DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY
Office for Scholarly Communication

Assessing indoor exposures to particulate matter and radon

Citation
Carrion-Matta, Aleshka. 2021. Assessing indoor exposures to particulate matter and radon. Doctoral dissertation, Harvard University Graduate School of Arts and Sciences.

Permanent link
https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37369482

Terms of Use
This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story .

Accessibility

HARVARD UNIVERSITY

Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the
Committee on Higher Degrees in Population Health Sciences,
have examined a dissertation entitled

"Assessing Indoor Exposures to Particulate Matter and Radon"

presented by

Aleshka Carrion-Matta

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

Dr. Petros Koutrakis, Ph.D., Committee Chair, Harvard T. H. Chan School of Public Health

Dr. Joel Schwartz, Ph.D., Harvard T. H. Chan School of Public Health

Dr. Philip Demokritou, Ph.D., Harvard T. H. Chan School of Public Health

In lieu of all Dissertation Advisory Committee members' signatures,
I, Tyler J. VanderWeele, Ph.D., appointed by the Ph.D. in Population Health Sciences,
confirm that the Dissertation Advisory Committee has examined the above dissertation,
presented by Aleshka Carrion-Matta, and hereby certify that it is worthy of acceptance as of 14 January 2021.

Date: 14 January 2021

ASSESSING INDOOR EXPOSURES TO PARTICULATE MATTER AND RADON

A dissertation presented

by

ALESHKA CARRION-MATTA

To

The Department of Population Health Sciences

In the Environmental Health Area of Study

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

In the subject of

Population Health Sciences

Harvard University

Cambridge, Massachusetts

January, 2021

© 2021 Aleshka Carrion-Matta Scholar

All rights reserved.

Assessing indoor exposures to particulate matter and radon

ABSTRACT

Indoor air pollutants are ranked in the top five environmental risks to public health. As the time spent indoors is large enough and increasing over time, we were interested in understanding sources, transport, physicochemical characteristics, influential factors, and trends to enhance our understanding of exposures in the indoor environment.

In Chapter 2, we analyzed $PM_{2.5}$ and its components to determine indoor and outdoor possible sources affecting indoor $PM_{2.5}$ in classrooms. $PM_{2.5}$ mass and its components were collected from 32 inner-city schools in the Northeastern U.S. from 2009-2013. We applied the USEPA PMF to the $PM_{2.5}$ components to estimate the source apportionment both indoors and outdoors. Classroom indoor concentrations of $PM_{2.5}$ (an average of 5.2 $\mu g/m^3$) were lower than outdoors (an average of 6.5 $\mu g/m^3$). The major sources (contributions) of indoor $PM_{2.5}$ were secondary pollution (41%) and motor vehicles (17%), followed by Calcium (Ca)-rich particles (12%), biomass burning (15%), soil dust (6%), and marine aerosols (4%). Likewise, the major sources of outdoor $PM_{2.5}$ were secondary pollution (41%) and motor vehicles (26%), followed by biomass burning (17%), soil dust (7%), road dust (3%), and marine aerosols (1%).

In Chapter 3, we were interested to find influential factors of radon levels in residential environments from different geographical areas. We analyzed factors from the soil, geology, topography, atmospheric variables, radiation, urbanization, community economic well-being, and monthly and year variations using random forest. We analyzed 802 zip codes from Massachusetts, Michigan, and Minnesota during 2005-2018. Factors that predict radon varied across the states, due possibly to a different soil composition in the states. Cross-validated R^2 between predictions

and measurements was 0.68 (RMSE = 47.81 Bq/m^3) in Minnesota, 0.67, (RMSE = 52.61 Bq/m^3) in Michigan, and 0.41 (RMSE = 52.57 Bq/m^3) in Massachusetts.

In Chapter 4, we assessed whether proximity to shale gas development areas affect indoor radon levels. We analyzed 35,442 monthly indoor radon observations from Pennsylvania between 2001-2015 using linear mixed effects model, with a random intercept for zip code. We found that shale gas development via hydraulic fracturing was associated with an increase in downwind indoor radon levels [slope = 1.93 (± 0.94); p-value 0.0393]. In addition, our analysis revealed that the contribution of having 100 wells in 1 km will increase indoor radon levels by 185 Bq/m^3 (5 pCi/L).

TABLE OF CONTENTS

ABSTRACT iii

TABLE OF CONTENTS v

LIST OF FIGURES vii

LIST OF TABLES ix

DEDICATION x

ACKNOWLEDGMENTS xi

CHAPTER 1. Introduction 1
1.1 Overview 2
1.2 Aims 5

CHAPTER 2. Classroom indoor $PM_{2.5}$ sources and exposures in inner-city schools 6
Authorship and Acknowledgments 7
Abstract 9
2.1 Introduction 11
2.2 Materials and methods 13
2.3 Results and discussion 16
2.4 Conclusions 31

CHAPTER 3. Predictors of indoor radon levels 34
Authorship and Acknowledgments 35
Abstract 36
3.1 Introduction 37
3.2 Materials and methods 38
3.3 Results 42
3.4 Discussion 48
3.5 Conclusions 52

CHAPTER 4. Indoor radon levels near areas of shale gas development via hydraulic fracturing 54
Authorship and Acknowledgments 55
Abstract 56
4.1 Introduction 57
4.2 Materials and methods 58
4.3 Results 60
4.4 Discussion 63
4.4 Conclusions 66

CHAPTER 5. Conclusion Remarks ……...…….. 68

SUPPLEMENTAL INFORMATION ……...….... 71
Supplemental Information for Chapter 2 ……. 72
Supplemental Information for Chapter 3 ……...…….. 79
Supplemental Information for Chapter 4 ……...…….. 87

REFERENCES ……...…… 88

LIST OF FIGURES

Figure 2.1 Source profiles contributing to indoor $PM_{2.5}$. The grey filled-in rectangle represents the estimated average concentration of each species by source factor estimated using PMF. The small closed (red) square represents the percentage of a specific species ascribed to the specific source (factor) by PMF. For example, using PMF, the mean concentration of indoor SO_4 ascribed to secondary pollution is 0.7 $\mu g/m^3$, which is 62% of the total sulfate concentration (See also Table 2.3). The small open square represents the estimated average concentration of each species by source factor estimated using DISP with the error bars indicating the confidence intervals.

Figure 2.2 Source profiles contributing to outdoor $PM_{2.5}$. The grey filled-in rectangle represents the estimated average concentration of each species by source factor estimated using PMF. The small closed (red) square represents the percentage of a specific species ascribed to the specific source (factor) by PMF. For example, using PMF, the mean concentration of indoor SO_4 ascribed to secondary pollution is 1.3 $\mu g/m^3$, which is 80% of the total sulfate concentration (See also Table 2.4). The small open square represents the estimated average concentration of each species by source factor estimated using DISP with the error bars indicating the confidence intervals.

Figure 2.3 Box plots of indoor-to-outdoor $PM_{2.5}$ and sulfur ratios by season. Green boxes represent indoor-to-outdoor $PM_{2.5}$ ratios and gray boxes represent indoor-to-outdoor sulfur ratios. The box plot shows the mean, 25th and 75th percentiles, while the end of the whiskers indicates the minimum and maximum of the data.

Figure 2.4 Relationship between classroom indoor $PM_{2.5}$ source contributions and I/O sulfur ratio by seasons: Blue dots (spring); Red dots (fall); and Green dots (winter).

Figure 3.1 SHAP dependence plot for A) soil temperature at 800 cm depth, B) elevation, C) upward long wave radiation flux, D) dew point, E) downward long wave radiation flux, and F) temperature in Massachusetts. The plotted functions are interpreted as the effect of the feature's values on the model's prediction. For example, the predicted radon level was high in areas where the elevation was above 100 m.

Figure 3.2 SHAP dependence plot for A) elevation, B) uranium, C) soil temperature at 800 cm depth, D) sand fraction, E) clay fraction, and F) land-surface slope in Michigan. The plotted functions are interpreted as the effect of the feature's values on the prediction. For example, in areas where the elevation was above 250 m, the predicted radon level was high.

Figure 3.3 SHAP dependence plot for A) clay fraction, B) sand fraction, C) land-surface slope, D) soil porosity @ 2-150 cm depth, E) organic matter, and F) soil moisture in Minnesota. The plotted functions are interpreted as the effect of the feature's values on the prediction. For example, in areas where the sand fraction was above 0.4, the predicted indoor radon level was low.

Figure 4.1 Contribution of wells to indoor radon levels.

Figure S2.C Correlation between sources contributing to indoor and outdoor $PM_{2.5}$ by season. Blue dots represent samples collected during the spring season. Red dots represent samples collected during fall season. Green dots represent samples collected during winter season.

Figure S2.D Source contribution by schools.

Figure S2.E.1 Observed vs Predicted Concentrations of indoor $PM_{2.5}$ mass.

Figure S2.E.2 Observed vs Predicted Concentrations of outdoor $PM_{2.5}$ mass.

Figure S3.1 SHAP summary plot of the top 20 influential factors of radon levels in Massachusetts. The plot shows the distribution of Shapley values per factor. Factors are ordered based on importance. The color represents the factor value (yellow indicates high factor value and purple indicates low factor value). For example, low soil temperatures at 800 cm depth increased the predicted radon levels.

Figure S3.2 SHAP summary plot of the top 20 influential factors of radon levels in Michigan. The plot shows the distribution of Shapley values per factor. Factors are ordered based on importance. The color represents the factor value (yellow indicates high factor value and purple indicates low factor value). For example, houses located at higher elevation had an increase in the predicted radon levels.

Figure S3.3 SHAP summary plot of the top 20 influential factors of radon levels in Minnesota. The plot shows the distribution of Shapley values per factor. Factors are ordered based on importance. The color represents the factor value (yellow indicates high factor value and purple indicates low factor value). For example, houses located in areas with higher clay fraction in the soil had an increase in the predicted radon levels.

LIST OF TABLES

Table 2.1 Summary of weekly indoor and outdoor concentrations of $PM_{2.5}$ mass and its components.

Table 2.2 Indoor and outdoor $PM_{2.5}$ concentrations and its possible sources from various schools.

Table 2.3 Source contributions to indoor $PM_{2.5}$ mass and elemental concentration (ng/m^3).

Table 2.4 Source contributions to outdoor $PM_{2.5}$ mass and elemental concentration (ng/m^3).

Table 2.5 Sources affecting $PM_{2.5}$ by season.

Table 3.1 Variable importance based on SHAP values.

Table 4.1 Linear mixed model for unconventional wells.

Appendix S2.A PMF input and diagnostics summary.

Appendix S2.B PMF base error estimation summary.

Table S3.1 Rationale to include explanatory variables to the model.

Table S3.2 Summary of zip codes level measurements.

Table S3.3 Range of values used for hyper-parameter tuning.

Table S3.4 Annual variation of indoor radon levels. Average and interquartile range (Bq/m^3).

Table S3.5 Monthly variation of indoor radon levels across states. Average and interquartile range (Bq/m^3).

Table S4.1 Annual variation of residential radon levels.

Table S4.2 Monthly variation of residential radon levels.

Table S4.3 Exposure distribution.

DEDICATION

To my Mom and Dad,

and to my grandmothers,

for always loving, supporting, and

giving me strength throughout this journey.

ACKNOWLEDGEMENTS

First, I would like to thank my advisor and committee chair Dr. Petros Koutrakis for giving me the opportunity to work with him at Harvard. His guidance, shared knowledge, skills, and experience have been of great help to my research. Also, I would like to thank my research committee members Dr. Philip Demokritou and Dr. Joel Schwartz for their time, wisdom, guidance, and for their courses, which were a great foundation for this work.

I would like to thank my collaborators who provided helpful comments, assistance and made this work possible. I would like to thank Dr. Diane Gold, for her help, guidance, and kindness; Mike Wolfson, for sharing his knowledge and wisdom in laboratory procedures; Dr. Ramon Sanchez, for his guidance on my teaching fellow work and for his great advice. A special thanks to Choong-Min Kang who genuinely helped, encouraged, and gave me great advice all these years. Also, I appreciate his willingness to help me anytime and revise my presentations and writings.

I would like to express my appreciation to the Department of Population Health Sciences for their support. A big thanks to Bruce Villineau and Matthew Boccuzzi for their help in navigating graduate school processes and good conversations.

I would like to thank my family and friends for their great support during this journey even from miles away. I would like to thank my salsa dance team peers for their laughs and great dances.

Finally, I would like to thank my parents and my grandmothers for their great support every single day. Their love and encouragement kept me going in tough times.

Thank you all.

Aleshka Carrion-Matta

CHAPTER 1

Introduction

Introduction

1.1. Overview

Indoor air pollution is one of the leading contributors to global burden diseases, of which 10-30% are attributable to exposure to indoor generated pollutants (Morawska et al. 2013). According to the National Human Activity Pattern Survey, people in US spend around 87% of their time indoors, of which 69% is attributable to residential occupancy and 5% to workplace environment (Klepeis et al. 2001). Time spend indoors varies with season, age, gender, type of work, health conditions of inhabitants and so on. However, in the US, approximately 20% of buildings have serious and 40% have somewhat serious indoor air quality problems (Cretescu et al. 2019). Researchers have shown that indoor air pollution levels often differ from outdoors, representing 2-5 times higher than outdoors. Consequently, the exposure to indoor air pollution could be 100 times greater than in outdoors, causing harmful health risks such as allergic reaction, asthma, and lung cancer (Gaikwad and Shivhare 2019).

Indoor air quality in schools have been of interest in recent times due to the rising incidence of respiratory conditions in children who spend a substantial amount of their time in school premises. In U.S., about 60% of the schools have one or more building defects that can lead to poor indoor air quality (GAO 1996) which can further cause various health effects ranging from allergies, respiratory infections, and asthma, affecting student's performance and attendance. Different sources (e.g., cooking, cleaning activities, traffic, industry) can contribute to poor indoor air quality in schools, being the outdoor sources, the ones contributing greatly to indoor air. However, the penetration of the outdoor particles depends on factors related to building features, ventilation, and their physicochemical characteristics and size distributions (Rivas et al. 2014).

Some studies have explored the extent of poor indoor air quality in schools (Godwin and Batterman 2007), with some accounting it to school proximity to local traffic (Forns et al., 2016; Green Rochelle et al., 2004) and the traffic encountered on the walk to school (Alvarez-Pedrerol et al., 2017). However, apart from studies on school proximity to traffic, little is known about children's exposure in schools and classrooms in U.S. since school-based studies are lacking. In Chapter 1, as a part of the School Inner City Asthma Study (SICAS), using measurements of $PM_{2.5}$ mass, trace elements and black carbon (BC), we provide detailed characterization of indoor and outdoor sources affecting indoor $PM_{2.5}$. This study helps us close the gap in the absence of information about children's personal exposures in these microenvironments.

Depending on location, pollutants emitted from different sources can entrain to the indoor environment through different mechanisms. There are three main mechanisms that allow the entrainment of outdoor pollutants to indoors (Leung 2015): mechanical ventilation, natural ventilation, and infiltration. Mechanical ventilation circulates fresh air drawn in from outdoor air by ventilation fans, air conditioning in a residence, or by air conditioning system of a building. Natural ventilation occurs whenever the windows and doors of a space are open and relies on pressure differences between the inside and the outside to move fresh air. Infiltration occurs through cracks and leaks in the building shell structure and relies on pressure differences. These mechanisms along with factors varying throughout the year such as temperature gradients, relative humidity, wind direction and speed; allow that consequently these pollutants dilute and accumulate in the indoor environment, exposing its inhabitants to adverse health effects.

Radon main mechanisms of entrainment to the indoor environment is through the cracks of the building. Radon is generated within the Earth's crust and rises towards the surface and into the air by diffusion and convection (Dueñas et al. 1997). Exposure to radon is associated to lung

cancer, even at levels below the action levels (Vogiannis and Nikolopoulos 2015). Radon exposure is a multifactorial problem and some studies (Branion-Calles et al. 2016; Dai et al. 2019; Steck 2009; Xie et al. 2015) have shown that indoor radon concentrations is related to factors such as: (1) atmospheric conditions, (2) occupant behavior, (3) soil, (4) building characteristics, (5) geology, and (6) spatial and time variations. However, there is no consensus about which factors are the most important. In Chapter 3, we analyzed residential radon levels in different geographical areas to determine which factors are the most influential.

The U.S. EPA estimated that approximately 1 in 15 homes exceed the threshold level at which mitigation is recommended (4 pCi/L) (EPA 2012). Pennsylvania is one of the states with more than 40% of the houses with levels above the remediation level. The recent technology advances in the oil and gas industry in the state brings environmental and public health concerns. One of the major concerns among those that oppose to oil and gas production are the harmful chemicals that are brought up during extraction. In the extraction process, toxic flowback water, production brine and naturally occurring radioactive materials (such as Uranium (^{238}U), and decay products such as radium (^{226}Ra) and radon (^{222}Rn)) are brought up to the surface (Finkel and Hays 2016). Literature relating indoor radon to these oil and gas practices is scarce and there is no consensus whether there is an exposure risk from living nearby oil and gas development sites. In Chapter 4, we assessed whether proximity to oil and gas areas affect indoor radon levels.

The assessment of outdoor pollution has proven to be essential in obtaining a framework of pollution at different spatial and temporal scales but there is a limitation in understanding the adverse health effects of air pollutants due to the inability to directly equate outdoor concentration to personal exposure. The correlation between personal exposure and outdoor concentration is complex, variable and not fully developed in the literature, and little is known about permissible

exposure levels for homes and nonindustrial occupational settings. Until 1970, it was believed that indoor air quality was solely influenced by outdoor air concentrations and since then, and due to energy crisis, construction and building renovations were focused on the tightness of the building (Śmiełowska et al. 2017). However, these changes have created and increase in indoor pollutant concentrations as confirmed by more than 2/3 of studies that investigates both the indoor and outdoor environment, which have found that indoor air concentration is higher than outdoors (Chen and Zhao 2011; Leung 2015). As the time spent indoors is large enough and increasing over time, understanding indoor sources, transport and the physicochemical characteristics of the air pollutants justifies the importance of conducting more studies to enhance our understanding of exposures in the indoor environment.

1.2 Aims

This dissertation is divided into two main parts. The first part includes Chapter 2 that seeks to identify indoor and outdoor $PM_{2.5}$ sources affecting indoor exposure to children attending inner-city schools to help close the gap in the absence of information about children's personal exposures in these microenvironments. In the study, we ran a source apportionment analysis in inner city schools in the Northeastern U.S. between 2008-2013. In the second part, which consists of Chapter 3 and 4, we aimed to study indoor radon pollution in residential environments. Chapter 3 focus on examining the relationship between indoor radon levels and atmospheric conditions, soil, geology, topography, radiation factors, urbanization, and economic well-being. In the study, we ran a random forest model in Massachusetts, Michigan, and Minnesota between 2005-2018 to determine the main predictors of indoor radon levels. In Chapter 4, we assessed whether proximity to oil and gas practices affect indoor radon levels. In the study, we ran linear mixed effects models with a random intercept for zip code in Pennsylvania between 2001-2015.

CHAPTER 2

Classroom indoor $PM_{2.5}$ sources and exposures in inner-city schools

Classroom indoor $PM_{2.5}$ sources and exposures in inner-city schools

Authors

Aleshka Carrion-Matta[a], Choong-Min Kang[a], Jonathan M. Gaffin[b,c], Marissa Hauptman[b,d], Wanda Phipatanakul[b,e], Petros Koutrakis[a], Diane R. Gold[a,b,f]

Affiliations

[a]Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA

[b]Harvard Medical School, Boston, Massachusetts, USA

[c] Division of Respiratory Diseases, Boston Children's Hospital, Boston, Massachusetts, USA

[d]Division of General Pediatrics, Boston Children's Hospital, Boston, Massachusetts, USA

[e]Division of Allergy and Immunology, Boston Children's Hospital, Boston, Massachusetts, USA

[f] Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA

Acknowledgments

This publication was made possible by the National Institutes of Health (NIH) grants (R01 AI 073964, K23AI106945, U01 AI 110397, K24 AI 106822, R01HL137192), Coway grant (ISRA-2017) and US EPA grant (RD-835872). Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the granters. Further, the granters do not endorse the purchase of any commercial products or services mentioned in the publication. We

acknowledge the efforts of Jack M. Wolfson, all other laboratory and field staffs for laboratory and field measurement support.

Abstract

Children spend over 6 h a day in schools and have higher asthma morbidity from school environmental exposures. The present study aims to determine indoor and outdoor possible sources affecting indoor $PM_{2.5}$ in classrooms. Weeklong indoor $PM_{2.5}$ samples were collected from 32 inner-city schools from a Northeastern U.S. community during three seasons (fall, winter and spring) during the years 2009 to 2013. Concurrently, daily outdoor $PM_{2.5}$ samples were taken at a central monitoring site located at a median distance of 4,974 m (range 1,065 –11,592 m) from the schools. Classroom indoor concentrations of $PM_{2.5}$ (an average of 5.2 $\mu g/m^3$) were lower than outdoors (an average of 6.5 $\mu g/m^3$), and these averages were in the lower range compared to the findings in other schools' studies. The USEPA PMF model was applied to the $PM_{2.5}$ components measured simultaneously from classroom indoor and outdoor to estimate the source apportionment. The major sources (contributions) identified across all seasons of indoor $PM_{2.5}$ were secondary pollution (41%) and motor vehicles (17%), followed by Calcium (Ca)-rich particles (12%), biomass burning (15%), soil dust (6%), and marine aerosols (4%). Likewise, the major sources of outdoor $PM_{2.5}$ across all seasons were secondary pollution (41%) and motor vehicles (26%), followed by biomass burning (17%), soil dust (7%), road dust (3%), and marine aerosols (1%). Secondary pollution was the greatest contributor to indoor and outdoor $PM_{2.5}$ over all three seasons, with the highest contribution during spring with 53% to indoor $PM_{2.5}$ and 45% to outdoor $PM_{2.5}$. Lower contributions of this source during fall and winter are most likely attributed to less infiltration indoors. In contrast, the indoor contribution of motor vehicles source was highest in the fall (29%) and winter (25%), which was presumably categorized by a local source. From the relationship between indoor-to-outdoor sulfur ratios and each source contribution, we also estimated the local and regional influence on indoor $PM_{2.5}$ concentration.

Overall, the observed differences to indoor $PM_{2.5}$ are related to seasonality, and the distinct characteristics and behavior of each classroom/school.

2.1 Introduction

Air pollution, including particulate matter with aerodynamic diameter < 2.5 μm ($PM_{2.5}$) has been associated with asthma exacerbation and cognitive impairment (Sunyer et al., 2015; Tétreault et al., 2016) in children but the $PM_{2.5}$ sources and components most responsible for these health effects are imperfectly understood. $PM_{2.5}$ sources and components vary by the microenvironments in which children live, learn and play. The school classroom, where children spend a significant portion of their day, has been identified as one of the most important microenvironments where $PM_{2.5}$ exposure with relevance to health may occur. School proximity to local traffic (Forns et al., 2016), and traffic encountered on the walk to school (Alvarez-Pedrerol et al., 2017), have been implicated in studies of pollution and poorer neurocognitive testing in school children. However, apart from studies on school proximity to traffic (Forns et al., 2016; Green Rochelle et al., 2004), limited data are available on the sources or components of $PM_{2.5}$ present in the classroom environment that may account for adverse health outcomes like asthma morbidity or neurocognition (Gaffin et al., 2017). One Barcelona study implicated indoor school PM from traffic, but not from other sources, as an influence on child neurocognition (Basagaña et al., 2016).

Key environmental and behavioral factors such as occupancy, ventilation rates, building characteristics and seasonality may influence the school indoor $PM_{2.5}$ concentration and composition, including the amount of traffic $PM_{2.5}$ that can come indoors. School buildings characteristics such as age, building materials, and ventilation; and factors such as occupant density per classroom volume, affect the transport and mixing of pollutants, which pose a risk to the health of the occupants, often affecting their performance in school (MacNaughton et al., 2017). Koo, et al. (Koo et al., 1997) found that student learning decreased in natural ventilation as

compared to air-conditioned classrooms. Building materials can act as reservoirs of irritants that can contribute to the gas-to-particle processes, affecting the indoor air quality of the classrooms (Smedje and Norbäck, 2008). Occupant activities can increase the concentration of pollutants by resuspension of previously deposited particles and by introducing new particles through clothing and shoes (Stranger et al., 2008a). Factors that vary throughout the year such as temperature gradients between indoor and outdoor, relative humidity, wind direction and speed have an effect on the penetration rate of outdoor pollutants to indoor settings (Chatzidiakou et al., 2012). Yearlong studies examining seasonal variability of PM levels in schools are scarce and the limited investigations have shown differences in PM concentrations between summer and winter where the temperature difference between indoor and outdoor might affect the infiltration of the pollutants (Fromme et al., 2007; Goyal and Khare, 2009).

Inner-city areas tend to have a higher level of asthma morbidity (Gergen and Togias, 2015). In particular, higher rates of hospitalizations and asthma morbidity were found in these impoverished areas compared to other urban neighborhoods (Weiss et al., 1992). People living and working in these environments experience high exposure to pollutants such as elemental carbon (EC) given their proximity to highways (Spira-Cohen et al., 2011).

