# DIRECT SENSITIVITY TECHNIQUES IN REGIONAL AIR QUAILTY MODELS: DEVELOPMENT AND APPLICATION

A Dissertation
Presented to
The Academic Faculty

by

Wenxian Zhang

In Partial Fulfillment
of the Requirements for the Degree
Doctoral Philosophy in the
School of Earth and Atmospheric Sciences

Georgia Institute of Technology
December 2013

# DIRECT SENSITIVITY TECHNIQUES IN REGIONAL AIR QUAILTY MODELS: DEVELOPMENT AND APPLICATION

Approved by:


Dr. Armistead G. Russell, Advisor
School of Civil & Environmental Engineering
*Georgia Institute of Technology*

Dr. Yuhang Wang
School of Earth &Atmospheric Sciences
*Georgia Institute of Technology*

Dr. Michael H. Bergin
School of Civil & Environmental Engineering and School of Earth & Atmospheric Sciences
*Georgia Institute of Technology*

Dr. Rodney Weber
School of Earth & Atmospheric Sciences
*Georgia Institute of Technology*

Dr. Athanasios Nenes
School of Chemical &Biomolecular Engineering and School of Earth & Atmospheric Sciences
*Georgia Institute of Technology*


Date Approved: September 6, 2013

*To Summer*

# ACKNOWLEDGEMENTS

I would like to thank the many people who have made this dissertation possible. First and foremost, I sincerely thank my advisor, Prof. Ted Russell, for his patient guidance and generous support over the past five years. I am deeply grateful to him for giving me such a valuable chance to work in his research group, for providing me numerous opportunities to work on interesting projects, and for his advice and support in determining a future career path. I would like to thank Prof. Thanos Nenes for bringing me into a number of interesting projects and for his insightful comments on my work. I would like to thank my Ph.D. thesis committee members, Prof. Michael Bergin, Prof. Yuhang Wang, and Prof. Rodney Weber, for taking time out from their busy schedules and providing valuable feedback on my research.

I am lucky to have the members of the Russell group as my friends and colleagues. I sincerely thank Dr. Yongtao Hu for sharing his broad knowledge and valuable experience, and for being patient and tremendously helpful with all issues related to air quality modeling. I would like to thank Dr. Shannon Capps and Dr. Heather Holmes for providing thoughtful comments on my research and being helpful in many ways. I would like to thank Peng Liu, Marcus Trail, Aika Davis, Sivaraman Balachandran, Fernando Garcia, Boris Galvis, Sunni Ivey, Mariel Friberg, , Emily Landtrip , Dr. Gretchen Goldman, Dr. Burcak Kaynak, Dr. Jaemeen Baek, Dr. Bo Yan, Dr. Di Tian, Dr. Sergey Napelenok, Dr. Soonchul Kwon, Dr. K.J. Liao, Dr. Farhan Akhtar, Dr. Talat Odman, and the many others for their important contributions to this dissertation.

I am also lucky to have had the opportunity to work with a number of excellent scientists and researchers. I would like to thank Dr. Erin Tullos from Phillips 66 for her continuous support on the study of flare emissions impact, which made possible the fourth chapter of this dissertation and for her valuable comments on my research as well. I would like to thank Prof. Valerie Thomas and her student Paul Kerl for organizing an excellent team and making possible the interesting project as well as the fifth chapter of this dissertation. I would like to thank Barry Exum from TCEQ for providing the valuable data for flare modeling and being patient with all my questions. I would also like to express my gratitude to Prof. Robert Dickinson for mentoring me in the beginning of my Ph.D. studies and giving me insightful advice on research.

This dissertation could never have been possible without the unconditional love and support of my family. I thank my husband, Peng, who has companied me through the ups and downs in this journey and fills my life with joy and fun. I thank my parents, who are always there when I need their help. I thank my daughter Summer, who brings me so much joy and courage.

Finally, I would like to acknowledge funding sources for this dissertation which include ConocoPhillips, Phillips 66, and the U.S. Environmental Protection Agency (RD83479901).

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# SUMMARY

Regional air quality models have been widely used in air quality management. The primary goal of an air quality model is to understand the relationship between trace contaminant levels and chemical and physical processes. Such relationship can be extracted from an air quality model by direct sensitivity analysis. This dissertation outlines the development and applications of an advanced sensitivity techniques, the decoupled direct method (DDM), in a regional air quality model. Given the complex chemical and physical processes involved in the formation of secondary particulate matter, this dissertation implemented high-order DDM (HDDM) sensitivity techniques for particulate matter in the Community Multiscale Air Quality (CMAQ) model and quantified spatially and temporally variable responses of particulate matter species to precursor emissions and the interaction of them. The implementation of the high-order sensitivity technique provides important insights in source apportionment of fine particulate matter and the trade-offs that may be involved in multi-species control strategies for secondary particulate matter.

Based on the high-order DDM sensitivity techniques, this dissertation then established a reduced form model of CMAQ, which keeps the pollutant-emissions response. The reduced form model is computationally efficient and has a flexible formulation. This dissertation demonstrates the applications of the reduced form model in the following three aspects: 1) Uncertainty analysis of air quality model - The reduced form model is applied to quantify the uncertainty of simulated particulate matter in the presence of uncertain emissions rates. 2) Emission assessment - The reduced form model is applied to assess the emissions impact of flares at different operating modes. 3)

Optimization of control strategies - The reduced form model is integrated into an electricity generation planning model to find the optimized solution by taking into account air quality-related health costs.

# CHAPTER 1

# INTRODUCTION

## 1.1. Context and Motivation

Particulate matter (PM) control is perhaps the most challenging aspect of current air quality management. PM is a complex mixture of extremely small particles and liquid droplets that are varying in size, chemical composition, and origin. These suspended particles can be inhaled by people and have been linked to a wide range of serious health effects, including premature death, heart attacks, acute bronchitis, and aggravated asthma among children (Sarnat et al., 2006; Peel et al., 2005; Dockery et al., 1993). They are also known to have adverse effects on the environment by forming acid deposition (Baptista-Neto et al., 2006) and impairing visibility (Latha and Badarinath, 2005). They can affect climate directly by absorbing and scattering of shortwave and longwave radiation and indirectly by acting as cloud condensation nuclei (CCN) to form droplets that are more effective in absorbing and scattering of radiation (IPCC, 2001). Due to its adverse impact on public health and public welfare, PM is regulated by the United States Environment Protection Agency (U.S. EPA) as one of the criteria pollutants under the National Ambient Air Quality Standards (NAAQS). PM regulation is usually based on its size because its potential to harm people's health depends on its ability to penetrate the lungs. PM pollution includes "inhalable coarse particles", those with aerodynamic diameters larger than 2.5 micrometers and smaller than 10 micrometers ($PM_{10}$) and "fine particles", those with aerodynamic diameters smaller than 2.5 micrometers ($PM_{2.5}$).The annual health (primary) standard for $PM_{2.5}$ has recently been tightened from 15 $\mu g\ m^{-3}$ to 12 $\mu g\ m^{-3}$, the health benefits of which are estimated to range from $4 billion to over $9 billion

per year. While the revised standard is expected to bring significant health benefits, it adds challenges to the already-complex task of PM management. Based on monitoring data from 2008 to 2010, 18 additional counties that are not currently designated as nonattainment don't meet the revised annual $PM_{2.5}$ standard (Figure 1.1).


Figure 1.1. Counties not meeting the January 2013 revised primary annual $PM_{2.5}$ NAAQS based on 2008-2009 monitoring data. (Source: U.S. EPA, http://www.epa.gov/pm/2012/mapb.pdf).

The complexity of PM formation has made the development of control strategies more difficult. PM can be directly emitted from primary sources such as elemental carbon (EC) from on-road mobile sources. PM can also be formed secondarily by the chemical transformation of reactive gaseous emissions such as $SO_2$, $NH_3$, $NO_x$, and VOC. Secondary PM constitutes a significant fraction of the PM mass. For example, most of sulfate in the atmosphere is formed by oxidation of $SO_2$ (Figure 1.2). In the daytime, $SO_2$

can be oxidized by the hydroxyl radical, OH, in gas phase. It can also be oxidized in the aqueous phase by a variety of oxidants such as hydrogen peroxide and ozone. Likewise, the formation of nitrate usually occurs through the oxidation of NO and $NO_2$ ($NO_x$) either during the daytime with OH or during the night with ozone and water (Figure 1.2). $NO_x$ and $SO_2$ tend to compete for oxidants, so the formation of sulfate and nitrate interact with each other. Another example is the formation of aerosol ammonium. The available ammonia in the atmosphere preferentially neutralizes $H_2SO_4$ and forms ammonium sulfate aerosol. Nitrate can also react with ammonia and form ammonium nitrate, depending on the availability of ammonia and favorable conditions of low temperature and high relative humidity. Because the interaction among secondary sulfate, nitrate , and ammonium aerosols, the increasing nitrate aerosol corresponding to the reduction in $SO_2$ emission is a significant issue for designing PM control strategies. Another issue in PM management is that PM shares precursors with other air pollutants such as ozone. VOC and $NO_x$ are precursors of both ozone and a fraction of PM (nitrate and secondary organic aerosol). Meng et al. (1997) found that in Riverside California, a 25 percent reduction in VOC emissions resulted in a 19 percent reduction in the peak ozone, but an 18 percent increase of $PM_{2.5}$ level. Organic compounds constitute about one-fifth to one-half the average annual $PM_{2.5}$ (NARSTO, 2004) and also interact with other $PM_{2.5}$ species through common precursors such as $NO_x$ and VOC. For instance, reduction in VOC may lead to an increase in $PM_{2.5}$ concentrations in that the available $NO_2$, instead of being converted to PAN, is converted to $HNO_3$, which reacts with available ammonia to form ammonium nitrate. At the same time, ozone concentrations are predicted to decrease with decreasing VOC (Meng et al., 1997).

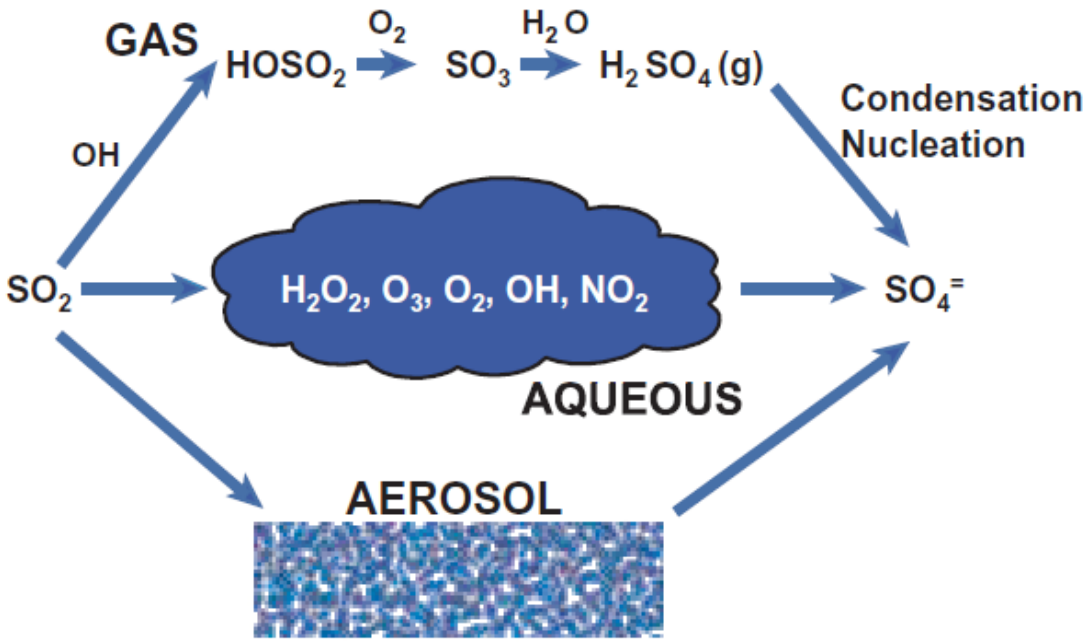


Figure 1.2. Schematic of the three pathways (reactions in gaseous phase, cloud, and condensed phase) for the formation of sulfate aerosol (NARSTO, 2004).

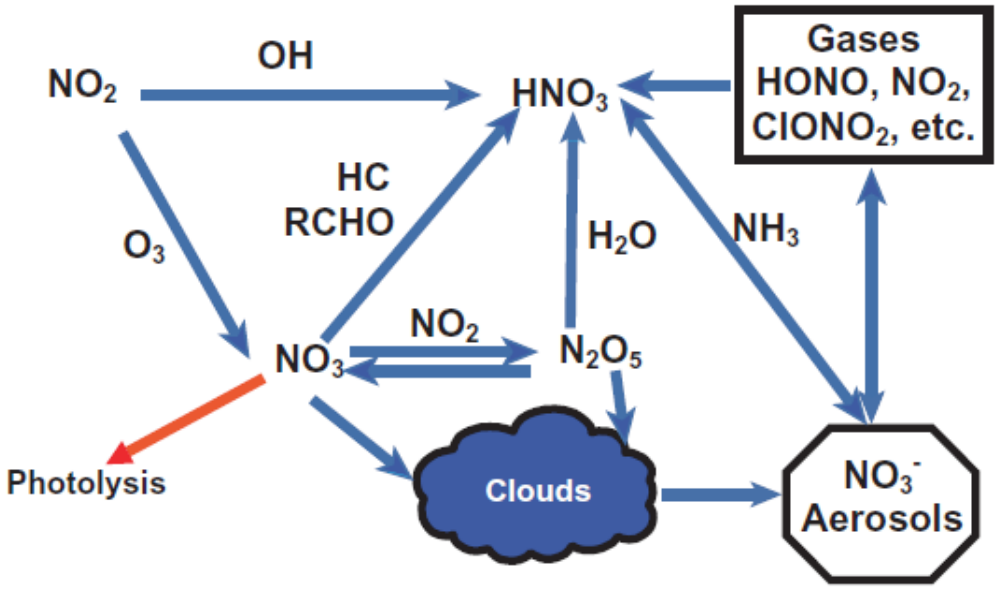


Figure 1.3. Schematic of the formation of $HNO_3$ and nitrate aerosol (NARSTO, 2004).

The above examples illustrate the intricate coupling among the different species of $PM_{2.5}$ and between $PM_{2.5}$ and ozone. Designing a control strategy for both $PM_{2.5}$ and ozone requires considering the coupling effects among the different species. One control strategy that is efficient at reducing a certain $PM_{2.5}$ species may cause an increase in another species. The optimal strategy for ozone controls may increase $PM_{2.5}$ levels. The multi-pollutant response to various emission control strategies is often investigated using chemical transport models (CTM). The Community Multiscale Air Quality (CMAQ) model (Byun et al., 2006) is one of the most widely used CTM in air quality management. It has the state-of-science representation of the formation of ozone, photochemical oxidants, acid deposition, and fine and coarse particulate matter and is suitable to be used as the base model for sensitivity analysis of pollutant response to emission changes.

### 1.2. Sensitivity Analysis

For policy-oriented applications, it is important to consider the response of PM concentrations to changes in emission rates. However, emission source impacts are difficult to directly measure. Such pollutant-emission responses can be quantified using model-based sensitivity.

In a CTM, a straightforward way to conduct sensitivity analysis is to calculate the differences between model outputs that are due to a certain amount of perturbation in emission rates (Figure 1.2). The model is run once in its base state, then it is re-run with emissions from a source perturbed up or down (potentially totally removed). This method is called the brute force method. Since it is a numerical approximation, its accuracy may

be affected by the perturbation size (Hakami et al. 2004). It is also subject to noise at the discontinuous point of a response surface. Studies also indicate that the brute force method exhibits more noisy behavior with increased orders of sensitivity (Hakami et al., 2003). Besides the accuracy issue, another issue is that brute force becomes computationally expensive when sensitivities to a large number of emission sources must be calculated.

The Decoupled Direct Method in Three Dimensions (DDM-3D) (Dunker, 1981; Yang et al., 1997) enables simultaneous calculation of sensitivities to multiple emission rates or other parameters within a single model simulation. The extension of DDM-3D to compute higher-order sensitivities of ozone (Hakami et al., 2003) has enabled the exploration of the nonlinearity of ozone response to precursors. One method that has been demonstrated for simulating ozone response to a wide range of perturbations has been to include first- and second-order sensitivity coefficients in Taylor expansions (Hakami et al., 2003). The use of higher-order DDM (HDDM) in scientific investigations of pollutant formation has been demonstrated in the analysis of ozone nonlinearity in Georgia and Texas (Cohan et al., 2005, and Xiao et al., 2010). The use of higher-order DDM (HDDM) in informing policy development has been demonstrated by Simon et al., as an approach to evaluate the ozone NAAQS (2012).

HDDM can be used to construct a reduced form model of the underlying air quality model. The reduced form model has a simple formulation that directly relates pollutant response to the change in emissions. The information on transport and chemical reaction is retained in sensitivity coefficients in the formulation. The reduced form model has flexibility in selecting the sensitivity parameters. It has been applied to quantify the

uncertainties in modeled ozone concentrations that are caused by uncertainties in emission rates, reaction rates, and model structures (e.g., Cohan et al. 2010; Tian et al., 2010; Pinder et al., 2011). Uncertainty analysis helps us understand the model's credibility and provides insight for model improvement. The reduced form model has also been incorporated into a cost optimization model to determine the least-cost approach to achieve attainment of the NAAQS. An example can be found in the study by Cohan et al. (2010), which investigated ways to optimize control measures to achieve attainment for ozone in Macon, Georgia.

DDM-3D for PM has been implemented in current chemical transport models (e.g., Napelenok et al., 2006; Koo et al., 2007). It enables assessment of source impact on different species of PM (Napelenok, 2006). It also has the advantage of capturing spatial and temporal variability of source impacts, which has been shown to be essential in health studies relating public health issues to emission source categories. Recent studies have combined DDM-3D for PM with a receptor model to generate an ensemble approach to conduct source apportionment studies for PM (Hu et al., 2013, Ivey et al., 2013). The results are intended to be used in epidemiological studies to explore the impact of emission source categories on human health. DDM-3D for PM can also be used to construct reduced form models to quickly estimate the response in PM species due to the change in precursor emissions. The power of reduced form model in evaluating control strategies and informing policy development may be of increasing attractiveness as air quality managers need to deal with the more stringent NAAQS for $PM_{2.5}$,. However, given the complex chemical and physical processes regarding the formation of PM as discussed in Section 1.1, it is of interest to investigate the nonlinearity in PM

formation and how a nonlinear response would affect how accurately PM response can be predicted by a reduced form model. It is also of value to investigate uncertainty in model simulations of $PM_{2.5}$ concentration and response to better utilize and improve current CTM.

### 1.3. Scope of This Work

Given the link between $PM_{2.5}$ and human health, a clear understanding of source impacts on $PM_{2.5}$ is important for developing effective control strategies. This dissertation demonstrates how the sources of PM can be determined using direct sensitivity analysis in a chemical transport model and explores the nonlinear response of $PM_{2.5}$ to its precursors. This dissertation also demonstrates how sensitivity techniques can be used to assess uncertainty in model simulation and to inform the cost-optimization of pollutant control strategies.

Specifically, the chapters are organized as follows.

- **Chapter 2, "Development of high-order direct sensitivity analysis of particulate matter in regional air quality models,"** implements high-order DDM sensitivity techniques for particulate matter in CMAQ, enabling the investigation of nonlinear responses of particulate matter to precursor emissions. The results are then compared to results attained by using the traditional brute form method. While the two methods are generally in agreement, DDM provides more accurate and stable values at the locations where the brute force method is subject to noise. This chapter also describes a case study that was conducted in Southeastern Texas for a summertime episode and then quantifies the source contributions to $PM_{2.5}$ from various precursor emissions through both linear and nonlinear responses are

quantified. This chapter also investigated how first-order sensitivity of $PM_{2.5}$ changes with emissions rates of various precursor gases.

- **Chapter 3, "Quantification of uncertainty in particulate matter simulation in the presence of uncertain emissions using CMAQ/HDDM-PM,"** explains the construction of a reduced form model to estimate the emission-associated model uncertainty in $PM_{2.5}$ simulation in an efficient manner. The reduced form model is based on the high-order DDM sensitivity analysis developed in Chapter 2. Uncertainty in a 24-hour average $PM_{2.5}$ simulation is the focus due to its policy relevance. The uncertainty ranges have been applied to five monitoring sites in the Houston area and the bias between simulation and observation of 24-hour average $PM_{2.5}$ is discussed.
- **Chapter 4, "Assessing the impact of flare emissions on air quality at variable operating conditions,"** explores the impact of flare VOC emissions on the formation of ozone in the Houston Ship Channel. Temporally variable flare VOC emissions are incorporated into the CTM simulation. This chapter estimates the difference between the ozone impact under a constant destruction and removal rate (DRE) (98%) and DRE that is estimated in a more realistic way. An important component of this chapter is applying DDM sensitivity analysis and reduced form model to efficiently assess flare impact assessment.
- **Chapter 5, "Integrating air quality-related health effects in a multi-decade electricity capacity planning model,"** demonstrates how an air quality model can be integrated into an electricity capacity planning model to account for the impact of the emissions from power plants on air quality and thus the associated health effects.

This chapter describes a case study that was conducted in Georgia and estimates the increase in the cost associated with the health impact of $PM_{2.5}$ per kilowatt-hour (kWh) of electricity generated by power plants at various locations. The integrated model provides a cost-optimized planning on power generation with air quality-related health costs taken into consider.

- **Chapter 6** provides conclusions and recommendations for future studies.

# CHAPTER 2

# DEVELOPMENT OF HIGH-ORDER DIRECT SENSITIVITY ANALYSIS OF PARTICULATE MATTER IN REGIONAL AIR QUALITY MODELS*

## 2.1. Introduction

Airborne particulate matter (PM), or aerosol, is a major pollutant in the atmosphere. Studies have shown that PM impairs visibility (Watson 2002), may cause harmful effects on ecosystems (Galloway et al., 2004), and affects human health (e.g., Zanobetti et al., 2000; Kaiser, 2005). In response, control strategies are designed to lower the concentrations of anthropogenic PM in the atmosphere (U.S. EPA, 2004). Historically, multiple air quality model simulations using different sets of emissions have been used to evaluate the expected benefit of different strategies (e.g., Bergin et al., 2008). This approach is resource-intensive (Dunker, 1984), and the numerical precision of models limits the size of emissions changes that can be actually evaluated (Hakami et al., 2004). An alternative approach is to use sensitivity analysis tools integrated in the simulation.

Sensitivity analysis reveals the relationship of model outputs (e.g., pollutant concentrations) to model input parameters (e.g., emissions rates, initial or boundary

---

* This Chapter is an extension of "Development of the high-order decoupled direct method in three dimensions for particulate matter: enabling advanced sensitivity analysis in air quality models," published in *Geoscientific Model Development* in March 2012. Co-authors are Shannon Capps, Yongtao Hu, Athanasios Nenes, Sergey Napelenok, and Armistead Russell.

conditions, and chemical reaction rates). Several different sensitivity analysis methods quantitatively express partial derivatives as the "sensitivity coefficients". One approach is the brute force (BF) approximation; using central finite difference approximation, first- and second-order sensitivities are expressed as:

$$S_{ij}^{(1),BF} = \frac{C_i|_{p_j+\Delta p_j} - C_i|_{p_j-\Delta p_j}}{2\Delta p_j} \tag{2.1}$$

$$S_{ijj}^{(2),BF} = \frac{C_i|_{p_j+\Delta p_j} - 2C_i|_{p_j} + C_i|_{p_j-\Delta p_j}}{(\Delta p_j)^2} \tag{2.2}$$

where $S_{ij}^{(1),BF}$ and $S_{ijj}^{(2),BF}$ represent the brute force first-order and second-order sensitivities, respectively of species *i* with respect to parameter $p_j$ (e.g., emissions, initial or boundary conditions, or reaction rates). $C_i$ represents the concentration of species *i*. $p_j$, $p_j + \Delta p_j$, and $p_j - \Delta p_j$ represent the values of the input parameter at which the concentrations are evaluated. Computational requirements for BF sensitivity analysis scale with the number of parameters investigated. Obviously, BF becomes resource-intensive with an increasing number of parameters of interest or with increasing order (e.g., second order or higher) of sensitivities. In addition to being computationally inefficient, the BF sensitivities are prone to considerable numerical noise. One reason for the numerical noise is the truncation errors, which are introduced by omitting the higher-order terms when deriving Equations (2.1) and (2.2) from the Taylor series expansion. The truncation error is a function of both the perturbation size ($\Delta p$) and the magnitude of higher-order sensitivities. If the system is highly nonlinear, even a small perturbation can cause sizable truncation error (Hakami et al., 2004). Another reason for the numerical noise of BF is due to the modeling accuracy and precision. For example, incomplete

convergence in iterative solvers will cause such errors. Both types of errors for second-order BF sensitivities are amplified compared to first-order BF sensitivities. Actually, as the order of sensitivities increase, BF approximations become significantly less accurate (Hakami et al., 2004).

An alternative approach to BF is the decoupled direct method in three dimensions (DDM-3D). This method operates integrally within a chemical transport model (CTM) and simultaneously computes local sensitivities of pollutant concentrations to perturbations in input parameters (Dunker, 1984; Yang et al., 1997; Cohan et al., 2005; Napelenok et al., 2006, Cohan et al., 2010). DDM-3D sensitivities are calculated by solving sensitivity equations that are the derivatives of the partial differential equations governing the CTM. DDM-3D is computationally efficient for three or more sensitivity parameters and is subject to considerably less numerical noise than BF. The difference in numerical cost has been studied by Napelenok et al. (2006). CPU time required by the two approaches to compute the same set of sensitivities is compared, with the number of sensitivity parameters ranging from 1 to 8. The CPU time needed by BF is almost twice that needed by DDM-3D if two or more parameters are considered. For 8 sensitivity parameters, the CPU time for BF is 27 minutes and DDM-3D 15 minutes. DDM-3D has been implemented in CTMs (e.g., CMAQ (Byun and Schere, 2006), CAMx (ENVIRON, 2005), URM (Boylan et al., 2002)) to conduct source impact analysis for ozone and PM (Yang et al., 1997; Mendoza-Dominguez and Russell, 2000; Odman et al., 2002; Napelenok et al., 2006; Koo et al., 2007). Initially, DDM-3D was applied to calculate first-order sensitivities, which are the locally linear responses of pollutant concentrations

to changes in model inputs and parameters (e.g., emissions, and initial and boundary conditions) at the conditions currently modeled.

DDM-3D has been extended to calculate high-order sensitivities of gaseous species by Hakami et al. (2003) within the Multiscale Air Quality Simulation Platform (MAQSIP) (Odman and Ingram, 1996). They calculated second- and third-order sensitivities using DDM-3D and showed that the approach could accurately capture the nonlinear response of ozone concentration to $NO_x$ and VOC emission changes. They also investigated the efficiency of DDM-3D in calculating high-order sensitivities. An important outcome of that work was that higher than second order sensitivities are not necessary for the majority of potential applications. More recently, the high-order approach for gaseous species has also been implemented in the Community Multiscale Air Quality (CMAQ) model (Cohan et al., 2005) and the Comprehensive Air Quality with extensions (CAMx) (Koo et al., 2010). High-order sensitivity calculations of gaseous species have been applied to source apportionment and air quality model uncertainty analysis (Cohan et al., 2005; Tian et al., 2010). Although nonlinear effects of aerosol precursors on aerosol concentrations have been of concern in the past decade (Ansari and Pandis, 1998; West and Pandis, 1999), developing HDDM for PM has not yet been undertaken due to the discontinuous, highly nonlinear solution surface of the inorganic aerosol thermodynamics. Only now has the challenging task of extending high-order, direct sensitivity analysis to particulate matter species been accomplished. HDDM-3D/PM is implemented in the Community Multidimensional Air Quality model, version 4.5 (CMAQ4.5).

## 2.2. Model Description

CMAQ is an Eulerian air quality model (Byun and Schere, 2006) that simulates emissions, deposition, transport and chemical transformation of atmospheric species primarily by solving the advection-diffusion-reaction equations:

$$\frac{\partial C_i}{\partial t} = -\nabla(uC_i) + \nabla(K\nabla C_i) + R_i + E_i \quad (2.3)$$

where $C_i$ is the concentration of the $i^{th}$ species, $u$ the fluid velocity, $K$ the turbulence diffusivity, $R_i$ the net chemical reaction rate of all chemical reactions that affect the concentration of the $i^{th}$ species, and $E_i$ the emission rate for the $i^{th}$ species (Seinfeld and Pandis, 2006). The chemicals species can be in gas phase or aerosol form.

In the modal treatment of aerosol in CMAQ, aerosol species are tracked based on their size using three modes: Aitken, accumulation, and coarse. The two smaller modes (noted as Aitken and accumulation modes, respectively) approximately represent $PM_{2.5}$, aerosols smaller than 2.5μm in aerodynamic diameter. CMAQ includes modeled processes of secondary inorganic aerosol (i.e., sulfate, nitrate, ammonium), anthropogenic secondary organic aerosol (SOA), and biogenic SOA formation as well as primary emissions of elemental carbon and sea salt in the Aitken and accumulation modes. $PM_{2.5}$ changes in response to new particle production from vapor phase precursors, coagulation of particles, growth by condensation from gaseous species, transport and deposition of particles, and emissions (Byun and Schere, 2006). The concentration of $PM_{2.5}$ is highly dependent on gas phase species concentrations because of the significant fraction of secondary aerosol in this size range. CMAQ4.5 assumes the secondary inorganic aerosols are in thermodynamic equilibrium with surrounding gases,

and uses ISORROPIAv1.7 (Nenes et al., 1998a; Fountoukis et al., 2007) to simulate their condensation and evaporation. A dynamic equilibrium approach has also been used by CMAQ4.7 to simulate the chemical interactions between coarse particles and gas-phase pollutants (Kelly et al., 2010). CMAQ4.5 partitions SOA between gas and condensed phase based on the two-product model of Odum et al. (1997) using empirically derived coefficients from chamber experiments (Schell et al., 2001). The algorithm to compute SOA concentrations is similar to that of photochemical reactions. Studies show that the thermodynamic coupling between SOA and the inorganic species can impact the total aerosol water content and the aerosol nitrate concentrations (Ansari and Pandis, 2000). This would result in a greater sensitivity of aerosol water content and nitrate concentrations to SOA precursors (e.g., monoterpenes and xylene). However, such a coupling is not parameterized in CMAQ4.5, so DDM sensitivities do not reflect these effects. Thus, this work mainly focuses on the sensitivities of inorganic aerosol species to $SO_2$, $NO_x$, and $NH_3$. The SOA representations in CMAQ are being updated (Edney et al., 2007 and Carlton et al., 2010), and further interactions between inorganic and organic aerosol fractions are likely to be included in future updates. The implementation of HDDM and DDM sensitivity analysis can be modified accordingly.

ISORROPIA assumes that equilibrium exists between gas phases and aerosol species and uses thermodynamics to calculate the composition of inorganic aerosols and concentrations of surrounding gases. Inputs to ISORROPIA include the total (gas and aerosol) concentrations of five inorganic precursor species (i.e., sulfate, nitrate, ammonium, sodium, and chloride), temperature, and relative humidity. To determine the aerosol composition at equilibrium, ISORROPIA first identifies the solution regime of

the given system based on sulfate ratio (i.e., the ratio of total ammonium and sodium to total sulfate). Then, the appropriate set of equilibrium and mass and charge conservation relationships are solved to calculate the phase state and equilibrium concentrations (Table 2. 1). Each of ten subcases has its own solution procedure and a distinct set of possible species at equilibrium.

### 2.3. Development of HDDM-3D/PM

HDDM-3D/PM directly computes the high-order DDM sensitivity coefficients of PM concentrations to input parameters, such as emissions rates, and initial or boundary conditions, by solving derivatives of the original equilibrium and conservation equations. First- and second-order sensitivity coefficients are defined as

$$s_{ij}^{(1)} = \frac{\partial C_i}{\partial p_j} \tag{2.4}$$

$$s_{ijk}^{(2)} = \frac{\partial^2 C_i}{\partial p_j \partial p_k} \tag{2.5}$$

where $s_{ijk}^{(2)}$ denotes second-order sensitivity of species *i* to parameters *j* and *k*; $C_i$ denotes the ambient concentration of species *i*; and $p_j$ and $p_k$ denote any two input parameters of interest.

HDDM-3D/PM calculates semi-normalized sensitivity coefficients, expressed in the same units as concentration and which allows for easier interpretation and application:

$$S_{ij}^{(1)} = \frac{\partial C_i}{\partial \varepsilon_j} \tag{2.6}$$

$$S_{ijk}^{(2)} = \frac{\partial^2 C_i}{\partial \varepsilon_j \partial \varepsilon_k} \tag{2.7}$$

where $\varepsilon_j$ and $\varepsilon_k$ are relative perturbations in parameters $p_j$ and $p_k$, and they are related to the absolute perturbation of a parameter by $\delta\varepsilon = \frac{\delta p}{p}$.

The fundamental steps to obtain high-order DDM-3D sensitivities for PM from CMAQ are similar to those for the gaseous species. Taking second-order derivatives of the governing equation results in a similar equation which can be solved for second-order sensitivity of PM:

$$\frac{\partial S_{ijk}^{(2)}}{\partial t} = -\nabla(uS_{ijk}^{(2)}) + \nabla(K\nabla S_{ijk}^{(2)}) + \vec{J}_i \vec{S}_{jk}^{(2)} + f(C_i, S_{ij}^{(1)}, S_{ik}^{(1)}, u, K, R_i, E_i) \tag{2.8}$$

$S_{ijk}^{(2)}$ is the second-order sensitivity of species *i* with respect to parameters $p_j$ and $p_k$; $S_{ij}^{(1)}$ and $S_{ik}^{(1)}$ are first-order sensitivities of species *i* to parameters $p_j$ and $p_k$, respectively; $\vec{J}_i$ is the $i^{\text{th}}$ row of Jacobian matrix defined as $J_{ik} = \partial R_{ik} / \partial C_k$. *k* is the $k^{\text{th}}$ species in the concentration vector. $\vec{S}_{jk}^{(2)}$ is the vector of second-order sensitivity coeffiients. *f* is a function primarily of $C_i$, $S_{ij}^{(1)}$, and $S_{ik}^{(1)}$. It can also be related to $u$, $K$, $R_i$, and $E_i$, depending on the types of sensitivity parameters. Details of $f$ can be found in Equation (2.9) in Hakami et al. (2003).

Equation (2.8) can be directly propagated through most of the processes associated with the formation and transport of PM species, such as the oxidation of reactive organic gases and the gas/particle partitioning of organic compounds (Schell et al., 2001). However, the secondary inorganic aerosol species are strongly coupled as they are assumed to be in thermodynamic equilibrium with their precursors (i.e., $NH_3$, $HNO_3$,

and HCl). The equilibrium is assumed to be reached instantaneously, so the direct use of Equation (2.8) is not appropriate. Thus, a different treatment for inorganic aerosol species is necessary to implement HDDM-3D/PM when using ISORROPIA.

The implementation of HDDM in ISORROPIA involves differentiation of the equilibrium reactions that are involved in determining the concentrations of each species. For example, the equilibrium reaction for the balance between nitric acid gas ( $HNO_{3,(g)}$ ) and nitrate ion ( $NO_3^-$ ) is

$$HNO_{3,(g)} \longleftrightarrow H^+ + NO_3^- \tag{2.9}$$

The corresponding equilibrium expression is

$$K = \frac{[H^+][NO_3^-]\gamma_{H^+}\gamma_{NO_3^-}}{P_{HNO_3} w_{H_2O}^2} \tag{2.10}$$

where $K$ is the equilibrium constant; [$A$] denotes the molar concentration of $A$; $\gamma_{H^+}$ and $\gamma_{NO_3^-}$ are the activity coefficients of $H^+$ and $NO_3^-$; $P_{HNO_3}$ is the partial pressure of $HNO_3$ gas; $w_{H_2O}$ is the water content.

