VARIATIONAL INVERSE MODELING OF REGIONAL NITROGEN OXIDES EMISSIONS

By

Farid Amid

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

**Doctor of Philosophy
In Environmental Engineering**

Department of Civil and Environmental Engineering
Carleton University, Ottawa, Canada

© 2015, Farid Amid

ABSTRACT

Inverse modeling methods have been widely used for model performance improvement and parameter estimation. For air quality studies, inverse modeling is often used for emission inversion as emissions are associated with significant amount of uncertainties. Emission inversion aims to find the optimized perturbations for emissions referred to as scaling factors that reduce the distance between model output and observations. Nitrogen oxides emissions are estimated through a four dimensional variational (4D-Var) inverse modeling approach using Ozone Monitoring Instrument (OMI) and ground-network observations. The modeling period was chosen July 2007 over North America domain.

The Community Multiscale Air Quality model is used along with Sparse Matrix Operator Kernel Emissions (SMOKE) and Weather Research and Forecasting model (WRF) for emissions and meteorological modeling.

In the non-temporal set-ups, NO_x emissions are adjusted by assimilating NO_2 columns from satellite (OMI) and NO_2 from ground based observations separately. The results indicate that the average scaling factors vary from 0.41 to 1.74 or from 0.45 to 2.52 when ground observations and OMI observations are used, respectively. The average scaling factors are in the range of 0.43 to 1.79 when both types of observations were used simultaneously. The total amount of emissions increase 3.5% when inversion is based on ground observations only, 17.3% when inversion includes only OMI observations and 13.6% when both observations are used simultaneously.

Under the temporal set-up, NO_x emissions are adjusted by using OMI and modified surface observations to estimate hourly profiles of emissions at each location. Hourly profiles show that

in large urban locations, CMAQ tends to overestimate the NO_2 column densities as an example at Los Angeles during the simulation period the emissions are reduced about %40

Lightning is one of the more uncertain sources of NO_2 . While inverted emissions included point, mobile, biogenic, and other sources in the CMAQ, lightning as one of the most significant natural sources of NO_x was excluded in previous inversion setups. To account for the impact of this source, lightning emissions are parameterized as inputs for the CMAQ model from the observations of Lightning Detection Networks flash rates. Places and observations that were dominated by anthropogenic sources were filtered. However, scarcity of data did not allow us to extend emission inversions for lightning events across the entire domain and during all episodes.

ACKNOWLEDGMENT

I deeply appreciate the tremendous support and help I have received from many individuals during my Ph.D. career. First of all, I would like to thank my supervisor Dr. Amir Hakami, for his continuous support and guidance. He has been an inspiration for me with his terrific scientific enthusiasm. Secondly, I would like to thank my Ph.D. thesis committee members Dr. Xiaohong XU, Dr. Dariush Motazedian, Dr. Abhijit Sarkar, and Dr. Ousmane Seidou, for their constructive comments and suggestions during my research study. Their visions and insights not only helped me to better address the scientific problems, but also guided me how to think more critically and creatively as a scientist. My thanks go to all the members of the Carleton Atmospheric Modeling Group, especially Dr. Morteza Mesbah and Matthew Russell.

While writing this acknowledgement, I remember many days I have spent researching, debugging, plotting, writing, and preparing this thesis. I enjoyed every minute of it and feel proud for this fantastic experience and the knowledge I gained.

Last but not least, I would like to express my sincerest gratitude to my family, for their tremendous love, blessing, understanding, and support throughout these years.

Farid Amid

TABLE OF CONTENTS

CHAPTER ONE: Introduction and objectives	1
CHAPTER TWO: Background	7
2.1 Inverse modeling	7
2.1.1 Optimal interpolation	13
2.1.2 Sequential methods	14
2.1.3 Variational methods	17
2.2 Sensitivity Analysis	29
2.2.1 Forward sensitivity analysis	30
2.2.2 Backward sensitivity analysis.....	34
2.3 Observations.....	37
2.3.1 Satellites	38
2.3.2 The ozone monitoring instrument (OMI).....	41
2.3.3 Retrieval approach for OMI	44
2.3.4 DOMINO Level 2 product	47
2.3.5 Ambient air monitoring networks.....	49
CHAPTER THREE: Methods and modeling tools.....	53
3.1 Community Multiscale Air Quality model	54
3.2 Adjoint of CMAQ	57
3.3 Sparse Matrix Operator Kernel Emissions	58

3.4	Weather Research and Forecasting	59
3.5	Inverse modeling.....	60
CHAPTER FOUR: Variational inverse modeling of regional NO _x emissions based on satellite and ground based observations		
		66
4.1.	Introduction.....	66
4.2.	Methods	70
4.3	Inversion set-up.....	78
4.4.	Results.....	80
4.4.1	Identical twin experiment using pseudo observational data.....	80
4.4.2	Global weighting factor.....	84
4.4.3	Emission inversion.....	85
4.5.	Summary and conclusion.....	104
CHAPTER FIVE: Large scale inverse modeling of NO _x emissions and their time dependent profiles.....		
		106
5.1	Introduction.....	106
5.2	Methodology	109
5.2.1	Chemical transport model and its adjoint.....	110
5.2.2	Observations	111
5.2.2.1	Ground NO ₂ measurements	111
5.2.2.2	OMI measurements.....	113

5.2.3	Background error covariance matrix	114
5.2.4	Experiment set-up.....	116
5.3	Results & Discussion.....	118
5.3.1	Correlation length	118
5.3.2	Emission inversion.....	120
5.3.3	Variability.....	122
5.3.4	Forecasting	139
5.4	Conclusions.....	141
CHAPTER SIX: Inversion of lightning NO emissions.....		143
6.1	Introduction.....	143
6.2	Methodology	146
6.2.1	LNO	146
6.2.2	CMAQ.....	149
6.2.3	OMI observations	149
6.2.4	Filtering	149
6.2.5	Inversion set-up	151
6.3	Results.....	152
6.3.1	Identical twin experiment.....	152
6.3.2	Emission inversion.....	154
6.3.3	Effect of LNO on ozone and nitrogen dioxide concentrations.....	161

6.4	Conclusions.....	162
	CHAPTER SEVEN: Summary and future work	163
	REFERENCES	170

LIST OF FIGURES

Figure 2-1. Schematic procedures of variational (left) and sequential (right) inversion methods	18
Figure 2-2. Steepest – descent schematic illustration	29
Figure 2-3. OMI measurement design	43
Figure 2-4. Size of OMI ground pixels	43
Figure 2-5. Schematic representation of slant path (red line) and vertical path (blue line) of the light	45
Figure 2-6. U.S. EPA Ozone ground network distribution	51
Figure 2-7. U.S. EPA Nitrogen Dioxide ground network distribution	52
Figure 3-1. North American domain.....	54
Figure 3-2. Model-3 components	56
Figure 3-3. Comparing the pressure layers of CMAQ and OMI (left), comparing the coverage of each CMAQ grid cell with OMI observations (right).....	62
Figure 3-4. Schematic process of emission inversion.....	64
Figure 4-1. U.S. EPA Nitrogen Dioxide ground network distribution	72
Figure 4-2. Schematic form of 4D-Var system	78
Figure 4-3. Identical twin experiment results, scaling factors to calculate false state (a), optimized scaling factors recovered from inversion (b), rate of cost function reduction during the iterative inversion process (c)	82

Figure 4-4. NO ₂ concentrations in identical twin experiment initial (a), observations (b) and optimized case (c)	83
Figure 4-5. Percentage of observational cost function reduction with different global weighting factor (Case-G)	84
Figure 4-6. Percentage of observational cost function reduction with different global weighting factor (Case-OG).....	85
Figure 4-7. The mean value of daily scaling factors from inversion of ground observations only (Case-G)	88
Figure 4-8. Relationship between scaling factors and emissions (a) and scaling factors and their standard deviation (b) (Case-G).....	89
Figure 4-9. Mean of scaling factors from inversion of OMI observations (Case-O)	91
Figure 4-10. Relationship between scaling factors and emissions (a) and scaling factors and their standard deviation (b) (Case-O).....	92
Figure 4-11. Mean of scaling factors from inversion of OMI observations (Case-OG).....	93
Figure 4-12. Relationship between scaling factors and emissions (a) and scaling factors and their standard deviation (b) (Case-O).....	96
Figure 4-13. Scaling factors of inversion set ups for Austin,Texas (a) and Los Angeles (b), California.....	98
Figure 4-14. Comparison of concentrations from inverted emissions, Austin (a), Los Angeles (b)	100

Figure 4-15. Weekday (a) and weekend (b) scaling factors from inversion of OMI and ground observations	102
Figure 4-16. Comparison of scaling factors of two different set ups as monthly scaling factors; monthly assimilation window (a) and average monthly (b).....	103
Figure 5-1. NO ₂ concentrations before (a) and after (b) applying correction factors for the observations of 10 ground stations	113
Figure 5-2. Grid cells with perturbed emissions (a), and corresponding difference between NO _x concentrations before and after perturbation; 1% (b), 5% (c), 10% (d).....	119
Figure 5-3. Error correlations for a grid cell when correlation length is 36 km.....	120
Figure 5-4. Comparison of NO _x concentrations (a) before and after optimization and related scaling factors (b) in Los Angeles; weekday (July 2nd 2007).	121
Figure 5-5. Weekend and weekdays scaling factors L.A., Kansas City and Washington	123
Figure 5-6. Comparison of before and after applying scaling factors on emissions (LA, Kansas City and New York).....	124
Figure 5-7. Comparison of before and after applying scaling factors on emissions (Average of 31 days, Austin, Texas), average of weekdays (a), average of weekends (b) and average of all 31 days (c)	126
Figure 5-8. Comparison of before and after applying scaling factors on emissions (Average of 31 days, Los Angeles, California), average of weekdays (a), average of weekends (b) and average of all 31 days (c).....	127
Figure 5-9. Mean (a) and standard deviation (b) of the scaling factors at the time of OMI overpass during weekends.....	129

Figure 5-10. Mean (a) and standard deviation (b) of the scaling factors at the time of OMI overpass during weekdays	130
Figure 5-11. Mean (a) and standard deviation (b) of the scaling factors at the time of OMI overpass during July.....	131
Figure 5-12. Median (a) and 1- σ inter-percentile variability (half of the difference between 16 and 84 percentiles) (b) of scaling factors on the entire domain during July	133
Figure 5-13. Percentage of reduction in observational cost function during July 2007	134
Figure 5-14. Average of the NO _x emissions at 5 a.m. (a) and 2 p.m. (b)	136
Figure 5-15. Difference between concentrations before (a) and after (b) inversion at 10 a.m....	137
Figure 5-16. Difference between concentrations before (a) and after (b) inversion at 3 p.m.	138
Figure 5-17. Difference between RMSE before and after inversion for all days	139
Figure 6-1. The percentage of NO observed emissions distributed based on local time (a) and CMAQ layer (b).....	148
Figure 6-2. Grid cells containing OMI observations and lightning during July 2007	151
Figure 6-3. Results of identical twin experiment for absolute emissions, true state (a), false state (b) and optimized (c).....	153
Figure 6-4. Comparison between Vertical Column densities before (a) and after inversion (b), OMI VCDs (c) for July 1st 2007	154
Figure 6-5. Improvement of column densities, before inversion (a) after inversion (b), July 1st 2007.....	155

Figure 6-6. Comparison between Vertical Column densities before (a) and after inversion (b), OMI VCDs (c) for July 26th 2007	156
Figure 6-7. Vertical change of the emissions after inversion, over the ocean (a), over the land (b)	157
Figure 6-8. Column of emissions (6 hour average), after inversion (a) and before inversion (b)	160

LIST OF TABLES

Table 2-1. Chemical species and related products from sensors.....	39
Table 2-2. List of major current satellites in air quality studies.....	40
Table 2-3. OMI instrument characteristics.....	42
Table 4-1. Statistical parameters regarding CMAQ simulations and comparison with data.....	87
Table 4-2. Percentage of distribution of scaling factors at different layers	90
Table 4-3. Percentage of change in total emissions for different inversion set ups.....	94
Table 4-4. Statistical parameters regarding daily inversions for ground observations (Case-OG).....	95
Table 4-5. Percentage of distribution of scaling factors at different layers	101
Table 4-6. Percentage of change in total emissions for different inversion set ups.....	102
Table 5-1. Percentage of scaling factors greater and smaller than one in West and East.....	135
Table 6-1. Statistical results from NO inversion	159

LIST OF ABBREVIATIONS

3D-Var	Three Dimensional Variational
4D-Var	Four Dimensional Variational
ADE	Atmospheric Diffusion Equation
AMF	Air Mass Factor
AQM	Air Quality Model
AQS	Air Quality System
BF	Brute-Force
BFGS	Broyden–Fletcher–Goldfarb–Shanno
CCTM	CMAQ Chemical Transport Model
CMAQ	Community Multiscal Air Quality Model
CTM	Chemical Transport Model
DDM	Decoupled Direct Method
DOMINO	Dutch OMI NO ₂
ENKF	Ensemble Kalman Filter
EPA	Environmental Protection Agency
ICARTT	International Consortium for Atmospheric Research on Transport and Transformation

NAAQS	National Ambient Air Quality standards
NLDN	National Lightning Detection Network
SAPRC	Statewide Air Pollution Research Center
SMOKE	Sparse Matrix Operator and Kernel Emissions
TLM	Tangent Linear Model

CHAPTER ONE:

Introduction and objectives

Air pollution has harmful effects on human health (such as respiratory problems) and other living organisms on the planet Earth. Health concerns related to pollutants emitted into the atmosphere have resulted in the progressive development of chemistry transport models (CTMs), or as they are often referred to at the policy-relevant regional scale, air quality models (AQMs). Air quality related studies often rely on CTMs and AQMs as they can account for different complex chemical and physical processes occurring in the atmosphere. Despite the significant improvement in AQMs, their capability in predicting air quality remains a challenging problem.

National Ambient Air Quality Standards (NAAQS) are established by the U.S. Environmental Protection Agency (U.S. EPA) concerning six pollutants in order to achieve the aims of the Clean Air Act which intends to protect public health and the environment from the damages caused by air pollution. These pollutants are nitrogen dioxide (NO_2), particulate matter (PM), carbon monoxide (CO), ozone (O_3), lead (Pb) and sulfur dioxide (SO_2). Canada has adopted similar standards in the form of Canada-Wide Standards, and later Canadian Ambient Air Quality Standards (CCME, 2012). To ascertain/evaluate attainment of these standards, AQMs

are used to simulate concentrations of criteria pollutants across various locations. AQMs are also used in order to better understand the formation processes and transport of air pollutants to develop emission control strategies to achieve the goal of improving air quality (National Research Council, 2007). However, the outputs of AQMs are affected by various sources of uncertainties and thus it is critical for researches and policy makers to reduce these uncertainties.

There are two main types of uncertainties in air quality studies; parametric and structural uncertainties. Parametric uncertainties come from errors in various inputs, including those from upstream models such as emission inputs or meteorological inputs. Structural uncertainties come from inaccurate and inadequate representation of physical and chemical mechanisms.

There are large uncertainties associated with AQMs' inputs. One major source of input uncertainties is incomplete and/or inaccurate emission inventories (Hanna et al, 2001). There are two main approaches for constructing emission inventories, “bottom-up” and “top-down”. For the bottom-up method, the inventory data is collected from activity information, emission factors, speciation, and spatial and temporal allocations. For industries and agricultural sectors, inventories are estimated from in-situ data collection. For mobile sources, emission inventories are estimated from source-specific emission factors indicating the relationship between the amount of emissions and the fuel consumed for different vehicles based on the travelled distance. Under the bottom-up approach, emission inventories are evaluated for every single source category and then are aggregated, considering spatial and temporal allocations to obtain the total amount of AQM-ready gridded emissions. However, there is a large amount of errors and uncertainties associated in each step used for total emission calculation. Meanwhile, the top-down approach is based on total emission estimations. This method uses statistical information (e.g. the number of power plants and their total fuel consumption, etc) to assess the total

emissions across the whole simulation domain. At the next step, this total emission estimate will be disaggregated by location and time based on the information about individual emission sources.

Uncertainty in emission inventories can be addressed in two perspectives: lack of knowledge about the sources and size of the inaccuracy in the emission inventories; reliability of the collected information from different sources (van Aardenne, 2002). Sources of inaccuracy can be expressed as inaccuracy in the structure of the emission inventory and input value inaccuracy (values of emission factors and activity data). Uncertainty due to reliability can come from transparency (insufficient documentation about sources), consistency (data collection from different methods applied for different based years), completeness (omission of pollutants or sources) and comparability (incomplete reporting about sources and/or deviations in aggregation). These uncertainties can be reduced to some extent by using different methods; one of them is inverse modeling that is addressed in this work.

Inverse modeling is a mathematical approach for estimating model parameters (inputs) by minimizing the disparity between model output and observations. Uncertainties in model outputs are mostly due to uncertainties in model parameters (inputs), the mathematical representation of governing equations, and numerical solution methods. The main objective of inverse modeling is to use observations to identify optimal values for model parameters. Observations of species in the atmosphere (e.g., ozone concentrations) can be obtained from different sources such as ships, aircrafts, conventional ground-based networks, and more recently, from satellites with significant and continuous spatial and temporal coverage providing valuable information for inverse modeling of emissions.

In atmospheric studies, chemistry transport models are formulated to simulate the spatial and temporal distribution of concentrations for different pollutants known as the state of the atmosphere. However, the performance of these models is strongly dependent on the input parameters such as emissions which are known as one of the most uncertain input parameters. This research examines whether uncertainties in input parameters (in this thesis, emissions) can be reduced using an iterative inversion method. In most of the previous studies, the focus of the research was on the possibility of recovering the sources using inverse modeling. Therefore, the model runs were for a few days and the observations were averaged for the period. In this work, not only the inverse modeling of emissions for a longer period of time was addressed, but also the behavior and variability of updated emissions on a daily basis were studied. To investigate the hypothesis, different inversion setups were built for July 2007 as Strunk et al., (2010) suggested that the number of observations from satellite is limited in winter because of lower boundary layer heights comparing to summer. First, the effectiveness of the inversion setup is tested, and then different sets of observations are used in order to scale the input emissions and find the optimal emission rates. While previous research studies have often focused on constraining emissions at different locations, also the temporal inversion of emission rates was considered. The two main setups that are designed for emission inversion in this thesis are non-temporal and temporal emission inversions. In the non-temporal case, emissions at all times at each location are perturbed with one value (which is called a scaling factor), while in the temporal case, emissions at each location are perturbed with different scaling factors at each time. Based on the observations used for inverting emissions there are 3 different setups; only ground-based network observations, only satellite observations, and combining both ground-based and satellite observations.

The objective of this work is to develop and use state-of-the-art tools and mathematical frameworks to constrain available emissions. To accomplish this goal, observations from various platforms (satellite, ground-based, etc.) need to be collected, processed, and mapped/incorporated into the modeling process. The main focus of this research is to combine inverse modeling methods within an air quality modeling framework to achieve reliable emission adjustments. The results of this work would be daily and hourly emission scaling factors that can be used to correct emissions and consequently lead to an improvement in performance of CTMs.

This thesis is organized into 7 chapters as follows:

Chapter two contains the literature review of the subject. In this chapter inverse modeling and related studies regarding emission inversion are reviewed. Also, a brief introduction and description of inverse modeling materials and methods is included. Different inversion methods with more focus on variational approaches are discussed.

Chapter three covers the methodology for the inverse modeling approach used in this thesis. In this chapter, the big picture of the work is presented explaining the methodology, tools, and models that are used in this research. In this chapter, the different observation networks which are used in this thesis are introduced. Satellite data and their characteristics and related details are presented. Brief information about ground-based networks, their coverage in modeling domain, and their sampling methods are presented.

Chapter four addresses the effectiveness of emission inversion in reducing the discrepancy between model output and observations. Results of four different emission inversion scenarios are presented in this chapter. In this chapter, the behavior of updated emissions using ground based observations, satellite observations and combining the two sets of observations together on

a daily basis was studied. Last, the change in the updated emissions when the assimilation window increases from one day to the entire month was investigated. The results of time independent emission inversion are also described.

Chapter five shows the results of inverse modeling combining satellite observations and ground-based observations for time dependent scenarios. Updated emissions are treated as independent hourly variables at each grid and location, the performance of inversion and the variance of the scaling factors are investigated in this chapter.

Chapter six explores the applicability of variational inverse modeling for constraining the emissions of NO from a challenging emission source, the lightning, using satellite observations. Lightning can create a significant change in NO_x concentrations, in a region of the atmosphere that is less affected by anthropogenic sources, especially during summer. Despite its significance, lightning has traditionally neglected by air quality models, and only recently its impact on surface air quality has received attention. The main objective of this chapter is to improve the emission inventory by considering available lightning sources.

In chapter seven, a summary of findings of this thesis is provided followed by concluding remarks on how this thesis has contributed to various approaches that have provided top-down constraints on emission rates. This chapter is concluded by reviewing the limitations of the current work and with suggestions for potential future research.

CHAPTER TWO:

Background

Inverse modeling is widely used in different applications and different disciplines for model improvement and reducing uncertainties. In this chapter the main tools and methods used in this research are introduced, the current state of the scientific literature is discussed, and specifically the background and usage of these methods in atmospheric applications are reviewed.

2.1 Inverse modeling

Air quality modeling plays an important role in environmental decision making. Air quality forecasting has been improved with advancement in CTMs, which are developed to depict the fate and transport of the atmospheric chemical constituents. Despite significant improvements in our understanding of the atmosphere and different atmospheric processes, it is still a challenge to accurately describe the evolution of pollutant concentrations in time and space. This is mainly because of the uncertainties associated with different parts of CTMs including: incomplete and inaccurate emission information, initial or boundary conditions, and insufficiently parameterized processes in the atmosphere. Taking advantage of real-time observations, one can perform

inverse modeling and estimate the parameters whose values are associated with large uncertainties. Doing so, using the estimated parameters as model inputs would lead to better prediction of the future state of the atmosphere through the assimilation and fusion of observations into the model results.

One of the main advantages of inverse modeling techniques is the combination of models and measurements to draw a complete picture of the state of the atmosphere. Inverse modeling in atmospheric studies draws upon mathematical approaches of different levels of sophistication to produce more accurate predictions of the current state of the atmosphere. This approach can then be used to, ideally, reduce model errors and find the best estimate of the future state of the atmospheric system. Inverse modeling methods are generally derived from data assimilation techniques and therefore they resemble the formalism of those techniques. In general, the difference of terminology between inverse modeling and data assimilation is mostly semantic, and mostly revolves around the definition of the state and inverted parameters. The state parameters of a model are those parameters that are available as model outputs, e.g., meteorological conditions for weather prediction models or pollutants concentrations for AQMs. However, for models that integrate the evolution of the state in time, initial state of the model (i.e., initial conditions) are also input parameters, and can be the subject of mathematical inversion. In this context, the objective of data assimilation is often to find a more accurate (initial) state of the system while in inverse modeling applications the goal is to invert (i.e., find better values for) input parameters other than the state, and use them to improve the prediction of the future state of the system. In other words, while both methods aim to reduce the distance between model output and available observations, inverse modeling of input parameters leads to

adjustment of wider range of inputs and consequently results in a better understanding of the state of the atmosphere.

In the general sense of inverse modeling, the relationship between the model parameters and observations can be presented by Eq. 2-1.

$$\mathbf{y} = \mathbf{F}(\mathbf{x}, \mathbf{b}) + \varepsilon \quad (2-1)$$

Here, $\mathbf{y}$ is a vector of available observations, $\mathbf{F}$ is the (generally nonlinear) forward model, $\mathbf{x}$ is the model parameter that is being optimized, $\mathbf{b}$ represents the vector of all model parameters that are not affected during optimization and ε represents the error associated with both model and observations. $\mathbf{F}(\mathbf{x}, \mathbf{b})$ is used to link the model parameters to the observation vector. This can be expressed by a linear/nonlinear operator that maps the output from model space to the observation space. Hereafter, projection of $\mathbf{F}(\mathbf{x}, \mathbf{b})$ onto the observation domain is shown as $\mathbf{H}(\mathbf{F}(\mathbf{x}))$, or by combining the two operators, simply as $\mathbf{H}(\mathbf{x})$.

Inverse modeling tries to solve an ill-posed problem; therefore it is essential to have a first guess, which in atmospheric studies would be the background state of the model. The general method of solving an inverse problem can be addressed by the Bayesian approach as this method allows one to include the model parameters and the statistics of both parameter and observation errors. In the case of Eq. (2-1), the Bayesian approach can link the relationship between input parameter ($\mathbf{x}$) and the error in observation ($\mathbf{y} - \mathbf{H}(\mathbf{x})$).

There are different methods to solve the inverse problem such as least square methods, Monte Carlo methods, Bayesian methods, etc. The inverse problem can be formulated as an optimization problem with a scalar valued objective function $J(\mathbf{X})$ which is also known as a cost function, penalty function or fitness function.

Applying the Bayesian theorem, to solve the inverse problem there are two vectors as defined earlier: $\mathbf{X}$ (model parameter or input parameter) and $\mathbf{Y}$ (observation vector). The probability of these vectors can be shown as:

$$p(\mathbf{X}|\mathbf{Y}) = \frac{p(\mathbf{Y}|\mathbf{X})p(\mathbf{X})}{p(\mathbf{Y})} \quad (2 - 2)$$

In the above equation, $p(\mathbf{X})$ is the probability distribution function of the state vector (initial or a-priori or first guess, without any information regarding observations, only based on previous model simulations or other sources of data), $p(\mathbf{Y})$ is the probability distribution function of the observations vector which is not dependent on the model parameter, $p(\mathbf{Y}|\mathbf{X})$ is the conditional probability distribution function of the error at a given model parameter (which is provided by the instrument that recorded the observation) and $p(\mathbf{X}|\mathbf{Y})$ is the probability distribution function of the model parameter in the presence of observations. $p(\mathbf{Y})$ is the first term that will be ignored as it is not related to either model or input parameters.

To provide a Bayesian solution to an inverse problem, the probability distribution functions (pdfs) of the vectors and the error associated with them are essential. These pdfs can be chosen according to different types of distributions. In general, the pdfs can be considered to be Gaussian. In that case, considering vector $\mathbf{x}$ with prior information of $\mathbf{x}_b$ and the error covariance matrices of $\mathbf{B}$ and $\mathbf{R}$ as background and observational error covariance matrix, respectively, the pdfs can be written as follows:

$$p(\mathbf{X}) = \frac{e^{-\frac{1}{2}(\mathbf{x}-\mathbf{x}_b)\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b)^T}}{\sqrt{(2\pi)^n|\mathbf{B}|}} \quad (2 - 3)$$

$$p(\mathbf{Y}|\mathbf{X}) = \frac{e^{-\frac{1}{2}(\mathbf{y}-\mathbf{H}(\mathbf{X}))\mathbf{R}^{-1}(\mathbf{y}-\mathbf{H}(\mathbf{X}))^T}}{\sqrt{(2\pi)^m|\mathbf{R}|}} \quad (2-4)$$

where n and m are the dimensions of vectors $\mathbf{x}$ and $\mathbf{y}$. (Jacob 2007)

Combining Eq. (2-3) and Eq. (2-4) and use them in Eq. (2-2)

$$p(\mathbf{X}|\mathbf{Y}) = \frac{e^{-\frac{1}{2}(\mathbf{y}-\mathbf{H}(\mathbf{X}))\mathbf{R}^{-1}(\mathbf{y}-\mathbf{H}(\mathbf{X}))^T}}{\sqrt{(2\pi)^m|\mathbf{R}|}} \frac{e^{-\frac{1}{2}(\mathbf{x}-\mathbf{x}_b)\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b)^T}}{\sqrt{(2\pi)^n|\mathbf{B}|}} \quad (2-5)$$

There are different methods to solve the Eq. (2-5). One of these methods is Maximum a-posteriori (MAP) method which finds a value for $\mathbf{x}$ that maximizes the $p(\mathbf{X}|\mathbf{Y})$. Eq. (2-5) can be expressed as follows:

$$\begin{aligned} -\ln p(\mathbf{X}|\mathbf{Y}) &= \frac{1}{2}(\mathbf{y} - \mathbf{H}(\mathbf{X}))\mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}(\mathbf{X}))^T + \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b)^T \\ &\quad + \frac{1}{2}\ln((2\pi)^{m+n}|\mathbf{R}| |\mathbf{B}|) \end{aligned} \quad (2-6)$$

To find a value that maximizes the $p(\mathbf{X}|\mathbf{Y})$ means to minimize the right hand side of the Eq. (2-6) and that is the main purpose of inverse modeling. The right hand side of the above equation is called the cost function which is denoted by $J(\mathbf{x})$. Please note that the last part of the right hand side of the above equation is independent of $\mathbf{x}$ and can thus be ignored in the optimization process.

