsuoka Y. 2011. Analysis of indoor PM2.5 exposure in Asian countries using time use survey. Sci. Total Environ. 409:5243–52; doi:10.1016/j.scitotenv.2011.08.041.
Shmueli G. 2010. To explain or to predict? Stat. Sci. 25:289–310; doi:10.1214/10-STS330.
Smith K, Bruce N, Balakrishnan K, Adair-Rohani H, Balmes J, Dherani M, et al. 2014. Millions dead: how do we know and what does it mean? Methods used in the Comparative Risk Assessment of Household Air Pollution. Am. Rev. Public Heal. 35:185–206; doi:10.1146/annurev-publhealth-032013-182356.
Smith K, Pillarisetti A. 2012. A short history of woodsmoke and implications for Chile. Estud. Públicos.
Smith K, Pillarisetti A, Hill LD, Charron D, Delapena S, Garland C, et al. 2015. Proposed methodology: quantification of a saleable health product (aDALYs) from household cooking interventions.
Smith KR. 2000. National burden of disease in India from indoor air pollution. Proc. Natl. Acad. Sci. 97:13286–13293; doi:10.1073/pnas.97.24.13286.
Soldo D, Quarto A, Di Lecce V. 2012. M-DUST: an innovative low-cost smart PM sensor. 2012 IEEE I2MTC - Int. Instrum. Meas. Technol. Conf. Proc. 1823–1828; doi:10.1109/I2MTC.2012.6229647.
Speck. 2017. About us. Available: https://www.specksensor.com/about-us [accessed 15 March 2017].
Specker H. 1968. Problems and possibilities of modern trace analysis. Angew. Chemie Int. Ed. English 7:252–259; doi:10.1002/anie.196802521.
Statistics Department of Ulaanbaatar. 2013. Population and household census 2012.
Steinle S, Reis S, Sabel CE. 2013. Quantifying human exposure to air pollution —moving from static monitoring to spatio-temporally resolved personal exposure assessment. Sci. Total Environ. 443:184–193; doi:10.1016/j.scitotenv.2012.10.098.
Sumpter C, Chandramohan D. 2013. Systematic review and meta-analysis of the associations between indoor air pollution and tuberculosis. Trop. Med. Int. Heal. 18:101–108; doi:10.1111/tmi.12013.
Tcakhiur S, Narangerel G, Ganbat P. 2013. Ulaanbaatar 2020 master plan and development approaches for 2030.
The Harris Poll. 2015a. Ownership and awareness of select smart-home devices/appliances

among US internet users, May 2015.
The Harris Poll. 2015b. Smart-home devices/appliances owned by US internet users, by beneration, May 2015.
The World Bank. 2011. Air quality analysis of Ulaanbaatar: improving air quality to reduce health impacts.
Tiwari M, Sahu SK, Bhangare RC, Yousaf A, Pandit GG. 2014. Particle size distributions of ultrafine combustion aerosols generated from household fuels. Atmos. Pollut. Res. 5:145–150; doi:10.5094/APR.2014.018.
TSI Inc. 2013. Rationale for programming a photometer calibration factor (Pcf) of 0.38 for ambient monitoring, EXPMN-007 rev. A.
Ulzii D, Warburton N, Brugha RE, Tserenkh I, Enkhmaa D, Enkhtur S, et al. 2015. Personal exposure to fine-particle PM2.5 black carbon air pollution in schoolchildren living In Ulaanbaatar, Mongolia. Heal. Eff. Air Pollut. Nanoparticles 1: 67–74.
United Nations Children's Fund. 2016. Reducing impacts of air pollution on children's health in Ulaanbaatar, Mongolia.
United Nations Children's Fund Mongolia. 2016. Understanding and addressing the impact of air pollution on children's health in Mongolia.
United Nations Department of Economic and Social Affairs: Population Division. 2013. World population prospects the 2012 revision volume I: comprehensive tables. I.
United States Bureau of the Census. 2012. Households and families: 2010. Census 2010 Br. C2010BR-14; doi:C2010BR-14.
United States Environmental Protection Agency. 2006. 40 CFR part 60; standards of performance for electric utility steam generating units, industrial- commercial-institutional steam generating units, and small industrial-commercial-institutional steam generating units; final rule. Fed. Regist. 71: 9866–9886.
United States Environmental Protection Agency. 2016. Air Quality System data mart [monitor listing]. Available: http://www.epa.gov/ttn/airs/aqsdatamart [accessed 14 March 2017].
United States Environmental Protection Agency. 1995a. User's guide for the industrial source complex (ISC3) dispersion models. Rep. No. EPA-454/B-95-003a 1; doi:EPA-454/B-95-003a.
United States Environmental Protection Agency. 1995b. User's guide for the industrial source complex (ISC3) dispersion models. Rep. No. EPA-454/B-95-003b 2; doi:EPA-454/B-95-003b.
US Fire Adminsitration. 2017. Smoke alarm outreach materials. Available: https://www.usfa.fema.gov/prevention/outreach/smoke_alarms.html [accessed 22 June 2017].
US National Archives and Records Administration. 2016. Appendix D to oart 58, network design criteria for ambient air quality monitoring.
van der Laan MJ, Polley EC, Hubbard AE. 2007. Super learner. Stat. Appl. Genet. Mol. Biol. 6:Article25; doi:10.2202/1544-6115.1309.
Van Dingenen R, Raes F, Putaud J-P, Baltensperger U, Charron A, Facchini M-C, et al. 2004. A European aerosol phenomenology—1: physical characteristics of particulate matter

at kerbside, urban, rural and background sites in Europe. Atmos. Environ. 38:2561–2577; doi:10.1016/j.atmosenv.2004.01.040.

van Donkelaar A, Martin R V., Brauer M, Boys BL. 2014. Use of satellite observations for long-term exposure assessment of global concentrations of fine particulate matter. Environ. Health Perspect. 110:135–143; doi:10.1289/ehp.1408646.

