APPLICATION OF ADJOINT SENSITIVITY ANALYSIS FOR PERFORMANCE ENHANCEMENT OF POWER PLANTS' NITROGEN OXIDES CONTROL POLICIES

by

Seyyed Morteza Mesbah

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in Environmental Engineering

Carleton University
Ottawa, Ontario

ABSTRACT

Significant investments in nitrogen oxide (NO_x) emission controls in the U.S. have led to a substantial reduction in emissions. However, it is unclear whether these programs have optimally reduced ozone concentrations and their corresponding health damages. Current cap-and-trade program allocates emission quotas to participants and allows the trade of quotas on a one-to-one basis. However, it does not account for spatial and temporal differences in health damage of NO_x emissions. This shortcoming in the current U.S. NO_x control policy is explored in this research.

Spatial and temporal differences in NO_x emissions can be included in policy design if emission quotas are valued differently (exchange rate policy) or if polluters pay time- or location-specific emission fees (taxation policy). The main objective of this work is to develop a decision support system model for evaluating different policies. The proposed model includes an optimization platform to predict the polluters' behavior, and an air quality model and its adjoint (or backward) sensitivity model to calculate the derivatives of the environmental or health damage function with respect to NO_x emissions used for emission differentiations.

The results from a case study of U.S. power plants show that exchange rate trading outperforms current indiscriminate trading policies. These findings imply that by implementing exchange rate trading or taxation policies, current improvements in air quality could have been achieved at lower costs, or alternatively, more substantial improvements could have been reached at little to no additional costs. Furthermore, the results indicate that setting the emission fees on an hourly basis leads to a outcome

comparable to setting fees based on location. Moreover, the per ton health benefit of NO_x emission reductions is found to increase as emissions are reduced. This finding is particularly important from an environmental policy perspective as it impacts the optimal NO_x emission reduction target. Our results also indicate that power plants in the restructured electricity market are willing to pay more for emission quotas. Uncertainties involved in the proposed model, challenges for implementation of the proposed policies, and inclusion of health impacts caused by exposure to particulate matter are main directions for future research.

PREFACE

The following non-copyrighted articles have been reproduced in full in this thesis with permission from the co-authors:

- **Mesbah, S. M.**, Hakami, A., Schott, S. "Non-convexity in ozone-based NO_x health damage: an application of the adjoint of CMAQ" (Revised draft in preparation, chapter 7).

- **Mesbah, S. M.**, Hakami, A., Schott, S., "Optimal ozone control with inclusion of spatiotemporal marginal damages and electricity demand" (Final draft in preparation, chapter 6).

- **Mesbah, S. M.**, Schott, S., Hakami, A., "Short-term NO_x emissions control behavior in the regulated and restructured electricity market" (Final draft in preparation, chapter 8).

The following copyrighted articles have been reproduced in full in this thesis with permission from the co-authors and the journal, Environmental Science and Technology:

- **Mesbah, S. M.**, Hakami, A., Schott, S., (2013) "Optimal ozone reduction policy design using adjoint-based NOx marginal damage", *Environmental Science and Technology,* Vol. 47 (23), 13528–13535, (chapter 5).

- **Mesbah, S. M.**, Hakami, A., Schott, S., (2012) "Improving NOx cap-and-trade system with adjoint based emission exchange rate", *Environmental Science and Technology*, Vol. 46 (21), 11905–11912, (chapter 4).

All materials should be cited as outlined above.

ACKNOLEDGMENTS

This thesis is the outcome of my interdisciplinary PhD research on air pollution modeling and environmental economics at Carleton University. I had a wonderful time in Carleton working on my PhD with great faculty members and staff.

I would like to express my deepest thanks to many whom without their support and help the creation of this thesis was impossible. I would like to thank my supervisor, Dr. Amir Hakami, and co-supervisor, Dr. Stephan Schott, whose wise advice, passion, continuous support, critiques, comments, encouragement, and positive attitude kept me motivated in my PhD journey. I would also like to thank my advisory committee members: Mr. Michael Donohue, Dr. Deniz Karman, Dr. Leigh Mazany, and Dr. Daniel Cohan for their time and valuable advice and comments. I would like to thank the members of the atmospheric research group at Carleton University whose help at different stages of my thesis was greatly appreciated. I want to thank, Matthew Russell who devoted his time to assist the group in fixing computer related issues, which came up very often. I would also like to thank Farid Amid, Maryam Mirzajani, Amanda Pappin, and Dr. Shunliu Zhao who each helped me with different aspects of my thesis. Help from Dr. Yongtao Hu from Georgia Institute of Technology was also greatly appreciated.

I would like to extend my deepest thanks to my parents who have always been supportive. Words cannot express how grateful I am for their support. I would like to dedicate my thesis to my wife whose continuous support, understanding, and proof-reading was a great help in accomplishing this thesis.

S. Morteza Mesbah

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

LIST OF ABBREVIATIONS

ADE	Atmospheric diffusion equation
ADM8	Average daily maximum 8-hour
APS	Ambient-permit system
APEEP	Air Pollution Emission Experiments and Policy
BACT	Best Available Control Technology
BEIS	Biogenic Emission Inventory Systems
BCON	Boundary condition processor
BenMAP	Environmental Benefits Mapping and Analysis Program
CAA	Clean Air Act
CAMx	Air Quality Model with Extensions
CAIR	Clean Air Interstate Rule
CaC	Command-and-control
CaT	Cap-and-trade
CaT-EX	Cap-and-trade with an exchange rate
CaT-EXP	Cap-and-trade with an exposure-based exchange rate
CEM	Continuous Emission Monitoring
CMAQ	Community Multiscale Air Quality
CMIN	Cost minimization
CRDM	Climatological Regional Dispersion Model
CSAPR	Cross State Air Pollution Rule
CTM	Chemical transport model
DMIN	Damage minimization
DSS	Decision support system
DDM	Decoupled direct method
DM8	Daily maximum 8-hour

D-SCMIN	Demand-based social cost minimization
EGU	Electric generation utility
EPA	Environmental Protection Agency
ERCAM-NOx	Emission Reduction and Cost Analysis Model for Oxides of Nitrogen
ERS	Exchange-rate emission trading system
FIP	Federal Implementation Plan
F-SCMIN	Flexible social cost minimization
LAER	Lowest Available Emission Rate
GAD	Gross annual damage
GDP	Gross domestic production
ICON	Initial condition processor
IDA	Inventory Data Analyzer
IPM	Integrated Planning Model
KPP	Kenetic pre-processor
LNB	Low NO_x burner
MAC	Marginal abatement cost
MD	Marginal damage
MPO	Modified pollution-offset
MCIP	Meteorology Chemistry Interface Processor
NAAQS	National Ambient Air Quality Standards
NAM	North American Mesoscale
NBP	NO_x Budget Program
NDO	Non-degradation offset
NEI	National Emission Inventory
NO_x	Nitrogen oxides
NPRI	National Pollutant Release Inventory

NSR	New Source Review
OTC	Ozone Transport Commission
OPF	Optimal power flow
OFA	Over fire air
ORISPL	Office of Regulation Information System Plant Location
PHEV	Plug-in hybrid electric vehicles
PJM	Pennsylvania−New Jersey−Maryland
PM	Particulate matter
POS	Pollution-offset system
PSD	Prevention of Significant Deterioration
RACT	Reasonably Available Control Technologies
SIP	State Implementation Plan
SMOKE	Sparse Matrix Operator Kernel Emission
SCMIN	Social cost minimization
SPT-D-SCMIN	Spatial demand-based social cost minimization
SPT-F-SCMIN	Spatial flexible social cost minimization
TMP-SCMIN	Temporal social cost minimization
SCR	Selective catalytic reduction
SNCR	Selective non-catalytic reduction
TAC	Total abatement cost
TD	Total damage
TRS	Trading ratio system
TLM	Tangent linear model
TMP-D-SCMIN	Temporal demand-based social cost minimization
TMP-F-SCMIN	Temporal flexible social cost minimization
VOC	Volatile organic compound

VMT	Vehicle mile traveled
WRF	Weather Research and Forecasting
WPS	WRF Preprocessing System
WTP	Willingness to pay

LIST OF SYMBOLS

A	ambient quality (chapter 2), abetment activity (chapter 8)
A_j^*	target ambient quality at receptor j
A	background pollution
e_i	emission from source i
r_i	emission reduction from source i
t_{ij}	transfer coefficient from source i to receptor i
l_{ij}	number of ambient permits that source i holds at receptor j
c_i	abatement cost of source i
c_i'	marginal abatement cost of source i
ER_{ij}	exchange rate between sources i and j
$\overline{T_i}$	allocated permit to source i
T_{ij}	number of permits source i buys from source j
$\overline{TD}$	cap on total system-wide damage
C_i	concentration of species i
U	three-dimensional the wind field
P	air density
K	turbulent diffusivity tensor
R_i	chemical reaction rate for species i
E_i	emission rate of species i
F_i	ith row of the Jacobian of the chemical reaction rates
J	adjoint cost function
g	local cost function
λ	adjoint variable (or Lagrangian multiplier in an optimization framework)

φ	adjoint forcing term
$\mathcal{N}$	numerical operator for forward method
$\mathcal{L}$	numerical operators for forward sensitivity method
$\mathcal{L}^*$	numerical operators for adjoint sensitivity method
P_i	population at grid i
P_q	electricity price
Q	electricity generation level
R_{io}	rate of transformation between input and output
P_i	input price
c_p	power plant's capital and operating cost
c_c	control technology's capital and operating cost
R_{eo}	rate of transformation between the output and emissions
λ_i	ozone metric sensitivities to NO_x emissions from polluters i
α_{ij}	exchange rates between polluters i and j
e_i^0	allocated emission quotas to polluter i
$e_i^{\max}$	maximum possible emission level for polluter i
β	epidemiological concentration response factor
ΔM	change in mortality
M_0	baseline non-accidental mortality rate
V_{SL}	value of statistical life
ΔTD	change in total damage
D_i	damage function for source i
E_T	total ozone season cap on emissions
MD	marginal damage
Q_T	total electricity demand in the ozone season

G_i	generation capacity of polluter i
R_i	generation intensity (MWh/ton) for polluter i
Q_t	total electricity demand in the ozone season for hour t

CHAPTER 1:

INTRODUCTION

Surface ozone is a serious threat to public health even at very low concentrations because it can contribute to premature mortality resulting from short-term and long-term exposure (Bell et al. 2004; Jerrett et al. 2009). The first ozone regulation effort in the U.S. was promulgated in 1970 and introduced the National Ambient Air Quality Standards (NAAQS). NAAQS standards defined the limit for six main pollutants, called criteria pollutants by the U.S. EPA, including ozone.

Ozone is formed in the atmosphere through a series of photochemical reactions and is therefore considered a secondary pollutant. The reactions occur when there are adequate amounts of NO_x (NO_2 and NO), volatile organic compounds (VOCs), and sunlight. NO_x emission sources include mobile (emissions produced by vehicles) and point sources, as well as other combustion processes (e.g., biomass burning) and natural emissions (e.g., soil NO_x). Ozone concentrations are usually higher in hot seasons because the photochemical reaction depends often on the amount of sunlight. Accordingly, ozone control programs are usually in place during the ozone season from May 1st to September 30th. The first NO_x control program was started in the 1990s as a nationwide program to prevent acid rain (formed from either NO_x or SO_2 emissions); the program, however, was not designed to deal with surface ozone problem. There were two options for control of NO_x under the acid rain program: reduction of the emission rates for all units of a facility (power plant) to

below a defined level, or reduction of the average emission rates for a facility to below a certain level (EPA, 2007).

The Ozone Transport Commission (OTC) NO_x trading program was the first cap-and-trade program, and was established in 1999 in the northeastern U.S. Since then, a number of NO_x cap-and-trade programs have been implemented in the eastern U.S. The number of states participating in these cap-and-trade programs has also increased (EPA 2008).

Cap-and-trade programs are popular due to their flexibility. Participants in cap-and-trade programs are assigned emission quotas and have the option to trade in a free market, but need to keep enough quotas to cover their emissions for a specific period of time. The emission market motivates participants with different emission reduction costs to trade emission quotas in a way designed to reduce their expenses. Cap-and-trade programs are also attractive for regulators because they allow them to reduce the total emissions in the system by reducing the total number of quotas. The total NO_x emissions in the eastern U.S., where cap-and-trade programs have been in place for the past two decades, has been reduced drastically (EPA 2008). However, ground-level ozone is still in violation of the NAQS standards in several regions in the eastern U.S.

The current U.S. cap-and-trade NO_x program does not account for location and time specific effects of NO_x emissions on ozone formation. In the current NO_x cap-and-trade program, one unit of emissions at one location can be traded for one unit of emissions at any other location. Issuing trading permits on a one-to-one basis

(with no exchange rate) is not problematic for long-lived pollutants species, such as carbon dioxide (CO_2), because they exist long enough to be uniformly mixed. For ozone or its precursors, the shorter lifetime results in spatial variability in the transport path from sources to receptors. In particular, the response of ozone to NO_x emissions is very location-dependent and nonlinear (Hakami et al. 2004; Tong and Mauzerall 2006). The ozone formation potential of NO_x emissions depends on where and when they are emitted and what chemical atmospheric regime they are exposed to during their trajectory. Therefore, the effectiveness of NO_x emission reduction is affected by the time and location of the reduction. Inclusion of these temporal and spatial effects can improve the performance of the current system (Muller and Mendelsohn 2009; Nobel et al. 2001; Sun et al. 2012).

The first step to include the differences in ozone formation potentials in the policy making process is an accurate estimation of the impacts of NO_x for different times and locations. The estimation of individual source impacts is not an easy task when it comes to modeling the fate of pollutants in a complex atmospheric system with different chemical and physical processes. The variability associated with the emissions from different sources and varied meteorological conditions, contributes to the complexity of atmospheric pollution modeling.

Calculation of source impacts on health or other policy outcomes relies on simplified or formal sensitivity methods. Both approaches are used to calculate the relationships between sources and receptors, and to identify the contribution of the sources to the pollutant concentrations at the receptors. Simplified methods, such as

once-at-a-time perturbation approach (also called brute-force method), are straightforward but become computationally infeasible for large number of sources. Formal sensitivity approaches rely on calculation of derivatives of concentrations with respect to emission rates through mathematical differentiation of the governing equations. Simplified and formal sensitivity methods can be applied to atmospheric models with different levels of complexity from dispersion models to three-dimensional photochemical air quality models (Cho et al. 2012; Fann et al. 2009; Muller and Mendelsohn 2009; Tong et al. 2006; Hakami et al., 2004; Pappin and Hakami 2013).

Sensitivity analyses that rely on simplified models often do not account for physical and chemical processes occurring during the transport of pollutants from sources to receptors, and thus, have lower accuracy for calculation of source-receptor relationships particularly for secondary pollutants such as ozone. The traditional sensitivity methods, known as brute-force methods, account for physical and chemical processes in the atmosphere but are more expensive because they rely on computationally challenging simulations. The brute-force methods need conducting simulation by air quality models for multiple perturbed scenarios and therefore are only affordable for identifying the impacts from a limited number of sources (Nobel et al. 2001; Tong et al. 2006; Wang, Thompson, et al. 2007).

Formal sensitivity methods can calculate the source-receptor relationships efficiently with realistic computational costs. Formal sensitivity approaches are categorized into forward and backward (adjoint) sensitivity methods. Both methods

can establish source-receptor relationships but with different efficiencies. For a matrix of source-receptor relationships with M rows for sources and N columns for receptors, there are $M \times N$ source-receptor coefficients. A forward sensitivity model can calculate one row of the matrix in a single simulation, while a backward sensitivity model can estimate one column of the matrix in a single model run. Therefore, if of interest is the relationships between a few sources and all receptors, the forward method is more efficient, whereas the relationships between a few receptors and all sources are more efficiently addressed by the backward (adjoint) method.

This thesis employs a recently developed adjoint sensitivity tool for a state-of-the-art air quality model and investigates its applications in environmental policy making. We present a methodology for differentiating between NO_x emission sources which can, in turn, lead to the design of optimal control strategies and policies for various environmental and health objectives, and an environmentally improved emissions trading system. This work aims to provide decision makers with a decision support system that can assist quantification of the effect of different control programs on air quality, and help establish more effective environmental guidelines. In this study, the following questions will be addressed:

- What is the contribution of each power plant to the damage caused by the elevated levels of ambient ozone? How do the per ton benefits of NO_x emission reductions vary temporally and spatially? How can the inclusion of such information in policy design improve environmental performance?

- How can adjoint sensitivity analysis and targeted trading (trading that differentiates between sources) be used towards establishment of an ozone trading system that outperforms the current NO_x cap-and-trade system?
- How different is the environmental performance of a trading system when targeted trading is in place? What emission distribution would maximize environmental performance?
- How does the benefit per ton of NO_x emission reductions change when emission are reduced and what are the policy implications of these changes?
- How do power plants response to different types of electricity markets in the short-term? How does it impact their emission trading behavior?

This thesis is organized into 9 chapters. Five chapters are organized in a paper format and contain more detailed introductory sections. Chapters 4 to 8 have their own introduction and methodology as a paper, and as such chapters 2 and 3 are mainly provided to be complements, which may introduce some redundancy.

In chapter 2, the background and relevant literature are reviewed. This review covers chemical reactions in the atmosphere, ozone and NO_x control programs and regulations in the U.S., NO_x control technologies for power plants, methods for estimation of pollution health damages, and the theoretical basis for cap-and-trade systems and its variants. As this work uses adjoint sensitivity analysis for estimation health damages attributable to NO_x emissions, the adjoint sensitivity analysis is also reviewed.

Chapter 3 of this thesis describes the methods and tools used in this work. The components of the proposed decision support system are introduced and the process through which the proposed model can inform decision making process is discussed. The components of the decision support model including the air quality model, the sensitivity analysis model and the optimization tool are presented and discussed.

Chapter 4 presents the details of a theoretical but improved NO_x cap-and-trade system. Unlike the current policy in the U.S., which treats NO_x emissions from all sources equally, this system differentiates among NO_x emissions by accounting for their location-specific ozone formation potentials through use of emission exchange rates. Two different NO_x control policies (i.e., the cap-and-trade with and without exchange rates) are compared, and their efficiency is evaluated. We demonstrate that the decision support model can predict NO_x emission trading behavior and post-trade ozone concentrations. The environmental outcomes and abatement cost performances of the two policies are examined in a case study for 218 U.S. power plants.

Chapter 5 addresses the inclusion of spatially different benefit per ton of emission reductions (marginal damages) in policy design and setting location-specific emission fees. It demonstrates which distribution of emissions results in a socially optimal environmental performance. This chapter addresses the potential performance enhancement of switching from the current policy to a taxation policy. The benefits gained by redistributing of emissions are estimated for a case study similar to that presented in chapter 4. ..

We include temporal differences in the marginal damages of emissions in the study presented in Chapter 6. This chapter investigates how marginal damages vary with time, and how such variations can be incorporated into an economic instrument such as cap-and-trade. The benefits gained by redistributing electricity generation are also investigated. Under the proposed policies, the interaction of the electricity market and NO_x control policies is considered. This chapter identifies optimal times and locations for electricity generation resulting in minimized costs of emission reductions and health damages.

Marginal damage value can change in times as significant reductions in emissions take place. The study presented in Chapter 7 investigates how marginal damages change with nationwide reductions in emissions. Source-specific marginal damages of NO_x emissions are calculated when emissions from different sectors (i.e., mobile and point sources) are reduced. Using the calculated marginal damages for different baseline emissions, source-specific marginal damage curves are constructed and relevant policy insights are discussed.

Chapter 8 proposes a methodology to calculate the short-term willingness to pay for emission permits in the U.S. using plant specific data. This estimation is based on plant-level benefit maximization in both the electricity and emissions permit markets. In this chapter, using the proposed methodology, the behavior of power plants is estimated for the 218 coal-fired electric generation units that took part in the NO_x Budget Program (NBP) during the ozone season of 2007.

In Chapter 9, a summary of the findings of this thesis is provided and insights and conclusions from these results are discussed. The thesis is concluded with suggestions for potential areas for future research.

CHAPTER 2:

BACKGROUND

In this chapter, first ozone and NO_x chemistry will be explained which will clarify how ozone concentrations respond to NO_x controls for different atmospheric regimes. Then, the relevant regulations and programs in the U.S. and their success in NO_x and ozone control will be discussed. Next, costs and benefits of the U.S. cap-and-trade programs calculated by different studies will be reviewed. Finally, NO_x control technologies, theoretical emissions trading models for different types of pollutants, and sensitivity analysis methods that are used for the proposed improvements to current cap-and-trade systems will be elaborated.

2.1. Ozone formation and NO_x

Ozone and NO_x , one of its main precursors, are the two main species considered in this thesis. There is a significant spatial and temporal variability in response of ozone concentrations to NO_x emissions. Identifying differences in NO_x emissions are of importance in policy design and can be used for targeted emission reductions. Furthermore, the relationships between ozone and NO_x is very complex because an increase in NO_x emissions may lead to an increase or decrease in ozone concentrations depending on the governing atmospheric regime. This section aims to explain the chemistry behind the NO_x-ozone relationship which is essential information for chapters 4 to 7.

The ozone molecule includes three atoms of oxygen. It exists in both the stratosphere and troposphere. The stratospheric ozone protects the earth against ultraviolet wavelengths from the sun. In contrast, the tropospheric ozone, or surface ozone, is harmful to humans, plants, and animals. In this work, all the discussions are about tropospheric, or more accurately, surface ozone which will simply be referred to as ozone.

Ozone is a secondary pollutant, in that it is not directly emitted into the atmosphere. It is formed by a set of photochemical reactions that require volatile organic compounds (VOCs), nitrogen oxides or NO_x (which include NO and NO_2), and sunlight. Summer has the most favorable conditions for ozone formation because it is hot, dry and sunny. Correspondingly, ozone is regulated during the ozone season (May 1st to September 30th). The main source of NO_x is combustion, which mainly occurs in vehicle engines, industries, and electric generation utilities (power plants).

Ozone's response (increased/decreased concentration) to changes in its precursor (VOCs and NO_x) emissions, is nonlinear and differs from one location and/or time to another (Muller and Mendelsohn 2009; Tong et al. 2006). This response can be well represented by the ozone isopleth (Figure 2.1). All points on an isopleths contour line have the same ozone concentration.

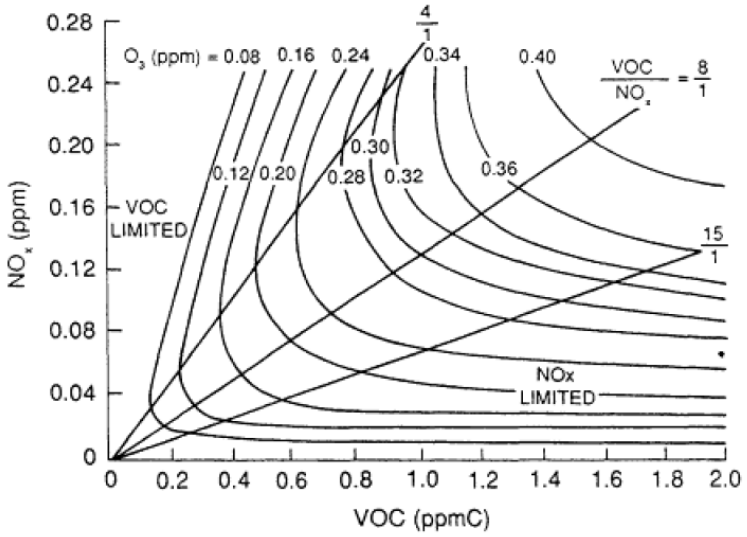

Figure 2.1. A typical ozone isopleth used in U.S. EPA Empirical Kinetic Modeling Approach (EKMA) (Doge, 1977).

There are two main regimes described by ozone isopleth: the NO_x-limited and the NO_x-inhibited (VOC-limited) regimes. In the NO_x-limited regime, the ratio of NO_x to VOC is low. This regime exists mainly in rural areas where there are higher (biogenic) VOC concentrations compared to NO_x. For this chemical regime, a reduction in the NO_x concentration reduces ozone formation. On the other hand, in the NO_x-inhibited (or VOC-limited) regime, the ratio of NO_x to VOC is high. Urban areas with high NO_x vehicular emissions or power plants are good examples of NO_x-inhibited regimes. For NO_x-inhibited conditions, NO_x emission reduction can cause increased ozone. For each predominant chemical regime, various chemical reactions dominate the ozone production/loss pathways.

Atmospheric chemical reactions take place in a state of dynamic balance where the species involved are consciously formed and consumed. The concentration

of a species increases when its rate of production is greater than its rate of consumption. The reactions below are a simplified representation of the ozone/NO_x chemical cycle (Seinfeld and Pandis 2006):

$$NO_2 \xrightarrow{hv\ (O_2)} NO + O_3 \tag{2-1}$$

$$NO + O_3 \rightarrow NO_2 + O_2 \tag{2-2}$$

$$O_3 \xrightarrow{hv} O(^1D) + O_2 \tag{2-3}$$

$$O(^1D) + H_2O \rightarrow 2OH \tag{2-4}$$

$$OH + RH + O_2 \rightarrow RO_2 + H_2O \tag{2-5}$$

$$RO_2/HO_2 + NO \rightarrow NO_2 + RO/OH \tag{2-6}$$

$$OH + NO_2 \rightarrow HNO_3 \tag{2-7}$$

$$OH + CO \rightarrow CO_2 + H \tag{2-8}$$

$$H + O_2 + M \rightarrow HO_2 + M \tag{2-9}$$

Among the above equations, the photolysis reactions, and those dependent on radicals formed trough photolysis reactions only occur in the presence of sunlight. During the day, ozone is formed by reaction 2-1 and is consumed by reaction 2-2. During night, reaction 2-2 consumes part of the ozone produced during the day by reaction 2-1. Reaction 2-6 can change the ozone formation/consumption balance by converting NO to NO_2 without consuming O_3 (i.e., reaction 2-2) which amounts to a net production of ozone. The rate of reaction 2-6 (which is essential for enhanced production of ozone during photochemical smog episodes) depends on the availability

of radicals. Radicals (i.e., species with a free electron such as OH, or RO_2) are key components in photochemical ozone production. Availability of peroxy radicals (i.e., HO_2 and RO_2) depends on hydroxyl radical availability. The production of HO_2 is also dependent on the availability of hydroxyl radical for oxidation of CO and VOCs (reaction 2-5 and 2-8). The hydroxyl radical is often referred to as the atmospheric cleansing agent as it initiates the majority of oxidation pathways in the atmosphere.

The availability of radicals (e.g., OH) and the rate of reaction 2-6 depends on the NO_x regime. The effectiveness of NO_x emissions in producing ozone is the result of ongoing competition between various reactions for reacting with active radicals taking part in the photochemical production of ozone. In particular, VOCs (shown as RH) and NO_x compete for OH (reaction 2-5 and 2-7) and it is the NO_x regime that determines which reaction wins. Note that while reaction 2-5 leads to eventual production of ozone, reaction 2-7 removes radicals from the system and slows down ozone production.

In a NO_x-inhibited regime, reaction 2-7 is the governing reaction because NO_x availability limits reaction 2-6. In this case, NO_x reduction causes increased ozone concentrations as the decrease in NO_x through reaction 2-7 provides more OH radicals for initiation of oxidation processes (reactions 2-5 and 2-8) and the subsequent enhancement in the rate of reaction 2-6.

In the NO_x-limited regime, the chemical system is constrained by the availability of VOCs, rather than NO_x, in the atmosphere. In this case, the rate of reaction 2-6 is controlled by availability of NO_x (NO) and decreases in NO_x emissions

leads to reduction in ozone formation. The different responses of ozone to NO_x are of interest in control policy design because they result in negative or positive benefit per ton of emission reductions. A negative benefit per ton usually occurs in large urban locations with high NO_x concentrations, and it means that emission reductions are not beneficial and increase the damages to human health. More discussions on NO_x health damages and benefit per ton of emission reductions will be provided in chapters 5, 6, and 7.

The transition region between NO_x-limited and NO_x-inhibited regimes is often referred to as the ozone ridge. This region represents a chemical regime where NO_x to VOC ratios are near optimal and ozone production occurs with high efficiency. The ozone response to NO_x reductions in NO_x-limited and NO_x-inhibited regimes is close to linear. Nonlinearity in the ozone response to NO_x emissions happens when changes in NO_x availability causes transition from one regime to another (Hakami et al., 2004).

2.2. U.S. ozone and NO_x programs

2.2.1. NO_x regulations and programs in the U.S.

NO_x regulations in the United States date back to 1970 when ozone was designated as one of the six criteria pollutants by Congress in the Clean Air Act (CAA). The 1970 amendment required the EPA to establish the NAAQS for criteria pollutants. It also required states to develop State Implementation Plans (SIPs) to

control their emission levels to satisfy NAAQS. The 1970 amendment focused on new sources and existing sources which were planning on major modification. These sources were subjected to an emissions standard called the New Source Performance Standard (NSPS). In 1970, The CAA mandated the EPA to develop NSPS emission standards. NSPS determined an emission rate (lb/MMBtu) based on the Best Available Control Technology (BACT) for the new sources.

In 1977, Congress added more provisions to the CAA to protect against degradation of air quality. The two main additions were the Prevention of Significant Deterioration (PSD) provision in attainment areas and the New Source Review (NSR) provision in nonattainment areas. Non-attainment areas are defined in CAA as the areas where air quality does not meet the standard. PSD was added to the CAA to prevent the degradation of the ambient quality and NSR was added for prevention of increased emissions in nonattainment areas. All sources in non-attainment areas were required to use Reasonably Available Control Technologies (RACT). According to the NSR, new sources in nonattainment areas were only permitted if they used the Lowest Available Emission Rate (LAER) and offset emissions from existing sources (Burtraw and Szambelan, 2009).

The 1977 CAA amendment did not lead to NO_x emissions reductions since they had no strict limitations for the existing stationary sources. Emissions restrictions for existing sources were added to the CAA in 1990. The changes imposed strict rules for emissions reductions, and they gave some authority to the EPA for regulating these sources. Title I and Title IV of the CAA 1990 amendments were two ways to

limit NO_x emissions. Title IV, which created the Acid Rain Program, required a 2 million ton yearly reduction in NO_x emissions from coal-fired power plants. Title I of the 1990 CAA amendment included provisions for attainment and maintenance of NAAQS. It also outlined a timetable for states to meet the NAAQS requirements (Burtraw et al., 2005).

The CAA revision of 1990 also created the Ozone Transport Committee (OTC). In 1994, the OTC was formed under the CAA to consider the ozone problem in the eastern U.S. In 1999, the OTC states signed a memorandum and the OTC Trading Program started in 11 states and the District of Columbia. The OTC states were grouped into outer, northern, and inner zones. The inner zone included the populated eastern states, most of which were among the ozone non-attainment regions, affected by emissions from up-wind emitters in the outer zone. To protect the ambient air quality in the inner zone, the original OTC trading program was designed based on zonal trading limits. Unlike a non-zonal emission trading which allows the exchange of emission quotas on a one-to-one basis between zones, a zonal trading sets a ratio for the exchange of quotas between zones to discourage the increase of emissions from up-wind states contributing to the concentrations in down-wind states. However, further studies suggested that non-zonal trading would not degrade the ambient air quality leading to the implementation of non-zonal trading by OTC states. The OTC was not under the EPA, but it was helped by the EPA monitoring system and its air quality models (EPA, 2003).

The ozone problem in the eastern states occurred in wider areas than OTC regions. States were not able to meet the 1997 NAAQS, because downwind states received sizeable contributions to their ambient air quality from up-wind out of state sources. According to the CAA amendment of 1990, states could send petitions to the EPA if upwind states contributed to their ambient air quality. Following receipt of such petitions, the EPA used its power under the CAA to revise SIPs in 1998. The rule is known as the NO_x SIP Call since it calls states to revise their implementation plans. According to the NO_x SIP Call rule, states could either follow an emission rate limit in the state or could take part in a regional cap-and-trade program. The trading program under the NO_x SIP Call (NO_x Budget Program (NBP)) started in 2003 after regulatory steps took place. 19 states and the District of Colombia took part in the NBP (EPA, 2008).

In 2005, the EPA used its power under the 1990 amendment and established the Clean Air Interstate Rule (CAIR). This rule is known as the CAA good neighbor rule since neighbor states must reduce their emissions cap if they have a significant negative contribution to other states' air quality. CAIR, which is the current trading program in place in the U.S., was promulgated in 22 eastern states in 2008. The rule was temporarily vacated in 2008 in a litigation brought before the D.C. circuit court before being reinstated by the appeal court later in the year.

In 2011, the EPA finalized the Cross State Air Pollution Rule (CSAPR) which was intended to replace CAIR. Under the CSAPR, if the contribution of upwind states to ozone formation in downwind states were more than a limit, the upwind states were

required to make additional emission reductions. Out of 28 states covered by CSAPR, 5 states were required to additionally reduce NO_x emissions. In 2012, the U.S. supreme court vacated CSAPR and ordered the EPA to keep CAIR in place while working on a replacement rule. There were two main reasons for the court's decision. First, the court found that the EPA mandated upwind states to reduce their emissions under the Federal Implementation Plan (FIP), and did not allow them to develop their own SIPs to meet emission reduction requirements. Second, the court concluded that the EPA forced upwind states to reduce more emissions than their contribution to downwind states. The court ruling was one of the main reasons for the significant drop in the NO_x permit price. The NO_x price was $825/ton in the ozone season of 2008, but dropped significantly after 2008. The price stayed below $30/ton during the ozone season of 2012. Other reasons for such a low NO_x price included the significant emission reductions and high availability of banked emission quotas (EPA 2013).

2.2.2. Achievements of NBP

NBP sources are about 20 percent of the total NO_x emissions in NBP regions. The other major sources of NO_x are natural (e.g., lightening, soil, and wildfire), mobile sources, area sources, and point sources that are not part of the NBP. There were 2569 units participating in the NBP in 2008. Among these units 88 percent were electric generation utilities (EGUs). Figure 2.1 shows the number of units participating in the NBP by type.

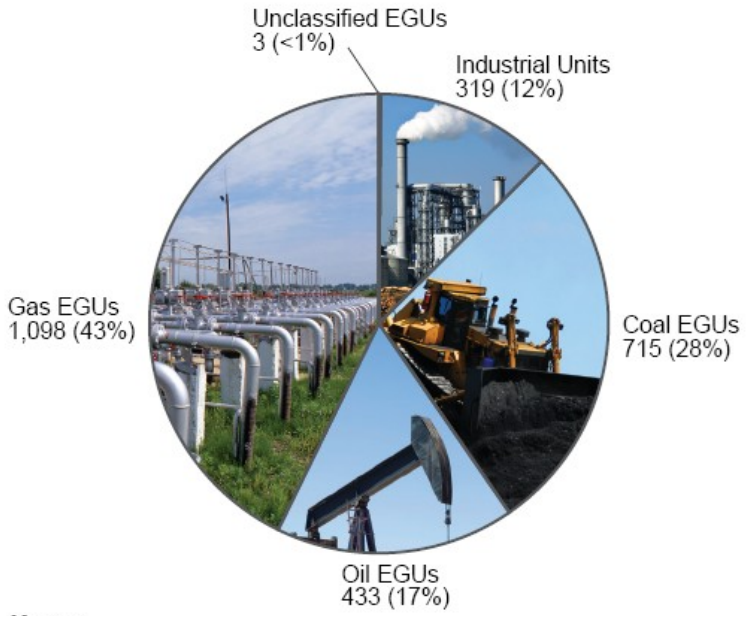

Figure 2.2. Number of NBP sources by type (EPA 2008).

The implementation of different programs in the U.S. has decreased the emissions from affected sources significantly. In 2008, the ozone season NO_x emissions from NBP sources decreased to 62 percent below the year 2000 level (before the implementation of NBP) and to 75 percent below the year 1990 emission level (before the implementation of CAA amendments). The total emissions in 2008 were 9% below the cap (Figure 2.3). The overall region-wide ozone concentration since the implementation of NBP has decreased between 11 to 14 percent (EPA, 2008). The seasonal average 8-hour ozone concentration in NBP regions is presented in Figure 2.4. Recent years (after CAIR) have witnessed an even more drastic reduction in NO_x emissions from EGUs.

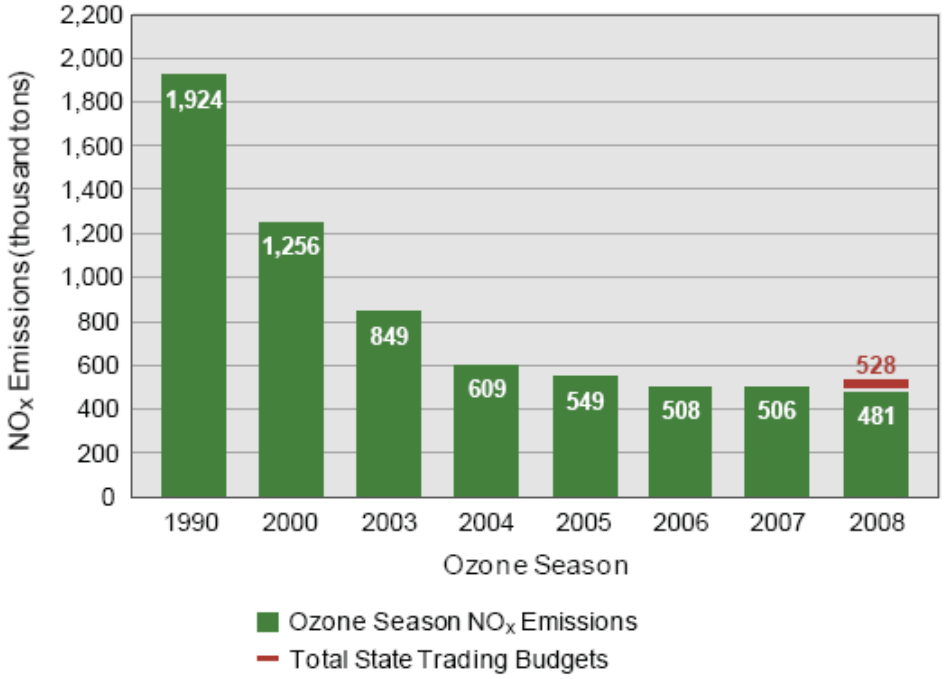

Figure 2.3. NO_x emissions from all NBP sources (EPA, 2008).

Since meteorology also influences ozone formation, in order to evaluate the effect of emission reduction on ozone concentrations in different years, equal meteorological conditions should be imposed. As shown in Figure 2.4, the average ozone concentration in 2004 (unadjusted for weather) is lower than 2007 although the total emissions in this year were higher than emissions in 2007. This was because the ozone season in 2007 was warmer than in 2004. Removing the meteorological effect results in a higher concentration in 2004 and lower concentration in 2007 (EPA, 2008).

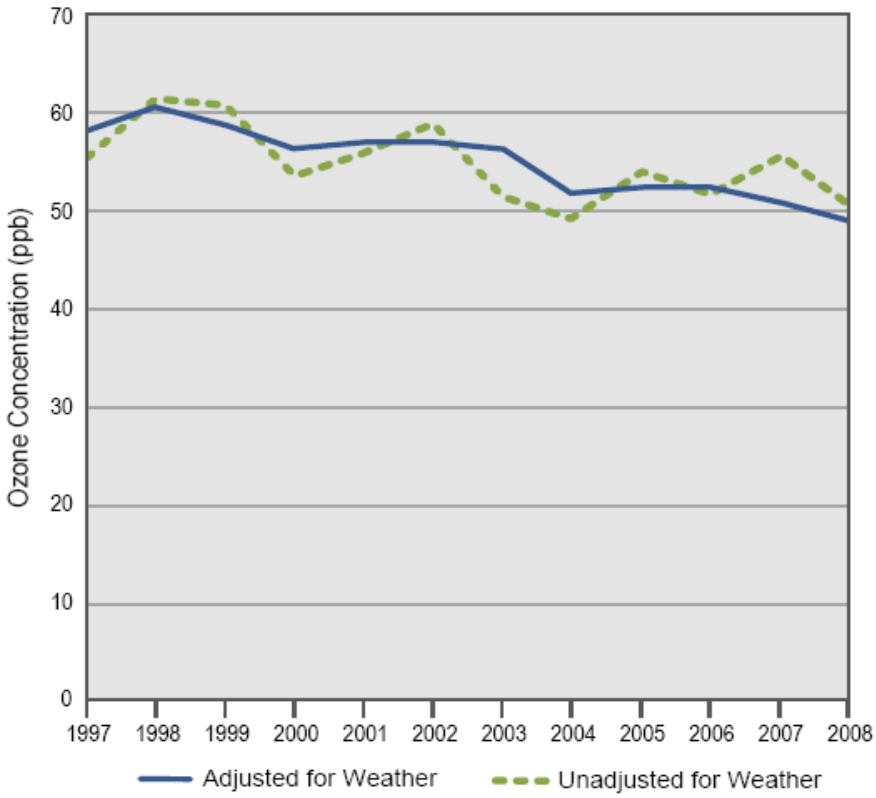

Figure 2.4. Seasonal average 8-hour ozone concentration in NBP region (EPA, 2008).

2.2.3. Performance of NO_x control programs in the U.S.

Despite the improvement in ambient air quality (ozone concentration) under the NO_x cap-and-trade programs in the U.S., the approach may not be optimally efficient because it neglects the source-specific ozone formation potentials. A socially optimal control program minimizes the social cost of emissions which is defined as the emission reduction costs plus the damage costs to the environment and health, known as external costs. The U.S. NO_x cap-and-trade is a market-based program that provides cost-saving incentive for participants to trade emission quotas. In the market, participants with high emission reduction costs purchase quotas and save in emission

reduction costs. In contrast, participants with low emission reduction costs reduce more emissions and benefit by selling their unused emission quotas. Hence, the cost of emission reduction is minimized in the system. However, the current U.S. emission market does not account for the external costs and may result in shifting emissions from low-damage to high-damage locations and lead to an increase in health damage costs. Therefore, the social cost of emissions is not necessarily minimized under a traditional cap-and-trade system. We will discuss social cost minimization policies in further details in chapters 5 and 6. Previous studies have estimated the cost saving obtained by an efficient NO_x emissions reduction in the U.S. Seskin et al. (1983) examined the cost of reducing NO_x emissions from stationary sources to the standard (250 $\mu g/m^3$) levels in the Chicago air quality control region. They estimated the annual control costs at 9 million dollars under the least cost strategy and 130 million dollars under the State Implementation Plant (93.1 % saving). Farrell et al. (1999) compared NO_x cap-and-trade systems with traditional command and control from 1999 to 2006 in the northeastern U.S. They predicted that the total cost decreased from $1930 million dollars to $1030 million dollars (47 %) by implementation of a cap-and-trade program.