The School Inner City Asthma Study (SICAS) investigates the effect of classroom-based environmental exposures on students with asthma in the northeast United States. In a recent SICAS publication, using the indoor-to-outdoor sulfur ratio as a tracer we showed that the infiltration of outdoor $PM_{2.5}$ varied by school and that there were indoor contributions to indoor $PM_{2.5}$ levels (Gaffin et al., 2017). As a part of the SICAS study, using measurements of $PM_{2.5}$ mass, trace elements and black carbon (BC), we provide detailed characterization of indoor and outdoor sources affecting indoor $PM_{2.5}$. Our aim is to identify indoor and outdoor $PM_{2.5}$ sources affecting

indoor exposure to children attending inner-city schools to help close the gap in the absence of information about children's personal exposures in these microenvironments

2.2 Materials and methods

Studied schools and central outdoor site

Indoor measurements were conducted in classrooms of inner-city schools during occupied and non-occupied periods during the school week but not on weekends. One or two samples per classroom were collected in different seasons (1 per season: either fall, winter, or spring) during the academic years 2009–2013. In total, 139 classrooms across 32 schools were studied. Details of school characteristics and locations are reported in Gaffin et al. (2017). The study was conducted in schools from a single urban metropolitan location. Briefly, the schools in this study varied widely in age (built between 1899 and 2002, with many built before the 1970's) and in the type of ventilation. Often the buildings within the school were of varying ages with varying types of ventilation. None of the schools had gas cooking in the building and all had recently serviced furnaces. Of the 32 schools studied in this source apportionment analysis, 4 had central HVAC systems, 15 had radiant heat “natural ventilation”, and the remainder had either classroom-based unit vents or a mix of ventilation types. Most of the classrooms (96%) were in the 2nd and ≥3rd levels of the building, and the flooring consisted of linoleum/tile (68%) and rugs (80%). For outdoor measurements, a central site was located on the roof of a 6-story building, which is not adjacent to local highways. The median distance between the central site and schools was 4974 m, ranging from 1065 to 11,592 m.

Measurements and analyses

Indoor samplers were used to collect weeklong indoor $PM_{2.5}$ sample on Teflon filters at a flow rate of 1.8 L/min (Demokritou et al., 2001). A total of 205 indoor samples were collected

during the study period. Concurrently, daily outdoor $PM_{2.5}$ concentrations were also measured at the central monitoring site. Indoor and outdoor $PM_{2.5}$ samples collected on Teflon filters including blanks were weighed with an electronic microbalance (MT-5 Mettler Toledo, Columbus, OH) prior to measurement the filters were equilibrated for a period of 48 h in a room of controlled temperature (22 ± 1.5 °C) and relative humidity (40 ± 5%). Following the weighing, the indoor filters were measured for BC concentrations using a Smokestain Reflectometer (Model EEL M43D, Diffusion Systems Ltd., United Kingdom). X-Ray Fluorescence (XRF) spectroscopy (model Epsilon 5, PANalytical, The Netherlands) was used to determine the elemental composition of indoor and outdoor $PM_{2.5}$ samples. A total of 48 elements were measured and quality Assurance/ Quality Control of XRF analysis is described in Kang et al. (2014). At the central site, hourly elemental (EC) and organic carbon (OC) were measured on pre-fired quartz fiber filters based on the thermal optical transmittance (TOT) technique using a semi-continuous OCEC field analyzer (Model 3, Sunset Laboratory Inc., Tigard, OR). Continuous BC concentrations were measured using a single (λ=880 nm) channel aethalometer (model AE-16, Magee Scientific, Berkeley, CA). Hourly and daily data from the central outdoor site were averaged and matched over indoor sampling periods each classroom. All analyses were conducted at the Harvard T.H Chan School of Public Health.

The indoor-to-outdoor sulfur concentration ratios were estimated for every sampling period and classroom. This ratio has been widely used in previous studies to approximate the infiltration fraction of $PM_{2.5}$ from outdoors (Sarnat et al., 2002; Wallace and Williams, 2005). Assuming negligible indoor sources of sulfur, the indoor-to-outdoor sulfur ratios were used as an infiltration factor for each classroom.

Source apportionment

The US Environmental Protection Agency Positive Matrix Factorization (USEPA PMF 5.0) model was used to determine the possible sources of indoor and outdoor $PM_{2.5}$. The PMF model has been described in detail elsewhere (Paatero and Tapper, 1994). The uncertainty was calculated using the equation below:

$$u_{ij} = s_{ij} + \frac{DL_{ij}}{3}$$

where s_{ij} is the analytical uncertainty, and DL_{ij} is the method detection limit (Polissar et al., 1998; Reff et al., 2007).

Three validation methods were used to assess the stability and variability of the results: (1) bootstrap (BS), which helps to identify the observations that could influence the solution; (2) displacement (DISP), which helps to understand the final solution and its sensitivity to small changes; and (3) BS-DISP analysis, which is a hybrid approach that combines BS and DISP methods. The results from these validation methods are included in Appendices A and B. The goodness of fit for the PMF solutions was examined using the percent Mean Relative Errors (MRE) (Masri et al., 2015):

$$MRE = \left|\frac{x_i - \widehat{x_i}}{x_i}\right| \times 100$$

where, x_i is the measured concentration and $\widehat{x_i}$ is the predicted concentration of species i.

Elements with >80% of their concentration values below the limit of detection were excluded from further analysis. Exception to this exclusion criteria for indoor analysis were Al which is a tracer of soil; Cu, and Pb, which are tracers of vehicle sources; and V and Ni, which are tracers of oil combustion. The exception for outdoor analysis was Cu. Consequently, 15 elements of 48 elements analyzed were included for further analysis. The PMF model categorizes the species based on their signal-to-noise ratio. Species categorized as ‘bad’ had a signal-to-noise ratio < 0.2,

'weak' species had 0.2 ≤ signal-to-noise ratio < 2, and 'strong' species had a signal-to-noise ratio ≥ 2. Species categorized as 'bad' were excluded from the analysis, 'strong' species were included in the analysis and remained unchanged, and 'weak' species were also included in the analysis but were down-weighted by tripling their uncertainty values (Paatero and Hopke, 2003). PMF was run multiple times using 4 to 8 factors for both indoor and outdoor $PM_{2.5}$. The results chosen were the ones more interpretable among the different solutions. Sources were assigned according to the abundance of key tracer elements.

2.3 Results and Discussion

$PM_{2.5}$ mass concentrations and chemical compositions

Table 2.1 represents the summary of weeklong concentrations of indoor and outdoor $PM_{2.5}$ mass and components. Outdoor concentrations from the central site were averaged over indoor sampling periods at each classroom. Sulfate was calculated stochiometrically from sulfur by multiplying by three, assuming that all sulfur is in the form of sulfate. On average, outdoor concentrations of $PM_{2.5}$ (6.5 μg/m^3) were higher than indoors (5.2 μg/m^3). As shown in Table 2.2, these averages are in the lower range compared to other studies performed in other cities of US (Bozlaker et al. 2017; Hochstetler et al. 2011; Keeler et al. 2002), Europe (Almeida et al. 2011; Amato et al. 2014; Borgini et al. 2011; Canha et al. 2014; Fromme et al. 2008; Janssen et al. 2001; Jovanović et al. 2014; Stranger M et al. 2008; Wichmann et al. 2010; Zwoździak et al. 2013), China (Hou et al. 2015; Xu et al. 2015), and India (Chithra and Nagendra 2012; Habil et al. 2013). In the other US cities, studies of Texas (Bozlaker et al. 2017) and Michigan (Keeler et al. 2002) schools reported that outdoor concentrations were 13.4 and 11.6–20.6 μg/m^3, respectively. In Europe, the schools in Spain (29 μg/m^3) (Amato et al. 2014), Germany (17.0 μg/m3) (Fromme et al. 2008), Belgium (63.0 μg/m^3) (Stranger M et al. 2008), Sweden (9.7 μg/m^3) (Wichmann et al.

2010), Italy (77.9 µg/m^3) (Borgini et al. 2011) and the Netherlands (24.8 µg/m3) (Janssen et al. 2001) also had on average higher outdoor concentrations than our study. In our study, outdoor daily $PM_{2.5}$ concentrations did not exceed the daily NAAQS $PM_{2.5}$ level of 35 µg/m^3. Among the 15 elements, Sulfate as S was the most abundant, followed by Na, Cl, Ca, Fe and Si. These elements were also abundant for both indoor and outdoor concentrations. The mean indoor-to-outdoor (I/O) $PM_{2.5}$ across the schools' studied was 0.8. The ratio was lower than 1.3 in Belgium (Stranger M et al. 2008) and 0.98 in Serbia (Jovanović et al. 2014) under natural ventilation. Also, these were lower than the schools studies in India where the I/O $PM_{2.5}$ was 1.4 in the study and had natural ventilation (Chithra and Nagendra 2012). In schools in China, the results were comparable to the ones in our study where the I/O $PM_{2.5}$ was 0.8 and also had natural ventilation (Xu et al. 2015). For the studies in the other US cities, the results were different: in Texas, the I/O $PM_{2.5}$ was lower with 0.2, but this school had a mechanical ventilation system; and in Michigan the I/O $PM_{2.5}$ ranged from 0.5 to 2.8 (Bozlaker et al. 2017; Keeler et al. 2002)

Table 2.1 Summary of weekly indoor and outdoor concentrations of $PM_{2.5}$ mass and its components

	Indoor			Outdoor		
Pollutant	N	Mean	SD	N	Mean	SD
Units: µg/m^3						
$PM_{2.5}$ Mass	205	5.2	2.3	205	6.5	2.7
BC	205	0.3	0.2	205	0.6	0.3
OC	NA	NA	NA	195	2.9	1.2
Sulfate	205	1.1	0.5	205	1.5	0.6
Units: ng/m^3						
Al	205	6.0	11.0	205	23.0	19.0
Ca	205	40.0	23.0	205	21.0	7.0
Cl	205	52.0	95.0	205	33.0	56.0
Cu	205	3.0	2.0	205	2.0	0.9
Fe	205	31.0	20.0	205	46.0	13.0
K	205	32.0	14.0	205	38.0	16.0
Mg	ND	ND	ND	205	22.0	17.0
Na	205	92.0	69.0	205	128.0	60.0

Ni	205	0.7	0.7	ND	ND	ND
Pb	205	3.0	3.0	205	3.0	1.0
Si	205	30.0	24.0	205	44.0	29.0
Ti	205	2.0	2.0	205	2.0	1.0
V	205	0.7	0.9	ND	ND	ND
Zn	205	6.0	4.0	205	8.0	4.0

NA: not available
ND: below lower detection limit

Table 2.2. Indoor and outdoor $PM_{2.5}$ concentrations and its possible sources from various schools

Reference	Location	# Schools	Concentration range (μg/m^3)		Ventilation system	Findings
			Indoor	Outdoor		
This study	Northeast, USA	32	0.8-16.9	3.3-16.6	17 natural, 3 central HVAC, 12 mixed	I/O $PM_{2.5}$ = 0.8; Sources: Ca-rich material, road dust, sea salt, vehicles, biomass burning, regional.
Hochstetler et al. (2011)	Ohio, USA	4	6.9-28.3	3.8-27.6	3 schools open windows	I/O $PM_{2.5}$ = 0.5-2.8
Keeler et al. (2002)	Michigan, USA	2	6.5-16.4	11.6-20.6	Opening windows, no air conditioning	I/O $PM_{2.5}$ = 0.5-1.4
Bozlaker et al. (2017)	Texas, USA	1	2.3-4.1	13.4[ǂ]	Mechanical	I/O $PM_{2.5}$ = 0.2; Sources: soil & road dust, vehicular emissions, petroleum refining, oil combustion, coal combustion, vegetative burning, sea salt, ca-rich material, incineration, steel plant
Amato et al. (2014)	Barcelona, Spain	39	7-105	1-192	Natural	Sources: organic/ textile/ chalk, heavy oil, metallurgy, sulfate & organics, nitrate, traffic, road dust, mineral, sea salt
Jovanović et al. (2014)	Serbia	1	26.9-63.9	--	Natural	I/O $PM_{2.5}$ = 1.0
Fromme et al. (2008)	Munich, Germany	1	19.3-105.9	5.1-67.4	Natural	Indoor PM consists mainly of crustal materials, detrition of the building materials and chalk
Canha et al. (2014)	Rural Portugal	1	100[ǂ]	--	Natural	Sources: bakery industry, wood burning process/ soil re-suspension/ chalk use, crustal, and marine contributions
Almeida et al. (2011)	Lisbon, Portugal	3	10[ǂ]	3 - 10	Natural	Elemental composition suggested sources: crustal materials, detritions of the building materials and chalk
Zwoździak et al. (2013)	Wroclaw, Poland	1	13.5-59.8	9.1-15.6	Natural	I/O $PM_{2.5}$ = 2.0 for winter, 4.1 for summer
Stranger et al. (2008)	Antwerp, Belgium	27	26-129	12-148	Opening windows	I/O $PM_{2.5}$ = 1.3
Borgini et al. (2011)	Milan, Italy	3	79.4[ǂ]	77.9[ǂ]	Natural	I/O $PM_{2.5}$ = 1.0
Janssen et al. (2001)	Netherlands	24	7.7-52.8	5.2-60.8	--	I/O $PM_{2.5}$ = 0.9

Wichmann et al. (2010)	Stockholm, Sweden	6	2.8-13.9	5.2-24.2	Mechanical	I/O $PM_{2.5}$ = 0.9
Hou et al. (2015)	Beijing, China	2	11 - 79	17 - 87	Natural	I/O $PM_{2.5}$ = 0.6 (closed), 0.9 (open) windows and doors
Xu et al. (2015)	Xi'an, China	1	141.8[ǂ]	167.8[ǂ]	Natural	I/O $PM_{2.5}$ = 0.8; sources: coal combustion, motor vehicle exhaust, other primary sources, and secondary formation
Habil et al. (2013)	Agra, India	10	71.3-90.1	--	Natural	I/O $PM_{2.5}$ = 1.0 for winter, 1.2 for summer, 1.1 for monsoon; Sources: vehicle, soil, metal processes, wind-blown dust, industrial
Chithra and Nagendra (2012)	Chennai, India	1	32-61	--	Natural	I/O $PM_{2.5}$ = 1.4 ± 0.7

Indoor and outdoor sources contributing indoor $PM_{2.5}$

The PMF receptor model identified six source factors that contributed to indoor $PM_{2.5}$ concentrations as shown in Table 2.3 and Figure 2.1. Percent mean relative errors (%MRE) were used to assess the goodness of fit for the PMF model solution. The results showed a good prediction of indoor $PM_{2.5}$ mass and elemental concentrations with low MREs (<15%), except for Al (27.6%), which had concentration values below the limit of detection.

Ca-rich particles, motor vehicles, secondary pollution, soil dust, biomass burning and marine aerosols were identified as sources affecting $PM_{2.5}$ in school classrooms, which account for 12, 17, 41, 6, 15, and 4% of the total mass, respectively. The first factor, Ca-rich particles was characterized by the high contributions of Ca, Fe, and Ti, which suggests the contribution of road-dust from outdoors, or possibly the wear-out of cement and dry-wall indoors. The second factor, soil dust, was characterized by the contributions of terrestrial elements Al and Si. Overall, we observed sulfate mass concentration in every factor, although its relative contribution to indoor sources was, as expected, low. Other studies (Song et al. 2001) in this region have shown the contribution of sulfate to their source profiles due to the contribution of coal-fired power plants located in the Midwest USA. For the Ca-rich particles and soil dust, the crustal particles can enter classrooms through windows or be released from the children's shoes and clothes, especially after playing outside. These two may be related to children's activities as has been reported by other school studies in Texas, Germany, Spain and Portugal (Almeida et al. 2011; Bozlaker et al. 2017; Canha et al. 2014; Fromme et al. 2008). The study of Bozlaker et al. (2017) in Texas found contributions of 2.1 and 10.2% from Ca-rich particles and soil/road dust, respectively. Canha et al. (2014) reported a 27% contribution from soil and a contribution of 53% from chalk/resuspension in a rural area of Portugal. In Barcelona, Amato et al. (2014) estimated contributions of 17% for a

mineral source and 45% for textile/chalk source. It is worth noting that in this study's school classrooms chalk is not used, the possible sources of Ca-rich particles might be due to cement or drywall wear-out processes. The third factor was identified as secondary pollution because it was associated with sulfate, which is a tracer of regional and long-range transport of particulate air pollution (Ling-Da et al. 2013; Viana et al. 2008). Nitrate and ammonium markers have been used in source apportionment studies to identify a secondary nitrate factor, but the nitrate concentration is low in the Northeast USA, with mean concentration of 0.79 $\mu g/m^3$, representing a $PM_{2.5}$ mass fraction of 9% during the study period. This factor was also associated with V and Ni, which are tracers of oil combustion. However, since oil combustion was not significant, it might not be identified as an individual factor. The fourth factor was identified as the motor vehicles with a high contribution of BC, which is a clear tracer of vehicular emissions (Viana et al. 2008). The fifth factor was related to marine aerosols because of its association with Na and Cl, which are major constituents of sea water (Mazzei et al. 2008; Viana et al. 2008). The sixth factor was identified as biomass burning because was characterized by contributions from K and Zn, which are tracers of biomass burning (Echalar et al. 1998).

In Figure S2.D in Appendix S2.D, we show the percentage contribution of these sources to indoor $PM_{2.5}$ in each school. The secondary pollution factor was the dominant source among all the schools followed by motor vehicles factor and biomass burning. Low contribution from the marine aerosols factor was also observed. The unaccounted fraction was defined as: total measured mass – total modeled mass. This mass fraction could be contribution from secondary nitrate or other secondary organic aerosols.

The source apportionment analysis of outdoor $PM_{2.5}$ concentrations identified six source factors shown in Table 2.4 and Figure 2.2. Goodness of fit showed a good prediction of outdoor

$PM_{2.5}$ mass and elemental concentrations with low (<15%) of MREs, except for Cl (18%). The factors included: biomass burning (17%), road dust (3%), soil dust (7%), secondary pollution (41%), motor vehicles (26%), and marine aerosols (1%). The first factor, biomass burning, was associated with K, which is a tracer of biomass burning. The second factor was identified as road dust because of the high contributions of Ca and Fe. Soil dust, the third factor, was characterized by the high contributions of Al and Si. The fourth factor was identified as secondary pollution due to high contributions of Sulfate and Na. The fifth factor was identified as the motor vehicles source for having high contributions from BC, OC, Cu, and Zn. The last factor was marine aerosols and identified with high contribution of Cl. Contrary to indoor environment, the ambient molar ratio of Na to Cl could not be estimated because a considerable amount of particle Cl could be converted to gas-phase HCl by the reactions with HNO_3 and H_2SO_4 in urban atmosphere (Seinfeld and Pandis 2016). For this reason, we used Na and Cl as a different source tracer for outdoor $PM_{2.5}$.

Table 2.3. Source contributions to indoor $PM_{2.5}$ mass and elemental concentration (ng/m^3)

	Ca-rich particles	Motor vehicles	Secondary pollution	Soil dust	Marine aerosols	Biomass burning	Estimate	Measured	%MRE
Mass	619.1	856.9	2,142.1	302.4	207.3	787.3	4,915.1	5,247.9	6.3
BC	5.0	212.7	13.8	0.0	4.9	21.9	258.3	260.0	0.7
Sulfate	0.0	211.9	668.6	0.0	8.6	184.8	1,073.9	1,074.6	0.1
Al	0.0	0.0	0.0	7.4	0.0	0.0	7.4	5.8	27.6
Ca	27.6	1.5	3.0	0.5	4.1	3.1	39.8	39.9	0.3
Cl	0.0	0.0	0.0	0.0	51.7	0.0	51.7	51.8	0.2
Cu	0.9	0.5	0.2	0.1	0.3	0.7	2.7	2.8	3.6
Fe	5.9	8.2	4.7	4.6	0.9	5.0	29.3	30.2	3.0
K	3.1	6.5	0.0	5.4	2.0	15.7	32.7	32.7	0.0
Na	9.2	4.5	23.1	2.7	36.8	13.0	89.3	92.7	3.7
Ni	0.2	0.1	0.2	0.0	0.0	0.1	0.6	0.7	14.3
Pb	1.1	0.0	0.1	0.5	0.2	1.3	3.2	3.5	8.6
Si	7.5	2.3	0.0	20.5	0.1	0.0	30.4	30.4	0.0
Ti	0.9	0.0	0.3	0.4	0.1	0.3	2.0	1.9	5.3
V	0.0	0.2	0.4	0.0	0.0	0.0	0.6	0.7	14.3
Zn	1.1	1.8	0.1	0.3	0.1	2.6	6.0	6.0	0.0

Table 2.4. Source contributions to outdoor $PM_{2.5}$ mass and elemental concentration (ng/m^3)

	Road dust	Motor vehicles	Secondary pollution	Soil dust	Marine aerosols	Biomass burning	Estimate	Measured	%MRE
Mass	209.2	1,698.7	2,630.9	434.8	75.0	1,107.2	6,155.8	6,497.3	5.3
BC	68.0	365.0	66.1	2.8	2.5	29.0	533.4	628.0	15.1
OC	400.3	846.7	485.8	0.0	13.3	811.3	2,557.4	2,868.6	10.8
Sulfate	0.0	287.3	1,314.8	0.0	9.4	40.0	1,651.5	1,508.4	9.5
Al	0.1	0.0	6.6	17.9	0.5	1.6	26.7	23.4	14.1
Ca	13.1	3.1	0.0	4.1	0.8	0.4	21.5	21.0	2.4
Cl	0.0	0.0	0.0	0.9	25.9	0.0	26.8	32.7	18.0
Cu	0.1	1.1	0.1	0.5	0.0	0.6	2.4	2.4	0.0
Fe	7.5	26.3	0.0	12.0	0.2	0.0	46.0	46.1	0.2
K	0.0	0.0	0.0	8.2	0.5	30.5	39.2	37.9	3.4
Mg	2.3	0.0	9.7	4.2	3.1	2.5	21.8	21.9	0.5
Na	25.5	13.2	49.6	0.0	24.7	1.4	114.4	128.7	11.1
Pb	0.3	1.3	0.4	0.5	0.0	0.5	3.0	2.9	3.4
Si	4.5	0.4	2.9	36.1	0.1	0.0	44.0	44.4	0.9
Ti	0.1	0.7	0.1	1.1	0.0	0.3	2.3	2.2	4.5
Zn	0.5	3.8	0.0	0.2	0.1	2.6	7.2	8.0	10.0

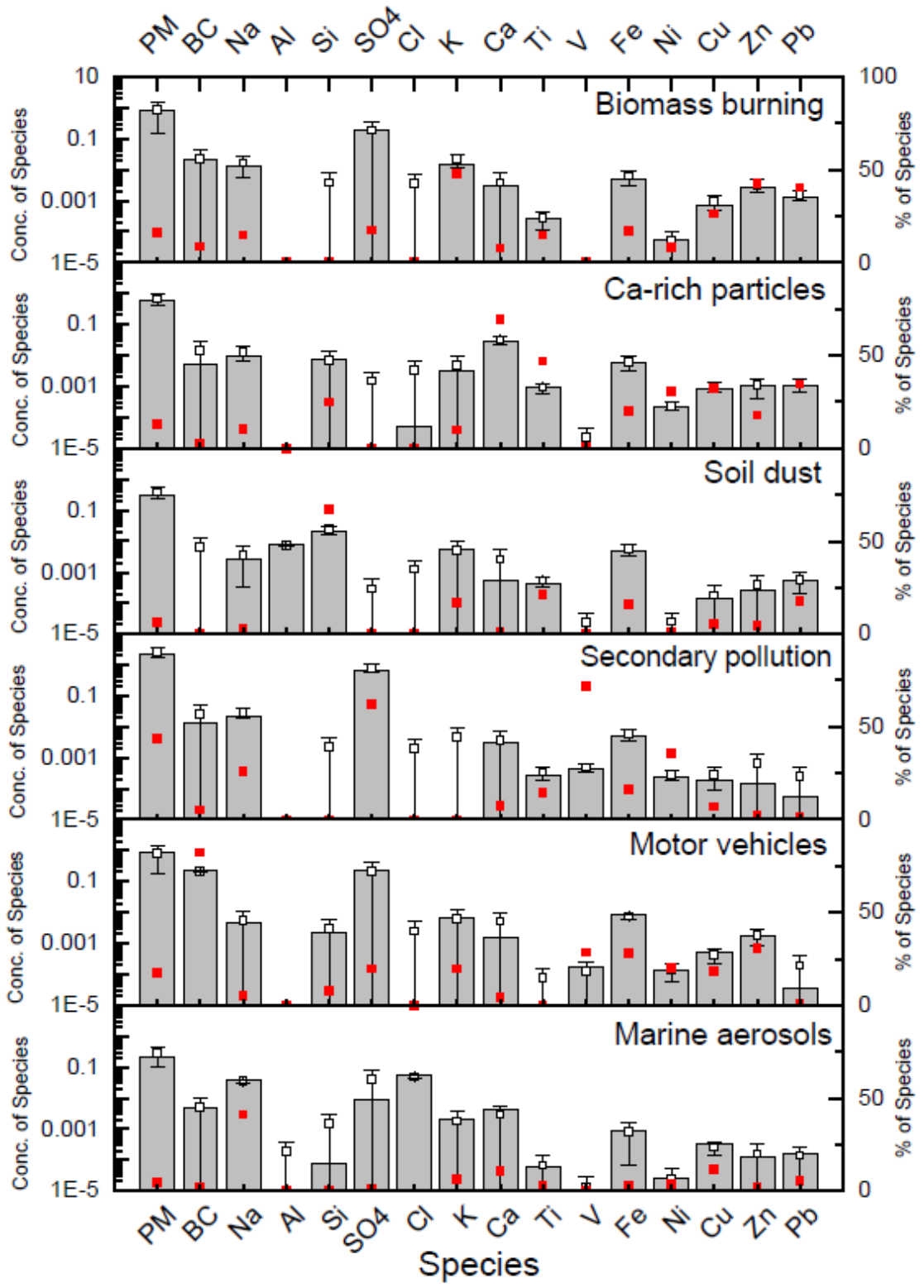

Figure 2.1. Source profiles contributing to indoor $PM_{2.5}$. The grey filled-in rectangle represents the estimated average concentration of each species by source factor estimated using PMF. The small closed (red) square represents the percentage of a specific species ascribed to the specific source (factor) by

PMF. For example, using PMF, the mean concentration of indoor SO_4 ascribed to secondary pollution is 0.7 $\mu g/m^3$, which is 62% of the total sulfate concentration (See also Table 2.3). The small open square represents the estimated average concentration of each species by source factor estimated using DISP with the error bars indicating the confidence intervals.