Venkataraman C, Sagar AD, Habib G, Lam N, Smith KR. 2010. The Indian National Initiative for Advanced Biomass Cookstoves: the benefits of clean combustion. Energy Sustain. Dev. 14:63–72; doi:10.1016/j.esd.2010.04.005.

Volckens J, Quinn C, Leith D, Mehaffy J, Henry CS, Miller-Lionberg D. 2017. Development and evaluation of an ultrasonic personal aerosol sampler. Indoor Air 27:409–416; doi:10.1111/ina.12318.

Wallace L a, Wheeler AJ, Kearney J, Van Ryswyk K, You H, Kulka RH, et al. 2011. Validation of continuous particle monitors for personal, indoor, and outdoor exposures. J. Expo. Sci. Environ. Epidemiol. 21:49–64; doi:10.1038/jes.2010.15.

Wang Y, Li J, Jing H, Zhang Q, Jiang J, Biswas P. 2015. Laboratory evaluation and calibration of three low-cost particle sensors for particulate matter measurement. Aerosol Sci. Technol. 49:1063–1077; doi:10.1080/02786826.2015.1100710.

Warburton D, Gilliland F, Dashdendev B. 2013. Environmental pollution in Mongolia: effects across the lifespan. Environ. Res. 124:65–6; doi:10.1016/j.envres.2013.04.002.

Ward T, Lange T. 2010. The impact of wood smoke on ambient PM2.5 in northern Rocky Mountain valley communities. Environ. Pollut. 158:723–729; doi:10.1016/j.envpol.2009.10.016.

Williams R, Suggs J, Rea A, Sheldon L, Rodes C, Thornburg J. 2003. The Research Triangle Park particulate matter panel study: modeling ambient source contribution to personal and residential PM mass concentrations. Atmos. Environ. 37:5365–5378; doi:10.1016/j.atmosenv.2003.09.010.

World Bank Asia Sustainable and Alternative Energy Program. 2009. Mongolia: heating in poor, peri-urban ger areas of Ulaanbaatar.

World Health Organization. 2010. Mongolian steps survey on the prevalence of noncommunicable disease and injury risk factors.

World Health Organization. 2014. WHO guidelines for indoor air quality: household fuel combustion.

World Health Organization Western Pacific Region. 2012. Smoking bans in all public areas, Mongolia. WHO Represent. Off. Mong. Available: http://www.wpro.who.int/mongolia/mediacentre/smoking/en/ [accessed 7 June 2017].

Zhang J, Smith KR. 2007. Household air pollution from coal and biomass fuels in China: measurements, health impacts, and interventions. Environ. Health Perspect. 115:848–855; doi:10.1289/ehp.9479.

Zhou Z, Liu Y, Yuan J, Zuo J, Chen G, Xu L, et al. 2016. Indoor PM2.5 concentrations in residential buildings during a severely polluted winter: a case study in Tianjin, China. Renew. Sustain. Energy Rev. 64:372–381; doi:10.1016/j.rser.2016.06.018.

Zou B, Wang M, Wan N, Wilson JG, Fang X, Tang Y. 2015. Spatial modeling of PM2.5

concentrations with a multifactoral radial basis function neural network. Environ. Sci. Pollut. Res. 22:10395–10404; doi:10.1007/s11356-015-4380-3.

Appendix A

List of acronyms from Chapters 1-5 and Appendix B

Acronym	Description
ACE-1	Africa clean energy stove
aDALY	Averted disability-adjusted life year
ADC	Analog to digital converter
ALRI	Acute lower respiratory infection, typically referring to disease in those < 5 years of age
API	Application program interface
BAU	Business as usual
CHP	Combined heat power plant
CIU	Clean indoor use (describing a heating source that is clean at the point of use)
COPD	Chronic obstructive pulmonary disease
CV	Cross-validated or cross-validation
DALY	Disability-adjusted life year
DHS	Demographic and health survey
E	East
EPA	Enivornmental Protection Agency
GBD 2010	Study of the 2010 global burden of disease
GHz	Gigahertz
GIS	Geographic information system
GM	Geometric mean
GPIO	General purpose input/output
GSD	Geometric standard deviation
HAP	Hosuehold air pollution
HAPIT	Household air pollution intervention tool
HECH	Household energy, climate, and health research group
HEPA	High efficiency particulate air
HOB	Heat only boiler(s)

HVAC	Heating, ventilation, and cooling
Hz	Hertz, or once per second
ID	Identification(s)
IHD	Ischemic heart disease
IR	Infrared
JICA	Japanese International Cooperation Agency
KAP	Kitchen air pollution (a form of HAP)
KEF	Kitchen exposure factor
km	Kilometer
L	Liter
Lao PDR	Lao People's Democratic Republic
LC	Lung cancer
LED	Light-emitting diode
LOD	Lower limit of detection
LPB	Low pressure boiler(s)
LPM	Liters per minute
LPS	Liters per second
LSIS	Lao social indicator survey
LUR	Land use regression
m	Meter
mA	Milliamps
MCA	Millennium Challenge Account
MHz	Megahertz
MicroPEM	Micro personal PM exposure monitor
mm	Millimeter
MMBtu	Million British thermal units
MSE	Mean squared error
mV	millivolt(s)
N	North
NE	Northeast
NO	Nitric oxide
NO_2	Nitrogen dioxide
NW	Northwest
OEM	Original equipment manufacturer
PAF	Population attributable fraction
PATS	Particle and temperature sensor