Krupnick (1986) estimated the potential cost savings from 200 of the largest NO_x emission sources in Baltimore. His work compared the least-cost strategy with a combined least-cost and command-and-control strategy based on reasonably available control technology (RACT) for three different NO_x ambient quality standards (250, 375, and 500 $\mu g/m^3$). Note that under the combined strategy, emission trading was

allowed after participants met the RACT requirements. His results showed that changing the strategy from combined command-and-control and least-cost strategy to the least-cost strategy leads to a significant cost saving in the system. Based on his results, the savings are 24% (from \$2.2 million to \$1.66 million per year) when the NO_x ambient air quality standard is 250 $\mu g/m^3$, 69% (from \$1.10 million to \$0.37 million per year) when the standard is 375 $\mu g/m^3$, and 96% (from \$1.52 million to \$0.07 million per year) when the standard is 500 $\mu g/m^3$.

In a study similar in scope to this one, Krupnick et al. (2000) examined the effectiveness of the NO_x cap-and-trade programs in the eastern U.S. They compared three different policies: command-and-control, cap-and-trade with a one-to-one exchange rate, and cap-and-trade with an ozone exposure exchange rate. Their exchange rate (the ratio at which various partners trade their emissions) was based on ozone exposure and was calculated by running an air quality model (Urban Airshed Model UAM-V) for three two-day typical episodes from the 1990 ozone season. They used the Emission Reduction and Cost Analysis Model for Oxides of Nitrogen (ERCAM-NO_x) (Pechan, 1997) to estimate the abatement costs for different categories of point sources (including but not limited to EGUs). They concluded that the EPA's policy (one-to-one ton exchange of NO_x) is efficient compared to other policies. They found that there is a large saving (50%) in migrating from a command-and-control approach to one-to-one trading policies, but that there is not a significant savings (only 2%) between one-to-one trading and the exchange rate policy. The reported costs from their work were 1369 million dollars per year for command-and-

control, 680 million dollars per year for one-to-one trading policy, and 666 million dollars per year for ozone exposure trading.

Burtraw et al. (2001) also calculated the cost and benefits from a NO_x budget program. They calculated the cost of post-combustion technologies (selective catalytic reduction (SCR) and selective non-catalytic reduction (SNCR)) in 2008 for three different policies: ozone season trading in SIP regions ($2,146 million), annual trading in SIP regions ($2,728 million), and annual trading throughout the U.S. ($4,434 million). They reported the necessity of implementation of an annual program by calculating a higher benefit ($1,777 million) for an annual trading program in SIP regions compared to seasonal program ($749 million). Their benefit analysis, however, only considered the health benefits from particulate matter (PM) reductions, and not improvement in ozone air quality. As such, their work underestimates the benefit that can be achieved in the ozone season but captures the more seasonally diverse impact from the PM exposure.

2.2.3.1. Marginal damage and its calculation

The benefits obtained by implementation of cap-and-trade programs in the U.S. has been achieved without consideration of spatial and temporal differences in the marginal damage of pollution. Marginal damage (MD) or the damage per ton caused by pollution represents the monetary value of damage caused by an additional unit of emissions. MDs, or benefits per ton of emission reductions, are important information, which can be used to calculate the total benefit gained by emission reductions under different policies. They are also essential for setting optimal

emissions reduction targets. MD information, which is not included in the current policy, is especially important for a pollutant whose damages are time and location dependent because they can be used by regulators to target emission reductions for high MD locations and times, which in turn results in a significant reduction in system-wide damages.

Estimation of health damages caused by air pollution is an inter-disciplinary research topic and involves economics, epidemiology, and air quality modeling. Epidemiological studies are used to link the change in air pollutant concentrations to change in mortality. Economic studies estimate the value of statistical life, expressing the monetary value of loss of one human life. Air quality models are used to construct source-receptor relationships which are used to identify how emissions from different sources contribute to pollutant concentrations at receptors. Through combination of economics, epidemiology, and air quality modeling, source-specific marginal damages, can be calculated. The calculation of source-receptor relationships for atmospheric pollutants is often difficult in a large domain without some simplification.

Sensitivity coefficients represent the change in concentrations at receptors as the result of a small change in emissions from sources. The brute-force method is traditionally used to model source-receptor relationships. Under this method, an air quality model is run multiple times (i.e., an additional time per additional source). Before each additional run, the emissions from the source of concern are perturbed. Then, the model is re-run and the resulting concentrations are subtracted from the

concentrations under the base case to identify the impact of source of concern on receptors across the modeling domain. One of the limitations of the brute-force method is that it requires numerous model runs to construct source-receptor relationships. This limitations results in using state-of-the-art air quality models only for a limited number of sources (e.g., Mauzerall et al. 2005; Tong et al. 2006) or using simplified dispersion models for a large number of sources (e.g., Levy et. al 1999; Rowe et al. 1995). However, as explained in chapter 1, the simplified models do not account for all physical and chemical process in the atmosphere and therefore may lack accuracy.

To estimate the source-receptor relationships, Levy et al. (1999) used three dispersion models to simulate the primary pollutant close to the source (ISCST3), primary pollutants at a certain distance from the source and secondary particles (SLIM3), and the secondary pollutant, ozone (SOMS). They compared three older methods for the estimation of health damages caused by power plants (European Commission 1995; Ridge 1994; Rowe et al. 1995), and suggested a method for the calculation of health damages. The three studies reported the damage costs for oil fueled power plants to be between $0.0001/kWh to $0.019/kWh, which was mainly due to human health damage (between $.000003/kWh and $0.017/kWh). Levy et al.'s method was an extension of the previous methods with some improvement in air quality modeling for ozone, an updated epidemiological section, and additional accounting for co-pollutant confounding. The species that were included in their damage function were CAA criteria pollutants: PM, ozone, carbon monoxide, NO_2,

SO_2, and lead. They found that the plant location was a key determining factor for the contribution of pollutants to total damage. They estimated marginal damages for the year 1997 for different species as follows: \$12,000/ton for PM_{10}; \$770/ton for NO; \$790/ton for SO_2; and \$8100/GWh for the total criteria pollutants.

In addition to the simplified dispersion models, reduced-form models are used to construct source-receptor relationships for a large number of sources (e.g., Levy et. al 2009; Fann et al. 2009; Muller and Mendelsohn 2007, 2009; Muller 2011). Reduced-form models rely on multiple response points from the base model to construct source-receptor relationships using the outputs of an air quality model such as Multiscale Air Quality (CMAQ) model (Fann et. al 2009) or the Climatological Regional Dispersion Model (CRDM)) (Levy et al 2009; Muller and Mandelson 2007, 2009). As these models do not fully account for photochemical processes in the atmosphere, they are not well-equipped to characterize the response of secondary pollutants such as ozone.

Levy et al. (2009) calculated health damages based on PM-based mortality using a reduced-form of CRDM. They investigated the uncertainty associated with the calculation of MDs for 407 coal-fired power plants in the U.S. The reduced form model was used to construct source-receptor relationships and calculate MDs for NO_x, SO_2, and $PM_{2.5}$. Their MD calculation was based on $PM_{2.5}$ related mortality. Their calculated MDs varied from \$500/ton to \$15k/ton for NO_x; from \$6k/ton to \$50k/ton for SO_2; and \$30k/ton to \$500k/ton for $PM_{2.5}$, which translated into \$0.02/kWh to \$1.57/kWh. The multivariable regression analysis indicated that the variability in

damage per ton was mainly due to variability in population and meteorology, and the variability in damage per kWh was due to population, meteorology, fuel, and control technologies. Their results also indicated that removing 4 million tons of SO_2 and 1.5 million tons of NO_x would result in about $100 billion in benefits.

Muller and Mendelsohn (2007) used reduced-form model based on the CRDM and developed an integrated assessment model, called the Air Pollution Emission Experiments and Policy (APEEP) model, to calculate the marginal damage from 10,000 sources in the U.S. The total damage of each source was estimated by multiplying the source's marginal damage and its total emissions. They named the aggregation of total damage caused by individual sources the gross annual damage (GAD) which was estimated as a parallel accounting for gross domestic product (GDP). They calculated the marginal damage for SO_2, NO_x, NH_3, VOC, $PM_{2.5}$, and PM_{10}. Their results showed that the GAD varies from 0.7 to 2.8% of the total U.S. GDP. Moreover, they found that the majority of the damage was to human health. They concluded that NO_x and coarse particles form half of the total emissions by mass but only 20% of the total damage. Their estimated MDs were lower than MDs estimated in later studies, which is partly due to the use of a simplified model that is limited in its representation of secondary pollutants. Their damage estimation was based on an assumption that NO_x MDs do not change when emissions are reduced. We will show in chapter 7 that this assumption underestimates the monetary value of damages caused by NO_x emissions.

In another study (Muller and Mendelsohn 2009), Muller and Mendelsohn used the APEEP model to examine the performance of a cap-and-trade system with exchange rates defined based on the ratio of marginal damages. They constructed a case study for SO_2 emissions from electric generating units that were regulated under title IV of the 1990 Clean Air Act Amendments (CAAA) in the year 2002. Their results indicated that MDs are higher in populated urban areas than low-populated rural areas, and the ratio of the highest MD to the lowest is about 150. Furthermore, they found that in urban areas, the elevation where the pollution is emitted is an important factor, and constructing tall smokestacks is an effective abatement method. Their results also showed that an annual gain of $310 to $940 million could be obtained by switching from the current SO_2 cap-and-trade to spatially explicit taxation or by a damage-based trading ratio cap-and-trade system.

Muller (2011) in a similar study to Levy et al. (2009) investigated the uncertainty associated with the marginal damage estimation of 565 electric utility power plants in the U.S for five pollutants including SO_2, NO_x, VOC, NH_3, and $PM_{2.5}$. He utilized a Monte Carlo simulation and APEEP model to estimate how MDs vary in magnitude and across power plants. His results showed that the statewide average NO_x MD varies from $620/ton (with a standard deviation of $1210/ton) to $3520/ton (with a standard deviation of $6140/ton). His results also indicated that power plants in urban areas were more variable than power plants in the rural areas. Moreover, he showed that adult mortality dose-response, mortality valuation, and air quality modeling were the three main sources of uncertainties for marginal damage

estimation. Furthermore, his results indicated that for all of the pollutants considered except NO_x, air quality modeling was the main source of uncertainty in calculation of exchange rates defined based on the ratio of marginal damages of pollutants. The main limitation of this study was the use of a dispersion model, which does not account for chemical and physical process occurring in the atmosphere for the estimation of MDs.

In another related study, Fann et al. (2009) used a reduced-form of an air quality model to predict marginal damage. They used the U.S. EPA's response surface model (RSM), generated from the EPA's Community Multiscale Air Quality (CMAQ) model results, and used the Environmental Benefits Mapping and Analysis Program (BenMAP) (Abt Associates 2009) to monetize the change in ambient air quality to dollar values. Their analysis estimated the benefits gained by $PM_{2.5}$ reduction as the result of reduction in NO_x, SO_2, NH_3, organic particles, and VOC emissions from 9 urban areas in the U.S. They used RSM and estimated the impact of pollution from nine urban areas on the continental U.S. They found that the variability in MD was due to differences in the type of pollution, source type, and the location of polluters. Their calculation showed that PM-based NO_x MDs range from -$4.5k/ton to $28k/ton for non-electricity generating units, and range from $1.1k/ton to $120k/ton for electricity generating units. Their results also indicated that the largest portion of the reduction of the $PM_{2.5}$ MD belongs to the direct reduction of $PM_{2.5}$, the second largest portion belongs to SO_2, and NO_x has the lowest contribution to the $PM_{2.5}$-related health damage.

Another alternative for the calculation of source-receptor relationships is the use of the adjoint (backward) sensitivity model. As explained in chapter 1, the adjoint model is an efficient tool if the relationships between a few receptors (or metrics) and all sources is of interest. The adjoint model can calculate the contribution of individual sources to one receptor or a metric, such as health damage function, in a single simulation and is therefore efficient method for calculation of MDs. This model does not face the limitations of the brute-force method (i.e., computationally expensive for a large number of sources) or reduced form models (i.e., not accounting for chemical and physical processes in the atmosphere). The adjoint method will be introduced in further detail in section 2.4.

Pappin and Hakami (2013) combined epidemiological information with a sensitivity version of the CMAQ model (adjoint-CMAQ) to calculate the marginal impacts of NO_x and VOC emissions from individual sources across North America on short-term mortality in Canada and the U.S. They estimated the health benefits based on 24-h average ozone and 1-h ozone concentrations. They found a significant variability in health benefits obtained by NO_x reductions for different locations across North America. For example, their study showed that a 10% reduction of NO_x in Hamilton, Canada, or in Detroit, U.S. results in a benefits in Canada and the U.S. equal to $253,000 per day, or $47,000 per day, respectively. They reported that NO_x MD could be as high as $75,000/ton in some locations in the U.S.

The MDs estimated by the studies presented in this short review vary (see Table 2.1) due to different assumptions and methods used. Different value of

statistical life, concentration response factors, baseline mortality numbers, modeling years, and most importantly different methods used for calculation of source-receptor relationships are the main reasons for differences in the MD estimations of NO_x emissions.

Table 2.1. MD estimations by different studies.

	NO_x MD ($/ton)	Damage function	Model	Base year	Comment
Levy et al. (1999)	770	PM_{10} and ozone mortality	Dispersion model	1997	Average MD is reported.
Levy et al. (2009)	500 to 15,000	$PM_{2.5}$ mortality	Reduced-form CRDM	1999	Median MDs for power plants are reported.
Muller (2011)	620 to 1210	$PM_{2.5}$ and ozone mortality and morbidity, agriculture, etc.	APEEP based on a reduced-form CRDM	2005	Average statewide MDs are presented.
Fann et al. (2009)	-4,500 to 120,000	$PM_{2.5}$ mortality	Reduced-form CMAQ	2015	Point source MDs from 9 large urban locations are presented.
Pappin and Hakami (2013)	-60,000 to 75,000	Ozone mortality	Adjoint-CMAQ	2007	MDs are based on 1-h max O_3 for Canada and average O_3 for the U.S.

Despite the fact that the ranges of the calculated MDs were different for the studies mentioned, they all reported a considerable heterogeneity in MDs and highlighted the importance for polices to distinguish between emissions by MDs or source impacts. While some studies provided a qualitative insight for potential benefits under emission differentiated polices (Pappin and Hakami 2013, Levy et al. 1999, 2009; Fann et al. 2009; Muller 2011; Levy et al. 2009, Mauzerall et al. 2005),

others quantified the costs and benefits of including source impact information within economic instruments (e,g., Muller and Mendelsohn 2009, Krupnick et al. 2000, Nobel et al. 2001). The objectives of chapters 4, 5, and 6 are to include source impact information within different economic instruments using the adjoint-CMAQ. The use of an adjoint model within economic instruments strengthens the air quality modeling side of the previous studies. It allows for the inclusion of a large number of sources, which was a major limitation for studies using the brute-force method (e.g., Krupnick et al. 2000, Nobel et al. 2001, Mauzerall et al. 2005, Tong et al. 2006). Moreover, the use of the adjont model improves the precision of MD estimation in comparison with studies using reduced-form or simplified models for a large number of sources (e.g., Muller and Mendelsohn 2009). As mentioned before, these simplified approaches/models oversimplify the complex atmospheric processes and may lack accuracy, especially for secondary pollutants such as ozone.

2.2.3.2. Re-dispatching of electricity generation

In addition to redistribution of emissions based on marginal damages, which lowers the system-wide damage, re-dispatching or shifting electricity generation from power plants with low to high generation intensity (i.e., the ratio of generation to emissions) is another strategy to reduce NO_x emissions and their corresponding impacts on ozone-based health damages. Re-dispatching strategies can be achieved by imposing an emission fee to all electricity generating units (Martin 2008; Alhajeri et al. 2011; Sun et al. 2012). In a competitive electricity market, units with low cost per unit of generation win the competition and take a higher share of the system-wide

generation. When units have to pay an emission fee, the cost per unit generation increases at a higher rate for units with a low generation intensity as compared to units with a high generation intensity. Therefore, if the emission fee is high enough, high generation intensity units can better compete in the market and take a higher share in generation, which results in a re-dispatching of generation and a reduction in system-wide emissions. Several studies have evaluated the impact of re-dispatching strategies on NO_x and corresponding ozone reduction.

Martin (2008) investigated the impacts of re-dispatching through a time and location differentiated NO_x control policy in the Pennsylvania−New Jersey−Maryland (PJM) region. Two different methods for modeling power networks were used: optimal power flow simulations and zonal model simulations. For both methods, the electricity network constraint, which limits the transmission of electricity between power plants, had little impact on re-dispatching. The findings showed that power plants could reduce region-wide hourly NO_x emissions from 15% (6 tons) on the highest electricity demand day to 30% (8 tons) on average demand days. The flexibility in region-wide emission reductions was due to the hourly electricity demand for specific days. Using WorldPower simulations, it was found that setting the NO_x price to $10k/ton, $20k/ton, $50k/ton, $100k/ton, or $125k/ton resulted in a NO_x reduction of 7%, 10%, 12%, 13%, or 14%, respectively. However, the linkage between NO_x reductions and the related change in ozone concentrations and the health damages is not addressed in the study conducted by Martin (2008).

Alhajeri et al. (2011) investigated the effect of re-dispatching strategies on NO_x reductions in the grid of the Electricity Reliability Council of Texas. They implemented an optimal power flow model to simulate the response of electricity generation to hourly demand. Different NO_x pricing scenarios ($0/ton, $2k/ton, $10k/ton, $25/ton, $50/ton) were tested while the SO_2 price was kept constant ($500/ton). Their results indicated that NO_x price setting strategies led to a 22.1% to 50.9% NO_x reduction as compared to the base case when the NO_x price was set to $0/ton. The co-benefits of re-dispatching strategies were also reported as 24.5% to 70.9% reductions in SO_2, 16% to 82% reductions in mercury, and 8.8% to 22% reductions in CO_2. There was also a 4.4% to 8.7% reduction in water consumption used by power plants for cooling the steam from boilers. The significant emissions reductions were due to a shift in generations from coal plants to natural gas plants which resulted in 4% to 13% increase in generation costs.

Pacsi et al. (2013) examined the costs and air quality impacts of re-dispatching to shift water use away from regions with drought in southern Texas. They implemented the PowerWorld model to simulate power plants' generations. The results showed that the system has the capacity to shift up to 10% of electricity generation from power plants in regions with droughts to other locations. The Comprehensive Air Quality Model with Extensions (CAMx) was used to calculate the impact of re-dispatching on air quality. Their results indicated that the 8-h ozone concentrations changed by 0.2 ppb to -0.55 ppb, SO_2 decreased by 3% to 21%, and NO_x changed by -2% to 8% compared to the base case (Pacsi et al. 2013). In their

study, the improvement in ozone concentration gained while the re-dispatching was targeted for drought issues. In chapter 7, we will present a re-dispatching strategy that is targeted to minimize ozone-based health damages.

Thompson et al. (2013) investigated the impact of replacing 20% of gasoline-based vehicle miles travelled by plug-in hybrid electric vehicles (PHEV) in major cities in Texas by the year 2018. The additional power for the replacement was assumed to be supplied by power plants in the Electricity Reliability Council of Texas grid. They examined three different scenarios based on maximizing battery life, maximizing driver's convenience, and overnight charges. They used a forward sensitivity model (CAMx-DDM) to quantify the impact of sources on 8-h ozone concentrations. Because their used approach had low efficiency for identifying the impacts from a large number of sources, they identified a subset of power plants (20, 28, and 45 sources under the three scenarios) for PHEVs' hourly power supply and calculated their impacts on ozone concentrations. Their results showed an increase in 8-h ozone concentrations overnight and a decrease in 8-h ozone concentrations during the day under all scenarios. They concluded that the positive impact of replacing gasoline cars with PHEVs outweighed the negative impacts of increased generation by power plants.

A recent study (Sun et al. 2012) examined a temporal emission differentiated policy focused on NO_x reduction on days with the highest ozone concentrations. This policy would set a higher price for NO_x when a high ozone day is forecasted, with no additional constraints for normal days. They implemented an optimal power flow

(OPF) model to simulate power plants' electricity generation and corresponding NO_x emissions when different prices are set for NO_x emissions in the PJM region. They found that setting a NO_x price of \$30k/ton, \$50k/ton, or \$100/ton resulted in 23%, 30%, and 41% lower ozone season NO_x emissions when compared to a base case with NO_x price of \$2k/ton. The significant NO_x reduction is obtained only by shifting electricity generation from high to low rate NO_x units while the demand was fixed and no power was imported from other regions. This was comparable to the reduction obtained by installation of SCR technologies for 32% of power plants in the region, which resulted in a 28% NO_x reduction or installation of SNCR technology resulting in a 34% NO_x reduction. They also used the Comprehensive Air Quality Model with Extensions (CAMx) model to estimate the impacts of NO_x control policies on ozone reduction. They found that setting the NO_x price to \$30k/ton, \$50k/ton, or \$100/ton lowers the maximum 8-h ozone concentrations by at least 1ppb over 14%, 23%, 33% of grids respectively. Installation of SNCR, or SCR lowers the maximum 8-h ozone concentration by 1 ppb over 14% and 55% of grids respectively.

In a similar study to those conducted by Martin (2008), Alhajeri et al. (2011), and Sun et al. (2012), we will examine the impact of emission fees on re-dispatching of generation. The aforementioned studies examined the use of a single emission fee for all participants under different hypothetical pricing scenarios, and reported a substantial reduction in NO_x emissions and corresponding ozone obtained by re-dispatching generations. However, one-fee based polices do not account for heterogeneity in MDs, as discussed in section 2.2.3.1 and may result in a suboptimal

ozone reduction. In chapter 6, we will take a different approach and examine the performance of several re-dispatching strategies under which regulators set location or time specific emission fees based on adjoint-driven marginal damages.

2.2.4. NO_x Control Technologies

Two main categories of NO_x control technologies are combustion modification and post-combustion control technologies. Combustion technologies aim to prevent the generation of NO_x emissions by modification of the combustions process, whereas post-combustion technologies aim to reduce NO_x emissions that have already been produced in combustion. Combustion modification technologies can reduce NO_x emissions by changing three main factors: peak flame temperature, the residence time of the peak flame, and the amount of air available for the flame. These technologies can also cause a reduction in a plant's efficiency because of incomplete combustion (Sloss et al., 1992).

The air-to-fuel ratio is an important factor that determines the amount of NO_x emissions. The optimum ratio corresponds to the maximum efficiency of the plant. Changing air-to-fuel ratio from the optimum can reduce NO_x emissions but also decreases efficiency. Combustion modification technologies provide different levels of air-to-fuel ratios in different flame regions to reduce NO_x emissions. This can be achieved by air or fuel staging. Air staging can occur within the burner (low NO_x burner technology) or in the furnace (over fire air technology).

A low NO_x burner (LNB) generates a thin and long flame at a lower temperature, hence generating less NO_x. LNB can cause a 50% reduction in NO_x

emissions. In an over fire air (OFA) furnace, unburned carbons are burnt and the combustion is completed. This technology can cause a 20% reduction in NO_x emissions. Fuel staging occurs within the furnace and usually on top of the burner (reburning technology). This technology generates a fuel rich region on top of the burner where some hydrocarbon radicals will be formed. The radicals reduce the NO_x generated in the burner to molecular nitrogen (EPA, 1999).

Two common post-combustion control technologies for NO_x are SNCR and SCR. Both of these control technologies are designed to convert NO_x to N_2 (EPA, 1999). For the SNCR technology, a reagent, ammonia or urea, is added to combustion gases. Ammonia reacts with NO_x and the products are water and molecular nitrogen (Reaction 2-10). The efficiency of SNCR is usually between 30 to 50%. The required temperature for the reaction 2-10 is about 1600-2000°F. The correct temperature is important because temperatures higher than 2000°F increase NO_x production and temperatures less than 1600°F cause the production of ammonia byproducts that can cause corrosion problems for the equipment.

$$4NH_3 + 4NO + O_2 \rightarrow 6H_2O + 4N_2 \qquad (2\text{-}10)$$

In SCR technologies, a reagent is mixed with the combustion gases and then the mixture goes to a catalytic reactor where a catalyst (e.g., titanium oxide) causes the formation of N_2 and water. The difference between SCR and SNCR is that in SCR, the catalyst can help reaction occur at a lower temperature (600-800°F). The efficiency of an SCR can be up to 90% but the trade-off is that the capital cost of a SCR is higher than a SNCR (EPA 1999a; Miller 2011).

2.3. Theoretical cap-and-trade systems

The design of an efficient cap-and-trade program depends on the nature of the pollutants being traded. In this context, pollutants are categorized into three main classes: uniformly mixed assimilative pollutants (such as biogenic VOCs over certain regions); uniformly mixed accumulative pollutants (e.g., CO_2); and non-uniformly mixed assimilative pollutants (NO_x) (Tietenberg 1985). For uniformly mixed pollutants, the location where the pollutant is emitted does not play a role in ambient air quality, whereas the location of emitters is very important for non-uniformly mixed pollutants. Accumulative pollutants are those that do not deteriorate easily, and last for a long time in the atmosphere. Note for the above categorization of pollutants, NO_x is a non-uniformly mixed pollutant. However, an efficient NO_x policy design has more complexity than that for other pollutants in this category (e.g., sulfur dioxide) because NO_x contributes to the formation of secondary pollutants such as ozone and particulate matter.

Uniformly mixed assimilative pollutants are the simplest to regulate since the pollutant's lifetime is substantially longer than its characteristic mixing time, but is short enough that there is no carry over from one regulation year to the next. The emission and ambient quality relationship for these pollutants is (Tietenberg, 1985):

$$A = a + \sum (e_i - r_i) \tag{2-11}$$

where A is the ambient quality, a is background pollution mainly from biogenic and natural sources, e_i is the baseline emission from source i, and r_i is its emission

reduction. To design a cap-and-trade system for these types of pollutants, an environmental authority can easily set a cap based on the environmental capacity and let traders exchange their emission on one-to-one unit bases.

Regulating uniformly mixed accumulative pollutants is more complicated since the pollutants accumulate with time. For this type of pollutant the relationship between ambient quality and emissions is as follow (Tietenberg, 1985):

$$A_t = \mathrm{A}_0 + \sum_i \sum_t (e_{it} - r_{it}) \tag{2-12}$$

where A_t is the ambient quality for year t, and A_0 is the initial ambient quality. e_{it} and r_{it} are the emissions and emissions reductions, respectively, from source i and year t. The implementation of cap-and-trade programs for this type of trade is more complicated since the emissions from one year affect the ambient quality in future years. Allowance allocation, using allowances from one year for another year, and setting a compliance date are some difficulties for the environmental authority when regulating uniformly mixed accumulative pollutants.

Regulating the non-uniformly mixed assimilative pollutants is further complicated since the ambient quality varies from one location to another. Also, the location where the pollutant is emitted has a key role in determining the ambient quality. The ambient quality for these types of pollutants can be defined as follows (Tietenberg 1985):

$$A_i = \mathrm{a_i} + \sum_j t_{ij}(e_j - r_j) \tag{2-13}$$

where A_i is the ambient quality for location i, $\mathrm{a_i}$ is the background pollution for location i, t_{ij} is the transfer coefficient from source j to receptor i. For this category, one unit of emissions in one location is not equal to one unit of emissions in another location. In theory, there are two main methods for trading these types of pollutants: ambient quality permit trading and emissions permit trading. Under ambient quality permit trading, a source needs to collect ambient permits from all receptors it affects. Under emissions permit trading, permits are assigned to sources and they can trade their permits as long as the ambient quality does not deteriorate. This can occur by defining an exchange rate among sources. The exchange rate for two sources is a ratio that determines the number of additional emission units the buyer can emit when buying one unit of emissions from the seller.

Major theoretical trading systems for non-uniformly mixed assimilative pollutants are: the ambient-permit system (APS) (Montgomery, 1972), the pollution-offset system (POS) (Krupnick et al., 1983), the modified pollution-offset (MPO) (McGartland and Oates, 1985), the non-degradation offset (NDO) (Atkinson and Tietenberg, 1984), the exchange-rate emission trading system (ERS) (Klaassen et al., 1994), the exposure based exchange rate (Krupnick et al., 2000), the trading ratio system (TRS) (Hung and Shaw, 2005), and marginal damage based exchange rates (Farrow et al., 2005; Muller and Mendelsohn, 2009). It should be noted that the current U.S. NO_x cap-and-trade program is not based on any of these theoretical

models for non-uniformly mixed pollutants. It applies a uniformly mixed concept (i.e., treats all emissions equally) to a non-uniformly mixed pollutant whose impact on ambient air quality is dependent on the time and location of emissions.

The ambient permit system (APS) is an ambient quality permit trading system. In this system, the environmental authority needs to define the ambient permits for all receptors and allocate the initial permits to sources. Then, sources can trade their ambient permits. They need to collect enough ambient permits from all affected receptors to cover their emissions. Under APS, when there are n receptors, the constraint on the emissions from source i is as follow (Montgomery, 1972):

$$e_i \leq \min \frac{\mathrm{l_{ij}}}{\mathrm{t_{ij}}} \text{ for } \mathrm{j} = 1, .., \mathrm{n} \tag{2-14}$$

where e_i is the emissions form source i, l_{ij} is the number of ambient permits that source i holds at receptor j and t_{ij} is the transfer coefficient from source i to receptor j. The implementation of this system is complicated since the environmental authority needs to assign many receptor sites, and allocate initial permits to sources for each receptor (Krupnick et al., 1983).

Modified pollution offset (MPO) is an extension to POS that is stricter about ambient quality (McGartland and Oates, 1985). Under MPO, the ambient quality must not degrade in any region after any trade even if the ambient quality is not violated in the region before trading occurs. In other words, the deterioration in ambient quality caused by one source must be offset by improvement of ambient quality caused by emissions reduction from another source.

Non-degradation offset (NDO) is another system that is based on two restrictions: there must be no violation in ambient quality standards, and no exchange rate higher than one. The latter restriction means that in any trade, the aggregate emissions from two traders after the trade cannot exceed their aggregate emissions before the trade (Atkinson and Tietenberg, 1984). The exchange rates under PO, MPO, and NDO are not fixed and can vary as long as the corresponding system's restriction is not violated.

The Exchange Rate System (ERS) offers a fixed exchange rate for trading emission permits. Under ERS, the exchange rates are calculated based on the ratio of polluters' marginal abatement costs (marginal abatement cost is the derivative of abatement cost) at the optimal emissions level. The optimal emission level is a set of emissions that maximizes the total net benefit. The following cost minimizing equations clarify how the exchange rate is defined under ERS (Klaassen et al., 1994):

$$Minimize \sum_{i} c_i(e_i^0 - e_i) \tag{2-15}$$

$Subject\ to$:

$$\sum_{i} t_{ij} e_i \leq A_j^* \ \ j = 1, .., n$$

where c_i is the cost of emission reductions for source i from e_i^0 to e_i, t_{ij} is the transfer coefficient from source i to receptor j, and A_j^* is the target ambient quality at receptor j.

The corresponding first-order condition is:

$$c_i^{'} - \sum_j t_{ij}\lambda_j = 0 \tag{2-16}$$

where c_i' is the marginal abatement cost of source i, and λ_j is the Lagrangian multiplier that represents the shadow price on the target ambient quality at receptor j. The shadow price for receptor j indicates the amount of cost-savings result from loosening the constraint on ambient quality at receptor j by one unit when the system-wide abatement cost is minimized. From (2-16), the marginal cost corresponding to the optimal solution for source i is: $c_i' = \sum_i t_{ij}\lambda_j$. The exchange rate ($ER_{mn}$) between source m and source n is then defined as the ratio of the marginal abatement costs (Klaassen et al., 1994):

$$ER_{mn} = \frac{c_m^{'}}{c_n^{'}} = \frac{\sum_j t_{mj}\lambda_j}{\sum_j t_{nj}\lambda_j} \tag{2-17}$$

Simply, λ_j represents the cost savings in the system if the target ambient quality constraint A_j^* shifts by one unit. If the ambient quality at receptor j is binding ($A_j = A_j^*$), the shadow price on receptor j is positive. But if the ambient quality at a receptor is not binding ($A_j < A_j^*$), the shadow price is zero. In other words, for binding receptors, releasing the constraint results in some cost savings for the system, and for non-binding receptors releasing the constraint does lead to cost savings. A disadvantage of ERS is that it needs the information of the abatement cost functions for all sources.

Krupnick et al. (2000) examined the effectiveness of the NO_x cap-and-trade programs in the eastern U.S. using a fixed exchange rate system. They defined exchange rates based on ozone exposure[1] and calculated them using source-receptor relationships (i.e., NO_x emissions from sources and ozone concentrations in receptors). The minimization problem under their proposed system is as follows:

$$Minimize \sum_{i}^{n} c_i(r_i) \tag{2-18}$$

$Subject\ to$:

$$\sum_{i}^{n} a_i r_i \geq L$$

where c_i is the cost of emission reduction for source i (r_i), L is the target level of ozone exposure reduction, and a_i is a source-receptor coefficient which converts emission reductions of source i (r_i) to ozone exposure. From the first order conditions of this problem, the exchange rate (ER_{ij}) of source i and source j can be calculated:

$$ER_{ij} = \frac{c_i^{'}}{c_j^{'}} = \frac{a_i}{a_j} \tag{2-19}$$

[1] Ozone exposure is defined as population times ozone concentration integrated over time.

The exchange rate for source i and source j under this system shows the ratio of the total effect of NO_x emissions on ozone exposure from source i to that of source j.

The Trading Ratio System (TRS) is another cap-and-trade system with fixed exchange rates that is designed specifically for trading water quality permits in rivers. The unidirectional flow to the lowest level in rivers simplifies the design of trading programs in these systems. The TRS system has three main characteristics: (1) the zonal effluent cap is set by accounting for the water pollutant loads transferred from upstream zones; (2) the trading ratios are set equal to the exogenous transfer coefficients between zones; and (3) permits are freely tradable among dischargers according to the unidirectional trading ratios.. The minimization problem under TRS are as follow (Hung and Shaw, 2005):

$$Minimize \sum_i c_i(e_i^0 - e_i) \tag{2-20}$$

$Subject\ to$:

$$e_i - \sum_{k=1}^{i-1} t_{ki}T_{ki} + \sum_{k>i}^{n} T_{ik} \leq \overline{T_i}, \qquad i = 1, \dots, n.$$

$$T_{ki}, T_{ki} > 0, \qquad 0 \leq e_i \leq e_i^0$$

where c_i is the cost of effluent reduction for source i from e_i^0 to e_i, $\overline{T_i}$ is the allocated permit to source i, t_{ki} is the transfer coefficient from source i (under this system there is only one source in each region) to the region j, T_{ki} is the number of permits source k buys from source i, and T_{ik} is the number of permits source k sells to source i.

From the first order conditions of this problem, the trading ratio can be calculated (Hung and Shaw, 2005):

$$c_i^{'} = \lambda_i, \qquad \forall i, \tag{2-21}$$

$$\lambda_i = t_{ik}\, \lambda_k \quad (i < k) \tag{2-22}$$

$$ER_{ij} = \frac{c_i^{'}}{c_j^{'}} = \frac{t_{ik}}{t_{jk}} \tag{2-23}$$

where t_{ij} is the transfer coefficient from source i to region j. k could be any region downstream of the two traders. The main difference between ERS and TRS is the method of exchange rate calculation (equations 2-17 and 2-23). TRS is designed for water quality and benefits from the unidirectional river flow but can be applied (with simplifying assumptions) to strictly downwind trading. TRS calculates the exchange rates using the ratio of transfer coefficients, which is equal to the ratio of marginal abatement costs. The main advantage of TRS over ERS is that the cost information is not needed for calculation of exchange rates. Our proposed optimization framework in chapter 4 is an extension of TRS, but is designed for air pollutants and accounts for relationships between NO_x emissions and ozone concentrations.

A similar study for water quality trading by Farrow et. al (2005) defines the exchange rate as the ratio of marginal damages for the sources. The marginal damage for a source is defined as the change in total damage in the system when the source increases its emission level by one unit. The total damage of each source is calculated using the transfer coefficients as well as household information (e.g., number of

people affected, their ages, their willingness to pay, etc.). The general optimization problem becomes (Farrow et al., 2005):

$$Minimize \sum_i c_i(e_i^0 - e_i) \tag{2-24}$$

$Subject\ to$:

$$\sum_i D_i(e_i) \leq \overline{TD}$$

$$0 \leq e_i \leq e_i^0$$

where c_i is the cost of emission reduction for source i from e_i^0 to e_i, $D_i(e_i)$ is the total damage from source i, and $\overline{TD}$ is the cap on total system wide damage.

From the first order conditions of this problem the trading ratio can be calculated:

$$ER_{ij} = \frac{c_i'}{c_j'} = \frac{\frac{\partial D_i}{\partial e_i}}{\frac{\partial D_j}{\partial e_j}} \tag{2-25}$$

where ER_{ij} is the exchange rate between sources i and j. This ratio represents the ratio of marginal damage from the two sources. The main difference between the system offered by Farrow et al. (2005) and TRS and ERS is the constraints on the cost minimization problem. Under TRS and ERS, there are many constraints that limit emissions in each zone, whereas under the system offered by Farrow et al. (2005)

there is only one general constraint that limits the total damage in the system. Muller and Mendelsohn (2009) introduced a system similar to Farrow et al. (2005) but for air quality emission permit trading. The exchange rates are defined based on the ratio of the marginal damage of sources (equation 2-25).

The design of cap-and-trade policies for air quality trading differs from that of water quality due to the inherent differences of the two media. For example, unidirectional flow to the lowest level in rivers simplifies the design of trading programs while the fate of pollutants in air is affected by multidirectional wind. Hence, calculating source-receptor relationships and exchange rates for the atmosphere is more challenging than water. To calculate source-receptor relationships, Krupnick et al. (2000) ran an air quality model for a few days with a limited number of sources, and Muller and Mendelsohn (2009) used simplified dispersion models that do not adequately represent the chemistry and other atmospheric processes. In most cases, calculating accurate source-receptor relationships requires Eulerian air quality model simulations that can be computationally demanding.

2.4. Sensitivity analysis

Sensitivity analysis refers to general methods for calculating derivatives. Formal sensitivity methods are those that solve sensitivity equations while approximate methods such as the brute-force approach estimate sensitivity coefficients by introducing one-at-a-time perturbations into the model. The brute-force approach is easy to implement but formal methods are more efficient and faster

for calculating derivatives of outputs (i.e., ambient quality) with respect to inputs (e.g., emissions) for applications that involve a large number of inputs and outputs.

Brute-force methods calculate derivatives by running the model twice (or more); once with the base-case emissions followed by perturbed emissions. The difference in outputs is then used to calculate the derivatives. The brute-force method has been widely used to develop source-receptor relationships for different applications such as cap-and-trade (e.g., Krupnick et al., 2000, Nobel et al., 2001), estimation of the health damage of pollution (e.g., Mauzerall et al., 2005), and intercontinental pollution transport estimation (e.g., Jacob et al., 1999, Fiore et al., 2009).

Forward and backward sensitivity analyses are two general categories of formal sensitivity analysis. They are both used to calculate the derivative of model outputs with respect to the model inputs. Forward sensitivities are source-based while backward (or adjoint) sensitivity methods are receptor-based. The efficiency of these two general categories of sensitivity analyses is critically related to the number of sources and receptors for which sensitivity information is desired. In one forward sensitivity model run, the derivative of all receptors with respect to one single source can be calculated, while in one single backward sensitivity model run, the sensitivity of a receptor or a concentration-based function (a function at a lumped group of receptors) with respect to all inputs can be estimated. Correspondingly, if the problem of interest is to determine the effect of a particular power plant on the ambient quality in different regions, forward sensitivity analysis is more efficient. However, if the

problem of interest is to find out the contribution of power plants to ambient air quality in different regions, then backward sensitivity analysis is the logical method of choice.

Assume a nonlinear operator $\boldsymbol{M}$**,** that projects the inputs vector X (a collection of inputs X_k for k-th input such as emissions from a power plant) into an output vector C (e.g., concentrations of different species, C_j) (Hakami et al., 2007):

$$C = \boldsymbol{M}X \tag{2-26}$$

Equation 2-26 is an operator form of the forward model. A perturbation in input such as emissions from a power plant, causes a perturbation in output such as the ozone concentration. The corresponding derivative equation is then:

$$\delta C = \boldsymbol{L}\delta X \tag{2-27}$$

where $\boldsymbol{L}$ is the Jacobian of model operator $\boldsymbol{M}$. Equation 2-27 is an operator form of the tangent linear model (TLM) or forward sensitivity model that is used to calculate sensitivities. The TLM is also referred to as the decoupled direct method (DDM; Dunker, 1981) in air quality modeling literature. To evaluate model output or use outputs for policy making purposes, one may define a metric (cost function) ($J = J(C)$) which is a function of concentrations (e.g., J could be the average ozone concentration over the domain or the square of the difference between the predicted and observed concentrations). The variation of the cost function J with respect to output C is:

$$\delta J = \langle \nabla_C J, \delta C \rangle \tag{2-28}$$

where $\langle .,. \rangle$ represents the inner product. Substituting equation 2-27 into 2-28 and applying the duality principle, equation 2-28 turns to (see Wang et al., 2001, Hakami et al., 2007):

$$\delta J = \langle \nabla_C J, \boldsymbol{L}\delta X \rangle = \langle \boldsymbol{L}^* \nabla_C J, \delta X \rangle \tag{2-29}$$

where $\boldsymbol{L}^*$ is the transpose of the TLM operator $\boldsymbol{L}$. This equation shows that the gradient of J with respect to input I is:

$$\nabla_I J = \boldsymbol{L}^* \nabla_C J \tag{2-30}$$

Equation 2-30 is the main adjoint equation in an operator form and can be used within air quality models to calculate the sensitivities.

The governing equation for the forward model in air quality models is the atmospheric diffusion equation (ADE) (Seinfeld and Pandis, 2006):

$$\frac{\partial C_i}{\partial t} = -\boldsymbol{u} \cdot \nabla C_i + \frac{1}{\rho} \nabla \cdot (\rho \boldsymbol{K} \nabla C_i) + R_i + E_i \tag{2-31}$$

where C_i is the mixing ratio of species i and varies spatially and temporally, $\boldsymbol{u}$ is the vector of the wind field and has components in three directions, ρ is the air density, $\boldsymbol{K}$ is the turbulent diffusivity tensor, R_i is the chemical reaction rate for species i, E_i is the emission rate of species i. The TLM and adjoint equations can also be developed starting from the ADE.