Since $\gamma_{HNO_3}^2 = \gamma_{H^+}\gamma_{NO_3^-}$ and $P_{HNO_3} = [HNO_3]RT$,

$$K = \frac{[H^+][NO_3^-]\gamma_{HNO_3}^2}{[HNO_3] w_{H_2O}^2 RT} \tag{2.11}$$

where $\gamma_{HNO_3}$ is the mean activity coefficient of $H^+$ and $NO_3^-$; $R$ is the universal gas constant; and $T$ is temperature. Taking the logarithmic derivative of Equation (2.11) with respect to the first parameter of interest ($p_1$, where for brevity, $T$ is assumed constant) leads to the expression of first-order sensitivity equation:

$$\frac{S^{(1)}_{H^+,p_1}}{[H^+]}+\frac{S^{(1)}_{NO_3^-,p_1}}{[NO_3^-]}-\frac{S^{(1)}_{HNO_3,p_1}}{[HNO_3]}+\frac{2S^{(1)}_{\gamma_{HNO3},p_1}}{\gamma_{HNO3}}-\frac{2S^{(1)}_{H_2O,p_1}}{w_{H_2O}}=0 \tag{2.12}$$

Differentiating Equation (2.12) with respect to the second parameter of interest ($p_2$) gives the equation for second-order sensitivity:

$$\begin{aligned}&\frac{S^{(2)}_{H^+,p_1,p_2}}{[H^+]}+\frac{S^{(2)}_{NO_3^-,p_1,p_2}}{[NO_3^-]}-\frac{S^{(2)}_{HNO_3,p_1,p_2}}{[HNO_3]}+\frac{2S^{(2)}_{\gamma_{HNO3,p_1,p_2}}}{\gamma_{HNO_3}}-\frac{2S^{(2)}_{H_2O,p_1,p_2}}{w_{H_2O}}\\&=\frac{S^{(1)}_{H^+,p_1}S^{(1)}_{H^+,p_2}}{[H^+]^2}+\frac{S^{(1)}_{NO_3^-,p_1}S^{(1)}_{NO_3^-,p_2}}{[NO_3^-]^2}-\frac{S^{(1)}_{HNO_3,p_1}S^{(1)}_{HNO_3,p_2}}{[HNO_3]^2}+\frac{2S^{(1)}_{\gamma_{HNO3,p_1}}S^{(1)}_{\gamma_{HNO3,p2}}}{\gamma^2_{HNO_3}}-\frac{2S^{(1)}_{H_2O,p_1}S^{(1)}_{H_2O,p_2}}{w^2_{H_2O}}\end{aligned} \tag{2.13}$$

Repeating the same process with the other equilibrium reactions involved in the system gives similar expressions to Equation (2.13). Combining them with mass and charge balance equations leads to a system of linear equations (Table 2.1) with which second-order sensitivities can be calculated. In this implementation, the available options for the two parameters $p_j$ and $p_k$ are emission rates, reaction rate constants, initial conditions, and boundary conditions. The approach can be extended to parameters in ISORROPIA such as equilibrium coefficients, which would require minor modification to the right hand side of Equations (2.12) and (2.13).

Calculating second-order DDM-3D sensitivities depends on the corresponding first-order sensitivities, so second-order sensitivities are computed sequentially following the first-order sensitivities in the same model run. Comparing Equations (2.12) and (2.13), identical coefficient terms multiplying the sensitivities are found on the left-hand sides, which reduces computational cost by allowing the two systems of linear equations to share the same coefficient matrix. Overall, the computational cost of second-order sensitivities is very close to that of first-order because the main computing processes (mainly transport) are the same for each sensitivity.

Table 2.1. Equilibrium Relations, Mass and Charge Balance of ISORROPIA

Equilibrium Reactions

| | |
|---|---|
| $HSO_4^- \leftrightarrow H^+ + SO_4^{2-}$ | $K_1 = \frac{[H^+][SO_4^{2-}]\gamma_{H^+}\gamma_{SO_4^{2-}}}{[HSO_4^-]\gamma_{HSO_4^-}}$ |
| $NH_3(g) + H_2O(aq) \leftrightarrow NH_4^+ + OH^-$ | $K_2 = \frac{[NH_4^+][OH^-]\gamma_{NH_4^+}\gamma_{OH^-}}{P_{NH_3}a_w}$ |
| $HCl(g) \leftrightarrow H^+ + Cl^-$ | $K_3 = \frac{[H^+][Cl^-]\gamma_{H^+}\gamma_{Cl^-}}{P_{HCl}}$ |
| $HNO_3(g) \leftrightarrow H^+ + NO_3^-$ | $K_4 = \frac{[H^+][NO_3^-]\gamma_{H^+}\gamma_{NO_3^-}}{P_{HNO_3}}$ |
| $H_2O(aq) \leftrightarrow H^+ + OH^-$ | $K_w = \frac{[H^+][OH^-]\gamma_{NH_4^+}\gamma_{OH^-}}{a_w}$ |
| $Na_2SO_4(s) \leftrightarrow 2Na^+ + SO_4^{2-}$ | $K_5 = [Na^+]^2[SO_4^{2-}]\gamma^2_{Na^+}\gamma_{SO_4^{2-}}$ |
| $NH_4Cl(s) \leftrightarrow NH_3 + HCl$ | $K_6 = P_{NH_3}P_{HCl}$ |
| $(NH_4)_2SO_4(s) \leftrightarrow 2NH_4^+ + SO_4^{2-}$ | $K_7 = [NH_4^+]^2[SO_4^{2-}]\gamma^2_{NH_4^+}\gamma_{SO_4^{2-}}$ |
| $NaCl(s) \leftrightarrow Na^+ + Cl^-$ | $K_8 = [Na^+][Cl^-]\gamma_{Na^+}\gamma_{Cl^-}$ |
| $NaNO_3(s) \leftrightarrow Na^+ + NO_3^-$ | $K_9 = [Na^+][NO_3^-]\gamma_{Na^+}\gamma_{NO_3^-}$ |
| $NH_4NO_3(s) \leftrightarrow NH_3(g) + HNO_3(g)$ | $K_{10} = P_{NH_3}P_{HNO_3}$ |
| $NaHSO_4(s) \leftrightarrow Na^+ + HSO_4^{2-}$ | $K_{11} = [Na^+][HSO_4^{2-}]\gamma_{Na^+}\gamma_{HSO_4^-}$ |
| $NH_4HSO_4(s) \leftrightarrow NH_4^+ + HSO_4^-$ | $K_{12} = [NH_4^+][HSO_4^-]\gamma_{NH_4^+}\gamma_{HSO_4^-}$ |
| $(NH_4)_3H(SO_4)_2(s) \leftrightarrow 3NH_4^+ + HSO_4^- + SO_4^{2-}$ | $K_{13} = [NH_4^+]^2[HSO_4^-][SO_4^{2-}]\gamma^3_{NH_4^+}\gamma_{HSO_4^-}\gamma_{SO2_4^-}$ |

Mass Balance

$$[tNa] = [Na^+] + 2[Na_2SO_4] + [NaCl] + [NaNO_3] + [NaHSO_4]$$

$$[tSO_4] = [SO_4^{2-}] + [HSO_4^-] + [Na_2SO_4] + [NaHSO_4] + [(NH_4)_2SO_4] + [NH_4HSO_4] + 2[(NH_4)_3H(SO_4)_2]$$

$$[tNH_4] = [NH_3] + [NH_4^+] + 2[(NH_4)_2SO_4] + [NH_4HSO_4] + 3[(NH_4)_3H(SO_4)_2] + [NH_4Cl] + [NH_4NO_3]$$

$$[tNO_3] = [HNO_3] + [NO_3^-] + [NaNO_3] + [NH_4NO_3]$$

$$[tCl] = [HCl] + [Cl^-] + [NaCl] + [NH_4Cl]$$

Charge Balance

$$[H^+] + [Na^+] + [NH_4^+] = [NO_3^-] + [Cl^-] + 2[SO_4^{2-}] + [HSO_4^-] + [OH^-]$$

∗ All quantities in [ ] denote molar concentrations, the unit is mol $m^{-3}$air .

In ISORROPIA, the mean activity coefficients are determined by Bromley's formula (Bromley 1973). Sensitivities of the mean activity coefficients, $S^{(1)}_{\gamma_{HNO3},p_1}$ and $S^{(2)}_{\gamma_{HNO3},p_1,p_2}$ in Equations (2.12) and (2.13), are calculated by directly differentiating Bromley's formulas. As the activity coefficients are functions of the ion concentrations, their sensitivities are finally expressed as the combinations of sensitivities of relevant ion concentrations.

The liquid water content of aerosols is computed by the Zdanovskii-Stokes-Robinson (ZSR) relationship (Stokes and Robinson, 1966):

$$[H_2O] = \sum_i \frac{E_i}{m_{0i}} \tag{2.14}$$

where $E_i$ is concentration of the $i^{\text{th}}$ electrolyte in the multicomponent solution; $m_{0i}$ is the molality of a solution with only the $i^{\text{th}}$ electrolyte and the same water activity as the multicomponent solution. Sensitivities of the liquid water content are obtained by differentiating Equation (2.14). Because the concentrations of electrolytes are calculated from the equilibrium ion concentrations, both first- and second-order sensitivities of liquid water content can be ultimately expressed as a function of ion sensitivities:

$$S^{(1)}_{H_2O,p_1} = \sum_i \frac{1}{m_{oi}} \sum_j \frac{\partial E_i}{\delta A_j} S^{(1)}_{A_j,p_1} \tag{2.15}$$

$$S^{(2)}_{H_2O,p_1,p_2} = \sum_i \frac{1}{m_{oi}} \sum_j \frac{\partial E_i}{\delta A_j} S^{(2)}_{A_j,p_1,p_2} \tag{2.16}$$

where $A_j$ represents the $j^{\text{th}}$ ionic species in the system.

ISORROPIA uses different algorithms to treat neutralized and acidic aerosol, so this work applied a case-specific approach when implementing HDDM-3D sensitivity

analysis. Depending on the acidity of the aerosol, each subcase has its own solution routine and assumptions. For example, the neutralized aerosol algorithm assumes that bisulfate ions are minor species, and its concentration is adjusted after solving the equilibrium reactions of nitrate, nitric acid gas, ammonium, and ammonia gas. Alternately, the acidic algorithm assumes that either ammonia or nitric acid gas is a minor species and resolves its final concentration after determining aerosol concentrations of their counterparts. This feature was usually neglected in previous implementations of DDM in ISORROPIA, which caused discrepancies between BF and DDM sensitivites. The problem is now solved by the case-specific approach, which exactly follows the treatment of ISORROPIA for different aerosols during HDDM implementation.

## 2.4. Results and Discussion

### 2.4.1 Evaluation of HDDM for ISORROPIA

The performance of HDDM-3D/PM is evaluated in both the stand-alone ISORROPIA and the CMAQ model for inorganic species. In the stand-alone ISORROPIA, the HDDM-3D/PM sensitivities were compared to brute-force sensitivities (first- and second- order) calculated by Equations (2.1) and (2.2), using a relative perturbation of 1%. The input concentrations of total sulfate, ammonium, and nitrate range from 0.1 to 10 $\mu mol \cdot m^{-3}$ with an incremental of 0.1 $\mu mol \cdot m^{-3}$. The input concentrations of total sodium and chloride are 0.5 and 1 $\mu mol \cdot m^{-3}$, respectively (Table 2.2). These inputs are consistent with the typical chemical composition of inorganic

aerosols (Nenes et al., 1998b) and are also over a wide range allowing each subcase in ISORROPIA to be tested. The inorganic aerosol species are assumed to be in metastable state in CMAQ4.5, so the aerosols with the same chemical composition but different relative humidities are treated using the same algorithm. Therefore, we used a fixed relative humidity of 95% for stand-alone testing.

Table 2.2. Input cases for testing of HDDM-PM using stand-alone ISORROPIA

| ***Parameters*** | ***Values ( μmol $m^{-3}$ )*** |
|---|---|
| Total Sulfate | 0.1 ~ 10 |
| Total Ammonium | 0.1 ~ 10 |
| Total Nitrate | 0.1 ~ 10 |
| Total Sodium | 0.5 |
| Total Chloride | 1.0 |
| Relative Humidity | 95% |
| Temperature | 298K |

We first compared the first-order DDM-3D and BF sensitivities of the five major ions (i.e., $H^+$, $NH_4^+$, $SO_4^{2-}$, $HSO_4^-$, $NO_3^-$) with respect to input total concentrations of sulfate, ammonium, and nitrate (Figure 2.1). Good agreement is found between first-order BF and DDM sensitivities for all species (slope = 1 and $R^2$ = 0.99), which is essential for evaluating the second-order sensitivities due to the dependence of second-order DDM-3D and BF sensitivities on the first-order counterparts.

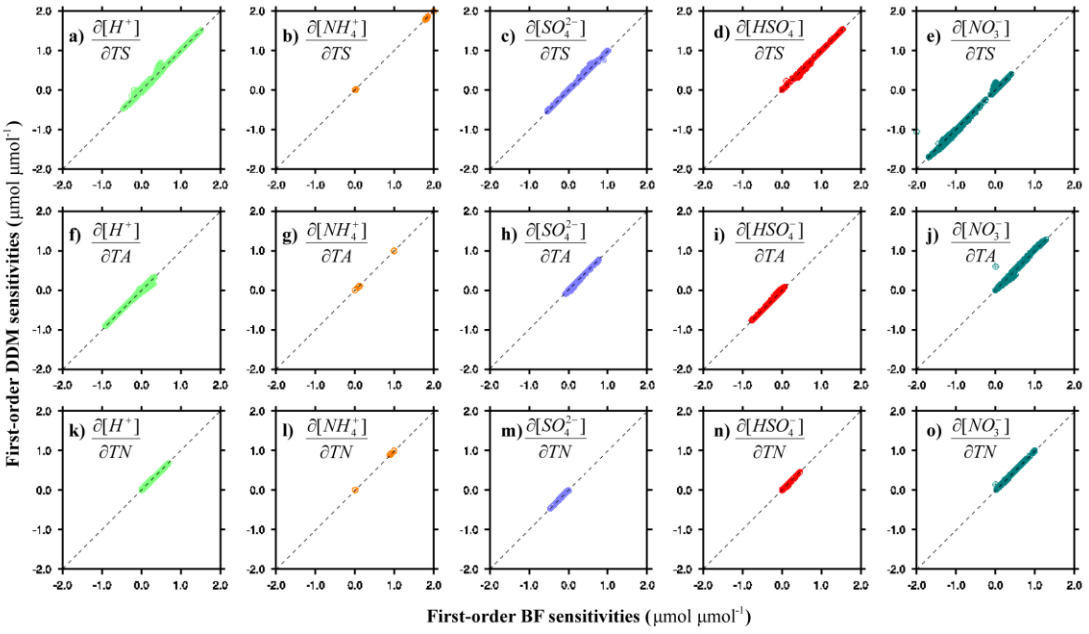


Figure 2.1 Comparison of first-order DDM and BF sensitivity coefficients of the five major ions (i.e., $H^+$, $NH_4^+$, $SO_4^{2-}$, $HSO_4^-$, and $NO_3^-$) to the change of total sulfate (*TS*), total ammonia (*TA*), and total nitrate (*TN*) in the stand-alone ISORROPIA. Each plot corresponds to the comparison of one sensitivity coefficient that is labeled on the up-left of the plot. For example, (a) shows the comparison of first-order sensitivity of hydrogen ion ($H^+$) to total sulfate predicted by DDM and BF. The dashed line is the one-to-one line.

The same comparison was conducted for second-order BF and DDM-3D sensitivities (Figure 2.2). Although most of the points fall on the one-to-one line (slope = 1, $R^2$ = 0.95), discrepancies were found for some second-order sensitivities (Figure 2.2). This is due to the noisy behavior of BF. As mentioned above, as the order of sensitivity coefficients increases, the two types of errors of BF approximations can become significantly larger. In other words, a lower degree of agreement between DDM-3D and BF are expected for second-order sensitivities. Our investigation into the noisy behavior of second-order BF sensitivities shows that second-order BF sensitivities vary dramatically with various sizes of perturbation ($\Delta p$) and the convergence criteria of the

ISORROPIA solution algorithm ( $\Delta\eta$ ) (Figure 2.3). This has also been demostrated by Capps et al. (2012). Further investigation into the charge balance for second-order BF and DDM-3D sensitivities revealed that the charge balance for BF sensitivities is not satisfied when they exhibit a noisy behavior. On the other hand, the charge balance is satisfied for DDM-3D sensitivities. These results strongly suggest that the HDDM-3D sensitivity coefficients are much more stable, while the BF second-order sensitivity coefficients are subject to significant numerical noise.

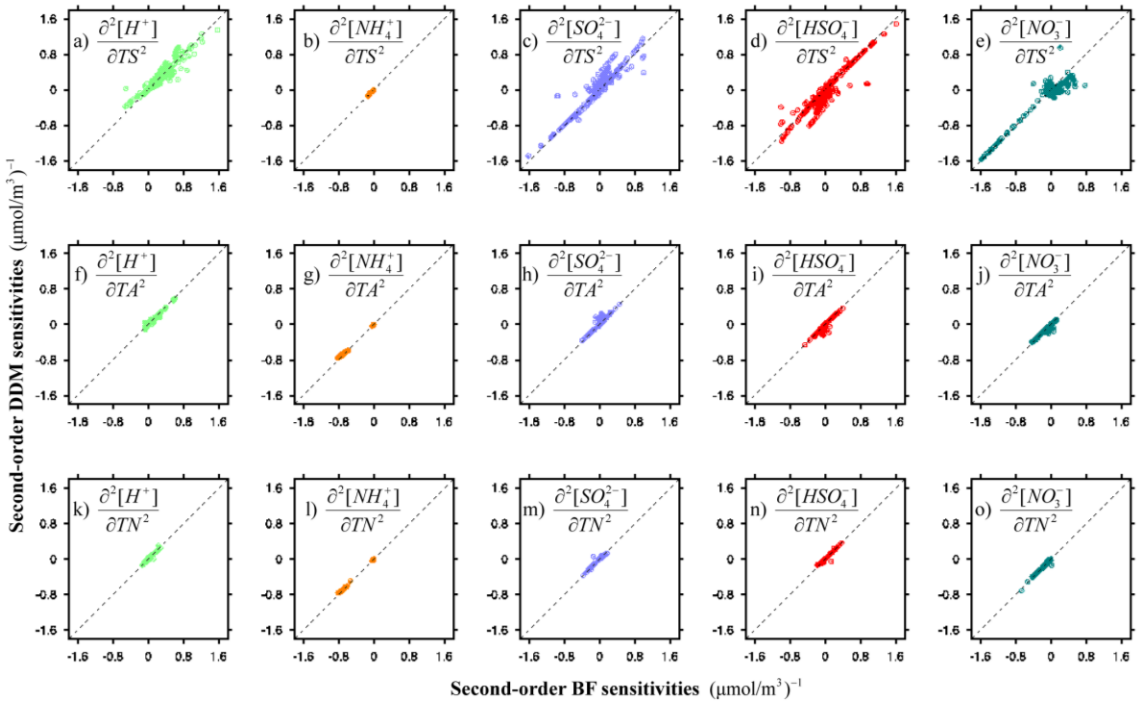


Figure 2.2. Comparison of second-order DDM and BF sensitivity coefficients of the five major ions (i.e., $H^+$, $NH_4^+$, $SO_4^{2-}$, $HSO_4^-$, and $NO_3^-$) to the change of total sulfate (*TS*), total ammonia (*TA*), and total nitrate (*TN*) in the stand-alone ISORROPIA. Each plot corresponds to the comparison of one sensitivity coefficient that is labeled on the up-left the plot. For example, (a) shows the comparison of second-order sensitivity of hydrogen ion ($H^+$) to total sulfate predicted by DDM and BF. The dashed line is the one-to-one line.

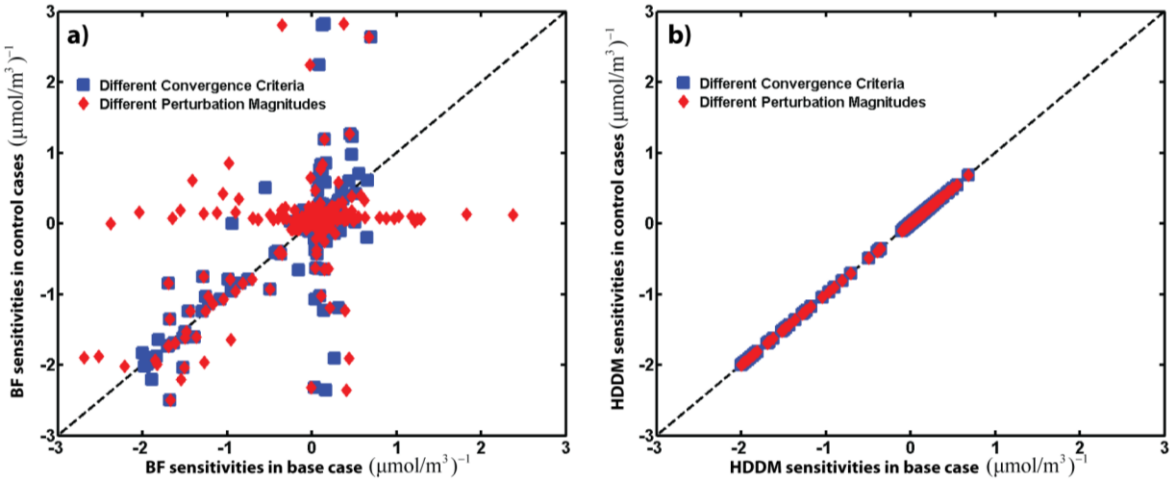


Figure 2.3. Second-order sensitivity coefficients of aerosol nitrate to total sulfate in stand-alone ISORROPIA calculated by (a) BF and (b) HDDM under three conditions: 1) base case, where the perturbation used by BF ($\Delta p$) = 1% and the convergence criteria of ISORROPIA ($\Delta \eta$) = $1\times10^{-10}$; 2) control case 1 (blue squares) with $\Delta p$ = 1% and $\Delta \eta = 1\times10^{-3}$; and 3) control case 2 (red diamonds) with $\Delta p$ =0.1% and $\Delta \eta$ = $1\times10^{-10}$. Results from the two control cases are compared to those from the base case. The dashed line is the one-to-one line.

### 2.4.2 Evaluation of HDDM for CMAQ

HDDM-3D/PM is applied to simulate a winter episode: Jan 1-7, 2004. Winter episodes have higher nitrate levels, which is a more stringent test of HDDM-3D/PM. The modeling domain covers the entire continental United States and portions of Canada and Mexico (Figure 2.4) using a 36-km horizontal grid-spacing and thirteen vertical layers extending about 16km above the ground. The meteorological fields were developed using the Fifth-Generation PSU/NCAR Mesoscale Model (MM5) (Grell, Dudhia et al. 1994). Emissions were prepared using the Sparse Matrix Operator Kernel Emissions (SMOKE) model (CEP 2003). SAPRC99_AE4_AQ was selected as the chemical mechanism (Carter, 2000; Binkowski and Roseelle, 2003).

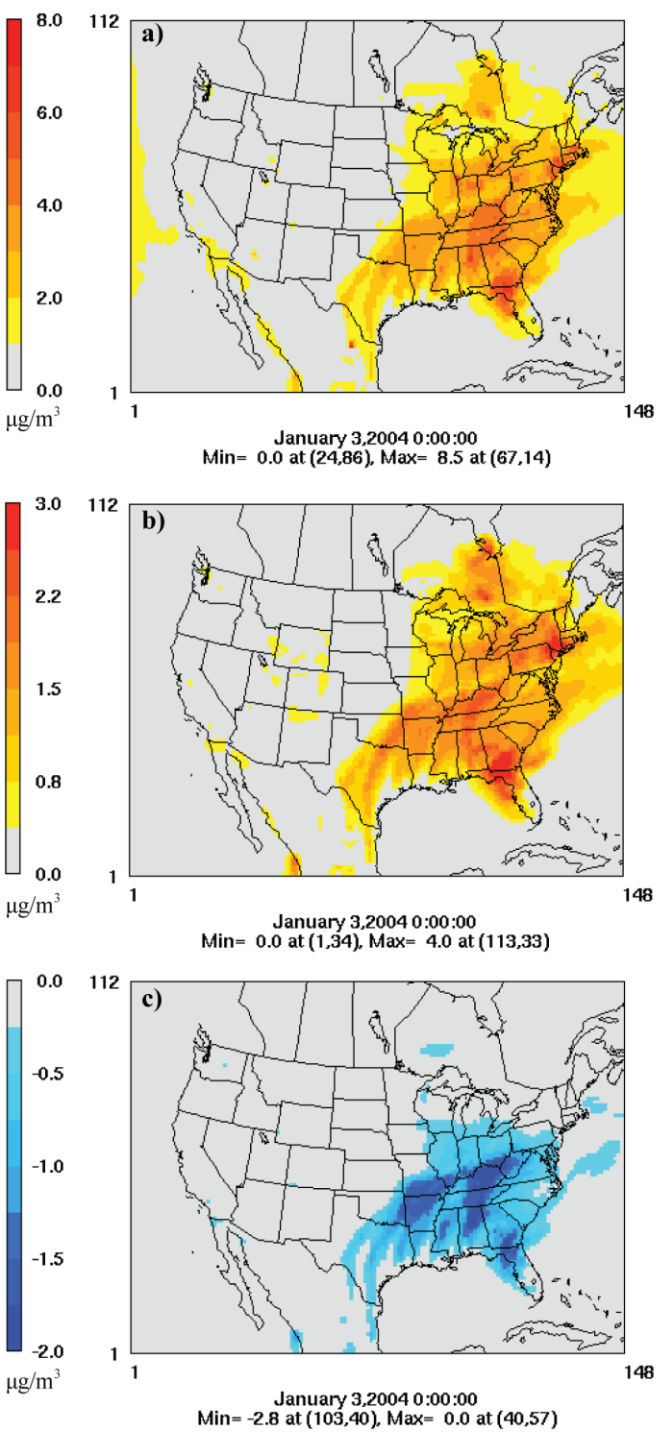


Figure 2.4. Spatial distribution of 24-hr averages of a) simulated concentration of sulfate, b) first- and c) second-order sensitivities of sulfate to $SO_2$ at surface layer on Jan 3, 2004.

The sensitivities of aerosol sulfate, nitrate, and ammonium to domain-wide $SO_2$, $NO_x$, and $NH_3$ emissions are studied in this simulation. During a single simulation, HDDM-3D/PM provides all the sensitivities of interest for each grid at each time step. The spatial patterns of first- and second-order DDM sensitivities of aerosol sulfate to $SO_2$ show that the most sensitive area is the Eastern US (Figure 2.4); since this region is the area with the highest $SO_2$ emissions, these sensitivities were expected. Spatial distributions of first- and second-order sensitivities are found to be consistent. The magnitudes of the second-order sensitivities are smaller and usually opposite in sign, but still indicate a significant contribution to the total response.

Comparison of first- and second-order BF and HDDM-3D/PM sensitivities of sulfate, nitrate, ammonium, and $PM_{2.5}$ to domain-wide $SO_2$, $NO_x$, and $NH_3$ emissions find similar results to the stand-alone version (Figures. 2.5 and 2.6). First- and second-order BF sensitivities are calculated using Equations (2.1) and (2.2) with a 50% reduction of each emission of interest, respectively. A choice of 50% is made to minimize the impact of noise for BF sensitivities when taking a small difference between two relatively large concentrations though it is expected that nonlinearities may be of some importance over this range. Using a smaller reduction leads to considerably larger error, which has been identified when testing HDDM-3D/PM in the stand-alone ISORROPIA. Most of the DDM-3D and BF first-order sensitivities are in good agreement with an overall slope of 0.9 and $R^2$ of 0.91 (Figure 2.5). The degree of agreement between DDM-3D and BF sensitivities of $PM_{2.5}$ to $NO_x$ and $NH_3$ emissions is improved from $R^2 = 0.63$ to $R^2 = 0.93$ by the case-specific DDM approach in ISORROPIA. Sensitivity of aerosol nitrate to $SO_2$ emissions is of concern to policy makers since the nitrate levels may be increased from

$SO_2$ emission controls (West et al., 1999). A relatively low degree of agreement was found between DDM-3D and BF sensitivities of nitrate to $SO_2$ (Figure 2.5c). However, nitrate concentrations are usually expected to increase with decreasing $SO_2$ emissions, so the first-order sensitivity should be negative, as is shown by DDM-3D. The BF, however, is producing a significant amount of positive sensitivities, which is due to the nonlinear dependence of nitrate on $SO_2$ emissions coupled with numerical noise. The comparison for sensitivity of sulfate to $NO_x$ has two branches that are slightly off the one-to-one line. These disagreements are caused by cloud processes as additional testing shows that the discrepancies disappear when the cloud module is turned off. The disagreement for the sensitivity of sulfate to $NH_3$ also comes from the cloud module where $SO_2$ is oxidized to sulfate. The oxidation process is highly affected by the pH value, and the response of sulfate to $NH_3$ is quite nonlinear. BF sensitivities of sulfate to $NH_3$ are strongly affected by this nonlinearity. Further investigation showed that they change dramatically with the perturbation sizes as well as the BF approaches (i.e., forward and central finite difference). Overall, first-order BF and DDM-3D sensitivities compared well. BF sensitivities become less accurate when the system is quite nonlinear. This also implies the significance of the monlinear response and the necessarity of performing high-order sensitivity analysis.

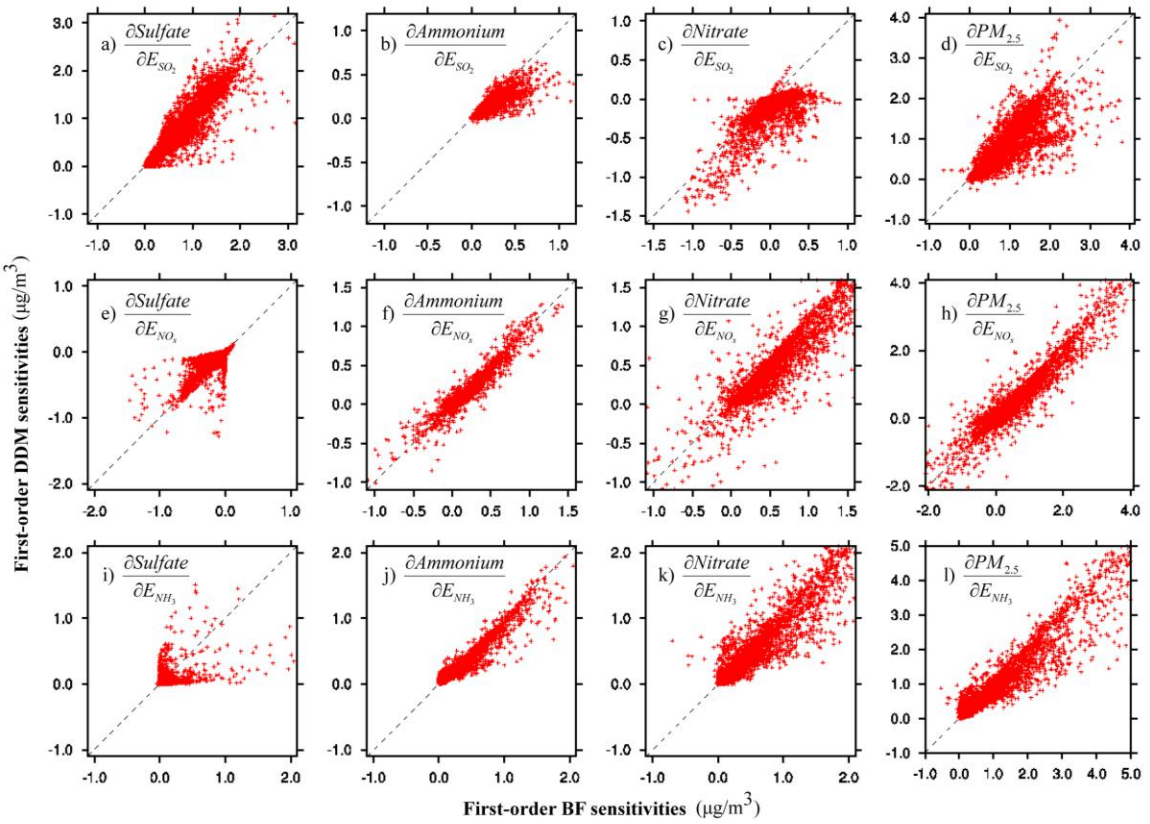


Figure 2.5. Comparison of first-order sensitivities of sulfate, ammonium, nitrate, and $PM_{2.5}$ to $SO_2$, $NO_x$, and $NH_3$ calculated by HDDM-3D/PM and BF at surface layer on Jan 2, 2004. Each plot represents one sensitivity coefficient that is labeled on the up-left of the plot. The dashed line is the one-to-one line.

Second-order DDM-3D sensitivities are also evaluated using BF. Good agreement is found for $S^{(2)}_{SO_4^=,SO_2,SO_2}$, $S^{(2)}_{SO_4^=,NO_x,NO_x}$, $S^{(2)}_{NH_4^+,NO_x,NO_x}$, and $S^{(2)}_{NO_3^-,NO_x,NO_x}$ (Figures. 6a, 6e, 6f, and 6g) while the correlations are relatively low for some sensitivities, such as $S^{(2)}_{NO_3^-,SO_2,SO_2}$ and $S^{(2)}_{SO_4^=,NH_3,NH_3}$ (Figures. 2.6c and 2.6i). As mentioned above, second-order BF sensitivities for stand-alone ISORROPIA are strongly affected by the size of the perturbation. Here, we also investigated the impact of perturbation size to second-order BF sensitivities. For each second-order sensitivity of interest, we compared the BF results with 10% and 50% emission reduction. The noisy behavior of second-order BF sensitivities is evident

(Figure 2.7). The two BF scenarios in particular show little consistency for second-order sensitivity of sulfate to $NH_3$, which suggests that BF sensitivities directly computed from Equations (2.1) and (2.2) may not be reliable (Figure 2.7g). The plot for second-order sensitivity of nitrate to $SO_2$ also shows that the BF results vary significantly (Figure 2.7c). Thus, BF is not able to accurately approximate second-order local sensitivities of PM in CMAQ. Given the good performance of HDDM in the stand-alone ISORROPIA, and the great scatter between implementing BF with different perturbations, the direct approach is expected to provide more reliable results.

The average computational cost of calculating one second-order sensitivity of PM is found to be very close to that of one first-order sensitivity. For one day simulation, the average model time needed by the aerosol module for one first-order and one second-order sensitivites are 9 and 11 minutes, respectively, given that the second-order sensitivity calculation uses the same solution algorithm as first-order sensitivity. Therefore, the time required by matrix factorization and transport-related computations is almost the same for first- and second-order sensitivities. An indirect cost associated with the second-order sensitivity calculation is that all relevant first-order sensitivities should also be calculated, which is generally of interest anyway in any application involving high-order sensitivity (Hakami, Odman et al. 2003). On the other hand, BF needs more than one simulation, and its computational cost increases directly with the order and the number of sensitivity parameters. HDDM-3D/PM provides an efficient approach to conduct high-order sensitivity analysis as it computes high-order sensitivities at a similar computational effort as first-order sensitivities.

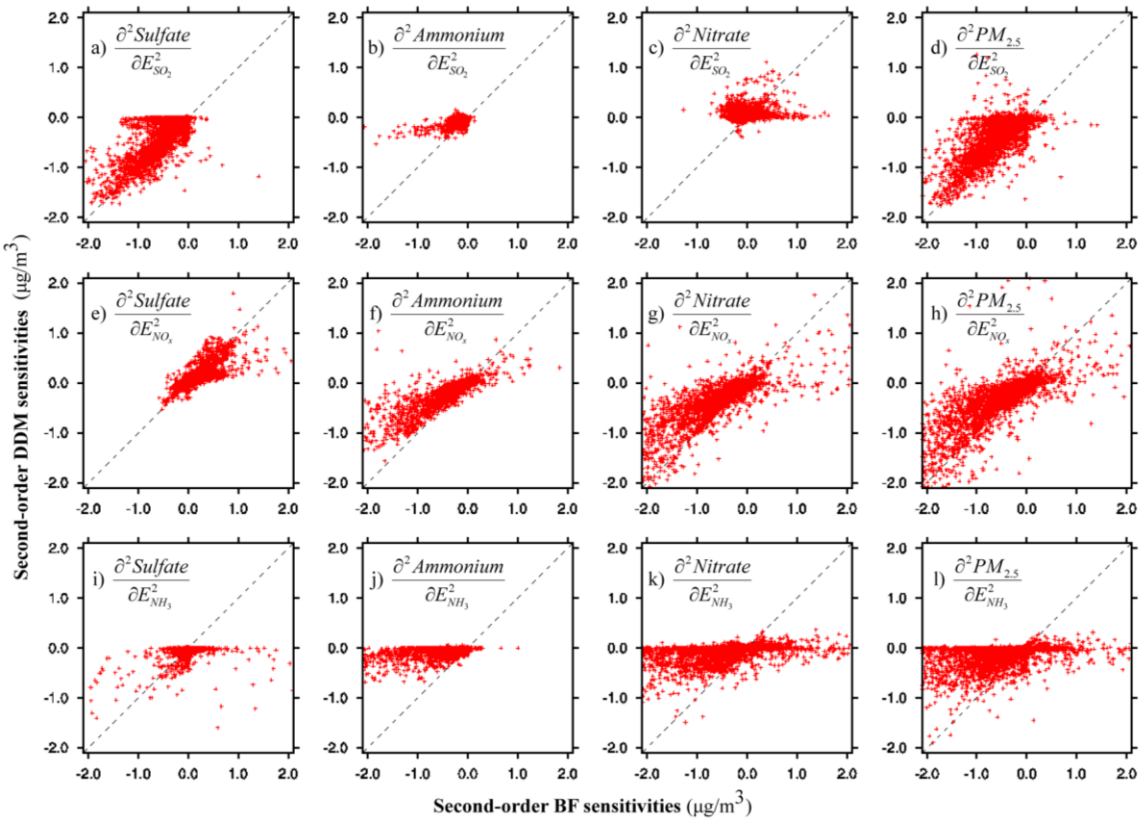


Figure 2.6. Comparison of second-order sensitivities of sulfate, ammonium, nitrate, and $PM_{2.5}$ to $SO_2$, $NO_x$, and $NH_3$ calculated by HDDM-3D/PM and BF at surface layer on Jan 2, 2004. Each plot represents one sensitivity coefficient that is labeled on the up-left of the plot. The dashed line is the one-to-one line.