PM	Particulate matter
PM_{10}	Particulate matter smaller than 10 μm in diameter
$PM_{2.5}$	Particulate matter smaller than 2.5 μm in diameter
QA	Quality assurance
RH	Relative humidity
RMSE	Root mean squared error
SD	Standard deviation
SE	Southeast
SHS	Second hand smoke
SLAMS	State and local air monitoring station(s)
SUMS	Stove use monitoring system
SW	Southwest
UB	Ulaanbaatar, the capital city of Mongolia
UC	University of California
UCB	University of California, Berkeley
US	United States
USD	United States dollar
V	Voltes
W	West
WHO	World Health Organization
μg	Microgram(s)
μs	Microsecond(s)

Appendix B [36]

Supplemental information for Chapter 2

B.1 Demographic conditions

Projecting population and household numbers

Citywide estimates of population and household size were used to calculate household numbers and disease burdens. Population projections (total population and population < 5 years old) for Ulaanbaatar (UB) for 2010, 2015, 2020, 2025, and 2030 were taken as the "medium growth" (version 1b) projections identified in the 2010 Population and Housing Census of Mongolia Report (National Statistics Office of Mongolia 2012). Annual population estimates for relevant interim years (2014-2024) were estimated by linearly interpolating the 5-year projections in Microsoft Excel for Mac 2011 and 2016. Detailed information on the spatial distribution of population and household number by household type was obtained from 2012 city census data (Statistics Department of Ulaanbaatar 2013).

Projections were made for the number of homes by type throughout Ulaanbaatar. Household types most relevant to the Ulaanbaatar context were identified as gers, single family houses, and multi-family apartments, as described in Chapter 2. Projections for the number of total households in each year were unavailable, and so were estimated from family size and total population, assuming an average of one family per household. Family size was extrapolated by fitting a curve to historical trends (2000-2010) obtained from Mongolia's Annual Statistical Yearbook series (Figure 1) (National Statistics Office of

[36] This appendix is a version of a supplemental text submitted for peer review as part of: "Hill, LD, Edwards, R, Turner, JR, Damdinsuren, Y, Olkhanud, P, Odsuren, M, Guttikunda, S, Ochir, C, Smith, K. Health assessment of future $PM_{2.5}$ exposures from indoor, outdoor, and environmental tobacco smoke concentrations under alternative policy pathways in Ulaanbaatar, Mongolia. *In review. 2017.*" The title was revised during review to (at the time of dissertation filing) to: "Health assessment of future $PM_{2.5}$ exposures from indoor, outdoor, and second hand tobacco smoke concentrations under alternative policy pathways in Ulaanbaatar, Mongolia." This also applies to all figures and tables therein. At the time of dissertation filing, a final response had not been received from reviewers.

Mongolia 2003, 2007, 2011). Because of Ulaanbaatar's recent rapid rate of decrease in family size, a linear extrapolation resulted in unrealistically low family sizes in future years. Thus, total fertility rates (TFR) were used to infer a reasonable lower limit.

TFR is defined by the United Nations (UN) as the average number of children a hypothetical cohort of women would have at the end of their reproductive period if they were subject during their whole lives to the fertility rates of a given period and if they were not subject to mortality. While modeling the UN's 5-year TFR estimates for Mongolia against the 10 available years of household size data yielded no discernable relationship, we assumed that trends in family size would follow the national TFR. Mongolia's TFR has experienced a dramatic decrease in recent decades, but is expected to level off. The UN suggests that Mongolia's TFR will become stable at near-Western levels by 2030 (United Nations Department of Economic and Social Affairs: Population Division 2013). This suggests that while the estimated rate of decline in household size observed between 2000 and 2010 was steep and linear, it is likely to level off in the near future. For this reason, a trigonometric curve was fit to past data that would approximate a near-term asymptotic approach of 2.6 persons/home, the average 2010 US household size (United States Bureau of the Census. 2012). Eureqa Formulize (Schmidt and Lipson 2013) was used to fit this curve. Household sizes for individual years during 2010-2030 were then taken from this curve with the assumption of one family per household (Figure 1). Citywide household number was estimated by dividing the expected size of Ulaanbaatar's population in each year by the average household size.