A perturbation in inputs (e.g., emissions) in equation 2-31 will perturb outputs (e.g., concentrations) and can be well represented by equation 2-32 (Hakami et al., 2004):

$$\frac{\partial \delta C_i}{\partial t} = -\boldsymbol{u} \cdot \nabla(\delta C_i) + \frac{1}{\rho} \nabla \cdot (\rho \boldsymbol{K} \nabla \delta C_i) + \boldsymbol{F}_i \delta \boldsymbol{C} + \delta E_i \tag{2-32}$$

where δC_i represents a perturbation in the concentration of species i, and $\boldsymbol{F}_i$ is the ith row of the Jacobian of the chemical reaction rates. If the adjoint cost function (J) is defined as the integration of the local cost functions (g) over different times (t) and locations (w) as follows:

$$J = \iint_{t\,w} g(\boldsymbol{C}, t, w)\, dt\, dw \tag{2-33}$$

Then, the governing adjoint equation 2-34 can be derived by applying the Lagrange multiplier to equation 2-32 followed by integration by parts (see Sandu et al., 2005; Elbern et al., 2000):

$$-\frac{\partial \lambda_i}{\partial t} = \nabla \cdot (\boldsymbol{u} \lambda_i) + \nabla \cdot \left(\rho \boldsymbol{K} \nabla \frac{\lambda_i}{\rho}\right) + \boldsymbol{F}_i^T \boldsymbol{\lambda} + \varphi_i \tag{2-34}$$

where λ_i is the adjoint variable for a specific time and location ($\frac{\partial J}{\partial C_i}$), $\boldsymbol{F}_i^T$ is the ith row of the transposed Jacobian of the chemical reaction rates (transpose of $\boldsymbol{F}_i$ in equation 2-32). φ_i, the forcing term, is the derivative of the local cost function with respect to the concentration ($\frac{\partial g}{\partial C_i}$) and is injected into the model for different times and locations in a manner equivalent to E_i in the forward ADE equation.

There are two approaches for the integration of the adjoint equation: the continuous adjoint and discrete adjoint. In the continuous approach, first the adjoint

equation is derived analytically and then the derived equation is solved numerically. In the discrete approach, an adjoint (transpose) model is developed for the discretized form of the forward equation.

In air quality modes, the ADE equation is usually solved numerically by a sequence of N time steps (Δt) through an operator splitting scheme. The result is an operator ($\mathcal{N}^{[t,t+1]}$) that relates the concentration at time step t (C^t) to the concentration at next time step (C^{t+1}) as shown in equation 2-35 (Sandu et al., 2005).

$$C^{t+1} = \mathcal{N}^{[t,t+1]} \circ C^t,\ C^N = \prod_{t=0}^{t=N-1} \mathcal{N}^{[t,t+1]} \circ C^0 \tag{2-35}$$

where C^0 is concentration at time 0 and $\mathcal{N}^{[t,t+1]}$ is the split numerical solution operator for the time interval $[t, t+1]$ and is calculated based on an splitting approach. In the operator splitting approach, various processes including chemistry, aerosols dynamics and thermodynamics, clouds, advection in three directions, horizontal diffusion, vertical diffusion, vertical advection, and emissions injections can be integrated separately and sequentially. The overall numerical solution for N time steps and m processes is then (Hakami et al., 2007):

$$\mathcal{N} = \left(\mathcal{N}_{p_1}^1 \circ \mathcal{N}_{p_2}^1 \circ \ldots \circ \mathcal{N}_{p_m}^1\right) \circ \ldots \circ \left(\mathcal{N}_{p_1}^N \circ \mathcal{N}_{p_2}^N \circ \ldots \circ \mathcal{N}_{p_m}^N\right) \tag{2-36}$$

where $\mathcal{N}_{p_i}^t$ is the numerical operator for process i (e.g., chemistry, horizontal diffusion, etc.) and time step t. Similarly, the forward sensitivity and adjoint model

with different governing equations are solved numerically. The numerical operators for forward sensitivity ($\mathcal{L}$) and adoint ($\mathcal{L}^*$) are:

$$\mathcal{L} = \left(\mathcal{L}_{p_1}^1 \circ \mathcal{L}_{p_2}^1 \circ \ldots \circ \mathcal{L}_{p_m}^1\right) \circ \ldots \circ \left(\mathcal{L}_{p_1}^N \circ \mathcal{L}_{p_2}^N \circ \ldots \circ \mathcal{L}_{p_m}^N\right) \tag{2-37}$$

$$\mathcal{L}^* = (\mathcal{L}_{p_m}^N)^T \circ (\mathcal{L}_{p_{m-1}}^N)^T \circ \ldots (\mathcal{L}_{p_1}^N)^T \circ \ldots \circ (\mathcal{L}_{p_m}^1)^T \circ (\mathcal{L}_{p_{m-1}}^1)^T \circ \ldots (\mathcal{L}_{p_1}^1)^T \tag{2-38}$$

The gradients of the cost function (J) with respect to the emissions (E) can be calculated from the adjoint variable (λ) during the integration (Hakami *et al*., 2007):

$$\frac{\partial J}{\partial E^t} = \left(\frac{\partial C^{t+1}}{\partial E^t}\right)^T \left(\frac{\partial J}{\partial C^{t+1}}\right) = \left(\frac{\partial C^{t+1}}{\partial E^t}\right)^T \lambda^{t+1} \tag{2-39}$$

Adjoint modeling has been widely used in meteorological studies since the 1970's (e.g., Marchuk, 1974) but not used in three-dimensional photochemical air quality models until the past decade. For air quality applications, an adjoint model was first developed for a Lagrangian model of the stratosphere (Fisher and Lary, 1995), and then for a Lagrangian model of the troposphere (Elbern et al., 1997). The first adjoint model for a chemical transport model (CTM) (i.e., European Air Pollution Dispersion Model, EURAD-CTM2) was developed by Elbern and Schmidt (1999). The adjoints of other air quality models have been developed recently including: HANK (Vukićević and Hess, 2000), CHIMERE (Vautard et al., 2000; Menut, 2003), STEM (Sandu et al., 2005), Polair (Mallet and Sportisse, 2005; Quélo et al., 2005), IMAGES (Müller et al. 2005), DRAIS (Nester et al., 2005), CIT (Martien and Harley, 2006), GOES-Chem (Henze et al., 2007), and CMAQ (Hakami et al., 2007).

CHAPTER 3:

METHODS AND MODELING SIMULATION TOOLS

In this chapter, the components of the proposed decision support system and the tools used therein are introduced and discussed. Note that this section is only a general overview of the methods and simulations tools used in this thesis, and the detailed methodology for various sections will be explained in greater detail in the respective chapters.

3.1. The decision support system (DSS)

A DSS will be proposed to evaluate the performance of different emission control policies. The DSS includes the CMAQ model, its adjoint, and an optimization model (KNITRO). The forward CMAQ is used to calculate the evolution of concentrations with time, the adjoint of CMAQ is used to calculate the sensitivities or marginal health damages, and the optimization model is used to estimate the polluters' emission response to different policies and determine the polluter's post-trade emission levels and related health damages. Figure 3.1 shows the components of the proposed DSS which can be configured differently. The first component of the proposed DSS is a module for selecting/defining an emissions control policy. The DSS is able to evaluate the performance of the following policies: 1) command-and-control (CaC); 2) cap-and-trade (CaT) with no exchange rate, which is the current policy in the U.S.; 3) cap-and-trade with an exchange rate (CaT-EX), which will be explained in detail in chapter 4; 4) damage minimization (DMIN); 5) and social cost

minimization (SCMIN), (4) and (5) will be explained in chapter 5; and 6) temporal social cost minimization (TMP-SCMIN) policies which will be discussed in chapter 6.

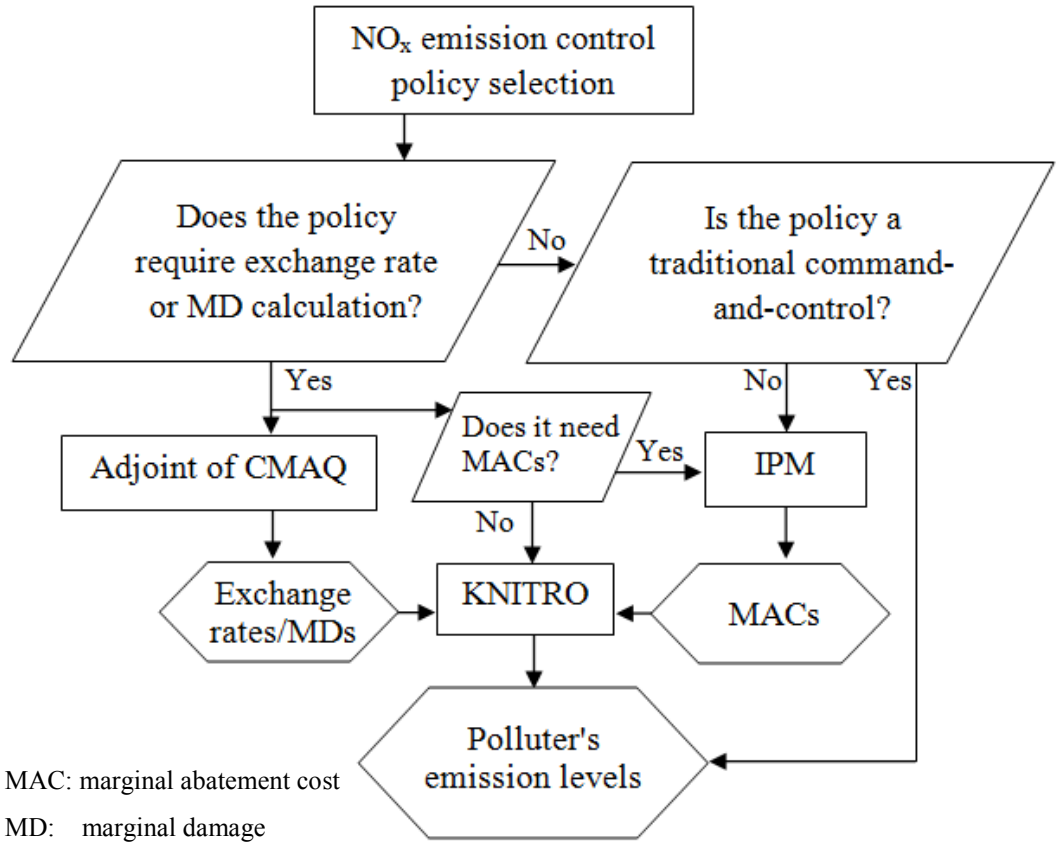

Figure 3.1. Components of the proposed decision support system.

Note that only the first two policies considered (i.e., CaC and CaT) do not account for exchange rates or marginal health damages. Under the CaC policy, emissions reduction goals are set by the environmental authority and polluters are required to keep their emissions at the regulation level. Therefore, there is no need to predict the polluters' emission behavior with an optimization model. Under a CaT

policy, allowances are allocated to the polluters, and it is the market that determines the post-trade level of emissions. For all other policies, emissions from different polluters are differentiated by exchange rates or marginal damages, and the adjoint model is used to estimate those values. The KNITRO model can be configured to find the optimal emission levels for different optimization objective functions and constraints. The details of the optimization problems and related formulations will be discussed in chapters 4, 5, and 6.

3.2. CMAQ and its adjoint

Chemical Transport Models (CTMs) have been widely used to predict ambient air quality for different emission control policies. CTMs (and regional CTMs or air quality models) can calculate the temporal and spatial concentrations of multiple species affected by transformation and transport processes. CTMs do not calculate the dynamics of the atmosphere (e.g., wind field or temperature) but use outputs of meteorological models for atmospheric dynamics. The model used in this study is the CMAQ developed by the U.S. EPA (Byun and Schere, 2006). CMAQ is the most widely used air quality model in the U.S. and worldwide.

The CMAQ modeling system (Models-3 modeling platform) has three main components (Figure 3.2): an emissions model that provides CMAQ-ready gridded and allocated emissions files from inventory data; a meteorological model that simulates the atmospheric state and motion; and a chemical transport model that includes advection, diffusion, photolysis rate computation, gas-phase chemical reactions,

aqueous-phase reactions, aerosol dynamics and thermodynamics, and dry and wet deposition (Byun and Schere, 2006).

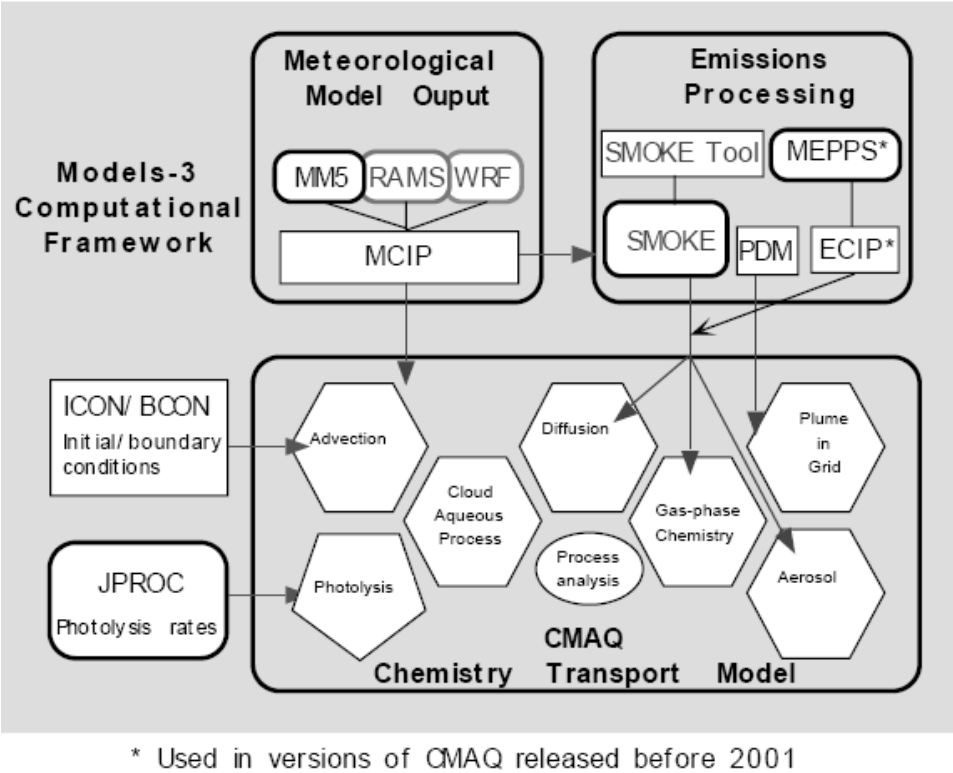

Figure 3.2. Model-3 components (Byun and Schere, 2006).

Other inputs for CTM include the initial and boundary conditions and photolysis rate coefficients. The initial condition and boundary conditions are processed in the initial condition processor (ICON) and boundary condition processor (BCON), respectively. The JPROC processor calculates the clear sky photolysis rate for different times and latitudes. The photolysis rates affect the rate of chemical reactions and depend on the time of day, season, atmospheric composition, and latitude.

CMAQ includes various chemical mechanisms including the Statewide Air Pollution Research Center, Version 1999 (SAPRC-99) photochemical mechanisms (Carter 2000) and the Carbon Bond version 5 mechanism (CB05) (Yarwood et al., 2005). Within each of these mechanisms, speciation and chemical reactions are defined. SAPRC-99 includes 214 reactions whereas CB05 includes 156. SAPRC-99 considers 80 different species whereas CB05 considers 59 species. The run time of SAPRC-99 is more than CB05 if all of the other conditions are equal (Luecken et al., 2008). The CMAQ model used in this thesis is the gas-phase variant of version 4.5.1 with the SAPRC-99 chemical mechanism.

The adjoint of CMAQ used in this study was developed for gas-phase processes[2] and was implemented in CMAQ version 4.5.1 (Hakami et al., 2007). As our application focuses on ozone concentrations, using the gas-phase version of the CMAQ adjoint is not problematic. The sequence of processes for the adjoint of CMAQ is similar to the forward CMAQ, but model integration is run backward in time. In the forward model, the concentrations of species are calculated for each time step by different processes, whereas for the CMAQ adjoint the sensitivities are calculated and written to the output files. The adjoint CMAQ also requires the concentration of species for nonlinear operators (for calculation of the transpose of the Jacobian in equation 2-34) such as the chemistry solver. As such, the forward CMAQ must save (i.e., write) the concentrations into checkpoint files before the adjoint run.

[2] The adjoint of aqueous-phase and the aerosol processes is under development.

However, for subsequent adjoint sensitivity runs under the same baseline conditions, checkpoint files from previous runs can be re-used. This results in significant savings in computational time if multiple adjoint runs are performed.

For the adjoint of CMAQ, a discrete adjoint scheme is employed for chemistry, diffusion, and vertical advection and a continuous adjoint is employed for horizontal advection. The CMAQ adjoint uses the kenetic pre-processor (KPP) version 2.2 for chemistry integration, (Damian et al. 2002; Daescu et al. 2000), the piecewise parabolic method (Colella and Woodward 1984) for the horizontal advection process, and an upwind first order finite deference method for vertical diffusion.

CMAQ can run in serial (i.e., with a single processor) or parallel (i.e., with multiple processors) mode. Parallelization of adjoint models is particularly important because adjoint simulations are inherently expensive. In this thesis, we take advantage of parallelization of the CMAQ adjoint model. In parallel mode, depending on the number of available processors, the domain is divided into several sub-domains among available processors. In order to control the communication between processors, CMAQ uses the MPICH message passing interface (MPI) which builds a connection between sub-domains (Bridges et al., 1995).

The CMAQ domain in our study is a North American domain with a 36 km grid size, and 34 vertical layers. The running period is form July 1st to September 30th (Chapter 4) or from June 1st to September 30th (chapters 5, 6, and 7) of 2007. The adjoint cost functions are ozone concentration, exposure to ozone (Chapter 4), and

short-term mortality health damages (chapters 5, 6, and 7) which will be discussed in further detail in corresponding chapters.

Interactions with CMAQ and adjoint is done through a shellscript which is used to run CMAQ within a Linux operating system. All variables required for running the model are set in the script. These variables include: the start date, end date, grid name, number of assigned processors for a parallel run, input data path, output directory path, species and layer range for integrated average concentrations, and a parameter to determine whether the user wants CMAQ to rewrite the existing outputs or keep them.

The shellscript used for CMAQ and its adjoint are very similar with some minor differences. For the adjoint model, runs go backward in time, and therefore, the user needs to set the end of the modeling period as the start of the adjoint model run. Also, the adjoint model requires some additional inputs whose names and paths should be set in the script. These additional inputs are checkpoints containing the concentrations of species and the adjoint forcing terms explained in section 3.2.1.

3.2.1. Adjoint cost function and forcing term

As explained in section 2.5, the adjoint model calculates the derivatives of a metric known as the cost function with respect to model inputs (equation 2-33). In this section, some examples for the cost function and the related forcing terms are presented. Note that the forcing term is the local derivative of the cost function with respect to concentrations. The forcing can be written by any code (e.g., Fortran code) after the forward CMAQ has generated the concentration files. This section does not

cover an explanation of how these codes are prepared. Instead, it details what the forcing terms are so that such code can be written in any programming language.

3.2.1.1. Spatial average ozone

Assume a simple modeling domain with only four grid cells with the ozone concentrations of C_1 to C_4:

C_1	C_2
C_3	C_4

If the cost function J is defined as the average of the ozone concentrations:

$$J = \frac{C_1 + C_2 + C_3 + C_4}{4} \tag{3-1}$$

Then, the forcing terms are:

$$\frac{\partial J}{\partial C_1} = \frac{\partial J}{\partial C_2} = \frac{\partial J}{\partial C_3} = \frac{\partial J}{\partial C_4} = \frac{1}{4} \tag{3-2}$$

In this case, one can simply make a forcing file with the following values:

1/4	1/4
1/4	1/4

3.2.1.2. Spatial and temporal averages of ozone concentrations

For a similar domain but over multiple time periods concentrations can be represented as:

C_{11}	C_{21}
C_{31}	C_{41}

C_{12}	C_{22}
C_{32}	C_{42}

where C_{it} represents the ozone concentration at grid i and time t, the average ozone cost function is defined as follows:

$$J = \frac{(C_{11} + C_{21} + C_{31} + C_{41})/4 + (C_{12} + C_{22} + C_{32} + C_{42})/4}{2} \tag{3-3}$$

Then, the corresponding forcing terms are:

$$\frac{\partial J}{\partial C_{11}} = \frac{\partial J}{\partial C_{21}} = \frac{\partial J}{\partial C_{31}} = \frac{\partial J}{\partial C_{41}} = \frac{\partial J}{\partial C_{12}} = \frac{\partial J}{\partial C_{22}} = \frac{\partial J}{\partial C_{32}} = \frac{\partial J}{\partial C_{42}} = \frac{1}{8} \tag{3-4}$$

In this case, one can make a forcing file with two time steps and the following values:

1/8	1/8
1/8	1/8

1/8	1/8
1/8	1/8

3.2.1.3. Spatially-averaged 1-hour maximum ozone

For the same modeling domain but over a 24 hour period, ozone concentrations are presented as:

C_{11}	C_{21}
C_{31}	C_{41}

C_{12}	C_{22}
C_{32}	C_{42}

...

$C_{1,23}$	$C_{2,23}$
$C_{3,23}$	$C_{4,23}$

$C_{1,24}$	$C_{2,24}$
$C_{3,24}$	$C_{4,24}$

where C_{it} is the ozone concentration at grid i and time t. For the average maximum 1-hour ozone, the cost function is:

$$J = \frac{Max1hr_1 + Max1hr_2 + Max1hr_3 + Max1hr_4}{4} \tag{3-5}$$

where $Max1hr_i$ is the maximum of 1-hour ozone concentration at grid i over 24 hours. The forcing terms for this example are as follows:

$$\frac{\partial J}{\partial C_{it}} = \begin{cases} 1/4 & t = t_{Max1hr} \\ 0 & t \neq t_{Max1hr} \end{cases} \tag{3-6}$$

where t_{Max1hr} is the time when the maximum 1-hour ozone occurs. For this case, one can make a forcing file with 24 time steps and assign 1/4 to the grids and times where and when the maximum 1-hour ozone occurs.

3.2.1.4. 8-hour maximum ozone

The process to make forcing files for 8-hour ozone is similar to that for 1-hour ozone. Assume the same example for 1-hour ozone, but this time for the following cost function:

$$J = \frac{Max8hr_1 + Max8hr_2 + Max8hr_3 + Max8hr_4}{4} \tag{3-7}$$

where $Max8hr_i$ is the maximum of 8-hour averaged ozone concentration for grid i, and is defined as follows:

$$Max8hr_i = Max\left(\frac{\sum_t^{t+7} C_{it}}{8}\right), \quad t = 1, \dots, 24 \tag{3-8}$$

The forcing terms for this example are as follows:

$$\frac{\partial J}{\partial C_{it}} = \begin{cases} 1/(4 * 8) & t \in t_{Max8hr}(i) \\ 0 & t \notin t_{Max8hr}(i) \end{cases} \tag{3-9}$$

where $t_{Max8hr}(i)$ is the time when the maximum 8-hour ozone occurs at grid i. For this example, a forcing file with 24 time steps can be created and 1/32 assigned to the grids and times where maximum 8-hour ozone occurs. Note the factor 1/32 in equation 3-9 comes from the factor 1/4 in equations 3-7 and the factor 1/8 in equation 3-8.

3.2.1.5. Exposure

Assume that a cost function is defined based on human exposure to ozone or population weighted ozone concentrations:

$$J = \sum_{t=1}^{24} P_i C_{it} \tag{3-10}$$

where P_i is the population at grid i, and C_{it} is the concentration at grid i and time t. In this case, the forcing terms are:

$$\frac{\partial J}{\partial C_{it}} = P_i \tag{3-11}$$

For this example, it is not necessary to create a forcing file, and the adjoint can simply be configured to read a population file within a subroutine called read forcing. Note that in equation 3-11, exposure is a linear function defined based on ozone concentrations. Linear cost functions do not require knowledge of model state, i.e. concentrations, whereas non-linear cost functions (e.g., maximum 8-h ozone) do require knowledge of concentrations from the forward run. For human exposure to 8-hour and 1-hour ozone, the process is similar but forcing terms only have non-zero values in hours when 8-hour and 1-hour ozone occur.

3.3. Other models

The proposed decision support system also includes other models including: a meteorological model, an emission processing model, and an optimization model as described below.

3.3.1. Meteorological model

The meteorological model used in this study is the Weather Research and Forecasting (WRF, version 3.1) model. WRF includes two main parts: the WRF Preprocessing System (WPS) and the dynamic simulation model. The WPS prepares inputs to the dynamic model and includes three sub-programs named geogrid, ungrib,

and metgrib. Geogrid defines the domain, ungrib extracts fields from the Gridded Information in Binary (GRIB) formatted data, and metgrib horizontally interpolates the extracted data. The dynamic model then vertically interpolates the data to the eta levels and integrates the state forward in time. An eta level is defined as a horizontal surface where the air pressure is constant.

In this thesis, the GRIB formatted North American Mesoscale (NAM) datasets are used (NCEP, 2005) for meteorological model initialization. The horizontal grid resolution is 36 km, and the domain uses a Lambert conformal projection with 165 columns, 129 rows, and 34 vertical layers (eta levels). The WRF output file cannot be directly used as input in CMAQ and it needs a conversion in format. The Meteorology Chemistry Interface Processor (MCIP) converts WRF outputs to a format that CMAQ can read (Byun and Schere, 2006). MCIP is able to change the domain into a smaller domain horizontally and vertically. In this thesis MCIP projects the WRF files into a smaller domain with 148 columns, 112 rows, and 34 layers which is the domain used for the CMAQ simulations. Note that a larger WRF domain allows for the calculation of boundary conditions for the smaller CMAQ domain.

3.3.2. Emission processing model

The emission model used in this thesis is the Sparse Matrix Operator Kernel Emission (SMOKE) version 2.4 (CEP, 2009). The SMOKE model converts the resolution of the emission inventory data to a resolution that matches that of the air quality model. Air quality models usually need hourly emission data for each species in each grid cell. In SMOKE, different sub-programs such as chemical specification,

temporal allocation, and spatial allocation convert the overall inventory data to the desired format. SMOKE can simulate toxic pollutants, particulate matter, and criteria pollutants. It can also provide air quality models with various outputs for various chemical mechanisms such as CB05 or SAPRC-99 chemical mechanisms.

SMOKE processes biogenic, area, non-road, mobile, and point sources separately and then merges them all into one file. Examples of biogenic and natural emissions include isoprene emissions from trees or NO_x emissions from lightening. Biogenic emissions in SMOKE are processed using the Biogenic Emission Inventory Systems 2 and 3 (BEIS2 and BEIS3) (CEP, 2009). Area sources such as forest fires or residential heating are spread over an area. Non-road emissions include emissions from boats, gardening equipment, and construction vehicles. For mobile sources SMOKE uses MOBILE6 (and in more recent versions MOVES) which calculates emission factors from activity inventories. The activity inventory includes vehicle type, speed, and vehicle mile traveled (VMT). The point sources in SMOKE can be generated separately for non-electric unit utilities (non-EGUs) and electric unit utilities (EGUs). The emissions from power plants are calculated in point-source processes and SMOKE recognizes a power plant by its ORISPL (Office of Regulation Information System Plant Location). Knowing ORISPL code from a plant and filtering the point source inventory file, the emission from a specific plant can be generated in SMOKE.

In this thesis, the Inventory Data Analyzer (IDA) emission inventory format is used. The U.S. EPA prepared the IPA files from the National Emission Inventory

(NEI). The IDA files include the yearly emission data, which SMOKE allocates into gridded and speciated hourly emission files. The IDA files were projected from 2005 to 2007 based on the change in population growth, industrial activities and land use patterns.

3.3.3. Optimization model

The optimization tool is used to optimize an objective function while the constraining criteria are met. The type of variables being optimized (i.e., real or integer), the type of objective function (i.e., linear, quadratic, nonlinear), and number of objective functions (single or multi-objective) and the form of constraints (linear, non-linear) have important impacts on the choice of the optimization technique. In this thesis, the KNITRO 8.0 optimization package is used to optimize a linear objective function.

KNITRO is an optimization software library for solving a variety of problems including: unconstrained and bound constrained problems; linear and non-linear least square problems; convex and non-convex quadratic programming problems; general convex and non-convex nonlinear non-linear constrained problems. The objective function and the constraints must be defined to run the KNITRO model. To solve a continuous nonlinear optimization problem, KNITRO uses a trust region approach based on the interior-point method (Byrd and Gilbert, 2000). In this method, the nonlinear programming problem is replaced by a series of barrier sub-problems. In each barrier sub-problem, one or more minimization steps are performed, then the barrier parameter is decreased and the process is repeated until the original problem

has been solved towards the global optimum and by the desired accuracy (Waltz and Plantenga, 2010).

There are three main methods which interact with KNITRO and provide the required inputs for defining an optimization problem within KNITRO. Firstly, an optimization problem can be defined through a modeling language such as AIMMS, GMAS, AMPL, or MPL. Secondly, the problem can be defined by a numerical computing environment such MATLAB. Lastly, a programming language such as Java, Fortran, C/C++, or Python can be used to define the optimization problem. For all three methods, the user needs to define the optimization objective function and constraints. Defining an optimization problem with a programming language has more complexity because the user must calculate first and second order derivatives (i.e., Jacobian and Hessian matrixes). However, this method provides more control over the techniques used for solving the optimization problem (Waltz and Plantenga 2010).

CHAPTER 4:

IMPROVING NO_X CAP-AND-TRADE SYSTEM WITH ADJOINT-BASED EMISSION EXCHANGE RATES[3]

4.1. Introduction

Acute and chronic exposure to surface ozone, even in low concentrations, is harmful to public health (Bell et al. 2004; Jerrett et al. 2009). Ozone is formed in the atmosphere through a series of photochemical reactions in presence of adequate amounts of NO_x (i.e., NO_2 and NO), volatile organic compounds (VOCs), and sunlight. The effectiveness of NO_x emissions in producing ozone is greatly affected by factors such as location, time of release, and the predominant chemical regime along the transport trajectory. In other words, NO_x emissions in different times and locations have different ozone formation potentials. As such, an environmentally effective NO_x cap-and-trade program needs to account for temporal and spatial variability in ozone formation potential of various emissions trading partners.

Several regional NO_x trading programs have been established in the past decade such as the NO_x Budget Trading Program (NBP) and the Clean Air Interstate Rule (CAIR), which replaced the NBP in 2008. Under NBP and CAIR, each state would issue a certain number of allowances (one allowance is one short ton of NO_x emissions), and then firms could trade their allowances, even between states. In NBP

[3] This chapter is a reformatted version of the following published article: (Mesbah et al. 2012).

or CAIR, one unit of emissions at one location can be traded for one unit of emissions at any other location. Trading permits on a one-to-one basis is not problematic for long-lived pollutants such as CO_2 because they live long enough to be uniformly mixed. However, the short lifetimes for ozone or its precursors result in spatial variability in net ozone formation along the transport path. In particular, the response of ozone to NO_x emissions is very location-dependent and nonlinear (Hakami et al. 2004), as NO_x can either help or hurt the production of ozone depending on the governing atmospheric regime. Within the current cap-and-trade framework, emissions are only differentiated based on their associated abatement costs. If all emissions are treated equally (like the CAIR program), trades may cause deterioration of ambient air quality despite a decreasing total emission cap. However, differentiating between emissions based on ozone formation potential has not been implemented by any NO_x trading program in the U.S. since previous studies did not find a significant benefit for such differentiation (EPA 1998; Krupnick et al. 2000).

One of the pioneer cap-and-trade methods is the ambient permit system (APS) where polluters trade the ambient quality rather than the emissions (Montgomery 1972). APS is hard to implement due to complexities associated with collecting ambient permits from different receptors. Other trading systems allow the exchange of emission permits either with variable (Atkinson and Tietenberg 1984; Krupnick et al. 1983; McGartland and Oates 1985) or fixed exchange rates (Farrow et al. 2005; Hung and Shaw 2005; Klaassen et al. 1994; Krupnick et al. 2000). The exchange rate is the ratio by which a polluter can increase its emissions after buying a unit of emission

from another polluter. Variable exchange rate systems are based on external restrictions on emissions trading (Atkinson and Tietenberg 1984; Krupnick et al. 1983; McGartland and Oates 1985). To ensure whether restrictions are met, these systems need air quality modeling for each trade and therefore are computationally expensive to design and maintain. Fixed exchange rate systems are easier to implement since the polluters trade their permits based on the exchange rates determined by the environmental authority. Fixed exchange rates are defined in different ways such as the ratio of marginal abatement costs of polluters (Klaassen et al. 1994), the ratio of polluters' impact on ambient quality (Hung and Shaw 2005), the number of people exposed to pollution (Krupnick et al. 2000), or the ratio of marginal damage of the pollutants (Farrow et al. 2005; Muller and Mendelsohn 2009; Muller 2011).

The calculation of exchange rates requires abatement cost information or the source-receptor relationships. As traditional approaches for quantifying source-receptor relationships require numerous simulations, calculating these relationships for a realistically long period has been infeasible. One way to overcome this high computational cost is using simplified versions of air quality models (Muller and Mendelsohn 2009; Muller 2011). Simplified models are computationally inexpensive but they do not account for physical and chemical processes in the atmosphere as well as the complexity involved in calculation of some parameters such as the photolysis rates (Binkowski et al. 2007). Another complication in using simplified models is that

ozone is a secondary pollutant, and therefore, the source-receptor relationships need to account for transformation of NO_x (and VOCs) to ozone.

In this work, we propose a sensitivity-based exchange rate system that employs an adjoint of the U.S. EPA's Community Multiscale Air Quality (CMAQ) model (Hakami et al. 2007). The efficiency of the proposed system is evaluated through a simulation-optimization model that can predict NO_x emissions trading behavior and post-trade ozone concentrations. The environmental outcomes and cost performances of the proposed and current policies are examined in a case study for 218 coal-fired electric generation units that participated in the NBP during the ozone season of 2007.

4.2. NO_x control decision support system

Figure 4.1 shows the components of a proposed decision support system consisting of three policies: 1) command-and-control (CaC), which assumes all sources keep their emissions at the initial allocated levels, 2) cap-and-trade (CaT) with no exchange rate which is the current policy in the U.S., and 3) cap-and-trade with an exchange rate, which will be explained in more detail later. The exchange rate policy has been examined for two different scenarios: a) for the average daily maximum 8-hour (ADM8) ozone (CaT-EX), and b) for ozone exposure (CaT-EXP). Exposure is calculated based on 8-hour ozone to preserve consistency with the National Ambient Air Quality Standard (NAAQS). Note that ozone exposure in this work is defined as the population weighted ozone concentrations. This simplified definition does not account for all the complexities of personal exposure assessment

such as the time spent indoors, activity level, or age. If the CaC policy is chosen, the emission distribution is predetermined, and the ambient air quality is calculated based on allocated emissions. If the selected policy is one of the two cap-and-trade policies, the optimization model is required for abatement cost minimization. The optimization model minimizes the total abatement cost in the system subject to different policy constraints by using plant-specific abatement cost information. For CaT-EX(P) policy, an environmental constraint limits the trade by fixed exchange rates. The exchange rates are calculated by the adjoint of gas-phase CMAQ. These simulations provide location-specific sensitivities that are used to establish exchange rates between trading partners based on their respective influences on ozone concentrations. The outputs from the optimization model are the total abatement cost and post-trade emissions distribution. The latter can be used for prediction of ozone concentrations.

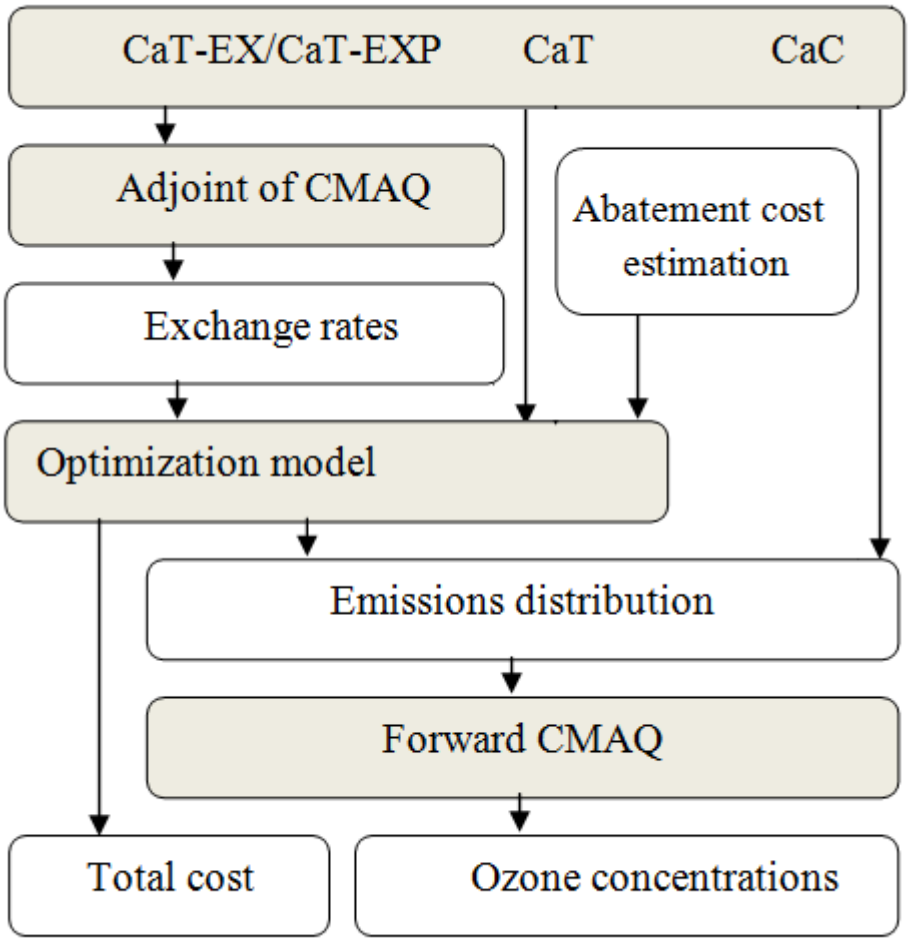

Figure 4.1. The components of the proposed decision support system.

4.3. Sensitivity analysis

Various sensitivity analysis methods can be used to calculate derivatives of concentrations with respect to emissions. Formal sensitivity methods entail integration of equations that carry sensitivity information. Two general categories of formal techniques for local sensitivity analysis are forward (such as the decoupled direct method (Dunker 1984) and backward (or adjoint) methods. Forward sensitivity

models integrate evolution of model sensitivities forward in time while adjoint simulations integrate an auxiliary set of model sensitivities backward in time and space. Both forward and adjoint sensitivity methods calculate the derivatives of outputs (e.g., ambient quality) with respect to inputs (e.g., emissions). Forward sensitivities map influences of a predetermined single source (or a scaling of multiple sources) on all receptors (Dunker 1981, 1984; Hakami et al. 2003) but do not provide influences from individual sources on the receptors. Conversely, adjoint sensitivities map influences on a predetermined single receptor (or on an integrated metric based on multiple receptors) from individual sources (Hakami et al. 2007; Henze et al. 2007; Sandu et al. 2005). Therefore, adjoint sensitivities are considered source-specific (as they provide sensitivity information about individual sources) but they lack receptor specificity because they do not calculate the response at each receptor (Hakami et al. 2006). The efficiency of these two methods is related to the number of sources and receptors of concern in a specific application. Adjoint sensitivity analysis is significantly more efficient when influences (i.e., sensitivities) from many sources (of different species) are sought on a predetermined single receptor or receptor-based metric.

The focus of this chapter is on differentiating between individual sources in terms of their potential influence on ozone concentrations. Therefore, adjoint sensitivity analysis is the natural method of choice. As mentioned before, adjoint analysis can also be carried out for an integrated metric (in space and/or time) instead of one single receptor. This metric is referred to as the adjoint cost function and

should not be mistaken with the "cost function" used in the abatement cost minimization within an emissions trading framework. The adjoint cost function can be defined as any function of concentrations of one or multiple species such as average ozone, maximum 8-hour ozone, ozone exposure, total mortality (e.g., from ozone and $PM_{2.5}$), etc, as long as the functionality is mathematically established and differentiable. This function can be integrated over receptors, times, model columns, or species. In this study, the adjoint sensitivity term λ_i is the sensitivity of the ozone metric (J) to NO_x emissions from polluter i, and the adjoint of the gas-phase CMAQ model is used to differentiate between NO_x emissions based on their adjoint sensitivities as a measure of influence. More details about the adjoint of CMAQ and its formulation and verification can be found elsewhere (Hakami et al. 2007).

4.4. Exchange rate enhanced cap-and-trade system

Establishing a NO_x trading system with exchange rates requires setting certain parameters (i.e., ozone target, emission quota, and exchange rates) by the environmental authority. Then, participants must meet the requirements by varying their emissions and trading permits. This system involves the following steps:

1. The target ozone air quality metric, a function of ozone concentrations (e.g., ADM8 or exposure), is defined by the environmental authority. This metric should be an integrated scalar across all receptors. We compare the change in environmental performance with a baseline taken as the CaC policy where all sources emit at their allocated quota. ΔJ is then the change in the environmental metric as the result of

implementation of a trading policy. Accordingly, ΔJ_i^0 is the change in the environmental metric from trades for polluter i ($\Delta J = \sum_{i=1}^{n} \Delta J_i^0$) and is set by the environmental authority. We assume ΔJ_i^0 of zero to ensure that no environmental deterioration results from emissions trading for any source. However, the methodology presented below is general and ΔJ_i^0 can be set as a negative value to ensure a certain level of post-trade improvement in environmental quality.

2. The environmental authority allocates emission quota (e_i^0) to each polluter, as well as the system-wide cap ($e_t = \sum_{i=1}^{n} e_i^0$). The relationship between emissions from polluter i (e_i) and the ozone metric is:

$$\lambda_i \Delta e_i = \lambda_i (e_i - e_i^0) = \Delta J_i \tag{4-1}$$

where Δe_i represents the change in emissions from polluter i compared to the allocated level, and ΔJ_i is the resulting change in the ozone metric caused due to Δe_i.

3. Exchange rates (α_{ij}) are set by the environmental authority. α_{ij} is the amount of emissions polluter j can increase by buying one unit of emissions from polluter i:

$$\alpha_{ij} = \frac{\lambda_i}{\lambda_j} \tag{4-2}$$

where λ_i and λ_j are the ozone metric sensitivities to NO_x emissions from polluters i and j, respectively.

In a CaC system, the environmental authority sets the emission quota, and all polluters must maintain their emissions below the allocated quota (i.e., $e_i \le e_i^0$) which will ensure a negative or zero ΔJ_i. However, under a CaT system, polluters can have a higher emission level than their allocated quota if they buy permits. Theoretically, a CaT policy will minimize the system-wide abatement costs by directing emission reductions to least-cost cases. Therefore, we simulate post-trade emissions distribution by minimizing the system-wide abatement costs:

$$Minimize \ \sum_{i=1}^{n} c_i (e_i^0 - e_i) \tag{4-3a}$$

Subject to :

$$e_i \le e_i^0 - \sum_{\forall j \ne i} T_{ij} + \sum_{\forall j \ne i} T_{ji}, \ i = 1,...,n; j = 1,...,n, \tag{4-3b}$$

$$T_{ij}, T_{ji} \ge 0, \tag{4-3c}$$

$$e_i \in [0, e_i^{\max}]. \tag{4-3d}$$

where c_i is the abatement cost for polluter i and T_{ij} is the amount of quota that polluter i sells to polluter j. The post-trade emission level, e_i, is the minimization variable, and the maximum possible post-trade emission for a plant, $e_i^{\max}$, is dictated by the plant's physical capacity and its available control technology. Note that equation 4-3b is a sufficient condition for maintaining the system-wide cap.

