GARCIA, VALERIE COVER. Examination of Inter-relationships among Atmospheric Transport Patterns, Ozone Concentrations, and Human Health Endpoints in New York State. (Under the direction of Dr. Heather M. Cheshire and Dr. Viney Aneja).

In its mission to protect human health and the environment, the United States (U.S.) Environmental Protection Agency (EPA) implemented the NO\textsubscript{x} Budget Trading Program (NBP) to reduce the emissions of nitrogen oxides (NO\textsubscript{x}) in the Eastern U.S., with the intent of reducing ambient concentrations of both NO\textsubscript{x} and the secondarily-formed ozone chemicals. These pollutants and their precursors can be transported downwind, contributing to pollutant levels at locations much farther from the emission sources, potentially impacting human health in downwind areas. This study investigates the health risks in New York State (NYS) from exposure to polluted air parcels transported from the Midwest. Back-trajectories are performed from several sites within NYS for ten summers (June, July and August from 1997 to 2006) to identify days that the air parcel passed through the Ohio River Valley (ORV) region within 48 hours back in time. The ORV zone is defined as a boundary encompassing relatively high-emitting power plants in the Midwest and is used as an indicator variable to represent the transport of ozone from this area into NYS in an epidemiology analysis. The ORV zone variable and the daily maximum 8-hr average ozone concentrations are then used as the main health effects in a Generalized Additive Model (GAM) to investigate potential associations between these two variables and respiratory-related hospital admissions. The results of the analysis indicate that the risk of being admitted to the hospital for a respiratory-related illness on days that air parcels are transported from the Midwest is elevated in NYS sub-regions 2, 3 and 6. In addition, the risk of respiratory-related hospital admission from exposure to ozone is elevated in Regions 2 and 4. Two time periods before the implementation of the NBP and after the implementation of the NBP (summers of 1997 – 2000 and 2004 – 2006, respectively) were also examined, but an analysis of the number of summers needed to conduct the analysis revealed that there are not enough years of data to discern a difference in the health signal between the two time periods. According to this analysis, the health effect measures become inconsistent (as
compared to applying the model using the data from all 10 summers) at about 5 summers (552 days) or less.
Examination of Inter-relationships among Atmospheric Transport Patterns, Ozone Concentrations, and Human Health Endpoints in New York State

by

Valerie Cover Garcia

A thesis submitted to the Graduate Faculty of North Carolina State University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Forestry & Environmental Resources

Raleigh, North Carolina

2010

APPROVED BY:

Dr. Heather M. Cheshire
Committee Chair

Dr. Viney P. Aneja
Committee Co-Chair

Dr. Stacy A.C. Nelson

Dr. S. Trivikrama Rao
DEDICATION

I dedicate this dissertation to my husband, Edward; my four children, Jessica, Tyson, Alannah and Christopher; their spouses, Matt and Laura; and my grandchildren (Jack, Ryan and Baby Meister) that for so many months handled my incessant studying in the background and the times when I “hit the wall” with loving hearts and words of encouragement. I love you all and without you, I could not have done this.
Valerie Garcia is a Branch Chief in the National Exposure Research Laboratory (NERL), within the U.S. Environmental Protection Agency (EPA), Office of Research and Development. In this position, she leads the planning and execution of applied research to better characterize air pollution exposure and link to ecosystem and human health through the use of models and other tools. Her diverse background in leading programs ranging from ecosystems to atmospheric science has laid the foundation for this current effort that requires bridging research disciplines. Past experience with Congressional and Agency budget planning and execution, science policy, and program management have further strengthened her foundation for leading scientific programs. This past experience includes being a Deputy Director in the Atmospheric Research Laboratory within the National Oceanic and Atmospheric Association (NOAA), where she was responsible for the efficient operations of the Division; and being a Lead Assistant Lab Director for NERL, where she was responsible for the strategic planning of the organization’s $10M research program.

Among other awards, Valerie received a Silver Medal in 2008 for work she did in linking air quality to human health. In addition, she received the Administrator’s Award from NOAA in 2008 for visionary leadership, the Government Executive Magazine Technology Leadership Award in 1997, and the Meritorious Civilian Award from the U.S. Navy in 1990. Valerie earned her Master’s of Science in Forestry with a minor in spatial analysis at North Carolina State University in 2004, and a Bachelor of Science in Information Systems Management from the University of San Francisco in 1995. She is the mother of four children and the Grammy of three grandchildren, and hopes to continue to dare, love and live life to its fullest.
ACKNOWLEDGMENTS

First and foremost, I acknowledge my Committee, Dr. Cheshire who has worked with me for over a decade in pursuit of my master’s and doctoral degrees, Dr Rao for his suggestion of the research problem and his tireless technical and educational guidance and support, and Dr. Aneja and Dr. Nelson for committing hours of their time towards my educational and professional growth. I thank Dr. Edith Porter for her technical support in several areas, Dr. James Godowitch, Adrienne Wooten and Dr. Rob Gilliam for help with the meteorology and back-trajectories; Dr. Jeff Young and Dr. Rohit Mathur for their help with core atmospheric science concepts; Dr. James Crooks, Dr. Kristen Foley and Dr. Jenise Swall for their help with the statistical aspects of the project and coding in R language; Wilma Jackson for her help in producing a decade worth of combined surfaces; and Sarah Mikols, Dr. Haluk Ozkaynak, Dr. Shao Lin and Dr. Cristian Pantea for their help with the epidemiology analysis.

Finally, I’d like to thank Steve Howard and Alfreida Torian for their help in collecting and processing data; and Dr. Tanya Otte, Dr. Vlad Isakov and Dr. Brian Eder for their consistent support every time it was asked for. A heartfelt thanks goes out to each and everyone of these individuals that were so instrumental to my success.
# TABLE OF CONTENTS

LIST OF TABLES .................................................................................................................. vii
LIST OF FIGURES ............................................................................................................... viii

1.0 Introduction .................................................................................................................. 1
  1.1 Background.................................................................................................................. 1
    1.1.1 Policy .................................................................................................................... 1
    1.1.2 Ozone Transport .................................................................................................. 6
    1.1.3 Ozone Health Studies .......................................................................................... 8
  1.2 Objectives and Impact of this Research ....................................................................... 12

REFERENCES ..................................................................................................................... 14

2.0 Refining Air Quality Characterization ........................................................................... 19
  2.1 “A Comparison of Statistical Techniques for Combining Modeled and Observed Concentrations to Create High-Resolution Ozone Air Quality Surfaces” ................................................................. 19
    2.1.1 Abstract .............................................................................................................. 19
    2.1.2 Implications ....................................................................................................... 19
    2.1.3 Introduction ....................................................................................................... 20
    2.1.4 Approach .......................................................................................................... 23
    2.1.5 Observations ..................................................................................................... 24
    2.1.6 Model Output .................................................................................................... 25
    2.1.7 Statistical Combination Techniques ................................................................. 25
    2.1.8 Comparison of different techniques ................................................................... 29
    2.1.9 Results and Discussion ..................................................................................... 31
    2.1.10 Summary ......................................................................................................... 38
  2.2 Comparison of Air Quality Surfaces in Epidemiology Model ....................................... 39

REFERENCES ..................................................................................................................... 43

3.0 Impact of Transported Pollution on Human Health ....................................................... 47
  3.1 “An evaluation of transported pollution and respiratory-related hospital admissions in the state of New York” ........................................................................................................... 47
    3.1.1 Abstract .............................................................................................................. 47
    3.1.2 Introduction ....................................................................................................... 48
    3.1.3 Air quality and health data ................................................................................. 51
    3.1.4 Air pollution transport ....................................................................................... 52
    3.1.5 Classification of transport directions ................................................................ 53
    3.1.6 Assessment of corresponding ozone concentrations and pollutant transport directions .............................................................................................................................. 55
    3.1.7 Assessment of hospital admissions and transport directions ............................ 56
    3.1.8 Examination of temperature ............................................................................. 56
    3.1.9 Results .............................................................................................................. 57
    3.1.10 Discussion ....................................................................................................... 60
    3.1.11 Summary ....................................................................................................... 62
3.2 Cluster Analysis.............................................................................................................. 63
REFERENCES ......................................................................................................................... 65
4.0 Assessing the Health Impacts of the NOx Budget Trading Program (NBP) ................................................................................................................................. 68
4.1 “Method to Track the Impact of Regional-Scale Air Quality Regulations on Human Health Applied in New York State” ................................................................. 68
  4.1.1 Abstract ....................................................................................................................... 68
  4.1.2 Introduction ................................................................................................................. 69
  4.1.3 Air quality and health data .......................................................................................... 71
  4.1.4 Approach .................................................................................................................... 73
  4.1.5 Results ....................................................................................................................... 76
  4.1.6 Discussion ................................................................................................................. 83
  4.1.7 Summary .................................................................................................................... 87
REFERENCES ......................................................................................................................... 88
5.0 Conclusions and Future Research .................................................................................. 92
  5.1 Conclusions .................................................................................................................... 92
  5.2 Future Research ............................................................................................................ 94
REFERENCES ......................................................................................................................... 96
LIST OF TABLES

Table 2.1.1: Cross-validation results (daily maximum 8-hr average ozone concentrations). ........................................................................................................32

Table 3.1.1: Odds ratio calculations resulting from the arc and bounded methods of determining exposure. ..........................................................59

Table 4.1.1: Risk Calculated for Ozone (8-hr DMA) and the ORV Variable ...............84
LIST OF FIGURES

Figure 1.1.1: Schematic of the photochemical pathways leading to the production and accumulation of ozone in the troposphere and the main termination reaction at high and low VOC NO\textsubscript{x} reactions. (Source: modification of Chinkin et al., 2002) .................................................................2

Figure 1.1.2: Typical ozone isopleths plot showing 1-hour maximum ozone concentrations (in ppb) calculated as a function of initial VOC and NO\textsubscript{x} concentrations and the regions of the diagram that are characterized as VOC- or NO\textsubscript{x}-limited. (Source: Chinkin et al., 2002) .........................................................4

Figure 1.1.3: Location of Major Point Sources of NO\textsubscript{x} emissions (tons yr\textsuperscript{-1}). Source: National Emissions Inventory, US EPA, 2001. .................................................................5

Figure 2.1.1: Ozone 8-hr DMA (ppb) for June 13, 2001: (a) modeled, (b) observed and (c) kriged observations. Kriging smoothes the hotspots whereas, the model may over-predict titration in NYC. .................................................................23

Figure 2.1.2: Domain and observations used in study. Solid circles show location of AQS sites, triangles show NAPS sites and squares show CASTNET sites. Open circles highlight sites clustered in urban areas. .........................................................24

Figure 2.1.3: Location of cross-validation sites. Circles denote NAPS sites, triangles denote CASTNET sites and diamonds denote AQS sites. .................................................................30

Figure 2.1.4: Observed (y-axis) and predicted (x-axis) 8-hr DMA zone (ppb) binned by percentile for cross-validation sites. Black = 0-50%, red = 51-75%, orange = 76-90%, blue = 91-95%, green = 96-100%. .................................................................33

Figure 2.1.5: Mean error (binned by percentile) between modeled, kriged and the four combination techniques versus mean observed concentrations for all cross-validation sites. Circles represent average mean for each binned percentile; lines are for identification of technique but do not represent linear relationships between averaged points. .................................................................34

Figure 2.1.6: Contribution of spatial information from model in combined surface of 8-hr DMA. Panel (a) is modeled surface, panel (b) is interpolated observations and panel (c) is multiplicative adjusted bias. Top panels are for June 7, 2001 and bottom panels are for July 19, 2001. .................................................................35

Figure 2.1.7: Mean 8-hr DMA ozone concentrations across all days for (a) modeled, (b) kriged observations, (c) multiplicative bias, (d) additive bias, (e) weighted average and (f) HBM. .................................................................36

Figure 2.1.8: Coefficient of variation calculated for 8-hr DMA ozone concentrations across all days for (a) modeled, (b) kriged observations, (c) multiplicative bias, (d) additive bias, (e) weighted average and (f) HBM .................................................................37

Figure 2.1.9: Maximum values of (a) observed-to-modeled ratios (eq 4) and (b) 8-hr DMA ozone concentrations resulting from the multiplicative adjusted bias approach. .................................................................38
Figure 2.2.1: Four New York City counties and associated population used in comparison study.

Figure 2.2.2: Boxplots of 3-day moving average of maximum daily 8-hr ozone concentrations using observations (red) and combined observations and model (Combined: green).

Figure 2.2.3: Risk calculated using same-day, lag 1 – 3 days, and 3-day moving average using ozone observations and combined observations and model (Combined).

Figure 3.1.1: Transport of ozone and precursor chemicals from Midwestern U.S. into Northeastern U.S. Source: OTAG 1997.

Figure 3.1.2: Location of eight NYS regions, back-trajectory origination sites (red squares), CASTNet monitoring sites (black circles) and AQS monitoring sites (yellow triangles).

Figure 3.1.3: Arc approach. Black triangles pointing up represent the direction of end points for back-trajectories originating from the North, purple squares represent the southeasterly direction of origin, blue triangles pointing down represent the southerly direction of origin, red circles represent the southerly direction of origin and green diamonds represent the westerly direction of origin. X-axis is longitude, y-axis is latitude.

Figure 3.1.4: Bounded zone approach. Panel (a) shows high NOx emissions (July 1997 shown as an example) originating from power plants in Midwestern U.S.; (b) shows boundaries drawn for identifying days when the air parcel passed through a zone in the Midwest versus days it passed through a zone in the North (red rectangles). Also shown is the southern zone used in the temperature analysis.

Figure 3.1.5: Boxplots displaying the daily maximum 8-hr ozone concentrations. Ozone concentrations for (a) northerly direction defined using the arc method; (b) southwesterly direction defined using the arc method; (c) northerly direction defined using the bounded zone method; and (d) southwesterly direction defined using the bounded zone method.

Figure 3.1.6: Odds ratio. Panel (a) boxplots show the interquartile range and 5th and 95th percentiles for maximum temperature in Region 2 on days when air transport is from the south and the southwest (as defined by the bounded zone method). Panel (b) shows the odds ratio calculations for these two exposure definitions.

Figure 3.1.7: Odds ratios with 95% confidence intervals for panel (a) arc approach and panel (b) bounded zone approach.

Figure 3.2.1: Clustered back-trajectories for (a) Region 2 and (b) Region 3 shown as examples. Each cluster is represented by a different color; dotted lines indicate daily trajectories and solid lines represent the average trajectory for each cluster.

