RICE UNIVERSITY

Incorporation of Satellite Observations into Texas Ozone Attainment Modeling

by

Wei Tang

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE

Dr. Daniel S. Cohan, Chair
Associate professor, Civil and Environmental Engineering

Dr. Robert L. Griffin
Professor, Civil and Environmental Engineering

Dr. Caroline A. Masiello
Associate professor, Earth Science

Dr. Arastoo Pour-Biazar
Scientist, Earth System Science Center, UAH

HOUSTON, TEXAS
May 2014
ABSTRACT

Incorporation of Satellite Observations into Texas Ozone Attainment Modeling

by

Wei Tang

Uncertain photolysis rates and nitrogen oxides (NO\textsubscript{x}) emission inventories impair the accuracy of ozone (O\textsubscript{3}) regulatory modeling. Satellite-observed clouds have been used to correct model-predicted photolysis rates, and satellite-constrained top-down NO\textsubscript{x} emissions have been used to identify and reduce uncertainties in bottom-up NO\textsubscript{x} emissions. However, studies on using multiple satellite-derived model inputs to improve O\textsubscript{3} State Implementation Plan (SIP) modeling are rare. In this thesis, observations of clouds from the Geostationary Operational Environmental Satellite (GOES) and of NO\textsubscript{2} from the Ozone Monitoring Instrument (OMI) are used to adjust the inputs to SIP modeling of O\textsubscript{3} in Texas. The discrete Kalman filter (DKF) inversion approach is successfully applied with decoupled direct method (DDM) sensitivities in the Comprehensive Air Quality Model with extensions (CAMx) model to adjust Texas NO\textsubscript{x} emissions in designated emission regions and categories to better match OMI NO\textsubscript{2} data. The NO\textsubscript{2} vertical column densities (VCD) difference between OMI and CAMx over rural areas is alleviated by adding missing lightning and aviation and underestimated soil NO\textsubscript{x} emissions to the base regulatory emission inventory and further reduced by increasing modeled NO\textsubscript{x} lifetime and adding an artificial NO\textsubscript{2} layer in the upper troposphere. The region-based DKF inversion using OMI NO\textsubscript{2} tends to scale up NO\textsubscript{x} emissions in most
regions, which conflicts with the inversion results using ground NO₂ measurements and fails to improve the ground-level O₃ simulations. The sector-based DKF inversion using OMI NO₂ suggests scaling down area and non-road NOₓ emissions by 50%, leading to approximately 2-5ppb decrease in ground 8-h O₃ concentrations, and improving both hourly ground-level NO₂ and O₃ simulations by reducing biases by 0.25 and 0.04 and errors by 0.13 and 0.04, respectively. Finally, using both GOES-derived photolysis rates and OMI-constrained NOₓ emissions reduces modeled bias and error by 0.05, and increases the model correlations in simulating ground O₃ measurements and makes O₃ more sensitive to NOₓ emissions in the O₃ nonattainment areas.
Acknowledgments

First and foremost, I would like to give my deepest appreciation to my advisor Dr. Daniel Cohan for providing me this wonderful opportunity to get into this exciting research field. His support, understanding, trust, encouragement, and insightful guidance make me finally go through this long journey.

I am also grateful to my committee members, Dr. Robert Griffin, Dr. Caroline Masiello, and Dr. Arastoo Pour Biazar, for being willing to serve on my thesis committee. Their suggestions and feedback helped me better understand my research and added immense value to this dissertation.

I sincerely appreciate the whole-hearted help given by my colleagues Dr. Xue Xiao, Dr. Wei Zhou, Dr. Antara Digar, Ben Lash, Erin Chavez-Figueroa, and Aditya Jani. Without their help and encouragement, it would be impossible for me to complete this journey. I also thank Dr. Lok Lamsal at NASA for providing OMI data, Jim McKay, Ron Thomas, Doug Boyer, and Mark Estes at TCEQ for providing emission inputs and insightful discussions about the TCEQ emission inventory, Gary Wilson, Greg Yarwood, and Jeremiah Johnson at ENVIRON for CAMx support, and Dr. Barron Henderson at University of Florida for providing processed DC-8 measurements data.

I want to give my sincerest gratitude here to my family, for their tremendous love, blessing, understanding, and support throughout these six years.

Finally, I gratefully acknowledge the financial support for this project provided by the U.S. NASA Research Opportunities in Space and Earth Sciences (ROSES) grant NNX10AO05G and by the NASA Air Quality Applied Science Team.
Contents

Acknowledgments ............................................................................................................ iv
Contents ............................................................................................................................. v
List of Figures ................................................................................................................. viii
List of Tables .................................................................................................................... xi
Nomenclature .................................................................................................................. xii

Introduction ..................................................................................................................... 19

1.1. Background and motivation ................................................................................... 19
  1.1.1. Tropospheric ozone ......................................................................................... 19
  1.1.2. Air quality in eastern Texas ............................................................................. 20
  1.1.3. Air quality modeling ........................................................................................ 21
  1.1.4. Uncertainty in air quality modeling ................................................................. 23
  1.1.5. Space-based measurements ............................................................................. 26
    1.1.5.1. Satellite-observed clouds .......................................................................... 26
    1.1.5.2. Satellite-observed NO₂ .............................................................................. 26
  1.1.6. Inverse modeling ............................................................................................. 29
1.2. Scope of this work .................................................................................................. 31
1.3. Research outline ..................................................................................................... 33

Inverse modeling of Texas NOₓ emissions using space-based and ground-based NO₂
observations ..................................................................................................................... 34

2.1. Introduction ............................................................................................................ 35
2.2. Methodology .......................................................................................................... 37
  2.2.1. Model inputs and configurations ..................................................................... 37
  2.2.2. Emission inventory ........................................................................................ 39
  2.2.3. Inversion Regions ............................................................................................ 41
  2.2.4. Inversion method ............................................................................................. 42
  2.2.5. NO₂ observations ............................................................................................ 45
    2.2.5.1. Satellite NO₂ measurements ...................................................................... 45
    2.2.5.2. Ground and other NO₂ measurements ....................................................... 46
2.3. Results and Discussion ........................................................................................... 48
  2.3.1. Pseudodata test for the DKF inversion with CAMx-DDM ............................. 48
Supporting Information for Chapter 2 ................................................................. 123
  A1. Meteorological parameters evaluations ......................................................... 123
  A2. Sensitivity analysis of the DKF inversions ...................................................... 126
  A3. Direct Scaling (DS) inversion ...................................................................... 127
    A3.1 Method .................................................................................................. 127
    A3.2 Results ............................................................................................... 129

Appendix B .......................................................................................................... 133

Supporting Information for Chapter 3 ................................................................. 133
  B1. CAMx modeled profile-based OMI retrieval .................................................. 133
  B2. Impact of increased NOx lifetime and artificial layer on modeled NO2 VCD ... 138
  B3. Sensitivity of DKF inversion to error covariance matrices ............................ 140
  B4. Top-down VOC emissions ........................................................................... 142
    B4.1 Base case VOC emission inventory ......................................................... 143
    B4.2 VOC observations .............................................................................. 144
    B4.3 Results ............................................................................................... 145

Appendix C .......................................................................................................... 150

List of publications and conference presentations ............................................ 150
  C1. Journal publications .................................................................................. 150
  C2. Conference presentations ...................................................................... 150
List of Figures

Figure 1-1. Diagram of the tropospheric O₃ formation process (Source: Digar 2012). ................................................................. 20

Figure 1-2. Air quality planning areas in Texas (Source: Gonzales and Williamson 2011). ................................................................. 21

Figure 1-3. Schematic of air quality modeling (Source: Digar 2012). ................................................................. 22

Figure 1-4. O₃ sensitivities to NOₓ emissions if NOₓ emissions were 50% smaller (left) or larger (right) than its base value (source: Xiao et al., 2010). ........................................ 25

Figure 1-5. Flow diagram of application of satellite data to ozone attainment planning. ................................................................. 32

Figure 2-1. 12-km CAMx modeling domain for eastern Texas (black square), inversion regions (shaded), ground AQS NO₂ monitoring sites (blue triangles), and Moody Tower (red circle). ................................................................. 39

Figure 2-2. Schematic diagram of Kalman filter inversion process. ................................................................. 45

Figure 2-3. Pseudodata test showing that the DKF inversion accurately adjusts the NOₓ emissions from the perturbed case (a) to the a posteriori case (b) to match the desired base NO₂ column densities. Numbers indicate perturbation factors in the legend (a) and adjustment factors in the legend (b). Similar performance is found for the August 13-22 test period. ................................................................. 49

Figure 2-4. Monthly averaged (June 3 to July 1) tropospheric NO₂ vertical columns at 1-2pm from a) OMI observations and from CAMx simulations using b) base case emissions inventory, c) a priori emission inventory (with additional lightning, aircraft, and soil NOₓ), and d) OMI-based inverted NOₓ emissions using the DKF method. ................................................................. 51

Figure 2-5. Same as Fig. 2-4 but for the August-September episode. ................................................................. 56

Figure 2-6. Daily variations of modeled (solid line) and observed (dashed line) ground NO₂ concentrations for the June (red) and August-September (blue) episodes. Note: NO₂ concentrations were taken by averaging monthly data for all sites. ................................................................. 57
Figure 2-7. Monthly 8-h (10am-6pm) averaged ground-level O₃ concentrations (top), O₃ sensitivity to NOₓ (middle), and O₃ sensitivity to VOC (bottom) for the a priori case (left column), and differences (a posteriori minus a priori) for the OMI-based (middle column) and ground-based (right column) DKF inversions in the June episode. The August-September episode shows similar results. 61

Figure 3-1. Seven designated inversion regions in eastern Texas (shaded) within 12-km CAMx modeling domain (black square) covered by ground NO₂ monitoring sites (blue triangles), VOC monitoring sites (green circles), and O₃ monitoring sites (red circles). 71

Figure 3-2. Differences between satellite-derived (GOES) and model predicted (MOD) JNO₂ (left) in simulating NO₂ (middle) and O₃ (right) at 13:00 on 2 September 2006. 82

Figure 3-3. Monthly 8-h (10:00-18:00LT) averaged differences between satellite-derived (GOES) and model predicted (MOD) (a) JNO₂ in simulating (b) NO₂, (c) O₃, and O₃ sensitivities to (d) NOₓ and (e) VOC. 83

Figure 3-4. Change in model performance (R², NMB, and NME) in simulating daily 8 hours (10:00-18:00LT) NO₂ (left) and O₃ (right) caused by satellite-derived photolysis rates. 84

Figure 3-5. Vertical column densities of NO₂ sensitivities to NOₓ emissions of (a) area, (b) non-road, (c) mobile, (d) biogenic, (e) lightning, (f) aviation, and (g) non-EGU points source sectors. 86

Figure 3-6. Pseudodata analysis for the sector-based DKF inversion (top), and its sensitivities to uncertainties in emissions (bottom left) and in OMI observations (bottom right). 87

Figure 3-7. Comparisons of modeled NO₂ vertical distributions with INTEX NASA DC-8 flight (left) and TexAQS 2006 NOAA P-3 aircraft (right) measurements. 87

Figure 3-8. Monthly averaged (16 August to 15 September) tropospheric NO₂ VCDs at 13:00-14:00LT from (a) OMI, (b) a priori simulation, (c) difference between OMI and a priori simulation, and simulations using a posteriori NOₓ emissions generated by (d) region-based DKF inversion, and sector-based DKF inversion (e) case I and (f) case II. 90

Figure 3-9. Monthly 8-h (10:00-18:00LT) averaged ground O₃ concentrations (top), O₃ sensitivity to NOₓ (middle), and O₃ sensitivity to VOC (bottom) for the a priori
case (left), and differences between a posteriori and a priori for the sector-based DKF inversions case I (middle) and case II (right).

Figure 3-10. Monthly 8-h (10:00-18:00LT) averaged differences in modeled (a) ground O$_3$ concentrations, (b) O$_3$ sensitivity to NO$_x$, and (c) O$_3$ sensitivity to VOC resulting from use of both satellite-derived photolysis rates and NO$_x$ emissions in place of a priori data.
List of Tables

Table 2-1. Categorized a priori NO\textsubscript{x} emission rates in inversion region for two modeling episodes.......................................................................................................................................................................................................................................................... 41

Table 2-2. Scaling factors for each region from different inversions.......................... 52

Table 2-3. Performance of CAMx in simulating OMI-observed NO\textsubscript{2} column densities...................................................................................................................................................................................................................................................... 54

Table 2-4. Performance of CAMx in simulating AQS Ground-level NO\textsubscript{2} \textsuperscript{a} .................. 58

Table 2-5. Performance of CAMx in simulating P-3 aircraft-observed NO\textsubscript{2} and NO\textsubscript{y}. .................................................................................................................................................................................................................................................................................. 58

Table 2-6. Performance of CAMx in simulating AQS hourly ground-level O\textsubscript{3}. ...... 60

Table 3-1. NO\textsubscript{x} emission rates for seven sectors in seven inversion regions.......... 77

Table 3-2. Scaling factors of region-based and sector-based inversions....................... 77

Table 3-3. Evaluation of CAMx modeled NO\textsubscript{2} using OMI NO\textsubscript{2}. ................................. 92

Table 3-4. Evaluation of CAMx modeled NO\textsubscript{2} using hourly AQS ground-measured NO\textsubscript{2}. .................................................................................................................................................................................................................................................................................. 92

Table 3-5. Evaluation of CAMx modeled NO\textsubscript{2} using P-3 aircraft-measured NO\textsubscript{2} and NO\textsubscript{y}. .................................................................................................................................................................................................................................................................................. 93

Table 3-6. Evaluation of CAMx modeled O\textsubscript{3} using hourly AQS ground-measured O\textsubscript{3}........................................................................................................................................................................................................................................................................................................... 97
Nomenclature

Abbreviations

AOD   Aerosol optical depth
AKs   Averaging Kernels
AMF   Air Mass Factor
AQS   Air Quality System
ARD   Acid rain database
Auto-GC Automatic Gas Chromatographs
BEHR  Berkeley High-Resolution product
BPA   Beaumont-Port Arthur
CB-05 Carbon Bond version 2005
CAMx  Comprehensive Air Quality Model with extensions
CEM   Continuous emissions monitoring
CMAQ  Community Multiscale Air Quality Modeling System
DDM   Decoupled Direct Method
DFW   Dallas Fort-Worth