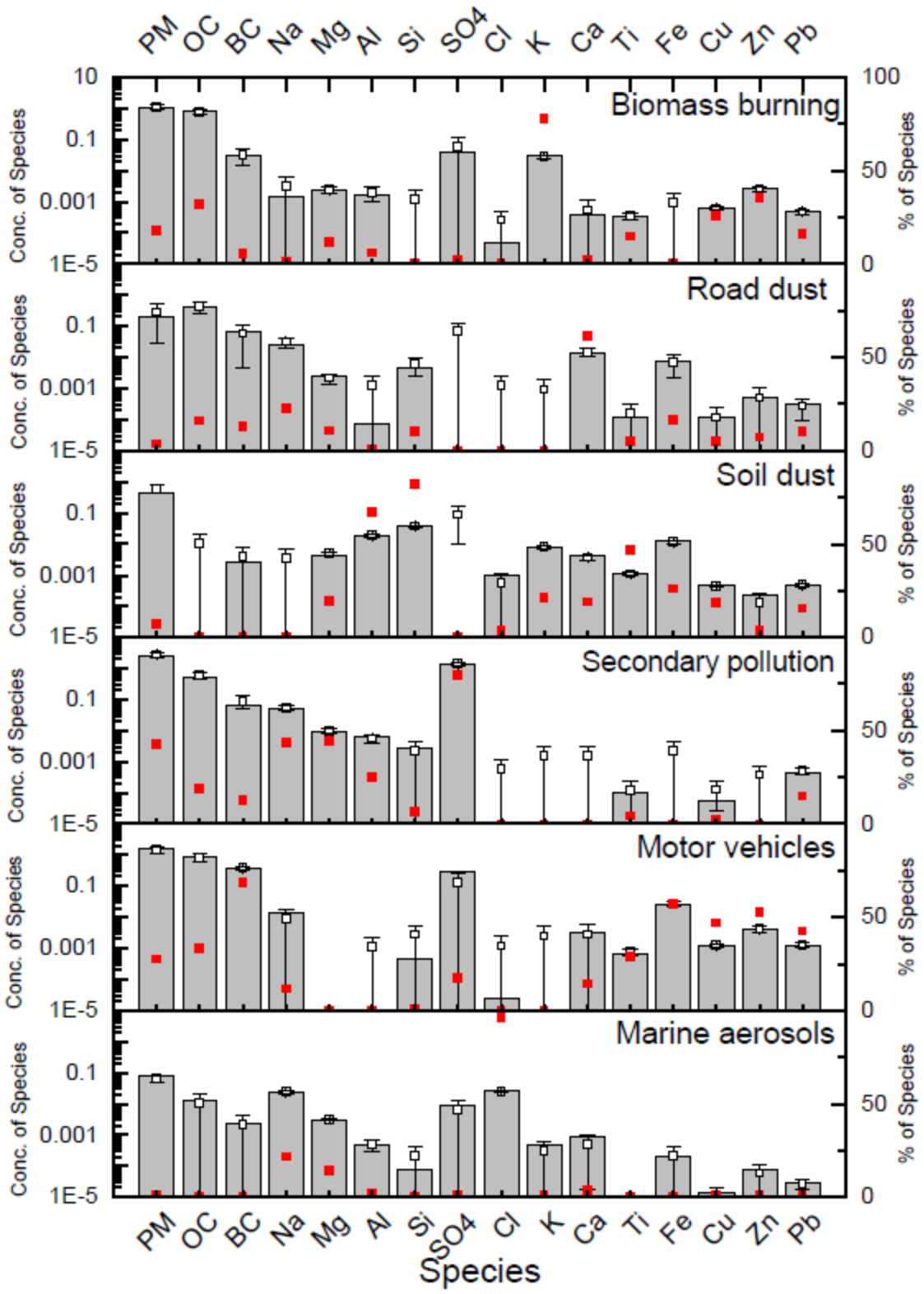

Figure 2.2. Source profiles contributing to outdoor $PM_{2.5}$. The grey filled-in rectangle represents the estimated average concentration of each species by source factor estimated using PMF. The small closed (red) square represents the percentage of a specific species ascribed to the specific source (factor) by

PMF. For example, using PMF, the mean concentration of indoor SO_4 ascribed to secondary pollution is 1.3 $\mu g/m^3$, which is 80% of the total sulfate concentration (See also Table 2.4). The small open square represents the estimated average concentration of each species by source factor estimated using DISP with the error bars indicating the confidence intervals.

Indoor-to-outdoor (I/O) ratios

Indoor-to-outdoor sulfur ratio is used widely to estimate the infiltration of $PM_{2.5}$ from outdoor to indoor assuming that there is non-significant indoor source of sulfur (Sarnat et al. 2002; Wallace and Williams 2005). For this study, four schools out of 32 schools had a sulfur ratio greater than 1. A ratio above 1 suggests that there might be indoor sulfur sources or possibly outdoor sulfur sources near the schools which do not impact the central site.

Figure 2.3 shows the I/O ratios for sulfur and $PM_{2.5}$ by season. The highest I/O sulfur ratios were observed during the spring (0.73). The highest I/O $PM_{2.5}$ ratios were observed during the spring (0.96). The opening of windows during the spring season might be the reason for the higher I/O ratios indicating a high infiltration. Higher I/O ratios were observed in schools in Agra, India (Habil et al. 2013): 1.0 for winter, 1.2 for summer, and 1.1 for the monsoon period. In Beijing, Hou et al. (2015) reported an I/O $PM_{2.5}$ ratio of 0.6 when the windows and doors were closed. This is comparable to those observed during the fall and winter seasons in Boston. In addition, the climatic conditions in Boston and Beijing are similar, where winter seasons are very cold. Furthermore, the I/O $PM_{2.5}$ was 0.9 in Hou et al. (2015) study when the windows and doors were open, which is similar to that we observed in Boston during the spring when windows and doors might more frequently open as compared to the other seasons.

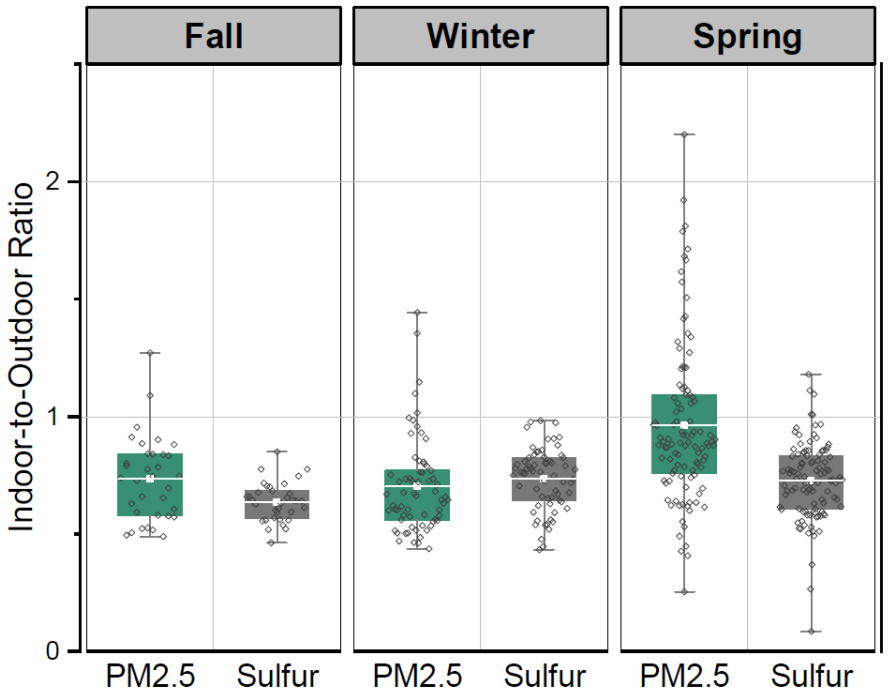

Figure 2.3. Box plots of indoor-to-outdoor $PM_{2.5}$ and sulfur ratios by season. Green boxes represent indoor-to-outdoor $PM_{2.5}$ ratios and gray boxes represent indoor-to-outdoor sulfur ratios. The box plot shows the mean, 25th and 75th percentiles, while the end of the whiskers indicates the minimum and maximum of the data.

Soil dust and Ca-rich particles sources can be categorized as indoor origins as described above, whereas motor vehicles, secondary pollution, biomass burning, and marine aerosols are categorized as outdoor origins. Seasonal relationships between source contributions and I/O sulfur ratios are shown in Figure 2.4. As it can be seen, the relationships of indoor origin differ from those observed for the outdoor ones. Indoor sources exhibit high variability between schools because the contributions are not related with infiltration (I/O sulfur ratio). Furthermore, motor vehicles and biomass burning sources have a different variation from the other two outdoor-origin sources (i.e., marine aerosols and secondary pollution) that have similar seasonal patterns, as also shown in Table 2.5 and Figure S2.C in Appendix S2.C. Higher contributions of motor vehicles in the winter and fall is more likely derived from local sources in the vicinity of the schools. Activities

in the school environment during winter such as using motor equipment/vehicles to clean the snow and spreading salt, and during the fall such as cleaning the leaves may have driven these differences to the other sources. In addition, higher contribution of biomass burning in winter is more likely associated with wood burning at the neighborhood of schools. Consequently, each source would have different distance scale: soil dust and Ca-rich particles are in the schools; motor vehicles and biomass burning are in the vicinity of schools; secondary and marine aerosols have a regional scale.

Table 2.5. Sources affecting $PM_{2.5}$ by season

Sources	Fall (*N* = 34)		Winter (*N* = 66)		Spring (*N* = 103)	
	$\mu g/m^3$	%*	$\mu g/m^3$	%	$\mu g/m^3$	%
Indoor						
Observed $PM_{2.5}$	4.38	-	4.94	-	5.73	-
Estimated $PM_{2.5}$	4.23	97	4.88	99	5.23	91
Motor vehicles	1.23	29	1.20	25	0.55	11
Secondary pollution	1.28	30	1.66	34	2.75	53
Marine aerosols	0.13	3	0.15	3	0.27	5
Soil dust	0.21	5	0.18	4	0.41	8
Ca-rich particles	0.56	13	0.51	10	0.71	14
Biomass burning	0.82	19	1.16	24	0.54	10
Outdoor						
Observed $PM_{2.5}$	6.17	-	7.50	-	5.97	-
Estimated $PM_{2.5}$	5.80	94	6.39	85	5.48	92
Motor vehicles	2.00	34	2.11	33	1.42	26
Secondary pollution	2.06	36	2.21	35	2.46	45
Marine aerosols	0.12	2	0.08	1	0.09	2
Soil dust	0.35	6	0.32	5	0.55	10
Road dust	0.16	3	0.14	2	0.24	4
Biomass burning	1.12	19	1.53	24	0.72	13

*Percentage of observed $PM_{2.5}$ mass

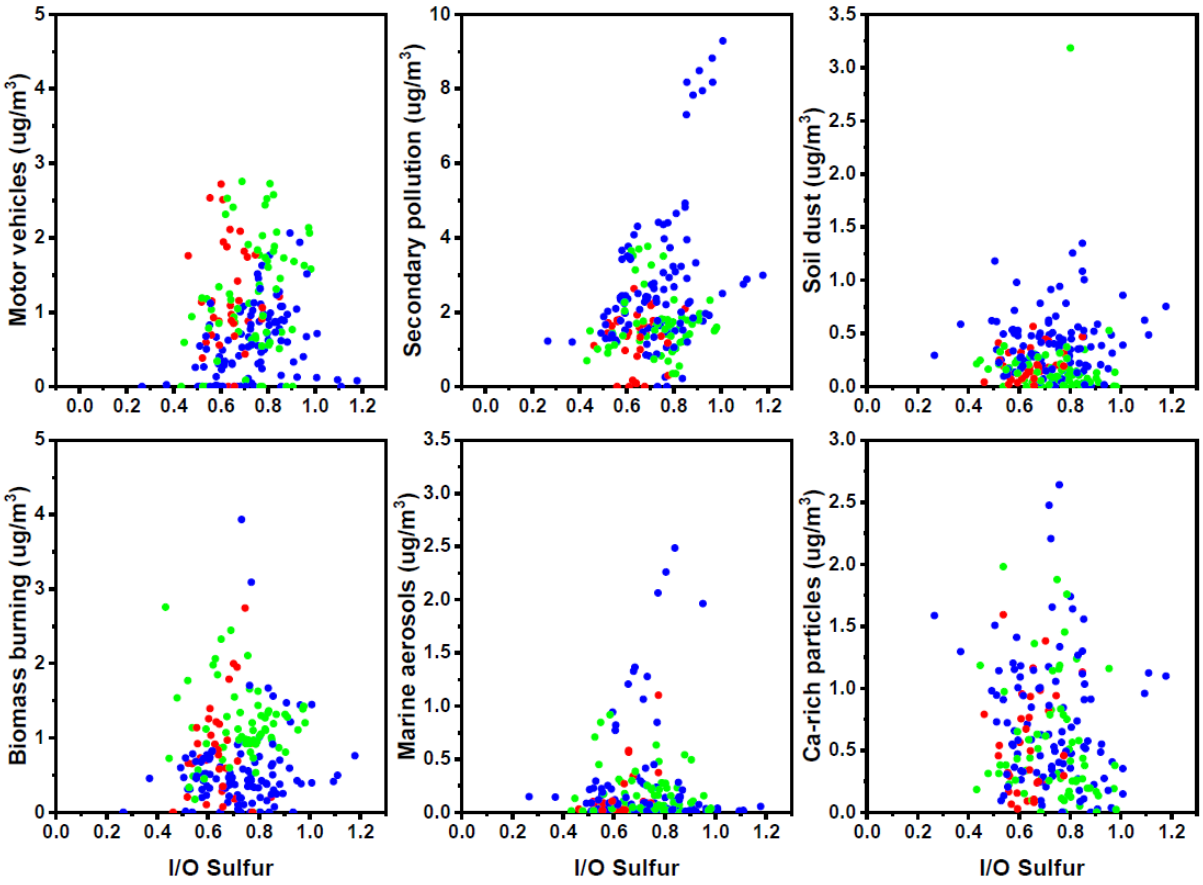

Figure 2.4. Relationship between classroom indoor $PM_{2.5}$ source contributions and I/O sulfur ratio by seasons: Blue dots (spring); Red dots (fall); and Green dots (winter).

2.4 Conclusions

Possible sources contributing to indoor $PM_{2.5}$ and seasonal variability were investigated at inner-city schools' environment. Indoor $PM_{2.5}$ concentrations were comparable to those observed outdoors, demonstrating penetration of outdoor pollution to indoors, but the concentrations were in the lower range compared to other school studies. Six source factors (Ca-rich particles, soil dust, motor vehicles, marine aerosols, secondary pollution and biomass burning) were identified to affect indoor $PM_{2.5}$ concentration. Ca-rich particles and soil factors were likely associated with child-related activities that transport materials from playgrounds. Concurrent outdoor $PM_{2.5}$ concentrations were estimated to be contributed by six source factors (motor vehicles, secondary

pollution, marine aerosols, soil dust, road dust, biomass burning), which were similar to the previous studies of the studied area. From the relationship between I/O sulfur ratio and each source contribution, seasonal patterns of motor vehicles and biomass burning were different with a lower in spring and a higher in winter from those of other outdoor source factors (i.e., secondary pollution and marine aerosols) with a higher in spring with a higher infiltration, which indicates difference in distance scales. Contrary to regional scale of secondary pollution and marine aerosols, motor vehicles and biomass burning have local scale in the vicinity of the schools.

Our study has some limitations. First, the use of a single outdoor measurement might somewhat misclassify the outdoor-to-indoor comparison of $PM_{2.5}$. However, we have previously demonstrated that ambient measurement at the ambient monitoring site reasonably represented the outdoor levels in the geographic area of this study. Also, the difference in distance to the ground between the schools and the 6-story building might affect the precision of the outdoor-to-indoor comparison. In addition, we did not measure the air exchange rate in each of the classrooms, and, given the variability of the classroom ventilation within as well as between schools, source apportionment by classroom or school ventilation type becomes difficult. Second, we did not measure secondary organic aerosols and nitrate indoor and outdoor, which might cause bias of our results. However, note that nitrate concentration is low in the northeast USA, with mean concentration of 0.79 $\mu g/m^3$ accounting for a 9% of the $PM_{2.5}$ mass during the study period.

Schools are complex microenvironments with different factors that can contribute to the transport, mixing and generation of indoor particles. Further investigations should incorporate indoor child behavior and activities to better assess exposure and the relevance of $PM_{2.5}$ in inner-city school indoor environments. If school-level exposures and their sources are assessed along

with their relative importance in influencing child health, then targeted interventions may improve children's health and well-being.

CHAPTER 3

Predictors of indoor radon levels

Predictors of indoor radon levels

Authors

Aleshka Carrion-Matta[a], Choong-Min Kang[a], Longxiang Li[a], Joel Schwartz[a], Philip Demokritou[a], Petros Koutrakis[a]

Affiliations

[a]Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA

Acknowledgements

We thank Spruce Environmental Technologies, Inc. for providing us with the residential radon measurement data. This publication was made possible by U.S. EPA grant **RD-835872**. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the U.S. EPA. Further, U.S. EPA does not endorse the purchase of any commercial products or services mentioned in the publication.

Abstract

Radon accumulation, mainly in low-ventilated underground floors and basements, expose its occupants to cause serious health diseases such as lung cancer, even at low levels. Prediction of radon exposure in enclosed spaces might be a big challenge due to several influential factors. Soil, geology, topography, atmospheric variables, radiation, urbanization, community economic well-being, and monthly and year variations are considered as influential factors. We determined which factors influence indoor radon levels. We analyzed 17,332 monthly indoor radon measurements of 802 zip codes from Massachusetts, Michigan, and Minnesota during 2005-2018 using a random forest model. Using Shapley Additive exPlanations (SHAP) values we investigated the contribution of the factors using variable importance and partial dependence plots. Factors that predict radon varied across the states, due possibly to a different soil composition in the states. Cross-validated R^2 between predictions and measurements was 0.68 (RMSE = 47.81 Bq/m^3) in Minnesota, 0.67, (RMSE = 52.61 Bq/m^3) in Michigan, and 0.41 (RMSE = 52.57 Bq/m^3) in Massachusetts. Our study result underlines the importance of soil structure. The differences across states also suggest that radon studies performing model development should also consider geographical variables, along with other influential factors. As indoor radon levels are multifactorial, an understanding of the factors that influence its emanation and build up indoors will help better assess spatial and temporal variations, which will further help better prevention and mitigation control strategies.

3.1 Introduction

In the U.S., approximately 21,800 lung cancer deaths per year are related to long-term exposure to radon (USEPA 2003). Furthermore, radon exposure might be a contributing factor to Non-Hodgkin's Lymphoma (Ha et al. 2017), multiple sclerosis (Bølviken et al. 2003), and skin (Vienneau et al. 2017), stomach, and brain cancer (López-Abente et al. 2018). Recent evidence from epidemiological studies has shown that exposure to radon at levels below the action levels can be associated to lung cancer (Vogiannis and Nikolopoulos 2015). The U.S. Environmental Protection Agency (EPA) estimates that approximately 7 million homes in the U.S. exceed the radon level of 148 Bq/m^3 (or 4 pCi/L), which is the action level for remediation (Gallaway et al. 2019; USEPA 2016). As people spend more of their time indoors, radon exposure is a problem in terms of radiological protection and health management.

Radon (^{222}Rn) is an inert and radioactive soil gas with decay products, which are significant contributors to the total radiation dose. Radon progeny can remain unattached (~2 nm diameter) or attach to ambient particles. Both fractions can pass the upper respiratory tract and can deposit onto the lungs where they irritate lung tissue leading to lung damage. Inhalation of radon progeny accounts for approximately 50% of the total natural radiation that humans are exposed (UNSCEAR 2000).

Radon, is generated within the Earth's crust and rises towards the surface and into the air by diffusion and convection (Dueñas et al. 1997). Indoor air radon levels are related to different factors such as: (1) atmospheric conditions (e.g., temperature, relative humidity, pressure, precipitation), (2) occupant behavior (e.g., opening/closing windows, (3) soil (e.g., porosity, permeability, temperature, moisture), (4) building characteristics (e.g., year structure built, ventilation), (5) geology (e.g., uranium content, rock type), and (6) topography (e.g., elevation)

(Vogiannis and Nikolopoulos 2015). Once the radon enters the indoor environment, it can interact with aerosol particles, build up and accumulate mostly in low-lying areas such as basements.

Predicting indoor radon is challenging due to the numerous factors that influence its levels. The use of machine learning methods represents an alternative to predict radon based on several parameters influencing its levels as it can analyze qualitative and quantitative predictors simultaneously (Khan et al. 2020). Even though machine learning methods have been used in indoor air quality, predominantly predicting particulate matter, carbon dioxide, and radon, to this date their use is scarce (Wei et al. 2019). As radon exposure is increasing over time in North America and has become a year-round problem (Stanley et al. 2019), it is important to know how exposure occurs.

In this context, the objective of this study is to examine the relationship between indoor radon levels and climate, soil, geology, topography, and radiation factors using a random forest model. Radon exposure is a multifactorial problem and the awareness related to its health effects remains low under general population. The results from this study may be used for exposure assessment and mitigation control strategies.

3.2 Materials and methods

Indoor radon

We used 422,329 indoor radon (Rn) measurements (Bq/m^3) conducted in Massachusetts (229,897), Michigan (86,748), and Minnesota (105,684) during the period 2005-2018. The data were obtained from the Spruce Environmental Technologies, Inc. State selection was based on data availability. Radon samples were collected in residential buildings under closed-house conditions at different house levels before mitigation, after mitigation, and during real estate property transactions. Sampling periods ranged from 2 to 98 hours. The data was collected using

Air Check and AccuStar test kits. The test kits use either activated charcoal (AC) or liquid scintillation (LS) methods. After collection, the samples were shipped to the Spruce Environmental Technologies laboratories for analysis.