Figure 4.1.1: Transport of ozone and precursor chemicals from Midwestern U.S. into Northeastern U.S. Source: OTAG, 1997.
Figure 4.1.2: Location of eight NYS Regions, back-trajectory origination sites (red squares), CASTNet monitoring sites (black circles) and AQS monitoring sites (yellow triangles) .................................................................72

Figure 4.1.3: Bounded zone approach. Panel (a) shows high NOx emissions (July 1997 shown as an example) originating from power plants in the Midwestern U.S.; (b) shows boundary drawn around for identifying days when the airmass passed through a zone in the Midwest ..................................................73

Figure 4.1.4: Map of 8-h DMA ozone concentrations averaged across all 10 summers for each NYS region; (a) mean ozone concentrations (ppb) and (b) 95th percentile mean ozone concentrations (ppb) ........................................................................77

Figure 4.1.5: Map of population-weighted respiratory-related hospital admissions for each region; inset is pie-chart of unweighted hospital admission counts. ........77

Figure 4.1.6: Boxplots of daily data averaged for all regions; (a) hospital admissions counts, (b) 8-hr DMA ozone concentrations (ppb), (c) maximum temperature (F°), and (d) dew point (1/10 F°). .........................................................78

Figure 4.1.7: 8-hr DMA Ozone Concentrations for (a) days when the air parcel passed through the ORV boundary (“ORV” Days), and (b) days when it did not (“Non-ORV” Days). ..................................................................................79

Figure 4.1.8: Risk calculated for the two explanatory variables; (a) ORV indicator variable and (b) 8-hr DMA ozone concentrations .......................................................80

Figure 4.1.9: The risk and 95th percentile confidence intervals for the ORV variable using the (a) Unadjusted Odds Ratio, and (b) GAM ..................................................................................81

Figure 4.1.10: The risk for ozone exposure (per 10 ppb increase) and for the ORV variable is calculated for (a) ozone (pre-NBP time period); (b) ozone (post-NBP time period); (c) ORV variable (pre-NBP time period); (d) ORV variable (post-NBP time period). ..................................................................................82

Figure 4.1.11: The correlation between temperature and ozone is 0.60 for all regions........85

Figure 4.1.12: Risk and 95th percentile confidence intervals calculated using decreasing levels of data. .................................................................86
1.0 Introduction

The Clean Air Act (CAA) requires that U.S. Environmental Protection Agency (EPA) set National Ambient Air Quality Standards (NAAQS) for pollutants considered harmful to human health and the environment. Previous research has shown that high ambient ozone levels are harmful to humans. While ozone is not directly emitted (i.e. it is secondarily formed), the formation and distribution of ozone in the atmosphere is driven by chemical interactions involving nitrogen oxides (NO\textsubscript{x}) and Volatile Organic Compounds (VOCs) in the presence of sunlight (SOS, 2004). These pollutants and their precursors can be transported downwind under certain meteorological conditions, contributing to ozone levels at locations much farther from the emission source region, potentially impacting human health in downwind areas. As a result, the NO\textsubscript{x} Budget Trading Program (NBP) was issued by the EPA to reduce the transport of ozone (U.S. EPA, 1998). This study investigated methods to best characterize air quality associated with the transport of pollution into a domain encompassing New York State (NYS). Back-trajectories were performed from several sites within NYS across ten summers (June, July and August from 1997 to 2006) to identify air parcels transported from a high NO\textsubscript{x} emissions area in the Midwest. This classification was used in an epidemiology analysis to examine associations between transported air parcels from the Midwest and respiratory-related hospital admissions in NYS. This analysis was repeated for two time periods selected before and after the implementation of the NBP to examine whether a change in the health signal could be discerned between the two time periods.

1.1 Background

1.1.1 Policy

The Clean Air Act (CAA) (42 USC §7401) establishes two major categories of air pollutants requiring different standard setting processes; criteria pollutants and hazardous air pollutants (HAPS). In general, criteria pollutants are considered to be more ubiquitous, pose a risk to a larger fraction of the general population, and have more widespread impacts on ecosystems and natural resources than HAPs (NRC, 2004). Criteria pollutants are regulated
primarily through the setting of National Ambient Air Quality Standards (NAAQS), which establish concentration thresholds that must be attained by each state. HAPs are regulated through standards that limit the emissions of such compounds (as opposed to their ambient concentrations; NRC, 2004). Ozone, the criteria pollutant of focus in this study, is not directly emitted into the atmosphere. The formation and distribution of ozone is controlled by chemical interactions involving NOx and VOCs (Figure 1.1.1) and prevailing meteorological conditions.

![Figure 1.1.1: Schematic of the photochemical pathways leading to the production and accumulation of ozone in the troposphere and the main termination reaction at high and low VOC NOx reactions. (Source: modification of Chinkin et al., 2002)](image)

As evident in, the relationship between ozone, NOx and VOCs is non-linear. VOCs are relatively short-lived and are emitted from vegetation (biogenic) and anthropogenic
sources. More than 50% of VOCs for some areas of the United States, such as in the Southeast, are produced by vegetation (SOS, 2004). Anthropogenic VOCs are emitted by mobile sources and industry (e.g., solvent use) and are generally considered an urban phenomenon (Jacobson, 2005; American Lung Association, 2004, USEPA, 1998). NOₓ is produced by mobile and area sources, and electric utilities (Jacobson, 2005; American Lung Association, 2004, USEPA, 1998). NOₓ is longer-lived and can be transported hundreds of kilometers downwind (Godowitch et al., 2008a; Olaguer et al., 2006; Civerolo et al., 2003; Biswas et al. 2001; Brankov et al., 1998; OTAG, 1997), particularly when emitted higher up in the atmosphere such as with the buoyant plumes released from the tall stacks associated with coal-fired power plants (Figure 1.1.3). Such plumes can enter the free troposphere during the evening when the boundary layer is compressed and can be carried aloft in the nocturnal jet stream where cooler temperatures allow precursor chemicals to exist for a longer period of time allowing for the generation of ozone. This, coupled with the minimization of removal processes due to the lack of surface layer friction, allows for the accumulation and regional transport of ozone to downwind areas where it is brought to the surface during the daytime expansion of the boundary layer.
Figure 1.1.2: Typical ozone isopleths plot showing 1-hour maximum ozone concentrations (in ppb) calculated as a function of initial VOC and NO\textsubscript{x} concentrations and the regions of the diagram that are characterized as VOC- or NO\textsubscript{x}-limited. (Source: Chinkin et al., 2002)
During the 1970s and 1980s, control strategies targeted VOC reductions only, as ozone was believed to be primarily an urban problem. Despite reduced VOC emissions, however, some areas, particularly in the Northeast, were still unable to comply with the NAAQS. Beginning in the 1990’s, a better understanding of the underlying meteorology and chemistry of ozone formation caused the regulatory community to rethink the VOC-only strategy and to consider NO\textsubscript{x} controls in addressing ozone transport (Olaguer et al., 2006; NRC, 1991; Chameides et al., 1988). The CAA Amendments of 1990 introduced the Acid Rain Program aimed at reducing SO\textsubscript{2} and NO\textsubscript{x} emissions from coal-fired power plants. In addition, the CAA Amendments created the Ozone Transport Commission (OTC) to address transported precursor and ozone pollution in the Northeastern U.S.—particularly during the summertime “ozone season” (May 1 through August 31).

The Ozone Transport Assessment Group (OTAG) was formed, consisting of members from EPA, the 37 Eastern States and the District of Columbia, industry representatives and environmental groups (OTAG, 1997). OTAG conducted a 2-year regional assessment of ozone transport that, among other findings, concluded that the central portion of the 37-state OTAG domain was characterized by persistent elevated ozone levels.
producing an "ozone pool." Transport in any direction from this region was associated with high ozone levels in neighboring areas. High ozone levels in the southern portion of the OTAG domain were associated with stagnant pollutant conditions, whereas high ozone levels in the northern portion of the OTAG domain were associated with higher wind speeds and persistent transport conditions from inside the OTAG domain, particularly from the Ohio River Valley where major NOx emitting power plants are located (OTAG, 1997).

As a result of control actions implemented by the OTC, modest reductions of NOx emissions began in 1996 in the Northeast. In 1998, the NBP was issued to further reduce the regional transport of ozone in the Eastern U.S. by reducing summertime NOx emissions from major sources—predominantly the electric utilities in the central-eastern portion of the United States. The NBP was the first regulation issued to address ozone transport across state boundaries.

Recent studies have shown that the NBP has contributed to improvements in reduced ozone concentrations at downwind sites (Gego, et al., 2007, US EPA, 2007). This study will investigate methods to capture the impact of transported pollution from the Midwest on human health in NYS. Methods to improve the characterization of air quality are investigated and techniques to isolate transported pollution from an area in the Midwest are identified. A transport variable is identified and applied in an epidemiology analysis to examine the risk of respiratory-related hospital admissions from exposure to transported air pollution. The method is then applied to two time periods before and after the implementation of the NBP to examine whether a change in the health effects measure can be discerned across the two time periods.

1.1.2 Ozone Transport

Several studies have used trajectory analysis to investigate the regional transport of pollutants. Wolff and Lioy (1980) described a distinct area of high ambient ozone concentrations associated with a synoptic scale high-pressure weather system that carried ozone in a “river” northeastward from the Gulf of Mexico using trajectory analysis. They examined three high-ozone episodes that occurred during July 1977 and found that ozone concentrations within this river were ~120 to 130 ppb, and in one case, traveled in excess of
2300 km in 48 hrs. Ashbaugh (1983) used a trajectory technique to examine the existence of clean and dirty transport corridors into the Grand Canyon National Park based on transport pathways of air masses that arrive at an observation site under given pollution conditions. Ashbaugh concluded that high sulfur concentrations were related to stagnant air and slow transport from southern California and northeastern Nevada, and that low sulfur concentrations were associated with good ventilation and transport from eastern Utah and western Colorado and high speed transport from southern California. Dorling and Davies (1995) clustered isobaric trajectories from 1981 to 1984 to relate weather patterns to precipitation chemistry at five northwest European monitoring stations. They found that while topography and heavy precipitation can confound the approach, trajectory clusters can be used as surrogates for synoptic patterns.

Ryan et al., (1998) used a mix of data (rural surface observations, sonde data and aircraft measures) and back-trajectory analysis to examine the meteorology, chemistry and source of a severe regional ozone event from July 12 - 15, 1995 over the Baltimore-Washington region. They found that the chemical composition aloft (high ozone, sulfur dioxide and total reactive nitrogen and very low nitric oxide concentrations) indicated photochemical aging of an air parcel transported some distance with, at least partially, a coal combustion source. Brankov et al. (1998) identified potential pollution source regions and atmospheric patterns associated with the long-range transport of air pollutants. In a separate study, Brankov et al. (1999) clustered a large number of back-trajectories from several observation sites in the Northeastern U.S. to examine the relationship between synoptic-scale atmospheric transport patterns and concentration levels of several toxic trace elements observed in the domain.

Taubman et al. (2004) used airborne observations of trace gases, particle size distributions and particle optical properties taken from New Hampshire to Maryland during a multi-day ozone-haze episode on August 14, 2002 to investigate a two-reservoir model, composed of the lower free troposphere, where photochemical processes are accelerated and removal from deposition does not occur, and the planetary boundary layer, where most precursor species are injected. Taubman et al., found that back-trajectories indicated source
regions in the Midwest and Mid-Atlantic urban corridor, with southerly transport up the urban corridor augmented by the Appalachian lee trough and nocturnal jet stream. Zhang et al. (2001) and Taubman et al. (2006) used cluster analysis of back-trajectories in conjunction with vertical profile data to identify source regions and characterize transport patterns during summertime pollution episodes. The greatest trajectory density for these high-pollutant events lay over the northern Ohio River Valley. The resulting chemical composition of the air masses (high ozone values, large SO²/CO ratios, highly scattering particles and large aerosol optical depths) supported this conclusion.

Finally, and specifically related to the NBP, Godowitch, et al. (2008a) and Godowitch et al. (2008b) used forward-trajectories and the Community Multiscale Air Quality (CMAQ) chemical transport model to demonstrate that southwesterly wind flows across the midwestern source region transport pollutants toward the Northeastern states, and that emission reductions resulting from the NBP dramatically decreased ozone production rates downwind of these sources. The authors also found evidence that the chemical regime at these downwind locations has shifted towards a more NOₓ-limited condition.

1.1.3 Ozone Health Studies

Several studies have also shown an association between ozone concentrations and health effects. Burnett, et al. (1994) investigated the associations between emergency daily respiratory admissions at 168 acute care hospitals in Ontario and ambient ozone concentration levels between January 1983 and December 1988. Positive and statistically significant associations were found between hospital admissions and ozone levels recorded on the day of admission and up to 3 days prior to the date of admission. The authors also found that temperature had no effect on the relationship between air pollution and hospital admissions.

Burnett, et al. (1997) studied the effects of ozone on lung function that were severe enough to require hospitalization after controlling for weather, considering gaseous and particulate air pollution at lower concentrations typically observed in Canada. To determine if low levels of ozone contribute to hospitalization for respiratory disease, air pollution data were compared to hospital admissions related to respiratory diseases for 16 cities across
Canada between April 1981 and December 1991. After controlling for sulfur dioxide, nitrogen dioxide, carbon monoxide, soiling index, and dew point temperature, the daily 1-hour maximum concentration of ozone recorded one day previous to the date of admission was positively associated with respiratory-related hospital admissions. These results suggest that ambient ozone at the relatively low concentrations observed in this study is associated with increased hospital admissions for respiratory diseases in populations experiencing diverse weather and air pollution profiles. To clarify the health effects of ozone exposure in young children, Burnett, et al. (2001) studied the association between air pollution and hospital admissions for acute respiratory problems in children less than 2 years of age between 1980 and 1994 in Toronto, Canada. After adjusting for other ambient air pollutants and weather variables, the authors found an increase in respiratory-related hospitalization rates associated with the daily 1-hour maximum average ozone concentrations of 45 parts per billion for May through August.

Thurston and Ito (2001) evaluated past epidemiological studies that assessed the short-term association of ozone with mortality, and investigated one possible reason for variations in their ozone effect estimate, i.e., differences in their approaches to the modeling of weather influences on mortality. The authors found that past time-series studies using linear temperature-mortality specifications may have under-predicted the premature mortality effects of ozone air pollution, and that mortality effect estimates derived by time-series analyses for ozone can be sensitive to the way that weather-induced variations are addressed in the model. Generally, the authors found that the ozone-mortality effect estimate increased in size and statistical significance when the nonlinearity and the humidity interaction of the temperature-health effect association are incorporated into the model weather specification.

Bell, et al. (2004) used analytical methods and databases developed for the National Morbidity, Mortality, and Air Pollution Study (NMMAPS) to estimate a national average relative rate of mortality associated with short-term (daily and weekly) exposure to ambient ozone for 95 large U.S. urban communities from 1987 to 2000. The authors adjusted for time-varying confounders (particulate matter, weather, seasonality, and long-term trends). The study results indicated a statistically significant association, on average, between short-
term changes in ozone and mortality for the 95 urban communities which include about 40% of the total U.S. population.

Ito, et al. (2005) conducted a review and meta-analysis of ozone mortality studies for 7 U.S. cities (Chicago, Detroit, Houston, Minneapolis-St. Paul, New York City, Philadelphia, and St. Louis), and found a positive short-term association between ozone and daily mortality in the majority of the cities, although the estimates were heterogeneous across cities. The authors also found that including PM in the model did not substantially reduce the ozone risk estimates, but that the difference in the weather adjustment model could result in a 2-fold difference in risk estimates. Bell, et al. (2005) performed a meta-analysis of 144 effect estimates from 39 time-series studies, and estimated pooled effects by lags, age groups, cause-specific mortality, and concentration metrics. The authors compared results with pooled estimates from the NMMAPS, a time-series study of 95 large U.S. urban centers from 1987 to 2000. Bell, et al. found that both the meta-analysis and NMMAPS results provided strong evidence of a short-term association between ozone and mortality and that the results were insensitive to adjustment for particulate matter and model specifications.

Schwartz (2005) examined the sensitivity of controlling for temperature in modeling the association between ozone and daily deaths. The use of Poisson regression, with complicated modeling strategies to control for season and weather, has raised concerns that the results may be sensitive to these modeling protocols. Schwartz used the case-crossover approach to convert the problem into a case-control study, where the control for each person is the same person on a day near in time, when he or she did not die. Schwartz used this method to set the control day to have the same temperature as the event day and applied this approach to study more than 1 million deaths in 14 U.S. cities. Schwartz found that the association between ozone and mortality risk is unlikely to be caused by confounding by temperature.