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{y} - \mathbf{H}(\mathbf{X}))\mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}(\mathbf{X}))^T + \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b)^T \quad (2-7)$$

The solution from minimizing Eq. (2-7) is equivalent to the solution of the maximum likelihood method (Bouttier et al., 1999).

$$\begin{aligned}
\nabla J(\mathbf{x}_a) &= -\mathbf{H}^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}(\mathbf{x}_a)) + \mathbf{B}^{-1}(\mathbf{x}_a - \mathbf{x}_b) \\
&= -\mathbf{H}^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}(\mathbf{x}_b + \mathbf{x}_a - \mathbf{x}_b)) + \mathbf{B}^{-1}(\mathbf{x}_a - \mathbf{x}_b) \\
&\approx -\mathbf{H}^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}(\mathbf{x}_b) + \mathbf{H}(\mathbf{x}_a - \mathbf{x}_b)) + \mathbf{B}^{-1}(\mathbf{x}_a - \mathbf{x}_b) \\
&= -\mathbf{H}^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}(\mathbf{x}_b)) - \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}(\mathbf{x}_a - \mathbf{x}_b) + \mathbf{B}^{-1}(\mathbf{x}_a - \mathbf{x}_b) = 0 \quad (2-8)
\end{aligned}$$

In the third line in Eq. (2-8) $\mathbf{H}$ represents the approximation of the second term as a Taylor series truncated after the linear term. In fact $\mathbf{H}$ is the tangent linear forward model operator with respect to the model state vector and evaluated at the model background state.

$$\mathbf{H} = \left. \frac{d\mathbf{H}}{d\mathbf{x}} \right|_{\mathbf{x}_b} \quad (2-9)$$

This operator helps to solve the Eq. (2-8) and find the analysis as follows (Bouttier et al., 1999):

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{B} \mathbf{H}^T (\mathbf{H} \mathbf{B} \mathbf{H}^T + \mathbf{R})^{-1} (\mathbf{y} - \mathbf{H}(\mathbf{x}_b)) \quad (2-10)$$

$$\mathbf{K} = \mathbf{B} \mathbf{H}^T (\mathbf{H} \mathbf{B} \mathbf{H}^T + \mathbf{R})^{-1} \quad (2-11)$$

$$\mathbf{A} = \mathbf{B} - \mathbf{B} \mathbf{H}^T (\mathbf{H} \mathbf{B} \mathbf{H}^T + \mathbf{R})^{-1} \mathbf{H} \mathbf{B} \quad (2-12)$$

The inverse modeling framework in atmospheric applications needs three main tools:

- Observations
- Chemical transport model
- Initial or a-priori knowledge of the state of the atmosphere

As mentioned earlier, inverse modeling tries to estimate the best future state of the atmosphere, which is called “forecast”, or “analysis”. This forecast state can be estimated by minimizing the

distance between observations and the “current” state of the atmosphere. Both observations and the CTM are associated by uncertainties. The forecast is obtained from minimizing a weighted average of the model and observation errors. If model errors are larger than the observation errors, the analysis would be closer to observations, and if model errors are smaller than the observations, the analysis would be closer to the background state (Bouttier et al., 1999).

Uncertainties in models are mostly due to estimations in the input parameters (emissions, meteorology conditions, initial and boundary conditions, etc.), mathematical equations dictating the relationships between parameters in the model, and methods for solving the governing differential equations. Observational inaccuracy can stem from instrument errors, sampling errors, representativeness error (i.e., in situ data vs. volume-averaged model values), or errors in processing the measured observations (Bouttier et al., 1999).

From the mid 1960s meteorologists have been developing and using data assimilation methods to estimate the three dimensional state of the global atmosphere and to predict weather conditions for a long term period (Bouttier et al., 1999). In other fields such as oceanography, geophysics, and atmospheric studies, scientists have used inverse modeling techniques in order to gain more knowledge about the future state of the dynamical systems. For example, more details about using inversion techniques in ocean models can be found in Anderson et al. (1996).

There are three main methods for inversion; optimal interpolation, variational, and sequential methods.

2.1.1 Optimal interpolation

Optimal interpolation is the earliest method for data assimilation (Denby et al., 2008). The advantage of using this method is in its ability to fill in the gaps in sparse datasets. This method

is simple and computationally inexpensive for a limited set of observations. The missing data can be calculated from a linear combination of available data and observations.

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}(\mathbf{y} - \mathbf{H}(\mathbf{x}_b)) \quad (2-13)$$

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} \quad (2-14)$$

Where $\mathbf{x}_a$ is forecast, $\mathbf{x}_b$ is the background or current state, $\mathbf{H}$ is the observation operator matrix, $\mathbf{y}$ is the observations vector, $\mathbf{K}$ is the linear operator or weight matrix, $\mathbf{B}$ is the background error covariance and $\mathbf{R}$ is the observation error covariance.

The earliest application of optimal interpolation was used for the analysis of meteorological fields (Gandin, 1963). Lorenc (1981) applied this method as a statistical multivariate interpolation scheme for wind, height, and atmospheric layer thickness prediction on a global scale (Lorenc, 1981). Flemming et al. (2001) employed an optimal interpolation method for assessing air quality prediction for continuous predictions by using conventional network information (Flemming et al., 2001).

2.1.2 Sequential methods

Sequential methods are based on a stochastic approach for inverse problems and data assimilation applications. These methods use a probabilistic framework in order to give an estimate of the whole system state (data assimilation) or input parameters (inverse modeling) by propagating information into the system sequentially. The earliest filtering method used for parameter estimation and system control theory was the Kalman filter (Kalman, 1960) technique (Hartly et al., 1993, Welch et al., 2001). The Kalman filter has been used in a wide range of applications such as, tracking objects (such as missiles), economics, navigations, hydrology,

oceanography and atmospheric studies. The Kalman filter estimates the state of the system in the presence of observations.

To demonstrate the formulation for the Kalman filter let's assume the following stochastic-dynamic system (Polavarapu, 2004):

$$\mathbf{x}_{k+1} = \boldsymbol{\varphi}_k \mathbf{x}_k + \mathbf{w}_k \quad (2 - 15)$$

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k \quad (2 - 16)$$

The state of the model at time t_k is $\mathbf{x}_k$, $\boldsymbol{\varphi}_k$ is the transition matrix for the linear (or linearized) model and $\mathbf{w}_k$ is the model error. $\mathbf{y}_k$ is the observed state with the dimension of m-vector. $\mathbf{H}_k$ is the observation operator which maps the state of the model to the observed state, with a dimension of $m \times n$. $\mathbf{v}_k$ is the observation error. If $\mathbf{Q}$ is the error covariance of the model and $\mathbf{R}$ is the error covariance of the observations, then the forecast and its error covariance ($\mathbf{P}$) can be calculated as:

$$\hat{\mathbf{X}}_{k+1}^f = \boldsymbol{\varphi}_k \hat{\mathbf{X}}_k^a \quad (2 - 17)$$

$$\mathbf{P}_{k+1}^f = \boldsymbol{\varphi}_k \mathbf{P}_k^a \boldsymbol{\varphi}_k^T + \mathbf{Q}_k \quad (2 - 18)$$

The state estimate and error covariance are then propagated to the next time step and the analysis will be:

$$\hat{\mathbf{X}}_k^a = \hat{\mathbf{X}}_k^f + \mathbf{K}_k (\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{X}}_k^f) \quad (2 - 19)$$

$$\mathbf{P}_k^a = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^f \quad (2 - 20)$$

where $\mathbf{K}_k$ is the weight or gain matrix, and is a measure of the weight that is given to the observations.

Evenson (1990) implemented an inverse theory and sequential method based on Monte Carlo simulations on non-linear ocean models (Evenson, 1990). Kalman filter was designed for linear models but in the case of very large systems and systems with the nonlinearity Kalman filter was not efficient. Evenson (1992) developed the Extended Kalman filter (EKF) in order to use filtering techniques in the oceanography field. This version of Kalman filter is efficient for weak nonlinear systems. To deal with large systems, the suboptimal Kalman filters were developed under which the dimension of the system was still a problem. To deal with highly non-linear with large error covariance matrix, Evenson (1994) developed an intriguing Kalman filter approximation which is called Ensemble Kalman Filter (EnKF).

Loon et al. (2000) studied the effect of uncertainties in NO_x , SO_x , VOC and CO on ozone concentrations by using EnKF (Loon et al., 2000). Filtering methods can estimate ozone concentrations better when the number of runs of the chemical transport model are increased (Hamil et al., 2001). Elbern (1992) developed an Ensemble Kalman filter for short-term forecasting of tropospheric ozone (Elbern et al., 2005). Hanea (2004) reviewed the Kalman filter performance on air quality applications (Hanea et al., 2004). Chemical tracer forecast uncertainties associated with meteorological initial conditions and the effect of observation error on the forecast is studied by Stuart using ensemble-Kalman filter (Stuart et al., 2007). Napelenok et al. (2008) applied a filtering approach (Discrete Kalman Filter) to the Community Multiscale Air Quality (CMAQ) model on the southeastern United States domain to estimate a top-down NO_x emission inventory using Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMCHY) satellite and ground level NO_2 observations. Errera et al. (2008)

compared the performance of using 4D-Var and EnKF method for data assimilation of ozone and nitrogen dioxide stratospheric concentrations using satellite data. They reported that the result from simulations showed a good agreement between variational and sequential approaches. The same behavior was reported in another study (Skachko et al., 2014). Wei et al. (2013) constrained NO_x emissions using Ozone Monitoring Instrument (OMI), applying direct scaling factors and the Discrete Kalman Filter approach on the same modeling domain.

Sequential methods are useful for applications with large number of real time observations which need good state estimations. Under these methods, the trajectory of the model is updated in the presence of available observations and the occasional jumps in the trajectory of the model have little importance. However, these methods become less efficient as the number of inverted parameters grows substantially. As the main purpose in this research is to update the emissions at all grid cells simultaneously variational approach was used.

2.1.3 Variational methods

Variational methods in inverse modeling are based on iterative minimization of a scalar value. This scalar value, which is called cost function, is defined by the disparity between the measured observation and the output of the CTM. Effective filtering methods such as EnKF are relatively new in comparison with variational methods that have been used in various applications and are known for their reliable results. Variational methods are employed in different assimilation applications in the context of an optimal control approach. Generally, for short time intervals, filtering methods are expected to provide more accurate results in comparison to variational methods, but for infrequent observations and non-linear behavior of error, variational methods

are more accurate (Kalnay et al., 2007). Figure 2-1, taken from Khattatov et al. (1999), graphically illustrates the difference between variational and sequential methods.

Figure 2-1. Schematic procedures of variational (left) and sequential (right) inversion methods (Khattatov et al., 1999)

Variational (4D-var) and sequential (extended Kalman filter) methods of data assimilation were implemented for photochemical active species in order to reduce the distance between model output and observations (Khattatov et al., 1999). As shown in Figure 2-1, in the variational method, initial conditions are adjusted to minimize the disparity between observation and model output. In variational methods, the simulated state variable is continuous and is pushed closer to the observations, while in the sequential methods, the simulated state variable is discontinuous and is shifted closer to the observed values intermittently and in the presence of observation (i.e., at discrete observation windows). Both methods bring the state variable closer to the observation; the difference is in continuity and physical consistency and the way the errors propagate. Kalman filters update the model at the presence of observation and go to the next observation and will carry the error information through the inversion process. 4D-Var method deals with errors all at

once and does not provide as much details on forecast error, but it is continuous and preserves the physical consistency of the inverted system.

Khattatov (1999) built a photochemical box model for data assimilation of O_3 , NO_2 , HNO_3 , $ClONO_2$, CH_4 and N_2O by using the Upper Atmosphere Research Satellite (UARS) data in the stratosphere. They reported that the extended Kalman filter was more advantageous when observation errors are small and its performance is better for prolonged periods of time. On the other hand, the variational method had better performance in the case of larger observation errors over a moderately long, finite period of time; they suggested a 48 hour analysis for assimilation window (Khattatov et al., 1999).

There are two main variational data assimilation methods: 3D variational data assimilation (3D-Var) and 4D variational data assimilation (4D-Var). 4D-Var is an extension of 3D-Var which considers evolution of the system in time.

In variational methods, the assimilated model state can be derived from minimization of the following cost function (Niu et al., 2007):

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}(\mathbf{y}^o - \mathbf{H}(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y}^o - \mathbf{H}(\mathbf{x}))$$

$$= J_b + J_o \quad (2-21)$$

where J is called the cost function of the analysis or the penalty function, J_b contains the background term and J_o is the observation term. Assuming that there are n variables in the model state with p number of observations, parameters in Eq. (2-21) can be defined as follows:

$\mathbf{x}_t$ is true model state (dimension n), $\mathbf{x}_b$ is the background model state (dimension n), $\mathbf{x}_a$ is the posterior model state, the optimal and close to true state is $\mathbf{x}_t$ (dimension n), $\mathbf{y}$ is the vector of observations (dimension p), $\mathbf{H}$ is the observation operator (from n to p). In cases where the space of the observation and model are not the same, using the observation operator is essential. It is quite straightforward to construct the cost function, if both observation and output of the model are concentrations or any other value from the same space but if model output and observation are of different kinds (e.g., concentrations vs. satellite vertical column density, as is the case in this research), it is critical to properly map the outputs to states comparable with observations. $\mathbf{B}$ is the covariance matrix of the background errors (dimension $n \times n$) and $\mathbf{R}$ is the covariance matrix of observation errors ($\mathbf{y} - \mathbf{H}[\mathbf{x}_t]$) (dimension $p \times p$).

This form of variational method is called 3D-Var and is widely used in different meteorological applications (Chouinard et al., 2001, Gauthier et al., 1999). In order to minimize the cost function (the distance between observations and modeled state/concentrations), the gradient of the function $\mathbf{J}$ at the analysis state should be equal to zero, therefore (Bouttier et al., 1999):

$$\nabla \mathbf{J}(\mathbf{x}_a) = 0 \quad (2-22)$$

The error terms in this formulation include:

- Background error (model error): $\boldsymbol{\varepsilon}_b = \mathbf{x}_b - \mathbf{x}_t$ with covariance $\mathbf{B}$; this part includes the estimation error of the background state with the true state of the model.
- Observation errors: $\boldsymbol{\varepsilon}_o = \mathbf{y}^0 - \mathbf{H}(\mathbf{x}_t)$ covariance $\mathbf{R}$; this part includes observational errors in the operator in comparison with the true state of the model.

A-priori or background ($\mathbf{x}_b$) represents the best estimated state of the atmosphere before using any observations; meanwhile the goal of using the 3D-Var assimilation technique is to change this prior state of the atmosphere and make it more consistent with observations ($\mathbf{y}^0$). The assimilated state, which is denoted as $\mathbf{x}_a$, will become closer to the true state ($\mathbf{x}_t$) as the result of inversion (Bouttier et al., 1999). The 3D-Var data assimilation method has been successfully used in assimilation applications for model analysis and prediction of aerosols over global oceans (Zhang et al., 2008).

The first 4D-Var data assimilation application for chemical species was implemented by Fisher et al. (1995). They used an explicit model for calculating the adjoint of the system's tangent linear model. They used the Upper Atmosphere Research Satellite (UARS) observation data and employed the limited memory Quasi-Newton scheme M1QN3 method for optimization. Adjoint equations are essential for the calculation of the gradient of the cost function. Derivation of these equations is possible in different ways (see Talagrand et al., 1987; Daley, 1991 for more details).

4D-Var is the extension of three dimensional variational methods that considers the time in its assimilation procedure. The first part is the same as that of 3D-Var Eq. (2-21), but the observation operator should be developed to take the time steps for model states and observations into consideration as below:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2} \sum_{i=0}^n (\mathbf{y}_i - \mathbf{H}_i[\mathbf{x}_i])^T \mathbf{R}_i^{-1}(\mathbf{y}_i - \mathbf{H}_i[\mathbf{x}_i]) \quad (2 - 23)$$

Elbern (1997) used 4D-Var data assimilation for three dimensional atmospheric chemistry models for tropospheric data assimilation (Elbern et al., 1997). They studied the potential of 4D-Var data assimilation for optimization of emission rates of precursors of ozone using ozone

observations. To evaluate the performance of the assimilation process, they used three steps. The first step was preparing the true state of NO emissions from the reference run. The next step was perturbing NO emission rates to make the first guess for concentrations. The third step was running the chemistry model (EURAD) in order to bridge the gap between the first guess and the true state of NO emission rates using ozone observations. Their results showed that if in the first guess NO emission rates are not too far from the true state, the performance of assimilation procedure is fairly good (Elbern et al., 2000).

Sandu et al. presented an overview of the mathematical aspects of the adjoint method in 3D air quality models (Sandu et al., 2005). They discussed the continuous and discrete methods for the development of the STEM adjoint model as an essential tool for data assimilation system. The discrete adjoint approach was employed for data assimilation on ozone concentration as a case study in the East Asia domain (Sandu et al., 2005).

Hakami (2005) used the STEM CTM and its adjoint for inverse modeling of the black carbon emissions (anthropogenic and biomass burning) during the Asian Pacific Regional Aerosol Characterization Experiment (ACE-ASIA). They used an inverse modeling method in the presence of reliable data and wide temporal and spatial coverage of measurements for recovering the emissions (Hakami et al., 2005).

Chai et al. (2007) used the STEM CTM for 4D-Var data assimilation using ICARTT ozone measurements. They used the observational platforms of ozone (aircraft, surface, and ozone sondes) separately and all together for assimilation into the STEM model. Assimilation results showed that adding all the different platforms of observations together will improve the model predictions; they suggested combining satellite data will further improve the model performance.

As initial ozone concentrations were the only control variable in their case study, they recommended that in order to get better forecast, adjusting initial concentration of other species could be helpful.

Data assimilation can be used for optimizing the initial condition and emissions to have better results. Zhang et al. (2008) combined ground-based data sets (AIRNow) of ozone with NO₂ satellite retrievals (SCHIAMACHY) for emission inversion (Zhang et al., 2008). The first part of their experiment was assimilating ground-based ozone concentration in a 24 hour long (4 am, July 1 to 4 am July 2) assimilation window for improvement in model predictions. To combine both ground-based and satellite data in the 4D-Var data assimilation of ground level ozone, they first found an optimal solution by using only ground level observations with the assimilation window of a 24 hour period (4 am July 16 to 4 am July 17) and then added the SCHIAMACHY info (NO₂ and HCHO) into the assimilation process. The correlation coefficient between ozone predictions and ground level ozone observations increased from ($R^2=0.37$) to ($R^2=0.69$). Also, there was an increase in the correlation coefficient between what the model predicted for NO₂ and the SCHIAMACHY NO₂ observation from ($R^2=0.19$) to ($R^2=0.36$). Unlike ozone and NO₂ predictions, there was a decrease in the correlation coefficient for HCHO from ($R^2=0.19$) to ($R^2=0.11$). Moreover, they used both datasets for emission inversion of NO₂ and found that when only satellite data was used, the results were consistent for short (8 hr) and long (32 hr) assimilation windows.

The potential and limitations of the 4D-Var inversions were assessed for estimating the ozone precursors, i.e., sulphur oxides and nitrogen oxides by Elbern et al. (2007). Their work showed that the emission inventory over Europe overestimates sulfur emissions, but the estimation is within an acceptable range. They mentioned that using the CTM with a finer grid size may lead

to better predictions, when compared to the coarse horizontal resolution that they had in their study (54 Km). 4D-Var is used for other species such as methane to reduce uncertainties in posterior emissions of methane (Meirink et al., 2008). A 4D-Var data assimilation method was developed for emission inversion of nitrogen dioxides (Chai et al., 2009). They adjusted NO_x emissions by using conventional ozone observations and NO_2 column densities of SCIAMACHY. They reported that changing the assimilation window from 24 to 48 hours results in identical outcomes. They also examined day-by-day assimilation for 2 months and found that adding ozone ground-based observation does not lead to better results due to uncertainties in VOCs. Lin et al (2012) employed GEOS-CHEM model combining with OMI observations to recover NO_x emission sources over China. They proposed a method to improve the model results from analyzing the sensitivity of modeled NO_2 columns to errors in meteorological and chemical parameters important to the nitrogen abundance. Their improvement led to 22% and 10% increases in the vertical column densities in July and January. They reported the decrease of 15 ppb on surface ozone as a result of inversion in July. In these studies the main limitation was that emissions were adjusted by multiplication of coefficient per day for all hours during the day and the improvement of the results were reported in a short period of time without addressing the behavior of the emissions. Resler et al (2010) extended this method to apply a diurnal time profile of coefficients on emissions using CMAQ model and combination of GOME-2, OMI and IASI observations and reported significant improvement on cost function reduction, bias and model error. Improving the emission sources has a direct effect on model calculation. Huang et al (2015) suggested a multi-scale satellite inversion to improve the ozone concentrations of Western U.S. (California-Nevada domain). As a part of the study, they estimated the NO_x

emission sources through 4D-Var emission inversion using OMI observations. They used the estimated scaling factors for both NO_x and NMVOCs as they are important precursors of ozone.

In this work the emission inversion took place for a relatively long period of time (one month) and the focus was the behavior of emissions in two general scenarios; the change in the emissions on daily and hourly basis.

2.1.3.1 Covariance modeling

An important part of inverse modeling is to construct proper error covariance matrices. Error covariance for both observations and background play an important role in variational data assimilation. More detailed understanding of a state of the art of 4D-Var system can be obtained in the presence of sufficient information about all sources (observations and background) with the associated errors. The accuracy of analysis or “forecast” is dependent on the accuracy of the observations and background. To have better predictions, adequate information about the covariance matrix of the background errors (**B**) and covariance matrix of observation errors (**R**) is essential.

- **Observation error covariance**

The observation error covariance matrix contains information about the vector of observations (**y**). Observation errors can be related to instrumental errors and errors that come from the observation operator (**H**). The matrix of observation error (**R**) are often considered as a diagonal matrix with one unique error variance (Chai et al., 2007, Zhang

et al., 2008). In this case, the main assumption is that the errors in observations are not temporally and spatially correlated.

- Background error covariance

Defining proper background error covariance matrix (**B**) is one of the main challenges in the variational data assimilation applications. Background error provides information about accuracy of both parameter errors and model errors. Background errors are especially important for inverse modeling where the main purpose is to obtain more information about model and input parameters. To construct a background error covariance matrix without having enough information about the errors is not practically possible. On the other hand, sufficient knowledge, if available, results in a large background covariance matrix which would be computationally expensive to process or invert. One of the first studies for constructing the background covariance matrix was conducted by Hollingsworth (Hollingsworth et al., 1986) who used an observational method for errors due to short term forecasts. In this method, background covariance matrix can be generated based on the difference between observations and model outputs. Another approach, which has been applied to chemical data assimilation, is the National Meteorological Center (NMC) method (Chai et al., 2006). In this method, the error covariance matrix is approximated from the differences between different forecasts at a specific time. Constantinescu et al. (2007) constructed a background error matrix from the autoregressive processes. However, NMC is only applicable to data assimilation applications, and cannot be used for inverse modeling where non-state parameters are inverted.

Elbern et al. (2007) used a diffusion operator method to construct the background error covariance matrix for an emission inversion application. The purpose of their work was to estimate precursors of ozone, sulphur dioxide and nitrogen oxides. Assuming the simulation domain is defined by (N_x, N_y, N_z) grid points, where x, y, z are indices for directions for the total number of chemical species of N_s , the dimension of the model state vector will be:

$$N = N_x \times N_y \times N_z \times N_s \quad (2-24)$$

The very large size of the background error covariance matrix makes it computationally expensive to construct. Another alternative would be to construct the same size error covariance matrix but when it comes to calculation of the inverse of that matrix, it will become difficult and computationally very expensive. These limitations are the main reasons why in most studies, the background error covariance matrix is assumed to be diagonal. Singh et al. (2011) proposed an effective method for constructing non-diagonal background covariance matrices. They suggested an exponential decay function to define the correlations between model grid cells. In the present study, both diagonal and non-diagonal error covariance matrices were used. In chapter four, error covariance matrices were assumed to be diagonal and in chapter five a non-diagonal covariance matrix is constructed based on the proposed method by Singh et al. (2011).

2.1.3.2 Optimization

As mentioned before, in 4D-Var application, gradients of the cost function with respect to the species of concern are required for the minimization of the cost function. The calculation of gradients often requires sensitivity analyses. There are a few optimization routines that can be

used in data assimilation and inverse modeling application. They differ based on the physics of the problem and sometimes the dimension of the unknowns. For example, direct-search methods are very efficient when it comes to a problem with merely a few parameters. When the problem has large dimensions/parameters, gradient-based methods are more efficient. In gradient-based methods the cost function can be minimized in the presence of the gradient and the Hessian (the second derivative) of the cost function.

To minimize the cost function in a 4D-Var inverse modeling framework, a commonly used optimization method (i.e., the L-BFGS method) is employed (Nocedal, 1980, Liu et al., 1989). The Limited memory of Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) method has been widely used for 3D and 4D-Var data assimilation in atmospheric applications. The BFGS method is an algorithm that can easily solve a large number of nonlinear optimization problems. Here the bound constrained extension of LBFGS which is called L-BFGS-B (Byrd et al., 1994) is used.

This limited memory method, which is a sub-category of Quasi-Newton methods, approximates the Hessian matrix of the objective function (second order partial derivatives of a function) instead of calculating the actual matrix which is computationally expensive. Briefly, in this method at different iterations, the quadratic form of the function can be estimated with a combination of a function values and a gradient of the function (Liu et al., 1989). The next iterate is estimated by the line search direction. Alternatively to the Newton line search which requires the full Hessian matrix, in this method the steepest descent algorithm is used to find the next iterate. The steepest descent algorithm only requires the approximate of the Hessian matrix therefore makes L-BFGS-B a suitable method for emission inversion applications with a large number of parameters. However, lack of full information about the Hessian matrix has a disadvantage that makes the L-BFGS-B method slow on convergence. This method does not

need any information about the structure of the function, and it is easy to implement with low iteration cost. Figure 2-2 shows an example of steepest-descent algorithm.

Figure 2-2. Steepest – descent schematic illustration (Jacob 2007)

2.2 Sensitivity Analysis

Sensitivity analysis is a powerful tool that enables modelers to estimate the rate of the change in the output of the model with respect to changes in model inputs. Sensitivity analysis can also be employed for studying the uncertainties of a model, as it helps modelers to determine the level of

accuracy required for a parameter in order to build a reliable model. Sensitivity analysis also can identify the parameters affecting model performance the most (Frey et al., 2000).

Sensitivity analysis is used in AQMs in order to fulfill two main purposes in modeling and uncertainty analysis:

- Source-receptor relationships
- Understanding the behavior of the model with regards to its parameters

Sensitivity analysis methods are used in order to achieve more understanding of models. In other words, sensitivity analysis, as a tool, is used to characterize the response of atmosphere. Generally, in AQMs, sensitivity analysis is categorized into forward and backward methods (Hakami et al., 2007) and both groups of methods can be used in inverse modeling applications.

2.2.1 Forward sensitivity analysis

In forward sensitivity analysis methods, a small perturbation in an input parameter is propagated forward (through space and time) into the model output. These methods are efficient when the sensitivities of all outputs (i.e., concentrations at all receptors) are needed with respect to perturbations in a few input parameters. The definition of a sensitivity coefficient, in this context, is the first-order derivative of the model output with respect to the specific model input(s). Sensitivity coefficients can be calculated by finite difference estimation or directly by differentiating the governing equations for the model.

Sensitivity coefficients are traditionally calculated by finite difference approximations based on the difference between model outputs with the perturbed input(s) and unperturbed input(s). This method, which is called the brute-force method, is easy to implement. In this method, one needs to run the model multiple times. After the base case run with no perturbation in model inputs,

one input or model parameter is perturbed, and then the sensitivities are calculated based on the differences in outputs of the two runs. This method becomes computationally expensive when the model is complex and the number of input parameters, whose impacts on model outputs need to be calculated, increases.

$$s_{ij} \cong \Delta C_i / \Delta p_j \quad (2-25)$$

Under the brute-force method, to calculate the derivative of model outputs with respect to M input parameters, one needs to run the model $M+1$ times. In this case, calculating a large number of coefficients in complex and nonlinear models with large perturbations becomes very inefficient and expensive (Dunker, 1980). An example of applications with brute force sensitivity is reported in research by Jacob (Jacob et al., 1999).

An alternative to the brute-force method is the decoupled direct method (DDM), also called the tangent linear model (TLM), which calculates the sensitivity coefficients directly. In fact, in this method, sensitivity equations are derived from the model equations and solved separately. In this source-based method, a perturbation of a single input or model parameter is propagated forward in time and will result in a sensitivity field of model outputs. The main advantage of this method over the brute-force method is that different perturbations in model inputs can be integrated simultaneously (Yang et al., 2007), and the calculated sensitivity coefficients are less prone to numerical noise and roundoff errors resulting from subtracting two close numbers.

Yang et al. (2007) developed a variant of the DDM as DDM-3D method for sensitivity analysis for the California Institute of Technology (CIT) airshed model. The formation and transport of

atmospheric components is described by atmospheric diffusion equation (ADE) (Yang et al., 2007):

$$\frac{\partial c_i}{\partial t} = -\nabla(\mathbf{u}c_i) + \nabla(\mathbf{K}\nabla c_i) + R_i[c_1, c_2, \dots, c_n; T, t] + E_i \quad i = 1, \dots, N \quad (2-26)$$

where c_i is the concentration of species i , $\mathbf{u}$ is the wind field, $\mathbf{K}$ is the turbulent diffusivity tensor, R_i is the net rate of chemical productions and E_i is the emission source rate of emissions of species i .