Explanatory Variables

Different parameters were used to predict indoor radon levels and are shown in Table S3.1. We obtained ambient temperature (°K), pressure (Pa), precipitation (kg/m^2), evaporation (kg/m^2), dew point (°K), wind speed (m/s), wind direction (°), relative humidity (%), planetary boundary layer height (m), soil moisture content (kg/m^2), soil temperature (°K), upward short- and long-wave, downward short- and long-wave radiation flux (W/m^2), and snow cover from the North American Regional Reanalysis (NARR) provided by the NOAA/OAR/ESRL PSL, Boulder, Colorado, USA, (https://psl.noaa.gov/). Wind speed and direction were calculated using the u- and v-wind components. Soil temperature data included five depth distances: 0, 10, 40, 100 and 800 cm. The normalized difference vegetation index (NDVI), was obtained from the MODIS/Terra from NASA EOSDIS Land Processes DAAC (Didan 2015). Equivalent concentration of Uranium-238 (ppm) and Thorium-232 (ppm) were obtained from the U.S. Geological Survey (USGS) (Duval et al. 2005). Depth to bedrock (m), land-surface slope (%), organic matter (wt %) and permeability rate (in/hr) were obtained from the State Soil Geographic (STATSGO) database. Daily total sunspot number (sunspots) data was obtained from the Solar Influences Data Analysis Center (SIDC), WDC-SILSO, Royal Observatory of Belgium, Brussels. Sand, clay and silt soil fraction, and soil porosity were obtained from the Global Land Data Assimilation System (GLDAS) (Rodell et al. 2004). Soil porosity was obtained from three distances ranges (0-2 cm (top), 2-150 cm (middle), and 150-350 cm (bottom)). Elevation was obtained from the Shuttle Radar Topography Mission (SRTM) (SRTM 2013). Urbanization data were obtained from the

Rural-Urban Commuting Area (RUCA) codes. RUCA classifies zip codes in 10 different codes based on population density, urbanization, and daily commuting. The 10 code-classifications are metropolitan (1-3), micropolitan (4-6), small towns (7-9), and rural areas (10). For this study, we dichotomized the codes into three categories: metropolitan (1-3), micropolitan (4-6), and rural areas (7-10). Community economic well-being was obtained from the Distressed Communities Index (DCI) score. DCI score was developed by the Economic Innovation Group and ranks communities based on poverty rate, median income ratio, change in employment, change in establishments, housing vacancy rate, high school diploma and employment. Communities are then grouped by quintiles: prosperous (0 to 20), comfortable (20 to 40), mid-tier (40 to 60), at risk (60 to 80), and distressed (80 to 100). For this study, we assigned the DCI score from 2007-2011 to years 2005-2011 and from 2012-2016 to years 2012-2018. In addition, monthly, and yearly variations were considered for our analysis to control for seasonality and temporal trends.

Data Processing

We used radon levels at underground/basement levels before mitigation. Radon levels were aggregated by zip code and averaged by month. We used zip code as a proxy for house location due to inaccessibility. As shown in Table S3.2, we only analyzed the zip codes that had at least a month with over 5 measurements; in total we analyzed 802 zip codes (17,332 observations): 179 (9,817) from Massachusetts, 288 (2,600) from Michigan, and 335 (4,915) from Minnesota.

Monthly averages of numerical explanatory variables were calculated. Variables were geocoded based on zip code latitude and longitude obtained using ArcGIS 10.6 (ESRI, Redlands, CA, USA). Categorical variables were converted to dummy variables.

Random Forest: Model Selection and Evaluation

Random forest is an ensemble of decision trees that are grown using bootstrap samples. The model can capture complex and non-linear relationships between predictors and the outcome (Breiman 2001). The machine-learning algorithm was fitted in Python using the statistical package scikit-learn. We used nested cross-validation to ensure our model selection and hyper-parameter tuning was independent of the test data (Scheffer 1999; Varma and Simon 2006). In a nested cross-validation, an inner cross-validation is used for hyper-parameter tuning and model selection, and an outer cross-validation is used to evaluate the best model optimized by the inner cross-validation. In this study, we used an outer 10-fold cross-validation, where in each iteration a different 10% of the data was used as an independent test and 90% of the data was used as the training set. In the training set, an inner 5-fold cross-validation was performed to tune the hyper-parameters and select the best model to apply to the independent test. Hyper-parameter tuning was done using grid search and the hyper-parameters that we tuned were number of trees, maximum depth, maximum features, and maximum leaf nodes (Table S3.3).

The estimates of feature importance of the predictor variables were assessed after model training using Shapley Additive exPlanations (SHAP) (Lundberg et al. 2020). SHAP values are a game-theoretic approach based on Shapley values,

$$\varphi_i = \sum_{S \subseteq N \setminus i} \frac{|S|!\,(|N| - |S| - 1)!}{|N|!} (f(S \cup \{i\}) - f(S)) \qquad (1)$$

where ϕi is the contribution (Shapley value) of feature i, N is a set containing all the features, S is a subset of features not containing feature i, f(S) is the total contribution of features in S, and f(S ∪ {i}) is the total contribution of features in S including the i feature. The first part of the equation considers the ordering and the second part the contribution of features. Thus, Shapley values

measure model performance with and without the feature, considering all possible ordering. SHAP approximates Shapley values as a conditional expectation function,

$$f_x(S) = E[f(x)|x_S] \quad (2)$$

where E[f(x)] is the model's base value. The SHAP values tells us the magnitude and direction of the contribution to the prediction of the target variable. This method aids in model interpretation and explanation. Using this approach, we identified the six most important factors that predicted indoor radon levels in each state. In addition, we calculated the correlation (R^2) between the predicted and the observed indoor radon levels, and the standard deviation of the prediction residual (root mean squared error (RMSE)).

3.3 Results

Radon levels: descriptive analysis

Indoor radon levels measured in Massachusetts, Michigan and Minnesota had a geometric mean of 111.98, 114.54, and 148.23 Bq/m^3, and an arithmetic mean of 124.61, 133.21, and 161.67 Bq/m^3, respectively. Radon levels did not have a yearly variation as shown in Table S3.4. On the other hand, we found some monthly variations with low values in July and August as shown in Table S3.5.

Main influential factors

We used SHAP values to look at the variable importance per model and visualize the relationship between the covariates and the impact on the prediction. The mean absolute Shapley values (|SHAP values|) are shown in Table 3.1 and indicate average impact on model output magnitude. High mean absolute Shapley values have the most impact to the final model, and thus higher importance. The six most important factors found in Massachusetts were soil temperature

at 800 cm depth, elevation, upward longwave radiation flux, downward longwave radiation flux, dew point, and temperature. They changed the predicted absolute radon levels on average by 7.39, 7.37, 4.57, 1.87, 1.86, and 1.67 percentage points, respectively. The six most important factors for Michigan model were elevation, uranium concentration, soil temperature at 800 cm depth, sand fraction, clay fraction, and land-surface slope. They changed the predicted absolute indoor radon levels on average by 19.23, 10.77, 7.32, 7.07, 5.78, and 5.14 percentage points, respectively. The six most important factors for Minnesota model were the clay fraction, sand fraction, land-surface slope, soil porosity at 2-150 cm depth, organic matter, and soil moisture. They changed the predicted absolute indoor radon levels on average by 11.94, 11.01, 4.14, 3.71, 3.68 and 3.47 percentage points, respectively.

The positive or negative contributions of each factor to the predicted indoor radon levels are shown in Figures S3.1-S3.3. Factors with high SHAP values (above zero) have a positive impact on model prediction, whereas low SHAP values (below zero) have a negative impact on model prediction. In Massachusetts, low soil temperature at 800 cm depth, high elevation, low upward and downward longwave radiation flux, low dew point, and low ambient temperatures, increased the predicted radon levels, as shown in Figure S3.1. In Michigan, high elevation, low sand fraction, high uranium concentration, high clay fraction, high land-surface slope, and high soil temperature at 800 cm depth increased the predicted indoor radon levels as shown in Figure S3.2. In Minnesota, high clay fraction, low sand fraction, high land-surface slope, low soil porosity at 2-150 cm depth, low soil moisture and low organic matter, increased the predicted indoor radon levels as shown in Figure S3.3.

Dependence plots allow us to see the exact form of the relationship between the variable and the prediction. In Massachusetts, the predicted radon level was high in areas where the soil

temperature at 800 cm depth was below 284 K, elevation was above 100 m, upward longwave radiation flux was less than 400 W/m^2, dew point was below 288 K, downward longwave radiation flux was less than 325 W/m^2, and the ambient temperature was below 290 K, as shown in Figure 3.1. In Michigan, the predicted indoor radon level was high in areas where the sand fraction was below 0.4, elevation was above 250 m, the uranium concentration was higher than 1.2 ppm, clay fraction was above 0.2, the land-surface slope was more than 3.5%, and the soil temperature at 800 cm depth was higher than 281 K as shown in Figure 3.2. In Minnesota, the predicted indoor radon level was high in areas where the clay fraction was above 0.23, sand fraction was below 0.4, soil porosity at 2-150 cm depth was below 0.49, soil moisture was below 450 kg/m^2, and organic matter was below 15 wt % as shown in Figure 3.3.

Prediction Performance

Model performance varied across states. Massachusetts, Michigan and Minnesota had a cross-validated R^2 of 0.41, 0.67, and 0.68, respectively. The cross-validated RMSE for Massachusetts was 52.57 Bq/m^3, 52.61 Bq/m^3 for Michigan, and 47.81 Bq/m^3 for Minnesota.

Table 3.1. Variable importance based on SHAP values.

Covariates	**Mean (|SHAP value|), SD, %**		
	Massachusetts	**Michigan**	**Minnesota**
Temperature	1.67 (1.23) (5%)	0.10 (0.24) (0%)	0.27 (0.60) (0%)
Dew Point	1.86 (1.34) (5%)	0.12 (0.25) (0%)	0.28 (0.59) (0%)
Pressure	0.18 (0.21) (0%)	0.92 (0.94) (1%)	0.81 (1.03) (1%)
Precipitation	0.20 (0.23) (1%)	0.42 (0.59) (1%)	0.11 (0.16) (0%)
Evaporation	1.05 (0.67) (3%)	0.48 (0.64) (1%)	0.39 (0.44) (1%)
Wind Speed	0.21 (0.23) (1%)	0.25 (0.34) (0%)	0.15 (0.20) (0%)
Wind Direction	0.12 (0.19) (0%)	0.39 (0.50) (1%)	0.15 (0.21) (0%)
Relative Humidity	0.21 (0.30) (1%)	0.25 (0.42) (0%)	0.16 (0.24) (0%)
HPBL	0.25 (0.30) (1%)	0.45 (0.61) (1%)	0.17 (0.28) (0%)
NDVI	0.19 (0.21) (1%)	0.34 (0.72) (0%)	0.93 (2.65) (2%)
Snow Cover	0.03 (0.09) (0%)	0.22 (0.51) (0%)	0.16 (0.22) (0%)
Soil Temperature at 0 cm	0.44 (0.41) (1%)	0.11 (0.21) (0%)	0.39 (0.94) (1%)
Soil Temperature at 10 cm	0.10 (0.11) (0%)	0.20 (0.31) (0%)	0.66 (0.72) (1%)
Soil Temperature at 40 cm	0.07 (0.11) (0%)	0.45 (0.69) (1%)	0.22 (0.28) (0%)

Soil Temperature at 100 cm	0.14 (0.27) (0%)	0.64 (1.52) (1%)	0.33 (0.40) (1%)
Soil Temperature at 800 cm	7.39 (2.93) (20%)	7.32 (5.36) (9%)	1.84 (1.90) (3%)
Soil Moisture	0.52 (0.60) (1%)	1.41 (1.24) (2%)	3.47 (3.44) (6%)
Sand Fraction	0.05 (0.07) (0%)	7.07 (3.96) (9%)	11.01 (7.76) (18%)
Clay Fraction	0.03 (0.05) (0%)	5.78 (3.42) (7%)	11.94 (9.46) (20%)
Silt Fraction	0.04 (0.05) (0%)	0.70 (0.61) (1%)	0.29 (0.39) (0%)
Soil Porosity at 0 – 2 cm	0.07 (0.09) (0%)	0.69 (0.79) (1%)	0.60 (0.59) (1%)
Soil Porosity at 2 – 150 cm	0.08 (0.11) (0%)	1.62 (2.16) (2%)	3.71 (2.09) (6%)
Soil Porosity at 150 – 350 cm	0.10 (0.14) (0%)	0.58 (0.72) (1%)	0.43 (0.76) (1%)
Organic Matter	1.63 (1.96) (4%)	4.59 (3.05) (6%)	3.68 (6.19) (6%)
Permeability Rate	0.81 (1.15) (2%)	1.94 (1.65) (2%)	1.02 (1.29) (2%)
Uranium	1.28 (1.78) (3%)	10.77 (7.21) (14%)	0.89 (1.01) (1%)
Thorium	0.94 (1.11) (3%)	1.25 (1.04) (2%)	2.72 (2.92) (5%)
Depth to Bedrock	0.64 (0.61) (2%)	0.14 (0.25) (0%)	0.30 (0.55) (1%)
Number of Sunspots	0.20 (0.17) (1%)	0.10 (0.22) (0%)	0.20 (0.25) (0%)
Upward short-wave radiation flux	0.41 (0.29) (1%)	0.56 (0.54) (1%)	0.15 (0.18) (0%)
Downward short-wave radiation flux	0.24 (0.23) (1%)	0.32 (0.53) (0%)	0.28 (0.58) (0%)
Upward long-wave radiation flux	4.57 (2.64) (12%)	0.08 (0.12) (0%)	0.32 (0.61) (1%)
Downward long-wave radiation flux	1.87 (1.22) (5%)	0.11 (0.16) (0%)	0.35 (0.66) (1%)
Elevation	7.37 (4.89) (20%)	19.23 (9.88) (25%)	1.98 (1.55) (3%)
Land-Surface Slope	0.78 (0.99) (2%)	5.14 (4.64) (7%)	4.14 (3.16) (7%)
DCI score	0.50 (0.52) (1%)	0.61 (0.81) (1%)	3.21 (2.86) (5%)
Urbanization (Dummy Variables)	0.00 (0.00) (0%)	1.60 (2.77) (2%)	1.61 (1.98) (3%)
Month (Dummy Variables)	0.12 (0.33) (0%)	0.18 (0.75) (0%)	0.09 (0.23) (0%)
Year (Dummy Variables)	0.25 (0.62) (1%)	0.69 (1.56) (1%)	0.53 (1.17) (1%)
Test Type (Dummy Variables)	0.05 (0.09) (0%)	0.00 (0.00) (0%)	0.02 (0.08) (0%)

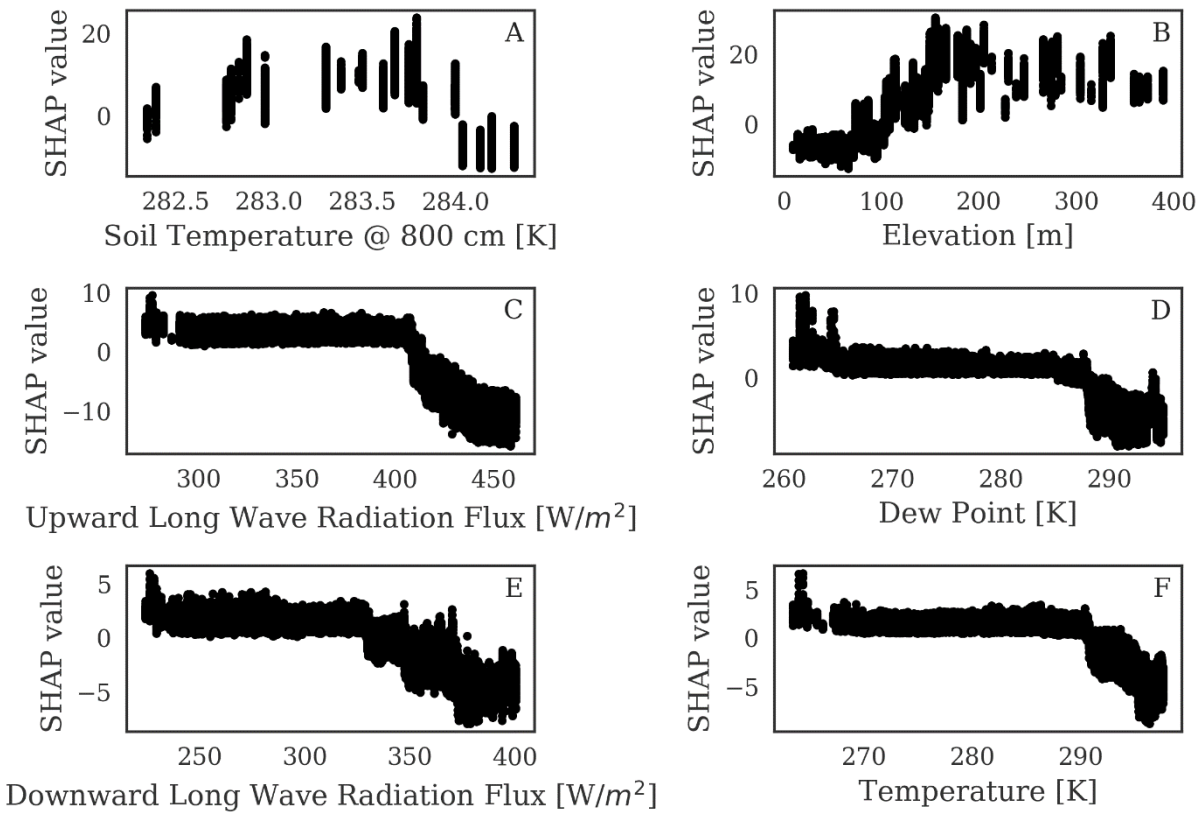

Figure 3.1 SHAP dependence plot for A) soil temperature at 800 cm depth, B) elevation, C) upward long wave radiation flux, D) dew point, E) downward long wave radiation flux, and F) temperature in Massachusetts. The plotted functions are interpreted as the effect of the feature's values on the model's prediction. For example, the predicted radon level was high in areas where the elevation was above 100 m.

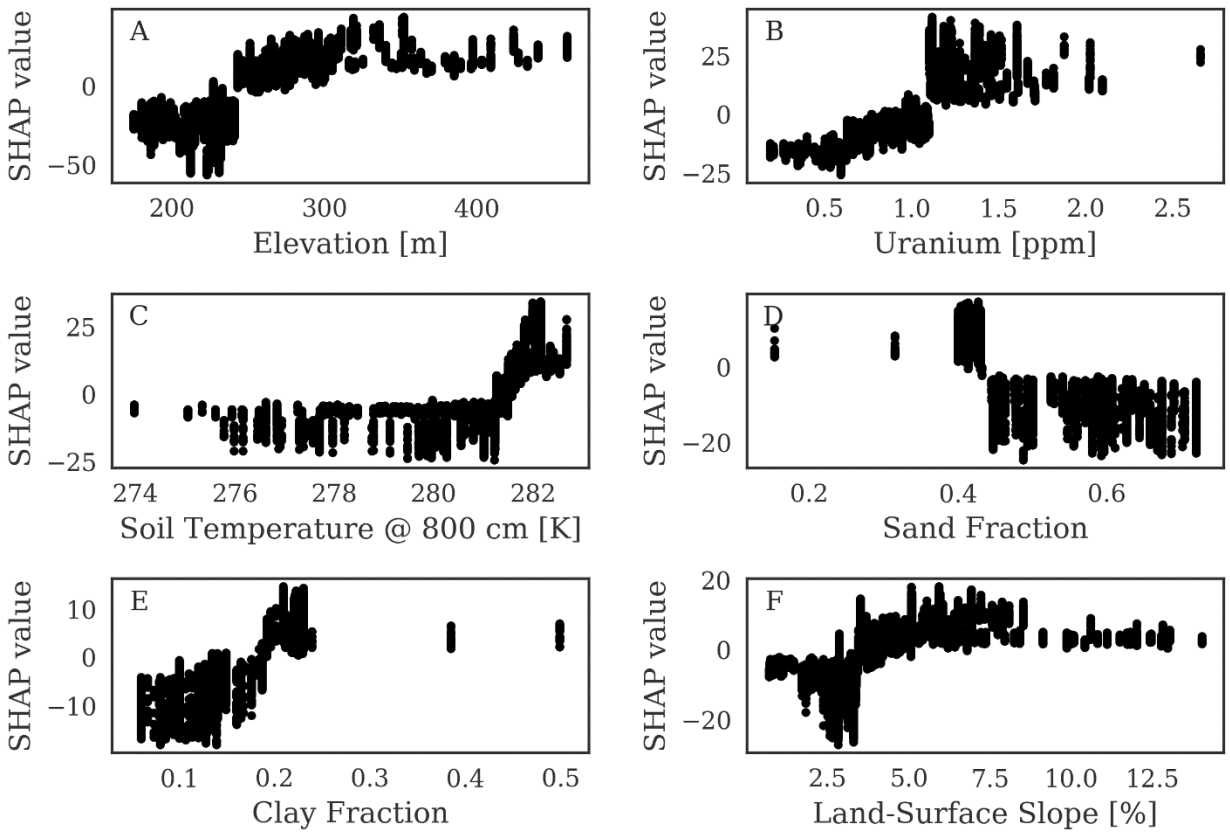

Figure 3.2 SHAP dependence plot for A) elevation, B) uranium, C) soil temperature at 800 cm depth, D) sand fraction, E) clay fraction, and F) land-surface slope in Michigan. The plotted functions are interpreted as the effect of the feature's values on the prediction. For example, in areas where the elevation was above 250 m, the predicted radon level was high.

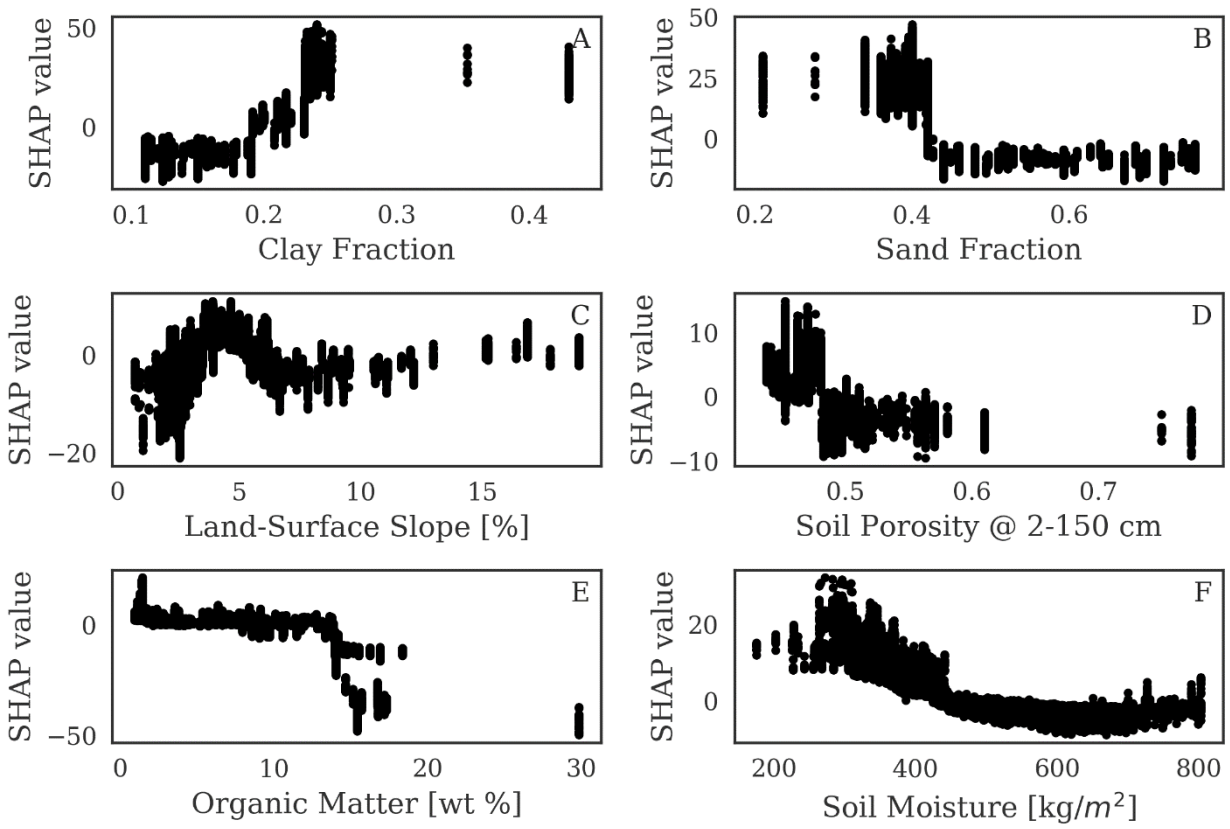

Figure 3.3 SHAP dependence plot for A) clay fraction, B) sand fraction, C) land-surface slope, D) soil porosity @ 2-150 cm depth, E) organic matter, and F) soil moisture in Minnesota. The plotted functions are interpreted as the effect of the feature's values on the prediction. For example, in areas where the sand fraction was above 0.4, the predicted indoor radon level was low.

3.4 Discussion

We determined important influential factors of radon levels in residential environments. We found that radon levels vary across states; in Minnesota, the mean levels (161.67 Bq/m^3) exceeded the remediation-action level of 148 Bq/m^3, and it was below the remediation-action level in Massachusetts (124.61 Bq/m^3) and Michigan (133.21 Bq/m^3). Also, we found that indoor radon levels were low in the summer months of July and August for the three states. Studies have mixed results about seasonality of indoor radon, some have shown that have seen higher values in winter

(Xie et al. 2015; Yousef and Zimami 2019) and recent studies have found high indoor radon levels in the summer months (Stanley et al. 2019).

Factors predicting indoor radon levels varied across states, where most soil factors explained the radon variability especially in Michigan and Minnesota. Whereas, in Massachusetts the atmospheric factors accounted the most for the variability of indoor radon levels.