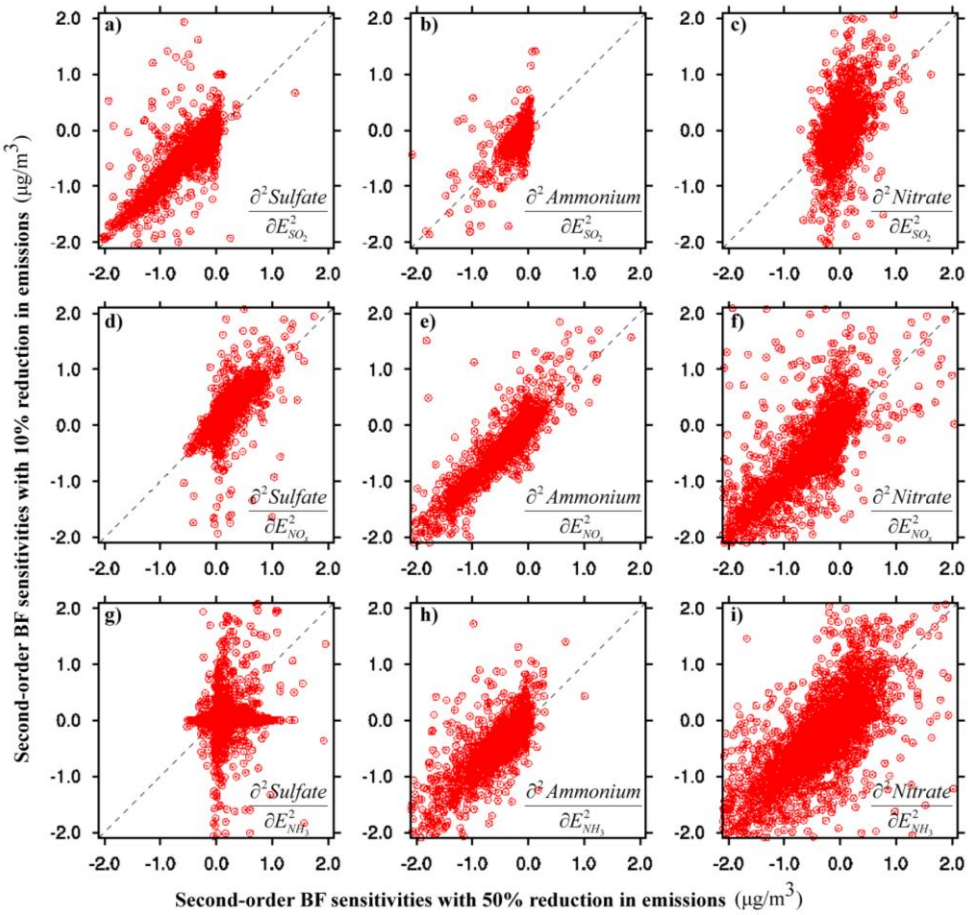


Figure 2.7. Comparison of second-order BF sensitivities calcualted with 10% and 50% perturbation in emissions using CMAQ simulation on Jan 2, 2004 at surface layer.

2.4.3 Evaluation of the reduced form model of CMAQ

HDDM-3D/PM has many practical applications, most of which are based on Taylor series expansion (Hakami, Odman et al. 2003):

$$C(\Delta\varepsilon) = C(0) + \Delta\varepsilon S^{(1)}(0) + \frac{\Delta\varepsilon^2}{2} S^{(2)}(0) + \text{higher order terms} \qquad (2.17)$$

where $C(0)$ stands for the pollutant concentration at base case emissions and $C(\Delta\varepsilon)$ with a perturbation of $\Delta\varepsilon$ in emissions. With Equation (2.17), one can quickly compute the impact of emission perturbations on the ambient concentrations of pollutants.

Including the second-order term (i.e., the third term on the right hand side of Equation (2.17)) is expected to reduce the error between the approximations using Taylor series expansion and the model simulation. For example, assuming 50% of domain-wide $NO_x$ emissions are reduced in the simulation above, we predicted the concentration of nitrate using first- and second-order Taylor series expansion (Equation (2.17)) and compared them with model simulation (Figure 2.8a). Predictions using second-order Taylor series expansions are closer to the model simulation than those using first-order Taylor series expansions (Figure 2.8a). A similar result is also found for nitrate concentration with a 50% reduction in $SO_2$ emissions (Figure 2.8b). Thus, including the second-order term in Taylor series approximation improves the accuracy of prediction.

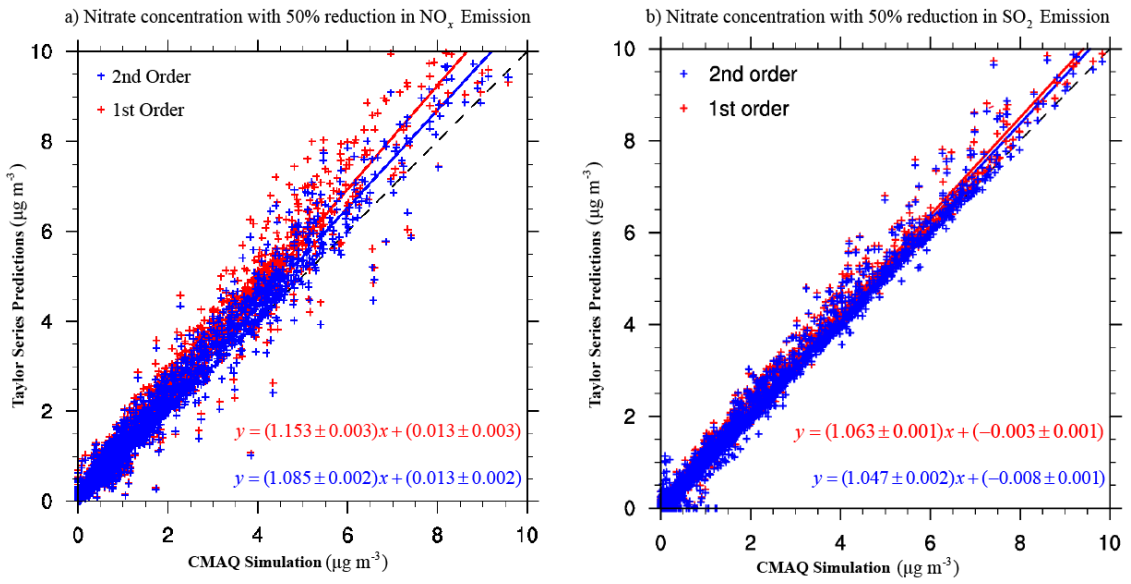


Figure 2.8. Comparisons of model simulation and predictions using Taylor Series Expansions for concentrations of nitrate at 16:00 EDT on Jan 2, 2004, with a 50% reduction in $NO_x$ and a 50% reduction in $SO_2$. The solid lines reflect the linear regression of the Taylor series predictions against the CMAQ simulation results; the dotted lines represent the area of perfect agreement.

Taylor series expansions derived using HDDM sensitivity coefficients enable efficient evaluation of emission control strategies. One CMAQ-HDDM simulation would be sufficient to estimate the changes in pollutant concentrations with respect to emission reductions. Predictions of nitrate concentrations with 20% and 100% reductions in total $SO_2$ emission using HDDM sensitivities compare well with the CMAQ model simulation. The slope from linear regression analysis is close to 1 (Figures. 2.9a and 2.9b). Predictions driven by BF sensitivities are close to the CMAQ simulation at 20% reductions and are a little off the one-to-one line for 100% reductions (Figure 2.9b). The BF sensitivities used here are results of a 50% perturbation. BF sensitivities prepared using a 10% perturbation were also tested (not shown here), but suffered from more numerical noise. Simulated sulfate concentrations with 20% and 100% reductions in total $NH_3$ emissions also exhibit good agreement with model simulation (Figures. 2.9c and 2.9d).

The reduction in concentrations that would occur if the sources of interest did not exist is called the zero-out source contribution (ZOC) (Cohan et al., 2005). The advantage of using Equation (2.17) to calculate ZOC is that it is based on an air quality model with relevent physical and chemical processes included. Indirect effects, such as source contributions of $SO_2$ emissions to nitrate and $NH_3$ emissions to sulfate (Figures. 2.9b and 2.9d), can be reasonablely evaluated. The ZOC can also be applied to a combination of source emissions. Consider two emission sources ( $p_j$ and $p_k$ ) that are perturbed simultaneously. The expression of ZOC of species $i$ ( $ZOC_i$ ) can be obtained from Equation (2.17) with multiple sensitivity parameters:

$$ZOC_i(p_j, p_k) \approx (S^{(1)}_{i,j} - 0.5S^{(2)}_{i,j,j}) + (S^{(1)}_{i,k} - 0.5S^{(2)}_{i,k,k}) - S^{(2)}_{i,j,k} \quad (2.18)$$

The cross sensivitity, the last term on the right-hand-side of Equation (2.18), is able to quantify the interactions between the two emissions.

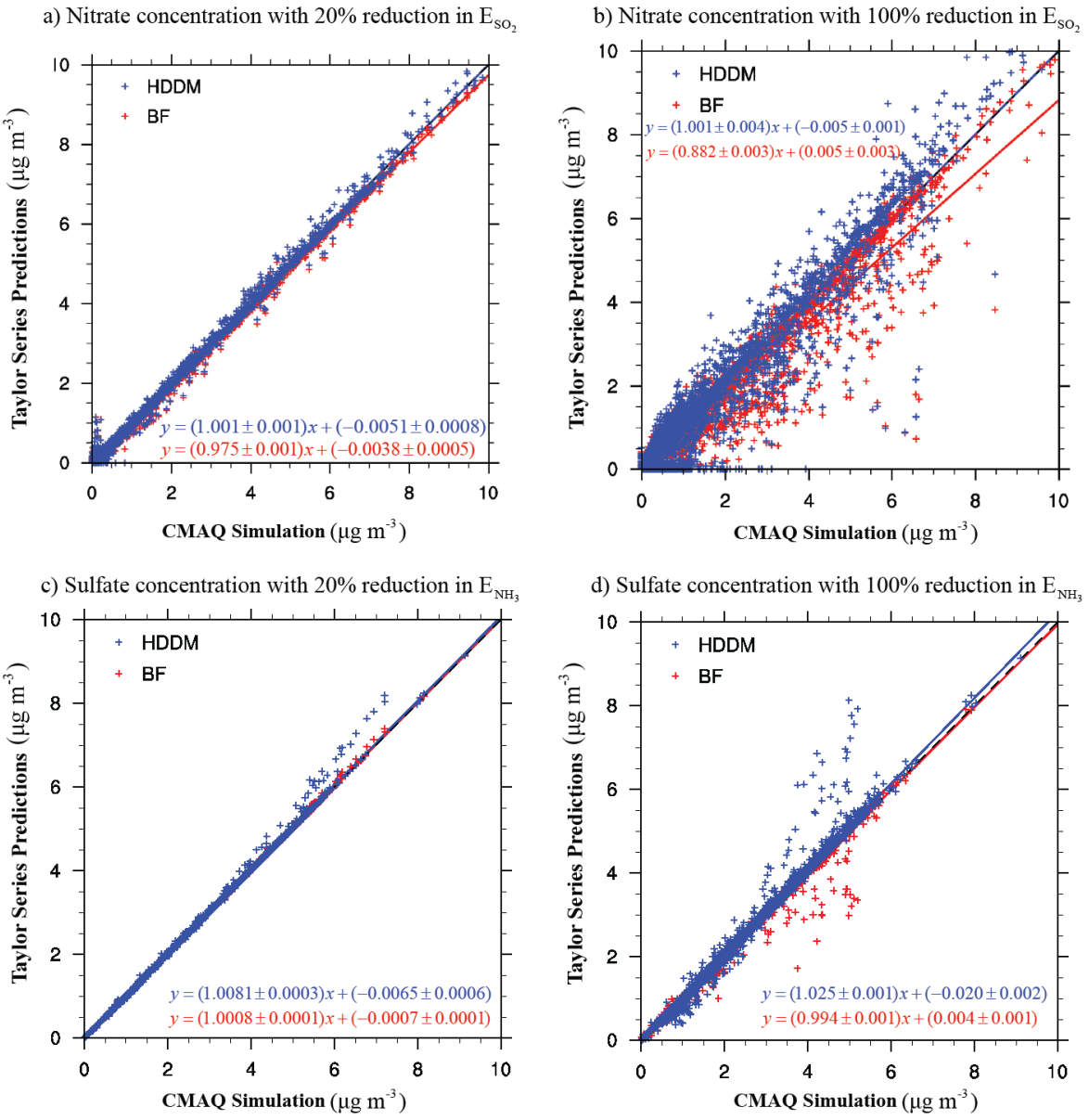


Figure 2.9. Comparisons of model simulation and predictions using Taylor series expansions with HDDM and BF sensitivities for concentrations of nitrate with 20% and 100% domain-wide reductions in $SO_2$ emissions rates and concentrations of sulfate with 20% and 100% domain-wide reductions in $NH_3$ emissions rates at 16:00 EDT on Jan 2, 2004. BF sensitivities are from a 50% perturbation. The solid lines reflect the linear regression of the Taylor series predictions against the CMAQ simulation results; the dotted lines represent the area of perfect agreement.

2.4.4. Source contributions to $PM_{2.5}$ in the Houston region

The newly implemented HDDM-3D for particulate matter was applied to the Houston region to investigate the source contributions from various emissions and the interaction of them. The modeling domain covers the southeastern Texas and a small fraction of the western Louisianan. The eight counties in Houston-Galveston-Brazoria area (the red area in Figure 2.10), is designated as nonattainment area of NAAQS for 8-hour ozone. A high ozone episode from August 10 to September 14 in 2006 was selected for study. This episode is one of the State Implementation Plan (SIP) modeling episodes of this area, and the nonlinear behavior of ozone was found to be significant by Xiao et al. (2010). As ozone and particulate matter interact with each other through complex chemical reactions, this episode is appropriate for conducting analysis for the nonlinear responses of particulate matter to emissions and the relationship to ozone changes.

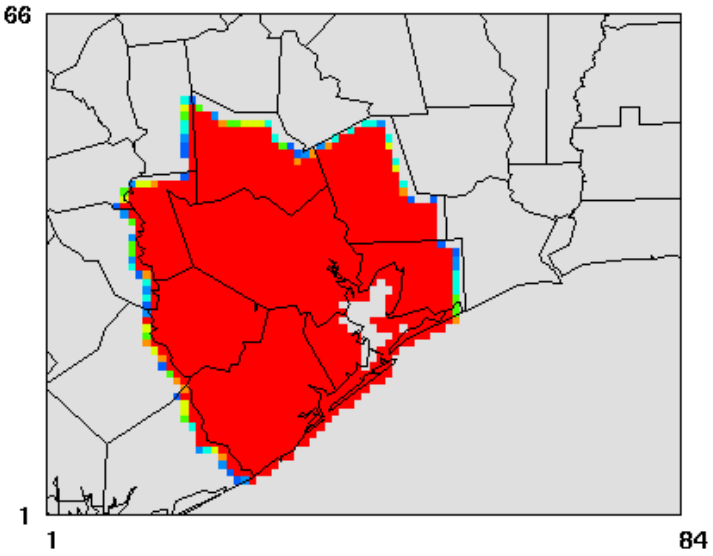


Figure 2.10. Modeling domain and the HGB area including the following eight counties: Brazoria, Chambers, Fort Bend, Galveston, Harris, Liberty, Montgomery, and Waller.

First of all, we examined the source contributions to daily average $PM_{2.5}$ in the HGB area from five major pollutant emissions including $NO_x$, VOC, $SO_2$, $NH_3$, and primary PM, as well as their interactions which are represented as cross sensitivities. The sums of source contributions compared well with the daily average $PM_{2.5}$ concentrations. Over the entire episode, the percentages of source contributions through linear responses ranged from 66.3% to 96.5%. Correspondingly, the percentages of source contributions that arise through nonlinear responses (either second order responses to the source, or interactions between sources) range from 3.5% to 33.7%. For 6 days of the month-long episode, source contributions through nonlinear responses constitute over 15% of the daily average $PM_{2.5}$. These results suggest that the response of $PM_{2.5}$ to emissions in the HGB area in summer time is mostly linear much of the time. Days with large fractions of nonlinear responses were concentrated in the beginning of September. During these days, light synoptic-scale winds were in the opposite direction of the sea breeze, and the interaction between them resulted in stagnant conditions (Banta et al., 2005). Such conditions allow the primary emissions to accumulate and react to form pollutants such as PM and ozone. $NO_x$ emissions became more reactive during the days with high nonlinearity than the other days. The contributions of $NO_x$ to daily $PM_{2.5}$ can reach 12% on days with high nonlinearity while they are typically below 5% for the other days. Besides the stagnation, there are two other reasons for the active role of $NO_x$ on these days: 1) These were the days with the highest ozone buildup in this episode (Rappengluck et al., 2009), so more $NO_2$ reacted with ozone to form $NO_3$ and then $HNO_3$, 2) The daily average temperature is lower on these days, which favored the

formation of $NH_4NO_3$. The contributions of $SO_2$ emissions were higher on these days as well because more ozone could oxidize $SO_2$ to form sulfate via aqueous-phase reactions.

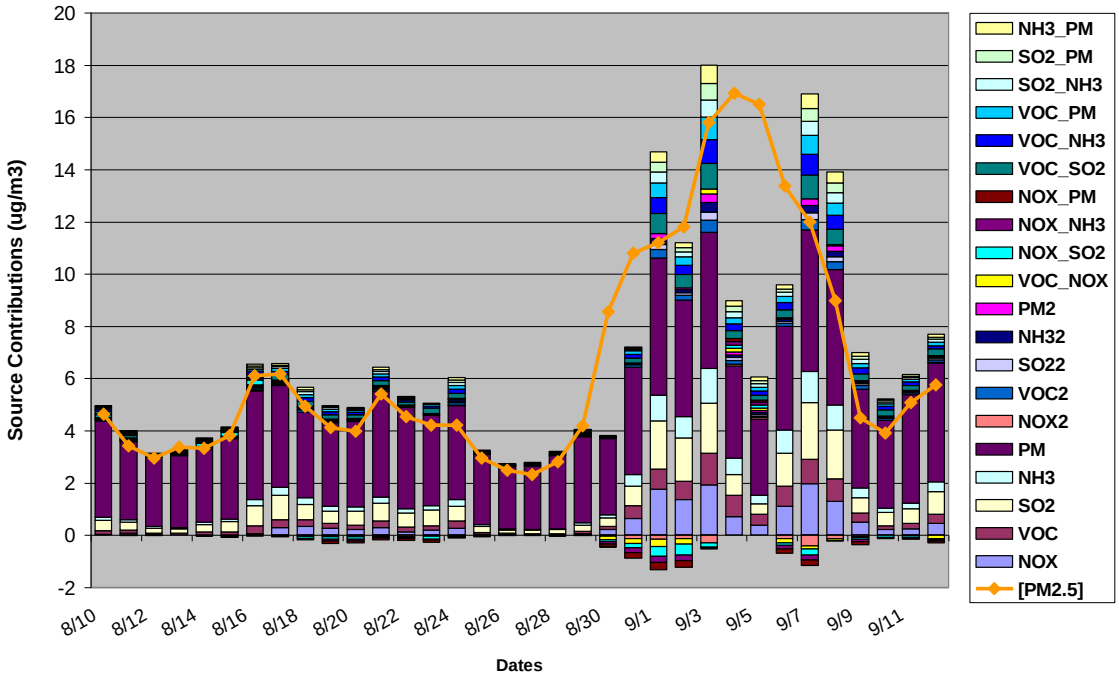


Figure 2.11. Source contributions to daily average $PM_{2.5}$ in the HGB area. The notations in the legend indicate different types of responses. For example, NOX denotes the contribution of $NO_x$ to $PM_{2.5}$ through first-order sensitivity (i.e., $\frac{\partial PM_{2.5}}{\partial NO_x}$), NOX2 denotes the contribution of $NO_x$ to $PM_{2.5}$ through second-order self sensitivity (i.e., $-0.5\times\frac{\partial^2 PM_{2.5}}{\partial NO_x^2}$), NOX_SO2 denotes the contribution of the interaction between $NO_x$ and $SO_2$ through the cross sensitivity (i.e., $-\frac{\partial^2 PM_{2.5}}{\partial NO_x \partial SO_2}$). The orange line denotes simulated daily average $PM_{2.5}$ concentrations.

For nonlinear responses, contributions from cross sensitivities were higher than those from second-order self sensitivities, which indicates that the interactions between various emissions can be more important for the formation of $PM_{2.5}$ in this region. This feature is different from that of ozone, which has a large portion of contributions from second-order self sensitivities of $NO_x$ and VOC. Adding more $NO_x$ may push the

chemical regime for ozone formation from $NO_x$-limited to VOC-limited, so the second-order self sensitivities are significant. However, PM precursors tend to react with each other to form more PM, so most of the source contributions from cross sensitivities are positive. For example, VOC can cause more ozone formation, and ozone can oxidize $SO_2$ in the aqueous phase to form more sulfate and can increase HO in the gas phase. Thus, the source contribution from the interaction between VOC and $SO_2$ is positive. The interaction between VOC and $NH_3$ is also positive in that ozone formed from VOC can increase both sulfate and nitrate concentrations, which can bring more $NH_3$ into the particle phase. Despite the small magnitude, negative contributions occur for the interaction between $NO_x$ and other precursors such $SO_2$, $NH_3$, and VOC. This can be illustrated by looking at the diurnal pattern of the source contributions to both $PM_{2.5}$ and ozone (Figure 2.12). $PM_{2.5}$ formed from midnight to early morning when the temperature was low and relative humidity was high and when a large amount of ozone had accumulated and reacted with $NO_x$. Then, adding $NO_x$ will reduce ozone, so less $SO_2$ will be oxidized by ozone to form sulfate, and thus less $NH_4SO_4$ will formed. The interaction between VOC and $NO_x$ is also through ozone. The opposite effects on ozone concentration lead to a negative cross-sensitivity contribution. The close interaction of ozone and PM sheds light on the importance of designing multi-species control strategies.

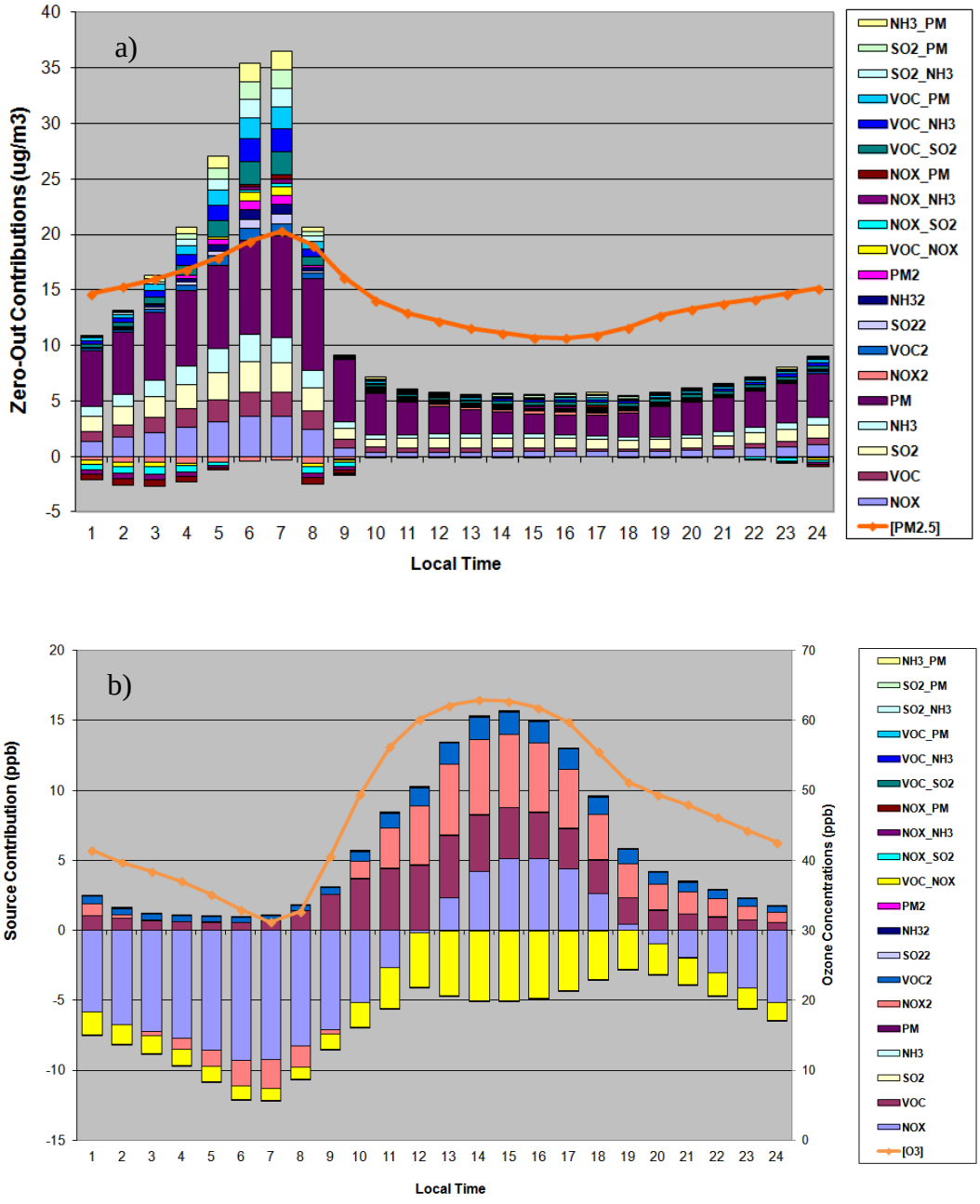


Figure 2.12. Diurnal patterns of source contributions to a) $PM_{2.5}$ and b) ozone over the HGB area. Results have been averaged for 1-5 September 2006. The notions in the legend indicate different types of responses. For example, NOX denotes the contribution of $NO_x$ to $PM_{2.5}$ through first-order sensitivity (i.e., $\frac{\partial PM_{2.5}}{\partial NO_x}$), NOX2 denotes the contribution of $NO_x$ to $PM_{2.5}$ through second-order self sensitivity (i.e., $-0.5\times\frac{\partial^2 PM_{2.5}}{\partial NO_x^2}$), NOX_SO2 denotes the contribution of the interaction between $NO_x$

and $SO_2$ through cross sensitivity(i.e., $-\frac{\partial^2 PM_{2.5}}{\partial NO_x \partial SO_2}$). The orange line denotes simulated daily average $PM_{2.5}$ concentrations.

## 2.5. Conclusions

The high-order decoupled direct method in three dimensions for particulate matter (HDDM-3D/PM) has been implemented in the Community Multiscale Air Quality (CMAQ) model. The implementation of HDDM-3D/PM into ISORROPIA applied a case-specific approach and explicitly computes the sensitivity of activity coefficients. Comparisons of the results with the traditional BF approach generally give good agreement. The BF sensitivities are found to be dependent on the perturbation sizes and the model accuracy, which leads to noisy behavior, especially for high-order sensitivities (Figures. 2.3 and 2.7). The direct assessment of second-order sensitivities with HDDM-3D/PM avoids the apparent pitfalls of the BF approach that cause this noise.

HDDM-3D/PM has similar computational cost to the previous DDM-3D/PM. The CPU time required by the aerosol module to conduct a one-day simulation with one first-order and one second-order sensitivity parameter are 9 and 11 minutes, respectively. This is another advantage over the BF approach, for which computational time increases more with the order of the sensitivities computed.

The case study in the HGB area demonstrated the application of HDDM-3D for assessing source contributions to $PM_{2.5}$. The nonlinear responses occurred when $NO_x$ became active in chemical reactions. Stagnant meteorological conditions, abundance of ozone during the night, low temperature, and high relativity humidity are the factors that lead to highly reactive role of $NO_x$. The abundance of ozone also brings more $SO_2$ and $NH_3$ into the aerosol phase. Diurnal patterns of PM and ozone exhibit opposite trends.

PM formed rapidly from midnight to early morning while ozone built up rapidly in the afternoon. This implies the importance of designing multi-species control strategies to consider the interaction between different emissions as well as already formed pollutants.

The implementation of HDDM-3D/PM provides a powerful extension to the CMAQ model, as allowing efficient assessment of control strategy effectiveness, source contribution quantification, and model uncertainty analysis. Initial studies show that Taylor series expansions with the second-order term predict the model response to various emission levels very well. HDDM-3D/PM can be easily implemented into other versions of CMAQ, as well as other chemical transport models that already include DDM.

# CHAPTER 3

# QUANTIFICATION OF UNCERTAINTY OF PARTICULATE MATTER SIMULATION IN THE PRESENCE OF UNCERTAIN EMISSION INVENTORIES USING CMAQ/HDDM-PM

## 3.1. Introduction

In the past few decades, effort has been expanded to improve air quality due to its influence on human health and the environment. The United States Environmental Protection Agency (U.S. EPA) sets National Ambient Air Quality Standards (NAQQS) to protect public health and the environment. Fine particles ($PM_{2.5}$) 2.5 micrometers in diameter and smaller, are currently regulated as one of the criteria pollutants by the NAQQS. Their contribution to acid deposition and visibility reduction has also been identified (e.g., Latha and Badarinath, 2005; Galloway et al., 2004; Watson 2002). The harmful effects of PM on human health have been a focus as fine particles can contain toxic substances which are associated with asthma and chronic obstructive pulmonary disease (Zanobetti et al., 2000; Ramachandran and Vincent, 1999; Brauer and Brook, 1997; Schwartz, 1994, Dockery et al., 1993). A recent study (Kaiser, 2005) found that fine particles are potentially of more concern than larger particles in causing respiratory diseases and premature death due to their ability to penetrate deep into the lung. In order to more effectively protect the public from adverse health effects due to exposure to fine particles, in December 2012, U.S. EPA strengthened the annual health NAAQS for fine particles from 15 $\mu g\ m^{-3}$ to 12 $\mu g\ m^{-3}$.

PM control is perhaps the most complex aspect of current air quality management. The complexity comes from the composition of PM and formation routes. Regional air quality models are frequently used in air quality management to evaluate the effectiveness of emissions control strategies that are designed to reduce the exposure of people to particulate matter (U.S. EPA, 2004). The accuracy of these models is limited by the simplified representation of the complex chemical and physical processes as well as the lack of accuracy in model parameters (e.g., chemical reaction rates) and inputs (e.g., emissions rates, meteorological conditions, and initial and boundary conditions). Previous studies have investigated model uncertainties on ozone, focusing on uncertainties due to emission estimates, model structure, initial and boundary conditions, grid sizes, and chemical reactions (e.g., Hanna. et al., 2001, 1998; Cohan et al., 2010; Pinder et al., 2009). Uncertainties in emission inventories remain a leading factor for discrepancies between models and observations (Xiao et al., 2010). Therefore, quantification of the influence of uncertain emission inventories on model output is useful for air quality management processes as well as to guide model improvement.

A framework to estimate the uncertainty of air quality models due to the uncertainties in input parameters has been established by previous research work. Most of the research work has used Monte Carlo simulations with randomly sampled model inputs according to their probability distributions and then quantified the uncertainties of model outputs (e.g. pollutant concentrations and sensitivities) by using the ensemble outputs obtained from the Monte Carlo simulations. Initially, studies conducted the Monte Carlo simulations by running the underlying air quality model multiple times (e.g., Deguillaume et al., 2008; Hanna et al., 2001; Bergin et al., 1999; Hanna et al., 1998).

However, this approach becomes very computationally intensive for three-dimensional time-dependent models applied over large domains. More recently, studies have employed a reduced from model (RFM) approach, which substantially reduces the computational cost (e.g. Napenelok et al., 2011; Tian et al., 2010; Digar and Cohan, 2010; Pinder et al., 2009). This approach constructs a reduced form model of the underlying air quality model by capturing pollutant-parameter responses of the original model. High order direct sensitivity analysis is efficient at extracting the pollutant-parameter response by simultaneously providing first- and second-order sensitivity coefficients along with concentration simulation. This advanced sensitivity technique has been implemented in air quality models [e.g., CMAQ (Hakami et al., 2003), CAMx (Cohan et al., 2010)] for gas species. It has also been applied in order to characterize the uncertainty of modeled ozone production (Napenelok et al., 2011; Tian et al., 2010; Digar and Cohan, 2010; Pinder et al., 2009) and to investigate the influence parameter uncertainties on ozone sensitivities (Xiao et al., 2010; Cohan et al., 2010). Similar systematic studies can also be carried out to evaluate the uncertainties of modeled $PM_{2.5}$ species by implementing high order decoupled direct sensitivity analysis in the CMAQ model (Zhang et al., 2012).

This chapter discusses application of the reduced form model based on CMAQ to efficiently quantify the uncertainties of the simulated $PM_{2.5}$ sensitivities for an air pollution episode in the Houston region. Concentrated petrochemical plants in this region, as well as the unique geographic and meteorological conditions, lead to complex pollutant-emissions responses that need further understanding. Highly nonlinear ozone formation has been found in this region, and the accuracy of emission inventories can

strongly influence prediction of ozone response to emission controls (Xiao et al., 2010). It was reported that the ethene and propene emissions from the point sources in the Houston Ship Channel had been underestimated by a factor of 10 or more in the original emission inventories (Ryerson et al., 2003; Wert et al., 2003). It was also found that mobile emissions of $NO_x$ had been underestimated (Cowling and Parrish, 2007). Given the large uncertainty in emission inventories of this region, this dissertation present an approach to evaluate the impact of uncertain emission inventories on particulate concentrations using high-order DDM sensitivity analysis.

## 3.2. Method

### 3.2.1 Modeling system

Air quality modeling is conducted using CMAQ version 4.7.1 (CMAQ v4.7.1) with the SAPRC 99 (Carter, 2000) chemical mechanism and the AE5 aerosol module. The AE5 aerosol module reflects state-of-the-art aerosol sciences (Foley et al., 2010; Carlton et al., 2010; Kelly et al., 2009; Davis et al., 2008), which is suitable for studying the emission-associated uncertainty of $PM_{2.5}$ in the sense that it helps decrease the uncertainties of model representation. CMAQ v4.7.1 has been equipped with the Decoupled Direct Method in Three Dimensions (DDM-3D) (Napelenok et al., 2008), which has been extended to high-order DDM-3D for particulate matter (HDDM-3D/PM) to account for any nonlinear response of pollutants to model inputs (Zhang et al., 2012)

The CMAQ model application here uses three one-way nested modeling domains. The outer-most domain covers the entire continental United States and portions of Canada and Mexico with 36- by 36-km horizontal grids; The middle domain covers the

East of Texas and the surrounding states of Oklahoma, Arkansas, and Louisiana with 12 by 12 kilometer grids; The inner-most domain covers Southeastern Texas which contains the Houston-Galveston-Brazoria region where intense emissions from petrochemical industries occur. The three domains have 13 vertical layers extending approximately 16 km above ground, with seven layers below 1 km.

The Weather Research and Forecasting (WRF) model is the state-of-the-art meteorological model that is widely used in atmospheric research and weather forecasting. This application used the WRF model to prepare the meteorological fields and is run with 35 levels using four-dimensional data assimilation (FDDA) techniques and the Noah land-surface model with MODIS landuse data. The Sparse Matrix Operator Kernel for Emissions (SMOKE) is used to process emissions to provide gridded, CMAQ-ready emissions. The inventory used is the U.S. National Emissions Inventory (NEI) of 2005 (ftp://ftp.epa.gov/EmisInventory/2005v4/). Figure 3.2 shows the emissions of six major air pollutants categorized by different emission sources from the 2005 NEI.

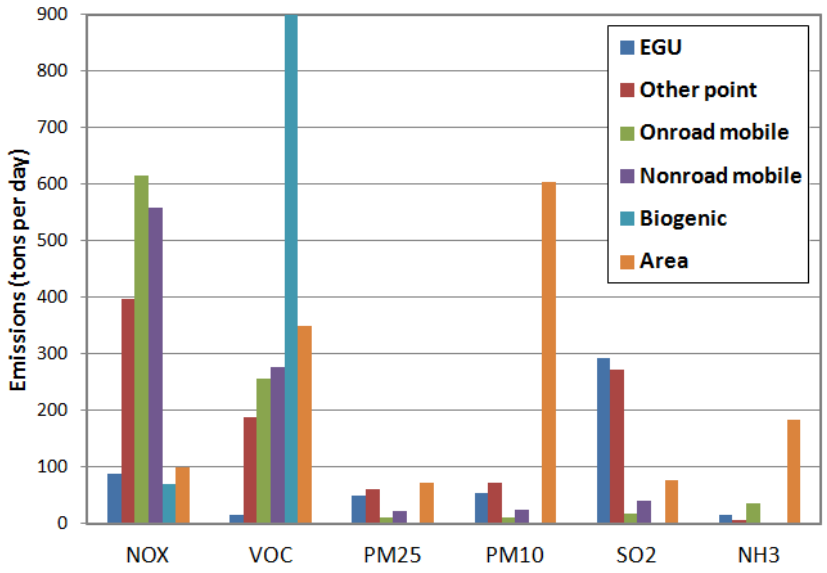


Figure 3.1: Emission rates of six major pollutants categorized by emission sources. The emission rates are the daily averages of the domain-wide emissions.