DISCOVER Deriving Information on Surface conditions from Column and
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DKF</td>
<td>Discrete Kalman filter</td>
</tr>
<tr>
<td>DOAS</td>
<td>Differential Optical Absorption Spectroscopy</td>
</tr>
<tr>
<td>DOMINO</td>
<td>Derivation of OMI tropospheric NO₂</td>
</tr>
<tr>
<td>DP</td>
<td>DOMINO Product</td>
</tr>
<tr>
<td>DS</td>
<td>Direct Scaling</td>
</tr>
<tr>
<td>EDGAR</td>
<td>Emission Database for Global Atmospheric Research</td>
</tr>
<tr>
<td>EGAS</td>
<td>Economic Growth Analysis System</td>
</tr>
<tr>
<td>EGUs</td>
<td>Electric generating units</td>
</tr>
<tr>
<td>ENVISAT</td>
<td>Environmental Satellite</td>
</tr>
<tr>
<td>EOS</td>
<td>Earth Observing System</td>
</tr>
<tr>
<td>EPA</td>
<td>Environmental Protection Agency</td>
</tr>
<tr>
<td>EPS3</td>
<td>Emission Processing System, version 3</td>
</tr>
<tr>
<td>ESA</td>
<td>European Space Agency</td>
</tr>
<tr>
<td>ETH</td>
<td>ethylene</td>
</tr>
<tr>
<td>ETHA</td>
<td>ethane</td>
</tr>
<tr>
<td>ERS-2</td>
<td>second European Remote Sensing</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------------------------------------------</td>
</tr>
<tr>
<td>GEOS</td>
<td>Goddard Earth Observing System</td>
</tr>
<tr>
<td>GC-FID</td>
<td>Gas Chromatographs-flame ionization detector</td>
</tr>
<tr>
<td>GloBEIS</td>
<td>Global Biosphere Emissions and Interactions System</td>
</tr>
<tr>
<td>GMI</td>
<td>Global Modeling Initiative</td>
</tr>
<tr>
<td>GOES</td>
<td>Geostationary Operational Environmental Satellite</td>
</tr>
<tr>
<td>GOME</td>
<td>Global Ozone monitoring Experiment</td>
</tr>
<tr>
<td>HGB</td>
<td>Houston Galveston Brazoria</td>
</tr>
<tr>
<td>HRVOC</td>
<td>highly reactive VOC</td>
</tr>
<tr>
<td>INTEX</td>
<td>Intercontinental Chemical Transport Experiment</td>
</tr>
<tr>
<td>ISOP</td>
<td>isoprene</td>
</tr>
<tr>
<td>LPAS</td>
<td>Laser Photoacoustic Spectroscopy</td>
</tr>
<tr>
<td>MISR</td>
<td>Multiple Imaging SepectroRadiometer</td>
</tr>
<tr>
<td>MM5</td>
<td>5th Generation Mesoscale Meteorological Model</td>
</tr>
<tr>
<td>MODIS</td>
<td>Moderate Resolution Imaging SpectroRadiometer</td>
</tr>
<tr>
<td>MOVES</td>
<td>Motor Vehicle Emission Simulator</td>
</tr>
<tr>
<td>MOZART</td>
<td>Model for Ozone and Related Chemical Tracers</td>
</tr>
<tr>
<td>NAAQS</td>
<td>National Ambient Air Quality Standards</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>----------------------------------------------------------</td>
</tr>
<tr>
<td>NASA</td>
<td>National Aeronautics and Space Administration</td>
</tr>
<tr>
<td>NCAR</td>
<td>National Center for Atmospheric Research</td>
</tr>
<tr>
<td>NLDN</td>
<td>National Lightning Detection Network</td>
</tr>
<tr>
<td>NMIM</td>
<td>National Mobile Inventory Model</td>
</tr>
<tr>
<td>NOAA</td>
<td>National Oceanic and Atmospheric Administration</td>
</tr>
<tr>
<td>NOx</td>
<td>Nitrogen Oxides</td>
</tr>
<tr>
<td>OMI</td>
<td>Ozone Monitoring Instrument</td>
</tr>
<tr>
<td>PAQMs</td>
<td>Photochemical air quality models</td>
</tr>
<tr>
<td>PAMS</td>
<td>Photochemical Assessment Monitoring Stations</td>
</tr>
<tr>
<td>PAN</td>
<td>peroxyacetyl nitrate</td>
</tr>
<tr>
<td>PBL</td>
<td>Planetary Boundary Layer</td>
</tr>
<tr>
<td>PSCF</td>
<td>Potential Source Contribution Function</td>
</tr>
<tr>
<td>PTRMS</td>
<td>Proton Transfer Reaction Mass Spectrometer</td>
</tr>
<tr>
<td>SCIAMCHY</td>
<td>Scanning Imaging Absorption Spectrometer for Atmospheric</td>
</tr>
<tr>
<td></td>
<td>Chartography</td>
</tr>
<tr>
<td>SIP</td>
<td>State Implementation Plan</td>
</tr>
<tr>
<td>SP</td>
<td>Standard Product</td>
</tr>
<tr>
<td>STARS</td>
<td>State of Texas Air Reporting System</td>
</tr>
</tbody>
</table>
**Symbols**

- $P$: emission error covariance matrix
- $R$: measurement error covariance matrix
- $E$: emissions
- $S$: sensitivity coefficient matrix
- $G$: Kalman gain matrix
- $C$: Concentration matrix
- $k$: smoothing parameter
\( K \) smoothing kernel

\( M \) transition matrix

\( \varepsilon \) model error

\( Q \) modeling process errors

\( S_{ij}^{(1)} \) first-order semi-normalized sensitivity coefficient

\( p_j \) a model parameter,

\( \tilde{p}_j \) a unperturbed field

\( \varepsilon_j \) scaling valuable with a nominal value of 1

\( \delta \) inversion convergence criterion

\( x \) perturbation factor

\( X_i \) CAMx modeled partial NO\textsubscript{2} sub-column density

\( \Omega \) NO\textsubscript{2} column density

\( J \) Photolysis rates

\( \sigma(\lambda) \) absorption cross-section

\( \lambda \) wavelength

\( \phi(\lambda) \) quantum yield

\( F(\lambda) \) actinic flux

\( f_c \) cloud fraction for a grid cell

\( tr_c \) cloud transmissivity at each model grid layer

\( \theta \) solar zenith angle
\( \tau_c \)  cloud optical depth

\( \beta \)  scattering phase-function asymmetry factor

ppb  parts per billion

ppm  parts per million

ppmC  parts per million carbon

MBE  mean bias error

RMSE  root mean square error

\( S_c \)  satellite slant column

\( V_c \)  satellite vertical column

O  high order terms

\( y \)  observation

\( F(x,b) \)  forward model

\( x_a \)  a prior emission that seeks to optimize

\( b \)  other model parameters that do not seek to optimize
Chapter 1

Introduction

1.1. Background and motivation

1.1.1. Tropospheric ozone

The U.S. Environmental Protection Agency (U.S. EPA) establishes National Ambient Air Quality Standards (NAAQS) for six criteria air pollutants as required by the Clean Air Act to protect public health and the environment. The six criteria air pollutants are carbon monoxide (CO), nitrogen dioxide (NO₂), sulfur dioxide (SO₂), ozone (O₃), particulate matter (PM), and lead (Pb). O₃ in the troposphere is a secondary pollutant that forms through the reactions between anthropogenic and biogenic emitted nitrogen oxides (NOₓ) and volatile organic compounds (VOCs) in the presence of sunlight and humidity (Fig. 1.1). The formation of nitric acid (HNO₃) and hydrogen peroxide (H₂O₂) occur simultaneously, terminating the NOₓ and HOₓ cycles in O₃ production (Seinfeld and Pandis, 2006).
1.1.2. Air quality in eastern Texas

Two metropolitan regions in eastern Texas, Houston-Galveston-Brazoria (HGB) and Dallas-Fort Worth (DFW) are designated by the U.S. EPA as O₃ non-attainment areas for violating the 1997 eight-hour ozone NAAQs of 0.08ppm (Fig. 1.2). The HGB region locates in proximity to the Gulf of Mexico with the population of six million and the largest concentration of petrochemical facilities in the United States. Large sources of NOₓ and VOC emissions combining with intensive solar radiation and humidity make HGB one of the worst O₃ pollution regions in the US for almost two decades (Kleinman et al., 2002; Ryerson et al., 2003; Daum et al., 2004; Rappengluck et al., 2008, TCEQ 2010). The DFW region located in northeast Texas has similar but less humid weather.
conditions, with NOx emissions from large mobile sources (TCEQ 2011). The recently tightened O3 standard of 75ppb makes several additional metropolitan regions in eastern Texas on the edge of violating the new standard and provides more challenges for air quality management.

Figure 1-2. Air quality planning areas in Texas (Source: Gonzales and Williamson 2011).

1.1.3. Air quality modeling

States that contain nonattainment areas violating the NAAQS are required by the U.S. EPA to develop attainment demonstrations to address air pollution abatement efforts (National Research Council, 2004). Photochemical air quality models (PAQMs) are widely used by scientists and policy makers to better understand the transport and
formation processes of air pollutants and to develop emission control strategies to improve air quality (e.g., Russell and Dennis, 2000; National Research Council, 2007). PAQMs simulate air pollutant concentrations, such as O₃, by taking emission, meteorology, and chemistry parameters as inputs (Fig. 1.3) and updating the concentration field via operator splitting of atmospheric and chemical processes (Eq. 1.1) using finite difference calculations (ENVIRON 2010; CMAS 2012).

\[
\frac{dC}{dt} = [\text{advection}] + [\text{diffusion}] + [\text{deposition}] + [\text{chemistry}] + [\text{emission}] \tag{1.1}
\]

The Community Multiscale Air Quality (CMAQ) model (Byun and Schere, 2006) and the Comprehensive Air Quality Model with extensions (CAMx) (ENVIRON 2010) are three-dimensional (3-D) Eulerian air quality models that are both approved by the U.S. EPA for regulatory modeling of tropospheric O₃ formations.

Figure 1-3. Schematic of air quality modeling (Source: Digar 2012).
1.1.4. Uncertainty in air quality modeling

Two types of uncertainties, structural and parametric, associate with air quality models. Structural uncertainty results from imperfect numerical representation of physical mechanisms, for example, modeled concentrations may vary using different vertical mixing schemes to represent vertical transport process in the model. Parametric uncertainty results from errors in the model inputs and parameterizations such as uncertainties in the emission inputs or model parameterized eddy diffusivity. These uncertainties may significantly impair the accuracy of model performance and misdirect control strategy development (Pinder et al., 2009; Fine et al., 2003). Hence, identifying and reducing these uncertainties in the model are essential for O₃ regulatory modeling. Structural uncertainties are typically characterized by an ensemble of model runs in simulating a common episode (Pinder et al., 2009; Mallet and Sportisse, 2006). Parametric uncertainties are usually quantified via sensitivity analysis to characterize the response of pollutant concentrations by perturbing the model inputs (Pinder et al., 2009; Tian et al., 2010; Digar and Cohan, 2010; Tang et al., 2011).

Decoupled Direct Method (DDM)-3D is a widely used sensitivity analysis tool that calculates semi-normalized sensitivity coefficients as shown in Eq. 1.2,

\[
S_{ij}^{(1)} = \frac{\partial C_i}{\partial p_j} = \frac{\partial C_i}{\partial (\varepsilon_j, \tilde{p}_j)} = \frac{\partial C_i}{\partial \varepsilon_j}
\]  (1.2)
where $S_y^{(1)}$ is the first-order semi-normalized sensitivity coefficient, $p_j$ represents a model parameter, $\tilde{p}_j$ represents a unperturbed field and $\varepsilon_j$ is a scaling valuable with a nominal value of 1.

DDM computes the semi-normalized sensitivity coefficients in the same way as pollutant concentration calculations in the model and updates the sensitivity field (Eq. 1.3) simultaneously with concentration field (Eq. 1.1); hence, it is more direct, efficient, and stable and less subject to numerical noise compared to the other sensitivity calculation methods.

$$\frac{dS_y^{(1)}}{dt} = \frac{\partial}{\partial \varepsilon} [\text{advection}] + \frac{\partial}{\partial \varepsilon} [\text{diffusion}] + \frac{\partial}{\partial \varepsilon} [\text{deposition}] + \frac{\partial}{\partial \varepsilon} [\text{chemistry}] + \frac{\partial}{\partial \varepsilon} [\text{emission}]$$

DDM can be used to determine the model responsiveness to variations of input parameters quantitatively (Dunker, 1984; Yang et al., 1997; Hakami et al., 2003, 2004) and has been implemented into both CMAQ (Cohan et al., 2005; Napelenok et al., 2006, 2008) and CAMx (Koo et al., 2007). Recent studies (e.g. Digar and Cohan, 2010; Xiao et al., 2010; Tang et al., 2011) applied DDM to investigate influences of model uncertainties on O3 concentrations and their sensitivities to NOx and VOC emissions which are used for determining O3 abatement strategies. For example, if computed first-order sensitivity coefficient (Eq. 1.2) of O3 concentration to NOx emission is positive in a region, controlling NOx emissions favors O3 reduction (NOx-limited zone); otherwise, controlling VOC emissions may be more efficient for decreasing O3 concentrations (VOC-limited zone). However, study from Xiao et al. (2010) indicates that the
uncertainty in NO\textsubscript{x} emission inventory may flip NO\textsubscript{x}-limited zone to VOC-limited zone and mislead control strategy development (Fig. 1.4).

**Figure 1-4.** O\textsubscript{3} sensitivities to NO\textsubscript{x} emissions if NO\textsubscript{x} emissions were 50% smaller (left) or larger (right) than its base value (source: Xiao et al., 2010).

Recent studies show that uncertain photolysis rates and NO\textsubscript{x} emission inventories are two leading uncertainties in the simulations of O\textsubscript{3} concentrations and their sensitivities to precursor emissions (Hanna et al., 2001; Koo et al., 2008; Xiao et al., 2010). The NO\textsubscript{x} emission inventory used in the current SIP modeling is developed by a bottom-up approach based on the estimated emission factors and activities. The photolysis rates are calculated via the model predicted cloud fractions and model parameterized transmissivity. Directly measuring the emissions and photolysis rates over
the 3-D modeling domain would be the best way to reduce these uncertainties, but it is impractical to do so with ground and aircraft measurements. However, space-based measurements are capable of observing some atmospheric parameters and air pollutants with great spatial coverage that other methods cannot achieve.

1.1.5. Space-based measurements

1.1.5.1. Satellite-observed clouds

The Geostationary Operational Environmental Satellite (GOES) operated by the National Oceanic and Atmospheric Administration (NOAA) is capable of measuring cloud properties with high spatial resolution down to 1-km for cloud optical reflectance and 4-km for cloud top heights, and with high temporal resolution down to an hour or less. Recent study (Guenther et al., 2012) suggests that modeled clouds are underpredicted by the current meteorological models, leading to more solar irradiance reaching the earth surface and thus causing more O₃ formation in the model. Pour-Biazar et al. (2007) used GOES satellite-based cloud information to correct the model photolysis rates in CMAQ, and the results showed large discrepancies occurred between model predicted and satellite-derived photolysis rates, NO₂, and O₃ concentrations.