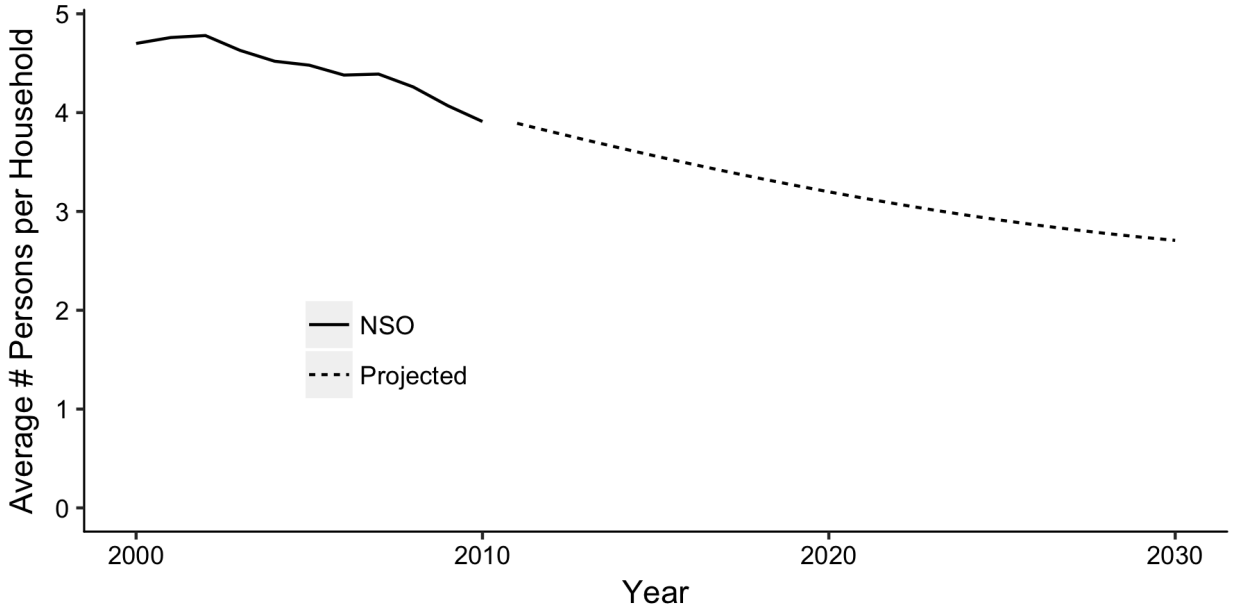

Figure 1. Trends in family size (persons per household).
Identified by the National Statistics Office of Mongolia for 2000 – 2010, and estimated using extrapolation and assumptions of the Total Fertility Rate for 2011 - 2030.

Gers and single-family houses are typically located in regions identified by the Statistics Department of Ulaanbaatar as "ger areas," while apartment households are typically in "apartment areas." Projections of the proportion of Ulaanbaatar residents living in the ger areas were provided by the Ulaanbaatar 2020 Master Plan and Development Approaches for 2030 report (Tcakhiur S et al. 2013). Annual trends were linearly interpolated in Microsoft Excel for Mac 2011 and 2016. The proportion of all households located in ger areas was taken directly from the population interpolations. All other households were assumed in the apartment areas. Approximately 99% of households in the apartment areas in 2012 were multi-unit apartments (Statistics Department of Ulaanbaatar 2013). The remaining 1% of households was classified as "luxury house" or "homeless"; their household heating emissions and exposures were not explicitly calculated, and were thus assigned the population-weighted averages of all other households. Statistics Department figures show that in the ger areas, families were about evenly split between gers and houses. In 2012, families living in gers accounted for 44.4% of ger area households, while families living in houses accounted for 54.4% of ger area households (Statistics Department of Ulaanbaatar 2013). The remaining ger area household types (~ 1%) were assigned the population-weighted average heating emissions and exposures of all other household types. Projections for the proportion of ger area households as gers vs. houses were unavailable, and so the 2012 proportion was assumed constant over time. The final projections of household number by type and location are shown in Figure 2.

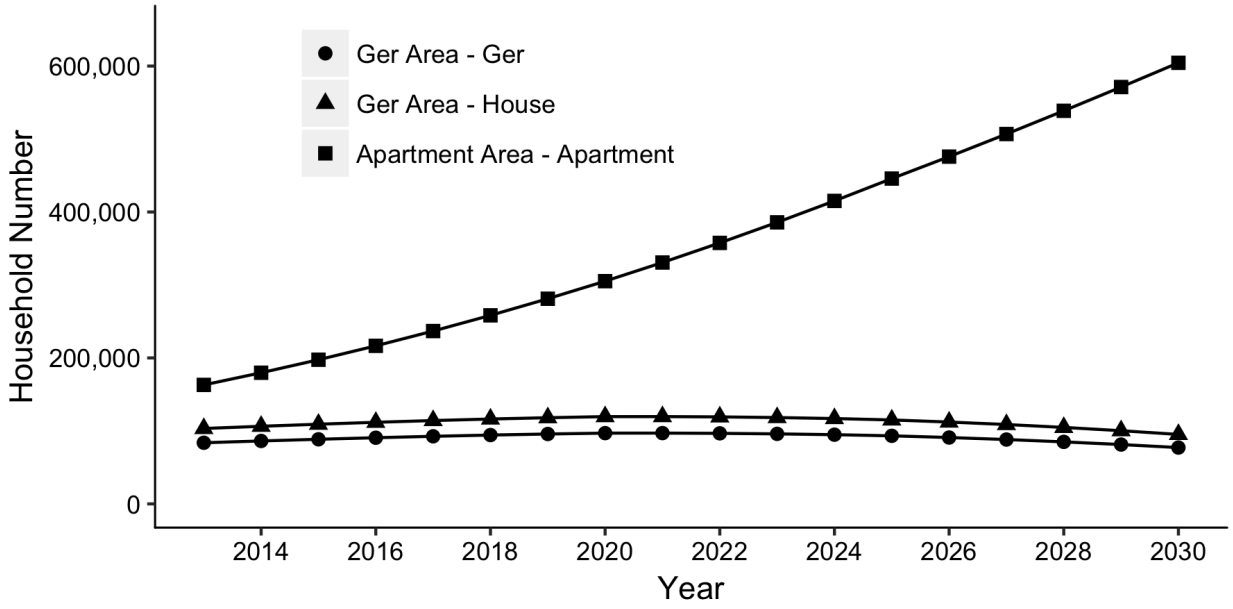

Figure 2. Projections of Ulaanbaatar household numbers stratified by area and home type.

Projecting background disease rates

During analysis, a considerable discrepancy was discovered between locally identified death rates and nationally identified death rates. Recent estimates of Mongolia-wide disease-specific mortality rates (Lim et al. 2012) show values that are between 50% and 400% higher than those identified in the data for Ulaanbaatar received from the Ministry of Health and Sports. Although national disease rates should not exactly mirror those of the capital city, they are expected to be similar. It is suspected these discrepancies arise from the inclusion of "garbage codes", or improperly coded deaths, in the raw Ministry of Health and Sports dataset. Garbage codes are a well-known phenomenon and occur frequently all over the world (Lozano et al. 2012). International teams like the Institute for Health Metrics and Evaluation (IHME) employ rigorous statistical and diagnostic methods to redistribute such deaths to their probable underlying causes. This results in disease-specific mortality estimates that more accurately represent true rates. Disease-specific data employed in this analysis were thus adjusted using IHME values (Lim et al. 2012), as discussed in greater detail below.

