HYBRID DISPERSION/ LAND USE REGRESSION MODELING FOR IMPROVING AIR POLLUTANT CONCENTRATION ESTIMATES

by

Andrew Ryan Michanowicz

BS, Juniata College, 2006

MPH, University of Pittsburgh, 2009

Submitted to the Graduate Faculty of

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Doctor of Public Health

University of Pittsburgh

2014

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This dissertation was presented

by

Andrew Ryan Michanowicz

It was defended on

December 1st, 2014

and approved by

Committee Chair: James Peterson, PhD, Associate Professor, Environmental and Occupation Health, Graduate School of Public Health, University of Pittsburgh

Ravi K. Sharma, PhD, Assistant Professor, Behavioral and Community Health Sciences, Graduate School of Public Health, University of Pittsburgh

James P. Fabisiak, PhD, Associate Professor, Environmental and Occupation Health, Graduate School of Public Health, University of Pittsburgh

Kyra Naumoff Shields, PhD, Project Manager, Center for Energy Development and Global Health, Colorado State University

Dissertation Advisor: Jane E. Clougherty, MSc, ScD, Assistant Professor, Environmental and Occupation Health, Graduate School of Public Health, University of Pittsburgh

Jane E. Clougherty, MSc, ScD
James Peterson, PhD

HYBRID DISPERSION/ LAND USE REGRESSION MODELING FOR IMPROVING POLLUTANT CONCENTRATION ESTIMATES

Andrew R. Michanowicz, DrPH

University of Pittsburgh, 2014

ABSTRACT

The overall objective of this dissertation was to examine the utility of incorporating source-meteorological interaction information from two commonly employed atmospheric dispersion models into the land use regression technique for predicting ambient NO_2 and $PM_{2.5}$. Ultimately, we are interested in obtaining highly resolved spatiotemporal pollutant estimates to examine the attenuation of health effect estimate bias that may result from exposure model misspecification. A multi-pollutant sampling campaign was conducted across six successive weekly sampling sessions in the summer and winter seasons of 2011-2013 in Pittsburgh, PA. As a preliminary investigation, predictions from a roadway dispersion model (Caline3) were included as an independent predictor in pre-constructed winter season LUR models for NO_2. Caline3 output improved out-of-sample model fitness and added an additional portion of unexplained variation (3-10% by leave-one-out cross-validated R^2) in NO_2 observations compared to the standard LUR models. Correspondingly, the AERMOD dispersion model was implemented to predict $PM_{2.5}$ from local and regional stationary sources in a similar hybrid framework. As per cross-validated R^2 and *RMSE*, AERMOD predictions improved overall model fitness and explained an additional 9-13% in out-of-sample variability in summer and winter $PM_{2.5}$ models. Both dispersion model output functioned similarly when incorporated into standard LUR models, effectively displacing

the respective GIS-based covariates, corroborating model interpretability, and capturing the greatest degree of improvements at nearby, high-density source locations. To examine the potential for spatially-differential exposure measurement improvement in health effect estimation studies, we applied LUR and hybrid LUR/ dispersion model $PM_{2.5}$ predictions to non-sampled locations and observed non-Berkson-type measurement error only when the modeling domain was restricted to a near-source (<1km) environment. By a simple stochastic simulation, we demonstrated that a well characterized dispersion-derived geographic covariate, defined by a robust variance about the monitoring locations, can theoretically result in less exposure measurement error and exposure misclassification. Therefore, highly refined spatiotemporal information can improve out-of-sample prediction accuracy; however, the statistical fidelity remains constrained by the degree of source contribution captured by monitoring locations. These findings have important public health implications for understanding air pollutant exposure measurement error derived from typical LUR studies. In the absence of a spatially dense monitoring network, we demonstrated that AERMOD can produce a spatiotemporally resolved prediction surface compared to typical GIS-based covariates across a large urban-to-suburban domain with pertinent pollutant sources and complex topography.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I am forever grateful to my advisor, Jane Clougherty, whose scientific prowess is the exemplar of human progress. Her rigorous curiosity and scrutiny of minutiae have been quintessential principles of my scientific tutelage.

This body of work would also not be possible without the intellectual contributions from Jessie Carr Shmool. Her clever perceptiveness was a steadfast resource of enlightened clarity.

Brett Tunno, Sara Gillooly, Sheila Tripathy, Ellen Kinnee, Courtney Roper, Lauren Chubb, Jeffrey Howell, and Leah Cambal were instrumental for their contributions to field work and collaboration within our laboratory.

I would also like to thank Ravi Sharma, James Petereson, James Fabisiak, and Kyra Naumoff Shields for their contributions as reading committee members.

I would particularly like to acknowledge the staff scientists at the Allegheny County Health Department. Portions of this project would not have been possible without both knowledge and data sharing.

I would also like to thank the individuals who have been particularly influential to my scientific progress including, Phillip Johnson, Daniel Bain, Dan Volz, Kyle Ferrar, John Graham, Bruce Pitt, and Bernard Goldstein.

DEDICATION

To all of my ancestors that successfully engender offspring, especially the few that I've had the pleasure of knowing…

ABBREVIATIONS

ACHD	Allegheny County Health Department
AERMOD	American Meteorological Society/Environmental Protection Agency Regulatory Model
AIC	Akaike Information Criterion
ASOS	Integrated Surface Observation System
AQS	Air Quality System
Caline3	Caline3QHCR line (roadway) source dispersion model
CALPUFF	Air quality puff dispersion model
CMAQ	Community Multi-scale Air Quality Model
CO	Carbon Monoxide
CO_2	Carbon Dioxide
FRM	Federal Reference Method
GIS	Geographic Information System
ISHD	Integrated Surface Hourly Data
IDW	Inverse Distance Weighting
LOOCV	Leave-One-Out-Cross-Validation
LUR	Land Use Regression
N_2	Nitrogen
NAAQS	National Ambient Air Quality Standards
NMSE	Normalized Mean Square Error

NO_2	Nitrogen Dioxide
NO_x	Oxides of Nitrogen
O_2	Oxygen
O_3	Ozone
PAHs	Polycyclic Aromatic Hydrocarbons
$PM_{2.5}$	Particulate Matter less than 2.5 microns in aerodynamic diameter
RMSE	Root Mean Square Error
SD	Standard Deviation
SIP	State Implementation Plan
SO_2	Sulfur Dioxide
TRAP	Traffic-Related Air Pollution
USEPA	United States Environmental Protection Agency
VIF	Variance Inflation Factor
VOCs	Volatile Organic Compounds

1.0 INTRODUCTION

1.1 ATMOSPHERIC POLLUTION

Earth's atmosphere is believed to have been formed following the accretion of an interstellar cloud of gas and dust where less dense materials coalesced farther from the core. Earth's current atmosphere is composed primarily of the gases N_2 (78%), and O_2 (21%), whose relative abundances have depended upon various physical forcings (e.g., uptake and release from crustal material) spanning approximately 4,567 million years. The remaining constituents, therefore, represent less than 1% of the atmosphere. Water vapor is highly variable and can reach a concentration abundance of 3% in the lower atmosphere depending upon evaporation and precipitation rates. Nonetheless, trace gases and aerosols play a vital role in regulating Earth's complex biosphere and trace gas abundances have changed dramatically over the past two centuries (Seinfeld and Pandis, 2012).

Atmospheric pollution follows a series of events where, the generation of pollutants is released from a source into the atmosphere; pollutants are transported and transformed; and effects from air pollution are defined at a receptor point (e.g., humans, vegetation, materials, and ecosystems). Airborne particles have increased dramatically since the Industrial Revolution, and have led to unforeseen consequences including the detrimental urban smog events in Donora, PA and London, UK, for example. In addition to processes that directly emit PM into the air (primary

PM), PM can also be formed when certain gaseous pollutants including sulfur dioxide (SO2), various oxides of nitrogen (NOx), volatile organic chemicals (VOCs), and ammonia (NH3) condense into particulates (secondary PM) after release from a source. The chemical fates of air pollutants are inextricably coupled with complex physical and chemical processes in the atmosphere, and depending on their functional lifetimes, pollutants can exhibit a tremendous degree of spatial and temporal variability.

Airborne particles or particulate matter (PM) is a term used to describe the sum of tiny solid and liquid particles suspended in the atmosphere. PM is a chemically, physically and biologically diverse mixture of materials including dusts, organic chemicals, smoke, soot, metals, acids, and liquid droplets that originate from numerous natural and man-made sources. Not surprisingly, PM produced by diesel combustion engines, coal-fired power plants, and volcanoes differs substantially in composition. A large contributor of anthropogenic air pollution is traffic-related air pollution (TRAP), which has become is a major concern in urban areas, where the majority the world's population now lives (HEI, 2010; Heilig, 2012). In addition to PM, TRAP also includes significant quantities of gaseous and aerosolized pollutants such as: nitrogen oxides (NO_x), carbon monoxide (CO), carbon dioxide (CO_2), volatile organic compounds (VOCs), and polycyclic aromatic hydrocarbons (PAHs). Thus, environmental and human health effects from atmospheric pollution are related to physical and chemical properties including airborne concentrations, PM particle size, and overall chemical and elemental compositions.

1.2 ADVERSE HUMAN HEALTH EFFECTS OF AIR POLLUTION

The average human adult takes about 20,000 breaths per day consisting of 10-25 m^3 of exchanged air (0.14-0.29 L/s) (Hinds, 2012). Although mechanisms are not fully known, one-in-eight global deaths is currently attributable to polluted air (World Health Organization, 2012). Exposures to high levels of air pollution over short periods of time, or lower levels over longer time periods, are both cause for concern and both short-term and long-term effects on health have been demonstrated (Brunekreef and Holgate, 2002). No evidence has been obtained for a threshold below which adverse effects do not occur (Pope, 2000).

Numerous human health studies and subsequent reviews have linked exposures to certain air pollution with increased hospitalization for cardiopulmonary (heart and lung) diseases, decreased lung function, respiratory symptoms, adverse reproductive effects and premature death. The references cited to document these effects are typical of a large body of accumulating scientific literature [for reviews see: (Anderson et al., 2012; Bell et al., 2013; Bernstein et al., 2004; Brunekreef and Holgate, 2002; Cohen et al., 2005; Dockery, 2009; Faustini et al., 2014; Hoek et al., 2013; Holland et al., 1979; Kampa and Castanas, 2008; KuÈnzli et al., 2000; Matus et al., 2012; Pope III, 2000; Pope III and Dockery, 2006; Rückerl et al., 2011; Samet et al., 2000; Spengler and Sexton, 1983; Wang et al., 2014; World Health Organization, 2012)].

Air pollution effects are not restricted to the respiratory system since small particles can be absorbed into the circulatory system, as deduced from markers of systemic inflammation and oxidative stress throughout the body (Araujo, 2011; Huttunen et al., 2012). It is likely such responses are linked with numerous health outcomes including asthma and chronic bronchitis; and triggering premature death from preexisting heart and lung disease. Therefore, accurate human exposure assessment to air pollution is fundamental to understanding the true global and local burden of air pollution-related disease.

1.3 EXPOSURE ASSESSMENT METHODOLOGIES

As it is not practical to measure personal exposures for all individuals in large cohort studies, exposure assessments that estimate proximal ambient air pollution, usually at the residential address, are commonly employed (Jerrett et al., 2005). These predicted exposures are then included as explanatory variables in a regression model to evaluate a health effect parameter of interest. However, the use of predicted air pollution levels as surrogates of true exposure, are inevitably affected by measurement error and uncertainty (Basagaña et al., 2013). Therefore, it has been assumed that exposure predictions with less measurement error relative to the unknown true exposures will result in improved health effect estimates (Jerrett et al., 2005). The degree to which exposure prediction, and subsequent exposure measurement error engenders uncertainty and bias in health-effect estimates has invoked research interests (Alexeeff et al., 2014; Basagaña et al., 2013; Szpiro et al., 2011a; Szpiro et al., 2011b).

The most straightforward approach of exposure prediction employed has been location-based methods, which rely on the degree of propinquity to an emission source to proxy for human exposure (Baccarelli et al., 2009; Brender et al., 2011; Hoek et al., 2002; Maheswaran and Elliott, 2003; Van Roosbroeck et al., 2007). Subsequent refinements and variations of methodologies have included statistical interpolation (Jerrett et al., 2001; Künzli et al., 2005; Sahu and Mardia, 2005; Wong et al., 2004), land use regression (Brauer et al., 2003; Briggs et al., 1997; Clougherty et al., 2013b; Jerrett et al., 2005), air quality models (Ainslie et al., 2008; Bell, 2006; Gulliver and

Briggs, 2011; KuÈnzli et al., 2000), and hybrid applications combining these methods (Arunachalam et al., 2014; Bekhor and Broday, 2013; Isakov et al., 2009; Johnson et al., 2010; Kloog et al., 2014; Kloog et al., 2012; Mölter et al., 2010b; Su et al., 2008; Van den Hooven et al., 2012). Attempts to resolve spatio-temporal concentrations of ambient $PM_{2.5}$ and NO_X over larger areas (e.g., Northeastern U.S) have leveraged satellite-derived aerosol optical depth (AOD) measurements (Chang et al., 2013; Chudnovsky et al., 2013; Kim et al., 2013; Kloog et al., 2014; Kloog et al., 2012; Lee et al., 2011; Lin et al., 2013; Nordio et al., 2013). Spatial resolution of satellite-based AOD measurements have improved substantially from 10 x 10 km^2 grid (Levy et al., 2007) to 1 x 1 km^2 (Chang et al., 2013; Chudnovsky et al., 2013) and recently to 200m x 200m localized daily predictions using a series of mixed effects models (Kloog et al., 2014).

Due to improved methods using geographic information systems (GIS), land use regression (LUR) has emerged as a standard tool for intra-urban exposure assessment (Jerrett et al., 2005). LUR models employ relatively simple inputs and provide significantly higher spatial resolution than proximity-based or purely statistical interpolation methods (Jerrett et al., 2005). The LUR process combines a relatively large number of systematically distributed air pollution measures with “land use” variables (e.g., population density) usually managed in GIS (Fig. 1). Statistical relationships between air pollutant measurements and land use predictor variables are derived using ordinary least squares multiple linear regression (Hoek et al., 2008). The resulting stochastic model is then applied to non-sampled areas by exploiting the observed pollutant variance explained by the statistically robust predictor (land use) variables. Exposure predictions are then included as explanatory variables, usually in linear or logistic regression models for a health outcome of interest. Therefore, the LUR method for epidemiological study relies upon the quantity and quality of pollutant measurements, fidelity of the GIS (e.g., variability represented by pertinent geographic covariates) (Madsen et al., 2011), and the variability of geographic covariates in the subject population of the study cohort (Szpiro et al., 2011a).

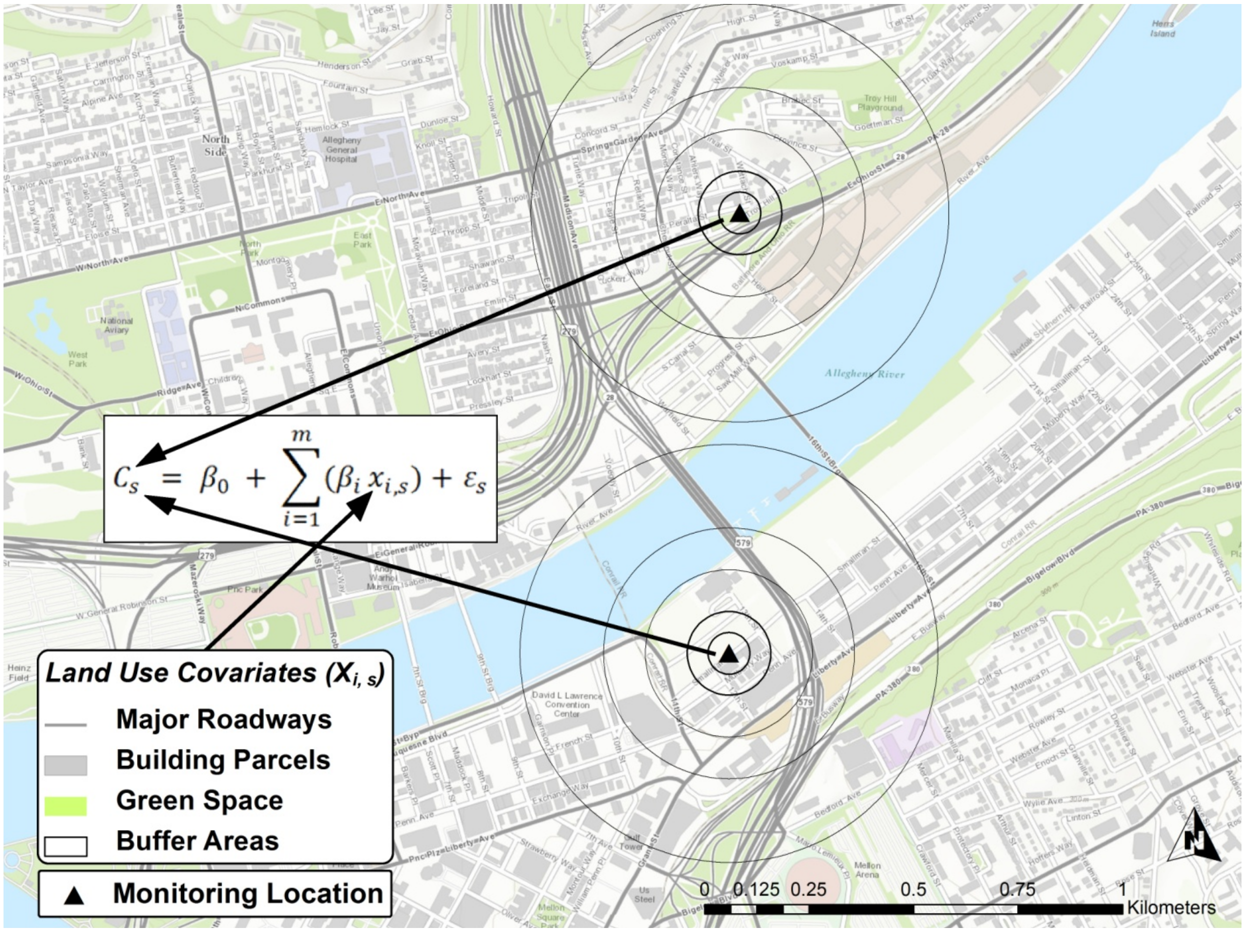

Figure 1. Components of a land use regression model with pollutant measures from monitoring locations as the dependent variable and land use characteristics within buffer areas as the independent predictor variables

The Health Effect Institute provided a critical review of traffic-related air pollution exposure models noting a fundamental limitation of LUR - its inability to represent the true contribution (associated variance) of traffic-related emissions (HEI, 2010). This phenomenon is exemplified when adjacent land-use and predictor variables in LUR are measured and summed as nearest distances from- or as densities within circular areas (Euclidean buffers) (Fig. 1). These isotropic areal units fail to capture small-scale spatiotemporal pollutant variability governed in part by interactions between emissions sources and meteorological processes (eg., upwind vs. downwind advective motion) (Ainslie et al., 2008; Jerrett et al., 2005; Su et al., 2008; Wilton, 2011).

In an attempt to better represent near-roadway source-concentration variance, prior LURs have built-in some measures of temporal variability by including meteorological covariates (e.g., wind speed or mixing heights) (Arain et al., 2007; Clougherty et al., 2009; Jerrett et al., 2007; Su et al., 2008), or by weighting source-concentration relationships by predominant wind direction (Clougherty et al., 2008; Mavko et al., 2008; Van den Hooven et al., 2012). Ainslie et al. (2008) and Su et al. (2008) attempted to capture atmospheric dispersion using a source-area concentration grid of distributed emissions under varying atmospheric conditions and three-dimensional wedge shaped buffers based on predominant wind fields. Likewise, Wilton (2011) incorporated meteorologically-varying covariates as volume sources in a CALPUFF Lagrangian puff model (Scire et al., 1990). Wilton et al. (2010) and Lindström et al. (2013) both attempted Caline3/LUR modeling efforts with each reporting inconsistent model improvement, albeit more parsimonious and interpretable models.

Ideally, estimation of ground-level concentration of air pollutants should include emissions characteristics, meteorologically-related dispersion, transformation and removal processes (Bekhor and Broday, 2013), along with a means of validation (Chang and Hanna, 2004). Mathematical models can be used simulate transport of pollutants deterministically, as a function of source characteristics (e.g., location, strength, size) and temporally-varying meteorological conditions (e.g., wind speed, direction, atmospheric stability) (Briant et al., 2013; Chang and Hanna, 2004). Modeling, therefore provides a supplement to air quality monitoring by providing information that cannot be provided by other means (Barratt, 2013). Of the many types of models employed, Gaussian-type plume dispersion models are the most widely developed and utilized regulatory atmospheric dispersion models (Ristic et al., 2014). Gaussian models assume a Gaussian distribution of the fluid plume in both the vertical and horizontal directions. Therefore, under steady-state conditions, by assuming the downwind velocity vector coincides with the *x* axis, the width of the plume in the *y* and *x* axes can be determined by the respective standard deviations σ_x and σ_y given sufficient averaging times. Dispersion models have been employed extensively in regulatory air quality management, and to a lesser degree in human exposure assessments (Jerrett et al., 2005; Johnson et al., 2010; Marshall et al., 2008; Mölter et al., 2010b; Nafstad et al., 2003; Nyberg et al., 2000; Van den Hooven et al., 2012). Wide adoption of air quality models has been hindered by relatively intensive data input requirements, high costs, and programming demands; however, recent Microsoft graphical user interfaces (e.g., Lakes Environmental, BREEZE Software) have benefitted ease of use.

In comparison with LUR approaches that can provide detailed spatial resolution, dispersion modeling offers high temporal variability with theoretically unlimited spatial resolution. Furthermore, it has also been demonstrated that LUR-derived exposure misclassification may depend more so on how much of the true spatial variability is explained by the geographic covariates in the exposure model, and not necessarily the accuracy of the predictions (Alexeeff et al., 2014; Szpiro et al., 2011a), especially when LUR models are constructed from a small number of measurement sites (Basagaña et al., 2013). Ergo, standard LUR could be strengthened by incorporating source-meteorology interaction information, thus producing theoretically- or physically-based exposure estimates as opposed to predictions derived purely from empirical relationships (Jerrett et al., 2005; Su et al., 2008; Wilton et al., 2010). Gaussian plume dispersion model output nested within LUR, therefore, offers a complementary framework – where spatio-temporal variability of pollutant source-concentration relationships are derived deterministically, thereby improving physical model interpretability and reliability of exposure estimates.

1.4 DISSERTATION OBJECTIVES

In acknowledging the emergence of land use regression modeling for exposure assessment in epidemiological studies, the overall objective of this dissertation is to examine the utility of incorporating source-meteorological interaction information from two commonly employed atmospheric dispersion models into the land use regression technique for both NO_2 and $PM_{2.5}$.

Chapter 2 of the dissertation specifically aims to better capture near-roadway source-concentration variability of NO_2 across Pittsburgh, PA by incorporating model output from the Caline3QHCR line- (roadway) source dispersion model into winter-only LUR models.

Chapter 3 examines the utility of incorporating industrial source-meteorological information from the AERMOD modeling system into an LUR predicting $PM_{2.5}$ across Pittsburgh, PA. In contrast to the Caline3 model, AERMOD can provide detailed resolution in the spatio-temporal variability of air pollutants emitted from stationary sources in both simple and complex terrain scenarios.

In Chapter 4, we examine the impact of measurement error on health effect estimates from LUR and hybrid AERMOD/ LUR models. We constructed two annual $PM_{2.5}$ prediction models by combining summer and winter measurements (presented in Chapter 3) with (1) local EPA AQS measures; and (2) local EPA AQS measures and annual long-term AERMOD predictions. Specifically, we examine AERMOD's potential to impact measurement error and subsequent acute and chronic health-effect bias. We used a simulated cohort of 5,000 residential addresses to examine the potential magnitude of bias and variance inflation in measurement error between annualized LUR and LUR/ AERMOD modeling frameworks.

The final portion of the dissertation summarizes the overall scientific contribution, and attempts to place the findings in the relative context of public health and risk assessment disciplines. The final summary includes a short description of planned epidemiologic studies utilizing the hybrid modeling framework presented here, and also provides suggestions for future research in the field of exposure assessment.

2.0 HYBRID CALINE3/ LUR MODEL FOR PREDICTING NO_2

2.1 INTRODUCTION

Land use regression (LUR) has emerged as a standard tool for intra-urban air pollution exposure assessment in recent years (Brauer et al., 2003; Briggs et al., 1997; Clougherty et al., 2013b; Jerrett et al., 2005). LUR, however, offers limited capability to incorporate source-meteorology interaction information, thereby producing estimates based on empirical relationships, rather than a theoretical-physical basis (Jerrett et al., 2005; Su et al., 2008; Wilton et al., 2010). Thus, there is now growing interest in incorporating principles of air dispersion modeling into LUR in the hopes of improving accuracy, interpretability and generalizability of such models (Gulliver and Briggs, 2011; Lindström et al., 2013; Mölter et al., 2010b; Wilton et al., 2010).

LUR quantifies statistical relationships between measured pollution concentrations and emission source indicators to estimate concentrations at non-sampled locations (Hoek et al., 2008). Significant traffic-source indicators have included total length of roadway (Henderson et al., 2007), distance from nearest roadway (Gilbert et al., 2005) and traffic count density (Ross et al., 2006) within various radial buffer distances. The statistical relationships derived from these metrics in LUR are based on observed values and statistical principles, and generally fail to account for short-term interactions between sources and atmospheric conditions (Wilton et al., 2010). Moreover, traffic-related pollution can lead to complex spatio-temporal patterns in air pollution, necessitating dedicated near-roadway sampling (Gulliver and Briggs, 2011; Mölter et al., 2010b), beyond the data obtained from fixed-site monitors (Jerrett et al., 2005), and refined spatial analysis.

Prior LURs have been attempted to incorporate some measure of temporal variance into source-concentration relationships by including meteorological covariates (e.g., mean wind speed or direction) (Arain et al., 2007; Clougherty et al., 2009; Jerrett et al., 2007; Su et al., 2008), or by weighting source-concentration relationships by predominant wind direction (Clougherty et al., 2009; Mavko et al., 2008; Van den Hooven et al., 2012). Ainslie et al. (2008) and Su et al. (2008) attempted to capture atmospheric dispersion using a source-area concentration grid of distributed emissions under varying atmospheric conditions. Likewise, Wilton (2011) incorporated meteorologically-varying covariates as volume sources in a CALPUFF Lagrangian puff model (Scire et al., 1990). To the best of our knowledge, only two other hybrid line-source dispersion/LUR modeling efforts have been attempted (Lindström et al., 2013; Wilton et al., 2010) with each reporting variable model improvement, albeit more parsimonious and interpretable models.

Ideally, estimation of ground-level concentration of air pollutants should include emissions characteristics, meteorologically-related dispersion, transformation and removal processes (Bekhor and Broday, 2013), along with a means of validation (Chang and Hanna, 2004). Of the many types of models employed, Gaussian-type plume dispersion models are the most widely developed and utilized regulatory atmospheric dispersion models (Ristic et al., 2014). Gaussian dispersion models have been employed extensively in regulatory air quality management, and to a lesser degree in human exposure assessments (Jerrett et al., 2005; Johnson et al., 2010; Marshall et al., 2008; Mölter et al., 2010b; Nafstad et al., 2003; Nyberg et al., 2000; Van den Hooven et al., 2012). Gaussian dispersion models can be used simulate transport of pollutants deterministically, as a function of source characteristics (e.g., location, strength, size) and temporally-varying meteorological conditions (e.g., wind speed, direction, atmospheric stability) (Briant et al., 2013;

Chang and Hanna, 2004). Therefore, standard LUR could be strengthened by incorporating source-meteorology interaction information from dispersion model output, thus producing theoretically- or physically-based exposure estimates as opposed to predictions derived purely from empirical relationships (Jerrett et al., 2005; Su et al., 2008; Wilton et al., 2010).