Rainham, et al. (2005) assessed whether meteorological conditions modified the relationship between short-term (daily) exposure to ambient air pollution and mortality, examining air pollution and human mortality associations using hybrid spatial synoptic classification procedures. Concentrations of air pollutants and human mortality from all non-
accidental and cardio-respiratory causes were examined according to typical winter and summer synoptic climatologies in Toronto, Canada, between 1981 and 1999. Air masses were derived using a hybrid spatial synoptic classification procedure associating each day over the 19-year period with one of six different typical weather types, or a transition between two weather types. Generalized linear models (GLMs) were used to assess the risk of mortality from air pollution within specific air mass type subsets. Mortality follows a distinct seasonal pattern with a maximum in winter and a minimum in summer. Average air pollution concentrations were similar in both seasons with the exception of elevated sulfur dioxide levels in winter and elevated ozone levels in summer. The authors found that subtle changes in meteorological conditions (e.g., wind speed, temperature, relative humidity) can alter the strength of pollutant associations with health outcomes, especially in the summer season. Although there does not appear to be any systematic patterning of modification, variation in pollutant concentrations seems dependent on the type of synoptic category present.

A study sponsored by the New York State Energy Research and Development Authority (NYSERDA, 2006) investigated the relationship between several pollutants and asthma emergency department (ED) visits in the Manhattan and South Bronx area within New York City between January 1999 and November 2000. This study found a correlation between asthma ED visits and daily maximum 8-hr average ozone and 24-hour average NO\textsubscript{2} in the Bronx study area. This study also found that ozone causes health effects independent of the other pollutants studied.

In general, these studies indicate that there is a positive association between ozone concentration levels and health effects, particularly respiratory diseases. In addition, Ito, et al., (2005), Bell, et al. (2005) and NYSERDA (2006) found that ozone was not confounded by other pollutants investigated in their respective studies. While the literature supports an established positive relationship between temperature and health effects (Gover, 1938; Bull and Morton, 1975; Larsen, 1990; Wolfe, et al., 2001; Curriero, et al., 2002; Basu and Samet, 2002), some studies (Burnett, et al., 1994; and Schwartz, 2005) found that temperature does not have an effect on the ozone-health signal. Other studies (Thurston and Ito, 2001; Ito, et
al., 2005; and Rainham, et al., 2005) indicate that there is evidence that temperature confounds the health signal. Hence, whether temperature and humidity, and related synoptic weather patterns confound the ozone-respiratory illness association is less clear.

1.2 Objectives and Impact of this Research

The CAA (amended 1990) requires that EPA set NAAQS for pollutants considered harmful to public health and the environment. The NBP mandated significantly reduced summertime emissions of NO\textsubscript{x} by 2004 for 21 states in the Eastern U.S. to help reduce the transport of ozone and its precursor chemicals. The promulgation and implementation of regulations such as these are costly. More importantly, understanding whether we have sufficiently protected the public is of critical concern. Thus, determining whether regulatory actions actually reduce air pollution and improve public health and the environment is an important step in environmental policy implementation.

Previous studies have established that NO\textsubscript{x} and ozone are transported under certain meteorological conditions affecting ozone concentration levels in the Northeastern U.S. In addition, several health studies have related ozone to adverse respiratory responses and mortality. This study is unique in that it applies an approach to include the risk of transported air pollution in investigating associations with health endpoints. Once verified, the approach is then applied to investigate whether a change in the health risk can be detected between the pre-NBP and post-NBP time periods.

Specifically, the objectives of this study are to: (1) investigate methods to best characterize the spatio-temporal variability in air quality; (2) investigate associations among ozone concentrations, meteorological patterns, and respiratory-related hospital admissions; and (3) assess differences in these associations before and after the implementation of the NBP. The overarching research questions to be addressed by this study are:

- Do air quality surfaces derived from various statistical data combination techniques help us to best characterize air quality and enable us to understand associations between transported air pollution and human health endpoints?
- Can atmospheric transport patterns associated with high-ozone days be linked to human health endpoints?
• Can we discern a difference in the health effects measure before and after the implementation of the NBP?

This dissertation is composed of three main sections that correspond to the three research questions outlined above. Each of the three sections represents a journal article in various stages of submission and publication. Therefore, the approach, results and discussion of the research are presented in the respective section coinciding with the questions and articles.
REFERENCES


2.0 Refining Air Quality Characterization

This section presents the manuscript entitled “A Comparison of Statistical Techniques for Combining Modeled and Observed Concentrations to Create High-Resolution Ozone Air Quality Surfaces”. The manuscript was published in the refereed Journal of the Air & Waste Management Association in 2010 (Garcia et al.). The article compares several methods for generating high quality ambient concentration surfaces for NYS and recommends an approach for use in this study. The sub-section after the paper describes the results from applying the approach in a timeseries analysis to investigate the change in the calculated risk using two characterizations of ozone exposure; average observations and the combination technique presented in the paper.

2.1 “A Comparison of Statistical Techniques for Combining Modeled and Observed Concentrations to Create High-Resolution Ozone Air Quality Surfaces”

2.1.1 Abstract

Air quality surfaces representing pollutant concentrations across space and time are needed for a multitude of applications, including tracking trends and relating air quality to human and ecosystem health. The spatial and temporal characteristics of these surfaces may reveal new information about the associations between emissions, pollution levels, and human exposure and health outcomes that may not have been discernable before. This paper presents four techniques, ranging from simple to complex, to statistically combine observed and modeled daily maximum 8-hr average ozone concentrations for a domain covering the greater New York State area for the summer of 2001. Cross-validation results indicate that, for the domain and time period studied, the simpler techniques (additive and multiplicative bias adjustment) perform as well as or better than the more complex techniques. However, the spatial analyses of the resulting ozone concentration surfaces revealed some problems with these simpler techniques in limited areas where the model exhibits difficulty in simulating the complex features such as those observed in the New York City area.

2.1.2 Implications

Linking emission control actions to human health impacts is important in determining whether the regulations that have been implemented are reducing air pollution as intended.
Measurements of pollutant concentration levels are often spatially sparse, and modeled outputs are only an estimate of the “true” pollutant concentration levels, hampering our ability to detect a relatively small signal of change embedded in ambient concentrations. This paper assesses four techniques to combine the strengths of modeled and observed data to provide high-resolution ozone concentration surface maps for use in human health studies and assessing whether regulatory control actions have had the intended impact.

2.1.3 Introduction

Air pollutant concentrations across space and time are used in a multitude of applications, including tracking trends and relating air quality to human and ecosystem health. Often, changes in air quality attributable to emission reductions stemming from control policies are weak signals within the overall changes in observed or modeled concentrations. This signal can be further confounded when investigating the impacts of emission reductions on human health. Although air quality observations taken at various locations represent the “ground truth”, these observations are often limited in terms of spatial and temporal coverage. Air quality models can predict pollutant concentrations over a given spatial domain, but the modeled values are uncertain due to model input errors and the model’s inability to perfectly simulate the various physical and chemical processes occurring in the atmosphere. To alleviate these problems, four techniques that statistically combine air quality measurements with model output to produce high-resolution ambient pollutant concentrations are considered here to better characterize air quality in a study area encompassing New York State for the 2001 ozone season. These improved air quality surfaces may reveal associations between pollution levels and health outcomes not discernable before.¹

Several investigators ²⁻⁷ have applied techniques, such as the Kalman Filter and ensemble approaches, to adjust forecasted meteorological variables and air pollution concentrations using observations. Gego et al.⁸ and Hogrefe et al.⁹ applied bias-adjustment approaches to predict air pollutant concentrations for use in health studies. Hirst et al.¹⁰ used a hierarchical modeling approach to combine measured and modeled deposition values to estimate long-range transport of air pollutants in Europe by modeling the underlying
(unobserved) "true" deposition process as a function of two stochastic components; one non-stationary and correlated over long distances, and the other describing variation within a grid square. Fuentes and Raftery \textsuperscript{11} used a Bayesian approach to combine observed and modeled dry deposition pollutant concentrations to improve spatial predictions, evaluate the model and remove the model bias. Similar to Hirst et al. \textsuperscript{10}, Fuentes and Raftery \textsuperscript{11} assumed the model output and available observations can be represented as a function of an unobserved ground truth plus error and bias terms. McMillan et al. \textsuperscript{12} defined an approach used in this study (discussed later in the paper) that again applies a hierarchical Bayesian approach to predict ozone concentrations, but extends the model to include a temporal dimension using an autoregressive structure.

Although varying in their approach, the intent of these statistical data combination techniques is to retain the strongest components of each data type (i.e., observations and model outputs) to best represent the pollutant of concern. For example, observations are the best representation of the "true" pollutant concentration value at a given site and time, however, depending on the network and the chemical being measured, the spatial and temporal extent may be limited. Classical interpolation of observed concentrations helps to "fill-in" the spatial and temporal gaps, but tends to produce overly smoothed results. Three-dimensional deterministic air quality models, such as the Community Multiscale Air Quality (CMAQ) model\textsuperscript{13}, can estimate pollutant concentrations across a uniform spatial and temporal scale. However, the accuracy of these estimates is based on our uncertain understanding of the physical and chemical processes underlying the formation, interaction and fate of atmospheric pollutants, and errors in the model input (e.g., emissions, meteorology, boundary conditions). Even a perfect model with perfect model input cannot reproduce the observations exactly since random variations embedded in the observations taken at individual monitoring locations are not explicitly estimated in current regional-scale models. Thus, an observation reflects a single event out of a population, whereas, the modeled concentration represents the population average. Moreover, the predictions from a deterministic model represents the volume-averaged concentration for the grid cell while observations at a given monitoring location reflect point measurements.\textsuperscript{14}
Figure 2.1.1 illustrates some of the strengths and weaknesses inherent in the modeled and observed data used in this study by displaying the modeled, observed and interpolated (kriged) surfaces for the domain encompassing New York State for June 13, 2001. On this day, the interpolated observations appear to be overly smooth and miss the “hot spots” generated by emissions from significant sources of ozone precursor chemicals captured by the model. The modeled surface shows the effects of titration in the area of New York City (NYC) where high nitrogen oxide (NO) emissions “scavenge” ozone creating areas of low ozone concentrations around the emission source. If these local features are measured, kriging tends to smooth them out. The model captures this important feature, but may overestimate the extent of the titration effect. Thus, capturing “hot spots” and other spatial gradients that may be smoothed out by kriging, yet correcting the bias that may exist in model estimates may produce improved air quality surface concentrations critical to detecting health impacts. In this study, we apply statistical combination techniques that integrate the observed concentrations with the model estimates to optimize the strengths of each dataset. We focus on the summer of 2001 for the New York State (NYS) domain as a pilot study to demonstrate the use of these enriched air quality data in an epidemiological health study and risk assessment for NYS.
2.1.4 Approach

Four techniques ranging from relatively simple to complex, are investigated here for combining observed and modeled ozone concentrations for a domain in the greater NYS area from June 1 through August 31, 2001 (Figure 2.1.2). The focus of this investigation is on providing the daily maximum 8-hr average ozone concentrations to state health assessors for their use in investigating relationships between air quality and human health endpoints (e.g., respiratory-related hospital admissions) across multiple years. The daily maximum 8-hr average ozone concentrations used in the study were calculated from observations and from CMAQ model predictions. Four statistical combination techniques were applied to the observed and modeled data to produce combined daily maximum 8-hr average ozone concentration surfaces on a 12 km x 12 km spatial grid for a total of 92 days in the summer of 2001. These combined surfaces were quantitatively compared through cross-validation.
and qualitatively compared through analysis of the spatial surfaces produced by each of the methods.

![Domain and observations used in study. Solid circles show location of AQS sites, triangles show NAPS sites and squares show CASTNET sites. Open circles highlight sites clustered in urban areas.](image)

**Figure 2.1.2:** Domain and observations used in study. Solid circles show location of AQS sites, triangles show NAPS sites and squares show CASTNET sites. Open circles highlight sites clustered in urban areas.

### 2.1.5 Observations

Hourly ozone observations for June 1 through August 31, 2001 were obtained from:

1. the Environmental Protection Agency’s (EPA) Air Quality System (AQS) database ([http://www.epa.gov/oar/data/aqsdb.html](http://www.epa.gov/oar/data/aqsdb.html)) and Clean Air Status and Trends Network (CASTNET) ([http://www.epa.gov/castnet/](http://www.epa.gov/castnet/)); and
2. the Canadian Environmental Assessment Agency’s National Air Pollution Surveillance Network (NAPS) ([http://www.etc-cte.ec.gc.ca/publications/napsreports_e.html](http://www.etc-cte.ec.gc.ca/publications/napsreports_e.html)).

All monitoring networks provided hourly concentrations for each day in the summer with a total of 200 sites (139 AQS sites, 52 NAPS sites and 9 CASTNET sites). Information on the quality assurance conducted for each network can be found on the websites provided above. The daily maximum 8-hr average
ozone concentrations were calculated by applying an 8-hr moving window to the hourly time series and selecting the 8-hr time window with the highest average ozone concentration (referred to as the daily maximum 8-hr average DMA throughout the paper). Only those days having greater than 20 hours of data were used for computing the DMA.

2.1.6 Model Output

Ozone 8-hr DMAs were calculated from the hourly concentration values simulated by EPA’s CMAQ model, version 4.5. Specifically, the simulated concentrations for June 1 through August 31 were extracted from the 2001 annual simulation with 12 km x 12 km horizontal grid cells. The meteorology and emissions inputs for this simulation were from the Fifth-Generation NCAR / Penn State Mesoscale Model (MM5) and EPA’s 2001 National Emissions Inventory, respectively. The 12-km simulation encompassed most of the Eastern United States (U.S.) and was nested within a 36 km x 36 km horizontal grid simulation covering the contiguous U.S. using the same model configuration as the 12-km nested simulation. Boundary conditions for the 36-km simulation were provided by a global chemical transport model (GEOS-CHEM). Several evaluations have been done of the CMAQ model. In particular, Appel et al. evaluated the CMAQ annual simulation used in this study and found that the median observed and predicted ozone concentrations correspond well between 10 a.m. and 6 p.m. which are the hours that typically makeup the 8-hr DMA. Appel et al. also report that the model consistently over-estimated low concentrations and under-estimated high concentrations of ozone. Additional details on the CMAQ simulation used in this study and the evaluation of the simulation results can be found in Byun and Schere and Appel et al., respectively.

2.1.7 Statistical Combination Techniques

The four techniques investigated for combining the observed and modeled values are: additive bias adjustment, multiplicative bias adjustment, weighted average, and hierarchical Bayesian. The application of each of these techniques resulted in a spatial surface (12 km horizontal resolution) of 8-hr DMA ozone concentrations (ppb) for each day of the summer in 2001. In general, the bias adjustment approaches modify the modeled surface concentration values by accounting for bias in the modeled ozone concentrations. The
weighted-average approach uses weights for the interpolated observations and modeled surfaces based on a relative accuracy of each surface at each grid cell, and the hierarchical Bayesian approach treats both the observations and the model concentrations as representing a true surface and calculates model parameters using a Markov Chain/Monte Carlo technique.