Mortality data for the capital city were modeled through 2024 using historical data for 2006-2012 provided by the Health Development Center of the Ministry of Health and Sports in conjunction with the Mongolian National University of Medical Sciences. Deaths for 2006-2012 matching the ICD-10 codes used to calculate $PM_{2.5}$-related illness in the IHME 2010 Global Burden of Disease Study were obtained (Lozano et al. 2012)— Ischemic Heart Disease was defined as mortality from ICD-10 codes I20-I25; Stroke, from I63, I65 - I67 (excluding I67.4), I69.3, I60-I62, I69.0 - I69.2, and I67.4; Lung Cancer, from C33- C34, D02.1-D02.2, and D38.1; Chronic Obstructive Pulmonary Disease, from J40 – J44, and J47; and Acute Lower Respiratory Infections in children < 5 years old, from J09 – J11 , J13, J14, J12.1, J12 (excluding J12.1), J15 – J22, J85, and P23. Linear models were chosen to produce consistent, parsimonious projections of total mortality in UB from each disease in each year of the study. Disease-specific mortality models were created in R (R Core Team 2016) and took the form of Equation B.1 S1, where $\beta_{0,k}$ is the y-intercept for disease "k" and $\beta_{1,k}$ is the regression coefficient for the effect of year "j" on the number of deaths from disease "k". The model was adjusted for the discrepancy between local and national mortality data using "ϕ", the ratio of IHME-reported national death rate to UB-specific death rate (Table B.1.1). The results of the model are shown in Table B.2. Ratios (ϕ) were created for 2010, which was the only year for which both IHME and Ministry of Health and Sports estimates were available. Values for the 2010 UB total population and child (0-4 year) population were taken as the "medium growth" estimates (version 1b) from the 2010 Population and Housing Census of Mongolia Report (National Statistics Office of Mongolia 2012) for consistency with previously discussed demographic estimation methods. Adjusted estimates of background disease values in each year are provided in Table B.3, and adjusted disease-specific mortality models are shown in Figure 3. This figure shows that not all mortality models had a significant fit, but the resulting estimates were reasonable given the low

sample sizes. Population was explored as a covariate, but overall did not produce a better set of models.

$$Adjusted\ Total\ Mortality\ _{j,k} = \left(\beta_{0,k}\ +\ \beta_{1,k} \cdot (Year\ j)\right) \cdot \phi_k \qquad \text{(B.1)}$$

Table B.1 . Rate adjustment factor (ϕ) and background mortality rates for 2010 as tabulated for UB from Ministry of Health and Sports data for UB and as taken from IHME (Lim et al. 2012) national rates

Disease	UB 2010 Deaths per 100,000	Mongolia 2010 Deaths per 100,000	Ratio of National Rate to UB rate (ϕ)
Lung Cancer	12.12	18.14	1.50
ALRI	78.22	276.82	3.54
COPD	4.45	21.66	4.87
Ischemic Heart Disease	86.13	164.19	1.91
Stroke	76.89	124.22	1.62

Table 2. Unadjusted background disease rate projection models, by disease

	β0	SE	β1	SE	p-value	ϕ
Lung Cancer	-7767	7267	3.93	3.62	0.33	*1.50*
ALRI	-10541	7341	5.29	3.65	0.21	*3.54*
COPD	-11509	3451	5.75	1.72	0.02	*4.87*
Ischemic Heart Disease	-89524	35962	45.0	17.9	0.05	*1.91*
Stroke	-22871	20113	11.80	10.01	0.29	*1.62*

Table B.3. Adjusted estimates of background disease in Ulaanbaatar, 2014-2024

	Lung Cancer		ALRI (0-4 years)		COPD		Ischemic Heart Disease		Stroke	
Year	*Deaths*	*DALYs*	*Deaths*	*DALYs*	*Deaths*	*DALYs*	*Deaths*	*DALYs*	*Deaths*	*DALYs*
2014	218	5,463	371	31,760	350	13,700	1,974	43,977	1,402	32,220
2015	224	5,610	389	33,364	378	14,796	2,060	45,890	1,421	32,659
2016	230	5,758	408	34,967	406	15,892	2,146	47,803	1,440	33,097
2017	236	5,905	427	36,570	434	16,987	2,232	49,716	1,459	33,536
2018	242	6,052	446	38,173	462	18,083	2,318	51,628	1,478	33,975
2019	248	6,200	464	39,776	490	19,179	2,404	53,541	1,497	34,414
2020	254	6,347	483	41,379	518	20,275	2,490	55,454	1,516	34,853
2021	260	6,494	502	42,983	546	21,370	2,575	57,367	1,536	35,291
2022	266	6,642	520	44,586	574	22,466	2,661	59,280	1,555	35,730
2023	272	6,789	539	46,189	602	23,562	2,747	61,193	1,574	36,169
2024	277	6,936	558	47,792	630	24,658	2,833	63,106	1,593	36,608