In this chapter, we aimed to improve prediction of NO_2 across Pittsburgh, PA, USA, by incorporating the Caline3QHCR line- (roadway) source dispersion model (Benson, 1992; Eckhoff and Braverman, 1995) output as an independent covariate into pre-constructed LUR models. Our multi-pollutant spatial saturation study was designed to disentangle impacts of multiple pollution sources (e.g., legacy industry, vehicle traffic), and to assess potential modifiers of source-concentration relationships (e.g., elevation) across an urban-to-suburban landscape (Shmool et al., 2014). We utilized two successive years of winter-season only NO_2 measurements. We evaluated improvements in model fit by adding Caline3 predictions as an additional term to three pre-constructed LUR models and observed changes in regression coefficients and covariate significance. Specifically, we tested (1) Caline's effectiveness given diurnal traffic variability in a weekday-only (year 1) vs. full-week (year 2) LUR models; (2) whether Caline's improvements in fitting accuracy differed across sampling intervals by including modeled predictions in a combined years LUR model (year 1 + year 2); and (3) Caline's effect on LUR predictions as a function of traffic density and distance from roadway in an attempt to better explain near-source variability.

2.2 METHODS

2.2.1 NO_2 Measurements for Pittsburgh

NO_2 was sampled across two successive winter seasons from early January through late March of 2012 and 2013. Year 1 comprised of six successive 5-day (Monday through Friday) sampling sessions and is hereafter referred to as the weekday model. Year 2 was comprised of six successive 7-day (Monday through Sunday) sampling sessions and is referred to hereafter as the full-week model. We employed a spatial saturation design to characterize intra-urban variability in multiple air pollutants (e.g., $PM_{2.5}$, NO_2, O_3, SO_2) across the greater Pittsburgh, PA metropolitan area, systematically allocating sampling sites across complex topography and emission source regimes, as detailed in Shmool et al. (2014).

NO_2 samples were collected using Ogawa passive badge samplers (Ogawa & Co. USA Inc., Pompano Beach, FL, USA) housed in weather-tight shelters and mounted three meters above street-level. Ogawa badges were analyzed via water-based extraction and spectrophotometry (Thermo Scientific Evolution 60S UV-Visible Spectrophotometer). Co-located NO_2 measurements were well correlated (r = 0.93) across eight (four per year) randomly-selected monitoring locations. Measurements were corrected for blank samples which ranged from 0.01 to 0.05 ppb.

2.2.2 Study Domain and Site Selection

Our study domain encompassed a contiguous 500 km^2 area containing the Pittsburgh metropolitan area and key local industrial sources, demarcated at census administrative boundaries to enable merging with socioeconomic and health data in future epidemiological applications. We used a geographic information system (GIS) to systematically allocate monitoring locations cross-stratified across important local pollution sources (e.g., traffic, steel manufacturing) and potential topographic modifiers of source-concentration interactions (e.g., elevation) using ArcMap 10.0-10.3 (ESRI, Redlands, CA, USA) and Geospatial Modeling Environment, V. 0.7.2 (Spatial Ecology, LLC).

Specifically, we anticipated variance in the local pollutant regime to be characterized by: 1) traffic density, 2) industrial density (weighted emissions: $PM_{2.5}$ + NO_X + SO_2 + VOCs), and 3) elevation at 30 m^2 grid resolution. We used stratified random sampling to select monitoring locations representing all possible combinations of high and low source intensities. Site selection and GIS-based covariate calculations are detailed elsewhere (Shmool et al., 2014). Notably, the traffic density metric used for site allocation was total daily vehicle counts from all primary roadways, and an estimated 500 vehicles/ day for secondary roadways, multiplied by road segment length (meters). Resultant traffic densities were extrapolated as a Gaussian decay function from roadway centerlines, producing a continuous kernel density surface. The dichotomization for high vs. low traffic density was chosen at the 70th percentile, given the left-skewed distribution and goal of over-sampling hypothesized high-pollution areas (Shmool et al., 2014).

Integrated NO_2 samples were collected across six successive sampling sessions with six randomly-selected sites per session, resulting in a total of 36 measurements per season. To minimize temporal confounding across sessions, sites were systematically allocated across sessions to balance emissions-indicator strata and spatial coverage. A randomly-selected subset of 12 sites, representing all possible combinations ($n=2^3$) of emissions source strata, were retained in both years (Fig. 2) for direct comparison. Thus, two winter-only sampling campaigns covered 60 unique locations with a total of 72 NO_2 measurements.

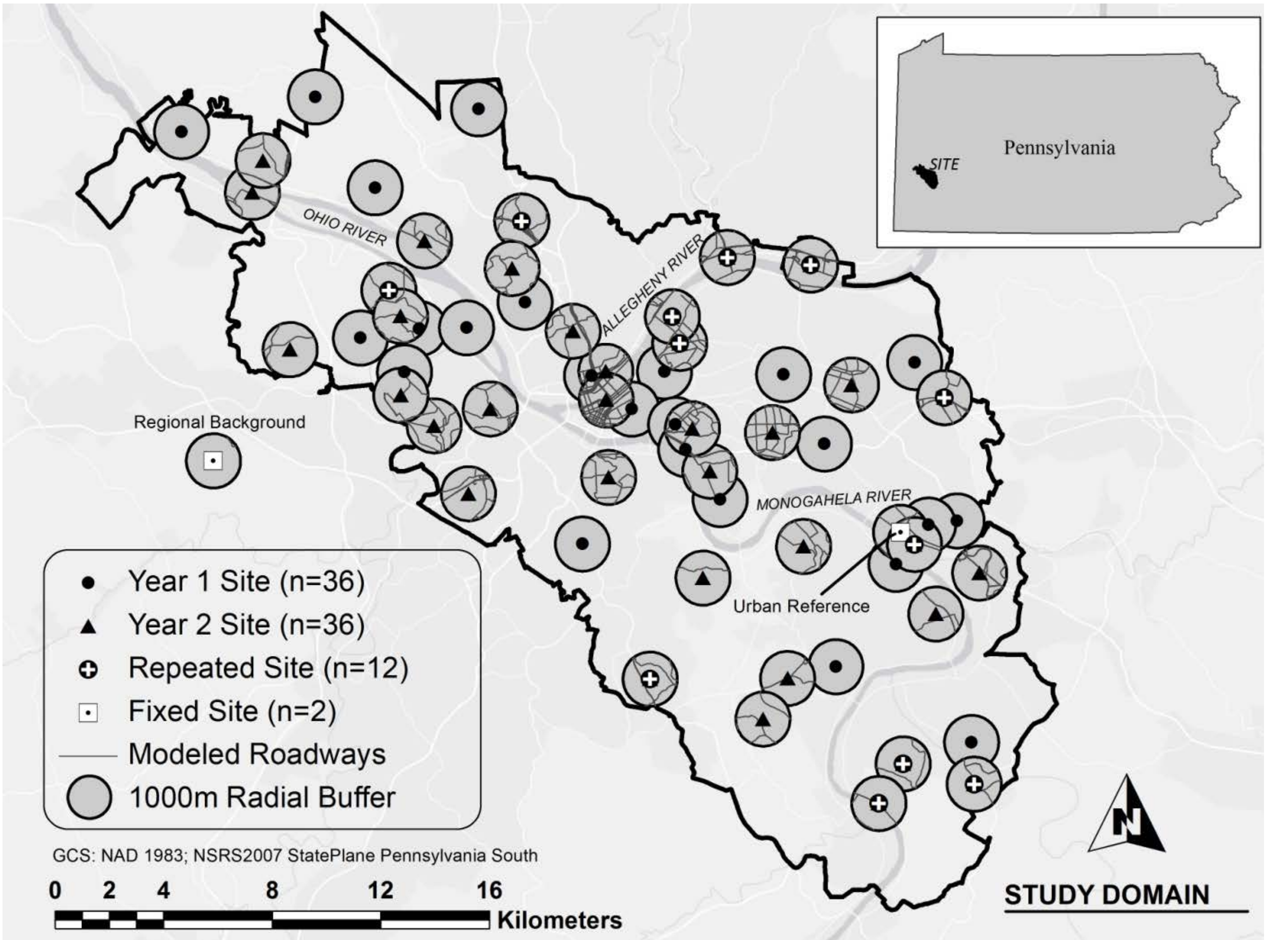

Figure 2. Study domain of Greater Pittsburgh Metropolitan Area and year 1 and 2 sampling locations and reference sites. Primary roadways modeled using Caline3 are shown in 1000 m radial buffers

2.2.3 Temporal Reference

Two continuous reference sites were sampled each weekly session to adjust for temporal variability in pollutant measures and to limit spatiotemporal bias in comparing measures across sessions (Brauer et al., 2003; Henderson et al., 2007; Hoek et al., 2008). A 'regional background' site was selected in a county park (Settler's Cabin Park) upwind from the study area and away from local sources, about 4.0 km west of the study domain (Fig. 2). The site was categorized in the hypothesized lowest-concentration source strata (low industry, low traffic, high elevation). The second reference site (Braddock, PA – in the eastern part of our domain) was designated an ‘urban reference’ site (high industry, high traffic, low elevation) (Fig. 2). From year 1 sampling, we found that the temporal reference adjustment method influenced observed source-concentration relationships, and the mean of the background and urban reference sites was more appropriate for temporally adjusting NO_2 given consistent near-zero concentrations at the background site (Shmool et al., 2014).

2.2.4 Caline3 Line-Source Dispersion Model

We implemented Caline3 (Caline3QHCR) line source dispersion model (Benson and Baishiki, 1980; Eckhoff and Braverman, 1995) using *CalRoads View* user interface (Lakes Environmental, Waterloo, Ontario, CA), to simulate primary vehicle emissions within 1000 m of sampling sites. Given the site-specific source characteristics and session-specific meteorological conditions, Caline3 uses a Gaussian, steady-state dispersion model to calculate transport of nonreactive aerosols, providing hourly concentration estimates at discrete receptors. The discrete modeling receptors were defined as the 60 unique sampling locations. We modeled a nonreactive gaseous pollutant environment by choosing *CalRoads*’ particulate matter designation with a settling velocity of 0.0 g/s to estimate total NO_X ($NO + NO_2$) similarly to Wilton et al. (2010). We assigned a fleet-wide-specific NO_X emission factor obtained for all mobile source types and all road types (excluding off-network) for Allegheny County, PA using the U.S. EPA’s Motor Vehicle Emission Simulator (MOVES) 2010a (USEPA, 2010), and derived a weighted average of 1.325 (g/vehicle-mile) of NO_X for all roadway segments.

Primary roadways within a 1,000 m radial distance of each sampling site were included in the Caline3 model, totaling 8,274 modeled straight-line, one-way traffic roadway links (Pennsylvania Department of Transportation, 2013) (Fig. 2). The 1,000 m radial buffer was chosen to capture all roadway emissions given an estimated 80-90% decrease in roadway NO_2 concentrations within 115-570 m (Karner et al., 2010), as evidence for roadway effects beyond 1000m is mixed (Jerrett et al., 2007; Su et al., 2009; Wilton et al., 2010). Caline3 output was calculated utilizing hourly meteorological data corresponding to the precise sampling session, encompassing an integrated average derived from 120 modeled hours for the weekday model, an integrated average from 168 modeled hours for the full-week model. Typical graphical model output is shown in Fig. 3. Surface characteristics (e.g., albedo, Bowen ratio) were estimated with AERSURFACE (Lakes Environmental, Waterloo, Ontario, CA) for an urban setting during winter conditions.

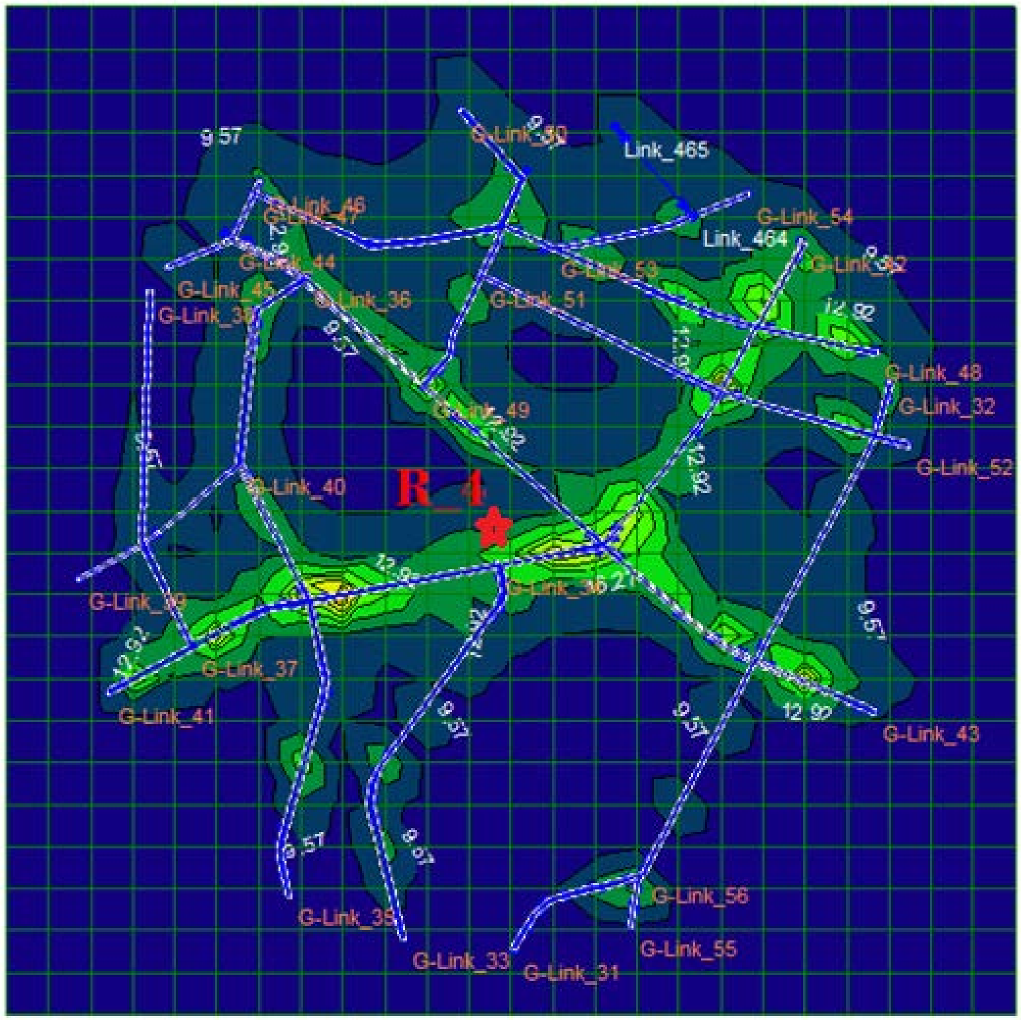

Figure 3. Typical Caline3 model output indicating estimated concentration contours from modeled roadway links within 1000m buffer area of receptor/sampling site (R_4)

2.2.5 Meteorological Data

Hourly meteorological data (e.g., wind speed, wind direction, temperature, precipitation, ceiling height) were downloaded from the National Climate Data Center (NCDC) in TD-3505 (ISHD – full archival) format, and used as both Caline3 inputs and as independent and interaction covariates in LUR model building. Radiosonde upper air data was collected at the Pittsburgh National Weather Service station located in Moon Township, PA, approximately 20 miles upwind of Pittsburgh and was obtained from the National Oceanic and Atmospheric Administration (NOAA). Surface and profile files were formatted in AERMET View 7.3.0 (Lakes Environmental, Waterloo, Ontario). Planetary boundary layer estimates were generated using both surface and profile data with AERMET View and were imputed into the RAMMET View 5.2.0 (Lakes Environmental, Waterloo, Ontario) mixing height estimator to produce hourly urban mixing height estimates and atmospheric stability categories.

2.2.6 LUR Model Building

LUR models were first constructed without Caline3 to test the marginal benefit of incorporating dispersion into a LUR modeling context, as a supplemental may be most applicable elsewhere. GIS-based covariates were calculated across a range of source indicator categories, each at monitoring location (Table 1). The following model-building approach similar to Clougherty et al. (2013b) was implemented: 1) candidate indicators were grouped by source category (e.g., traffic indicators, meteorology, industrial emissions) and ranked according to the nonparametric bivariate correlations (Spearman correlations, $p < 0.1$) with temporally-adjusted NO_2 concentrations by the formula:

$$\boldsymbol{adjConc_{sj} = \frac{Conc_{sj}}{[Ref_{mean}]_j} * [Ref_{mean}]_k}$$

(2.1)

Where $adjConc_{sj}$ is the temporally-adjusted pollutant concentration at monitoring site s during sampling session j, $Conc_{sj}$ is the pollutant concentration at monitoring site s during sampling session j, $[Ref_{mean}]_j$ is the mean of regional background and urban reference site concentration during sampling session j, $[Ref_{mean}]_k$ is the seasonal arithmetic average of the mean regional background/urban reference session values (n=6). 2) Temporal variability was accounted for in LUR models using the session-specific regional background measurement ($[Ref_{mean}]_j$ from eq. 2.1) as the first independent term. 3) Two terms from each source category were retained (if applicable) for linear regression given the strength of bivariate correlations with temporally-adjusted NO_2 (maximum p-values of 0.05) (Shmool et al., 2014). 4) Regression models were initially fit using forward stepwise selection and verified with backward stepwise selection to assess overall model improvement at each stage, using the coefficient of determination (R^2), and removing non-significant ($p > 0.05$) covariates in order of descending p-value. 5) Given the high potential for collinearity, covariates were removed if variance inflation factors (VIF) were greater than 2 and further sensitivity tests were performed including; 6) random forest decision trees and forward stepwise addition based on buffer size (largest to smallest and vice versa). LUR Model building was performed in STATA/SE 13.0 (StataCorp. 2013).

To evaluate the utility of Caline3 within a LUR framework, we first built standard yearly and combined years LUR models without Caline3 following the general form in Equation 2.2:

$$\boldsymbol{C_S = \beta_0 + \beta_1 TEMP_t + \sum_{i=1}^{m}(\beta_i\, x_{i,s}) + \varepsilon_s}$$

(2.2)

where C_s is the measured concentration of NO_2 at location s (µg/m^3), β_0 is the intercept (µg/m^3), $\beta_1 TEMP_t$ is the mean concentration of regional background and urban reference for session j, β_i is the regression coefficient of the i^{th} spatial variable (Table 1) in appropriate units, $x_{i,s}$ is the value of the i^{th} spatial variable at location s, m is the number of spatial covariate classes (Table 1) and $\mathcal{E}_s$ is the model prediction error at location s.

Weekday and full-week LUR models were built independently to allow for comparisons given varying weekend diurnal traffic patterns, and to better assess the contribution of Caline3 which includes both spatial and temporal information. Finally, LUR and subsequent LUR/ Caline3 models were constructed utilizing all 72 NO_2 measurements, hereafter referred to as the merged years model. This merged model increased model power and tested Caline3's effectiveness when combining temporally misaligned measurement data. Repeated measures were treated as random effects by including random intercepts for year sampled in a two-level mixed model with restricted maximum likelihood and an independent covariance structure.

Table 1. GIS-based spatial covariates at various buffer distances for LUR modeling building

Source category for LUR Modeling	**Covariates examined within (50, 100, 200, 300, 500, 750, 1000 m)**
Traffic density indicators	Mean density traffic (primary roads) Mean density traffic (primary and secondary roads) Number of signaled intersections
Road-specific measures	Average daily traffic on nearest primary road [a] Distance to nearest major road [a] Distance to roadways stratified by standard deviations greater than mean (e.g., urban, arterial, saturated) Summed length of primary roadways Summed length of primary and secondary roadways
Truck, Bus, and Diesel	Mean density of bus traffic Distance to nearest bus route [a] Outbound and inbound trip frequency per week summed by route Mean density of heavy truck traffic on nearest primary roadway
Population	Census population density
Land Use / Built Environment	Total area of industrial parcels Total area of industrial and commercial parcels
Industrial emissions	Distance to nearest industrial stationary source Summed density of total TRI pounds emitted per meter Summed density of total NEI pounds of $PM_{2.5}$, SO_2, NOx, and VOCs emitted per meter Summed density of total $PM_{2.5}$ emitted per meter Summed density of total SO_2 emitted per meter Summed density of total NOx emitted per meter Summed density of total VOCs emitted per metes
Transportation Facilities	Distance to nearest active railroad [a] Summed line length of active railroads Distance to nearest bus depot [a]

Table 1. cont.

Potential Modifying Factors	
Topography	Average elevation Elevation at receptor
Meteorology	Temperature/Relative humidity [a,b] Frequency of inversions [a] Wind direction and wind speed [a]

[a] area buffer not applicable
[b] temperature and humidity were collected on-site

2.2.7 Hybrid LUR/ Caline3 model framework

Modeled concentration predictions from Caline3 were incorporated as an independent covariate in LUR models for NO_2. Figure 4 provides a conceptualization of integrating meteorological and traffic volume information into LUR via Caline3, resulting in a hybrid LUR modeling framework.

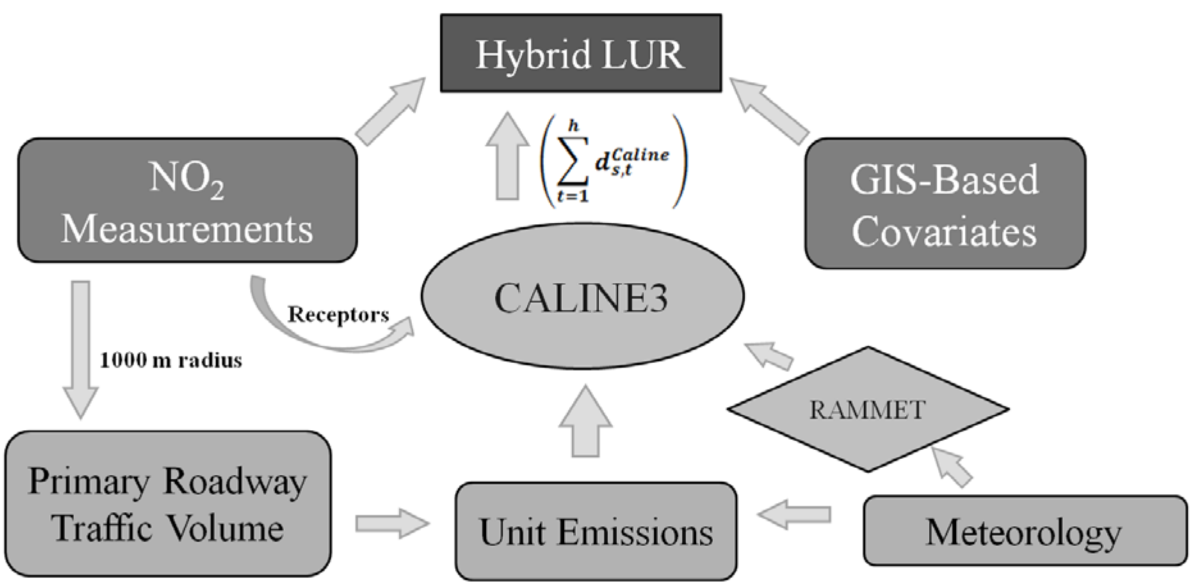

Figure 4. Conceptual framework for incorporating traffic-related emissions and meteorology information into Caline3 preceding addition to the land use regression model

To incorporate Caline3 information into LUR, session-specific Caline3 model predictions were added as an independent covariate to equation 2.2 and incorporated as shown in equation 2.3:

$$\boldsymbol{C_s = \beta_0 + \beta_1 TEMP_j + \sum_{i=1}^{m}(\beta_i\, x_{i,s}) + \alpha_{Caline}\left(\sum_{t=1}^{h} d_{s,t}^{Caline}\right) + \varepsilon_s}$$

(2.3)

Where; $\boldsymbol{\alpha_{Caline}}$ = regression coefficient for the Caline3 covariate

$\boldsymbol{d_{s,t}^{Caline}}$ = dispersion concentration (μg/m^3) predictions from Caline3 Gaussian dispersion model for site *s* for hour *t*

2.2.8 Model Performance Statistics

Model performance was evaluated by coefficient of determination (R^2), given by the equation 3:

$$\boldsymbol{R^2 = 1 - \frac{\sum_{i=1}^{n}(x_i - \hat{x}_i)^2}{\sum_{i=1}^{n}(x_i - \bar{x}_i)^2}}$$

(2.4)

Where; *n* is the number of data points, x_i are the measured values, $\hat{x}_i$ are the predicted values, and $\bar{x}_i$ is the mean of the measured values. Root-mean-square-error (RMSE) was also calculated as a measure of model performance, given by the formula:

$$\boldsymbol{RMSE = \sqrt{MSE} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{x}_i - x_i)^2}{n}}}$$

(2.5)

Where; x_i are the measured values, $\hat{x}_i$ are the predicted values. Instead of the RMSE for the merged model, the Akaike information criterion (AIC) was reported given the dependence on the maximum likelihood framework. Finally, standardized beta (β) coefficients were computed by transforming outcome and predictor variables to z-scores prior to regression. Standardized coefficients are measured in standard deviations, as opposed to the respective variable units. This allows for inter-comparison of predictors within each model by providing a relative impact when adding or removing terms.

Cross-validation: All models were evaluated using the leave-one-out cross-validation method where predictions from a regression model were built from n-1 measurement sites. The model estimated using n-1 sites is considered the training set, from which, the predicted value for the test site is obtained. This process is repeated n times, until a prediction value is generated for each site using its respective training set. Cross-validated R^2 (R^2_{CV}) and *RMSE* are computed by regressing the observed measures against the cross-validated predictions using the equations above. In evaluating highly resolved spatio-temporal information from dispersion output, this cross-validation process allows for an assessment of out-of-sample performance, which we are ultimately interested in.

2.3 RESULTS

2.3.1 Summary Statistics

Higher NO_2 concentrations, on average, were observed for weekday-only (year 1) samples and greater variability was observed in full-week (year 2) samples (Table 2). Measurement variability can also be observed between and within sessions as indicated by box-plots in Fig. 5. The 12 repeated sites were well correlated between years (Pearson's r=0.65, p=0.02). On average, higher concentrations were observed at high traffic, high industry, and valley sites. Of the three source indicators originally used for site selection, valley vs. non-valley produced the largest concentration differences, followed by traffic density, and industrial emissions. Moreover, all three source indicators were prominent in LUR models (Tables 4 and 5). Caline3 predictions stratified by low- and high-traffic sites produced means of 1.69 $\mu g/m^3$ (SD = 1.66, n = 37) and 4.48 $\mu g/m^3$ (SD = 3.6, n = 35), respectively. The maximum range in predictions at a repeated site was 4.24 µg/m3 signifying the potential impact of source/meteorological interaction information.

Table 2. Summary statistics of non-adjusted winter NO_2 measurements (PPB)

	Weekday[1]	Full-week[2]	Regional Background[3]	Urban Reference[3]
n	36	36	12	12
Min	8.9	6.4	3.9	11.5
Max	29.8	26.9	10.4	24.1
Mean	17.9	14.7	7.4	18.1
Median	18.4	13.7	7.88	18.5
SD	4.4	4.9	2.2	3.5

We observed consistent and stable covariance between the regional background and the urban reference site measurements in all sampling sessions (Table 2, and Fig. 5). Generally, the urban reference site captured above-mean concentrations during most sessions, while the regional background site recorded the lowest concentration during all sessions producing a mean reference value near the 25th percentile of distributed measures (Fig. 5).