1.1.5.2. Satellite-observed NO₂

Currently, there are three satellite instruments measuring atmospheric NO₂: Global Ozone monitoring Experiment (GOME), Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMCHY), and Ozone Monitoring Instrument (OMI). All three satellites follow sun-synchronous polar orbits and are
equipped with nadir-viewing UltraViolet-Visible (UV-Vis) spectrometers that measure solar backscattered irradiance from the Earth’s atmosphere. GOME on board the European Space Agency (ESA) second European Remote Sensing (ERS-2) satellite which launched in April 1995, measures atmospheric NO2 within the 425 to 450 nm spectral range at 10:30 local time (LT) with global coverage once every three days and with spatial resolution down to scales of 40×320 km² in the nadir view pixels (Burrows et al., 1999). SCIAMCHY deployed on ESA’s Environmental Satellite (ENVISAT) which launched in March 2002, measures atmospheric NO2 within the 425 to 450 nm spectral range at 10:00 LT with global coverage once every six days and with spatial resolution down to scales of 30×60 km² in the nadir view pixels (Bovensmann et al., 1999). The Dutch-Finnish Ozone Monitoring Instrument (OMI) aboard NASA’s Earth Observing System (EOS) Aura satellite, which was launched on 15 July 2004, retrieves atmospheric NO2 in the spectral range from 405nm to 465 nm at 13:30 LT with daily global coverage and with spatial resolution down to scales of 13×24 km² in pixels of nadir view (Levelt et al., 2006a, 2006b). Since OMI has the highest temporal and spatial resolution among three satellite measurements, OMI-retrieved NO2 is used in this study.

The OMI-retrieved tropospheric NO2 vertical column density (VCD) is used in this study and compared to the modeled NO2 VCD. The OMI tropospheric NO2 VCD retrieval contains three steps (Boersma et al., 2007). First, the NO2 slant column density (SCD) is derived by the Differential Optical Absorption Spectroscopy (DOAS) spectral fitting technique; second, the stratospheric NO2 SCD was separated from the tropospheric NO2 SCD; third, the tropospheric NO2 SCD is divided by an air mass factor (AMF), converting SCD to the tropospheric NO2 VCD. Currently, there are two independent OMI
NO$_2$ products: OMI_NASA standard product, version 2 (SP2) (Bucsela et al., 2013), available at NASA Goddard Earth Science Data Active Archive Center (http://disc.sci.gsfc.nasa.gov/data/datapool/OMI/), and OMI_DOMINO product, version 2 (DP2) (Boersma et al., 2011), available at Tropospheric Emission Monitoring Internet Service (http://www.temis.nl/). The differences between these two products are mainly from the last two post-processing SCD steps described above. In step two, stratospheric NO$_2$ SCD in the SP2 product is separated using local analyzed stratospheric field information, while the DP2 product derives stratospheric NO$_2$ slant column using a global chemical transport model, TM4. In step three, the AMF used in the SP2 product for converting NO$_2$ SCD into VCD is generated by a radiative transfer model, TOMRAD, with a monthly mean NO$_2$ profile derived from the multiannual (2005-2007) simulations using the Global Modeling Initiative (GMI) model, whereas the AMF used in the DP2 product is computed by the DAK radiative transfer model using TM4 simulated tropospheric NO$_2$ vertical profiles at the time of measurements. The surface reflectivity used in SP2 is from OMI-based monthly climatology, while that in DP2 is from Total Ozone Mapping Spectrometer (TOMS)- and GOME-based climatologies.

A Berkeley High Resolution (BEHR) OMI NO$_2$ product (Russell et al., 2011) was recently developed from SP2. Compared to the SP2 product, the AMF calculated in BEHR uses finer resolution terrain and a priori NO$_2$ profiles, resulting in lower NO$_2$ retrieval in rural areas and higher NO$_2$ in urban areas. However, the a priori NO$_2$ profile in BEHR was generated from the regional WRF-Chem model without lightning and fire emissions included.
Several studies have used satellite-retrieved NO$_2$ to adjust the bottom-up NO$_x$ emission inventory in both global scale (Martin et al., 2003; Lin et al., 2009) and regional scale modeling (Quélo et al., 2005; Konovalov et al., 2006, 2008; Kurokawa et al., 2009; Mendoza-Dominguez and Russell, 2000; Zhang et al., 2008). The adjustments in the NO$_x$ emission inputs are inferred by the difference between satellite-retrieved and modeled NO$_2$ concentrations. Thus, inverse modeling is required to constrain the emission inputs based on the difference between satellite data and model output concentrations.

1.1.6. Inverse modeling

Inverse modeling combined with PAQMs has been used to estimate model variables that cannot be measured directly (Gilliland and Abbitt, 2001), such as NO$_x$ emissions. Applying inverse modeling techniques with independent NO$_2$ observations optimizes the NO$_x$ emission inputs and reduces the uncertainties in the emission inventories.

The relationship between model parameters and observations ($y$) can be described via a forward model ($F$) as shown in Eq. 1.4

$$ y = F(x, b) + \varepsilon $$

(1.4)

where $x$ is a model parameter that we seek to optimize, $b$ are the other model parameters that we do not seek to optimize, and $\varepsilon$ represents the errors in observations and in model parameters. The forward model in Eq. 1.4 can be further demonstrated in a linearized form using Taylor expansion as shown in Eq. 1.5,
\[ y = F(x_a, b) + S^{(1)}(x - x_a) + O((x - x_a)^2) + \varepsilon \]  

where \( x_a \) is the a priori estimate of \( x \), \( S^{(1)} = \frac{\partial F}{\partial x} \), and \( O \) represents the high order terms and usually can be ignored. \( S^{(1)} \) is the first-order sensitivity coefficient that can be calculated using DDM (Napelenok et al., 2008) or an adjoint method (Kurokawa et al., 2009).

In air quality modeling, NO\(_x\) emission is a model input, and NO\(_2\) concentrations are generated by the model as output. By contrast, inverse modeling is driven by the difference between observed and modeled NO\(_2\) concentrations, and seeks the optimal NO\(_x\) emissions by minimizing errors in a weighted least squares sense (Jacob, 2007) as shown in Eq. 1.6.

\[ y - F(\tilde{E}_{NO_x}, b) = S^{(1)}(E_{NO_x} - \tilde{E}_{NO_x}) + \varepsilon \]  

where \( y \) is observed NO\(_2\), \( F(\tilde{E}_{NO_x}, b) \) is a priori modeled NO\(_2\), \( E_{NO_x} \) is the optimized NO\(_x\) emission, and \( \tilde{E}_{NO_x} \) is the a priori NO\(_x\) emission (the bottom-up NO\(_x\) emission inventory). Non-linearity is usually associated with the relationship between NO\(_x\) emissions and NO\(_2\) concentrations in air quality models. Hence, inversions are conducted iteratively.

Discrete Kalman filter (DKF) (Prinn, 2000) is an inverse modeling technique that is well established and has been used in several atmospheric studies to constrain air pollutant emissions, such as CO, isoprene, and ammonia (Mulholland and Seinfeld 1995; HaasLaursen et al., 1996; Gilliland and Abbitt, 2001). Napelenok et al. (2008) recently
has conducted DKF inversions on NO\textsubscript{x} emissions for southeastern US using the SCIAMACHY satellite NO\textsubscript{2} observations.

### 1.2. Scope of this work

The main scope of this study is to test the capabilities of using satellite observations to reduce the two largest model uncertainties, photolysis rates and NO\textsubscript{x} emissions, in Texas ozone attainment modeling (Fig. 1.4). Although using satellite data either to correct photolysis rates or constraining NO\textsubscript{x} emissions has been performed in previous studies, applying multiple satellite retrievals to enhance ozone regulatory modeling in Texas has not been explored.

In this study, O\textsubscript{3} modeling is focused on two Texas SIP modeling episodes, 31 May to 2 July and 13 August to 15 September 2006 in eastern Texas, when highest O\textsubscript{3} concentrations occurred in DFW and HGB. The GOES-retrieved clouds and transmissivity have been implemented into the CAMx model and replace the model standard procedure to reproduce the photolysis rates. The DKF inversions with OMI-observed NO\textsubscript{2} are conducted to constrain NO\textsubscript{x} emissions in designated regions and sectors. SP2 and a newly developed OMI NO\textsubscript{2} product based on SP2 but retrieved with a finer resolution NO\textsubscript{2} profile from nested GEOS-Chem model are both applied in the inversions. As the ground NO\textsubscript{2} monitoring produces a dataset with better temporal resolutions than satellite observations, the EPA Air Quality System (AQS) ground monitoring NO\textsubscript{2} data are also used to constrain NO\textsubscript{x} emissions, and the results are compared with satellite-based inversions. In order to reduce errors in the inversion
results, missing lightning and aviation NO\textsubscript{x} emissions are added into the base case emission inventory, and underestimated soil NO\textsubscript{x} emission is doubled from its base value; furthermore, the NO\textsubscript{x} lifetime is manually increased in the model by reducing the reaction rate constant of the reaction NO\textsubscript{2} + OH, and a homogeneous NO\textsubscript{2} layer is added into the model top.

Figure 1-5. Flow diagram of application of satellite data to ozone attainment planning.

In this work, all model inputs for CAMx simulations are provided by TCEQ. The GOES-retrieved clouds and transmissivity are provided by University of Alabama at Huntsville, and the 0.1×0.1 gridded NASA standard OMI NO\textsubscript{2} data are provided by Dr. Lok N. Lamsal at NASA Goddard Space Flight Center. The National Lightning Detection Network (NLDN) lightning data are purchased from the Vaisala Ltd., and the aviation
NO\textsubscript{x} emissions are downloaded from Emission Database for Global Atmospheric Research (EDGAR). The measurement data from the Second Texas Air Quality Study (TexAQS II) (Parrish et al., 2009) and the Texas Radical and Aerosol Measurement Program (TRAMP) (Luke et al., 2010) are used to evaluate the model performance.

Although the inversions performed in this study are only for Texas NO\textsubscript{x} emissions, the method can be extended to constrain NO\textsubscript{x} in other states or emissions from other species that could be observed from space.

1.3. Research outline

This thesis contains:

**Chapter 2:** Inverse modeling of Texas NO\textsubscript{x} emissions using space-based and ground-based NO\textsubscript{2} observations

**Chapter 3:** Influence of satellite-derived photolysis rates and NO\textsubscript{x} emissions on Texas ozone modeling

**Chapter 4:** Summary of findings in this research and suggestions for the future works

**Appendix A:** Supplementary information for Chapter 2

**Appendix B:** Supplementary information for Chapter 3

**Appendix C:** List of publications and conference presentations
Chapter 2

Inverse modeling of Texas NO$_x$ emissions using space-based and ground-based NO$_2$ observations

This chapter demonstrates the capabilities of using different inversion methods or observations to improve ground-level NO$_2$ and O$_3$ simulations. The NO$_x$ emissions are constrained by both direct scaling and DKF methods using SP2 OMI-observed NO$_2$, and the DKF inversion is also conducted using ground-observed NO$_2$. The DKF inversions yield conflicting results: the OMI-based inversion scales up the a priori NO$_x$ emissions in most regions by factors of 1.02 to 1.84, while the ground-based inversion indicates the a priori NO$_x$ emissions should be scaled down by factors of 0.34 to 0.57 in each region. However, none of the inversions improve the model performance in simulating ground-level O$_3$ concentrations.
2.1. Introduction

Nitrogen oxides (NO\textsubscript{x} = NO + NO\textsubscript{2}) in the troposphere are primary air pollutants, emitted from both anthropogenic sources like fossil-fuel combustion and biomass burning, and natural sources such as soil microbial processes and lightning. NO\textsubscript{x} also acts as a precursor of a secondary air pollutant, tropospheric O\textsubscript{3}, when it reacts with the oxidation products of volatile organic compounds (VOCs) in the presence of sunlight. Oxidation with hydroxyl (OH) radical is the dominant sink of NO\textsubscript{x}, leading to atmospheric nitric acid (HNO\textsubscript{3}) formation. The atmospheric lifetime of tropospheric NO\textsubscript{x} varies from a few hours in summer to a couple of days in winter (Seinfeld and Pandis, 2006).

NO\textsubscript{x} emission inventories used in air quality modeling are typically developed by a bottom-up approach based on estimated activity rates and emission factors for each category. Due to inaccuracies in determining these rates and factors, the uncertainty in NO\textsubscript{x} emission inventories has been suggested to be as high as a factor of two and classified as one of the top uncertainties in ozone simulations and sensitivity analysis (Hanna et al., 2001; Xiao et al., 2010).

Inverse modeling techniques can be used with atmospheric models to estimate model variables that may not be directly measurable (Gilliland and Abbitt, 2001). Inverse modeling generates an optimized ‘top-down’ NO\textsubscript{x} emission inventory for air quality models by minimizing the difference between observed and modeled NO\textsubscript{2} concentrations, providing an opportunity to identify possible biases in the bottom-up NO\textsubscript{x} emission inventory (Napelenok et al., 2008). However, as uncertainties also may be associated
with the measurement data and the inverse methods themselves, inverse modeling has its own limitations. Hence, it is valuable to compare both bottom-up and top-down NOx emission inventories in order to improve the understanding of NOx emissions.

Several inverse modeling studies have used surface NO2 measurements (Mendoza-Dominguez and Russell, 2000; Quélo et al., 2005; Pison et al., 2007) or aircraft NO2 measurements (Brioude et al., 2011) to constrain NOx emissions. Compared to ground and aircraft measurements, satellite-based observations generate greater spatial coverage of NO2. Studies on combining satellite NO2 measurements with inverse modeling techniques to create the top-down NOx emission inventories also have been conducted recently on both global (Martin et al., 2003; Müller and Stavrakou, 2005; Jaeglé et al., 2005; Lin et al., 2010) and regional scales (Konovalov et al., 2006, 2008; Deguillaume et al., 2007; Napelenok et al., 2008; Kurokawa et al., 2009; Zhao and Wang, 2009; Chai et al., 2009).

Discrete Kalman filter (DKF) (Prinn, 2000) is an inverse modeling method that solves the inverse problem iteratively and can be applied to the cases with linear or weakly non-linear relationships between emissions and pollutants. It has been used in several studies to constrain emissions of carbon monoxide (Mulholland and Seinfeld, 1995), chlorofluorocarbons (Haas-Laursen et al., 1996), isoprene (Chang et al., 1996) and ammonia (Gilliland et al., 2003). Most recently, Napelenok et al. (2008) applied the DKF method to the regional Community Multiscale Air Quality (CMAQ) model, generating a top-down NOx emission inventory for the southeastern United States using Scanning
Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMCHY) (Bovensmann et al., 1999) satellite NO₂ data.

Despite the growing number of scientific studies conducting satellite-based inversions of NOₓ emissions, the applicability of these methods to state-level regulatory attainment modeling has not been widely explored. In this work, the DKF method introduced by Napelenok et al. (2008) is applied with finer resolution satellite NO₂ data now available from the Ozone Monitoring Instrument (OMI) as well as ground-level NO₂ observations to constrain NOₓ emissions for an actual regulatory modeling episode in Texas. Lightning and aircraft NOₓ emissions are added to the base case NOₓ emission inventory to address the bias noted by Napelenok et al. (2008) of regional models underestimating upper tropospheric NOₓ. The DKF inverted a posteriori emissions are compared to the base case emissions, the a priori emissions, and a posteriori emissions derived by the inversion method of Martin et al. (2003).