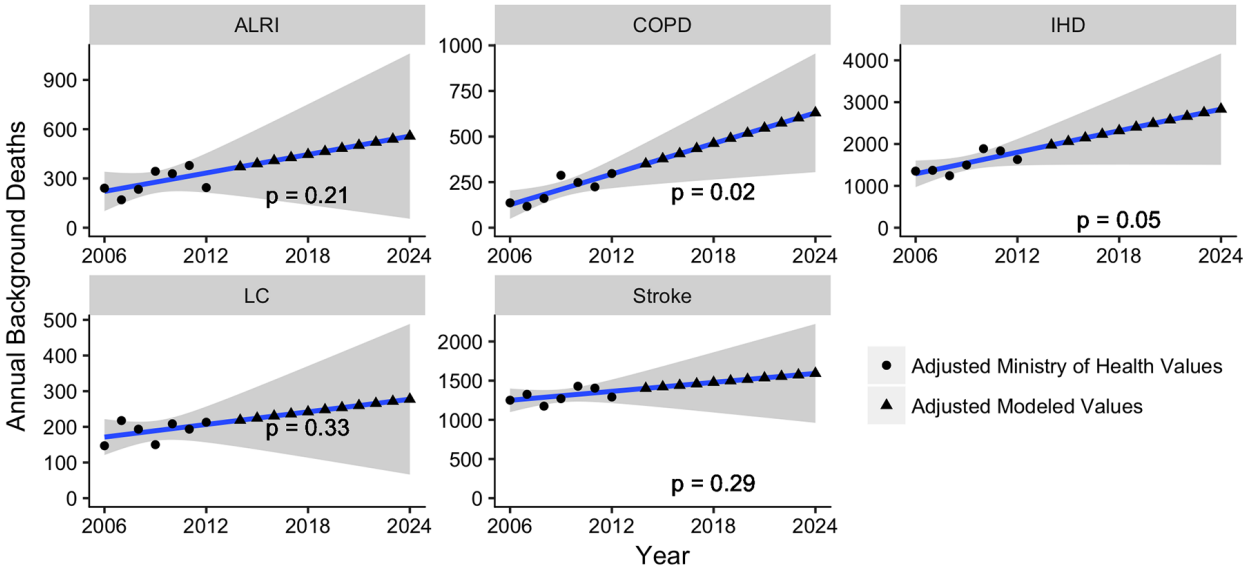

Figure 3. Adjusted disease-specific annual background mortality data (2006-2012) and projections (2014-2024).
Projection models, model fits (α = 0.05), and model 95% Confidence Intervals are shown.

B.2 Indoor air quality estimates

Indoor air quality for MCA stove users

Linear models were used to estimate indoor concentrations in stove-heated homes from data collected as part of the impact evaluation of the Millennium Challenge Account (MCA) Mongolia Energy and Environment Project Energy Efficient Stove Subsidy Program conducted by Social Impact (SI) for the Millennium Challenge Corporation (MCC) (Greene et al. 2014b, 2014a, 2014c). Details of the larger study can be found in that report (Greene et al. 2014a). Briefly, overnight indoor $PM_{2.5}$ concentrations were measured using filter-based techniques in gers and houses throughout UB during the winter of 2012-2013. Measurements were spread over three distinct winter phases representing early, mid, and late winter. The linear model was created in R (R Core Team 2016) and took the form of Equation B.2 S2, where Indoor $PM_{2.5}$ is the overnight average indoor concentration of $PM_{2.5}$ (mg/m^3), β_0 is the intercept representing a non-smoking house with an MCA project stove during the first measurement phase of the emissions study, β_1 is the impact of using a traditional stove rather than an MCA stove, β_2 is the impact of the presence of a smoker in the home, β_3 is the impact of a ger environment rather than a house, and β_4 and β_5 are the additional impacts of the measurement being taken during the second and third measurement phases, respectively.

$$\text{Log}(Indoor\ PM_{2.5}) = \beta_0 + \beta_1 * Traditional + \beta_2 * ETS + \beta_3 * Ger + \beta_4 * Phase2 + \beta_5 * Phase3 \quad \text{(B.2)}$$

These variables were chosen as evidence suggests they should have a considerable impact on indoor concentrations at the population level (Balakrishnan et al. 2013; Chowdhury et al. 2013; Li et al. 2016). A log-transformation was made to indoor concentration, as required by the log-normal distribution of the data (Shapiro-Wilk $p < 0.0001$). Four data points from the original MCC dataset were excluded because of inconsistency between various household type indicators, and sixteen observations were excluded because at least one of the variables of interest was missing, leaving a total model sample size of 196 for the modeling. The results of Equation B.2 S2 are detailed in Table B.4. Average indoor wintertime concentrations by home, stove, and second hand tobacco smoke (SHS) as applied in the exposure model discussed in Chapter 2 were calculated by averaging model results across the three wintertime study phases.

Table B.4. Household indoor log-$PM_{2.5}$ concentration model

	Estimate	Std. Error	p-value
β_0	-2.21	0.16	< 0.0001
β_1 (Traditional Stove)	0.057	0.090	0.53
β_2 (SHS)	0.16	0.082	0.059
β_3 (Ger Dwelling)	-0.10	0.082	0.22
β_4 (Study Phase 2)	0.33	0.15	0.031
β_5 (Study Phase 3)	-0.15	0.16	0.34
Model Adjusted R^2: 0.132		Model p-value: < 0.0001	

<u>Indoor air quality in homes using low pressure boilers, heat only boilers, and other stoves</u>

Low pressure boilers and semi-coke coal stoves are widely heralded by the public for their improvements in efficiency and functionality over traditional coal stoves, but little data on their contributions to indoor $PM_{2.5}$ exist. It is unreasonable to suggest completely clean function. For the sake of simplicity and a lack of data, we assigned the same indoor $PM_{2.5}$ concentrations to homes with low pressure boilers and semi-coke coal stoves as were assigned to those with MCA stoves, by home type. As discussed elsewhere, Future Tech stoves were assumed to produce a 20% reduction in indoor $PM_{2.5}$ concentrations over MCA stoves, by home type. For BAU and both alternative policy pathways, the contribution of SHS to indoor concentrations in gers was assumed the same as the contribution modeled

from SHS in gers using MCA stoves (18.1 μg/m^3). The contribution of SHS to indoor concentrations in apartments and houses was assumed the same as that modeled in MCA stove houses (20.0 μg/m^3), as our model did not provide data on SHS in apartments and because apartments are structurally more similar to houses than gers.