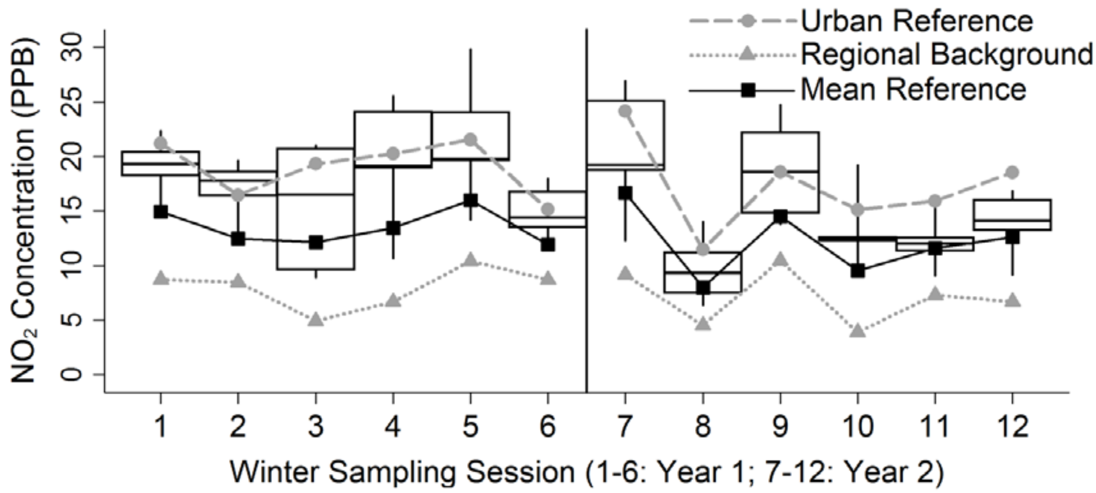

Figure 5. Boxplots of NO_2 measurements from distributed sites with urban reference and regional background continuous sites as plotted lines by session

2.3.2 Summary of Model Performance

Pre-constructed LUR models without Caline3 produced final cross-validated R^2_{CV} values of 0.57, 0.76 and 0.73 (Snijders/Bosker R^2) for weekday, full-week, and merged years, respectively (Table 3). The addition of the Caline3 term improved R^2_{CV} values to 0.67 and 0.79 for both yearly models each doing so with one fewer predictor. The cross-validated R^2 improved to 0.78 for the merged years model following the addition of the Caline3 term (Table 3). Cross-validated RMSE values also demonstrated improvements following the addition of Caline3.

Table 3. Summary LUR and LUR + Caline3 model results. R^2, and RMSE leave-one-out cross-validated

Model	Weekday only - Year 1			Full-week - Year 2			Merged Years		
	n terms	R^2_{CV}	RMSE	*n* terms	R^2_{CV}	RMSE	*n* terms	R^2_{CV}	AIC
LUR	4	0.57	2.51	4	0.76	2.48	5	0.73	379.72
LUR + Caline3	3	0.67	2.31	3	0.79	2.21	5	0.78	362.15

2.3.3 Weekday LUR + Caline3

The pre-constructed weekday (Year 1) LUR model included *distance to nearest industrial source, mean traffic density within 50m radius,* and *average wind speed*. The temporal term explained approximately 22% of NO_2 in-sample variability across sampling sessions. The addition of the *Caline3* term to the pre-constructed model effectively displaced the *mean traffic density (50 m)* (p = 0.28) and *average wind speed* (p = 0.14) terms, while improving overall model fit as per cross-validated R^2 and RMSE (Table 3). Following the addition of Caline3, changes in standardized β coefficients show a decrease in relative strength for all three spatial predictors, with the most significant decrease occurring for the mean traffic density term (Table 4).

Table 4. Weekday LUR (*n* = 36) with addition of Caline3 covariate

Covariates Predicting Weekday NO_2	LUR		LUR + Caline3		
	NO_2 β (*p*-value)	Seq. R^2	NO_2 β (*p*-value)	Seq. R^2	Change in std. β
Intercept	11.31		3.66		--
Mean temporal NO_2	0.99*	0.22	1.08*	0.41	+0.006
Distance to nearest industrial stationary source	-6.0×10^{-4} **	0.49	-5.5×10^{-4}**	0.59	-0.03
Mean traffic density (50m)	0.03**	0.66	NA *(0.31)* Ŧ	--	-0.26
Average wind speed	-1.68*	0.71	NA *(0.14)* Ŧ	--	-0.08
Caline3	--	--	0.84**	0.75	--

Ŧ Covariate removed due to *p* > 0.05
***significant: p <0.05; **significant p <.0001**

2.3.4 Full-week LUR + Caline3

The pre-constructed full-week (7-day) LUR model differed substantially in comparison to the weekday model. The temporal term explained approximately 50% of in-sample variability of NO_2 compared to only 22% in the weekday model. Spatial predictors included *mean elevation within 300m*, *number of traffic-signaled intersections within 750m* and *total area of industrial and commercial land use parcels within 1,000m* (Table 5). Elevation was tested with various interaction terms, but was not significant. Similarly to the weekday model, the *signaled intersections (750m)* (p = 0.11), and *total industrial and commercial parcels (1000 m)* (p = 0.27) terms were displaced by the addition of the Caline3 term in the full-week model. Standardized beta coefficients decreased for the two displaced terms and increased for the temporal and elevation terms. Therefore, after accounting for temporal variability, the 7-day LUR model with only mean elevation within 300m and Caline3 output explained 83% of in in-sample variability in NO_2, with a LOOCV R^2 of 0.79. Thus, Caline3 provided slightly greater model improvement for the weekday-only model compared to the full-week model.

Table 5. Year 2 (full-week) LUR (n=36) with addition of Caline3 output

Covariates Predicting Full-week NO_2	LUR		LUR + Caline3		
	NO_2 β (p-value)	Seq. R^2	NO_2 β (p-value)	Seq. R^2	Change in std. β
Intercept	6.38		8.83		--
Mean temporal NO_2	1.12^{**}	0.50	1.24^{**}	0.50	+0.03
Mean elevation (300m)	-0.03^{*}	0.69	-0.04^{*}	0.69	+0.04
Signaled intersections (750m)	0.18^{*}	0.78	NA *(0.11)* ₮	--	-0.12
Total area of industrial and commercial parcels (1000m)	$2.57\text{x}10^{-7\,*}$	0.82	NA *(0.29)* ₮	--	-0.11
Caline3	--	--	0.53^{**}	0.83	--

₮ Covariate removed due to *p* > 0.05
***significant: p <0.05; **significant p <.0001**

2.3.5 Merged Years LUR + Caline3

The merged years (weekday + full-week) model included all winter-season NO_2 measures and followed identical model building methods to preceding models. Repeated measured were accounted for by a random intercept in a mixed effects modeling structure utilizing restricted maximum likelihood (p < 0.0001). All covariates significant in the weekday-only model were retained in the merged model with the addition of the *mean elevation (300 m)* term (Table 6). Following the addition of the Caline3 term, the *mean traffic density (50 m)* term was displaced. In contrast to the weekday-only model, the *mean wind speed* term remained significant (p = 0.017) following the addition of Caline3. Variance inflation factors were 1.56 and 1.02 for the *mean wind speed* and *Caline3* terms, respectively. The merged model had an intra-class correlation coefficient of 0.41 due to repeated site variation. AIC and cross-validated values are shown in Table 3, and indicated an improved model fit for the model containing Caline3. Similarly to yearly models, Caline3 was effective in improving overall prediction accuracy for a model that combined measurements of varying averaging times.

Table 6. Merged years LUR (n=72) with addition of Caline3

Covariates Predicting Merged Years NO_2	LUR		LUR + Caline3		
	NO_2 β (p-value)	Seq. R^2	NO_2 β (p-value)	Seq. R^2	Change in std. β
Intercept	15.43		15.31		--
Mean reference NO_2	1.01**	0.41	1.04**	0.41	+ 0.01
Distance to nearest industrial stationary source	-4.1 x10^{-4} **	0.59	-3.5x10^{-4}**	0.59	- 0.03
Mean traffic density (50m)	0.03**	0.72	NA *(.15)* Ŧ	--	- 0.20
Elevation (300m)	-0.02*	0.74	-0.02 **	0.64	+ 0.01
Mean wind speed	-1.42*	0.77	-1.39 *	0.66	- 0.007
Caline3	--	--	0.58 **	0.81	--

Ŧ Covariate removed due to *p* > 0.05
***significant: p <0.05; **significant p <.0001**

To examine Caline's effectiveness in capturing spatial variability in model fit in relation to near-source gradients, residuals from pre-constructed LUR and LUR/ Caline3 models were examined as a function of distance to the nearest roadway. Fig. 6 displays the absolute value residual differences from the LUR/Caline3 residual minus the pre-constructed LUR residual, matched by site. Residual value differences in Fig. 4 are dichotomized by high and low traffic sites defined by the 70th percentile of traffic density, originally defined in site selection. In Fig. 6, smaller residuals derived from the LUR/ Caline3 model compared to the LUR model result in greater magnitude differences, and therefore, larger absolute values. Whereas, residuals from each model that were more similar in magnitude, resulted in smaller differences, and therefore, produced smaller absolute values. Thus, the largest differences in modeled residuals occurred at the high traffic sites (> 70th %) and at locations most proximal to primary roadways (Fig. 6), and produced a negligible effect on low traffic sites beyond 300m. Therefore, the marginal improvements observed in model fits, may be decomposed to near-source/high traffic locations.

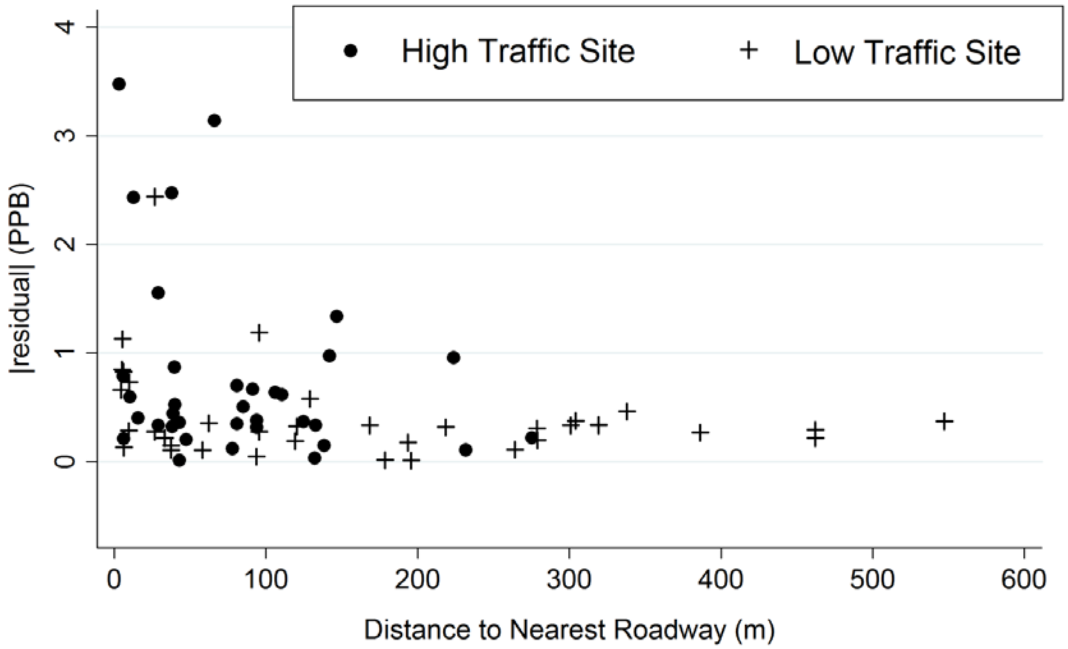

Figure 6. **Absolute value residual differences of combined years LUR vs. LUR/Caline3 model predictions with linear fit and 95% CI as a function of distance to nearest roadway and distinction of traffic density**

2.4 DISCUSSION

Here, we presented a method to incorporate output from a spatio-temporal line source dispersion model into LURs predicting NO_2 across two successive winter seasons, across a large urban-to-suburban area. As expected, Caline3 provided greater model improvement for the weekday-only model as per cross-validated *RMSE* and R^2. Moreover, Caline3 displaced the GIS-based traffic-related term in each model, corroborating the interpretability of each. Perhaps more importantly, we found greater improvements in predictions at higher-concentration locations near roadways, which may have important bearing towards accurately characterizing exposures in near-source locations for epidemiological studies.

Comparability of results to other hybrid models: Wilton et al. (2010) observed similar improvements in model fit with a Caline3/LUR hybrid model for summer-only NO_2 and NO_X, utilizing data from a 2-week snapshot sampling campaign designed to capturing near-road gradients outside of metropolitan areas. Our efforts differed by: (1) measurement sites were allocated systematically across a metropolitan area - not specifically to capture near-road gradients, and (2) we modeled all primary roadways within 1000m of each sampling site in Caline3. Corroboratory Wilton et al. (2010), we observed the greatest degree of model improvement when model output from high-traffic density roadways (i.e., > 100,000 vehicles per day) was included and was proximate to receptor locations (25 – 300m). Lindström et al. (2013) extended the hybrid work presented by (Wilton et al., 2010), but did not observe a similar degree of model improvement within their spatio-temporal modeling framework.

Temporal adjustment in LURs for NO_2: Because our measures were collected over a series of six sampling weeks each season, LUR models required adjustment for temporal variance using reference site data. Further, accurately characterizing temporal variance for reactive pollutants, such as NO_2, remains an important challenge. Given consistent near-zero concentrations at our regional background site, we needed to average this with an urban reference site to provide a useful temporal signal. More variability was explained by the reference term in the full-week model (R^2 = 0.49) than in the weekday-only model (R^2 = 0.20), which may be explained by substantial differences in weekday and weekend traffic, both incorporated in full-week samples, with some variation across weeks in the relative proportion of each (i.e., federal holidays).

Spatial vs. temporal variability in Caline3: Because Caline3 incorporates both spatial and temporal (meteorological) information, it is challenging to assess the relative contribution of each in the hybrid model, and retaining a reference site term from LUR in hybrid models may diminish some of the potential explanatory power of the Caline3 predictions. Lindström et al. (2013) noted that the LUR portion of a hybrid model may serve to over-emphasize the temporal (vs. spatial) contribution from Caline3. This may be a particular concern in our dataset, as our study design maximized our ability to capture spatial variance by cross-stratifying on confounded sources and modifiers (i.e., vehicular traffic, industry, and elevation). Indeed, indicators from each of these three source categories were significant in final LUR models. Finally, the Caline3 term also displaced one industrial term in the full-week model [*industrial and commercial area*], hypothesized, in part, to capture industrial vehicular truck traffic. This result may highlight the utility of source dispersion models to improve upon the physical interpretability of empirical LURs. Nonetheless, novel spatio-temporal modeling frameworks applied by Lindström et al. (2013) and Keller et al. (2014) may help to further disentangle interpretation of spatio-temporal explanatory variables, though application here was beyond the scope of this work.

Caline3 and meteorological data: Caline3 incorporates hourly meteorological data directly into source dispersion estimates, as is not the case for other source terms in LUR, and thus the hybrid likely more accurately captures roadway emissions relative to other sources. Further, the displacement of *mean wind speed* in the weekday-only model may point to this improved temporal information introduced via the Caline3 term, although these two terms were not collinear (*VIF* = 1.13). *Mean wind speed* was retained in the combined years model, however, again not collinear with the Caline3 term. This could be the result of the implicit temporal variability provided by this predictor given the temporal misalignment in combining two separate seasons, albeit controlling for season.

Limitations: Numerous limitations of the Caline3/ LUR framework were addressed in Wilton et al. (2010). The *CalRoads*' particulate matter (PM) pollutant designation option more appropriately estimated total NO_X (NO + NO_2). Ideally, to best capture the influence from combustion sources such as motorized traffic, NO should also be measured along with NO_2.

However, high correlations between NO_2 and NO_X have been reported in prior near-road studies (Karner et al., 2010; Su et al., 2009; Wang et al., 2011). All meteorological data (except temperature and humidity) were obtained from the National Weather Station at the Pittsburgh International Airport, approximately 20 miles west of our modeling domain.

Strengths and Implications: Incorporating Caline3 output into LUR displaced GIS-based traffic covariates in two separate models, and improved overall cross-validated model performance while corroborating model interpretability. The greatest degree of model improvement was observed with weekday-only measures, at high traffic density sites, and at locations closest to primary roadways (<300m), indicating the utility of our hybrid approach towards better capturing pertinent source intensity exposures for epidemiological applications. Finally, because Caline3 accounts for hourly meteorological variability and source-meteorology interactions, the hybrid approach may substantially improve interpretability of source terms, and ultimately may prove more reliable for model extrapolation.

2.5 SUMMARY

The model framework described in chapter 2 helped to explain an additional portion of variation in NO_2 observations than a standard LUR model, especially proximal to roadways. Differential variability explanation near sources was a hypothesized result in incorporating source/ meteorological interaction information in LUR via atmospheric dispersion principles. Moreover, given the sharp concentration decay gradients of NO_2 as a function of distance from roadways, a spatiotemporally-varying explanatory variable from deterministic dispersion information can benefit intra-urban pollutant variability studies over short temporal scales (e.g., quarterly, seasonally, daily). Ambient $PM_{2.5}$, however, tends to vary more so at a regional scale as opposed to the local-type scale of NO_X, though fine PM has been associated with a much larger wealth of adverse human health outcomes, usually derived through population-level epidemiological studies. The number of oxides of nitrogen LUR models greatly outnumbers $PM_{2.5}$ models considering low-cost passive NO_X samplers vs. more intensive monitoring efforts required for $PM_{2.5}$. In Chapter 3, we apply the same hybrid modeling framework; however, the pollutant of interest is $PM_{2.5}$, and the sources of interest are industrial stationary sources as opposed to traffic-related sources. We modeled all $PM_{2.5}$ sources across the Greater Pittsburgh, PA Region with the AERMOD Gaussian plume modeling system and similarly examined the utility of AERMOD predictions with LUR for estimating $PM_{2.5}$. In contrast to the Caline3 model, AERMOD incorporates planetary boundary layer turbulence and scaling algorithms for predicting dispersion from stationary sources in both simple and complex terrain environments.

3.0 HYBRID AERMOD/ LUR MODEL FOR PREDICTING $PM_{2.5}$

Land use regression (LUR) is a standard method used to explain the spatial distribution of ambient air pollution for use in epidemiological studies (Brauer et al., 2003; Briggs et al., 1997; Clougherty et al., 2013b; Jerrett et al., 2005). LUR for exposure assessment, however, can be constrained by the spatial variability expressed by the pertinent geographic predictors in relation to the locations of the monitoring sites, and the true underlying pollutant variability (Alexeeff et al., 2014; Basagaña et al., 2013). Therefore, there is growing interest in incorporating spatio-temporally varying geographic covariates in LUR, such as Gaussian dispersion output, in the hopes of better simulating pollutant variability while improving accuracy, interpretability, and transferability of such models.

Empirically-based LUR models employ relatively simple inputs and provide significantly higher spatial resolution than proximity-based, or purely statistical interpolation methods (Jerrett et al., 2005). The LUR process combines a relatively large number of systematically distributed air pollution measures with “land use” variables usually managed in GIS. Variables used to explain intra-urban $PM_{2.5}$ variability have included surrogates for automobile traffic emissions, population density, household density, industrial and commercial land use, land cover and open space, elevation and primary $PM_{2.5}$ emissions density (Hoek et al., 2008). Geographic variables are generally measured as nearest distances from sources or as densities within circular areas. These Euclidean metrics and isotropic areal units fail to capture small-scale spatiotemporal pollutant variability, governed, in part, by interactions between emissions and meteorological processes (e.g., upwind vs. downwind advection) (Jerrett et al., 2005; Su et al., 2008; Wilton, 2011).

Prior LURs have been attempted to incorporate some measure of temporal variance into source-concentration relationships by including meteorological covariates (e.g., mean wind speed or direction) (Arain et al., 2007; Clougherty et al., 2009; Jerrett et al., 2007; Su et al., 2008), or by weighting source-concentration relationships by predominant wind direction (Clougherty et al., 2009; Mavko et al., 2008; Van den Hooven et al., 2012). Vienneau et al. (2009) originally presented a GIS-based method using distance weighted emissions and monitoring data that was improved by Gulliver and Briggs (2011) through the incorporation of meteorological dispersion principles enabling daily and annual PM_{10} predictions at $1km^2$ resolution. Ainslie et al. (2008) and Su et al. (2008) attempted to capture atmospheric dispersion using a source-area grid of distributed emissions under varying atmospheric conditions. Likewise, Wilton (2011) incorporated meteorologically-varying covariates as volume sources derived from the CALPUFF Lagrangian puff model. To our knowledge, only two hybrid line-(traffic) source dispersion/LUR modeling efforts have been attempted with each reporting variable model improvement, albeit more parsimonious and interpretable models (Lindström et al., 2013; Wilton et al., 2010).

To further refine small-scale (e.g., intra-urban) spatial concentration gradients, techniques to combine spatially-scalable models to better capture near-source variability have been employed (e.g., localized traffic demand modeling for emissions factor estimation) (Cook et al., 2008; Isakov et al., 2007; Kinnee et al., 2004). Isakov et al. (2009) combined a regional background model (CMAQ) capable of photochemical reactions with more localized predictions from AERMOD to produce hourly air pollutant predictions at block-group resolution. Other hybrid approaches have utilized dispersion output as the dependent variable to develop LUR models with refined spatial (Isakov et al., 2009; Johnson et al., 2010) and spatio-temporal (Johnson et al., 2010; Mölter et al., 2010a) estimates for NO_2 and PM_{10}. Recently, Dionisio et al. (2013) demonstrated refined spatial and temporal estimates of multiple pollutants using AERMOD predictions to disentangle regional background and localized spatio-temporal variability. In a complementary study, Sarnat et al. (2013) observed stronger heath effect estimate associations with the spatially-refined exposure metrics compared to a central site exposure scenario.

Atmospheric dispersion models have been employed extensively in regulatory air quality management but only more recently for exposure assessments (Jerrett et al., 2005; Johnson et al., 2010; Marshall et al., 2008; Mölter et al., 2010b; Van den Hooven et al., 2012). Dispersion models simulate transport of pollutants, as a function of source characteristics and temporally-varying meteorological conditions (Briant et al., 2013; Chang and Hanna, 2004). In comparison with LUR approaches that can provide detailed spatial resolution, dispersion modeling offers high temporal variability with theoretically unlimited spatial resolution. Furthermore, it has also been demonstrated that LUR-derived exposure misclassification may depend more so on how much of the true spatial variability is explained by the geographic covariates in the exposure model, and not necessarily the accuracy of the predictions (Alexeeff et al., 2014; Szpiro et al., 2011a), especially when LUR models are constructed from a small number of measurement sites (Basagaña et al., 2013). Therefore, standard LUR could be improved by incorporating deterministic source-meteorology interaction information, especially in highly industrialized areas. Thus, producing theoretically-physically based estimates, as opposed to purely empirically-derived estimates that rely upon the quantity and quality of measurement data (Jerrett et al., 2005; Su et al., 2008; Wilton et al., 2010).

In this chapter, we incorporate modeled $PM_{2.5}$ predictions with AERMOD into an LUR model for predicting $PM_{2.5}$ in a region of relatively intense industrial-source activity. The study domain covers an urban-to-suburban landscape with varying terrain and many legacy industrial sources situated within river valleys. Our multi-pollutant spatial saturation study was designed to disentangle impacts of multiple pollution sources (e.g., industry, vehicle traffic), and to assess potential modifiers of source-concentration relationships (e.g., elevation) (Shmool et al., 2014). We examined $PM_{2.5}$ measures collected during successive summer and winter sampling campaigns. We evaluated the utility of AERMOD with LUR by adding session-specific AERMOD predictions as an independent covariate to seasonal LUR models and observed changes in modeling diagnostics and accuracy of predictions using cross-validated methods. Additionally, to decompose AERMOD at near-source settings, we focused on area of intense industrial activity within a valley to examine differential prediction accuracy derived from LUR models containing a GIS-based industrial covariate vs. AERMOD predictions at a 100m x 100m grid resolution.

3.1 METHODS

3.1.1 $PM_{2.5}$ Measurements

$PM_{2.5}$ sampling took place from June 5th to July 26th 2012, and was repeated in the winter from January 8th through March 10th 2013. A total of six successive weekly (7-day) sessions of 6-7 distributed sites per session comprised a sampling season. Samplers operated for an integrated 24-hour, 7-day sample of 15 minutes per hour equating to 42 total hours of sampling per session. Further detail is available in Shmool et al. (2014).

Sampling instruments included stainless-steel Harvard Impactors (Air Diagnostics and Engineering Inc.) with 37mm Teflon filters and a data logger (HOBO - Onset Computer Corporation), which were contained in waterproof Pelican cases. Sampling units were customdesigned to capture integrated street-level (~3m height above ground) measurements of $PM_{2.5}$ (Clougherty et al., 2013a). Instruments were programmed to sample during the first 15 minutes of each hour using a chrontroller interface (ChronTrol Corporation). A tetraCal volumetric air flow calibrator (BGI Instruments) was used to calibrate intake flow to approximately 4.0 LPM. Concurrently, an on board HOBO data logger recorded temperature and relative humidity at fifteen minute intervals. Prior to field deployment, 37mm Teflon filters (Pall Life Sciences) were equilibrated for 48 hours and then pre-weighed using an ultramicrobalance (Mettler Toledo Model XP2U) using a temperature (20°C) and relative humidity (35%) controlled glove box (PlasLabs Model 890 THC). Filters were post-weighed under identical conditions and concentrations were derived from time-integrated mass calculations.

3.1.2 Study Domain and Site Selection

Our study domain encompassed a contiguous 500 km^2 area containing the Pittsburgh metropolitan area and key local industrial sources, demarcated at census administrative boundaries to enable merging with socioeconomic and health data in future epidemiological applications (Fig. 7). We used a GIS to systematically allocate monitoring locations cross-stratified across important local pollution sources (e.g., traffic, steel manufacturing) and potential topographic modifiers of source-concentration interactions (e.g., elevation) using ArcMap 10.0-10.3 (ESRI, Redlands, CA, USA) and Geospatial Modeling Environment, V. 0.7.2 (Spatial Ecology, LLC).