2.2. Methodology

2.2.1. Model inputs and configurations

Base case model inputs were taken from episodes developed by the Texas Commission on Environmental Quality (TCEQ) for Texas ozone attainment planning. CAMx version 5.3 (ENVIRON, 2010) was used in this study to simulate two modeling episodes in 2006 with high ozone concentrations in the Dallas-Fort Worth (DFW) region, from May 31 to July 1, hereafter referred as the June episode, and in the Houston-Galveston-Brazoria (HGB) region, from August 13 to September 15 (Fig. 2.1), hereafter
referred as August-September episode. The NCAR/Penn State (National Center for Atmospheric Research/Pennsylvania State University) Mesoscale Model, Version 5, release 3.7.3 (MM5v.3.7.3) (Grell et al., 1994), conducted with the asymmetric convective model, version 2 (ACM2) scheme for the June episode and the Eta scheme for the August-September episode, was used to generate the meteorological fields with 43 vertical layers. The preprocessor MM5CAMx was used to convert MM5 outputs into CAMx-ready meteorology inputs. The modeled meteorological parameters temperature, wind speed, wind direction, and planetary boundary layer (PBL) height in both episodes are evaluated as shown in the Appendix A, section 1. The vertical configuration of CAMx modeling consists of 17 vertical layers for the August-September modeling episode, whereas 28 vertical layers were used for the June modeling episode. Modeling was conducted with the Carbon Bond version 2005 (CB-05) chemical mechanism, PPM advection scheme, and K-theory vertical diffusion scheme (TCEQ, 2010, 2011). Boundary conditions for the 36-km eastern US domain were generated by the Model for Ozone and Related Chemical Tracers (MOZART) global model (ENVIRON, 2008).
Figure 2-1. 12-km CAMx modeling domain for eastern Texas (black square), inversion regions (shaded), ground AQS NO\textsubscript{2} monitoring sites (blue triangles), and Moody Tower (red circle).

2.2.2. Emission inventory

Base case emission inventories were provided by TCEQ (Table 2.1). The point source emissions were from the State of Texas Air Reporting System (STARS) database which collects emission information from approximately 2000 point sources annually and
from the EPA acid rain database (ARD) which contains emissions from electric generating units (EGUs). The on-road mobile emission inventory was generated by Motor Vehicle Emission Simulator 2010a (MOVES2010a), and the non-road mobile inventory was developed by National Mobile Inventory Model (NMIM) and the Texas NONROAD (TexN) mobile source model. The area source inventory was projected by the EPA Economic Growth Analysis System (EGAS) model based on 2005 emissions from the Texas Air Emissions Repository (TexAER) database. The area emission in the June episode is larger than what shown in the Aug-Sep episode (Table 2.1) because a few oil and gas production sources and drill rig emissions were missed in the area sector of the Aug-Sep episode. The Emission Processing System, version 3 (EPS3) (ENVIRON, 2007) was used for processing the point, mobile, and area emissions to the model-ready format (TCEQ, 2010, 2011). Biogenic emissions were generated by the Global Biosphere Emissions and Interactions System (GloBEIS) biogenics emissions model, version 3.1 (Yarwood et al., 1999), with soil NO\textsubscript{x} emissions estimated by the Yienger and Levy method (Yienger and Levy, 1995).

Lightning and aircraft NO\textsubscript{x} emissions in the upper troposphere were missing in the base case emission inventories and had to be added before conducting inversions. In this study, lightning NO emissions were developed based on National Lightning Detection Network (NLDN) data obtained from Vaisala Inc., following the approach of Kaynak et al. (2008). Intra-cloud lightning flashes were treated as three times of cloud-to-ground lightning flashes with 500 moles NO emission per flash. Lightning NO was placed into the model to match the time and location of NLDN flashes and then distributed vertically based on the profile obtained from the mean April to September
2003-2005 vertical distribution of VHF sources from the Northern Alabama Lightning Mapping Array (Koshak et al., 2004). Global aircraft NO$_x$ emissions of year 2005 in 0.1°×0.1° resolution were obtained from the Emission Database for Global Atmospheric Research (EDGAR) v4.1
(http://edgar.jrc.ec.europa.eu/datasets_grid_list41.php?v=41&edgar_compound=NOx) and mapped to our modeling domain and placed at 9km altitude.

Table 2-1. Categorized a priori NO$_x$ emission rates in inversion region for two modeling episodes.

<table>
<thead>
<tr>
<th>Modeling episodes</th>
<th>Area (tons/day)</th>
<th>Mobile (tons/day)</th>
<th>Non-road (tons/day)</th>
<th>Biogenic (tons/day)</th>
<th>Aircraft (tons/day)</th>
<th>Lightning (tons/day)</th>
<th>Elevated points (tons/day)</th>
<th>Total (tons/day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>June</td>
<td>453</td>
<td>760</td>
<td>374</td>
<td>474</td>
<td>172</td>
<td>434</td>
<td>543</td>
<td>3211</td>
</tr>
<tr>
<td>Aug-Sep</td>
<td>290</td>
<td>766</td>
<td>402</td>
<td>464</td>
<td>171</td>
<td>226</td>
<td>547</td>
<td>2866</td>
</tr>
</tbody>
</table>

2.2.3. Inversion Regions

Five urban areas: Houston-Galveston-Brazoria (HGB), Dallas-Fort Worth (DFW), Beaumont-Port Arthur (BPA), Northeast Texas (NE Texas), and Austin and San Antonio; plus two surrounding rural areas: North Rural area (N rural) and South Rural area (S rural) (Fig. 2.1) were designed as inversion regions for the DKF inversions of NO$_x$ emissions. The five urban regions are all air quality planning areas included in Texas SIP development (Gonzales and Williamson, 2011). HGB and DFW were classified by U.S. EPA as ozone nonattainment areas for violating the 1997 ozone National Ambient Air
Quality Standard (NAAQS) of 84ppb. BPA was designated as an ozone maintenance area, and NE Texas, Austin and San Antonio were designated as ozone early action compact areas under that standard. However, the recent tightening of the NAAQS to 75ppb has heightened interest in ozone reduction in all of these regions. The sensitivities of NO₂ concentrations to boundary conditions and to NOₓ emissions from each inversion region and the border region (the area between model boundary and inversion regions) were computed through DDM. The border region minimizes the impacts from boundary conditions on the inversion regions to the level of only 2%. The DDM sensitivities show that NOₓ emissions from each urban region has the most impact on NO₂ concentrations within that region, and has less than 10% influence on other regions.

2.2.4. Inversion method

Two methods were applied for inverse modeling: a direct scaling method introduced by Martin et al. (2003), and the DKF method. However, the direct scaling method creates spatial smearing errors when applied to regional models with fine resolution. It also assumes concentrations scale proportionally with emissions; hence, the nonlinearity between NO₂ concentrations and NOₓ emissions becomes problematic because NOₓ may influence its own lifetime by influencing concentrations of OH radicals (Martin et al., 2003). Thus, we present the DS method and results in the Appendix A (section 3), and focus our attention on the DKF inversion.

The DKF inversion (Fig. 2.2) solves the spatial smearing problem by taking the spatial relationship between NO₂ concentrations and NOₓ emissions directly from model simulations, and also reduces the non-linearity issue by performing the inversion
iteratively. To constrain NO\textsubscript{x} emissions, the DKF inversion includes two processes at each time step: the measurement update (correction) process and the time update (prediction) process (Rodgers, 2000; Welch and Bishop, 2001). In the measurement update process at time step k (Eqs. 2.1-2.3), the inversion corrects the predicted NO\textsubscript{x} emission (\(E_{NOx,k}^{-}\)) and error covariance (\(P_{NOx,k}^{-}\)) by incorporating the measurement data (\(C_{NO2,k}^{measured}\)) and Kalman Gain (\(G_k\)), and then generates the corrected emission (\(\hat{E}_{NOx,k}\)) and error covariance (\(\hat{P}_{NOx,k}\)).

\[
G_k = P_{NOx,k}^{-} S_k^T (S_k P_{NOx,k}^{-} S_k^T + R_k)^{-1}
\]  

(2.1)

\[
\hat{E}_{NOx,k} = E_{NOx,k}^{-} + G_k (C_{NO2,k}^{measured} - C_{NO2,k}^{modeled})
\]  

(2.2)

\[
\hat{P}_{NOx,k} = (I - G_k S_k) P_{NOx,k}^{-}
\]  

(2.3)

\(S\) represents the NO\textsubscript{2} sensitivity to NO\textsubscript{x} emissions. \(R\) is the measurement error covariance, and it relates to the uncertainties in OMI and ground NO\textsubscript{2} measurements. In here, the uncertainty for the AQS ground NO\textsubscript{2} measurements was set to 0.15 (U.S. EPA, 2006) and for the NASA standard OMI NO\textsubscript{2}, version 2, was set to 0.3 (Bucsela et al., 2013) for all diagonal elements in \(R\). The error covariance (\(P\)) relates to the uncertainty in the NO\textsubscript{x} emission inventory, and the uncertainty value of 2.0 (Napelenok et al., 2008) was chosen here for all diagonal elements in \(P\). To simplify, off-diagonal elements in \(R\) and \(P\) were set to zero, because we assume each inversion region is an independent element.
In the time update process at time step k, the inversion process predicts the emission \( E_{NOx,k+1}^− \) and the error covariance \( P_{NOx,k+1}^− \) for the measurement update process at time step k+1, based on the corrected emission \( \hat{E}_{NOx,k} \) and error covariance \( \hat{P}_{NOx,k} \) from the measurement update process at time step k (Eqs. 2.4-2.5).

\[
E_{NOx,k+1}^− = M_k \hat{E}_{NOx,k} + \epsilon_k
\]  
\[
P_{NOx,k+1}^− = M_k \hat{P}_{NOx,k} M_k^T + Q_k
\]  

\( M \) represents a transition matrix; \( \epsilon \) and \( Q \) are process errors which relate to errors in modeling processes, and are difficult to estimate. Since we assume the bias between modeled and measured NO\(_x\) is mostly from errors in NO\(_x\) emissions (Prinn, 2000; Napelenok et al., 2008), \( \epsilon \) and \( Q \) were set to zero.

CAMx-DDM (Koo et al., 2007) calculates a semi-normalized NO\(_2\) sensitivity to NO\(_x\) emissions (unitless), as shown in Eq. (2.6), replacing sensitivity elements in \( S \) in Eq. (2.1),

\[
S_{NO_2,NOx} = \frac{\partial C_{NO_2}}{\partial E_{NOx}} = \frac{\partial C_{NO_2}}{\partial ((1+x)\hat{E}_{NOx})} = \frac{\partial C_{NO_2}}{\partial (1+x)} = \frac{\partial C_{NO_2}}{\partial x}
\]  

where \( \hat{E} \) represents the unperturbed NO\(_x\) emission field and \( x \) is the perturbation factor. Hence, in this study, the DKF inversion actually seeks the optimal perturbation factor \( x \) at each iteration. The inversion processes will repeat iteratively until the perturbation
factor for each emission region converges within a prescribed criterion, $\delta$ (Fig. 2.2), for which the value of 0.01 was chosen in this study.

![Schematic diagram of Kalman filter inversion process.](image)

**Figure 2-2. Schematic diagram of Kalman filter inversion process.**

### 2.2.5. NO$_2$ observations

#### 2.2.5.1. Satellite NO$_2$ measurements

The Dutch-Finnish Ozone Monitoring Instrument (OMI) aboard NASA’s EOS Aura satellite, launched on 15 July 2004 is a nadir-viewing UV-Vis spectrometer that measures solar backscattered irradiance in the range of 270nm to 500nm. It has been utilized to retrieve atmospheric NO$_2$ in the spectral range from 405nm to 465nm with spatial resolution down to scales of 13×24 km$^2$ at nadir view point (Levelt et al., 2006a, b). The EOS Aura satellite follows a Sun-synchronous polar orbit at approximately
705km altitude with local equator crossing time around 13:40 (Levelt et al., 2006b; Boersma et al., 2007). In this study, the NASA standard product, version 2 (Bucsela et al., 2013) retrieval of OMI NO₂, gridded at 0.1°×0.1° resolution, was obtained from NASA Goddard Space Flight Center and mapped to the 12-km CAMx modeling domain. OMI pixels with cloud radiance fraction greater than 0.5 and sizes of more than 20×63km² at swath edges were excluded in the dataset. The OMI averaging kernels (Eskes and Boersma, 2003) were interpolated into each CAMx model layer and then applied to the modeled NO₂ column density (Eq. 2.7) to account for the influence of the a priori NO₂ vertical profile used in the OMI retrieval and the OMI measurement sensitivities at each altitude:

\[ C_{\text{mod, NO}2} = \sum A_i * X_i \]  

(2.7)

where \( A_i \) is the averaging kernel at pressure level \( i \) and \( X_i \) is the CAMx modeled partial NO₂ sub-column density at the corresponding pressure level.

In order to reduce the OMI measurement uncertainties and effects from invalid data points, monthly averaged OMI NO₂ column densities were used in the DKF inversions.

### 2.2.5.2. Ground and other NO₂ measurements

The U.S. EPA Air Quality System (AQS) NO₂ ground monitoring network data (Fig. 2.1) (http://www.epa.gov/ttn/airs/airsaqs/) were also used for inverse modeling. AQS monitors are equipped with a heated molybdenum catalytic converter that first transforms NO₂ to NO, and then measures the resultant NO using a chemiluminescence
analyzer. NO\textsubscript{2} is then calculated by subtracting NO measured in a separate NO mode from the resultant NO (U.S. EPA, 1975). Studies (U.S. EPA, 1975; Demerjian, 2000; Lamsal et al., 2008) indicate that the catalytic converter also converts fractions of other reactive nitrogen species (e.g., HNO\textsubscript{3}, PAN) into NO during this measurement. Therefore, correction factors computed from CAMx modeled concentrations by the method of Lamsal et al. (2008) (Eq. 2.8) are applied before deploying the AQS NO\textsubscript{2} data in the DKF inversion:

\[
CF = \frac{NO_{2}}{NO_{2} + \sum AN + (0.95PAN) + (0.35HNO_{3})}
\] (2.8)

In Eq. (2.8), \(\sum AN\) represents the sum of all alkyl nitrates and PAN is peroxyacetyl nitrate. The CAMx model with CB05 mechanism does not output alkyl nitrates specifically, so the difference between modeled total organic nitrates and PAN was used to represent modeled alkyl nitrates.

The NOAA P-3 aircraft NO\textsubscript{2} data (http://www.esrl.noaa.gov/csd/tropchem/2006TexAQS/) and the Texas Radical and Aerosol Measurement Program (TRAMP) NO\textsubscript{2} data, measured at Moody Tower (Fig. 2.1), (http://geossun2.geosc.uh.edu/web/blefer/TRAMP/Final%20data/) were used to evaluate the inverse modeling results. The Moody Tower measurement site located at the University of Houston campus is approximately 70m above the ground (Luke et al., 2010), corresponding to the CAMx model layer 2, with hourly averaged NO\textsubscript{2} data available for the whole August-September episode, but no coverage for the June episode. The P-3 aircraft measurement was made from ground level to around 5000m height with
1-second resolution, but data are only available on 4 days (August 31, September 11, September 13, and September 15, 2006) during our modeling period. Hourly averaged aircraft NO$_2$ data was used to compare with the hourly modeled NO$_2$ at corresponding grid cells. Both P-3 aircraft and Moody Tower NO$_2$ measurements were made using a photolytic converter and did not require corrections via Eq. (2.8).