For a model parameter p_j , sensitivity coefficients can be calculated as the partial derivative:

$$s_{ij}(t) = \frac{\partial c_i(t)}{\partial p_j} \quad (2-27)$$

To make the equation useful for all parameters with different ranges in time and space, one can use semi-normalized sensitivity coefficients. For this purpose, every parameter variation is defined by:

$$p_j(x, t) = \epsilon_j P_j(x, t) \quad (2-28)$$

where

$P_j(x, t)$; unperturbed field of parameter

ϵ_j ; scaling variable with nominal value of 1

Semi-normalized sensitivity coefficients then become:

$$s_{ij}^*(t) = P_j \cdot \frac{\partial c_i(t)}{\partial p_j} = P_j \cdot \frac{\partial c_i(t)}{\partial \epsilon_j P_j} = \frac{\partial c_i(t)}{\partial \epsilon_j} \quad (2-29)$$

By substituting the semi-normalized coefficient in the ADE, the DDM sensitivity equation turns into (Hakami et al., 2003):

$$\frac{\partial s_{ij}^*}{\partial t} = -\nabla(us_{ij}^*) + \nabla(K\nabla s_{ij}^*) + J_{ik}s_{kj}^* + \frac{\partial R_i}{\partial \epsilon_j} + \frac{\partial S_i}{\partial \epsilon_j} - \nabla(uc_i)\delta_{ij} + \nabla(K\nabla c_i)\delta_{ij} \quad (2-30)$$

where

$$J; \quad \text{Jacobian Matrix} \left(\frac{\partial R_i}{\partial c_k} \right)$$

δ_{ij} ; binary variable in Kronecker delta function

As forward sensitivity analysis provides model-based gradients, it has been used as a tool in inverse modeling applications. Mendoza et al. (2000) employed forward sensitivity analysis to recover the emissions from known perturbations to the model. Their goal was to evaluate the method and the results for 5 different cases on pseudo observations. In all cases (with different pollutant and perturbation percentage), the results were acceptable. This method was employed for inverse modeling and emission adjustment of CO, NO_x and VOC over Atlanta (Mendoza et al., 2001). Inverse modeling of black carbon emissions with the help of CMAQ model and forward sensitivity analysis were used to calibrate the emission inventory of black carbon in the continental United States (Hu et al., 2009).

2.2.2 Backward sensitivity analysis

Conceptually, forward sensitivity analysis is source-based; this method is efficient for analysis and study of the effect of a few inputs on various model outputs (Hakami et al., 2006). On the other hand, backward or adjoint sensitivity analysis is a receptor-oriented method (receptor-based), and more efficient when the impact of a large number of model parameters or inputs (sources such as emissions) on a few number of target outputs is under investigation (Sandu et al., 2005). In adjoint sensitivity analysis, a perturbation at the receptor is propagated backward in both time and space with respect to all sources and input parameters (Hakami et al., 2007). This method is widely used in meteorology, oceanography, and nuclear physics for various applications such as parametric sensitivity analysis, data assimilation, inverse modeling, and parameter estimation.

Hakami et al. (2007) developed the adjoint version of the Community Multiscale Air Quality (CMAQ) model which is the most widely used regional air quality model worldwide. Unlike air quality models which output concentrations, the adjoint and tangent linear models output sensitivities. To develop the adjoint or tangent linear version of an air quality model, one needs to derive related differential equations that explain how sensitivities change over time and locations. The tangent linear model and adjoint equations can be derived using the mixing ratio form of the atmospheric diffusion equation:

$$\frac{\partial c_i}{\partial t} = -\mathbf{u} \cdot \nabla c_i + \frac{1}{\rho} \nabla \cdot (\rho \mathbf{K} \nabla c_i) + R_i + E_i \quad i = 1, \dots, N$$

(2-31)

In operator form the tangent linear model (or DDM model) can be written as:

$$\delta \mathbf{C} = \mathbf{L} \delta \alpha \quad (2-32)$$

Eq. (2-32) is an operator form of Eq. (2-24) which represents the linearized form of the forward model or the TLM (Hakami et al., 2007).

Applying a Lagrange variable to the TLM and integration by parts results in the adjoint equation (Hakami et al., 2007):

$$-\frac{\partial \lambda_i}{\partial t} = \nabla \cdot (\mathbf{u} \lambda_i) + \nabla \cdot \left(\rho \mathbf{K} \nabla \frac{\lambda_i}{\rho} \right) + \mathbf{F}_i^T \lambda + \varphi_i \quad (2-33)$$

where λ_i is the adjoint variable $\left(\partial J / \partial c_i \right)$ and a negative sign shows backward integration in time, ρ is the air density, $\mathbf{F}_i^T$ is the i^{th} row of the transposed Jacobian of the chemical reaction rates, and φ_i is the forcing term for the adjoint equations $\left(\partial g / \partial c_i \right)$.

In this formulation, J is referred to as the cost function. The cost function that is defined and calculated in this research will be discussed in the methodology chapter;; various cost functions will be used in these thesis for different applications. Further details of each cost function will be provided in corresponding chapters.

The adjoint of a linear operator $\mathbf{L}$ can be derived from the duality principle (Wang et al., 2001)

$$\langle \mathbf{u}, \mathbf{L} \mathbf{v} \rangle = \langle \mathbf{L}^* \mathbf{u}, \mathbf{v} \rangle \quad (2-34)$$

where $\mathbf{L}^*$ is the adjoint operator and $\langle ., . \rangle$ denotes the inner product. It can be shown that for a linear (or linearized operator such as the TLM) $\mathbf{L}^* = \mathbf{L}^T$. Defining an adjoint cost function

$J=G(y)$ that is a function on model outputs, and applying the duality principle on the inputs and outputs of the TLM results in:

$$\delta J = \langle \nabla_y J, \delta y \rangle = \langle \nabla_y J, L \delta x \rangle = \langle \delta x, L^* \nabla_y J \rangle \quad (2-35)$$

which defines the adjoint-based gradients as $\nabla_x J = L^* \nabla_y J$.

Discrete and continuous methods can be applied for integration of the adjoint equations. In the discrete method the adjoint model is developed directly from the discretized TLM by transposing the linearized operator and employing an operator splitting scheme for backward integration of adjoint variables (Carmichael et al., 2008). On the other hand, in the continuous method, the adjoint equation is discretized and solved numerically (Hakami et al., 2007; Carmichael et al., 2008).

For emission inversion applications, calculating the gradient of the cost function with respect to emissions is essential. In this case, the gradient of the cost function with respect to emissions is integrated over time as the adjoint model marches back in time steps (Hakami et al., 2007):

$$\frac{\partial J}{\partial \mathbf{E}^n} = \left(\frac{\partial C^{n+1}}{\partial \mathbf{E}^n} \right)^T \times \frac{\partial J}{\partial C^{n+1}} \quad (2 - 36)$$

where $\mathbf{E}^n$ is the emission vector at the time step of n , and $\frac{\partial J}{\partial C^{n+1}}$ is gradient of the cost function with respect to concentration at the last time step (going backward).

While adjoint models have been historically used for data assimilation and inverse modeling applications, the method can be used for different purposes such as population exposure analysis,

environmental exposure, effect of meteorological conditions on air pollution levels, policy evaluation, etc. (Hakami et al., 2007).

In the context of chemical data assimilation or inverse modeling, the adjoint version of a CTM such as CMAQ is essential. One of the first attempts for adjoint modeling was by Fisher et al. (1995) based on an adjoint for a stratospheric Lagrangian model. At the tropospheric level, the adjoint of a Lagrangian model was developed by Elbern et al. (1997). Sandu et al. (2005) developed the adjoint version of STEM for ozone data assimilation in East Asia. The adjoint version of air quality models such as CHIMERE (Vautard et al., 2000), IMAGES (Muller et al., 2005), GEOS-CHEM (Henze et al., 2007) were developed in the past decade.

The adjoint version of CHIMERE was used for recovery of boundary ozone values (Vautard et al., 2000). Initial concentrations of several species were optimized by using the adjoint of EURAD-CTM2 over central Europe (Elbern et al., 2001). CO and NO_x emissions were inverted simultaneously over Europe using the adjoint of IMAGES model (Muller et al. 2005). Adjoint version of STEM-2k1 was used for inverse modeling of black carbon (Hakami et al. 2005). The adjoint version of GEOS-CHEM was applied for inverse modeling of CO using satellite observations (Kopacz et al., 2009).

2.3 Observations

Remote sensing techniques rely on using electromagnetic radiation to acquire information from the atmosphere without any physical contact with the surface of the earth. Aerial photography from balloons was one of the first methods to acquire information about the Earth's surface in the mid eighteenth century. Nowadays, after years of research and development in technology,

continuous and long-term data acquisition is possible from a sensor in orbit around the Earth. Today, with the remote sensing techniques and available sensors covering the Earth, scientists are able to assess different environmental properties such as ground temperature, pressure profiles, concentration of tropospheric and stratospheric trace gases, surface albedo, change in the land and ocean elevation, etc. Beside satellites, there are other ways to get observations such as getting samples from ships, aircrafts or balloons, or collecting data from conventional stations; however, those methods have their own advantages and disadvantages. For example, observations from ships may be good for testing limited parameters at specific points, but they cannot provide enough temporal and spatial coverage to study dynamics of the ocean or atmosphere. Aircraft samples cannot be used on a consistent basis for regional and intercontinental movement of trace gases. Ground measurements are available with higher temporal resolutions with a limited number of stations. Satellites have better spatial resolution with limited temporal resolution and their view can be obstructed by clouds or over specific land covers. For this reason, combining a variety of observation platforms offers an opportunity to take advantage of the specific strengths of various observing platforms.

2.3.1 Satellites

In the early nineteenth century, Robbert H. Goddard used rockets for weather photography, but officially, by launching Sputnik 1 (supported by the USSR) on October 4th of 1957 the new age of satellite usage began (Hunley., 1995). After that satellite, Sputnik 2, Explorer 1, Explorer 7, TIROS 1, ATS 1 and the NIMBUS series were launched to get more accurate data and obtain better understanding of the Earth's properties. After about 50 years, satellites provide a wide range of global geophysical and biological parameters for scientists. The first weather satellite was launched by NASA on April 1st 1960, which was called TIROS 1. The main mission of that

satellite was to identify the storm areas, snow and ice over land and water, and temperature profiles over land and ocean. The first satellite, which measured Earth radiations based on Earth's energy budget, was Explorer 7. Satellite measurements are great tools for understanding the chemical and dynamical behavior of pollutants in the atmosphere. There are different kinds of satellites with different characteristics and usages that are being employed in atmospheric studies. Table 2-1 shows a number of chemical species with space borne observations and their related sensors, while Table 2-2 shows the main characteristics of these satellites.

Table 2-1. Chemical species and related products from sensors (Vijayaraghava et al., 2008)

Product	Sensor
Bromine oxide (BrO)	OMI
Carbon monoxide (CO)	AIRS,MLS,MOPITT,TES
Dinitrogen pentoxide (N ₂ O ₅)	HIRDLS
Formaldehyde (HCHO)	OMI
Glyoxal (OCHCHO)	OMI
Nitric Acid (HNO ₃)	MLS
Nitrogen dioxide (NO ₂)	HIRDLS,OMI
Ozone (O ₃)	AIRS,HIRDLS,MLS,MODIS,OMI,SBUV/2,TES,TOMS
Aerosol properties such as aerosol optical depth (AOD)	CALIOP,HIRDLS,MISR,MOIS,OMI

Table 2-2. List of major current satellites in air quality studies (Martin, 2008)

Instrument	Platform	Measuring period	Typical nadir (km)	Equator crossing time	Global coverage (days)	Spectral range
GOME	ERS-2	1995-2003	320*40	10:30d	3	0.23-0.79
MOPITT	Terra	2000-current	22*22	10:30d	3.5	4.7
MISR	Terra	2000- current	18*18	10:30d	7	0.45-0.87
MODIS	Terra	2000- current	10*10	10:30d	2	
	Aqua	2002- current	10*10	1:30a	2	0.41-14.2
AIRS	Aqua	2002- current	14*14	1:30a	1	3.7-16
SCHIAMACHY	ENVISAT	2002- current	60*30	10:00d	6	0.23-2.3
OMI	Aura	2004- current	24*13	1:45a	1	0.27-0.50
TES	Aura	2004- current	8*5	1:45a	-	3.3-15.4
PARASOL	PARASOL	2004- current	18*16	1:30a	1	0.44-1
CALIOP	CALIPSO	2006- current	40*40	1:30a	-	0.53,1.06
GOME-2	MetOp	2006- current	80*40	9:30d	1	0.24-0.79
IASI	MetOp	2006- current	12*12	9:30d	0.5	3.6-15.5

Nowadays, most of the industrialized countries have ground-based air quality measurement stations to observe the surface-level concentrations of pollutants. In some cases industries have their own stations to detect their emissions emitted to the atmosphere. These stations are good enough to provide basic estimation of the concentration of pollutants on limited number of points, but they cannot be useful in large-scale studies of pollutants and their transport. For air

quality studies from 1995, scientists have tried to extract useful information from satellite data to cover a larger area (global coverage) (Baum et al., 2006). In the atmospheric chemistry studies related to NO_2 , using satellite data is useful as this pollutant has a short lifetime, is highly variable, both spatially and temporally, and its concentrations is closely linked to emission sources.

Until recently the two main instruments that observed NO_2 from space were UV-Visible spectrometers GOME and SCIAMACHY. The advantage of using GOME and SCIAMACHY was to observe and record yearly and seasonal data of NO_2 , but there were some limitations such as their spatial resolution and coverage period which were the main reasons for launching new satellites. GOME had a coarse resolution ($320 \times 40 \text{ km}^2$) and SCIAMACHY had a better but still coarse spatial resolution ($60 \times 40 \text{ km}^2$) which could only cover the emissions of large cities. Moreover, while SCLAMACHY had a better temporal resolution with a global coverage of once a month (Boresma et al., 2004), providing a global coverage was still a challenge.

2.3.2 The ozone monitoring instrument (OMI)

The Ozone Monitoring Instrument (OMI) was built by the Dutch Space/TNO-TPD (Netherlands) in co-operation with VTT and Patria Finavitec (Finland) (Boresma et al., 2004). This instrument was launched via the EOS-AURA satellite in July 2004. Figure 2-3 shows the design on OMI. OMI is a new generation of UV/VIS solar backscatter spectrographs that has two main features: high spatial resolution ($13 \times 24 \text{ km}^2$ (along $\times$ across track)) and daily nadir-viewing. The resolution at nadir viewing is $13 \times 24 \text{ km}^2$; this resolution gradually increases for pixels far from the nadir because of the curvature of the surface of the Earth and the position of the sensors (Figure 2-4). This resolution can cover most cities and major sources for assessment of individual sources impacts. Furthermore, it provides a wide range of global coverage leading to

more cloud-free pixels compared to the older instruments (e.g., GOME and SCIAMACHY) (Boresma et al., 2004). The daily coverage is particularly important for studying pollutants' transport and the daily assessment of emission sources because it provides more observations. OMI measures several parameters used for atmospheric chemistry (in tropospheric and stratospheric) climate research and air pollution including: ozone, nitrogen dioxide, sulfur dioxide, formaldehyde, aerosols and cloud properties.

OMI covers the ultraviolet-visible direct and atmosphere-backscattered sunlight, which ranges from 270 nm to 500 nm (Levelt et al., 2006). This range is categorized in three spectral channels: two ultraviolet channels (UV1 (264-311 nm) and UV2 (307-383 nm)) and one visible channel (349-504nm) (OMI User's Guide, 2008). OMI traces polar orbit at 705 km altitude with local equator crossing time between 13:40 and 13:50 local time (Boresma et al., 2007). For one orbit, OMI observes 2600 km wide spatial swath, which is quite a wide field of view in contrast to the older instruments (GOME and SCHIAMACY).

Table 2-3. OMI instrument characteristics (Boresma 2004)

Mass	65 kg
Power	66 watts
Size	50 cm × 40 cm × 35 cm
Data rate	0.8 Mbps (average)
Orbit	Polar, sun-synchronous
Average altitude	705 km (438 miles)
Orbit period	98 minutes and 53 seconds
Ascending node local time	13:42 (1.42 PM)
Telescope swath Instantaneous Field Of View (IFOV)	115 degrees (2600 km on the ground)
Telescope flight Instantaneous Field Of View (IFOV)	1.0 degree (12 km on the ground)
Operational CCD temperature	265.0 K (UV and VIS)
Duty cycle	60 minutes on dayside (Earth and sun measurements), 10-30 minutes on eclipse side (calibration measurements)

Figure 2-3. OMI measurement design (OMI User's Guide, 2008)

Figure 2-4. Size of OMI ground pixels (OMI User's Guide, 2008)

2.3.3 Retrieval approach for OMI

The National Aeronautics and Space Administration (NASA) in corporation with the Royal Netherlands Meteorological Institute (KNMI, Koninklijk Nederlands Meteorologisch Instituut) developed a research plan to derive information related to NO_2 from OMI. This product is developed under the name of Dutch OMI NO_2 (DOMINO). Some features that are derived from DOMINO are:

- Averaging kernels: Averaging kernel is a vector that depends on the sensitivity of the retrieval to concentrations for each pixel and each pressure layer.
- Profile shape: Tropospheric profile shapes of NO_2 are extracted by using a global chemistry transport model (TM4). Vertical distribution of NO_2 is very important because of the variability of NO_2 in different altitudes, which may create errors in the estimation of NO_2 column density.
- Clouds: one of the main problems in older instruments, besides their coverage period, was the lack of information obtained in the presence of clouds. Cloud algorithms are improved in OMI to get more information in partially cloudy conditions by improving cloud fraction and cloud top height algorithms. This will improve the retrievals of the OMI NO_2 product at high altitudes.

The near-real time NO_2 retrieval algorithm is based on the retrieval-assimilation-modeling approach (RAM). This approach is based on two main steps: spectral fitting and estimation of air mass factors (AMF) (Boresma et al., 2007). Slant column density of a trace gas is defined as the concentration of that trace gas integrated over the path of the light in the atmosphere from the surface to the sensor. This value consists of the tropospheric and stratospheric amounts observed

by the satellite and therefore this value is reported as column amount. Slant column density of the trace gas is determined from radiance spectral fitting by the Differential Optical Absorption Spectroscopy (DOAS) technique. The DOAS technique defines slant column density as the integrated abundance of NO₂ along the average photon path from the Sun to the sensor (Lamsal et al., 2010). The stratospheric column density is calculated from assimilating OMI NO₂ slant columns into TM4. In order to use the measured value by the satellite as an observation in air quality applications, slant column density has to be converted into vertical values, i.e., vertical column density. Figure 2-5 shows the schematic representation of the slant path and the vertical path.

Figure 2-5. Schematic representation of slant path (red line) and vertical path (blue line) of the light

The tropospheric air mass factors (AMFs) are essential to convert a tropospheric slant column to a tropospheric vertical column. These air mass factors are derived as a function of a NO₂ vertical profile calculated by TM4.

$$V_{tr} = \frac{(S - S_{st})}{M_{tr}(x_{a,tr}, b)} \quad (2 - 37)$$

Where:

- S is the total slant column
- S_{st} is the slant column related to the stratosphere
- M_{tr} or AMF_{tr} is the tropospheric air mass factor
- $x_{a,tr}$ is the a-priori tropospheric NO_2 profile
- b contains information about cloud fraction cloud pressure, surface albedo, aerosols and viewing geometry

Based on Boresma et al. (2004) tropospheric air mass factors can be calculated as follows:

$$AMF_{tr} = \frac{\sum_l m_l(\hat{b}) x_{a,l} c_l}{\sum_l x_{a,l}} \quad (2 - 38)$$

Where:

- l is atmospheric layer index
- m_l are box air mass factors which are altitude dependent
- $x_{a,l}$ is the a priori tropospheric NO_2 profile
- c_l are layer correction terms

Box air mass factors are computed from DAK radiative transfer model, while the a priori tropospheric NO_2 profile is calculated from TM4 model (Zhou et al., 2009). Averaging kernels can be formulated as a function of air mass factors as follows (Boresma et al., 2004):

$$Ax_i = \frac{\sum_l m_l(\hat{b}) x_{a,l}}{AMF_{tr}(x_{a,tr}, b)} \quad (2 - 39)$$

The tropospheric column is dependent on the NO₂ vertical profile shape and this profile shape is calculated from a global chemistry transport model (TM4). Lamsal et al. (2008) suggested a new formulation for tropospheric column density to remove the influence of the TM4 model.

$$DP_{GC} = \frac{\pi_s \sum_k X_k^{GC}}{AMF \sum_k A_k X_k^{GC}} \quad (2 - 40)$$

where

- π_s is tropospheric slant column density (difference between total slant and stratospheric slant column density)
- A_k is averaging kernels for each layer
- k is the pressure layers of OMI
- X_k^{GC} is NO₂ profile shape calculated by GEOS-CHEM

The benefit of using this formulation is that the column density calculated this way is more consistent with the GEOS-CHEM model. The same approach can be used to estimate column densities that are consistent with another CTM such as CMAQ.

2.3.4 DOMINO Level 2 product

There are different OMI products for different applications. In this thesis, the Level 2 product of DOMINO is employed as the product developed for NO₂ studies. This DOMINO product contains geolocated NO₂ columns and other related parameters such as geophysical information related to each pixel observed by the instrument. Only a handful of parameters from this product are used which are listed below (Boresma et al., 2009):

- TM4PressurelevelA : A pressure level from TM4 model (Pa)
- TM4PressurelevelB : B pressure level from TM4 (Pa)
- TM4SurfacePressure : surface pressure level from TM4 (hPa)
- TM4TropopauseLevel: TM4 level where tropopause occurs (no unit)
- AveragingKernel : averaging kernels (no unit, 0.001) (Eskes and Boersma 2003)
- TroposphericVerticalColumn : NO₂ tropospheric vertical column
- VCDTropErrorUsingAvKernel : errors in NO₂ tropospheric vertical column
- CloudFraction : effective cloud fraction

There are a large number of applications for satellite observations in atmospheric studies. Martin et al. (2003) studied the relationship between ground concentrations of NO₂ and the GOME satellite observations. Muller et al. (2005) used GOME observations for inverse modeling of CO and NO_x emissions. Satellite observations of SCIAMACHY are employed to recover and improve the NO_x emissions sources (Martin et al., 2006). Konovalov et al. (2006) improved the estimates of NO_x emissions on regional scale using SCIAMACHY observations. Inverse modeling of NO_x emissions in southern United States have been conducted using SCIAMACHY observations (Napelenok et al., 2008). The sources of SO₂ emissions were estimated from SCIAMACHY and OMI satellite observations using GEOS-CHEM model (Lee et al., 2011). Temporal and spatial relationship between ground-based network observations operated by the Ontario Ministry of Environment and vertical column densities of NO₂ observed from OMI were studied over the Great Lakes region (Lee et al., 2011). Using the global chemistry model GEOS-CHEM and over the African Sahel during the monsoon onset (April-June) employing OMI satellite observations, the contribution of

soils as a major source of NO_x were estimated (Hudman et al., 2012). Torres et al. (2012) proposed a method to detect the aerosol optical depth both above and under the clouds. The recent work in the same field by Tang et al. (2013) examined NO_x emission inversion over Texas using sequential methods with the help of OMI observations and ground-based observations. Global emission sources of SO_2 were estimated from combining satellite observations of OMI, SCHIMACY and GOME-2 (Fioletov et al., 2013). Vinken et al. (2014) estimated the worldwide emissions sources of biogenic soil NO_x sources using GEOS-CHEM and OMI observations. NO_x emissions from ships over Europe were constrained using OMI observations and nested GEOS-CHEM (Vinken et al., 2014). Lu et al. (2015) employed OMI NO_2 retrievals to estimate NO_x emissions from 35 major urban cities in the U.S. in a long term from 2005-2014.

2.3.5 Ambient air monitoring networks

The ambient air monitoring networks are designed for three main reasons: to provide air pollutant data in a specific timely coverage, to be used as a support tool for setting ambient air quality standards and corresponding emission control strategy development and finally, to be used as a support data base for air pollution research studies (AQS fundamentals, 2006).

The Air Quality System (AQS) is a U.S.-based network for storing ambient outdoor airborne pollutants data. These data are collected by the Environmental Protection Agency (EPA) stations with the help of some states and local air pollution control agencies. It is then provided for the development of the AQS, which has been in place since 1957. AQS is not a real-time database; it has a delay of 3 months for data processing and quality assurance/quality control. The collection and maintenance of air quality data by this network

and provisions of a national database are among the main requirements of the U.S. Clean Air Act (AQS user guide, Air Quality System 2009).

These network stations record the data of concentration of criteria pollutants (SO_2 , NO_2 , O_3 , CO , PM_{10} , $\text{PM}_{2.5}$, and Pb), air toxins and related meteorological data in a timely manner. AQS networks are mainly located in 3 land use categories; suburban, urban and center city, and rural. Sampling duration varies from 5 minutes to 28 days and depends on the usage of data. The sampling method varies based on the type of the pollutant. The two main methods in sampling photochemical criteria pollutants are gas-phase chemiluminescence and ultraviolet photometry methods.

The monitoring stations of the two main species, ozone and nitrogen dioxide, employed in this thesis are described as below:

- Ozone measurements

Ground level ozone is a production of nitrogen oxide (NO_x) and volatile organic compounds (VOCs). Ozone has harmful effects on vegetation and human health. Under the Clean Air Act regulations, ozone concentrations are to be monitored and maintained under the certain level defined by the National Ambient Air Quality Standards (NAAQS). In the AQS database, ozone concentration (Figure 2-6) is stored in two main units: ppmv and ppbv. Ultraviolet photometry is a method of analysis for ozone concentration in the AQS database. This method is based on the absorption band in the ultraviolet region at 254 nm (U.S. EPA, 1996).

Shaded states have monitors

Monitor Location: ▽ O3 (1,213)

Source: US EPA Office of Air and Radiation, AQS Database

Tuesday, December 7, 2010

Figure 2-6. U.S. EPA Ozone ground network distribution

- Nitrogen dioxide measurements

Nitrogen oxides are major toxic pollutants which play an important role in ozone formation and acid deposition chemistry. Assessment of NO_2 concentration in the atmosphere is required for developing effective strategies to control emissions and improve air quality. Hourly data of NO_2 is monitored (Figure 2-7), processed and then stored in the AQS in ppmv unit. Gas-phase chemiluminescence method is an analysis method for measuring NO_2 concentrations, which is different from O_3 analysis methods (Wang et al., 2005, U.S. EPA, 2002).

Figure 2-7. U.S. EPA Nitrogen Dioxide ground network distribution

In this method NO and NO_x (NO + NO₂) concentrations are measured separately and then the NO₂ concentration is obtained from the difference between NO and NO_x measurements.

One main limitation of the chemiluminescence method is the reaction between NO and other nitrogen oxides during the measurement process which in turn results in an overestimation of NO₂ concentrations at ground stations (Demerjian et al., 2000). In the case of using ground measurements, the aforementioned limitation, which creates a bias in measured NO₂ concentrations, should be adjusted by applying a correction factor.

CHAPTER THREE:

Methods and modeling tools

In this chapter, the general methodology used in this thesis as related to emission inversion is presented. As each following chapter discusses specifics of the methods used for the segment of research, this chapter only provides a brief overview of the general methods and tools that are common to all segments. Emission inversion for the research conducted in this thesis will be carried out using the gas-phase adjoint version of U.S. EPA's CMAQ model (Hakami et al., 2007) model (gas-phase processes). The chemical mechanism for this version of CMAQ (Byun and Schere, 2006) was SAPRC-99 (Carter, 2000) and simulations were set up over the domain of North America. Domain designed based on a 36 km horizontal resolution and 13 vertical layers.

Figure 3-1. North American domain

Two main groups of inputs for CMAQ are emissions and meteorological data. Emission data used in this research are provided by Sparse Matrix Operator Kernel Emissions (SMOKE) (CEP, 2009) model. Meteorological fields are simulated by The Weather Research and Forecasting (WRF) (NCAR, 2010) model. OMI Vertical Column Densities (VCDs) are used as observations in this work.

3.1 Community Multiscale Air Quality model

The CMAQ model developed by U.S. EPA is an advanced community model that has become the most widely used air quality model in the world. CMAQ is a state-of-the-science multi-pollutant and multi-scale air quality model for simulating atmospheric and land process. It

simulates the relationships between pollutants on continental, regional, and urban scales. CMAQ solves the following Atmospheric Diffusion Equation (ADE):

$$\frac{\partial c_i}{\partial t} = -\nabla(\mathbf{u}c_i) + \nabla(\mathbf{K}\nabla c_i) + R_i[c_1, c_2, \dots, c_n; T, t] + S_i \quad i = 1, \dots, N \quad (2-31)$$

where c_i is the concentration of species i , $\mathbf{u}$ is the wind field, $\mathbf{K}$ is the second order turbulent diffusivity tensor, R_i is the net rate of chemical productions and S_i is the emission source rate of emissions of species i . CMAQ is part of Models-3 modeling platform that includes three modeling systems: meteorological, emission, and chemical transport models. As CMAQ is an offline model, meteorological and emission inputs are needed to be prepared by other models such as WRF and SMOKE separately. There are different interface processors in the modeling system in order to incorporate meteorological and emission outputs with input information of the initial condition and boundary condition and the chemical processors. These interface processors are: emission-chemistry, meteorology-chemistry, initial and boundary conditions and photolysis processors. CMAQ chemical transport model (CCTM) simulates main processes in the atmosphere, such as chemistry, transport and deposition. CCTM simulates processes through a computational framework that allows for domain decomposition and parallel processing. This framework is based on a one-atmosphere dynamics modeling system, chemistry mechanism, specialized numerical solvers, horizontal and vertical advection and diffusion, photolysis rate estimation, and photochemistry. Science processes in CCTM are advection and diffusion, gas phase chemistry, plume-in-grid (PinG) modeling, particle modeling and visibility, cloud processes, and photolysis rates (Figure 3-2).