Under a CaT-EX system, equation 4-3b can be modified to include the exchange rates (equation 4-4b) and change in environmental metric:

$$Minimize \ \sum_{i=1}^{n} c_i (e_i^0 - e_i) \tag{4-4a}$$

Subject to :

$$\lambda_i (e_i + \sum_{\forall j \neq i} T_{ij} - \sum_{\forall j \neq i} \alpha_{ji} T_{ji} - e_i^0) \leq \Delta J_i^0, \ i = 1,...,n; j = 1,...,n, \tag{4-4b}$$

$$\sum_{i=1}^{n} e_i \leq e_t, \tag{4-4c}$$

$$T_{ij}, T_{ji} \geq 0, \tag{4-4d}$$

$$e_i \in [0, e_i^{\max}]. \tag{4-4e}$$

where e_t is the total cap on emissions and is added to the system to ensure that the total emissions remain below the total cap (equation 4-4c). The outputs of this optimization model are the optimal trading pattern, the corresponding post-trade emissions distribution, and abatement costs. Note that while equation 4-3b defines the total constraint in a traditional cap-and-trade system, equation 4-4b effectively imposes a cap on the ozone concentration metric. In other words, incorporation of exchange rates into the existing NO_x emissions trading would transform the system into an ozone trading system. As CAIR is mainly concerned with ozone, this transformation is more likely to result in improved target air quality.

4.5. Case study

We evaluate the efficiency of the proposed decision support system in a case study of 218 coal-fired electric generation units with post-combustion control

technologies (either SCR or SNCR) that participated in the NBP program in 2007. Emissions trading has long-term impacts and can affect strategic planning of generating units. Plants would plan to employ control technologies to properly position themselves in the trading framework. The decision of the plant will depend on a multitude of factors such as the cost of control technology, environmental constraints, and electricity and fuel prices. In the short-term (i.e., within the ozone season), it is reasonable to assume that capital is fixed. Firms then only have the choice to buy or sell emission permits and to adjust output (i.e., electricity generation). We assume that output adjustments due to emissions trading does not incur any opportunity cost from foregone electricity generation, which is consistent with other studies (EPA 2010a; Fowlie 2010). In the short-term, therefore, differences in marginal abatement costs between plants are merely a function of given control technologies and operating costs. In other words, we evaluate the trading system with the assumption that no new control technology can be introduced during the ozone season of a given year. The short-term variable abatement cost for the units are calculated using the EPA Integrated Planning Model (EPA 2010a).

For air quality modeling, the adjoint of the CMAQ model version 4.5.1 is used to calculate sensitivities of the ozone metric to NO_x emissions. Forward and adjoint CMAQ models used in this study only include gas-phase processes because an adjoint for aerosol processes in CMAQ is not currently available. The simulation domain is the continental U.S. with 36 km grid resolution and 34 vertical layers, and the simulated period is from July 1st to September 30th of 2007. This coarse resolution

is chosen for reduced computational cost over the continental domain even though finer resolution may be required to better resolve local features such as NO_x titration in power plant plumes. The adjoint cost function for the CaT-EXP policy is defined as the ozone exposure based on the daily maximum 8-hour (DM8) ozone. The cost function for the CaT-EX policy is based on the ADM8 ozone for critical grids where DM8 ozone is higher than 60 ppb. SMOKE version 2.4 (emission model) and WRF version 3.1 (metrological model) and MCIP version 3.4 (meteorological post-processing) are used to generate inputs for CMAQ. Power plant emissions in SMOKE are augmented by Continuous Emission Monitoring (CEM) data[4]. For the abatement cost minimization modeling, a global optimization package, KNITRO version 8.0, is used (Waltz and Plantenga 2010). Our simulations have mean fractional error and bias of 16.5% and +5.5%, respectively, for a concentration cut-point of 60 ppb which is the same threshold used in the adjoint cost function definition.

4.6. Exchange rates

Exchange rates in this study are defined as the ratios of location-specific sensitivities (i.e., derivatives of the adjoint cost function with respect to emissions). These sensitivities are local derivatives and provide slopes. However, in a nonlinear

---

[4] Power plant emissions are from the CEM data for the year 2007, and emissions from other sources are based on the 2005 National Emissions Inventory (NEI) for the U.S., and the 2006 National Pollutant Release Inventory (NPRI) for Canada which are projected to the year 2007. This footnote was not included in the published paper.

response surface, such as that expected in atmospheric chemistry, these slopes can change as the system shifts from one state to another. For large perturbations in emissions or in a highly nonlinear system, these local sensitivities, and accordingly the exchange rates, may need to be recalculated as trades take place.

In our study we assume fixed exchange rates, but examine how this assumption is affected by the existing nonlinearities. Previous studies have indicated that NO_x emission changes up to about 30% result in a near-linear response (Cohan et al. 2005; Dunker 1981; Dunker et al. 2002). The exchange rates are calculated and compared for pre-trade and post-trade emission distributions (Figure 4.2). Due to computational costs associated with adjoint calculations, this comparison is performed for one randomly chosen week.

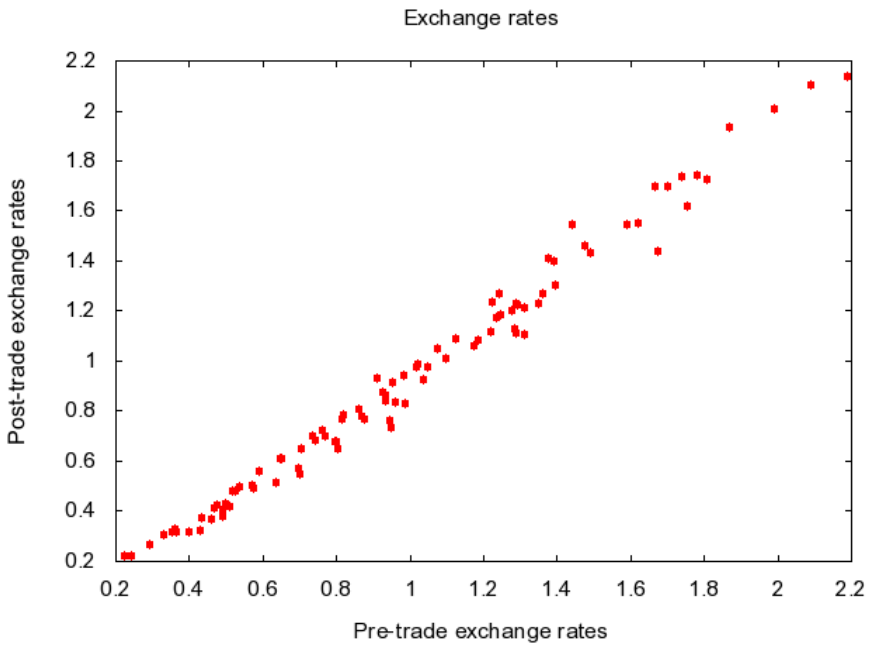

Figure 4.2. Exchange rates for all plants before and after trade (CaT-EX policy). All exchange rates are for trades with the plant with the average pretrade sensitivity.

Some plants shown in Figure 4.2 experience significant changes in their emissions as the result of trading. The ratio of post-trade to allocated emissions ranges between zero to 6.4 under various policies. Despite large changes in emissions, the exchange rates do not exhibit much sensitivity to changes in emissions distribution. Nonlinearity in atmospheric response occurs when there is sufficient change to trigger a shift in the chemical regime of the atmosphere (Hakami et al. 2004). Previous studies of nonlinearity in response to emission changes have used forward sensitivities where a domain-wide perturbation in emissions was employed. Domain-wide changes in emissions are much more likely to shift the atmospheric state and trigger nonlinear behavior than changes in individual sources. In other words, while some plants may have large trade volumes, these changes, while influential along the plume path, are not large enough to significantly affect the overall chemical state of the atmosphere. Therefore, a fixed set of exchange rates seems to be adequate for capturing differences in source influences on the adjoint cost function in our case study.

Results shown in Figure 4.2 only address nonlinearity and justify the use of fixed exchange rates within a single episode; exchange rates are likely to vary temporally across multiple episodes or years. Also, these results are calculated for a coarse resolution (36 km), which may include multiple power plants in one grid cell. A finer grid resolution is required to better delineate nonlinear chemistry in individual plumes. Furthermore, consistency of exchange rates shown in Figure 4.2 should not be interpreted as a linear response to emission changes. Exchange rates are ratios of

local derivatives and as long as two sources exhibit similar nonlinear behavior towards emissions redistribution, the ratio may remain unaffected while the sensitivities change. Finally, our case study of 218 power plants is not a close representation of the actual trading system in place. As the number of trading partners and the trading volume becomes larger than our case study, a more nonlinear behavior than that seen in Figure 4.2 is possible.

4.7. Results and discussion

Here, we contrast the current trading policy (CaT) to the proposed policies (CaT-EX and CaT-EXP) with regards to their impact on air quality (ozone concentrations) in the U.S., as well as the costs and benefits associated with their implementation. We examine each policy's impact on the average and the 99th percentile DM8 ozone over 92 days in the summer of 2007 (Figure 4.3). The 99th percentile is used as a proxy for the regulatory attainment dependence on the 4th highest concentration in the year. We assume that all 4 highest concentrations occur during the 5-month ozone season, and that our 3-month simulation captures 60% of those occurrences.

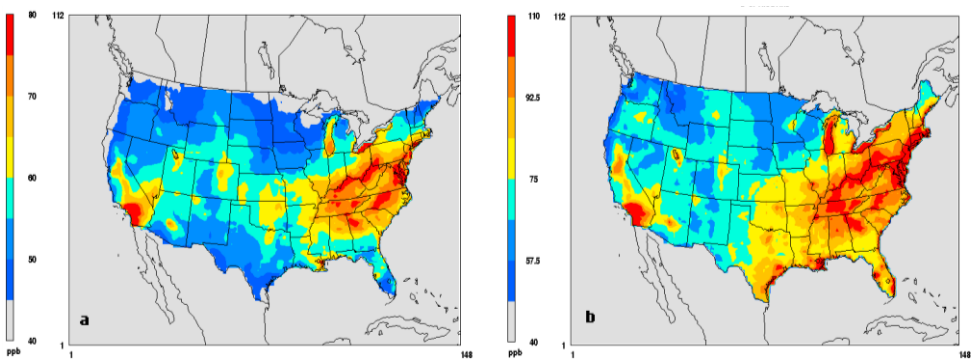

Figure 4.3. a) Average daily maximum 8-hour, and b) 3rd highest daily maximum 8-hour ozone for the base case simulation.

The choice of the adjoint cost function has some impact on the calculated exchange rates (Figure 4.4). The differences between exchange rates under CaT-EX and CaT-EXP policies are driven by the inclusion of population in the exposure-based adjoint cost function. While some differences exist, consistent features persist across both policies. For instance, lower than average exchange rates along the Ohio River Valley, particularly for CaT-EXP policy, reflect the existing chemical regime in this region. The high density of NO_x-rich plumes along the river creates a more persistent NO_x-inhibited chemical regime, resulting in smaller emissions reduction benefits (or even dis-benefits) immediately downwind of the source region before larger benefits materialize further downwind. As a result, the overall ozone improvement from NO_x emission control from this region is subdued.

To compare the environmental performance of various policies, post-trade ozone concentrations from CaT-EX and CaT-EXP policies are compared to that of CaT policy (Figure 4.5). CaT-EX and CaT-EXP have different influences on ozone concentrations as they are driven by somewhat different exchange rate distributions (Figure 4.4). Under the CaT-EX policy, the ADM8 ozone decreases in the eastern and southeastern U.S., while concentrations in some areas in the northeast increase (Figure 4.5a). Under the CaT-EXP policy, the ADM8 ozone decreases in the southeast without any significant increase in the northeast. The largest improvement in the ADM8 ozone (averaged over 92 days) is 1.1 ppb under CaT-EX and 2 ppb under CaT-EXP. Larger daily improvements are observed on numerous days.

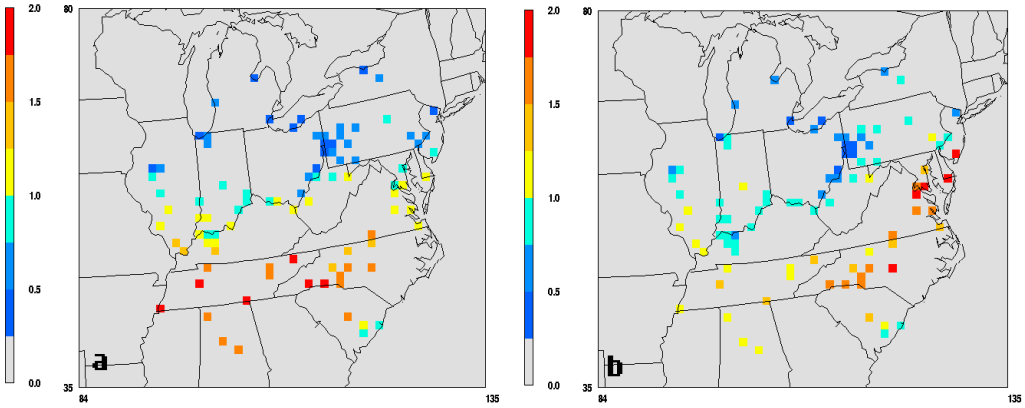

Figure 4.4. Normalized exchange rates for CaT-EX (a) and CaT-EXP (b) policies. The normalized exchange rate represents the ratio by which a power plant can exchange its emissions with the plant with the average sensitivity in the base case simulation.

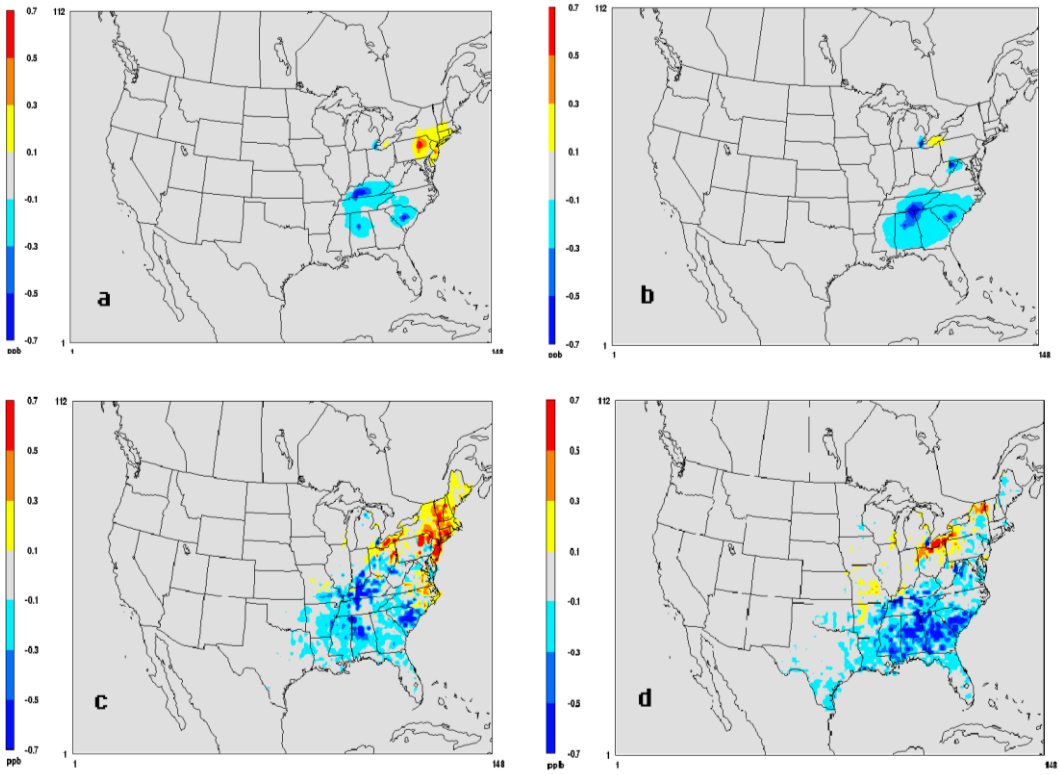

Figure 4.5. Differences between CaT and CaT-EXP (left) policies. The panels show differences in ADM8 ozone (top) and 99th percentile DM8 ozone (bottom). Negative values correspond to air quality improvements. Values in each plot are calculated (as average or percentile) from 92 daily values.

Under both CaT-EX and CaT-EXP policies, the 99th percentile DM8 ozone decreases in most areas in comparison to the CaT policy (Figure 4.5c and 4.5d). These improvements are spread over a wider area under CaT-EXP than CaT-EX, which is consistent with the simulated improvement in ADM8 ozone (Figure 4.5a and 4.5b). Like the ADM8 ozone, 99th percentile DM8 ozone concentrations under the CaT-EX policy increase in some areas of the northeast, but such instances of air quality deterioration are fewer under the CaT-EXP policy.

Various explanations can be offered for air quality deteriorations seen in Figure 4.5. In the case of CaT-EX policy, the adjoint cost function (which defines an environmental constraint for the trading system) only includes grid cells where DM8 ozone is larger than 60 ppb. A number of grid cells with deteriorating ozone are not included in the adjoint cost function (see Figure 4.3), and therefore, their environmental deterioration is not penalized in the optimization. A no-threshold adjoint cost function is likely to prevent these deteriorations but at the expense of reduced improvement in other regions. Also, both CaT-EX and CaT-EXP exchange rates are calculated for the whole episode, and therefore, environmental performances in individual days are not considered. Furthermore, ADM8 ozone and 99th percentile ozone performances are not explicitly part of the adjoint-based exposure constraint in the CaT-EXP policy. Finally, and most importantly, the adjoint-based exchange rates constrain the trades by their overall environmental impact. Therefore, a trade that results in deteriorated environmental performance in one region, but larger improvement in another, is considered permissible. This reflects an inherent limitation

in using adjoint sensitivity analysis for the development of exchange rates. While the adjoint method provides source specificity required for calculation of exchange rates for individual sources, it lacks receptor specificity to allow for inclusion of receptor-specific constraints.

Daily environmental performance of exchange rate policies is further examined in Figure 4.6 where the change in DM8 ozone for each grid is depicted. Under CaT-EX, largest improvements occur in areas where the DM8 ozone concentration under CaT is between 60 and 110 ppb (Figure 4.6a). Also, the deterioration outweighs the improvement in air quality for grids with DM8 ozone concentrations higher than 120 ppb. The pattern of change in DM8 ozone under the CaT-EXP policy is different than that of CaT-EX (Figure 4.6b) and most improvements occur when the DM8 ozone concentration under CaT is between 40 to 120 ppb. Like the CaT-EX policy, deterioration in grids with DM8 ozone concentrations above 120 ppb persists in the CaT-EXP policy. Frequency distribution plots (Figure 4.6c and 4.6d) for both policies suggest that improvements outweigh deteriorations, particularly for the CaT-EXP policy.

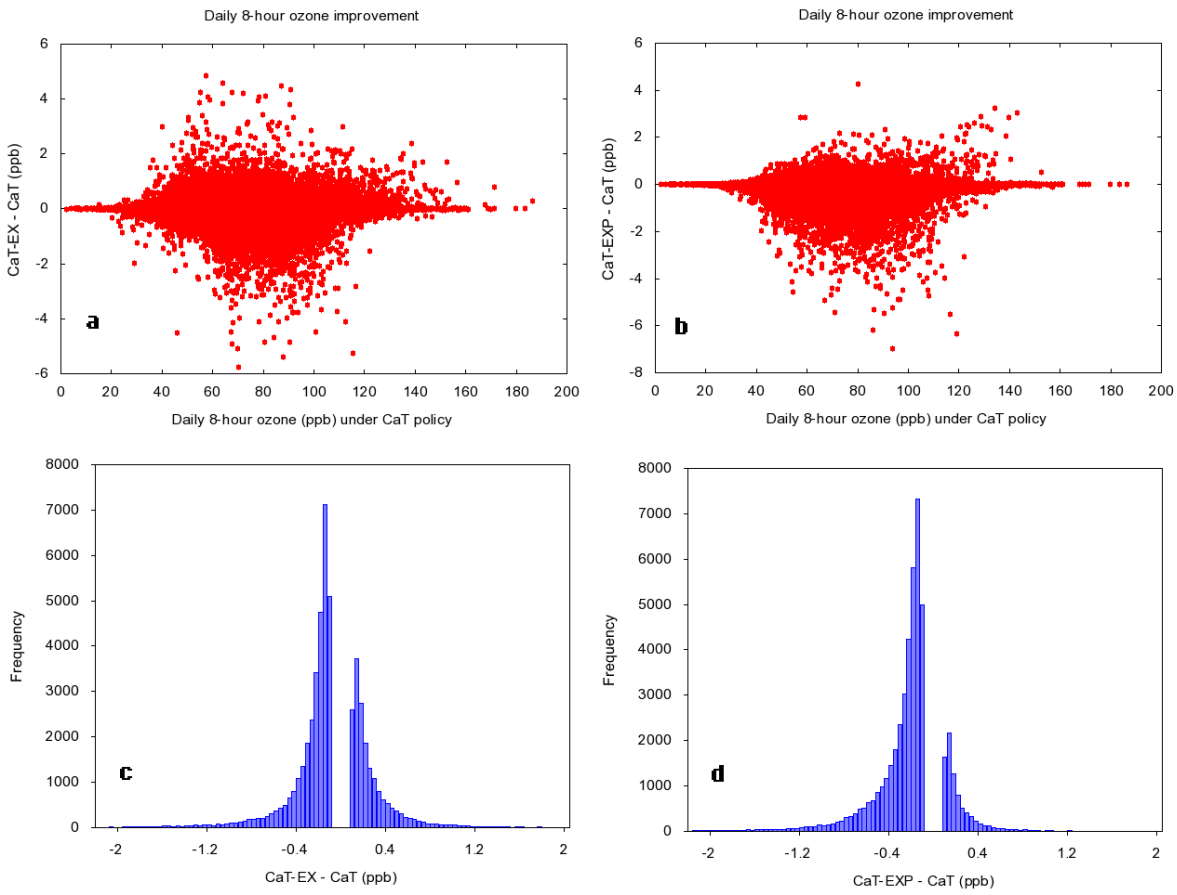

Figure 4.6. Change in DM8 ozone between CaT and CaT-EX (left) and CaT-EXP (right) policies. Each data point is the DM8 ozone in one grid-day. Days with small changes ([-0.1, 0.1]) are not included in the frequency distributions. Negative values indicate air quality improvements.

The abatement costs for various policies are compared in Table 4.1 with reference to the CaC policy. As expected, the CaT policy has the lowest cost as it is designed to minimize the system-wide abatement costs without any environmental constraint except the total imposed cap. The total abatement costs for the exchange rate policies are about 0.6 % greater than the CaT policy. It should be noted that the total emissions under all policies are the same and only the distribution of emissions change from one policy to another.

Abatement costs of various policies can also be compared to valuation of benefits resulting from changes in concentrations. Here we consider health benefits associated with reduced mortality due to short-term exposure to ozone in each policy. We assume the following relationship between changes in mortality (ΔM) and concentrations (EPA 1999b):

$$\Delta M = M_0 \, P(1 - e^{-\beta \Delta C}) \tag{4-5}$$

where M_0 is the baseline non-accidental mortality rate, β is an epidemiological concentration response factor, ΔC is the change in ozone concentration, and P is the population. M_0 is taken as a nationwide average of 0.7448 % for non-accidental mortality rate (Xu et al. 2010). The response factor (β = 0.00051 ppb^{-1}) is taken from Zanobetti and Schwarts (Zanobetti and Schwartz 2008) since they report values based on DM8 ozone concentrations. Mortality benefits are then monetized using the value of statistical life (VSL), taken to be $6.8 million in 2007 dollars (EPA 2010b). The high benefit-cost ratio for CaT-EXP in Table 4.1 indicates the effectiveness of this exchange rate policy compared to the current system (CaT). A move from CaT to CaT-EX does not result in net health benefits as CaT-EX imposes a constraint on ADM8 ozone rather than exposure. CaT-EX results in increased concentrations in more populous northeast regions compared to improvements in the southeast. The costs in Table 4.1 do not include administrative and transaction costs of trading. On the other hand, benefits are likely to be underestimated, as they do not include morbidity and mortality from long-term exposure to ozone. An analysis that includes all sources in the NBP is likely to result in correspondingly larger benefits and costs.

Table 4.1. Abatement costs and health benefits of CaT, CaT-EX, and CaT-EXP policies (million $).

	CaT	CaT-EX	CaT-EXP
Abatement cost savings (compared to CaC[1])	411	405	406
Health benefits for the season (switching from CaT)	-	-21	177

[1] Total abatement cost for the CaC policy in this case study is 1409.

It is important to emphasize that all trade-based policies in our case study, including those with exchange rates, are designed to minimize system-wide abatement costs; however, policies based on exchange rates do so while imposing an environmental constraint. As a result, cost performances of all trade-based policies are very similar. If emission redistributions were driven entirely or partly by environmental (or health) damages rather than abatement costs, benefits would be substantially higher but at the expense of higher system-wide abatement costs.

Our benefit-cost comparisons differ from those reported by Krupnick et al. (Krupnick et al. 2000). They examined the effectiveness of the NO_x cap-and-trade programs in the Eastern U.S. by calculating exposure-based exchange rates through source-receptor relationships extracted from air quality model simulations (Urban Airshed Model UAM-V) for three two-day episodes in 1990. They concluded that the EPA's CaT policy was considerably more cost-effective than a CaC approach but they did not find any discernible benefit in using a spatially differentiated (based on exchange rates) trading approach. There are important differences between Krupnick

et al. and our work that could explain the discrepancy in findings. First, emissions in the reference years are very different due to significant reductions in power plant emissions in recent years. Second, Krupnick et al. assigned significantly lower mortality benefits to reduced ozone exposure than this work due to the overall lack of consensus about the topic in epidemiological literature and the EPA at the time. Finally, and most importantly, Krupnick et al. used source-receptor matrices developed by forward sensitivity analysis, which limited their ability for spatial differentiation to 6 source regions. By contrast, using adjoint sensitivity analysis models, this study is able to differentiate between all 218 individual sources.

4.8. Conclusions

The results presented here are based on a number of simplifying assumptions and are meant as a proof-of-concept study rather than a conclusive endorsement of exchange rate enhanced emissions trading systems. Our study only included a fraction of all power plants in the U.S. and did not consider long-term abatement costs for the installation of new control technologies and the opportunity cost of electricity output reduction. Various sources of uncertainty, including those associated with air quality modeling, emission inventory, cost estimations, epidemiological factors, or the VSL, can affect our findings. Inter-annual consistency of exchange rates, multi-pollutant (including $PM_{2.5}$) performance of the proposed system, and the impact of model resolution are among topics that require further research. Finally, important practical issues such as details of exchange rate calculations, methods for exchange of credits, and monitoring of trades should also be addressed before implementation is feasible.

These limitations notwithstanding, our findings suggest that potentially significant benefits may arise from introduction of exchange rates into the existing NO_x emissions trading systems.

CHAPTER 5:

OPTIMAL OZONE REDUCTION POLICY DESIGN USING ADJOINT-BASED MARGINAL DAMAGE INFORMATION[5]

5.1. Introduction

Ozone is formed through photochemical reactions involving NO_x and volatile organic compounds (VOCs). Ozone, even in low concentrations, can inflict damage on human health (Bell et al. 2004), ecosystems, and agriculture (Mauzerall and Wang 2001). NO_x cap-and-trade (CaT) programs were established in the eastern U.S. to reduce surface ozone (EPA 2008). These programs have expanded over time and have led to substantial reductions in emissions because they were designed to control ozone indirectly by lowering the cap on total NO_x emissions. While ozone concentrations have decreased in much of the country, it is unclear whether the CaT system has resulted in environmentally optimal performance. The ozone formation potentials from NO_x emissions vary spatially depending on the governing atmospheric regime. Correspondingly, the damage to human health caused by NO_x emissions changes spatially. A suboptimal performance in the NO_x CaT program is possible since it does not account for the spatial heterogeneity in ozone-related damage. Recent studies have suggested that inclusion of spatial variability in NO_x health damage can improve the

[5] This chapter is a reformatted version of the following published article: (Mesbah et al. 2013).

performance of NO_x CaT programs (Mesbah et al. 2012; Muller and Mendelsohn 2009).

The socially acceptable emission level of a particular pollutant from an emitter depends on the marginal abatement cost (MAC) and the marginal damage (MD) caused by the polluter. The MAC for a polluter is the cost of reducing an additional ton of emissions, whereas the MD is the dollar value of (health, ecosystem, or climate) damage caused by an additional ton of emissions from the polluter. In a CaT system, tradable emission allowances are allocated to polluters and emissions must be kept lower than the available allowances for each polluter. Low-MAC polluters (i.e., polluters with a MAC lower than the market-based permit price) reduce their emissions and benefit by selling unused allowances. On the other hand, high-MAC polluters emit more than their allocated allowances and benefit from buying allowances at a lower price than their MACs. Therefore, under a CaT system, emissions are differentiated based on MACs through market driven cost-saving incentives, and the system-wide abatement cost is minimized with no consideration for source-specific MDs. Under a CaT system, the number of emission allowances is the only constraint on total damage (TD).

Whether the impacts of emissions are spatially homogeneous or location-dependent is an important aspect in emission control policy design. For pollutants for which the damages are not location-dependent, the TD does not change as the result of trading and redistribution of emissions. Therefore, a cap on total emissions can effectively reduce the TD. However, for pollutants whose damages are location-

dependent, the trading impact on TD is unknown. To limit the TD for these pollutants, a CaT system can be reformed by inclusion of exchange rates that are set by the environmental authority. An exchange rate for two polluters adjusts the exchange of allowances between the two polluters by a rate that can be defined as the ratio of their MDs.

In chapter 4, we showed that the proposed exchange rate CaT system could reduce ozone concentrations at little extra cost (Mesbah et al. 2012). While chapter 4 provided a formal framework for some consideration of MDs through exchange rates, it mainly aimed to minimize system-wide abatement costs and thus remained a CaT system in concept. When emissions are instead differentiated based on damages and the emitters are assessed fees based on their MDs, high-damage polluters have higher incentives to reduce their emissions. These incentives can be provided to high-damage polluters by either subsidizing their emission reduction, or by imposing a higher price on their emissions (Banzhaf 2004; Mauzerall et al. 2005). In this chapter, we investigate the environmental performance of two other NO_x control policies under which polluters' emissions are differentiated based on their impacts on public health. In these policies, the socially effective emissions distribution is found based on minimizing the inflicted damage rather than minimizing the system-wide abatement cost alone. To do so, we employ a recently developed approach for estimation of source-specific health damages through backward or adjoint sensitivity analysis in air quality models (Pappin and Hakami 2013).

5.2. Methodology

In a mathematical framework, MD for a polluter is defined as the derivative of the total damage function with respect to emissions from the polluter. To estimate source-specific MDs, the traditional brute-force approach is the most commonly used method in the literature. In this method, an air quality model (or a simplified representation of it) needs to be run twice for each individual polluter. The first run simulates the concentrations using baseline emissions from all sources, while in the second run emissions from the source of concern are increased (or decreased) by a certain amount. The MD for the polluter is the finite difference approximation of the total damage calculated in conjunction with epidemiological coefficients, which relate the change in concentrations to damage. The brute-force method for the calculation of MDs requires an additional simulation per polluter and therefore becomes computationally infeasible as the number of polluters increase. To overcome this computational barrier against the estimation of source-specific impacts, previous studies have used advanced air quality models but only for a limited number of sources (Mauzerall et al. 2005; Tong et al. 2006), or regionally grouped sources (Fann et al. 2009; Krupnick et al. 2000) or employed simplified or reduced form models for large number of sources (Fann et al. 2009; Muller and Mendelsohn 2009; Muller 2011). Grouping of sources assumes a prescribed regional homogeneity in MDs that is not supported by the existing literature. On the other hand, simplified models address the issue of computational cost but do not account for complex transport and

transformation processes in the atmosphere, and therefore lack accuracy in physical representations of source-receptor relationships.

A more recent method for the calculation of derivatives is backward or adjoint sensitivity analysis. The adjoint method is an effective approach for estimation of derivatives with respect to multiple parameters such as emissions at various locations. Unlike the brute-force method and the forward sensitivity analysis techniques (Dunker 1984; Yang et al. 1997), the adjoint method traces influences and sensitivities back to individual sources in a single simulation (Hakami et al. 2006). The brute-force and forward sensitivity methods provide receptor by receptor sensitivity information about the influence of a single source, or a single group of sources. As such, they can differentiate between the impacts exerted on different receptors but they cannot feasibly distinguish between individual source impacts. Adjoint sensitivity analysis, on the other hand, provides information about influences of individual sources on a collective metric and is therefore an ideal choice for differentiation between sources in policy or economic instruments. A more detailed description and formulation of adjoint sensitivity analysis in air quality models can be found elsewhere (Hakami et al. 2007; Henze et al. 2007; Sandu et al. 2005).

The critical point in an adjoint application is the policy metric for which sensitivity information is desired. This metric is known as the adjoint cost function (not to be mistaken with abatement cost). The adjoint cost function must be a scalar function of model outputs (i.e., concentrations) but can be integrated in space, time, or across multiple species. Examples of adjoint cost functions include average or

maximum domain-wide concentrations, attainment, total mortality, deposition onto a lake, crop damage caused by ozone, etc. In other words, so long as the policy concern can be expressed as a single number, the adjoint method is an ideal approach to provide source-specific sensitivity information. The total damage function is a good example of an adjoint cost function. In this work, we define the adjoint cost function as the total health damage, and therefore, our calculated adjoint sensitivities with respect to NO_x emissions are used as MDs. Our definition of the total health damage is limited and only refers to mortality due to short-term exposure to ozone and does not include morbidity, mortality from other pollutants, and ecosystem or climate related damage[6]. Change in total damage (ΔTD) is defined as (Anenberg et al. 2010):

$$\Delta TD = V_{SL} M_0 P (1 - e^{-\beta \Delta C}) \tag{5-1}$$

where V_{SL} is the value of statistical life, M_0 is the baseline non-accidental mortality rate, P is the population, β is the epidemiological concentration response factor which correlates the air pollution mortality to ozone concentrations, and ΔC is the change in ozone concentrations from a baseline. As β is usually a very small number, equation 5-1 can be linearized using a Taylor expansion (i.e., $\Delta TD = V_{SL} M_0 P \beta \Delta C$). The adjoint cost function is introduced into the adjoint model through the adjoint forcing term (φ) defined as (Pappin and Hakami 2013):

---

[6] Note that the damage function in this work is defined as the damage to human health in the U.S., and Canada and Mexico are not included. This footnote was not included in the published paper.

$$\varphi = \partial TD / \partial C = V_{SL} M_0 P \beta \quad (5\text{-}2)$$

The forcing terms in adjoint models can be regarded as the source of influences on the adjoint cost function and its role is similar to that of emissions in a forward model. In a forward model, the emissions are injected into the model and the concentrations are integrated after each atmospheric process (i.e., advection, diffusion, chemical reactions). Similarly, in an adjoint model, forcing terms are injected into the adjoint model, and derivatives evolve backward in time through each process and toward the originating sources. More details on using the adjoint model for sensitivity analysis of ozone mortality (or its cost valuation) can be found elsewhere (Pappin and Hakami 2013).

5.3. Optimization framework

An optimization framework is used as the prototype for a decision support system (DSS) under which the performance of damage-based emission-differentiated policies can be evaluated. The components of the proposed system are shown in Figure 5.1. The considered policies include: (1) a CaT policy which minimizes the system-wide abatement costs, (2) a damage minimization (DMIN) policy which minimizes damages for a given allowance allocation, (3) and a social cost minimization (SCMIN) policy which minimizes the summation of damages and abatement costs (i.e., the social cost) for a given cap. MDs and MACs for polluters are calculated by the adjoint of the Community Multiscale Air Quality (CMAQ) model and the U.S. EPA's Integrated Planning Model (IPM) (EPA 2010a), respectively. Under the CaT policy, only MACs are used in the optimization process

to minimize the abatement costs, and to predict the post-trade emissions distribution. On the other hand, under the DMIN policy, only MDs of polluters are used to differentiate between emissions for the minimization of the total damage. Under the SCMIN policy, the DSS uses both MACs and MDs to provide optimal emissions distribution under which the social cost is minimized. For all policies, the post-optimization emissions distribution can be used by CMAQ to simulate the spatial and temporal ozone concentration distribution, and the corresponding health benefit/cost due to adoption of that policy.

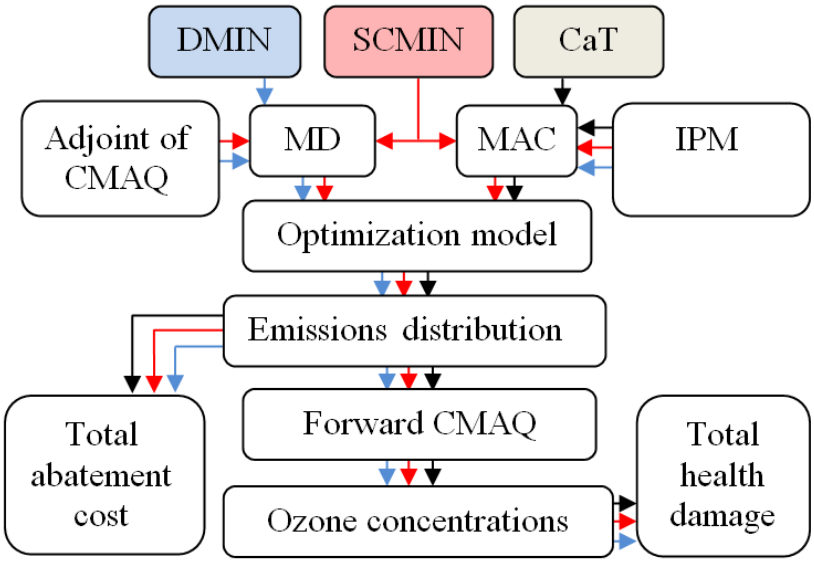

Figure 5.1. Components of the proposed decision support system for emission differentiation.

5.4. Optimization model

The optimization framework defined in this chapter is different from chapter 4 in two ways. First, the optimization problem in chapter 4 aimed to model the market in a CaT system and, therefore, minimized the total abatement cost for a given

aggregate emission target, whereas the optimization problem in this chapter minimizes the system-wide health or social cost. Second, the optimization problem in chapter 4 had an additional constraint for exchange rates, whereas such a constraint is not necessary in this chapter.

To estimate how power plants change their emissions under various policies, different optimization problems are constructed. Under the CaT policy the total abatement cost is minimized:

$$Minimize: \quad \sum_{i=1}^{n} c_i (e_i^0 - e_i), \tag{5-3a}$$

$$Subject\ to: \quad \sum_{i=1}^{n} e_i \le E_t, \tag{5-3b}$$

$$e_i \in [0, e_i^{\max}]. \tag{5-3c}$$

where c_i is the abatement cost of source i, n is the number of sources, e_i^0 is the allocated allowances for source i, e_i is a variable representing the ozone season emissions from source i, $e_i^{\max}$ is the maximum possible emissions from source i, and E_t is the total ozone season (May to September, inclusive) cap on emissions.

Under the DMIN policy, the optimization problem minimizes the total damage in the system:

$$Minimize: \quad \sum_{i=1}^{n} D_i (e_i), \tag{5-4a}$$

$$Subject\ to:$$
$$\sum_{i=1}^{n} e_i = E_t, \tag{5-4b}$$

$$e_i \in [0, e_i^{\max}]. \tag{5-4c}$$

where D_i is the damage function for source i which is defined as the MD of source i times its emissions ($D_i(e_i) = MD_i e_i$). Note that equation 5-4b is not the same as 5-3b because, if the summation of emissions from all sources has no lower bound, damage minimization will result in zero emissions for all polluters. The solution of the problem 5.4a to 5.4c leads to an emissions distribution that minimizes the damage. On the other hand, under the SCMIN policy, the social cost minimization problem considers both abatement costs and health damages:

$$Minimize:$$
$$\sum_{i=1}^{n} (D_i(e_i) + c_i(e_i^0 - e_i)) \tag{5-5a}$$

$$Subject\ to:$$
$$\sum_{i=1}^{n} e_i = E_t, \tag{5-5b}$$

$$e_i \in [0, e_i^{\max}]. \tag{5-5c}$$

5.5. Case study

The potential improvement in the current CaT system is examined through simulations for a case study of 218 coal-fired electric generation units. These are chosen as the major units with selective catalytic reduction (SCR) or selective non-catalytic reduction (SNCR) post-combustion control technologies that took part in the NO_x budget trading program in 2007. The allocated allowances to these 218 units

were about 37% of the total cap (534,000 tons) allocated to all 2594 participating units in 2007. The cost estimation in this study is for the short-term (i.e., one ozone season) when the capital costs are fixed and power plants do not get an opportunity to switch to other control technologies. Therefore, only variable costs (i.e., cost of operation and maintenance) are taken into consideration. In the short term, power plants trade allowances or adjust electricity generation to meet their allocation requirements. The short-term abatement costs are estimated using the U.S. EPA's IPM (EPA 2010a) and are based on the control technology and the capacity of the power plants.

For air quality simulation and sensitivity analysis, the gas-phase CMAQ version 4.5.1 and its adjoint are used (Hakami et al. 2007). The modeling period is the ozone season (May to September, inclusive) of 2007, and the modeling domain is the contiguous U.S. with a 36 km horizontal grid resolution and 34 vertical layers. The Sparse Matrix Operator Kernel Emissions (SMOKE) version 2.4 is used for emissions processing. The emission inventory files used by SMOKE are the 2006 National Pollutant Release Inventory (NPRI) for Canada, and the 2005 National Emissions Inventory (NEI) for the U.S., which are projected to 2007 based on the change in population growth, industrial activities and land use patterns.[7] The Weather Research Forecasting Model (WRF) version 3.1 is used to generate meteorological inputs. The

[7] The emissions for the power plants are based on the Continuous Emission Monitoring (CEM) data for 2007. This explanation was not included in the published paper.

CMAQ-ready meteorological inputs are processed by the Meteorology-Chemistry Interface Processor (MCIP) version 3.6. The mean normalized error and bias for the simulations for ozone are 18% and +1%, respectively.

For the optimization, a global optimization package (KNITRO, version 8.0) is used (Byrd et al. 2006). For the calculation of MDs, the value of statistical life (V_{SL}) is taken as $6.8 million in 2007 (EPA 2010b), and the baseline non-accidental mortality rate is calculated using the International Classification of Disease (ICD)-10 codes A─R (Bell et al. 2004).[8] The epidemiological concentration response factor (β) in this work is taken as 0.051% based on the daily maximum 8-h ozone (Zanobetti and Schwartz 2008) to conform with the averaging time of the current U.S. standard for ozone. We note again that our damage calculation is only based on short-term ozone mortality and does not include long-term effects, morbidity, particulate matter (PM) mortality, or environmental and climate damages.