### 2.3. Results and Discussion

#### 2.3.1. Pseudodata test for the DKF inversion with CAMx-DDM

To evaluate the performance of the DKF inversion technique, a controlled pseudodata test was performed for 10 modeling days (May 31 to June 9, and August 13 to August 22) for each modeling episode. The 10-day averaged modeled NO$_2$ columns at 1-2pm from the base case were used as pseudo-observations, and the model was rerun with NO$_x$ emissions from each region perturbed by known factors ranging from 0.5 to 2.0 (Fig. 2.3). Applying the DKF inversion successfully adjusted the perturbed NO$_x$ emissions from each region back to their base values, converging in 4 iterations (Fig. 2.3). The robustness of the DKF inversion was tested by varying the uncertainty parameters, which were set to 2.0 for emissions and 0.3 for observations in the initial pseudodata test. While higher values of the emission uncertainty parameter and lower values of the observation uncertainty parameter led to more rapid adjustments, the final results of the DKF inversion were insensitive to the assumed uncertainty parameters and also to the off-diagonal elements in the error covariance matrix. Similar results were
found by adjusting the assumed uncertainty parameters and error covariance matrix in the actual simulations (Appendix A, Figure A3).

![Graph showing pseudodata test](image)

Figure 2-3. Pseudodata test showing that the DKF inversion accurately adjusts the NO\(_x\) emissions from the perturbed case (a) to the a posteriori case (b) to match the desired base NO\(_2\) column densities. Numbers indicate perturbation factors in the legend (a) and adjustment factors in the legend (b). Similar performance is found for the August 13-22 test period.

### 2.3.2. Additional NO\(_x\) emissions

Since DKF inversions scale emissions from their original levels, an appropriate a priori NO\(_x\) emission inventory is essential for obtaining reasonable results. The NASA Intercontinental Chemical Transport Experiment (INTEX-A) air quality study (Singh et al., 2006) found large discrepancies between aircraft measurements and CMAQ simulations of NO\(_2\) concentrations in the upper troposphere. Possible explanations could
be upper tropospheric NO\textsubscript{x} sources such as lightning and aircraft NO\textsubscript{x} emissions that are often neglected in emission inventories. Missing NO\textsubscript{x} sources in the upper troposphere may bias the inversion on the remaining emissions (Napelenok et al., 2008). At ground level, Hudman et al. (2010) found that the soil NO\textsubscript{x} emissions estimated by the widely used Yienger and Levy method (Yienger and Levy, 1995) were underestimated by a factor of 2 over the United States. Therefore, in this study, the lightning and aircraft NO\textsubscript{x} emissions were added in the upper troposphere as described in the section 2.2.2, and the soil NO\textsubscript{x} emissions were doubled from base case levels (Table 2.1). The emission inventory with added lightning and aircraft NO\textsubscript{x}, and doubled soil NO\textsubscript{x} (hereafter referred to as the a priori emission inventory) was used for the following inversion studies. Inclusion of these NO\textsubscript{x} sources improves the performance of the model in simulating satellite observed NO\textsubscript{2} column densities, especially in the rural areas (Figs. 2.4c and 2.5c), and reduces the bias and error by around 0.15 (Table 2.3).
Figure 2-4. Monthly averaged (June 3 to July 1) tropospheric NO₂ vertical columns at 1-2pm from a) OMI observations and from CAMx simulations using b) base case emissions inventory, c) a priori emission inventory (with additional lightning,
aircraft, and soil NOx), and d) OMI-based inverted NOx emissions using the DKF method.

Table 2-2. Scaling factors for each region from different inversions.

<table>
<thead>
<tr>
<th>Source Region</th>
<th>June 3 to July 1, 2006</th>
<th>August 16 to September 15, 2006</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Base NOx emission (tons/day)</td>
<td>Priori NOx emission (tons/day)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HGB</td>
<td>374</td>
<td>455</td>
</tr>
<tr>
<td>DFW</td>
<td>335</td>
<td>435</td>
</tr>
<tr>
<td>BPA</td>
<td>81</td>
<td>97</td>
</tr>
<tr>
<td>NE Texas</td>
<td>141</td>
<td>164</td>
</tr>
<tr>
<td>Austin and San Antonio</td>
<td>252</td>
<td>319</td>
</tr>
<tr>
<td>N rural</td>
<td>522</td>
<td>823</td>
</tr>
<tr>
<td>S rural</td>
<td>472</td>
<td>728</td>
</tr>
</tbody>
</table>

a. Adds lightning and aircraft NOx, and doubled soil NOx emissions to the base case
b. Conducted with 24-h averaged ground-level NO2 data

2.3.3. Top-down NOx emissions using OMI NO2

DKF inversion using the OMI NO2 measurements was conducted to constrain NOx emissions from the seven designated regions. The monthly averaged (June 3 to July 1, and August 16 to September 15) OMI and CAMx NO2 column densities at 1-2pm were used in the inversion. All modeling grids in the inversion area were covered by the OMI NO2 measurement data. The DKF inversions were performed with 2116 data points in
one time step (1-2pm). The scaling factors generated by inversion for each region were applied to the NO\textsubscript{x} emission inventory hourly, because we assume that the 1-2pm NO\textsubscript{2} column density is contributed by the NO\textsubscript{x} emissions from all previous hours and the uncertainty in the bottom-up NO\textsubscript{x} emission inventory should be the same for every time step. The satellite-based DKF inversions scale a priori NO\textsubscript{x} emissions by factors ranging from 1.02 to 1.84 in almost all regions in both episodes (Table 2.2), adhering to the specified uncertainty range of 0.5 to 2.0 (Hanna et al., 2001). The scaling factors tend to be larger over the rural and small urban regions than over the urban DFW and HGB O\textsubscript{3} nonattainment regions, where the inversions scale up emissions only slightly (factors of 1.02 to 1.14). This results from the inversion attempts to compensate for the large gap between higher observed than modeled NO\textsubscript{2} over rural regions, despite varied patterns over urban grid cells. One exception occurs in the NE Texas region in the August-September episode (Table 2.2), which shows downward scaling (factor of 0.56). This reflects the inversion shifting emissions between NE Texas and the much larger surrounding N rural region (Fig. 2.1); taken together, the net scaling factor for the two regions in the August-September episode is 1.72, consistent with the upward scaling of rural emissions throughout the two episodes. Apart from this anomaly, scaling factors for most regions were consistent across the two episodes, varying by less than 15%.

CAMx modeled NO\textsubscript{2} column densities with the inverted NO\textsubscript{x} emissions (Figs. 2.4d and 2.5d) are increased by 3-55% in all regions, but the increments are much more moderate compared to the DS method inversion (Fig. A4). The statistical results (Table 2.3) indicate that the DKF inverted NO\textsubscript{2} are closer to OMI observations than the a priori case in terms of 0.2 less in bias and 0.1 less in error but without improvements in the
spatial distribution. The DS method scales up NO\textsubscript{x} emissions more than the DKF inversion (Table A2), making the inversed NO\textsubscript{2} concentrations have slightly less bias and error (Table A3). However, the DKF inverse NO\textsubscript{2} has better R\textsuperscript{2} than the DS method, indicating the DKF inversion method has better ability to retain the spatial structure of NO\textsubscript{x} emissions. Each of the inversions using OMI NO\textsubscript{2} data actually worsens the model performance in simulating ground level NO\textsubscript{2} concentrations (Table 2.4), since the modeled ground NO\textsubscript{2} using the base case emission inventory had already been overestimated (Fig. 2.6). Similarly, since the base model already overestimated P-3 aircraft observations of NO\textsubscript{2} and NO\textsubscript{y}, the DKF inversion worsens model bias relative to these measurements (Table 2.5). Greater deterioration resulted from the DS inversion (Tables A3-A6).

### Table 2-3. Performance of CAMx in simulating OMI-observed NO\textsubscript{2} column densities.

<table>
<thead>
<tr>
<th>Statistical Parameters</th>
<th>June 3 to July 1, 2006</th>
<th>August 16 to September 15, 2006</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Base case</td>
<td>Priori\textsuperscript{c}</td>
</tr>
<tr>
<td>R\textsuperscript{2}</td>
<td>0.62</td>
<td>0.61</td>
</tr>
<tr>
<td>NMB\textsuperscript{a}</td>
<td>-0.47</td>
<td>-0.30</td>
</tr>
<tr>
<td>NME\textsuperscript{b}</td>
<td>0.48</td>
<td>0.32</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Normalized mean bias
\textsuperscript{b} Normalized mean error
\textsuperscript{c} Adds lightning and aircraft NO\textsubscript{x} and doubled soil NO\textsubscript{x} emissions to the base case
2.3.4. Top-down NOₓ emissions using ground AQS NO₂

Ground-level AQS NO₂ measurements were also used to drive DKF inversions of NOₓ emissions for the two modeling episodes. There are 37 ground measurement sites in the designated inversion regions (Fig. 2.1), mostly located in the urban cores. The N rural and S rural regions were excluded in this case because they contain too few measurement sites. Correction factors from Eq. (2.8) were applied to the ground NO₂ before using the data in the inversion.

The base case simulations strongly overpredicted observed NO₂ in the early morning and late afternoon during both modeling episodes (Fig. 2.6), when the model may underestimate PBL heights (Kolling et al., 2013). To alleviate the influence from PBL heights, daily 24-h averaged NO₂ levels were used in the inversions.

To address the overprediction of ground-level NO₂, the ground-based inversions sharply reduce a priori NOₓ emissions by applying scaling factors of 0.30 to 0.57 (Table 2.2). The reductions in NOₓ emissions reduce model error relative to the AQS (Table 2.4) and Moody Tower NO₂ observations on an hourly basis, as well as NO₂ and NOₓ observed by the P-3 aircraft (Table 2.5), but the reductions may be too sharp, as they lead to negative bias in predicting NO₂ from the AQS monitors (Table 2.4) and the P-3 aircraft NO₂ measurements (Table 2.5). More moderate scaling factors are obtained if the inversion is conducted with data only from a midday window (9am-2pm) when PBL heights are less problematic (not shown). However, scaling factors still remain far below 1.0 and show up to factor of two inconsistencies between the two episodes.
Figure 2-5. Same as Fig. 2-4 but for the August-September episode.
Figure 2-6. Daily variations of modeled (solid line) and observed (dashed line) ground NO2 concentrations for the June (red) and August-September (blue) episodes. Note: NO2 concentrations were taken by averaging monthly data for all sites.
Table 2-4. Performance of CAMx in simulating AQS Ground-level NO$_2^a$.

<table>
<thead>
<tr>
<th>Statistical Parameters</th>
<th>June 3 to July 1, 2006</th>
<th>August 16 to September 15, 2006</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Base case</td>
<td>Priori</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.56</td>
<td>0.56</td>
</tr>
<tr>
<td>NMB</td>
<td>0.89</td>
<td>0.98</td>
</tr>
<tr>
<td>NME</td>
<td>1.01</td>
<td>1.09</td>
</tr>
</tbody>
</table>

$^a$. Hourly AQS data was used to compare with modeled NO$_2$ at corresponding locations.

Table 2-5. Performance of CAMx in simulating P-3 aircraft-observed NO$_2$ and NO$_y$.

<table>
<thead>
<tr>
<th>Statistical Parameters</th>
<th>NO$_2^a$</th>
<th>NO$_y^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Base case</td>
<td>Priori</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.23</td>
<td>0.34</td>
</tr>
<tr>
<td>NMB</td>
<td>0.10</td>
<td>0.65</td>
</tr>
<tr>
<td>NME</td>
<td>0.99</td>
<td>0.97</td>
</tr>
</tbody>
</table>

$^a$. Comparison available for only four days (August 31, September 11, September 13, and September 15, 2006).

2.3.5. Impacts on O$_3$ simulations

O$_3$ concentrations and their sensitivities to changes in emissions are calculated for both modeling episodes using the a priori and each of the a posteriori emission inventories. The scaled up NO$_x$ emissions from the satellite-based DKF inversion (Table 2.2) lead to 1-7 ppb higher modeled 8-h (10 am-6 pm) O$_3$ concentrations over most of the domain in the June episode (Fig. 2.7, top row). Largest increases occur over NE Texas.
and N rural regions (Fig. 2.1), where the a priori simulation shows O₃ to be most sensitive to NOₓ (Fig. 2.7, middle row) and where the satellite-based DKF inversion scaled up emissions by large amounts.

The a priori simulation shows O₃ to be primarily sensitive to NOₓ over most of the domain, but VOC-limited in the core of the Houston region and with joint sensitivity to NOₓ and VOC in Dallas, Austin, and San Antonio (Fig. 2.7, left column). The satellite-based inversion increases NOₓ emissions and thus shifts the O₃ formation chemistry toward being more VOC sensitive (Fig. 2.7, middle column). Over much of the domain, O₃ sensitivity to VOC increases by a factor of about 1.5. The slight increases in O₃ sensitivity to NOₓ occur because the semi-normalized sensitivity coefficients represent the local slope of O₃-emissions response scaled to a 100% change in emissions. As the satellite-based inversion scales up NOₓ emissions, these semi-normalized coefficients increase, even though the impacts per ton of NOₓ decrease.

The ground-based DKF inversion leads to O₃ reductions of 3-8 ppb over urban regions (Fig. 2.7, top right), where it scales down emissions (Table 2.2) and smaller changes over rural regions where emissions were left unchanged due to lack of NO₂ monitors. The reduction in urban NOₓ makes O₃ less sensitive to VOC emissions as expected (Fig. 2.7, bottom right). However, the impact on sensitivity to NOₓ is mixed. In urban areas which are transitional between NOₓ-limited and NOₓ-saturated conditions, the reduction in NOₓ emissions pushes the chemistry toward more NOₓ-limited conditions and thus increases the sensitivities. In downwind regions which are already
NO\textsubscript{x}-limited, the sensitivities decline because there are now less NO\textsubscript{x} emissions contributing to the semi-normalized coefficients.

Model performance in simulating hourly AQS ground-level observations of O\textsubscript{3} indicates that the bias and error slightly worsened when each of the a posteriori inventories are used in place of the a priori inventory (Table 2.6). The largest deterioration comes from the DS inversion as the bias and error increase by around 0.1 (Table A5), likely because this inversion method does not retain the spatial structure of emissions from the a priori inventory. For the other inversions, the changes in bias and error are too slight to determine if performance is meaningfully impacted.