### 3.2.2 Reduced-form model of CMAQ

The uncertainty analysis of this work is based on a reduced-form model of CMAQ. The reduced-form model represents the relationship between pollutant concentrations and the model inputs in a straightforward way and can be used as an efficient method to propagate uncertainties from model inputs to outputs. Constructing the reduced-form model involves Taylor series expansion of the pollutant concentration at a given time and location, C(x, t), for fractional perturbations in sensitivity parameters of interest (Equation 3.1). The sensitivity parameters could be emissions rates, chemical reaction rates, or initial and boundary conditions. Only emissions rates are considered in this work for the purpose of studying emission-associated model uncertainties.

$$C_i^* = C_{i,0} + \sum_{j=1}^{J} \Delta\varepsilon_j S_{i,j}^{(1)} + \frac{1}{2}\sum_{j=1}^{J} \Delta\varepsilon_j^2 S_{i,j,j}^{(2)} + \sum_{j\neq k} \Delta\varepsilon_j \Delta\varepsilon_k S_{i,j,k}^{(2)} + H.O.T. \quad (3.1)$$

In Equation 3.1, $C_i^*$ and $C_{i,0}$ denote the concentration of pollutant $i$ with and without perturbations in sensitivity parameters, respectively. $S_{i,j}^{(1)}$, $S_{i,j,j}^{(2)}$, and $S_{i,j,k}^{(2)}$ are semi-normalized sensitivity coefficients (Equation 2.7). $i$ denotes the $i^{th}$ species, $j$ and $k$ denotes the $j^{th}$ and $k^{th}$ emissions rates. *H.O.T.* stands for higher order terms.

CMAQ-HDDM-3D is used to calculate local first- and second-order semi-normalized sensitivity coefficients. This approach is efficient in that it simultaneously computes sensitivity coefficients with respect to all chosen parameters along with concentrations in one model run. The controlling equations for sensitivity coefficients are derived by differentiating governing equations for the concentrations with respect to the sensitivity parameters. Thus equations involving sensitivities and concentrations have a similar structure and are calculated using the same numerical algorithms. First- and

second-order sensitivity coefficients calculated by CMAQ-HDDM-3D have been evaluated by comparing them with traditional finite differences and good agreement has been observed for both gas-phase species and particulate matter (Zhang et al., 2012; Hakami et al., 2003).

### 3.2.3 Quantification of emission-associated uncertainties

Monte Carlo simulations using the reduced-form CMAQ are applied to quantify the emission-associated uncertainties of modeled $PM_{2.5}$. Three steps are involved in the Monte Carlo simulations. The first step is to sample the emissions rates of interest, given their uncertainties and the associated probability distributions. The second step is to propagate uncertainties through the reduced-form model and then construct an ensemble of model outputs. The third step is to quantify model uncertainties from the output ensemble.

Emissions rates are assumed to be log-normal probability distributions, as is approximately found for many environmental geographical variables that are constrained to be positive (Hanna et al., 1998). The uncertainties for lognormal distributions are usually expressed as an uncertainty factor which would include 95% of the possible values. If $X$ is log-normally distributed with an uncertainty factor of $Y$, the standard deviation of $\ln(X)$ (i.e., the $\sigma$ for the lognormal distribution) equals to $0.5\ln(Y)$ (Hanna et al., 2001). For example, if the daily average $PM_{2.5}$ concentration, $C_{PM_{2.5}}$, is assumed to have an uncertainty factor of 3, the standard deviation of $\ln(C_{PM_{2.5}})$ is $0.5\ln(3) = 0.55$. This study focuses on the emissions of five major pollutants that can raise atmospheric PM levels: $SO_2$, $NO_x$, VOC, $NH_3$, and primary PM. The uncertainty factors of these

emissions are obtained by reviewing previous studies. Hanna et al. (2001) summarized the estimates of uncertainty factors for $NO_x$ and VOC emissions from major point, mobile, biogenic, and area sources. The NARSTO PM assessment (2004) provides the confidence levels of national emission inventory for $SO_2$, $NO_x$, VOC, $NH_3$, and primary PM from various source categories. Combining the confidence levels in the NARSTO assessment with the uncertainty factors in Hanna et al., the uncertainty factors of the emissions rates of the five pollutants of interest can be estimated. Table 3.1 summarizes the uncertainty factors developed for the five pollutant emissions.

Table 3.1. Uncertainty factors and associated σ (standard deviations of log-transformed data) of emission rates of five major pollutants.

| Emissions | Uncertainty Factor | σ |
|---|---|---|
| SO2 | 1.62 | 0.243 |
| NOx | 1.67 | 0.258 |
| VOC | 2.11 | 0.373 |
| NH3 | 2.74 | 0.505 |
| Primary PM | 2.71 | 0.500 |

Simple random sampling is applied to produce multiple sets of possible input emissions rates given their probability distributions and uncertainties. This study selected a sampling size of 1000, which has been demonstrated to achieve sufficient convergence in the uncertainty analysis on ozone simulation conducted by Pinder et al. (2009). The results of the sampling are shown in Figure 3.2 in the form of relative uncertainty.

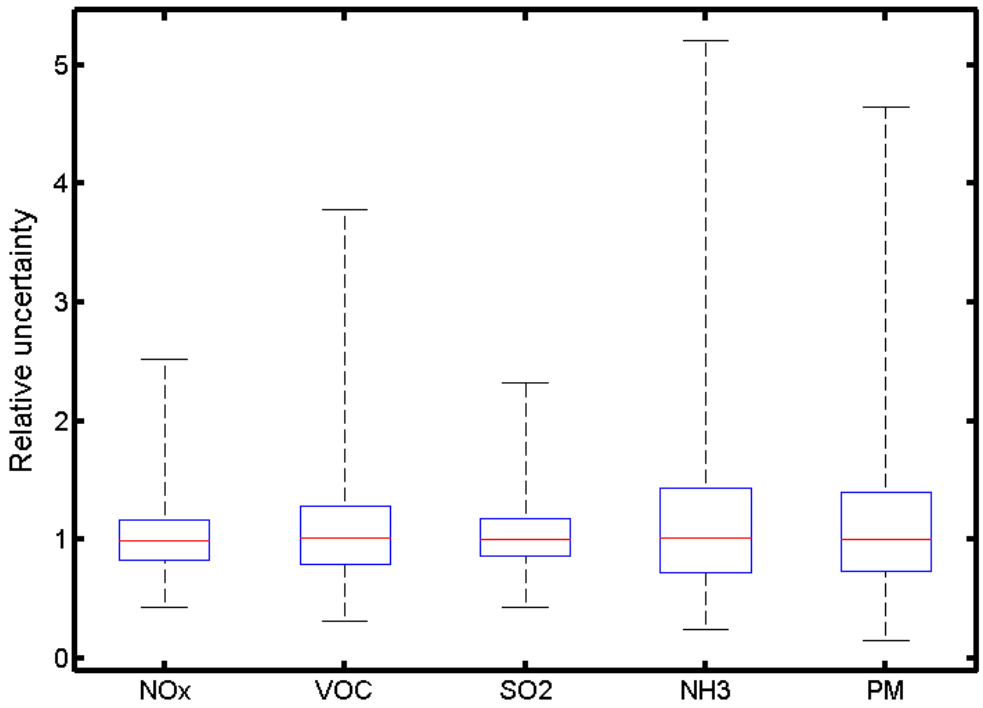


Figure 3.2 Ratios of sampling results to baseline emissions. The red line in the center of the box indicates the median of samples. The top and bottom of the box indicate the 75% and 25% percentiles, respectively. The whiskers indicate the maximum and minimum values in the corresponding bins.

The samples are used to drive the reduced-form model and obtain an ensemble of Monte Carlo simulation results. For every grid cell and time step, ensembles with 1000 values of PM concentrations are generated. The inferred coefficient of variance (ICOV) defined by Tian et al. is used to quantify the uncertainty of the simulated PM concentrations (2010). It is defined as a half of the 68.3% confidence interval divided by the median of the corresponding ensemble (Equation 3.2). This definition is based on the confidence interval and using it is a more robust method than directly using the coefficient of variance since applying the inferred coefficient can exclude outliers due to the bad performance of the reduced form model with extremely large emission changes.

$$ICOV = \frac{68.3\%CI}{2 \times median} = \frac{84.15^{th}\ percentile - 15.85^{th}\ percentile}{2 \times median} \qquad (3.2)$$

## 3.3. Results and Discussion

### 3.3.1 Model evaluation

Surface meteorological fields simulated by WRF are evaluated by using hourly surface observations in the U. S. and Canada. The bias and root mean square errors (RMSE) for the three domains are within the range considered to be acceptable for air quality modeling (Table 3.1) (Emery et al., 2001; Hanna and Yang, 2001). Good agreement between the WRF simulation and observations helps minimize uncertainty due to input meteorological fields. However, although the episode average performance is good, the WRF simulation had a higher RMSE for temperature, wind direction, and relative humidity in the last week of the episode. This may lead to higher uncertainty in $PM_{2.5}$ simulation in that period. As described in Chapter 2, the formation of $PM_{2.5}$ in this region is favored by low temperature, high relative humidity, and weak northerly winds, so all those three meteorological variables can strongly impact the formation of $PM_{2.5}$.

CMAQ performance is evaluated using the Air Quality System (AQS) observational data. Simulated daily averaged $PM_{2.5}$ concentrations are compared with the Air Quality System (AQS) monitoring data inside the 4 km domain. $PM_{2.5}$ simulation is commonly evaluated by using mean fractional bias (MFB) and mean fractional error (MFE) which are -30% and 54% for this simulation, within the acceptable range according to the guidance provided by EPA (2007). Daily averages of sulfate, nitrate, ammonium, and organic carbon (OC) are underestimated by 64.7%, 122%, 59.3%, and 25%, respectively (Table 3.3). EC is slightly overestimated, with a MFB of 46.8%.

Table 3.2. Evaluation of WRF-generated meteorological fields from Aug 10 to Sep 14, 2006 with the Techniques Development Laboratory (TDL) surface observations.

| | **Surface Wind Speed** | | **Surface Wind Direction** | | **Surface Air Temperature** | | **Surface Humidity** | |
|---|---|---|---|---|---|---|---|---|
| **Model Domain Resolution** | **Bias ($m\ s^{-1}$)** | **RMSE* ($m\ s^{-1}$)** | **Bias (deg.)** | **Gross Error (deg.)** | **Bias (K)** | **RMSE (K)** | **Bias ($g\ kg^{-1}$)** | **Gross Error ($g\ kg^{-1}$)** |
| **36km** | 0.23 | 1.76 | 2.49 | 32.78 | -0.43 | 1.96 | 0.44 | 0.92 |
| **12km** | 0.53 | 2.27 | 14.90 | 52.08 | -0.92 | 3.34 | -0.13 | 1.18 |
| **4km** | 0.47 | 1.91 | 10.50 | 53.17 | 0.63 | 2.29 | 0.29 | 1.46 |

*RMSE: root mean square error

Table 3.3. Evaluation of CMAQ-simulated concentrations of $PM_{2.5}$ species by comparison with the AQS observational data from August 10 to September 14 in 2006.

| **PM Species** | **Number of comparison pairs** | **Mean of concentration ($\mu g\ m^{-3}$)** | **Mean Bias ($\mu g\ m^{-3}$)** | **Normalized Mean Bias (%)** | **Mean Error ($\mu g\ m^{-3}$)** | **Normalized Mean Error (%)** | **Mean Fractional Bias (%)** | **Mean Fractional Error (%)** |
|---|---|---|---|---|---|---|---|---|
| **PM25_daily** | 106 | 15 | -2.5 | -16.93 | 6.9858 | 46.55 | -30 | 54.37 |
| **SO4_daily** | 15 | 5.5 | -2.5 | -46.3 | 2.6367 | 48.03 | -64.68 | 66.08 |
| **NO3_daily** | 11 | 0.5 | -0.4 | -68.27 | 0.4269 | 79.21 | -121.98 | 131.64 |
| **NH4_daily** | 15 | 2.4 | -1.2 | -52.59 | 1.2742 | 52.89 | -59.35 | 60.16 |
| **EC_daily** | 16 | 0.5 | 0.3 | 62.5 | 0.3294 | 66.33 | 46.84 | 50.59 |
| **OC_daily** | 16 | 3.2 | -1.1 | -34.95 | 1.6129 | 50.48 | -25.08 | 53.85 |
| **PM25_hourly** | 10710 | 14.2 | -3.6 | -25.28 | 8.0929 | 56.91 | -39.52 | 69.06 |

### 3.3.2 Uncertainty of modeled concentrations

Using the sampling results in Figure 3.2 to drive reduced-form CMAQ gives an ensemble of pollutant concentrations for each grid cell at every time step. In this study, daily average of $PM_{2.5}$ concentrations is assessed because that is the metric used for air quality standards and policy making. Since the response of $PM_{2.5}$ to precursor emissions varies spatially and temporally, the median and uncertainty (i.e., ICOV) are calculated for 171864 ensembles ranging across the entire modeling domain over 30 days in the episode. The ratios of the medians and the corresponding simulated $PM_{2.5}$ concentrations are close to one, indicating that the simulations of $PM_{2.5}$ by the deterministic model (i.e.,

CMAQ) agree well with the most probable values given uncertainty in emissions rates. The average of all the uncertainties is 36%, which is to say, that with the current estimate of uncertainties in the emissions inventory, on average, there is a 36% uncertainty in $PM_{2.5}$ simulations due to uncertain emission rates. The medians of uncertainties for different levels of $PM_{2.5}$ fall into a narrow range from 42% to 52%, with a slight increase in $PM_{2.5}$ levels (Figure 3.3). Bins with lower $PM_{2.5}$ concentrations have more variability since there are more data points and the chance that they have high sensitivities is higher. Higher sensitivity leads to a larger standard deviation and thus the uncertainty becomes high after the standard deviation is divided by a low concentration value. This has been confirmed by comparing the spatial distribution of uncertainty and concentration. Although the gradient in $PM_{2.5}$ concentrations can be clearly seen, most areas have uncertainties around 40% (Figure 3.4). Hot spots in the concentration simulation do not correspond to high uncertainty. Instead, uncertainty distributes more evenly over the entire domain, indicating that CMAQ has similar accuracy over a wide range of $PM_{2.5}$ concentrations.

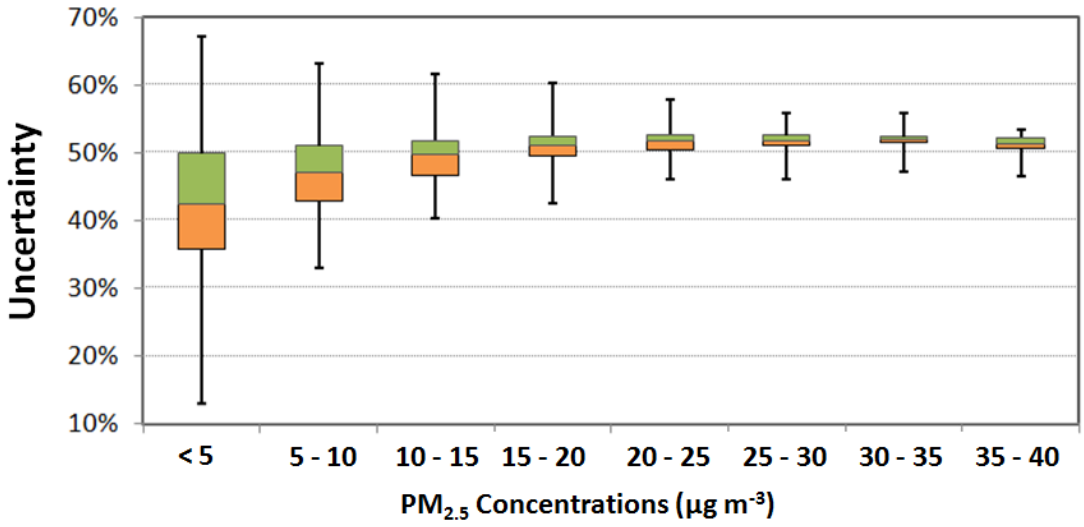


Figure 3.3. Uncertainty (i.e., ICOV) in simulated $PM_{2.5}$ concentrations for the entire modeling domain. The box shows median, 25% and 75% percentiles. The line between the green and orange boxes represents the median. The whiskers indicate a 95% confidence interval.

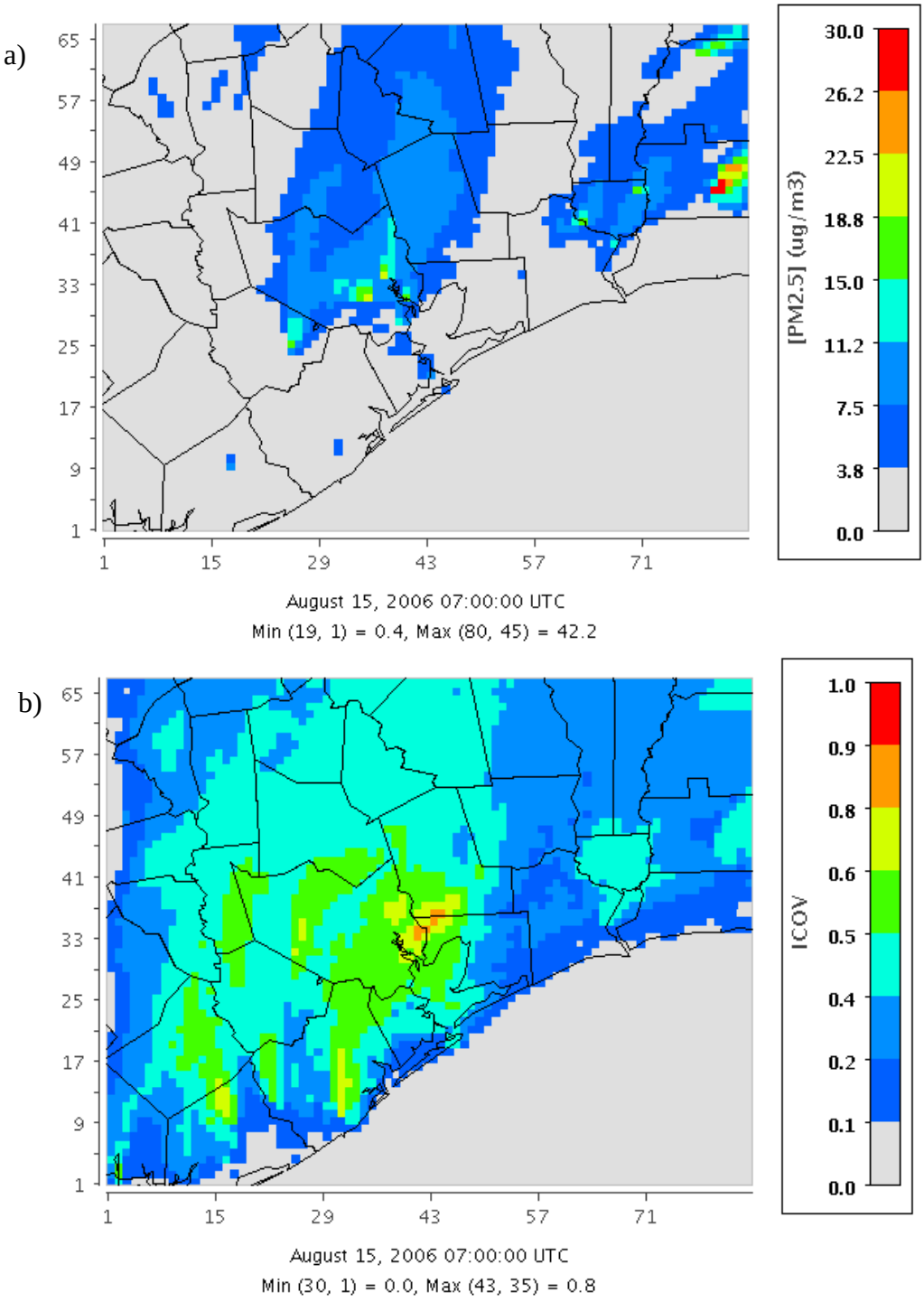


Figure 3.4. Spatial distribution of a) daily average $PM_{2.5}$ concentrations and b) uncertainty (ICOV) on August 15, 2006.

The individual impacts of the five major emissions rates on simulated $PM_{2.5}$ concentration are calculated by propagating the uncertainty of only one emission rate and

keeping the others unchanged. The dominant contributor is the uncertainty associated with primary PM, with an ICOV of 38%, which means the estimated uncertainties in emissions of primary PM alone can lead to a 38% uncertainty in simulated $PM_{2.5}$ averaged over the domain. The other emission rates have relatively less impact, with ICOVs of 3%, 2%, 2%, and 2% for $NH_3$, $SO_2$, $NO_x$, and VOC emissions, respectively. One reason for the high impact from primary PM emissions is their high uncertainty factor, which is the second largest of the five emissions rates. Another reason is the high sensitivity of $PM_{2.5}$ to primary PM emissions. The domain-wide contribution of primary PM to 24-hr average $PM_{2.5}$ concentrations is 50%. Primary PM emissions in Southeast Texas come from various sources, such as mobile, industrial, biomass burning, and residential sources. Their significant contribution to $PM_{2.5}$ mass in this region has been summarized by Allen and Fraser (2006). Reducing the uncertainty in primary PM emissions would be an effective way to bring down the uncertainty in modeled $PM_{2.5}$ concentrations. For example, assuming the uncertainty factor of primary PM emissions drops from 2.71 to 2, the mean of episode and domain-wide uncertainty in daily $PM_{2.5}$ simulation drops from 36% to 24%.

### 3.3.3 Comparison with observations

The uncertainty associated with uncertain emission rates is used to investigate the bias between model simulation and ground measurement of $PM_{2.5}$ concentrations. Five observation sites were selected in the Houston Ship Channel region (Figure 3.5). The five sites have continuous $PM_{2.5}$ monitoring and a diversity of land use types, including urban, suburban residential, agricultural, and industrial. They are located to the south, north, and

west of the Ship Channel, so they are able to represent various impacts from the areas with intense industrial emissions.

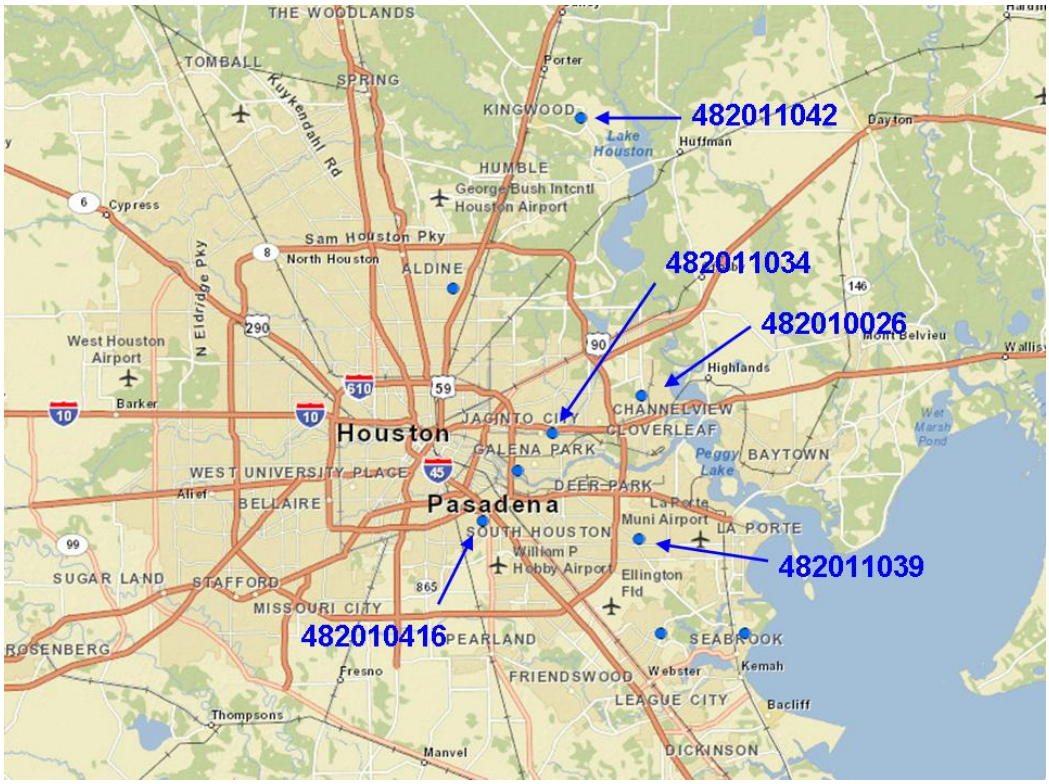


Figure 3.5. Five AQS monitoring sites selected for comparison with $PM_{2.5}$ simulation.

Daily averages of modeled and observed $PM_{2.5}$ concentrations at the five sites are compared for 34 days. The time series of the comparisons are shown in Figure 3.6. The error bars represent the emission associated uncertainties, expressed as the standard deviation obtained in Section 3.3.2. The dashed lines correspond to the 95% confidence interval (CI), which are obtained by calculating the 2.5th and 97.5th percentiles of the ensemble results. For all five sites, the percentage of observations that fall in the range of the standard deviations is 60%, and the percentage of observations that fall in the 95% CI is 85% (Table 3.4). Both Houston East (AQS#482011034) and Channelview

(AQS#482010026) have over 90 percent of the observed concentrations falling in the 95% CI, implying that uncertainties in the emission rates can explain much, but not all, of the difference between the simulated and observed $PM_{2.5}$ concentrations at the two sites (Figure 3.6 a and b). The two sites which are south of the Houston Ship Channel (Figure 3.6 c and d) have about 70% of the observations in the 95% CI. Modeled $PM_{2.5}$ concentrations are consistently biased high during the last ten days of the episode. During the same period, meteorological fields experience higher errors and bias than on the other days: Temperature and humidity are biased high, and wind direction shows larger deviation from observations compared to the first half of the episode. The correlations between the root mean square error of the meteorological fields and that of $PM_{2.5}$ concentrations indicate that the error in $PM_{2.5}$ simulation is related more to errors in wind direction ($R^2 = 0.2$) and temperature ($R^2 = 0.03$) than to errors in wind speed and relative humidity. In addition, the low bias in the $PM_{2.5}$ simulation for August 28-30 at all the five sites is due, in part, to errors associated with meteorological fields. The system simulated a precipitation event occurred from August 28 to 30, which came from north and swept the Houston Ship Channel and the heavy rainfall decreased the simulated $PM_{2.5}$ concentration. However, the same reduction is not found in the observed $PM_{2.5}$ concentrations at the five monitoring sites, so the biased-low $PM_{2.5}$ may be attributable to the error in the simulated precipitation and estimated scavenging. The Kingwood site (Figure 3.6 e) exhibits a low bias for $PM_{2.5}$ simulation. The reason may be due to its location, which, unlike the other four sites near the Houston Ship Channel, is 23.7 miles northeast of Houston's downtown and is located in an area with substantial biogenic VOC emissions. Studies have shown that the current air quality model tends to underestimate

secondary organic aerosols at monitoring sites. Here, the comparison between model simulated and observed concentrations also indicates a low bias in simulation. 19 out of 31 days of the observed daily average $PM_{2.5}$ are higher than the upper bound of the standard deviation of the model simulation. This is probably due to model representation of the chemical reactions and thermodynamic partitioning related to the formation of secondary organic aerosols.

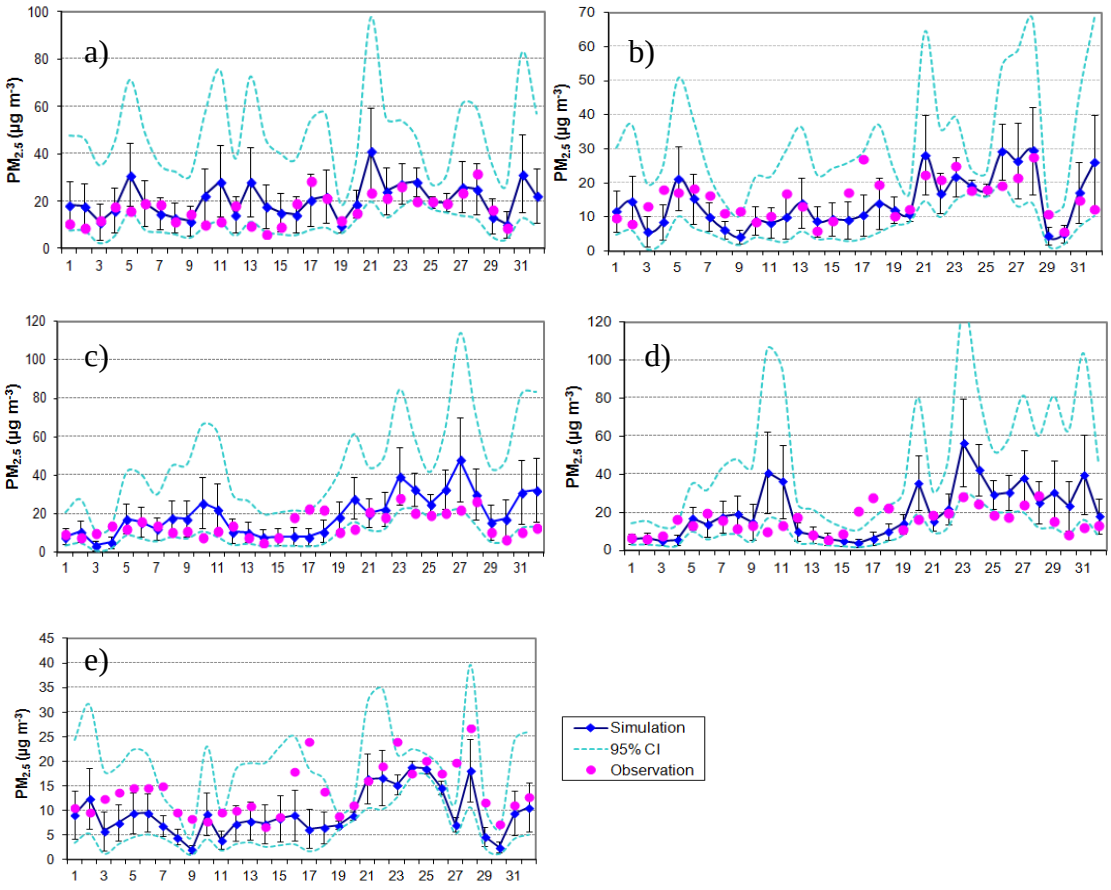


Figure 3.6. Time series of daily average $PM_{2.5}$ concentrations for the five AQS sites: a) Houston East (AQS#: 482011034), b) Channelview (AQS#: 482010026, c) Deer Park (AQS#: 482011039), d) Park Place (AQS#: 482010416), e) Kingwood (AQS#: 482011042).

Table 3.4 Summary of the comparison between simulated and observed daily average $PM_{2.5}$ concentrations.

| Site name | Site number | Number of days within one standard deviation | Fraction of days within one standard deviation | Number of days within 95% CI | Fraction of days within 95% CI |
|---|---|---|---|---|---|
| **Houston East*** | 482011034 | 25 | 83% | 28 | 93% |
| **Channel View** | 482010026 | 22 | 69% | 31 | 97% |
| **Deer Park** | 482011039 | 19 | 59% | 27 | 84% |
| **Park Place** | 482010416 | 16 | 50% | 24 | 75% |
| **Kingwood** | 482011042 | 12 | 38% | 24 | 75% |
| **All Sites** | | 94 | 60% | 134 | 85% |

* This site has 30 days of observations available.

### 3.4. Conclusions

This study quantified the emission-associated uncertainty in an air quality model simulation for $PM_{2.5}$. A reduced form model of CMAQ is constructed based on high-order DDM sensitivity analysis. One thousand possible combinations of emissions rates of five major pollutants are sampled based on log-normal distribution and uncertainty factors estimated from a literature search. The ensemble output from the reduced form model is used to quantify the model uncertainty associated with emission rates. The uncertainty (i.e., ICOV) of modeled 24-hour average of $PM_{2.5}$ concentration is 36% averaged over the modeling domain and episode. The medians of uncertainties for different levels of $PM_{2.5}$ concentrations are around 50%, and have a fairly uniform spatial distribution over the modeling domain.

The emission-associated uncertainty in $PM_{2.5}$ simulation has been used to interpret the comparison between ground measurement and model simulation of daily average $PM_{2.5}$. The simulation captures the trend of the observation. 85% of the observations at five monitoring sites fall into the 95% CI of the corresponding simulation

due to the uncertainty in emission rates. This suggests that much of the difference between the observed and simulated concentrations can be attributed to the uncertainty in emission rates. Besides the uncertain emission rates, meteorological conditions play an important role in air quality modeling in this region due to the frequently changing wind direction near the coast. The biased-low $PM_{2.5}$ simulation north of Houston suggests that there is either uncertainty in the model representation of the formation of secondary organic aerosol or bias in biogenic VOC emission in the inventory.

# CHAPTER 4

# ASSESSING THE IMPACT OF FLARE EMISSIONS ON REGIONAL AIR QUALITY AT LOW OPERATING RATES

## 4.1. Introduction

Flare emissions have recently been of concern to air quality management in the Houston region. Air quality studies in this region have indicated that highly reactive volatile organic compounds (HRVOCs) are associated with rapid ozone formation in the Houston-Galveston-Brazoria (HGB) area (e.g., Ryerson et al., 2003; Murphy and Allen, 2005; Webster, et al., 2007; Pavlovic et al., 2012). HRVOCs include ethylene, propylene, butenes, and 1,3-butadiene. The primary sources of the HRVOCs are petrochemical refineries and chemical manufacturing plants, which produce flare emissions. It is estimated that 60% of the HRVOCs come from flares (Chen et al., 2012). HRVOC emissions from flares are currently calculated using a simple mass reduction method, which converts the vent gas flow ratio to emissions rates by using the destruction and removal efficiency (DRE). In most cases, 98% is used as the DRE for flares operating at their designed conditions and complying with 40 CFR §60.18 (TCEQ 2009b). However, the DRE may drop below 98% when the flare operation is under 40 CFR §60.18 (Allen, 2011).

The DRE is dependent on the operating conditions of a flare. Processes such as start-up, turndown, steaming, and aeration have the potential to reduce the DRE to below 98%. Recent evaluation of flare operations by the Texas Commission on Environmental Quality indicated that the DRE for steam-assisted flares drops dramatically when

combustion zone heating values fall below 250 BTU/scf (equivalent to 9315 KJ/$m^3$; scf: standard cubic foot) and that the DRE for air-assisted flares drops linearly with the air flow rate (Torres et al., 2012). As the HRVOC emissions from a flare are sensitive to the DRE, understanding the relationship between the operating processes of a flare and the local air quality are important for designing efficient emissions control strategies.

To assess the emissions impact of flares on local air quality conditions, this study conducted three-dimensional chemical transport modeling in the Houston region using the Texas 2006 special inventory (SI). Selecting the SI ensures that the temporal variability of flares is well captured. The SI includes hourly monitoring data of $NO_x$ and VOC emissions collected from over 140 industrial point sources in Texas during the second Texas Air Quality Study (TexAQS II) (Parrish et al., 2009) from August 15 through September 15, 2006. Currently, most of the chemical transport models assign a uniformly-distributed temporal profile to industrial point sources, so using the SI better serves the goal of quantifying the impact of flare operating processes on air quality. The chemical transport modeling (CTM) applies the Community Multiscale Air Quality (CMAQ) model, which is one of the most widely used three-dimensional CTM for air quality management. The CMAQ model has embedded a direct sensitivity technique (i.e., the Decoupled Direct Method in Three Dimensions (DDM-3D)), which enables efficient assessment of pollutant response to emissions change. This dissertation uses CMAQ-DDM3D to investigate the impact of combustion efficiency on the flare VOC impact on ozone under different flare operating modes and then constructs a reduced form model of CMAQ by using the results from CMAQ-DDM3D simulations. The reduced form CMAQ was evaluated in previous studies to be consistent with the original CMAQ

model. The reduced form model has the advantage of substantially reducing the computational cost and selecting sensitivity parameters in a flexible manner. In order to reflect the impact of the temporally variable DRE on flare VOC emissions as well as ambient ozone concentrations, the reduced form model includes sensitivities of ozone concentrations to flare VOC emissions at intervals of every two hours of a day.

## 4.2. Method

### 4.2.1 Modeling system

Air quality modeling is conducted using CMAQ version 4.7.1 (CMAQ v4.7.1) with an SAPRC 99 (Carter, 2000) chemical mechanism and the AE5 aerosol module. The AE5 aerosol module reflects state-of-the-art aerosol sciences (Foley et al., 2010; Carlton et al., 2010; Kelly et al., 2009; Davis et al., 2008). CMAQ v4.7.1 can operate with the Decoupled Direct Method in Three Dimensions (DDM-3D) (Napelenok et al., 2008) which has been extended to account for the nonlinear response of pollutants to model inputs (Zhang et al., 2012).