Figure 3-2. Model-3 components (Byun and Schere, 2006)

Other than the meteorological and emission inputs, CTM needs the initial and boundary conditions and photolysis rate coefficients. The initial condition and boundary conditions are processed in the initial condition processor (ICON) and boundary condition processor (BCON), respectively. The clear sky photolysis rates for different times and latitudes are calculated by JPROC processor. Photolysis rates are affected by factors such as the location, time of the day, day of the season and composition of the atmosphere. There are different chemical mechanisms developed for running CMAQ. For example, the Statewide Air Pollution Research Center, Version 1999 (SAPRC-99) used in our study includes 214 reactions and 80 different species while Carbon Bond version 5 (CB05) mechanism includes 156 reactions for 59 species. The CMAQ model used in this thesis is the gas-phase variant of version 4.5.1 with the SAPRC-99 chemical mechanism.

3.2 Adjoint of CMAQ

The implementation of the four-dimensional variational (4D-Var) requires the cost function and the gradient of the objective cost function. In the presented study, the cost function calculated from the CMAQ forward run, and the gradients are provided through backward simulations which are carried out using the adjoint version of CMAQ (CMAQ-ADJ, gas-phase, Hakami et al., 2007).

The sequence of the process for CMAQ-ADJ is very similar to its forward and the main difference is that model is integrated backward in time. In the processes such as chemistry, diffusion and vertical advection the discrete adjoint scheme is used but for horizontal advection continuous adjoint is employed. The CMAQ adjoint uses the kinetic pre-processor (KPP) version 2.2 for chemistry integration, (Damian et al. 2002; Daescu et al. 2000). The piecewise parabolic

method (Colella and Woodward 1984) is used for the horizontal advection process, and an upwind first order finite difference method integrates vertical diffusion and emission injection.

As discussed in chapter 2, adjoint sensitivity analysis is a receptor-oriented approach that amounts to propagating perturbations in a receptor-based cost function backward in time and space, and into individual emission sources. Backward integration of the adjoint model results in estimation of the sensitivities of the cost function with respect to the input parameter(s) such as emissions. In other words, the integration of the adjoint equation results in the adjoint variables (i.e., derivatives of the cost function with respect to state) and adjoint gradients (i.e., derivatives of the cost function with respect to emissions). These adjoint variables and gradients contain sensitivity information with respect to input parameter(s) at each time and in each grid cell. The adjoint cost function for our applications is calculated based on the difference between observations and model output after the forward run. Adjoint variables are calculated using the adjoint equation fed by adjoint forcing terms which are the local derivative of the cost function with respect to the concentrations. In that sense, the adjoint forcing term drives generation of the adjoint values in the same fashion that emissions lead to generation of concentrations in the forward model simulations.

3.3 Sparse Matrix Operator Kernel Emissions

The Sparse Matrix Operator Kernel Emissions (SMOKE) (CEP, 2009) modeling system was developed for emission processing by MCNC Environmental Modeling Center (EMC). Not only can SMOKE process pollutants such as nitrogen oxides (NO_x), carbon monoxide (CO), sulfur dioxide (SO_2), ammonia (NH_3), volatile organic compounds (VOCs), and particulate matters (PM_{10} and $\text{PM}_{2.5}$), but it can also process toxic pollutants such as benzene, cadmium, mercury

and formaldehyde. SMOKE as an emission processor is used to translate the emission inventory data into spatial and temporal gridded data readable by air quality models. In general, emission inventory data includes emissions in annual-total or average-day emissions while air quality models need data on an hourly basis.

The SMOKE model creates emission inputs for air quality models by considering area, mobile and point source emissions, as well as consideration for natural and biogenic emissions. For emission processing in SMOKE, inventories are classified into several source categories: stationary area/nonpoint sources, non-road mobile sources, on-road mobile sources, point sources, wildfire sources, and biogenic sources which contain land use data.

Inventory Data Analyzer (IDA) is the format of the emission inventory input files for SMOKE processing in this work. The IDA files include the yearly emission data, which SMOKE allocates into gridded hourly emission files. The IDA files were taken from the U.S. EPA's National Emission Inventory (NEI) and projected from the year 2005 to 2007 based on the change in population growth, industrial activities and land use patterns.

3.4 Weather Research and Forecasting

The Weather Research and Forecasting model is a mesoscale numerical weather system designed to provide meteorological data for atmospheric research and weather forecasting. WRF is a non-hydrostatic meteorological model with the terrain-following pressure vertical coordinate, and is a proven model for resolving the atmospheric dynamics at regional scales (NCAR, 2010). The WRF model is used in this study to generate the meteorological fields required for the simulations for summer 2007. This model prepares the meteorological fields, suited for the emission model (SMOKE) and CMAQ, which are temporal and gridded. The modeling domain

of WRF for this simulation covers North America with a horizontal resolution of 36 km. It has 13 vertical layers, with the first layer of about 20 m height, extending from the surface to 100 hPa. The 13 pressure layers are 1, 0.9975, 0.995, 0.99, 0.98, 0.96, 0.92, 0.875, 0.82, 0.65, 0.475, 0.32 and 0.16 in sigma coordinates. Sigma coordinate is the traditional name for the terrain-following hydrostatic-pressure vertical coordinate. WRF version 3.0 used in this study has a domain that includes 165 columns and 129 rows and 35 vertical layers. However, WRF outputs are not gridded the same way as the CMAQ outputs and they should be converted to become readable by CMAQ. The Meteorology Chemistry Interface Processor (MCIP) (Byun et al., 1999a) is used to convert WRF outputs. This processor generates the required meteorological input data for CMAQ with the resolution of (148 columns $\times$ 112 rows $\times$ 13 layers) from WRF outputs (165 columns $\times$ 129 rows $\times$ 35 layers).

3.5 Inverse modeling

Emission inversion is done through an iterative assimilation process using the adjoint version of CMAQ. The 4 D-Variational processes that is used for inversion includes four main iterative steps:

(1) Forward run

In this step, concentration files while the model is fed by emissions and scaling factors will be generated. In the first run, the scaling factors which are responsible for perturbing NO and NO₂ are assumed to be 1. In other words, under the first forward run, which is called call the reference run, concentrations at all points and layers were calculated without any changes in emission rates of NO and NO₂.

(2) Observations

OMI Vertical Column Densities (VCDs) are used in this work as the observation datasets because of their resolution and daily coverage of the whole domain. In order to compare the observations with model output, vertical column densities of CMAQ are calculated and mapped onto the OMI domain. In this step, CMAQ-VCDs are calculated and mapped horizontally and vertically to be comparable with OMI-VCDs. By definition, VCDs are the number of molecules in the unit area ($\text{molecules}/\text{cm}^2$) and this value for OMI can be extracted directly from OMI files. Calculation of the cost function requires computing the difference between column density obtained from OMI-VCDs and CMAQ-VCDs which have different vertical and horizontal resolutions. Two operators are defined for mapping column density: the vertical operator and the horizontal operator.

- Vertical operator:

OMI has 34 pressure layers while CMAQ has 13 layers vertically. This operator is responsible for calculation of the number of molecules per unit of area based on difference between pressure layers of OMI and CMAQ. In other words, the mass of air for each OMI layer located within a CMAQ layer(s) is calculated using the OMI data and then vertically added together in order to find the total air mass for a CMAQ layer.

- Horizontal operator:

The resolution of CMAQ grids and OMI observations are different. While the resolution of CMAQ grid cells are $36 \times 36 \text{ km}^2$, OMI observations have finer resolution of $13 \times 24 \text{ km}^2$ at nadir. This operator is responsible to find the area of CMAQ grid cell(s) which is covered by each OMI observation.

Figure 3-3. Comparing the pressure layers of CMAQ and OMI (left), comparing the coverage of each CMAQ grid cell with OMI observations (right)

(3) Backward run

In the backward run, the adjoint version of CMAQ is used to calculate the gradient of the cost function with respect to NO_x emissions. Backward simulations require concentrations values at each model time steps and are therefore conducted after forward simulations.

(4) Optimization

Inversion is an iterative method which needs cost function value and gradients. For inverse modeling of the emissions, NO_2 column densities that are obtained from OMI, after applying observation operators, are used in the 4D-Var data assimilation cost function (Hakami et al., 2006):

$$J = \frac{1}{2} (\mathbf{VCD}_{\text{CMAQ}} - \mathbf{VCD}_{\text{OMI}})^T \mathbf{R}^{-1} (\mathbf{VCD}_{\text{CMAQ}} - \mathbf{VCD}_{\text{OMI}}) + \frac{1}{2\mu} (1 - \mathbf{x})^T \mathbf{B}^{-1} (1 - \mathbf{x}) \quad (3-6)$$

In Eq. (3-6) $\mathbf{VCD}_{\text{CMAQ}}$ is the vector of model column density, $\mathbf{VCD}_{\text{OMI}}$ is the vector of satellite-based column densities, μ is the regularization parameter (global weighting factor), $\mathbf{x}$ is the vector of emission scaling factors and $\mathbf{R}$ and $\mathbf{B}$ are observation and background error covariance matrices.

The three-dimensional concentration field of CMAQ is mapped onto each equivalent OMI column density using vertical and horizontal operators as well as OMI's averaging kernel. The column density of the model corresponds to each OMI observation (Hakami et al., 2006):

$$\mathbf{VCD}_{\text{CMAQ}} = \mathbf{K}^T \mathbf{H}_v \mathbf{H}_h \mathbf{C}_{\text{CMAQ}} \quad (3-7)$$

In Eq. (3-7), $\mathbf{K}$ is the averaging kernel vector, as described before, which shows the relation of the retrieved information from satellites to the true state of atmosphere. $\mathbf{H}_v$ and $\mathbf{H}_h$ are vertical and horizontal observation operators, respectively. The vertical operator accounts for calculating the number of the molecules in the comparable column of OMI and the horizontal operator calculates the covered area of the CMAQ grid cell for each OMI observation. After applying averaging kernels, vertical column densities of OMI and CMAQ can be compared.

These operators are also used to calculate the forcing term required for backward simulations. For all observations with diagonal observation covariance matrix, the adjoint forcing term is defined as (Hakami et al., 2006):

$$\varphi_k = \frac{1}{R_k} \left(\left(\frac{\partial \text{VCD}_{\text{CMAQ},k}}{\partial C} - \frac{\partial \text{VCD}_{\text{OMI},k}}{\partial C} \right) (\text{VCD}_{\text{CMAQ},k} - \text{VCD}_{\text{OMI},k}) \right) \quad (3-8)$$

Once the gradients of the cost function with respect to emissions are evaluated, those gradients will be used in a Quasi-Newton L-BFGS-B optimization algorithm for optimization and updating

emission scaling factors. The updated scaling factors will be used again to perturb NO_x emissions and calculate the concentrations in the forward mode.

Figure 3-4. Schematic process of emission inversion

The aim of this work is to use inverse modeling method to reduce the uncertainties in the inputs (emissions). To do so, emission inversion is done in two main categories: time independent and time dependent. In the time independent case, emissions are inverted for the entire domain on a daily basis, while in the time dependent case; emissions are adjusted on an hourly basis. These adjustments are defined by scaling factors. The mission of these scaling factors is to perturb emissions of NO and NO_2 in a way that model outputs become closer to the real time observations. The difference between time dependent and time independent emission inversion is the way that scaling factors are designed. For the time independent case, emissions at each grid cell are perturbed by only one number at all times, whereas under the time dependent case, emissions at each time are perturbed by different scaling factors.

It is necessary to mention that preparation of the WRF and SMOKE outputs was a task that took place as a group work based on partial contributions from researchers other than author of this thesis. Extracting the observations from both satellite and ground-network platform and running the simulations were conducted by the author.

CHAPTER FOUR:

Variational inverse modeling of regional NO_x emissions based on satellite and ground based observations

4.1. Introduction

Inverse modeling has been widely used as an effective tool to estimate the input parameters of a system. In atmospheric applications, inverse modeling techniques can be used to reduce the uncertainty associated with model inputs such as initial condition, boundary condition, and particularly emissions. The uncertainty in emissions used by air quality models stems from emission processing models that calculate spatial and temporal emission rates using emission inventory data and some simplified assumptions. Inverse methods reduce the overall uncertainty by constraining emissions and through minimizing the distance between observed and simulated model outputs (i.e., concentrations). In general, uncertain emission information leads to uncertainty in the concentration calculated by the model. Based on the difference between the concentrations and available observations, inverse modeling provides a better estimation of

emissions that would lead to more accurate (realistic) concentrations. However, the uncertainty in the concentrations is not only because of emission information and other sources such as grid cell resolution, chemical mechanism, atmospheric reaction and/or removal process can create errors in concentrations. The focus of this work is on emissions as the largest source of uncertainty.

Inverse methods are derived from data assimilation algorithms (Talagrand and Courtier, 1987) and have been used in diverse fields including meteorology and atmospheric chemistry. In atmospheric chemistry, inverse modeling methods have been widely applied to estimate the sources and sinks of various airborne pollutants such as methane (Houweling et al., 1999; Wang et al., 2002), carbon dioxide (Rodenbeck et al., 2003; Gurney et al., 2005), and carbon monoxide (Bergamaschi et al., 2000; Petron et al., 2004; Muller et al., 2005).

Nitrogen Oxides ($\text{NO}_x = \text{NO}_2 + \text{NO}$) have an adverse effect on human health and play an important role in atmospheric photochemistry. NO_x emissions have a direct impact on ozone (O_3) production; influence all important OH concentrations and aerosol formations. The main sources of NO_x emissions are anthropogenic activities, primarily combustion of fossil fuels, along with biomass burning and natural processes such as soil and lightning. Slight changes in the NO_x emission inventory information can have a significant impact on ozone prediction (Xiao et al., 2010). From a health perspective, there is a connection between short-term NO_2 exposure and adverse respiratory effects (EPA, 2009).

Ground stations of the current NO_2 monitoring networks in the U.S. are sparsely distributed and their coverage is limited. Surface measurements of NO_2 have been used to improve the emission inventory via several case studies (Mendoza-Dominguez et al., 2001; Quélo et al., 2005), all of which showed a significant improvement in optimized emissions. In recent years, satellite

measurements of NO₂ have been used in inverse modeling applications to complement ground-based observations by providing a wider global coverage. Satellite observation of NO₂ has a history of almost two decades with the earliest observations from Global Ozone Monitoring Experiment (GOME-1) (Burrows et al., 1999) and Scanning Image Absorption spectrometer for Atmospheric CHartography (SCIAMACHY) (Bovensmann et al., 1999). GOME-1 was observing the globe in 3-day temporal and spatial resolution with a nadir resolution of 320 × 40 km² and SCIAMACHY had a global coverage of 6 days with a nadir resolution of 60 × 30 km². With GOME and SCIAMACHY out of commission, Ozone Monitoring Instrument (OMI) (Levelt et al., 2006b) provides daily information of vertical column of NO₂ with a better nadir resolution (24 × 13 km²).

Tropospheric NO₂ columns retrieved from satellite measurements have been widely used to provide top-down estimates of NO_x emissions through inverse modeling methods at the global and continental scale. At the global scale, inverse modeling of surface NO_x emissions: has paired GOME-1 (Martin et al., 2003a) or SCIAMACHY (Martin et al., 2006) with the GEOS-CHEM chemistry model; has combined ground-based observations with GOME-1 and aircraft observations using the IMAGES model (Müller et al., 2005); has combined GOME and SCIAMACHY using the IMAGES model (Stavrakou et al., 2008); and has combined GOME-2 and OMI using GEOS-CHEM (Lin et al., 2010). At the continental scale, inverse modeling of NO_x emissions has combined GOME-1 and SCIAMACHY satellite measurements with the CHIMERE CTM model (Konovalov et al., 2006); OMI satellite data with the REAM model (Zhao et al., 2009); and GOME satellite data with the CMAQ model (Kurokawa et al., 2009).

Inverse modeling of NO_x emissions has used two main methods: sequential and variational. The variational method minimizes the distance between model output and observations during the

whole modeling period and minimizes to improve the model inputs while the sequential method optimizes model inputs at multiple times when the observation is available (Khattatove et al., 1999). Both methods have advantages and disadvantages. The variational method requires the adjoint version of the forward model for gradient calculations while the sequential analysis results in a discontinuity whenever a new observation is encountered during the forward model run. Advances in NO_x emissions inversion using sequential methods were done by Napelenok et al. (2008) (i.e., CMAQ model with SCHIAMACY observations for the South Eastern United States) and Tang et al. (2013) (i.e., CAMx model with OMI and ground-based observations for Eastern Texas).

Four dimensional variational inverse modeling (4D-Var; 3-D spatial dimensions plus time) has also been used to estimate ground level NO_x emission sources. For example, (Zhange et al., 2008) used the STEM model with SCIAMACHY satellite observations and (Chai et al., 2009) assimilated two-month average NO_2 observations columns into STEM . They used different set ups to invert the emissions and reported a reduction of between 8.9% to 11.4% in all simulations. All their cases show the reduction of emissions in Ohio valley and Washington. To study the changes in the emissions, Huang et al. (2014) inverted NO_x emissions using different set-up observations and STEM in May 2010 on a $12 \times 12 \text{ Km}^2$ horizontal resolution. The results from inversions only by NO_2 column observations from OMI, only by ground-based network observations and by both indicated the reduction of 26%, 29% and 30% from base emissions on 2005. The effect of this change in emissions over California reported as a reduction in the mean error in surface ozone over the Western U.S. and the uncertainties in urban areas increased because of NO_x saturated regime.

The main objective of this chapter is to adjust NO_x emissions on a daily basis. Four different modeling cases were carried out to adjust predicted NO_x emissions using:

- (i) ground-based observations (Case-G);
- (ii) satellite observations (Case-O);
- (iii) combined ground-based and satellite-based observations (Case-OG); and,
- (iv) a monthly inversion for all observations using an inversion window of one month (case-MOG).

4.2. Methods

4.2.1 CMAQ modeling

In this study the Community Multi-scale Air Quality Model (CMAQ) (Byun and Schere, 2006) is used to simulate the concentrations of all species. CMAQ is a regional chemical transport model that is widely used in regulatory and research applications. The modeling domain covers the continental United States, Southern Canada and Northern Mexico with a horizontal grid resolution of 36 Km × 36 Km and 13 vertical layers. The CMAQ used in this research is the gas-phase variant of version 4.5.1 with the SPARC-99 (Carter 2000) chemical mechanism in parallel mode. The 4D-Var inverse modeling framework needs the gradients of the cost function with respect to the model input. The adjoint of CMAQ (CMAQ-ADJ) used in this study was developed for gas phase pollutants in version 4.5.1 (Hakami et al, 2007).

CMAQ requires emissions and meteorological inputs. Emission inputs are provided by the Sparse Matrix Operator Kernel Emission (SMOKE) version 2.4 (CEP, 2009) which references the National Emission Inventory (NEI); emissions processing was projected from 2005 to 2007.

Meteorological information is generated by Weather Research and Forecasting (WRF) model version 3.1 and was mapped to a modeling domain compatible with CMAQ using 148 columns, 112 rows and 13 layers.

4.2.2 Observations

Ground-based NO₂ observations were obtained from the U.S. EPA Air Quality System (AQS), which is sampled on an hourly basis (i.e., dataset available online: <http://www.epa.gov/ttn/airs/airsaqs/>). Details about sampling and analysis methods and calibration are provided by EPA (U.S. EPA, 1975). AQS provides hourly concentrations at surface level and current monitoring networks are sparsely distributed. Figure (4-1) shows the distribution of ground-based NO₂ observations and is taken from AQS website.

Figure 4-1. U.S. EPA Nitrogen Dioxide ground network distribution

Compared to ground-based sampling, satellite observations can provide information over a wider range, as well as the vertical distribution profile of species in the atmosphere. For this research, satellite column observations of NO₂ were obtained from Ozone Monitoring Instrument (OMI) data (i.e., OMI DOMINO (Dutch OMI NO₂) product version 2.0, DP-version 2; Boersma et al., 2011; available from Tropospheric Emission Monitoring Internet Service, TEMIS, at <http://www.temis.nl/airpollution/no2.html>). OMI is a nadir-viewing UV–VIS spectrometer that

measures solar backscattered irradiance in the range of 270 nm to 500 nm. Installed on NASA's EOS Aura satellite and launched on 15 July 2004, OMI has a local Equator crossing time around 13:40 with a spatial resolution of $13 \times 24 \text{ Km}^2$ at nadir viewpoint (Boersma et al., 2007).

The DOMINO algorithm to extract NO_2 vertical column densities is described by Boersma et al. (2007). In brief, the near-real time NO_2 retrieval algorithm is based on a three-step procedure:

- (i) computation of a slant column density from a spectral fit to the Earth reflectance spectrum using Differential Absorption Spectroscopy Approach (DOAS);
- (ii) Determination of the stratospheric NO_2 contribution of the slant column by assimilating the slant column into a global chemistry and transport model (TM4).
- (iii) Quantifying the tropospheric contribution of the slant column and converting it to an equivalent vertical column using tropospheric Air Mass Factors (AMFs).

Tropospheric NO_2 vertical columns are retrieved as follow:

$$\text{VCD}_{\text{trop}} = \frac{S - S_{\text{strop}}}{\text{AMF}_{\text{trop}}} \quad (4 - 1)$$

where S is the slant column density from the first step, S_{strop} is the stratospheric contribution of the slant column from the second step.

The tropospheric AMF is dependent on the a priori NO_2 profile, cloud fraction, cloud pressure, surface albedo and viewing geometry and can be computed by DAK radiative transfer model (Stammes, 2001, Boersma et al 2004). Huijnen et al. (2010) evaluated the DOMINO NO_2 product (version 1.0.2) and reported that DOMINO tropospheric columns are much larger than model simulation results, with a bias as large as 40%). To remove the influence of TM4 NO_2

profile shape, Lamsal et al. (2010) presented a new data set (DP_GC) compatible with GEOS-CHEM that combines averaging kernels from DOMINO product with the NO₂ profile shape from GEOS-CHEM simulations. Latest DOMINO product data set is available at TEMIS with an updated approach for AMF calculations (Boersma et al., 2011). In this study similar to (DP_GC) approach, a compatible data set with CMAQ is provided (DP_CMAQ).

In order to remove the influence of TM4 and substitute the effect with CMAQ, the air mass factor has to be generated based on altitude-dependent air mass factor and CMAQ concentrations. Averaging kernels, which are layer-dependent weighting factors explaining the relative contributions to the retrieval from true state, are available within DOMINO product at each vertical layer. By multiplying the air mass factor computed by TM4 into averaging kernels they will be dependent to in the forward model. Averaging kernels are calculated as follow (Eskes and Boersma, 2003):

$$A_i = \frac{m_i}{M} \quad (4 - 2)$$

$$M = \frac{\sum_i m_i C_i}{\sum_i C_i} \quad (4 - 3)$$

where A_i is averaging kernel at layer i , m_i is altitude-dependent air mass factor calculated by radiative transfer model, M is total air mass factor and C is a priori trace gas profile. Combining Eq. (4-2) and Eq. (4-3) and substitute in Eq. (4-1):

$$DP_{CMAQ} = \frac{(S - S_{strop}) \sum_i C_i^{CMAQ}}{AMF_{trop} \sum_i A_i C_i^{CMAQ}} \quad (4 - 4)$$

Observations with cloud fraction less than 50% were considered reasonable in this study. To improve model interpretability, pixels with bad cloud flag were filtered and pixels with a dimension larger than 36 km (edges of the satellite swath) were not considered.

4.2.3 Variational inverse modeling

Variational methods solve the inverse problem in an optimal control framework. The objective is to reduce the distance between model output and related observations by finding the proper control values (model parameter) through a minimization process. The problem of finding the most proper control values can be treated as a generalization of the 4D-Var approach which is widely applied to estimate the state space variables. Inverse modeling combines information from three main sources:

- (i) A model explaining the chemical and physical laws that govern the evolution of the system. The model is essential to calculate the future state values of the system. In general the model can compute the future state from initial state($\mathbf{x}_0$).

$$\mathbf{x}_i = \mathbf{F}_{t_0 \rightarrow t_i}(\mathbf{x}_0) \quad (4-5)$$

- (ii) The best estimate of the current state of the system. This estimate is generated from the model considering all known chemical and physical processes in the system. The aim is to improve this state and for that real observations explaining the behavior of the system are vital. The current state of the system before considering observations is called a-priori or background.
- (iii) Observations of the current state. If the model outputs and observations are similar in form and dimension (e.g. concentrations), the comparison will be

straightforward. In cases where model output is not similar to observations (e.g. satellite observations) for comparison observation operators are needed to map the state space into observation space.

$$\mathbf{y}_i^{\text{obs}} = \mathcal{H}(\mathbf{x}) + \boldsymbol{\varepsilon}_i^{\text{obs}} \quad (4-6)$$

$$\langle \boldsymbol{\varepsilon}^{\text{obs}} \rangle = 0 \quad \langle (\boldsymbol{\varepsilon}^{\text{obs}})(\boldsymbol{\varepsilon}^{\text{obs}})^T \rangle = \mathbf{R} \quad (4-7)$$

where $\mathbf{y}_i^{\text{obs}}$ is the observation vector at specific time, $\mathcal{H}$ is the observation operator to map the state space into observation space, $\boldsymbol{\varepsilon}_i^{\text{obs}}$ is the observational error. Errors in observations are assumed to have a normal distribution with zero mean with $\mathbf{R}$ as the covariance matrix. The error in one observation is uncorrelated to other observations at different time and space. Measurements, recording observations, are usually sparsely distributed through the system and the number of available observations is smaller than the dimension of the system.

With appropriate information from these three sources, inverse modeling framework can be a valuable tool to estimate a more accurate state of the system by reducing the discrepancy between model output and observations. This discrepancy is defined as the “cost function” and has two main terms, the background term and the observational term. The background term defines the departure of the initial state from background state; the observational term is the difference between model outputs and observations in all times and locations. In other words, the 4D-Var inverse method applied in this research tries to minimize the cost function by bringing the model output closer to observations with a penalty for departing too far from background (a-priori) state. The general mathematical form of the 4D-Var cost function is:

$$J(\mathbf{x}_0) = \frac{1}{2\mu} (\mathbf{x}_0 - \mathbf{x}_0^b) \mathbf{B}^{-1} (\mathbf{x}_0 - \mathbf{x}_0^b)^T + \frac{1}{2} \sum_{i=1}^N (\mathcal{H}(\mathbf{x}_i) - \mathbf{y}_i^{obs}) \mathbf{R}^{-1} (\mathcal{H}(\mathbf{x}_i) - \mathbf{y}_i^{obs})^T \quad (4-8)$$

where $J(\mathbf{x}_0)$ is the cost function based on absolute background values, μ is the global weighting factor, $\mathbf{x}_0$ is the initial state, $\mathbf{x}_0^b$ is the background state, $\mathbf{B}$ is the background error covariance matrix, $\mathcal{H}(\mathbf{x}_i)$ model output which is mapped into the observation state, $\mathbf{y}_i^{obs}$ is the observation vector and $\mathbf{R}$ is the observation error covariance matrix. The background term is weighted by the inverse of background covariance error matrix and the observational term is weighted by the inverse of observation error covariance matrix in order to be comparable with each other value wise.

The optimal state estimated by 4D-Var method is called posterior ($\mathbf{x}^a$) and is computed as:

$$\mathbf{x}_0^a = \text{argmin } J(\mathbf{x}_0) \quad (4-9)$$

The optimization problem is solved numerically using a gradient-based technique and thus requires that the derivative of the cost function with respect to the input parameters under investigation, making this a challenging step in inverse modeling application. Gradients are calculated using the adjoint version of the model being used for cost function calculation (i.e., CMAQ-ADJ in this research). CMAQ-ADJ calculates appropriate gradients using observations increments backward in time. The numerical optimization used in this experiment is limited memory bound-constrained L-BFGS (Zhu et al., 1997).

Figure 4-2. Schematic form of 4D-Var system

4.3 Inversion set-up

The emission inversion set up in this study is based on 4D-Var system using CMAQ. The distance between the available observation and model output is calculated in the cost function and the optimal solution for the emission scaling factors are obtained by finding new scaling factors which minimize the cost function. The sensitivity of the cost function with respect to the emissions is essential to solve the minimization problem efficiently. To invert the emissions, a scaling factor of the background emission at each day is used:

$$E_{\text{opt}}^{\text{NO}_x}(x_i, y_i, l_i, t) = \alpha_{i,j,l} E_B^{\text{NO}_x}(x_i, y_i, l_i, t) \quad (4 - 10)$$

where $E_{\text{opt}}^{\text{NO}_x}$ is the optimized NO_x emission after inversion, $E_B^{\text{NO}_x}$ is the initial NO_x emission and α is the scaling factor. The scaling factors are time independent and for each x_i, y_i they are applied to all the time profiles. In the base case, scaling factors are equal to one at all locations.