**Table 2.6. Performance of CAMx in simulating AQS hourly ground-level O\textsubscript{3}.**

<table>
<thead>
<tr>
<th>Statistical Parameters</th>
<th>June 3 to July 1, 2006</th>
<th>August 16 to September 15, 2006</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Priori</td>
<td>Posteriori OMI-based DKF inversion</td>
</tr>
<tr>
<td>R\textsuperscript{2}</td>
<td>0.61</td>
<td>0.63</td>
</tr>
<tr>
<td>NMB</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>NME</td>
<td>0.29</td>
<td>0.30</td>
</tr>
</tbody>
</table>
Figure 2-7. Monthly 8-h (10am-6pm) averaged ground-level O₃ concentrations (top), O₃ sensitivity to NOₓ (middle), and O₃ sensitivity to VOC (bottom) for the a priori case (left column), and differences (a posteriori minus a priori) for the OMI-based (middle column) and ground-based (right column) DKF inversions in the June episode. The August-September episode shows similar results.
2.4. Conclusions

Inverse modeling has been performed using either NO$_2$ column densities observed by OMI satellite or ground-level NO$_2$ concentrations observed by AQS monitors to constrain the NO$_x$ emissions for two regulatory attainment modeling episodes in Texas. Two inversion methods, DS and DKF, are applied to the OMI NO$_2$ data, and the DKF method also is applied to the ground-level NO$_2$ data. Pseudodata test results validate that the DKF method effectively captures known perturbations in CAMx simulations.

Two missing NO$_x$ sources in the upper troposphere, lightning and aircraft NO$_x$ emissions, are added into the base case NO$_x$ emission inventory, contributing 14% and 6% to the total NO$_x$ emissions for the June episode, and 7% and 6% for the August-September episode, respectively. The underestimated soil NO$_x$ emissions are doubled from the base case, adding an additional 8% NO$_x$ emission to the base case for both episodes. The additional NO$_x$ emissions increase the modeled NO$_2$ column densities mostly in rural areas and improve the inversion performance with the OMI NO$_2$ but not with the ground NO$_2$.

The DS method originally was pursued to provide an alternate approach featuring more spatial heterogeneous adjustments to emissions. However, it tends to overshoot the OMI-observed NO$_2$ column densities because this linear inversion method ignores the nonlinear influence of NO$_x$ on its own lifetime. The iterative approach of the DKF inversion avoids this problem but fails to substantially improve the spatial
correlation of modeled and observed NO₂ levels because it applies only a single scaling
factor to each inversion region.

The overall tendency of the model to underpredict OMI observed NO₂ column
densities and to overpredict AQS observed ground NO₂ concentrations leads to
conflicting results between the inversions. It is difficult to determine which observations
provide a more reliable basis for the inversions because none of the inversions improves
model performance against independent data such as aircraft-observed NO₂ or ground-
level O₃ concentrations. Whether this indicates that the a priori inventory is the best
available representation of NOₓ emissions or that tuning of the base model led to its
better performance is impossible to determine. Nevertheless, this suggests that inverse
modeling of NOₓ emissions should for now remain a complement to SIP modeling efforts
rather than a substitute for traditional bottom-up inventories.

The AQS ground NO₂ measurements face limitations due to the inaccuracies of
the molybdenum converter method. Furthermore, the mostly urban locations of
measurement sites may be unrepresentative of the entire region and do not capture the
rural areas where OMI observations suggest NO₂ is underestimated. In addition, model
shortcomings in simulating PBL heights in the early morning and late afternoon may
contribute to the low scaling factors in the ground-based inversions.

For the satellite data, several factors could explain the more spatially smeared
and higher rural NO₂ in the satellite observations than the base model that drove the
upward scaling of emissions. Our inclusion of lightning and aircraft NOₓ emissions and
doubling of soil NOₓ emissions narrowed but did not eliminate the discrepancy. A higher
resolution OMI NO$_2$ product (retrieved with small pixels and high resolution a priori profile) has been shown to enhance NO$_2$ column densities in urban areas and reduce them in rural areas (Russell et al., 2011), which would more closely resemble the modeled distribution. Lin et al. (2012) highlighted several uncertain model parameterizations that impact model predictions of NO$_2$ column density for a given emission inventory. For example, lowering the rate constant of the NO$_2$ + OH reaction to match the rate of Mollner et al. (2010) would lead to a longer NO$_x$ lifetime and reduce the gap between modeled urban and rural NO$_2$ concentrations. Henderson et al. (2011) suggested that better representation of acetone and organic nitrates in the CB05 mechanism could help address its underprediction of NO$_2$ in the remote upper troposphere. Future work could explore how combinations of these adjustments influence satellite-based inversions.

The DISCOVER-AQ campaign by NASA in fall 2013 provides vertically resolved measurements of NO$_x$ from repeated aircraft spirals in the Houston region. This may help resolve some of the discrepancies noted here between inversions driven by ground-based and satellite-based NO$_2$ observations. The future Tropospheric Emissions: Monitoring of Pollution (TEMPO) mission using a geostationary satellite with high spatial and temporal measurement capabilities could provide a richer data source to drive NO$_x$ inversions. Future work also could conduct inversions based on emission categories rather than emission regions to explore potential errors in the emission inventory on a component rather than location basis.
Chapter 3

Influence of satellite-derived photolysis rates and NO$_x$ emissions on Texas ozone modeling

The study in this chapter investigates the impact of satellite-derived photolysis rates and NO$_x$ emissions on ground-level NO$_2$ and O$_3$ simulations. First of all, GOES-retrieved cloud fractions and transmissivity are used to correct modeled photolysis rates. Results show that GOES-derived photolysis rates reduce photochemical activity over most of the domain and improve the ground-level O$_3$ simulations. Secondly, the DKF inversions are performed to constrain NO$_x$ emissions in different regions and sources using high resolution SP2 OMI-observed NO$_2$. The sector-based DKF inversion decreases the area and nonroad NO$_x$ emissions in the bottom-up NO$_x$ emission inventory and improves model performance in simulating ground-level NO$_2$ and O$_3$. In addition, the direct inversion of five primary VOC species emissions is also explored. However, inversion results show no improvements in ground-level O$_3$ modeling.
3.1. Introduction

Tropospheric O₃ is a secondary air pollutant formed via the reactions between nitrogen oxides (NOₓ = NO + NO₂) and volatile organic compounds (VOCs) with heat and sunlight (Seinfeld and Pandis, 2006). Eastern Texas is one of the most populous areas in the United States and has been suffering from O₃ pollution for decades due to large anthropogenic emission sources such as motor vehicles, petrochemical facilities, and coal-burning power plants with unique meteorological conditions of extended heat and humidity and intense solar radiation (Kleinman et al., 2002; Ryerson et al., 2003; Daum et al., 2004; Rappenglück et al., 2008; Kim et al., 2011; Zhou et al., 2012).

In eastern Texas, several regions require careful air quality planning for O₃ reductions. First and foremost, the Houston-Galveston-Brazoria (HGB) region and the Dallas-Fort Worth (DFW) region are both O₃ non-attainment areas classified by U.S Environmental Protection Agency (U.S EPA) with O₃ design values exceeding 1997 O₃ National Ambient Air Quality Standard (NAAQS) of 84ppb. Besides that, Beaumont-Port Arthur (BPA) is designated as an O₃ maintenance area, and Northeast Texas (NE Texas), Austin and San Antonio are designated as O₃ early action compact areas by Texas Commission on Environmental Quality (TCEQ) based on their being close to violating the 84ppb O₃ standard (TCEQ, 2010, 2011; Gonzales and Williamson, 2011). In addition, the recent tightening of O₃ standard from 84ppb to 75ppb makes ozone attainment in eastern Texas more challenging.

To comply with the O₃ NAAQS, TCEQ is required by U.S EPA to develop attainment strategies to show emission control efforts for reducing O₃ in two non-
attainment areas using photochemical air quality models. However, model uncertainties may impair the accuracy of model performance and misdirect emission control strategies (Fine et al., 2003; Digar and Cohan, 2010; Simon et al., 2012). Recent studies show that uncertain bottom-up emission inventories and modeled photolysis rates are two leading uncertainties in O₃ modeling (Deguilaume et al., 2007; Digar et al., 2011) and can significantly impact simulated O₃ concentrations and their sensitivities (Cohan et al., 2010; Xiao et al., 2010). Hence, identifying and reducing these uncertainties are essential to ensuring reliability of regulatory decision making.

Direct measurements of emissions and photolysis rates are spatially limited and impractical to perform covering the entire modeling domain. However, satellite-based measurements provide an invaluable opportunity to observe some atmospheric parameters and air pollutants from space and generate a rich measurement dataset with great spatial coverage. Pour-Biazar et al. (2007) used the GOES-based cloud information to reproduce photolysis rates in the Community Multiscale Air Quality (CMAQ) model. Results showed large differences between model-predicted and satellite-derived photolysis rates, leading to significant changes in modeled O₃ concentrations. Guenther et al. (2012) found that Weather Research and Forecasting (WRF) and MM5, which are usually used to generate meteorological fields for CAMx or CMAQ, underpredict cloud fractions and lead to more modeled solar radiations reaching the ground, causing overestimation of modeled photolysis rates.

Studies using satellite NO₂ measurements to create top-down NOₓ emissions for atmospheric modeling also have been conducted and show promising results (Streets et
Most recently, Tang et al. (2013) performed region-based DKF inversions using OMI NO2 data to adjust NOx emission inventory used in Texas SIP modeling; however, results show that the region-based DKF inversions with National Aeronautics and Space Administration (NASA) OMI NO2 standard product, version 2, tended to scale up the NOx emission inventory by factors of 1.02 to 1.84 and deteriorated the model performance of ground NO2 and O3 simulations. Tang et al. (2013) found that the NO2 gap (difference in NO2 VCDs) between OMI and the model that occurred in rural areas was alleviated but not eliminated by adding lightning, aviation, and extra soil NOx emissions, which limited the capabilities of the region-based DKF inversions to adjust accurately the ground anthropogenic NOx emissions.

The underprediction of upper tropospheric NO2 in the model contributes largely to that NO2 gap in the remote areas. Several studies (Hudman et al., 2007; Henderson et al., 2011; Allen et al., 2012; ENVIRON, 2013) have demonstrated that the upper tropospheric NO2 underestimation still remains after adding lightning and aviation sources. It also could result from uncertainties in model chemistry of simulating NOx sinks (Mollner et al., 2010; Henderson et al., 2012, Lin et al., 2012, Stavrakou et al., 2013). ENVIRON (2013) recently has done substantial work to address the low bias in the modeled upper tropospheric NO2 by increasing the vertical transport, adding transport of NOy from the stratosphere to the upper troposphere, and revising NOx sinks chemistry suggested by the above references; however, comparisons between modeled NO2 and the NASA DC-8 flight measured NO2 from the Intercontinental Chemical Transport Experiment-North America (INTEX-NA) field campaign (Singh et al., 2006) showed the low bias still remains.
Besides improving the NO\textsubscript{2} simulations in the model, high resolution OMI NO\textsubscript{2} retrievals (Russell et al., 2011) have been shown to yield greater contrast between NO\textsubscript{2} in urban and rural areas, which may further reduce the gap between models and satellite observations. In addition, the spatial relationship between modeled NO\textsubscript{2} concentrations and NO\textsubscript{x} emissions in each emission sector can be tracked by sensitivity analysis methods such as DDM. Hence, a sector-based DKF inversion could potentially serve as an alternative approach to adjust NO\textsubscript{x} emissions in order to avoid applying a single scaling factor to an entire inversion region.

In addition to improving NO\textsubscript{x} emissions, an accurate VOC emission inventory also is important for Texas O\textsubscript{3} modeling and NO\textsubscript{x} inversion studies. In eastern Texas, VOC emissions contribute significantly to O\textsubscript{3} formation, especially in the HGB region with high petrochemical activities (Kleinman et al., 2002; Murphy and Allen, 2005; Nam et al., 2006; Webster et al., 2007; Vizuete et al., 2008). However, large uncertainties were found in the Texas VOC emission inventory during two intensive measurement campaigns, Texas Air Quality Study (TexAQS) 2000 and 2006, that reported highly reactive VOC (HRVOC) emissions were underestimated up to an order of magnitude (Ryerson et al., 2003; Wert et al., 2003; Jiang and Fast, 2004; Gilman et al., 2009; de Gouw et al., 2009; Parrish et al., 2009; Washenfelder et al., 2010). Byun et al. (2007) directly multiplied Texas HRVOC inventory values by factors of 3 to 12, and Kim et al. (2011) reconstructed HRVOC emissions in the 2005 National Emission Inventory using Solar Occultation Flux measurements, with both showing improved O\textsubscript{3} simulations over the Houston area.
In this work, first, the influence of GOES-derived photolysis rates on the modeled NO₂ and O₃ is investigated. Second, the model shortcomings of underestimating upper tropospheric and rural NO₂ demonstrated in Tang et al. (2013) are further improved by comparing with aircraft measurements and increasing modeled NOₓ lifetime. Finally, the applicability of the sector-based DKF inversion using the recently developed NASA high resolution OMI product to Texas NOₓ emissions is explored. In addition, emissions of five VOC species in the Texas VOC emission inventory are adjusted via directly comparing modeled concentrations with ground observations.

3.2. Methodology

3.2.1. CAMx modeling

CAMx version 5.3 (ENVIRON, 2010) with the Carbon Bond version 2005 (CB-05) chemical mechanism was used to simulate a SIP modeling episode developed by TCEQ for the HGB O₃ attainment demonstration (Fig. 3.1) from 13 August to 15 September 2006, which was covered by the intensive measurement campaign TexAQS 2006. The meteorology fields were modeled by the NCAR/Penn State (National Center for Atmospheric Research/Pennsylvania State University) Mesoscale Model, Version 5, release 3.7.3 (MM5v.3.7.3) (Grell et al., 1994), and the boundary conditions were taken from the Model for Ozone and Related Chemical Tracers (MOZART) global model (ENVIRON, 2008). The base case emission inventory for HGB SIP modeling was provided by TCEQ (TCEQ, 2010). Lightning and aviation NOₓ emissions were added into the base emission inventory, and soil NOₓ emission was doubled from its base value
because the Yienger and Levy method (YL95) (Yienger and Levy, 1995) has been found to underpredict soil NO\textsubscript{x} by around a factor of 2 over the United States (Hudman et al., 2010). More details about the model inputs and configurations, the emission inventory development, and evaluations of model meteorological inputs can be found in Tang et al. (2013).

Figure 3-1. Seven designated inversion regions in eastern Texas (shaded) within 12-km CAMx modeling domain (black square) covered by ground NO\textsubscript{2} monitoring sites (blue triangles), VOC monitoring sites (green circles), and O\textsubscript{3} monitoring sites (red circles).
3.2.2. GOES-derived photolysis rates

The photolysis rate calculations in CAMx include two steps (ENVIRON, 2010). First, a Tropospheric Ultraviolet and Visible (TUV) Radiation Model developed by the National Center for Atmospheric Research (NCAR) is used to generate a multi-dimensional table of clear sky photolysis rates (Madronich, 1987) as inputs for the CAMx model as shown in Eq. (3.1).

Clear sky photolysis rates \((s^{-1})\) are calculated as:

\[ J = \int_{\lambda_i}^{\lambda_f} \sigma(\lambda) \phi(\lambda) F(\lambda) d\lambda \] \hspace{2cm} (3.1)

where \(\sigma(\lambda)\) (\(m^2/molecule\)) is the absorption cross-section, \(\lambda\) is the wavelength (\(\mu m\)), \(\phi(\lambda)\) is the quantum yield (\(molecules/photon\)), and \(F(\lambda)\) is the actinic flux (\(photons/m^2/s/\mu m\)).