Soil and geology factors. Soil and geology factors had a strong association with indoor radon levels in Michigan and Minnesota, probably due to the variability in glacial deposits in the area, which has an obscured association with indoor radon, or positive in some areas. These glacial deposits are poorly sorted and are derived from uranium-bearing granites and metamorphic rocks (Schumann et al. 1994). The type, composition, permeability, and thickness of these deposits is what controls the radon emanation (Gundersen 1993). Sand and clay fractions were important influential factors of indoor radon levels in Michigan and Minnesota. Even though it is known that radon moves rapidly in permeable soils such as sand and gravel, it is important to look at the underlying soil parent material. The northern Midwest of USA have silt and clay soils, which are derived from uranium-bearing shales and glacial lake deposits. These soils have high emanation radon coefficients due to high specific surface area and high permeability because are mixed with coarser sediments (Schumann et al. 1994). High uranium concentration was associated with elevated indoor radon levels in Michigan and Massachusetts. This agrees with other studies that have included uranium concentration as influential factor of indoor radon (Drolet et al. 2013; Kline et al. 1989). Among the three soil porosity depths, the one between 2-150 cm depth was the most important for the models. High organic matter was associated with high indoor radon levels in Massachusetts and Michigan, whereas it was the opposite in Minnesota. The factors that

contributed less than 5% to the models were silt fraction, depth to bedrock, permeability rate, and soil porosity at 0-2 and 150-350 cm depths.

Atmospheric factors. Studies (Aquilina 2019; Kitto 2005) have shown some association between atmospheric factors (pressure, wind speed, temperature, precipitation, and relative humidity) and radon levels. Atmospheric factors contributed more to indoor radon levels in Massachusetts. Among the five soil temperature depths, the soil temperature at 800 cm depth seemed to be the most important. High soil temperatures were associated with high indoor radon in Minnesota and Michigan, the opposite association was observed in Massachusetts. Studies have shown that when there is a temperature difference between air (low) and soil (high), it promotes convection of the soil gas (Klusman and Jaacks 1987), which probably was driving the emanation and transport of radon. Low ambient temperature was associated with elevated indoor radon levels in Massachusetts, which is typical of winter months. In addition, ambient temperature indirectly determines the resident behavior in terms of opening and closing windows, which will affect the entrainment of radon. Dew point had an inverse relationship with indoor radon in Massachusetts, this agrees with the previous study (Xie et al. 2015). Low soil moisture was associated with high indoor levels in all states. For example, in Minnesota a threshold value of approximately 400 kg/m^2 determined radon transport to indoors, where values above 400 kg/m^2 decreased radon emanation, causing low indoor radon levels. In this study, the factors that contributed less than 5% to the models were pressure, precipitation, wind speed, wind direction, relative humidity, HPBL, NDVI, snow cover, evaporation, ambient temperature, and soil temperature at 0, 10, 40, and 100 cm depth.

Topographic factors. High elevation and land-surface slope were associated with high indoor radon levels in all states. Some studies (Mose and Mushrush 1997) have shown an inverse relationship with indoor radon levels. Nevertheless, probably houses constructed in higher

elevation and land-surface slope might have had elevated indoor radon levels due to the permeability of the underlying soil, allowing higher emanation rates.

Urbanization and community economic well-being. Urbanization did not have an impact on indoor radon levels in Massachusetts. Whereas, in Michigan and Minnesota, a high radon level was observed in the micropolitan and non-metropolitan areas (i.e., micropolitan and rural areas). Distressed communities in Michigan and Minnesota showed a higher indoor level, implying the need to allocate mitigation strategies given that these communities have presented health disparities. In Massachusetts, it was the opposite, prosperous communities had high indoor radon levels.

Radiation factors. Longwave radiation contributed 12% (upward) and 5% (downward) to the random forest model in Massachusetts. To the best of our knowledge, longwave radiation has not been used to predict indoor radon levels. Outgoing (or upward) longwave radiation (OLR) has been used to study earthquakes anomalies, in which abnormal variations of radon and OLR occurred at the same time (Jing et al. 2013). It seems that there is some correlation between these variables that should be further explored in the context of indoor environments. Number of sunspots and shortwave radiation contributed less than 1% to radon prediction across all states.

Monthly and yearly variations. In our study, monthly and yearly variations did not have an impact on predict indoor radon levels (0-1% contributions, respectively).

Indoor radon studies have demonstrated many-sided conclusions (Borgoni et al. 2014; Dai et al. 2019; Hauri et al. 2013), due probably to the environmental complexity with various parameters. Our model included the majority of the variables used across studies and could predict reasonably well the indoor radon levels in Michigan and Minnesota. However, our model

performance for Massachusetts was moderate, representing that further research would be necessary for a better prediction. Overall, we concluded that predicting indoor radon levels by region and zip code level had a good performance.

Our results highlight the importance of geographical diversity in assessing indoor radon levels. An understanding of the factors that influence its emanation and build up indoors will help better assess spatial and temporal variations. Therefore, more studies combining different parameters to predict indoor radon levels are needed to better determine exposure for epidemiological studies. In addition, we draw attention to increase indoor radon testing to improve future analysis, especially in rural areas, where testing is usually low given that families underestimate the risk of radon exposure (Zahnd et al. 2018). In addition, our results can help to build the specific prevention and mitigation strategies in different regions. It would estimate the influential factors pertaining to its own area to raise effectiveness in decision-making process.

Our study has some limitations. First, the study falls short in having individual housing characteristics and occupant behavior, as these can explain some of the indoor radon variation (Borgoni et al. 2014; Dai et al. 2019; Hauri et al. 2013), but we tried to account for that using zip code-level values. Second, we do not have indoor pressures and temperature to see to what extent differences between indoor and outdoor was driving high indoor radon levels. Third, we do not have soil gas radon measurements, which has been shown to be correlated to indoor radon levels (Chen et al. 2009; Gundersen 1993). However, despite these limitations, our model could predict reasonably well indoor radon levels as well as obtain important influential factors.

3.5 Conclusions

Plausible influential factors of indoor radon levels were determined at residential environments in Massachusetts, Michigan, and Minnesota using a random forest model. Factors

of indoor radon varied across states, with soil and topography factors mainly accounting for these variations in Michigan (main factors: elevation, uranium concentration, sand fraction, soil temperature at 800 cm depth, clay fraction, and land-surface slope) and Minnesota (main factors: clay fraction, sand fraction, land-surface slope, soil porosity at 2-150 cm depth, organic matter, and soil moisture). On the other hand, in Massachusetts, topography and atmospheric factors accounted for these variations (main factors: soil temperature at 800 cm depth, elevation, longwave (upward and downward) radiation flux, dew point, and ambient temperature). Increase in indoor radon levels was associated with low soil temperature at 800 cm depth, high elevation, low upward and downward longwave radiation flux, low dew point, and low ambient temperatures in Massachusetts; in Michigan it was associated with high elevation, low sand fraction, high uranium concentration, high clay fraction, high land-surface slope, and high soil temperature at 800 cm depth; and in Minnesota it was associated with high clay fraction, low sand fraction, high land-surface slope, low soil porosity at 2-150 cm depth, low soil moisture and low organic matter. The factors used in this study explained well the variability of indoor radon levels, especially in Michigan and Minnesota, and provide useful information for epidemiological and exposure studies. However, as noted in other studies, indoor radon is a multifactorial problem and future studies should include soil gas radon measurements, resident behavior, and specific housing characteristics to account for the unexplained variability of indoor radon. In addition, the difference across states shows that studies about indoor radon should be targeted geographically, as different factors explain its variability per region. Mitigation and prevention control strategies should incorporate all the possible influential factors to better target interventions.

CHAPTER 4

Indoor radon levels near areas of shale gas development via hydraulic fracturing

Indoor radon levels near areas of shale gas development via hydraulic fracturing

Authors

Aleshka Carrion-Matta[a], Longxiang Li[a], Joel Schwartz[a], Philip, Demokritou[a], Petros Koutrakis[a]

Affiliations

[a]Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA

Acknowledgements

We thank Spruce Environmental Technologies, Inc. for providing us with the residential radon measurement data. This publication was made possible by U.S. EPA grant RD-835872. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the U.S. EPA. Further, U.S. EPA does not endorse the purchase of any commercial products or services mentioned in the publication.

Abstract

Hydraulic fracturing expanded U.S. natural gas industry in the mid 2000's by providing an alternative to energy use. However, its development and practices have brought environmental and public health concerns given the proximity of communities to shale gas development areas. Therefore, using Linear Mixed Effects (LME) models, we tried to assess whether proximity to shale gas development areas affect indoor radon levels. We analyzed 35,442 indoor radon monthly observations from Pennsylvania between 2001-2015 obtained from Spruce Environmental Technologies, Inc. We used the inverse distance weighted method incorporating the prevailing winds to determine the effect of the distance to unconventional wells on indoor radon levels. We looked at the upwind and downwind inverse distances from the wells to assess their effect. Secondarily, we investigated the contribution of 10, 50, and 100 wells at different distances. We found that shale gas development via hydraulic fracturing were associated with an increase in indoor radon levels in communities located downwind from the wells [slope = 1.93 (± 0.94); p-value 0.0393]. Our analysis revealed that the contribution of having 100 wells in 1 km will increase indoor radon by 185 Bq/m^3 (5 pCi/L). Our findings suggest that there is distance and direction effect of wells on indoor radon levels. In addition, the findings contribute to evidence on effects of exposure to radon on communities living near shale gas development areas.

4.1 Introduction

Hydraulic fracturing is a technique that involves the injection of water, a proppant (sand and/or silica), and chemicals into dense shale formations to extract oil and gas. This technology expanded U.S. natural gas industry in the mid 2000's by providing an alternative in energy use, which reduced the dependence of foreign oil and gas, and created job opportunities for people living near the sites (Finkel and Hays 2016). However, it has been in the middle of controversy due to environmental and public health concerns. Scientific evidence has shown that 84% of public health studies have found public health hazards, elevated risks, and adverse health outcomes; 69% of water quality studies have found water contamination; and 87% of air quality studies have found elevated atmospheric concentrations and high air pollutant emissions (Hays and Shonkoff 2016).

One of the major concerns among those that oppose to oil and gas production are the harmful chemicals that are brought up during extraction. In the extraction process, toxic flowback water, production brine and naturally occurring radioactive materials (such as Uranium (^{238}U), and decay products such as radium (^{226}Ra) and radon (^{222}Rn)) are brought up to the surface (Finkel and Hays 2016). Radon is an odorless and radioactive gas that is in the decay chain of Uranium. It is the second leading cause of lung cancer worldwide, after smoking, and contributes approximately to 21,000 lung cancer deaths annually in the U.S. (EPA 2003). Exposure to radon in the indoor environment can increase the chances of lung cancer from 3 to 14% (WHO 2009).

Radon's primary pathway to enter indoors is through cracks in basements and foundations. Shale gas development via hydraulic fracturing can promote new pathways for radon to both escape to the atmosphere and brought indoors as it can create routes for the gas to move to the horizontal pipes and consequently to the surface (Burkhart et al. 2013). Also, the gas can leak from the pipes and from the separation/storage tanks; as well as outgassing from the water brought to

the surface. Different studies have shown evidence following these processes as wells as an association between radon and shale gas operations. In Colorado, (Burkhart et al. 2013) found elevated ambient radon concentration during shale gas operations. Also, in Australia, radon concentration was found to be 20% higher in the center of the gas field relative to outside of it (Tait et al. 2013). In Pennsylvania (Casey et al. 2015) and Ohio (Xu et al. 2019), studies have shown an association between indoor radon levels and fracking. However, a recent re-analysis of Casey et al. 2015 study showed that the distance to the wells has zero effect on indoor radon levels (Black et al. 2019).

The scientific evidence finding an association between shale gas development via hydraulic fracturing areas and radon levels is still scarce. Moreover, studies relating radon to the whole range of processes associated with shale gas development via hydraulic fracturing are just emerging, with studies mainly performed at Pennsylvania, Colorado, and Ohio, where most of the production is taking place. Radon is subject to different influential factors in the indoor environment, that affects its accumulation. However, in areas where shale gas development has occurred, the potential of exposure to radon in indoor environments might increase and might vary based on the proximity to development areas as well as the route of exposure. In this context, our aim is to determine the association between indoor radon and shale gas development activities in Pennsylvania. We hypothesize that communities located downwind from the shale gas development areas will see an increase in indoor radon levels.

4.2 Methods and materials

Indoor radon

We used 61,333 indoor radon (Rn) measurements (Bq/m^3) conducted in Pennsylvania during the period 2001-2015. The data were obtained from the Spruce Environmental

Technologies, Inc. Radon samples were collected in residential buildings under closed-house conditions at basement levels before mitigation. The data was collected using Air Check and AccuStar test kits. The test kits use either activated charcoal (AC) or liquid scintillation (LS) methods. After collection, the samples were shipped to the Spruce Environmental Technologies laboratories for analysis. We aggregated samples by month, having a total of 35,442 monthly observations.

Exposure to wells

To determine the exposure of homes to unconventional wells, we used the inverse-distance-weighted (IDW) method incorporating prevailing winds. Details of the calculation are shown in Li et al. 2020. Briefly, a population weighted average of the inverse distance to unconventional wells were calculated for each zip code. Then, the prevailing winds were incorporated to determine the upwind and downwind inverse distance to zip code center from the wells. The exposure is calculated within a 15-km buffer, thus distances above 15-km are not included in this study. In addition, exposure is averaged by year.

Statistical Analysis

We used linear mixed effects (LME) models to characterize the association between indoor radon and shale gas development activities, with a random intercept for zip code. To determine the effect of the wells on nearby communities we looked at the exposure distribution. Three types of analyses were performed for unconventional wells using SAS (version 9.4). First (model 1), we analyzed the effect of the downwind inverse distance from the wells. Second (model 2), we analyzed the effect of the upwind inverse distance from the wells. Third (model 3), we analyzed the effect of both the upwind and downwind inverse distance from the wells. We considered results significant when $p < 0.05$.

To determine the contribution of wells to indoor radon levels we used the following equation,

$$Contribution = \frac{\beta_{invdistance} \times (\#\ wells)}{distance}$$

where we tried 10, 50, and 100 wells over distance.

4.3 Results

Summary statistics

In this study, we analyzed 35,442 monthly indoor radon measurements. Indoor radon had a mean value of 296.42 ± 465.77 Bq/m^3 (mean ± SD). These values are higher than the remediation level established by the U.S. EPA of 148 Bq/m^3 (4 pCi/L). As summarized in Table S4.1, indoor radon has been slightly decreasing over the years, probably because of mitigation strategies. However, these values are still above the remediation level, which is a cause of concern. The monthly variation of indoor radon levels is shown in Table S4.2. We found that indoor radon levels were lower in the summer months, with levels in July been the lowest 214.46 ± 377.44 Bq/m^3 (mean ± SD).

Exposure distribution

To determine the zip codes that were affected by the wells we ran a univariate model to see their distribution. As shown in Table S4.3, around 10% of the zip code observations were within a 15-km distance and affected by both upwind and downwind wells. We used these observations to run our models.

Association between indoor radon levels and proximity to the wells

To study the association between indoor radon and proximity to oil and gas practices, we ran linear mixed effect models using the downwind (model 1), upwind (model 2), and upwind +

downwind (model 3) inverse distances, as shown in Table 4.1. For model 1, we found an effect of the downwind inverse distance from the wells on indoor radon levels for unconventional oil and gas practices (slope = 1.93 , SE = 0.94, p-value 0.0393). For model 2, we did not find that upwind indoor radon levels were statistically significant associated with unconventional gas practices. For model 3, we did not find an effect of the upwind and downwind inverse distance from the wells on indoor radon levels for the unconventional oil and gas practices.

Table 4.1. Linear mixed model for unconventional wells.

Fixed Effect	**Model 1**		**Model 2**		**Model 3**	
	Coefficient (SE)	p-value	Coefficient (SE)	p-value	Coefficient (SE)	p-value
Intercept	244.62 (48.16)	<0.0001	218.52 (46.65)	< 0.0001	213.61 (55.30)	0.0001
Downwind inverse distance	1.93 (0.94)	0.0393	-	-	2.52 (1.36)	0.0652
Upwind inverse distance	-	-	1.60 (0.86)	0.0618	-0.19 (1.33)	0.8857
Year: 2001	86.30 (54.23)	0.1117	50.47 (55.57)	0.3638	85.18 (65.41)	0.1930
Year: 2002	40.29 (54.76)	0.4619	38.13 (53.97)	0.4799	77.57 (67.32)	0.2493
Year: 2003	-73.54 (63.59)	0.2476	-58.75 (61.85)	0.3423	-79.21 (74.51)	0.2878
Year: 2004	53.45 (54.39)	0.3258	94.07 (54.94)	0.0869	111.74 (66.74)	0.0942
Year: 2005	-27.56 (48.12)	0.5669	17.12 (49.75)	0.7308	-16.65 (60.99)	0.7848
Year: 2006	99.43 (47.16)	0.0351	87.97 (45.34)	0.0524	113.12 (54.46)	0.0379
Year: 2007	84.57 (47.76)	0.0767	61.52 (45.42)	0.1757	102.28 (54.00)	0.0583
Year: 2008	120.34 (40.94)	0.0033	120.25 (39.57)	0.0024	140.71 (44.57)	0.0016
Year: 2009	45.52 (43.19)	0.2920	23.08 (41.92)	0.5820	10.66 (47.23)	0.8214
Year: 2010	35.46 (43.03)	0.4099	15.60 (40.80)	0.7023	28.14 (46.38)	0.5441
Year: 2011	57.17 (40.39)	0.1570	41.89 (39.13)	0.2845	68.74 (44.87)	0.1256
Year: 2012	-31.55 (40.02)	0.4305	-29.78 (38.60)	0.4404	-28.72 (44.14)	0.5153
Year: 2013	-49.65 (40.30)	0.2180	-16.49 (38.49)	0.6683	-36.09 (44.21)	0.4145
Year: 2014	40.01 (38.97)	0.3046	38.31 (37.95)	0.3128	42.21 (42.70)	0.3230
Month: 1	31.87 (43.84)	0.4673	57.84 (43.22)	0.1809	58.64 (50.43)	0.2450
Month: 2	1.97 (43.32)	0.9638	8.75 (42.18)	0.8357	-1.08 (49.40)	0.9825
Month: 3	-25.99 (43.75)	0.5525	-18.69 (42.73)	0.6618	-23.91 (49.80)	0.6311
Month: 4	-58.68 (43.63)	0.1788	-8.28 (42.43)	0.8453	-32.71 (49.43)	0.5083
Month: 5	-88.07 (44.28)	0.0468	-22.65 (42.81)	0.5968	-63.48 (50.24)	0.2065
Month: 6	-99.74 (46.31)	0.0313	-65.60 (45.55)	0.1499	-76.72 (52.71)	0.1456
Month: 7	-123.27 (46.81)	0.0085	-106.23 (44.89)	0.0180	-119.06 (51.85)	0.0217
Month: 8	-79.52 (47.35)	0.0931	-73.44 (45.99)	0.1104	-79.51 (53.01)	0.1338
Month: 9	0.39 (46.96)	0.9935	17.41 (45.32)	0.7009	24.45 (53.10)	0.6453
Month: 10	47.23 (42.36)	0.2650	51.12 (41.34)	0.2163	54.33 (48.31)	0.2609
Month: 11	18.94 (42.54)	0.6562	55.01 (41.76)	0.1879	42.16 (48.29)	0.3827

Effect of seasonal and year variations

In our three models, we adjusted for seasonal and year variations. We found seasonal effects, where the summer months of June [slope = -99.74, SE = 46.31, p-value 0.0313] and July caused a decrease in indoor radon levels [slope = -123.27, SE = 46.81, p-value 0.0085] in Model 1. In 2008, we found that the indoor radon levels increased [slope = 120.34, SE = 40.94, p-value 0.0033]. The sudden increase in drilling of unconventional wells around 2008 could have been the cause of these findings.

Contribution of wells to indoor radon levels

The contribution of wells to indoor radon is shown in Figure 4.1. We found that having 100 wells located at 1 km from the communities will increase downwind indoor radon levels by 5 pCi/L. We found that the contribution of 10 wells to downwind indoor radon levels was not strong compared to the contribution of 50 and 100 wells. However, if there are more than 10 wells, we see contributions to indoor radon levels at short distances. Also, at larger distances we do not see an effect (less than 0.5 pCi/L).

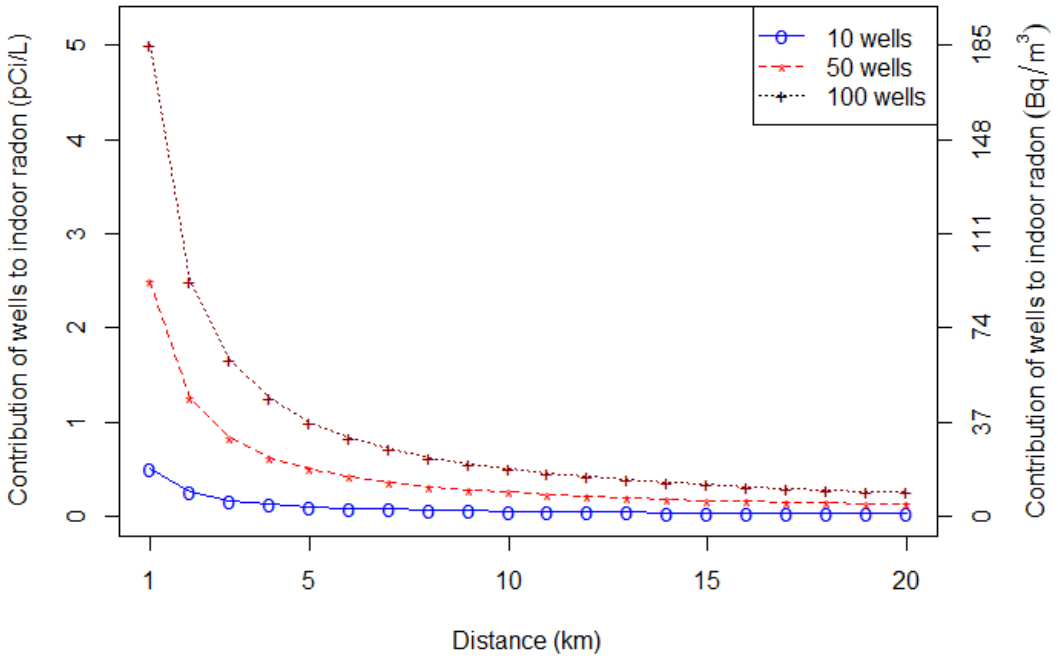

Figure 4.4 Contribution of wells to indoor radon levels.

4.4 Discussion

Studies associating indoor radon with shale gas development practices via hydraulic fracturing are scarce, to date there are only studies performed in Pennsylvania and Ohio. In Pennsylvania, the study by Casey et al. 2015, found that there was an increase in indoor radon levels with proximity to wells. However, in a re-analysis of the former study by Black et al. (2019), they did not find an effect of distance to the wells on indoor radon levels. Whereas, in Ohio, the study by Xu et al. 2019, found that indoor radon levels were higher near unconventional wells. These studies show conflicting findings about the association of shale gas development via hydraulic fracturing with indoor radon levels, which demonstrate the need for more studies to obtain a comprehensive understanding of the processes and transport of radon to indoor

environments. In a recent study by Li et al. (2020), the authors found higher levels of particle radioactivity downwind from the wells and provided insights of possible transport of radon to homes nearby unconventional wells when they incorporated prevailing winds fields to their exposure assessment. In our study, using the prevailing wind fields in our exposure assessment, we found that there is an effect of the proximity to wells on indoor radon levels, specifically an effect of downwind distance from the wells. Thus, these studies provide evidence of potential transport of radon from unconventional wells and activities to residences nearby the sites. The fact that our observations were taken at the basement level and we could see an effect downwind from the wells, tells us that probably the effect is even larger.

There is evidence of air pollutant transport to communities close to the wells (Adgate et al. 2014; Litovitz et al. 2013; Moore et al. 2014; Srebotnjak and Rotkin-Ellman 2014). Exposure to diesel exhaust emissions, road dust, and nitrogen oxides (NOx) occur from traffic related to industrial operations. Benzene, alkanes, VOCs, silica, hydrogen sulfide, and aromatic hydrocarbons have been measured during well development and completion. An increase in ozone levels have been observed in Texas, Oklahoma, and Colorado, with levels above the U.S. EPA 8 h National Ambient Air Quality Standard (NAAQS). However, the transport of radioactive materials such as radon to the surroundings is not fully developed in the literature and more work needs to be done to assess exposures.

It is worth discussing that some shale gas development practices could promote the release of radioactive material and thus its transport to nearby communities. For example, in the flowback, wastewater is often put into frack ponds, where this wastewater could contain radioactive materials that could outgas and release radium and its decay products, which will be transported to the surroundings. In a study by the New York's Department of Environmental Conservation, they

found that wastewater from the Marcellus Shale had radium levels as high as 267 times the limit for discharge (Kargbo et al. 2010; Lustgarten 2009). Usually, the assessment of radioactive materials in frack ponds has focused on radium, which underestimate the total concentration of radioactive material as wells as the increase in levels over time (Nelson Andrew et al. 2015). As mentioned in Nelson Andrew et al. 2015, radon, due to their long half-life of 3.82 days, promotes the migration and separation of radium decay products in frack ponds, which shows the need to include radon in future assessments. Other shale gas development practices and events such as leakage from storage tanks (Burkhart et al. 2013), from abandoned wells, from cement and gas well casing failures, and solid waste (Swiedler et al. 2019), expose communities and the workers to high concentrations of radon, and should be further studied.