Three one-way nested modeling domains are used. The outer-most domain covers the entire continental United States and portions of Canada and Mexico with 36- by 36-km horizontal grids; The middle domain covers eastern Texas and the surrounding states with 12- by 12-km grids. The inner-most domain covers southeastern Texas which includes the Houston-Galveston-Brazoria (HGB) area where intense emissions from petrochemical plants are located (Figure 4.1). The three domains have 13 vertical layers extending approximately 16 km above ground with seven of the layers below 1 km.

The Weather Research and Forecasting (WRF) model is used to prepare the meteorological fields and is run with 35 levels using four-dimensional data assimilation

(FDDA) techniques and the Noah land-surface model, which uses the MODIS landuse data. The Sparse Matrix Operator Kernel for Emissions (SMOKE) is used to process the emissions to provide gridded, CMAQ-ready emissions. The inventory used is the U.S. National Emissions Inventory (NEI) of 2005 (ftp://ftp.epa.gov/EmisInventory/2005v4/). The hourly VOC and $NO_x$ emissions from the Texas 2006 SI are used to replace the NEI emissions for petrochemical plants and chemical plants.

### 4.2.2 Texas 3006 Special inventory

Hourly emissions of industrial sources from the Texas 2006 special inventory provide an improved database for assessing the impact of flare emissions. The SI was based on the TexAQS II measurements and analyses for a high-ozone episode from August 15 to September 15, 2006. The assembled inventory provides hourly emissions rates of VOC and $NO_x$ from 141 industrial sites, with 49 sites in the HGB area. The reported industrial sources include flares, stacks, cooling towers, and fugitive sources. The episode average of the SI is 37 tons per day, which is comparable to 33 tons per day in the 2005 ozone season day (OSD) point source emission inventory from the NEI. Comparison of the two inventories in other aspects, such as composition and industrial sectors, suggests reasonable agreement between the two (Pavlovic et al., 2009).

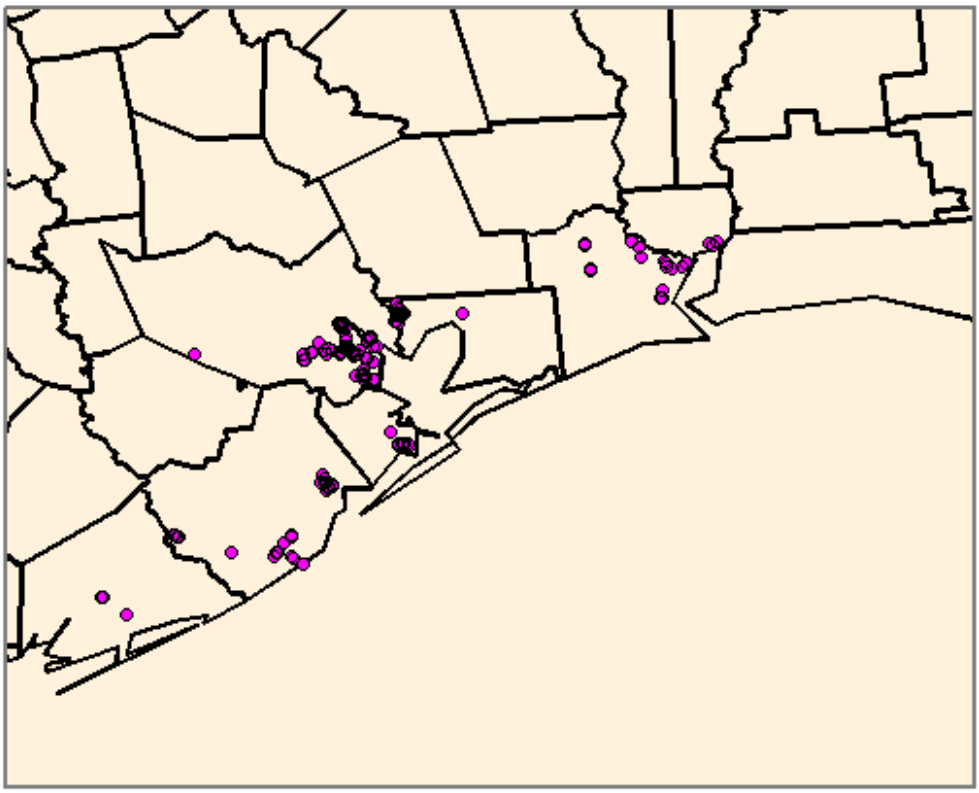

Figure 4.1. Locations of recorded flares in the 2006 Texas special inventory.

In 2006 SI, HRVOC emissions originated primarily from flares. The contributions of flares, stacks and cooling towers, and fugitive sources which make up the total HRVOC emissions are 77%, 18%, and 5%, respectively. Note that the 2006 SI represents only a segment of industrial emissions in the HGB area and does not include some significant sources for HRVOC such as fugitives from tanks and valves. Most of the flares were located around the Houston Ship Channel. Flare emissions are primarily associated with three industrial sectors: chemical manufacturing (55%), plastic materials and resins (26%), and petroleum refining (16%). Those three sectors were responsible for 97% of the HRVOC emissions for the SI and 82% of those for the 2005 OSD.

The emission species and temporal patterns of flares depend on specific industrial sectors and the associated processes. Based on the industrial sectors, flares can be categorized into two major groups: refinery flares and chemical manufacturing flares. The three primary categories for refinery flares are 1) fuel fired equipment, process and natural gas flares, which have emissions of light gases (propane) and gasoline range volatiles (isobutane and n-butane) ; 2) fluid catalytic cracking unit (FCCU) flares, which typically have emissions of propylene, ethylene, propane, isobutene, and isopentane; and 3) unclassified flares, those with no distinct identification with a particular industrial process, which have emissions of propane, pentane, n-butane, isobutene, propylene, and unclassified VOC. Table 4.1 summarizes the classifications of flares and how they are further grouped into sub-categories based on their temporal variability. As will be discussed in Section 4.2.4, the reduced-form air quality model will be used to assess the emission impact and its formulation depends on the model input parameters of interest (here, the VOC emissions rates from flares). Since flares operate at different time schedules, using the domain-wide flare emissions as one parameter does not fulfill the goal of modifying DRE only at low-flow rate conditions. Including every flare as a parameter in the reduced-form model will substantially complicate the formulation as well as increase the computational cost. The parameters in the reduced-form model are on a two-hour basis, so adding one flare corresponds to an extra 24 parameters in the formulation (See further discussion in Section 4.2.4). Thus, here, the focus is on one presenting flare with significant amount of VOC emissions and appropriate location and temporal variability.

The Deer Park refinery and chemical plant is the 11th largest refinery in the United States. Located in the center of the Houston Ship Channel, it is about 20 miles east of downtown Houston and its emission impact on air quality in the urban area is expected to be large. One of the flares located in this refinery (RN100211879_OP3GRFLA) has the highest hourly average VOC emissions (79 lbs/hr) among all the flares located in the Houston Ship Channel. The VOC emission time series of this flare indicate that the time it operates above and below the average mass flow is about half and half (Figure 4.1) and that the hourly mass flow rate can reach over ten times the average. This flare is steam assisted and has been reported to be improperly operated so that excess VOCs are emitted (http://ens-newswire.com/2013/07/10/shell-oil-agrees-to-spend-100-million-on-flare-gas-recovery/). Therefore, this flare serves as a good sample to include in a case study on the impact of flare operating conditions on VOC emissions and thus the regional air quality.

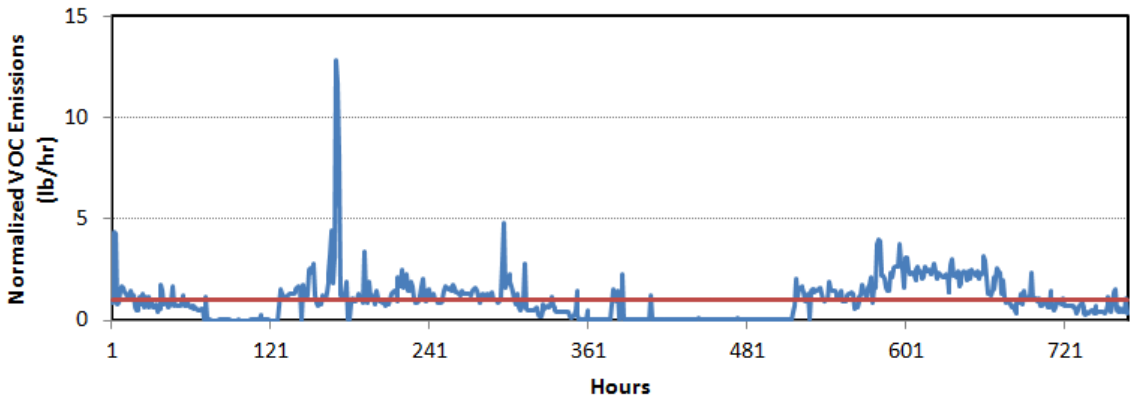


Figure 4.2 Time series of normalized hourly VOC emission rates for flare RN100211879_OP3GRFLA at the Deer Park plant. The labels on the x-axis represent hours starting from 0:00 CST on August 15, 2006. The blue line indicates the normalized VOC emissions. The red line indicates where the normalized hourly emissions equal one.

Table 4.1. Summary of flare categories

| Industrial Sectors | Industrial Process | Temporal Variability | % of the time above the average | Sub Category |
|---|---|---|---|---|
| Refinery flares | Fuel fired equipment, process and natural gas flares | Below 95% | 43-61 | Sub-category I |
| | | Between 95% and 160% | 25-44 | Sub-category II |
| | | Above 160% | 19-28 | Sub-category III |
| | Fluid catalytic cracking unit (FCCU) flares | Below 95% | 32-56 | |
| | Unclassified flares | Below 150% | 31 - 65 | Sub-category I |
| | | Above 150% | 3 - 33 | Sub-category II |
| Chemical manufacturing | Thermal cracking | Below 95% | 36-60 | Sub-category I |
| | | Between 95% and 160% | 16-41 | Sub-category II |
| | | Above 160% | 8-15 | Sub-category III |

4.2.3 Flare VOC emissions and operating conditions

*Flare VOC emissions and Destruction and removal efficiency*

Destruction and removal efficiency (DRE) is defined by Texas Commission on Environmental Quality (TCEQ) to calculate the VOC emissions from a flare (Fortner et al., 2012). The flare VOC emissions are calculated using the DRE defined by the TCEQ as

$$DRE(\%) = \left(1 - \frac{E_{VOC}}{VG}\right) \times 100, \quad (4.1)$$

where $E_{VOC}$ is the mass flow rate of the VOC species found in the flare plume, that is, the emissions rate of VOC species from the flares; $VG$ is the mass flow rate of the vent gas. Equation 4.1 can be used to convert flare VOC emissions into the mass flow rate of the vent gas with the assumed 98% or 99% DRE. This equation will also be used to adjust the flare VOC emissions according to the changes in DRE for different flare operating conditions (Equation 4.2).

$$E_{VOC} = VG \times (1 - \frac{DRE}{100}), \quad (4.2)$$

Another value used to calculate flare VOC emissions is combustion efficiency which is defined as the percentage of the total hydrocarbon stream entering the flare that burns completely to form only carbon dioxide and water. Combustion efficiency is assumed to be lower than DRE since DRE is simply the destruction of the starting hydrocarbon mass. Both combustion efficiency and DRE will be used in this study.

*DRE of steam- and air-assisted flares at low flow conditions*

Recent studies have indicated that the DRE of a flare is affected by steaming and air injection. Torres et al. (2012) carried out measurement and modeling work to characterize the VOC emissions for steam-assisted and air-assisted flares. A series of tests with full-scale industrial flares indicated that the DRE for steam-assisted flares drops dramatically when combustion zone net heating values (CZNHV) fall below 250 BTU/scf. The CZNHV is a function of the vent gas mass flow rate (VG):

$$CZNHV = \frac{(VG \times LHV_{VG})(385.3/MW_{VG}) + (PG \times LHV_{PG})(385.3/MW_{VG})}{(VG \times LHV_{VG})(385.3/MW_{VG}) + (PG \times LHV_{PG})(385.3/MW_{VG}) + S \times (385.3/18.02)} \quad (4.3)$$

where $LHV_{VG}$ is the vent gas lower heating value; $MW_{VG}$ is the vent gas molecular weight; $PG$ is the pilot gas mass flow rate, which is assumed to be constant as 10 lb/hr; $LHV_{PG}$ is the pilot gas lower heating value; $MW_{VG}$ is the pilot gas molecular weight; $S$ is the total steam mass flow rate.

The CZNHV is the key value to determine the DRE. In a presentation to Southeast Texas Photochemical Modeling Technical Committee, Smith (2012) proposed the following fitted model of the relationship between DRE and CZNHV:

$$DRE = (1 + 288.8e^{-0.0417 CZNHV})^{-1} \quad \text{for steam flare with center steam} \quad (4.4)$$

$$DRE = (1 + 5.0181e^{-0.0229 CZNHV})^{-1} \quad \text{for steam flare without center steam} \quad (4.5)$$

Unlike with the steam-assisted flares, the DRE of air-assisted flares drops linearly with the stoichiometric ratio, which is the ratio of the actual amount of air assist to the amount of stoichiometric air required for combustion of the vent gas. The relationship between DRE and the stoichiometric can be expressed as the following linear model:

$$DRE = 1.00842 - 0.003485 \times Ex\_air - 0.0000947 \times Ex\_air^2 + 0.000001447 \times Ex\_air \times LHV \quad (4.6)$$

where $Ex\_air$ is the stoichiometric ratio and $LHV$ is the lower heating value of the vent gas.

4.2.4 Emission scenarios for different flare operating conditions

The purpose of designing emissions scenarios is to simulate the impact of DRE/combustion efficiency on regional air quality. As summarized in Table 4.1, the mass flow of vent gas to a flare is subject to temporal variability. A flare can achieve the assumed DREs (i.e., 98% for C4 HRVOC and 99% for C2 and C3 HRVOC) if the flares operate at certain designed values. However, depending on the associated industrial processes, flares may have routinely or episodic time periods operating with low mass flow rates. For those periods of time, the DREs should be adjusted to be consistent with the low operating rates.

In order to assess the impact combustion on regional air quality, we first assume that the combustion efficiency of a flare is more dependent on its operating conditions (i.e., the amount of steam and inert gas added) under high turndown ratio than under low turndown ratio. We then design three hypothetical flares to represent three typical operating modes of flares. A single combustion efficiency curve is applied to the three hypothetical flares. Flare 1 represents the type of flares that are in continuous use without gas recovery (Figure 4.3a). The flare gas recovery system can minimize flaring by recovering and recycling waste gases and limit flare operation to emergency releases and scheduled maintenance. Flare 2 represent the type of flares are in continuous use but with partial gas recovery or other project that limits the flow to flare (Figure 4.3b). Flare 3 represents the type of flares that are in intermittent use with full gas recovery and flaring 5-10% of the time when the capacity of flare gas recovery unit is exceeded due to start-up, shut-down, or a malfunction (Figure 4.3c).

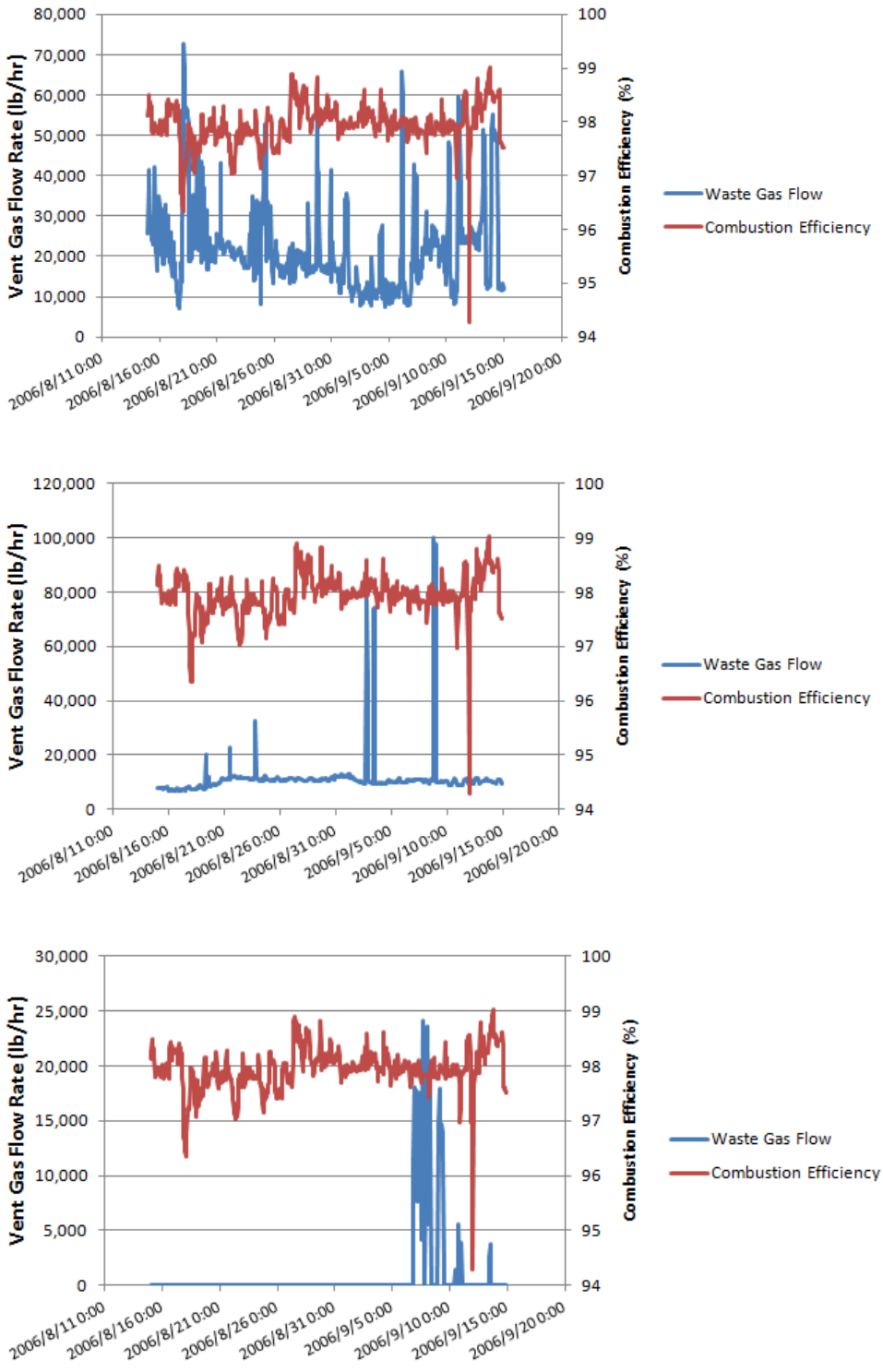


Figure 4.3. Vent gas mass flow rate (blue line) and combustion efficiency (red line) for a) flare 1, b) flare 2, and c) flare 3.

For each of the three hypothetical flares, the base case is designed with the flare VOC emissions calculated using the constant 98% combustion efficiency whereas the control case is designed with flare VOC emissions calculated using the presumed combustion efficiency as shown in Figure 4.3. For each hour, the flare VOC emissions are modified if the combustion efficiency for that hour falls below 98%. Next, the flare VOC emissions of the base and control cases are applied to the air quality modeling to examine the difference in ozone concentration due to the difference in combustion efficiency. In order to minimize numerical noise, we adapt a DDM approach by directly computing the source contribution of the difference in flare VOC emissions to ozone concentration (Figure 4.4).

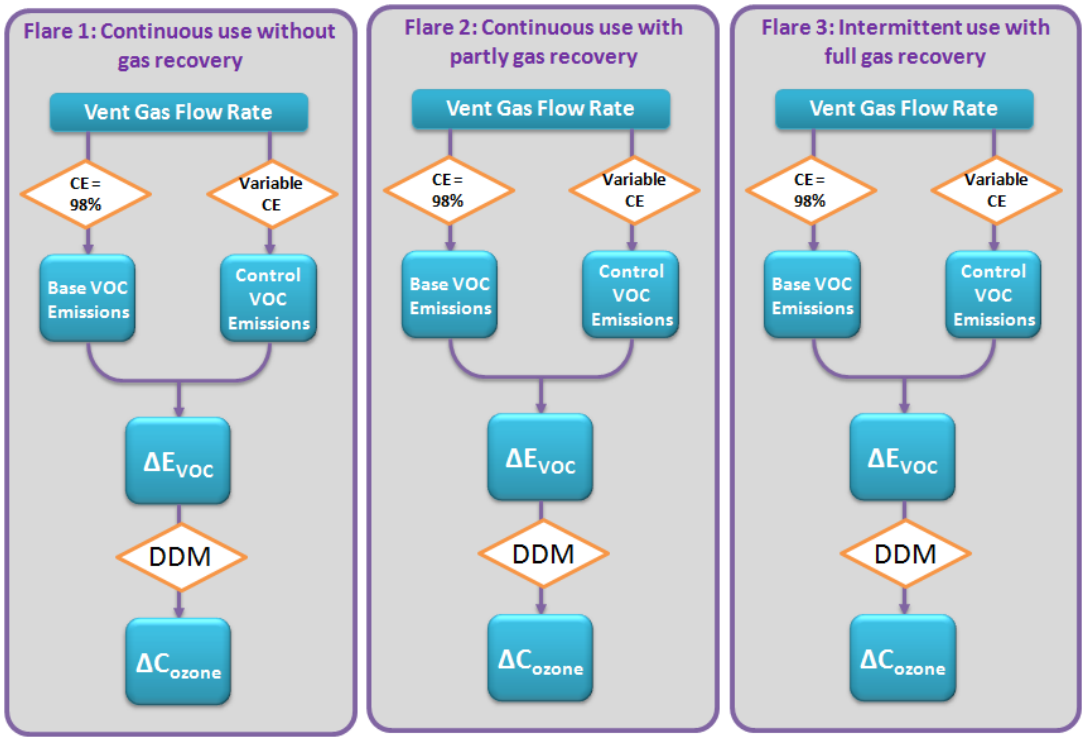


Figure 4.4. Emission scenarios and steps to use CMAQ-DDM3D to assess the impact of combustion efficiency on ozone concentration.

Beside the hypothetical flares, another set of emissions scenarios is designed based on the operation data of a real flare. As discussed in Section 4.2.2, the flare located in the Deer Park plant (Special inventory flare ID: RN100211879_OP3GRFLA) is selected due to its location and the amount of VOC emitted during the modeling episode. Two emission scenarios are designed for this flare. The base case uses the original flare VOC emissions from the 2006 Texas special inventory, which assumes that the DRE is constant at 98%. The control case estimates the DRE using the flow rate of assist steam obtained from a recently developed emission inventory by TCEQ for flare study (Figure 4.5), and the flare VOC emissions at every hour are modified if the re-estimated DRE falls below 98%. The response of ozone concentration to the change in DRE at each hour is estimated by a reduced form model of CMAQ which accounts for the contribution of flare VOC emissions to ozone sensitivities at every two hours. The detailed steps of designing emission scenarios for this part are shown in Figure 4.6.

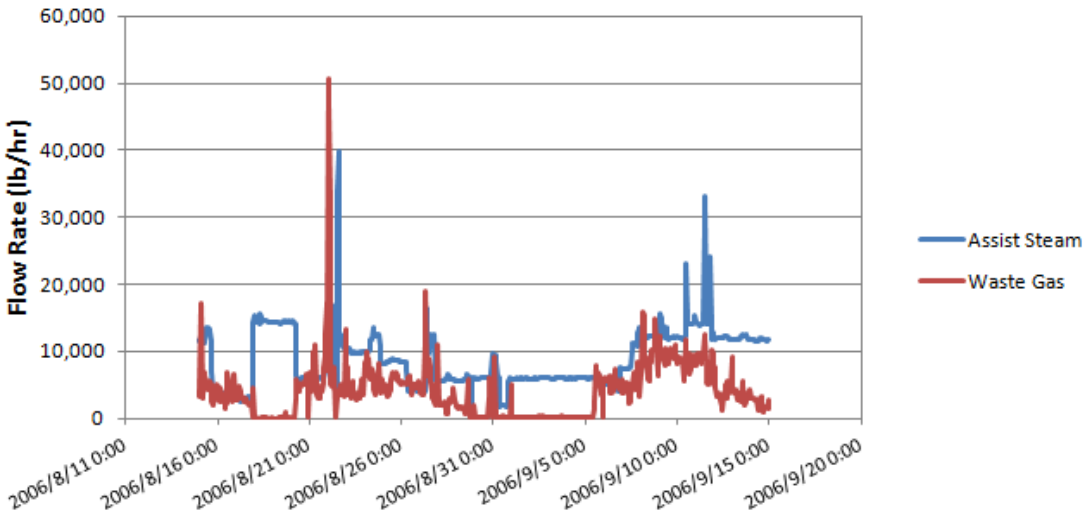


Figure 4.5. Time series of flow rates of waste gas (red) and assist steam (blue) to flare RN100211879_OP3GRFLA in TCEQ special inventory 2006.

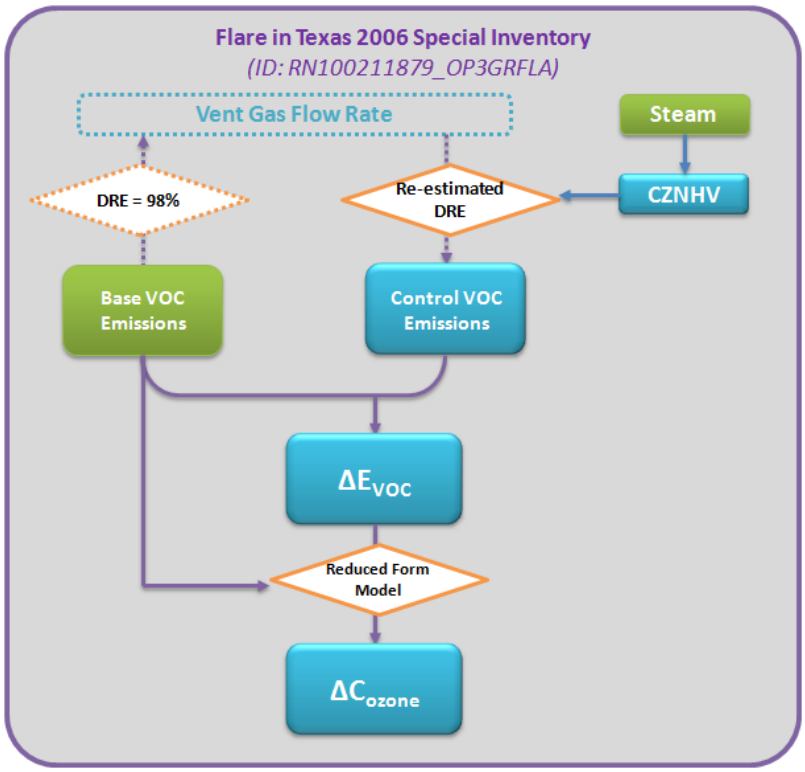


Figure 4.6. Emissions scenarios and steps to assess the impact of DRE on ambient ozone concentrations using the flare VOC emissions from 2006 Texas special inventory and reported flow rate of assist steam for the flare located at Deer Park plant (SI ID: RN100211879_OP3GRFLA).

4.2.5 Formulation of reduced-form CMAQ

The reduced form CMAQ is based on the sensitivity parameters that are obtained from DDM sensitivity analysis. Sensitivity parameters represent the response of model output (i.e., pollutant concentrations) to changes in input parameters, such as emissions rates, initial and boundary conditions, and chemical reaction rates. DDM sensitivity analysis is an advanced technique which is able to simultaneously calculate the local sensitivity coefficients with respect to every input parameter of interest, along with the simulation of pollutant concentrations. Conveniently used in air quality modeling, the semi-normalized sensitivity coefficients are defined as

$$S_{i,j} = P_j \frac{\partial C_i}{\partial p_j} = P_j \frac{\partial C_i}{\partial (\varepsilon_j P_j)} = \frac{\partial C_i}{\partial \varepsilon_j}, \tag{4.7}$$

where $S_{i,j}$ denotes first-order semi-normalized sensitivity coefficients of pollutant $i$ to input parameter $p_j$, $P_j$ denotes the unperturbed base value of $p_j$, $C_i$ denotes concentration of pollutant $i$, and $\varepsilon_j$ is a unitless scaling factor. The advantage of using this definition is that sensitivity parameters have the same unit as concentrations, which better supports the application of sensitivity coefficients in emission assessment. With the semi-normalized sensitivity coefficients, the reduced form model can be written as

$$C_i^* = C_{i,0} + \sum_{j=1}^{J} \Delta\varepsilon_j S_{i,j}^{(1)} + 0.5 \sum_{j=1}^{J} \left(\Delta\varepsilon_j\right)^2 S_{i,j}^{(2)}, \tag{4.8}$$

where $C_i^*$ and $C_{i,0}$ denote the concentrations of pollutant $i$ with and without perturbations in all interested sensitivity parameters , respectively. $\Delta\varepsilon_j$ is the fractional perturbation in $p_j$.

In this study, the reduced form model is formulated by using sensitivities of ozone concentrations to VOC emissions from the Deer Park flare at every two hours. As shown in Figures 4.3 and 4.5, the flare operation does not follow a periodic pattern. Attempting to keep the reduced-form model in a flexible formulation, since changing the formulation of the reduced-form model requires substantial computational cost, leads to including sensitivities of pollutant concentrations to flare VOC emissions every two hours in a day. This formulation allows flexible modification of emissions during any short period of time. Changes in the emission rates are applied to the corresponding time periods. The formulation also allows attribution of pollutant concentrations back to flare emissions at any past time. We can identify, for example, which time period during a day the flare emissions have the largest contribution to the daily maximum 8-hour average ozone concentration. Cutting the day into 2-hour short periods fulfills the goal of capturing the temporal variability and is more computational efficient, saving half of the simulation time required by computing sensitivities for every hour. Thus, there are 24 sensitivity parameters in Equation 4.8 and $\Delta\varepsilon_j$ represents the fractional change of the flare VOC emissions during the corresponding two-hour period. $\Delta\varepsilon_j$ is prepared in the emission scenarios discussed in Section 4.2.4.

## 4.3. Results and Discussion

### 4.3.1 Air quality impact of variable combustion efficiency on three flare operating modes

Three hypothetical flares are studied here to represent the three primary flare operating modes: 1) Continuous use with routine fluctuation, 2) Continuous use with some flare gas recovery, and 3) Intermittent use with full flare gas recovery and flares 5-

10% due to start-up/shut-down/maintenance. A single combustion efficiency curve is applied to the three hypothetical flares, presuming that the CZNHV is affected more by the way the flare is operated (i.e. the amount of the gas or steam that is added) than by the flow rate of the vent gas under high turndown (low mass flow) conditions. The emissions rates of flare VOC is calculated using Equation 4.2. The base case flare VOC emissions are calculated using the constant 98 % combustion efficiency. The control case modified flare VOC emissions at the hours with lower than 98 % combustion efficiency according to the presumed curve. Figure 4.7 shows the time series of flare VOC emission rates of the base and control cases and the differences between the two for the three hypothetical flares. The chance that lowered combustion efficiency has impact on the intermittent use flare is lower than that change that it has an impact on the two continuous use flares.

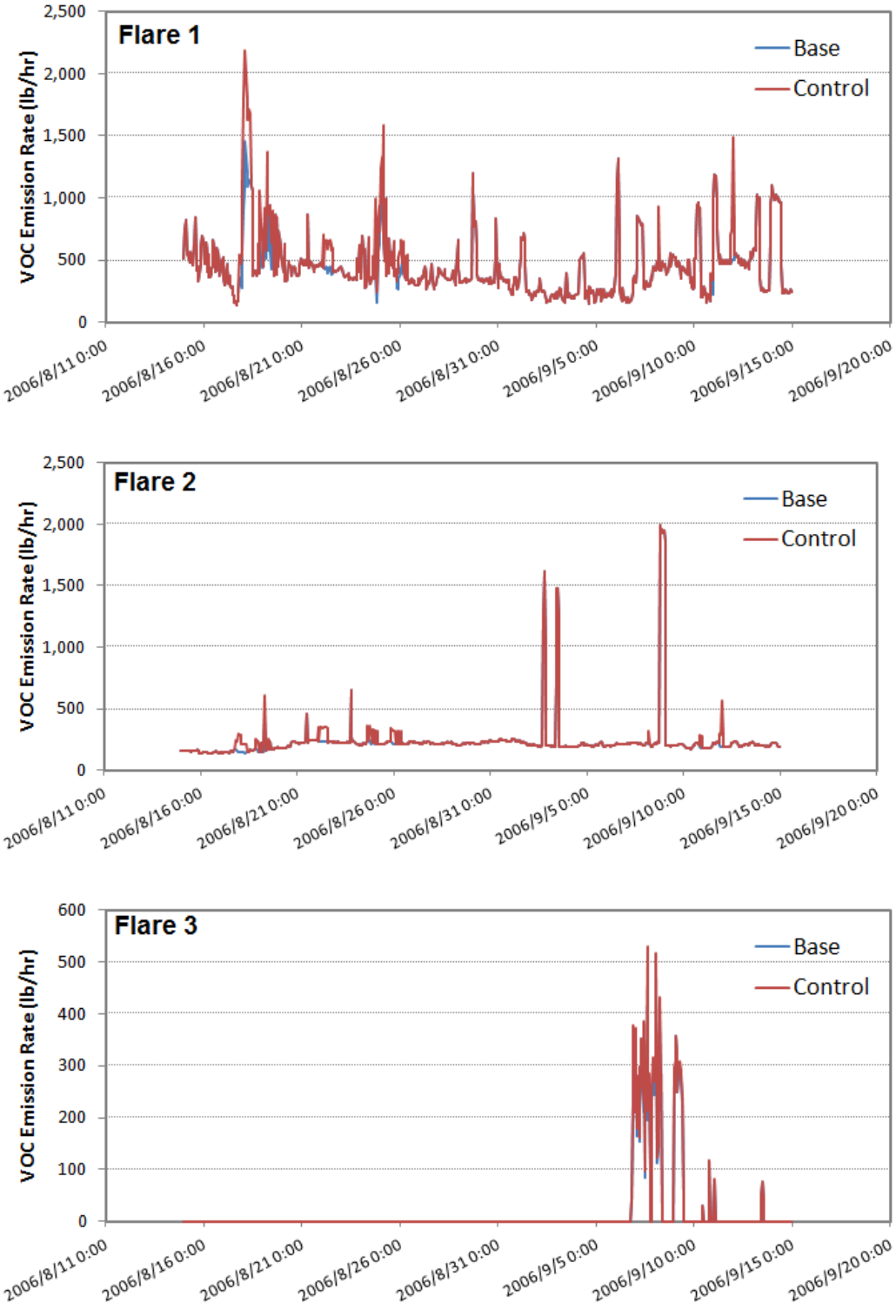


Figure 4.7. Flare VOC emissions before (blue) and after (red) adjustment of the combustion efficiency for three hypothetical flares.

The impact of the variable combustion efficiency is investigated using CMAQ-DDM3D. For each of the three hypothetical flares, its base VOC emissions are used to substitute for the real flare at the Deer Park plant in the special inventory. DDM sensitivities of 8-hour ozone concentrations to the change in flare VOC emissions are computed for each of the three hypothetical flares. The largest differences in 8-hour average ozone concentration for the base and control cases of hypothetical flares 1, 2, and 3 are 0.88 ppb, 0.55 ppb, and 0.01 ppb, respectively (Figure 4.8). The corresponding base 8-hour average ozone concentrations at the locations with the maximum differences are 63.1 ppb, 69.4 ppb, and 60.5 ppb. For the two continuous use flares, the increase in ozone concentration due to the variability of combustion efficiency occurs more frequently compared to that during the intermittent use flare. The intermittent use flare has zero mass flow for 90-95% of the time, so the impact of combustion efficiency is limited to the time with non-zero mass flow. If the operating conditions of this type of flare are controlled to achieve higher combustion efficiency during the non-zero mass flow time, the impact on air quality can be minimized.

The deviated combustion efficiency is able to increase ozone concentration in various areas around the flare. The meteorological condition is the key factor in determining the areas which are impacted by the flare VOC emissions. For instance, the maximum add-on 8-hour average ozone concentrations for flares 1 and 2 both occurred when strong westerly winds brought the flare plume down to the southeast and a significant amount of ozone formed near the coast (Figures 4.8a and 4.8c). The second largest add-on 8-hour average ozone for flare 1 occurred when easterly winds brought the

flare plume to the areas in the west, including the southern part of Houston's downtown (Figure 4.9a). The second largest add-on 8-hour average ozone for flare 2 occurred when the flare plume was brought southwest, which raised the ozone levels in the northern part of Brazoria County (Figure 4.9b). Throughout the episode, the areas that were under frequent impact of the flare VOC emissions were the four counties of Harris, Galveston, Brazoria, and Ford Bend. This is due to the frequent sea breeze along the shore of Galveston Bay. The frequency with which the flare plume hits the eastern counties is lower, but strong westerly winds can lead to stagnation along the shore of Galveston Bay and form a significant amount of extra ozone.