**Interpolated Observations.** In addition, 8-hr DMA ozone observations were interpolated with a kriging approach that applies a Matern spatial correlation function to produce daily 12 km x 12 km maps. For each day of the study, the parameters for the Matern covariance function were estimated using the restricted maximum likelihood estimation technique. Kriging was performed to estimate the concentration at the grid center-point rather than block kriging since the results of the two techniques were similar. Differences in the spatial correlation structure along different directions (anisotropy) was also accounted for in the model. This same kriging approach was used in the bias-adjustment approaches and the weighted-average technique discussed below.

**Additive Bias Adjustment.** The additive bias was calculated by subtracting the modeled value from the observed value at each observation site for each day. This bias was then interpolated to a 12 km x 12 km grid structure using the kriging method explained above to match the model grid structure. To derive the final corrected ozone concentration value, the interpolated bias fields were added to the modeled values as follows:

\[ O_{i}^{corr} = \hat{C}_j + O_j^{mod} \]  

(1)

\( \hat{C}_j \) is \( C_i \) kriged to estimate the grid-cell center bias values. \( C_i \) is calculated as follows:

\[ C_i = O_i^{obs} - O_i^{mod} \]  

(2)

Where \( i \) refers to ozone monitor \( i \), \( j \) refers to the grid-cell center point of a 12 km x 12 km horizontal grid structure (variables are also indexed by day, but this designation is omitted for simplification); \( \hat{C} \) refers to the estimated kriged bias; \( O^{obs} \) refers to observed ozone concentration value; \( O^{mod} \) refers to modeled ozone concentration value; and \( O^{corr} \) refers to the corrected ozone concentration value.
**Multiplicative Bias Adjustment.** The multiplicative bias was calculated by dividing the observed value by the modeled value at each observation site for each day. Similar to the additive bias approach, the bias ratio was interpolated to a 12 km x 12 km grid structure using the kriging technique described above. However, because large ratios can result in those cases where the modeled value is small in comparison to the observed value, the ratios were log-transformed before interpolating and then back-transformed before multiplying the ratios by the model surface. The corrected ozone concentration values were calculated as follows:

\[ O_{j}^{corr} = \hat{C}_{j} \times O_{j}^{mod} \]  

(3)

\( \hat{C}_{j} \) is \( C_{i} \) kriged to estimate the grid-cell center bias values. \( C_{i} \) is calculated as follows:

\[ C_{i} = \frac{O_{i}^{obs}}{O_{i}^{mod}} \]  

(4)

Neither the additive or multiplicative bias adjustment approaches calculate an error estimate.

**Weighted-average.** This technique used the kriging interpolation method described above to calculate a gridded surface based on the observed values. The final estimated ozone concentration was calculated using a weighted average of the observation-based estimate, \( O_{j}^{kri} \), and the CMAQ output value, \( O_{j}^{mod} \). The following statistical model was used to combine these two sources of information about the true (unknown) ozone concentration, \( O_{j}^{true} \), at grid cell \( j \):

\[ O_{j}^{kri} = O_{j}^{true} + \varepsilon_{j}; \quad \varepsilon_{j} \sim (0, \sigma_{\varepsilon_j}^{2}) \]  

(5)

\[ O_{j}^{mod} = O_{j}^{true} + \eta_{j}; \quad \eta_{j} \sim (0, \sigma_{\eta_j}^{2}) \]  

(6)

This statistical model does not rely on any assumptions about the distribution of the errors, \( \varepsilon_{j} \) and \( \eta_{j} \), except that they each have mean zero and known variances, \( \sigma_{\varepsilon_j}^{2} \) and \( \sigma_{\eta_j}^{2} \), respectively. The final weighted average estimate at each grid cell is:
\[ O_j^{\text{corr}} = w_j O_j^{\text{krig}} + (1 - w_j) O_j^{\text{mod}} \] (7)

The weights are determined by minimizing the mean square error of \( O_j^{\text{corr}} \). Using this least squares approach, the weight factor is defined as:

\[ w_j = \frac{\sigma_{\eta_j}^2}{\sigma_{\epsilon_j}^2 + \sigma_{\eta_j}^2}. \]

The kriging analysis provides an estimate for the error variance \( \sigma_{\eta_j}^2 \), for each grid cell based on the covariance structure of the observed ozone. This estimated error variance accounts for measurement error in the observations and for uncertainty in the kriging prediction, due to sparseness of the monitoring network and the heterogeneity of the underlying unobservable ozone field. Since the uncertainty in the model output is more difficult to characterize, the error variance for the model values is held constant across all N grid cells and is set equal to the maximum kriging error variance for a given day:

\[ \sigma_{\eta_j}^2 = \max_j (\sigma_{\epsilon_j}^2) \] for \( j = 1, \ldots, N \). This choice for the model error variance was made because it produced the following properties for the final estimate. At locations where the kriging error variance is large (e.g. in regions of very few monitors), the kriging estimate will be given less weight than at other locations, but its weight is never less than half that of the model. In grid cells that contain a monitoring site, the kriging estimate will be weighted more heavily as compared to the modeled value. Thus this approach uses the error variances to quantify the relative quality or accuracy of the observation-based gridded concentrations compared to the model output. In addition, the error variance of the final estimate is the ratio:

\[ \frac{\sigma_{\epsilon_j}^2 \sigma_{\eta_j}^2}{\sigma_{\epsilon_j}^2 + \sigma_{\eta_j}^2}. \]

**Hierarchical Bayesian Modeling (HBM).** A Bayesian hierarchical space-time fusion modeling approach\(^\text{12}\) has been developed for integrating various sources of air quality data. This flexible model was developed to provide daily pollutant predictions over the continental U.S. for multiple years. In this application, the HBM model was applied to estimate ozone concentration values for the greater NYS domain (Figure 2.1.1) for the summer of 2001.
Bayesian analysis decomposes the modeling problem into linked stages: 1) air quality monitoring data; 2) CMAQ output; 3) measurement errors and CMAQ bias; and 4) the underlying “true” concentration surface. A Bayesian approach incorporates ‘prior knowledge’ of the unknown parameters which results in improved estimation of the uncertainty of the ‘true’ pollutant surface at any location in space and time. This model assumes that both monitoring data and CMAQ output provide good information about the same underlying pollutant surface, but with different measurement error structures. Discussion of the choice of parameters used and additional details on the overall HBM approach can be found in McMillan et al.12

2.1.8 Comparison of different techniques

The interpolated observations and combined surfaces resulting from each of the techniques described above were evaluated using cross-validation. Selection of cross-validation sites used in an evaluation can present many challenges. In this study, the number and location of the ozone monitors was relatively dense. However, because monitors are sometimes placed to determine compliance with regulatory exceedance thresholds, the monitors tend to be clustered around urban areas (Figure 2.1.2). As a result, random selection of monitors can result in a relatively large number of urban sites. This tendency can bias the results of the evaluation to favor the interpolated observations as interpolation will always perform best in those areas where there are many monitors. In addition, the clustering of monitoring sites in urban areas can result in under-representation of rural areas. In order to ensure that rural areas as well as urban areas were represented in the selection, the observation sites were overlaid on the 2000 Census Bureau urban metropolitan area boundaries using a Geographical Information System to determine whether the observations were in a rural or urban environment. Cross-validation sites were then selected in two steps: (1) 7 rural CASTNET monitoring sites were used for cross-validation; and (2) 20 sites were randomly selected from the AQS and NAPS networks for a total of 27 sites and 2,454 observations (Figure 2.1.3). Four of the 27 randomly selected AQS and NAPS sites were designated as rural, resulting in a total of 11 rural sites (7 CASTNET sites + 4 AQS/NAPS
sites) and 16 urban sites. These observations were set aside for cross-validation and all methods utilized the remaining observations to generate the combined surfaces.

![Image](image.png)

Figure 2.1.3: Location of cross-validation sites. Circles denote NAPS sites, triangles denote CASTNET sites and diamonds denote AQS sites.

The coefficient of determination ($R^2$), mean bias, and root mean square error (RMSE) were calculated between the observed value at the cross-validation site and the modeled output, interpolated observation and each of the three combination techniques for all 92 days of the summer. In addition, time series plots were generated to evaluate the error for each day at all sites. $R^2$, mean bias and RMSE were also compared for urban sites versus rural sites, and by network (AQS, NAPS and CASTNET), however, these latter analyses are not shown as they did not result in substantial differences among the different data combination techniques.

In order to assess how the various combination techniques compared across different percentiles, the observed concentrations used in the cross-validation were ranked by concentration level, and then binned by non-uniform percentiles ($0-50^{th}$, $50^{th}-75^{th}$, $75^{th}-90^{th}$,
90th-95th, and 95th-100). The matching cross-validation surface results for each technique were also binned, and the binned values were compared to the binned observations through scatterplots. In addition, the error (predicted – observed) for each technique was calculated and averaged for each bin, allowing for all techniques to be shown on one line plot. Finally, the spatial features of the combined surfaces were qualitatively assessed by comparing the spatial maps (i.e., concentration estimates at each 12 km x 12 km grid cell) produced by the three combination techniques, the model and the interpolated observations. For the spatial analysis, maps displaying the mean, median, various percentiles, standard deviation, coefficient of variation (standard deviation/mean) were examined. Only the most relevant of these maps are included in the paper.

2.1.9 Results and Discussion

Of the four combination techniques, the additive and multiplicative bias adjustment approaches were the easiest methods to use. The HBM approach was the most complex model, requiring specification of prior distributions for all model parameters. Related to this requirement, estimating the model error parameter for both the weighted-average and HBM approaches was problematic as this value is unknown. In the near future, however, use of ensemble runs may improve our ability to estimate the model variance. It should also be noted that the weighted-average and HBM approaches were the only combination techniques compared in the study that provided an estimate of predicted error. This estimate of predicted error can be important for some applications.

For the standard metrics examined, the cross-validation produced similar results for all four combination techniques and the kriged observations (Table 2.1.1). The RMSE for the four combination techniques and kriged observations was within 1.4 ppb of each other, and the mean bias was within 1.05 ppb of each other. R² ranged from 85 to 88 percent. The metrics also indicated that all four combination techniques substantially improved the modeled surface (R² of 0.66).
Table 2.1.1: Cross-validation results (daily maximum 8-hr average ozone concentrations).

<table>
<thead>
<tr>
<th>Method</th>
<th>RMSE (ppb)</th>
<th>Mean Bias (ppb)</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMAQ Only</td>
<td>11.70</td>
<td>0.88</td>
<td>0.60</td>
</tr>
<tr>
<td>Kriged Observations</td>
<td>6.40</td>
<td>0.49</td>
<td>0.88</td>
</tr>
<tr>
<td>Additive Bias Adjustment</td>
<td>6.80</td>
<td>0.33</td>
<td>0.86</td>
</tr>
<tr>
<td>Multiplicative Bias Adjustment</td>
<td>6.80</td>
<td>0.05</td>
<td>0.86</td>
</tr>
<tr>
<td>Weighted average</td>
<td>6.70</td>
<td>0.50</td>
<td>0.88</td>
</tr>
<tr>
<td>Hierarchical Bayesian</td>
<td>7.80</td>
<td>1.10</td>
<td>0.85</td>
</tr>
</tbody>
</table>

The percentile rank-ordered analysis, however, revealed interesting differences among the combination techniques and the interpolated observations. Figure 2.1.4 displays scatterplots (predicted versus observed), highlighted by color code to depict the percentile range. Note that the weighted-average and HBM techniques correct the model bias fairly well at the lower percentiles, but follow the scatterplot pattern of the raw modeled output at the higher percentiles. The additive and multiplicative bias techniques follow the one-to-one line closely, indicating that a simple correction of the model bias may be effective for improving the spatial characterization of ozone concentrations.
Figure 2.1.4: Observed (y-axis) and predicted (x-axis) 8-hr DMA zone (ppb) binned by percentile for cross-validation sites. Black = 0-50%, red = 51-75%, orange = 76-90%, blue = 91-95%, green = 96-100%.

Although cross-validation often favors kriging of the ozone observations (due to the concept discussed earlier of randomly selecting cross-validation data from clustered monitoring sites), the bias-adjustment techniques produce slightly better results than kriging at the higher percentiles. This same difference is evident in the error plots in Figure 2.1.5. The tendency of the CMAQ model to overestimate low ozone values and underestimate high ozone values can be clearly seen. Similar to the scatterplots, the weighted-average and HBM techniques appear to correct this overestimation at the lower ozone concentrations, but do not do as well at reproducing the observed ozone concentration levels at the higher percentiles (it should be noted that the HBM technique was designed to provide Bayesian predictions over large national spatial scales rather than the small regional domain of this study). The additive and multiplicative bias adjustment approaches appear to perform best across all percentiles, slightly out-performing the kriging of the observations at the higher percentiles as noted earlier. Although it is recognized that inferences from the smaller sample sizes in the higher percentile bins must be done with caution (sample size ranges from 370 to 123 site-days for the three highest percentile bins), the sample sizes range from 5% - 15% of the total sample size of 2,454 site-days, providing credence to the results discussed above.
Figure 2.1.5: Mean error (binned by percentile) between modeled, kriged and the four combination techniques versus mean observed concentrations for all cross-validation sites. Circles represent average mean for each binned percentile; lines are for identification of technique but do not represent linear relationships between averaged points.

In addition, the spatial texture seen with the combined surfaces indicates that model-based spatial information seen in CMAQ is retained with the combined surfaces. Figure 2.1.6 shows 8-hr DMA ozone values for the (a), CMAQ model (b) kriged observations and (c) multiplicative bias adjusted values for two representative days on June 7, 2001 and July 19, 2001. Note the overly smooth surface inherent in the interpolated observations and the influence of the model in the spatial texture of the combined surface.
Figure 2.1.6: Contribution of spatial information from model in combined surface of 8-hr DMA. Panel (a) is modeled surface, panel (b) is interpolated observations and panel (c) is multiplicative adjusted bias. Top panels are for June 7, 2001 and bottom panels are for July 19, 2001.

Figure 2.1.7 shows the mean concentration values for each grid cell. The model estimates of the ozone titration effect near NYC and Boston can clearly be seen in the modeled surface. The measurements in these same areas also show low ozone concentration values; however, these low values are averaged out by the interpolation. In addition, high ozone values are predicted by the model over the Great Lakes and the Atlantic Ocean due to lower planetary boundary layer heights, stable atmospheric conditions, and reduced turbulence and deposition; all physical processes known to exist over large waterbodies. Ozone measurements taken over Lake Michigan and aircraft observations over the coastal areas of the Northeast indicate the presence of high ozone concentrations over large waterbodies which the model seems to capture. Similar to titration in the urban core, the interpolated observations do not show this physical phenomenon. Since the purpose of this study is to provide improved air quality data for health studies, the difference in the estimation of titration is particularly relevant as this physical phenomenon can occur in highly-populated areas.
Figure 2.1.7: Mean 8-hr DMA ozone concentrations across all days for (a) modeled, (b) kriged observations, (c) multiplicative bias, (d) additive bias, (e) weighted average and (f) HBM.

The coefficient of variation (Figure 2.1.8) calculated for all days at each grid cell reveals a problem introduced by the large standard deviation values (relative to the mean) produced by the titration effect in both the additive and multiplicative bias adjustment approaches.
Figure 2.1.8: Coefficient of variation calculated for 8-hr DMA ozone concentrations across all days for (a) modeled, (b) kriged observations, (c) multiplicative bias, (d) additive bias, (e) weighted average and (f) HBM.