The cost function is redefined by inclusion of the scaling factors as:

$$J = \frac{1}{2\mu} (1 - \alpha) \mathbf{B}_{\text{scal}}^{-1} (1 - \alpha)^T + \frac{1}{2} (\mathbf{C}_{\text{CMAQ}} - \mathbf{C}_{\text{Obs}}) \mathbf{O}^{-1} (\mathbf{C}_{\text{CMAQ}} - \mathbf{C}_{\text{Obs}})^T \quad (4 - 11)$$

$$J = \frac{1}{2} (\mathbf{VCD}_{\text{CMAQ}} - \mathbf{VCD}_{\text{OMI}}) \mathbf{R}^{-1} (\mathbf{VCD}_{\text{CMAQ}} - \mathbf{VCD}_{\text{OMI}})^T + \frac{1}{2\mu} (1 - \alpha) \mathbf{B}_{\text{scal}}^{-1} (1 - \alpha)^T \quad (4 - 12)$$

$$J = \frac{1}{2} (\mathbf{VCD}_{\text{CMAQ}} - \mathbf{VCD}_{\text{OMI}}) \mathbf{R}^{-1} (\mathbf{VCD}_{\text{CMAQ}} - \mathbf{VCD}_{\text{OMI}})^T + \frac{1}{2\mu} (1 - \alpha) \mathbf{B}_{\text{scal}}^{-1} (1 - \alpha)^T + \frac{1}{2} (\mathbf{C}_{\text{CMAQ}} - \mathbf{C}_{\text{Obs}}) \mathbf{O}^{-1} (\mathbf{C}_{\text{CMAQ}} - \mathbf{C}_{\text{Obs}})^T \quad (4 - 13)$$

$\mathbf{R}$, $\mathbf{B}_{\text{scal}}$ and $\mathbf{O}$ are error covariance matrices for OMI observations, scaling factors of emissions, and ground observations. $\mathbf{VCD}_{\text{CMAQ}}$ is the total number of molecules of NO_2 calculated by model in a vertical layer corresponding to each OMI observation. The OMI averaging kernels are applied to column density of CMAQ. $\mathbf{VCD}_{\text{OMI}}$ is calculated from Eq. (4-4). $\mathbf{C}_{\text{CMAQ}}$ is the NO_2 concentration calculated by CMAQ and $\mathbf{C}_{\text{Obs}}$ is the observation gathered from AQS network. Eq. (4-11) is in the case of only inverting emission using ground based (hereafter, case-G) observations, Eq. (4-12) is for the case of only inverting emissions using OMI observations (hereafter, Case-O) and Eq. (4-13) is in the case of emission inversion for both ground based and OMI (i.e., Case-OG) observations at the same time. The challenging part is to start the inversion process with an appropriate error representation for the emissions. To build the background error covariance matrices, there is no reference for comparison, and thus the distribution of daily

emissions is considered as a reference and calculated the variances based on the variability of emissions for each day during summer of 2007. In this chapter, the error covariance matrices are considered to be diagonal for scaling factors of emissions and ground observations. The error covariance matrix for OMI observations is taken from the DOMINO product and is considered to be diagonal.

A large-scale bound-constrained limited-memory Quasi-Newton code, L-BFGS-B (Zhu et al., 1997), is used for the inversion in this research. The upper and lower bounds of scaling factors during inversion were assigned as 4 and 0.25 in order to allow them increase or reduce 4 times the original values. For all inversion cases (i.e., ground-based observation data; satellite observation data; etc.) the assimilation window was set to 24 hours and the number of iterations was set to 15. The inversion took place for all the grid cells for which the average emission during the whole 31 days was more than %1 of the average of the total emission of the entire domain.

4.4. Results

4.4.1 Identical twin experiment using pseudo observational data

The identical twin experiment is designed to evaluate the performance of the emission inversion set up. Under this experiment, two sets of concentrations, or model states, known as true and false states are generated by the model. The “false state” represents concentrations when false emissions (i.e., emissions with errors) are unperturbed. The “true state” represents hypothetical observed concentrations (pseudo observations), which are generated by the model and known scaling factors. The main advantage of building the observation state by the model is that all model errors are neglected and true emissions with no error (i.e., emissions that inverse models

aim to achieve) are known. In order to build the observations or “true state”, the scaling factors of 0.25, 0.5, 1.5, 2, 2.5, 3, 3.5 and 4 are applied to emissions at different locations.

The retrieved scaling factors from inversion and the cost function reduction are presented in Figure (4-3). Results show that in less than 7 iterations, cost function reduces by 90% and then the rate of the reduction is very slow for further iterations. Note that the inverse problem is ill-posed, full recovery of the scaling factors is not possible. In other words, the optimized scaling factors represent one possible solution, as in such ill-posed problems uniqueness of the solution is not achieved. Figure (4-4) shows the “false state” or initial state where there is no perturbation in emissions, “true” state or pseudo observations and concentrations after inversion. During the inversion, the scaling factors are optimized such that the calculated state is closer to observations.

Scaling factors applied to emissions to build observation state

Min (25, 40) = 0.25, Max (112, 56) = 4.0

Optimized scaling factors after inversion

Min (25, 40) = 0.25, Max (82, 56) = 4.0

Observational cost function reduction

Figure 4-3. Identical twin experiment results, scaling factors to calculate false state (a), optimized scaling factors recovered from inversion (b), rate of cost function reduction during the iterative inversion process (c)

Initial state of NO₂ concentrations

Pseudo observations of NO₂ concentrations

Min (134, 39) = 0.000, Max (135, 74) = 0.069

Min (143, 20) = 0.000, Max (135, 74) = 0.076

Optimized state of NO₂ concentrations

Min (109, 101) = 0.000, Max (135, 74) = 0.076

Figure 4-4. NO₂ concentrations in identical twin experiment initial (a), observations (b) and optimized case (c)

4.4.2 Global weighting factor

The first step for emission inversion is to define an appropriate global weighting factor, or μ in Eq. (4-8). Apart from considering the observational and background error covariance matrices, it is essential to have a balance weighted observational and background term in Eq. (4-8). In order to choose a realistic number for μ , a number of simulations for one day is performed to find the best value. Figures (4-5) and (4-6) shows the reduction in cost function for different μ values of 1, 25, 50, 75, 100, 150, 200, 250, 400 and 500 with 15 iterations for each case. The results of simulations for two cases (case G and case OG) show that the range of suitable and reliable value for μ is 75-100. A μ of 100 is chosen for inverse modeling as the optimal reduction in cost function for observation term occurs at such μ .

Figure 4-5. Percentage of observational cost function reduction with different global weighting factor (Case-G)

Figure 4-6. Percentage of observational cost function reduction with different global weighting factor (Case-OG)

4.4.3 Emission inversion

After defining the global weighting factor, the daily inversion was performed using different sets of observations. In the case of ground observations, emissions were inverted with the same scaling factors for the day and during 24 hours. In the case of inverting OMI observations, only the observations between 13:00 and 14:00 pm local time (OMI overpass time) were present and still the assimilation window was considered as 24 hours. A similar set up was used when both observations were combined.

4.4.3.1 Daily inversions:

Ground observations (Case G)

In this section, the results of daily inversions of emissions using only ground-based network observations (AQS) are presented. Table (4-1) presents the statistical results of each day of the month. Figure 4-5 shows the average of scaling factors for the 31 days in July 2007. As NO_2 is a fairly short-lived pollutant, the structure of inverted sources remains highly local, and only major sources appear in map of inverted scaling factors. In the case of ground observations, scaling factors range from 0.41 to 3.25 for the month of July. Mean bias and RMSE is reduced significantly after inversion for every day.

Table 4-1. Statistical parameters regarding daily inversions using ground-based network observations

Day	Mean Bias (ppb)		RMSE (ppb)	
	Before Inversion	After Inversion	Before Inversion	After Inversion
July 1	5.34	3.94	9.32	5.49
July 2	5.83	3.68	8.11	5.03
July 3	6.12	4.06	10.68	6.12
July 4	5.91	3.74	10.13	4.96
July 5	6.02	4.13	9.92	5.08
July 6	5.86	3.22	8.17	4.85
July 7	4.76	3.08	7.33	4.06
July 8	5.37	4.14	8.14	4.73
July 9	6.08	4.43	7.52	5.89
July 10	5.71	3.86	10.86	6.17
July 11	5.45	3.69	8.81	5.76
July 12	5.72	3.67	7.39	4.65
July 13	5.64	3.55	6.94	5.15
July 14	5.83	3.52	7.39	6.1
July 15	5.18	3.41	8.25	5.26
July 16	5.23	3.12	9.31	6.43
July 17	5.27	3.65	7.08	5.75
July 18	5.45	3.21	8.49	5.09
July 19	5.61	3.24	6.73	4.18
July 20	5.95	4.27	6.92	3.23
July 21	6.22	4.84	7.67	5.03
July 22	5.18	3.91	6.72	4.12
July 23	4.95	4.25	8.39	6.31
July 24	6.44	3.83	8.13	5.28
July 25	5.91	3.15	7.35	4.16
July 26	5.65	3.72	5.89	3.32
July 27	6.14	4.49	8.21	5.86
July 28	5.29	3.68	7.59	4.75
July 29	4.87	3.04	6.63	4.37
July 30	5.54	2.47	7.28	5.81
July 31	5.49	3.31	9.1	4.92

Figure 4-7. The mean value of daily scaling factors from inversion of ground observations only (Case-G)

Figure 4-8. Relationship between scaling factors and emissions (a) and scaling factors and their standard deviation (b) (Case-G)

Figure (4-8) provides information on the statistical behavior of scaling factors during 31 days. The grids with largest emissions are more reliable and remain fairly unchanged or slightly decreased, while large scaling factors exist for some of the smaller sources. Furthermore, it appears that the day-to-day variability in scaling factors is proportional to the magnitude of the factor.

Satellite observations (Case O)

In this section the results of daily inversions of emissions using only OMI observations are presented. Table (4-2) presents the statistical results of each day of the month. Figure 4-7 shows the average of scaling factors for the 31 days in July 2007.

Table 4-2. Statistical parameters regarding daily inversions using OMI observations

Day	Bias (1×10^{15})		RMSE (1×10^{15})	
	Before Inversion	After Inversion	Before Inversion	After Inversion
July 1	0.82	0.59	1.43	0.88
July 2	0.96	0.61	1.22	0.79
July 3	0.61	0.48	1.79	0.97
July 4	0.52	0.37	1.44	0.92
July 5	0.63	0.41	1.31	0.85
July 6	0.68	0.47	1.17	0.59
July 7	0.81	0.53	1.61	0.62
July 8	0.66	0.42	1.49	0.71
July 9	0.78	0.55	1.91	0.61
July 10	0.62	0.47	1.36	0.73
July 11	0.61	0.52	1.27	0.78
July 12	0.54	0.36	1.33	0.57
July 13	0.62	0.45	1.29	0.62
July 14	0.61	0.37	1.55	0.79
July 15	0.78	0.53	1.59	0.83
July 16	0.52	0.31	1.48	0.83
July 17	0.67	0.44	1.51	0.88
July 18	0.62	0.39	1.58	0.91
July 19	0.74	0.52	1.26	0.81
July 20	0.67	0.46	1.31	0.79
July 21	0.56	0.39	1.71	0.82
July 22	0.79	0.31	1.33	0.65
July 23	0.63	0.37	1.12	0.77
July 24	0.69	0.29	1.37	0.59
July 25	0.71	0.42	1.76	0.93
July 26	0.56	0.37	1.23	0.76
July 27	0.58	0.35	1.12	0.68
July 28	0.63	0.48	1.31	0.53
July 29	0.86	0.51	1.59	0.85
July 30	0.52	0.28	1.26	0.66
July 31	0.64	0.38	1.81	0.97

Scaling factors from inversion of OMI observations only (Avg)

Min (130, 68) = 0.45, Max (40, 37) = 2.72

Figure 4-9. Mean of scaling factors from inversion of OMI observations (Case-O)

As it can be noticed in Figure (4-9), OMI has a wider spatial coverage and the behavior of scaling factors is different from the previous simulation. In this case, there are significantly more scaling factors that increase the emissions (Figure 4-10). The statistical results show that the inversion was successful every day. Unlike the previous case (G), majority of scaling factors are greater than one; however, the common result similar to case (G) is that majority of grid cells with high emissions are effected by smaller than one scaling factors.

Figure 4-10. Relationship between scaling factors and emissions (a) and scaling factors and their standard deviation (b) (Case-O)

Combined ground and satellite observations (Case OG)

Combining ground and satellite observations is the final set of simulations for the daily inversions set. Figure (4-11) shows the average of the scaling factors.

Figure 4-11. Mean of scaling factors from inversion of OMI observations (Case-OG)

Table 4-3. Statistical parameters regarding daily inversions for ground observations (Case-OG)

Day	Mean Bias (ppb)		RMSE (ppb)	
	Before Inversion	After Inversion	Before Inversion	After Inversion
July 1	5.34	2.86	9.32	4.31
July 2	5.83	2.64	8.11	5.29
July 3	6.12	4.11	10.68	4.14
July 4	5.91	3.66	10.13	5.38
July 5	6.02	4.02	9.92	4.77
July 6	5.86	3.13	8.17	3.92
July 7	4.76	3.06	7.33	3.93
July 8	5.37	4.16	8.14	3.71
July 9	6.08	4.47	7.52	4.91
July 10	5.71	3.81	10.86	5.93
July 11	5.45	3.63	8.81	5.32
July 12	5.72	3.68	7.39	4.63
July 13	5.64	3.57	6.94	5.13
July 14	5.83	3.54	7.39	4.81
July 15	5.18	3.32	8.25	5.39
July 16	5.23	3.05	9.31	6.72
July 17	5.27	3.62	7.08	5.39
July 18	5.45	3.25	8.49	5.26
July 19	5.61	3.21	6.73	3.22
July 20	5.95	4.22	6.92	4.85
July 21	6.22	4.81	7.67	4.47
July 22	5.18	3.92	6.72	3.34
July 23	4.95	3.76	8.39	5.32
July 24	6.44	4.28	8.13	4.72
July 25	5.91	4.11	7.35	4.38
July 26	5.65	3.49	5.89	3.14
July 27	6.14	4.12	8.21	5.33
July 28	5.29	3.37	7.59	4.68
July 29	4.87	3.18	6.63	4.41
July 30	5.54	3.49	7.28	5.72
July 31	5.49	3.61	9.1	6.55

Table 4-4. Statistical parameters regarding daily inversions for OMI observations (Case-OG)

Day	Bias (1×10^{15})		RMSE (1×10^{15})	
	Before Inversion	After Inversion	Before Inversion	After Inversion
July 1	0.82	0.53	1.43	0.92
July 2	0.96	0.68	1.22	0.83
July 3	0.61	0.49	1.79	0.97
July 4	0.52	0.39	1.44	0.97
July 5	0.63	0.44	1.31	0.85
July 6	0.68	0.49	1.17	0.62
July 7	0.81	0.48	1.61	0.55
July 8	0.66	0.37	1.49	0.69
July 9	0.78	0.51	1.91	0.58
July 10	0.62	0.49	1.36	0.72
July 11	0.61	0.49	1.27	0.78
July 12	0.54	0.35	1.33	0.57
July 13	0.62	0.49	1.29	0.63
July 14	0.61	0.42	1.55	0.79
July 15	0.78	0.47	1.59	0.81
July 16	0.52	0.23	1.48	0.54
July 17	0.67	0.42	1.51	0.84
July 18	0.62	0.45	1.58	0.96
July 19	0.74	0.57	1.26	0.81
July 20	0.67	0.49	1.31	0.80
July 21	0.56	0.39	1.71	0.81
July 22	0.79	0.42	1.33	0.78
July 23	0.63	0.31	1.12	0.76
July 24	0.69	0.36	1.37	0.62
July 25	0.71	0.53	1.76	0.95
July 26	0.56	0.38	1.23	0.75
July 27	0.58	0.39	1.12	0.68
July 28	0.63	0.35	1.31	0.49
July 29	0.86	0.64	1.59	0.88
July 30	0.52	0.34	1.26	0.61
July 31	0.64	0.44	1.81	1.02

Figure 4-12. Relationship between scaling factors and emissions (a) and scaling factors and their standard deviation (b) (Case-O)

In general because of the wider range that OMI can observe, OMI provides observations from vast areas of rural and suburban while the AQS networks are most likely located close to urban areas. Looking at scaling factors obtained from inversion of any observations, there are more decreasing scaling factors (i.e., scaling factors smaller than one) at urban areas than increasing.

The joint inversion (GO) shows a clear pattern of decreasing scaling factors for larger sources (Figure 4-12). The behavior of scaling factors is different for each day and location. Mostly at urban locations, scaling factors tend to decrease the emissions. The trends of scaling factors for a few days at Austin, Texas are shown by Figure (4-13). Scaling factors from inverting emissions using OMI observations are totally different from the set up when ground observations are inverted. While at almost all the days scaling factors from ground observations are less than one, scaling factors estimated using OMI observations are greater than one. However, as it is shown in Figure (4-14) in another location (Los Angeles), scaling factors from inverting ground observations show a similar behavior to scaling factors from inverting OMI observations, in a sense that they both decrease the emissions. In general, there is an observation for each grid cell at each day, however, as the location of the satellite sensors are not fixed, there is a possibility that during filtering a number of observations with large pixels were omitted. For example, during July 2007 after filtering the pixels, observations for 15 days over Los Angeles could be recovered. For all other days, observations had large pixel sizes and were filtered out in this study. Here the scaling factors are presented for 15 days as OMI has 15 overpasses at these two locations.

Figure 4-13. Scaling factors of inversion set ups for Austin,Texas (a) and Los Angele (b), California

Figure 4-14 shows the performance of emission inversions when emissions were inverted using ground observations only and inversion for ground combined by OMI observations for every day during the simulation period. Overall, inclusion of OMI into the inversion process seems to have a positive effect on the quality of inversion, particularly for Austin, TX.

Figure 4-14. Comparison of concentrations from inverted emissions, Austin (a), Los Angeles (b)

Table 4-5 shows the percentage of scaling factors in each set up. The scaling factors that were very close to one were filtered out. 55% of scaling factors related to ground level observations inversion were less than 1. This further explains why concentrations were slightly overestimated. Inverting OMI observations resulted in larger emissions or 76% of scaling factors being greater than one, indicating the underestimated emissions. Notably, third set up resulted in a reduction in the percentage of scaling factors greater than one from 76% from the OMI only observations to 64% when both observations were combined.

Table 4-5. Percentage of distribution of scaling factors at different layers

	Scaling factors	Ground	OMI	Ground and OMI
Layer 1	>1	45	76	64
	<1	55	24	36
All Layers	>1	-	61	72
	<1	-	39	28

For all three set ups mentioned, the simulation period was two days but the inversion was applied for the first day in order to study the effect of the inversion on the future day.

4.4.3.2 Monthly inversion

The 3 set ups discussed so far considered daily variability of emissions for the whole month of July 2007. Accordingly, the scaling factors were calculated on a daily basis. From the daily information, monthly scaling factors, scaling factors for weekends and weekdays can be generated. Back to the concept of the inversion, the adjoint equation is driven by the distance between the simulated and observed output at the times and the locations of the observations.

The temporal and spatial distribution of the adjoint variables exemplifies the area of the influence on the overall objective cost function. Figure 4-15 shows the average of each set up.

Table 4-6. Percentage of change in total emissions for different inversion set ups

	Ground			OMI			Ground and OMI		
	Weekday	Weekend	Total	Weekday	Weekend	Total	Weekday	weekend	Total
Layer 1	3.6	2	3.5	18.5	15.2	17.3	14.4	9.3	13.6
All Layers	-	-	-	19.2	16.3	18.7	15.4	7.6	14.4

Figure 4-15. Weekday (a) and weekend (b) scaling factors from inversion of OMI and ground observations

During weekdays, the magnitude of the scaling factors is higher compared that during weekends (Figure 4-15). This is due to the fact that the concentrations of NO_x decreases from 6:00 a.m. to 3:00 p.m. (Blanchard et al, 2008) during weekends compared to weekdays which also covers the overpass time of OMI. Also, weekend emissions are more closely related to mobile NO_x emissions from heavy-duty vehicles, while for weekdays light-duty vehicles may play a more significant role. As such, weekday-weekend contrast in inverted scaling factors may indicate a distinction in the errors associated with mobile emission sectors.

Figure 4-16. Comparison of scaling factors of two different set ups as monthly scaling factors; monthly assimilation window (a) and average monthly (b)

For the monthly average scaling factors, the change in emissions can be calculated by accounting for the change in emissions during weekend/weekday and the whole month. In the case of ground observations, the change in emissions is not substantial. The change in total emissions by applying ground emission scaling factors is 3.5%. For the OMI observations (Case-O) the

change in total emissions is almost 5 times larger than that of the inversion with ground observations (Case-O). This is due to higher scaling factors under the Case-O and the fact that scaling factors were applied to all layers of the CMAQ under the Case-O, while they were applied only to the first layer under the Case-G. In the third case (Case-OG) where both sets of observations are used, the change in total emission is less than that of OMI observation case (Case-O). The reason is that combining ground observations with satellite observations results in a more balanced set of scaling factors. The change in total emissions under this simulation is an 8.5% increase.

The last set up which was the monthly inversion with assimilation window of 31 days showed a similar spatial pattern in scaling factors with a smaller range (0.782 to 1.339) compared to the monthly average scaling factors (0.429 to 1.791). Monthly assimilation window appears to provide a more stable inversion result; however, daily inversion set up provides a measure of variability that can be used as a proxy for uncertainty in posterior emissions.

4.5. Summary and conclusion

In this chapter, the variational inverse modeling approach was applied for the recovery of NO_x emissions from different sets of observations with different assimilation windows. Most recent version of OMI observations and AQS ground network observations are employed on a North American domain for emission inversion using CMAQ and CMAQ-ADJ. Optimizing location dependent emission scaling factors reduced the bias and error between model and both set of observations.

In order to study the daily variability of the scaling factors three different frameworks in emission inversion were demonstrate. The range of scaling factors is different depending on the type of observations. The test cases show that the emission inversion results are also sensitive to the type of observations. The reduction of emission for all set of inversions over the Ohio Valley and Washington and District of Colombia agrees with the results of Chai et al. (2009). In addition to providing a measure of variability, daily emission inversion offered the advantage of averaging the scaling factors for monthly, weekend, and weekday series. The weekend average scaling factors were smaller than weekday scaling factors.

There are two factors that can lead to better inversion results; firstly, defining a more accurate error covariance matrix which has a key role in inversion, and secondly, applying hourly scaling factors for emissions. Future research can investigate the daily and time dependent emission scaling factors as well as the non-diagonal background covariance matrix to improve background term.

CHAPTER FIVE:

Large scale inverse modeling of NO_x emissions and their time dependent profiles

5.1 Introduction

Inverse modeling techniques are efficient methods that can use satellite-derived information to estimate uncertain input parameters (e.g., emissions). Emission inversion studies have been carried out for pollutant species such as CO and NO_x (Muller and Stavrakou, 2005; Chai et al., 2009), black carbon (Hakami et al., 2005), mercury (Pan et al., 2007), PM_{2.5} precursor (Henze et al., 2008), methane (Bergamaschi et al., 2007) and CO₂ (Nassar et al., 2011). In earlier applications, ground-based observations were used to constrain the sources and sinks of chemical species, but in the last two decades, satellites have provided a wide range of valuable observations that can be used in air quality studies. In the case of NO₂, satellite observation has a history of almost two decades with the earliest observations from Global Ozone Monitoring Experiment (GOME-1) [Burrows et al., 1999] and Scanning Image Absorption spectrometer for Atmospheric CHartographY (SCIAMACY) [Bovensmann et al., 1999]. One of the most recent

satellites currently providing daily observations for NO_2 is the Ozone Monitoring Instrument (OMI, from 2004) (Lavett et al., 2006b).

Most of the recent studies applying inversion of NO_x emissions have sought better agreement between observations and model output through time-independent emission inversion: the difference between an average of one or two months of observations and the average modeled output is minimized to estimate one scaling factor (i.e., a factor by which emissions are under/overestimated) per episode and per grid cell. In other words, only one scaling factor is applied to all days and all hours of the day at each location. Tang (2013) reported that in the case of using time-independent factors separately, the estimated scaling factor will differ based on the observation set, finding that inverting the emissions using satellite observation resulted in a scaling factor greater than one while using ground observation resulted in a scaling factor value less than one. Resler (2010) presented a version of a time-dependent 4D-Var system to optimize the diurnal time profiles of NO_x emissions in Europe. In this case, for each hour at each location a scaling factor is applied to the emissions to explain the behavior of emissions at each hour. They reported that the cost function reduction under hourly optimization is more than that of daily optimization and concluded that hourly optimization leads to improvement in identifying emissions and better forecasts.

The goal of inverse modeling is to decrease the difference between the model output, which is associated with much uncertainty, mainly from emissions, and the observations by appropriate emissions estimations. The difference between model input and output is known as the cost function and has two main terms: (i) observation term; and, (ii) background term. One of the main challenges in the process of inverse modeling is to properly build these two terms comparable to each other. To improve the reliability of inverse modeling, each error term is

weighted by the error covariance matrix. The background error covariance matrix normalizes the magnitude of the background errors and their correlations, and its construction is a key step to achieve reliable outcomes.

There are several methods to construct the background covariance matrices. Observational method is one of the earliest methods suggested by Hollingsworth (1986) for construction of covariance matrix. Under this method, the error is extracted from the difference between observations and background. Another alternative is to estimate the background covariance error from several model outputs simultaneously. This method is referred to as the National Meteorological Center (NMC) (Parrish et al., 1992). Under this method, error covariance is calculated by substituting the model errors with the difference between different forecasts. Constantinescu et al. (2007a) suggested autoregressive models, based on perturbing the model itself, as being more efficient than the NMC method.

Chai et al. (2007) applied the NMC method successfully in 4D-Var data assimilation using different sets of observations (Airnow, aircrafts) to improve the predicted ozone distribution. In a recent study using a non-diagonal error covariance matrix, Singh et al. (2011) suggested the use of an exponential function and proposed an efficient way to calculate the background error covariance matrix using a one-dimensional correlation matrix. In order to define the correlation length for the error correlation matrix, Singh et al. (2011) suggested a simple adjoint sensitivity analysis to detect the number of grid cells in each direction across which the errors are correlated and found that performance of the data assimilation is highly sensitive to the correlation length.

The objective of this chapter is to adjust NO_x emissions using time-dependent scaling factors on an hourly basis. Unlike the study presented in Chapter 4, the goal here is to perturb emissions of

each hour separately from other hours at each location. Considering that the background error correlations decay exponentially in space, a non-diagonal error covariance matrix for the 4D-Var inversion system was constructed to achieve this objective.

5.2 Methodology

The process of finding the most probable and optimized emissions can be treated as a generalization of the 4D-variational (4D-Var) inverse modeling approach. The discrepancy between model predictions and observations at all simulation time steps together with the departure of the initial state from the background state are measured by the 4D-Var cost function which is to be minimized. In order to calculate the cost function, three ingredients are needed: (i) set of observations with corresponding covariance matrix; (ii) a chemical transport model that evolves an initial state (e.g., concentrations) at the initial time to future state values; and, (iii) a background state which represents the best estimate of the true state before any comparison with observations. In most cases, an observation operator is also needed to map the state space onto the observation state for comparison between the model's prediction and actual observations. For example, in the case of using satellite observations, the satellites do not directly measure the concentrations, but they measure column densities, which can become comparable with concentrations through an observation operator.

The mathematical form of the 4D-Var cost function in general is:

$$J(\mathbf{x}_0) = \frac{1}{2\mu} (\mathbf{x}_0 - \mathbf{x}_0^b) \mathbf{B}^{-1} (\mathbf{x}_0 - \mathbf{x}_0^b)^T + \frac{1}{2} \sum_{i=1}^N (\mathcal{H}(\mathbf{x}_i) - \mathbf{y}_i^{\text{obs}}) \mathbf{R}^{-1} (\mathcal{H}(\mathbf{x}_i) - \mathbf{y}_i^{\text{obs}})^T \quad (5-1)$$

where $J(\mathbf{x}_0)$ is the cost function, μ is the global weighting factor, $\mathbf{x}_0$ is the initial state, $\mathbf{x}_0^b$ is the background state, $\mathbf{B}$ is the background error covariance matrix, $\mathcal{H}(\mathbf{x}_i)$ is the model output which

is mapped into the observation state, $\mathbf{y}_i^{\text{obs}}$ is the observation vector and $\mathbf{R}$ is the observation error covariance matrix. In fact, the background term is weighted by the inverse of the background covariance matrix, and the observational term is weighted by the inverse of the observation covariance matrix.