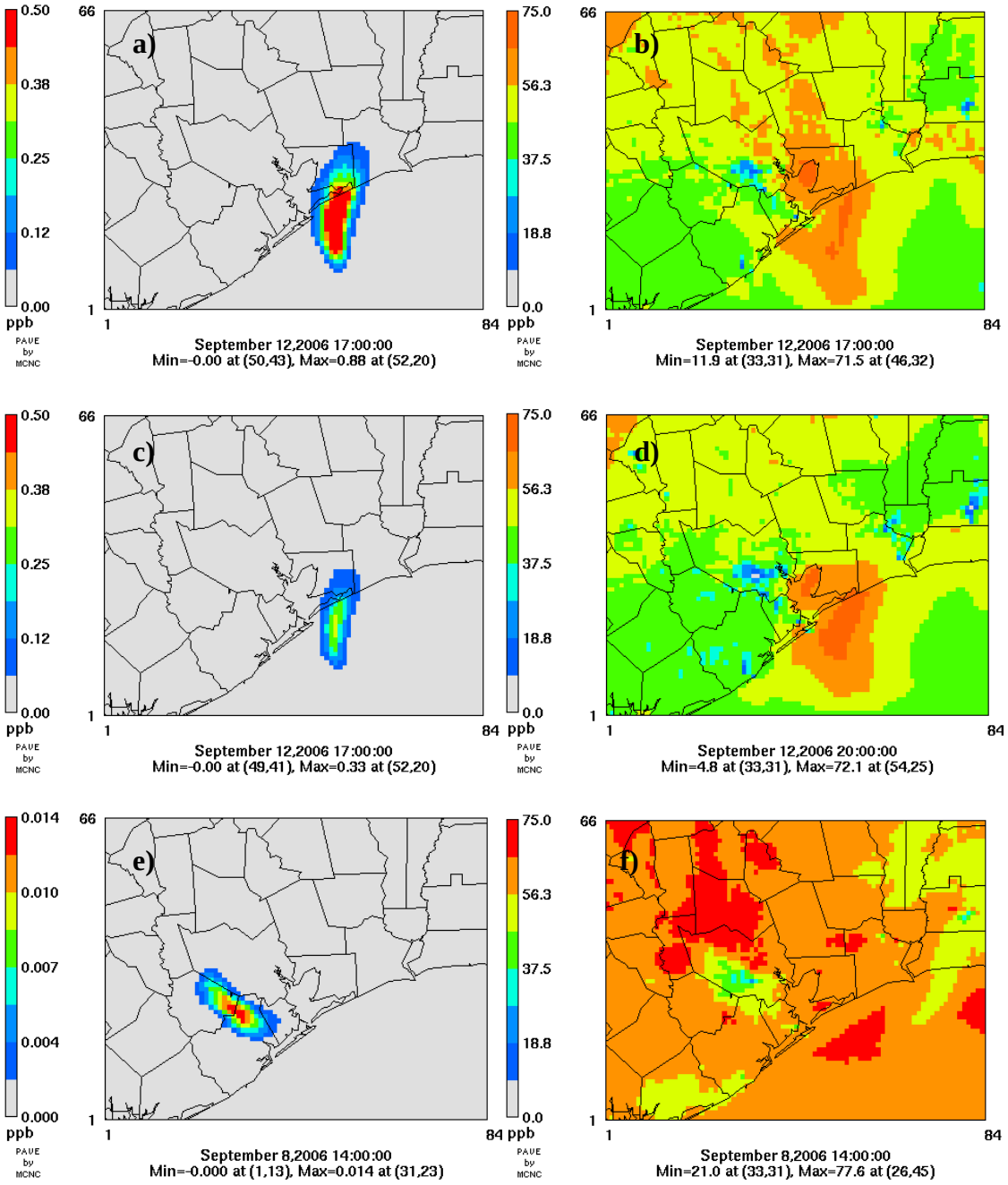


Figure 4.8. Spatial distribution of the differences in 8-hour average ozone concentration at the hours when the maximum difference in 8-hour average ozone occurred for a) hypothetical flare 1, c) flare 2, and e) flare3. b), d), and f) are the spatial distribution of the 8-hour average ozone concentrations at the time when the maximum differences occurred for flares 1, 2, and 3, respectively. Note the change of scale in e).

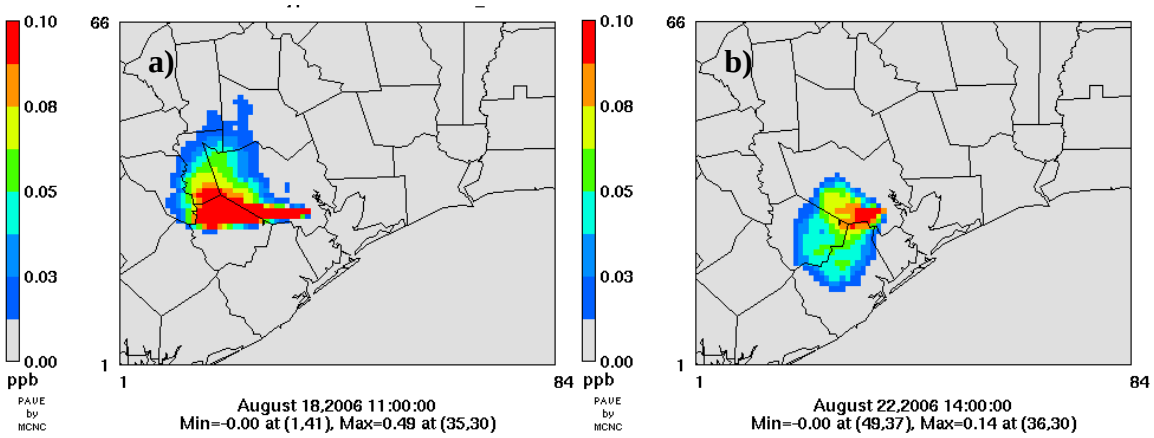


Figure 4.9. Source impact of flare VOC emissions in surrounding counties for a) flare 1 and b) flare 2.

While the maximum increase in ozone concentration reflects the potential impact of flare combustion efficiency on ozone formation, the difference in daily maximum 8-hour average ozone provides more insight for designing emission control strategies. Here we studied the difference in daily maximum 8-hour ozone concentrations at 30 observation sites in the modeling domain. For hypothetical flares 1 and 2, the change in combustion efficiency leads to an increase in 8-hour average ozone concentration at monitoring sites mostly in the three counties: Harris, Jefferson, and Brazoria. For both flares, the maximum differences in flares 1 and 2 are 0.26 ppb and 0.09 ppb, respectively (Figure 4.10). Both of the maximum differences occurred south of downtown Houston and west of the flare. The magnitudes of the differences in daily maximum 8-hour ozone concentrations for flare 1 are larger than those for flare 2 because the episode average hourly VOC emissions for flare 1 are about twice those for flare 2. With the single combustion efficiency curve assumption, the hours with higher base VOC emissions will emit more extra VOC than those with low base VOC emissions. The magnitude of the source impact from the intermittent use flare increases less than 0.01 ppb, which is much

lower than the impact from the two continuous use flares. Furthermore, fewer monitoring sites in the modeling domain are under impact of the intermittent use flare than are the cases under continuous use flares, in that the VOC emissions are limited to a short time period for the intermittent use flare and the chance to cause wide geographic impact is lower. Therefore, among the three types of flare operating modes, the DRE of the two continuous use flares has more significant impact on regional ozone concentrations than that of the intermittent use flare. The DRE of the continuous flare without gas recovery has the largest impact on daily maximum 8-hour average ozone concentrations due to its higher episode average VOC emissions and larger fluctuations in the mass flow of the vent gas.

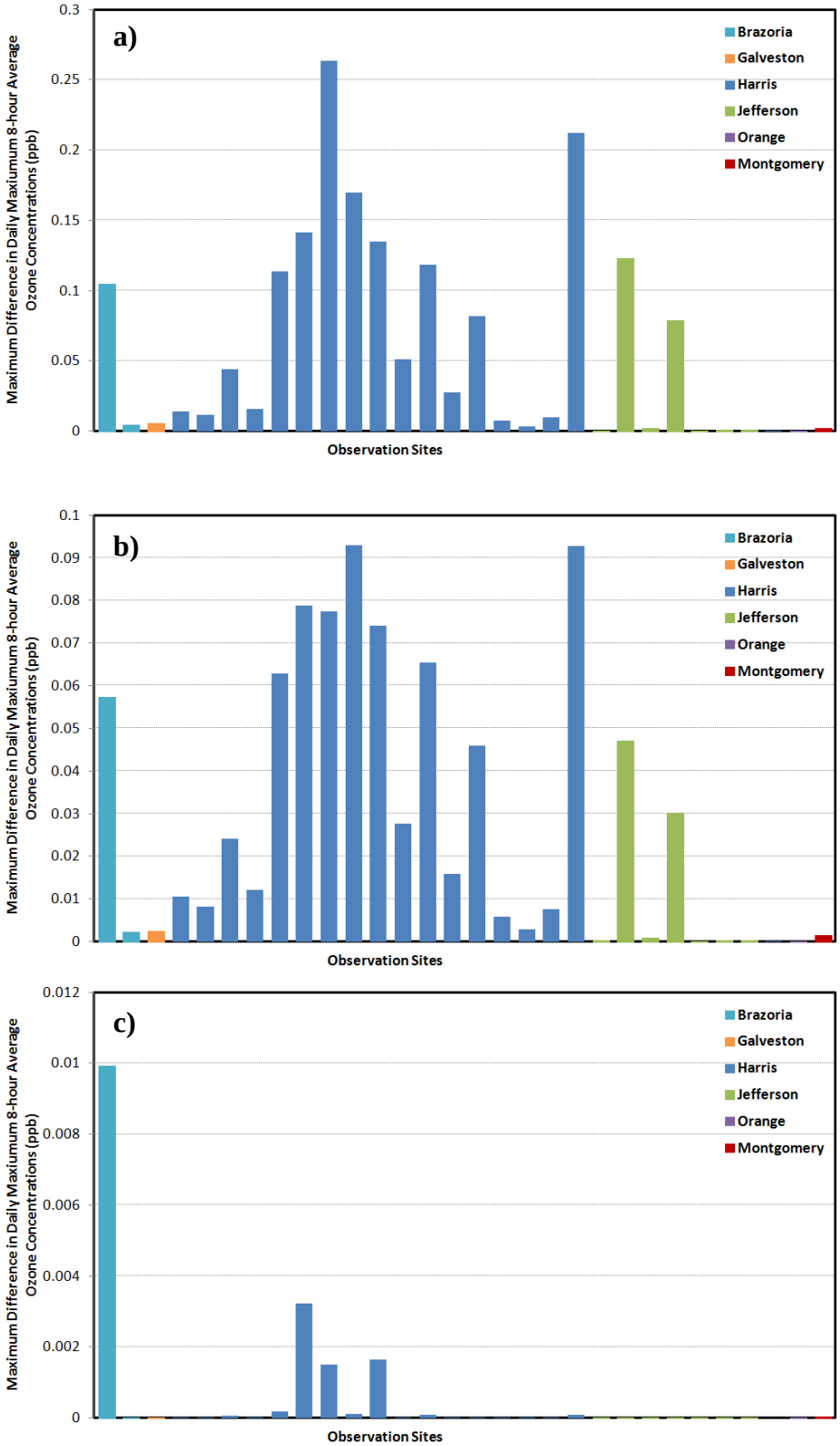


Figure 4.10. Maximum differences in 8-hour average ozone at 30 observations sites. Bar colors represent different counties. The differences are between the base and control

cases for hypothetical flares 1, 2, and 3, which corresponds to a), b), and c), respectively. Note the scale change in y axis.

### 4.3.2 Impact of flare operating conditions on flare VOC emissions

The flare operating conditions are taken into account in the calculation of VOC emissions from the Deer Park flare. As discussed in Section 4.2.2, this flare was selected because it is located in the Houston Ship Channel and is close to the downtown Houston and also has a significant amount of VOC emissions with the highest hourly average VOC emissions among the flares in the Houston Ship Channel. This flare is steam assisted, and the flow rate of the assist steam and the LHV of the vent gas are available from the recent flare study data from TCEQUATION The mass flow rate of the vent gas is obtained from the 2006 Texas special inventory. Figure 4.5 shows the time series of the mass flow rates of the vent gas and the assist steam. The corresponding CZNHV and DRE for each hour is calculated using Equations 4.2 and 4.3, respectively, assuming there is center steam. The relationship of the steam to vent gas ratios, CZNHV, and DRE is shown in Figure 4.11. CZNHV decreases as the steam to vent gas ratios increase, and DRE drops significantly once CZNHV falls below 250 BTU/scf. DRE decreases as increasing steam to vent gas ratios increase.

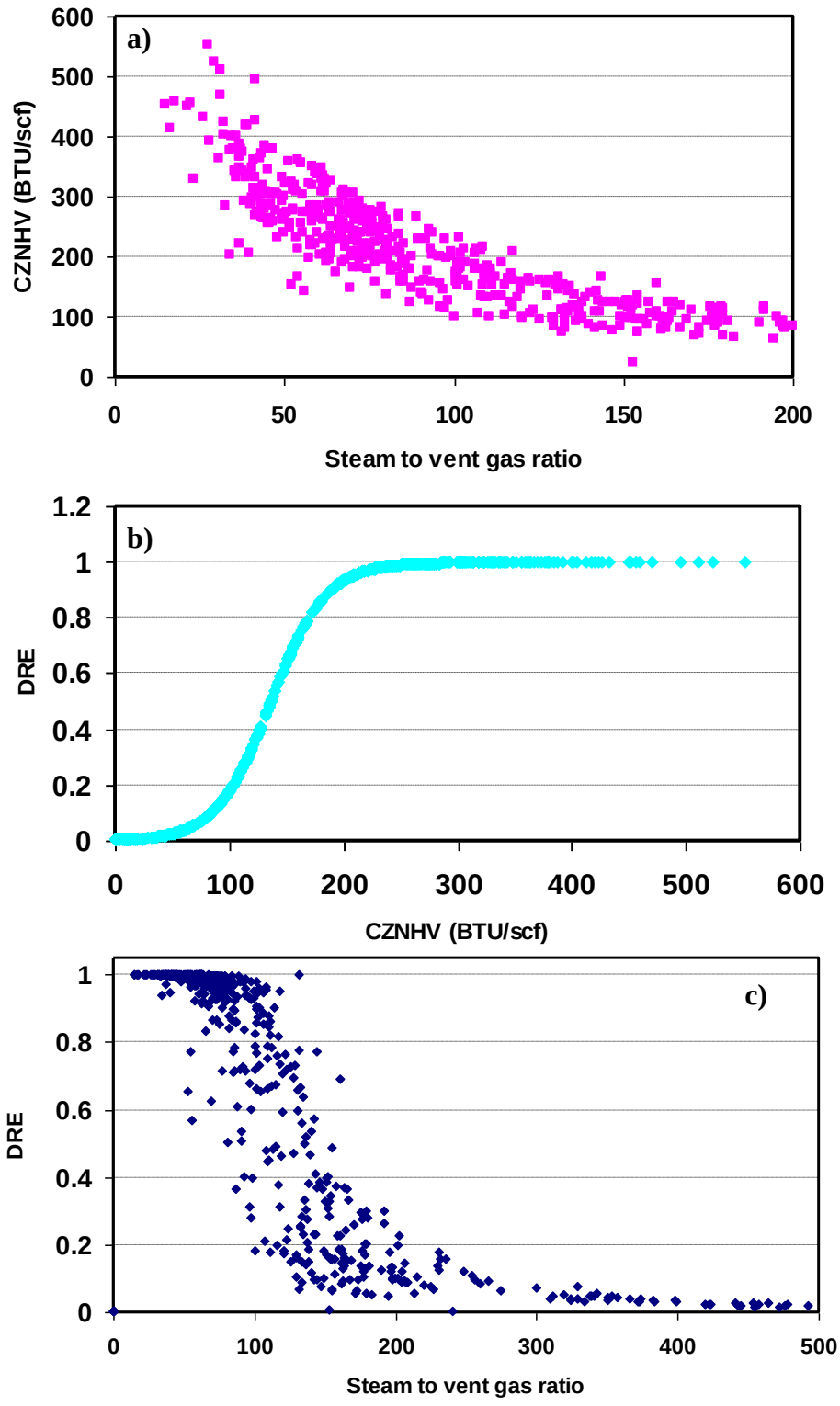


Figure 4.11. The relationship between a) CZNHV and steam to vent gas ratio, b) DRE and CZNHV, c) DRE and steam to vent gas ratio.

The adjusted DRE's are used to calculate flare VOC emissions with Equation 4.8. For a given hour, if the adjusted DRE is below 98%, the VOC emission for the corresponding hour is elevated according to the adjusted DRE. Figure 4.12 shows the time series of VOC emissions from the Deer Park plant before and after adjusting the DRE. A logarithm scale is used here to amplify the details at low mass flow time periods. As expected, the flare VOC emissions at time periods with high steam to gas ratios increase. Some time periods show a significant increase in VOC emissions (an increase of more than 40 times) such as during hours around 200, hours 390-520, and hours 670-768.

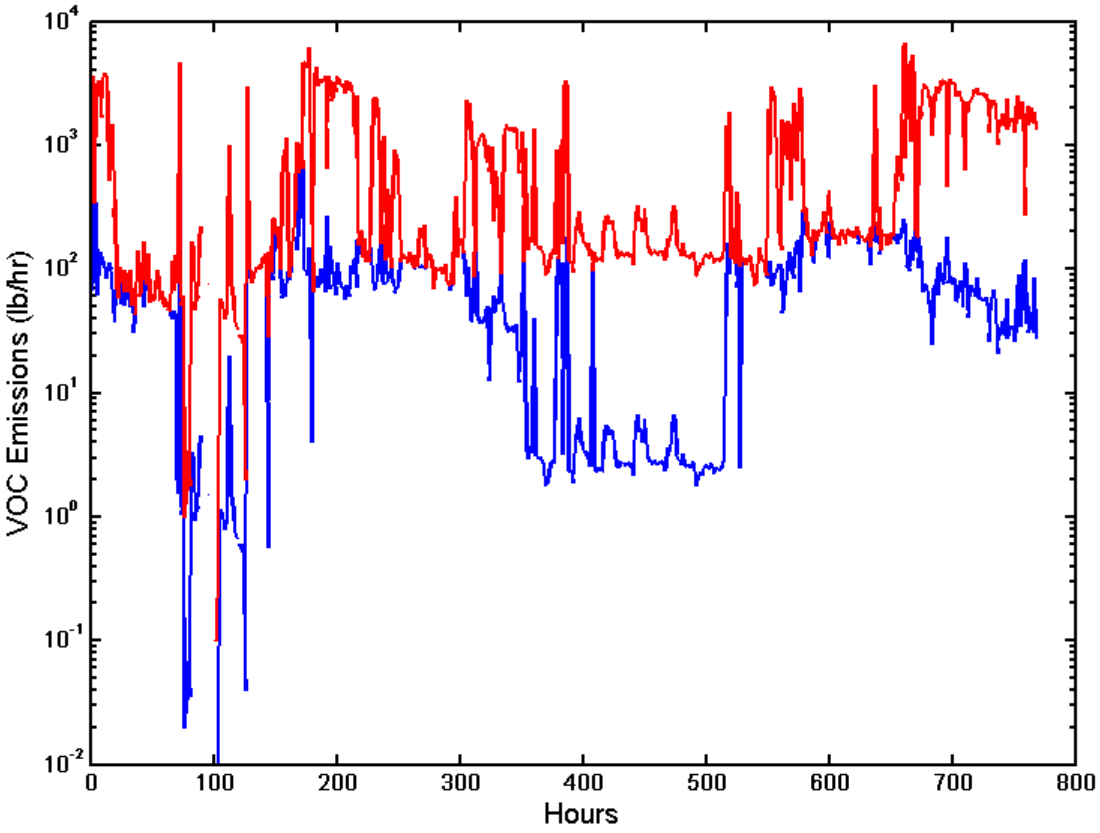


Figure 4.12. Time series of baseline (blue) and controlled (red) VOC emission from Deer Park plant. The y-axis represents the number of hours starting on August 15, 0:00 Central Standard Time (CST).

### 4.3.3 Impact of flare VOC emissions on ozone concentrations

The impact of flare VOC emissions is assessed using the difference in 8-hour average ozone concentrations. The difference in 8-hour average ozone concentrations is estimated using the reduced form model of CMAQ, which is based on the sensitivity of 8-hour average ozone concentration with respect to the flare VOC emissions of every two hours. The fractional change in flare VOC emissions at every two hours due to the DRE change is used as input for the reduced form model (Equation 4.8). The add-on 8-hour average ozone at each grid and each time step is estimated by subtracting the base concentration from the re-estimation. The maximum add-on 8-hour average ozone during the episode is found on August 23, 2006 at 20:00 CST at grid (37, 30) (Figure 4.13). It is located in the same grid as the flare. Although the magnitude of the add-on 8-hour average ozone can reach as high as 10.5 ppb, this high value occurs at the location with a low base concentration of ozone, which is 29ppb. The two grids east of the one with the maximum add-on 8-hour average ozone, grids (38, 30) and (39, 30), also have a significant increase in 8-hour average ozone concentration, with 6.5 ppb and 2.2 ppb, respectively. Besides the local impact, the increased flare VOC emissions lead to a sub-ppb increase in the 8-hour average ozone of the downwind area. The episode peak 8-hour average ozone of the base case is 99.4 ppb, which occurred on September 6, 2006 at 13:00 CST in grid (34, 9). The episode peak 8-hour average ozone in the control case remains the same as that in the base case when the effect of the DRE is taken into account.

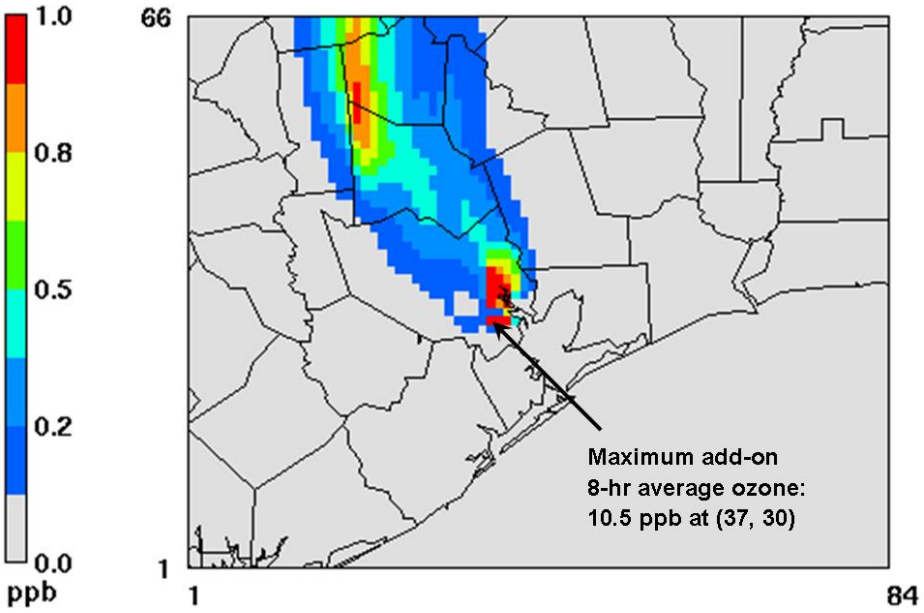


Figure 4.13. Maximum add-on 8-hour average ozone concentration during the episode. It occurred at grid (37, 30) on August 23, 2006 at 20:00 CST.

We investigate further the impact of flare DRE in the downwind area. For 30 monitoring sites in six counties around the Houston Ship Channel, the differences in daily maximum 8-hour average ozone are calculated for each day during the modeling episode. The difference for each site and day is positive. The maximum increase in daily maximum 8-hour average ozone concentration for each site is plotted in Figure 4.14. The adjustment in flare DRE has more than a 1 ppb impact on most of the sites in the three surrounding counties of Harris, Galveston, and Brazoria. The change in DRE can also impact the daily maximum 8-hour average ozone in counties further east, such as Jefferson, Orange, and Montgomery. Although most of the increases in daily maximum 8-hour average ozone in these counties were less than 0.5 ppb, the increase can reach as high as 2.9 ppb when strong westerly winds bring the plume from the flare to the east. The five sites that had an increase in 8-hour average ozone around 2.5 ppb are all located in the southwest of downtown Houston. Therefore, the variation of DRE is estimated to

increase the daily maximum 8-hour average ozone concentrations in both the urban and residential areas and is likely to increase ozone exposure due to the high density of the population in these areas.

Also investigated are the time series of the add-on daily maximum 8-hour average ozone concentration averaged over the 30 monitoring sites (Figure 4.15). Since the 30 monitoring sites are located in the six counties of interest in southeast Texas, each point in Figure 4.15 can represent the spatial averaged increase in daily maximum 8-hour average ozone concentration in southeastern Texas due to the variable DRE of the Deer Park flare. The highest increase is estimated to be 0.6 ppb on the $30^{th}$ day (September 12, 2006). The days with increase in daily maximum 8-hour average ozone are consistent with the hours that exhibit increased flare VOC emissions due to the lowered DRE, so the increased flare VOC emissions had an impact on the daily maximum 8-hour average ozone within the same day. The impact did not last for several days due to the transport and chemical destruction of ozone.

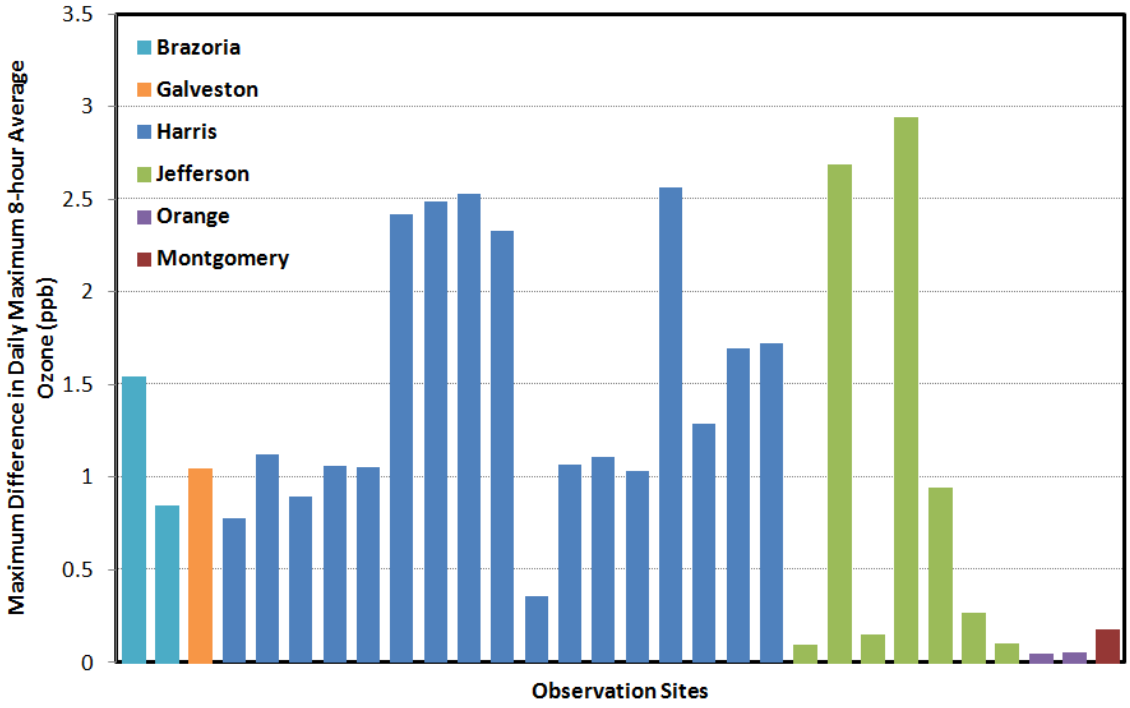


Figure 4.14. Maximum increase in daily maximum 8-hour average ozone concentrations at 30 observation sites in six Texas counties around the Houston Ship Channel.

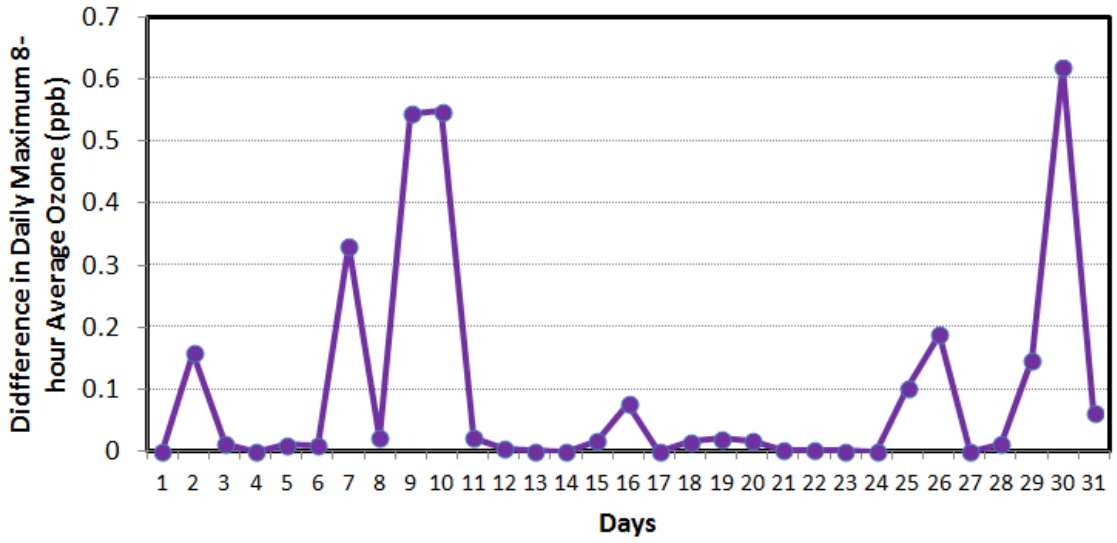


Figure 4.15. Average increase in daily maximum 8-hour average ozone concentration over all the observation sites in Harris, Galveston, Brazoria, and Jefferson counties for each day during the modeling episode.

Since the impact of the flare VOC emissions is a same day effect (i.e., the impact on ozone is due to emissions from the same day), we explore further the hourly response of ozone concentration to the change in flare VOC emissions due to the lowered DRE. Two observation sites are studied here: AQS#482010051 and AQS#482450011. One reason that these two sites are selected is that they both had significant increase in daily maximum 8-hour ozone concentration as shown in Figure 4.14. Another reason is that their locations are in different downwind directions of the flare so the different pattern of flare impact can be investigated. One observation site, AQS#482010051, is located southwest of downtown Houston, which is about 20 miles west of the Deer Park flare. The other site, AQS#482450011, is located in a remote area about 50 miles east of the Deer Park flare. Time series of the sum of sensitivities of every two hours at the two observation sites are compared in Figure 4.10. More frequent impact is found at the observation site near downtown Houston because the meteorological conditions frequently carry the flare plume to this region (Figure 4.16a). The highest impact of this episode is estimated to be 0.16 ppb, which means the 8-hour average ozone concentration at this observation site would drop by 0.16 ppb if the VOC emissions from the flare were completely removed. Although less frequent impact is found at the remote observation site, the maximum impact on 8-hour average ozone at this site reaches 0.11 ppb, which is comparable to that of the other site which is closer to the flare (Figure 4.16b). However, the contributions of the flare VOC emissions at different periods of a day to the ozone concentration at the two observation sites do not have the same pattern. For the observation site closer to the flare (AQS#482010051), the time period with a significant impact from the flare VOC emissions on 8-hour average ozone is from 6am to 10am local

time. The impact lasts for several hours during the day typically from 6am to 12pm. The impact can also be continuous for several days (e.g., impact around the 600$^{th}$ and 700$^{th}$ hours), depending on the meteorological fields and the operating mode of the flare. For the observation site in Jefferson County, the contribution of the flare VOC emissions to 8-hour average ozone concentration is mainly from those emitted during night hours. The contributions from flare VOC emissions between 20:00 and 2:00 were about 0.06 ppb. Thus, while the flare VOC emission during night hours has only a trivial impact on ozone concentration at observation sites near the flare, it is possible that ozone concentrations increase at the remote sites under favorable meteorological conditions. Similar to the findings in the hypothetical flares, a larger impact is found in the counties including or near the flare.

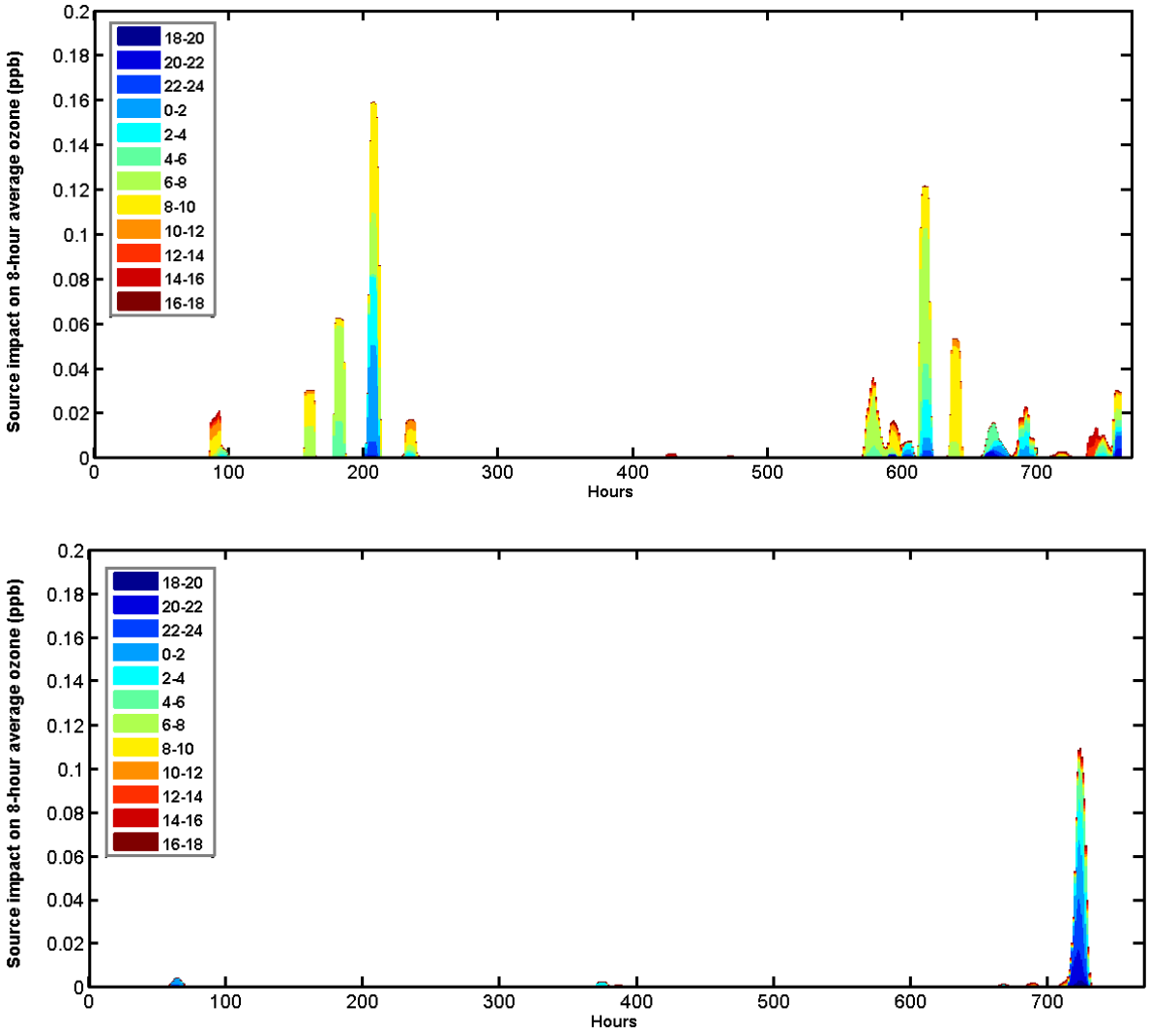


Figure 4.16. Time series of source impact of flare VOC emissions on 8-hour average ozone concentrations at two observation sites: a) 482010051 and b) 482450011.

## 4.4. Conclusions

The impact of combustion efficiency/DRE on ambient ozone concentration has been investigated using CMAQ-DDM3D. Among the three flare operating modes, i.e., the continuous use without gas recovery, the continuous use with partial gas recovery, and the intermittent use with full gas recovery, the two continuous use flares have a more significant impact on regional ozone concentrations than does the intermittent use flare in

terms of the magnitude of source impact and the area of the regions impacted. The maximum add-on concentration to the 8-hour average ozone was 0.88 ppb in the first flare operating mode. It occurred on the shore of the Galveston Bay when strong westerly winds carried the flare plume to the southeast. The difference in daily maximum 8-hour average ozone due to variable combustion efficiency has been studied at 30 observation sites. The sites in Harris, Brazoria, and Jefferson counties experienced a larger impact of the variable DRE than more remote sites. The largest impact on daily maximum 8-hour ozone at any monitoring site was 0.26 ppb, which was found in the first operating mode at the site southwest of downtown Houston and west of the flare.