For example, the coefficient of variation is high for the multiplicative bias surface near Staten Island and Boston for which the model predicts low ozone concentration values due to titration. These high coefficient of variation values are the result of large differences between the modeled and observed concentrations that result in very large observed-to-modeled ratios (Eq 4; Figure 2.1.9a). Kriging the ratios and multiplying them by the model surface creates high ozone concentrations in the non-titrated area surrounding the titrated area (Figure 2.1.9b). While selection of a cross-validation site in one of the impacted areas may have changed the results, this effect occurs for less than 0.03% of the total concentrations and for only 3 days during the 92-day time period. In addition, the high observed-to-modeled ratios do not produce excessively high ozone concentrations except for over the ocean outside of Staten Island where the population is low.
2.1.10 Summary

The cross-validation results of this pair-wise comparison using standard statistical metrics did not reveal a large difference among the four combination techniques, but did reveal that all techniques provide improved estimates of 8-hr DMA ozone concentrations as compared to the model surface alone. The percentile analysis of the cross-validation results revealed interesting results not discerned by the all-days/all-sites metrics alone. The percentile analysis indicated that the additive and multiplicative bias adjustment techniques tended to improve the combined 8-hr DMA ozone concentrations at the higher percentiles as compared to the other techniques, including kriging the observations. Further analysis of the resulting spatial surfaces, however, revealed problems with the additive and multiplicative bias adjustment approaches introduced by the modeled titration effect, yielding artificially high ozone concentration values in adjacent cells. This problem, though, occurred for less than 0.03% of the total concentrations on only 3 days of the total summer, and primarily over waterbodies where the population is low. The qualitative spatial analysis performed supported that the combination techniques added spatial information from the model as compared to kriging the observations alone. In the case of this study, the intended application of the combination approach is to provide improved air quality surface maps for
conducting epidemiology studies in NYS. The additive and multiplicative bias adjustment approaches are considered appropriate for this application because; (1) accurately representing days of high-ozone concentrations is important for the health study of interest and the additive and multiplicative bias adjustment approaches out-performed the other methods at the higher ozone concentration percentiles, (2) the additive and multiplicative bias adjustment approaches are relatively simple and can readily be applied by the state health community, and (3) to date, estimates of predicted error produced by the HBM and weighted average approaches are not generally used in health studies. However, as epidemiology studies move towards the use of predictive distributions, more complex approaches such as HBM may be needed to estimate prediction error. Finally, these results are limited in applicability to the domain, pollutant and time period studied.

2.2 Comparison of Air Quality Surfaces in Epidemiology Model

The multiplicative adjusted bias approach outlined in the article above was applied to modeled and observed ozone concentrations to produce a combined surface of maximum daily ozone concentrations for 5 summers (2001 – 2005). The ozone concentrations were averaged for four counties in the New York City Metro Area (Bronx, Kings, New York and Queens counties; Figure 2.2.1). The study time period and domain were selected based on the availability of CMAQ output and because of the large number of hospital admissions available for these four counties. The data were used in the epidemiology model described in Section 4.0 and compared to using maximum daily 8-hr ozone concentrations averaged for each county (Garcia et al., 2010).
Figure 2.2.1: Four New York City counties and associated population used in comparison study.

The boxplots in Figure 2.2.2 (center line is mean, box ends are 25th and 50th percentile respectively, and the tails show the extent of the distribution and outliers) indicates that the combined surface is more variable than the distribution derived from the raw observations. Based on the analysis presented in the paper above, this is due to the increased spatial information provided by the model. In epidemiology studies, exposure misclassification can bias the results (explained in terms of risk) away from the null. Hence, capturing this additional spatial gradient may reduce exposure misclassification and result in a more accurate estimate of risk.
Figure 2.2.2: Boxplots of 3-day moving average of maximum daily 8-hr ozone concentrations using observations (red) and combined observations and model (Combined; green).

The risk was calculated for five metrics (same-day, lag 1 day, lag 2 days, lag 3 days and 3-day moving average) for each data type (averaged observations and combined). The results in Figure 2.2.3 indicate that the combined air quality data does add information relevant for discerning associations between ozone and respiratory-related hospital admissions. Note that the lag 1 day, lag 2 days and 3-day moving average all produce significant risk (confidence intervals are above the 1.0 line) as compared to using observations alone. Hence, for most metrics, using the combined ozone concentration surface increased the health effect measure, and in the case of the lag 1 day, lag 2 days, and 3-day moving average metrics, resulted in a significant finding.
Figure 2.2.3: Risk calculated using same-day, lag 1 – 3 days, and 3-day moving average using ozone observations and combined observations and model (Combined).
REFERENCES


3.0 Impact of Transported Pollution on Human Health

The article entitled “An Evaluation of Transported Pollution and Respiratory-related Hospital Admissions in the State of New York” is presented in this section to investigate approaches for capturing the potential effect of transported air pollution from the Midwest on respiratory-related hospital admissions. The paper has been accepted for publication in *Atmospheric Pollution Research* and is scheduled for publication in January, 2011.

3.1 “An evaluation of transported pollution and respiratory-related hospital admissions in the state of New York”

3.1.1 Abstract

Human exposure to air pollution transported from the Midwest is evaluated in eight New York State (NYS) regions over ten summers (1997 - 2006) for association with respiratory-related hospital admissions. Days when pollution is transported into the Northeastern United States (U.S.) are identified by using back-trajectories from the eight regions. These back-trajectories help identify predominant meteorological patterns associated with “polluted” air parcels (originating in the Midwest where power plant emissions are known to be relatively high) and “clean” air parcels (originating from the North where pollution is known to be relatively low). Ambient ozone concentration measurements are used to validate the classification of “polluted” and “clean” air parcels. These classifications are then used to define the days of high- versus low-exposure for populations residing within each region. The results of this analysis indicate that the risk of being hospitalized for respiratory-related illness in NYS is greater on those days when air is transported from the Midwest as compared to days when air is transported from the North. Using a refined method to examine air parcels moving through a boundary drawn around high-emitting power plants in the Midwestern U.S. resulted in stronger associations across more regions (significant odds ratios ranging from 1.06 to 1.16 for the entire study time period for six of the eight NYS regions). An assessment of temperature and its impact on the odds ratio calculation in the New York City metropolitan region indicates that temperature alone does not explain the increased association between air pollution and respiratory-related hospital admissions.
3.1.2 Introduction

The Clean Air Act (CAA) requires that United States (U.S.) Environmental Protection Agency (EPA) set National Ambient Air Quality Standards (NAAQS) for pollutants considered harmful to human health and the environment. Previous research has shown that high ambient ozone levels have harmful effects on humans (among others, Burnett et al. 2001; Medina-Ramon and Schwartz, 2006; Tolbert et al., 2007; Lin et al, 2008; Moore et al., 2008). The formation and distribution of ozone is driven by chemical interactions involving nitrogen oxides (NOx) and Volatile Organic Compounds (VOCs) in the presence of sunlight, as well as prevailing meteorological conditions. These pollutants and their precursors can be transported downwind, contributing to pollutant levels at locations much farther from the location of emission sources, potentially impacting human health in downwind areas.

The Ozone Transport Assessment Group (OTAG) conducted a two-year regional assessment of ozone transport (1997) and concluded that the central portion of the 37-state OTAG domain (Figure 3.1.1) was characterized by persistent elevated ozone levels producing an "ozone pool." High ozone levels in the southern portion of the OTAG domain were associated with near-stagnant pollutant conditions whereas high ozone levels in the northern portion of the OTAG domain were associated with fast-moving weather systems and persistent transport conditions from inside the OTAG domain--particularly in the Midwestern U.S. where major NOx emitting power plants are located (OTAG, 1997; Rao et al., 1997). In such cases, ozone is advected into the backside of a high-pressure system from the North causing the rapid southwesterly transport of elevated ozone levels from the Gulf to the Northeast (Wolff and Lioy, 1980).
Since the issuance of the OTAG Report, the implementation of the NOx Budget Trading Program (fully implemented in 2004) has significantly reduced NOx emissions from the power-generating sector in the Midwest (Gego et al., 2007; U.S. EPA, 2005). In its most recent evaluation of the nation’s air quality, the U.S. EPA found that the daily maximum 8-hr average ozone concentrations in the Eastern U.S. have declined by 10 percent since 2001 (U.S. EPA, 2010). While the predominant source of NOx emissions is now from on-road vehicles (36%), the power industry still accounts for 23% of the total emitted NOx in the Northeastern U.S. (U.S. EPA, 2005). Distinguishing transported pollution from locally emitted pollution is critical in improving our understanding of the health outcomes resulting from regulatory actions.

Previous studies have examined the transport of polluted air parcels from the Midwest into the Northeastern U.S. Ryan et al., (1998) used various monitoring data and back-trajectory analysis to examine the meteorology, chemistry and source of a severe regional ozone event during July 12 - 15, 1995 over the Baltimore-Washington region. They found that the chemical composition aloft (high ozone, sulfur dioxide and total reactive nitrogen and very low nitric oxide concentrations) indicated photochemical aging of an air parcel transported some distance with, at least partially, a coal combustion source. Brankov et al.
(1998) examined the role of synoptic-scale weather patterns in pollution transport and its influence on observed ozone concentrations at three locations in NYS, New Jersey and Massachusetts. They found that consistently high ozone levels in the Northeast tend to occur when air arriving at the monitoring sites has previously traveled over the Southeast and Midwest.

Taubman et al. (2004) used airborne observations of trace gases, particle size distributions and particle optical properties taken from New Hampshire to Maryland during a multi-day ozone-haze episode on August 14, 2002 and found that backward trajectories indicated source regions in the Midwest and Mid-Atlantic urban corridor. In a separate study, Taubman et al. (2006) used cluster analysis of back-trajectories in conjunction with vertical profile data to identify source regions and characterize transport patterns during summertime pollution episodes. The largest number of trajectories for these high-pollutant events lay over the northern Ohio River Valley. The resulting chemical composition of the air masses (high ozone values, large SO$_2$/CO ratios, highly scattering particles and large aerosol optical depths) supported this conclusion. Rainham et al. (2005) assessed whether meteorological conditions modified the relationship between short-term (daily) exposure to particulate matter and mortality using a hybrid spatial synoptic classification system. The authors found that although there does not appear to be any systematic patterning of modification, variation in pollutant concentrations seems dependent on the type of synoptic category present.

This study used two approaches to successively classify the transport of polluted air parcels for ten consecutive summer seasons (June, July and August) between 1997 and 2006. The objective of the study was to determine if air parcels transported from the Midwest into NYS lead to differentiable ozone concentrations and, if so, whether or not these transported air parcels were also associated with distinguishable daily respiratory-related hospital admissions. The study complements previous studies in two ways; first, it presents a methodology for identifying polluted air parcels versus relatively clean air parcels, and second, it uses these air parcel classifications to identify exposed and unexposed groups.
within eight NYS regions for purposes of evaluating the prevalence of respiratory-related hospital admissions between the two groups.

3.1.3 Air quality and health data

The daily maximum 8-hr averaged ozone concentrations for June 1 through August 31 for the years 1997 through 2006 were calculated from the hourly measurements archived at the U.S. EPA’s Air Quality System (AQS) database (http://www.epa.gov/oar/data/aqsdb.html) and Clean Air Status and Trends Network (CASTNET) (http://www.epa.gov/castnet/). The locations of the monitors used in the study are shown in Figure 3.1.2. The daily maximum 8-hr ozone concentrations were calculated by applying an 8-hr moving window to the hourly time series and selecting the 8-hr time window with the highest averaged ozone concentration value. Only those days having greater than 20 hours of data were used for computing the daily maximum.

Figure 3.1.2: Location of eight NYS regions, back-trajectory origination sites (red squares), CASTNet monitoring sites (black circles) and AQS monitoring sites (yellow triangles).

Hospital admission information was obtained from the NYS Department of Health Statewide Planning & Research Cooperative (SPARCS), which collects inpatient information for all NYS hospitals, excluding psychiatric and federal hospitals. SPARCS is a legislatively
mandated discharge database that is known to include at least 95% of acute care hospitalizations. Respiratory diseases were based on the International Classification of Diseases, 9th Revision (ICD-9 code; US Department of Health and Human Services, 1991), and included: asthma (ICD-9 code 493), chronic bronchitis (491), emphysema (492), and chronic obstructive pulmonary disease (COPD; 496).

3.1.4 Air pollution transport

This study involved the successive classification of transport directions into a limited number of relevant categories to target air parcels transported from the Midwest into NYS. The method implements various statistical and epidemiological techniques to examine the differences in ambient ozone concentrations and daily respiratory-related hospital admissions in correspondence with the selected transport categories.

The calculation of the transport direction was performed with the Hybrid Single-Particle Lagrangian Integrated Trajectory model (HYSPLIT; Draxler and Hess, 1997). HYSPLIT was set to calculate back trajectories, i.e., to perform the calculation back in time to determine the origin of the transported air parcel. The model was run repetitively to simulate transport every day during the 10 summers considered (total of 920 days). In addition, eight starting locations approximately coinciding with the center of the eight regions in NYS (http://www.dec.ny.gov/chemical/34985.html) were utilized for each simulated day. A total of 7360 trajectories (920 days x 8 starting locations) were examined, however, eight days on average for each region were eliminated from further interpretation because simulations ended prematurely (less than 48 hours simulated) due to missing meteorology data.

For all simulations, the starting elevation for the trajectory was fixed at 500 m above ground level to reproduce transport across the Appalachian Mountains in the atmospheric residual layer aloft at night, yet represent surface-level ozone concentrations through vertical mixing during the day. Vertical motion was reproduced using the vertical velocity fields that are produced by the meteorological model rather than the more restrictive isentropic or isobaric displacement options also available to simulate transport. Initial time was set in the afternoon (1800 UTC) to correspond to peak ozone times when pollutants are mixed.
uniformly throughout the daytime atmospheric boundary layer (Ryan et al., 1998; Rao et al., 2003). Simulated time was set to 48 hours to allow enough time to capture regional transport patterns, but minimize errors that accumulate with the length of simulation (Taubman, 2006; Stohl and Seibert, 1998). The meteorology fields used to run HYSPLIT included hourly pressure, vertical height, wind speed and wind direction. These data were obtained from the North American Regional Reanalysis (NARR) dataset (http://nomads.ncdc.noaa.gov/data.php#narr_datasets).

3.1.5 Classification of transport directions

It is well known that air flow directions vary continuously in time and space, covering the entire 360 degree spectrum about each starting location. Yet, numerous statistical methods require all data to be classified into a limited number of categories (the major transport directions) prior to being processed. Deciding on the number of categories to form as well as setting their limits contains some subjectivity. Two options were used to perform this task. The first option, hereafter referred to as the arc method, relies on the calculation of the azimuth; the angle measured from 0 degrees of the starting trajectory point to its endpoint location 48 hours back in time.