Second, the tabular clear sky photolysis rates are interpolated into each grid cell in the modeling domain and adjusted based on cloud information generated by the meteorology model, as shown in Eqs. (3.2) and (3.3). Within and below the cloud, photolysis rates are adjusted as (Chang et al., 1987):

\[ J_{\text{within/below}} = J_{\text{clear}} \left[1 + f_c (1.6 \cdot t_r \cos(\theta) - 1)\right] \] \hspace{2cm} (3.2)

Above the cloud, photolysis rates are modified as:
\[ \begin{align*}
J_{\text{above}} &= J_{\text{clear}} \left[ 1 + f_c \cos(\theta)(1 - tr_c) \right] \\
\end{align*} \]

(3.3)

where \( f_c \) is the cloud fraction for a grid cell, \( tr_c \) is cloud transmissivity at each model grid layer, and \( \theta \) is the solar zenith angle.

In CAMx, the \( tr_c \) is calculated using Eq. (3.4) (Stephens, 1978),

\[ tr_c = \frac{5 - e^{-\tau_c}}{4 + 3\tau_c(1 - \beta)} \]

(3.4)

where \( \tau_c \) is the cloud optical depth simulated in the model and \( \beta \) is the scattering phase-function asymmetry factor assumed to be 0.86 (Chang et al., 1987). The \( f_c \) in each grid cell is predicted by the MM5 model.

GOES-observed cloud properties can be used to recover \( f_c \) and broadband \( tr_c \) to adjust photolysis rates used in Eqs. (3.2) and (3.3). This technique bypasses the need for estimating \( tr_c \) by Eq. (3.4) and uses the observations directly. GOES is capable of measuring cloud properties with spatial resolution down to 1-km and temporal resolution down to an hour or less (Haines et al., 2004), ensuring the sufficient spatial and temporal data coverage for the modeling episode. In this study, hourly GOES observations with integrated 12km cloud properties have been used. However, due to the satellite data availability, satellite-retrieved \( f_c \) and broadband \( tr_c \) may not be available in the early morning and late afternoon. In such cases, the \( f_c \) and \( tr_c \) generated by standard operational procedures in CAMx will be used. The photolysis rates within cloud are adjusted via the interpolation of values between satellite-retrieved cloud top and model-estimated cloud
base. More details regarding satellite retrievals of \( f_c \) and \( tr_c \) can be found at Pour-Biazar, et al. 2007.

### 3.2.3. Emission regions and sectors for the inversion

As in Tang et al. (2013), an inversion region inside the 12km model domain is designed for both region-based and sector-based DKF inversions, including five urban areas HGB, DFW, BPA, NE Texas, and Austin and San Antonio surrounded by a north rural area (N rural) and a south rural area (S rural) (Fig. 3.1).

Six separate NO\(_x\) emission sectors, area, non-road mobile (non-road), on-road mobile (mobile), biogenic, electric generating units (EGU) and non-EGU point sources are provided by TCEQ. Lightning and aviation NO\(_x\) emission sectors were developed in Tang et al. (2013) and added into base emission inventory as independent elevated sources (Table 3.1). Area sources, including small-scale industry and residential sources such as oil and gas production, gas stations and restaurants, contribute 10% of total emissions in the entire inversion region and 25% in NE Texas. Non-road sources, including construction equipment, locomotives and commercial marine, contribute 14% of total emissions. Mobile source emissions by on-road vehicles contribute 27% of total NO\(_x\) emissions and dominate in the cities such as HGB and DFW. The biogenic NO\(_x\) source is from soil emissions, which contribute 16% of total NO\(_x\) emissions but dominate in remote regions. Lightning and aviation sources contribute 8% and 6% to the total emission, respectively. Non-EGU point sources such as refineries, big boilers and flares, contribute 40% of NO\(_x\) emissions in BPA and 21% in HGB, the two regions with most of the petrochemical industries. EGU point emissions are from major power plants with the
hourly NO\textsubscript{x} emissions measured by continuous emissions monitoring (CEM) systems, which are considered the most accurate NO\textsubscript{x} emission source in the bottom-up emission inventory. Thus, in this study, EGU NO\textsubscript{x} emissions are assumed to be correct and are not adjusted by DKF inversions.

NO\textsubscript{2} sensitivities to NO\textsubscript{x} emission in each emission sector used in the following sector-based DKF inversions are calculated through DDM (Fig. 3.5). The biogenic, lightning, and non-EGU point sources have their own spatial patterns that differ from the other emission sectors. For example, the aviation source shows strong sensitivity centered from the DFW and HGB regions and slowly spreading elsewhere. The sensitivities from the area, non-road and mobile sources have similar spatial patterns concentrated in the urban areas due to strong anthropogenic activities, while the mobile source can be distinguished by the strong highway emissions. Previous studies (Rodgers, 2000; Curci et al., 2010) indicated that the inversion results would be ill-conditioned to estimate strongly overlapped sources. Therefore, in this study, the area and non-road sources are treated together as the same sector in the DKF inversions.

3.2.4. DKF inversion

Two inversion approaches, region-based and sector-based DKF inversions, are applied in this study to create top-down NO\textsubscript{x} emissions for Texas. The procedure of incorporating DKF method into the CAMx-DDM model was described in detail in Tang et al. (2013).
The DKF inversion process (Prinn 2000), driven by the difference between the measured NO$_2$ ($C_{NO_2}^{\text{observed}}$) and the modeled NO$_2$ ($C_{NO_2}^{\text{predicted}}$) seeks the optimal emission perturbation factors ($\hat{x}$) (a posteriori) by adjusting NO$_x$ emissions in each designated emission region or sector iteratively until each a priori emission perturbation factor ($x^-$) converges within a prescribed criterion, 0.01.

$$\hat{x}_{NO_x} = x_{NO_x}^- + P_{NO_x}^T (S P_{NO_x}^T S + R_{omi})^{-1} (C_{NO_2}^{\text{observed}} - C_{NO_2}^{\text{predicted}} - S x_{NO_x}^-)$$  (3.5)

$S$ in Eq. (3.5), calculated via DDM in this study, is the first-order semi-normalized NO$_2$ sensitivity matrix to either region-based or sector-based NO$_x$ emissions. The uncertainty value in the measurement error covariance matrix ($R$) for the OMI observed NO$_2$ is set to 30% (Bucsela et al., 2013) for all diagonal elements. The uncertainties adopted from Hanna et al. (2001) provide the values for each of the diagonal elements in the emission error covariance matrix ($P$). A value of 100% is assigned to each emission region, and to the area, non-road, mobile, and biogenic emission sectors, but a value of 50% is assigned to the non-EGU point emission sector. The uncertainty for aviation emission is set equal to that for other non-road sources. The lightning NO$_x$ emission is developed based on the measured National Lightning Detection Network (NLDN) data with intra-cloud flashes assumed to be three times of cloud-to-ground flashes and 500 moles NO emissions per flash (Kaynak et al., 2008). The uncertainty of lightning NO$_x$ emissions was estimated in recent studies, ranging from 30% (Martin et al., 2007) to 60% (Schumann and Huntrieser, 2007) on a global scale; thus, the uncertainty value in the
lightning sector is set to 50% here. The off-diagonal elements in $P$ are set to zero since each emission component is assumed to be independent.

Table 3-1. NO$_x$ emission rates for seven sectors in seven inversion regions.

<table>
<thead>
<tr>
<th>Source Region</th>
<th>Area (tons/day)</th>
<th>Mobile (tons/day)</th>
<th>Non-road (tons/day)</th>
<th>Biogenic (tons/day)</th>
<th>Aviation (tons/day)</th>
<th>Lightning (tons/day)</th>
<th>Non-EGU points (tons/day)</th>
<th>EGU (tons/day)</th>
<th>Total (tons/day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HGB</td>
<td>28 (6%)</td>
<td>159 (36%)</td>
<td>71 (16%)</td>
<td>10 (2%)</td>
<td>28 (6%)</td>
<td>21 (5%)</td>
<td>92 (21%)</td>
<td>29 (7%)</td>
<td>438</td>
</tr>
<tr>
<td>DFW</td>
<td>35 (8%)</td>
<td>152 (37%)</td>
<td>77 (19%)</td>
<td>60 (14%)</td>
<td>44 (11%)</td>
<td>23 (6%)</td>
<td>19 (5%)</td>
<td>6 (1%)</td>
<td>416</td>
</tr>
<tr>
<td>BPA</td>
<td>8 (8%)</td>
<td>24 (24%)</td>
<td>7 (7%)</td>
<td>2 (2%)</td>
<td>3 (3%)</td>
<td>8 (8%)</td>
<td>40 (40%)</td>
<td>8 (8%)</td>
<td>101</td>
</tr>
<tr>
<td>NE Texas Austin and San Antonio</td>
<td>43 (25%)</td>
<td>34 (20%)</td>
<td>28 (16%)</td>
<td>2 (1%)</td>
<td>3 (2%)</td>
<td>14 (8%)</td>
<td>9 (5%)</td>
<td>41 (24%)</td>
<td>174</td>
</tr>
<tr>
<td>N Rural</td>
<td>9 (3%)</td>
<td>113 (37%)</td>
<td>37 (12%)</td>
<td>72 (24%)</td>
<td>12 (4%)</td>
<td>5 (2%)</td>
<td>21 (7%)</td>
<td>34 (11%)</td>
<td>303</td>
</tr>
<tr>
<td>S Rural</td>
<td>85 (13%)</td>
<td>123 (18%)</td>
<td>79 (12%)</td>
<td>176 (26%)</td>
<td>30 (4%)</td>
<td>61 (9%)</td>
<td>57 (8%)</td>
<td>672</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>290 (10%)</td>
<td>766 (27%)</td>
<td>402 (14%)</td>
<td>464 (16%)</td>
<td>171 (6%)</td>
<td>226 (8%)</td>
<td>281 (10%)</td>
<td>266 (9%)</td>
<td>2866</td>
</tr>
</tbody>
</table>

Note: percentage indicates the apportionment of each emission sector to the regional total.

Table 3-2. Scaling factors of region-based and sector-based inversions.

<table>
<thead>
<tr>
<th>Region-based inversion</th>
<th>Sector-based inversion I</th>
<th>Sector-based inversion II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emission region</td>
<td>Scaling factor (unitless)</td>
<td>Scaling factor (unitless)</td>
</tr>
<tr>
<td>HGB</td>
<td>1.11</td>
<td>Area</td>
</tr>
<tr>
<td>DFW</td>
<td>0.97</td>
<td>Non-road</td>
</tr>
<tr>
<td>BPA</td>
<td>1.49</td>
<td>Mobile</td>
</tr>
<tr>
<td>NE Texas Austin and San Antonio</td>
<td>1.10</td>
<td>Biogenic</td>
</tr>
<tr>
<td>N rural</td>
<td>1.24</td>
<td>Aviation</td>
</tr>
<tr>
<td>S rural</td>
<td>0.98</td>
<td>Lightning</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-EGU points</td>
</tr>
</tbody>
</table>
3.2.5. NO$_2$ observations

3.2.5.1. Satellite NO$_2$ observations

The Dutch-Finnish OMI aboard the NASA Aura satellite measures daily atmospheric NO$_2$ signal at around 13:40 LT with the highest spatial resolution of 13×24 km$^2$ at nadir viewpoint (Levelt et al., 2006a, b; Boersma et al., 2007). Tang et al. (2013) used the NASA OMI standard, version 2 (Bucsela et al., 2013) NO$_2$ retrieval with an a priori profile generated from the Global Modeling Initiative (GMI) model to conduct inverse modeling. That retrieval may overestimate NO$_2$ levels in rural areas based on comparisons with the Berkeley High-Resolution (BEHR) product (Russell et al., 2011), which uses an a priori profile generated from the 4km × 4km WRF-Chem model. Most recently, a high resolution OMI NO$_2$ retrieval is developed based on the NASA standard product, version 2.1, but using an a priori NO$_2$ profile generated from nested GEOS-Chem simulations (0.5°×0.666°) with a 2005 emission inventory. Because the emission inventory used in GEOS-Chem simulations includes lightning and other elevated sources, it may better represent the upper tropospheric NO$_2$ in the retrieval; hence, in this study, the high resolution NASA retrieval is chosen for the following DKF inversions. In the high resolution NASA product, only the OMI pixels with sizes less than 16×40km$^2$ (scan position 10-50) in the clear-sky condition (cloud radiance fraction < 0.5) are selected in creating the gridded data at 0.1°×0.1° resolution and then mapped to the 12km CAMx modeling domain. Because the a priori profile used in the high resolution retrieval is generated from a nested GEOS-Chem simulation with updated emissions, the OMI
uncertainty that arises from the a priori profile is reduced and less of concern. The OMI averaging kernels (Eskes and Boersma, 2003) are not applied to the CAMx modeled NO2 VCD here (Appendix B sect. 1).

3.2.5.2. Ground and P-3 aircraft NO2 observations

The CAMx simulated NO2 is evaluated by both ground and aircraft measurements. The ground-level NO2 measurements data are taken from the U.S. EPA Air Quality System (AQS) NO2 ground monitoring network (Fig. 3.1) (http://www.epa.gov/ttn/airs/airsaqs/). The correction factors (Lamsal et al., 2008; Tang et al., 2013) are applied to the ground measured NO2 before comparing with the modeled results due to the measurement artifacts in the heated molybdenum catalytic converter used by AQS NO2 monitors.

The NOAA P-3 aircraft measurements (http://www.esrl.noaa.gov/csd/tropchem/2006TexAQS/) are available on 31 August, 11 September, 13 September, and 15 September 2006 in our modeling period. The NO2 was measured by UV photolysis converter-chemiluminescence (Ryerson et al., 2000), and NOy was measured by Au converter-chemiluminescence (Ryerson et al., 1999) aboard the P-3 aircraft, from ground to approximately 5km aloft and with a time resolution of 1-second; thus, hourly averaged P-3 NO2 and NOy are calculated to compare with the modeled data at corresponding time and grid cells.
3.2.5.3. INTEX-NA NASA DC-8 flight NO₂ observations

The NO₂ measured by NASA DC-8 flights (http://www-air.larc.nasa.gov/cgi-bin/arcstat) during the INTEX-NA field campaign in 2004 (Singh et al., 2006) is used in this study to validate the modeled NO₂ vertical profile, especially in the upper troposphere. The DC-8 flight NO₂ measurements were made on a total of 18 days from 1 July to 14 August 2004, spanning from 7:00 to 20:00 CST with 1-second resolution. The NO₂ was measured by Thermal Dissociation-Laser Induced Fluorescence (TD-LIF) instrument which is interfered by methyl peroxy nitrate (CH₃O₂NO₂) and HO₂NO₂, depending on temperature; thus, corrections based on the method of Browne et al., (2011) are applied before comparing with the modeled profile. The modeled NO₂ in grid cells within the 36km domain are used to match the measurement data in space, and then all measurement data at each model layer are averaged over all measurement time to compare with the monthly 12-h (7:00-20:00LT) averaged modeled data at the corresponding layer. Although the measurements took place in 2004 and our modeling period is in 2006, we ignore inter-annual variation of upper tropospheric NO₂.