The potential impacts to human health due to the proximity of communities to the wells had cause great concerns given that studies have found an association between fracking and infant mortality (Apergis et al. 2019; Busby and Mangano 2017), asthma exacerbations (Rasmussen et al. 2016), and gonorrhea and chlamydia (Beleche and Cintina 2018). Moreover, health impact assessments including distance to the wells have been conducted to provide an improved and broad understanding of the effects of hydraulic fracturing on public health (Meng and Ashby 2014). A health impact assessment study by McKenzie et al. 2012, reported that residents living nearby unconventional natural gas wells in Colorado were at a greater risk of health impacts from chronic and subchronic non-cancer. Moreover, a distance-based risk assessment by Meng 2015 showed that populations living up to 3 km from the wells in Pennsylvania were at risk of natural gas fracking. In addition, a study by Coons and Walker (2008), found benzene emissions near fracking wells in Colorado. However, epidemiological studies are still scarce, in part mostly to the difficulty in obtaining information about the chemicals used in the development activities, which the drilling

companies are not mandated to disclose chemicals and quantities (Finkel and Hays 2016). This makes it difficult to have a complete understanding of the impacts of shale gas development via hydraulic fracturing.

Certainly, Marcellus shale region has high radioactive signatures which makes homes in Pennsylvania prone to higher radon levels than other areas. But an increase in drilling will increase the risk of exposure to radon as it creates new pathways for radon to migrate to the surface and thus transported to residences. Furthermore, shale gas development practices will contribute to the emanation of radon, which is often overlooked or not fully discussed in the literature. In the early days of oil and gas development, drilling happened in remote areas, far from population centers. However, this have changed in some states due to regulations that have lessen restrictions, which pose a threat to the public. It is important to mention that not only communities are exposed, but also workers in these sites are at major risks of exposure to radon. Occupational exposures include radon and dust inhalation, and skin exposure. The handling of materials such as brine, flowback water with high levels of radioactive material, put workers at great risk. Even though workers can use dosimeters, they do not track what each person have inhaled and sometimes these dosimeters do not track alpha particles, which are the ones emitted by naturally occurring radioactive materials. A better understanding of their exposures could protect them.

Our study has some limitations. First, we do not have specific address for each observation, but we tried to account for that using zip code level measurements, which is far more precise than using state or county levels. Second, the data might have gaps in years and missing information from zip codes, as the main purpose for the collected samples was not for research purposes. Third, we do not know the specific stages in the shale gas development that are causing the observed indoor radon concentrations. Fourth, we do not account for other potential sources of radon that

have a downwind effect. However, despite these limitations we found statistically significant effects of proximity of wells.

4.5 Conclusions

The effect of proximity to wells on indoor radon levels were investigated in Pennsylvania from 2001-2015, where we found that proximity to unconventional wells were associated with an increase in indoor radon levels. Moreover, the increase in indoor radon occurred in residences located downwind from the wells. In addition, we found that season and yearly variations influenced indoor radon levels. The contribution of wells varied with distance, where we found that 100 wells will increase radon levels at residences within 1 km. Our findings suggest that there is a proximity and direction effect of unconventional wells on indoor radon levels.

Shale gas development via hydraulic fracturing brought technological advances and economic growth to the Marcellus Shale region. Despite these advances, there are critical drawbacks regarding environmental and public health concerns. The proximity of communities to these practices expose residents to radon as wells as the workers in the sites. These practices have increased over the years and are expected to continue increasing and occurring closer to population centers, where residents and workers are at higher risk of exposure to radon and thus developing diseases. A better understanding of the transport and fate of radon will help target better mitigation strategies.

CHAPTER 5

Conclusion Remarks

Conclusion Remarks

As people spend most of their time indoors, they are exposed to different air pollutants that could affect their health and well-being. Indoor environments are highly complex, and different factors and sources contribute to the transport, mixing and generation of ambient particles and radon. In our studies, we assessed exposures to $PM_{2.5}$ and radon at the residential and school levels and showed the different factors that contribute to the exposure to these pollutants.

Indoor $PM_{2.5}$ concentrations in schools were comparable to outdoor $PM_{2.5}$, demonstrating the penetration of outdoor pollution to indoors. Sources that affected indoor $PM_{2.5}$ were Ca-rich particles, soil dust, motor vehicles, marine aerosols, secondary pollution, and biomass burning. These sources had different distance scales: regional (secondary and marine aerosols), in the vicinity (motor vehicles and biomass burning), and inside the schools (soil dust and Ca-rich particles). Also, these sources varied by season. Overall, the observed differences to indoor $PM_{2.5}$ were related to seasonality, and the distinct characteristics and behavior of each classroom/school.

Meanwhile, indoor radon concentrations in residential buildings varied across states, where soil and topography factors mainly accounted for these variations in the Midwestern states of Michigan (main factors: elevation, uranium concentration, sand fraction, soil temperature at 800 cm depth, clay fraction, and land-surface slope) and Minnesota (main factors: clay fraction, sand fraction, land-surface slope, soil porosity at 2-150 cm depth, organic matter, and soil moisture). Whereas topography and atmospheric factors accounted for these variations in Massachusetts (main factors: soil temperature at 800 cm depth, elevation, longwave (upward and downward). The difference across states shows that studies about indoor radon should consider geographical differences, as different factors explain its variability per region. Also, we found that indoor radon concentrations varied near shale gas development via hydraulic fracturing areas, where we showed

that communities living downwind and close to the wells had an increase in indoor radon levels. Shale gas development practices have increased over the years and are closer to population centers, where residents and workers are at higher risk of exposure to radon and thus developing diseases.

Our findings provide evidence for future exposure and epidemiological studies. In modern environments, the pollutants and their sources must be well-understood to better target appropriate mitigation and prevention control strategies. As we strive for healthier indoor environments, developments of comprehensive assessments are needed.

SUPPLEMENTAL INFORMATION

Supplemental Information for Chapter 2

Appendix S2.A. PMF input and diagnostics summary

A.1. Indoor $PM_{2.5}$

**** Input Data Statistics ****

Species	Category	S/N
Mass	Weak	9
BC	Strong	5.6
Na	Weak	1.3
Mg	Bad	0
Al	Weak	0.1
Si	Strong	4.8
Sulfate	Strong	10
Cl	Strong	7.4
K	Strong	9.9
Ca	Strong	9.7
Ti	Weak	0.9
V	Weak	0.6
Cr	Bad	0
Mn	Bad	0
Fe	Weak	6.5
Ni	Weak	0.1
Cu	Weak	0.1
Zn	Weak	2.5
Br	Bad	0.1
Pb	Weak	0.3

**** Constrained Run Summary ****

Base model run number:	5
Base user-selected seed:	67
Number of factors:	6
Extra modeling uncertainty (%):	0

**** Constraints ****

Factor	Element	Type	Value	dQ	% dQ
Factor 2	Sulfate	Pull Down Maximally	NA	3.48	1
Factor 5	Sulfate	Pull Down Maximally	NA	3.48	1

**** Constrained Run Summary Table ****

Constrained #	dQ(Robust)	Q(Robust)	Q(Aux)	Q(True)	Converged	# Steps
1	1.6	349.7	1.7	348	Yes	1,079

**** Constrained Run Statistics ****

Species	Intercept	Slope	SE	r^2	KS test stat	KS test p value
Mass	1.17	0.71	0.92	0.76	0.06	0.42
BC	0	0.99	0.01	1	0.19	0
Na	0.01	0.88	0.03	0.84	0.04	0.95
Al	0	0.56	0	0.67	0.15	0
Si	0	1	0	1	0.05	0.6
Sulfate	0	1	0	1	0.07	0.3
Cl	0	1	0	1	0.12	0.01
K	0	1	0	1	0.05	0.7
Ca	0	1	0	1	0.03	0.97
Ti	0	0.41	0	0.4	0.13	0
V	0	0.15	0	0.15	0.18	0
Fe	0.01	0.61	0.01	0.51	0.09	0.08
Ni	0	0.16	0	0.17	0.10	0.02
Cu	0	0.15	0	0.12	0.12	0
Zn	0	0.55	0	0.50	0.08	0.13
Pb	0	0.1	0	0.04	0.06	0.47

A.2. Outdoor $PM_{2.5}$

**** Input Data Statistics ****

Species	Category	S/N
Mass	Weak	9
OC	Strong	4.5
EC	Bad	0.9
BC	Strong	7.2
UVC	Bad	7.2
DC	Bad	3.5
Na	Strong	3.1
Mg	Weak	0.5
Al	Strong	2.5
Si	Strong	7.1
Sulfate	Strong	10
Cl	Strong	5.1
K	Strong	10
Ca	Strong	9.7
Ti	Strong	1.9
V	Bad	1.9

Cr	Bad	0.1
Mn	Bad	0.1
Fe	Strong	9.2
Ni	Bad	0.4
Cu	Weak	0.1
Zn	Strong	4.2
Br	Bad	1
Ba	Bad	0.3
Pb	Weak	0.9

**** Constrained Run Summary ****

Base model run number:	15
Base user-selected seed:	86
Number of factors:	6
Extra modeling uncertainty (%):	0

**** Constraints ****

Factor	Element	Type	Value	dQ	% dQ
Factor 1	Sulfate	Pull Down Maximally	NA	139.87	0.5
Factor 6	Sulfate	Pull Down Maximally	NA	139.87	0.5

**** Constrained Run Summary Table ****

Constrained #	dQ(Robust)	Q(Robust)	Q(Aux)	Q(True)	Converged	# Steps
1	71.1	28044.2	70	63321.1	Yes	2,028

**** Constrained Run Statistics ****

Species	Intercept	Slope	SE	r^2	KS test stat	KS test p value
Mass	1.15	0.74	1.6	0.81	0.05	0.01
OC	0.84	0.58	1.22	0.33	0.06	0
BC	0.15	0.65	0.17	0.64	0.05	0.01
Na	0.02	0.78	0.07	0.65	0.06	0
Mg	0.01	0.49	0.01	0.44	0.06	0
Al	0	0.88	0.01	0.88	0.03	0.26
Si	0	1	0	0.99	0.06	0
Sulfate	0	1	0	1	0.08	0
Cl	0	1	0	1	0.13	0
K	0	1	0	1	0.07	0
Ca	0	0.99	0	1	0.02	0.52

Ti	0	0.64	0	0.69	0.1	0
Fe	0	0.96	0	0.98	0.04	0.02
Cu	0	0.47	0	0.43	0.05	0
Zn	0.01	0.09	0	0.08	0.23	0
Pb	0	0.27	0	0.18	0.03	0.14

Appendix S2.B. PMF base error estimation summary

B.1 Indoor $PM_{2.5}$

**** Constrained BS-DISP Diagnostics ****						
# of Cases Accepted:	101					
%dQ:	-2.63126					
Swaps by Factor:	0	0	0	0	0	0

**** Constrained DISP Diagnostics ****						
Error Code:	0					
%dQ:	-0.39317					
Swaps by Factor:	0	0	0	0	0	0

**** Constrained BS Mapping ****

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Unmapped
Boot Factor 1	100	0	0	0	0	0	0
Boot Factor 2	0	100	0	0	0	0	0
Boot Factor 3	0	0	100	0	0	0	0
Boot Factor 4	0	0	0	100	0	0	0
Boot Factor 5	0	0	0	0	100	0	0
Boot Factor 6	0	0	0	0	0	100	0

B.2. Outdoor $PM_{2.5}$

**** Constrained BS-DISP Diagnostics ****						
# of Cases Accepted:	90					
%dQ:	-0.07431					
Swaps by Factor:	0	0	0	0	0	0

**** Constrained DISP Diagnostics ****						
Error Code:	0					
%dQ:	-0.0398					
Swaps by Factor:	0	0	0	0	0	0

**** Constrained BS Mapping ****

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Unmapped
Boot Factor 1	100	0	0	0	0	0	0
Boot Factor 2	0	100	0	0	0	0	0
Boot Factor 3	0	0	100	0	0	0	0
Boot Factor 4	0	0	0	100	0	0	0
Boot Factor 5	0	0	0	0	100	0	0
Boot Factor 6	0	0	0	0	0	100	0

Appendix S2.C. I/O Correlation

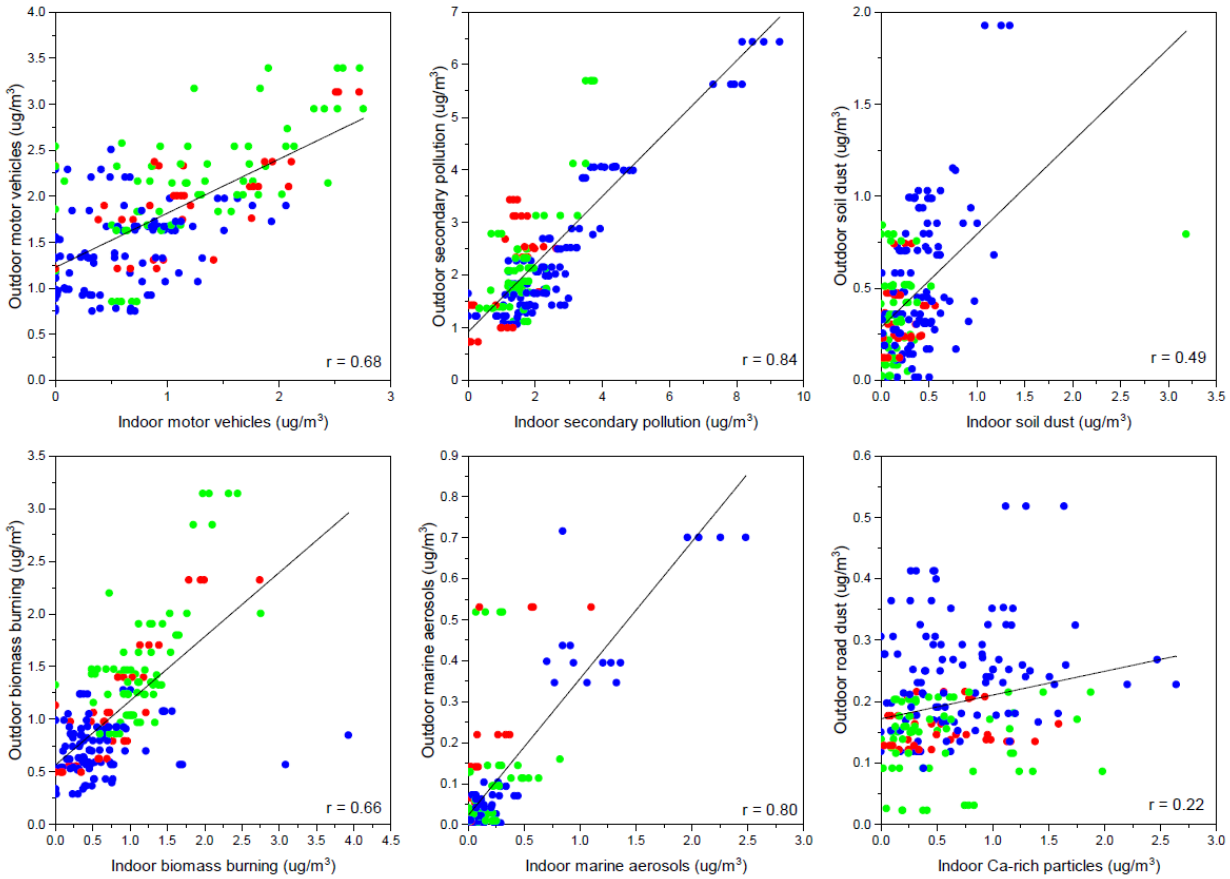

Figure S2.C. Correlation between sources contributing to indoor and outdoor $PM_{2.5}$ by season. Blue dots represent samples collected during the spring season. Red dots represent samples collected during fall season. Green dots represent samples collected during winter season.

Appendix S2.D. Source contribution by school

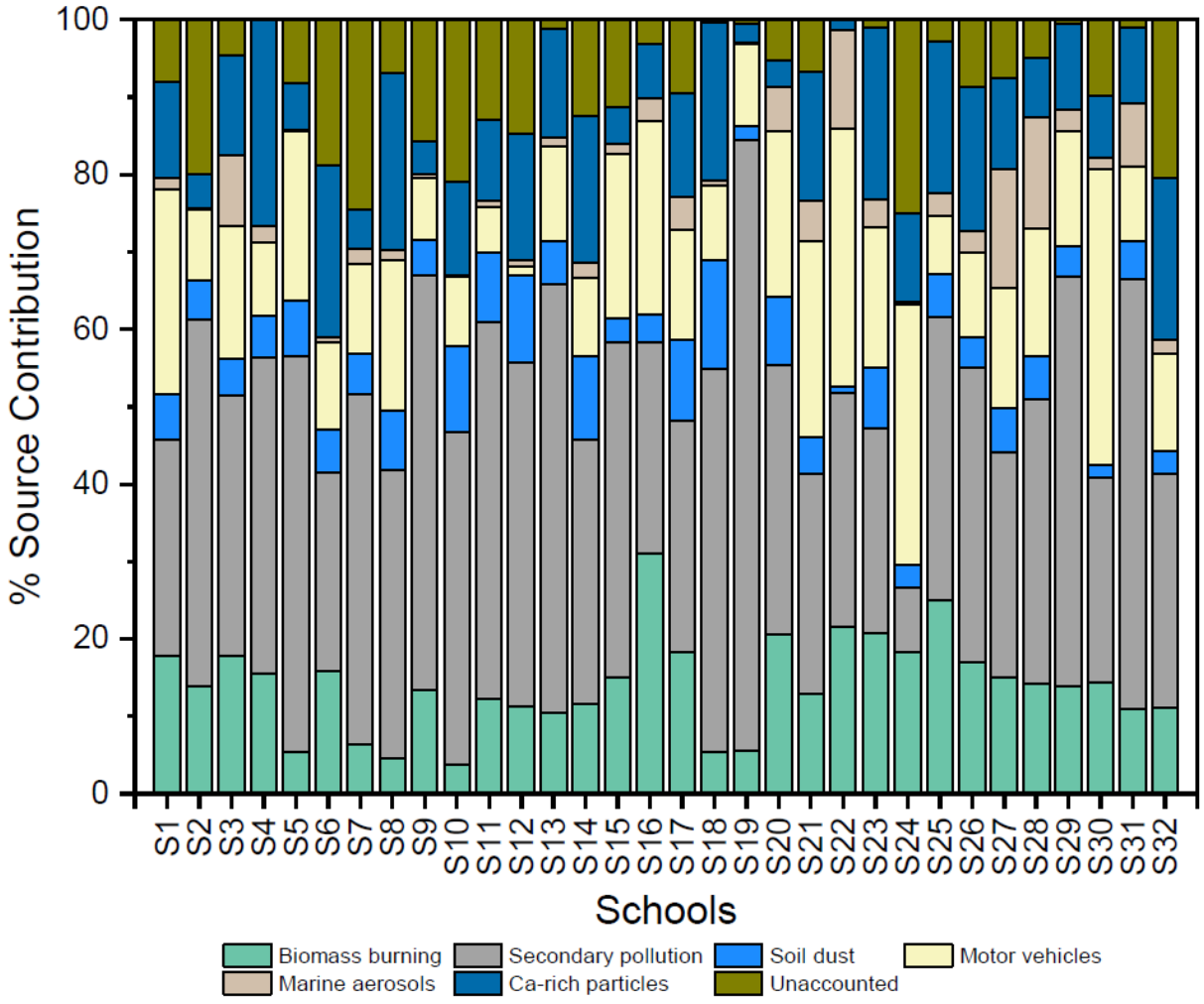

Figure S2.D. Source contribution by schools.

Appendix S2.E. Observed vs Predicted Concentrations

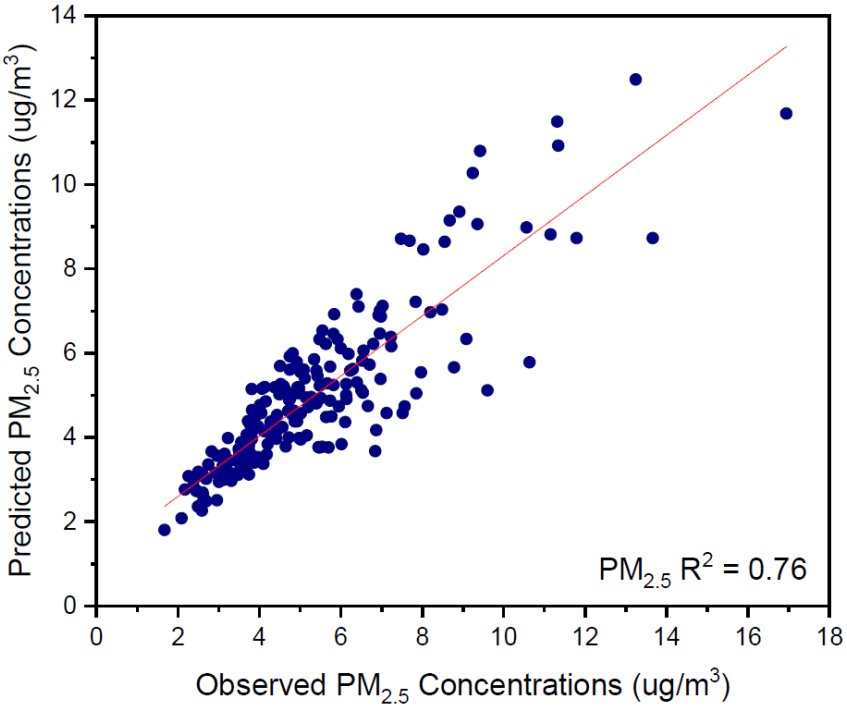

Figure S2.E.1. Observed vs Predicted Concentrations of indoor $PM_{2.5}$ mass.

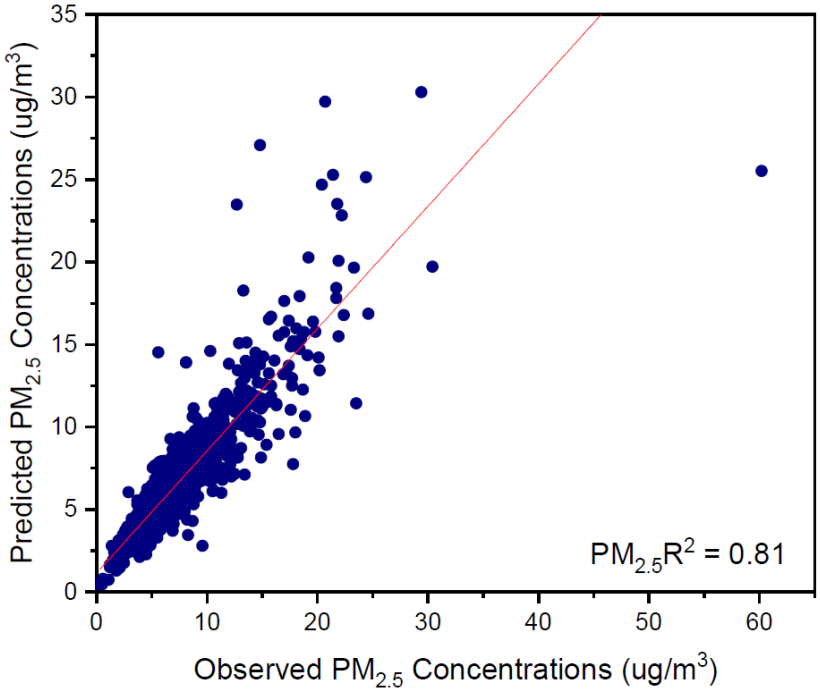

Figure S2.E.2. Observed vs Predicted Concentrations of outdoor $PM_{2.5}$ mass.

Supplemental Information for Chapter 3

S3.A. Summary of explanatory variables

Table S3.1. Rationale to include explanatory variables to the model.