Infiltration efficiency

Data on the infiltration efficiencies of houses and apartments in the context of Ulaanbaatar were lacking, so local infiltration efficiencies were estimated from geographically similar regions. Seasonal average infiltration efficiencies for $PM_{2.5}$ in apartment buildings and houses were taken from EXPOLIS, a study that included the calculation of infiltration efficiency in several major cities (Hänninen et al. 2004; Jantunen et al. 1998). We used the average seasonal infiltration efficiencies specifically reported for buildings in urban Helsinki, because it is the EXPOLIS city with a climate most similar to that of UB. From these data, the average of the infiltration efficiencies reported for the spring (March-May) and summer (June-August) seasons (64%) was used in our summertime (April - September) calculations, and the average of the infiltration efficiencies reported for the autumn (September - November) and winter (December - February) seasons (53%) was used in our wintertime (October-March) calculations. These values are consistent with infiltration efficiencies found in similar home types, climates and seasons (Long et al. 2001).

Unlike houses and apartments, the ventilation characteristics of which are more limited, gers have ceiling flaps and large doorways that are left open for much of the summer months. This is likely to result in high ventilation rates and a virtual elimination of filtration related to airflow through building casings. Thus, a summer infiltration efficiency of 100% was applied to gers. A wintertime infiltration efficiency of 70% was estimated from blower door tests performed as part of the United Nations Development Programme – Global Environment Facility Commercialization of Super-Insulated Buildings in Mongolia project (MON/99/G35), the results of which were communicated to us by a collaborator on the project, Munkhbayar B. at the Mongolian University of Science and Technology. The tests provided the number of winter time air changes per hour at 50 Pascals of pressure difference (n50) for gers with modest insulation and fly cover: 45 air changes per hour. The n50 air change rate was converted to a natural air change rate of 2.25 air changes per hour using the methods described in (Sherman 1987). This natural air change rate was translated into an estimate of wintertime PM2.5 infiltration efficiency of 70% using a curve presented in (Williams et al. 2003) for the translation of general air exchange rates into PM2.5 infiltration rates.

B.3 Residential heating stoves emissions field

Figure 4 shows the 2014 Base scenario wintertime (September through March) average emission rate from coal-fired residential heating stoves as projected onto 1 km × 1 km grids.

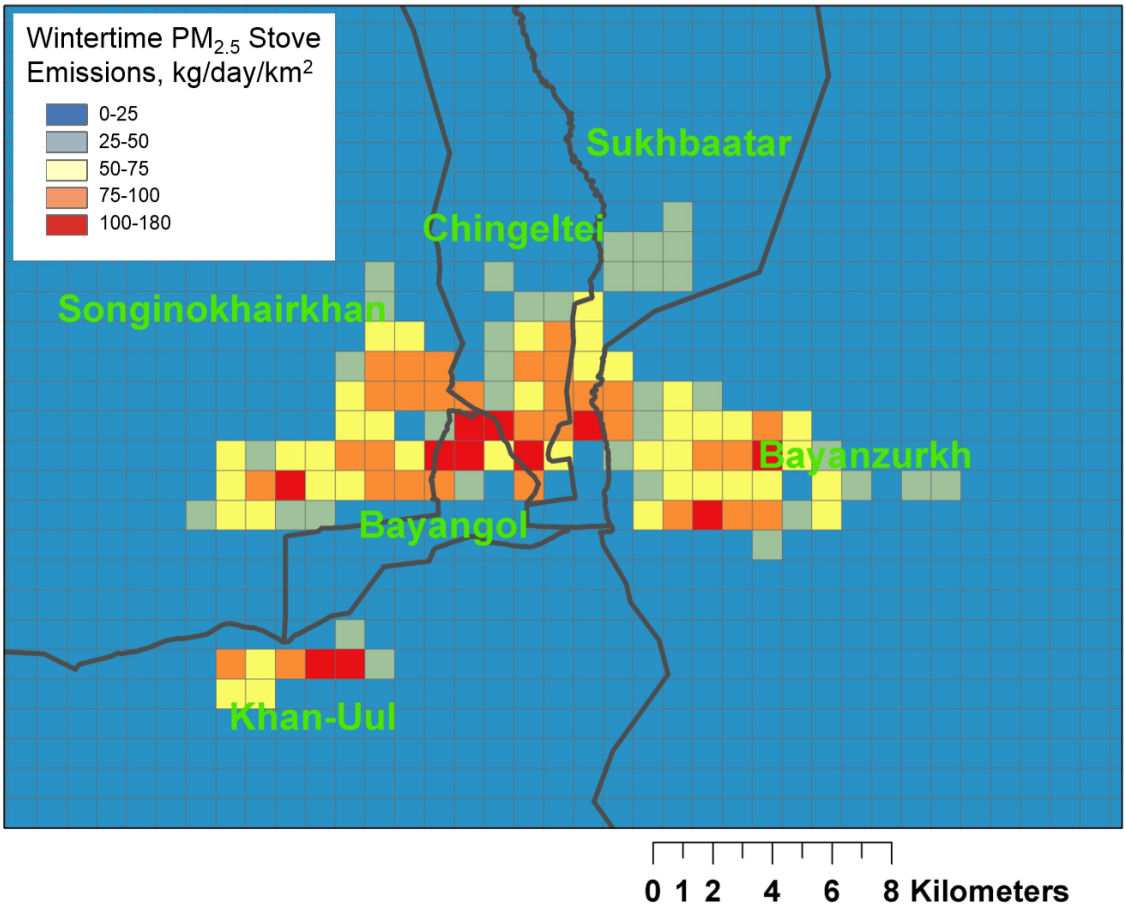

Figure 4. Wintertime (September through March) average emissions from residential heating stoves. Figure created by Dr. Jay Turner, who performed the outdoor modeling.