5.2.1 Chemical transport model and its adjoint

This work applied the Community Multi-scale Air Quality Model (CMAQ) which is a regional chemical transport model (Byun and Schere, 2006); this research applied the gas-phase variant of CMAQ version 4.5.1 with the SPARC-99 (Carter et al. 2000) chemical mechanism in parallel mode. CMAQ is widely used for air quality studies and solves the atmospheric diffusion equation:

$$\frac{\partial C_i}{\partial t} = -\mathbf{u} \cdot \nabla C_i + \frac{1}{\rho} \nabla \cdot (\rho \mathbf{K} \nabla C_i) + R_i + \alpha E_i \quad (5-3)$$

where C_i is the mixing ratio of species i , $\mathbf{u}$ is the vector wind field, $\mathbf{K}$ is the diffusivity tensor, ρ is the air density, R_i is the net chemical reaction rate for the species and E_i represents elevated emissions. α is the emission scaling factor which is equal to one for the baseline forward CMAQ. When it comes to emission inversion, optimization process aims to estimate scaling factors by minimizing the distance between model output and observations. This study aims to estimate hourly profiles of emissions and thus the scaling factor (α) is defined as:

$$\alpha(i, j, l, t) \quad (5-4)$$

where the first three indices indicate the location of the grid cell in the modeling domain: i is the column, j is the row, l is the layer, and the last index t is time.

CMAQ receives input parameters from the following two models:

- (i) Meteorological model: WRF (Weather Research and Forecasting) 3.1 (Skamarock et al. 2005) provides meteorological input parameters. WRF output is mapped to a CMAQ-compatible modeling domain (i.e., 148 columns, 112 rows and 13 layers) by the MCIP model.
- (ii) Emission model: SMOKE version 2.4 (Sparse Matrix Operator Kernel Emission) (CEP 2007) is utilized to generate the emission data from all the different sources influencing the study region such as stationary sources, mobile sources, biogenic sources, etc. The modeling domain covers the continental United States, Southern Canada and Northern Mexico with the horizontal grid resolution of $36 \text{ Km} \times 36 \text{ Km}$ and 13 vertical layers.

The 4D-Var inverse modeling framework needs the gradients of the cost function with respect to the model inputs. The adjoint of CMAQ (CMAQ-ADJ) used in this study was developed for the gas phase in CMAQ version 4.5.1 (Hakami et al., 2007).

5.2.2 Observations

Observations used in this study include combined ground-based observations of NO_2 with column observations of NO_2 from satellite.

5.2.2.1 Ground NO_2 measurements

Hourly in situ measurements of NO_2 were obtained from the U.S. EPA Air Quality System (AQS). AQS stations use molybdenum converters heated to $300\text{-}400^\circ\text{C}$ to reduce NO_2 to NO with subsequent analysis by chemiluminescence to measure NO_x (Steinbacher et al., 2007); NO_2 is then calculated as the difference between NO and NO_x . Previous studies have shown that during the conversion process from NO_2 to NO other oxidized nitrogen compounds such as

peroxyacetyl nitrate (PAN), nitric acid (HNO₃) and other NO_z species are also reduced to NO, possibly resulting in an overestimation of the NO₂ concentration.

To reduce this uncertainty in AQS data, the correction factor suggested by Lamsal et al. (2008) is applied before inverse modeling:

$$CF = \frac{NO_2}{NO_2 + \sum AN + (0.95 \text{ PAN}) + (0.35 \text{ HNO}_3)} \quad (5 - 5)$$

where in Eq. (5-5), $\sum AN$ is the sum of all alkyl nitrates (e.g. isopropyl nitrate, n-propyl nitrate, isobutyl nitrate, isoprene nitrates). In order to calculate the correction factors CMAQ was run in forward mode for 36 days starting from June 26th 2007. Figure 5-1 shows the difference before and after applying correction factors. These correction factors are applied before deploying the AQS data in inversion simulations.

Figure 5-1. NO₂ concentrations before (a) and after (b) applying correction factors for the observations of 10 ground stations

5.2.2.2 OMI measurements

The Dutch-Finnish OMI, launched on 15 July 2004 aboard NASA's EOS Aura satellite, is a nadir-viewing UV-Vis spectrometer that measures solar backscattered irradiance. This instrument operates at an altitude of 705 km and crosses the equator approximately around 13:40 local time (Levelt et al., 2006b). OMI has a local spatial resolution of $13 \times 24 \text{ Km}^2$ at nadir

viewpoint (Boersma et al., 2007). The DOMINO product (DP-version 2) is available from Tropospheric Emission Monitoring Internet Service (TEMIS).

Satellite observations were screened before being used in the assimilation process: first, images with more than 50% cloud coverage were not used; second, pixels larger than 36 km were not considered. To avoid a biased distribution for column densities, the negative values were included in the analysis.

The vertical column density is defined as below (Lamsal et al. 2010):

$$DP_{CMAQ} = \frac{(S - S_{strop}) \sum_i C_i^{CMAQ}}{AMF \sum_i A_i C_i^{CMAQ}} \quad (5 - 6)$$

where S is the slant column density from the first step, S_{strop} is the stratospheric contribution of the slant column, AMF is the air mass factor, A_i is averaging kernel at layer i , and C is a priori trace gas profile.

5.2.3 Background error covariance matrix

The goal of inverse modeling is to integrate observations and models in order to obtain simulation results closer to reality. The information provided by observations can be employed to adjust the initial conditions and model parameters. One of the factors affecting the quality of inversion results is the available knowledge on background errors. As mentioned above, background covariance matrices are assumed to be diagonal in most chemical data assimilation and inverse modeling applications. However, a few studies have employed non-diagonal

covariance matrices. The diagonal terms of the error covariance matrix contain variances and the off-diagonal terms are the correlation between two observations or parameters (i.e., depending on the application). A non-diagonal background error covariance matrix allows the information from a parameter in a grid cell to spread out on the domain and contribute to corrections in neighboring locations.

In this study, I assumed that the background error correlations decay exponentially in space, based on the covariance localization technique used in ensemble Kalman filtering (Ott et al., 2004). Singh et al. (2011) suggested the idea of using the exponential decay function as a successful approach to construct a background error covariance matrix for global chemical data assimilation.

For a uniform domain with grid points (x_i, y_j) where $i = 1, \dots, n_x$ and $j = 1, \dots, n_y$, the one dimensional correlation between errors at two locations (x_i, y_k) and (x_j, y_k) is modeled as (Singh et al., 2011) :

$$(C_x^k)_{ij} = \text{corr}((x_i, y_k), (x_j, y_k)) = e^{-\frac{\text{dist}((x_i, y_k), (x_j, y_k))^2}{l_x^2}} \quad (5 - 7)$$

$$i, j = 1, \dots, n_x; \quad k = 1, \dots, n_y$$

where l_x is the correlation distance in the x direction. The same function can be modeled for the y direction of the domain.

$$(C_y^k)_{ij} = \text{corr}((x_k, y_i), (x_k, y_j)) = e^{-\frac{\text{dist}((x_k, y_i), (x_k, y_j))^2}{l_y^2}} \quad (5 - 8)$$

$$i, j = 1, \dots, n_y; \quad k = 1, \dots, n_x$$

The full background error correlation matrix, which accounts for both x and y correlations, is constructed as:

$$\mathbf{C} = \left(\mathbf{C}_x \otimes \mathbb{I}_{n_y \times n_y} \right) \cdot \left(\mathbb{I}_{n_x \times n_x} \otimes \mathbf{C}_y \right) \quad (5 - 9)$$

where $\otimes$ is a Kronecker product symbol. More details about this method can be found at Singh et al. (2011). The size of the tensor product matrix produced is very large. The diagonal values for this tensor are equal to unity. The main part is to define a proper value for correlation distance values at x (l_x) and y (l_y) directions.

The full covariance matrix then becomes:

$$\mathbf{B} = \boldsymbol{\sigma} \cdot \mathbf{C} \cdot \boldsymbol{\sigma} \quad (5 - 10)$$

where $\boldsymbol{\sigma}$ is a diagonal matrix with all standard deviations at all grid points.

The inverse of the non-diagonal covariance matrix is:

$$\mathbf{B}^{-1} = \boldsymbol{\sigma}^{-1} \cdot \left(\mathbb{I}_{n_x \times n_x} \otimes \mathbf{C}_y^{-1} \right) \cdot \left(\mathbf{C}_x^{-1} \otimes \mathbb{I}_{n_y \times n_y} \right) \cdot \boldsymbol{\sigma}^{-1} \quad (5 - 11)$$

5.2.4 Experiment set-up

The emission inversion set-up in this study is based on the CMAQ 4D-Var system. The distance between the available observation and model output is calculated in the cost function. The optimal solution for the emission scaling factors is obtained by minimizing the cost function. The sensitivity of the cost function with respect to the emissions is essential to solve the minimization problem efficiently. To invert the emissions, the scaling factors for emissions are used:

$$E_{\text{opt}}^{\text{NO}_x}(x_i, y_i, l_i, t) = \alpha_{i,j,l,t} E_B^{\text{NO}_x}(x_i, y_i, l_i, t) \quad (5 - 12)$$

where $E_{\text{opt}}^{\text{NO}_x}$ is optimized NO_x emission after inversion, $E_{\text{B}}^{\text{NO}_x}$ is initial NO_x emission and α is the hourly scaling factor. The scaling factors are time dependent and for each x_i, y_i they are applied as hourly coefficients to emissions at each location. In the base case, scaling factors are equal to unity at all locations.

The cost function is defined as:

$$J = \frac{1}{2} (\mathbf{VCD}_{\text{CMAQ}} - \mathbf{VCD}_{\text{OMI}}) \mathbf{R}^{-1} (\mathbf{VCD}_{\text{CMAQ}} - \mathbf{VCD}_{\text{OMI}})^T + \frac{1}{2\mu} (1 - \alpha) \mathbf{B}^{-1} (1 - \alpha)^T + \frac{1}{2} (\mathbf{C}_{\text{CMAQ}} - \mathbf{C}_{\text{Obs}}) \mathbf{O}^{-1} (\mathbf{C}_{\text{CMAQ}} - \mathbf{C}_{\text{Obs}})^T \quad (5 - 13)$$

$\mathbf{R}$, $\mathbf{B}$ and $\mathbf{O}$ are error covariance matrices for OMI observations, background and ground observations, respectively. $\mathbf{VCD}_{\text{CMAQ}}$ is the vector of the total number of molecules of NO_2 simulated by CMAQ in a vertical layer of a similar OMI observation. The OMI averaging kernels are applied to the column density of CMAQ. $\mathbf{C}_{\text{CMAQ}}$ is the NO_2 concentration as calculated by CMAQ and $\mathbf{C}_{\text{Obs}}$ is the observation gathered from the AQS networks. In the case of emissions, the variances are generated based on the difference between the optimized emissions from the chapter 4 of this thesis and daily emissions and the correlations are calculated based on the exponential decay function. The diagonal covariance matrix for OMI observations are extracted from DOMINO product. A large-scale bound-constrained limited-memory Quasi-Newton code L-BFGS-B (Zhu et al., 1997) is used for the inversion. The upper and lower bounds of scaling factors during inversion were assigned as 4 and 0.25, respectively. For all inversion set-ups, the assimilation window was considered to be 24 hours and the number of iterations was set to 15.

5.3 Results & Discussion

5.3.1 Correlation length

A 4D-Var inverse modeling system was set up to constrain NO_x emissions and estimate their time dependent profiles. The first step before inversion was to find an appropriate correlation length for the experiment. The correlation length is an important parameter and impacts the inversion process when the covariance matrix is not diagonal. The correlation length varies by the resolution of the grid cells of the domain, meteorological parameters such as wind speed and wind direction, vertical layers of the domain, and the characteristics of the species such as their lifetime, reactivity, etc.

In the current study, the correlation length of NO_x emissions was calculated using several forward model runs. In order to identify the grid cells where the errors are correlated, two different runs are performed: (i) first, the model was run in base mode to calculate the concentrations; (ii) second, NO_x emissions were perturbed at specific sample points. The differences in NO_x concentrations under the two runs were then used to determine the number of the grid cells influenced by that perturbation in the vicinity of cells with perturbed emissions.

Separate forward runs perturbed the NO_x emissions of randomly selected grid cells by 1%, 5% and 10%. The results are shown in Figure 5-2.

Figure 5-2. Grid cells with perturbed emissions (a), and corresponding difference between NO_x concentrations before and after perturbation; 1% (b), 5% (c), 10% (d)

The difference between NO_x concentrations was averaged over a time window of 24 hours for all the perturbed points. Figure 5-2 shows the number of neighbor grid cells affected by the perturbation in one grid cell in each direction. Nitrogen oxides have a lifetime of a few hours near the surface and more in the higher troposphere. Considering the results from Figure 5-2 and the lifetime of NO_x I chose to construct the error covariance matrix related to ground observations (O) using $l_x = l_y = 36$ km in each direction, shown in Figure 5-3.

		0.018		
		0.367		
0.018	0.367	1	0.367	0.018
		0.367		
		0.018		

Figure 5-3. Error correlations for a grid cell when correlation length is 36 km

5.3.2 Emission inversion

After construction of the background error covariance matrix, the daily inversion was performed using both OMI and AQS observations. Emissions were inverted for the whole day and unlike the previous chapter under which one scaling factor was used for all hours at each grid cell, this study estimates hourly scaling factors.

An example of emission inversion at a polluted area (Los Angeles) is presented for a weekday (July 2nd 2007) in Fig 5-4.

Figure 5-4. Comparison of NOx concentrations (a) before and after optimization and related scaling factors (b) in Los Angeles; weekday (July 2nd 2007).

Emission inversion improves the agreement between the model output and observations. For example, for a specific day (2 July 2007), the bias for ground observations was reduced from 6.1 to 2.2 and the root mean square error (RMSE) of satellite and CMAQ column densities was reduced from 7.2×10^{14} to 2.8×10^{14} .

5.3.3 Variability

In the work presented in this chapter, emissions were inverted for the month of July 2007 (31 days). In all locations, the estimated hourly scaling factors imposed on emissions lead to optimized model states closer to observations. Scaling factors are very dependent on the situation of the day, the first guess from the model, and the observations. As emissions, observations, meteorology and model predictions vary from day to day, a long term inversion is essential to study the behavior of scaling factors. The following figures show an example of scaling factors in different locations in different days. Figure 5-5 shows the different scaling factors for Los Angeles, Kansas City and Washington during a weekend and a weekday, demonstrating that inventory information is more greatly overestimated during weekends. The hourly profiles of scaling factors also differ for each day and applying them to the emissions improves accuracy of the model's calculated concentrations. Figure 5-6 shows the change in the emissions before (base) and after (optimized) applying scaling factors.

Figure 5-5. Weekend and weekdays scaling factors L.A., Kansas City and Washington

Figure 5-6. Comparison of before and after applying scaling factors on emissions (LA, Kansas City and New York)

Figures 5-7 and 5-8 show the effect of hourly scaling factors on emissions. Results of inversion at Austin show scaling factors greater than one as well as scaling factors less than one. On average, weekend and weekday scaling factors for hours between 0:00 to 10:00 in the morning and 18:00 to 24:00 in the afternoon are less than 1. On the other hand, scaling factors are greater than one from noon through the afternoon, which can be because of the OMI overpass. During the simulation period at this grid cell, estimated scaling factors resulted in 19% increase of NO_x emissions. Weekdays have larger emissions compared to weekends and the rate of the increase in emissions for weekends and weekdays are 20% and 18%, respectively. Studying the results at Los Angeles show that the estimated scaling factors are less than one every day during the simulation period, inversion resulted in more than 40% emission reductions at that grid cell. The overall results for the domain during month of July entail an increase of 12.1% in NO_x emissions.

Figure 5-7. Comparison of before and after applying scaling factors on emissions (Average of 31 days, Austin, Texas), average of weekdays (a), average of weekends (b) and average of all 31 days (c)

Figure 5-8. Comparison of before and after applying scaling factors on emissions (Average of 31 days, Los Angeles, California), average of weekdays (a), average of weekends (b) and average of all 31 days (c)

Comparing the standard deviation of the scaling factors on the entire domain, Figures 5-9 through 5-11, show that the variability is not very high at rural places where only OMI observations were available; most of the high standard deviations can be seen in the Eastern United States. The mean of the scaling factors are also higher in the East compared to those in the West.

Min (1, 1) = 0.000, Max (84, 63) = 3.548

Min (14, 75) = 5.144E-5, Max (122, 51) = 1.604E0

Figure 5-9. Mean (a) and standard deviation (b) of the scaling factors at the time of OMI overpass during weekends

Figure 5-10. Mean (a) and standard deviation (b) of the scaling factors at the time of OMI overpass during weekdays

Min (1, 1) = 0.000, Max (84, 63) = 3.613

Min (74, 99) = 7.153E-7, Max (122, 77) = 1.875E0

Figure 5-11. Mean (a) and standard deviation (b) of the scaling factors at the time of OMI overpass during July

The total number of estimated scaling factors for each grid cell varies across the domain because different number of observations is available at each grid cell and its surrounding grid cells. To study the behavior of the scaling factors, all of the estimated scaling factors at each grid cell are sorted and then 16th percentile, median, and 84th percentile are presented in Figures 5-12 to provide an approximation of $\pm 1\sigma$ uncertainty.

Min (1, 1) = 0.000, Max (84, 63) = 3.268

Min (1, 1) = 0.000, Max (32, 51) = 1.875

Figure 5-12. Median (a) and 1- σ inter-percentile variability (half of the difference between 16 and 84 percentiles) (b) of scaling factors on the entire domain during July

The rate of cost function reduction was different for each day. In most cases, the cost function for weekends was smaller than weekdays and the rate of the cost function reduction for weekends were smaller than weekdays. This is possibly due to the fact that concentration values and VCDs during weekends are smaller than those during weekdays.

Figure 5-13. Percentage of reduction in observational cost function during July 2007

Figure 5-13 shows the percentage of reduction in observational cost function during the simulation period. Last, study of the variability of the scaling factors over the entire domain, where emissions are overestimated or underestimated, reveals additional patterns. I studied the eastern and western parts of the domain (Western and Eastern U.S.) and observed whether the scaling factors served to increase or decrease the emissions. In general, during weekdays and weekends the majority of scaling factors are greater than unity in the western part of the domain, while the eastern part of the domain has more cases of reduced emissions.

Table 5-1. Percentage of scaling factors greater and smaller than one in West and East

Day	West		East	
	Scaling factor <1	Scaling factor >1	Scaling factor <1	Scaling factor >1
Weekday	27%	73%	37%	63%
Weekend	30%	70%	41%	59%

The outcome of this chapter is time profiles of scaling factors that update the hourly emissions to improve the model outputs. Figure (5-14) shows the average of the emissions at two different hours, one in the early morning and one at the time of the OMI overpass. Changes in emissions near the time of overpass are more substantial, mainly owing to the availability of observations for the hours close to the overpass. For the same reason, scaling factors for these hours that are close to the overpass time should be considered more reliable than hours around which observations are scarce. Figures (5-15) and (5-16) show the error before and after inversion at two different times. As these figures show, the difference (error) between the model output and observations is reduced significantly as the result of inversion.

Min (1, 1) = 0.000, Max (23, 46) = 36.835

Min (1, 1) = 0.000, Max (130, 68) = 77.345

Figure 5-14. Average of the NO_x emissions at 5 a.m. (a) and 2 p.m. (b)

Figure 5-15. Difference between concentrations before (a) and after (b) inversion at 10 a.m.

Figure 5-16. Difference between concentrations before (a) and after (b) inversion at 3 p.m.

Figure 5-17. Difference between RMSE before and after inversion for all days

5.3.4 Forecasting

The inversion was set to optimize daily emissions during July 2007. Correspondingly, the scaling factors were calculated on a daily basis. The daily information can be used to calculate monthly scaling factors or separate weekday and weekend scaling factors. These weekend/weekday or even monthly scaling factors can be employed for forecasting purposes as they update the emissions leading to better model outputs (e.g., concentrations). Weekday scaling factors were found to be higher than weekend, similar to the results presented by Chai et al. (2009), scaling factors due to the uneven temporal distribution of anthropogenic activities over various days of the week and this shows the same trend with the results from non-temporal inversion (Figure 4-15). In order to evaluate the performance of the optimized scaling factors, I applied the average weekday/weekend scaling factors from this chapter on 4 random days on August 2007 that were not in the simulation period. The resulting reduction in cost function was significant for all four days, but greater reduction was obtained for weekdays compared to weekdays. Using the scaling

factors for Aug 8th and 20th (weekday) reduced the cost function (observational) by 32% and 27% and for Aug 4th and 12th (weekend) by 21% and 18%, respectively. I conclude that optimized scaling factors were stable enough to be used outside of the inversion episode for which they were developed. This carryover of the improvement in the model's predictive skills also points to potential systematic bias in gridded emissions that carry from one month over to another.

5.4 Conclusions

This chapter presents a variational inverse modeling approach to estimate NO_x emissions using OMI satellite observations and AQS ground network observations. The emissions were inverted based on observations from both OMI and AQS to have more accurate predictions. Satellite observations were derived using TM4 and corrections were made to remove the effect of TM4.

Chemiluminescence, which is used to estimate the concentration of the nitrogen dioxide, has limitations when it comes to distinguishing NO and other nitrogen oxides and this limitation makes measuring nitrogen dioxide inaccurate. The correction factors were applied to ground observations to reduce this known error. Applying the correction factors led to better agreement between observation and concentrations calculated by the model. In this work, a non-diagonal background covariance matrix was constructed using an exponential decay function and emission inversion was performed on the North America domain using CMAQ and CMAQ-ADJ for 31 days.

The challenging part of the construction of the non-diagonal background is the correlation length which was estimated by several ‘brute force’ attempts. Considering the atmospheric lifetime of nitrogen oxides, the correlation length is assumed to be equal to the size of one grid cell. The value of correlation length for different species and/or different domains can be different than that used here.

Time-dependent scaling factors estimated through the emission inversion resulted in significant cost function reduction, and reduction in bias and error between observations and model predictions. These scaling factors vary from one day to another day and from one location to another. Our results show that the cost functions during weekdays are slightly higher than those

during weekends and the reduction in cost functions show the same trend. Overall, the scaling factors are more likely to increase the emissions. The outcome of this research can be useful for environmental decision making and policy. In the presence of more accurate information on the state of the atmosphere, decision makers can establish more effective environmental guidelines.

CHAPTER SIX:

Inversion of lightning NO emissions

6.1 Introduction

Incomplete knowledge of emission sources is one the main sources of uncertainty in air quality studies. The outcomes of such studies are highly dependent on the quality of emission inputs calculated by emission models providing temporal and spatial gridded emissions from emission inventory information. However, inventory information related to the time, location, and the magnitude of natural events such as wildfire and lightning is hard to obtain. Lightning is the largest natural source of nitrogen oxide, and more importantly it releases NO in the middle and upper troposphere where few other sources are available. By comparison, inventory information for anthropogenic sources of nitrogen oxides, such as vehicle traffic, are often easier to quantify and process.

Nitrogen oxides have a short lifetime at the lower troposphere as they react with other species and radicals. Lower tropospheric removal reactions include NO₂ reacting with OH and NO reacting with ozone (Seinfeld and Pandis, 2006). In the upper troposphere, however,

temperature decreases and reaction rates slow, yielding longer atmospheric lifetimes for NO_x (i.e., 1-2 weeks), in turn contributing to greater ozone production in the middle and upper troposphere.

The NO_x contribution from lightning is significant yet often neglected in air quality studies (Zhang et al., 2003, Singh et al., 2007). For instance, inclusion of lightning NO_x increases the amount of NO_x in the middle and upper troposphere by 60%-90% and in turn contributes to increased middle and upper tropospheric ozone by 20%-30% during summertime conditions (Zhang et al., 2003). In another study, Allen et al. (2010) reported that lightning NO (hereafter abbreviated as LNO) contributed 60%-90% of NO_x and 15%-35% of O_3 in the upper troposphere over North America in summer.

The atmospheric source of LNO is associated with large uncertainty. Traditionally, most studies have employed the bottom-up approach to estimate the lightning NO_x source. Levy et al (1996) introduced top-down approach to estimate the lightning source in the range of 2-6 Tg (N) per year. They compared the airborne observations of NO_x with the global model simulations. Beirle et al. (2004) used observations from GOME satellite to estimate the production rate of lightning over Australia and reported the range of 0.8-14 Tg N per year for lightning NO_x . They found a correlation between flash counts from the Lightning Imaging Sensor (LIS) and monthly mean of GOME observations. Boersma et al. (2005) found a correlation between lightning flash rates and tropospheric vertical column densities and thus were able to estimate the rate of LNO production as 1.1-6.4 Tg (N) per year. Martin et al. (2007) estimated the mean annual lightning production to 4-8 Tg N per year combining tropospheric columns from SCIAMACHY, tropospheric ozone columns from OMI and MLS and tropospheric HNO_3 with GEOS-CHEM. The production range

of global lightning is estimated to be in the range of 2 to 8 Tg (N) per year (Schumann and Huntrieser, 2007).

Detecting lightning NO from satellites is a challenging task that has received increased attention recently. It is not easy to distinguish the portion of lightning in total column of nitrogen oxides as surface sources and free atmosphere sources contribute to the total column of nitrogen oxides. Most importantly, lightning often coincides with locations and times with large cloud coverage when satellites are least useful. Most regions with considerable amount of lightning also contain other sources, however, altitude is the factor that separates lightning sources from other sources. Skamarock et al. (2003) suggested that lightning NO_x are dominant above 7 km. Pickering et al. (2014) developed an algorithm to retrieve the lightning component of NO_2 and produce a column density of lightning. Similar to tropospheric column density calculation, they proposed that the column density of lightning can be calculated by subtracting stratospheric column and background column from the total column density. Here, the background is an estimate of NO_2 column density produced by all sources except lightning. This estimation is a mean of NO_2 tropospheric column density. The main disadvantage of this method is uncertainty in estimation of the background column; this is evidenced by the method yielding negative column density for a number of days, which Pickering et al. (2014) suggested was due to error in estimation of the background column over populated cities. Despite this uncertainty, LNO emissions have been added to regional air quality models in some recent studies (Koo et al., 2010; Allen et al., 2010), requiring information on lightning flash rates from ground-based networks so as to estimate NO emissions.

In this study, a filtering method was proposed to find OMI satellite observations that are not affected by anthropogenic emissions. These observations, on a relative basis, contain more

information regarding LNO emissions and are used for improved estimation of NO₂ emission sources. At the same time, for the purpose of comparison with inversion results, the LNO emissions provided by Allen et al. (2012) were added to our chemistry transport model. The goal of this chapter is to find the appropriate emission sources of LNO that leads to better agreement between the OMI observations and model outputs.

6.2 Methodology

The goal of this section is to estimate LNO emissions that reduce the distance between observations and model output. The CTM that is used in this research is CMAQ and observations are taken from the OMI satellite.

6.2.1 LNO

LNO is triggered by lightning flash causing a sudden heat increase of up to 30,000 K in the middle and upper troposphere, which is sufficient thermal energy to split, N₂ and O₂ molecules and form NO. The chemical processes governing the production of NO in the presence of lightning are presented in Eq. (6-1) through Eq. (6-3) and is referred to as the Zeldovich mechanism (Zeldovich et al., 1947).

However, these equations do not fully represent the physical processes and the magnitude of the production of NO. In order to add the LNO to the chemistry transport models, the flash rates of

lightning require conversion to equivalent emissions. The LNO emissions can be estimated by the frequency of the flash (LF), the energy that flash contains (E) and the rate of the production of NO per unit energy (P) (Allen et al., 2010) as shown in Eq. (6-4).

$$\text{LNO} = k \times \text{LF} \times E \times P \quad (6-4)$$

where k is a conversion factor equal to the molecular weight of Nitrogen divided by Avogadro's number.

The energy of the flash in NO_x production is dependent on the energy per stroke and path length. There are two main flash types, intracloud (IC) and cloud-to-ground (CG). In an early study, Price et al. (1997) suggested that the flash production rate of IC is approximately 10% of that of CG. However in recent studies it has been shown that IC flashes are as energetic as CG and Eq. (6-4) is derived upon the latter assumption.

Eq. (6-4) converts the monthly averaged flashes that are observed by the National Lightning Detection Network (NLDN) (Orville et al., 2002) to NO emissions. The main assumption here is based on the previous studies, the mean production value of mid latitude lightning flash NO of 500 moles and tropical lightning flash NO mean production of 250 moles (Allen et al., 2010) and thus estimates spatial and temporal emissions. The LNO calculated from the Eq. (6-4) is the aggregated LNO over all layers and is reported as column number. To estimate layer-specific LNO, vertical profiles of LNO emissions can be used (Koshak et al., 2010). For vertical redistribution of LNO emissions, it is assumed that the rates of emissions are proportional to pressure convolved by the segment altitude of flashes from the North Alabama Lightning Mapping Array (Koshak et al., 2010). More information regarding the assumptions and methods used for calculating vertical profile of LNO can be found in Allen et al. (2010).

Figure 6-1 shows the temporal (local hour) and vertical distributions of LNO. On average the highest LNO is estimated at late afternoon and upper layers (between layer 6 and 10 out of 13 layers).