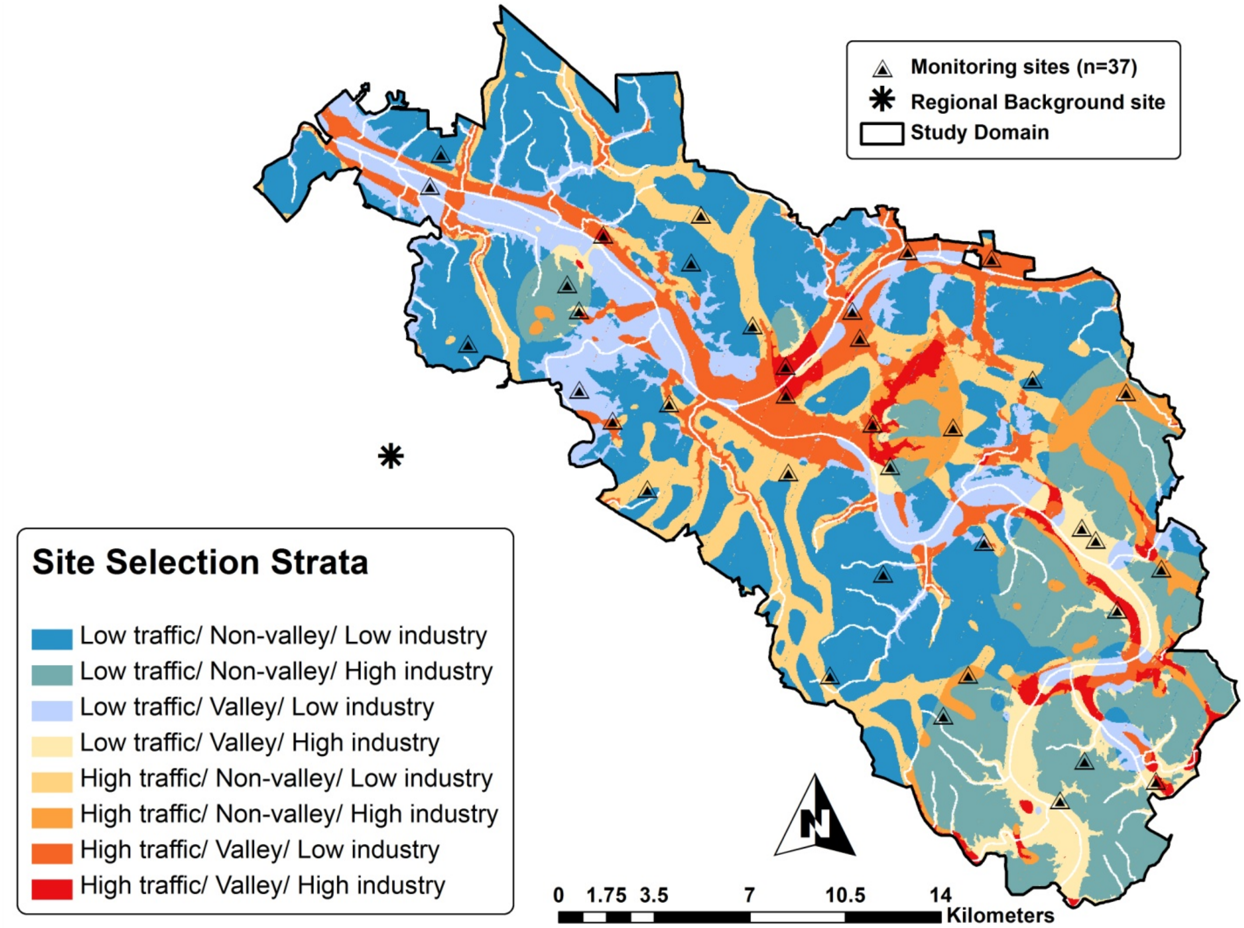

Figure 7. Study domain of Greater Pittsburgh Metropolitan Area with monitoring locations, temporal background reference site location and stratified sampling classifications

Specifically, we anticipated variance in the local pollutant regime to be characterized by: (1) traffic density, (2) industrial density (weighted emissions: $PM_{2.5}$ + NO_X + SO_2 + VOCs), and (3) elevation at 30 m^2 grid resolution. We used stratified random sampling to select monitoring locations representing all possible combinations of high and low source intensities. Site selection and GIS-based covariate calculations are detailed elsewhere (Shmool et al., 2014). Notably, the industry density metric used for site allocation originated from a simple inverse distance weighted (IDW) interpolation of multiple pollutants $PM_{2.5}$ (filterable and condensable), nitrogen oxides (NO_X), sulfur dioxide (SO_2), and volatile organic compounds (VOCs) – from reporting facilities in Allegheny County, PA. We then used inverse-distance interpolation to calculate an emission weighted proximity to industry indicator for each 100 m^2 grid cell centroid, drawing emissions information from facilities within an 80 km radial buffer threshold. The dichotomization for high vs. low industrial source density was chosen at the 70th percentile, given the left-skewed distribution and goal of over-sampling hypothesized high-pollution areas (Shmool et al., 2014). To minimize temporal confounding across sessions, sites were systematically allocated across sessions to balance emissions-indicator strata and spatial coverage. Integrated $PM_{2.5}$ samples were collected across six successive sampling sessions with six randomly-selected sites per session, resulting in a total of 36 measurements per season. Thus, two seasonal sampling campaigns covered 36 unique sites, resulting in 72 total $PM_{2.5}$ measurements.

3.1.3 Temporal Reference

A continuous reference site was monitored each weekly session to adjust for temporal variability in pollutant measures and to limit spatio-temporal bias in comparing measures across sessions (Brauer et al., 2003; Henderson et al., 2007; Hoek et al., 2008). A ‘regional background’ site was selected in a county park (Settler’s Cabin Park) upwind from the study area and away from local sources, about 4.0 km west of the study domain (Fig. 7). The site was categorized in the hypothesized lowest-concentration source strata (low industry, low traffic, high elevation). From pilot sampling, we found that the temporal reference adjustment method influenced observed source-concentration relationships, and the regional background site alone was appropriate for temporally adjusting $PM_{2.5}$ (Shmool et al., 2014).

3.1.4 AERMOD – Gaussian Plume Air Dispersion Model

AERMOD is a steady-state Gaussian plume atmospheric dispersion model that was co-developed by the American Meteorological Society and EPA (Cimorelli et al., 2005). Model development began in 1991 and was designed to capture near-source concentration gradients (<50km) by incorporating planetary boundary layer concepts. As of December, 9, 2006, AERMOD was fully promulgated within the Guideline on Air Quality Models for regulatory application of air quality models for assessing criteria pollutants under the clean air act (U.S.E.P.A., 2005). Treatment of simple and complex terrain is incorporated following the concept of dividing streamline (Snyder et al., 1985) from surface and elevated point, area and volume sources.

3.1.4.1 AERMET – Meteorological Preprocessing

Three separate meteorological datasets were utilized as inputs for AERMET preprocessing and were obtained from the National Oceanic and Atmospheric Administration's (NOAA) National Climate Data Center (NCDC): (1) sequential hourly integrated surface data (ISHD) format[1]; (2) automated surface observation systems (ASOS) 1-minute format[2]; and (3) upper air radiosonde data managed by Earth System Research Laboratory (ESRL)[3]. Surface data selected was utilized from two National Weather Stations located at local airports within the Greater Pittsburgh Area. Both stations recorded ASOS 1-minute wind data via Ice Free Wind sonic

[1] ftp://ftp.ncdc.noaa.gov/pub/data/noaa/

[2] ftp://ftp.ncdc.noaa.gov/pub/data/asos-onemin/

[3] http://www.esrl.noaa.gov/raobs/

anemometers and was preprocessed with AERMINUTE allowing for wind speeds truncation and nonrandomized wind directions. Surface and upper air meteorological data were combined with land cover data (USGS NLCD92 – 30m^2) in AERSURFACE to obtain surface parameters for albedo, Bowen ratio and surface roughness length. Maximum sectors were selected and surface characteristics were derived for the respective summer and winter modeled runs.

3.1.4.2 $PM_{2.5}$ Source Categories

AERMOD requires a detailed emissions inventory profile to model the pollutant or chemical of concern. Information on stack parameters for point sources included ground level elevation, height above ground level, stack exit velocity, stack exit temperature, stack diameter, and $PM_{2.5}$ emissions in g/s. Where applicable, coordinates of the specific stack release points within a facility's grounds were included. Area and volume sources included all of the above parameters in addition to physical dimensions of the emissions surface (e.g., fugitive emissions from an open conveyer). A partial source input file for major sources of $PM_{2.5}$ primary emissions was obtained from the Allegheny County Health Department (ACHD) Air Quality/ Pollution Control Program Division. Minor source stack parameters for additional sources within 100km of the sampling domain were obtained through subsequent ACHD permit applications which included AERMOD input data from Class I and Class II modeling analyses. Emissions rates were obtained from 2011-2012 ACHD emissions inventories and were converted to g/s, resulting in a total of 207 individual point, volume, and areas sources as shown in Fig. 8.

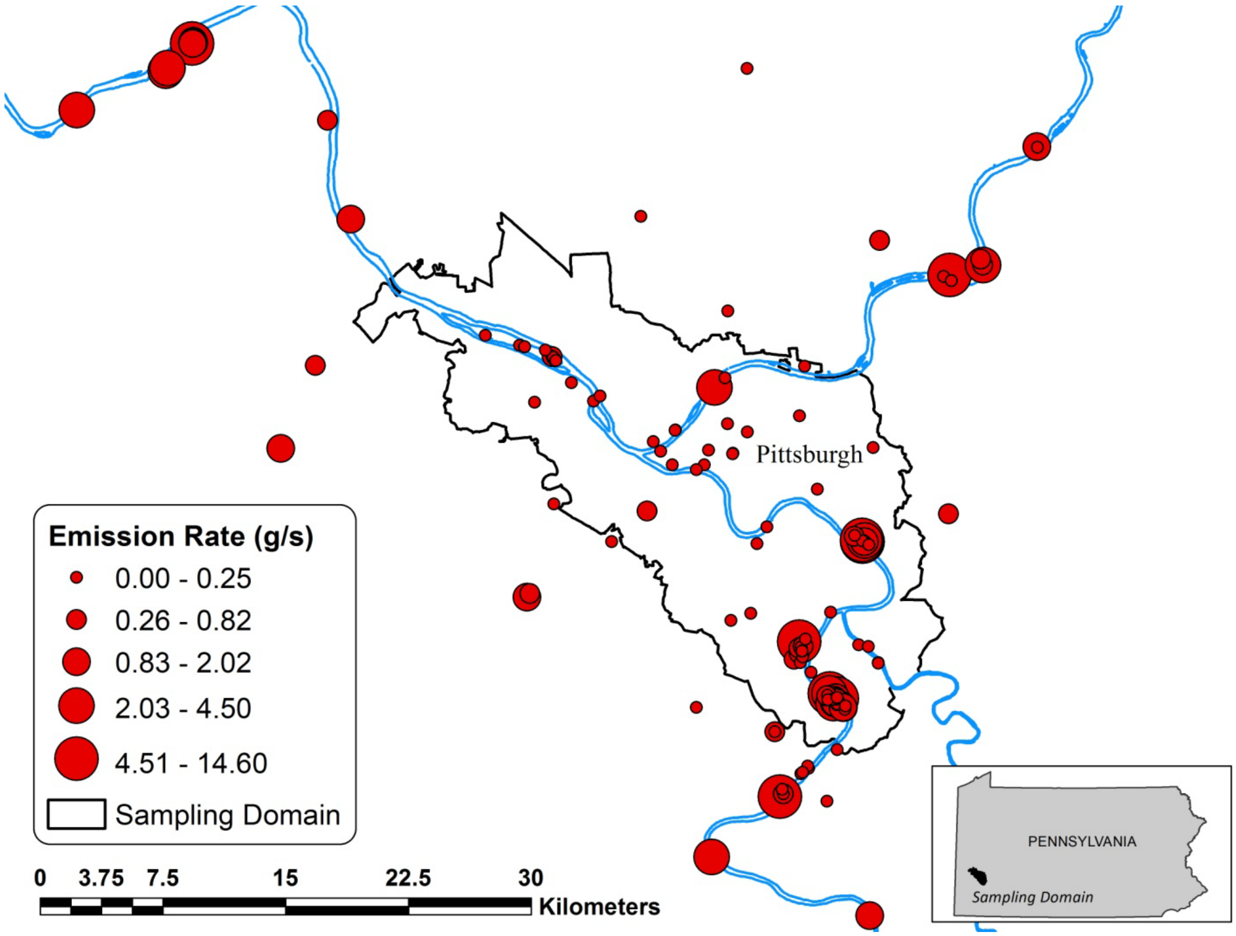

Figure 8. AERMOD modeled stationary $PM_{2.5}$ emissions sources (2011-2012) symbolized by emission rate surrounding sampling domain within Pittsburgh, PA

3.1.4.3 AERMOD Predictions as Geographic Covariate Predictor

To produce an independent covariate in seasonal LUR models, model receptor locations were defined at the monitoring locations (Fig. 7). To account for complex terrain (e.g., river valleys) effects, a $1km^2$ uniform Cartesian receptor grid was included in addition to discrete receptors in all model runs. To coincide with sampling sessions timeframes (7-day week), we produced mean AERMOD predictions utilizing the meteorological data corresponding to the respective weekly sampling session. To examine the spatio-temporal sensitivity of AERMOD predictions within LUR, we also modeled seasonal (corresponding to total sampling time across six sessions), and annual averaging times at each sampling receptor.

3.1.5 LUR Model Building

Separate summer and winter LUR models were pre-constructed without AERMOD to test the marginal benefit of incorporating dispersion into an LUR modeling context, as a supplemental addition may be most applicable elsewhere. The following model-building approach, similar to Clougherty et al. (2013b) was used: 1) candidate indicators were grouped by source category (e.g., traffic indicators, meteorology, industrial emissions) and ranked according to the nonparametric bivariate correlations (Spearman correlations, $p < 0.1$) with temporally-adjusted NO_2 concentrations (Shmool et al., 2014). Sampled pollutant concentrations were temporally adjusted by:

$$\boldsymbol{adjConc_{sj} = \frac{Conc_{sj}}{\left[Ref_{regional}\right]_j} * \left[Ref_{regional}\right]_k}$$

(eq. 3.1)

Where, $adjConc_{sj}$ is the temporally-adjusted pollutant concentration at monitoring site s during sampling session j, $Conc_{sj}$ is the pollutant concentration at monitoring site s during sampling session j, $[Ref_{regional}]_j$ is the regional background reference site concentration during sampling session j, $[Ref_{regional}]_k$ is the seasonal arithmetic average of the regional background site concentration (n=6). 2) Temporal variability was accounted for in LUR models using the session-specific regional background measurement ($[Ref_{regional}]_j$ from eq. 2.1) as the first independent term. 3) Two terms from each source category were retained (if applicable) for linear regression given the strength of univariate correlations with temporally-adjusted $PM_{2.5}$ (maximum p-values of 0.05) (Shmool et al., 2014). 4) Regression models were initially fit using forward stepwise selection and verified with automated backward stepwise selection to assess overall model improvement at each stage, using the coefficient of determination (R^2), and removing non-significant ($p > 0.05$) covariates in order of descending p-value. 5) Given the high potential for collinearity, covariates were removed if variance inflation factors (VIF) were greater than 2 and further sensitivity tests were performed including; 6) random forest decision trees and forward stepwise addition based on buffer size (largest to smallest and vice versa). LUR Model building was performed in STATA/SE 13.0 (StataCorp. LP, College Station, TX, 2013).

LUR seasonal models followed the general form:

$$\boldsymbol{C_s = \beta_0 + \beta_1 TEMP_j + \sum_{i=1}^{m} (\beta_i \, x_{i,s}) + \varepsilon_s}$$

(eq. 3.2)

Where, C_s is the measured concentration of $PM_{2.5}$ at location s (µg/m^3), β_0 is the intercept (µg/m^3), $\beta_1 TEMP_t$ is regional background concentration from session j, β_i is the regression coefficient of the i^{th} spatial variable in appropriate units, $x_{i,s}$ is the value of the i^{th} spatial variable at location s, m is the number of spatial covariate classes and $\mathcal{E}_s$ is the model prediction error at location s.

Spatial autocorrelation across the residuals of the distributed sites was determined using Moran’s I, and spatial correlations were evaluated using generalized additive models (GAMs). Sensitivity to covariate selection was assessed using different temporal adjustment methods including LUR models constructed from temporally adjusted $PM_{2.5}$ concentrations to assess associated spatial variability explained by significant covariates.

3.1.6 HYBRID LUR/ AERMOD MODEL FRAMEWORK

Modeled concentration predictions from AERMOD were incorporated as an independent covariate in LUR models for $PM_{2.5}$. Figure 9 provides a conceptualization of integrating meteorological data, $PM_{2.5}$ source emissions, and terrain information into LUR via AERMOD, resulting in a hybrid modeling framework.

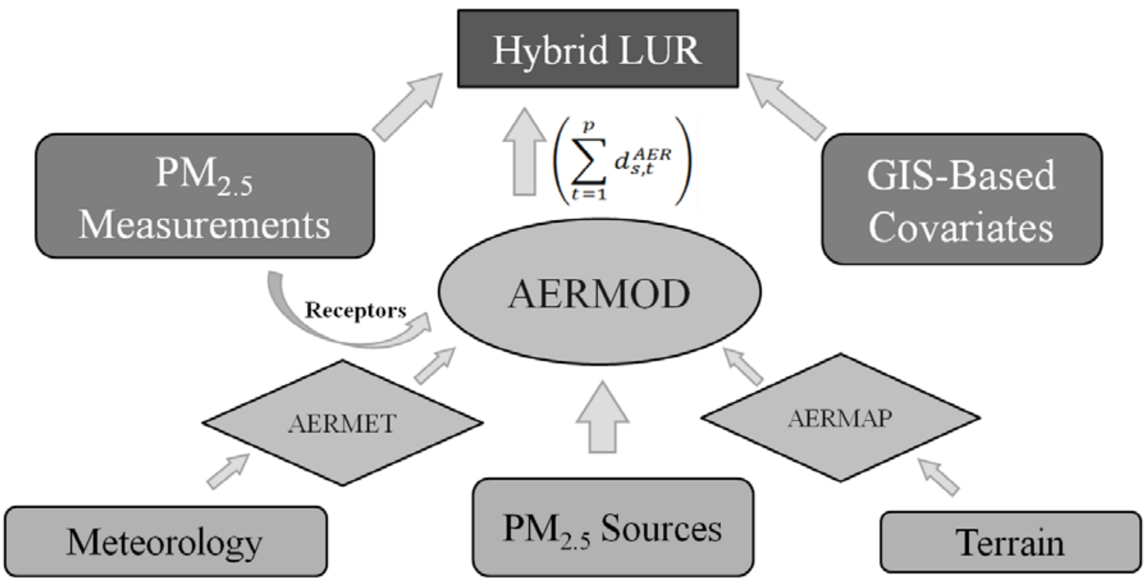

Figure 9. Conceptual framework for incorporating stationary PM emissions, meteorology and terrain information into AERMOD preceding addition to the land use regression model

To incorporate AERMOD information into LUR, session-specific AERMOD model predictions were added as an independent covariate to equation 3.1 and incorporated as shown in equation 3.2:

$$\boldsymbol{C_s = \beta_0 + \beta_1 TEMP_j + \sum_{i=1}^{m}(\beta_i x_{i,s}) + \alpha_{AER}\left(\sum_{t=1}^{h} d_{s,t}^{AER}\right) + \varepsilon_s}$$

(3.2)

Where,

$\boldsymbol{\alpha_{AER}}$ = regression coefficient for the AERMOD covariate

$\boldsymbol{d_{s,t}^{AER}}$ = dispersion concentration (μg/m^3) modeled from AERMOD for site s for hour t

Since C_s is measured in only select locations, the LUR model, based on the resolved subset of potential predictors is used to predict $\hat{C}_s$, the predicted concentration at non-sampled locations within the modeling domain.

3.1.7 Model Performance Statistics

Models were evaluated using the coefficient of determination (R^2), given by the equation 3.4:

$$\boldsymbol{R^2 = 1 - \frac{\sum_{i=1}^{n}(x_i - \hat{x}_i)^2}{\sum_{i=1}^{n}(x_i - \bar{x}_i)^2}} \tag{3.4}$$

Where, n is the number of data points, x_i are the measured values, $\hat{x}_i$ are the predicted values, and $\bar{x}_i$ is the mean of the measured values. Root-mean-square-error (RMSE) was also calculated as a measure of model performance, given by the formula:

$$\boldsymbol{RMSE = \sqrt{MSE} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{x}_i - x_i)^2}{n}}} \tag{3.5}$$

Where, x_i are the measured values, $\hat{x}_i$ are the predicted values. Finally, standardized beta (β) coefficients were computed by transforming outcome and predictor variables to z-scores prior to regression. Standardized coefficients are measured in standard deviations, as opposed to the respective variable units. This allows for inter-comparison of predictors within each model by providing a relative impact when adding or removing terms.

Cross-validation: All models were evaluated using the leave-one-out cross-validation method where predictions from a regression model were built from *n*-1 measurement sites. The model estimated using *n*-1 sites is considered the training set, from which, the predicted value for the test site is obtained. This process is repeated *n* times, until a prediction value is generated for each site using its respective training set. Cross-validated R^2 (R_{CV}^2) and *RMSE* are computed by regressing the observed measures against the cross-validated predictions using the equations above. In evaluating highly resolved spatio-temporal information from dispersion output, this cross-validation process allows for an assessment of out-of-sample performance, which we are ultimately interested in.

3.2 RESULTS

3.2.1 Summary Statistics

Higher $PM_{2.5}$ concentrations, on average, were observed during the summer (mean = 13.83, SD = 2.80) season compared to winter (mean = 11.18, SD = 3.04). Measurement variability was observed between and within sessions across both seasons as shown by box-plots in Fig. 10 that displays six measurements per session, repeated by season (i.e., session 1 measurements = session 7; session 2 = session 8, etc.). The regional background site consistently recorded the lowest measurements with the exception of one session in each season. Therefore, the session concentrations captured from the regional background site were utilized to control for temporal variability in all LUR models (see eq. 1).

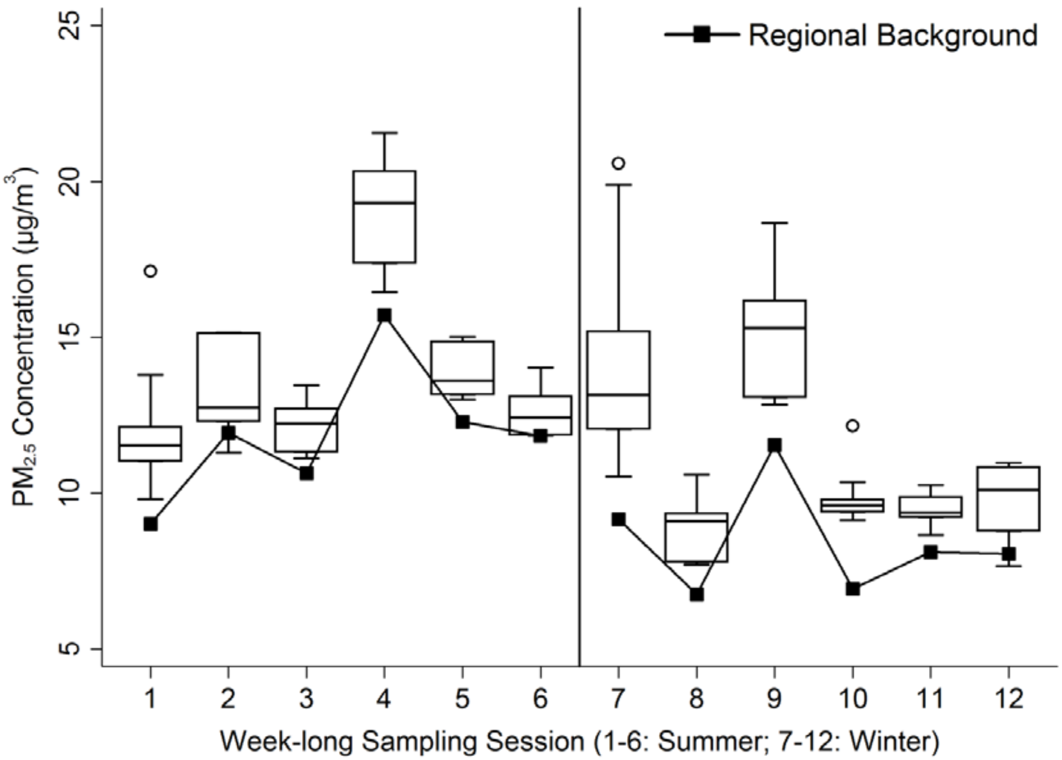

Figure 10. Summer and winter boxplots of $PM_{2.5}$ measurements from distributed sites with linear plot of regional background continuous measures

3.2.2 Summary of Model Performance

LUR models without AERMOD produced final cross-validated R^2 values of 0.73, 0.62 for summer, winter models respectively (Table 7). The summer model explained more variability overall than the winter model with one less covariate. The addition of AERMOD output improved cross-validated R^2 values to 0.82 and 0.75 for each season model, respectively. Cross validated RMSE values also improved across seasons following the addition of AERMOD.

Table 7. Summary LUR and LUR + AERMOD model results with cross-validated R^2 and RMSE values

Model	Summer			Winter		
	n terms	R^2_{CV}	RMSE	n terms	R^2_{CV}	RMSE
LUR	3	0.73	1.15	4	0.62	1.24
LUR + AERMOD	3	0.82	1.09	4	0.75	1.08

3.2.3 Summer LUR + AERMOD for $PM_{2.5}$

LUR modeling results from summer 2012 $PM_{2.5}$ samples are summarized in Table 7. In addition to the temporal term (*Temporal Background $PM_{2.5}$*), the pre-constructed summer LUR model included a kernel density covariate for $PM_{2.5}$ emissions within 50m area (*Density of $PM_{2.5}$ Emissions*) and a modifying binary wind direction term (*Blowing from NW/W*) that produced an overall in-sample R^2 of 0.82. The addition of the AERMOD covariate effectively displaced the $PM_{2.5}$ emissions term (p = 0.69); however, only a slight in-sample improvement in R^2 was observed. Standardized beta coefficients decreased for both spatial and temporal terms following the addition of AERMOD.

Table 8. Summer season standard LUR (n=37) with AERMOD predictions added as an independent covariate with sequential R^2 and change in standardized beta values

Covariates Predicting Summer (June – Aug) $PM_{2.5}$	LUR		LUR + AERMOD		
	$PM_{2.5}$ β (p-value)	*Seq.* R^2	$PM_{2.5}$ β (p-value)	*Seq.* R^2	Δ in std. β
Intercept	1.14		3.31		
Temporal background $PM_{2.5}$	1.17 **	0.62	1.02 *	0.62	-0.06
Density $PM_{2.5}$ emissions (50m)	1.90 **	0.74	NA (0.69) Ŧ	--	-0.15
Wind direction (binary)	--	--	--	--	--
Blowing from NW/W	-1.49 *	--	-1.96 **	--	-0.05
Blowing from SW/W	--	0.82	--	0.70	--
AERMOD	--	--	0.77 **	0.83	NA

Ŧ Covariate removed due to $p > 0.05$
***significant: p <0.05; **significant p <.0001**

3.2.4 Winter LUR + AERMOD

Table 8 summarizes LUR modeling results from winter 2013 $PM_{2.5}$ samples. In comparison to the summer LUR model, slightly less in-sample variability was explained by the temporal term in the winter pre-constructed model (R^2 = 0.54 vs. 0.62 for summer). The winter model similarly included the *$PM_{2.5}$ emissions density* term in addition to the *number of traffic signaled-intersections* and *industrial parcel area* both within 750m^2 buffer areas. The standard LUR model produced an in-sample R^2 value of 0.80 and RMSE of 1.42, respectively. Similarly, the addition of the AERMOD term displaced the static $PM_{2.5}$ density covariate (p = 0.75) in the winter model and resulted in moderate in-sample statistical improvement (R^2 = 0.85). Likewise, standardized beta coefficients decreased for all terms following the addition of AERMOD.

Table 9. Winter-season standard LUR (n=37) with AERMOD predictions added as an independent covariate with sequential R^2 and change in standardized beta values

Covariates Predicting Winter (Jan-March) $PM_{2.5}$	LUR		LUR + AERMOD		
	$PM_{2.5}$ β (p-value)	*Seq. R^2*	$PM_{2.5}$ β (p-value)	*Seq. R^2*	Δ in std. β
Intercept	-1.47		-1.32		
Temporal background $PM_{2.5}$	1.27 *	0.54	1.20 *	0.54	-0.02
Traffic signals (750m)	0.13 **	0.63	0.13 **	0.63	-0.004
Industrial parcel area (750m)	-5.8×10^{-6} *	0.77	5.0×10^{-6} *	0.77	-0.04
Density of $PM_{2.5}$ emissions	1.36 *	0.80	NA (0.75) Ŧ	--	-0.19
AERMOD	--	--	0.79 *	0.85	--

Ŧ Covariate removed due to $p > 0.05$
***significant: p <0.05; **significant p <.0001**

3.2.5 $PM_{2.5}$ Emissions Density vs. AERMOD at Near-source Gradients

To decompose AERMOD information within LUR, we focused our modeling efforts on an area of relatively intense industrial activity to specifically examine source-proximal differential concentration predictions derived from an isotropic industrial covariate (kernel density of $PM_{2.5}$ emissions within $50m^2$ radial distance) vs. AERMOD predictions at 100m x 100m grid resolution. Fig. 11 displays the spatial pattern of the mean *$PM_{2.5}$ emissions density within 50m* covariate (the smallest buffer distance tested) in the immediate area surrounding the United States Steel Clairton Coke Works Facility in Clairton, PA containing 129 point, area, and volume sources obtained from EPA's NEI, 2011. The 'sampling site' depicted in Fig. 11, was one of the 36 randomly selected distributed monitoring locations. The simple density surface in Fig. 11 was created using inverse distance weighted (IDW) interpolation of $PM_{2.5}$ emissions sources from the EPA's NEI 2011, followed by 'extract values to points' and 'spatial join' manipulations to obtain estimated mean tons emitted within varying radial distances surrounding respective sampling locations. The spatial pattern depicted in Fig. 11 highlights one of the intrinsic limitations of isotropic geographic predictors within LUR; where, low spatial variability is expressed and distributions fail to represent predominant upwind vs. downwind pollutant tendencies as indicated by the wind rose in Fig. 13. The frequency histogram in Fig. 12 further exhibits the limited spatial variance expressed across the distribution; however, this term was significant in both seasonal models following covariate selection processes.