5.6. Result and discussion

The two derivatives (i.e., MACs and MDs) used for distinguishing between emissions show significant spatial variability (Figure 5.2). The spatial variation in MACs is due to the variation in abatement costs for power plants depending on their

[8] The baseline mortality rates are spatially variable, which is different from the study conducted by Anenberg et al. (2010) which used a constant baseline mortality. This footnote was not included in the published paper.

location, electricity generation capacity, and emission control technologies. Power plants can change their MACs by switching to other control technologies in the long term, but they have little control over their MDs because they are determined by meteorology, governing atmospheric regime, baseline mortality rates, and population distribution. The outputs of the adjoint model are temporal and spatial MDs (i.e., the damage caused by 1 ton NO_x in any day at different locations). To calculate source-specific ozone season MDs, an emission-weighted aggregation of the daily MDs is used:

$$MD_{(x,y)} = \frac{1}{\sum_t \sum_k e_{kt}(x,y)} \sum_t \sum_k (e_{kt}(x,y) MD_{kt}(x,y)) \qquad (5\text{-}6)$$

where $MD(x,y)$ is the ozone season MD at grid (x,y); $e_{kt}(x,y)$ is the emissions at grid (x,y), layer k, and day t; and $MD_{kt}(x,y)$ is the daily MD at grid (x,y), layer k, and day t. MDs are different for various days and effective stack heights and the equation 5-6 preserves the relative spatiotemporal emission pattern of each source in the overall MD calculation. Non-weighted MDs over the ozone season are lower than those shown in Figure 5.2. Under the current CaT program, power plants have no restriction on their daily emissions and can adjust them to meet the daily electricity demand so long as their total ozone season emissions meet the requirements. Power plants are likely to generate more electricity (and emit more NO_x) on hot summer days when the electricity demand and the ozone formation potential are both likely to be larger (Martin 2008). Since high-emission days are also high-damage days, the weighted average results in higher seasonal MDs.

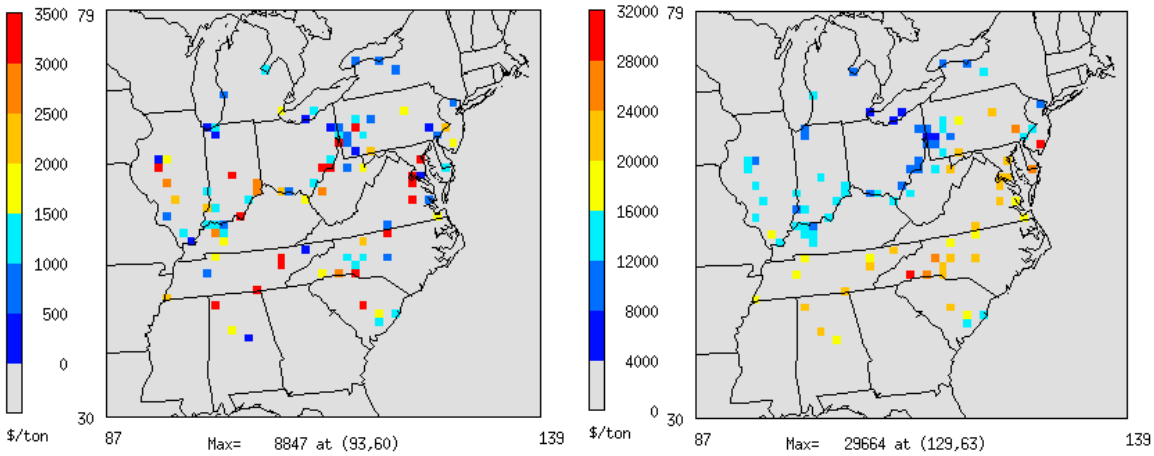

Figure 5.2. MACs (left) and MDs (right) for the power plants in the case study.

Ozone-based health damage from NO_x emissions can be negative (Fann et al. 2009; Pappin and Hakami 2013) but none of the power plants in this study had an overall negative seasonal MD. The negative MDs occur in NO_x-inhibited atmospheric regimes with high NO_x to VOC ratios. In such regimes, decreasing NO_x emissions results in increased ozone and consequently causes increased damage to human health. The power plants located in the Ohio River Valley have lower MDs compared to power plants elsewhere (Figure 5.2). In this region, numerous power plants emit a considerable amount of NO_x in a confined area which results in a NO_x inhibited atmospheric regime. As such, any decrease in NO_x emissions from these power plants leads to increased ozone (i.e., negative influence) in the valley but decreased ozone (positive influence) further downwind where the atmospheric regime becomes NO_x-limited. The overall damages from power plants in the valley are still positive because their downwind positive influences overweigh their local negative influences.

However, local negative influences cause MDs for these power plants to be lower compared to those in other locations.

Largest MDs are calculated for power plants in the eastern U.S. where a number of densely populated cities exist in a small region. Calculated MDs are also large for the power plants in southeastern U.S. where the governing atmospheric regime is predominantly NO_x-limited (Duncan et al. 2010). It should be noted that if a high-damage power plant is far removed from populous areas, it can be inferred that the damage occurs somewhere further downwind, but where it occurs is not known from the adjoint-based MDs as the adjoint method lacks receptor specificity (Pappin and Hakami 2013). To calculate the spatial distribution of the damage from a specific power plant or a group of power plants, forward sensitivity analyses are required.

To investigate the potential improvement in the system's performance, two damage-based emissions-differentiated policies (DMIN and SCMIN) are compared with the current CaT policy. The CaT policy is formulated to model the current trading system. The initial allocation of allowances under the CaT policy is based on the actual data from the U.S. EPA clean air market for the NO_x trading program. Under all three policies, the total emission cap is the same. Note that no new technology is introduced under different policies, and redistribution of emissions is the only reason for potential health benefits. The average daily maximum 8-hour (ADM8) ozone and the 4th highest seasonal daily maximum 8-hour (DM8) ozone for the CaT policy can be found in Figure 5.3. These metrics are of particular policy importance as the ADM8 ozone is the basis for health damage estimation (equation 5-

1) and the 4th highest DM8 (4DM8) ozone (averaged over three years) is the regulatory indicator of attainment. Our results suggest that the 4DM8 ozone for large areas in the U.S. were above the standard of 75 ppb for the simulated ozone season in 2007. Note that in this study, we assume that all of the four highest DM8 ozone days occur during the ozone season.

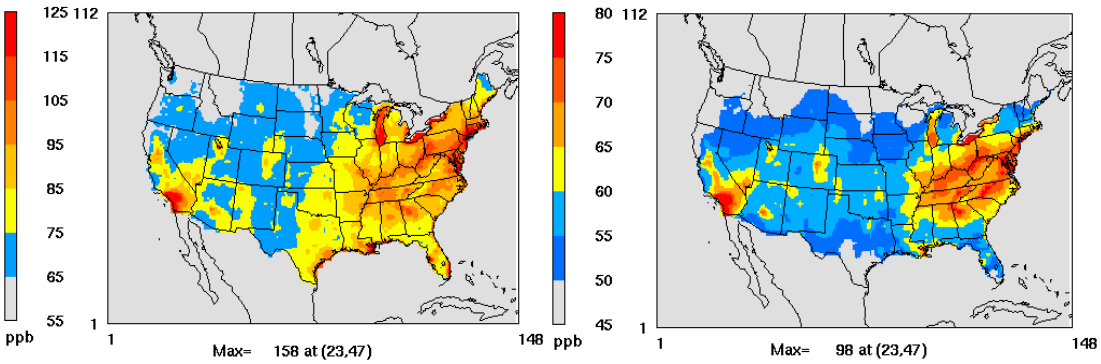

Figure 5.3. a) The ADM8 ozone (left) and b) the 4th highest DM8 ozone (right) for a 5-month modeling period for the CaT policy.

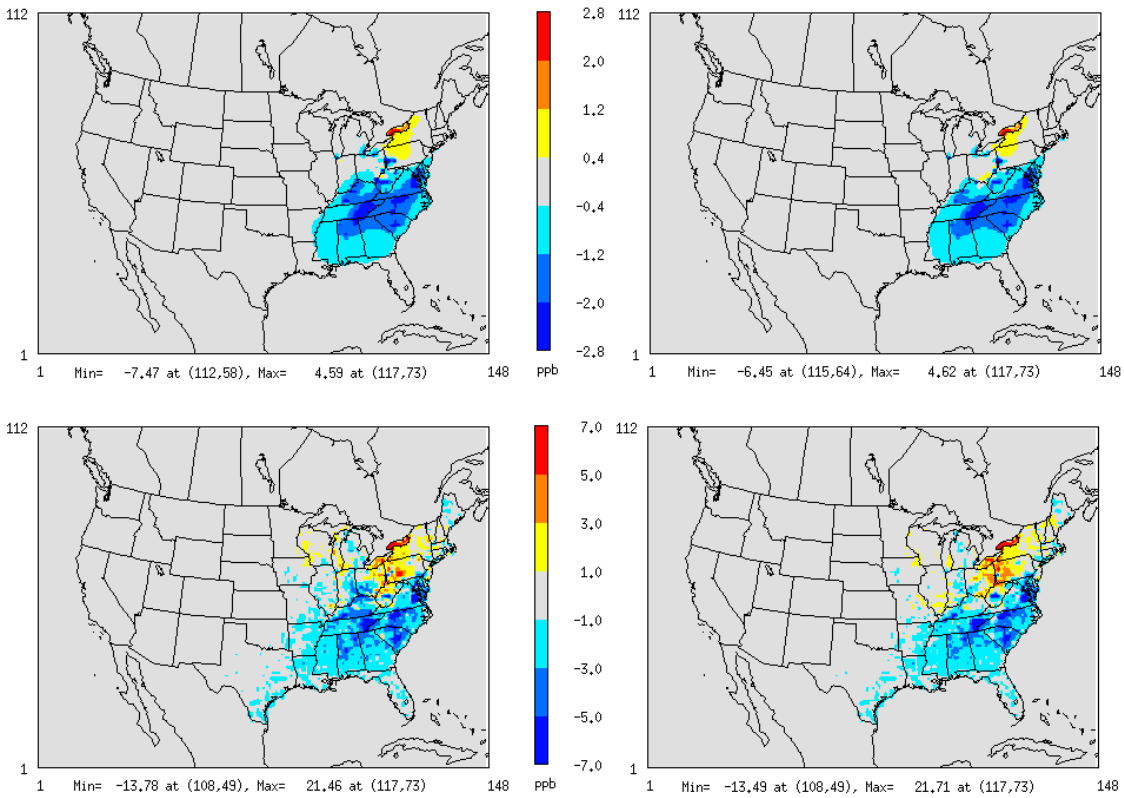

Figure 5.4. Differences in ozone concentrations by switching from the CaT policy to the DMIN policy (left), and to the SCMIN policy (right). The panels show change in the ADM8 ozone (top) and the 4th highest DM8 ozone (bottom) for the 5-month modeling period. Negative values represent improvements.

The ADM8 ozone and the 4DM8 ozone under the DMIN and SCMIN policies are evaluated in comparison with the CaT policy (Figure 5.4). The patterns of decreasing ozone concentrations are similar for the DMIN and SCMIN policies. The DMIN policy minimizes the total damage and the SCMIN policy minimizes the summation of damages and abatement costs. Since the MDs are higher compared to MACs (Figure 5.2), the dominant gradients in the minimization process for the SCMIN policy are MDs rather than MACs, hence the similarity of the impact induced by both policies. A consistent and significant decrease in both the 4DM8 ozone and

the ADM8 ozone occur in wide regions in the eastern and southeastern U.S. under both policies, while smaller increases are simulated in the Northeast and the Great Lakes region. It should be noted that the increase in ozone concentrations does not necessarily represent an increase in the total damage because some of the areas with increased ozone (e.g., over the Lake Ontario) are not populated and therefore have no contribution to the total health damage. Furthermore, air quality deterioration and increased damage in some locations is acceptable in our formulation, so long as this increased damage is outweighed by reduced damage elsewhere. Achieving air quality improvement in all locations would require multi-objective optimization, which is not straightforward with the adjoint approach.

The air quality improvements for the DMIN and SCMIN policies occur in a wide area in the eastern U.S. where ozone concentrations are high under the CaT policy (Figure 5.3). Under both damage-based policies, the maximum improvements in the ADM8 ozone and the 4DM8 ozone are about 7 and 14 ppb, respectively. The changes in air quality under different policies in Figure 5.3 occur when emissions are redistributed while constraining them by the same total cap. One of the limitations on health damage minimization is the maximum emission from low-MD units (i.e., equations 5-5c and 5-6c). If new power plants are built in low-impact damage regions, the possible improvement can be more than what is shown in Figure 5.4. Furthermore, the results presented in this chapter are based on spatial emissions differentiation and the temporal effect is not included. Shifting power plant emissions from high-damage to low-damage hours can also reduce ozone concentrations (Sun et al. 2012).

The air quality improvement in ADM8 ozone can result in increased health benefits if it takes place in locations with high baseline mortality rates and/or population. The decreased damages under different policies are calculated using equation 5-2 and presented in Figure 5.5. The decreased ADM8 ozone concentrations for the differentiated policies (shown in Figure 5.4) result in a corresponding decrease in damages (Figure 5.5). For both policies the benefits obtained by switching from the CaT policy to the DMIN or SCMIN occur in densely populated regions in the eastern U.S. Maximum benefits are estimated in Detroit, Michigan at $39 million (5.8 deaths/season) and $33 million (4.9 deaths/season) for the DMIN and SCMIN policies, respectively. Maximum increases in damage are $4.1 million under the DMIN and $4.4 million under the SCMIN policy, both in Buffalo and its vicinity.

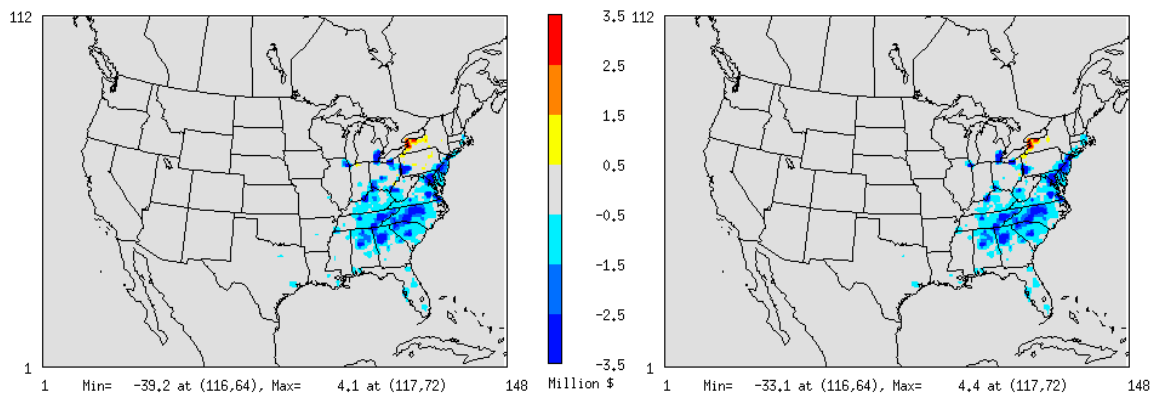

Figure 5.5. Differences in damages caused by the CaT and DMIN (left) or the SCMIN (right) policies.

The improvements in air quality under the DMIN and SCMIN policies are greater than those under an exchange rate CaT (CaT-EX), which was examined in chapter 4. Under CaT-EX policy, emission allowances for different power plants are

not treated equally and have different values. The exchange rate between two polluters determines a ratio under which their allowances can be traded. When these ratios are defined based on the MDs of polluters (or their contribution to exposure), the allowances allocated to high-damage power plants are assigned a higher value, which in turn motivates them to reduce their emissions and sell their allowances. Exchange rate inclusion can improve the environmental performance of a traditional CaT system but does not necessarily achieve the minimum social cost because different initial allocations of allowances under exchange rate systems can result in different outcomes (Førsund and Nævdal 1998; Krupnick et al. 1983). For pollutants whose damages are location-dependent, a CaT-EX policy limits the TD. This limit on the TD corresponds to the damage caused by the initial emissions allocation. To achieve an optimal state, a CaT-EX approach would lower the limit on the TD by inclusion of MD information in initial emission allocation, which in turn results in a lower pre-trade limit on TD. However, an exchange rate system is still a CaT concept designed to minimize abatement costs rather than the total damage. The improvements (in the 4DM8 and spatial health damage) under the CaT-EX policy in comparison to the CaT policy for the ozone season are presented in Figure 5.6. The maximum improvement under the CaT-EX policy as compared to the CaT policy is about 1.8 ppb for the ADM8 ozone and 4.1 ppb for the 4DM8 ozone which are considerably smaller than the corresponding values for the DMIN and SCMIN policies (Figure 5.4).

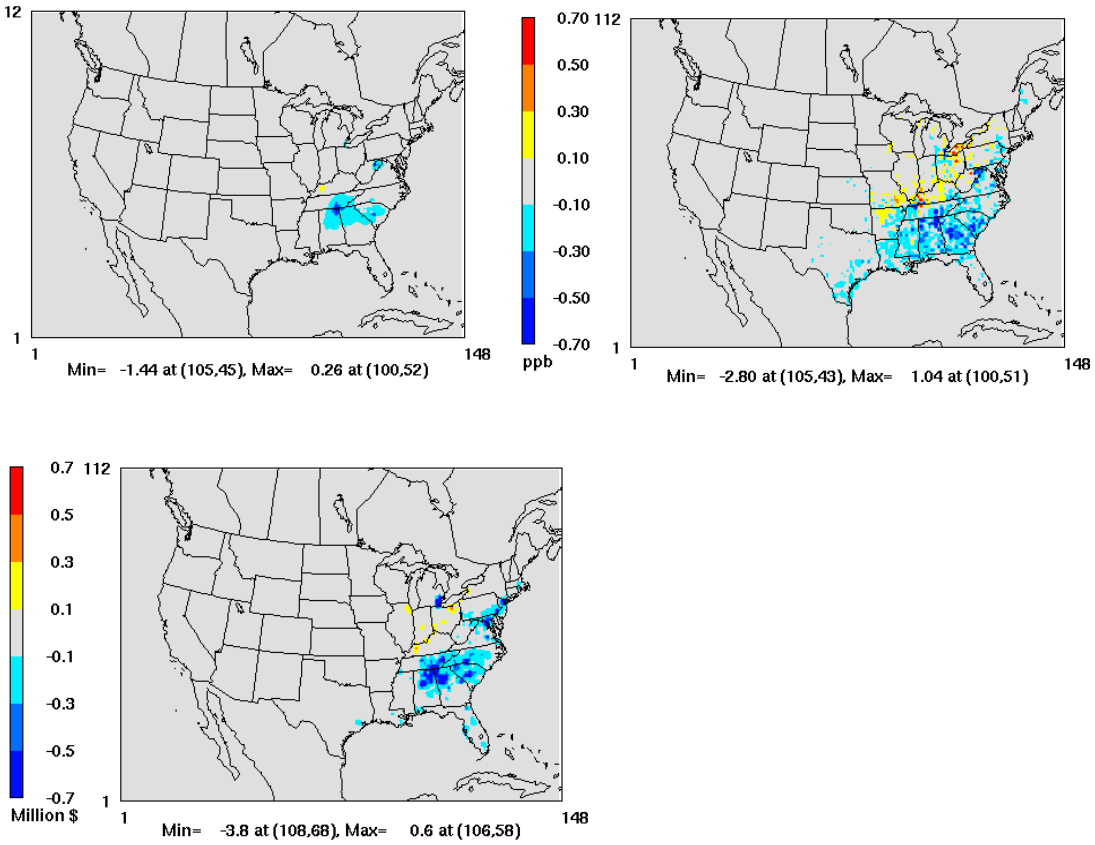

Figure 5.6. The spatial differences in the ADM8 ozone (up-left), the 4th highest DM8 ozone (up-right), and the spatial differences in health damage (bottom-left) comparing the CaT-EXP and CaT policies. The negative values represent improvement.

To better compare the heath damages under different policies with their associated costs, a summary of system-wide costs and benefits (i.e., reduction in the total damage) for all these policies are presented in Table 5.1. To examine the nonlinearity involved in the calculation of the total damage, the change in the total damage in Table 5.1 is calculated in two different ways: first, by using equation 5-1, i.e., the spatial summation of values presented in Figure 5.5, and second by using a linear estimation based on the adjoint MDs and optimization induced change in emission behavior:

$$\Delta TD = \sum_i MD_i \, \Delta e_i \tag{5-7}$$

where Δe_i is the change in emissions from polluter i occurring under different policies. The first method (i.e., equation 5-1) calculates the change in health damage from the change in ozone concentrations obtained by re-running CMAQ with post-optimization emissions distribution. As this method accounts for the nonlinearity involved in concentration response to changes in emissions, we refer to it as a global projection of health benefits. The second method (i.e., equation 5-7) assumes that MDs are constant and estimates the change in health damage by a linear extrapolation based on MDs, and therefore, we refer to it as the local projection of the health benefits.

Table 5.1. Abatement costs and health benefits for different policies ($ million).

Policy	Abatement costs compared to the CaT policy*	Health benefits compared to the CaT policy		Net benefits (Global health benefit minus abatement cost)
		Local projection	Global projection	
DMIN	520	1522	1189	669
SCMIN	372	1440	1089	717
CaT-EX	7	196	128	121

* The total abatement cost of the CaT policy is $411 million less than a command-and-control policy.

The values reported in Table 5.1 are short-term and do not represent long-term and implementation costs. Similarly, the benefits do not account for morbidity, reduction in secondary PMs due to reduced NO_x emissions, and long-term health benefits achieved by NO_x control, nor do they include environmental and climate benefits. The lowest costs are realized in the CaT policy which minimizes the total abatement cost, but the minimum total damage occurs in the DMIN policy which is designed to minimize the total damage with no regards for costs. Under the DMIN policy, reduction in health damage is higher, but the net benefit is lower, than that of the SCMIN policy.

Projection estimates in Table 5.1 suggest that MDs are not constant and vary with emission levels, as one would expect to see in a nonlinear system. If the MDs were constant, both methods for the calculation of change in the total damage would lead to the same result, but the difference between the results of the two methods is a measure of the degree of nonlinearity in damage estimations. Our benefit estimations indicate that there is up to 35% difference between the local and global estimations of damages, suggesting that MDs are nonlinear in emissions. This is a limitation of the methods that rely on MDs for projection over ranges that are large enough to induce a nonlinear response. One solution to avoid this shortcoming is to perform a stepwise optimization where changes in each polluter's emissions at each step are limited such that near-linear response is assured (Mesbah et al. 2012).

The nonlinearity in the calculation of MDs can be further investigated to establish source-specific MD curves. An MD curve represents how MD varies with

increasing emissions from a power plant. To establish such curves, multiple adjoint runs are required to update the value of the MD at different emission levels. Note that nonlinearity is different under the DMIN policy and the SCMIN policy. The emission behavior under the SCMIN policy is driven by both MDs and MACs. Since abatement costs are linear with emissions in our analysis, the overall nonlinearity driven by damage estimations is more heavily influenced by damage estimations and relative slope of MD curves for low- and high-MD power plants. Furthermore, the lack of a cost constraint in the DMIN policy allows for more substantial changes in emissions than under the SCMIN policy where both MDs and MACs control the optimal distribution of emissions. This, in turn, results in larger deviations from the baseline emissions and different nonlinear behavior.

This work should be treated as a proof-of-concept study as its practical reach is limited by a number of simplifying assumptions and limitations. Different types of uncertainties associated with air quality modeling, emission inventories, meteorological conditions, abatement cost estimations, valuation of statistical life, mortality response factors, etc, are not considered in this study. MACs are assumed to be constant for various emission levels and days, while in reality they can change on a daily basis depending on the electricity price and demand in the electricity market. Furthermore, MDs in this study are averaged over the ozone season resulting in overall positive values. However, daily MDs exhibit significant variability and are sometimes negative. MDs can also vary from year to year, and multi-year simulations are required to gain confidence in their robustness. Finally, our analysis is performed

for a fraction of power plants in the trading market. We believe emission differentiation becomes more effective as more units are included in the analysis because a wider range of differences in MACs and MDs can be exploited in determining the optimal distribution of emissions.

5.7. Conclusions

The calculated benefits in this study are based on the short-term mortality due to ozone exposure alone; PM exposure and long-term effects are not taken into account. The latter health impacts are believed to be substantial (Fann et al. 2012; Jerrett et al. 2009), and as such the MDs presented here are likely to be underestimates. Nevertheless, our results indicate that marginal damages in the form of mortality valuations significantly outweigh marginal abatement costs. Based on our results the average value of MD from the studied power plants was about $14700 per ton of NO_x in the ozone season, approximately 10 times larger than average MAC estimates. This disparity in damages and costs underlines the need for more refined policy instruments that formally include damage information. To compare estimated MDs with the electricity price, we use power plants' emissions and electricity generation in the ozone season and converted the average value of MD from $/ton to $/MWh. This simple conversion results in an average MD of $7.8/MWh. In comparison to the typical spot prices for the study region, this average MD constitutes 5 to 14 % of the retail price, that is, a sizable but modest fraction of the electricity price. This would only reinforce what we have demonstrated in this study, that is, that MDs can play an important role in making environmentally effective policy choices.

We note that the implementation of the SCMIN policy requires source-specific MD information, the calculation of which is affected by different uncertainties. These include uncertainty in epidemiological concentrations response factors, economic valuations, emissions, meteorology, model formulation, grid resolution, inter-annual and seasonal variability, etc.

Based on our results, the exchange rate policy, which allocates allowances based on historical electricity generation of units, leads to a lower improvement in environmental performance as compared to the two other damage-based policies considered. Our findings suggest that it is prudent to pay close attention to the location-specific MDs of power plants when devising pollution control strategies, as inclusion of damage information allows for development of more targeted policy instruments. Policies that rely on damage-based emissions differentiation can be implemented in different ways by creating higher incentives for high-MD polluters to reduce their emissions. One possible approach within a CaT system is to consider power plants' MDs when the initial allowances are allocated under an exchange rate CaT system. Imposing a tax based on a power plant's MD is another alternative for inclusion of damage information. The effects of initial allowance allocation on the performance of an exchange rate CaT system, regional health impacts, and the burden on the local or state utilities should be further explored in future work.

CHAPTER 6:

OPTIMAL OZONE CONTROL WITH INCLUSION OF SPATIOTEMPORAL MARGINAL DAMAGES AND ELECTRICITY DEMAND

6.1. Introduction

In the U.S., 37% of electricity is generated by coal-burning power plants (EIA 2010), placing them among the major contributors to the formation of ground level ozone. Ozone is a secondary pollutant, which is dependent on the amount of nitrogen oxides (NO_x), volatile organic compounds (VOC), and sunlight available. The ozone formation potential of these precursors may vary significantly by location and time. Due to these temporal and spatial differences, health impacts of NO_x emissions can vary up to 6 times in magnitude (Mauzerall et al. 2005) on a regional scale or up to 68 times on a national scale (Muller and Mendelsohn 2009). Inclusion of such differences in policy design can result in improved public health (Mauzerall et al. 2005; Mesbah et al. 2012; Muller and Mendelsohn 2009) simply by redistributing emissions without a need for emission reduction. Furthermore, re-dispatching or shifting electricity generation from high to low NO_x rate (i.e., emissions per unit generation) plants can reduce NO_x emissions significantly and result in a reduction in ozone concentrations while meeting the electricity demand (Martin et al. 2007; Sun et al. 2012). A Previous study suggests that dispatching based on local meteorological conditions and electricity demand may be an efficient method for control of surface

ozone, comparable to other emission control technologies such as installation of selective catalytic reduction (SCR) or selective non-catalytic reduction (SNCR) (Sun et al. 2012). Re-dispatching has also been considered for introducing electric vehicles into the fleet (Thompson et al. 2011), or for reducing water consumption of power plants at regions affected by drought (Alhajeri et al. 2011).

An effective NO_x emissions control policy aims to minimize the system-wide social cost which includes emission abatement costs for all polluters and the damage they impose on the environment and human health (external costs). The cost of damages caused by a polluter (hereafter damage) is calculated by the polluter's marginal damage (MD), defined as the dollar value of damage per ton of emissions (Muller and Mendelsohn 2007). Cap-and-trade programs are designed to achieve lower social costs by providing cost-saving incentives for participants while capping the total emissions. For uniformly mixed pollutants with long atmospheric lifespans (e.g., carbon), an optimal cap on total emissions can limit the system-wide damage and minimize the social costs. However, a cap on total emissions does not limit the system-wide damage for short-lived pollutants such as NO_x, whose MDs vary by location and time. By differentiating emissions by MDs through exchange rate cap-and-trade policies (Førsund and Nævdal 1998; Mesbah et al. 2012; Muller and Mendelsohn 2009; Nobel et al. 2001) or taxation (Montgomery 1972; Tietenberg 1995; Tong and Mauzerall 2006), the system-wide impact of NO_x emissions can be reduced.

Both exchange rate and taxation policies provide higher emission reduction incentives for high-MD polluters than low-MD emitters. Under an exchange rate policy, emission quotas are valued based on polluters' contributions to total damage, and under the taxation policy, polluters are charged fees based on their damage. A fee per ton of emissions provides emission reduction incentives for high-MD polluters because their marginal abatement costs (MACs), or costs per additional ton of emission reduction, is less than the imposed fee.

The accurate determination of the MD for a NO_x polluter then becomes crucial for determining the exchange rates or taxes required to minimize ozone concentration and corresponding damages. The MD of a polluter is the derivative of the system-wide damage function with respect to emissions from the polluter. There have been various efforts to calculate MDs with differing level of complexity. Some studies have used traditional sensitivity analysis methods which require running an air quality model for each additional source. This method is computationally expensive and has only been used for a limited number of sources (Nobel et al. 2001; Tong et al. 2006; Wang et al. 2007). Simplified or reduced-form models have been used to estimate MDs for a large number of sources (Fann et al. 2009; Levy et al. 2009; Muller and Mendelsohn 2007). However, these simplified approaches do not account for all different physical and chemical processes influencing the fate of pollutants in the atmosphere. Recent studies have used backward (adjoint) sensitivity method to calculate MDs for a large number of sources (Mesbah et al. 2013; Pappin and Hakami 2013). The adjoint method accounts for different physical and chemical process

included in the atmospheric model and computes MDs at comparatively low computational expense. The adjoint model is efficient for calculating the sensitivity of a desired function of outputs, such as ozone induced mortality, known as the adjoint cost function, with respect to individual inputs, such as emissions from sources at different times and locations (Pappin and Hakami 2013). Further details on the adjoint model and its mathematical formulations can be found elsewhere (Errico 1997; Hakami et al. 2007; Sandu et al. 2005; Wang et al. 2001).

In this work, the adjoint method is used to calculate source-specific MDs at various times. The calculated MDs are then used within an optimization framework to investigate how setting variable NO_x prices based on MDs would impact the re-dispatching strategy and the system-wide health damage. In comparison to previous studies on re-dispatching (Alhajeri et al. 2011; Martin et al. 2007; Sun et al. 2012; Thompson et al. 2011), this work does not model the electricity network but accounts for electricity demands. However, this study includes the spatiotemporal MD information to investigate the impact of re-dispatching strategies on ozone concentrations and the corresponding damage to human health. As such, MD information is used to set different spatial or temporal emission fees rather than setting one fee under different hypothetical scenarios used in previous studies.

6.2. Methodology

An adjoint model is combined with an optimization tool to include the MDs in the design of an improved NO_x emissions control policy. The adjoint model is used to calculate the MDs, which are then used as inputs into the optimization model to

predict the emission levels under different policies. The MDs in this work are defined based on a linearized form of the damage function which only includes the short-term ozone mortality in the U.S. (Anenberg et al. 2010).

$$MD = \frac{\partial TD}{\partial E} = V_{SL} M_0 P \beta \lambda \tag{6-1}$$

where $\frac{\partial TD}{\partial E}$ is the derivative of total damage (TD) with respect to emissions (E); V_{SL} is a value of statistical life; M_0 is the baseline non-accidental mortality rate; P is the population; β is the concentration response factor; and λ represents the adjoint gradient ($\frac{\partial C}{\partial E}$) which relates concentration (C) to emissions and is the main source of non-linearity in the damage function (Mesbah et al. 2013).

The optimization framework in this work is an extension of our previous study (Mesbah et al. 2013) with a few notable differences. Firstly, the MDs in this work are location- and time-specific, whereas the previous work did not consider temporal variability in MDs. Secondly, in the current optimization framework, the relationship between the power plants' electricity generation and emissions has been included as an additional constraint within the optimization framework. This constraint has been added to account for power plants' emission behavior in the electricity and emission markets where they have to supply electricity while meeting the emission requirements.

7.2.1. Polluters' behavior

Cap-and-trade systems are designed to minimize the system-wide emission reduction costs and this governs the polluters' behavior under such systems. The following optimization problem can be used to predict the distribution of polluters' emissions leading to the minimum of system-wide abatement costs:

$$Minimize:$$
$$\sum_{t=1}^{24}\sum_{i=1}^{n} c_i(e_{it}) \tag{6-2-a}$$

$$Subject\ to:$$
$$\sum_{t=1}^{24}\sum_{i=1}^{n} e_{it} \le E_T, \tag{6-2-b}$$

$$\sum_{t=1}^{24}\sum_{i=1}^{n} q_{it} = \sum_{t=1}^{24}\sum_{i=1}^{n} e_{it}R_i = Q_T, \tag{6-2-c}$$

$$q_{it} \in [0, G_i]. \tag{6-2-d}$$

where equation 6-2-a is the objective function of the optimization problem and is defined as the summation of the source-specific abatement cost functions (c_i) of n polluters; e_{it} is a variable representing the emissions from polluter i at hour t integrated over the ozone season; E_T is the system-wide cap on emissions; q_{it} is a variable representing electricity generation; R_i is a constant representing the generation intensity (i.e., the ratio of generations to emissions) for polluter i; Q_T is the total electricity demand in the ozone season, and G_i is the hourly generation capacity of polluter i. Equation 6-2-b ensures that the total emissions are less than the

total cap while equation 6-2-c ensures that the total electricity demand in the system is supplied by the power plants.

Equations 6-2-a through 6-2-d extend our previous application of this approach, which included no consideration of temporal effects (Mesbah et al. 2013). Our previous study also did not address electricity demand and therefore the emission constraint was set equal to the total cap (as opposed to less than or equal as in equation 6-2-b) to avoid a value of zero for the optimized emissions for all polluters. In this work; however, power plants must supply electricity based on demand (equation 2-b) and, therefore, their emission levels cannot be zero when minimizing the system-wide abatement costs.

Equations 6-2-a through 6-2-d are designed to find a distribution of emissions regardless of their spatial and temporal impacts on human health. To account for spatial and temporal effects of the emissions, a spatiotemporal social cost minimization problem can be defined as follows:

$$Minimize:$$
$$\sum_{t=1}^{24}\sum_{i=1}^{n} c_i(e_{it}) + D_{it}(e_{it}) \tag{6-3-a}$$

$$Subject\ to:$$
$$\sum_{t=1}^{24}\sum_{i=1}^{n} e_{it} \le E_T, \tag{6-3-b}$$

$$\sum_{t=1}^{24}\sum_{i=1}^{n} q_{it} = \sum_{t=1}^{24}\sum_{i=1}^{n} e_{it}R_i = Q_T, \tag{6-3-c}$$

$$q_{it} \in [0, G_i]. \tag{6-3-d}$$

where (6-3-a) represents the spatiotemporal social cost of emissions and accounts for

both abatement costs and health damage costs ($D_{it}(e_{it}) = MD_{it}e_{it}$) from individual polluters. Equation 6-3-c guarantees that power plants supply the total electricity demand, but it does not provide any restriction on hourly generation and therefore equations 6-3-a to 6-3-d are referred to as flexible social cost minimization, and allow power plants to change their hourly generation from one hour to another. This assumes flexibility in demand, meaning that consumers are willing to adjust the revised damage-driven electricity demand. Without such an assumption, the hourly electricity generation must meet the hourly electricity demand, which we refer to as demand-based social cost minimization. This optimization problem is similar to equations 6-3-a to 6-3-d, but 6-3-c is replaced with 6-3-e, shown below:

$$\sum_{i=1}^{n} e_{it} R_i = \sum_{i=1}^{n} q_{it} = Q_t \quad ,t = 1,..,24 \qquad \text{(6-3-e)}$$

Equation 6-3-e guarantees that the hourly system-wide electricity generation is equal to the electricity demand for hour t (Q_t).

7.2.2. Case study

A case study of 218 coal-fired electricity generating units in the eastern U.S. is conducted to examine the spatiotemporal effects of NO_x emissions under different policies. These policies include: 1) the cost minimization (CMIN) policy, similar to cap-and-trade, which only accounts for abatement costs (equations 6-2-a to 6-2-d); 2) a social cost minimization policy (equations 6-3-a to 6-3-d) under which power plants are flexible in hourly electricity generation (F-SCMIN); and 3) a social cost minimization policy based on equations 6-3-a, 6-3-b, 6-3-d, and 6-3-e, under which

power plants must supply hourly electricity demand (D-SCMIN). For each policy, a global optimization package (KNITRO 8.0) (Byrd et al. 2006) is used to solve the corresponding optimization problem and find the post-optimization emissions distribution.

The adjoint of the gas-phase Community Multiscale Air Quality (CMAQ) model version 4.5.1 is used to calculate the source-specific MDs (Byun and Schere 2006, Hakami et al. 2007). The CMAQ model is driven by meteorological inputs generated by the Weather Research and Forecasting (WRF) model (Skamarock et al. 2005) and emission inputs generated by the Sparse Matrix Operator Kernel Emissions (SMOKE) model (CEP 2007). The emission inventory used for the SMOKE model is based on the 2006 National Pollutant Release Inventory (NPRI) for Canada and the 2005 National Emissions Inventory (NEI) for the U.S., which are projected to the year 2007 based on population and economic growth factors. The power plant emissions are based on the Continuous Emission Monitoring (CEM) data for 2007. Modeling is conducted over a North American domain with a 36 km grid resolution with 34 vertical layers, for a period corresponding to the ozone season (May to September, inclusive) of 2007. The performance evaluation of ozone simulations results in an 18% mean normalized error and a 1% mean normalized bias. The hourly electricity generation and emissions are taken from U.S. EPA clean air market (EPA). The generation intensity for each power plant is calculated based on the ratio of the plant's total electricity generations to its total emissions during the ozone season. Cost estimation is for the short-term and is based on the U.S. EPA Integrated Planning

Model (EPA 2010a). The damage function in this study is for short-term ozone mortality and is defined based on a value of statistical life of 6.8 million (EPA 2010b), and an epidemiological response factor of 0.051% estimated for average 8h-ozone (Zanobetti and Schwartz 2008). The International Classification of Disease (ICD)-10 codes A-R (Bell et al. 2004) is used to calculate spatially variable baseline non-accidental mortality rates.

6.3. Result and discussion

The location of selected power plants and their average location-specific MDs are shown in Figure 6.1. The outputs of the adjoint model are location- and hour-specific MDs for the running period. For the five-month simulation period, each location has 153×24 hour-specific MDs. To calculate the average location- and hour-specific MD, the model outputs are integrated as follows:

$$MD_t(x,y) = \frac{1}{\sum_k \sum_d e_{kdt}(x,y)} \sum_k \sum_d (e_{kdt}(x,y) MD_{kdt}(x,y)) \tag{6-5}$$

where $MD_t(x,y)$ is the location- and hour-specific MD at the grid location (x,y) and hour t; $e_{kdt}(x,y)$ and $MD_{kdt}(x,y)$ are emissions and MD, respectively, at layer k, day d, hour t, and grid location (x,y). Note that $MD_t(x,y)$ is the MD at hour t averaged over 153 days.

To investigate the temporal and spatial effects separately, two average MDs are introduced. The first is the average time-specific MD ($\overline{MD_t}$), which is calculated by averaging MDs over N grids (N is the number of grids in color shown in Figure

6.1) for a particular hour (equation 6-6). The second is the average location-specific MD ($\overline{MD}(x,y)$), which is calculated by averaging hourly MDs over 24 hours for a specific location (equation 6-7).

$$\overline{MD}_t = \frac{\sum_{x,y} MD_t(x,y)}{N} \tag{6-6}$$

$$\overline{MD}(x,y) = \frac{\sum_t MD_t(x,y)}{24} \tag{6-7}$$

The average location-specific MDs (Figure 6.1) exhibit specific patterns. Power plants in the southeast region generally have high MDs because of the governing NO_x-limited atmospheric regime where ozone concentrations are highly and positively sensitive to NO_x reduction; power plants on the east coast also have high MDs because this region is densely populated; and power plants in the Ohio River Valley have low MDs due to an abundance of NO_x in this region.

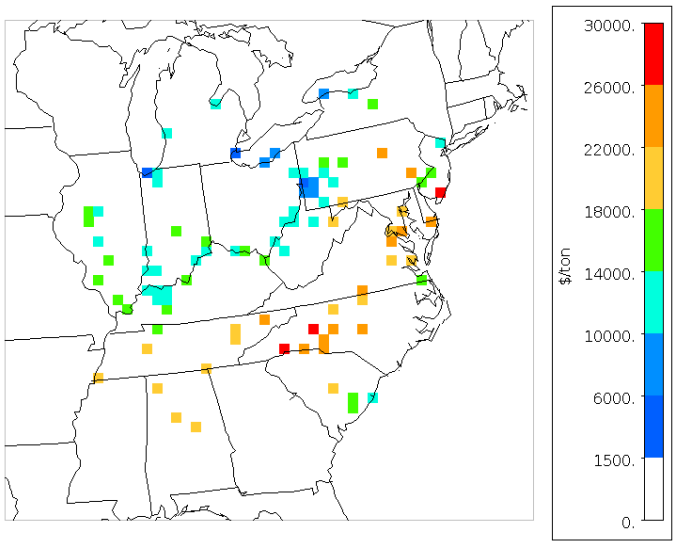

Figure 6.1. Average location-specific MDs for the selected power plants.

Location and hour-specific MDs are shown in a color-coded matrix in Figure 6.2. Each row of this matrix represents the MD at a specific location for different hours. While there is similarity in average location-specific MDs within some states, which supports the calculation of one state-wide MD used in some studies (Fann et al. 2009; Muller 2011), there is significant variability in average location-specific MDs for other states. For example, New Jersey contains three grid cells with average location-specific MDs of $11500/ton, $15100/ton, and $29700/ton.