Figure 6-1. The percentage of NO observed emissions distributed based on local time (a) and CMAQ layer (b)

6.2.2 CMAQ

The Community Multiscale Air Quality (CMAQ) version 4.5.1 (Byun and Schere 2006) model was applied over a North America domain with a 36 km grid cell resolution and 13 vertical layers for a period of one month (July 2007). The adjoint version of the gas-phase CMAQ was used to calculate the gradients. CMAQ input parameters are gathered from meteorological and emission processing models: emission inputs are generated by the Sparse Matrix Operator Kernel Emissions (SMOKE) model (CEP 2007); meteorological inputs are calculated by the Weather Research and Forecasting (WRF) model (Skamarock et al. 2005).

6.2.3 OMI observations

Observations were provided by the Ozone monitoring Instrument (OMI) aboard the NASA Aura satellite (launched July 2004) which measures direct and back-scattered sunlight in the ultraviolet and visible range. The nadir resolution of the satellite is 13 km $\times$ 24 km with daily coverage. DOMINO product version 2 was used in the research. As mentioned in Chapter 5, the effect of TM4 on this product is reduced by calculating the air mass factors using CMAQ and applying that on the difference between total and stratospheric slant column densities. As lightning is more likely to occur in the presence of clouds, only observations with cloud coverage between 10%- 50% were used. Moreover, observations with pixel size larger than 36 km were not considered. Any negative values were included in the analysis so as to avoid a biased distribution of column densities.

6.2.4 Filtering

Measuring NO_x produced by lightning is not an easy task because anthropogenic sources are dominant. To overcome this problem, first the grid cells without dominant anthropogenic

emissions were identified. Then, the contribution of anthropogenic sources to the total of emissions is defined as:

$$\beta = \frac{VCD_{NO_2,anthropogenic,101\%} - VCD_{NO_2,100\%}}{VCD_{NO_2,all,101\%} - VCD_{NO_2,100\%}} \quad (6 - 5)$$

where, $VCD_{NO_2,100\%}$ is the vertical column density when there is no perturbations, $VCD_{NO_2,anthropogenic,101\%}$ is the vertical column density when all the anthropogenic emission sources are perturbed by 1%, $VCD_{NO_2,all,101\%}$ is the vertical column density when all the emission sources are perturbed by 1%. This has been done for all layers of each grid cell.

The β value represents the ratio of anthropogenic emissions to the emissions from all sources in a grid cell and thus indicates whether or not a grid cell is dominated by anthropogenic emissions. In this study, the grid cells were selected in which the value of β was less than 0.1, indicating that anthropogenic emissions were minor relative to natural emissions. Figure 6-2 shows the overlap between the grid cells containing LNO emissions and OMI observations in July 2007.

6.2. Grid cells containing OMI observations and lightning during July 2007

6.2.5 Inversion set-up

CMAQ 4D-Var was used for emission inversion of LNO sources. The difference of inversion in this chapter with research presented in Chapters 4 and 5 is that instead of using scaling factors to perturb emissions, the absolute value of the emissions are inverted. This change enables us to invert sources with zero emissions. Similar to the methods presented in prior chapters, the difference between the observation and output of the model is defined in the cost function, shown in Eq. (6-6).

$$J = \frac{1}{2} (\text{VCD}_{\text{CMAQ}} - \text{VCD}_{\text{OMI}}) \mathbf{R}^{-1} (\text{VCD}_{\text{CMAQ}} - \text{VCD}_{\text{OMI}})^T + \frac{1}{2\mu} (\mathbf{E} - \mathbf{E}_a) \mathbf{B}^{-1} (\mathbf{E} - \mathbf{E}_a)^T \quad (6-6)$$

$\mathbf{R}$ and $\mathbf{B}$ are error covariance matrices for OMI observations and emissions. The variances are calculated from the updated emission from the time dependent inversions in previous chapters.

VCD_{CMAQ} is the total number of molecules of NO_2 simulated by CMAQ in vertical layers similar to the vertical layers of the OMI observation. The OMI averaging kernels are applied to column density of CMAQ. VCD_{OMI} is calculated from the difference between total and stratospheric slant column densities. E is the emission as an output from SMOKE and E_a is the optimized value from the inversion process. As a small number of observations were inverted, and based on multiple experiments, μ was considered to be equal to 1000. The optimized emissions are hourly and location-specific emissions. The lower and upper bounds for emissions were set to 0 and 300 moles per second, respectively and ten 4D-Var iterations were performed.

6.3 Results

6.3.1 Identical twin experiment

The identical twin experiment method was used to evaluate the performance of the inversion with absolute emissions. In this method, hypothetical observations are generated by the model with specific emissions with the assumption of zero modeling error. This hypothetical state, termed the false state, has observations generated from the average of ten days of NO emissions (July 1st to 10th). The goal is to see how fast inversion can lead the results from the initial model state, termed the true state, to the false state. Figure 6-3 shows the true state, false state and optimized state and demonstrates that the performance of inversion using absolute emissions versus inverting using adjusted emissions is nearly the same. In this case, cost function is reduced by more than 90 percent after 6 iterations.

Figure 6-3. Results of identical twin experiment for absolute emissions, true state (a), false state (b) and optimized (c)

6.3.2 Emission inversion

Emissions of NO are inverted at the places where observations were available. Pickering (2014) proposed that the lightning flash rates occurring 6 hours prior to OMI overpass contribute to the LNO derived from that grid cell. I made the same assumption and simulation window was considered to be six hours prior to OMI overpass.

Figure 6-4. Comparison between Vertical Column densities before (a) and after inversion (b), OMI VCDs (c) for July 1st 2007

Figures (6-4) and (6-6) show the result for an example of emission inversion for two overpasses. Emissions are optimized during the inversion process to bring the VCDs calculated by CMAQ closer to the OMI observations. Figure (6-5) shows the scatter plot of comparison before and after inversion for the same day.

Figure 6-5. Improvement of column densities, before inversion (a) after inversion (b), July 1st 2007

Figure 6-6. Comparison between Vertical Column densities before (a) and after inversion (b), OMI VCDs (c) for July 26th 2007

After filtering the grid cells not dominated by anthropogenic emissions ($\beta < 0.1$), most of the remaining grid cells had either small amount of emissions or no emissions (i.e., observed mostly over the ocean). Figure (6-7) shows the result of inversion of NO emissions for two grid cells, one over the ocean and one over the land.

Figure 6-7. Vertical change of the emissions after inversion, over the ocean (a), over the land (b)

NO emission inversions were done for every day in the month of July 2007. Table (6-1) shows the result of inversion for July 2007. This experiment provided more information for the grid cells that contained a small amount of emissions or zero emissions. Inclusion of LNO in

emissions increased NO_x emissions in some grid cells by as much as 300 times compared with the emissions solely calculated by SMOKE. For example, in a grid cell over the rural southern U.S. the total emissions increased from 0.15 to 22.85 moles per second after adding LNO emissions; after emission inversion, this total amount decreased to a still significant 16.78 moles per second. Over the ocean where there was no initial NO emissions, 3.73 moles per second of LNO-related emissions were added to one grid cell; comparing the columns with observation at this grid cell ended up increasing this total amount to 8.82 moles per second. On average, after inversion, LNO emissions increased over the ocean and decreased (or did not change) over the land during the modeling period. Figure 6-8 shows the total column amount of LNO before and after inversion. Please note that these values are the average of 6 hours. Table 6-1 shows the statistical results from NO inversion for July 2007.

Table 6-1. Statistical results of inversion for each day

Day	Mean Bias (10^{15} #/area)		RMSE (10^{15} #/area)	
	Before Inversion	After Inversion	Before Inversion	After Inversion
July 1	-0.73	-0.16	4.95	0.76
July 2	-1.02	-0.23	5.11	0.75
July 3	-0.81	-0.21	3.12	0.52
July 4	-0.97	-0.36	5.13	1.38
July 5	-0.82	-0.43	4.92	1.17
July 6	-1.13	-0.39	4.17	0.92
July 7	-0.79	-0.36	3.33	0.65
July 8	-0.84	-0.42	3.14	0.47
July 9	-0.82	-0.38	4.52	0.73
July 10	-0.98	-0.31	3.86	0.63
July 11	-1.09	-0.36	5.81	0.76
July 12	-0.96	-0.38	4.39	0.69
July 13	-0.77	-0.35	4.94	0.81
July 14	-0.89	-0.34	5.39	1.21
July 15	-0.76	-0.33	3.25	0.54
July 16	-0.83	-0.29	4.31	0.78
July 17	-0.84	-0.32	4.08	0.87
July 18	-0.95	-0.35	5.49	1.05
July 19	-1.06	-0.41	5.73	0.92
July 20	-0.87	-0.37	4.92	1.04
July 21	-1.09	-0.36	4.67	0.95
July 22	-0.96	-0.32	4.72	0.76
July 23	-0.77	-0.36	3.39	0.61
July 24	-0.84	-0.42	3.13	0.63
July 25	-0.91	-0.41	4.35	0.72
July 26	-0.88	-0.34	3.89	0.59
July 27	-1.21	-0.41	5.21	1.08
July 28	-1.19	-0.39	4.59	1.09
July 29	-0.82	-0.31	3.63	0.68
July 30	-0.82	-0.34	4.28	0.76
July 31	-0.95	-0.36	4.09	0.65

Figure 6-8.Column of emissions (6 hour average), after inversion (a) and before inversion (b)

6.3.3 Effect of LNO on ozone and nitrogen dioxide concentrations

Incorporating LNO in modeling simulations influenced both ozone and nitrogen oxide concentrations. Allen et al. (2012) used CMAQ to study the amount of LNO in eastern United States during summer 2006 and reported mean contribution of 0.5 – 2.5 ppbv in the North and 2.5 – 4.5 ppbv in the South. In this work, I found that inclusion of LNO resulted in a mean increase (after inversion) of 1.28 ppbv in O₃ on the surface layer and 1.91 ppbv in all layers comparing to the base case. However, there is a 0.027 ppbv decrease in the first layer and 0.04 ppbv in all layers in the case of comparing the inversion results with the results from the case when only NO emissions were added. The increase in NO₂ was 0.23 ppbv at the surface layer, much less than the increase in ozone, and 0.24 ppbv in all layers. A small fraction of observations and a small number of observation points and the value of LNO for the first two or three layers were very small and inverting the emissions at those places made no significant change in the first layer. A more complete picture can only be obtained by having a better spatial and temporal coverage of satellite observations, which may not be possible for lightning due to existence of clouds.

6.4 Conclusions

This study inverted the emissions of NO using OMI observations of LNO in order to improve model estimates. I proposed a method to identify and filter grid cells and observations that were dominated by anthropogenic emissions. The advantage of this filtering method is that the grid cell and the observations that are dominated by non-anthropogenic emissions are most likely influenced by natural sources.

One limitation of this method is the time of the OMI overpass, which is in the afternoon (local time) and does not provide ideal information for inversion since LNO mostly happens in the early evening. The possibility of lightning occurrence is also highly dependent on the existence of clouds, and I only considered OMI observations with 10% - 50% cloud coverage. OMI observations of less than 10% cloud cover suggest a very low probability of lightning occurrence, while greater than 50% cloud cover makes the OMI-reported observations highly uncertain.

Acknowledging these limitations, results of inversion showed significant cost function reduction and reduction in the bias and RMSE calculations. The process of filtering the grid cells also reduced the significant number of observations. Overall, observation limitations compromise the practical usefulness of our approach. Lightning happens sporadically, and filtering out locations that are influenced by anthropogenic sources, as well as those locations where satellite views are blocked by clouds (which are most likely locations for lightning), makes this problem particularly challenging.

CHAPTER SEVEN:

Summary and future work

Nitrogen oxides (NO_x) are among the precursors of secondary aerosols and ozone with adverse impact on human health. Thus, it is important to accurately quantify NO_x emission sources used for air quality modeling. However, there are usually large uncertainties associated with their estimations that are based on current emission inventories. Utilization of satellite-derived information within an advanced inverse modeling framework has proven to be an efficient tool to reduce the uncertainty associated with emission estimations.

In this research, I applied a variational inverse modeling approach to estimate NO_x emissions from different sets of observations. Most recent version of OMI observations combined with AQS ground network observations were employed on a North American domain for emission inversion using CMAQ and CMAQ-ADJ.

In order to study the variability and uncertainty in emission adjustments, I applied two approaches, non-temporal and temporal scaling factors. In the non-temporal approach, emissions at each location were adjusted by a dimensionless scaling factor. The location-specific scaling factors were estimated through an iterative optimization method to uniformly perturb the

emissions at all the hours to reduce the distance between model output and available observations. In temporal approach, location-specific emissions were optimized by an hourly vector of scaling factors, meaning that at each location, emissions scaling factors vary by the time of the day.

In non-temporal approach, three different inversion set-ups were demonstrated under which emissions were inverted using 1) only ground observations, 2) only satellite observations and 3) both ground and satellite of observations. The range of scaling factors was varied by the type of observations. The advantage of temporal inversion compared to non-temporal inversion is that the former can be used to adjust emissions on weekends, weekdays, or on a monthly basis by applying the scaling factors.

The results also indicate that emissions are overestimated more often during the weekends, while weekday emissions are mostly underestimated. In most urban regions, CMAQ tends to overestimate the ground concentrations and column densities and that can be found in some recent studies (Irie et al., 2013; Choi, 2014; Tong et al., 2015). Applying these adjustments to emissions at episodes outside of the inversion window shows marked improvement in the predictive skills of CMAQ.

The main disadvantage in the non-temporal simulations was assigning one scaling factor for all the hours at each grid cell. This has influence on optimizing the scaling factor especially when both ground observations and satellite column densities are present at that specific grid cell. Ground observations are available on an hourly basis, whereas OMI observations are available at one hour. Therefore, assigning one scaling factor to describe the change of both sets of observations may not be ideal. That is the main reason that in chapter 5 I applied hourly scaling

factors for emissions. Error covariance matrices were another limitation of the study in chapter 4 in which I assumed that error covariance matrix for both the observations and model to be diagonal.

In temporal inversion, similar to non-temporal inversion, I employed the variational inverse modeling approach and estimated NO_x emissions using OMI satellite observations and AQS ground network observations. The first step in the temporal inversion was to apply the correction factors calculated from CMAQ to ground network observations to lower the bias in observation errors. These factors were calculated by the CMAQ model, and led to a better agreement between improved observations and model outputs. The second improvement in the temporal approach was due to implementing a non-diagonal covariance matrix instead of commonly used diagonal representation. To construct a non-diagonal background covariance matrix, an exponential decay function was used. The main challenge of defining a non-diagonal background matrix was the calculation of correlation length, which was estimated by several brute force attempts.

Temporal scaling factors were estimated through a variational inverse modeling framework for July 2007 in the North America domain. Emission inversion in the case of temporal scaling factors resulted in more cost function reduction compared to the reduction under non-temporal inversion. The profile of hourly scaling factors varied from one day or location to another. The results showed that the cost functions during weekdays were slightly larger than those of weekends, but the reduction in cost functions had a similar trend for both weekends and weekdays. Based on the optimized hourly emission profiles at both east and west side of the domain, emissions were more underestimated than overestimated. However, at urban regions emissions were more overestimated. Results of applying hourly scaling factors optimized from

Chapter five on two days during weekday (Aug 8th and 20th) and weekend (4th and 12th) showed that taking the average of hourly scaling factors can be useful for forecasting the state of the model in the future.

The last part of this thesis was focused on lightning NO, the most uncertain source of nitrogen oxides in the middle and upper layers of atmosphere. The chemistry and the process of lightning NO generation is generally well understood but the physical process and magnitude of this natural source of nitrogen oxides is associated with large uncertainties. Previous studies were focused more on the surface layer than upper layers. The effect of upper layers of the atmosphere on the surface is shown to be non-negligible. In chapter 6, I proposed a method to estimate the sources of lightning NO using OMI observations covering LNO at both upper and lower atmosphere.

Anthropogenic sources of nitrogen oxides are dominant and it is not an easy task to quantify natural sources of these species in the presence of the much more dominant anthropogenic contribution. For that reason I proposed a method to filter the grid cells which are not dominated by anthropogenic sources by using brute force sensitivity analysis. The grid cells and observations not dominated by anthropogenic emissions were considered to be influenced by natural sources alone. Because lightning mostly occurs where clouds exist, only OMI observations with 10%-50% cloud coverage were considered. For observations with less than 10% coverage, the likelihood of lightning occurrence is low and for more than 50% coverage, the observations are unavailable or highly uncertain. Overall, our inversion exercise indicates that scarcity of data points that result from filtering anthropogenic source influences and appropriate cloud conditions, severely limits the ability to retrieve NO_x emissions for a large part of the domain.

There are two main limitations regarding lightning NO which is described in chapter 6. The major limitation arises from the availability of observations, high possibility of lightning occurrence early in the evening and away from OMI overpass. Another challenge in this study was the lack of data to ascertain the vertical behavior of ozone.

The results of this research are the proof of concept. The focus of this work was on the emissions however the effect of other sources in creating uncertainty in the concentrations should be considered in the future studies. In previous inverse modeling applications the error in concentrations due to chemical mechanism used in the CTM, the resolution of the model, removal processes were neglected. Recently, the effect of chemical uncertainty, based on the new developments from both field and laboratory studies, has come under investigation. It has been shown that considering more accurate removal processes in the atmosphere (Stavrakou et al., 2013) or more accurate photolysis rates (Tang et al., 2015) in the CTM can reduce the uncertainty in the concentrations calculated by the CTM.

Influence of the model grid cell resolution on the concentrations and vertical column densities, is another factor which was not studied in this research. Change in the grid resolution leads to change in the vertical column densities. Yamaji et al. (2015) reported that the change in the CMAQ vertical column densities are greater when the resolution changes from 40 Km to 20 Km (bias ~ 36%) comparing to the change from 80 Km to 40 Km (bias ~ 18%), and therefore finer resolution does not necessary mean better agreement between OMI and CMAQ vertical column densities.

In summary, the main focus of this research was to study the behavior of emissions. The aim was to update the emissions in order to calculate more accurate concentrations. Unlike most of the

available studies were the adjustment to emissions took place in a short period of time (few days), here longer period of time (a month) was considered. As mentioned before, most of the studies in this field were to apply a non-temporal scaling factor to emissions; first the non-temporal scaling factors were applied on different set-ups (Chapter 4). The next step was to invert emissions using temporal scaling factors (Chapter 5). The bias in the ground observations were reduced by applying correction factors and to have a better understanding of the effect of the emissions from one grid cell to the adjacent grid cells the error covariance matrix was extended from diagonal to non-diagonal. The focus of the last part of this research was on one uncertain source which is not always considered in air quality model employed in this study. A filtering method was proposed to detect the lightning sources and corrections to those sources were applied through inverse modeling of emissions. This proposed method for lightning was the first step to detect that neglected source and should be investigated more in the future studies in the presence of more useful observations (Chapter 6).

The following research topics are suggested for further investigation:

- **Long-term emission inversion**

In this thesis all the inversion set-ups were for one month (July 2007), however, it will be valuable to study the behavior of scaling factors in the long term. Generating scaling factors for a longer period than a month, season, and year can be used for potential corrections in available emission information and in air quality forecasting.

- **Background error covariance matrix**

The results of data assimilation or inverse modeling applications are sensitive to the structure of error covariance matrices. In the presented work, the assumption was that the

correlation length is equal to the size of one grid cell. Grid-specific correlation lengths can be estimated and used for emission inversion in a future study.

- **The effect of nitrogen oxides on PM and Ozone**

The focus of this research was only on emission inversion of NO_x. Any Change in emissions of NO_x affects the formation of particulate matter and ozone. As we use the gas-phase version of CMAQ, to study the effect of NO_x emissions on aerosols, a full version of the adjoint model (under development) is required. Emission inversion can be applied not only to reduce the distance between model outputs and observations of nitrogen dioxide, but also it can account for errors in concentrations of ozone and particulate matter in a multi-pollutant inversion platform.

- **Lightning NO**

Lightning NO inversion can be better informed by including satellite ozone observations, as well as higher resolution products as they become available.

REFERENCES

Allen, D. J., K. E. Pickering, B. Duncan, and M. Damon, “Impact of lightning NO emissions on North American photochemistry as determined using the Global Modeling Initiative (GMI) model”, *Journal of Geophysical Research*, 115, D22301, doi: 10.1029/2010JD014062, 2012.

Allen, D. J., Pickering, K. E., Pinder, R. W., Henderson, B. H., Appel, K. W., and Prados, A., “Impact of lightning-NO on eastern United States photochemistry during the summer of 2006 as determined using the CMAQ model”, *Atmos. Chem. Phys.*, 12, 1737–1758, doi:10.5194/acp-12-1737-2012, 2012.

Anderson, J. Sheinbaum, and K. Haines, “Data assimilation in ocean models”. *Rev. Prog. Phys.*, 59, 1209–1266, 1996.

Beirle, S., Platt, U., Wenig, M., and Wagner, T, “NO_x production by lightning estimated with GOME”, *Adv. Space Res.*, 34(4), 793– 797, 2004a.

Beirle, S., Huntrieser, H., and Wagner, T, “Direct satellite observation of lightning-produced NO_x”, *Atmos. Chem. Phys.*, 10, 10965–10986, doi:10.5194/acp-10-10965-2010, 2010.

Bergamaschi, P., R. Hein, M. Heinmann, and P. Crutzen, “Inverse modeling of the global CO cycle: 1. Inversion of CO mixing ratio”, *J. Geophys. Res.*, 105(D2), 1909– 1927, 2000.

Bergamaschi, P, “Satellite cartography of atmospheric methane from SCIAMACHY on board ENVISAT: 2. Evaluation based on inverse model simulations”, *J. Geophys. Res.*, 112, D02304, doi:10.1029/2006JD007268, 2007.

Blanchard, C. L.; Tanenbaum, S. J, “Differences between weekday and weekend air pollutant levels in southern California”. *J. Air Waste Manage. Assoc.* 53, 816–828, 2003.

Brookner, E. “Tracking and Kalman Filtering Made Easy”, Wiley Interscience, 1998.

Boersma, K. F., Eskes, H. J., and Brinksma, E. J. “Error analysis for tropospheric NO₂ retrieval from space”, *J. Geophys. Res.*, 109, D04311, doi:10.1029/2003JD003962, 2004.

Boersma, K. F., H. J. Eskes, E. W. Meijer, and H. M. Kelder, “Estimates of lightning NO_x production from GOME satellite observations”, *Atmos. Chem. Phys.*, 5, 2311–2331, 2005.

Boersma, K. F., Eskes, H. J., Veefkind, J. P., Brinksma, E. J., van der A, R. J., Sneep, M., van den Oord, G. H. J., Levelt, P. F., Stammes, P., Gleason, J. F., and Bucsela, E. J. “Near-real time retrieval of tropospheric NO₂ from OMI”, *Atmos. Chem. Phys.*, 7, 2103–2118, doi:10.5194/acp-7-2103-2007, 2007.

Boersma, K. F., Dirksen, R. J. Veefkind, J. P. Eskes, H. J. and van der A, R. J. “Dutch OMI NO₂ (DOMINO) data product, HE5 data file user manual”, 2009.

Boersma, K. F., Eskes, H. J., Dirksen, R. J., Van der, R. J. A., Veefkind, J. P., Stammes, V., Huijnen, Q. L., Kleipool, M., Sneep, J., Claas, J., Leitao, A., Richter, Zhou, Y., D. Brunner. “An improved retrieval of tropospheric NO₂ columns from the Ozone Monitoring Instrument” Atmos. Meas. Tech., 4, 1905-1928, 2011.

Bouttier, F., Courtier, P. “Data assimilation concepts and methods, Training Course Lecture Series”, ECMWF, 1999.

Bovensmann, H., J. P. Burrows, M. Buchwitz, J. Frerick, V. V. Rozanov, K. V. Chance, and A. P. H. Goede, “SCIAMACHY: Mission objectives and measurement modes”, J. Atmos. Sci., 56, 127– 150, 1999.

Burrows, J. P., ‘The Global Ozone Monitoring Experiment GOME: Mission concept and first scientific results”, J. Atmos. Sci., 56, 151– 175, 1999.

Byun, D. and Schere, K. L.: “Review of the governing equations, computational algorithms, and other components of the models-3 community multiscale air quality (CMAQ) modeling system”, Appl. Mech. Rev., 59, 51–77, 2006.

Carter, W. P. L.: “Implementation of the saprc-99 chemical mechanism into the models-3 framework”, United States Environmental Protection Agency, 2000.

CCME, “GUIDANCE DOCUMENT ON AIR ZONE MANAGEMENT”, Canadian Council of Ministers of the Environment, 2012.

CEP, ‘Sparse Matrix Operator Kernel Emissions Modeling System, (SMOKE) v2.4 User’s Manual’. Environment , Carolina Environmental program, Chapel hill, NC, 2007.

Carmichael, G. R, Sandu, A, Chai, T, Daescu, D., Constantinescu, E.M. and Tang, Y.: "Predicting Air Quality: Improvements through Advanced Methods to Integrate Models and Measurements". Journal of Computational Physics (Special issue on predicting weather, climate and extreme events), Vol. 227, Issue 7, p. 3540-3571, 2008.

Chai, T., Carmichael, G. R., Sandu, A., Tang, Y., and Daescu, D. N.: Chemical data assimilation of transport and chemical evolution over the Pacific (TRACE-P) aircraft measurements, J. Geophys. Res., 111, D02301, doi: 10.1029/2005JD005883, 2006.

Chai, T., Carmichael, G. R., Tang, Y., Sandu, A., Hardesty, M., Pilewskie, P., Whitlow, S., Browell, E. V., Avery, M. A., Thouret, V., Nedelec, P., Merrill, J. T. and Thomson, A. M., “Four dimensional data assimilation experiments with ICARTT (International Consortium for Atmospheric Transport and Transformation) ozone measurements”, Journal of Geophysical Research, Vol. 112, D12S15, doi:10.1029/2006JD007763, 2007.

Chai, T., Carmichael, G. R., Tang, Y., Sandu, A., Heckel, A., Richter, A., and Burrows, J. P.: “Regional NO_x emission inversion through a four-dimensional variational approach using SCIAMACHY tropospheric NO₂ column observations”. Atmos. Environ., 43, 5046–5055, 2009.

Choi, Y. “The impact of satellite-adjusted NO_x emissions on simulated NO_x and O₃ discrepancies in the urban and outflow areas of the Pacific and Lower Middle US”, *Atmos. Chem. Phys.*, 14, 675-690, doi:10.5194/acp-14-675-2014, 2014.

Chouinard, C. Charette, Hallé, J., Gauthier, P. Morneau J. and Sarrazin, R. “The Canadian 3D-Var analysis scheme on model vertical coordinate”. 14th AMS conference on numerical weather prediction, Fort Lauderdale, 2001.

Constantinescu, E.M., Chai, T., Sandu, A. and Carmichael G.R.: “Autoregressive Models of Background Errors for Chemical Data Assimilation”. *Journal of Geophysical Research*, Vol. 112, D12309, doi:10.1029/2006JD008103, 2007.

Constantinescu, E. M., Sandu, A., Chai, T. and Carmichael, G. R. “Ensemble-based chemical data assimilation. I: general approach”, *Quarterly Journal of the Royal Meteorological Society*, Volume 133, Issue 626, p. 1229–1243, Online ISSN: 1477-870X, Print ISSN: 0035-9009, July 2007 Part A.

Committee on Scientific Accomplishments of Earth Observations from Space, National Research Council, “Earth Observations from Space: The First 50 Years of Scientific Achievements”, National Academies Press, ISBN: 0309110955, 2008.

Daley, R., “Atmospheric data analysis”, Cambridge Univ. Press. 1991.

Denby, J., “Comparison of two data assimilation methods for assessing PM10 exceedances on the European scale”, *Atmos. Environ.*, v42i30 7122-7134, 2008.

Elbern, H., Schmidt H. and Ebel, A. “Variational data assimilation for tropospheric chemistry modeling”. *J. Geophys.Res.*102-D13, 1997.

Elbern, H., Schmidt, H., Talagrand, O., Ebel, A., “4D-variational data assimilation with an adjoint air quality model for emission analysis”. *Environ.Modeling Software*. v15. 539-548, 2001.

Elben, K., Juras, P., Resler, J., “An ensemble Kalman filter for short-term forecasting of tropospheric ozone concentrations”, *Q. J. R. Meteorol. Soc.*, 131, 3313–3322, 2005.

Elbern, H., Strunk, A, Schmidt, H. Talagrand, O. “Emission rate and chemical state estimation by 4-dimensional variational inversion”, *Atmos. Chem. Phys.*, 7, 3749-3769, doi:10.5194/acp-7-3749, 2007.

Environmental Protection Agency (EPA), “Primary national ambient air quality standard for nitrogen dioxide”, *Tech. Rep.*, Environmental Protection Agency, *Federal Register*, Vol. 74, No. 134, 2009.

Errera, Q., Daerden, F., Chabrillat, S., Lambert, J. C., Lahoz, W. A., Viscardy, S., Bonjean, S., and Fonteyn, D.: “4D-Var assimilation of MIPAS chemical observations: ozone and nitrogen dioxide analyses”, *Atmos. Chem. Phys.*, 8, 6169–6187, doi: 10.5194/acp-8-6169-2008, 2008.