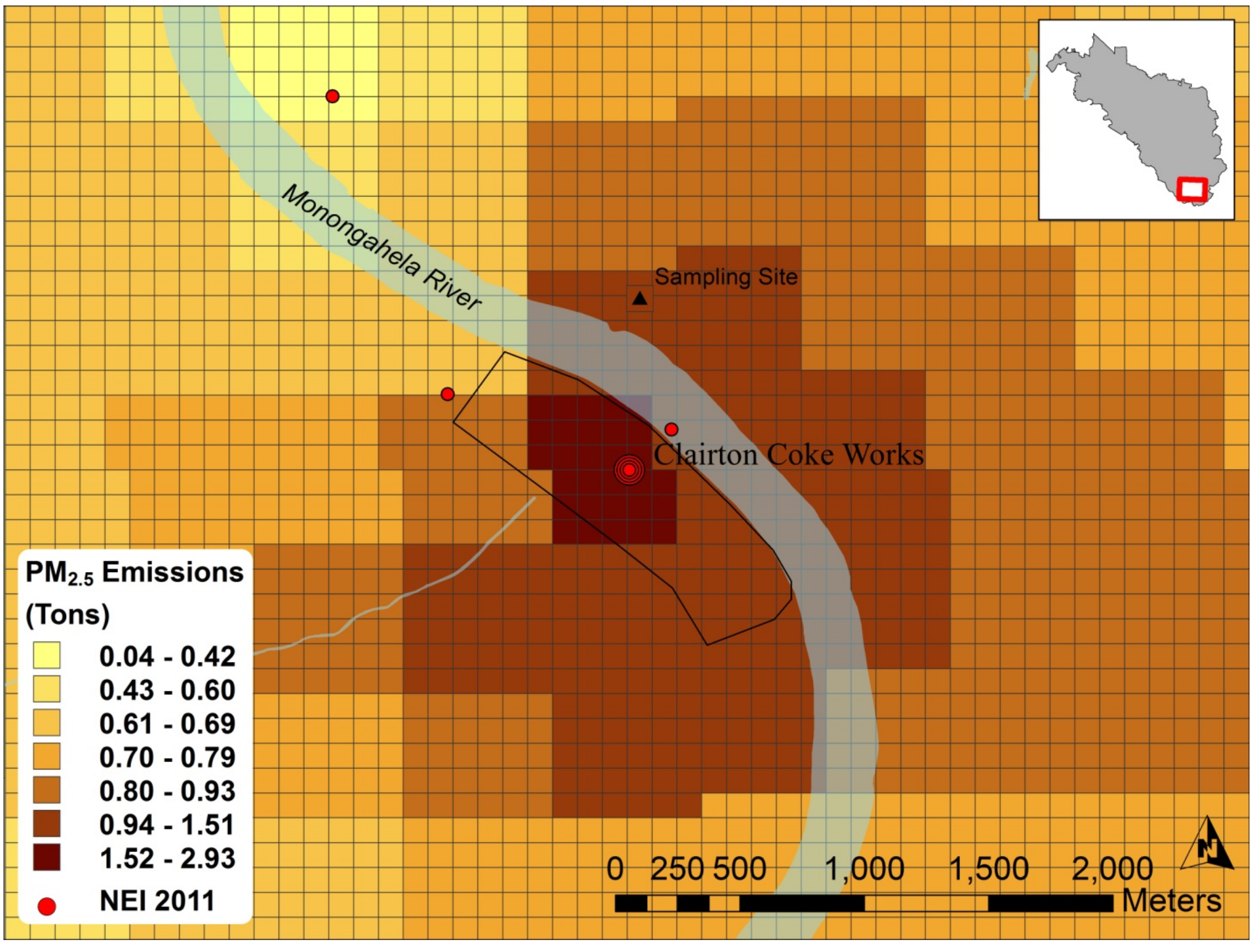

Figure 11. IDW Mean $PM_{2.5}$ emissions density (tons) at 100m x 100m grid resolution near the United States Steel Clairton Coke Works Facility in Clairton, PA (outlined in black). Surface derived from interpolated the EPA's 2011 National Emissions Inventory of $PM_{2.5}$ stationary sources as shown in red (NEI 2011)

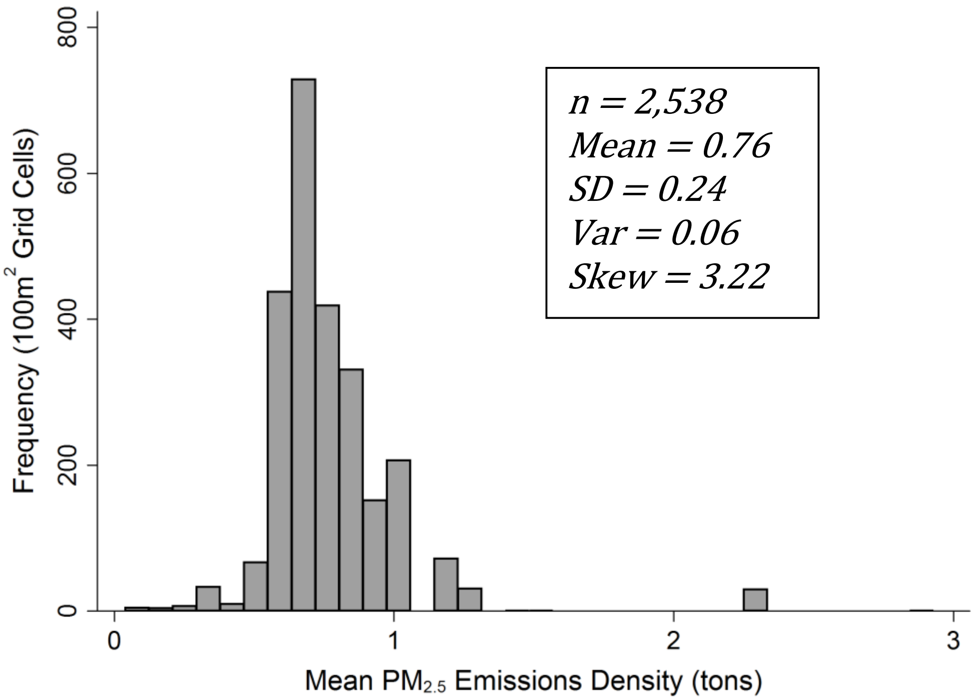

Figure 12. Frequency histogram with descriptive statistics of $PM_{2.5}$ emissions density in tons from spatial extent depicted in Fig. 11

The wind rose in Fig. 13 integrates the corresponding 1,488 modeled/sampling hours from the winter sampling season (Jan. 8th – March 10th, 2013), resulting in a predominant wind vector blowing from the west/south-west (255°). In contrast to the mean $PM_{2.5}$ emissions density surface displayed in Fig. 11, AERMOD predictions observed at the same spatial extent around the Clairton Coke Works, exhibited a more highly variable spatial pattern (mean = 2.54, var = 2.37) that includes source/ meteorological interaction information such as wind speed and direction. The 129 unique sources were aggregated to 27 unique sources with stack-specific geographic location within the facility.

Incorporation of dispersion principles resulted in a distinct delineation of upwind vs. downwind concentration gradients in proximity to the emissions sources. Furthermore, AERMOD predictions follow an exponential distance-decay pattern, which is more akin to observed air pollutant behavior (Whitlow et al., 2011). Additionally, the effect of varying terrain on pollutant behavior is captured by AERMOD and can be observed in Fig. 15, where the plume deposition centerline (dark brown) traverses diagonally and parallel to the opposing river valley hillside.

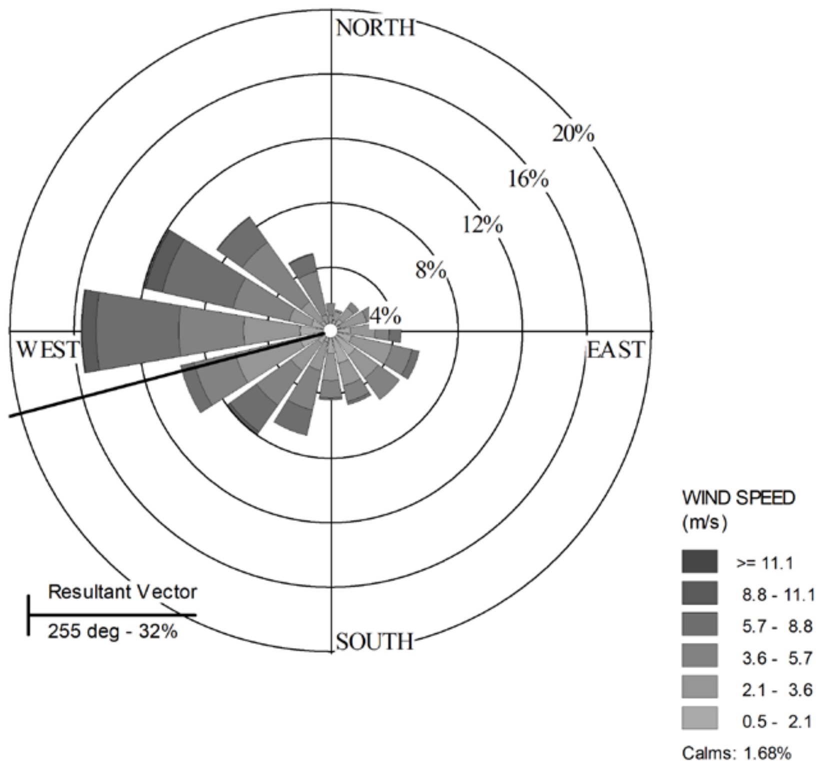

Figure 13. Wind rose displaying average speed (m/s) and direction (deg.) with resultant vector across all winter season $PM_{2.5}$ sampling/AERMOD modeled hours (1,488) from the IFW ASOS 1-minute (hourly averaged) data obtained from the NWS station at the Pittsburgh International Airport (40.5° N, 80.217° W)

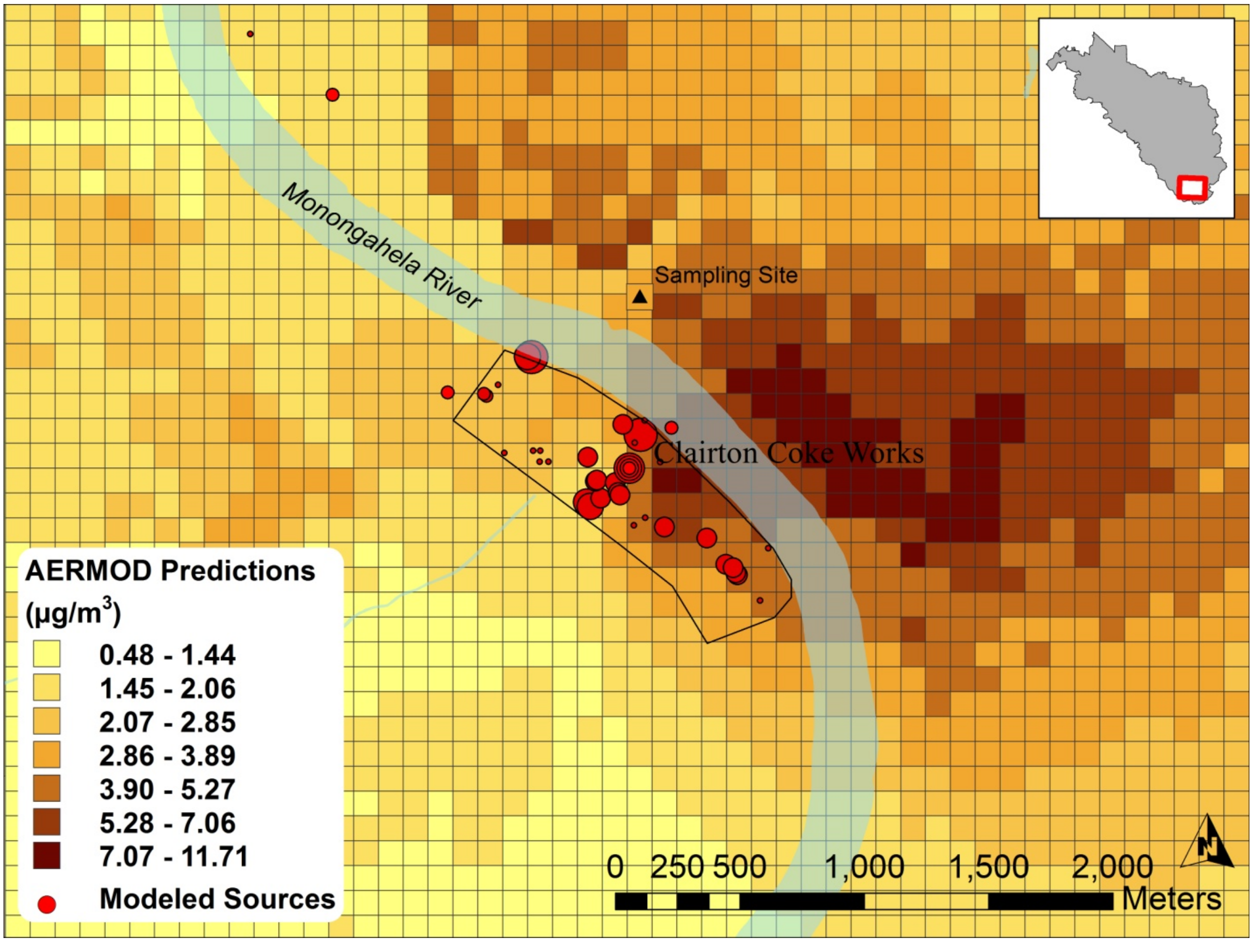

Figure 14. Choropleth map of winter (Jan 8th – March 10th, 2013) mean $PM_{2.5}$ AERMOD modeled concentration estimates at 100m x 100m grid resolution near the United States Steel Clairton Coke Works Facility in Clairton, PA (outlined in black). Red circles represent modeled $PM_{2.5}$ sources weighted by emissions factor (classification not shown)

Fig. 16 displays full model $PM_{2.5}$ predictions from the winter-season LUR-only model subtracted from the LUR/ AERMOD $PM_{2.5}$ model predictions at the non-sampled locations near the Clairton, PA area. The blue-shaded grid cells indicate areas where LUR overpredicted concentrations compared to the LUR/AERMOD hybrid model. Likewise, brown-shaded grid cells indicated areas where LUR alone underpredicted concentrations compared to LUR/ AERMOD predictions. Within this subset 5 x 5 km^2 area, the overall mean concentration difference did not differ substantially (+0.40 $\mu g/m^3$, SD = 1.17 $\mu g/m^3$). The maximum concentration difference between model predictions at the same $100m^2$ grid cell was +6.98 µg/m3, and was directly downwind from the facility. A minimum concentration difference between model predictions at the same $100m^2$ grid cell was of -2.88 μg/m^3 and was observed directly upwind from the facility.

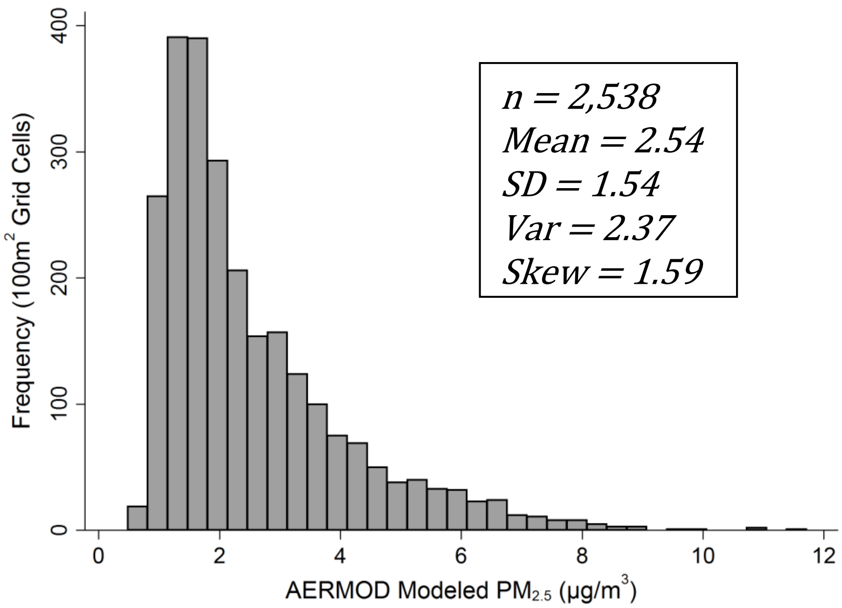

Figure 15. Frequency histogram with descriptive statistics of winter-season AERMOD $PM_{2.5}$ predictions in μg/m^3 from spatial extent depicted in Fig. 14

A complementary bar graph displaying the identical classifications to the choropleth map of Fig. 16 is included in Fig. 17; where, modeled concentration differences are plotted against the distance from the centroid of the industrial facility for each 100m x 100m grid cell. The maximum range in model prediction difference was 9.86 μg/m3, and was observed in area of less than 200m from the centroid of the facility. The areas of LUR overprediction (blue palette) exhibited a step-wise distance-decay pattern <400m from the facility and exhibited a near zero distance-decay ratio beyond 400m from the facility until a separate source was reached at over 2,400m. In contrast, the areas underpredicted by LUR (brown palette), exhibited a highly variable distribution with the most underpredicted areas (dark brown) closer to the facility and the less underpredicted areas (light brown) farther from the facility.

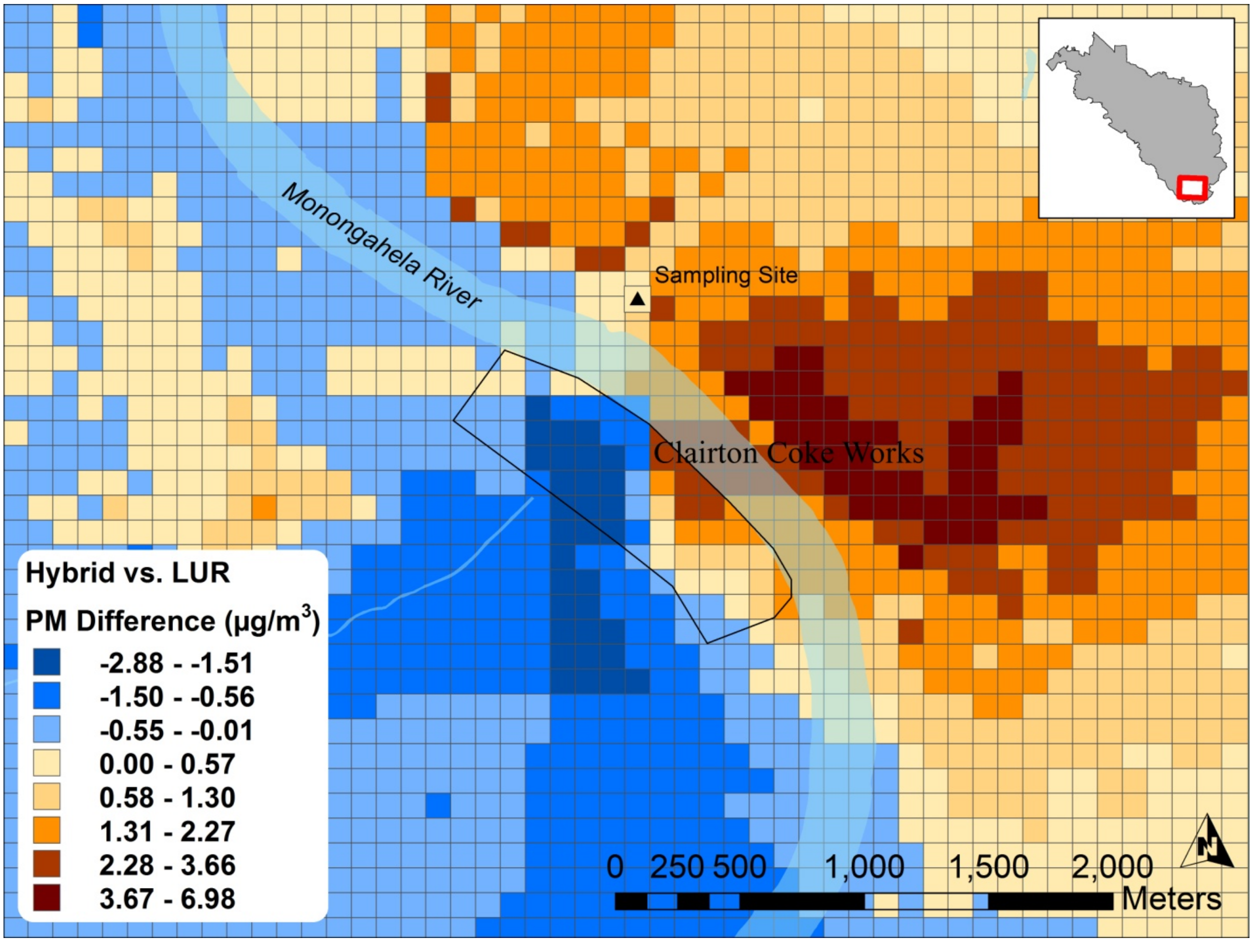

Figure 16. Concentration difference (Hybrid – LUR) in final winter-season model predictions for $PM_{2.5}$ at the 100m x 100m grid resolution in the area surrounding the United States Steel Clairton Coke Works Facility in Clairton, PA (outlined in black)

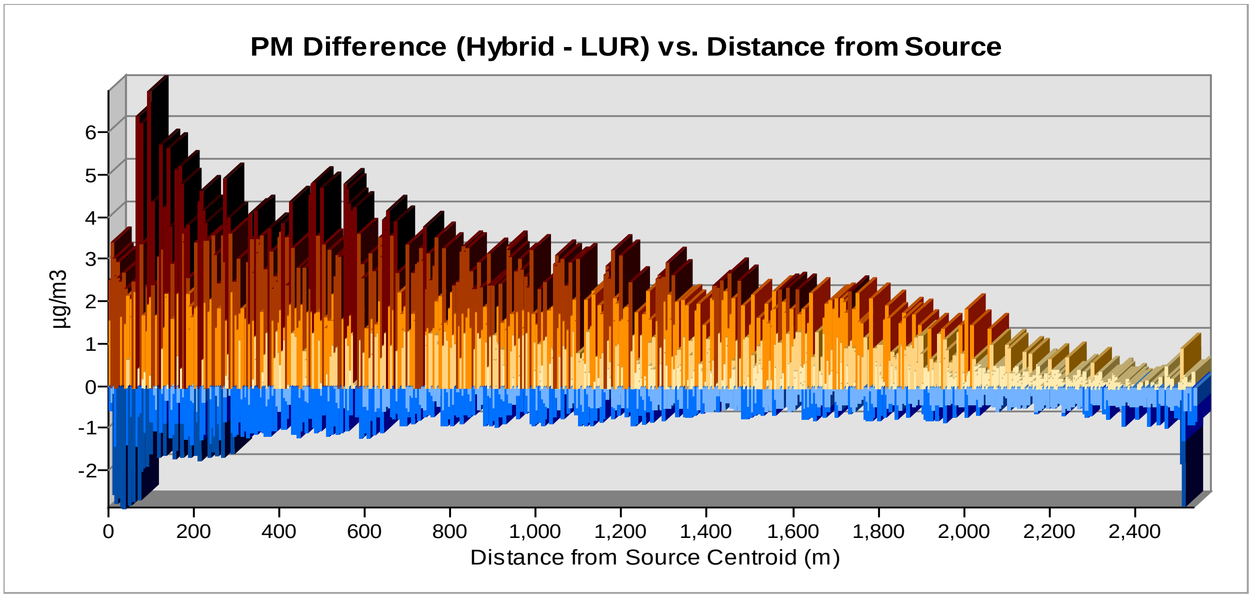

Figure 17. Concentration differences (Hybrid minus LUR) from Fig. 16 plotted as a function of distance from the centroid of the Clairton Coke Works facility. Color ramp classification values adhere to classification values presented in Fig. 16

3.3 DISCUSSION

We demonstrated the utility of adding stationary source dispersion output to LUR for predicting $PM_{2.5}$ across summer and winter seasons. To the best of our knowledge, this was the first attempt to explicitly add AERMOD predictions into a preexisting LUR as an independent predictor for estimating intra-urban $PM_{2.5}$. Overall, our LUR models built from 37 distributed measures performed reasonably well as per cross-validated R^2 values in comparison to similar efforts performed elsewhere. Summer and winter models differed by the degree of temporal variability observed and subsequently differed in explanatory variable structure. Our attempt to allocate monitoring locations to maximize variability by our three *a priori* source/modifying strata, may have influenced covariate selection and overall model prediction accuracies.

Temporal Adjustment: Because our measures were collected over a series of six sampling weeks each season, LUR models require adjustment for temporal variance using reference site data. Following the addition of AERMOD, slightly better statistical improvements were observed in the winter model compared to the summer model. This could partially be explained by the difference of variance explained by the temporal terms between the seasonal models. Given the regionally-varying nature of PM, the effect of long-range transport is indicated by the co-variance of distributed site measurements (box-plots) with regional background measurements (line-plot) in Fig. 10. The up-front adjustment for temporal variance in LUR could potentially handicap the intrinsic utility of AERMOD, effectively limiting the temporal variability resolved from meteorological information. This is may be evident by the slight decreases in standardized beta coefficients of the temporal terms following the addition of AERMOD to LUR.

Physical Model Interpretability vs. Statistical Fidelity: Minimal prediction accuracy improvement following the addition of a deterministic dispersion term to LUR has been reported (Lindström et al., 2013). The authors acknowledge the challenge in disentangling spatial and temporal contributions to a spatio-temporal model framework. Methods to decompose these facets within air quality modeling have been demonstrated, though application here is beyond the scope of this effort (Keller et al., 2014; Lindström et al., 2013). Marginal statistical improvement in terms of variance explained, could be attributable to the relatively large averaging area represented in by our modeling domain. For instance, areas that exhibit divergent urban-to-suburban gradients with diverse source regimes may necessitate less specific and more generalizeable pollutant surrogates (e.g., population density). Yet, specific source/meteorological interaction information can improve physical interpretability of concentration predictions especially in particular near-source gradients as was presented here and by others (Cook et al., 2008; Isakov et al., 2009; Wilton et al., 2010).

Therefore, an evaluation of statistical fidelity and physical model interpretability should be considered, especially in areas of distinct source regimes.

Transferability of LUR models is also desirable; however, attempts to transfer LUR models across space (e.g., intercity) and time commonly resulted in a loss in explanatory power and increased uncertainty (Allen et al., 2011; Poplawski et al., 2008; Vienneau et al., 2010). Success of LUR transferability may depend more so on between-city consistency of input data rather than geographical differences (Poplawski et al., 2008), therefore, universal air quality models could satisfy data input misalignment across study areas. Because AERMOD accounts for hourly meteorological variability and source-meteorology interactions, the hybrid approach may substantially improve interpretability of source terms, and ultimately may prove more reliable for model portability, though this was not explicitly tested.

Limitations: Though we observed moderate improvement in model predictions by adding AERMOD predictions, the applicability to other areas remains uncertain. Our sampling domain contained numerous steel- and coke-related industrial sources that emit particles at near ground-levels (e.g < 100m). We also acknowledge that evaluating a spatio-temporal explanatory variable with temporally misaligned measures is challenging. Furthermore, 37 distributed monitoring locations across our sampling domain may not be sufficient to resolve properly specified empirical models (Basagaña et al., 2013; Basagaña et al., 2012). From our analyses, it was beyond the scope to evaluate the relative contribution from smaller point sources for short-term pollutant predictions. Though, model predictions appeared to be overly sensitive to stack height, and low exit velocity (e.g., fugitives) input parameters.