Five major directions (Figure 3.1.3) were identified with this method. The northerly direction was more broadly defined as one group because ozone concentration levels transported from the north (over Canada) are, in general, low. Higher (but more variable) ozone concentration levels are associated with transported air masses from the westerly, southwesterly, and southerly directions, making it important to further delineate these wind patterns in defining exposure to pollution—particularly since this study seeks to investigate exposure from transported pollution from power plant emissions in the Midwest (Ryan et al., 1998; Brankov et al., 1998; Taubman et al., 2004; Taubman et al., 2006). In general, the westerly arc class isolated emissions from major industries across the Great Lakes, the southwesterly arc class targeted major power plant emissions in the Midwest, and the southerly arc class targeted mobile source emissions transported north along the Eastern seaboard transport corridor. There was no particular rationale for defining the southeasterly arc except that it was the remaining arc after classifying the other critical directions.
The second classification method refines the methodology described above to further isolate transported pollution from the Midwest. It relies on the definitions of two longitude-latitude bounded zones representing high power plant emissions in the Midwest and relatively cleaner air originating in the North (Figure 3.1.4). A given day is said to correspond to transport from the Midwest if the trajectory calculated for that day enters the bounded zone at any time during the 48-hr trajectory simulation. Note that the arc method investigates directions calculated after 48-hrs exactly (and not continuously). The two options will not lead to completely independent results as each zone is part of a single broader arc classification. However, the zone and arc classifications will not lead to the same results for every simulated day. Slow moving air masses may fit the arc method requirement to be classified in a given class but may not be included in the matching zone classification because they did not reach back far enough to enter the zone limits. Conversely, a trajectory that curves during the 48 simulated hours may not be included in a given transport class as
defined by the arc method but may enter the matching zone boundaries later on during the 48-hr simulated period.

Figure 3.1.4: Bounded zone approach. Panel (a) shows high NOx emissions (July 1997 shown as an example) originating from power plants in Midwestern U.S.; (b) shows boundaries drawn for identifying days when the air parcel passed through a zone in the Midwest versus days it passed through a zone in the North (red rectangles). Also shown is the southern zone used in the temperature analysis.

For this study, using the two successive approaches to define the origin of the polluted air parcel was important because it allowed us to examine the levels of pollution transported from all directions using the arc method first before applying the more targeted bounded zone method. While the boundaries for the zone approach were established a priori, this first evaluation was necessary to substantiate that relatively polluted air was transported into NYS from the southwest as opposed to other directions.

3.1.6 Assessment of corresponding ozone concentrations and pollutant transport directions

Daily maximum 8-hr ozone concentrations were averaged for each region and matched to their respective air transport direction for each day for each of the eight NYS regions. For the arc method, the mean ozone concentrations for those days classified as southeasterly, southerly, southwesterly or westerly were compared to the mean ozone concentration for those days classified as northerly using the two-tailed t-test with a 95% confidence level. The same analysis was repeated for the zone method except that only those
days classified as passing through the southwesterly zone versus the northerly zone were used. The purpose of this examination was to identify polluted air parcels (relative to the cleaner air parcels from the North) and to substantiate that air parcels from the Midwest represented relatively higher ozone for determining days of exposure for assessing associations with respiratory-related hospital admissions.

3.1.7 Assessment of hospital admissions and transport directions

Daily respiratory-related hospital admissions were summed for each region and matched to their respective air parcel direction for each day for each of the eight NYS regions. The daily respiratory-related hospital admissions associated with air parcels originating from the Midwest and the North were used to define the exposed and unexposed groups, respectively, for calculating an unadjusted odds ratio. An odds ratio calculation indicates the odds of an event occurring in the exposed group relative to the odds of it occurring in the unexposed group (Rothman, 2002). “Unadjusted” odds ratio indicates that the calculation does not adjust for other variables that may impact the results (see Discussion Section). For this study, the calculation estimates the odds that a respiratory-related hospital admission will occur on days when air is transported from the southwest relative to the odds of it occurring on days when air is transported from the north. An odds ratio greater than 1 indicates that hospital admissions are more likely to occur on days when the air parcel originates from the southwest, and an odds ratio of less than 1 indicates that hospital admissions are less likely to occur on days when the air parcel originates from the southwest. The 95% confidence interval (95% CI) was also used to describe the precision of each odds ratio.

While the unit of comparison was daily, the odds ratio was calculated for all days (total number of days that polluted air was transported into the region versus the total number of days that relatively clean air was transported into the NYS region). Since the number of exposed days varied from the number of unexposed days, the summed hospital admissions for each group were normalized by the total number of days.

3.1.8 Examination of temperature
The effect of temperature was further investigated for Region 2 by defining days of high exposure using two different bounded zone classes (southwesterly and southerly), both of which had relatively high temperatures, and then comparing the odds ratios for these two groups. The intent of this analysis was to determine whether temperature, rather than pollution or some other factor, could be responsible for an odds ratio greater than 1.0. The boundaries for the southerly zone were drawn to represent the relatively longer-range transport from the southerly direction; the southwesterly zone was defined as described earlier (Figure 3.1.4). The temperature was compared between the southwesterly and southerly zones by averaging the maximum temperature for all days that the air parcel passed through the respective zone. The odds ratio was then calculated for both the southwesterly and southerly exposure classifications relative to the bounded northerly zone.

3.1.9 Results

The boxplots in Figure 3.1.5 display the ozone concentrations averaged for the northerly and southwesterly transport directions (defined by the arc and bounded zone methods) for each of the eight NYS regions. The center-line of the box identifies the median ozone concentration, the lower and upper extent of the box represents the 25th and 75th percentile, the upper and lower whiskers represent the 5th and 95th percentiles, and the dashed line represents the 10-year summer average for all eight NYS regions. Regardless of whether the southwesterly direction was defined by the arc or bounded zone method, the mean ozone concentration for the southwesterly direction was significantly higher than the ozone concentration for the northerly direction for all eight regions (p-value < 0.0001 for each region regardless of the method).
Figure 3.1.5: Boxplots displaying the daily maximum 8-hr ozone concentrations. Ozone concentrations for (a) northerly direction defined using the arc method; (b) southwesterly direction defined using the arc method; (c) northerly direction defined using the bounded zone method; and (d) southwesterly direction defined using the bounded zone method.

Table 3.1.1 shows the results of the odds ratio calculations comparing the exposed group with the unexposed group. Table 3.1.1 also shows that the odds ratios (including 95% confidence intervals) are statistically significant for Regions 2, 7 and 8 when the exposed and unexposed groups are defined using the arc approach. The odds ratios are also statistically significant for Regions 1, 2, 3, 6, 7 and 8 in comparing the exposed group to the unexposed group using the bounded zone approach.
Table 3.1.1: Odds ratio calculations resulting from the arc and bounded methods of determining exposure.

<table>
<thead>
<tr>
<th>Region</th>
<th>Odds Ratio</th>
<th># Days</th>
<th># Hospital Admissions</th>
<th># Days</th>
<th># Hospital Admissions</th>
<th>Total Population</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(95% Confidence Intervals)</td>
<td>(Exposed)</td>
<td>(Exposed)</td>
<td>(Unexposed)</td>
<td>(Unexposed)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1.02 (0.98 to 1.07)</td>
<td>228</td>
<td>4,070</td>
<td>581</td>
<td>10,165</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.04 (1.02 to 1.06)</td>
<td>197</td>
<td>17,660</td>
<td>599</td>
<td>51,691</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.05 (0.97 to 1.13)</td>
<td>169</td>
<td>1,407</td>
<td>601</td>
<td>4,050</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.02 (0.96 to 1.08)</td>
<td>262</td>
<td>2,423</td>
<td>567</td>
<td>5,151</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1.08 (0.98 to 1.18)</td>
<td>308</td>
<td>914</td>
<td>508</td>
<td>1,398</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>1.05 (0.99 to 1.12)</td>
<td>261</td>
<td>1,969</td>
<td>541</td>
<td>3,480</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>1.04 (1.00 to 1.09)</td>
<td>257</td>
<td>4,031</td>
<td>532</td>
<td>6,926</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>1.07 (1.02 to 1.13)</td>
<td>318</td>
<td>3,999</td>
<td>504</td>
<td>5,018</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Region</th>
<th>Odds Ratio</th>
<th># Days</th>
<th># Hospital Admissions</th>
<th># Days</th>
<th># Hospital Admissions</th>
<th>Total Population</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(95% Confidence Intervals)</td>
<td>(Exposed)</td>
<td>(Exposed)</td>
<td>(Unexposed)</td>
<td>(Unexposed)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1.09 (1.02 to 1.17)</td>
<td>100</td>
<td>1,836</td>
<td>172</td>
<td>2,889</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.06 (1.03 to 1.09)</td>
<td>114</td>
<td>10,374</td>
<td>161</td>
<td>13,825</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.16 (1.04 to 1.29)</td>
<td>92</td>
<td>701</td>
<td>196</td>
<td>1,287</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.06 (0.98 to 1.18)</td>
<td>114</td>
<td>1,098</td>
<td>262</td>
<td>2,555</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1.13 (1.04 to 1.36)</td>
<td>76</td>
<td>238</td>
<td>369</td>
<td>1,022</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>1.13 (1.03 to 1.24)</td>
<td>130</td>
<td>937</td>
<td>303</td>
<td>1,832</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>1.08 (1.02 to 1.14)</td>
<td>173</td>
<td>2,427</td>
<td>202</td>
<td>2,624</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>1.10 (1.03 to 1.16)</td>
<td>211</td>
<td>2,312</td>
<td>138</td>
<td>1,378</td>
</tr>
</tbody>
</table>

The minimum, medium and maximum temperatures for the southerly arc approach were the same or higher than for the southwesterly zone approach (22° C, 26° C and 29° C versus 18° C, 24° C and 29° C, respectively). The odds ratio, however, is lowest for the southerly zone as compared to the odds ratio calculated for the southwesterly zone (Figure 3.1.6).
3.1.10 Discussion

In this study, the population for both the exposed and unexposed groups is the same because the exposure is defined by day. Therefore, several variables that may confound or modify the effect of the exposure (e.g., socioeconomic status, gender, age, day-of-week, holidays) are controlled by the design of the analysis. The study does not, however, investigate the potential time lag between the exposure and the health endpoint. This weakness in the study design is somewhat compensated for, though, because the time scale for synoptic weather changes is usually 3 – 5 days (Rao et al., 1997). Since the unadjusted odds ratio is a measure of prevalence (number of respiratory-related hospital admissions for the entire study time period), the lag effect is partially incorporated into the study design because of the length of time that the exposure exists. Long-term trends (e.g., climate change), however, may be a confounder (related to both the exposure of interest and the health endpoint), but it is unlikely that the long-term trend signal embedded in the days of

Figure 3.1.6: Odds ratio. Panel (a) boxplots show the interquartile range and 5th and 95th percentiles for maximum temperature in Region 2 on days when air transport is from the south and the southwest (as defined by the bounded zone method). Panel (b) shows the odds ratio calculations for these two exposure definitions.
high exposure would be different from the trend embedded in the days of low exposure. Local pollution (versus pollution transported over relatively long distances) may also influence the results of the study, especially when using the arc approach. While the bounded zone approach targets transported pollution more effectively than the arc method, local pollution may still be associated with wind flow from a particular direction. However, the likelihood of this exposure misclassification occurring across all eight NYS regions is unlikely.

The analysis of the ozone concentrations associated with each arc or bounded zone direction supports that air parcels originating in the Midwest are associated with higher ozone concentration levels. For some regions, ozone concentrations were also significantly different between air parcels originating from the south or west as compared to air parcels originating from the north. However, air parcels originating from the southwest were significantly different for all eight NYS regions, regardless of the method used. Thus, the use of the southwesterly and northerly air parcel origin classification to define the exposed and unexposed groups for calculating the odds ratio is substantiated by this analysis because the ozone concentrations are significantly higher for the exposed group as compared to the unexposed group.

The unadjusted odds ratio calculation for the arc approach indicates that respiratory-related hospital admissions were elevated on southwesterly wind flow days for all regions over the entire study time period, with significant associations for regions 2, 7 and 8 (Figure 3.1.7a). The number of exposed days was less for the bounded zone approach than for the arc approach because there are fewer days that the air parcel passes through the Midwest zone as compared to originating from the southwesterly direction. Despite the wider confidence intervals (particularly for Region 5), the odds ratios increased for all regions when using the bounded zone approach to define exposure as compared to the arc approach, and became significant for regions 1, 3 and 6 in addition to regions 2, 7 and 8 (Figure 3.1.7b). These results indicate that respiratory health is impacted in regions 1, 2, 3, 6, 7 and 8 when air parcels are transported from the Midwest zone into NYS.
Figure 3.1.7: Odds ratios with 95% confidence intervals for panel (a) arc approach and panel (b) bounded zone approach.

Despite the similarity in the minimum, medium and maximum temperatures for the southwesterly and southerly zones, the odds ratio calculations revealed a lower risk for exposure to air parcels originating from the south versus air parcels originating from the southwest (Figure 3.1.5). This analysis indicates that there is more influencing the odds ratios calculations than temperature alone, providing evidence that polluted air parcels transported from the Midwest zone contribute to the incidence of respiratory-related hospital admissions in NYS.

3.1.11 Summary

Focusing on NYS and its inhabitants, this study applied two successive methodologies (arc and bounded zone) to classify polluted air transported from the Midwest and relatively clean air transported from the North for ten consecutive summers between 1997 and 2006. Daily averaged ambient ozone measurements were used to validate the classification methods, and the classes were used to define exposed and unexposed groups for purposes of calculating the odds of respiratory-related hospital admissions occurring in the respective groups.
The study results revealed spatial differences in associations between the origin of the air parcel and daily maximum 8-hr ozone concentrations. However, the mean ozone concentration and the direction of the air parcel origin were consistently and significantly different for all eight regions for the southwesterly versus northerly directions. The bounded zone methodology was used to isolate transported air from the Midwest, and the unadjusted odds ratio using this method indicated positive associations between respiratory-related hospital admissions and exposure to polluted air parcels for all regions. The results were significant (lower extent of confidence interval > 1.0) for regions 1, 2, 3, 6, 7 and 8. These results indicate that exposure to air parcels transported from the Midwest into NYS (as compared to air parcels that originate in the North) result in excessive risk for increased respiratory-related hospital admissions over the entire 10-year study time period for regions 1, 2, 3, 6, 7 and 8. These findings are consistent with the southwesterly transport patterns seen in NYS. The temperature analysis supported that temperature alone does not explain this increased association for Region 2 during the time period studied.

3.2 Cluster Analysis

The paper above presents two methods for grouping the back-trajectories used in the study (arc and bounded zone approach). In addition to these two methods, the trajectories were also clustered using Ward’s hierarchical clustering method available through the R language for statistical computing (R Development Core Team, 2007; Figure 3.2.1). In general, hierarchical clustering groups the trajectories into clusters, so that the trajectories within each cluster are more closely related to one another than the trajectories assigned to different clusters. The agglomerative method starts by joining the two closest objects together. Each subsequent pass of the newly created clusters joins the clusters nearest together until the maximum number of clusters is achieved. The selected cluster threshold was determined by examining the cluster-number at which there was a large percentage of change in the RMSE. In addition, previous trajectory studies were consulted (Taubman et al., 2006, Brankov, et al., 1998). Based on this evaluation, a threshold of 8-9 clusters was selected.
After examining the associations with ozone concentration levels and respiratory-related hospital admissions, however, it was found that the clustering technique did not classify distinct air parcels with regard to ozone concentration. This may be because the clustering technique tended to combine stagnant and faster moving wind flows into one cluster. Increasing the number of clusters resulted in better delineation of these wind flow patterns, but produced too many classifications for a meaningful analysis. Thus, the results from applying the arc and bounded zone method for classifying the trajectories was retained and the clustering technique was dropped from further analysis.
REFERENCES


Ozone Transport Assessment Group, (1997). Air Quality Analysis Workgroup’s Volume 1: Executive Summary, (Guinnup, D., Collom, B. (Ed)).