Eskes, H. J., “ Combined retrieval, modeling and assimilation approach to GOME NO₂”, in GOA Final Report, European Commission Fifth Framework Programme 1998– 2002, EESD-ENV-99-2, pp. 116–122, Eur. Comm., De Bilt, Netherlands, April, 2003.

Eskes, H. J. and Boersma, K. F.: “Averaging kernels for DOAS totalcolumn satellite retrievals”, *Atmos. Chem. Phys.*, 3, 1285–1291, doi:10.5194/acp-3-1285-2003, 2003.

Evenson, G., “Using the extended Kalman filter with a multilayer quasi-geostrophic ocean model”, *J. Geophys. Res.*, 97(C11), 17,905–17,924, 1992.

Evensen, G. “ Inverse methods and data assimilation in nonlinear ocean models”. *Physica D*, 77:108–129, 1994.

Fisher, M., Lary, D. J., “Lagrangian four-dimensional variational data assimilation of chemical species”, *Q.J.R. Meteorology*, 121:1681–1704, 1995.

Fioletov, V. E., C. A. McLinden, N. Krotkov, K. Yang, D. G. Loyola, P. Valks, N. Theys, M. Van Roozendaal, C. Nowlan, K. Chance, X. Liu, C. Lee, and R. V. Martin, “Application of OMI, SCIAMACHY, and GOME-2 satellite SO₂ retrievals for detection of large emission sources”, J. Geophys. Res., 118, doi: 10.1002/jgrd.50826, 2013.

Flemming, J., Reimer, E., Stern R. “Data Assimilation for CT-Modelling based on Optimum Interpolation”. 25th ITM on Air Pollution Modeling and its Application, Brussels, Belgium, 2001.

Frey, H. C., and Patil, S. R., “Identification and review of sensitivity analysis methods” Journal of Risk Analysis, Vol. 22. No. 3, 2002.

Gandin, J.L., “The Objective Analysis of Meteorological Fields, Israel Program for Scientific Translations”, Jerusalem, 1963.

Gauthier, P. C., Charette, L., Fillion Koclas, P., Laroche, S. “Implementation of a 3D variational data assimilation system at the Canadian Meteorological Centre”, Part I: the global analysis. Atmosphere -Ocean, 37, 103- 156, 1999a.

Gurney, K., Y.-H. Chen, T. Maki, S. Kawa, A. Andrews, and Z. Zhu, “Sensitivity of atmospheric CO₂ inversions to seasonal and interannual variations in fossil fuel emissions”, J. Geophys. Res., 110, D10308, doi:10.1029/2004JD005373, 2005.

Hakami, A., Odman, M. T., Russell, A. G., “High-order, direct sensitivity analysis of multidimensional air quality models”. Environ. Sci. Technol. 37, 2442-2452, 2003.

Hakami, A., Henze, D., Seinfeld, J.H., Chai, T., Tang, Y., Carmichael, G.R., Sandu, A., “Adjoint inverse modeling of black carbon during the Asian Pacific Regional Aerosol Characterization Experiment”, J. Geophys. Res. 110 (D14301), 2005.

Hakami, A., Seinfeld, J.H., Chai, T., Tang, Y., Carmichael, G.R. & Sandu, D. "Adjoint Sensitivity Analysis of Ozone Nonattainment over the Continental United States", Environmental science & technology, vol. 40, no. 12, pp. 3855-3864, 2006.

Hakami, A., Henze, D.K., Seinfeld, J.H., Singh, K., Sandu, A., Kim, S., Byun, D. & Li, Q. "The Adjoint of CMAQ", Environmental science & technology, vol. 41, no. 22, pp. 7807-7817, 2007.

Hamill, T.M., Whitaker, J., Snyder, C. “Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter”, Mon. Wea. Rev., 129, 2776–2790, 2001.

Hanna, S.R., Z. Lu, H.C. Frey, N. Wheeler, J. Vukovich, S. Arunachalam, M. Fernau, and D.A. Hansen, “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, 2001.

Hanea, R. G. M. Velders, G. J. Heemink, A. “Data assimilation of ground-level ozone in Europe with a Kalman filter and chemistry transport model”, J. Geophys. Res., 109, D10302, doi:10.1029/2003JD004283, 2004.

Heland, J., Schlager, H., Richter, A., and Burrows, J. P, “First comparison of tropospheric NO₂ column densities retrieved from GOME measurements and in situ aircraft profile measurements”, Geophys. Res. Lett., 29(20), 1983,doi:10.1029/2002GL015528, 2002.

Henze, D.K., Seinfeld, J.H., Shindell, D.T., “Inverse modeling and mapping U.S. air quality influences of inorganic PM_{2.5} precursor emissions using the adjoint of GEOS-CHEM”, Atmos. Chem. Phys. Discuss. 8, 15031–15099, 2008.

Hollingsworth, A. and LŽnnberg, P.: “The statistical structure of short-range forecast errors as determined from radiosonde data”. Part I: The wind field. Tellus, 38A, 111-136, 1986.

Houweling, S., T. Kaminski, F. Dentener, J. Lelieveld, and M. Heinmann , “Inverse modeling of methane sources and sinks using the adjoint of a global transport model”, J. Geophys. Res., 104(D21), 26,137 – 26,160, 1999.

Hu, Y.T.; Odman, M.T.; Russell, A.G. “Top-down analysis of the elemental carbon emissions inventory in the United States by inverse modeling using Community Multiscale Air Quality model with decoupled direct method (CMAQ-DDM)”, J. Geophys. Res. Atmos. 2009.

Huang, M., Bowman, K. W., Carmichael, G. R., Chai, T., Pierce, R. B., Worden, J. R., Luo, M., Pollack, I. B., Ryerson, T. B., Nowak, J. B., Neuman, J. A., Roberts, J. M., Atlas, E. L., Blake, D. R. “Changes in nitrogen oxides emissions in California during 2005–2010 indicated from top-down and bottom-up emission estimates”, *Journal of Geophysical Research*, 10.1002/2014JD022268, 2014.

Huang, M.; Bowman, K.W.; Carmichael, G.R.; Lee, M.; Chai, T.; Spak, S.N.; Henze, D.K.; Darmenov, A.S.; da Silva, A.M. “Improved western U.S. background ozone estimates via constraining nonlocal and local source contributions using Aura TES and OMI observations”. *Journal of Geophysical Research: Atmospheres* , 120, 3572, 2015.

Hudman, R.C., N.E. Moore, R.V. Martin, A.R. Russell, A.K. Mebust, L.C. Valin, and R.C. Cohen, “Steps towards a mechanistic model of global soil nitric oxide emissions: implementation and space based-constraints”, *Atm. Chem. Phys.* , 12, 7779-7795, doi:10.5194/acp-12-7779-2012, 2012.

Huijnen, V., Eskes, H. J., Poupkou, A., Elbern, H., Boersma, K. F., Foret, G., Sofiev, M., Valdebenito, A., Flemming, J., Stein, O., Gross, A., Robertson, L., D’Isidoro, M., Kioutsioukis, I., Friese, E., Amstrup, B., Bergstrom, R., Strunk, A., Vira, J., Zyryanov, D., Maurizi, A., Melas, D., Peuch, V.-H., and Zerefos, C.: “Comparison of OMI NO₂ tropospheric columns with an ensemble of global and European regional air quality models”, *Atmos. Chem. Phys.*, 10, 3273–3296, doi:10.5194/acp-10-3273-2010, 2010.

Irie, H., K. Yamaji, K. Ikeda, I. Uno, S. Itahashi, T. Ohara, and J. Kurokawa, “An evaluation of the CMAQ reproducibility of satellite tropospheric NO₂ column observations at different local times over East Asia”, *Atmospheric Chemistry and Physics Discussion*, 13, 14037-14067, May 2013.

Jacob, D.J., Logan, J.A. & Murti, P.P. "Effect of rising Asian emissions on surface ozone in the United States", *Geophysical Research Letters*, 26, 14, pp. 2175-2178, 1999.

Kalnay, E., Miyoshi, T., Yang, S., Ballabrera-Poy, J., “4D-Var or ensemble Kalman filter”, *Tellus*, 59A, 758–773, 2007.

Jacob, J.D., “Lectures on Inverse Modeling”, 2007.

Khattatov, B.V., Gille, J.C., Lyjak, L.V., Brasseur, G.P., Dvortsov, V.L., Roche, A.E. and Waters, J.W. “Assimilation of photochemically active species and a case analysis of UARS data”. *Journal of Geophysical Research* 104: doi: 10.1029/1999JD900225. issn: 0148-022, 1999.

Konovalov, I. B., Beekmann, M., Richter, A., and Burrows, J. P.: “Inverse modelling of the spatial distribution of NO_x emissions on a continental scale using satellite data”, *Atmos. Chem. Phys.*, 6, 1747–1770, 2006, <http://www.atmos-chem-phys.net/6/1747/>, 2006.

Koo, B., Chien, C.-J., Tonnesen, G., Morris, R., Johnson, J., Sakulyanontvittaya, T., Piyachaturawat, P., and Yarwood, G.: “Natural emissions for regional modeling of background ozone and particulate matter and impacts on emissions control strategies”, *Atmos. Environ.*, 44, 2372–2382., doi:10.1016/j.atmosenv.2010.02.041, 2010.

Koshak, W. J., Peterson, H. S., McCaul, E. W., and Biazar, A.: “Estimates of the lightning-NO_x profile in the vicinity of the North Alabama Lightning Mapping Array”, 21st International Lightning Detection Conference (ILDC)/Vaisala, 18–22 April 2010, Orlando, FL, United States, NASA Technical Report Document 20100021055, 2010.

Kurokawa, J., Yumimoto, K., Uno, I., and Ohara, T.: “Adjoint inverse modeling of NO_x emissions over eastern China using satellite observations of NO₂ vertical column densities”, *Atmos. Environ.*, 43, 1878–1887, 2009.

Lamsal, L. N., Martin R. V., van Donkelaar A., Steinbacher M., Celarier E. A., Bucsela E., Dunlea E. J., Pinto J. P. “Ground-level nitrogen dioxide concentrations inferred from the satellite-borne ozone monitoring instrument”. *J. Geophys. Res.* 113:D16308, doi:10.1029/2007JD009235, 2008.

Lamsal, L. N., Martin R. V., van Donkelaar A., Celarier, E. J. Bucsela, K. F. Boersma, R. Dirksen, C. Luo, and Y. Wang. “ Indirect validation of tropospheric nitrogen dioxide retrieved from the OMI satellite instrument: insight into the seasonal variation of nitrogen oxides at northern midlatitudes”, *J. Geophys. Res.*, 115, D05302, doi:10.1029/2009JD013351, 2010.

Lamsal, L.N., R.V. Martin, A. Padmanabhan, A. van Donkelaar, Q. Zhang, C.E. Sioris, K. Chance, T.P. Kurosu and M.J. Newchurch, “Application of satellite observations for timely updates to global anthropogenic NO_x emission inventories”, *Geophys. Res. Lett.*, 38, L05810, doi:10.1029/2010GL046476, 2011.

Lee C., R.V. Martin, A. van Donkelaar, H. Lee, R.R. Dickerson, J.C. Hains, N. Krotkov, A. Richter, K. Vinnikov, J. J. Schwab, “SO₂ emissions and lifetimes: Estimates from inverse modeling using in situ and global, space-based (SCIAMACHY and OMI) observations”, *J. Geophys. Res.*, 116, D06304, doi:10.1029/2010JD014758, 2011.

Lee, C.J., J.R. Brook, G.J. Evans, R.V. Martin, and C. Mihele, “Novel application of satellite and in-situ measurements to map surface-level NO₂ in the Great Lakes region”, *Atmos. Chem. Phys.*, 11, 11761-11775, doi:10.5194/acp-11-11761-2011, 2011.

Levelt, P. F., G. H. J. van den Oord, M. R. Dobber, A. Malkki, H. Visser, J. de Vries, P. Stammes, J. O. V. Lundell, and H. Saari (2006b), “The Ozone Monitoring Instrument”, *IEEE Trans. Geosci. Remote Sens.*, 44, 1093–1101.

Levy H., Moxim, W. J., and Kasibhatla, P. S.: “A global three-dimensional time-dependent lightning source of tropospheric NO_x”, *J. Geophys. Res.*, 101(D17), 22 911–22 922, doi:10.1029/96JD02341, 1996.

Lewis, J. Lakshmivarahan, S and Dhall, S. “Dynamic Data Assimilation: A Least Squares Approach”, Cambridge University Press, 2006.

Li, Z., Navon, M. “Optimality of 4D-Var and its relationship with the Kalman filter and Kalman smoother”, Quart. J. Roy.Meteor.Soc, 1999.

Lin, J.-T., McElroy, M. B., and Boersma, K. F.: “Constraint of anthropogenic NO_x emissions in China from different sectors: a new methodology using multiple satellite retrievals”, Atmos. Chem. Phys., 10, 63–78, doi:10.5194/acp-10-63-2010, 2010.

Liu, D. C., Nocedal J., “On the limited memory BFGS method for large scale optimization”, Math. Programming, 45, 503–528, 1989.

Lorenz, A.C., “A global three-dimensional multivariate statistical interpolation scheme”, Mon. Weath.Rev., Vol. 109, pp. 701-721, 1981.

National Research Council: “Models in environmental regulatory decision making”. The National Academies Press, Washington, D.C., 2007.

Martin, R. V., Jacob, D. J., Chance, K., Kurosu, T. P., Palmer, P. I., and Evans, M. J.: “Global inventory of nitrogen oxide emissions constrained by space-based observations of NO₂ columns”. J. Geophys. Res., 108, 4537, doi:10.1029/2003JD003453, 2003.

Martin, R.V., Sioris, C.E., Chance, K., Ryerson, T.B., Bertram, T.H., Wooldridge, P.J., Cohen, R.C., Neuman, J.A., Swanson, A., Flocke, F.M., “Evaluation of spacebased constraints on global nitrogen oxide emissions with regional aircraft measurements over and downwind of eastern North America” .Journal of Geophysical Research 111, D15308, 2006.

Martin, R. V., Sauvage, B., Folkins, I., Sioris, C. E., Boone, C., Bernath, P., and Ziemke, J.: “Space-based constraints on the production of nitric oxide by lightning” , J. Geophys. Res., 112, D09309, doi:10.1029/2006JD007831, 2007.

Martin, R.V., “Satellite remote sensing of surface air quality”, Atmos. Environ., 42, 7823-7843, 2008.

Mendoza-Dominguez, A. and Russell, A. G.: “Iterative inverse modeling and direct sensitivity analysis of a photochemical air quality model”, Environ. Sci. Technol., 34, 4974–4981, 2000.

Meirink, J. F. Bergamaschi, P. Krol, M. C. (2008), “Four-dimensional variational data assimilation for inverse modelling of atmospheric methane emissions: method and comparison with synthesis inversion”, Atmos. Chem. Phys., 8, 6341–6353.

Muller, J.-F., and T. Stavrakou, “Inversion of CO and NO_x emissions using the adjoint of the IMAGES model”, Atmos. Chem. Phys., 5, 1157–1186, 2005.

Napelenok, S. L., Pinder, R. W., Gilliland, A. B., and Martin, R. V.: “A method for evaluating spatially-resolved NO_x emissions using Kalman filter inversion, direct sensitivities, and spacebased NO₂ observations”, *Atmos. Chem. Phys.*, 8, 5603–5614, doi:10.5194/acp-8-5603-2008, 2008.

Nassar, R., Jones, D. B. A., Kulawik, S. S., Worden, J. R., Bowman, K. W., Andres, R. J., Suntharalingam, P., Chen, J. M., Brenninkmeijer, C. A. M., Schuck, T. J., Conway, T. J., and Worthy, D. E.: “Inverse modeling of CO₂ sources and sinks using satellite observations of CO₂ from TES and surface flask measurements”, *Atmos. Chem. Phys.*, 11, 6029–6047, doi:10.5194/acp-11-6029-2011, 2011.

NCAR, 2010, Wang, W., Bruyère, C., Duda, M., Dudhia, J., Gill, D., Lin, H. C., Michalakes, J., Rizvi, S., & Zhang, X., “Weather Research and Forecasting, ARW Version 3 Modelling System User Guide”, Mesoscale Microscale Meteorology Division, National Center for Atmospheric Research (NCAR), <http://www.mmm.ucar.edu/wrf/users/>, 2010.

Niu, T., Gong, S. L., Zhu, G. F., Liu, H. L., Hu, X. Q., Zhou, C. H., Wang, Y. Q., and Zhang “Data Assimilation of Dust Aerosol Observations for CUACE/Dust Forecasting System”, *Atmos. Chem. Phys. Discuss*, 2007.

Nocedal, J. “Updating Quasi-Newton Matrices With Limited Storage”, *Math. Comput.*, 35, 773–782, 1980.

OMI User's guide, "Ozone Monitoring Instrument (OMI) Data User's Guide", 2008.

Orville, R. E., Huffines, R. E., Burrows, G. R., Holle, W. R., and Cummins, K. L., "The North American Lightning Detection Network (NALDN)-first results: 1998–2000", *Mon. Weather Rev.*, 130, 2098–2109, 2002.

Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A.V., Kostelich, E. J., Kostelich, M., Corazza, M., Sauer, T., Kalnay, E., Patil, D. J. and Yorke, J. A: "A local ensemble Kalman Filter for Atmospheric Data Assimilation", *Tellus*, Vol. 56A, pp. 415-428, 2004.

Pan, L., Chai, T., Carmichael, G., Tang, Y., Streets, D., Woo, J., Friedli, H.R., Radke, L.F., "Top-down estimate of mercury emissions in china using four-dimensional variational data assimilation". *Atmos. Environ.* 41 (13), 2804–2819, 2007.

Parrish, D. F. and Derber, J. C.: "The national meteorological center's spectral statistical-interpolation analysis system", *Monthly Weather Review*, (120), p. 1747–1763, 1992.

Petron, G., C. Granier, B. Khatatov, V. Yudin, J.-F. Lamarque, L. Emmons, J. Gille, and D. Edwards, "Monthly CO surface sources inventory based on the 2000– 2001 MOPITT satellite data", *Geophys. Res. Lett.*, 31, L21107, doi:10.1029/2004GL02056, 2004.

Pickering, K., Allen, D., A. Ring, A. Nag, E. Buscela, “Lightning NO_x Production Per Flash based on OMI NO₂ Observations and Ground Network Lightning Data”, ILDC/ILMC, Tucson, AZ, 2014.

Polavarapu, S., “Lecture notes on atmospheric data assimilation”, 2004.

Price, C., J. Penner, and M. Prather, “NO_x from lightning 1. Global distribution based on lightning physics”, J. Geophys. Res., 102 (D5), 5929-5941, 1997.

Quélo, D., Mallet, V., and Sportisse, B.: “Inverse modeling of NO_x emissions at regional scale over northern France: preliminary investigation of the second-order sensitivity”. J. Geophys. Res., 110, D24310, doi:10.1029/2005JD006151, 2005.

Resler, J., Eben, K., Jurus, P., and Liczki, J.: “Inverse modelling of emissions and their time profiles”, Atmos. Poll. Res., 1, 288–295, doi:10.5094/APR.2010.036, 2010.

Rodenbeck, C., S. Houweling, M. Gloor, and M. Heimann, “CO₂ flux history 1982– 2001 inferred from atmospheric data using a global inversion of atmospheric transport”, Atmos. Chem. Phys., 3, 1919–1964, 2003.

Sandu, A., Daescu, D. N. Carmichael, G. R. Chai, T. “Adjoint sensitivity analysis of regional air quality models”, Journal of Computational Physics, v.204 n.1, p.222-252, 2005.

Schumann, U. and Huntrieser, H.: “The global lightning-induced nitrogen oxides source”, *Atmos. Chem. Phys.*, 7, 3823–3907, doi:10.5194/acp-7-3823-2007, 2007.

Singh, H. B., Salas, L., Herlth, D., Kolyer, R., Czech, E., Avery, M., Crawford, J. H., Pierce, R. B., Sachse, G. W., Blake, D. R., Cohen, R. C., Bertram, T. H., Perring, A., Wooldridge, P. J., Dibb, J., Huey, G., Hudman, R. C., Turquety, S., Emmons, L. K., Flocke, F., Tang, Y., Carmichael, G. R., and Horowitz, L. W.: “Reactive nitrogen distribution and partitioning in the North American troposphere and lowermost stratosphere”, *J. Geophys. Res.*, 112, D12S04, doi:10.1029/2006JD007664, 2007.

Singh, M. Jardak, A. Sandu, K. Bowman, M. Lee, and D. Jones “Construction of Non-diagonal Background Error Covariance Matrices for Global Chemical Data Assimilation” *Geoscientific model development*, v.4, p. 299, 2011.

Seinfeld, J. H. and S. N. Pandis “Atmospheric Chemistry and Physics: From Air Pollution to Climate Change”, Wiley-Interscience, ISBN: 978-0-471-72018-8, 2006.

Skachko, S., Errera, Q., Ménard, R., Christophe, Y., and Chabrillat, S.: “Comparison of the ensemble Kalman filter and 4D-Var assimilation methods using a stratospheric tracer transport model”, *Geosci. Model Dev.*, 7, 1451-1465, doi:10.5194/gmd-7-1451-2014, 2014.

Skamarock, W. C., J. E. Dye, E. Defer, M. C. Barth, J. L. Stith, B. A. Ridley, and K. Baumann
“Observational- and modeling-based budget of lightning-produced NO_x in a continental
thunderstorm”, J. Geophys. Res., 108(D10), 4305, doi: 10.1029/2002JD002163, 2003.

Skamarock, W., Klemp, J., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W., and Powers, J. G.
“A description of the advanced research WRF version 2.” (June), 2005.

Stammes, P.: “Spectral radiance modeling in the UV-Visible range, IRS2000: in: Current
problems in atmospheric radiation”, edited by: Smith, W. L. and Timofeyev, Y. J., 385–388, A.
Deepak, Hampton, Va, USA, 2001.

Stavrakou, T., Müller, J.-F., Boersma, F., Smedt, I.D., van der A, R.J., “Assessing the
distribution and growth rates of NO_x emission sources by inverting a 10-year record of NO₂
satellite columns”. Geophys. Res. Lett. 35 (L10801). doi:10.1029/ 2008GL033521, 2008.

Stavrakou, T., Müller, J.-F., Boersma, K. F., van der A, R. J., Kurokawa, J., Ohara, T., and
Zhang, Q.: “Key chemical NO_x sink uncertainties and how they influence top-down emissions of
nitrogen oxides”, Atmos. Chem. Phys., 13, 9057–9082, doi:10.5194/acp-13-9057-2013, 2013.

Steinbacher, M., Zellweger, B., Schwarzenbach, S., Bugmann, B., Buchmann, C., Ordoñez, A. S. H.
Prevot, and C. Hueglin “Nitrogen oxides measurements at rural sites in Switzerland: Bias of
conventional measurement techniques”, J. Geophys. Res., D11307, doi:10.1029/ 2006JD007971,
2007.

Strunk, A., Elbern, H., Ebel, A., “Using Satellite Observations for Air Auality Assessment with an Inverse Model System”, Numerical Methods and Applications, 2010.

Stuart, A. L., Aksoy, A., Zhang, F., and Nielsen-Gammon, J.W., “Ensemble-based data assimilation and observation of a chemical tracer in a sea breeze model”. *Atmos. Env.*, 41, 3082-3094, 2007.

Talagrand, O., and P. Courtier, “Variational assimilation of meteorological observations with the adjoint vorticity equations”, part I - theory, *Q. J. R. Meteorol. Soc.*, 113, 1311– 1328, 1987.

Tang, W., D.S. Cohan, L.N. Lamsal, X. Xiao, and W. Zhou..“Inverse modeling of Texas NO_x emissions using space-based and ground-based NO₂ observations”. *Atmospheric Chemistry and Physics*, 13, 11005-11018. doi:10.5194/acp-13-11005-2013, 2013.

Tang, W., Cohan, D. S., Pour-Biazar, A., Lamsal, L. N., White, A. T., Xiao, X., Zhou, W., Henderson, B. H., and Lash, B. F.: “Influence of satellite-derived photolysis rates and NO_x emissions on Texas ozone modeling”, *Atmos. Chem. Phys.*, 15, 1601-1619, doi:10.5194/acp-15-1601-2015, 2015.

Tong, D. Q., Lamsal, L., Pan, L., Ding, C., Kim, H., Lee, P., Chai, T., Pickering, K., Stajner, I., “Long-term NO_x trends over large cities in the United States during the great recession: Comparison of satellite retrievals, ground observations, and emission inventories”, *Atmospheric Environment*, Volume 107, Pages 70–84, 2015.

Van Aardenne, J.A., “Uncertainties in emission inventories”. Thesis (PhD). Wageningen University, The Netherlands, 2002.

Vinken, G. C. M., K. F. Boersma, A. van Donkelaar, and L. Zhang, “Constraints on ship NO_x emissions in Europe using GEOS-Chem and OMI satellite NO₂ observations”, *Atmos. Chem. Phys.*, 14, 1353–1369, 2014, doi:10.5194/acp-14-1353-2014.

Vinken, G.C.M., K.F. Boersma, J.D. Maasakkers, M. Adon, and R.V. Martin, “Worldwide biogenic soil NO_x emissions inferred from OMI NO₂ observations”, *Atmos. Chem. Phys.*, 14, 10363-10381, 2014.

U.S. EPA, “Technical assistance document for the chemiluminescence measurement of nitrogen dioxide, Tech. Rep., Environmental Monitoring and Support Laboratory”, U.S. EPA, Research Triangle Park, NC, EPA-600/4-75-003, 1975.

U.S. EPA, “Air Quality Criteria for ozone and related photochemical oxidants”, EPA/600/P-93/004q-cF, pages 3-90 to 3-102, 1996.

U.S. EPA, “Reference Method for the Determination of Nitrogen Dioxide in the Atmosphere (Chemiluminescence)”, Office of air quality planning and standards, 2002.

U.S. EPA, “AQS Fundamentals”, Environmental Protection Agency, Office of air quality planning and standards, 2006.

U.S. EPA, “AQS user guide, Air Quality System”, Environmental Protection Agency, Office of air quality planning and standards, 2009,

Wang, Y., and S. Bentley, “Development of a spatially explicit inventory of methane emissions from Australia and its verification using atmospheric concentration data”, *Atmos. Environ.*, 36, 4965– 4975, 2002.

Welch, G., Bishop, G., “An Introduction to the Kalman Filter”, ACM SIGGRAPH, 2001.

Xiao, X., Cohan, D. S., Byun, D. W., and Ngan, F.: “Highly nonlinear ozone formation in the Houston region and implications for emission controls”. *J. Geophys. Res.*, 115, D23309, doi:10.1029/2010JD014435, 2010.

Yamaji, K., Ikeda, K., Irie, H., Kurokawa, J., Ohara, T. “Influence of model grid resolution on NO₂ vertical column densities over east asia”. *J Air Waste Manag Assoc* 64:436–444, 2014.

Zhang, L., Constantinescu, E. M., Sandu, A., Tang, Y., Chai, T., Carmichael, G. R., Byun, D., Olaguer, E.: “An adjoint sensitivity analysis and 4D-Var data assimilation study of Texas air quality”, *Atmospheric Environment*, Vol. 42, Issue 23, p. 5787–5804, 2008.

Zeldovich, Y. B., P. Y. Sadovnikov, and D. A. Frank-Kamenetskii “Oxidation of Nitrogen in Combustion”, M. Shelef. Trans., Academy of Sciences of USSR, Institute of Chemical Physics, Moscow, Leningrad, 1947.

Zhang, J., Reid, J. S. Westphal, D. L., Baker, N. L., Hyer, E. J., “A system for operational aerosol optical depth data assimilation over global oceans”, *J. Geophys. Res.*, 113, D10208, doi:10.1029/2007JD009065, 2008.

Zhang, R., X. Tie, and D. W. Bond “Impacts of anthropogenic and natural NO_x sources over the U. S. on tropospheric chemistry, *Proc. Natl. Acad. Sci. U. S.A.*, 100(4), 1505-1509, doi:10.1073/pnas.252763799, 2003.

Zhang, L. Constantinescu, E.M. Sandu, A. Tang, Y. Chai, T. Carmichael, G.R. Byun, D. Olaguer , E. “An Adjoint Sensitivity Analysis and 4D-Var Data Assimilation Study of Texas Air Quality”. *Atmospheric Environment*, Vol. 42, Issue 23, p. 5787-5804, 2008.

Zhao, C. and Wang, Y.: “Assimilated inversion of NO_x emissions over east Asia using OMI NO₂ column measurements”. *Geophys. Res. Lett.*, 36, L06805, doi:10.1029/2008GL037123, 2009.

Zhou, Y. Brunner, D. Boersma, K. F. Dirksen, R. Wang P., “An improved tropospheric NO₂ retrieval for OMI observations in the vicinity of mountainous terrain, *Atmospheric Measurement Techniques*”, Vol. 2, pp 401-416, 29-7-2009, 2009.

Zhu, C.; Byrd, Richard H.; Lu, Peihuang; Nocedal, Jorge. “L-BFGS-B: Algorithm 778: L-BFGS-B, FORTRAN routines for large scale bound constrained optimization”. *ACM Transactions on Mathematical Software* 23 (4): 550–560.doi:10.1145/279232.279236, 1997.