Wide adoption of air quality models has been hindered by relatively intensive data input requirements, high costs, and programming demands; however, recent Microsoft graphical user interfaces (e.g., Lakes Environmental, BREEZE Software) have benefitted ease of use. A major limitation in resolving reasonable predictions from deterministic models is the degree of accuracy of input data. Therefore, we greatly benefited from the expert collaboration with the Allegheny County Health Department's (ACHD) Air Quality/Pollution Control Program personnel. An emissions input data file for AERMOD was assembled by ACHD staff, and corroborated following updates. These data exist, in part, through the regulatory standing of the ACHD, and as a result of the USEPA's air quality designations for the $PM_{2.5}$ National Ambient Air Quality Standard (NAAQS) standard for the Pittsburgh-Beaver Valley and the Liberty-Clairton areas. As part of section 189(a)(2)(B) of the Clean Air Act, state and local governing bodies are required to submit State Implementation Plans (SIP) to demonstrate plans for attainment that usually entail detailed modeling efforts. Furthermore, new source permits in air quality designated areas, such as Pittsburgh and many other urban areas, must demonstrate emissions scenarios to be amenable with SIP NAAQS attainment goals, from which, verified AERMOD source input information can be obtained. Nonetheless, prediction measurement error due to modeling error can introduce additional uncertainty in the final exposure surfaces and therefore requires thoughtful consideration.

AERMOD and meteorological data: Meteorological data is also a source of potential error, and we found that meteorological data obtained from the National Weather Service station near the Pittsburgh International Airport provided more accurate predictions than data obtained from the weather station at the Allegheny County Airport, even though the former station was located approximately 20 miles west of our sampling domain, compared to the latter station located within our sampling domain. We also tested model runs with and without ASOS 1-minute data collected from ice-free anemometers from each meteorological station to examine the impact of missing hourly wind data. Formatted hourly wind speeds produced from non ASOS 1-minute data resulted in approximately 17% missing values annually, compared to <1% missing values for wind speeds derived from ASOS 1-minute sonic anemometers. This is partly due to the sensitivity to calm wind speeds (<1.76 m/s) programmed into AERMOD, and the subsequent randomization of wind speeds and wind speed truncation algorithms. These adjustments were in place to overcome the uncertainties of low wind speeds obtained from hemispherical cup anemometers, and have since been reconciled with the adoption of sonic anemometers and AERMOD's capability to integrate ASOS 1-minute wind data via AERMINUTE.

Based on best use practices as determined by the EPA for AERMOD, multiple years of meteorological data are recommended to obtain more robust modeled estimates (U.S.E.P.A., 2005). However, since our sampling sessions spanned a 7-day week, we modeled 7-day, seasonal, and annual averaging times to test the sensitivity to meteorological data in producing a significant covariate across the monitoring locations. Not surprisingly, slightly more variability was expressed in the 7-day averaging time period compared to the seasonal and annual model runs. Notably, the impact of longer averaging times was most noticeable at the monitoring locations proximal to larger industrial sources, where longer averaging times tended to reduce predicted concentrations. A combination covariate was also tested, where monitoring locations near major emissions sources (n = 3) were modeled annually and low industry sites were modeled according to the 7-day averaging time. While winter LUR models were less sensitive to variations of modeled $PM_{2.5}$ from AERMOD, the 7-day or session-specific averaging times most improved model fits across both seasons, potentially indicating the contribution of apropos source/ meteorology interaction.

Strengths and Implications: AERMOD moderately improved overall model fits as per cross-validated performance statistics, and effectively displaced the GIS-based $PM_{2.5}$ emissions density term in each season, corroborating the interpretability of each. The efficacy of AERMOD as a covariate for LUR ultimately resides in its ability to represent a high degree of spatio-temporal variability that spans the relevant exposure environments that may not be captured by the monitoring locations (e.g., sparse regulatory monitors). Therefore, it is preferable to design exposure assessments that maximize variability in apropos geographic covariates across both monitoring sites and subjects within a cohort (Szpiro et al., 2011a).

We demonstrated that AERMOD can produce a physically-realistic prediction surface compared to typical GIS-based covariates, especially in an area of high pollutant-source intensity. Notably, the $PM_{2.5}$ density variable was almost five times less variable (σ^2 = 0.25) across all 37 distributed monitoring locations, compared to variances of 1.18 and 1.45 for summer and winter AERMOD terms, respectively, which may result in more appropriate exposure measurements. This may have an important bearing in better understanding exposure measurement error approximated from invariable geographic covariates in LUR for epidemiological studies.

3.4 SUMMARY

Incorporating AERMOD into LUR models improved model predictions as per cross-validated coefficient of determination and RMSE, and explained an additional 9-13% in out-of-sample variability in $PM_{2.5}$. Following the addition of AERMOD output, the industrial geographic term in both summer and winter models was no longer significant. AERMOD provides a beneficial tool for exploring the spatio-temporal nature of the pollutant measurements for model building, especially in areas of high industrial-source intensity and complex terrain. Furthermore, if model improvement is confirmed, AERMOD predictions could be modeled directly at the subjects' residential addresses, and tailored to the averaging times of interest in an epidemiology setting.

In Chapter 4, we utilize AERMOD predictions to supplement an annual $PM_{2.5}$ prediction model by combining summer and winter measurements with annual AERMOD estimates for epidemiological application. We then simulate a theoretical cohort of 5,000 within our modeling domain to examine the potential magnitude of bias and variance inflation in health-effect estimates between LUR and LUR/AERMOD using a Monte Carlo simulation framework. Explicitly, we examine the potential for health estimate bias that may result from spatial model misspecification, and ultimately how much of the *true spatial variability* is explained by the model.

4.0 EVALUATING MEASUREMENT ERROR IN HEALTH EFFECT ESTIMATION USING HYBRID AERMOD/ LAND USE REGRESSION

With the advent of more sophisticated exposure prediction models, assessing measurement error is worthwhile given the increasing evidence for small-scale (e.g., intra-urban) pollutant variability, implying that the most meaningful exposure gradients may occur at very small (e.g., <50m) spatial gradients (Brauer et al., 2003; Clougherty et al., 2013b; Clougherty et al., 2008; Cook et al., 2008; Hoek et al., 2002; Jerrett et al., 2005; Kheirbek et al., 2012; Marshall et al., 2008). As it is not practical to measure personal exposures for all individuals in large cohort studies, exposure assessments that estimate proximal ambient air pollution, usually at the residential address, are commonly employed (Jerrett et al., 2005). These predicted exposures are then included as explanatory variables in a regression model to evaluate a health effect parameter of interest. However, the use of predicted air pollution levels as surrogates of true exposure, are inevitably affected by measurement error and uncertainty (Basagaña et al., 2013).

To sufficiently capture temporal variation annual average concentrations it is necessary to sample during the majority of a year at a large number of sites (Hoek et al., 2002). Most LUR studies are developed over a limited sampling period with varying numbers of measures, and are extrapolated to specific time periods of interest. Thus, it has been assumed that exposure predictions with less measurement error relative to the unknown true exposures will result in improved health effect estimates (Jerrett et al., 2005). LUR for exposure assessment, however, can be constrained by the spatial variability expressed by the pertinent geographic predictors in relation to the locations of the monitoring sites, and the true underlying pollutant variability (Alexeeff et al., 2014; Basagaña et al., 2013). The degree to which exposure prediction, and subsequent exposure measurement error engenders uncertainty and bias in health-effect estimates has invoked research interests (Alexeeff et al., 2014; Basagaña et al., 2013; Dionisio et al., 2013; Szpiro et al., 2011a; Szpiro et al., 2011b) especially for imminent multipollutant modeling frameworks (Dionisio et al., 2014) .

LUR and dispersion models are thought to perform similarly given optimum conditions (Dijkema et al., 2011). Though, high spatial correlations between models suggest reliability of overall long-term effect estimation derivation, small-scale refined information can lead to spatially differential estimates in effect estimates. Thus, for population-dense urban areas, small differences in measurement error and subsequent risk estimates can have important results, especially in spatially stratified analyses (Sarnat et al., 2013). Moreover, spatial refinement in exposure estimates may allow for more accurate source-concentration interpretability and in identifying subsequent associations among population subgroups for environmental justice intervention.

In this chapter, we explore the impact of measurement error on health effect estimates using LUR and hybrid AERMOD/ LUR models. We constructed two annual $PM_{2.5}$ prediction models by combining summer and winter measurements (presented in Chapter 3) with (1) local EPA AQS measures; and (2) local EPA AQS measures and annual long-term AERMOD predictions. Specifically, we examine AERMOD's potential to impact measurement error and subsequent acute and chronic health-effect bias. We used a simulated cohort of 5,000 residential addresses to examine the potential magnitude of measurement error between annualized LUR and AERMOD/ LUR modeling frameworks. We also apply a generic Monte Carlo simulation utilizing statistical properties from a GIS-based predictor and the AERMOD predictions to demonstrate the impact of distributional variance on heath effect estimation and bias.

4.1 METHODS

$PM_{2.5}$ measures, study design, site selection, and LUR model building methods were presented in detail in Chapter 3. Here, we construct and evaluate an annual $PM_{2.5}$ prediction model utilizing multi-season distributed measures and temporal trends from routine regulatory monitors for epidemiology application. To further supplement temporally misaligned measurement data, we included a long-term average of AERMOD dispersion output predictions and examined model improvement. We examine model prediction efficacy by applying exposure estimates to a theoretical cohort of 5,000 individuals. Finally, we explicitly compare the $PM_{2.5}$ emissions density covariate to AERMOD predictions in a Monte Carlo simulation to demonstrate the effect of explanatory covariate variability on health effect estimation.

4.1.1 Merged Season LUR Model

To produce a spatially-refined model for temporal extrapolation (e.g., daily, annual), a merged seasonal model was constructed by combining summer and winter $PM_{2.5}$ measures, resulting in 74 total dependent observed values, repeated over two seasons. To control for repeated measures across seasons, a random intercept with an independent unstructured covariance was applied (p = 0.003) in a mixed model framework with restricted maximum likelihood estimation. A merged season LUR was first constructed utilizing the study-deployed regional background measures to corroborate spatial covariate structure before applying temporal adjustment schemes (e.g., daily PM measures from routine regulatory monitors) necessary for temporally extending spatial LUR estimates. Explanatory variable selection procedures were followed as presented previously in Section 3.1.5.

4.1.2 Temporal Model Extrapolation

To temporally extend the spatial variability explained by the LUR models to various time scales (e.g., daily, annual), we examined regionally-located daily $PM_{2.5}$ measures from EPA's regulatory Air Quality System (AQS). The temporal stability of $PM_{2.5}$ measures across a greater six-county region of southwestern PA was examined through time series application of routine regulatory monitors from 2000-present. Three criteria were followed to extrapolate a temporal trend from nearby regulatory monitoring data: (1) agreement with regional background measures (two summer; two winter season) obtained during dedicated sampling campaigns, to allow for model validation; (2) data quality (e.g., sampling method, co-located monitors, non-systematic missing); (3) representativeness of a greater regional trend of Southwestern PA from 2000-present; and, (4) interpretability.

In following these criteria, a single 24-hr AQS (Thermo Scientific TEOM single point monitor) monitor demonstrated the most robust and representative temporal trend (Fig. 18). The selected AQS site (hereafter called central AQS) is located centrally located, and functions as designated NCore station consisting of multiple co-located $PM_{2.5}$ measures (e.g., FRM filter-based, FEM continuous Met One BAM) which greatly reduced the uncertainty in supplementing missing values. Though, data quality from this monitor is robust, with only 176 missing days over 11 years (2003-2013). In respect to our modeling domain, the monitor is located outside of the urban core, in a mixed commercial/residential area.

Daily measures from the selected AQS sites were matched and averaged to our dedicated weekly sampling sessions. These values were then substituted into the pre-existing seasonal and merged season LUR models to examine the changes in explanatory variables, similarly to when we added Caline3 and AERMOD. Though the selected monitor may capture a different nearby source regime in comparison to the regional background site, all prior explanatory variables were retained ($p < 0.05$) when the central AQS measured were used as temporal controlling term. Therefore, we did not reconstruct the LUR models with the AQS adjustment, as we assumed the geographic covariates chosen best represented the spatial variability in intraurban PM.

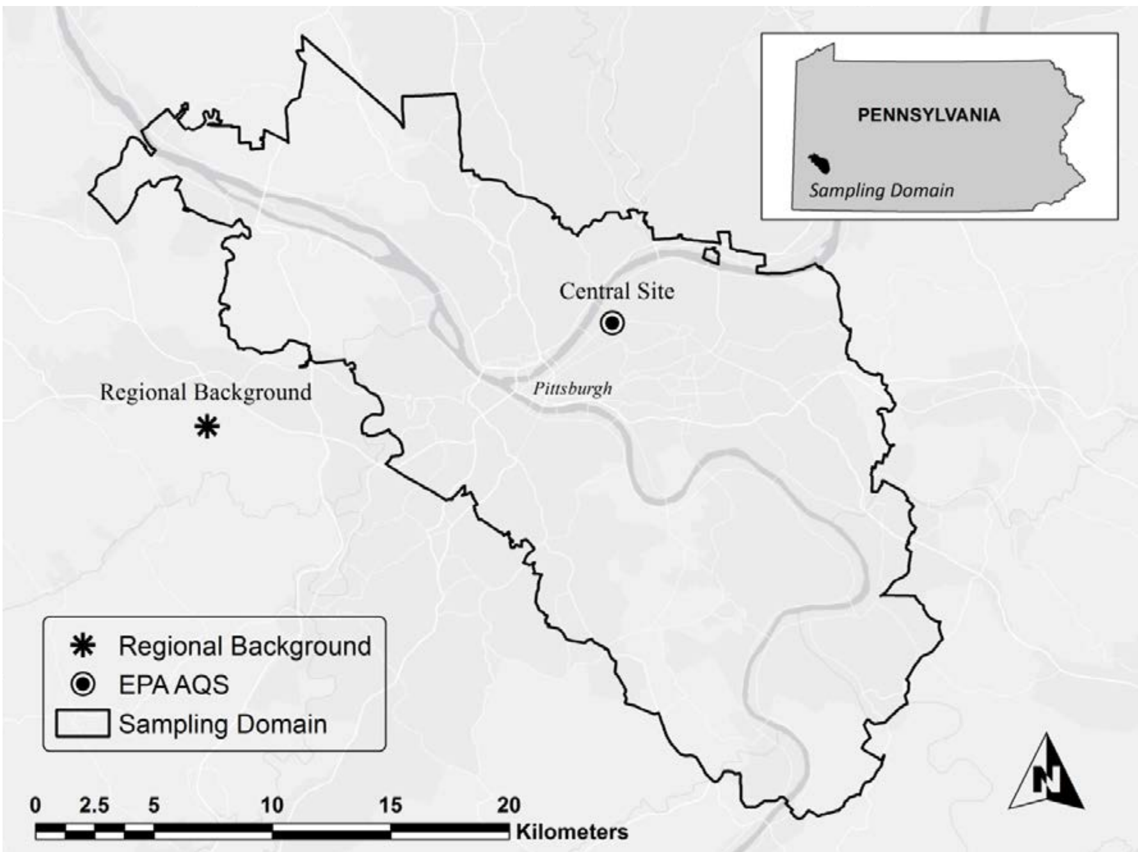

Figure 18. Sampling domain with designated regional background and EPA AQS central sites

4.1.3 Hybrid LUR/AERMOD $PM_{2.5}$ Prediction

To further supplement temporally misaligned measurement data, we included a long-term average of AERMOD dispersion output predictions and examined model difference. In contrast to the previous hybrid model framework described in Chapter 3, AERMOD predictions were approximated using a full year (2012) of hourly meteorological data as opposed to sampling session-specific averaging times. AERMOD predictions, therefore, capture long-term source/ meteorological interaction information across the entire modeled year. Similar to prior methods presented, the dispersion output was included as an independent covariate in the combined season model and model fits were assessed. Likewise, to produce an independent covariate in the merged seasonal LUR model, AERMOD receptor locations were defined at the monitoring locations (Fig. 7). To account for complex terrain (e.g., river valleys) effects, a 1km^2 uniform Cartesian receptor grid was included in addition to discrete receptors in all model runs. The resulting modeled predictions were added separately to the merged LUR model according to the formula:

$$\boldsymbol{C_s = \beta_0 + \beta_1 TEMP_j + \sum_{i=1}^{m}(\beta_i\, x_{i,s}) + \alpha_{AER}\left(\sum_{t=1}^{h} d_{s,t}^{AER}\right) + \varepsilon_s}$$

(4.1)

Where,

$\boldsymbol{\alpha_{AER}}$ = regression coefficient for the AERMOD covariate

$\boldsymbol{d_{s,t}^{AER}}$ = dispersion concentration (μg/m^3) modeled from AERMOD for site *s* for hour *t*

Following the model building/validation procedures, the explanatory variables derived in equation 3 were used to solve for concentrations predictions at non-sampled locations at the 100 x 100m grid resolution for the entire modeling domain.

4.1.4 Randomized Cohort Simulation

To examine whether AERMOD predictions attenuate exposure measurement error, a randomized theoretical cohort of 5,000 point locations was generated. To maximize spatial coverage and limit clustering, neighboring point locations were set at 100m distance intervals. Predictions from both annualized models were made at the 5,000 point locations and were compared spatially and temporally (e.g., daily).

4.1.5 Health Effect Estimation for Epidemiological Application

Health effect estimation can be derived from association-type studies, where statistical relationships are resolved typically by linear or logistic probabilistic models. Considering an association-type linear health effect model with the general form:

$$\boldsymbol{Y = \beta_0 + \beta_x X + \varepsilon}$$

(4.2)

Where, Y is the observed health outcome, X is the true pollutant exposure, and β_x is the effect estimate of interest. Since X is not measured at all residential locations of the N study participants, but at $n < N$ locations, the LUR model is constructed from n measures and a subset of r potential predictors are used to predict exposure $\hat{z}$ at the N residential locations. Thus, it is common practice to obtain the predicted health effect estimate $\hat{\beta}_z$ from a regression of Y on $\hat{z}$, also referred to as the naïve plug-in estimator. Therefore, there is interest in understanding the effect on $\hat{\beta}_z$ from factors of $\hat{z}$ estimation using LUR models (e.g., measurement error, model specification, variable selection, sample size).

4.1.6 Monte Carlo Simulation

We adapted the stochastic model simulation framework developed by Szpiro et al. (2011a) to examine the health effect estimate difference between two study-generated geographic covariates. The statistical theory within the model simulation is described in detail elsewhere (Gryparis et al., 2009; Szpiro et al., 2011a; Szpiro et al., 2011b). Briefly, the stochastic simulation performed by Szpiro et al. (2011a) assumed a well-characterized spatial model, from which exposure surfaces were generated using 100 theoretical pollutant measures and three geographic covariates for 10,000 subjects. The covariates were assumed to be independent of each other at all locations and between subjects. The first two covariates were distributed as $N(0,1)$, but the third as $N(0, \sigma^2)$, where σ^2 represents the degree of variability at the monitoring locations. $\hat{\beta}_z$ was then obtained by regressing a randomized distribution of a hypothetical linear health outcome with $\beta_0 = 1$, $\beta_x = 2$, $\sigma_\varepsilon = 25$ characteristics against the resolved exposure predictions for each cohort individual. This process was repeated 80,000 times to obtain information on the health effect estimate given various degrees of variability explained by the third geographic covariate in each linear LUR model.

Our simulation was designed to compare the variability explained between the two geographic covariates of interest obtained from our LUR model building process utilizing the 37 monitoring locations. These included: (1) $PM_{2.5}$ emissions density within 50m that varied about the 37 monitoring locations with a mean 0.52 and standard deviation of 0.54, and; (2) 2012 annual $PM_{2.5}$ AERMOD predictions that varied about the 37 monitoring predictions with a mean of 1.49 and variance of 1.45. To test the impact on health effect estimates using these two study-specific covariates, we utilized the standard deviations of each covariate to define the random distributions to produce exposure estimates for each of the theoretical 5,000 cohort members in separate simulations. We restricted the number of monitor values to 40 and number of cohort subjects to 5,000, and repeated the process 50,000 times. We compared the mean and standard deviations of $\widehat{\beta}_z$, and mean R^2 and RMSE between the two simulations.

4.2 RESULTS

4.2.1 EPA Air Quality System Measures

Weekly average measures from both the regional background site and the central AQS site are included in Table 9 and Fig. 19. On average, the central AQS site recorded higher concentrations within both seasons compared to the regional background site previously utilized for temporal LUR adjustment; however, a larger degree of difference in concentrations were observed in the winter season. The central site was efficient in capturing the temporal trend across sampling sessions as evident by the covariance structure shown in Fig. 19.

Table 10. Summary statistics comparing $PM_{2.5}$ temporal adjustment measures in μg/m^3

	Summer Background	Summer Central AQS	Winter Background	Winter Central AQS
n	6	6	6	6
Min	9.0	11.8	6.8	9.0
Max	15.7	17.3	11.5	15.1
Mean	11.9	12.9	8.4	11.4
Median	11.9	12.1	8.1	10.4
SD	2.2	2.2	1.8	2.5

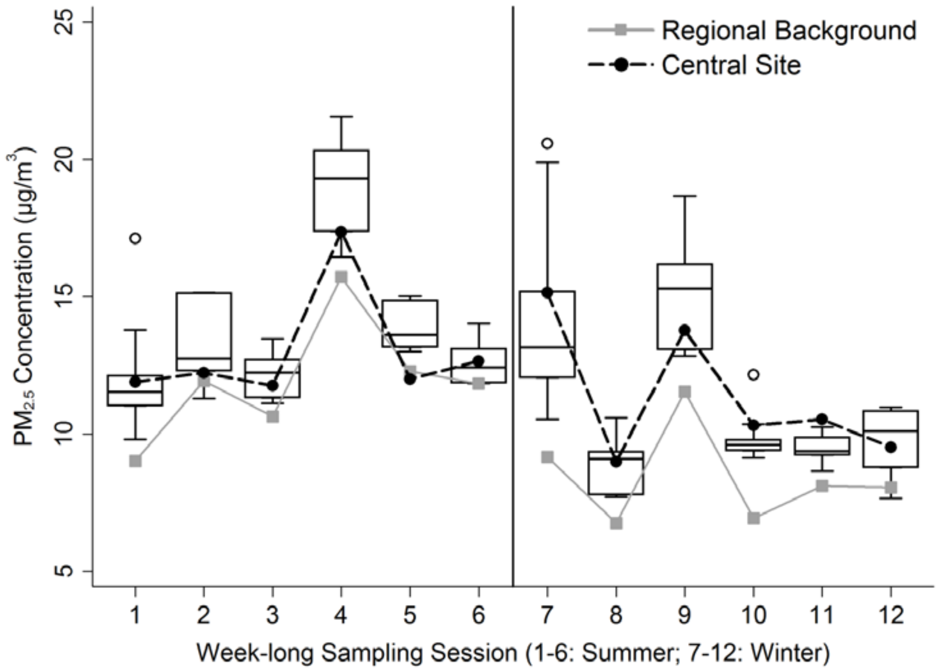

Figure 19. Summer and winter boxplots of $PM_{2.5}$ measurements from distributed sites with linear plot of regional background and central site measures (EPA AQS)

4.2.2 Merged Season LUR $PM_{2.5}$ Predictions

All prior explanatory variables were retained ($p < 0.05$) following the replacement of the regional background term with the central AQS term. The merged season LUR model with the central AQS term was identical in covariate structure to the winter-only model presented in the Chapter 3, and produced a final R^2 value of 0.76 (Snijders/Bosker Level 1) and AIC of 319 with the AQS adjustment. Final LUR $PM_{2.5}$ predictions for 2012 are shown in Fig. 20 in deciles with two addition classification breaks added at 12.0 and 15.0 to coincide with current and former national ambient air quality standards for the annual arithmetic mean of $PM_{2.5}$.

Table 11. Merged-season standard LUR (n=72) with sequential R^2 and AIC

Covariates Predicting Summer + Winter $PM_{2.5}$	LUR		
	$PM_{2.5}$ β (p-value)	*Seq.* R^2	AIC
Intercept	-1.25		
Central AQS $PM_{2.5}$	1.03 **	0.71	307
Industrial parcel area (750m)	$3.3x10^{-6}$ *	0.74	322
Traffic signals (750m)	0.07 *	0.76	321
Density of $PM_{2.5}$ Emissions (50m)	0.81 *	0.77	319

***significant: p <0.05; **significant p <.0001**

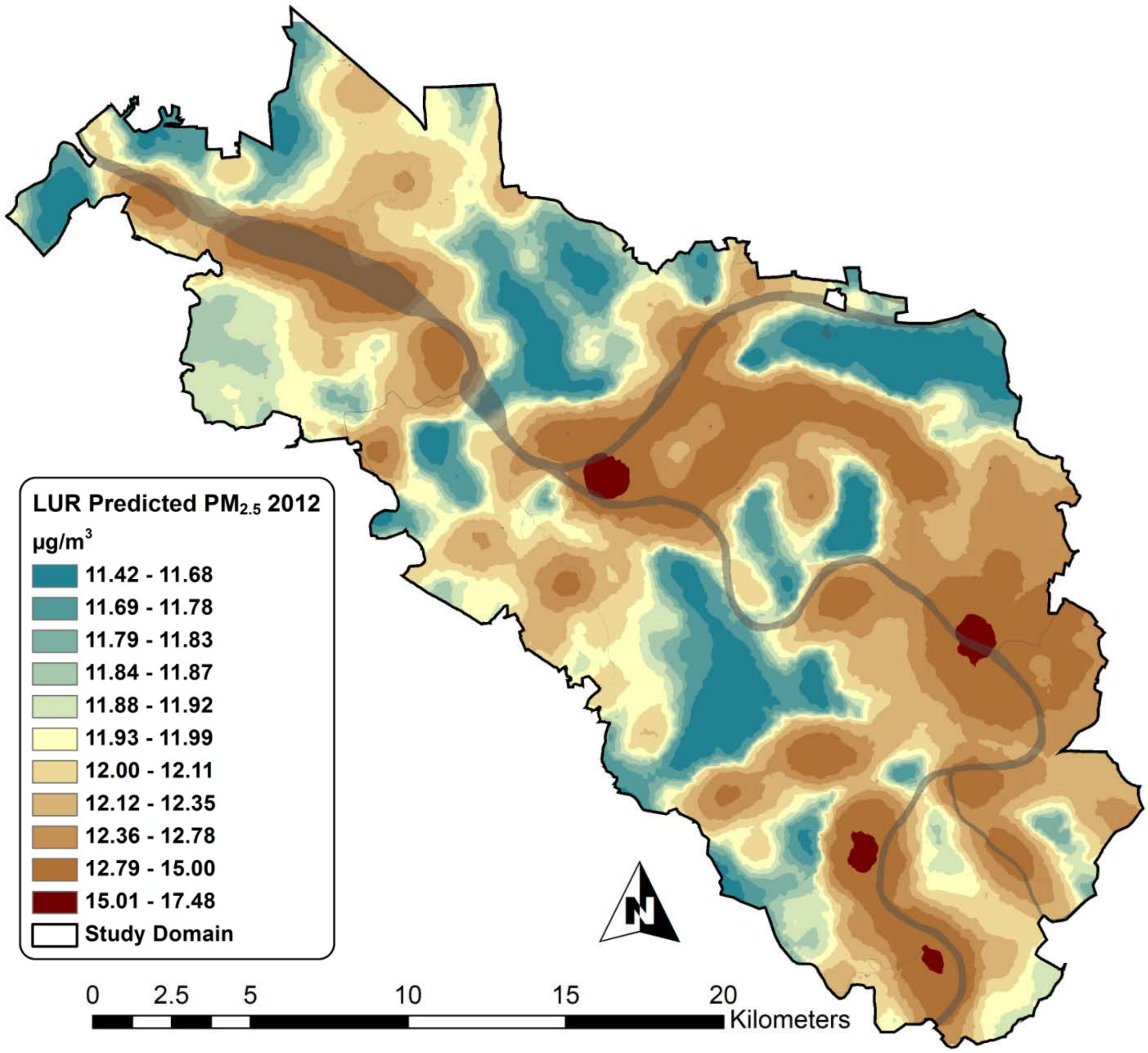

Figure 20. Annual 2012 LUR $PM_{2.5}$ predictions across the study domain

4.2.3 Merged Season Hybrid AERMOD/LUR

The hybrid LUR/ AERMOD model is shown in Table 11. Similarly to the seasonal models presented in Chapter 3, the addition of AERMOD output replaced the *density of $PM_{2.5}$ emissions* term and slightly increased the overall R^2 value to 0.77 and improved the AIC to 287. Notably, the AERMOD output utilized here was derived from an annual AERMOD $PM_{2.5}$ prediction model.