4.0 Assessing the Health Impacts of the NOx Budget Trading Program (NBP)

For the unadjusted odds ratio calculation applied in Section 3.0, the population for both the exposed and unexposed groups is the same because only the exposure to relatively polluted or clean air masses changes for each day. Therefore, several variables that may confound or modify the estimated effect measure (e.g., socioeconomic status, gender, age) are controlled by the design of the unadjusted odds ratio analysis. However, temperature and dew point are variables that are both associated with hospital admissions and with each other (confounding) that cannot be accounted for in an unadjusted odds ratio analysis. Since the time periods for assessing the NBP are relatively short (4 and 3 summers, respectively) and the inter-annual variability of the health effects (hospital admissions) and explanatory variables (e.g., ozone, temperature) is relatively high, a Generalized Additive Model (GAM) is applied to control for these and other unknown variables. The results are compared to the unadjusted odds ratio calculation, and are then applied to the pre- and post-NBP time periods to examine whether respiratory-related hospital admissions have gone down between the two time periods. The manuscript below describes the approach and the analysis in detail.

4.1 “Method to Track the Impact of Regional-Scale Air Quality Regulations on Human Health Applied in New York State”

4.1.1 Abstract

In its mission to protect human health and the environment, the United States (U.S.) Environmental Protection Agency (EPA) implemented the NOx Budget Trading Program (NBP) to reduce the emissions of nitrogen oxides (NOx) and the secondarily formed ozone. These pollutants and their precursors can be transported downwind, contributing to pollutant levels at locations much farther from the emission sources, potentially impacting human health in downwind areas. This study investigated the health impacts in New York State (NYS) from exposure to polluted air parcels transported from the Midwest. Back-trajectories were performed from several sites within NYS for ten summers (June, July and August from 1997 to 2006) to identify days that an air parcel passed through the boundaries of a defined Ohio River Valley (ORV) zone within a previous 48-hr time period. The ORV zone is a boundary encompassing relatively high-emitting power plants in the Midwest and was used
as an indicator variable to represent the transport of ozone and its precursors from this area into NYS. The ORV zone variable and the maximum daily 8-hr average ozone concentrations were used as the main effects in a Generalized Additive Model (GAM) to investigate potential associations between these two variables and respiratory-related hospital admissions. The results indicate that the risk of being admitted to the hospital for a respiratory-related illness on those days that air parcels are transported from the Midwest is elevated in Regions 2, 3 and 6 and that the risk is elevated for ozone (per 10 ppb increase) in Regions 2 and 4. Two time periods before the implementation of the NBP and after the implementation NBP (summers of 1997 – 2000 and 2004 – 2006, respectively) were also examined, but an analysis of the number of summers needed to obtain consistent results revealed that there is not enough data to discern a difference in the health signal between the two time periods. According to this analysis, data spanning at least 5 – 6 summers (460 – 552 days) is needed to produce consistent results from the GAM.

4.1.2 Introduction

Through the Clean Air Act (CAA; 42 USC §7401) and the National Ambient Air Quality Standards (NAAQS), the United States (U.S.) Environmental Protection Agency (EPA) establishes thresholds that must be obtained by each State to control the ambient concentration levels of criteria pollutants. In general, criteria pollutants are those atmospheric contaminants considered to be more ubiquitous, pose a risk to a larger fraction of the general population, and have more widespread impacts on ecosystems and natural resources as compared to other hazardous air pollutants (NRC, 2004). Several studies have shown a positive association between ozone (the criteria pollutant of interest in this study) and respiratory-related health effects (among others, Burnett et al., 1997; Burnett et al. 2001; Schwartz, 2005; Medina-Ramon et al., 2006; Tolbert et al., 2007; Lin et al, 2008; Moore et al., 2008). In addition, Ito et al. (2005), Bell et al. (2005) and NYSERDA (2006) found that accounting for co-pollutants (e.g., particulate matter) did not reduce the estimated effect measure of ozone on respiratory-related disease. Rainham, et al. (2005) assessed whether meteorological conditions modified the relationship between short-term (daily) exposure to particulate matter and mortality using a hybrid spatial synoptic classification system. The
authors found that subtle changes in meteorological conditions (e.g., wind speed, temperature, relative humidity) can alter the strength of pollutant associations with health outcomes, especially in the summer season.

While ozone is not directly emitted, the formation of ozone is driven by chemical interactions that occur in the presence of sunlight involving nitrogen oxides (NOx) and Volatile Organic Compounds (VOCs). NOx emitted into the upper atmosphere (particularly at night) from sources such as electric utilities with tall stack heights can be transported aloft hundreds of kilometers downwind (Godowitch, et al., 2008; Civerolo, et al., 2003; Brankov, et al., 1998; OTAG, 1997). The Ozone Transport Assessment Group (OTAG) conducted a 2-year regional assessment of ozone transport and concluded that the central portion of the 37-state OTAG domain was characterized by persistent elevated ozone levels producing an "ozone pool" (Figure 4.1). Transport in any direction from this region was implicated with high ozone levels in neighboring areas. High ozone levels in the southern portion of the OTAG domain were associated with stagnant pollutant conditions whereas high ozone levels in the northern portion of the OTAG domain were associated with higher wind speed and persistent transport conditions from inside the OTAG domain--particularly the Ohio River Valley where major NOx emitting power plants are located (OTAG, 1997). In such cases, ozone is advected into the backside of a high-pressure system from the North causing the rapid southwesterly transport of elevated ozone levels from the Gulf to the Northeast (Wolff and Lioy, 1980).
The NO$_x$ Budget Trading Program (NBP), issued by the EPA in 1998 (also known as the NOx State Implementation Plan Call), was the first regulation to specifically address the transport of pollutants in the Eastern U.S. by reducing summertime NO$_x$ emissions from major sources (predominantly the electric utility sector). Partially implemented in 2001 and fully implemented in 2004, several studies have already shown that the NBP has reduced ozone concentrations at downwind sites (Gego, et al., 2007, US EPA, 2010). This study further evaluates the impact of the NBP by first examining the associations between transported air pollution and respiratory-related hospital admissions by defining an indicator variable for use in a Generalized Additive Model (GAM), and then applying this indicator variable to examine the change in the association before and after implementation of the NBP. Ozone was also examined as a main health effect in the study.

### 4.1.3 Air quality and health data

The daily maximum 8-hr average (8-hr DMA) ozone concentrations for June 1 through August 31 for the years 1997 through 2006 were calculated from the hourly measurements archived at the U.S. EPA’s Air Quality System (AQS) database (http://www.epa.gov/oar/data/aqsdb.html) and Clean Air Status and Trends Network.
(CASTNET) (http://www.epa.gov/castnet/). The locations of the monitors used in the study are shown in Figure 4.1.2. The 8-hr DMA ozone concentrations were calculated by applying an 8-hr moving window to the hourly time series and selecting the 8-hr time window with the highest averaged ozone concentration value for each day. Only those days having greater than 20 hours of data were used for computing the daily maximum.

Figure 4.1.2: Location of eight NYS Regions, back-trajectory origination sites (red squares), CASTNet monitoring sites (black circles) and AQS monitoring sites (yellow triangles).

Hospital admission information was obtained from the NYS Department of Health Statewide Planning & Research Cooperative (SPARCS), which collects inpatient information for all NYS hospitals, excluding psychiatric and federal hospitals. SPARCS is a legislatively mandated discharge database that is known to include at least 95% of acute care hospitalizations. Respiratory diseases were based on the International Classification of Diseases, 9th Revision (ICD-9 code; US Department of Health and Human Services, 1991), and included: asthma (ICD-9 code 493), chronic bronchitis (491), emphysema (492), and chronic obstructive pulmonary disease (COPD; 496). Two extreme hospital admissions days (August 14 and 15, 2003) were removed and replaced with average values as these dates represent a widely distributed loss in power across the City during a heat wave (high hospital admissions were experienced due to excessive heat). In addition, the first day of every
summer month for 2002, 2003 and 2004 was replaced with average values as these dates had duplicate entries.

4.1.4 Approach

Back-trajectories were simulated from centrally-located locations within 8 NYS regions (http://www.dec.ny.gov/chemical/34985.html; Figure 4.1.3) to identify whether an air parcel passed through the coordinates of a boundary drawn around major NOx emitters in the Ohio River Valley (ORV) 48 hrs back in time. The Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT; Draxler and Hess, 1997) model was set to calculate back trajectories, i.e., to perform the calculation 48 hours back in time to determine the origin of the transported air parcel. The model was run repetitively to simulate transport every day during the 10 summers considered (total of 920 days). Out of a total of 7,360 trajectories (920 days x 8 starting locations), eight days on average for each region were eliminated from further interpretation because simulations ended prematurely (less than 48 hours simulated) due to missing meteorology data.

Figure 4.1.3: Bounded zone approach. Panel (a) shows high NOx emissions (July 1997 shown as an example) originating from power plants in the Midwestern U.S.; (b) shows boundary drawn around for identifying days when the airmass passed through a zone in the Midwest.
The trajectory data were matched in time to 8-hr DMA ozone concentration measurements and daily hospital admission counts for respiratory-related illness. The trajectory data were used to determine whether an air parcel passed through a boundary drawn around large NOx emitters in the Midwest (bounded zone approach; Figure 4.1.3). This, in turn, defined an indicator variable for use in a multi-variate regression model to represent exposure from transported air parcels from the Midwest. This process of calculating back-trajectories and identifying days when the air parcel was transported from the Midwest was repeated for the 8 NYS regions. A three-day moving average of ozone 8-hr DMA was also used as a main effects variable in the GAM. More details on the back-trajectories, the bounded zone approach and matching of the hospital admissions, ozone concentrations and ORV variable can be found in Garcia et al. (2011).

Calculating Risk. The GAM design was selected to examine associations between respiratory-related hospital admissions and air parcels from the ORV because this model design allows for covariates to have different distributions (e.g., binomial, Gaussian, log, etc.) and relationships (linear, non-linear). Specifically the ‘mgcv package’ available through the R Project for Statistical Computing (http://www.r-project.org/) provides functions for GAMs. The GAM used in this study applies a quasi-likelihood maximization, with automatic smoothness selection by generalized cross-validation (GCV; Wood, 2010). This flexible model type was used in this study because some variables included in the model are known to have a non-linear relationship with the health endpoint (Eq. 1):

\[
Y_t \sim \text{Poisson}(\mu_t) \\
\ln(\mu_t) = \beta_0 + \beta_1 \text{Moving avg} + \beta_2 \text{ORV}_t + s_3(\text{maxtemp}_t; df \sim 2.5) + s_4(\text{dewpoint}_t; df \sim 3) + \beta_5 \text{holiday}_t + \beta_6 \text{dow}_t + \beta_7 \text{year}_t + s_8(\text{Date}_t; df = 3 \times \# \text{ years})
\]

(Eq. 1)

In Eq. 1, \( \beta_0 \) is the natural log of the number of hospital admissions per day per region, given the other variables in the model; \( \beta_1 \text{O}_3\text{SMA}_t \) is the 3-day moving average of the 8-hr DMA ozone concentration for each day and region; \( \beta_2 \text{ORV}_t \) is a dichotomous variable used to identify days that the air parcel passed through the ORV bounded zone versus days it did not;
\( s_3(\text{maxtemp}_t) \) is a spline function applied to the average maximum temperature across each day and region (degrees of freedom [df] optimized by model at \( \sim 2.5 \) per year); \( s_4(\text{dewpoint}_t) \) is a spline function applied to the average dew point for each day and region (df optimized by GCV at \( \sim 3.0 \) per year); \( \beta_3 \text{holiday}_t \) is a dichotomous variable used to indicate holidays; \( \beta_6 \text{dow}_t \) is the day-of-week included to account for the delay in admissions over the weekend and other weekday effects; the year variable (\( \beta_7 \text{year}_t \)) accounts for inter-annual variability; and \( s_8(\text{Date}_{t, \text{year}}) \) is a spline function applied across month year\(^{-1}\) to account for the seasonal and long-term variability in the model (df set at 3.0 per year). The ozone variable (\( \beta_1 \text{O}_3 \text{Moving avg} \)) is a 3-day simple moving average (SMA) calculated in R as the average of the same day, lag 1 day previous, and lag 2 days previous. This variable (as opposed to same day, lag 1 day, lag 2 days and lag 3 days) provided the best fit of the model and lowest autocorrelation. Fine particulate matter ambient concentrations were also considered as a covariate in the model, but did not have a significant health effect and did not change the results substantially, so were left out of the final model.

The coefficient for the two main effects (ozone concentration and ORV zone variable) are used to derive an estimate of risk. Ozone risk is calculated by multiplying the natural log coefficient by 10 to report risk as “per 10 ppb increase”, and then transforming the value from natural log. For this study, the calculated risk can be interpreted as the risk that a respiratory-related hospital admission will occur per 10 ppb increase in ozone exposure. The risk of exposure to the transported air parcels (the ORV indicator variable) is interpreted as the risk that a respiratory-related hospital admission will occur when exposed to air transported from the ORV. Because a natural logarithm distribution is assumed for hospital admissions, a risk greater than 1.0 indicates that hospital admissions are more likely to occur as a result of the exposure, and a risk less than 1.0 indicates that hospital admissions are less likely to occur as a result of the exposure. The 95% confidence interval (95% CI) was also used to describe the precision of each risk value.

**Analysis of all summers:** Descriptive statistics (mean, interquartile range, coefficient of variation) were performed to generally analyze the data used in the study. Informative results were selected for presentation in this paper. Hospital admissions were normalized by
population to better examine the differences in rates due to other factors besides population. The risk using the ozone variable and the ORV indicator variable was calculated for the entire time period of 10 summers using the method described above and is compared to the unadjusted odds ratio method described in Garcia et al. (2011).

_Pre-NBP versus post-NBP implementation time periods:_ The GAM was applied to all data (summers of 1997 through 2006) and to data for a pre-NBP time period (summers of 1997 through 2000) and a post-NBP time period (summers of 2004 through 2006). [The summers of 2001 through 2003 were not used in the pre-NBP versus post-NBP analysis because the power plants were partially implementing emission controls mandated by the NBP during these years.] The GAM was run separately for each of the two time periods and the difference in risk between them was examined. Residuals were checked for autocorrelation and other patterns indicating problems with colinearity. Temperature and ozone were relatively highly correlated ($r = 0.6$). No other variables had this high of a correlation (including temperature and the ORV indicator or ozone and the ORV indicator). Even though temperature is relatively highly correlated with ozone, it is a strong explanatory variable for hospital admissions and was included in the model consistent with past studies.

Since the number of samples is reduced when considering the pre-NBP time period (4 summers of data) and post-NBP time period (3 summers of data), an inter-annual analysis was conducted to assess the impact of the sample size on the calculation of the health endpoint. The model was run with successively less data (10 summers, 9 summers, 8 summers...2 summers) to assess how stable the model performed as data were removed.

### 4.1.5 Results

**Analysis of all summers:** The maps in Figure 4.1.4 and Figure 4.1.5 depict the ozone concentrations and population-weighted hospital admissions averaged for each NYS region across all 10 summers. For ozone, NYS Regions 1 and 3 have relatively high ozone concentrations compared to NYS Regions 5 and 8, especially when examining extreme ozone values ($\geq 95^{th}$ percentile; Figure 4.1.4).
Figure 4.1.4: Map of 8-h DMA ozone concentrations averaged across all 10 summers for each NYS region; (a) mean ozone concentrations (ppb) and (b) 95th percentile mean ozone concentrations (ppb).

After accounting for population differences, NYS Regions 2, 3 and 4 have relatively high respiratory-related hospital admissions as compared to NYS Regions 1, 6 and 8 (Figure 4.1.5).