B.4 Scaling outdoor ambient $PM_{2.5}$ concentration models

Ambient air quality model performance was evaluated using model-to-monitor comparisons. Limited outdoor $PM_{2.5}$ data were available for this comparison. For example, during the 2012-2013 heating season $PM_{2.5}$ mass concentration data were collected with high data completeness by the National Agency for Meteorology, Hydrology, and Environment Monitoring at one location – air quality monitoring station #2 (UB02). This site was next to a major roadway and likely suffered high impacts from local traffic that could not be resolved by the model. Thus, outdoor $PM_{2.5}$ data collected by Ecography and Ecoworld under contract from MCA-Mongolia were used for the comparison. The sampling locations, methodology (Greene et al. 2014b, 2014c), and key results are detailed in the full SI project report (Greene et al. 2014a). Their data from January 22 to March 2, 2013 were used for the model-to-monitor comparison with 19 samples per site. The 2014 projected inventory was used except that residential stove emissions were calculated under the assumption of full-penetration of MCA stoves as defined in Chapter 2.

The measured average $PM_{2.5}$ concentrations at each site, shown by the single-crossed bars in Figure 5, demonstrate high spatial variability with up to a 50% difference between sites. The sample time period from the full MCA project was also modeled; average concentrations including all days between January 22 and March 2 are shown by the solid black bars in Figure 5 (Greene et al. 2014a). Modeled concentrations were much lower than the measured values and were less variable between sites. There were several possible reasons for these differences including, but not limited to, the emissions for these sources being underestimated and the model not being able to account for the trapping and accumulation of emissions from one hour to the next. The model was reconciled to the measurement data by increasing the residential stove, HOB, and motor vehicle emissions by a factor of 2.85, which was the value of the four measured-to-modeled concentration ratios. The cross-hatched bars in Figure 5 show the modeled $PM_{2.5}$ concentrations after this scaling. Assuming the only error was in the emissions inventory, the nearly threefold increase of the projected JICA 2010 inventory was still lower than the inventory projected by Guttikunda et al. for 2010 for each of these source categories (Guttikunda et al. 2013; Japan International Cooperation Agency 2013). While the scaling increases the emission inventory for these sources by about a factor of three, this places the effective emissions between those projected from the year 2010 inventories prepared by JICA and Guttikunda et al. Thus, the scaled emissions were deemed reasonable because they were bounded by the best available inventories. Power plant emissions were not scaled because the JICA and Guttikunda et al. inventories are relatively similar and emissions from tall stacks are less likely to be trapped and accumulate at ground level. This residential stove, HOB, and motor vehicle emissions scaling was applied during modeling to BAU and the two alternative pathways.

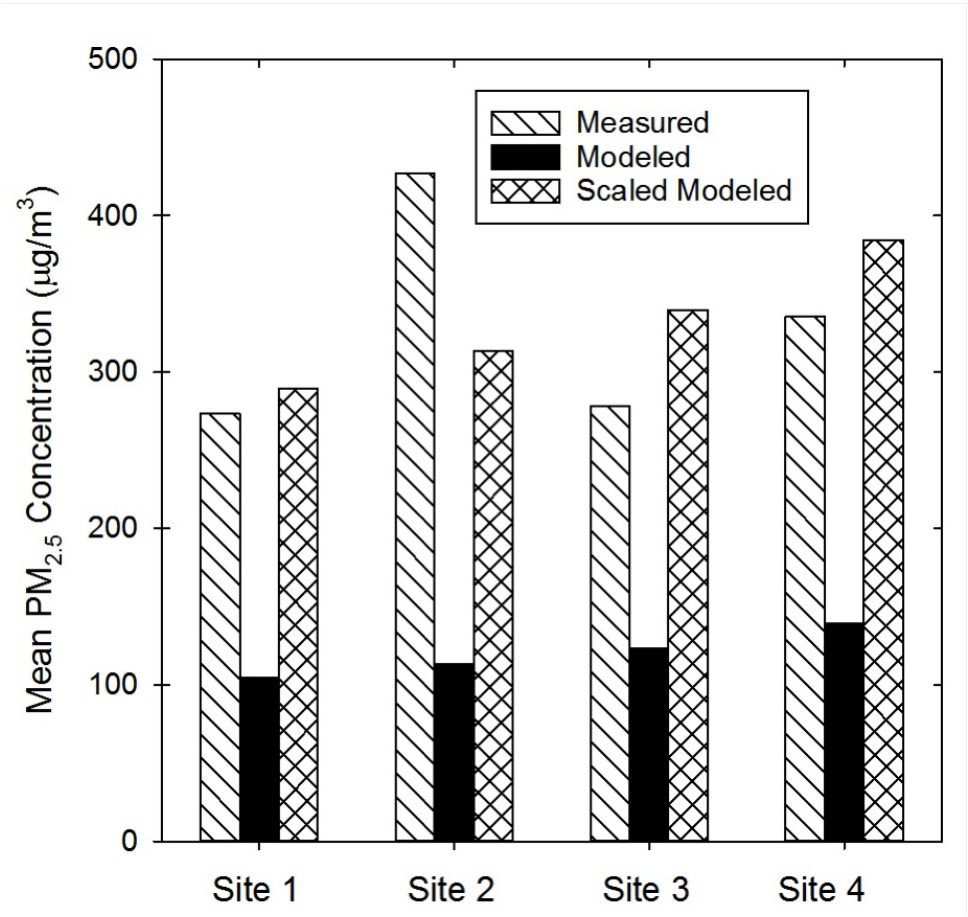

Figure 5. Measured and modeled $PM_{2.5}$ concentrations at four sites, January 22 – March 2, 2013. 10 µg/m^3 was subtracted from each of the observed concentration values to adjust for sources not included in the modeling. Figure created by Dr. Jay Turner, who performed the outdoor modeling.