Final LUR/ AERMOD $PM_{2.5}$ predictions for 2012 are shown in Fig. 20 in deciles with two addition classification breaks added at 12.0 and 15.0 to coincide with current and former national ambient air quality standards for the annual arithmetic mean of $PM_{2.5}$.

Table 12. Merged-season hybrid AERMOD/LUR (n=72) with sequential R^2 and AIC

Covariates Predicting Summer + Winter $PM_{2.5}$	**Hybrid AERMOD/LUR**		
	$PM_{2.5}$ β (p-value)	***Seq. R^2***	**AIC**
Intercept	-0.93		
Central AQS $PM_{2.5}$	0.98 **	0.71	307
AERMOD 2012	0.50 *	0.75	307
Traffic signals (750m)	0.08 *	0.76	294
Industrial parcel area (750m)	3.0×10^{-6} *	0.78	316

***significant: $p < 0.05$; **significant $p < .0001$**

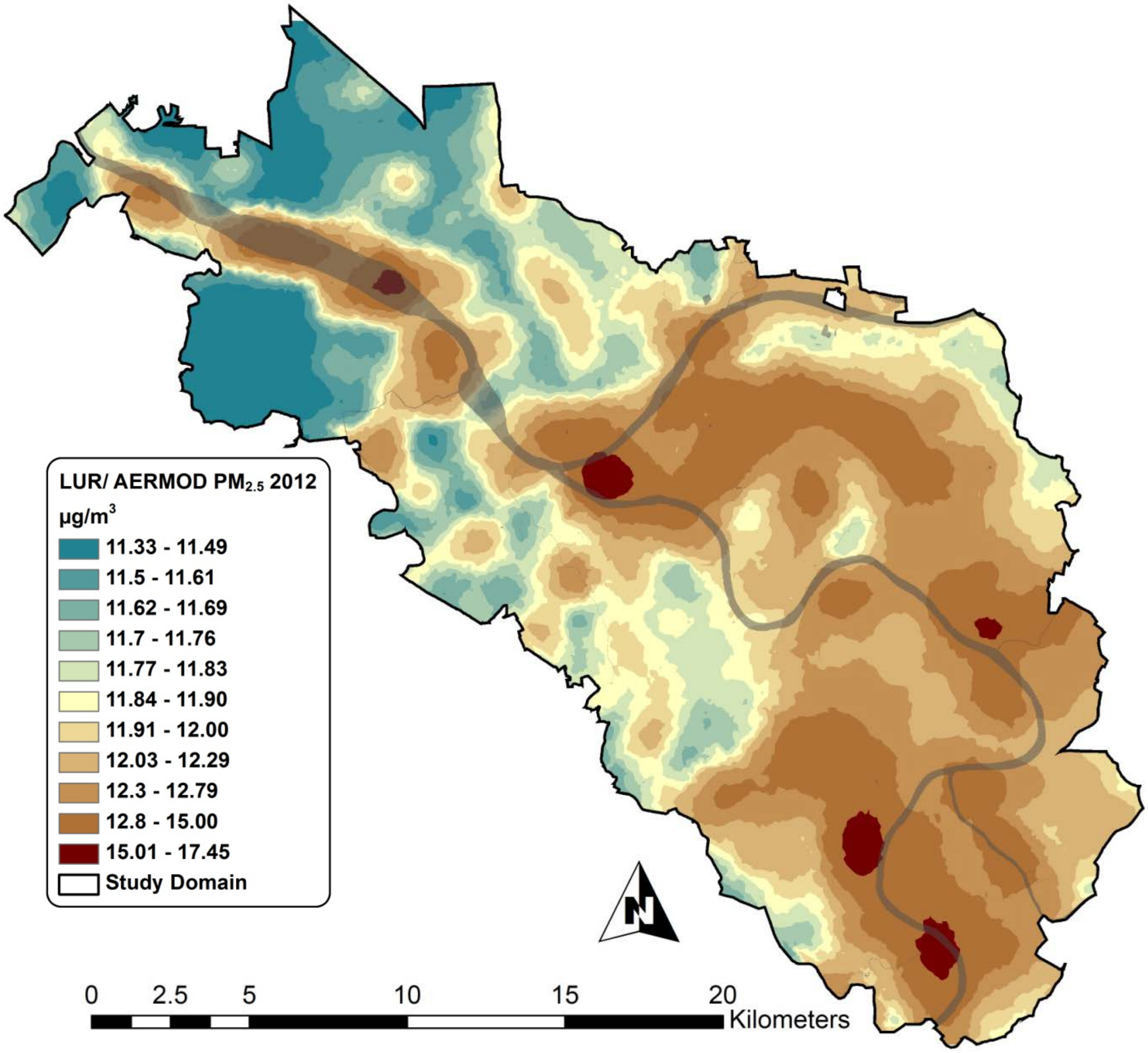

Figure 21. Annual 2012 LUR/ AERMOD $PM_{2.5}$ predictions across the study domain

4.2.4 Long-term Spatial Variability

After producing final prediction models across our modeling domain for the Greater Pittsburgh Area, we predicted exposures using each model at a randomized hypothetical cohort of 5,000 point locations. The prediction differences (hybrid – LUR) are depicted in Fig. 23 and descriptive statistics are shown in Table 13. In Fig 23, blue-to-green color gradients indicate locations where LUR predictions were higher compared to LUR/ AERMOD predictions. Conversely, yellow-to-red color gradients indicate areas where LUR underpredicted concentrations compared to LUR/ AERMOD exposure predictions.

Table 13. Summary statistics of model difference in μg/m³ corresponding to coordinate-level predictions displayed in Fig. 23

Exposure Model	*n*	Min	25th percentile	*Mean*	75th percentile	Max	Var
LUR	5,000	11.42	12.15	12.68	12.95	19.19	0.53
LUR/AERMOD 2012	5,000	11.26	11.72	12.27	12.54	19.13	0.77

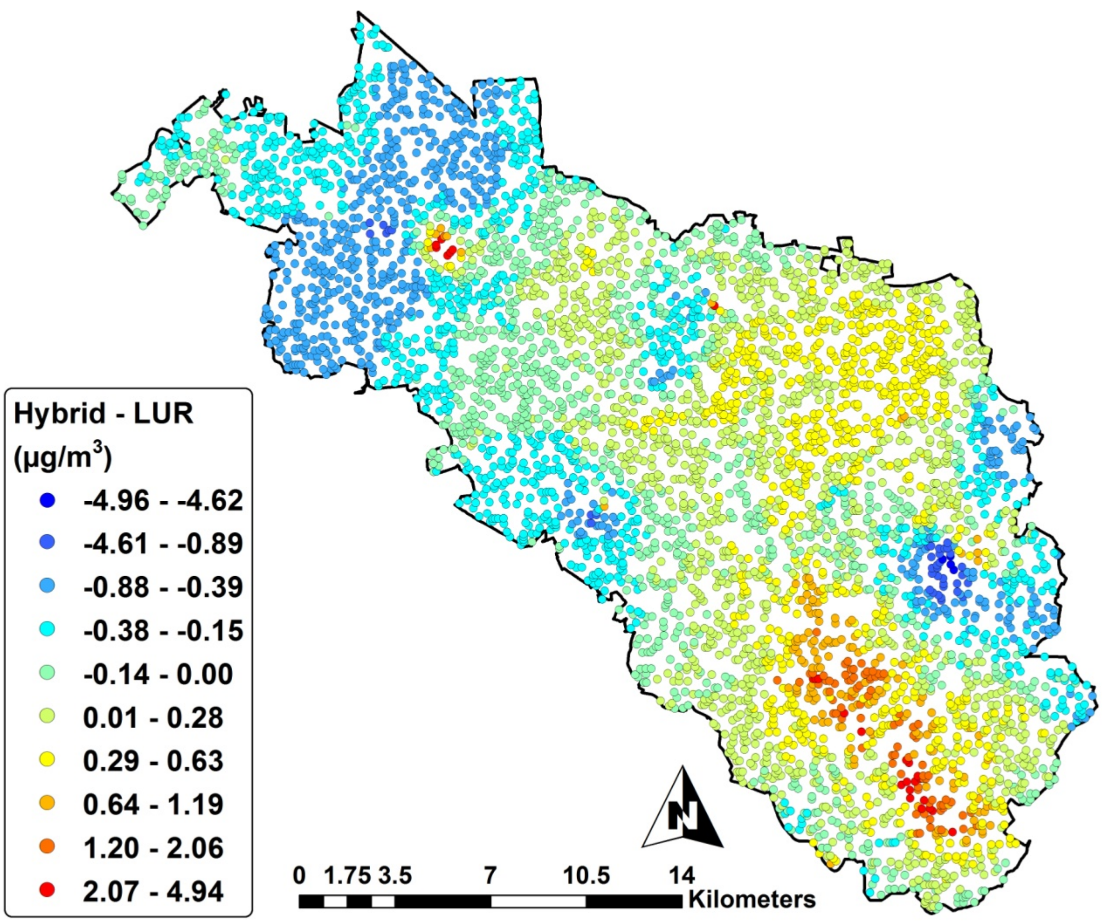

Figure 22. Predicted concentration difference (Hybrid minus LUR) defined at the residential level coordinates (latitude-longitude) from 2012 mean estimates

4.2.5 Daily Temporal Variability

LUR can produce robust predictions of long-term, fine-scale spatial variation in pollutant concentrations. Dispersion modeling, however, is capable of estimating fine-scale spatial resolution in addition to short-term averaging times. Fig. 23 exhibits differences by box-plots in daily exposure predictions for a 7-day week snapshot in January, 2013 at the 5,000 locations displayed in Fig. 22. Both models used the daily central AQS daily concentration to calibrate the daily exposure predictions. The differences in distributions between days (height of box-plots), indicates the differential prediction ability in AERMOD predictions, and indicates the impact of source-meteorological interaction information at small time scales. A maximum daily prediction difference of 16.47 was observed at a single location during the week snapshot. Generally, the two models estimated mean concentrations well across a relatively large, non-clustered cohort.

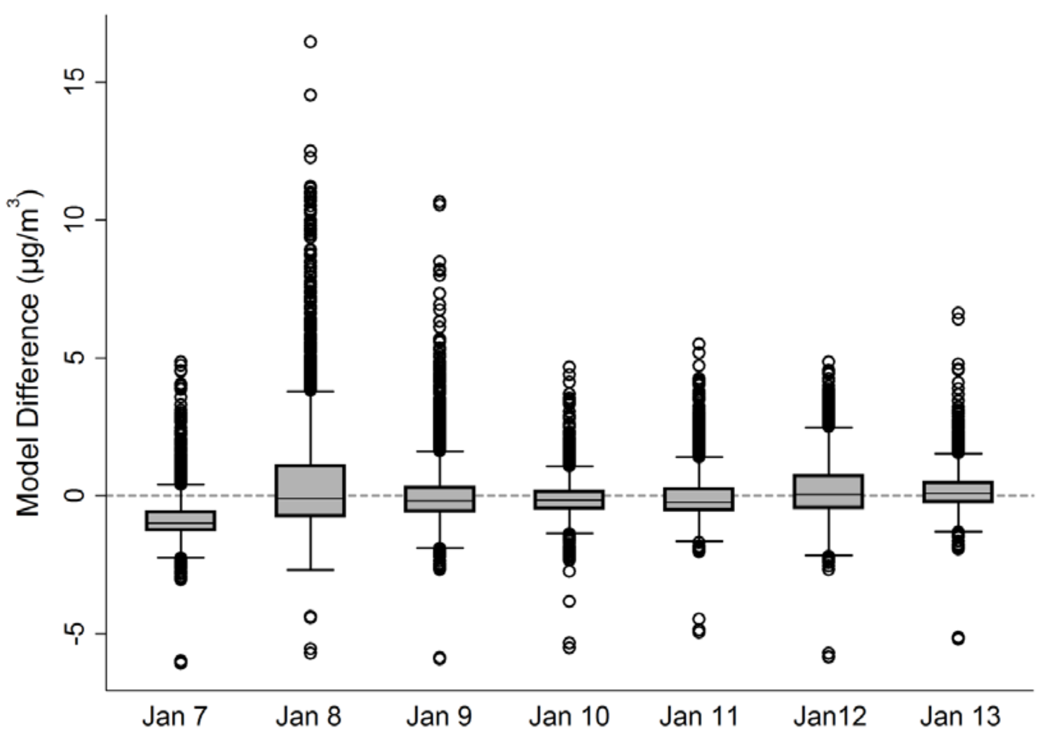

Figure 23. Difference in hybrid LUR/ AERMOD predictions and LUR predictions at the daily time scale

4.2.6 Model Simulation

Table 14 displays the results of the Monte Carlo simulation comparing two models that each contained distribution parameters from either the $PM_{2.5}$ density covariate, or the AERMOD covariate. The results of the simulations demonstrate the mathematical function of the geographic covariate variance and its resulting effect on a generic health effect estimate $\hat{\beta}_x$. Thus, a geographic covariate with a larger variance about the monitoring locations resulted in improved health effect estimate efficiency, though this relationship was not resolved by the model prediction accuracy as per mean coefficient of determination denoted by $\bar{R}^2$.

Table 14. Results from Monte Carlo simulations

Geographic Covariate	$\bar{R}^2$	SD $\hat{\alpha}_3$	Mean $\hat{\beta}_x$	SD $\hat{\beta}_x$
$PM_{2.5}$ Emissions Density (50m)	0.73	0.75	1.89	0.16
AERMOD 2012	0.74	0.28	1.99	0.11

4.3 DISCUSSION

In this chapter, we developed and evaluated an annual LUR model for $PM_{2.5}$, supplemented with yearly AERMOD $PM_{2.5}$ predictions and routine monitoring in Pittsburgh, PA in an attempt to enhance the spatial resolution of ambient air pollution data for long-term exposure estimation. We also demonstrated the utility of AERMOD with LUR for producing daily concentration estimates for acute exposure settings, and evaluated the model differences. These evaluations add to the limited number of studies that have compared spatial exposure techniques using real-world pollution measurements. Overall, the mean difference between models equated to a slight overestimation in LUR predictions compared to the hybrid model, though both models appear to estimate the underlying mean similarly. Though we only applied our model to a one weekly snapshot of daily predictions, these results indicate potential non-systematic differential predictions when including short-term AERMOD model output. However, we were unable to validate the daily estimates; nonetheless the daily estimates leverage AERMOD's temporal estimation flexibility and demonstrate a means to include meteorological processes for sources of interest.

We demonstrated that AERMOD can produce a highly variable prediction surface compared to typical GIS-based covariates across a large urban-to-suburban domain with relatively intense industrial sources. Notably, the $PM_{2.5}$ density variable was almost five times less variable ($\sigma^2 = 0.25$) across all 37 distributed monitoring locations, compared to variances of 1.18 and 1.45 for summer and winter AERMOD terms, respectively. In applying a quantitative comparison of exposure measurement error to a generic health outcome model, we were under the assumption that refining spatio-temporal resolution of exposure predictions would result in less exposure measurement error and less bias in estimating the health effect estimate. If exposure measurement error is non-differential with respect to a health outcome, a mis-specified exposure model containing error would result in bias towards the null hypothesis. Under this assumption, a properly specified exposure model with attenuated measurement error should result in less bias towards the null. Our simple Monte Carlo simulation demonstrated that the range in covariate values can theoretically impact exposure measurement error, and result in less bias towards the null, while improving efficiency. Moreover, prediction model accuracy assessed by the in-sample

R^2 value, may not provide adequate model evaluation conclusions. We acknowledge these results are based on an indirect means of examining exposure measurement error, and caveat our conclusions on health effect estimation as cursory.

Relatively few studies have explicitly compared LUR and dispersion models under epidemiological settings (Chang et al., 2012; de Hoogh et al., 2014; Sarnat et al., 2013; Sellier et al., 2014; Wu et al., 2011). Generally, higher correlations have been shown for traffic-related pollutants (e.g., NO_x, CO, $PM_{2.5}$ - EC) than for more regionally-varying pollutants (e.g., O_3, $PM_{2.5}$ - SO_4) (Sarnat et al., 2013; Sellier et al., 2014). Our attempt to model $PM_{2.5}$ was attempted given the presence of legacy industrial sources that exist in river valleys and emit pollutants near ground-level producing source-meteorological interaction events of interest.

Recently, Dionisio et al. (2013) produced refined spatial and temporal estimates of multiple pollutants using AERMOD predictions to disentangle regional background and localized spatio-temporal variability. In a complementary study, Sarnat et al. (2013) reported stronger heath effect estimate associations with the spatially-refined exposure metrics compared to less refined exposure techniques. Several simulation studies have been attempted to quantify exposure measurement error and related bias in the resulting risk assessment (Gryparis et al., 2009; Kim et al., 2009; Lopiano et al., 2010; Madsen et al., 2008; Szpiro et al., 2011b). These simulations have typically demonstrated that well specified spatial models and subsequent smoothing procedures produce very little bias in health effect estimates as measurement error in these contexts has a Berkson-like component as opposed to classical error.

Berkson error behaves similarly to the random ε in the disease model, where variance of the estimated coefficients in the health model increases, but is not biased (Szpiro et al., 2011b). Nonetheless, bi-directional health effect-bias was observed by Alexeeff et al. (2014) in comparisons of kriging and LUR models across various study design simulations. Basagaña et al. (2013) reported LUR associated measurement error and health-effect bias resulting from underpowered models (e.g., many predictor variables with few measurement sites: *n*=20,40,80). Therefore, potential for health estimate bias may result from spatial model misspecification, and ultimately how much of the *true spatial variability* is explained by the model which is ultimately unknown.

5.0 OVERALL SUMMARY

The objective of this dissertation was to examine the utility of incorporating source-meteorological interaction information from two commonly employed atmospheric dispersion models into the land use regression technique for both NO_2 and $PM_{2.5}$. Ultimately, we were interested in obtaining highly resolved spatio-temporal pollutant estimates using the popular LUR modeling framework, while providing a method to attenuate health effect estimate bias that may result from spatial model misspecification. We caveat our conclusions in respect to the diverse source regime within our study domain setting that is further confounded by complex topography and complex atmospheric processes. We also acknowledge that our temporally misaligned sampling design was not particular conducive for effective validation of our spatiotemporal deterministic modeling output. Our conclusions therefore are highly contingent upon internal cross validation measures and elementary mathematical deductions. While our simple hybrid methodology provided improved model predictions across our study domain, it is important to note that different exposure metrics apply to different aspects of air quality.

To investigate the efficacy of a hybrid land use regression/ atmospheric dispersion modeling framework, we began by examining output from a roadway dispersion output to predict NO_2 given the small-scale variability of NO_x. Our hybrid framework can more aptly be described as an LUR model supplemented by source-meteorological interaction information via Gaussian dispersion output from sources of interest. We simply added dispersion output as an independent covariate to pre-constructed LUR models. We attempted a validation of dispersion output from the Caline3 model that is shown in Appendix A, and observed robust correlations between measured and predictions, albeit appropriate background concentration derivation was not trivial.

The model framework described in chapter 2 helped to explain an additional portion of out-of-sample variation (3-10% LOOVC R^2) in NO_2 observations compared to the standard LUR model, Correspondingly, in Chapter 3, the AERMOD dispersion model was implemented to predict $PM_{2.5}$ from local and regional stationary sources in a similar hybrid framework. As per cross-validated R^2 and RMSE, AERMOD predictions and explained an additional 9-13% in out-of-sample variability in $PM_{2.5}$. Both dispersion models behaved similarly when added to the standard LUR models, effectively displacing GIS-based covariates, corroborating model interpretability and providing the greatest degree of model fitness for nearby, high-density source categories.

In the absence of a spatially dense monitoring network, we demonstrated that AERMOD can produce a highly variable prediction surface compared to typical GIS-based covariates across a large urban-to-suburban domain with relatively intense industrial sources. Our simple Monte Carlo simulation demonstrates that the range in covariate values can impact exposure measurement error in epidemiological studies, and prediction model accuracy assessed by the in-sample R^2 value, may not provide adequate model evaluation conclusions. We acknowledge these results are based on an indirect means of examining exposure measurement error, and caveat our conclusions on health effect estimation as preliminary. We intend to further investigate the assumption that spatiotemporally refined exposure predictions result in attenuated health effect bias by association-type epidemiological study.

APPENDIX: OBSERVED NO_2 VS. PREDICTED CALINE3 + BACKGROUND

Across distributed sites, Caline3 predictions stratified by low- and high-traffic sites produced means of 1.73 μg/m^3 (SD = 1.68, n = 74) and 4.63 μg/m^3 (SD = 3.54, n = 70), respectively. Figs. 24-26 display winter season scatter-plots of log-transformed measured NO_2 vs. modeled Caline3 added to: (a) regional background; (b) urban reference; and (c) mean of regional background & urban reference. Caline3 + regional background under-predicted measured NO_2 by 5.78 ppb, on average. From the geometric mean (m_g) values, Caline3 + regional background under-predicted measured NO_2 across both seasons. Conversely, Caline3 + urban reference over-predicted measured NO_2. Caline3 + mean reference produced the lowest geometric means, standard deviations and fractional bias values. Therefore, Caline3 + mean reference produced the least biased estimates of NO_2 across winter seasons, compared to either continuous site alone (Fig 26). A mean of both temporal measures was subsequently chosen to temporally control for misaligned measures in all LUR models predicting NO_2.

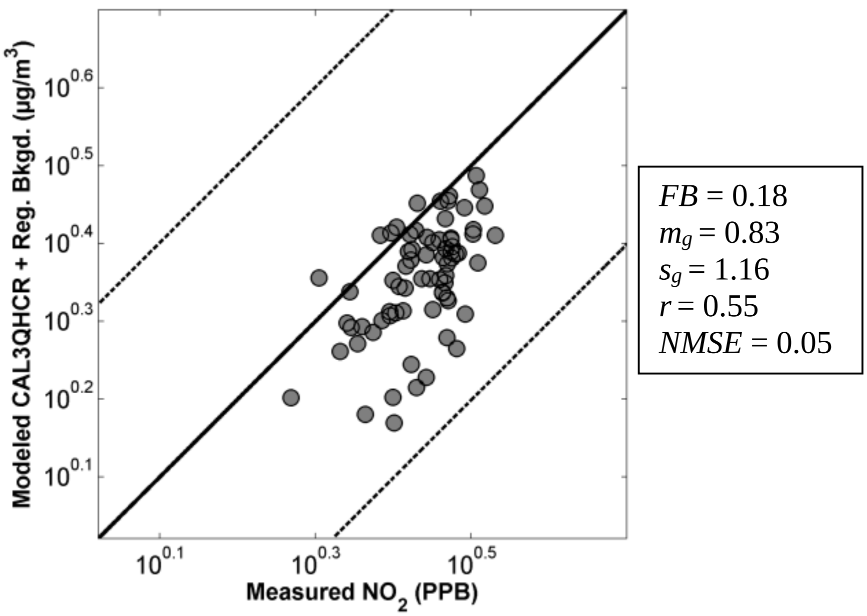

Figure 24. log-transformed scatter plot of measured NO_2 vs. Caline3 + regional background site measurements as background concentration with performance statistics

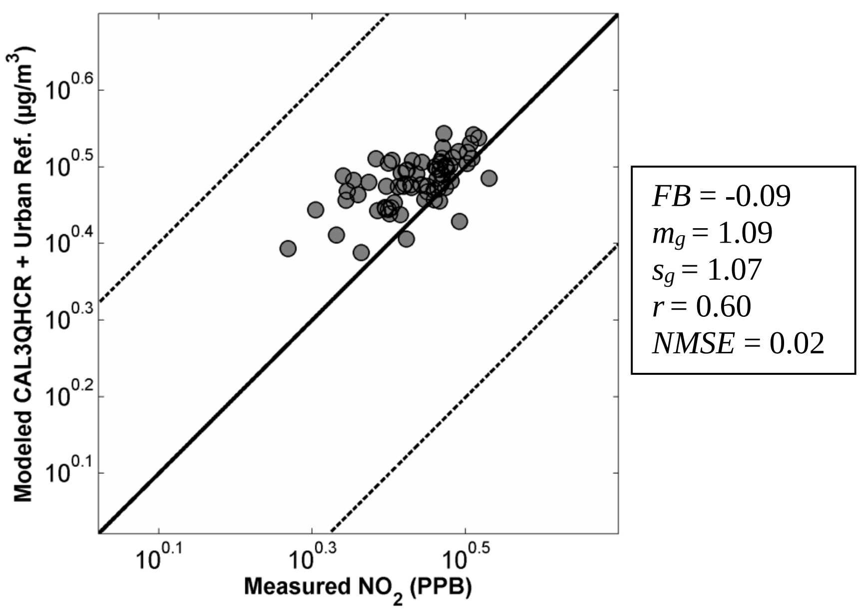

Figure 25. log-transformed scatter plot of measured NO_2 vs. Caline3 + urban reference site measurements as background concentration with performance statistics

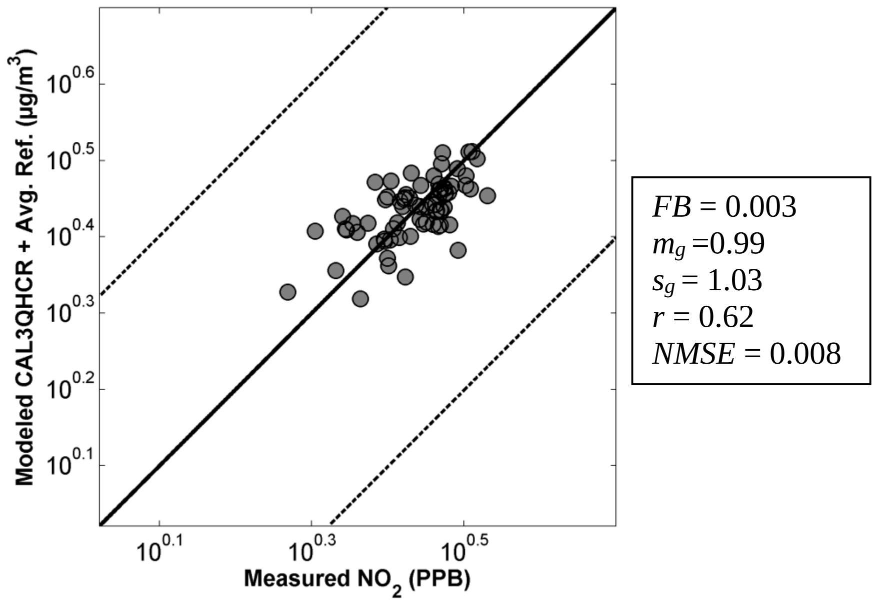

Figure 26. log-transformed scatter plot of measured NO_2 vs. Caline3 + mean of regional background & urban reference site measurements as background concentration with performance statistics

BIBLIOGRAPHY

Ainslie, B., Steyn, D., Su, J., Buzzelli, M., Brauer, M., Larson, T., Rucker, M., 2008. A source area model incorporating simplified atmospheric dispersion and advection at fine scale for population air pollutant exposure assessment. Atmospheric Environment 42, 2394-2404.