Figure 4.1.5: Map of population-weighted respiratory-related hospital admissions for each region; inset is pie-chart of unweighted hospital admission counts.
Annual boxplots of daily averaged ozone concentrations, maximum temperature, dew point and hospital admissions (Figure 4.1.6) indicate an unusually warm, humid summer in 2002 and relatively high hospital admissions in 2003.

Figure 4.1.6: Boxplots of daily data averaged for all regions; (a) hospital admissions counts, (b) 8-hr DMA ozone concentrations (ppb), (c) maximum temperature (F°), and (d) dew point (1/10 F°).

Figure 4.1.7 supports that ozone concentrations are higher for days when the air parcel passed through the ORV (mean ozone concentration = 63 ppb) compared to when it did not (mean ozone concentration = 51 ppb).
The results from applying the GAM to the entire study time period indicate that the risk was elevated for populations in NYS Regions 1, 2, 3, 5, 6, 7 and 8 on days when the air parcel passed through the ORV zone compared to days that it did not. This elevated risk is significant for NYS Regions 2, 3 and 6 for the ORV variable. The results were similar (except for Region 4) for ozone, where NYS Regions 2 and 4 were significantly associated with respiratory-related hospital admissions (Figure 4.1.8).
A comparison of risk produced by the GAM used in this study and the unadjusted odds ratio calculated in Garcia et al. (2011) revealed consistent results between the two studies but with fewer NYS Regions having significant associations for the ORV variable than when applying the unadjusted odds ratio calculation (Figure 4.1.9).
Figure 4.1.9: The risk and 95\textsuperscript{th} percentile confidence intervals for the ORV variable using the (a) Unadjusted Odds Ratio, and (b) GAM.

Pre-NBP versus Post-NBP implementation time periods: Figure 4.1.10 shows the risk values for exposure to ozone and air parcels from the ORV for before and after the NBP time periods.
The risk for ozone exposure (per 10 ppb increase) and for the ORV variable is calculated for (a) ozone (pre-NBP time period); (b) ozone (post-NBP time period); (c) ORV variable (pre-NBP time period); (d) ORV variable (post-NBP time period).

Table 4.1.1 summarizes the risk coefficients calculated for ozone and the ORV variable for the entire 10 summers, and before and after the NBP.
Table 4.1: Risk Calculated for Ozone (8-hr DMA) and the ORV Variable

### Exposure to Ozone and Transported Pollution

<table>
<thead>
<tr>
<th>Region</th>
<th>β (Ozone Variable)</th>
<th>β (&quot;ORV&quot; Variable)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Relative Risk (per 10 ppb)</td>
<td>Lower 95th Confidence Interval</td>
</tr>
<tr>
<td>Ten Summers (1997 - 2006)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.01</td>
<td>0.99</td>
</tr>
<tr>
<td>2</td>
<td>1.02</td>
<td>1.00</td>
</tr>
<tr>
<td>3</td>
<td>1.00</td>
<td>0.97</td>
</tr>
<tr>
<td>4</td>
<td>1.05</td>
<td>1.02</td>
</tr>
<tr>
<td>5</td>
<td>1.04</td>
<td>0.99</td>
</tr>
<tr>
<td>6</td>
<td>1.03</td>
<td>1.00</td>
</tr>
<tr>
<td>7</td>
<td>1.01</td>
<td>0.98</td>
</tr>
<tr>
<td>8</td>
<td>1.00</td>
<td>0.98</td>
</tr>
<tr>
<td>Pre-NBP Summers (1997 - 2000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.00</td>
<td>0.98</td>
</tr>
<tr>
<td>2</td>
<td>1.01</td>
<td>1.00</td>
</tr>
<tr>
<td>3</td>
<td>0.98</td>
<td>0.94</td>
</tr>
<tr>
<td>4</td>
<td>1.03</td>
<td>0.99</td>
</tr>
<tr>
<td>5</td>
<td>1.14</td>
<td>1.05</td>
</tr>
<tr>
<td>6</td>
<td>1.01</td>
<td>0.97</td>
</tr>
<tr>
<td>7</td>
<td>0.99</td>
<td>0.95</td>
</tr>
<tr>
<td>8</td>
<td>1.00</td>
<td>0.95</td>
</tr>
<tr>
<td>1</td>
<td>1.02</td>
<td>0.99</td>
</tr>
<tr>
<td>2</td>
<td>1.02</td>
<td>1.01</td>
</tr>
<tr>
<td>3</td>
<td>1.04</td>
<td>0.97</td>
</tr>
<tr>
<td>4</td>
<td>1.04</td>
<td>1.00</td>
</tr>
<tr>
<td>5</td>
<td>0.93</td>
<td>0.82</td>
</tr>
<tr>
<td>6</td>
<td>1.05</td>
<td>0.96</td>
</tr>
<tr>
<td>7</td>
<td>0.95</td>
<td>0.91</td>
</tr>
<tr>
<td>8</td>
<td>1.03</td>
<td>0.97</td>
</tr>
</tbody>
</table>

4.1.6 Discussion

Analysis of all summers. Applying the full GAM shown in Eq. 1 reveals a significant association between respiratory-related hospital admissions and both main effects (ozone and ORV indicator). The risk calculated from the GAM indicates that the risk for populations in NYS Regions 2, 3, and 6 is significantly elevated for respiratory-related hospital admissions on those days when the air parcel passed through the ORV zone compared to those days that it did not. Similarly, the risk for populations in NYS Regions 2 and 4 is significantly
elevated after exposure to ozone (calculated as a 3-day moving average). The results for the ORV variable are consistent with findings presented in Garcia et al. (2011) which indicates that the risk is significantly elevated for regions 1, 2, 3, 6, 7 and 8 when an unadjusted odds ratio formula is applied, but with fewer NYS regions showing significant association when the GAM was applied (Figure 4.1.9). The GAM accounts for several variables known to be associated with respiratory-related hospital admissions (e.g., day-of-week, holiday, temperature and dew point). The unadjusted odds ratio calculation compares the same population on exposed and unexposed days and does not need to account for some covariates, such as day-of week and holidays. The unadjusted odds ratio, however, does not account for temperature and humidity—both of which are known to be associated with each other and respiratory-related hospital admissions. The Pearson’s correlation coefficient (r) is 0.60 for temperature and ozone (Figure 4.1.11). Regardless of this colinearity between temperature and ozone, the standard in the epidemiology community is to include both of these variables in the timeseries model since both variables are associated with the health endpoint. Including confounded variables in the model can affect the detection of a health signal since information is removed as the variability associated with the confounding variable is removed. Because these confounded variables are included in the model, the GAM is a more conservative estimate of risk than the odds ratio analysis reported in Garcia et al. (2011).
Figure 4.1.11: The correlation between temperature and ozone is 0.60 for all regions.

Pre-NBP versus Post-NBP implementation time periods. The pre-NBP versus post-NBP time period analysis produced inconsistent results. Some time periods for some regions were associated with decreasing coefficients and some were not with no distinguishable pattern among them (Figure 4.1.10). The inter-annual analysis (the model was run with successively less annual data). (Figure 4.1.12) revealed that the coefficients for the main effects (ozone and ORV) are relatively consistent in applying the epidemiology model for 10-summers, 9-summers, 8-summers, 7-summers and 6-summers.
At around five summers, however, the 95th percentile confidence intervals widen and the risk estimates change more substantially. At four summers and less, the pattern of the risk estimates for each year begin to look very different (Figure 4.1.12). Region 4 is no longer significant and the relative risk for Region 7 moves below the 1.0 line. According to this
analysis, 5-summers is the least number of summers needed (given the decadal historic record examined) to estimate a reasonable risk and associated 95th percentile confidence intervals for associations with respiratory-related hospital admissions.

The NBP before and after time periods, however, have only 4-summers (1997 – 2000) and 3-summers (2004 – 2006) of data, respectively. Given the annual variability inherent in this data and the annual variability analysis discussed above, this is not enough years of data to discern a change in the health signal.

4.1.7 Summary

Including a variable to capture the transport of polluted air parcels from the Midwest into NYS Regions did reveal significant associations with respiratory-related hospital admissions for three NYS regions (2, 3 and 6). Exposure to ozone was also found to be associated with respiratory-related hospital admissions for two NYS regions (2 and 4). The analysis to determine whether respiratory-related hospital admissions decreased as a result of the NBP did not produce consistent results. Some NYS regions showed decreasing risk after implementation of the NBP, but some NYS regions showed increases. Based on an inter-annual analysis, the lack of data before the implementation of the NBP (4 summers) and after the implementation of the NBP (3 summers) does impact the validity of the results. At a minimum, data covering at least 5-6 summers (460 – 552 days) in each of the two time periods being examined (e.g., before and after the NBP) is needed.
REFERENCES


Ozone Transport Assessment Group, (1997). *Air Quality Analysis Workgroup’s Volume 1: Executive Summary*, (Guinnup, D., Collom, B. (Ed)).


Tolbert, P.E., M. Klein, J.L. Peel, S.E. Sarnat and J.A. Sarnat, (2007). “Multipollutant modeling issues in a study of ambient air quality and emergency department visits in Atlanta,” *Journal of Exposure Analysis and Environmental Epidemiology* 17(S2), S29-S35.


5.0 Conclusions and Future Research

5.1 Conclusions

This research study examined methods for characterizing exposure to polluted air parcels transported from the Midwest, and then applied a selected method in an epidemiology analysis to assess the health impact of transported pollution. In addition, two time periods were examined before and after the implementation of the NBP to assess whether a change in the health effect measure could be discerned as a result of the NBP emission controls. These research objectives were framed by posing three primary research questions. The key findings based on these questions are presented below.

1. Do air quality surfaces derived from various statistical data combination techniques help us to best characterize air quality and enable us to understand associations between transported air pollution and human health endpoints? Despite remaining challenges, the comparison of the data combination techniques indicated that the additive and multiplicative bias adjustment techniques can capture the extremes in the ozone concentrations better than the other methods compared (weighted average and hierarchical Bayesian model). Thus, the multiplicative bias approach was used to produce combined model and observed surfaces of daily maximum 8-hr average ozone concentrations for 5 summers (2001 through 2005). This combined surface was used in an epidemiology model to estimate risk and was compared to the same process that was applied to averaged ozone concentrations for same-day, lag 1 to 3 days, and 3-day moving averages. The calculated associations were generally stronger for the combined surface as compared to the surface derived from observations alone. In particular, the associations were significant using the lag 2 days, lag 3 days and 3-day moving average exposure metrics, whereas no association was discerned using raw observations alone.

2. Can atmospheric transport patterns associated with high-ozone days be linked to human health endpoints? Three methodologies (arc, bounded zone, and clustering) were examined to classify polluted air transported from the Midwest and relatively clean air transported from the North for ten consecutive summers between 1997 and 2006. These
definitions of exposure and non-exposure, respectively, were used to calculate an unadjusted odds ratio to examine associations between the transported air parcels and respiratory-related hospital admissions in NYS. The results revealed that the mean ozone concentration and the direction of the air parcel origin were consistently and significantly different for all eight regions for the southwesterly versus northerly directions. Of the three techniques, the bounded zone approach produced stronger associations with respiratory-related hospital admissions. The results of the study indicate that exposure to air parcels transported from the Midwest into NYS (as compared to air parcels that are transported from the North) result in elevated risk for respiratory-related hospital admissions over the entire 10-summer study period for NYS Regions 1, 2, 3, 6, 7 and 8. While the unadjusted odds ratio calculation does not control for temperature (that is also correlated with hospital admissions), a separate analysis of temperature in Region 2 provided evidence that these associations are not the result of temperature alone.

3. Can we discern a difference in the health effects measure before and after the implementation of the NBP? This phase of the study applied the transport variable defined using the bounded zone approach in a multi-variate model to account for variables such as temperature that cannot be accounted for with an unadjusted odds ratio calculation. Including the transport variable in the model did reveal significant associations with respiratory-related hospital admissions for three NYS regions (2, 3 and 6). Exposure to ozone was also found to be associated with respiratory-related hospital admissions for two NYS regions (2 and 4). The analysis to determine whether respiratory-related hospital admissions decreased as a result of the NBP did not produce consistent results. Some NYS regions showed decreasing risk after implementation of the NBP, but some NYS regions showed increases. Based on an inter-annual analysis, the lack of data before the implementation of the NBP (4 summers) and after the implementation of the NBP (3 summers) does impact the validity of the results. At a minimum, data covering at least 5-6 summers (460 – 552 days) in each of the two time periods being examined (e.g., before and after the NBP) is needed.
Overall, this study reveals that refining air pollution estimates through techniques that combine air quality model output and observed data do result in significant and stronger associations than using observations alone. This supports that the air quality model (CMAQ) provides important spatial information that may improve exposure misclassification and result in a discernable health signal where it would otherwise not be discernable. The bounded zone methodology was used to isolate transported air parcels from the Midwest, and both the unadjusted odds ratio and the multi-variate model revealed significant positive associations between the transported air parcels and respiratory-related hospital admissions in NYS Regions 2, 3 and 6. The comparison of the two time periods before and after the NBP did not produce consistent results because of the lack of data available for the two time periods. However, the transport variable shows promise as a viable indicator for tracking the health signal that may result from emission controls implemented to curb transported pollution.

5.2 Future Research

The use of the enhanced air quality surfaces generated using the multiplicative bias adjustment approach indicates that better characterizing air quality may reveal a health effect where it otherwise would not be evident. A next-step in this research is to examine whether the results are consistent when extending the time period and applying the enriched ozone surfaces across all 8 NYS regions.

The research described in this dissertation can be used as a baseline for applying varying exposure metrics (e.g., interpolated observations, spatial micro-environmental emissions data, averaged air exchange rates) to examine whether and how the calculated risk is altered. Depending on these results, the refined characterization of exposure can be applied in a post-processing step to generate an archive of pollutant surfaces for use in health studies. While measurements of PM2.5 and its components are limited, accurately characterizing PM2.5 and its components may be critical in evaluating human exposure and health effects. Future research is needed in applying refined combination approaches (e.g., Hogrefe et al., 2009) to generate an archive of these pollutant constituents for multiple years.
In summary, the following research activities can be pursued to further the work presented in this dissertation:

- Transform pollutant concentration surfaces to exposure surfaces to better discern air quality-human health signals.
- Apply intervention modeling approaches to concentration and exposure surfaces segregated by weather patterns to see if the impact of the NBP can be better discerned.
- Extend the analysis to bias-corrected PM2.5 and its species using data from air quality models.
- Apply synoptic typing methods to delineate pre- and post-NBP impacts on air quality and human health.
- Examine methods to address colinearity between ozone and temperature (e.g., de-seasonalize/de-trend ozone and hospital admissions data before applying the GAM).
- Investigate associations with outliers (spike in hospital admissions in 2003 after adjusting for power outage, particularly hot summer in 2002 and cool summer in 2004) and other metrics, such as coordinates of back-trajectory end points as a surrogate variable for exposure.
- Conduct screening analysis to investigate whether diurnal pattern of high nighttime ozone concentrations is associated with transport or health endpoints.
- Use historical dataset and results from this study to compare to using the exposure model output and examine the consistency of results across regions and counties.
- Perform sensitivity analysis and cross-validation by iteratively removing data.

The application of techniques that improve the characterization of exposure to atmospheric pollution may reveal associations between air quality and human health that were not discernable before. These approaches may reveal improvements in human health resulting from emission control activities such as the NBP as such techniques better refine exposure estimates and are applied over longer time periods.
REFERENCES