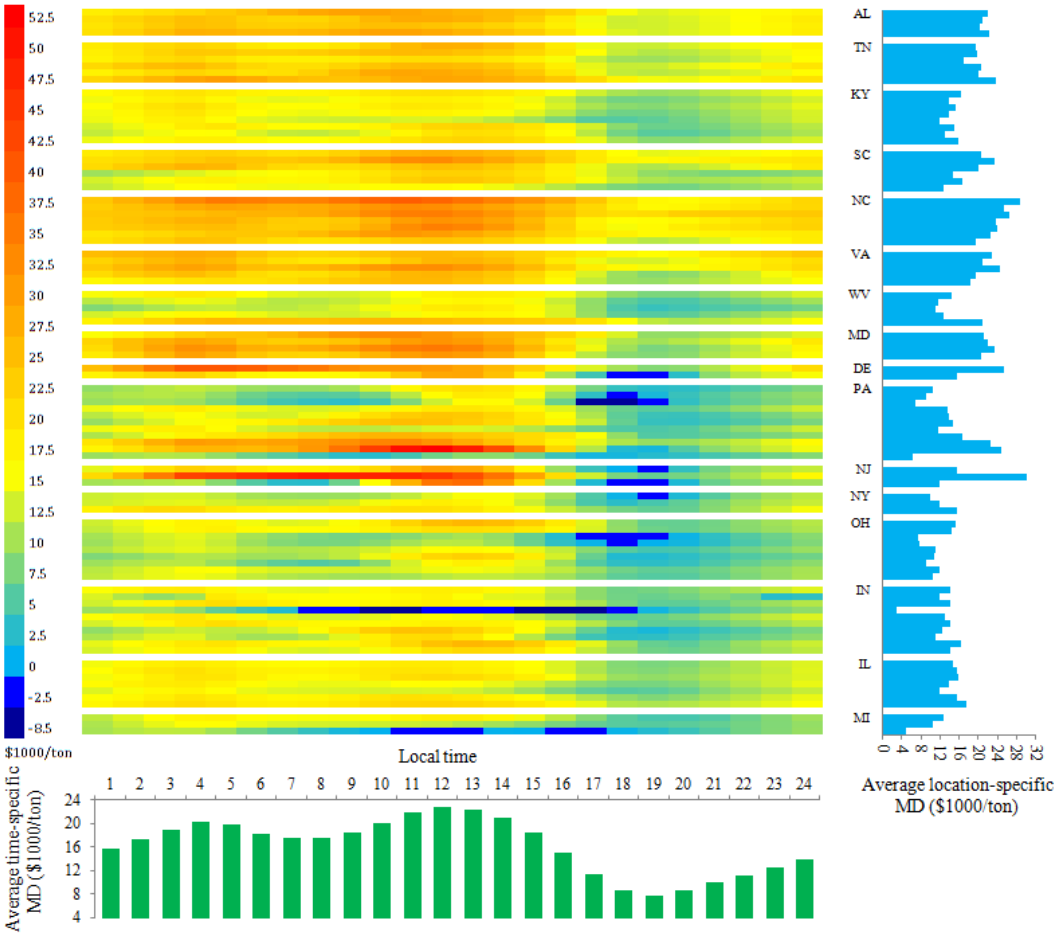

Figure 6.2. Color-coded location- and hour-specific MDs for the studied power plants. The dark blue cells of the matrix are locations and times with negative MDs. The graph on the right panel is the average of each row and represents the average location-specific MDs ($\overline{MD}(x,y)$). The graph at the bottom is the average of each column and represents the average time-specific MDs ($\overline{MD}_t$).

The lowest value of the time and location specific MD shown in the color coded matrix in Figure 6.2 is -$8500/ton and belongs to a grid cell in Clarion, Pennsylvania at hour 18. The highest value, ($52700/ton) occurs for Phoenixville, Pennsylvania at hour 12. This grid cell also contains the highest ratio of maximum MD (hour 12) to minimum MD (hour 19). This discrepancy in hourly MDs amounts

to an avoided mortality per season for transfer of 132 tons of NO_x emissions from hour 12 to hour 19 during the ozone season.

The general pattern of change for hourly MDs is also consistent for different locations. The highest average time-specific MD is $22300/ton for hour 12, and the lowest average time-specific MD is $7400/ton at hour 19. Overall, hours 11 to 14 are high-MD hours (with MDs greater than $20000/ton), and hours 18 to 21 appear as low-MD hours (with MDs less than $10000/ton). Hourly MDs increase from morning to noon when the highest MDs occur in most locations, and reach a minimum in the evening. This pattern for hourly MDs exists because of the dependency of photochemical ozone formation reactions on the intensity of sunlight and temperature. Different hourly MD-based fees on power plants' generation affect the generation pattern. However, hourly generation is also a function of electricity demand, which is usually higher in the late afternoon and evening when people are at home and engage in cooking and other energy intensive activities. High demand for electricity in the evening is beneficial for damage minimization because it shifts generation and the corresponding emissions to evening hours when emissions are less harmful. Shifting emissions to low-MD hours lowers health damages and thus favors damage minimization. In contrast, seasonal electricity demand fluctuations are not in favor of damage minimization because demand is high on hot summer days (usually late summer) when MD is also high. It should be noted that the current cap-and-trade program in the eastern U.S. has no limitations on hourly or daily emissions and only caps the total NO_x emissions during the ozone season when NO_x are more harmful.

Therefore, higher demand on late summer days results in disproportional and more use of emission quotas on days when emissions are more harmful resulting in an increase in health damages (Martin 2008). One approach to alleviate this problem would be to allocate total emission allowances separately for peak and off-peak ozone season.

It should be mentioned that there is no negative average time-specific or average location-specific MD for the power plants studied. This finding is consistent with the previous finding (Fann et al. 2009) that some urban polluters had negative NO_x MDs, but electricity generating power plants did not. Note that the hour and location specific MDs reported are averaged over 153 days and the daily values have more variability and occasionally include negative values. In fact, some of the studied electricity generating units, with positive average location-specific MDs, have negative MDs for some hours (see the MD matrix in Figure 6.2). The negative hourly values are observed in ten locations. For the majority of these locations, the negative MDs occur in hours 17 to 19. Exceptions are power plants in Indiana and Michigan, which have negative values for some hours in the morning and around noon. A negative MD indicates a NO_x-inhibited atmospheric regime along the emission trajectory where the ratio of ambient NO_x to VOCs is high. Note that NO_x availability in a particular location is not necessarily limited to NO_x emissions in the same location and can include NO_x emissions in upwind locations.

The three sets of gradients shown in Figure 6.2 can be used to evaluate the temporal and spatial effects of NO_x emissions redistribution. As explained before,

social costs can be minimized using the location-specific and time-specific MDs, both with and without limitations on hourly electricity generation (D-SCMIN and F-SCMIN, respectively). To consider only spatial effects, the two social cost minimization problems are conducted with one average location-specific MD for all 24 hours. Since these optimizations only account for spatial differences in MDs, they are referred to as the SPT-D-SCMIN and SPT-F-SCMIN policies. Similarly, to account for only temporal effects of NO_x emissions, social cost minimizations are carried out with the same average time-specific MD in all locations regardless of where the power plants are located. These social cost minimizations account only for temporal effects and are referred to as the TMP-D-SCMIN and TMP-F-SCMIN policies. Table 6.1 summarizes different aspects of the policies examined in this study.

Table 6.1. Differences under considered cost minimization policies.

	CMIN	SPT-D SCMIN	TMP-D SCMIN	D-SCMIN	F-SCMIN	SPT-F SCMIN	TMP-F-SCMIN
Temporal specificity	No	No	Yes	Yes	Yes	No	Yes
Spatial specificity	No	Yes	No	Yes	Yes	Yes	No
Hourly electricity demand	Yes	Yes	Yes	Yes	No	No	No

Social cost minimization, results in different total damages and costs under each policy (Table 6.2). All policies are then compared with the CMIN policy which minimizes the abatement costs (not social costs) for the short term.

Table 6.2. Abatement costs and health benefits (million dollar) for different policies compared to the CMIN policy.[1]

	SPT-D SCMIN	TMP-D SCMIN	D-SCMIN	F-SCMIN	SPT-F SCMIN	TMP-F-SCMIN
Total emissions (1000 ton)	108.2	100.8	108.2	109.9	108.2	102.3
Increase in abatement cost ($ million)	93.9	88.8	93.5	94.2	88.8	88.5
Decrease in seasonal damage ($ million)	1263	1154	1285	1315	1248	1202
Net benefit ($ million)	1169.1	1065.2	1191.5	1220.8	1159.2	1113.5

[1]Under the CMIN policy, total abatement cost is $1515 million and total emissions are 181 thousand tons

In all cases, total emissions are lower than the total cap on emissions (i.e., 204 thousand ton). This implies that the actual emission cap was not stringent enough as additional net gains could be realized by reducing aggregate emissions. Hourly electricity generation under demand-based optimizations (SPT/TMP-D-SCMIN, D-SCMIN) and total generation under flexible demand optimizations (SPT/TMP-F-SCMIN, F-SCIM) is, however, a binding constraint. Note that the total electricity generation under the CMIN policy and the SCMIN policies are the same but generation and corresponding emissions are transferred from low-efficiency to high-efficiency power plants. A high-efficiency power plant has a higher generation

intensity and thus generates electricity with lower emissions. The main reason for differences in total emissions under the CMIN and SCMIN policies is the imposed additional external damage on power plants under the SCMIN policies. Such extra costs will provide high-efficiency power plants in low-MD locations an advantage over low-efficiency units in high-MD locations. In other words, power plants with lower damage per unit of electricity generation ($/GWh) are among those that generate more electricity when the social cost is minimized. Shifting emissions from high to low emission intensity power plants results in a significant emissions reduction and corresponding health benefits under the SCMIN policies (Table 6.2).

One important finding from the comparison of different social cost minimizations is that the net benefit of policies that account for averaged spatial or temporal effects are comparable with those that account for spatiotemporal effects combined. Although using average MDs under the SPT/TMP-D/F-SCMIN policies results in a lower net benefit compared to the D/F-SCMIN policies, the additional benefits from inclusion of spatiotemporal damages are relatively small (i.e., 2 to 11% additional benefits) (Table 6.2). This is because under social cost minimization, emission reductions are beneficial as long as the MAC is less than the additional benefit of emission reduction (MD). The MDs used under the SPT/TMP-D/F-SCMIN policies are average location-specific and time-specific MDs. Based on our previous estimations, the average MAC is approximately 10 times smaller than the average location specific MDs (Mesbah et al. 2013). Because these MDs are still higher than the MACs, social cost minimization under these policies results in more or less

similar emissions reduction and net benefits when compared to the CMIN policy. It should be noted that the costs of emissions reduction for different hours are assumed to be the same, and therefore the differences in hourly abatement costs are not reflected in the results. This is one of the limitations of this study because the opportunity cost of emissions reductions by reducing electricity level changes with fluctuation in electricity prices. The difference in outcomes for different SCMIN policies is also dependent on the variation in MDs under different policies. The ratio of the 95 percentile to the 5 percentile values are 3.5 for the average location-specific MDs and 2.7 for the average time-specific MDs, while this ratio is 6.7 for the spatiotemporal MDs. The variability in MDs is important because it provides more flexibility for transfer of emissions from high-MD to low-MD locations or times, resulting in more substantial reduction in total damage. Results shown in Table 6.2 also suggest that although the assumption of flexible demand increases the net benefit, the increase is not significant. The TMP-D-SCMIN and TMP-F-SCMIN policies encourage a shifting of emissions from high-MD hours to low-MD hours, and differ only in the constraint on hourly electricity generation. This difference changes the distribution of hourly emissions (Figure 6.3c) from a shape similar to that of hourly electricity generation under base case (Figure 6.3a) to a shape with peaks at high and low MD hours (Figure 6.3b). The resulting difference in net benefits ($48 million) is not significant because low-MD hours, which emissions are shifted to, are also high-demand hours (Figure 6.3a). The difference in total emissions between the two policies at hour 19 is about 2000 ton and at hour 13 is about 1000 ton (Figures 6.3b and c). Note that for the most part, the net benefit under both the TMP-D-SCMIN and

TMP-F-SCMIN policies is a result of a 44% reduction in emissions as compared to the CMIN policy (an average of 3000 ton per hour).

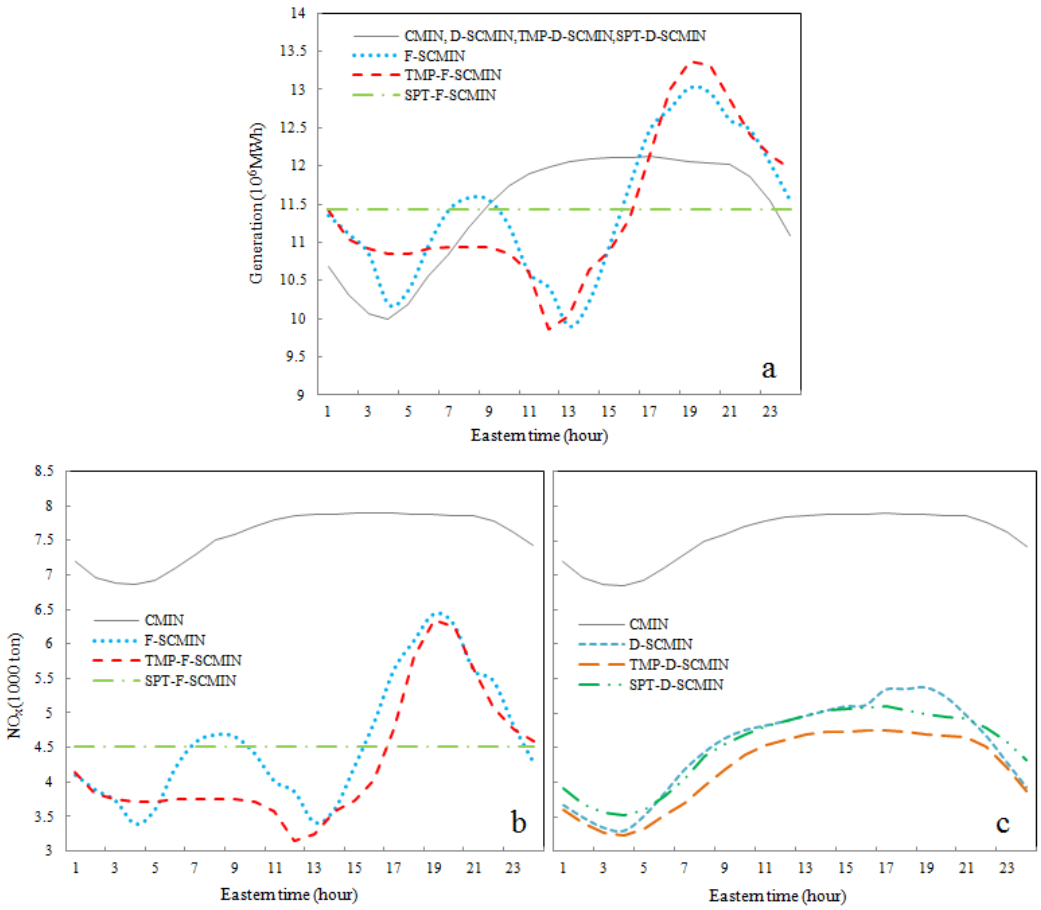

Figure 6.3. Hourly electricity generation (a) and hourly NO_x emissions from the power plants studied during the ozone season under flexible demand policies (b) and demand-based policies (c).

6.4. Conclusions

The cost-benefit analysis and sensitivity models used have restrictions and limitations, such as being subject to uncertainty in the air quality model, abatement and damage cost estimations (Mesbah et al. 2012, 2013) However, our study provides further evidence in support of the inclusion of time-based damage information in

economic instruments for ozone control. Another limitation of the current work is that the optimization framework presented does not account for the electricity generation dynamics and network transmission constraints, which limits the extent to which re-dispatching is possible. However, this constraint is not believed to have a significant impact on the NO_x redistribution patterns (Martin 2008). A previous study on the subject suggests that the insignificant influence from transmission constraints is possibly due to spatial heterogeneity caused by low and high generation intensity units within a single electricity generating zone, allowing for re-dispatching within the zone without a significant increase in the use of transmission lines (Martin 2008).

The SCMIN policies considered in this study result in a decrease in total damage by 44% to 50%, comparable reductions in total emissions, and a 4% increase in abatement costs as compared to the baseline CMIN policy. For all of the policies considered, the total electricity generation is the same, and the only difference is that power plants supply different amounts of electricity under each policy. The benefits obtained under the SCMIN policies are mainly due to the redistribution of electricity generation so that a higher portion of the generation is assigned to high-efficiency power plants with higher generation intensities located in low-damage regions. A key policy insight is that more power generation during peak hours may be beneficial even if by units with lower specific generation intensities are used.

Flexibility in demand can improve the performance of the system, but with less significance than temporal or spatial effects. The effect of flexibility in demand increases the net benefit by 5% under a TMP-F-SCMIN policy as compared to a non-

flexible temporal SCMIN policy (TMP-D-SCMIN). Flexibility in demand can be created to some extent by setting higher costs for electricity in high-MD hours so that consumers use less electricity during those times. The impact is not very significant since the demand for electricity is not price-elastic, i.e., the demand does not decrease by the same rate as the price increases. In the U.S., the range of price elasticity for residential electricity demand is between −0.20 and −0.35 in the short-term and between −0.3 to −0.8 in the long-term, (Alberini and Filippini 2011) meaning that for a 1 % decrease in electricity consumption for a particular hour, the price needs to be increased by 3 to 5%.

Consideration of temporal effects under the TMP-D/F-SCMIN policies is almost as effective as consideration of spatial effects under the SPT-D/F-SCMIN policies or both effects together under the D/F-SCMIN policies. For example, the increase in net benefit by switching from a TMP-D-SCMIN policy to a SPT-D-SCMIN policy is 10% and to a D-SCMIN policy is 12%. Policymakers need to address if the additional net benefit provides adequate motivations to switch to spatial policies considering the complexity added to the system under spatial or spatiotemporal policies. Under temporal policies, participants and decision makers are only exposed to 24 unique MDs as compared to many spatial MDs under spatial or spatiotemporal policies. To moderately increase system performance, inclusion of regional hourly MDs can also be considered. On an operational basis, emission redistribution can be achieved through electricity rate adjustments and shifting to different energy sources based on the time of day. Besides, spatial differentiation of

emissions from power plants in different regions can lead to equity considerations and subsequent legal complications and political conflicts between states. Temporal differentiation of emissions does not cause the same level of complexity and interregional inequities. In addition, the temporal MD regulations can be tailored to the specific consumption patterns of a region or state. This advantage can ease the path towards the implementation of damage-based economic instruments and more effective internalization of external costs of health damage.

CHAPTER 7:

NON-CONVEXITY IN OZONE-BASED NO_X HEALTH DAMAGE: AN APPLICATION OF THE ADJOINT OF CMAQ

7.1. Introduction

Estimation of external costs, or damage costs caused by emissions, are critical in the environmental decision making process as they can be used to set the optimal emission reduction targets. Damage functions represent the total monetary value of health and environmental damages caused by a pollutant change with emissions or with ambient concentrations. The derivative of the damage function is often referred to as the marginal damage (MD) of the emission (or marginal benefit due to its reduction) and is defined as the damage (or benefit) associated with a ton of emission. Damage functions are usually considered as convex in environmental economics, meaning that the monetary value of the damage increases with an increasing (or non-decreasing) rate when emissions increase. This convexity amounts to decreasing MD values with decreasing emissions or positively sloped MD curves. This behavior is due to the fact that the assimilative capacity of the environment deteriorates as the emissions increase. However, this work examines this widely accepted assumption and explores if it is correct for secondary pollutants such as ozone.

Various forms of MD curves are considered in environmental economics literature: positively sloped MD curves, flat MD curves (e.g., damages caused by carbon dioxide), MD curves with a threshold (e.g., toxic pollutants with a health-

based threshold), and negatively sloped MD curves. Generally, the negatively sloped MD curves are associated with aesthetic valuation of pollution (Anderson and Francois 1997; Randall et al. 1974). For example, an additional unit of pollution lowers the price of houses in a clean environment at a higher rate than houses in a polluted environment (Cropper and Arriaga-Salinas 1980). Similarly, first units of pollution disturb the beauty of a tourist attraction (and lowers their willingness to pay) more than subsequent units (Buhyoff and Leuschner 1978; Schultze and Brookshire 1983). Another example of a recognized non-convex damage function is that of acid rain in lakes. Initially, the MD increases with a positive rate because the acidification causes damages the lake ecosystem, but once the habitats has been destroyed, an increase in the acidity of the lake does not lead to any additional damage and thus the slope of MD curve becomes zero (Crocker and Forster 1981).

The ozone damage function has been implicitly recognized as being non-convex for some time (Goodstein 1995; Pappin and Hakami 2013; Starrett 1972). A particular case of non-convexity of ozone MD curves with respect to precursor emissions is widely known in the literature. The response of ground-level ozone to nitrogen oxides (NO_x) emissions depends on the atmospheric regime. In a NO_x-limited regime, where the ratio of NO_x-to-VOCs is low, a decrease in NO_x emissions slows down ozone production and thus leads to reduced ozone concentrations, i.e., a positive MD value. Conversely, under a NO_x-inhibited regime, ozone concentrations increase with reduction in NO_x availability, indicating a negative MD. The transition

from negative to positive NO_x MDs through reduced emissions is an indication of non-convexity in the ozone-based NO_x damage function (Fraas and Lutter 2011).

While negative MD values for NO_x emissions due to atmospheric chemistry have been recognized (Hall and Hall 1997; Repetto 1987) and calculated (Hakami et al. 2004), the shape of the NO_x MD damage curve has only been discussed briefly and by a few sources (Goodstein 1995; Perman et al. 2011). However, designation of a general shape for NO_x MD curves by these references is often speculative, qualitative, and incompatible with the chemical behavior of NO_x in producing ozone. Quantitative estimation of NO_x MD values requires use of atmospheric models that account for chemical transformations that lead to production of ozone from NO_x emissions. However, such calculation by traditional methods is computationally expensive and is only feasible for a limited number of sources (Mauzerall et al. 2005; Tong et al. 2006). Reduced-form (response surface) or simplified (dispersion) models have been used for NO_x MD estimations (Fann et al. 2009; Levy et al. 2009; Muller 2011). However, these approaches are not able to account for the nonlinearities involved in calculation of source-receptor relationships; a simplification that can lead to sizeable changes in estimated health damages (Mesbah et al. 2013).

An alternative for calculation of MDs is the use of the adjoint of an air quality model. Adjoint (or backward) sensitivity analysis allows for simultaneous calculations of derivatives with respect to a large number of parameters (Hakami et al. 2006; Sandu et al. 2005), making it an ideal tool for the estimation of source-specific MDs (Mesbah et al. 2013; Pappin and Hakami 2013). In this work, the adjoint of a state of

the art air quality model is used to estimate health-based NO_x MD values. MD calculations by the adjoint model are conducted at various domain-wide emission levels to construct MD curves for several urban and rural source locations in the U.S.

7.2. Methodology

We consider the damage function to be the monetary value of premature mortality in the U.S. due to short-term exposure to ozone exposure, and as such we do not account for environmental or non-fatal health damages. This simplification does not cause any loss of generality in our approach, as the shape of NO_x damage curve is dictated by the ozone-NO_x relationship rather than the damage end point. Our damage estimation is based on the following equation (Anenberg et al. 2010):

$$\Delta D = V_{SL} M_0 P(1 - e^{-\beta \Delta C}) \approx V_{SL} M_0 P \beta \Delta C \qquad (7\text{-}1)$$

where ΔD is the change in the damage function; V_{SL} is the value of a statistical life taken as \$6.8 million in 2007 (EPA 2010b); M_0 is the spatially variable baseline non-accidental mortality rate which is estimated based on the International Classification of Disease (ICD)-10 codes A-R (Bell et al. 2004); P is the population; β is a concentration response factor that correlates air pollution mortality to non-accidental death; and ΔC is the change in ozone concentration. A β of 0.051% was taken from an epidemiological study of 48 U.S. cities based on an 8-hour ozone averaging time in summer (Zanobetti and Schwartz 2008).

The adjoint model calculates derivatives of a scalar metric with respect to emissions at various locations. This scalar metric is referred to as the adjoint cost

function and for this study is considered to be nation-wide damage (D) in equation 7-1. The adjoint cost function is introduced to the ajdoint model via the adjoint forcing term or the derivative of the adjoint cost (damage) function with respect to concentration. The forcing term (φ) for the damage function is (Pappin and Hakami 2013):

$$\varphi = \frac{\partial D}{\partial C} = V_{SL} M_0 \, P \, \beta \tag{7-2}$$

The role of the forcing term in the adjoint model is similar to the role of emissions in the forward air quality models. Emissions are injected into the forward model and propagate forward in time to impact concentrations. Similarly, forcing terms are injected into the adjoint field and propagate backward in time to "force" influences on the damage function. A single adjoint run can calculate source-specific MDs in a large domain, but does not provide any information about how MDs change with emissions. To construct source-specific MD curves, multiple adjoint runs are required to calculate MDs at different emission reduction levels. The steps required to construct source-specific MD curves are as follows:

1. A meteorological model (WRF 3.1 (Skamarock et al. 2005)) is used to prepare meteorological input files, which are converted to the air quality model input format by the MCIP model.
2. An emission processing model (SMOKE 2.4(CEP 2007)) is utilized to create emission files from all sectors (e.g., point, mobile, biogenic, wildfires, area, non-road).

3. An air quality model (CMAQ 4.5.1 (Byun and Schere 2006)) is run with the meteorological and emission files. The forward CMAQ creates spatial and temporal concentration files (known as checkpoints) that are required for backward simulations. This requirement is due to the fact that the derivatives of nonlinear processes (e.g., chemical process) are functions of concentrations.
4. Spatial and temporal forcing files are created using equation 7-2 and checkpoint files.
5. The adjoint of gas-phase CMAQ (Hakami et al. 2007) is run using checkpoint and forcing files. The adjoint model creates temporal and spatial MDs, which indicate damages caused by an additional ton of NO_x for a particular day and location.

To calculate MDs for different emission levels, the emissions are reduced uniformly. Then, steps 3 to 5 are repeated to create new sets of MDs corresponding to the reduced emissions. In addition to MDs for baseline emissions, we generate MDs for several additional scenarios. In these scenarios, emissions from all point or mobile sources, or combined point and mobile sources are reduced uniformly by 20%, 40%, 60%, 80%, and 100%. The difference between these scenarios is that the CMAQ model uses different emission files in step 3, which results in different concentrations checkpoints and forcing files used by the adjoint model. Differences in concentrations reflect the change in the atmospheric regime that is induced by various NO_x emission reduction levels. In our simulations, the adjoint model is configured to create daily

MDs over the domain. To calculate the source-specific seasonal MDs, an emission weighted aggregation of daily MDs is used:

$$MD(x,y) = \frac{1}{\sum_d \sum_k e_{dk}(x,y)} \sum_d \sum_k e_{dk}(x,y) MD_{dk}(x,y) \tag{7-3}$$

where $MD(x,y)$ is the seasonal MD at grid (x,y); $e_{dk}(x,y)$ and $MD_{dk}(x,y)$ are the emissions and MD at grid (x,y), layer k, and day d. Note that for averaging daily MDs, the emissions in equation 7-3 can be either point source (PNT-MD) or mobile source (MB-MD) emissions. The daily MDs are calculated over a North American domain with 36 km grid resolution and 34 vertical layers. The simulation period is from May 1st to September 30th (the ozone season) of 2007.

7.3. Result and discussion

The outputs of the adjoint model, i.e., MB-MDs and PNT-MDs, are calculated under different emission reduction scenarios and presented in this section. Note that PNT-MDs are weighted averages for all 34 vertical layers (Figure 7.1), while the MB-MDs represent the damages caused by ground-level emissions (Figure 7.2). A value of zero for the PNT-MDs in Figure 7.1 indicates that no point source emissions were present in a location. If a new point source is established in those locations, an average daily MD (not weighted by emissions) can be used to estimate its potential MD. These zero-MD locations do not exist for MB-MDs as all locations (grids) in the domain have non-zero mobile NO_x emissions.

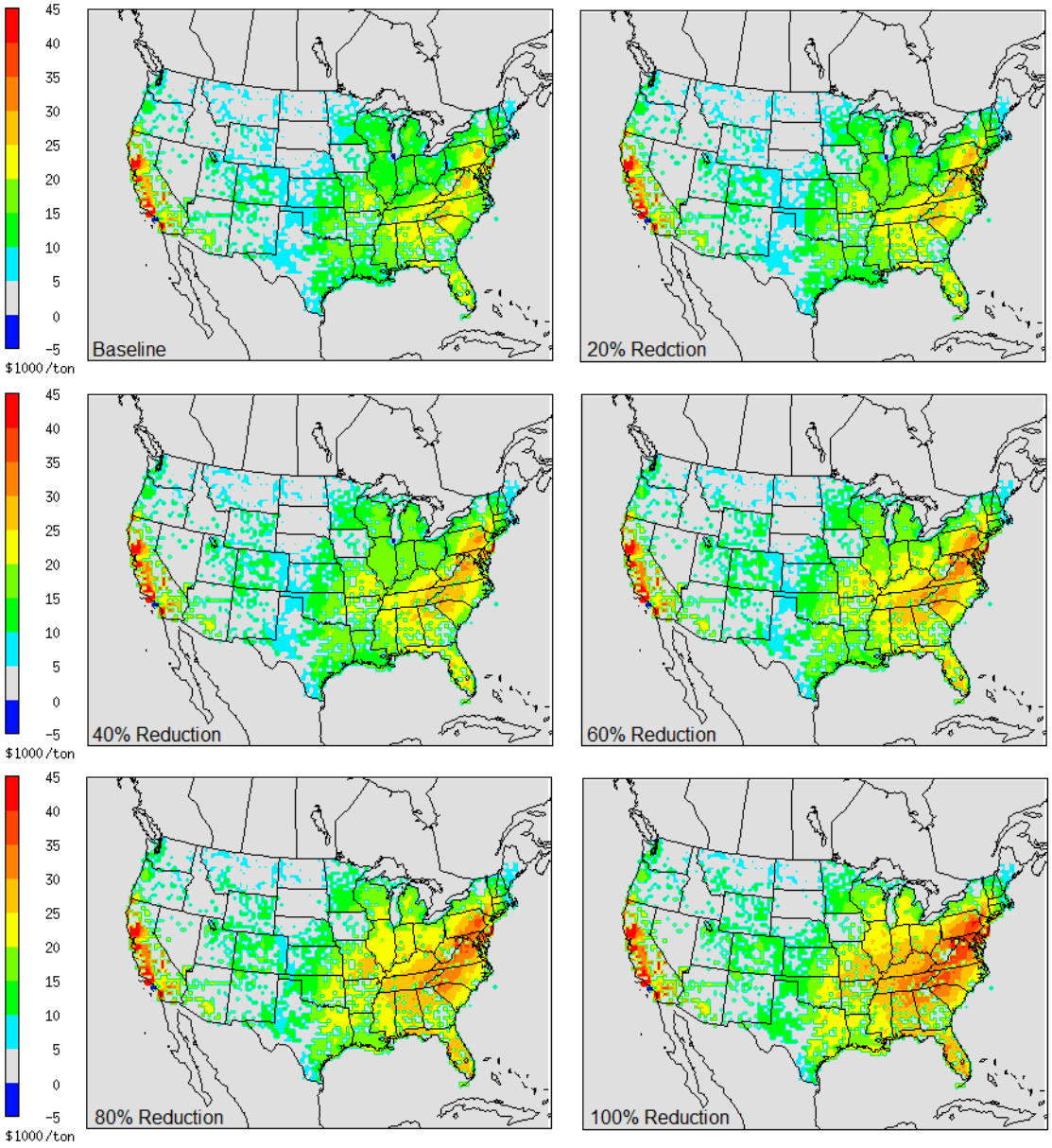

Figure 7.1. PNT-MDs for different emission reduction scenarios when point source emissions are reduced uniformly across the domain. Each panel shows PNT-MDs as average daily NO_x MDs weighted by point-source emissions in various model layers.

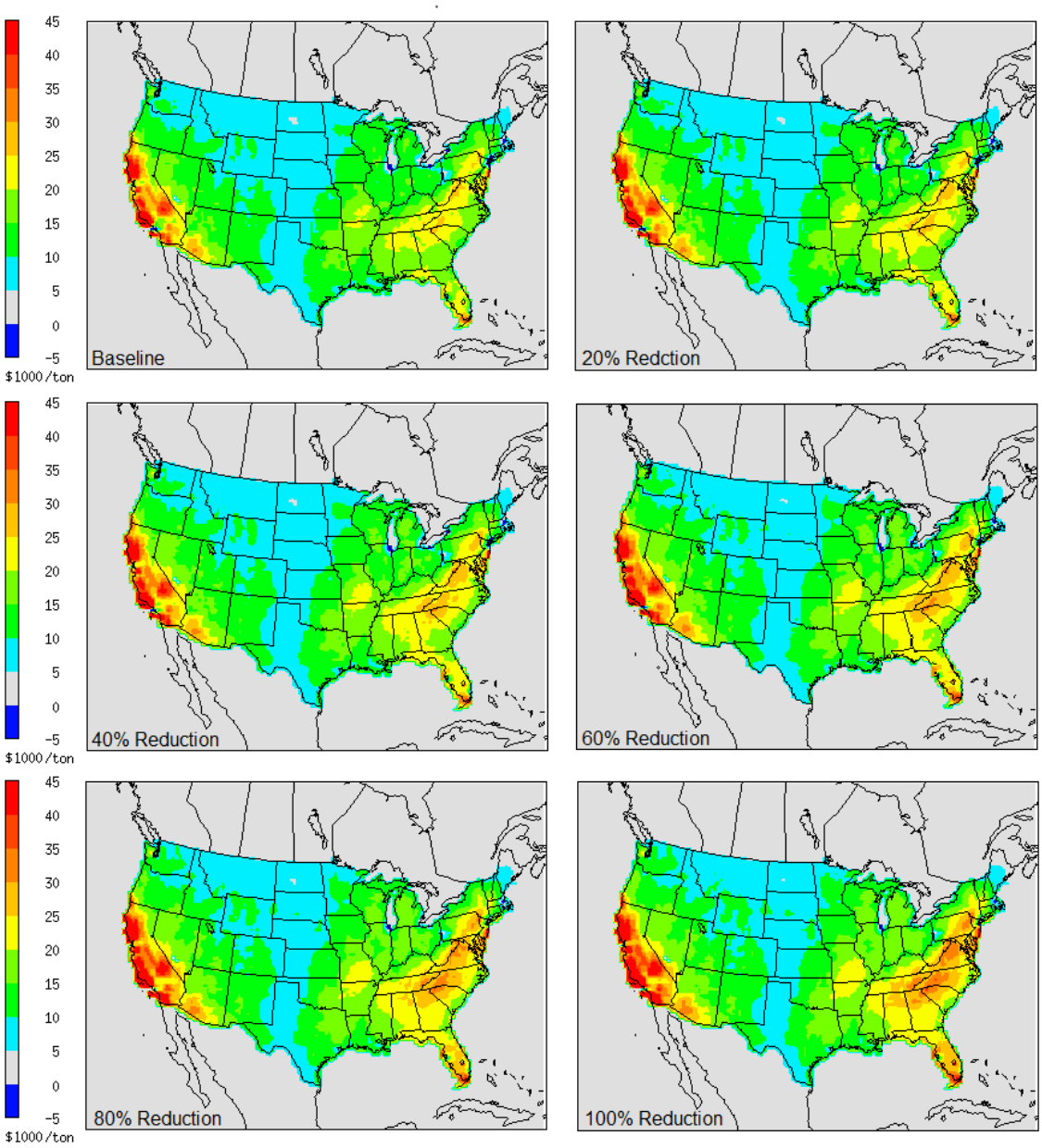

Figure 7.2. MB-MDs when mobile emissions are reduced uniformly across the domain. MB-MDs are average daily NO_x MDs weighted by mobile-source emissions.

For both PNT-MDs and MB-MDs, high-MD locations (shown in yellow, orange, and red in Figures 7.1 and 7.2) grow with emission reductions. This important finding indicates that the same unit of emission reduction carries larger benefits at

lower emission levels (i.e., MDs are higher). Note that positive MDs become more positive and negative MDs become less negative with emission reductions. This is due to the nature of atmospheric chemical reactions contributing to ozone production. The availability of NO_x is a determining factor in these reactions. When NO_x is abundant, there is great competition for additional NO_x molecules, and therefore, the impact of increased NO_x availability on the ozone concentration is small or even negative (in NO_x-inhibited regimes). When little NO_x is available, additional NO_x molecules face little competition in producing ozone, leading to high ozone formation efficiency and larger MD.

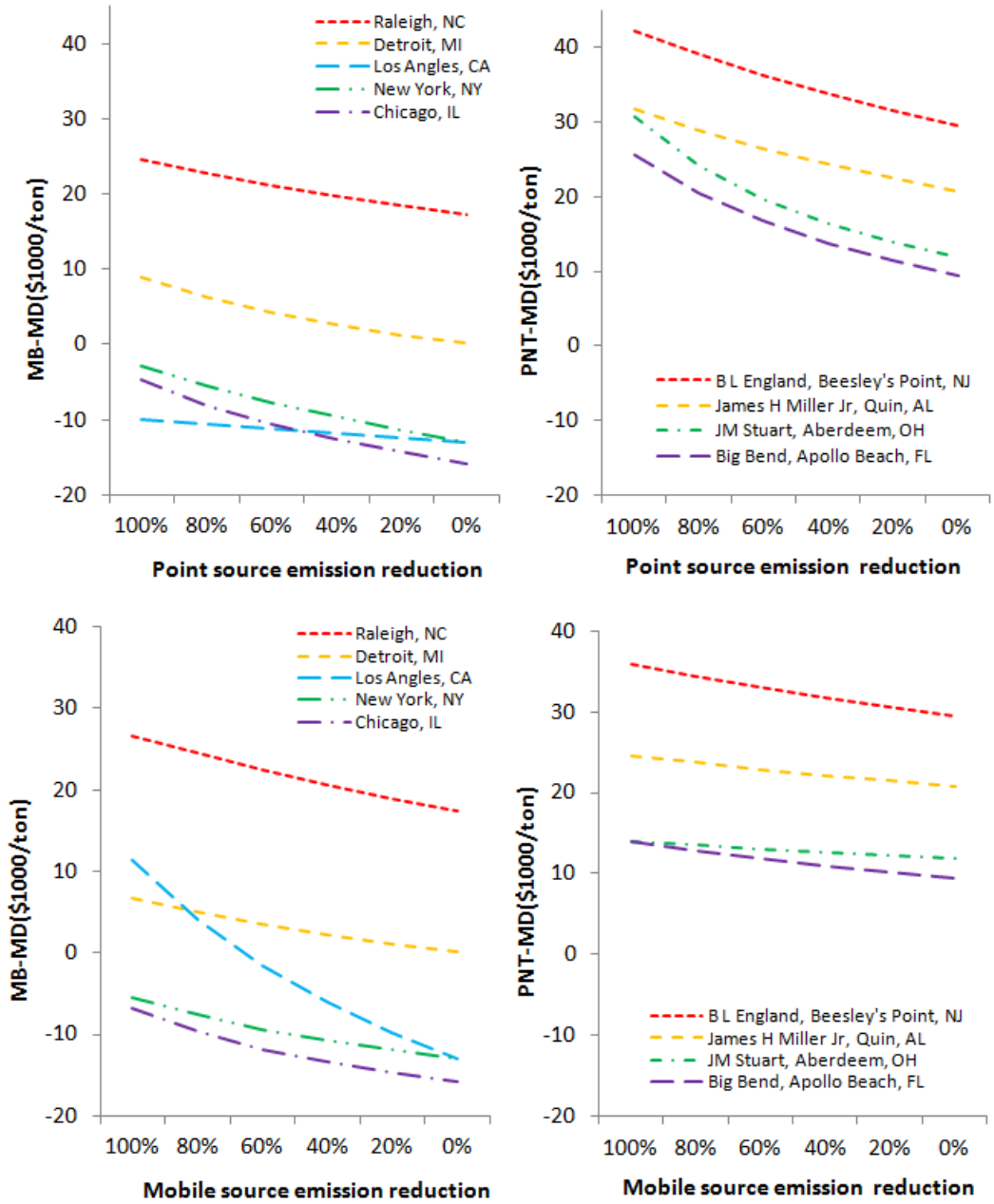

Figure 7.3. MDs at different locations when point (top) or mobile source emissions (bottom) are reduced. MD values are MB-MDs for few major cities (left) and PNT-MDs for selected power plants (right).

The change in MDs with emissions can be better depicted if source-specific MD curves are extracted and presented for power plants and cities (Figures 7.3).

Three of four selected power plants in Figure 7.3, are large power plants and were among the top facilities with highest ozone season emissions in 2007. These three power plants are: the James H Miller Power Plant in Alabama with a capacity of 2640 MW, the JM Stuart Power Plant in the Ohio River Valley with a capacity of 2340 MW, and the Big Bend Power Plant in Florida with a capacity of 1823 MW. The fourth power plant, the B L England Power Plant, with a capacity of 450 MW was selected because it is located in a high-MD region of New Jersey. As shown in Figure 7.3, all four power plants have negatively sloped MD curves.

Large urban regions such as New York City, Los Angles, and Chicago have negative MDs at baseline emissions (i.e., no emission reductions), whereas smaller cities have positive MDs. This is because of the governing NO_x-inhibited atmospheric regime under which NO_x emission reductions have negative influence on ozone formation (Pappin and Hakami 2013).

For both mobile and point source emission reductions, MDs are very similar for all five cities expect for Los Angles where MDs increase drastically with mobile source emission reductions, but are not sensitive to point source emission reductions. Elimination of mobile source emissions in Los Angeles changes the MDs from -$13000/ton to $11500/ton. This significant change is due to the large amount of mobile source emissions in the western U.S. For Los Angeles, change in the MD sign occurs for a mobile source emission reductions of about 70%. A change in the MD sign from negative to positive indicates that emission reductions increase ozone concentrations initially (negative MD) but decrease ozone concentrations eventually

because the atmospheric regime turns from NO_x-inhibited to NO_x-limited and MDs become positive. Note that the adjoint method is not able to predict the locations at which these transitions in the chemical regimes occur. Instead, the adjoint results indicate that emissions generated in that location (Los Angeles) experience an overall change of behavior somewhere along their trajectory in the domain as reflected in the estimated damage.

The change in sign of the MD does not occur for either New York or Chicago when emissions are reduced from point or mobile sources separately. However, the MDs became less negative in both cases. Removing both sources simultaneously results in a greater change in MDs and a change the sign of the MDs (Figure 7.4). Note that MD curves for point and mobile source emission reductions are constructed through different adjoint runs when emissions are simultaneously and uniformly reduced for both sources.