Variable	Rationale	Reference
Temperature	Temperature differential will cause radon entry into the building. High indoor temperatures cause upward flow, which decrease basement pressure, leading to radon entry.	**(Kitto 2005)**
Pressure	Pressure differentials affect convective flow and radon entry to indoors. Low pressure causes an increase of radon flux from the soil to the atmosphere.	**(Chen et al. 1995; Sextro et al. 1987)**
Precipitation	Precipitation saturates the soil, creating a capping effect, preventing radon to escape to the atmosphere.	**(Walia et al. 2005)**
Dew Point	Dew point has an inverse relationship with radon.	**(Xie et al. 2015)**
Evaporation	Evaporation can cause radon to escape from the soil to the atmosphere. During cold months, soil dries more slowly, so evaporation is diminished, thus increasing indoor radon.	**(Schumann et al. 1994)**
Wind Speed	Wind speed affects radon entry rates. High wind speed increases building's ventilation rate and flushes radon from the soil near the building, thereby reducing radon entry rates.	**(Riley et al. 1996)**
Wind Direction	Wind direction can enhance subslab concentration, thus increasing radon entry rates. These changes will depend on the orientation of the building with respect to the prevailing wind direction.	**(Riley et al. 1999)**
HPBL	Height affects convective mixing in the atmosphere. Low HPBL will not allow mixing, concentrating radon near the surface, which will lead to an increase in indoor radon.	**(Porstendörfer et al. 1991)**
Soil Moisture	Soil moisture can increase/decrease the radon emanation coefficient depending on a threshold value, causing an increase/ decrease in radon exhalation. Also, soil moisture reduces permeability to air.	**(Hosoda et al. 2007; Sextro et al. 1987)**
Soil Temperature	Soil temperature affects the emanation of radon from the soil grains. Higher soil temperatures enhance the emanation.	**(Yang et al. 2019)**
Soil Porosity	Porosity affects the diffusion of radon through the soil. Higher porosity soil has higher permeabilities, increasing the rate of diffusion, thus leading to higher indoor radon levels.	**(Sextro et al. 1987; Yusuff et al. 2019)**

Soil Properties (sand, silt, clay)	The type of soil determines the permeability, which will affect the convective flow. Sand has larger permeability, followed by silt and clay. Convective flow is enhanced in sand, while molecular diffusion predominates in silt and clay soils.	**(Sextro et al. 1987)**
Uranium concentration	Radon-222 is part of the uranium-238 natural decay chain.	**(Sextro 1994)**
Thorium concentration	Thorium has a strong correlation with uranium due similar ionic radii. Thorium can be used as a reference for determine the original concentration of uranium as its mobility is extremely low in comparison with uranium.	**(Al-Hilal and Al-Ali 2016)**
Number of sunspots	Number of sunspots is used as a proxy for solar activity. Long-term radon levels have been shown to match solar cycle activity.	**(Yan et al. 2017)**
Longwave (upward, downward) and shortwave (upward, downward) radiation flux	Components of the surface radiation balance that controls weather and drives environmental changes.	**(Fujinami and Esaka 1988)**
Depth to bedrock	The distance to the bedrock affects soil radon exhalation. Radon flow is decreased if it comes from deep bedrock.	**(Saad et al. 2018)**
Permeability Rate	Permeability affects the convective flow of radon. Radon transports easily in permeable soils, offering more space for radon to move.	**(Sextro et al. 1987)**
Snow cover	Snow cover thickness can create a capping effect in the top layer of the soil, which will not allow radon flow to the atmosphere, thus redirecting the flow to indoors.	**(Fujimoto 1986; López-Coto et al. 2013)**
NDVI	Plants can absorb radon from the soil and groundwater via their root system. The release of radon to the atmosphere is higher in summer months.	**(Megonigal et al. 2020)**
Urbanization	Radon testing is associated with perceived risk and varies among levels of urbanization. Rural areas tend to have a lower testing rate due to an underestimation of the risk.	**(Hill et al. 2006; Poortinga et al. 2008; Zahnd et al. 2018)**
DCI score	Radon awareness and testing have been found to be less in communities with lower education levels, home value, and family income.	**(Halpern and Warner 1994; Hill et al. 2006; Zahnd et al. 2018)**
Organic Matter	Organic matter can absorb radon in sediments and tend to have high emanation coefficients.	**(Wong et al. 1992)**
Land-surface slope and Elevation	Hilltops and hillsides usually have thinner and rockier soils and are more permeable, which allows faster movement and greater emanation of radon.	**(Cho et al. 2015; Janssen et al. 1991)**
Relative Humidity	Relative humidity can affect radon exhalation. High relative humidity enhances radon exhalation.	**(Walia et al. 2005)**

S3.B. Zip code measurements

Table S3.2. Summary of zip codes level measurements.				
	Massachusetts	**Michigan**	**Minnesota**	**Total**
Number of zip codes	681	1,158	963	2,802
Total number of zip codes (observations)	441 (66,177)	881 (48,541)	811 (47,984)	2,133 (162,702)
Number of zip codes (observations) with at least 1 month with over 5 measurements	284 (16,084)	398 (5,532)	389 (9,189)	1,071 (30,805)
Number of zip codes (observations) with complete cases	179 (9,817)	288 (2,600)	335 (4,915)	802 (17,332)

S3.C. Hyper-parameter tuning

Table S3.3. Range of values used for hyper-parameter tuning.	
Parameter	**Values**
Number of trees	[100, 500, 1000]
Maximum depth	[6, 10, 20]
Maximum leaf nodes	[10, 25, 50]
Maximum features	[None, 0.9, sqrt]*
Random state	42

* None means that the model will include the total number of factors, sqrt = square root (factors)

S.3.D. Monthly and yearly indoor radon variations

Table S3.4. Annual variation of indoor radon levels. Average and interquartile range (Bq/m^3).			
Year	**Massachusetts**	**Michigan**	**Minnesota**
2005	119.74 (75.89 – 152.05)	115.51 (59.94 – 159.47)	153.63 (110.06 – 189.36)
2006	125.88 (80.78 – 159.10)	128.88 (68.88 – 172.46)	160.09 (120.04 – 191.74)
2007	123.39 (78.93 – 152.41)	117.75 (70.30 – 152.88)	177.47 (131.69 – 220.15)
2008	126.72 (82.76 – 159.84)	123.64 (66.88 – 163.73)	177.71 (133.82 – 214.97)
2009	118.88 (78.44 – 145.53)	122.55 (69.29 – 160.95)	187.48 (130.85 – 235.18)
2010	123.08 (81.40 – 153.85)	138.38 (82.79 – 180.56)	186.51 (143.77 – 228.99)
2011	121.39 (77.02 – 152.69)	151.84 (92.84 – 209.42)	188.69 (131.40 – 241.09)
2012	128.08 (82.28 – 163.64)	143.26 (82.94 – 196.10)	181.33 (132.46 – 228.48)
2013	129.01 (85.76 – 161.98)	148.99 (91.95 – 185.00)	163.50 (116.00 – 202.06)
2014	129.04 (85.96 – 164.73)	129.71 (73.46 – 167.03)	160.74 (113.92 – 198.32)
2015	124.47 (82.88 – 155.75)	136.39 (88.02 – 170.61)	149.23 (103.07 – 185.00)

2016	125.40 (82.86 – 157.53)	136.73 (85.10 – 168.97)	149.45 (104.22 – 185.93)
2017	127.03 (85.10 – 156.14)	143.93 (91.39 – 176.37)	158.02 (117.66 – 189.23)
2018	121.28 (79.86 – 154.58)	142.79 (95.17 – 178.47)	156.73 (109.52 – 190.29)

Table S3.5. Monthly variation of indoor radon levels across states. Average and interquartile range (Bq/m^3).

Month	**Massachusetts**	**Michigan**	**Minnesota**
January	130.08 (79.18 – 167.12)	131.01 (74.93 – 172.42)	169.95 (122.84 – 208.13)
February	137.31 (89.90 – 175.79)	132.82 (76.10 – 174.07)	163.67 (119.63 – 202.72)
March	131.50 (86.03 – 166.50)	127.80 (71.78 – 165.76)	164.42 (117.69 – 201.53)
April	132.88 (88.15 – 164.58)	132.94 (80.54 – 174.71)	160.73 (113.11 – 198.32)
May	124.21 (84.57 – 154.17)	154.78 (92.87 – 199.37)	156.71 (111.29 – 191.01)
June	114.54 (75.70 – 143.20)	128.26 (86.46 – 154.32)	144.94 (98.42 – 179.64)
July	98.73 (64.36 – 119.21)	114.41 (71.13 – 136.75)	120.46 (81.71 – 147.61)
August	106.75 (69.74 – 130.03)	120.56 (76.50 – 152.82)	130.09 (89.70 – 165.86)
September	122.15 (84.51 – 149.85)	153.88 (95.37 – 206.85)	149.91 (112.23 – 176.75)
October	138.06 (94.30 – 177.94)	143.52 (82.14 – 193.10)	171.70 (126.23 – 211.52)
November	135.21 (90.21 – 172.53)	148.77 (93.61 – 203.50)	154.25 (105.45 – 195.47)
December	136..41 (88.80 – 174.44)	136.50 (90.65 – 171.19)	157.91 (117.41 – 195.79)

S3.E. Variable Importance

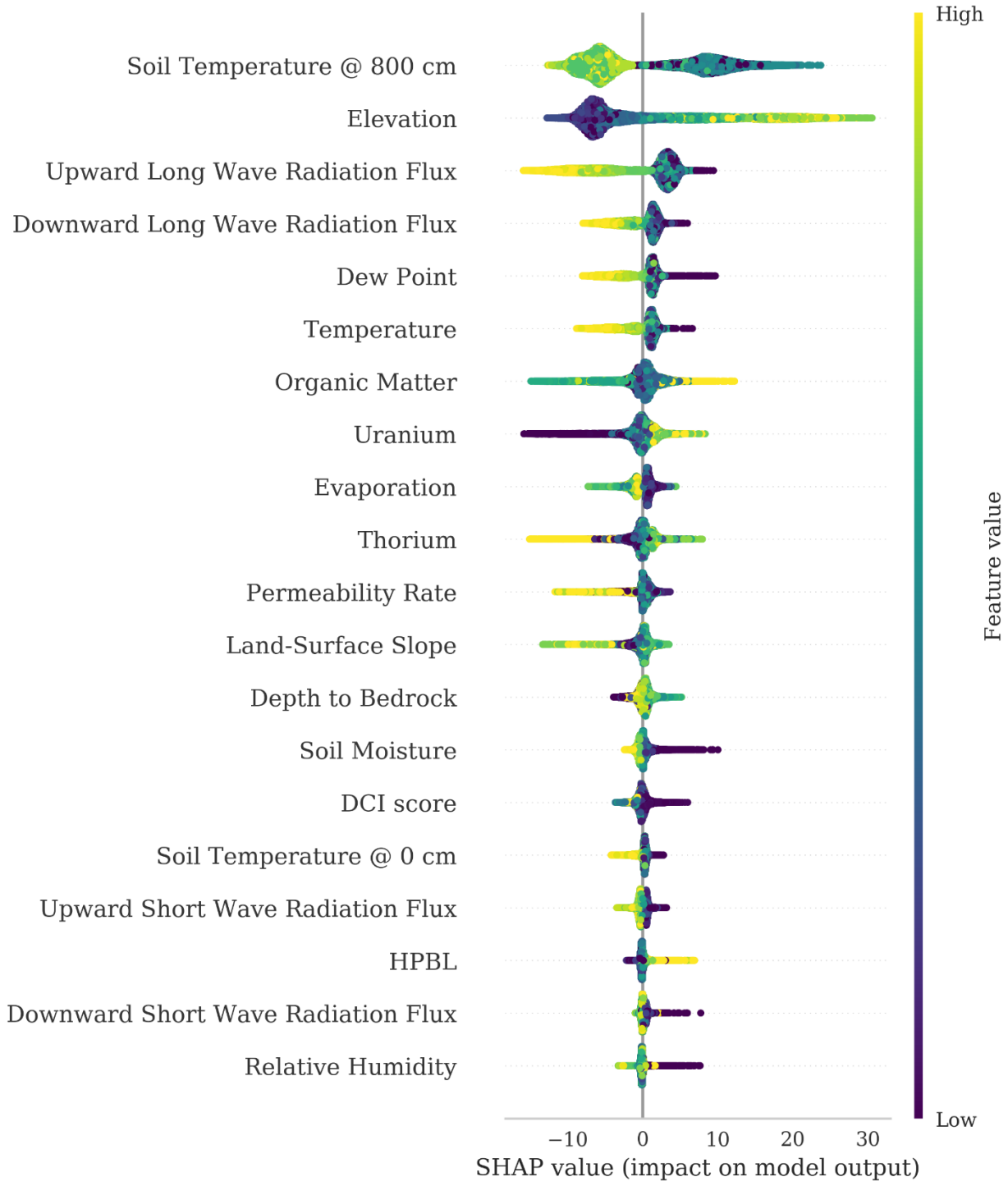

Figure S3.1 SHAP summary plot of the top 20 influential factors of radon levels in Massachusetts. The plot shows the distribution of Shapley values per factor. Factors are ordered based on importance. The color represents the factor value (yellow indicates high factor value

and purple indicates low factor value). For example, low soil temperatures at 800 cm depth increased the predicted radon levels.

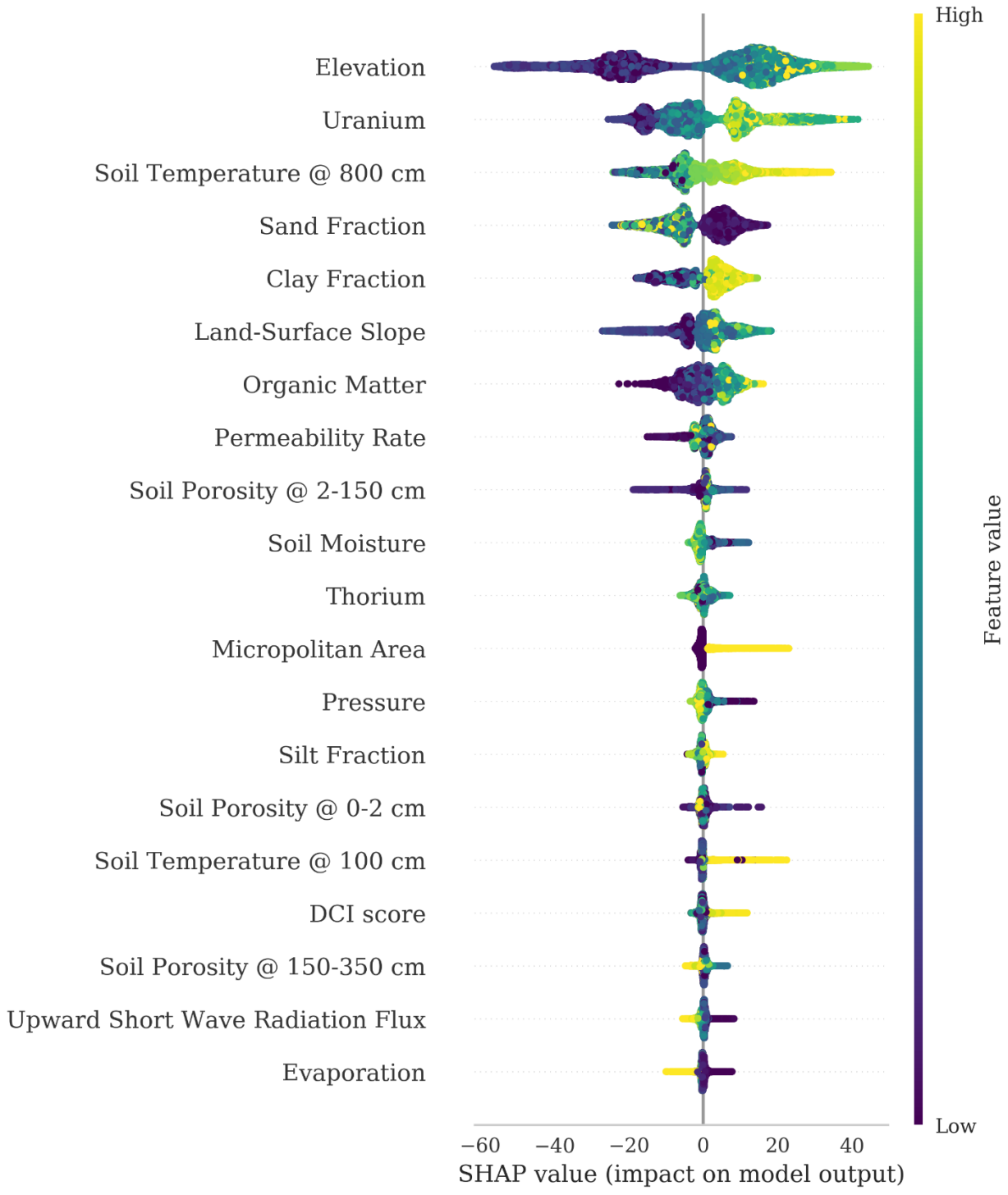

Figure S3.2 SHAP summary plot of the top 20 influential factors of radon levels in Michigan. The plot shows the distribution of Shapley values per factor. Factors are ordered based on importance. The color represents the factor value (yellow indicates high factor value and purple indicates low factor value). For example, houses located at higher elevation had an increase in the predicted radon levels.

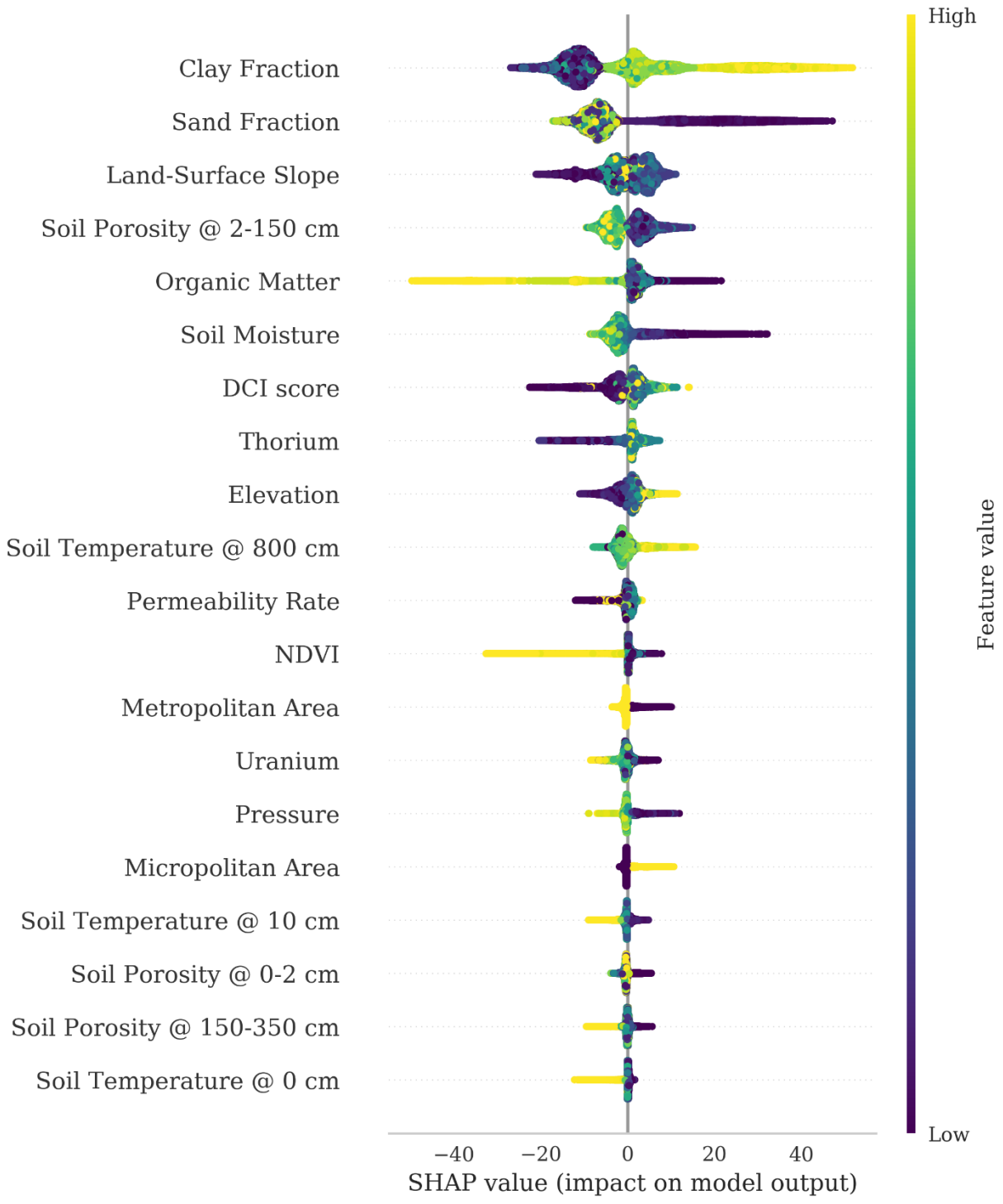

Figure S3.3 SHAP summary plot of the top 20 influential factors of radon levels in Minnesota. The plot shows the distribution of Shapley values per factor. Factors are ordered based on importance. The color represents the factor value (yellow indicates high factor value and purple indicates low factor value). For example, houses located in areas with higher clay fraction in the soil had an increase in the predicted radon levels.

Supplemental Information for Chapter 4

Table S4.1. Annual variation of residential radon levels.

Year	Mean	SD	Median	Min	Max
2001	381.36	534.20	185.00	0.00	3626.00
2002	355.38	496.37	175.75	0.00	3666.70
2003	311.84	500.99	133.20	0.00	3670.40
2004	270.31	442.88	118.40	0.00	3692.60
2005	282.10	469.67	111.00	0.00	3670.40
2006	281.40	432.46	129.50	0.00	3681.50
2007	303.31	474.56	133.20	0.00	3685.20
2008	306.16	470.87	136.90	0.00	3674.10
2009	304.60	471.90	140.60	0.00	3663.00
2010	312.83	479.85	140.60	0.00	3614.90
2011	288.32	448.11	129.50	0.00	3696.30
2012	279.28	449.29	125.80	0.00	3626.00
2013	257.40	412.71	114.70	0.00	3677.80
2014	247.19	417.41	99.90	0.00	3496.50
2015	294.30	481.12	114.70	0.00	3681.50

Table S4.2. Monthly variation of residential radon levels.

Month	Mean	SD	Min	Median	Max
1	336.44	499.42	0.00	151.70	3674.10
2	337.69	507.02	0.00	144.30	3626.00
3	288.22	461.00	0.00	122.10	3692.60
4	304.27	475.94	0.00	133.20	3696.30
5	280.15	448.68	0.00	122.10	3670.40
6	251.90	420.01	0.00	107.30	3640.80
7	214.46	377.44	0.00	92.50	3659.30
8	230.38	383.19	0.00	99.90	3496.50
9	278.03	439.80	0.00	125.80	3596.40
10	326.72	491.48	0.00	148.00	3681.50
11	320.90	473.20	0.00	144.30	3655.60
12	316.95	501.64	0.00	136.90	3589.00

Table S4.3. Exposure distribution.

	Min	10th Percentile	25th Percentile	Median	75th Percentile	90th Percentile	99th Percentile	Max
Downwind inverse distance	0.00	0.00	0.00	0.00	0.12	3.89	36.28	129.84
Upwind inverse distance	0.00	0.00	0.00	0.00	0.17	5.36	38.00	104.53

REFERENCES

Adgate JL, Goldstein BD, McKenzie LM. 2014. Potential public health hazards, exposures and health effects from unconventional natural gas development. Environmental Science & Technology 48:8307-8320.

Al-Hilal M, Al-Ali A. 2016. Investigation for uranium dispersion adjacent to cretaceous phosphatic outcrops in al-nassrieh basin, southern palmyrides, syria. Journal of Earth Science 27:786-795.

Almeida SM, Canha N, Silva A, do Carmo Freitas M, Pegas P, Alves C, et al. 2011. Children exposure to atmospheric particles in indoor of lisbon primary schools. Atmospheric Environment 45:7594-7599.

Alvarez-Pedrerol M, Rivas I, López-Vicente M, Suades-González E, Donaire-Gonzalez D, Cirach M, et al. 2017. Impact of commuting exposure to traffic-related air pollution on cognitive development in children walking to school. Environmental Pollution 231:837-844.

Amato F, Rivas I, Viana M, Moreno T, Bouso L, Reche C, et al. 2014. Sources of indoor and outdoor pm2. 5 concentrations in primary schools. Science of the Total Environment 490:757-765.

Apergis N, Hayat T, Saeed T. 2019. Fracking and infant mortality: Fresh evidence from oklahoma. Environmental Science and Pollution Research 26:32360-32367.

Aquilina N. The influence of meteorological parameters on indoor and outdoor radon concentrations: A preliminary case study. 2019.

Basagaña X, Esnaola M, Rivas I, Amato F, Alvarez-Pedrerol M, Forns J, et al. 2016. Neurodevelopmental deceleration by urban fine particles from different emission sources: A longitudinal observational study. Environmental Health Perspectives 124:1630-1636.

Beleche T, Cintina I. 2018. Fracking and risky behaviors: Evidence from pennsylvania. Economics & Human Biology 31:69-82.

Black KJ, McCoy SJ, Weber JG. 2019. Fracking and indoor radon: Spurious correlation or cause for concern? Journal of Environmental Economics and Management 96:255-273.

Bølviken B, Celius E, Nilsén R, Strand T. 2003. Radon: A possible risk factor in multiple sclerosis. Neuroepidemiology 22:87 - 94.

Borgini A, Tittarelli A, Ricci C, Bertoldi M, De Saeger E, Crosignani P. 2011. Personal exposure to pm2. 5 among high-school students in milan and background measurements: The eurolifenet study. Atmospheric environment 45:4147-4151.

Borgoni R, De Francesco D, De Bartolo D, Tzavidis N. 2014. Hierarchical modeling of indoor radon concentration: How much do geology and building factors matter? Journal of Environmental Radioactivity 138:227-237.