Alexeeff, S.E., Schwartz, J., Kloog, I., Chudnovsky, A., Koutrakis, P., Coull, B.A., 2014. Consequences of kriging and land use regression for PM2. 5 predictions in epidemiologic analyses: insights into spatial variability using high-resolution satellite data. Journal of Exposure Science and Environmental Epidemiology.

Allen, R.W., Amram, O., Wheeler, A.J., Brauer, M., 2011. The transferability of NO and NO_2 land use regression models between cities and pollutants. Atmospheric Environment 45, 369-378.

Anderson, J.O., Thundiyil, J.G., Stolbach, A., 2012. Clearing the air: a review of the effects of particulate matter air pollution on human health. Journal of Medical Toxicology 8, 166-175.

Arain, M., Blair, R., Finkelstein, N., Brook, J., Sahsuvaroglu, T., Beckerman, B., Zhang, L., Jerrett, M., 2007. The use of wind fields in a land use regression model to predict air pollution concentrations for health exposure studies. Atmospheric Environment 41, 3453-3464.

Araujo, J.A., 2011. Particulate air pollution, systemic oxidative stress, inflammation, and atherosclerosis. Air Quality, Atmosphere & Health 4, 79-93.

Arunachalam, S., Valencia, A., Akita, Y., Serre, M.L., Omary, M., Garcia, V., Isakov, V., 2014. A Method for Estimating Urban Background Concentrations in Support of Hybrid Air Pollution Modeling for Environmental Health Studies. International Journal of Environmental Research and Public Health 11, 10518-10536.

Baccarelli, A., Martinelli, I., Pegoraro, V., Melly, S., Grillo, P., Zanobetti, A., Hou, L., Bertazzi, P.A., Mannucci, P.M., Schwartz, J., 2009. Living near major traffic roads and risk of deep vein thrombosis. Circulation 119, 3118-3124.

Barratt, R., 2013. Atmospheric dispersion modelling: an introduction to practical applications. Routledge.

Basagaña, X., Aguilera, I., Rivera, M., Agis, D., Foraster, M., Marrugat, J., Elosua, R., Künzli, N., 2013. Measurement error in epidemiologic studies of air pollution based on land-use regression models. American journal of epidemiology 178, 1342-1346.

Basagaña, X., Rivera, M., Aguilera, I., Agis, D., Bouso, L., Elosua, R., Foraster, M., de Nazelle, A., Nieuwenhuijsen, M., Vila, J., 2012. Effect of the number of measurement sites on land use regression models in estimating local air pollution. Atmospheric Environment 54, 634-642.

Bekhor, S., Broday, D.M., 2013. Data-driven nonlinear optimisation of a simple air pollution dispersion model generating high resolution spatiotemporal exposure. Atmospheric Environment 79, 261-270.

Bell, M.L., 2006. The use of ambient air quality modeling to estimate individual and population exposure for human health research: a case study of ozone in the Northern Georgia Region of the United States. Environment international 32, 586-593.

Bell, M.L., Zanobetti, A., Dominici, F., 2013. Evidence on vulnerability and susceptibility to health risks associated with short-term exposure to particulate matter: a systematic review and meta-analysis. American journal of epidemiology, kwt090.

Benson, P., Baishiki, R., 1980. CALINE3-A versatile dispersion model for predicting air pollutant levels near highways and arterial streets.

Benson, P.E., 1992. A review of the development and application of the Caline3 and 4 models. Atmospheric Environment. Part B. Urban Atmosphere 26, 379-390.

Bernstein, J.A., Alexis, N., Barnes, C., Bernstein, I.L., Nel, A., Peden, D., Diaz-Sanchez, D., Tarlo, S.M., Williams, P.B., Bernstein, J.A., 2004. Health effects of air pollution. Journal of Allergy and Clinical Immunology 114, 1116-1123.

Brauer, M., Hoek, G., van Vliet, P., Meliefste, K., Fischer, P., Gehring, U., Heinrich, J., Cyrys, J., Bellander, T., Lewne, M., 2003. Estimating long-term average particulate air pollution concentrations: application of traffic indicators and geographic information systems. Epidemiology 14, 228-239.

Brender, J.D., Maantay, J.A., Chakraborty, J., 2011. Residential proximity to environmental hazards and adverse health outcomes. Journal Information 101.

Briant, R., Seigneur, C., Gadrat, M., Bugajny, C., 2013. Evaluation of roadway Gaussian plume models with large-scale measurement campaigns. Geoscientific Model Development 6, 445-456.

Briggs, D.J., Collins, S., Elliott, P., Fischer, P., Kingham, S., Lebret, E., Pryl, K., van Reeuwijk, H., Smallbone, K., Van Der Veen, A., 1997. Mapping urban air pollution using GIS: a regression-based approach. International Journal of Geographical Information Science 11, 699-718.

Brunekreef, B., Holgate, S.T., 2002. Air pollution and health. The lancet 360, 1233-1242.

Chang, H.H., Hu, X., Liu, Y., 2013. Calibrating MODIS aerosol optical depth for predicting daily PM2. 5 concentrations via statistical downscaling. Journal of Exposure Science and Environmental Epidemiology.

Chang, H.H., Reich, B.J., Miranda, M.L., 2012. Time-to-event analysis of fine particle air pollution and preterm birth: results from North Carolina, 2001–2005. American journal of epidemiology 175, 91-98.

Chang, J., Hanna, S., 2004. Air quality model performance evaluation. Meteorology and Atmospheric Physics 87, 167-196.

Chudnovsky, A.A., Kostinski, A., Lyapustin, A., Koutrakis, P., 2013. Spatial scales of pollution from variable resolution satellite imaging. Environmental Pollution 172, 131-138.

Cimorelli, A.J., Perry, S.G., Venkatram, A., Weil, J.C., Paine, R.J., Wilson, R.B., Lee, R.F., Peters, W.D., Brode, R.W., 2005. AERMOD: A dispersion model for industrial source applications. Part I: General model formulation and boundary layer characterization. Journal of Applied Meteorology 44, 682-693.

Clougherty, J.E., Houseman, E.A., Levy, J.I., 2009. Examining intra-urban variation in fine particle mass constituents using GIS and constrained factor analysis. Atmospheric Environment 43, 5545-5555.

Clougherty, J.E., Kheirbek, I., Eisl, H.M., Ross, Z., Pezeshki, G., Gorczynski, J.E., Johnson, S., Markowitz, S., Kass, D., Matte, T., 2013a. Intra-urban spatial variability in wintertime street-level concentrations of multiple combustion-related air pollutants: the New York City Community Air Survey (NYCCAS). Journal of exposure science & environmental epidemiology 23, 232-240.

Clougherty, J.E., Kheirbek, I., Eisl, H.M., Ross, Z., Pezeshki, G., Gorczynski, J.E., Johnson, S., Markowitz, S., Kass, D., Matte, T., 2013b. Intra-urban spatial variability in wintertime street-level concentrations of multiple combustion-related air pollutants: The New York City Community Air Survey (NYCCAS). Journal of Exposure Science and Environmental Epidemiology.

Clougherty, J.E., Wright, R.J., Baxter, L.K., Levy, J.I., 2008. Land use regression modeling of intra-urban residential variability in multiple traffic-related air pollutants. Environ Health 7, 17.

Cohen, A.J., Ross Anderson, H., Ostro, B., Pandey, K.D., Krzyzanowski, M., Künzli, N., Gutschmidt, K., Pope, A., Romieu, I., Samet, J.M., 2005. The global burden of disease due to outdoor air pollution. Journal of Toxicology and Environmental Health, Part A 68, 1301-1307.

Cook, R., Isakov, V., Touma, J.S., Benjey, W., Thurman, J., Kinnee, E., Ensley, D., 2008. Resolving local-scale emissions for modeling air quality near roadways. Journal of the Air & Waste Management Association 58, 451-461.

de Hoogh, K., Korek, M., Vienneau, D., Keuken, M., Kukkonen, J., Nieuwenhuijsen, M.J., Badaloni, C., Beelen, R., Bolignano, A., Cesaroni, G., 2014. Comparing land use regression and dispersion modelling to assess residential exposure to ambient air pollution for epidemiological studies. Environment international 73, 382-392.

Dijkema, M.B., Gehring, U., van Strien, R.T., van der Zee, S.C., Fischer, P., Hoek, G., Brunekreef, B., 2011. A comparison of different approaches to estimate small-scale spatial variation in outdoor NO2 concentrations. Environmental health perspectives 119, 670.

Dionisio, K.L., Baxter, L.K., Chang, H.H., 2014. An Empirical Assessment of Exposure Measurement Error and Effect Attenuation in Bipollutant Epidemiologic Models. Environ Health Perspect.

Dionisio, K.L., Isakov, V., Baxter, L.K., Sarnat, J.A., Sarnat, S.E., Burke, J., Rosenbaum, A., Graham, S.E., Cook, R., Mulholland, J., 2013. Development and evaluation of alternative approaches for exposure assessment of multiple air pollutants in Atlanta, Georgia. Journal of Exposure Science and Environmental Epidemiology.

Dockery, D.W., 2009. Health effects of particulate air pollution. Annals of epidemiology 19, 257-263.

Eckhoff, P.A., Braverman, T.N., 1995. Addendum to the user’s guide to CAL3QHC version 2.0 (CAL3QHCR user’s guide). Technical Support Division, Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina.

Faustini, A., Rapp, R., Forastiere, F., 2014. Nitrogen dioxide and mortality: review and meta-analysis of long-term studies. European Respiratory Journal, erj01147-02013.

Gilbert, N.L., Goldberg, M.S., Beckerman, B., Brook, J.R., Jerrett, M., 2005. Assessing spatial variability of ambient nitrogen dioxide in Montreal, Canada, with a land-use regression model. Journal of the Air & Waste Management Association 55, 1059-1063.

Gryparis, A., Paciorek, C.J., Zeka, A., Schwartz, J., Coull, B.A., 2009. Measurement error caused by spatial misalignment in environmental epidemiology. Biostatistics 10, 258-274.

Gulliver, J., Briggs, D., 2011. STEMS-Air: A simple GIS-based air pollution dispersion model for city-wide exposure assessment. Science of the total environment 409, 2419-2429.

HEI, 2010. Traffic-related air pollution: a critical review of the literature on emissions, exposure, and health effects. Health Effects Institute.
Heilig, G.K., 2012. World urbanization prospects the 2011 revision. United Nations, Department of Economic and Social Affairs (DESA), Population Division, Population Estimates and Projections Section, New York.
Henderson, S.B., Beckerman, B., Jerrett, M., Brauer, M., 2007. Application of land use regression to estimate long-term concentrations of traffic-related nitrogen oxides and fine particulate matter. Environmental science & technology 41, 2422-2428.
Hinds, W.C., 2012. Aerosol technology: properties, behavior, and measurement of airborne particles. John Wiley & Sons.
Hoek, G., Beelen, R., de Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P., Briggs, D., 2008. A review of land-use regression models to assess spatial variation of outdoor air pollution. Atmospheric Environment 42, 7561-7578.
Hoek, G., Brunekreef, B., Goldbohm, S., Fischer, P., van den Brandt, P.A., 2002. Association between mortality and indicators of traffic-related air pollution in the Netherlands: a cohort study. The lancet 360, 1203-1209.
Hoek, G., Krishnan, R.M., Beelen, R., Peters, A., Ostro, B., Brunekreef, B., Kaufman, J.D., 2013. Long-term air pollution exposure and cardio-respiratory mortality: a review. Environ Health 12, 43.
Holland, W.W., Bennett, A., Cameron, I., Florey, C.d.V., Leeder, S., Schilling, R., Swan, A., Waller, R., 1979. Health effects of particulate pollution: reappraising the evidence. American Journal of Epidemiology 110, 527-527.
Huttunen, K., Siponen, T., Salonen, I., Yli-Tuomi, T., Aurela, M., Dufva, H., Hillamo, R., Linkola, E., Pekkanen, J., Pennanen, A., 2012. Low-level exposure to ambient particulate matter is associated with systemic inflammation in ischemic heart disease patients. Environmental research 116, 44-51.
Isakov, V., Irwin, J.S., Ching, J., 2007. Using CMAQ for exposure modeling and characterizing the subgrid variability for exposure estimates. Journal of Applied Meteorology and Climatology 46, 1354-1371.
Isakov, V., Touma, J.S., Burke, J., Lobdell, D.T., Palma, T., Rosenbaum, A., kÖzkaynak, H., 2009. Combining regional-and local-scale air quality models with exposure models for use in environmental health studies. Journal of the Air & Waste Management Association 59, 461-472.
Jerrett, M., Arain, A., Kanaroglou, P., Beckerman, B., Potoglou, D., Sahsuvaroglu, T., Morrison, J., Giovis, C., 2005. A review and evaluation of intraurban air pollution exposure models. Journal of Exposure Science and Environmental Epidemiology 15, 185-204.
Jerrett, M., Arain, M., Kanaroglou, P., Beckerman, B., Crouse, D., Gilbert, N., Brook, J., Finkelstein, N., Finkelstein, M., 2007. Modeling the intraurban variability of ambient traffic pollution in Toronto, Canada. Journal of Toxicology and Environmental Health, Part A 70, 200-212.
Jerrett, M., Burnett, R.T., Kanaroglou, P., Eyles, J., Finkelstein, N., Giovis, C., Brook, J.R., 2001. A GIS-environmental justice analysis of particulate air pollution in Hamilton, Canada. Environment and Planning A 33, 955-974.
Johnson, M., Isakov, V., Touma, J., Mukerjee, S., Özkaynak, H., . 2010. Evaluation of land-use regression models used to predict air quality concentrations in an urban area. Atmospheric Environment 44, 3660-3668.

Kampa, M., Castanas, E., 2008. Human health effects of air pollution. Environmental pollution 151, 362-367.

Karner, A.A., Eisinger, D.S., Niemeier, D.A., 2010. Near-roadway air quality: synthesizing the findings from real-world data. Environmental science & technology 44, 5334-5344.

Keller, J.P., Olives, C., Kim, S.-Y., Sheppard, L., Sampson, P.D., Szpiro, A.A., Oron, A.P., Lindström, J., Vedal, S., Kaufman, J.D., 2014. A Unified Spatiotemporal Modeling Approach for Predicting Concentrations of Multiple Air Pollutants in the Multi-Ethnic Study of Atherosclerosis and Air Pollution. Environ Health Perspect.

Kheirbek, I., Johnson, S., Ross, Z., Pezeshki, G., Ito, K., Eisl, H., Matte, T., 2012. Spatial variability in levels of benzene, formaldehyde, and total benzene, toluene, ethylbenzene and xylenes in New York City: a land-use regression study. Environ Health 11, 51.

Kim, M., Zhang, X., Holt, J.B., Liu, Y., 2013. Spatio-Temporal Variations in the Associations between Hourly PM 2.5 and Aerosol Optical Depth (AOD) from MODIS Sensors on Terra and Aqua. Health 2013.

Kim, S.-Y., Sheppard, L., Kim, H., 2009. Health effects of long-term air pollution: influence of exposure prediction methods. Epidemiology 20, 442-450.

Kinnee, E., Touma, J., Mason, R., Thurman, J., Beidler, A., Bailey, C., Cook, R., 2004. Allocation of onroad mobile emissions to road segments for air toxics modeling in an urban area. Transportation Research Part D: Transport and Environment 9, 139-150.

Kloog, I., Chudnovsky, A., Just, A., Nordio, F., Koutrakis, P., Coull, B.A., Lyapustin, A., Wang, Y., Schwartz, J., 2014. A new hybrid spatio-temporal model for estimating daily multi-year PM2.5 concentrations across northeastern USA using high resolution aerosol optical depth data. Atmospheric Environment 95, 581-590.

Kloog, I., Nordio, F., Coull, B.A., Schwartz, J., 2012. Incorporating local land use regression and satellite aerosol optical depth in a hybrid model of spatiotemporal $PM_{2.5}$ exposures in the Mid-Atlantic states. Environmental science & technology 46, 11913-11921.

KuÈnzli, N., Kaiser, R., Medina, S., Studnicka, M., Chanel, O., Filliger, P., Herry, M., Horak Jr, F., Puybonnieux-Texier, V., Quenel, P., 2000. Public-health impact of outdoor and traffic-related air pollution: a European assessment. The Lancet 356, 795-801.

Künzli, N., Jerrett, M., Mack, W.J., Beckerman, B., LaBree, L., Gilliland, F., Thomas, D., Peters, J., Hodis, H.N., 2005. Ambient air pollution and atherosclerosis in Los Angeles. Environmental health perspectives, 201-206.

Lee, H., Liu, Y., Coull, B., Schwartz, J., Koutrakis, P., 2011. A novel calibration approach of MODIS AOD data to predict $PM_{2.5}$ concentrations. Atmos. Chem. Phys 11, 7991-8002.

Levy, R.C., Remer, L.A., Mattoo, S., Vermote, E.F., Kaufman, Y.J., 2007. Second-generation operational algorithm: Retrieval of aerosol properties over land from inversion of Moderate Resolution Imaging Spectroradiometer spectral reflectance. Journal of Geophysical Research: Atmospheres (1984–2012) 112.

Lin, G., Fu, J., Jiang, D., Hu, W., Dong, D., Huang, Y., Zhao, M., 2013. Spatio-temporal variation of PM2. 5 concentrations and their relationship with geographic and socioeconomic factors in China. International journal of environmental research and public health 11, 173-186.

Lindström, J., Szpiro, A.A., Sampson, P.D., Oron, A.P., Richards, M., Larson, T.V., Sheppard, L., 2013. A flexible spatio-temporal model for air pollution with spatial and spatio-temporal covariates. Environmental and Ecological Statistics, 1-23.

Lopiano, K.K., Young, L.J., Gotway, C.A., 2010. A comparison of errors in variables methods for use in regression models with spatially misaligned data. Statistical methods in medical research.

Madsen, C., Gehring, U., Håberg, S.E., Nafstad, P., Meliefste, K., Nystad, W., Lødrup Carlsen, K.C., Brunekreef, B., 2011. Comparison of land-use regression models for predicting spatial NO_x contrasts over a three year period in Oslo, Norway. Atmospheric Environment 45, 3576-3583.

Madsen, L., Ruppert, D., Altman, N., 2008. Regression with spatially misaligned data. Environmetrics 19, 453-467.

Maheswaran, R., Elliott, P., 2003. Stroke mortality associated with living near main roads in England and Wales a geographical study. Stroke 34, 2776-2780.

Marshall, J.D., Nethery, E., Brauer, M., 2008. Within-urban variability in ambient air pollution: comparison of estimation methods. Atmospheric Environment 42, 1359-1369.

Matus, K., Nam, K.-M., Selin, N.E., Lamsal, L.N., Reilly, J.M., Paltsev, S., 2012. Health damages from air pollution in China. Global environmental change 22, 55-66.

Mavko, M.E., Tang, B., George, L.A., 2008. A sub-neighborhood scale land use regression model for predicting NO_2. Science of the Total Environment 398, 68-75.

Mölter, A., Lindley, S., de Vocht, F., Simpson, A., Agius, R., 2010a. Modelling air pollution for epidemiologic research–Part II: Predicting temporal variation through land use regression. Science of the total environment 409, 211-217.

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

Nafstad, P., Håheim, L., Oftedal, B., Gram, F., Holme, I., Hjermann, I., Leren, P., 2003. Lung cancer and air pollution: a 27 year follow up of 16 209 Norwegian men. Thorax 58, 1071-1076.

Nordio, F., Kloog, I., Coull, B.A., Chudnovsky, A., Grillo, P., Bertazzi, P.A., Baccarelli, A.A., Schwartz, J., 2013. Estimating spatio-temporal resolved PM_{10} aerosol mass concentrations using MODIS satellite data and land use regression over Lombardy, Italy. Atmospheric Environment 74, 227-236.

Nyberg, F., Gustavsson, P., Järup, L., Bellander, T., Berglind, N., Jakobsson, R., Pershagen, G., 2000. Urban air pollution and lung cancer in Stockholm. Epidemiology 11, 487-495.

Pennsylvania Department of Transportation, B.o.P.a.R., Geographic Information Division, 2013. PennDOT - Pennsylvania Stateroads Pennsylvania Department of Transportation, Harrisburg, PA.

Pope, C.A., 2000. Invited commentary: particulate matter-mortality exposure-response relations and threshold. American Journal of Epidemiology 152, 407-412.

Pope III, C.A., 2000. Review: epidemiological basis for particulate air pollution health standards. Aerosol Science & Technology 32, 4-14.

Pope III, C.A., Dockery, D.W., 2006. Health effects of fine particulate air pollution: lines that connect. Journal of the Air & Waste Management Association 56, 709-742.

Poplawski, K., Gould, T., Setton, E., Allen, R., Su, J., Larson, T., Henderson, S., Brauer, M., Hystad, P., Lightowlers, C., 2008. Intercity transferability of land use regression models for estimating ambient concentrations of nitrogen dioxide. Journal of exposure science and environmental epidemiology 19, 107-117.

Ristic, B., Gunatilaka, A., Gailis, R., 2014. Achievable accuracy in parameter estimation of a Gaussian plume dispersion model, Statistical Signal Processing (SSP), 2014 IEEE Workshop on. IEEE, pp. 209-212.

Ross, Z., English, P.B., Scalf, R., Gunier, R., Smorodinsky, S., Wall, S., Jerrett, M., 2006. Nitrogen dioxide prediction in Southern California using land use regression modeling: potential for environmental health analyses. Journal of Exposure Science and Environmental Epidemiology 16, 106-114.

Rückerl, R., Schneider, A., Breitner, S., Cyrys, J., Peters, A., 2011. Health effects of particulate air pollution: a review of epidemiological evidence. Inhalation toxicology 23, 555-592.

Sahu, S.K., Mardia, K.V., 2005. A Bayesian kriged Kalman model for short-term forecasting of air pollution levels. Journal of the Royal Statistical Society: Series C (Applied Statistics) 54, 223-244.

Samet, J.M., Zeger, S.L., Dominici, F., Curriero, F., Coursac, I., Dockery, D.W., Schwartz, J., Zanobetti, A., 2000. The national morbidity, mortality, and air pollution study. Part II: morbidity and mortality from air pollution in the United States Res Rep Health Eff Inst 94, 5-79.

Sarnat, S.E., Sarnat, J.A., Mulholland, J., Isakov, V., Özkaynak, H., Chang, H.H., Klein, M., Tolbert, P.E., 2013. Application of alternative spatiotemporal metrics of ambient air pollution exposure in a time-series epidemiological study in Atlanta. Journal of Exposure Science and Environmental Epidemiology 23, 593-605.

Scire, J.S., Strimaitis, D.G., Yamartino, R.J., 1990. Model formulation and user's guide for the CALPUFF dispersion model. Sigma Research Corp., Concord, MA.

Seinfeld, J.H., Pandis, S.N., 2012. Atmospheric chemistry and physics: from air pollution to climate change. John Wiley & Sons.

Sellier, Y., Galineau, J., Hulin, A., Caini, F., Marquis, N., Navel, V., Bottagisi, S., Giorgis-Allemand, L., Jacquier, C., Slama, R., 2014. Health effects of ambient air pollution: do different methods for estimating exposure lead to different results. Environ. Int 66, 165-173.

Shmool, J.L., Michanowicz, D.R., Cambal, L., Tunno, B., Howell, J., Gillooly, S., Roper, C., Tripathy, S., Chubb, L.G., Eisl, H.M., 2014. Saturation sampling for spatial variation in multiple air pollutants across an inversion-prone metropolitan area of complex terrain. Environmental Health 13, 28.

Snyder, W.H., Thompson, R.S., Eskridge, R.E., Lawson, R.E., Castro, I.P., Lee, J., Hunt, J.C., Ogawa, Y., 1985. The structure of strongly stratified flow over hills: dividing-streamline concept. Journal of Fluid Mechanics 152, 249-288.

Spengler, J.D., Sexton, K., 1983. Indoor air pollution: a public health perspective. Science 221, 9-17.

Su, J.G., Brauer, M., Ainslie, B., Steyn, D., Larson, T., Buzzelli, M., 2008. An innovative land use regression model incorporating meteorology for exposure analysis. Science of the total environment 390, 520-529.

Su, J.G., Jerrett, M., Beckerman, B., Wilhelm, M., Ghosh, J.K., Ritz, B., 2009. Predicting traffic-related air pollution in Los Angeles using a distance decay regression selection strategy. Environmental research 109, 657-670.

Szpiro, A.A., Paciorek, C.J., Sheppard, L., 2011a. Does more accurate exposure prediction necessarily improve health effect estimates? Epidemiology (Cambridge, Mass.) 22, 680.

Szpiro, A.A., Sheppard, L., Lumley, T., 2011b. Efficient measurement error correction with spatially misaligned data. Biostatistics, kxq083.

USEPA, 2005. Revision to the Guideline on Air Quality Models: Adoption of a Preferred Gweneral Purpose (Flat and Complex Terrain) Dispersion Model and Other Revisions; Final Rule in: United States Environmental Protection Agency, 40 CFR Part 51, Washington D.C.

USEPA, 2010. Motor Vehicle Emissions Simulator. MOVES 2010a User Guide, in: Agency, E.P. (Ed.). Office of Transportation and Air Quality.

Van den Hooven, E.H., Pierik, F.H., Van Ratingen, S.W., Zandveld, P.Y., Meijer, E.W., Hofman, A., Miedema, H.M., Jaddoe, V.W., De Kluizenaar, Y., 2012. Air pollution exposure estimation using dispersion modelling and continuous monitoring data in a prospective birth cohort study in the Netherlands. Environmental Health 11, 1-11.

Van Roosbroeck, S., Jacobs, J., Janssen, N.A., Oldenwening, M., Hoek, G., Brunekreef, B., 2007. Long-term personal exposure to $PM_{2.5}$, soot and NO_x in children attending schools located near busy roads, a validation study. Atmospheric Environment 41, 3381-3394.

Vienneau, D., De Hoogh, K., Beelen, R., Fischer, P., Hoek, G., Briggs, D., 2010. Comparison of land-use regression models between Great Britain and the Netherlands. Atmospheric Environment 44, 688-696.

Vienneau, D., De Hoogh, K., Briggs, D., 2009. A GIS-based method for modelling air pollution exposures across Europe. Science of the Total Environment 408, 255-266.

Wang, M., Beelen, R., Stafoggia, M., Raaschou-Nielsen, O., Andersen, Z.J., Hoffmann, B., Fischer, P., Houthuijs, D., Nieuwenhuijsen, M., Weinmayr, G., 2014. Long-term exposure to elemental constituents of particulate matter and cardiovascular mortality in 19 European cohorts: Results from the ESCAPE and TRANSPHORM projects. Environment international 66, 97-106.

Whitlow, T.H., Hall, A., Zhang, K.M., Anguita, J., 2011. Impact of local traffic exclusion on near-road air quality: Findings from the New York City “Summer Streets” campaign. Environmental Pollution 159, 2016-2027.

Wilton, D., Szpiro, A., Gould, T., Larson, T., 2010. Improving spatial concentration estimates for nitrogen oxides using a hybrid meteorological dispersion/land use regression model in Los Angeles, CA and Seattle, WA. Science of the Total Environment 408, 1120-1130.

Wilton, D.C., 2011. Modelling Nitrogen Oxides in Los Angeles Using a Hybrid Dispersion/Land Use Regression Model, Civil and Environmental Engineering. University of Washington, p. 128.

Wong, D.W., Yuan, L., Perlin, S.A., 2004. Comparison of spatial interpolation methods for the estimation of air quality data. Journal of Exposure Science and Environmental Epidemiology 14, 404-415.

World Health Organization, 2012. Ambient Air Quality and Health Fact Sheet No. 313. Accessed on October 12, 2014.

Wu, J., Wilhelm, M., Chung, J., Ritz, B., 2011. Comparing exposure assessment methods for traffic-related air pollution in an adverse pregnancy outcome study. Environmental research 111, 685-692.