Sensitivity analysis has also been applied to the real flare at the Deer Park plant. The hourly VOC emissions of the flare are obtained from the 2006 Texas special inventory and the flow rates of the assist steam are obtained from a more recent inventory. The sensitivities of 8-hour average ozone concentration to flare VOC emissions at every two hours during a day are calculated using CMAQ-DDM3D and a reduced form model of CMAQ is established using those sensitivities. The changes in VOC emissions at two hours are first estimated by using the flow rate of assist steam and then used to drive the reduced form model to reconstruct the 8-hour average ozone concentrations at each grid in the modeling domain. The differences in 8-hour average ozone concentrations at the 30 observation sites are investigated. The largest increase at any monitoring site is 2.9 ppb, which is found at a monitoring site in Jefferson County to the east of the flare. The time series of average increase in 8-hour average ozone concentration at the 30 monitoring sites indicate consistence between the increase in ozone concentration and the decrease in DRE. The impact from flare VOC emissions typically lasts 6 to 12 hours. Further

investigations in the hourly contributions of flare VOC emissions indicate that flare VOC emissions during the morning hours contribute significantly to 8-hour average ozone concentration at locations close to the flare and that night hour flare VOC emissions contribute more to the 8-hour average ozone concentration at remote locations than those emitted during the daytime.

# CHAPTER 5

# INTEGRATING AIR QUALITY-RELATED HEALTH EFFECTS IN ELECTRICITY CAPACITY PLANNING MODEL

## 5.1. Introduction

This chapter, on human health impacts of a transition to natural gas, is part of a broader strategic study of natural gas, and is also contributing to the development of advanced energy modeling capability. Production of electricity from fossil fuels released pollutants that can form $PM_{2.5}$ (fine particulate matter <= 2.5 μm in diameter) and $O_3$ (ozone); their negative health effects include premature mortality and increases in emergency room visits and hospitalization rates (Krewski et al. 2009). Due to temporal and spatial dynamics of pollutant transport, changes in emissions likely have varying effects on pollutant concentrations in different locations. Moreover, a kWh of electricity generation produces different emissions at different electricity generating units (EGUs) depending on fuel type and pollutant controls installed at the EGU. In addition, mortality rates and population density can vary greatly county-to-county (CDC NCHS, 2012). From a given EGU, certain downwind areas may have high pollutant concentrations, while others may be virtually unaffected. Given this, a variety of important questions emerge. How can we model the health costs of electricity production more effectively, particularly given geospatial varying effects and affected populations? How can we identify which EGUs emissions are most worth reducing? Where will emissions reductions have the greatest impacts?

To address these questions, we first estimate both pollutant concentrations and health impacts geospatially and then measure the value of pollutant concentration changes due to emissions reductions of fossil fuel EGUs in the state of Georgia. In order to do this, we utilize a three-dimensional chemical transport model to simulate the spatial distribution of pollutant concentrations as well as their sensitivities to emission rates. The embedded direct sensitivity technique in the chemical transport model is able to assess the emissions impact of one or a selected set of power plants in the mean time of simulating the pollutant concentration. The sensitivities are then used to construct a reduced form model to efficiently estimate the response of pollutant concentrations to changes in the emissions of power plants. The reduced form model has been applied to control strategy optimization and evaluation for compliance with the National Ambient Air Quality Standards (NAAQS) for ozone (Cohan et al., 2006; Simon et al., 2013). Here we extend the reduced form model to account for the response of $PM_{2.5}$ to change in precursor emissions.

Next, we use emissions reductions and associated health impact estimates in an electricity capacity planning model, which examines decisions from the present through the year 2050. Using this model we can measure the air-quality related health impacts on a plant-by-plant basis, and demonstrate its ability to resolve health impacts and implied costs with unprecedented spatiotemporal resolution.

Previous studies have quantified health impacts using a $/ton metric across a wide region for specific pollutants, such as $SO_2$ (sulfur dioxide). However, damage can vary by source location owing to variability in atmospheric conditions and population density downwind of emissions. For this, the $/ton metric may either undervalue or overvalue

emissions depending on location. To address these challenges, we instead compare health costs with a $/kWh metric in order to compare the impact of a single unit of power production.

In this study we measure the costs from the health impact of emissions for two months, January and July of 2007. These results can help not only in making more informed EGU dispatching decisions, but also in planning future capacity decisions. Increases and decreases in EGU usage dependent on EGU location and emissions characteristics of that EGU could help alleviate health effects, as well as assist in achieving current and future NAAQS (National Ambient Air Quality Standards) compliance (EPA, 2013a). In addition, future policy decisions may be influenced by modeled spatial health impacts due to emission locations, downwind population and population demographics. This research could inform the pricing of emissions, as well as more informed placement of future EGUs. The work involves a collaboration, and work conducted as part of this thesis project involve air quality modeling. Air quality modeling results are then used in an electricity generation planning model being applied by another Ph.D. student. The formulation of the electricity generation planning model is presented here for completeness and to provide context of how the air quality modeling results are being used.

## 5.2. Methods

### 5.2.1 Electricity generation planning model

This study develops a deterministic mixed-integer linear programming (MILP) model that incorporates the features of individual power plants and downwind emission

impact. The indices, variables, and parameters used in the model are summarized in Table 5.1.

The objective function can be written as

$$\min \sum_{t \in T} R_t \sum_{i \in I} (c_i y_{i,t} + f_i x_{i,t} + v_{i,t} \sum_{s \in S} \sum_{h \in H} z_{i,s,h,t}) + \text{health effects} \quad (5.1).$$

Several constraints are used in the electricity generation planning model. The first is the change in capacity constraint from year to year (Equation 5.2). The next is the demand capacity constraint (the demand that must be met) which includes a reserve margin $\alpha$ (Equation 5.3). To ensure enough capacity is available during peak capacity hours, we have a constraint ensuring that capacity is above peak demand for each season (Equation 5.4). The capacity for each of the various power generation types is ensured through the last three constraints for solar (Equation 5.5), wind (Equation 5.6), and remaining power generation technologies (Equation 5.7). Finally, we have a renewable electricity standard constraint ensuring that the generated electricity within the regulated specific percentage of renewable electricity to be used (Equation 5.8). To estimate the cost of EGU emissions, we simulated representative months of January and July of 2007 (winter and summer respectively). When combined with other population and mortality estimates, the data are then used to estimate the health impacts of other years with varying electricity production and emissions.

$$x_{i,t} = x_{i,t-1} + y_{i,t} - q_{i,t}, \forall i \in I, t \in T \quad (5.2)$$

$$\sum_{i \in I} z_{i,s,h,t} \geq (1+\alpha) d_{s,h,t}, \forall s \in S, h \in H, t \in T \quad (5.3)$$

$$\sum_{k \in K} x_{k,t} \geq \delta_s \hat{d}_{s,h,t}, \forall t \in T \quad (5.4)$$

$$z_{i,s,h,t} \leq \phi_i x_{s,t}, \forall i \in I / (L \cup M), h \in H, s \in S, t \in T \quad (5.5)$$

$$z_{l,s,h,t} \leq \phi_{solar,s,h} x_{s,t}, \forall l \in L, h \in H, s \in S, t \in T \quad (5.6)$$

$$z_{m,s,h,t} \leq \phi_{wind,s,h} x_{s,t}, \forall m \in M, h \in H, s \in S, t \in T \quad (5.7)$$

$$\sum_{j \in J} \sum_{s \in S} \theta_s \sum_{h \in H} z_{j,s,h,t} \geq RES_t \sum_{i \in I} \sum_{s \in S} \theta_s \sum_{h \in H} z_{i,s,h,t,} \forall t \in T \quad (5.8)$$

Table 5.1. Indices, parameters, and variables in the model

| | |
|---|---|
| *Sets and indices* | |
| $i \in I = \{1,...,I\}$ | All power generating plants |
| $j \in J \subset I$ | Renewable energy generating plants |
| $k \in K \subset I$ | Hydro, solar and wind generating plants |
| $l \in L \subset I$ | Solar generating plants |
| $m \in M \subset I$ | Wind generating plants |
| $s \in S = \{1,...,3\}$ | Seasons, s = 1 (summer), 2 (winter), 3 (intermediate) |
| $t \in T = \{0,...,T\}$ | Years (T year time horizon) |
| $h \in H = \{1,...,24\}$ | Hours (24 hours per day) |
| *Parameters* | |
| $r$ | Risk adjusted discount rate |
| $R_t$ | Discount factor for year t |
| $d_{s,h,t}$ | Electricity demand (MW) in season s, hour h, year t |
| $\hat{d}_{s,t}$ | Maximum hourly demand (MW) in season s, year t |
| $\theta_s$ | Number of days in season s |
| $\phi_i$ | Maximum capacity factor of plant i (%) |
| $\phi_{solar,s,h}$ | Maximum capacity factor of solar in season s, hour h (%) |
| $\phi_{wind,s,h}$ | Maximum capacity factor of wind in season s, hour h (%) |
| $u_i$ | Upper bound on plant i generation (MW) in one hour |
| $\delta_s$ | Peak demand multiplicative factor in season s |
| $\alpha$ | Reserve margin (MWh) |
| $RES_t$ | Renewable electricity standard (%) in year t |
| $x_{i,0}$ | Starting capacity (MW) of plant i in year 0 |
| *Costs* | |
| $c_i$ | Capital investment costs (upgrades in capacity) for plant i ($/MW) |
| $f_i$ | Annual fixed costs (per year capacity costs) of plant i ($/MW) |
| $v_{i,t}$ | Variable fuel costs of plant i in year t ($/MWh) |
| *Decision variables* | |
| $x_{i,t}$ | Capacity (MW) of plant i in year t |
| $y_{i,t}$ | Increased capacity (MW) of plant i in year t |
| $q_{i,t}$ | Decreased capacity (MW) of plant i in year t |
| $z_{i,s,h,t}$ | Electricity generated (MWh) from plant i, in season s, hour h year t |

5.2.2 Air quality model

Geospatial pollutant concentration estimates are simulated by the Community Multiscale Air Quality (CMAQ) model, which is one of the most widely used chemical transport models in current air quality management (Byun et al., 2006). The modeling domain covers the southeast U.S., and a 12 km by 12 km grid resolution in this simulation. The meteorological fields are simulated by the Weather Research and Forecasting (WRF) model with four-dimensional data assimilation techniques. The gridded emissions rates are prepared by the Sparse Matrix Operator Kernel for Emissions (SMOKE) model using the 2008 National Emissions Inventory (NEI) and 2007 continuous emissions monitoring (CEM) system for nitrogen oxides ($NO_x$) and sulfur dioxide ($SO_2$) emissions from EGUs. The model performance is evaluated using the air quality system (AQS) observational data. The performance metrics for 8-hour average ozone and 24-hour average $PM_{2.5}$ concentrations for the modeling domain, which are summarized in Table 5.2, are within the acceptable range according to the guidance by EPA (2007).

Table 5.2: Performance metrics for 8-hour average ozone concentrations and 24-hour average $PM_{2.5}$ concentrations

| **Pollutants** | **Months** | **Mean Bias** | **Mean Error** | **Normalized Mean Bias (%)** | **Normalized Mean Error (%)** |
|---|---|---|---|---|---|
| Ozone | January | -1.04 ppb | 6.88 ppb | -5.39 | 35.7 |
| | July | 6.49 ppb | 12.58 ppb | 20.29 | 39.36 |
| $PM_{2.5}$ | January | 3.30 μg $m^{-3}$ | 4.94 μg $m^{-3}$ | 32.63 | 48.88 |
| | July | -1.67 μg $m^{-3}$ | 5.09 μg $m^{-3}$ | -11.61 | 35.37 |

A reduced form model of CMAQ is developed using the sensitivities calculated by a direct sensitivity technique, CMAQ-DDM3D (Yang et al. 1997; Hakami et al., 2003; Napenelok et al., 2006; Zhang et al., 2012) integrated in to CMAQ. Sensitivities quantify the pollutant-emission response.

$$S_{i,j} = \frac{\partial C_i}{\partial \varepsilon_j}, \tag{5.9}$$

where $S_{i,j}$ is the sensitivity of pollutant *i* to emission rate *j*, $C_i$ is the concentration of pollutant *I*, and $\varepsilon_j$ represents the fractional change in emission rate *j*. Both $S_{i,j}$ and $C_i$ vary with time and location (but are not noted in the notion for simplicity). CMAQ-DDM3D simultaneously calculates the sensitivities to all the emissions rates of interest along with the simulation of the pollutant concentrations. The reduced form model can be formulated as

$$C_i^* = C_i^0 + \sum_{j=1}^{N} S_{i,j} \Delta\varepsilon_j \,, \tag{5.10}$$

where $C_i^0$ is the baseline concentration of pollutant *I*, $C_i^*$ is the concentration of pollutant *i* with perturbation in emission rates of interested, and $\Delta\varepsilon_j$ is the fractional change in emission rate *j*. The number of sensitivity parameters, *N*, depends on how many emission sources are of interest. The estimation of the resulting reduced form model has been evaluated using the original CMAQ model and the reduced form model has been shown to well capture the pollutant-emission response (Hakami et al., 2003, Zhang et al., 2012). In this study, in the attempt to capture the geospatial difference in source impact of power plants at various locations, we selected seven emission sources including individual and aggregated EGU point sources to represent the power generation system in Georgia.

Their $SO_2$ and $NO_x$ emissions are treated as the sensitivity coefficients in the reduced form model (Equation 5.10). Table 5.3 summarizes the emissions of the seven EGU point sources on an annual basis for year 2007, and Figure 5.1 shows their locations. Since 2007, emissions at plants have changed. Notably, Plant McDonough has modernized the plant and switched from coal to natural gas. The July $SO_2$ emissions of Plant McDonough dropped from 2432.18 tons in 2007 to 2.25 tons in 2012, and the July $NO_x$ emissions of Plant McDonough dropped from 372.65 tons in 2007 to 12.79 tons in 2012.


Figure 5.1. Locations of power plants in Georgia, including the seven EGU point sources selected for the reduced form model. Different symbols represent fuel types.

Table 5.3. Summary of annual net generation of power and emissions of $SO_2$ and $NO_x$ for the seven point sources in Georgia in 2007.

| Point Sources | Annual Net Generation (MWh) | Annual $SO_2$ Emissions (tons) | Annual $NO_x$ Emissions (tons) |
|---|---|---|---|
| Bowen | 22,972,903.0 | 196,841.24 | 18,478.95 |
| Scherer | 25,053,667.0 | 76,456.25 | 18,840.75 |
| McDonough | 3,772,754.0 | 28,538.81 | 4,604.71 |
| Harllee | 10,373,819.0 | 98,362.71 | 20,928.80 |
| Wansley[a] | 21,970,857.0 | 171,067.10 | 6,614.40 |
| GA North[b] | 31,708,430.2 | 49,713.38 | 9,454.73 |
| GA South[b] | 29,097,886.0 | 35,570.08 | 13,223.90 |

[a] Combo source including Wansley, Yates, and Chattahoochee Energy Facility
[b] Dose not include the five sources listed above

Table 5.4. Summary of net generation of power and emissions of $SO_2$ and $NO_x$ for five EUG point sources in Georgia in July 2007 and July 2012.

| Point Sources | July 2007 | | | July 2012 | | |
|---|---|---|---|---|---|---|
| | Net Generation (MWh) | $SO_2$ Emissions (tons) | $NO_x$ Emissions (tons) | Net Generation (MWh) | $SO_2$ Emissions (tons) | $NO_x$ Emissions (tons) |
| Bowen | 2,262,944.0 | 17,277.58 | 592.44 | 1,259,879.0 | 246.3 | 371.3 |
| Scherer | 2,311,025.0 | 6,682.90 | 1,536.77 | 2,263,990.0 | 4,002.24 | 1,218.49 |
| McDonough | 348,640.3 | 2,432.18 | 372.65 | 1,090,280.0 | 2.25 | 12.79 |
| Harllee | 1,043,051.0 | 9,389.52 | 1,960.72 | 636,177.8 | 6,029.22 | 1,287.50 |
| Wansley[a] | 981,255.5 | 7,706.82 | 1,018.43 | 594,732.0 | 2,609.62 | 322.48 |

[a] Combo source including Wansley, Yates, and Chattahoochee Energy Facility

5.2.2 Air-quality related health effects

We use the EPA BenMAP CE (EPA, 2013b) tool to combine health impacts and cost estimates for both January and July of 2007, given a change in 1kWh of generation at the observed plants in the state of Georgia. The health-related cost is estimated by multiplying the value of statistical life (VSL) and the health response calculated using the

linear approximation of the concentration-response (CR) function. The linear approximation of the $PM_{2.5}$ related CR function can be written as

$$\Delta u_d \approx u_d \beta_{PM_{2.5}} \Delta C_{d,PM_{2.5}} P, \tag{5.11}$$

where $\Delta u_d$ is the change in mortality rate on day *d*; $u_d$ is the baseline mortality rate, $\beta_{PM_{2.5}}$ is the log-relative mortality rate associated with a unit increase in $PM_{2.5}$ and takes the unit of $(\mu g\ m^{-3})^{-1}$, $\Delta C_{d,PM_{2.5}}$ is the change in $PM_{2.5}$ concentrations, and $P$ is population. The cost of an increase in one microgram per cubic meter change in $PM_{2.5}$ at grid $\vec{x}$ can be defined as

$$c_{PM_{2.5}}(\vec{x}) = VSL \times \Delta u_d, \tag{5.12}$$

We use $7.4 million for the VSL according to EPA (2006).Observed mortality rates are obtained from the CDC-WONDER database (CDC NCHS, 2012), which has a strong geospatial dependence (All cause mortality in Georgia ranges from 300 to 1500 per 100,000). The population is estimated from demographic data and block level resolution data from the 2010 Census (U.S. Census Bureau, 2010).

The air quality related health cost is estimated by combining the reduced form model (Equation 5.10) and the cost associated with unit increase in $PM_{2.5}$ concentrations (Equation 5.12):

$$C = \sum_{\vec{x}\in X}\sum_{t\in T}\sum_{i\in I}\sum_{j\in J}\sum_{k\in K}\left(\frac{z_{t,k}e_{j,k} - E^0_{j,k}(t)}{E^0_{j,k}(t)}\right) S_{i,j,k}(\vec{x},t)c_i(\vec{x}), \tag{5.13}$$

where $C$ is the air quality-related health cost over a certain region and time period and $z_{t,k}$ is the megawatt hours (MWh) generated in hour *t* at plant *k*.

## 5.3. Results and discussion

### 5.3.1 Emission impact of EGUs

Seven EGU point sources are selected to represent the power generation system in Georgia. Their emissions impact on regional air quality is estimated using CMAQ-DDM3D. Sensitivities of $PM_{2.5}$ to $SO_2$ emissions of these power plants are highly dependent on their locations (Figure 5.3). North Georgia is more strongly impacted by plant emissions than is in South Georgia since most of the power plants are concentrated in the northern part of Georgia. Plant Bowen has an impact of 2 $\mu g\ m^{-3}$ on $PM_{2.5}$ concentration due to its $SO_2$ emissions, which is the highest among the seven EGU sources. The second largest impact is 0.9 $\mu g\ m^{-3}$, from Plant Harllee. The highest impact found in the aggregation of all the other power plants in North Georgia is 0.9 $\mu g\ m^{-3}$, which is even lower than the individual contribution from Plant Bowen. $NO_x$ emissions from power plants also contribute significantly to PM formation. Sensitivities of $PM_{2.5}$ to $NO_x$ emissions from six of the selected EGU sources are shown in Figure 5.3. Plant McDonough is not shown here since the magnitude of its sensitivity is too small to distinguish it from numerical noise. The contributions of $NO_x$ to $PM_{2.5}$ are smaller than the contributions of $SO_2$ to $PM_{2.5}$. This is due to the relatively smaller magnitude of $NO_x$ emissions. The geospatial distribution of the emissions impact of each EGU source is similar to that of the $SO_2$ emissions. Similarly, the emissions impact is concentrated in North Georgia in downwind areas of the power plants.

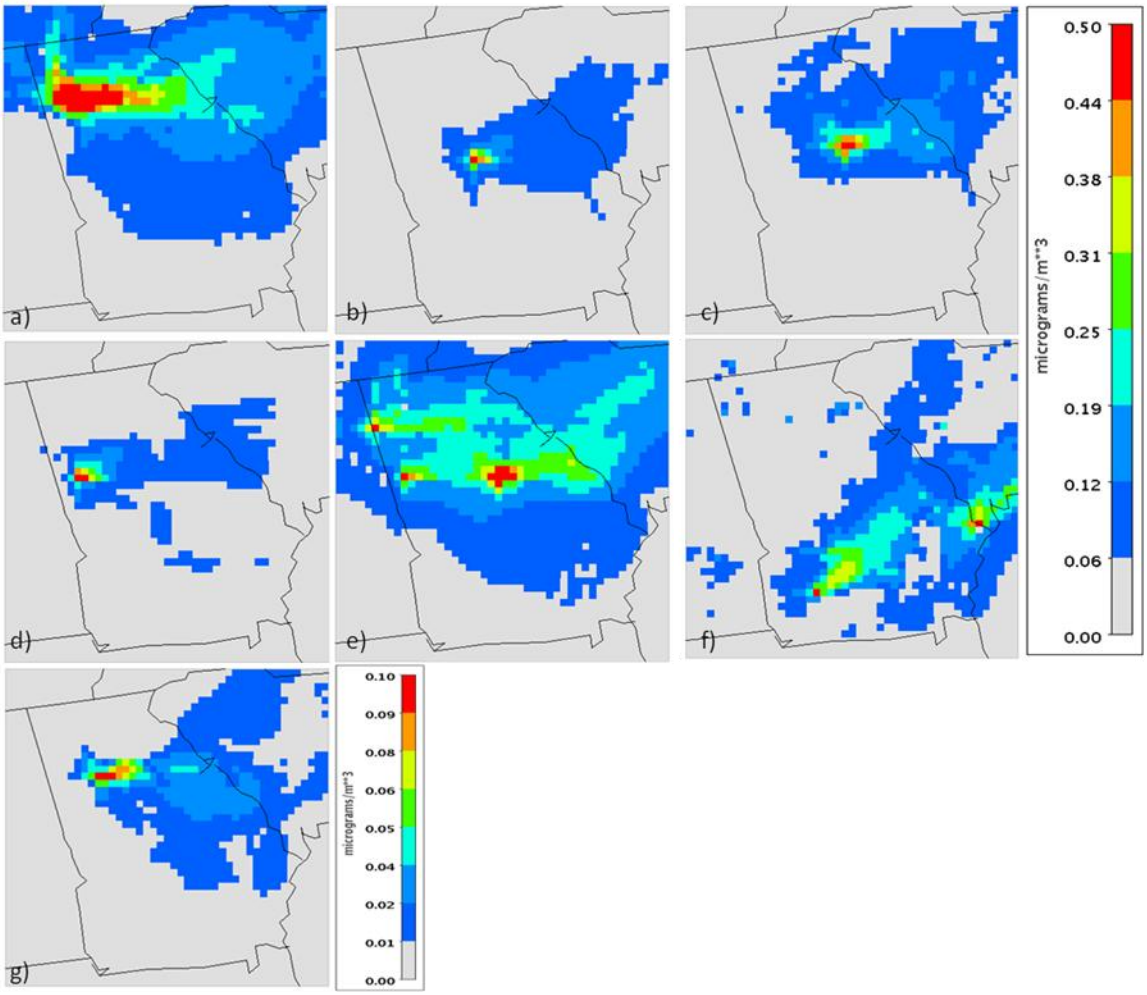


Figure 5.2. Sensitivities of $PM_{2.5}$ to $SO_2$ emissions from Plants a) Bowen, b) Scherer, c) Harllee, d) Wansley, Yates, and Chattahoochee, g) McDonough, and the other power plants in e) North and f) South Georgia. Note the change in scale for Figure 5.2g. The sensitivities are averaged over July.

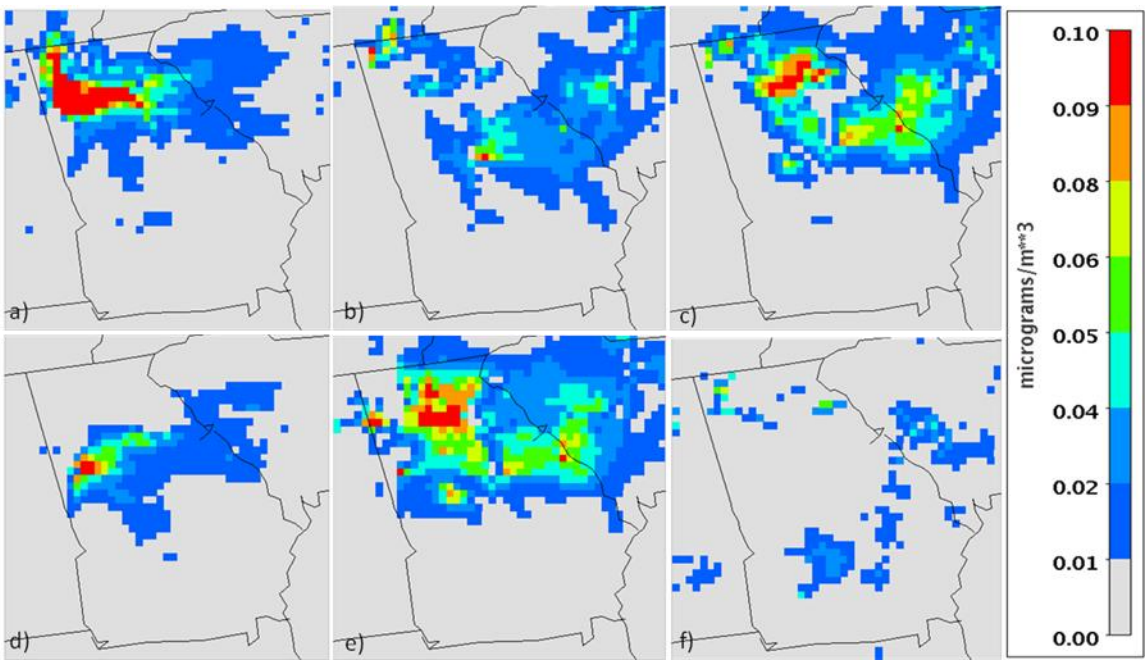


Figure 5.3. Sensitivities of $PM_{2.5}$ to $NO_x$ emissions from Plants a) Bowen, b) Scherer, c) Harllee, d) Wansley, Yates, and Chattahoochee, and the other power plants in e) North and f) South Georgia. The sensitivities are averaged over July.

### 5.3.2 Estimation of health cost

We first explored the health cost per MWh in July of 2007 due to secondary sulfate-based $PM_{2.5}$ generated from EGU $SO_2$ emissions. The sum of $PM_{2.5}$-associated health costs over Georgia due to $SO_2$ emissions from Plant Bowen, McDonough, and Scherer are $89.9/MWh, $123.3/MWh, and $3.0/MWh, respectively. Although the $SO_2$ emissions of Plant McDonough are lower than those of Plants Bowen and Scherer, Plant McDonough was estimated to have the largest health cost because it is located in Atlanta which has a high population density. The costs in Table 5.5 are comparable to those of previous expert estimates (Committee on Health, Environmental & National Research Council, 2010; Siler-Evans et al., 2013), but also capture useful geospatial variability that can be utilized to make informed decisions about plant locations and operations. The

geospatial dependence of the health cost for Plants Bowen, McDonough, and Scherer is displayed in the maps in Figure 5.4.

Table 5.5: Total $PM_{2.5}$ associated health cost due to $SO_2$ emissions for Georgia in July 2007

| **Plant Name** | **July 2007 $SO_2$ based $PM_{2.5}$ health cost ($/MWh)** |
|---|---|
| Scherer | 3.0 |
| Bowen | 89.9 |
| McDonough | 123.3 |

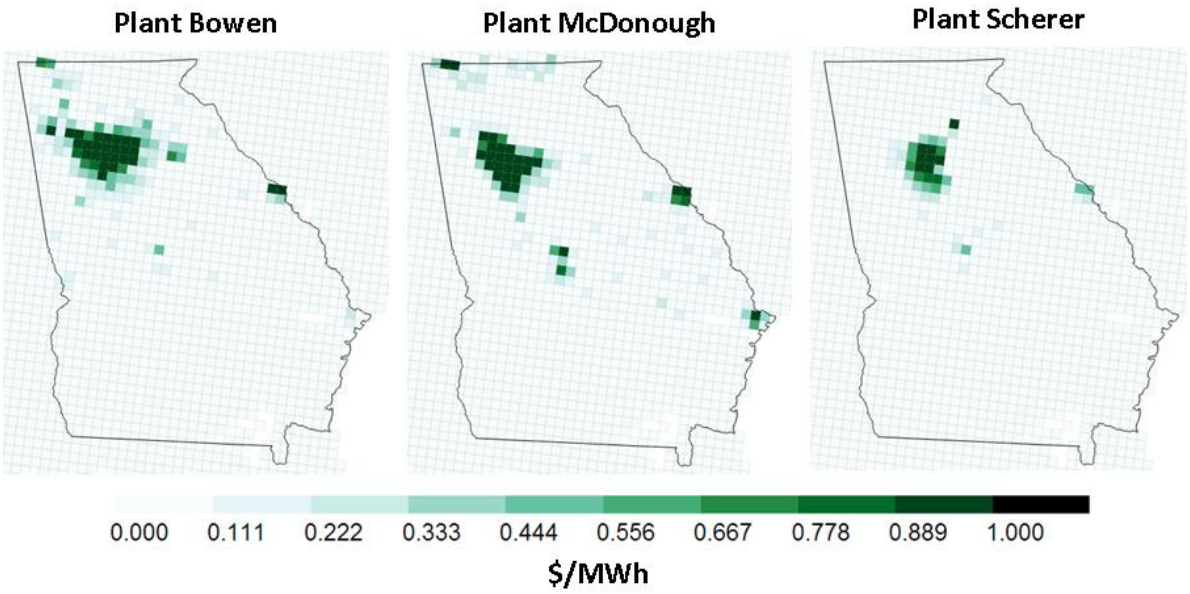


Figure 5.4. Monthly averaged geospatial costs in dollars per MWh for three power plants in July 2007.

### 5.4. Summary and future work

The reduced form air quality model has been implemented in the electricity generation model to represent geospatial dependent emissions impacts and air quality-related health costs of power plants at different locations. The reduced form model is established using sensitivities obtained from CMAQ-DDM3D. $SO_2$ and $NO_x$ emissions from seven EUG point sources in Georgia are used as sensitivity parameters in the

reduced form model. The sensitivities well capture the geospatial difference in emissions impacts and show distinct spatial patterns for power plants at various locations. The emissions impact of $SO_2$ and $NO_x$ emissions are mainly to the east of the corresponding point sources. The EGU emissions have a larger impact on North Georgia due to the abundance of EGUs in the northern part of the state.

Health costs based on $PM_{2.5}$ concentrations are estimated using the CR function. For the three power plants studied, McDonough has the largest health cost per MWh generated ($123.3/MWh in July). This is because Plant McDonough is located in Atlanta which has a high population density (Note: the calculations conducted were for 2007 when Plant McDonough was using coal. Plant McDonough recently switched to natural gas, and has much lower emissions on a MWh basis). The geospatial dependency of health cost is well captured, enabling further applications for optimization and policy evaluation.

Our next steps are to include several scenarios within an electricity capacity planning model which seeks the optimal solution to minimize the sum of operating costs and health costs due to a change in emissions. Combined with the above cost estimates, we have developed an innovative view of the power future of Georgia, and potentially other regions, with predictions of changing spatial health impacts across the state of Georgia. Our objectives and corresponding scenarios are outlined below:

- Evaluate the health impacts from increased use of natural gas for power generation in Georgia, and of shutting down (planned and suggested) coal plants owned by Georgia Power.

- Inform decisions on new construction of EGUs, including type and location decisions, by balancing geospatial health impacts and transmission losses with increasing demand for power.
- Validate emissions sensitivities by using observed changes in $PM_{2.5}$ concentrations and changes in mortality from 2007 through 2012.

# CHAPTER 6

# CONCLUSTION

This dissertation explored the applications of the direct sensitivity technique in both scientific investigation and policy guidance, and demonstrated that the direct sensitivity approach yields important insights into the formation of pollutants such as particulate matter and ozone. It also demonstrated that the direct sensitivity technique can be linked with uncertainty analysis and economic analysis to inform the development of control strategies. In detail, below are given the major findings of this dissertation along with directions for further study.

## 6.1. Major findings

### 6.1.1 High-order sensitivity analysis of particulate matter

Chapter 2 of this dissertation described the implementation of the high-order Decoupled Direct Method in Three Dimensions (HDDM-3D) for particulate matter in the Community Multiscale Air Quality (CMAQ) model. The key step of the implementation was accurately capturing both first- and second-order sensitivities of ISORROPIA, the embedded thermodynamic equilibrium model in CMAQ. This dissertation improved the accuracy of the sensitivity calculation for ISORROPIA by using a case-specific approach to avoid the discontinuity when the chemical regime changes from one case to another, and an explicit representation of the sensitivities of activity coefficients and water content, which have been proved to have a significant impact on the sensitivities of particulate matter. The new implementation significantly improved the accuracy of first-order DDM sensitivities of ammonium and nitrate to ammonia and $NO_x$ emissions,

compared to the initial implementation of DDM for particulate matter in CMAQ. Comparisons of both first- and second-order sensitivities with the traditional brute force method generally showed good agreement. The brute force sensitivities are found to be dependent on the perturbation sizes and the model configuration of computing accuracy, which leads to noisy behavior, especially for second-order sensitivities. For example, different size perturbations and convergence criteria of ISORROPIA lead to dramatic discrepancies in second-order brute force sensitivities of nitrate to total sulfate, different size perturbations also lead to large discrepancies in second-order brute force sensitivities of both nitrate to $SO_2$ emissions and sulfate to $NH_3$ emissions. In contrast to the noisy behavior of the brute force method, the direct assessment of second-order sensitivities with HDDM-3D is more consistent and avoids the apparent pitfalls of the brute force method that causes the noise. Furthermore, the direct sensitivities are more reasonable in some cases and can be explained by real physical and chemical processes while the brute force results are dramatically influenced by numerical noise that accumulates during model simulation. For example, it is expected that $SO_2$ emissions control would slightly increase nitrate concentrations, which is confirmed by negative sensitivities of nitrate to $SO_2$ emissions throughout the modeling domain and episode. However, the brute force results have a significant number of sensitivities of nitrate to $SO_2$ that are positive, which would not logically follow from the real chemical reactions. An additional important benefit of using HDDM-3D to compute sensitivities is computational efficiency. The computational time required by one additional second-order sensitivity parameter is comparable to the time required by one additional first-order sensitivity parameter, and both first- and second-order sensitivities to all parameters are calculated in a single

simulation. On the other hand, one additional second-order sensitivity parameter requires three simulations of the base CMAQ, and even more simulations are needed when calculating cross sensitivities.

The implementation of HDDM-3D for particulate matter enhances the ability of the CMAQ model to assess source contribution and control strategy effectiveness. These applications are based on a reduced form model that directly relates changes in emissions rates and the responses of pollutant concentrations by using Taylor Series Expansions. Including second-order sensitivities of particulate matter in the reduced form model improves the accuracy of the estimation of pollutant response to emissions controls. For example, after the second-order sensitivities of nitrate to $NO_x$ emissions are included in the reduced form model, the estimations of nitrate aerosol concentrations with a 50% reduction in $NO_x$ emissions are closer to the results simulated by the original CMAQ model compared to the estimations achieved by using a reduced form model that includes only first-order sensitivities of nitrate to $NO_x$ emissions. This reflects the nonlinearity in the formation of nitrate aerosols from $NO_x$. This dissertation also tested the ability of the reduced form model to estimate pollutant responses to a wide range of emissions control strength. The reduced form model based on first- and second-order DDM sensitivities provides results that are in good agreement with those simulated by the original model when precursor emissions are reduced by 20%, 50%, and 100%.

This dissertation explores the application of HDDM-3D for estimating source contributions by conducting a case study in the Houston area. Source contributions to daily average of $PM_{2.5}$ concentrations from five major pollutant emissions through first- and second-order relationships indicate that the dominant contributor to $PM_{2.5}$ formation

in this region was the primary PM emissions, which were responsible for about 27% to 85% of daily average $PM_{2.5}$ concentrations throughout the episode. The second largest contributor to $PM_{2.5}$ formation in this region throughout this episode was $SO_2$ emissions, which contributed 5% to 15% to the daily $PM_{2.5}$ concentrations. The contributions of $NO_x$ emissions to daily $PM_{2.5}$ concentrations varied in a larger range than those of $SO_2$ emissions. The contribution of $NO_x$ emissions can vary between slightly negative to about 12%. The contributions from the interaction of various emissions sources were smaller than those from a single source, but some of them had noticeable magnitudes, such as the interactions between VOC and $SO_2$, VOC and $NH_3$, VOC and primary PM, $SO_2$ and $NH_3$. The variation of the magnitudes of those cross sensitivities reflects the change in the nonlinearity of the chemical and physical system, which was shown to relate to the meteorological conditions in this region. Investigations in the diurnal pattern of the source contributions of various emissions found high contributions to $PM_{2.5}$ concentrations during night hours (i.e., between 1 am to 8 am local time). Source contributions during the night hours from $NO_x$, $SO_2$, VOC, $NH_3$, and source interactions increased dramatically compared to those during the day time. Comparison with the diurnal pattern of source contributions to ozone reflects that the accumulated ozone during the day time led to rapid formation of particulate matter during the night by reacting with $NO_x$ and $SO_2$. These reactions enhanced the interactions between various pollutant emissions such as VOC and $SO_2$, VOC and $NH_3$, and $SO_2$ and $NH_3$.