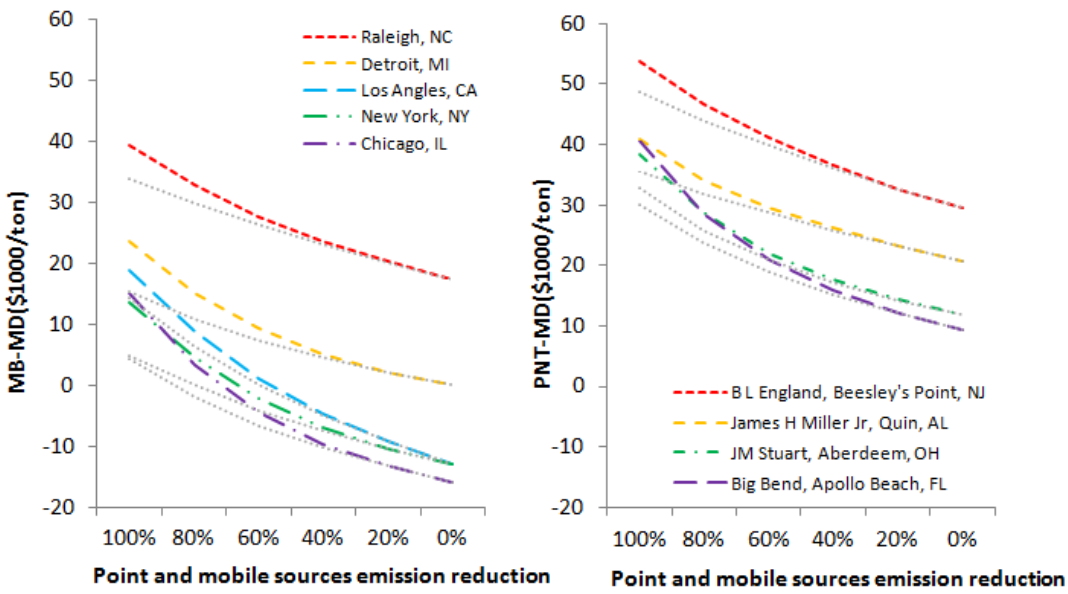

Figure 7.4. MDs at different locations when point and mobile source emissions are reduced. The MD values are MB-MDs for a few major cities (left) and PNT-MDs for selected power plants (right). The grey dashed lines are the aggravation of MDs from point source and mobile source emission reductions presented in Figure 7.3.

Note that the aggregation of MDs when emissions from point sources are reduced and the MDs when emissions from mobile sources are reduced (grey lines in Figure 7.4) are lower than the combined MDs when point and mobile sources are reduced simultaneously. This indicates not only that MDs increase when emissions are reduced, but also the rate of increase in MDs (or the slope of MD curves) increases when emissions are reduced.

The shape of the MD curve also impacts the estimation of total damage as the area underneath the MD curves. Some studies (Levy et al. 2009; Muller and Mendelsohn 2007; Muller et al. 2011) assume that source-specific MDs do not change with NO_x emissions and estimate the total damage by multiplying fixed MDs and

emissions. This linear approximation of total damage is prone to underestimation as it neglects the negative slope of NO_x MD curves.

An MD curve is important from a policy perspective because it is often used along with a marginal abatement cost (MAC) curve to devise economically optimal emission control policies. MAC curves are usually negatively sloped, meaning that the cost per ton of emission reduction increases as emission level decreases. Under an ideal emission control policy, polluters operate at an emission level where their MAC is equal to their MD (Montgomery 1972; Tietenberg 1995). This is due to the fact that the cost of removing an additional unit of emissions is justified by larger external benefits up to the point where the MD curve intersects with the MAC curve. After this point of economic equilibrium, further reduction in emissions is not economically justified. However, for a negatively sloped MD curve, the intersection point will occur at a lower emission level than that suggested by the traditional view in environmental economics literature (see Figure 7.5). Note that MDs at baseline emission levels are often much higher than MACs (Mesbah et al. 2013). If the MDs were lower than the MACs for baseline emissions, the negatively sloped MD curve could have resulted in a higher or lower optimal emission level than the traditional optimal level, depending on the relative slope of the MAC and MD curves. This finding has implications for the current U.S. cap-and-trade program. The target system-wide emission cap for a cap-and-trade system is often set based on the perceived equilibrium point where MAC and MD are equal. Negatively sloped MD

curves, as those found in this study, suggest a lower system-wide emission cap as that suggested by the traditional literature.

The particular shape of the NO_x MD curves calculated in this work are due to the fact that ozone is a secondary pollutant. For primary pollutants an MD curve with positive (or zero) slope that conforms to the traditional view is expected. However, production of ozone from NO_x occurs through nonlinear chemistry that is inherently non-convex in precursor concentration or availability. Note that our definition of MDs calculated in this paper are based on only short-term mortality of ozone exposure, and the damages caused by particulate matter (PM) formed by NO_x emissions are not included. The overall form of NO_x MD curve with inclusion of PM related damages may be different than those presented here. However, as contribution of NO_x to PM is also due to secondary inorganic constituent formed through chemical equilibrium, one would expect persisting, or even enhanced non-convexity with inclusion of PM in the damage function. Nevertheless, overall NO_x MD curves require further investigation to explore this hypothesis.

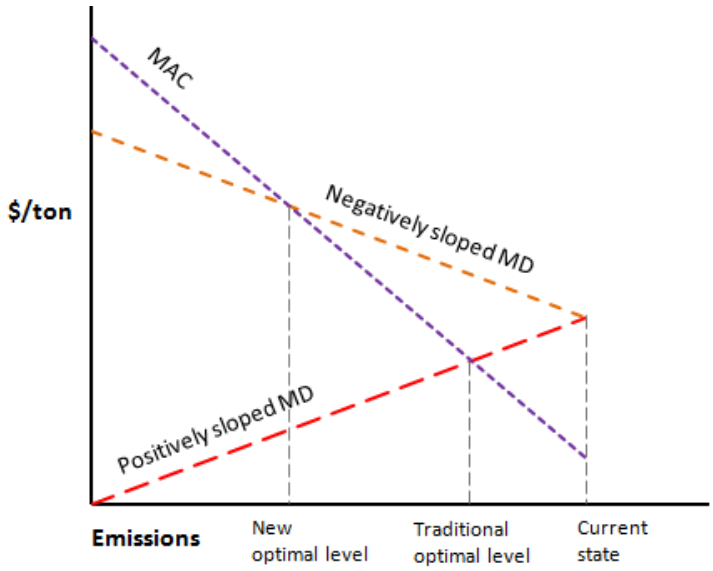

Figure 7.5. Schematic of traditional and negatively sloped MD curves in an economic equilibrium framework.

It should be noted the MD curves presented are based on a uniform emission reduction over the entire domain, not emission reductions only from the selected power plant or urban mobile emissions while other emissions are kept constant. As such, our MD curves should be regarded as projection of response to global rather than local changes in emission patterns. The change in emissions from one source in a large domain is small compared to available ambient NO_x in the system, and therefore one can assume emissions from one source have no tangible impact on the atmospheric regime. Such an assumption results in a flat MD curves when other sources do not change their emission levels. Changes in sectoral emissions such as mobile or electricity generating sources seldom happen in isolation and commonly materialize in a broader nationwide context. As such, we believe that our depiction of MD curves as response to global changes provides a more realistic and relevant view.

7.4. Conclusions

This work demonstrates that unlike the traditional form of MD curves, NO_x MD curves are negatively sloped. This finding is in favor of policies supporting more aggressive emission reductions because further reductions increase marginal benefit or benefit per ton. The benefit of NO_x emission reductions has been debatable in some cities where the atmospheric regime is NO_x-inhibited and the initial MD is negative (Pappin and Hakami 2013). While for these locations, emission reductions may not be beneficial in the short-term, consideration of MD curves along the emission control path provides support for continued NO_x control in urban environments. Inclusion of information garnered from MD curves into policy design is beneficial as it influences the determination of the optimal cap for the cap-and-trade program. In presence of negatively sloped MD curves, such as those for NO_x damage through ozone, lower emission targets than previously thought would be advisable.

CHAPTER 8:

SHORT-TERM NO_X EMISSIONS CONTROL BEHAVIOR IN THE REGULATED AND RESTRUCTURED ELECTRICITY MARKET

8.1. Introduction

Coal-fired electric generating power plants are among the main sources of nitrogen oxides (NO_x) emissions. The amount of electricity (output) that a power plant generates is highly dependent on the price of electricity and the restrictions on emissions. NO_x emissions are among the main precursors of surface ozone (a major constituent of smog), and are usually regulated during hot seasons, which are referred to as ozone seasons (May to September, inclusive), when NO_x emissions have higher ozone formation potentials. Under NO_x cap-and-trade programs, facilities receive allocated emission quotas which determine how much they are permitted to emit; one quota allows a facility to emit one ton of NO_x during the ozone season. Power plants can adjust their electricity generation or buy quotas in the short-term to meet emission requirements (i.e., holding enough quotas to cover their ozone season emissions). The main objective of this chapter is to predict how power plants behave in the short-term to comply with emission requirements while facing different electricity regulations and opportunity costs. Fowlie (2010) compared power plants' outputs for two years (before and after a NO_x trading program was in place in the eastern U.S.), and concluded that output reduction is not a compliance option for power plants in the long-term. This study finds that output adjustment is a short-term compliance option

for power plants and investigates how this option is affected by control technology and electricity market structure.

The United States divided into two different types of electricity markets: regulated and restructured. In a typical regulated market, power plants receive a pre-determined rate of return on their investments, whereas power plants in a restructured market are not guaranteed such returns. This difference in electricity markets provides more of an incentive for power plants in the regulated market to invest in abatement technologies in the long-term (Fowlie 2010). In this chapter, we investigate the short-term behavior of power plants and discuss whether they choose to adjust output or buy emission quotas. To predict the short-term behavior of power plants, we develop an analytical method to estimate plant level marginal abatement costs (MAC) for purchasing emission quotas that accounts for unit-level emissions control technologies as well as the opportunity cost[9] of output reduction.

The MAC has been calculated in three main ways (Vijay et al. 2010): the microeconomic method, the econometric method, and the engineering-economic method. Microeconomic methods are based on cost function and distance function approaches. In the cost function approach emissions are considered as one of many inputs contributing to cost for plants (Kolstad and Turnovsky 1998). The cost function approach minimizes the plant's cost and calculates the change in that cost when the emission levels change. The distance function approach (Färe et al. 1993) treats

[9] Opportunity cost is defined as the foregone benefit of selling electricity.

emissions as bad outputs and calculates the MAC as the shadow price[10] of reducing one unit of output using the duality between the revenue function and distance function. The econometric method (e.g., Becker 2005) uses data on capital and operation costs from plants and estimates the MAC. The engineering-economic method is based on technical information on control technologies and their corresponding costs (Wanger and Schopp 2007). The microeconomic and econometric methods do not consider the effect of control technologies on the abatement, whereas the engineering-economic method does not consider the opportunity cost of abatement (Vijay et al. 2010). In this work, we propose a method for the calculation of MAC that includes both control technology costs and the opportunity costs of output reduction. Then, we assume different output adjustment behavior (and different opportunity cost) for power plants under the regulated and restructured markets, and estimate MACs for U.S. coal-fired power plants. Next, we perform a regression analysis to examine how power plants' output adjustment behavior differs by electricity market structure. Subsequently, a firm's predicted behavior based on the estimated MACs is compared to actual emissions compliance in the NO_x trading market. Finally, we perform another regression analysis to identify the significant parameters affecting power plants' emission trading behaviors and construct a supply and demand function for emission quotas.

---

[10] Shadow price in an optimization problem (e.g., abatement cost minimization) is the change in the objective function by strengthening the constraint (e.g., external cap on emissions) by one unit.

8.2. NO_x control technologies and regulations

Both control technologies and regulations can affect abatement costs. Stricter emission standards lead to higher costs, and more cost-effective control technologies can lead to lower abatement costs for a power plant. Two main categories of NO_x control technologies are combustion modification and post-combustion control. Combustion technologies aim to prevent the generation of NO_x emissions by modifying the combustion processes, whereas the post-combustion technologies aim to reduce the NO_x emissions that have already been produced. Combustion modification can also reduce the plant's efficiency due to sub-optimal combustion.[11] Two common post-combustion control technologies for NO_x are selective non-catalytic reduction (SNCR) and selective catalytic reduction (SCR). Both of these technologies are designed to reduce NO_x using a chemical reaction.[12]

The Clean Air Act (CAA) amendment of 1990 is the main air quality regulation in the U.S. and has led to significant efforts to control NO_x emissions. Title

---

[11] More information about NO_x combustion modification technologies can be found elsewhere (e.g., Sloss et al. 1992).

[12] The reaction converts NO_x into water and molecular nitrogen (N_2) which is a harmless gas and can be released to the atmosphere. In both technologies, a reagent, ammonia or urea, is added to the NO_x stream. The difference between the two technologies is that in SNCR, the reaction occurs at high temperature (1145-1365°K), whereas in SCR, adding a catalyst allows the reaction to occur at a lower temperature (590-700°K). The efficiency of SNCR is usually between 30% to 50% reduction, whereas the efficiency of SCR is about 90% (Ziadi and Kumar 1995).

I and title IV of the 1990 CAA amendment both relate to NO_x emissions control. Title IV led to creation of the acid rain program that required a 2 million ton per year reduction in NO_x emissions from coal-fired power plants. Title I included provisions for attainment and maintenance of the National Ambient Air Quality Standard (NAAQS). It outlined a timetable for states to meet the NAAQS (Burtraw et al. 2005).[13]

The CAA also gave authority to the U.S. EPA to create an emissions trading program. The emissions trading program caps the total emissions of NO_x during the ozone season by allocating quotas to plants. The allocation is based on a unit's type (e.g., voluntary[14], industrial, electric utility, etc.) and their average net output (in MWh) during the two ozone seasons prior to the compliance year. The plants can trade their emission quotas, but they must have enough allowances at the trading deadline (December 31) to cover their emissions. The clean air interstate rule (CAIR) which included provisions for an emissions trading system was promulgated in 2005 and replaced the NO_x budget program (NBP) in 2008 (EPA 2008) after a series of litigations. CAIR is similar to NBP but it covers more states. CAIR remains in effect while the U.S. EPA reviews the now vacated cross-state air pollution rule (CSAPR).

---

[13] A comprehensive regulatory history of NO_x regulations can be found elsewhere (Burtraw and Szambelan 2009).

[14] "Opt in" units voluntarily agree to participate in NBP. They are subject to all the requirements for NBP sources.

Power plants that participate in the U.S. cap-and-trade program face different electricity regulations. Under a regulated electricity market, the regulators set rates for power plants and allow them to gain a reasonable (and sometimes inflated) rate of return on their investments and costs. Rate of return regulation would, therefore, favor investment into cleaner technologies compared to restructured competitive markets that compete on marginal cost pricing. In the 1990s, there was a large difference in rate of return across U.S. states (Warwick 2002). By 2001, 19 out of 50 U.S. states switched to a restructured market system, and 12 of these 19 states were among the 19 states participating in the NBP program.[15]

8.3. Methodology

Power plant emissions can be controlled either by reducing electricity generation or through abatement activities (e.g., combustion modification or post-combustion technologies). In the short-term, when power plants do not have enough time to install a new control technology, they meet their emission requirements, imposed by a cap-and-trade program, by output reduction or by purchasing emission quotas. Therefore, the amount that power plants are willing to pay for the emission quota depends on the opportunity cost of output reduction. Here we present a method to calculate the unit-level MACs. The benefit maximization problem for a plant participating in both the electricity and emissions market is:

---

[15] Among NBP states, CT, DE, IL, MA, MD, MI, NJ, NY, OH, PA, RI, and VA were under restructured electricity markets, and AL, IN, KY, NC, SC, TN, and WV were under a regulated system.

$$Maximize: \; P_q q - c(q, A)$$

$$Subject\ to: \tag{8-1}$$

$$e(q, A) \leq \bar{e}$$

where P_q is the exogenous electricity price, q is the electricity generation level, c is the cost function that depends on output (q) and abatement activity (A), e is the emission level of the plant that is also function of q and A, and $\bar{e}$ is the exogenous constraint on emissions.

The Lagrangian problem function associated with equation 8-1 is:

$$L = P_q q - c(q, A) - \lambda(e(q, A) - \bar{e}) \tag{8-2}$$

The first-order conditions are:

$$\frac{\partial L}{\partial q} = P_q - \frac{\partial c}{\partial q} - \lambda \frac{\partial e}{\partial q} = 0 \tag{8-3}$$

$$\frac{\partial L}{\partial A} = -\frac{\partial c}{\partial A} - \lambda \frac{\partial e}{\partial A} = 0 \tag{8-4}$$

$$\frac{\partial L}{\partial \lambda} = -e(q, A) + \bar{e} = 0 \tag{8-5}$$

From (8-4) and (8-5), λ is:

$$\lambda = \frac{-\frac{\partial c}{\partial A}}{\frac{\partial e}{\partial A}} = -\frac{\partial c}{\partial e} = \frac{P_q - \frac{\partial c}{\partial q}}{\frac{\partial e}{\partial q}} \tag{8-6}$$

where λ is the Lagrangian multiplier that is equal to the marginal cost or shadow

price of e. Equation 8-6 is valid if the constraint on emissions (equation 8-1) is binding. If there is a market in place and the quota price is not zero, the total emissions cap in the system and caps for firms are binding (Farrow et al. 2005). To estimate unit-level MAC, we use the following cost function:

$$c = (P_i R_{io} + c_p + c_c)\, q \tag{8-7}$$

which is a function of:

$q =$ Electricity generation (MWh)

$P_i =$ Input (coal) price ($/MMBtu)

$c_p =$ Power plant's capital and operating cost ($/MWh)

$c_c =$ Control technology's capital and operating cost ($/MWh)

$R_{io} =$ Rate of transformation between input and output (MMBtu/MWh)

From equations 8-7 and 8-8, the unit-level MAC can be calculated as follows:

$$MAC = (P_q - P_i R_{io} - c_p - c_c) R_{oe} \tag{8-8}$$

where P_q is the electricity price and R_{oe} is the rate of transformation between output and emissions (MWh/ton).

Equation 8-8 is used to estimate MAC for power plants in regulated and restructured electricity markets. We assume that P_q=average regulated retail market price and that P_q=minimum retail market price in a restructured market. We base our reasoning on the fact that power plants in restructured markets have an incentive to

meet the ozone season emission requirements by reducing emissions (giving up outputs) on days when the electricity price is lower (Martin 2008; Martin et al. 2007). As a rate of return is set based on power plants' average costs, the average electricity price is used to approximate short-term benefit under the regulated market. We will perform a regression analysis later to examine this assumption and investigate how power plants' output adjustment behavior differs by market structure. The MAC for a plant is important since it is one of the main factors for a plant's decisions to exchange emission quotas in the emissions market. When electricity is demanded, a plant operates with full capacity (or maximum emission level) if its MAC is higher than the emission quota price. In this case, the plant prefers not to reduce any output to control emissions because the additional cost for the plant is more than the price of the quota. As such, the plant prefers to pay for buying emission quotas as needed. A plant operates at its minimum emission level when its MAC is less than the emission quota price. For this case, the price of a quota is higher than the additional cost of one unit of emissions reduction. As a result, the plant would prefer to reduce its electricity level to the minimum possible level and sell its emission quotas to other plants. Note that the amount of emissions reduction and trading on the emissions market depends on the initial allocation of quotas.[16]

[16] All explained behaviors are only valid with the assumption of a perfect market in place.

8.4. Data and cost estimation

The proposed method is used to calculate MACs for coal-fired electric generation units with SCR or SNCR control technologies that participated in the NBP program during the ozone season of 2007. The capital and operating costs for the control technologies are calculated using the U.S. EPA Integrated Planning Model (EPA 2010). The power plants' capital and operating costs, the coal price, and the electricity price (minimum monthly retail price, and average monthly retail price) for the ozone season are taken from the U.S. Energy Information Administration that provides prices by location (EIA 2010, 2008, n.d.). To calculate rate of transformations, the emission-output and output-input rates are estimated using the daily unit-level data from the U.S. EPA clean air market (EPA n.d.) and a standard ordinary least square regression.

In the following section, an application of the proposed methodology for the unit-level MAC calculation is demonstrated for the Rochester Power Plant (RPP), which is a good representative of an emissions quota buyer in the system. We proceed by estimating MACs for the selected units and contrast them by technology and location. The RPP is located in New York and has four units with SNCR control technology. The emission-output and output-input relationship for these units are demonstrated in Figures 8.1 and 8.2 respectively.

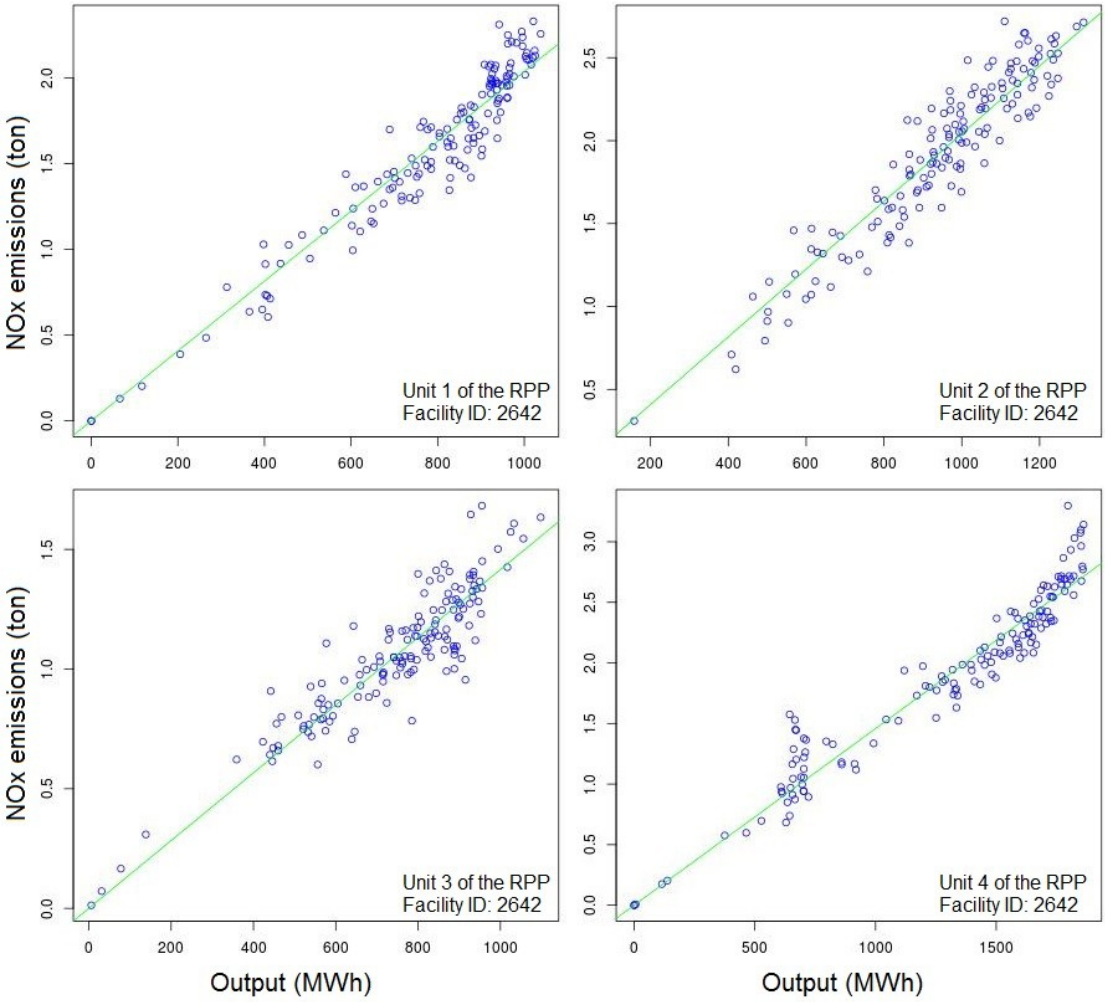

Figure 8.1. Output-emission relationship for units of the Rochester Power Plant. Each point represents the daily emissions and the daily output.

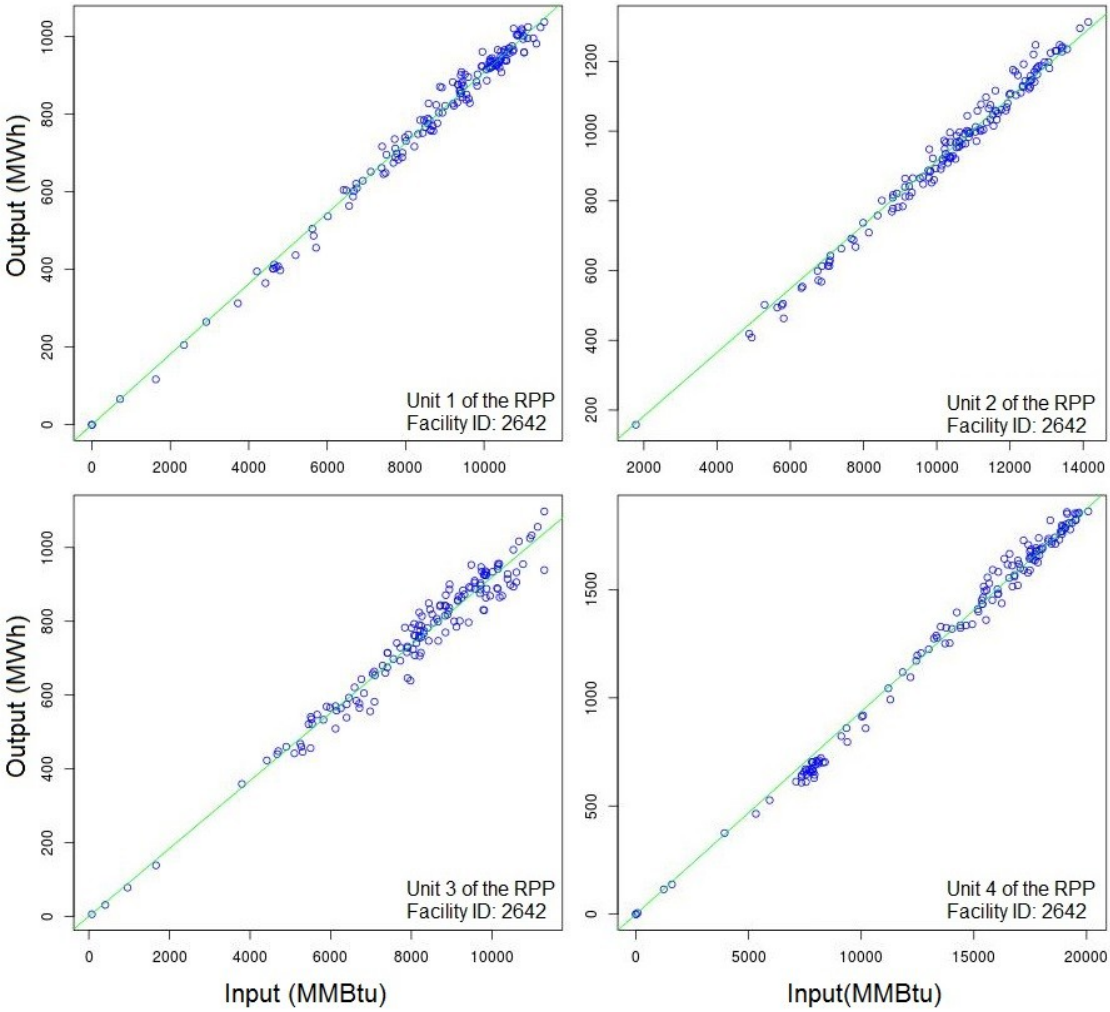

Figure 8.2. Input-output relationship for units of the Rochester Power plants. One point represents daily input and daily output.

The ozone season MACs for all units of the RPP are presented in Table 8.1. All MACs are higher than the average quota price ($825/ton) for the ozone season (EPA, 2008) meaning that this power plant should always try to buy quotas. The compliance data indeed confirms that the power plant was a quota buyer.

\

Table 8.1. MAC for the Rochester Power Plant in the ozone season of 2007.

	Unit 1	Unit 2	Unit 3	Unit 4
MAC (thousand $/ton)	19.5	20.1	29.2	29.4

Compliance data for the RPP for the ozone season are presented in Table 8.2.

Table 8.2. Compliance data for the Rochester Power Plant in the ozone season of 2007 (Clean air market U.S. EPA).

	Unit 1	Unit 2	Unit 3	Unit 4	Overdraft	Total
Allocated	96	131	128	276	0	531
Banked	66	3	4	5	0	78
Current held	3	4	4	33	1049	1093
Total held	69	7	8	38	1049	1171
Required deduction	264	263	219	219	0	965
Current deduction	3	4	4	33	872	916
Banked deduction:						
1 to1	16	1	1	1	0	19
2 to 1	50	2	3	4	0	59
Total deduction	69	7	8	38	872	994
Carried over	0	0	0	0	177	177

The RPP's overall quota withdrawal in 2007 was 994, whereas its required deduction was 965 (Table 8.2). The required deduction is determined based on the emissions of the power plant in the ozone season. The extra deduction of quotas was

because of the use of banked quotas. Overall, the plant used all its 78 banked quotas, and 59 of them were subject to a 2-to-1 exchange rate and could cover 30 ton of NO_x. The plant ended up with more quotas than needed in the current year, and 177 of them (total held - total deducted) were carried to the next year. From the compliance data, the plant was a quota buyer and bought 542 (current (1093) – allocated (531) – carried from last year (38)) quotas in 2007.

In 2007, 2589 units took part in the NBP. Among them, 2018 units were electric utility generations (EGUs), and 640 units were coal-fired. Only 250 of the latter had SCR or SNCR control technologies. These 250 units generated more than 165 kilotons (kt) of NO_x emissions in the ozone season of 2007, which accounted for about 40% of the total coal-fired EGU emissions (418 kt), and 37% of total EGU emissions (442 kt), or 32% of the total emissions of all facilities (including non-EGUs) in the NBP (506 kt). In this study, 218 of these 250 units were selected based on the available electricity generation data for the units.

The same procedure (as that applied to the RPP) is repeated for all other selected units. To show how the MAC changes from one location to another, 4 different regions were chosen (Figure 8.3). Note that some of the states in Figure 8.3 did not participate in the NBP (e.g., Georgia). Also, to show how the MAC changes for different technologies, units with SCR and SNCR were considered separately. The

control technology of the selected units in different regions, as well as the type of electricity market (i.e. regulated or restructured) are also shown in Figure 8.3.[17]

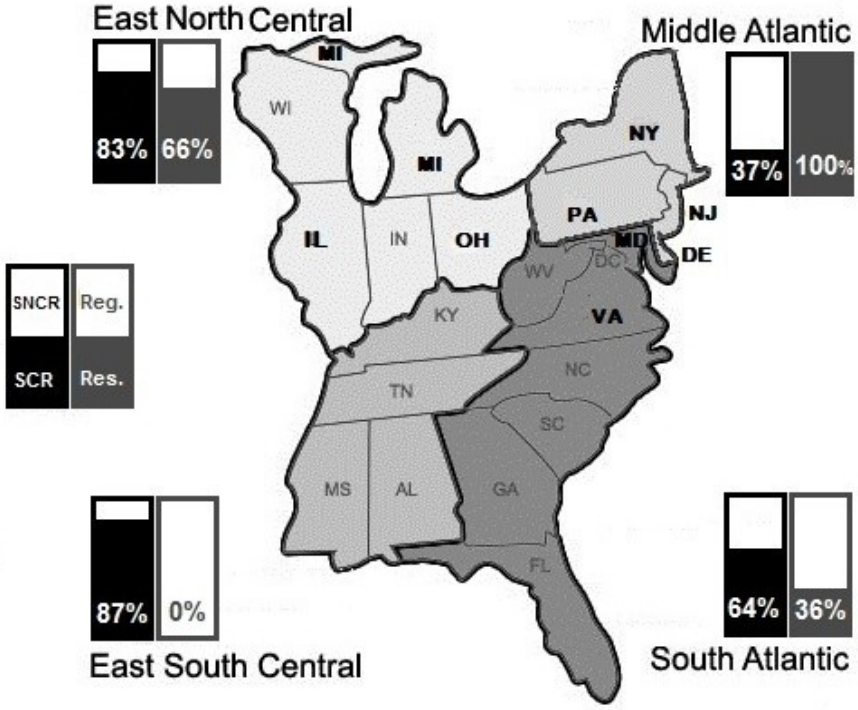

Figure 8.3. Selected units by region. The restructured states are shown in bold text. The columns represent the percentage of selected units in the restructured market and the percentage of units with SCR control technologies at each region.

A summary of unit-level information including the estimated MACs are presented in Table 8.3.

---

[17] 71 units were selected in the East North Central region, 38 in the Middle Atlantic region, 45 in the South Central region, and 64 in the South Atlantic region.

Table 8.3. Unit-level descriptive statistics by region.[a]

	East North Central	Middle Atlantic	East South Central	South Atlantic
Capacity (MW)	436(257)	346(297)	429(288)	444(293)
Ozone season generation (GWh)	1297(866)	1028(1035)	1438(1003)	1384(977)
Ozone season emissions (ton)	755(667)	794 (558)	574(407)	635(486)
Emission-output rate (lb/MWh)	1.53(1.56)	2.76(1.68)	0.98(0.96)	1.15(0.81)
Avg. electricity price ($/MWh)	77.3(8)	109.7(26.2)	68.6(6.8)	75.8(17.7)
Min. electricity price ($/MWh)	75.2(7.6)	105(22.7)	65.4(7.7)	71.2(14.5)
Input cost ($/MWh)	17.4(7)	31.6(7.3)	20.8(1.5)	27.7(4.9)
Power plant cost ($/MWh)	67.6(3.6)	68.9(4.2)	67.8(4)	67.5(4.1)
Control technology cost ($/MWh)	11.4(4.3)	6.8(4.1)	12(4.2)	9.1(4.8)
MAC (1000 $/ton)	-43(31)	-6.7 (22)	-102 (60)	-89(78)
Number of quotas traded b	437(956)	5(634)	669(757)	487(1259)
Percent of buyers in the region	24	58	13	31

[a] Values are average over the region. Standard deviations are in parentheses.

[b] Values are from the actual compliance data, the negative values represent the number of quotas a power plant bought.

Among the regions studied, all states in the middle Atlantic are part of the restructured electricity market (Figure 8.3). Power plants in the middle Atlantic region have the highest average electricity price and highest predicted MACs compared to the other regions (Table 8.3). In our model a high-MAC plant prefers to buy quotas

instead of controlling its emissions by output reductions. This predicted behavior is consistent with the actual market data which indicates that there are more quota buyers in the middle Atlantic region (Table 8.3) as compared to other regions.

Power plants in the middle Atlantic region also have the highest average emission-output rate because they mainly use lower efficiency and cheaper control technologies (i.e., SNCR) to meet regulations. 63 percent of studied units in the middle Atlantic region have SNCR control technologies and the rest are equipped with SCR control technologies. The majority of plants in this region had not invested in SCR. Being in a restructured electricity market under which a return on power plants' investment is not guaranteed could have been a contributing factor. Therefore, power plants are less likely to invest in higher efficiency, and thus more expensive technologies in the long-term (Fowlie, 2010).

In the east south central region, where all states are in the regulated electricity market, the electricity price and the estimated MACs are the lowest among all regions (Table 8.3). The low predicted MACs show that plants are not willing to buy quotas. This prediction is consistent with the actual market data which indicates that power plants in the south central region sold more quotas on average than other regions. In this region, most power plants have SCR control technologies, and thus the emission-output rate is the lowest, and the control technology is the most expensive among all regions.

We first conduct a simple OLS regression to investigate how power plants' short-term output adjustment is affected by the type of technology and electricity

market.[18] For this regression, the dependent variable is a ratio which represents average used capacity. The latter is defined as the ratio of actual seasonal electricity generation to the full seasonal generation capacity. The independent variables are a constant, a dummy variable for control technology (i.e., 1=SCR technology, and 0=SNCR technology), and a dummy variable for the type of electricity market (i.e., the 1=regulated market and 0=restructured market). Table 8.4 summarizes the regression coefficients and related statistics.

Table 8.4. Regression statistics and estimated regression variables.[a]

Variable	Coefficient	Std. error	t stat
Intercept	0.730	0.021	35.08
Control technology	0.095	0.024	4.01
Type of market	0.037	0.022	1.70

[a] The dependent variable is a ratio representing the used capacity.

The Intercept in the regression model represents the average capacity used for the power plants studied. From coefficients presented in Table 8.4, it is inferred that only 73 % of capacity is used on average, but SCR increases capacity used by 9.5 % and a regulated market by a further 3.7 %. Power plants with SCR technology make better use of capacity because of higher efficiency. The cleaner control technologies provide an opportunity for power plants to generate more electricity while being

[18] The SPPS Statistics version 22 is used for the regression analysis.

constrained by emission regulations. The regulated market also encourages power plants to use a higher percentage of their generation capacity. This is because in the regulated market, the profit of power plants is less sensitive to fluctuation in electricity price and power plants generate based on pre-determined rates of return. A lower use of capacity by power plants in the regulated market indicates that the output adjustment behavior differs by market structure and supports our assumption presuming different opportunity costs for power plants in the two types of electricity markets. Note that the regression results presented in Table 8.4 are based on seasonal output adjustment and indicate different seasonal output adjustment behaviors. However, they do not indicate that power plants reduce their output on days when electricity price is high. We did not perform a regression analysis based on daily output adjustment due to lack of daily electricity price data in the regions studied.

To evaluate the accuracy of the estimated unit-level MACs, they are compared with the actual compliance data. Sellers are expected to have a MAC lower than the quota price. On the other hand, the number of quotas that quota sellers hold at the beginning of a trading timeline are expected to be higher than the number of quotas that they hold at the end of trading timeline. The number of quotas at the beginning of the timeline is allocated quotas plus those carried from the previous year, and the required quotas for a unit at the end of the timeline is equal to its aggregate ozone season emissions. According to the compliance data from the EPA, 153 units out of 218 (70 %) are quota sellers. Overall, for 154 units (71%), the compliance data and predicted MACs are in agreement. The MAC is compared with the traded quotas

(allocated quotas plus quotas carried from last year minus actual emissions) in Figure 8.4.

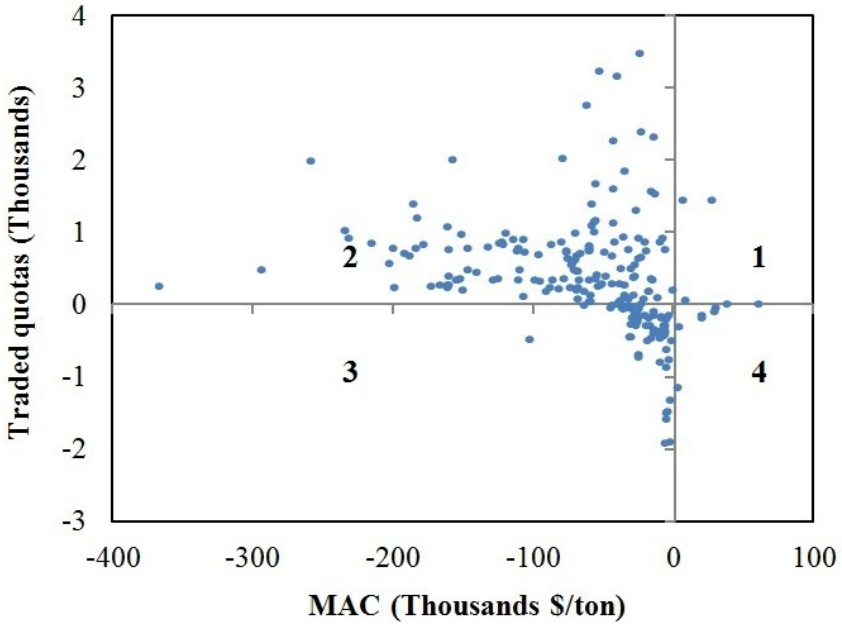

Figure 8.4. Comparison of the units' behavior (seller/buyer) based on the estimated MACs and actual compliance data. A positive number of traded quotas indicates the number of quotas sold. Note that the vertical axis intersects with the horizontal axis where MAC is equal to quota price (i.e., $825/ton).

The vertical axis in Figure 8.4 represents the number of traded quotas based on compliance data in 2007. The positive values on the vertical axis represent units that had more quotas than their actual emissions (quota seller), and the negative values are quota buyers. Correspondingly, the points on the right hand side of the vertical axis suggests that units are quota buyers because their MACs are more than the quota price, while the points on the left side of the vertical axis are units with a MAC smaller than the quota price (likely quota buyers). Region 2 (68% of points) and 4 (2% of points) are areas where estimated MAC predicts the units' behavior (seller/buyer) correctly and region 1 (3% of points) and 3 (27% of points) are the area

where estimated MAC does not match actual trading behavior. An ordinary least square fit for the presented data in Figure 8.4 leads to a p-value of 1.83×10^{-5} R^2, which indicates that the estimated MACs are good predictors of the trading behaviors.

The discrepancy in trading behavior in Figure 8.4 can be attributed to a number of reasons. The assumption that units would operate at their maximum benefit level may not be completely accurate especially when the plant is unsure about the future year's market conditions and regulations. The possibility to bank quotas into the future years is an incentive for potential sellers to operate at an unfavorable shadow price in anticipation that the quota price would increase in the future. This motivation caused the accumulations of unused quotas under NBP. The U.S. EPA used the progressive flow control provisions to resolve this issue that existed under NBP.

More importantly, the initial quota allocation method is a major source of discrepancy. The allocation of the allowances based on plants' historic output levels is a strong incentive for plants to maintain outputs to secure more allowances for future years. This incentive can affect plant's behavior in the market, and can prevent potential quota sellers from selling their quotas. Important future research extension is to predict a plant's behavior in a multi-year time frame using a dynamic optimization model which accounts for the effect of future year quota allocation on the plant's current behavior.

Besides banking and grandfathering quotas, internal quota transaction is another possible reason for potential sellers to deviate from the predicted behavior. A

power plant is responsible for all of its units' compliance and emissions. For example, if one unit of the plant has a high MAC and another unit has a low MAC, the plant can use the quotas from one unit to cover the emissions from the other, and that causes discrepancy in the estimated unit-level behavior. To explore this discrepancy, the Cross Power Plant located in South Carolina is examined in more detail. The plant has three units and its behavior from compliance is not correctly predicted by the estimated MAC. From MAC estimation, all 3 units are quota sellers, whereas from compliance data one of the units (unit 3) is not a quota seller. Units 1 and 2 have a high number of allocated quotas, whereas unit 3 has a low number of quotas (Table 8.5). The reason of low allocation for unit 3 is that the unit was shut down in 2006 (EPA clean air market) and as quotas are grandfathered based on the output level from two previous years, unit 3 received a low number of quotas.

Table 8.5. Compliance data for Cross Power Plant in the ozone season of 2007 (Clean air market U.S. EPA).

	Unit 1	Unit 2	Unit 3	Overdraft	Total
Allocated	1433	1450	120	0	3003
Quotas held at deadline	943	970	120	487	2520
Emissions (ton)	943	970	597		2510
Total deduction	943	970	120	477	2510
Carried over to next year	0	0	0	10	10

Low quota allocation to unit 3 created a high demand for this unit to buy quotas. On the other hand, unit 1 and 2 had excess quotas because they received significantly higher number of quotas. The plant sold 483 quotas (3003-2520) of the

spare quotas from unit 1 and 2, used 477 quotas to cover part of emissions form unit 3 (internal transfer of quotas), and saved 10 quotas for next year.

As discussed above, MAC is not the only factor affecting the number of emission quotas traded. To construct a supply and demand function for emission quotas, we perform a multivariable linear regression, which estimates the traded quotas as a function of different parameters including the number of allowances transferred from last year, capacity, control technology, electricity market structure, and MAC. A summary of the regression variables, the estimated coefficients, and relevant statistics are listed in Table 8.6.