Bozlaker A, Peccia J, Chellam S. 2017. Indoor/outdoor relationships and anthropogenic elemental signatures in airborne pm2. 5 at a high school: Impacts of petroleum refining emissions on lanthanoid enrichment. Environmental science & technology 51:4851-4859.

Branion-Calles MC, Nelson TA, Henderson SB. 2016. A geospatial approach to the prediction of indoor radon vulnerability in british columbia, canada. Journal of Exposure Science & Environmental Epidemiology 26:554-565.

Breiman L. 2001. Random forests. Machine Learning 45:5-32.

Burkhart JF, Drescher A, Huber TP. Potential radon release during fracking in colorado. 2013.

Busby C, Mangano JJ. 2017. There'sa world going on underground—infant mortality and fracking in pennsylvania. Journal of Environmental Protection 8:381-393.

Canha N, Almeida SM, do Carmo Freitas M, Wolterbeek HT, Cardoso J, Pio C, et al. 2014. Impact of wood burning on indoor pm2. 5 in a primary school in rural portugal. Atmospheric environment 94:663-670.

Casey JA, Ogburn El Fau - Rasmussen SG, Rasmussen Sg Fau - Irving JK, Irving Jk Fau - Pollak J, Pollak J Fau - Locke PA, Locke Pa Fau - Schwartz BS, et al. 2015. Predictors of indoor radon concentrations in pennsylvania, 1989-2013. Environmental Health Perspectives 123:1130-1137.

Chatzidiakou L, Mumovic D, Summerfield AJ. 2012. What do we know about indoor air quality in school classrooms? A critical review of the literature. Intelligent Buildings International 4:228-259.

Chen C, Thomas DM, Green RE. 1995. Modeling of radon transport in unsaturated soil. Journal of Geophysical Research: Solid Earth 100:15517-15525.

Chen C, Zhao B. 2011. Review of relationship between indoor and outdoor particles: I/o ratio, infiltration factor and penetration factor. Atmospheric Environment 45:275-288.

Chen J, Falcomer R, Bergman L, Wierdsma J, Ly J. 2009. Correlation of soil radon and permeability with indoor radon potential in ottawa. Radiation Protection Dosimetry 136:56-60.

Chithra V, Nagendra SS. 2012. Indoor air quality investigations in a naturally ventilated school building located close to an urban roadway in chennai, india. Building and Environment 54:159-167.

Cho B-W, Choo CO, Kim MS, Hwang J, Yun U, Lee S. 2015. Spatial relationships between radon and topographical, geological, and geochemical factors and their relevance in all of south korea. Environmental Earth Sciences 74:5155-5168.

Coons T, Walker R. 2008. Community health risk analysis of oil and gas industry impacts in garfield county. Saccomanno Research Institute and Mesa State College: Grand Junction, CO.

Cretescu I, Isopescu DN, Lutic D, Soreanu G. 2019. Indoor air pollutants and the future perspectives for living space design. In: Indoor environment and health:IntechOpen.

Dai D, Neal FB, Diem J, Deocampo DM, Stauber C, Dignam T. 2019. Confluent impact of housing and geology on indoor radon concentrations in atlanta, georgia, united states. Science of The Total Environment 668:500-511.

Demokritou P, Kavouras IG, Ferguson ST, Koutrakis P. 2001. Development and laboratory performance evaluation of a personal multipollutant sampler for simultaneous measurements of particulate and gaseous pollutants. Aerosol Science & Technology 35:741-752.

Didan K. 2015. Mod13c2 modis/terra vegetation indices monthly l3 global 0.05deg cmg v006. NASA EOSDIS Land Processes DAAC:https://doi.org/10.5067/MODIS/MOD13C2.006.

Drolet J-P, Martel R, Poulin P, Dessau J-C, Lavoie D, Parent M, et al. 2013. An approach to define potential radon emission level maps using indoor radon concentration measurements and radiogeochemical data positive proportion relationships. Journal of Environmental Radioactivity 124:57-67.

Dueñas C, Fernández MC, Carretero J, Liger E, Pérez M. 1997. Release of 222rn from some soils. Annales Geophysicae 15:124-133.

Duval JS, Carson JM, Holman PB, Darnley AG. 2005. Terrestrial radioactivity and gamma-ray exposure in the united states and canada. US geological survey open-file report 1413:2005.

Echalar F, Artaxo P, Martins JV, Yamasoe M, Gerab F, Maenhaut W, et al. 1998. Long-term monitoring of atmospheric aerosols in the amazon basin: Source identification and apportionment. Journal of Geophysical Research: Atmospheres 103:31849-31864.

EPA. 2003. Epa assessment of risks from radon in homes. Office of Radiation and Indoor Air, United States Environmental Protection Agency, Washington, DC.

EPA. 2012. A citizen's guide to radon. The guide to protecting yourself and your family from radon. Washington, DC.

Finkel ML, Hays J. 2016. Environmental and health impacts of 'fracking': Why epidemiological studies are necessary. Journal of Epidemiology and Community Health 70:221.

Forns J, Dadvand P, Foraster M, Alvarez-Pedrerol M, Rivas I, López-Vicente M, et al. 2016. Traffic-related air pollution, noise at school, and behavioral problems in barcelona schoolchildren: A cross-sectional study. Environmental Health Perspectives 124:529-535.

Fromme H, Twardella D, Dietrich S, Heitmann D, Schierl R, Liebl B, et al. 2007. Particulate matter in the indoor air of classrooms—exploratory results from munich and surrounding area. Atmospheric Environment 41:854-866.

Fromme H, Diemer J, Dietrich S, Cyrys J, Heinrich J, Lang W, et al. 2008. Chemical and morphological properties of particulate matter (pm10, pm2. 5) in school classrooms and outdoor air. Atmospheric Environment 42:6597-6605.

Fujimoto K. 1986. Shielding effect of snow cover on terrestrial gamma exposure rate. Japanese Journal of Health Physics 21:3-8.

Fujinami N, Esaka S. 1988. Influence of solar radiation and heat emission on radon 222 daughter concentrations in surface air. Journal of Geophysical Research: Atmospheres 93:12627-12629.

Gaffin JM, Petty CR, Hauptman M, Kang C-M, Wolfson JM, Awad YA, et al. 2017. Modeling indoor particulate exposures in inner city school classrooms. Journal of exposure science & environmental epidemiology 27:451-457.

Gaikwad A, Shivhare N. 2019. Indoor air pollution-a threat. INDOOR AIR 7.

Gallaway MS, Berens AS, Puckett MC, Foster S. 2019. Understanding geographic variations of indoor radon potential for comprehensive cancer control planning. Cancer Causes & Control 30:707-712.

GAO. 1996. School facilities: America's schools report differing conditions. Washington, DC.

Gergen PJ, Togias A. 2015. Inner city asthma. Immunology and Allergy Clinics 35:101-114.

Goyal R, Khare M. 2009. Indoor–outdoor concentrations of rspm in classroom of a naturally ventilated school building near an urban traffic roadway. Atmospheric Environment 43:6026-6038.

Godwin C, Batterman S. 2007. Indoor air quality in michigan schools. Indoor Air 17:109-121.

Green Rochelle S, Smorodinsky S, Kim Janice J, McLaughlin R, Ostro B. 2004. Proximity of california public schools to busy roads. Environmental Health Perspectives 112:61-66.

Gundersen LC. The correlation between bedrock geology and indoor radon: Where it works and where it doesn't-some examples from the eastern united states. In: Proceedings of the The 1993 international radon conference, 1993.

Ha M, Hwang S-s, Kang S, Park N-W, Chang B-U, Kim Y. 2017. Geographical correlations between indoor radon concentration and risks of lung cancer, non-hodgkin's lymphoma, and leukemia during 1999–2008 in korea. International Journal of Environmental Research and Public Health 14.

Habil M, Massey DD, Taneja A. 2013. Exposure of children studying in schools of india to pm levels and metal contamination: Sources and their identification. Air Quality, Atmosphere & Health 6:575-587.

Halpern MT, Warner KE. 1994. Radon risk perception and testing: Sociodemographic correlates. Journal of Environmental Health 56:31-35.

Hauri DD, Huss A, Zimmermann F, Kuehni CE, Röösli M, for the Swiss National C. 2013. Prediction of residential radon exposure of the whole swiss population: Comparison of model-based predictions with measurement-based predictions. Indoor Air 23:406-416.

Hays J, Shonkoff SBC. 2016. Toward an understanding of the environmental and public health impacts of unconventional natural gas development: A categorical assessment of the peer-reviewed scientific literature, 2009-2015. PLOS ONE 11:e0154164.

Hill WG, Butterfield P, Larsson LS. 2006. Rural parents' perceptions of risks associated with their children's exposure to radon. Public Health Nurs 23:392-399.

Hochstetler HA, Yermakov M, Reponen T, Ryan PH, Grinshpun SA. 2011. Aerosol particles generated by diesel-powered school buses at urban schools as a source of children's exposure. Atmospheric environment 45:1444-1453.

Hosoda M, Shimo M, Sugino M, Furukawa M, Fukushi M. 2007. Effect of soil moisture content on radon and thoron exhalation. Journal of Nuclear Science and Technology 44:664-672.

Hou Y, Liu J, Li J. 2015. Investigation of indoor air quality in primary school classrooms. Procedia Engineering 121:830-837.

Janssen I, Stebbings Jh Fau - Essling MA, Essling Ma Fau - Dimino KM, Dimino Km Fau - Rodgers DL, Rodgers DL. 1991. Prediction of 222rn from topography in pennsylvania. Health Physics 61.

Janssen NA, van Vliet PH, Aarts F, Harssema H, Brunekreef B. 2001. Assessment of exposure to traffic related air pollution of children attending schools near motorways. Atmospheric environment 35:3875-3884.

Jing F, Shen XH, Kang CL, Xiong P. 2013. Variations of multi-parameter observations in atmosphere related to earthquake. Natural Hazards and Earth System Sciences 13:27.

Jovanović M, Vučićević B, Turanjanin V, Živković M, Spasojević V. 2014. Investigation of indoor and outdoor air quality of the classrooms at a school in serbia. Energy 77:42-48.

Kargbo DM, Wilhelm RG, Campbell DJ. 2010. Natural gas plays in the marcellus shale: Challenges and potential opportunities. Environmental Science & Technology 44:5679-5684.

Keeler GJ, Dvonch T, Yip FY, Parker EA, Isreal BA, Marsik FJ, et al. 2002. Assessment of personal and community-level exposures to particulate matter among children with asthma in detroit, michigan, as part of community action against asthma (caaa). Environmental Health Perspectives 110:173.

Khan SM, Taron JM, Goodarzi AA. 2020. Machine learning as a next-generation tool for indoor air radon exposure prediction. SAGE Research Methods Cases.

Kitto ME. 2005. Interrelationship of indoor radon concentrations, soil-gas flux, and meteorological parameters. Journal of Radioanalytical and Nuclear Chemistry 264:381-385.

Klepeis NE, Nelson WC, Ott WR, Robinson JP, Tsang AM, Switzer P, et al. 2001. The national human activity pattern survey (nhaps): A resource for assessing exposure to environmental pollutants. Journal of Exposure Science & Environmental Epidemiology 11:231-252.

Kline SW, Mose DG, Zietz I. 1989. The use of gamma aeroradioactivity in regional indoor radon prediction: A case study from fairfax county, virginia, USA. Journal of Radiological Protection 9:121-128.

Klusman RW, Jaacks JA. 1987. Environmental influences upon mercury, radon and helium concentrations in soil gases at a site near denver, colorado. Journal of Geochemical Exploration 27:259-280.

Koo L, Luk M, Mok M, Yuen J, Yuen T. 1997. Health effects from air conditioning: Epidemiologic studies on schools and offices in hong kong. Proceedings of indoor and built environment problems in Asia at Kuala Lumpur, Malaysia 45.

Leung DYC. 2015. Outdoor-indoor air pollution in urban environment: Challenges and opportunity. Frontiers in Environmental Science 2.

Li L, Blomberg AJ, Spengler JD, Coull BA, Schwartz JD, Koutrakis P. 2020. Unconventional oil and gas development and ambient particle radioactivity. Nature Communications 11:5002.

Ling-Da Y, Guang-Fu W, Ren-Jian Z, Lei-Ming Z, Yu S, Bing-Bing W, et al. 2013. Characterization and source apportionment of pm2.5 in an urban environment in beijing. Aerosol and Air Quality Research 13:574-583.

Litovitz A, Curtright A, Abramzon S, Burger N, Samaras C. 2013. Estimation of regional air-quality damages from marcellus shale natural gas extraction in pennsylvania. Environmental Research Letters 8:014017.

López-Abente G, Núñez O, Fernández-Navarro P, Barros-Dios JM, Martín-Méndez I, Bel-Lan A, et al. 2018. Residential radon and cancer mortality in galicia, spain. Science of The Total Environment 610-611:1125-1132.

López-Coto I, Mas JL, Bolivar JP. 2013. A 40-year retrospective european radon flux inventory including climatological variability. Atmospheric Environment 73:22-33.

Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. 2020. From local explanations to global understanding with explainable ai for trees. Nature Machine Intelligence 2:56-67.

Lustgarten A. 2009. Is new york's marcellus shale too hot to handle?‖. ProPublica [Online].

MacNaughton P, Eitland E, Kloog I, Schwartz J, Allen J. 2017. Impact of particulate matter exposure and surrounding "greenness" on chronic absenteeism in massachusetts public schools. International journal of environmental research and public health 14:207.

Masri S, Kang C-M, Koutrakis P. 2015. Composition and sources of fine and coarse particles collected during 2002–2010 in boston, ma. Journal of the Air & Waste Management Association 65:287-297.

Mazzei F, D'Alessandro A, Lucarelli F, Nava S, Prati P, Valli G, et al. 2008. Characterization of particulate matter sources in an urban environment. Science of The Total Environment 401:81-89.

McKenzie LM, Witter RZ, Newman LS, Adgate JL. 2012. Human health risk assessment of air emissions from development of unconventional natural gas resources. Science of The Total Environment 424:79-87.

Megonigal JP, Brewer PE, Knee KL. 2020. Radon as a natural tracer of gas transport through trees. New Phytologist 225:1470-1475.

Meng Q, Ashby S. 2014. Distance: A critical aspect for environmental impact assessment of hydraulic fracking. The Extractive Industries and Society 1:124-126.

Meng Q. 2015. Spatial analysis of environment and population at risk of natural gas fracking in the state of pennsylvania, USA. Science of The Total Environment 515-516:198-206.

Morawska L, Afshari A, Bae GN, Buonanno G, Chao CYH, Hänninen O, et al. 2013. Indoor aerosols: From personal exposure to risk assessment. Indoor Air 23:462-487.

Mose DG, Mushrush GW. Variable spacial and seasonal hazards of airborne radon. 1997.

Nelson Andrew W, Eitrheim Eric S, Knight Andrew W, May D, Mehrhoff Marinea A, Shannon R, et al. 2015. Understanding the radioactive ingrowth and decay of naturally occurring radioactive

materials in the environment: An analysis of produced fluids from the marcellus shale. Environmental Health Perspectives 123:689-696.

Paatero P, Tapper U. 1994. Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. Environmetrics 5:111-126.

Paatero P, Hopke PK. 2003. Discarding or downweighting high-noise variables in factor analytic models. Analytica Chimica Acta 490:277-289.

Polissar AV, Hopke PK, Paatero P, Malm WC, Sisler JF. 1998. Atmospheric aerosol over alaska: 2. Elemental composition and sources. Journal of Geophysical Research: Atmospheres 103:19045-19057.

Poortinga W, Cox P Fau - Pidgeon NF, Pidgeon NF. 2008. The perceived health risks of indoor radon gas and overhead powerlines: A comparative multilevel approach. Risk Analysis 28:235-248.

Porstendörfer J, Butterweck G, Reineking A. 1991. Diurnal variation of the concentrations of radon and its short-lived daughters in the atmosphere near the ground. Atmospheric Environment Part A General Topics 25:709-713.

Rasmussen SG, Ogburn EL, McCormack M, Casey JA, Bandeen-Roche K, Mercer DG, et al. 2016. Association between unconventional natural gas development in the marcellus shale and asthma exacerbations. JAMA Internal Medicine 176:1334-1343.

Reff A, Eberly SI, Bhave PV. 2007. Receptor modeling of ambient particulate matter data using positive matrix factorization: Review of existing methods. Journal of the Air & Waste Management Association 57:146-154.

Riley WJ, Gadgil AJ, Bonnefous YC, Nazaroff WW. 1996. The effect of steady winds on radon-222 entry from soil into houses. Atmospheric Environment 30:1167-1176.

Riley WJ, Robinson AL, Gadgil AJ, Nazaroff WW. 1999. Effects of variable wind speed and direction on radon transport from soil into buildings: Model development and exploratory results. Atmospheric Environment 33:2157-2168.

Rivas I, Viana M, Moreno T, Pandolfi M, Amato F, Reche C, et al. 2014. Child exposure to indoor and outdoor air pollutants in schools in barcelona, spain. Environment International 69:200-212.

Rodell M, Houser PR, Jambor U, Gottschalck J, Mitchell K, Meng CJ, et al. 2004. The global land data assimilation system. Bulletin of the American Meteorological Society 85:381-394.

Saad AF, Abdallah RM, Hussein NA. 2018. Physical and geometrical parameters controlling measurements of radon emanation and exhalation from soil. Applied Radiation and Isotopes 137:273-279.

Sarnat JA, Long CM, Koutrakis P, Coull BA, Schwartz J, Suh HH. 2002. Using sulfur as a tracer of outdoor fine particulate matter. Environmental Science & Technology 36:5305-5314.

Scheffer T. 1999. Error estimation and model selection. Künstliche Intell 13:46-48.

Schumann RR, Gundersen LCS, Tanner AB. 1994. Geology and occurrence of radon. Radon Prevalence, Measurements, Health Risks and Control.

Seinfeld JH, Pandis SN. 2016. Atmospheric chemistry and physics: From air pollution to climate change:John Wiley & Sons.

Sextro RG, Moed BA, Nazaroff WW, Revzan KL, Nero AV. 1987. Investigations of soil as a source of indoor radon. In: Radon and its decay products, Vol. 331:American Chemical Society, 10-29.

Sextro RG. 1994. Radon and the natural environment. Radon: Prevalence, Measurement, Health Risks and Control ASTM MNL 15:9-32.

Smedje G, Norbäck D. 2008. Irritants and allergens at school in relation to furnishings and cleaning. Indoor Air 11:127-133.

Śmiełowska M, Marć M, Zabiegała B. 2017. Indoor air quality in public utility environments—a review. Environmental Science and Pollution Research 24:11166-11176.

Song X-H, Polissar AV, Hopke PK. 2001. Sources of fine particle composition in the northeastern us. Atmospheric Environment 35:5277-5286.

Spira-Cohen A, Chen LC, Kendall M, Lall R, Thurston GD. 2011. Personal exposures to traffic-related air pollution and acute respiratory health among bronx schoolchildren with asthma. Environmental Health Perspectives 119:559.

Srebotnjak T, Rotkin-Ellman M. 2014. Fracking fumes: Air pollution from hydraulic fracturing threatens public health and communities. Natural Resources Defense Council.

SRTM. 2013. Shuttle radar topography mission (srtm) global. OpenTopography:https://doi.org/10.5069/G9445JDF.

Stanley FKT, Irvine JL, Jacques WR, Salgia SR, Innes DG, Winquist BD, et al. 2019. Radon exposure is rising steadily within the modern north american residential environment, and is increasingly uniform across seasons. Scientific Reports 9:18472.

Steck DJ. 2009. Annual average indoor radon variations over two decades. Health Physics 96.

Stranger M, Potgieter-Vermaak SS, Van Grieken R. 2008. Characterization of indoor air quality in primary schools in antwerp, belgium. Indoor Air 18:454-463.

Stranger M, Potgieter-Vermaak SS, R. VG. 2008. Characterization of indoor air quality in primary schools in antwerp, belgium. Indoor Air 18:454-463.

Sunyer J, Esnaola M, Alvarez-Pedrerol M, Forns J, Rivas I, López-Vicente M, et al. 2015. Association between traffic-related air pollution in schools and cognitive development in primary school children: A prospective cohort study. PLoS Medicine 12:e1001792.

Swiedler EW, Muehlenbachs LA, Chu Z, Shih J-S, Krupnick A. 2019. Should solid waste from shale gas development be regulated as hazardous waste? Energy Policy 129:1020-1033.

Tait DR, Santos IR, Maher DT, Cyronak TJ, Davis RJ. 2013. Enrichment of radon and carbon dioxide in the open atmosphere of an australian coal seam gas field. Environmental Science & Technology 47:3099-3104.

Tétreault L-F, Doucet M, Gamache P, Fournier M, Brand A, Kosatsky T, et al. 2016. Severe and moderate asthma exacerbations in asthmatic children and exposure to ambient air pollutants. International Journal of Environmental Research and Public Health 13:771.

UNSCEAR. 2000. Sources and effects of ionizing radiation unscear 2000 report to the general assembly, with scientific annexes volume i: Sources. United Nations.

USEPA. 2003. Epa assessment of risks from radon in homes. Office of Radiation and Indoor Air, United States Environmental Protection Agency, Washington, DC.

USEPA. 2016. Consumer's guide to radon reduction. How to fix your home.

Varma S, Simon R. 2006. Bias in error estimation when using cross-validation for model selection. BMC Bioinformatics 7:91.

Viana M, Kuhlbusch TAJ, Querol X, Alastuey A, Harrison RM, Hopke PK, et al. 2008. Source apportionment of particulate matter in europe: A review of methods and results. Journal of Aerosol Science 39:827-849.

Vienneau D, de Hoogh K, Hauri D, Vicedo-Cabrera Ana M, Schindler C, Huss A, et al. 2017. Effects of radon and uv exposure on skin cancer mortality in switzerland. Environmental Health Perspectives 125:067009.

Vogiannis EG, Nikolopoulos D. 2015. Radon sources and associated risk in terms of exposure and dose. Frontiers in Public Health 2.

Walia V, Virk HS, Yang T, Mahajan S, Walia M, Bajwa BS. 2005. Earthquake prediction studies using radon as a precursor in n-w himalayas, india: A case study. Terrestrial Atmospheric and Oceanic Sciences 16:775-804.

Wallace L, Williams R. 2005. Use of personal-indoor-outdoor sulfur concentrations to estimate the infiltration factor and outdoor exposure factor for individual homes and persons. Environmental science & technology 39:1707-1714.

Wei W, Ramalho O, Malingre L, Sivanantham S, Little JC, Mandin C. 2019. Machine learning and statistical models for predicting indoor air quality. Indoor Air 29:704-726.

Weiss KB, Gergen PJ, Crain EF. 1992. Inner-city asthma: The epidemiology of an emerging us public health concern. Chest 101:362S-367S.

WHO. 2009. Who handbook on indoor radon: A public health perspective:World Health Organization.

Wichmann J, Lind T, Nilsson M-M, Bellander T. 2010. Pm2. 5, soot and no2 indoor–outdoor relationships at homes, pre-schools and schools in stockholm, sweden. Atmospheric Environment 44:4536-4544.

Wong CS, Chin Y-P, Gschwend PM. 1992. Sorption of radon-222 to natural sediments. Geochimica et Cosmochimica Acta 56:3923-3932.

Xie D, Liao M, Kearfott KJ. 2015. Influence of environmental factors on indoor radon concentration levels in the basement and ground floor of a building – a case study. Radiation Measurements 82:52-58.

Xu H, Guinot B, Shen Z, Ho KF, Niu X, Xiao S, et al. 2015. Characteristics of organic and elemental carbon in pm2. 5 and pm0. 25 in indoor and outdoor environments of a middle school: Secondary formation of organic carbon and sources identification. Atmosphere 6:361-379.

Xu Y, Sajja M, Kumar A. 2019. Impact of the hydraulic fracturing on indoor radon concentrations in ohio: A multilevel modeling approach. Frontiers in Public Health 7.

Yan R, Woith H, Wang R, Wang G. 2017. Decadal radon cycles in a hot spring. Scientific Reports 7:12120.

Yang J, Busen H, Scherb H, Hürkamp K, Guo Q, Tschiersch J. 2019. Modeling of radon exhalation from soil influenced by environmental parameters. Science of The Total Environment 656:1304-1311.

Yousef AMM, Zimami K. 2019. Indoor radon levels, influencing factors and annual effective doses in dwellings of al-kharj city, saudi arabia. Journal of Radiation Research and Applied Sciences 12:460-467.

Yusuff IM, Adagunodo TA, Omoloye MA, Olanrewaju AM. 2019. Interdependency of soil-gas radon-222 concentration on soil porosity at different soil-depths. Journal of Physics: Conference Series 1299:012099.

Zahnd WE, Mueller-Luckey GS, Ratnapradipa K, Smith T. 2018. Predictors and spatial variation of radon testing in illinois, 2005-2012. Journal of Public Health Management and Practice 24.

Zwoździak A, Sówka I, Krupińska B, Zwoździak J, Nych A. 2013. Infiltration or indoor sources as determinants of the elemental composition of particulate matter inside a school in wrocław, poland? Building and Environment 66:173-180.