In summary, the implementation of HDDM-3D for particulate matter in CMAQ is a powerful extension to the model, as it allows efficient estimations of both first- and second-order sensitivities of various species of particulate matter to different emission

sources, as well as the interactions between them. The results of HDDM-3D were evaluated by brute force sensitivities, and the estimations of the reduced form model which is constructed by using HDDM-3D sensitivities were evaluated by the results of the original CMAQ model. Good agreement is found in both evaluations. The case study in the Houston region demonstrated the ability of HDDM-3D to investigate the formation of particulate matter from various emissions sources and the interactions between different emissions sources.

6.1.2 Emissions-associated model uncertainty of particulate matter

Chapter 3 of this dissertation explored the emissions-associated uncertainties in model simulation of particulate matter. This chapter applied a reduced form model of CMAQ to conduct Monte Carlo simulations to assess the uncertainties in model simulation of $PM_{2.5}$ associated with uncertainties in emissions rates of $NO_x$, $SO_2$, VOC, $NH_3$, and primary PM in the Houston area. The medians of the uncertainties of simulated daily average of $PM_{2.5}$ concentrations at different levels ranged from 42% to 52%. The dominant contributor of emissions uncertainties to the uncertainties in simulated $PM_{2.5}$ was primary PM, which alone was estimated to cause 38% uncertainty in $PM_{2.5}$ simulations. One reason is that primary PM was the dominant contributor to the formation of $PM_{2.5}$ in the area, with a contribution over 80% to daily average $PM_{2.5}$ under certain meteorological conditions. The other reason is that primary PM was estimated to have relatively larger uncertainty among the five major pollutant emissions. This implies that reducing the uncertainty in the estimation of primary PM emissions may be an efficient method to improve the model simulation of $PM_{2.5}$ concentrations in the Houston region.

Chapter 3 also incorporated the ground measurement of $PM_{2.5}$ concentrations to investigate how much of the bias between model simulation and ground measurement of $PM_{2.5}$ is attributable to the uncertainties in emissions rates. At five observation sites in urban and suburban in the Houston area, 85% of the measured daily $PM_{2.5}$ concentrations fell into the 95% confidence interval associated with the uncertainties in emissions rates, which implies that the uncertainties in emissions rates can explain most of the bias between observations and simulations of $PM_{2.5}$ concentrations. Besides the uncertainties in estimated emissions rates, the uncertainties in meteorological fields were also responsible for the bias in $PM_{2.5}$ simulations. The model simulations were biased high on the days with poorer performance of the meteorological model. The errors in simulated $PM_{2.5}$ concentrations were more associated with the errors in simulated wind directions and temperature than with those in relative humidity and wind speed. Furthermore, the estimations of emissions rates were also dependent on meteorological conditions, so it is essential to conduct further research to explore more aspects of the impact of meteorological conditions on air quality model simulations. The comparison between observed and simulated $PM_{2.5}$ concentrations at the monitoring site that was close to high biogenic emissions shows a constant low bias in model simulation, which indicates that either the current emissions inventory underestimates the biogenic VOC emissions or the representations of the formation of $PM_{2.5}$ in the model underestimate the production of $PM_{2.5}$ from the biogenic emissions.

In summary, Chapter 3 demonstrated how DDM sensitivities can be applied to uncertainty analysis of model simulations. The reduced form model that is based on DDM sensitivities can efficiently estimate the model uncertainties caused by the

uncertainties in emissions rates. This chapter also shed light on the impact of the model uncertainties caused by the uncertainties in meteorological conditions and the uncertainties in model representations.

6.1.3 Assessment of emissions impact of flare VOC

Chapter 4 of this dissertation explored the application of DDM sensitivity to assess the emissions impact of a single point source. This chapter focused on the flare VOC emissions from the Deer Park plant located in the Houston Ship Channel and applied DDM sensitivities to address the impact of the variable destruction and removal efficiency (DRE) on ozone concentrations. This chapter designed two independent studies and conducted DDM sensitivity analysis with different types of sensitivity parameters. The first study focused more on the impact of three flare operating modes (i.e., continuous use without gas recovery, continuous use with partial gas recovery, and intermittent use with full gas recovery) on ozone concentrations and compared the difference in ozone concentrations caused by considering the decrease in the DREs of different operating modes. The differences in daily maximum 8-hour average ozone concentrations at 30 monitoring sites in the Houston area indicate that the first mode has the largest impact on ozone concentration. Among the 30 monitoring sites, the largest impact of DRE on daily maximum 8-hour average ozone concentrations, 0.26 ppb, occurred at the site located southwest of downtown Houston. The largest impact of decreased DREs on ozone concentrations at monitoring sites for the second mode (continuous use with partial gas recovery) was 0.09 ppb and that for the third mode (intermittent use with full gas recovery) was 0.01 ppb. Unlike the first and second modes, the third mode had an impact on only a few monitors because it flares only 5%-10% of

the time during start-up, shut-down, or maintenance. The comparison of the three flare operating modes indicates that the DREs of the intermittent use flare with full gas recovery have the least impact on ozone concentrations and implies the effectiveness of the flare gas recovery system in reducing the impact of flare VOC emissions on ozone formation.

The second study simulated the impact of a real flare of the Deer Park plant on ozone formation by using VOC emissions from the 2006 Texas special inventory and steam flow rates from a more recent inventory that was especially prepared for flare studies. This study constructed a reduced form model based on first- and second-order sensitivities of ozone to flare VOC emissions at intervals of every two hours of a day. The reduced form model has flexibility in modifying the flare VOC emissions at desired hours. The flare VOC emissions increased when the flare was improperly steamed, and the impact of the extra emissions of flare VOC on ozone was assessed by using the reduced form model. Increases of daily maximum 8-hour average ozone at 30 monitoring sites in the modeling domain indicated that decreased DREs had an impact on most of the monitoring sites in the three counties of Harris, Galveston, and Brazoria, as well as on some of the monitoring sites in Jefferson County. The largest increase in daily maximum 8-hour ozone concentrations at the monitoring sites, 2.9 ppb, occurred at a monitoring site in Jefferson County when strong westerly wind carried the flare plume to the east. This chapter also compared the source contributions of flare VOC emissions at different hours at two monitoring sites, one about 20 miles west of the flare and the other about 50 miles east of the flare. At the site closer to the flare, flare VOC emissions at morning hours (6:00 – 10:00) contributed more to ozone concentration than the other hours. At the

site in the more remote area, flare VOC emissions at night hours (22:00 – 2:00) contributed more to ozone concentration at this site than did flare VOC emissions at the other hours.

In summary, Chapter 4 assessed the impact of flare DRE's on ozone formation in the Houston area by using HDDM-3D sensitivity analysis. The differences between daily maximum 8-hour average ozone concentrations at constant and temporally variable DREs indicate significant impact of DRE on ozone concentrations in this region. The most significant impact was found when a flare was operated continuously without gas recovery. The impact of assist steam on flare VOC emissions and ozone concentrations can be large when both the vent gas flow rates and the steam to gas ratios are high.

6.1.4 Integration of an air quality model with an electricity generation planning model

Chapter 5 of this dissertation explored the application of the DDM sensitivity technique to integrate air-quality related health effects into an optimization model for electricity generation planning. The objective function of the optimization model is the total cost of electricity generation, including the air quality-related health costs. Incorporating the reduced form air quality model into the electricity generation planning model enables estimation of geospatial dependent health costs for different power plants and consideration of air quality-related policies in decision making. This chapter introduced the construction of a combined model of air quality and electricity planning and conducted a case study in Georgia to demonstrate the ability of the combined model to account for geospatial dependent air quality-related health costs.

The reduced form model of CMAQ was constructed based on sensitivities of $PM_{2.5}$ concentrations to $SO_2$ and $NO_x$ emissions from seven EGU point sources,

including five major power plants and aggregated sources in the northern and southern parts of the state. The different spatial patterns of the sensitivities reflect the dependency of the emissions source impact of power plants on locations. $SO_2$ emissions from Plant Bowen have the largest impact on $PM_{2.5}$ concentrations, which is 2 $\mu g\ m^{-3}$, followed by Plant Harllee with the second largest impact of 0.9 $\mu g\ m^{-3}$. The source impact of EGU $NO_x$ emissions is smaller in magnitude compared to that of EGU $SO_2$ emissions.

The air-quality related health costs were calculated by using the concentration-response (CR) function. Both the population density and mortality rates in this function were considered to vary with different geographic areas. The air-quality related costs for Plant McDonough due to its $SO_2$ emissions were estimated to be $123.3/MWh in July of 2007, which was the highest among the five EUG point sources. The reason is that Plant McDonough is located in Atlanta which has a high population density.

In summary, Chapter 5 of this dissertation established a framework for integrating the air quality model with the electricity generation planning model. The geospatial dependence of the air-quality related health costs were well represented by the integrated model. This integration provided an innovative view of the future of power generation and public health for the state of Georgia, and potentially other regions, and may make possible further research work on air quality-related health costs, air quality policies, and electricity generation planning.

### 6.2. Recommendations for future research

This dissertation has developed an efficient tool that provides important insights regarding source apportionment of particulate matter. The reduced form model established by this dissertation should have great potential in scientific investigation and

policy guidance. The following recommendations suggest the directions for future research.

6.2.1 Application of HDDM-3D of PM to other regions and episode

The high-order sensitivity technique for particulate matter developed in this dissertation is readily applicable to other locations and seasons. Given different emissions characterizations and meteorological conditions, the composition and concentrations of particulate matter may vary dramatically. West et al. (1999) indicated that in winter, a large area in the eastern U.S. is in the nonlinear regime for formation of particulate matter. A more recent study by Goldstein et al. (2009) suggested that the observed high aerosol optical thickness in summer time over the southeastern U.S. could be explained by secondary aerosol formed by the interaction between biogenic emissions and anthropogenic emissions such as $NO_x$ and $SO_2$. Assessments of cross-sensitivity interactions analogous to those conducted in Chapter 2 could be applied to the southeastern U.S. in both summer and winter time to shed light on the source contributions to secondary aerosol formation in this region.

6.2.2 Understanding the influence of uncertainty in meteorological conditions

As discussed in Chapter 3, meteorological inputs in an air quality model can strongly influence the simulated pollutant concentrations. The case study in Houston showed that the bias in simulations of fine particle concentrations was more associated with wind speed and temperature. It could be worthwhile to extend the DDM sensitivity technique to include sensitivity parameters from meteorological variables such as temperature and relative humidity. Even in the absence of DDM sensitivities to these meteorological inputs, analysis could be conducted using an ensemble of brute force

simulations. For example, we could perturb relative humidity by 5%, 10%, and 15% and compare the model simulations of particulate matter concentrations to quantify the extent to which different meteorological variables influence the simulated pollutant concentrations.

### 6.2.3 Assessing the emissions impact from multiple flares

The case study in Houston described in Chapter 4 demonstrated that the decrease in combustion efficiency of a single flare has the potential to cause over a 1 ppb increase in daily maximum 8-hr ozone at monitoring sites. Given the abundance of flares in this region, it could be valuable to extend the analysis to other flares located at different counties or all the flares in the modeling domain. Analogous analysis could be readily applied to several other flares. However, it would be difficult to analyze all the flares using the methods introduced in Chapter 4 because the flow rates of vent gas, assist steam, and assist gas are different from one flare to another. Moreover, detailed information of operating conditions of every flare is not available because some is confidential (B. Exum at TCEQ, personal communication). Profiles of operating conditions could be assigned to different types of flares, which can be determined by the Source Classification Codes from the emission inventory. DDM sensitivity analysis could then be conducted to assess the emissions impact of all the flares in this region.

### 6.2.4 The impact of air quality regulations on electricity generation planning

The integrated model developed in Chapter 5 is able to account for geospatial dependency of air quality-related health costs. The case study for power generation planning in Georgia serves as a good starting point to explore the potential of the integrated model to inform decision making. Future work has been planned to use this

model to evaluate the health effects of different fuel types (e.g., coal and natural gas) and provide advice on locations and types of EUGs that are planned to be constructed or shut down. Constraints could be added to address the impact of air quality regulations on electricity generation planning. For example, one constraint could be to bring an ensemble of monitors into attainment with the NAAQS.

### 6.3. Closing Remarks

While much work remains, important insights have been provided by this dissertation regarding the formation and interaction of particulate matter and ozone in southeastern Texas, the uncertainty in simulated particulate matter, and the applications of DDM sensitivities in cost optimization models. Efficient tools such as high-order DDM sensitivity analysis for PM and reduced form models of CMAQ have been developed and are readily applicable to other conditions. The methods developed by this dissertation have great potential in informing policy and scientific investigation and the findings of this dissertation provide a valuable reference for future research.

# REFERENCES


Allen, D. T., & Fraser, M. (2006). An overview of the Gulf Coast Aerosol Research and Characterization Study: The Houston Fine Particulate Matter Supersite. *Journal of the Air & Waste Management Association, 56*(4), 456-466.

Allen, D. T. and Torres, V. M. (2011). TCEQ 2010 Flare Study Final Report, TCEQ.

Ansari, A. S., & Pandis, S. N. (1998). Response of inorganic PM to precursor concentrations. *Environmental Science & Technology, 32*(18), 2706-2714. doi: 10.1021/es971130j.

Ansari, A. S., & Pandis, S. N. (2000). Water absorption by secondary organic aerosol and its effect an inorganic aerosol behavior. *Environmental Science & Technology, 34*(1), 71-77. doi: 10.1021/es990717q.

Banta, R. M., C. J. Seniff, J. Nielsen- Gammon, L. S. Darby, T. B. Ryerson, R. J. Alvarez, S. R. Sandberg, E. J. Williams, and M. Trainer (2005), A bad air day in Houston, *Bull. Am. Meteorol. Soc*., 86(5), 657–669, doi:10.1175/BAMS-86-5-657.

Baptista-Neto, J. A., Smith, B. J., McAllister, J. J., Silva, M. A. M., & Castanheira, F. S. (2006). Surface modification of a granite building stone in central Rio de Janeiro. *Anais Da Academia Brasileira De Ciencias, 78*(2), 317-330. doi: 10.1590/s0001-37652006000200011.

Bergin, M. S., Noblet, G. S., Petrini, K., Dhieux, J. R., Milford, J. B., & Harley, R. A. (1999). Formal uncertainty analysis of a Lagrangian photochemical air pollution model. *Environmental Science & Technology, 33*(7), 1116-1126. doi: 10.1021/es980749y

Bergin, M. S., Russell, A. G., Odman, M. T., Cohan, D. S., & Chameldes, W. L. (2008). Single-Source Impact Analysis Using Three-Dimensional Air Quality Models. *Journal of the Air & Waste Management Association, 58*(10), 1351-1359. doi: 10.3155/1047-3289.58.10.1351.

Binkowski, F. S., and Roselle, S. J. (2003). Models-3 community multiscale air quality (CMAQ) model aerosol component - 1. Model description, *Journal of Geophysical Research-Atmospheres*, 108.

Boylan, J. W., Odman, M. T., Wilkinson, J. G., Russell, A. G., Doty, K. G., Norris, W. B., & McNider, R. T. (2002). Development of a comprehensive, multiscale "one-atmosphere" modeling system: application to the Southern Appalachian Mountains. *Atmospheric Environment, 36*(23), 3721-3734. doi: 10.1016/s1352-2310(02)00356-4

Brauer, M., & Brook, J. R. (1997). Ozone personal exposures and health effects for selected groups residing in the Fraser Valley. *Atmospheric Environment, 31*(14), 2113-2121. doi: 10.1016/s1352-2310(96)00129-x

Bromley, L. A. (1973). Thermodynamic properties of strong electrolytes in aqueous solutions, *AIChE Journal*, 19, 313-320.

Byun, D., & 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*(1-6), 51-77. doi: 10.1115/1.2128636.

Carter, W.P.L. (2000). Documentation of the SAPRC99 chemical mechanism for VOC and reactivity assessment, Air Pollution Research Center and College of Engineering, Center for Environmental Research and Technology, University of California, Riverside, CA.

Center for Disease Control and Prevention, National Center for Health Statistics (CDC NCHS) (2012). Underlying Cause of Death 1999-2010. *CDC WONDER Online Database, released 2012*. Retrieved from http://wonder.cdc.gov/ucd-icd10.html

CEP (2003). Sparse Matrix Operator Kernel Emissions Modeling System (SMOKE) User manual; Carolina Environmental Program.

Cohan, D. S., Hakami, A., Hu, Y. T., & Russell, A. G. (2005). Nonlinear response of ozone to emissions: Source apportionment and sensitivity analysis. *Environmental Science & Technology, 39*(17), 6739-6748. doi: 10.1021/es05664m.

Cohan, D. S., Tian, D., Hu, Y. T., & Russell, A. G. (2006). Control strategy optimization for attainment and exposure mitigation: Case study for ozone in Macon, Georgia. *Environmental Management, 38*(3), 451-462. doi: 10.1007/s00267-005-0226-y

Cohan, D. S., Koo, B., & Yarwood, G. (2010). Influence of uncertain reaction rates on ozone sensitivity to emissions. *Atmospheric Environment, 44*(26), 3101-3109. doi: 10.1016/j.atmosenv.2010.05.034.

Committee on Health, Environmental, and Other External Costs and Benefits of Energy Production and Consumption, National Research Council. (2010). *Hidden Costs of Energy: Unpriced Consequences of Energy Production and Use*.

Cowling, E. B., Furiness, C., Dimitriades, B., Parrish, D., and Estes, M.: Final Rapid Science Synthesis Report: Findings from the Second Texas Air Quality Study Southern Oxidants Study Office of the Director at North Carolina State University Location: Raleigh, NC, USA, 2007

Deguillaume, L., Beekmann, M., & Derognat, C. (2008). Uncertainty evaluation of ozone production and its sensitivity to emission changes over the Ile-de-France region during summer periods. *Journal of Geophysical Research-Atmospheres, 113*(D2). doi: 10.1029/2007jd009081

Digar, A., & Cohan, D. S. (2010). Efficient Characterization of Pollutant-Emission Response under Parametric Uncertainty. *Environmental Science & Technology, 44*(17), 6724-6730. doi: 10.1021/es903743t

Dockery, D. W., Pope, C. A., Xu, X. P., Spengler, J. D., Ware, J. H., Fay, M. E., . . . Speizer, F. E. (1993). An Association Between Air-Pollution and Mortality in Six United-States Cities. *New England Journal of Medicine, 329*(24), 1753-1759. doi: 10.1056/nejm199312093292401.

Dunker, A. M. (1984). THE DECOUPLED DIRECT METHOD FOR CALCULATING SENSITIVITY COEFFICIENTS IN CHEMICAL-KINETICS. *Journal of Chemical Physics, 81*(5), 2385-2393. doi: 10.1063/1.447938.

Edney, E. O., Kleindienst, T. E., Lewandowski, M., and Offenberg, J. H. (2007) Updated SOA chemical mechanism for the Community Multi-Scale Air Quality model, EPA 600/X-07/025, U.S. EPA, Research Triangle Park, NC.

Emery, et al. (2001) Enhanced meteorological modeling and performance evaluation for two Texas ozone episodes, ENVRION, Novato, CA

Foley, K. M., Roselle, S. J., Appel, K. W., Bhave, P. V., Pleim, J. E., Otte, T. L., . . . Bash, J. O. (2010). Incremental testing of the Community Multiscale Air Quality (CMAQ) modeling system version 4.7. *Geoscientific Model Development, 3*(1), 205-226.

Fortner, E. C., Brooks, W. A., Onasch, T. B., Canagaratna, M. R., Massoli, P., Jayne, J. T., . . . Herndon, S. C. (2012). Particulate Emissions Measured During the TCEQ Comprehensive Flare Emission Study. *Industrial & Engineering Chemistry Research, 51*(39), 12586-12592. doi: 10.1021/ie202692y

Fountoukis, C., & Nenes, A. (2007). ISORROPIA II: a computationally efficient thermodynamic equilibrium model for K+-Ca2+-Mg2+-Nh(4)(+)-Na+-SO42--NO3--Cl--H2O aerosols. *Atmospheric Chemistry and Physics, 7*(17), 4639-4659.

Galloway, T. S., Brown, R. J., Browne, M. A., Dissanayake, A., Lowe, D., Jones, M. B., & Depledge, M. H. (2004). A multibiomarker approach to environmental assessment. *Environmental Science & Technology, 38*(6), 1723-1731. doi: 10.1021/es030570.

Goldstein, A. H., Koven, C. D., Heald, C. L., & Fung, I. Y. (2009). Biogenic carbon and anthropogenic pollutants combine to form a cooling haze over the southeastern United States. *Proceedings of the National Academy of Sciences of the United States of America, 106*(22), 8835-8840. doi: 10.1073/pnas.0904128106

Grell, G., Dudhia, J., and Stauffer, D.(1994). A description of the Fifth-Generation Penn State/NCAR Mesoscale Model (MM5), NCAR, NCAR/TN 398+STR.

Hakami, A., Odman, M. T., & Russell, A. G. (2003). High-order, direct sensitivity analysis of multidimensional air quality models. *Environmental Science & Technology, 37*(11), 2442-2452. doi: 10.1021/es020677h.

Hakami, A., Odman, M. T., & Russell, A. G. (2004). Nonlinearity in atmospheric response: A direct sensitivity analysis approach. *Journal of Geophysical Research-Atmospheres, 109*(D15). doi: 10.1029/2003jd004502.

Hanna, S. R., Chang, J. C., & Fernau, M. E. (1998). Monte Carlo estimates of uncertainties in predictions by a photochemical grid model (UAM-IV) due to uncertainties in input variables. *Atmospheric Environment, 32*(21), 3619-3628. doi: 10.1016/s1352-2310(97)00419-6

Hanna, S. R., Lu, Z. G., Frey, H. C., Wheeler, N., Vukovich, J., Arunachalam, S., . . . Hansen, D. A. (2001). Uncertainties in predicted ozone concentrations due to input uncertainties for the UAM-V photochemical grid model applied to the July 1995 OTAG domain. *Atmospheric Environment, 35*(5), 891-903. doi: 10.1016/s1352-2310(00)00367-8

Hu Y, Balachandran S, Pachon J, Baek J, Ivey C, Holmes H, Odman MT, Mulholland JA and Russell AG. Fine Particulate Matter Source Apportionment Using a Hybrid Chemical Transport and Receptor Model Approach. *In Preparation.*

IPCC, 2001. Climate change 2001: the scientific basis. Intergovernmnetal Panel on Climate Change. Cambridge University Press, Cambridge.

Ivey C, Holmes H, Hu Y, Mulholland JA, Russell AG. Spatial Extension of a Novel Hybrid Source Apportionment Model. *In preparation.*

Kaiser, J. (2005). Mounting evidence indicts fine-particle pollution. *Science, 307*(5717), 1858-1861.

Kelly, J., Bhave, P., Nolte, C., Shankar, U., and Foley, K. (2010) Simulating emission and chemical evolution of coarse sea-salt particles in Community Multiscale Air Quality (CMAQ) model, *Geoscientific Model Development*, 3, 257-273.

Koo, B., Dunker, A. M., & Yarwood, G. (2007). Implementing the decoupled direct method for sensitivity analysis in a particulate matter air quality model. *Environmental Science & Technology, 41*(8), 2847-2854. doi: 10.1021/es0619962.

Koo, B., Wilson, G., and Yarwood, G.(2010) High-order Decoupled Direct Method (HDDM) Radical Output Improvements in the Comprehensive Air-quality Model with Extensions (CAMx), 582-7-84005-FY10-27.

Krewski, D., Jerrett, M., Burnett, R.T., Ma, R., Hughes, E., Shi, Y., Turner, M.C., Pope III, A., Thurston, G., Calle, E.E., Thun, M.J. (2009). *Extended follow-up and spatial analysis of the American Cancer Society study linking particulate air pollution and mortality*. Health Effects Institute, Research Report 140. May 2009.

Latha, K. M., & Badarinath, K. V. S. (2005). Shortwave radiative forcing efficiency of urban aerosols - a case study using ground based measurements. *Chemosphere, 58*(2), 217-220. doi: 10.1016/j.chemosphere.2004.09.013.

Mendoza-Dominguez, A., & Russell, A. G. (2000). Iterative inverse modeling and direct sensitivity analysis of a photochemical air duality model. *Environmental Science & Technology, 34*(23), 4974-4981. doi: 10.1021/es991040.

Meng, Z., Dabdub, D., & Seinfeld, J. H. (1997). Chemical coupling between atmospheric ozone and particulate matter. *Science, 277*(5322), 116-119. doi: 10.1126/science.277.5322.116.

Murphy, C. F., & Allen, D. T. (2005). Hydrocarbon emissions from industrial release events in the Houston- Galveston area and their impact on ozone formation. *Atmospheric Environment, 39*(21), 3785-3798. doi: 10.1016/j.atmosenv.2005.02.051.

Napelenok, S. L., Cohan, D. S., Hu, Y. T., & Russell, A. G. (2006). Decoupled direct 3D sensitivity analysis for particulate matter (DDM-3D/PM). *Atmospheric Environment, 40*(32), 6112-6121. doi: 10.1016/j.atmosenv.2006.05.039.

Napelenok, S. L., Foley, K. M., Kang, D. W., Mathur, R., Pierce, T., & Rao, S. T. (2011). Dynamic evaluation of regional air quality model's response to emission reductions in the presence of uncertain emission inventories. *Atmospheric Environment, 45*(24), 4091-4098. doi: 10.1016/j.atmosenv.2011.03.030

NARSTO (2004) Particulate Matter Science for Policy Makers: A NARSTO Assessment. P. McMurry, M. Shepherd, and J. Vickery, eds. Cambridge University Press, Cambridge, England. ISBN 0 52 184287 5.

Nenes, A., Pandis, S. N., & Pilinis, C. (1998a). ISORROPIA: A new thermodynamic equilibrium model for multiphase multicomponent inorganic aerosols. *Aquatic Geochemistry, 4*(1), 123-152. doi: 10.1023/a:1009604003981.

Nenes, A., Pilinis, C., and Pandis, S. N. (1998b). Continued Development and Testing of a New Thermodynamic Aerosol Module for Urban and Regional Air Quality Models, Atmospheric Environment, 33, 1553-1560.

Odman, M. T., and Ingram, C. L. (1996) Multiscale Air Quality Simulation Platform (MAQSIP): Source code documentation and validation, Environmental Programs, North Carolina Supercomputing Center, Research Triangle Park, NC.

Odman, M. T., Boylan, J. W., Wilkinson, J. G., Russell, A. G., Mueller, S. F., Imhoff, R. E., . . . McNider, R. T. (2002). Integrated modeling for air quality assessment: The Southern Appalachians mountains initiative project. *Journal De Physique Iv, 12*(PR10), 211-234. doi: 10.1051/jp4:20020461.

Odum, J. R., Jungkamp, T. P. W., Griffin, R. J., Forstner, H. J. L., Flagan, R. C., and Seinfeld, J. H. (1997) Aromatics, reformulated gasoline, and atmospheric organic aerosol formation, Environmental Science & Technology, 31, 1890-1897.

Parrish, D. D., Allen, D. T., Bates, T. S., Estes, M., Fehsenfeld, F. C., Feingold, G., . . . Williams, E. J. (2009). Overview of the Second Texas Air Quality Study (TexAQS II) and the Gulf of Mexico Atmospheric Composition and Climate Study (GoMACCS). *Journal of Geophysical Research-Atmospheres, 114*. doi: 10.1029/2009jd011842

Pavlovic, R. T., Al-Fadhli, F. M., Kimura, Y., Allen, D. T., & McDonald-Buller, E. C. (2012). Impacts of Emission Variability and Flare Combustion Efficiency on Ozone Formation in the Houston-Galveston-Brazoria Area. *Industrial & Engineering Chemistry Research, 51*(39), 12593-12599. doi: 10.1021/ie203052w

Peel, J. L., Tolbert, P. E., Klein, M., Metzger, K. B., Flanders, W. D., Todd, K., Frumkin, H. (2005). Ambient air pollution and respiratory emergency department visits. *Epidemiology*, *16*(2), 164-174. doi: 10.1097/01.ede.0000152905.42113.db.

Pinder, R. W., Gilliam, R. C., Appel, K. W., Napelenok, S. L., Foley, K. M., & Gilliland, A. B. (2009). Efficient Probabilistic Estimates of Surface Ozone Concentration Using an Ensemble of Model Configurations and Direct Sensitivity Calculations. *Environmental Science & Technology*, *43*(7), 2388-2393. doi: 10.1021/es8025402.

Ramachandran, Q. and Vincent, J.H. (1999). A Bayesian approach to retrospective exposure assessment. *Applied Occupational and Environmental Hygiene*. 1999;14(8):547-557.

Rappenglück, B., et al. (2009). Air quality modeling of TexAQS- II episodes with data assimilation, final report, Univ. of Houston, Houston, Tex. (Available at http://projects.tercairquality.org:80/AQR/H098).

Ryerson, T. B., Trainer, M., Angevine, W. M., Brock, C. A., Dissly, R. W., Fehsenfeld, F. C., . . . Sueper, D. T. (2003). Effect of petrochemical industrial emissions of reactive alkenes and $NO_x$ on tropospheric ozone formation in Houston, Texas. *Journal of Geophysical Research-Atmospheres, 108*(D8). doi: 10.1029/2002jd003070

Sarnat, S. E., Coull, B. A., Schwartz, J., Gold, D. R., & Suh, H. H. (2006). Factors affecting the association between ambient concentrations and personal exposures to particles and gases. *Environmental Health Perspectives, 114*(5), 649-654. doi: 10.1289/ehp.8422.

Schell, B., Ackermann, I. J., Hass, H., Binkowski, F. S., and Ebel, A.(2001) Modeling the formation of secondary organic aerosol within a comprehensive air quality model system, *Journal of Geophysical Research-Atmospheres*, 106, 28275-28293.

Seaman, N. L. (2000). Meteorological modeling for air-quality assessments. *Atmospheric Environment, 34*(12-14), 2231-2259. doi: 10.1016/s1352-2310(99)00466-5.

Seinfeld, J. H., and Pandis, S. N. (2006) Atmospheric Chemistry and Physics: From Air Pollution to Climate Change, 2 ed., Wiley, New York.

Siler-Evans, K., Azevedo, I. L., Morgan, M. G., & Apt, J. (2013). Regional variations in the health, environmental, and climate benefits of wind and solar generation. *Proceedings of the National Academy of Sciences*, 110(29). doi:10.1073/pnas.1221978110

Simon, H., Baker, K. R., Akhtar, F., Napelenok, S. L., Possiel, N., Wells, B., & Timin, B. (2013). A Direct Sensitivity Approach to Predict Hourly Ozone Resulting from Compliance with the National Ambient Air Quality Standard. *Environmental Science & Technology*, *47*(5), 2304-2313. doi: 10.1021/es303674e.

Smith, J. (2012) Modeling Flare Destruction and Removal Efficiency. A Presentation to Southeast Texas Photochemical Modeling Technical Committee. Available at http://www.tcEquationtexas.gov/assets/public/implementation/air/am/committees/pmt_set/20121023/20121023_Flare_DRE_Modeling.pdf.

Stokes, R. H., and Robinson, R. A. (1966). Interactions in aqueous nonelectrolyte solutions I: Solute-solvent equilibria, J Phys Chem-Us, 70, 2126.

Schwartz, J. (1994). $PM_{10}$, Ozone, and Hospital Admissions for the Elderly in Minneapolis-Paul, Minnesota. *Archives of Environmental Health, 49*(5), 366-374.

TCEQ (2009a) Emissions Modeling for the HGB Attainment Demonstration SIP Revision for the 1997 Eight-Hour Ozone Standard. Available at http://www.tceq.state.tx.us/assets/public/implementation/air/sip/hgb/hgb_sip_2009/09017SIP_ado_Appendix_B.pdf.

TCEQ (2009b) Title 40 Code of Federal Regulations: Protection of Environment Part 60: Standards of Performance for New Stationary Sources Subpart A: General Provisions. Available at http://www.tceq.state.tx.us/assets/public/implementation/air/rules/Flare/AppendixD_FTF_Draft_Report.pdf.

Tian, D., Cohan, D. S., Napelenok, S., Bergin, M., Hu, Y. T., Chang, M., & Russell, A. G. (2010). Uncertainty Analysis of Ozone Formation and Response to Emission Controls Using Higher-Order Sensitivities. *Journal of the Air & Waste Management Association, 60*(7), 797-804. doi: 10.3155/1047-3289.60.7.797.

Torres, V. M., Herndon, S., & Allen, D. T. (2012). Industrial Flare Performance at Low Flow Conditions. 2. Steam- and Air-Assisted Flares. *Industrial & Engineering Chemistry Research, 51*(39), 12569-12576. doi: 10.1021/ie202675f

US Census Bureau. (2010). *Census 2010*. Washington, D.C.

U.S. EPA (2004) Air quality criteria for particulate matter. EPA/600/P-99/002aF.

U.S. EPA (2006) EPA Mortality Risk Evaluation. Available at http://yosemite.epa.gov/ee/epa/eerm.nsf/vwAN/EE-0568-22.pdf/$file/EE-0568-22.pdf

U.S. EPA. (2007). Guidance on the Use of Models and Other Analyses for Demonstrating Attainment of Air Quality Goals for Ozone, $PM_{2.5}$, and Regional Haze. *EPA -454/B-07-002*.

U.S. EPA. (2013a). National Ambient Air Quality Standards (NAAQS). Retrieved from http://www.epa.gov/air/criteria.html

US EPA. (2013b). US EPA BenMAP. Washington, D.C. Retrieved from http://www.epa.gov/airquality/benmap/download.html

Watson, J. G. (2002). Visibility: Science and regulation. *Journal of the Air & Waste Management Association, 52*(6), 628-713.

Webster, M., Nam, J., Kimura, Y., Jeffries, H., Vizuete, W., & Allen, D. T. (2007). The effect of variability in industrial emissions on ozone formation in Houston, Texas. *Atmospheric Environment, 41*(40), 9580-9593. doi: 10.1016/j.atmosenv.2007.08.052.

Wert, B. P., Trainer, M., Fried, A., Ryerson, T. B., Henry, B., Potter, W., . . . Wisthaler, A. (2003). Signatures of terminal alkene oxidation in airborne formaldehyde measurements during TexAQS 2000. *Journal of Geophysical Research-Atmospheres, 108*(D3). doi: 10.1029/2002jd002502.

West, J. J., Ansari, A. S., & Pandis, S. N. (1999). Marginal $PM_{2.5}$: Nonlinear aerosol mass response to sulfate reductions in the Eastern United States. *Journal of the Air & Waste Management Association, 49*(12), 1415-1424.

Xiao, X., Cohan, D. S., Byun, D. W., & Ngan, F. (2010). Highly nonlinear ozone formation in the Houston region and implications for emission controls. *Journal of Geophysical Research-Atmospheres, 115*. doi: 10.1029/2010jd014435.

Yang, Y. J., Wilkinson, J. G., & Russell, A. G. (1997). Fast, direct sensitivity analysis of multidimensional photochemical models. *Environmental Science & Technology, 31*(10), 2859-2868. doi: 10.1021/es970117w.

Zanobetti, A., Schwartz, J., & Dockery, D. W. (2000). Airborne particles are a risk factor for hospital admissions for heart and lung disease. *Environmental Health Perspectives, 108*(11), 1071-1077. doi: 10.1289/ehp.001081071.

Zhang, W., Capps, S. L., Hu, Y., Nenes, A., Napelenok, S. L., & Russell, A. G. (2012). Development of the high-order decoupled direct method in three dimensions for particulate matter: enabling advanced sensitivity analysis in air quality models. *Geoscientific Model Development, 5*(2), 355-368. doi: 10.5194/gmd-5-355-2012

# VITA

## Wenxian Zhang

Wenxian Zhang was born in Leshan, China, in February 1984. She received her B.A. in Atmospheric Sciences in 2006 from Peking University, where she was introduced to atmospheric modeling by Professor Benkui Tan. In the same year, she began graduate studies at Georgia Institute of Technology's School of Earth and Atmospheric Sciences (EAS) under the advisement of Professor Robert Dickinson, studying the interaction between land and atmosphere. After Wenxian earned her M.S. in EAS in August of 2008 from Georgia Institute of Technology, she joined the research group of Professor Ted Russell to work on the development of advanced sensitivity techniques in air quality models. Wenxian married Peng Xu in 2011, and they have a lovely daughter named Summer.