Table 8.6. Regression statistics and estimated regression variables. [a]

Variable	Coefficient	Std. error	t stat	Significance
Intercept	-361.0	86.94	-4.15	0.000
Allowance form last year	0.934	0.044	21.36	0.000
Capacity (MW)	-0.757	0.372	-2.04	0.043
Capacity squared (MW^2)	0.002	0.000	4.91	0.000
Technology (SCR=1, SNCR=0)	543.0	92.60	5.86	0.000
Type of market (Reg.=1, Res.=0)	148.5	69.29	-2.14	0.033
MAC ($/ton)	-0.002	0.001	-2.71	0.007

[a] The dependent variable is traded quotas and the R square for this regression is 0.786.

All parameters reported in Table 8.6 significantly impact the emission trading behavior and all have the expected signs. MAC is negatively related to the number of

quotas sold. We included the a second-order term (capacity squared) in our regression to account for the scale effect. A coefficient of -0.757 for the capacity and 0.002 for the capacity squared represent a convex function with negative starting value which turns to a positive effect at a capacity of 378.5 MW. This indicates that smaller power plants are more likely to buy emission quotas, whereas larger plants are more likely to sell emission quotas. Larger plants also can generate more electricity, which results in more allocated emission quotas for the next compliance year. Due to the multicolinearity between capacity and allocated quotas, we only included capacity in our regression analysis.

We also included two dummy variables in our regression analysis to distinguish between SCR and SNCR control technologies, and between regulated and restructured electricity markets. The positive coefficient for the control technology indicates that units with SCR sell more quotas. Units with SCR have lower emission-output rates and can generate more electricity (and obtain more emission quotas) than units with SNCR if they face the same emission caps, and therefore these units are more likely to sell their unused quotas.

The positive coefficient for the market structure suggests that power plants in the regulated market are more likely to sell quotas. This short-term behavior of power plants is consistent with their long-term behavior in the regulated market where plants have more incentive to invest in SCR (Fowlie 2010). In other words, power plants in the regulated market are more willing to invest in SCR rather than buying quotas to comply with the requirements on their emissions. In the restructured market, on the

other hand, power plants are more likely to buy quotas when opportunity costs are high and reduce output when opportunity costs are low, rather than investing in SCR to comply with their emission requirements. Power plants in the regulated market do not need to adjust their generation according to market conditions since they can factor emission costs in their average cost price or rate of return regulation.

8.5. Conclusions

Unit-level MACs are key decision making factors for power plants as they can help them operate at optimal emissions/output levels. Emissions abatement can occur by output reduction, abatement activities, or purchasing emission quotas. The study of U.S. coal-fired power plants with SCR or SNCR post-combustion technologies shows that the MAC for NO_x quotas vary significantly from one region to another because of differences in the structure of the electricity market, power plant's emission-output rates, their control technologies and the capacity for electricity generation.

Our results suggest that power plants' short-term output adjustment behavior is affected by control technology and the type of electricity market. The control technology is a significant factor with a large impact on capacity use while the market structure has significant but smaller impact. Power plants with more efficient control technologies have lower emissions and will use their electricity generation capacity more. Furthermore, units in the regulated electricity market are more likely to sell quotas in the short-term. Units in the restructured market show a different behavior; they are not as willing to invest in the more expensive SCR technology, but instead buy more quotas in order to comply with emission requirements. Our results also

indicate that the strategic behavior of a plant is significantly dependent on a number of factors including MAC, capacity, structure of the electricity market, control technology, and the number of quotas allocated or transferred from previous year.

Quota allocation under current regulations is based on historic output levels. When units install a high efficiency control technology, they can have lower emissions with the same level of output and quotas. The motivation for plants to switch to higher efficiency control technologies to gain spare quotas for trading in the market is more pronounced in the regulated markets where power plants have a regulated rate of return on their investments.

Unit-level MACs can be used by regulators for the prediction of plant's responses to different regulations. The prediction of short-term behavior of power plants can provide some insights into where emission reductions occur, and how the total target emission reductions are distributed spatially. This prediction is particularly important for NO_x emissions because the location where they are emitted matters in their contribution to the formation of surface ozone (a secondary pollutants that forms most of smog) and corresponding health damage (Pappin and Hakami 2013). Future research can compare the location specific marginal damage of NO_x emissions from power plants and their MACs, and investigate whether the current sets of regulatory incentives, created by electricity and emissions markets, are in favor of health damage minimization. Under a social cost minimization policy, power plants with higher marginal damages emit less as compared to low-marginal damage power plants. For example, power plants located in the East coast of the U.S. have relatively higher

marginal damages (Mesbah et al. 2013; Muller 2011) as compared to other regions. Most of these power plants are in the restructured electricity market which is not likely to provide a high incentive for emission reductions.

Knowledge of the unit-level responses can provide information on how emissions distribution changes spatially in a system as a result of a new policy, and where the priorities for innovation and investment into cleaner technologies are. The effect of system-wide emissions redistribution on ambient air quality and human health can be quantified using air quality models. This, in turn, would help decision makers to predict both the impact of different policies and various emissions trading scenarios on human health and the reaction of firms to different quota allocation frameworks.

CHAPTER 9:

CONCLUSIONS AND FUTURE WORK

Cap-and-trade programs have proven to be effective instruments for achieving environmental goals while incurring minimum costs. The nature of the pollutant, however, affects the design of these programs. NO_x, an ozone precursor, is a non-uniformly mixed pollutant with a short atmospheric lifetime. NO_x cap-and-trade programs in the U.S. are successful in reducing total NO_x emissions but may result in sub-optimal environmental performance because location-specific ozone formation potentials are neglected. Addressing this shortcoming was the main motivation for this thesis. To do so, a model-based decision support system was developed combining economic instruments with an advanced engineering modeling tool. The model was used to evaluate the costs and benefits of different NO_x emission control policies distinguishing emissions based on their temporal and spatial differences.

The current U.S. program for NO_x control is a cap-and-trade system which assigns tradable emission quotas to participants, and controls the total emissions by the number of quotas allocated to participants. A reformed cap-and-trade system can distinguish between emissions by assigning different values to emission quotas at different times and locations. This discrimination between emission quotas assigns higher values for quotas in locations with high ozone formation potential and therefore provides emission reduction motivations for units with higher damages. The results in chapter 4 indicated that better environmental performance at negligibly

higher system-wide abatement costs can be achieved through inclusion of emission exchange rates. Exposure-based exchange rates resulted in better environmental performance than those based on concentrations.

Another approach for distinguishing between emissions is to assign location specific emission fees for polluters. Assigning the fee based on the damage per ton of NO_x results in targeted emission reduction for locations where emissions are more harmful. In chapter 5, we showed that inclusion of location-specific damage information in policy design could significantly enhance public health performance of the current cap-and-trade system. The results presented in this chapter also indicated that the net benefit under the policy that minimizes the social cost (i.e. health costs plus abatement costs) was six times larger than that of an exchange rate cap-and-trade policy. This difference shows the importance of utilizing damage information in allocation of emission quotas under an exchange rate system.

The emission differentiated policies considered in chapters 4 and 5 were designed to shift emissions from high impact to low impact locations without accounting for electricity supply in the system. However, power plant emission adjustment is also a factor of electricity demand. Re-dispatching or shifting electricity generations to high efficiently units (i.e., units with low generations to emissions ratio) is an efficient method of emission reduction while the demand in the system is supplied. Re-dispatching occurs in the system when a high price for NO_x emission quotas is set which provides high incentives for electricity generation by low emission rate units, and in turn results in lower emissions and correspondingly lower health

damage in the system. Lower emissions; however, do not ensure that reductions occur at high MD locations and times. Therefore, charging polluters fees based on their temporal (as well as spatial) MD information results in further reduction in system-wide health damage. In chapter 6, we investigated the impact of inclusion of temporal and spatial marginal damages in emission and electricity markets. The results from a case study of U.S. power plants indicated that time-specific MDs were high around noon and low in the evening in all locations. Furthermore, a substantial emission reduction and net benefits could be gained if high-efficiency power plants in low-damage hours/locations supplied a greater portion of the electricity. The results also indicated that inclusion of temporal effects in NO_x control policies results in a comparable net benefit as compared to policies which included only spatial or spatiotemporal effects. This is an important finding as strategies and trading markets that are based on temporally differentiated MDs are far less complicated than instruments that rely on spatial differentiation.

The MD information can also be used to set optimal emission levels. However, the specific shape of the MD curve affects the optimal emission level. The results presented in chapter 7 indicated that MDs in all locations are higher for lower baseline emissions, which means source-specific NO_x MD curves are negatively sloped. This finding was in contrast with the traditional concept in environmental economics, which assumes a positive slope for source-specific MD curves. This unconventional behavior of NO_x MD curves was due to the atmospheric photochemical reactions of NO_x and ozone. This finding is important from a policy

perspective as using a negatively sloped MD curve instead of a positively sloped curve results in a lower emission reduction targets.

Power plants participating in cap-and-trade programs are faced with requirements from both emission and electricity markets. The type of electricity market plays an important role in power plants' behavior and how they adjust their emission level and trade emission quotas. In a regulated electricity market, power plants sell electricity based on a pre-determined rated of returns, whereas such rates do not exist under a restructured electricity market. In chapter 8, we investigated the short-term NO_x emission behavior in the regulated and restructured electricity market and found that power plants in the restructured electricity market are willing to pay more for emission quotas. These behaviors were due to differences in the opportunity cost of output reduction.

All emission differentiation policies presented in this thesis are dependent on the accurate calculation of source-specific impacts or marginal damages and the accuracy of the adjoint model. For a normal air quality model, the performance evaluation can be done through comparison of concentrations calculated by the model and observation data. For an adjont model; however, there is no observation to be compared with because the outputs are derivatives. This shortcoming is a general limitation of sensitivity methods. The adjoint model used in this thesis has been validated extensively against the forward model (Hakami et al., 2007). As is often the case with sensitivity studies, consistency with the underlying model (i.e. forward

CMAQ) is the main form of validation but the uncertainties and limitations existing in CMAQ remain unaddressed in the adjoint of CMAQ as well. MDs calculated by the adjoint model in this thesis were used within an optimization framework to reduce the system-wide damage. The fact that the examined policies resulted in a significantly lower damage is an indicator of the consistency of the adjoint and forward models. However, these findings are only as reliable as the forward model.

It is also worth mentioning that the modeling domain in this study was a North American domain with a 36 km grid resolution. This coarse resolution may overestimate some potential health damages (Thompson and Sin 2012) but it was chosen due to the computational limitations for continental simulations. While grid resolution can affect the results presented in the thesis, a larger domain and inclusion of other continents is likely to have little impact on the estimated MDs because NO_x has a very short life-time in the atmosphere and cannot travel on a continental scale.

Furthermore, we did not account for uncertainties associated with economic valuation, epidemiological concentration response factors, baseline mortality rates, or source-receptor relationships, all of which affect the estimated MDs and the subsequent outcomes under different economic instruments. Incorporating uncertainty in MDs within economic instruments is of interest of policymakers from several perspectives. Firstly, the uncertainty involved with source-receptor relationships, caused by large uncertainty in emissions and meteorology, affects the relative magnitude of source-specific MDs, which are important for performance

enhancement under exchange rate policies. Moreover, using the probability distribution function of source-specific MDs, policymakers can determine the likelihood of MDs being greater than the MACs, which is essential for sizeable net benefits under taxation policies. In other words, policymakers would be aware of the magnitude of net benefits as well as the likelihood of achieving such benefits. Lastly, information on uncertainty in MDs can be used by policymakers to set the optimal emission level with a reasonable level of confidence. Failure to identify and set the optimal level accurately may result in a substantial loss in system-wide costs.

It should be noted that the NO_x damage function in our study was based on ozone exposure in the U.S. only (Canada and Mexico not included). Additionally, mortality due to exposure to NO_2 was not included in our damages function as in the U.S., there is not enough evidence to link exposure to NO_2 with mortality. Inclusion of NO_2-based damages for calculation of MDs results in larger NO_x MDs and subsequently larger benefits obtained under differentiated policies. Inclusion of NO_2-based MDs is particularly important in large cities where ozone-based NO_x MDs are initially negative and emissions reduction is not beneficial. Accounting for NO_2-based damages results in an upward shift in the MD curve, indicating that emissions reduction will be beneficial earlier than expected when only ozone-based damages are included. This, in turn, further supports policies in favor of emissions reduction in large cities.

It should be noted that all results presented in this thesis are proof-of-concept and are not recommended for direct policy applications. There are several limitations in the proposed decision support system. Firstly, the calculated MDs were the integration of temporal and spatial MDs over one summer season. Modeling of longer time intervals is required to investigate how MDs vary from one year to another. Secondly, the modeling grid cells were large due to computational limitations, and sometimes included more than one power plant. Thirdly, various uncertainties are associated with the calculation of MDs, which we did not account for. Furthermore, the calculated MDs were only short-term ozone-based mortality damages; long-term effects, morbidity, and PM related damages were not included, and the analysis only covered short-term behaviors. Also, modeling electricity markets and emissions as discussed in chapter 6, did not include different types of electricity markets. Lastly, the case study presented in this thesis considered only a portion (albeit sizable portion) of the emission market; including all participants is an area of future research development. The following research topics are suggested for further investigation.

- **Uncertainty in exchange rates**

One of the main concerns about the implementation of location differentiated control policies is the uncertainty associated with calculation of exchange rates. Prediction of the exchange rates with a known degree of confidence can help the decision makers to set exchange rates with a safety factor to decrease the impact of uncertain estimation of exchange rates and prevent potential litigation. Using a Monte

Carlo simulation, the proposed DSS can be propagate the uncertainties from inputs to the estimated MDs in an effort to account for uncertainty in exchange rates.

- **Long-term cap-and-trade with exchange rate**

Our current work did not consider the long-term market behavior. Future work can potentially extend to prediction of the long-term behavior of firms in the market. The capital cost of different control technologies and the effect of future year permit allocation on the current year should be considered for the estimation of the long-term marginal abatement cost curves. Then, using the estimated long-term costs, firms can decide whether they should install a new control technology, or keep their old technology. The behavior of firms in the emission market can be predicted using the long-term marginal abatement cost curves and an optimization model.

- **Multi-objective (ozone and particulate matter) NO_x cap-and-trade system**

In the current work, NO_x emission trading was examined based on ozone formation potential of emitters. NO_x also contributes to the formation of PM particularly in certain regions in the U.S. and Canada and also at times other then the ozone season. The adjoint of CMAQ for PM is under development and can be used for development and examination of multi-objective exchange rate or social cost minimization systems.

- **Hypothetical Canada- U.S. emission cap-and-trade**

Given that there is significant cross-border pollutant transport between Canada and the U.S., especially for the states and provinces located near the border, there is a logical argument to be made for coordinated regulatory mechanisms that can benefit both countries in a cost-effective fashion. This is particularly true as regulations and approaches in Canada tend to be made in coordination with the U.S.

- **Real-time decision support system**

The proposed decision support system aimed to enable decision makers to compare the environmental performance of different NO_x control policies. Our approach did not aim to provide real-time decision support information. However, the proposed framework can be extended to a real-time decision support system under which an adjoint model combined with an air quality forecasting model predict the next-day source-specific marginal damages. Such information then can be used for setting the short-term exchange rates or emission fees.

REFERENCES

Abt Associates, I. (2009). *Environmental benefits mapping and analysis program (Version 3.0)*. NC.

Alberini, A., and Filippini, M. (2011). "Response of residential electricity demand to price: The effect of measurement error." *Energy Economics*, 33(5), 889–895.

Alhajeri, N. S., Donohoo, P., Stillwell, A. S., King, C. W., Webster, M. D., Webber, M. E., and Allen, D. T. (2011). "Using market-based dispatching with environmental price signals to reduce emissions and water use at power plants in the Texas grid." *Environmental Research Letters*, 6(4), 044018.

Anderson, S., and Francois, P. (1997). "Environmental cleanliness as a public good: welfare and policy implications of nonconvex preferences." *Journal of Environmental Economics and Management*, 34(3), 256–274.

Anenberg, S. C., Horowitz, L. W., Tong, D. Q., and West, J. J. (2010). "An estimate of the global burden of anthropogenic ozone and fine particulate matter on premature human mortality using atmospheric modeling." *Environmental Health Perspectives*, 118(9), 1189–95.

Atkinson, S. E., and Tietenberg, T. H. (1984). "Approaches for reaching ambient standards in non-attainment areas: financial burden and efficiency considerations." *Land Economics*, 60(2), 148–159.

Banzhaf, H. S. (2004). "Efficient emission fees in the US electricity sector." *Resource and Energy Economics*, 26(3), 317–341.

Becker, R. a. (2005). "Air pollution abatement costs under the Clean Air Act: evidence from the PACE survey." *Journal of Environmental Economics and Management*, 50(1), 144–169.

Bell, M. L., Mcdermott, A., Zeger, S. L., and Samet, J. M. (2004). "Ozone and short-term mortality in 95 US urban communities, 1987-2000." *Journal of the American Medical Association*, 292(19), 2372–2378.

Binkowski, F. S., Arunachalam, S., Adelman, Z., and Pinto, J. P. (2007). "Examining photolysis rates with a prototype online photolysis module in CMAQ." *Journal of Applied Meteorology and Climatology*, 46(8), 1252–1256.

Bridges, P., Doss, N., Gropp, W., Karrels, E., Lusk, E., and Skjellum, A. (1995). *Users ' Guide to mpich , a portable implementation of MPI*. Argonne National Laboratory, Lemont, IL.

Buhyoff, G., and Leuschner, W. (1978). "Estimating psychological disutility from damaged forest stands." *Forest Science*, 24(3), 424–432.

Burtraw, D., Palmer, K., Paul, A., Bharvirkar, R., and Paper, A. P. (2001). *Cost-effective reduction of NOx emissions from electricity generation*. Resources for the Future, Washington, DC.

Burtraw, D., Evans, D. A., Krupnick, A. J., Palmer, K., and Toth, R. (2005). *Economics of pollution trading for SO2 and NOx*. Resources for the Future, Washington, DC.

Burtraw, D., and Szambelan, S. J. (2009). "US emissions trading markets for SO2 and NOx." *Resources for the Future Discussion Paper*, Resources for the Future, Washington, DC.

Byrd, R., and Gilbert, J. (2000). "A trust region method based on interior point techniques for nonlinear programming." *Mathematical Programming*, 89(1), 149–185.

Byrd, R. H., Nocedal, J., and Waltz, R. A. (2006). "KNITRO: an integrated package for nonlinear optimization. " *Large-scale Nonlinear Optimization*, 83, 35-39.

Byun, D., and Schere, K. L. (2006). "Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling system." *Applied Mechanics Reviews*, 59(2), 51–77.

Carter, W. (2000). *Documentation of the saprc-99 chemical mechanism for VOC reactivity assessment volume 1 of 2. Final Report to California Air Resources Board.* Center for Environmental Research and Technology University of California, Riverside, CA.

CEP. (2007). *Sparse matrix operator kernel emissions modeling system (SMOKE) v2.4 User ' s manual*. Carolina Environmental Program, Chapel hill, NC.

Cho, S., Morris, R., McEachern, P., Shah, T., Johnson, J., and Nopmongcol, U. (2012). "Emission sources sensitivity study for ground-level ozone and PM2.5 due to oil sands development using air quality modelling system: Part II – Source apportionment modelling." *Atmospheric Environment*, 55, 542–556.

Cohan, D. S., Hakami, A., Hu, Y., and Russell, A. G. (2005). "Nonlinear response of ozone to emissions: Source apportionment and sensitivity analysis." *Environmental Science & Technology*, 39(17), 6739–6748.

Colella, P., and Woodward, P. R. (1984). "The piecewise parabolic method (PPM) for gas-dynamical simulations." *Journal of Computational Physics*, 54(1), 174–201.

Crocker, T. D., and Forster, B. a. (1981). "Decision problems in the control of acid precipitation: Nonconvexities and irreversibilities." *Journal of the Air Pollution Control Association*, 31(1), 31–37.

Cropper, M., and Arriaga-Salinas, A. (1980). "Inter-city wage differentials and the value of air quality." *Journal of Urban Economics*, 254, 236–254.

Daescu, D., Carmichael, G. R., and Sandu, A. (2000). "Adjoint implementation of Rosenbrock methods applied to variational data assimilation problems." *Journal of Computational Physics*, 165(2), 496–510.

Damian, V., Sandu, A., Damian, M., Potra, F., and Carmichael, G. R. (2002). "The kinetic preprocessor KPP-a software environment for solving chemical kinetics." *Computers & Chemical Engineering*, 26(11), 1567–1579.

Duncan, B. N., Yoshida, Y., Olson, J. R., Sillman, S., Martin, R. V., Lamsal, L., Hu, Y., Pickering, K. E., Retscher, C., Allen, D. J., and Crawford, J. H. (2010). "Application of OMI observations to a space-based indicator of NOx and VOC controls on surface ozone formation." *Atmospheric Environment*, 44(18), 2213–2223.

Dunker, A. M. (1981). "Efficient calculation of sensitivity coefficients for complex atmospheric models." *Atmospheric Environment*, 15(7), 1155–1161.

Dunker, A. M. (1984). "The decoupled direct method for calculating sensitivity coefficients in chemical kinetics." *The Journal of Chemical Physics*, 81(5), 2385.

Dunker, A. M., Yarwood, G., Ortmann, J. P., and Wilson, G. M. (2002). "Comparison of source apportionment and source sensitivity of ozone in a three-dimensional air quality model." *Environmental Science & Technology*, 36(13), 2953–2964.

EIA. (n.d.). *Electrictricity data browser, average retail price of electricity. US energy information administration, available online at <http://www.eia.gov/electricity/data/browser>*.

EIA. (2008). *Annual coal report, DOE/EIA-0584*. US Energy Information Administration, Office of Coal, Nuclear, Electric, and Alternate Fuels, Washington, DC.

EIA. (2010). *Annual energy outlook 2010 with projections to 2035*. DOE/EIA-0383, US Energy Information Administration, Office of Integrated Analysis and Forcasting, Washington, DC.

Elbern, H., Schmidt, H., and Ebel, A. (1997). "Variational data assimilation for tropospheric chemistry modeling." *Journal of Geophysical Research*, 102(D13), 15967–15985.

Elbern, H., Schmidt, H., Talagrand, O., and Ebel, a. (2000). "4D-variational data assimilation with an adjoint air quality model for emission analysis." *Environmental Modelling & Software*, 15(6-7), 539–548.

EPA. (n.d.). "Clean air market, data and maps. US Environmental Protection Agancey, Office of Air and Radiation, available online at <http://www.epa.gov/airmarkets/>."

EPA. (1991). *Guideline for regulatory application of the urban airshed model*. EPA, EPA-450/4-91-013, Research Triangle Park, NC.

EPA. (1998). *Regulatory impact analysis for the NBP*. EPA, Office of Air and Radiation, Washington DC.

EPA. (1999a). *Nitrogen oxides (NOx) why and how they are controlled*. EPA-456/F-99-006R, Research Triangle Park, NC.

EPA. (1999b). *The benefits and costs of the Clean Air Act, 1990 to 2010: EPA report to Congress*.

EPA. (2003). *Ozone transport commission NOx budget program 1999-2002 progress report*. EPA-430-R-03-900, Office of Air and Radiation, Washington, DC.

EPA. (2007). *Acid rain and related programs 2006*. EPA-430-R-07-011,Office of Air and Radiation, Washington, DC.

EPA. (2008). *NOx budget trading program: compliance and environmental results*. EPA-430-R-08-008, US Environmental Protection Agency, Office of Air and Radiation, Washington, DC.

EPA. (2010a). *Integrated planning model, chapter 5: emission control technologies*. US Environmental Protection Agency, Office of Air and Radiation, Washington, DC.

EPA. (2010b). *Guidelines for preparing economic analyses*. EPA 240-R-10-001, National Center for Environmental Economics, Office of Policy, Washington, DC.

EPA. (2013). *Clean air interstate rule, acid rain program, and former NOx budget trading program, 2012 progress report*. US Environmental Protection Agency, Office of Air and Radiation, Washington, DC.

Errico, R. (1997). "What is an adjoint model?" *Bulletin of the American Meteorological Society*, 78(11), 2577–2591.

European Commission. (1995). *ExternE: External Costs of Energy, Vol. 4: Oil and Gas*.

Fann, N., Fulcher, C. M., and Hubbell, B. J. (2009). "The influence of location, source, and emission type in estimates of the human health benefits of reducing a ton of air pollution." *Air Quality, Atmosphere, & Health*, 2(3), 169–176.

Fann, N., Lamson, A. D., Anenberg, S. C., Wesson, K., Risley, D., and Hubbell, B. J. (2012). "Estimating the national public health burden associated with exposure to ambient PM2.5 and ozone." *Risk Analysis*, 32(1), 81–95.

Färe, R., Grosskopf, S., Lovell, C. A. K., and Yaisawarng, S. (1993). "Derivation of shadow prices for undesirable outputs: a distance function approach." *The review of Economics and Statistics*, 75(2), 374–380.

Farrell, A., Carter, R., and Raufer, R. (1999). "The NOx Budget: market-based control of tropospheric ozone in the northeastern United States." *Resource and Energy Economics*, 21, 102–124.

Farrow, R. S., Schultz, M. T., Celikkol, P., and Houtven, G. L. Van. (2005). "Pollution trading in water quality limited areas : use of benefits assessment and cost-effective trading ratios." *Land Economics*, 81(2), 191–205.

Fiore, a. M., Dentener, F. J., Wild, O., Cuvelier, C., Schultz, M. G., Hess, P., Textor, C., Schulz, M., Doherty, R. M., Horowitz, L. W., MacKenzie, I. a., Sanderson, M. G., Shindell, D. T., Stevenson, D. S., Szopa, S., Van Dingenen, R., Zeng, G., Atherton, C., Bergmann, D., Bey, I., Carmichael, G., Collins, W. J., Duncan, B. N., Faluvegi, G., Folberth, G., Gauss, M., Gong, S., Hauglustaine, D., Holloway, T., Isaksen, I. S. a., Jacob, D. J., Jonson, J. E., Kaminski, J. W., Keating, T. J., Lupu, a., Marmer, E., Montanaro, V., Park, R. J., Pitari, G., Pringle, K. J., Pyle, J. a., Schroeder, S., Vivanco, M. G., Wind, P., Wojcik, G., Wu, S., and Zuber, a. (2009). "Multimodel estimates of intercontinental source-receptor relationships for ozone pollution." *Journal of Geophysical Research*, 114(D4), 1–21.

Fisher, M., and Lary, D. (1995). "Lagrangian four-dimensional variational data assimilation of chemical species." *Quarterly Journal of the Royal Meteorological Society*, 121, 1681–1704.

Førsund, F., and Nævdal, E. (1998). "Efficiency gains under exchange-rate emission trading." *Environmental and Resource Economics*, 12(4), 403–423.

Fowlie, M. (2010). "Emissions trading, electricity restructuring, and investment in pollution abatement." *The American Economic Review*, 100(3), 837–869.

Fraas, A., and Lutter, R. (2011). "Do some NOx emissions have negative environmental damages? evidence and implications for policy." *Environmental Science & Technology*, 45, 7613–7614.

Goodstein, E. S. (1995). *Economics and the environment*. Prentic-Hall, Inc., Englewood Cliffs, New Jersy, 575.

Hakami, A., Odman, M. T., and Russell, A. G. (2003). "High-order, direct sensitivity analysis of multidimensional air quality models." *Environmental Science & Technology*, 37(11), 2442–2452.

Hakami, A., Bergin, M. S., and Russell, A. G. (2004). "Ozone formation potential of organic compounds in the eastern United States: a comparison of episodes, inventories, and domains." *Environmental Science & Technology*, 38(24), 6748–6759.

Hakami, A., Odman, M. T., and Russell, A. G. (2004). "Nonlinearity in atmospheric response: A direct sensitivity analysis approach." *Journal of Geophysical Research*, 109(D15303), 1–12.

Hakami, A., Seinfeld, J. H., Chai, T., Tang, Y., Carmichael, G. R., and Sandu, A. (2006). "Adjoint sensitivity analysis of ozone nonattainment over the continental United States." *Environmental Science & Technology*, 40(12), 3855–64.

Hakami, A., Henze, D. K., Seinfeld, J. H., Singh, K., Sandu, A., Kim, S., Byun, D., and Li, Q. (2007). "The adjoint of CMAQ." *Environmental Science & Technology*, 41(22), 7807–17.

Hall, D. C., and Hall, J. V. (1997). "Estimating the benefits of emissions reductions in complex atmospheres." *International Journal of Global Energy Issues*, 9(4), 286–298.

Henze, D. K., Hakami, A., and Seinfeld, J. H. (2007). "Development of the adjoint of GEOS-Chem." *Atmospheric Chemistry & Physics*, 7, 2413–2433.

Hung, M. F., and Shaw, D. (2005). "A trading-ratio system for trading water pollution discharge permits." *Journal of Environmental Economics and Management*, 49, 83–102.

Jacob, D. J., Logan, J. A., and Murti, P. P. (1999). "Effect of rising Asian emissions on surface ozone in the United States." *Geophysical Research Letters*, Citeseer, 26(14), 2175–2178.

Jerrett, M., Burnett, R. T., Pope, C. A., Ito, K., Thurston, G., Krewski, D., Shi, Y., Calle, E., and Thun, M. (2009). "Long-term ozone exposure and mortality." *The New England Journal of Medicine*, 360(11), 1085–1095.

Klaassen, G. A. J., Førsund, F. R., and Amann, M. (1994). "Emission trading in Europe with an exchange rate." *Environmental and Resource Economics*, 4(4), 305–330.

Kolstad, C. D., and Turnovsky, M. H. L. (1998). "Cost functions and nonlinear prices: estimating a technology with quality-differentiated inputs." *Review of Economics and Statistics*, 80(3), 444–453.

Krupnick, A. J., Oates, W. E., and Van De Verg, E. (1983). "On marketable air-pollution permits: The case for a system of pollution offsets." *Journal of Environmental Economics and Management*, 10, 233–247.

Krupnick, A. J. (1986). "Costs of alternative policies for the control of nitrogen dioxide in Baltimore." *Journal of Environmental Economics and Management*, 13, 189–197.

Krupnick, A. J., Mcconnell, V., Cannon, M., Stoessell, T., and Batz, M. (2000). *Cost-effective NOx control in the eastern United States*. Resources for the Future, Washington, DC.

Levy, J. I., Hammitt, J. K., Yanagisawa, Y., and Spengler, J. D. (1999). "Development of a new damage function model for power plants: methodology and applications." *Environmental Science & Technology*, 33(24), 4364–4372.

Levy, J. I., Baxter, L. K., and Schwartz, J. (2009). "Uncertainty and variability in health-related damages from coal-fired power plants in the United States." *Risk analysis*, 29(7), 1000–14.

Luecken, D. J., Phillips, S., Sarwar, G., and Jang, C. (2008). "Effects of using the CB05 vs. SAPRC99 vs. CB4 chemical mechanism on model predictions: Ozone and gas-phase photochemical precursor concentrations." *Atmospheric Environment*, 42(23), 5805–5820.

Mallet, V., and Sportisse, B. (2005). "A comprehensive study of ozone sensitivity with respect to emissions over Europe with a chemistry-transport model." *Journal of Geophysical Research*, 110(D22), 1–15.

Marchuk, G. I. (1974). *Numerical methods in weather prediction*. Academic Press, 227.

Martien, P. T., and Harley, R. a. (2006). "Adjoint sensitivity analysis for a three-dimensional photochemical model: application to Southern California." *Environmental Science & Technology*, 40(13), 4200–10.

Martin, K. C., Joskow, P. L., Ellerman, A. D., and others. (2007). *Time and location differentiated NOx control in competitive electricity markets using cap-and-trade mechanisms*. MIT Center for Energy and Environmental Policy Research, MI.

Martin, K. C. (2008). *Implementing a time-and location-differentiated cap-and-trade program: flexible nitrogen oxide abatement from power plant in the eastern united states.* PhD thesis, Department of Technology, Management, and Policy, Massachusetts Institute of Technology, MI.

Mauzerall, D., and Wang, X. (2001). "Protecting agricultural crops from the effects of tropospheric ozone exposure: reconciling science and standard setting in the United States, Europe, and Asia." *Annual Review of Energy and the Environment*, 26, 237–268.

Mauzerall, D., Sultan, B., Kim, N., and Bradford, D. (2005). "NOx emissions from large point sources: variability in ozone production, resulting health damages and economic costs." *Atmospheric Environment*, 39(16), 2851–2866.

McGartland, A. M., and Oates, W. E. (1985). "Marketable permits for the prevention of environmental deterioration." *Environmental Economics and Management*, 12, 207–228.

McKitrick, R. (1999). "A derivation of the marginal abatement cost curve." *Journal of Environmental Economics and Management*, 37(3), 306–314.

Menut, L. (2003). "Adjoint modeling for atmospheric pollution process sensitivity at regional scale." *Journal of Geophysical Research*, 108(D17), 1–17.

Mesbah, S., Hakami, A., and Schott, S. (2012). "Improving NOx cap-and-trade system with adjoint-based emission exchange rates." *Environmental Science & Technology*, 46(21), 11905–12.

Mesbah, S. M., Hakami, A., and Schott, S. (2013). "Optimal ozone reduction policy design using adjoint-based NOx marginal damage information." *Environmental Science & Technology*, 47(23), 13528–35.

Miller, B. G. (2011). *Clean coal engineering technology*. Butterworth-Heinemann, 696.

Montgomery, W. D. (1972). “Market in licences and efficient pollution control program.” *Journal of Economic Theory*, 5, 395–418.

Müller, J. F., Stavrakou, T., and others. (2005). “Inversion of CO and NO x emissions using the adjoint of the IMAGES model.” *Atmospheric Chemistry and Physics*, 5(5), 1157–1186.

Muller, N. Z., and Mendelsohn, R. (2007). “Measuring the damages of air pollution in the United States.” *Journal of Environmental Economics and Management*, 54, 1–14.

Muller, N. Z., and Mendelsohn, R. (2009). “Efficient pollution regulation : getting the prices right.” *American Economic Review*, 99(5), 1714–1739.

Muller, N. Z. (2011). “Linking policy to statistical uncertainty in air pollution damages.” *The BE Journal of Economic Analysis & Policy Contributions*, 11(1).

Muller, N. Z., Mendelsohn, R., and Nordhaus, W. (2011). “Environmental accounting for pollution in the United States economy.” *The American Economic Review*, 101(August), 1649–1675.

NCEP. (2005). “National center for environmental predictions, GRIB data.” <http://www.nco.ncep.noaa.gov/pmb/docs/grib2/grib2_doc.shtml>.

Nester, K., Panitz, H. J., and others. (2005). “Sensitivity analysis by the adjoint chemistry transport model DRAIS for an episode in the Berlin ozone (BERLIOZ) experiment.” *Atmospheric Chemistry and Physics Discussions*, 5(5), 8715–8754.

Nobel, C. E., Mcdonald-buller, E. C., Kimura, Y., and Allen, D. T. (2001). “Policy analysis accounting for spatial variation of ozone productivity in NOx emission trading.” *Environmental Science & Technology*, 35(22), 4397–4407.

Pacsi, A. P., Alhajeri, N. S., Webster, M. D., Webber, M. E., and Allen, D. T. (2013). “Changing the spatial location of electricity generation to increase water availability in areas with drought: a feasibility study and quantification of air quality impacts in Texas.” *Environmental Research Letters*, 8(3), 035029.

Pappin, A. J., and Hakami, A. (2013). “Source attribution of health benefits from air pollution abatement in Canada and the United States: an adjoint sensitivity analysis.” *Environmental Health Perspectives*, 121(5), 572–9.

Pechan, E.H., and Associates, I. (1997). *NOx, The emission reduction and cost analysis model for (ERCAM-NOx)*. EPA-68-D3-0035, Ozone Policy and Strategies Group, Research Triangle Park, NC.

Perman, R., Ma, Y., Common, M., Maddison, D., and McGilvray, J. (2011). *Natural Resource and Environmental Economics*. Pearson Education Limited, Harlow, 712.

Quélo, D., Mallet, V., and Sportisse, B. (2005). "Inverse modeling of NOx emissions at regional scale over northern France: Preliminary investigation of the second-order sensitivity." *Journal of Geophysical Research*, 110(D24), 1–18.

Randall, A., Ives, B., and Eastman, C. (1974). "Bidding games for valuation of aesthetic environmental improvements." *Journal of Environmental Economics and Management*, 1, 132–149.

Repetto, R. (1987). "The policy implications of non-convex environmental damages : a smog control case study." *Journal of Environmental Economics and Management*, 14, 13–29.

Ridge, O. (1994). *Estimating fuel cycle externalities : analytical methods and issues*. Washington, DC.

Rowe, R. D. ., Lang, C. M. ., Chestnut, L. G. ., Latimer, D. A. ., Rae, D. A. ., Bernow, S. M. ., and White, D. E. (1995). *The New York electricity externality study volume I: introduction and methods*. Empire State Electric Energy Research Corporation, New York.

Sandu, A., Daescu, D. N., Carmichael, G. R., and Chai, T. (2005). "Adjoint sensitivity analysis of regional air quality models." *Journal of Computational Physics*, 204(1), 222–252.

Schultze, W., and Brookshire, D. (1983). "Economic benefits of preserving visibility in the national parklands of the southwest." *Natural Resources Journal*, 23, 149–173.

Seinfeld, J. H., and Pandis, S. N. (2006). *Atmospheric chemistry and physics: from air pollution to climate change*. John Wiley and Sons, New York, 1232.

Seskin, E. P., Anderson, R. J., Reid, R., Christofaro, A., Clark, T., Hoffman, J., Ingle, B., Luken, S., and Smith, W. (1983). "An empirical analysis of economic strategies for controlling air pollution." *Journal of Environmental Economics and Management*, 112–124.

Skamarock, W., Klemp, J., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W., and Powers, J. G. (2005). *A description of the advanced research WRF version 2*. NCAR Tech Notes, NCAR/TN-468+STR; National Center for Atmospheric Research: Boulder, CO.

Sloss, L. L., Hjalmassson, A. K., Hermine, N. S., Campbell, L. M., Stone, D. K., Shareef, G. S., Emmel, T., Maibodi, M., Livengood, C. D., and Markussen, J. (1992). *Nitrogen oxides control technology fact book*. Noyes Data Corporation, New Jersy, 635.

Starrett, D. (1972). "Fundamental nonconvexities in the theory of externalities." *Journal of Economic Theory*, 120(3), 253–290.

Sun, L., Webster, M., Mcgaughey, G., Mcdonald-buller, E. C., Thompson, T., Prinn, R., Ellerman, A. D., and Allen, D. T. (2012). "Flexible NOx abatement from power plants in the eastern United States." *Environmental Science & Technology*, 46, 5607–5615.

Thompson, T. M., King, C. W., Allen, D. T., and Webber, M. E. (2011). "Air quality impacts of plug-in hybrid electric vehicles in Texas: evaluating three battery charging scenarios." *Environmental Research Letters*, 6(2), 024004.

Thompson, T. M., and Selin, N. E. (2012). "Influence of air quality model resolution on uncertainty associated with health impacts." *Atmospheric Chemistry and Physics*, 12(20), 9753–9762.

Tietenberg, T. H. (1985). *Emission trading, an excersie in reforming pollution policy*. Resources for the Future, Washington, DC.

Tietenberg, T. (1995). "Tradeable permits for pollution control when emission location matters: What have we learned?" *Environmental and Resource Economics*, 5, 95–113.

Tong, D. Q., and Mauzerall, D. L. (2006). "Spatial variability of summertime tropospheric ozone over the continental United States: Implications of an evaluation of the CMAQ model." *Atmospheric Environment*, 40(17), 3041–3056.

Tong, D. Q., Muller, N. Z., Mauzerall, D. L., and Mendelsohn, R. O. (2006). "Integrated assessment of the spatial variability of ozone impacts from emissions of nitrogen oxides." *Environmental Science & Technology*, 40(5), 1395–1400.

Vautard, R., Beekmann, M., and Menut, L. (2000). "Applications of adjoint modelling in atmospheric chemistry: sensitivity and inverse modelling." *Environmental Modelling & Software*, 15(6-7), 703–709.

Vijay, S., DeCarolis, J. F., and Srivastava, R. K. (2010). "A bottom-up method to develop pollution abatement cost curves for coal-fired utility boilers." *Energy Policy*, 38(5), 2255–2261.

Vukićević, T., and Hess, P. (2000). "Analysis of tropospheric transport in the Pacific Basin using the adjoint technique." *Journal of Geophysical Research*, 105(D6), 7213–7230.

Waltz, R., and Plantenga, T. (2010). *KNITRO user's manual*. Ziena Optimization Inc., Evanston, IL.

Wang, K. Y., Lary, D., Shallcross, D., Hall, S., and Pyle, J. (2001). "A review on the use of the adjoint method in four-dimensional atmospheric-chemistry data assimilation." *Quarterly Journal of the Royal Meteorological Society*, 127(576), 2181–2204.

Wang, L., Thompson, T., Mcdonald-buller, E. C., Webb, A., and Allen, D. T. (2007). "Photochemical modeling of emissions trading of highly reactive volatile organic compounds in Houston, Texas. 1. reactivity based trading and potential for ozone hot spot formation." *Environmental Science & Technology*, 41(7), 2095–2102.

Wang, L., Mcdonald-buller, E., and Allen, D. T. (2007). "Photochemical modeling of emissions trading of highly reactive volatile organic compounds in Houston, Texas. 2. Incorporation of chlorine emissions." *Environmental Science & Technology*, 41(7), 2103–7.

Wanger, F., and Schopp, W. (2007). *Comparison of the RAIN emission control cost curves for air pollutants with emission control costs computed by the GAINS model*. IR-07-008, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Warwick, W. M. (2002). *A primer on electric utilities, deregulation, and restructuring of US electricity markets*. prepared for the US Department of Energy, Federal Energy Management Mrogram, Office of Eneregy and Renewable Energy, Washington, DC.

Xu, J., Kochanek, K. D., Murphy, S. L., and Tejada-Vera, B. (2010). *National vital statistics reports births: Final data for 2007*. National Center for Health Statistics, Hyattsville, MD.

Yang, Y.-J., Wilkinson, J. G., and Russell, A. G. (1997). "Fast, direct sensitivity analysis of multidimensional photochemical models." *Environmental Science & Technology*, 31(10), 2859–2868.

Yarwood, G., Rao, S., Yocke, M., and Whitten, G. (2005). *Updates to the carbon bond chemical mechanism: CB05. Final report to the US EPA*, RT-04-00675, Research Triangle Park, NC.

Zanobetti, A., and Schwartz, J. (2008). “Mortality displacement in the association of ozone with mortality: an analysis of 48 cities in the United States.” *American Journal of Respiratory and Critical Care Medicine*, 177(2), 184–9.

Ziadi, A., and Kumar, S. (1995). “Combustion control techniques to reduce NOx emissions from industrial boliers: what’s available and how to compare cost effectiveness.” *Pulp and Paper Canada*, 96(6), 84–88.