3.3. Results and discussion

3.3.1. Impact of GOES-derived photolysis rates on modeled NO₂ and O₃

The GOES-retrieved cloud fractions and broadband transmissivity as described in sect 3.2.2 are used to adjust the photolysis rates in CAMx. To investigate the impact from GOES-derived photolysis rates, the differences of modeled ground-level NO₂ photolysis rate (J_{NO2}), NO₂, and O₃ between CAMx modeling with or without the GOES-retrieved...
cloud fractions and transmissivity are calculated. Using GOES data, $J_{NO2}$ decreases over most of the domain, and has most impact on the places with high NO$_x$ emissions (Figs. 3.2 and 3.3). The largest discrepancy of 80ppb in modeled O$_3$ occurs at 13:00 on 2 September 2006 over the DFW region during the modeling period. The differences in $J_{NO2}$ and NO$_2$ at 13:00 on 2 September 2006 reach up to 6 times base value and 10ppb, respectively (Fig. 3.2), but these differences are not extremes as in modeled O$_3$ in the episode. The differences in modeled $J_{NO2}$, NO$_2$, and O$_3$ are much more moderate on monthly 8-h (10:00-18:00) averaged values, reaching only up to 3s$^{-1}$ for $J_{NO2}$, 0.6ppb for NO$_2$, and 3ppb for O$_3$, and the largest discrepancies occur in the HGB region.

The general impact of using GOES observations is that where the $J_{NO2}$ decreases, modeled NO$_2$ increases, and O$_3$ decreases (Figs. 3.2 and 3.3), indicating that slower photochemical activity inhibits O$_3$ formation and thus consumes less NO$_2$ and vice versa. However, an exception occurs at places close to the Houston Ship Channel, showing that although the $J_{NO2}$ decreases, modeled NO$_2$ still decreases (Fig. 3.3b) and O$_3$ slightly increases (Fig. 3.3c). This probably because the availability of other pathways for consuming NO$_x$ due to high concentrations of VOCs and organic radicals, and the inhibition of NO regeneration due to reduction in photochemical activities. Approximately 6% less $J_{NO2}$ over most of the domain makes modeled O$_3$ overall less sensitive to NO$_x$ emissions (Fig. 3.3d) and more sensitive to VOC emissions (Fig. 3.3e). In the urban core of the DFW region, stronger photochemistry related to increased $J_{NO2}$ (Fig. 3.3a) consumes more NO$_x$ and VOC to form more O$_3$ (Fig. 3.3c) and make O$_3$ more sensitive to both NO$_x$ and VOC emissions, and vice versa. For example, opposite effects occur in the urban core of the HGB region (Figs. 3.3d and 3.3e).
The modeled daily 8 hours (10:00-18:00LT) NO$_2$ and O$_3$ using either satellite-derived or base model photolysis rates are evaluated by AQS measured data for the entire modeling period, and the positive changes in R$^2$ and negative changes in NMB and NME indicate the improvements of model performance caused by using satellite-derived photolysis rates (Fig. 3.4). For O$_3$ simulations (Fig. 3.4 right), the difference in R$^2$ increases 1% on average and reaches up to 7% on 26 August, while the differences in biases and errors decrease 1% on average and reach up to 10% on 11 September, suggesting the satellite-corrected photolysis rates improve the model performance in simulating ground O$_3$. However, for NO$_2$ simulations (Fig. 3.4 left), in most of days, although the difference in R$^2$ increases, same trends are observed for the modeled biases and errors, indicating no improvements are made in modeling ground NO$_2$ with satellite corrections for photolysis rates, probably because the other model uncertainties, such as NO$_x$ emissions, may have larger impact on the accuracy of NO$_2$ simulations.

Figure 3-2. Differences between satellite-derived (GOES) and model predicted (MOD) JNO$_2$ (left) in simulating NO$_2$ (middle) and O$_3$ (right) at 13:00 on 2 September 2006.
Figure 3-3. Monthly 8-h (10:00-18:00LT) averaged differences between satellite-derived (GOES) and model predicted (MOD) (a) JNO₂ in simulating (b) NO₂, (c) O₃, and O₃ sensitivities to (d) NOₓ and (e) VOC.
Figure 3-4. Change in model performance ($R^2$, NMB, and NME) in simulating daily 8 hours (10:00-18:00LT) NO$_2$ (left) and O$_3$ (right) caused by satellite-derived photolysis rates.

3.3.2. Pseudodata test for the sector-based DKF inversion

A controlled pseudodata test was performed in Tang et al. (2013) to test the applicability of the DKF inversion to adjust the NO$_x$ emission in each inversion region with the CAMx-DDM model. This showed that the DKF method adjusted the perturbed NO$_x$ emission in each region accurately back to its base case. In this study, a similar controlled pseudodata test is conducted to test the applicability of the sector-based DKF inversion with CAMx-DDM.

The pseudodata test for the sector-based DKF inversion is conducted on 10 modeling days (13 August to 22 August), but the modeling results from the first 3 days are discarded to eliminate the model initialization error. A 7-day (16 August to 22
August) averaged modeled NO$_2$ VCD at 13:00-14:00LT with the base case NO$_x$ emission inventory is treated as a pseudo-observation, and the one using perturbed NO$_x$ emissions in six emission sectors with known perturbation factors ranging from 0.5 to 2.0 (Fig. 3.6) is used as a priori case. As described above, the area and non-road emission sources are considered as one sector (ARNR), and EGU point source is excluded from the inversion. The emission uncertainties are set to 50% for the non-EGU and lightning sectors and to 100% for the others. The measurement error for the pseudo-observation is set to 30%.

The pseudodata test results (Fig. 3.6 top) show that the a posteriori modeled NO$_2$ closely matches the base case modeled value, indicating the DKF inversion is capable of correcting the perturbed NO$_x$ emissions in each emission sector. The sensitivity analysis results (Fig. 3.6 bottom) illustrate that the inversions are insensitive to both emission and observation error covariance matrices for the pseudo-cases.
Figure 3-5. Vertical column densities of NO₂ sensitivities to NOₓ emissions of (a) area, (b) non-road, (c) mobile, (d) biogenic, (e) lightning, (f) aviation, and (g) non-EGU points source sectors.
Figure 3-6. Pseudodata analysis for the sector-based DKF inversion (top), and its sensitivities to uncertainties in emissions (bottom left) and in OMI observations (bottom right).

Figure 3-7. Comparisons of modeled NO$_2$ vertical distributions with INTEX NASA DC-8 flight (left) and TexAQS 2006 NOAA P-3 aircraft (right) measurements.
3.3.3. A priori NO$_2$ VCD

The a priori NO$_x$ emission inventory used in this study is based on the TCEQ base case emission inventory with added lightning and aviation and doubled soil NO$_x$ emissions (Tang et al., 2013). The reaction rate constant of the reaction NO$_2$ + OH in CB05 chemical mechanism is reduced by 25% based on Mollner et al. (2010); this tends to increase NO$_x$ lifetime and transport to rural regions.

To evaluate the extent to which the addition of lightning and aviation NO$_x$ closes the gap between observed and modeled NO$_2$ in the upper troposphere noticed by Napelenok et al. (2008), the modeled NO$_2$ vertical profile is compared with INTEX-NA DC-8 measured NO$_2$ profiles from the ground to the free troposphere. The comparison (Fig. 3.7 left) shows that CAMx with the adjusted a priori inventory strongly overestimates NO$_2$ near the ground, reasonably agrees with DC-8 NO$_2$ measurements from 1km to 5km, slightly overestimates NO$_2$ from 6km to 9km, and slightly underestimates NO$_2$ from 10km to 15km. The modeled NO$_2$ profile is further evaluated by the P-3 measured NO$_2$ from ground to 5km (Fig. 3.7 right), showing the same pattern of the overestimated surface NO$_2$ and good agreement with aircraft observations from 1km-5km. The injection of all aviation NO$_x$ emissions, obtained from the Emission Database for Global Atmospheric Research (EDGAR), into a single model layer at altitude 6km to 9km rather than more broadly distributed vertically probably causes the overestimation of modeled NO$_2$ compared to DC-8 at that altitude (ENVIRON, 2013). A low bias of modeled NO$_2$, approximately 40ppt, exists in the upper troposphere, from 10km to 15km altitude, which is the CAMx model top layer. Similar low bias of the
modeled NO$_2$ in the upper troposphere compared to the DC-8 measurement also has been found in Allen et al. (2012). Because the low bias in the upper troposphere may arise from model uncertainties other than those associated with emissions (Henderson et al., 2011; 2012), we follow the adjustment approach of Napelenok et al. (2008) and add 40ppt NO$_2$ homogeneously to the top layer of the model results when computing the CAMx NO$_2$ VCDs.

Although the revised CB05 chemical mechanism and artificially added upper tropospheric NO$_2$ increase modeled NO$_2$ VCD in the inversion region by an average of 13% (Appendix B sect. 2), CAMx modeled NO$_2$ VCDs remain an average of 2×10$^{14}$ molecules/cm$^2$ less than OMI observations in rural regions (Fig. 3.8c).

3.3.4. Top-down NO$_x$ emissions constrained by DKF inversions

The DKF inversions with OMI NO$_2$ are performed to constrain NO$_x$ emissions in each designated emission region and emission sector. To ensure sufficient spatial coverage, a monthly averaged OMI NO$_2$ VCD (13 August to 15 September) is calculated and paired with the corresponding modeled NO$_2$ VCD at satellite passing time (13:00-14:00LT). The DKF inversions are then conducted with 2116 data points covering every grid cell in the inversion region, and the hourly a priori NO$_x$ emissions are adjusted iteratively until the inversion process converges.
Figure 3-8. Monthly averaged (16 August to 15 September) tropospheric NO₂ VCDs at 13:00-14:00LT from (a) OMI, (b) a priori simulation, (c) difference between OMI and a priori simulation, and simulations using a posteriori NO₅ emissions generated by (d) region-based DKF inversion, and sector-based DKF inversion (e) case I and (f) case II.
3.3.4.1. Region-based DKF inversion

The region-based DKF inversion is conducted to adjust the NOx emissions in each inversion region. The inversion results suggest to moderately adjust the a priori NOx emissions in most regions with scaling factors ranging from 0.97 to 1.49 (Table 3.2) and increases NO2 VCD by 8% toward OMI measurement over the inversion region (Fig. 3.8d). Because this inversion is based on a new OMI retrieved and an improved a priori NO2 VCDs, the required adjustments in each inversion region are much lower compared to the results in Tang et al. (2013) with scaling factors ranging from 0.56 to 1.98 and 30% increased NO2 VCD.

The model performance is then evaluated by the ground and aircraft measurements. The DKF inversion adjusts DFW NOx emissions by only 3%, while it adds 50% to BPA emissions and less than 15% to other urban regions. The bias and error of the a posteriori modeled NO2 VCD decrease in every urban area and are reduced from -0.11 to -0.05 and from 0.17 to 0.16 overall compared to OMI. The spatial correlations between OMI and CAMx NO2 VCDs ($R^2$) are improved only in the BPA and Austin and San Antonio areas, but the overall region-wide performance is improved (Table 3.3). The modeled NO2 with a priori NOx emissions overpredicts ground-level NO2 (Table 3.4); hence, the increase in NOx emissions at most urban places suggested by the inversion actually deteriorates the ground-level NO2 simulations in all urban areas except in the DFW region. The modeled bias and error of ground O3 are reduced in the HGB and BPA regions, but not in DFW, probably because the increased NOx in the first two regions titrates more ground O3 at night and inhibits O3 formation during the day, decreasing the
O₃ concentrations which are already overestimated in the a priori simulation (Table 3.6).

No improvements of model performance are found in simulating P-3 observed NO₂ and NOₓ using the inverted NOₓ emissions.

Table 3-3. Evaluation of CAMx modeled NO₂ using OMI NO₂.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R²</td>
<td>NMB</td>
<td>NME</td>
<td>R²</td>
</tr>
<tr>
<td>HGB</td>
<td>0.57</td>
<td>-0.25</td>
<td>0.36</td>
<td>0.57</td>
</tr>
<tr>
<td>DFW</td>
<td>0.74</td>
<td>-0.21</td>
<td>0.29</td>
<td>0.72</td>
</tr>
<tr>
<td>BPA</td>
<td>0.40</td>
<td>-0.46</td>
<td>0.47</td>
<td>0.45</td>
</tr>
<tr>
<td>NE Texas Austin and San Antonio</td>
<td>0.24</td>
<td>-0.40</td>
<td>0.44</td>
<td>0.24</td>
</tr>
<tr>
<td>Overall</td>
<td>0.45</td>
<td>-0.25</td>
<td>0.35</td>
<td>0.47</td>
</tr>
</tbody>
</table>

a. Compare to OMI observations in all inversion regions

Table 3-4. Evaluation of CAMx modeled NO₂ using hourly AQS ground-measured NO₂.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R²</td>
<td>NMB</td>
<td>NME</td>
<td>R²</td>
</tr>
<tr>
<td>HGB</td>
<td>0.51</td>
<td>0.46</td>
<td>0.67</td>
<td>0.51</td>
</tr>
<tr>
<td>DFW</td>
<td>0.49</td>
<td>0.43</td>
<td>0.66</td>
<td>0.49</td>
</tr>
<tr>
<td>BPA</td>
<td>0.45</td>
<td>0.92</td>
<td>1.02</td>
<td>0.45</td>
</tr>
<tr>
<td>NE Texas Austin and San Antonio</td>
<td>0.70</td>
<td>0.86</td>
<td>0.93</td>
<td>0.70</td>
</tr>
<tr>
<td>Overall</td>
<td>0.46</td>
<td>0.60</td>
<td>0.87</td>
<td>0.47</td>
</tr>
</tbody>
</table>

a. Compare to all ground sites
Table 3-5. Evaluation of CAMx modeled NO$_2$ using P-3 aircraft-measured NO$_2$ and NO$_y$.

<table>
<thead>
<tr>
<th>Statistical parameters</th>
<th>NO$_2$</th>
<th>NO$_y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^2$</td>
<td>0.22</td>
<td>0.23</td>
</tr>
<tr>
<td>NMB</td>
<td>0.09</td>
<td>0.15</td>
</tr>
<tr>
<td>NME</td>
<td>0.99</td>
<td>1.03</td>
</tr>
</tbody>
</table>

a. Comparison available for only four days (August 31, September 11, September 13, and September 15, 2006).

3.3.4.2. Sector-based DKF inversion

The sector-based DKF inversion is first conducted on six NO$_x$ emission sectors: area and nonroad (ARNR), mobile, biogenic, aviation, lightning, and non-EGU points (Case I). The scaling factors generated by the inversion ranges from 0.54 to 4.10, with the largest scale-down in the ARNR sector and the largest scale-up in the aviation sector. The inversion reduces NO$_x$ emission in the biogenic sector by 30% from the a priori inventory which had doubled soil NO$_x$ from the base model. The inversion leaves mobile, lightning, and non-EGU points sectors nearly unchanged, applying less than 4% adjustments (Table 3.2). The NO$_2$ VCD is increased by only 6% toward OMI measurement over the inversion region in this case. Most of the increase in NO$_2$ VCDs occurs in rural areas, and some declines occur in urban areas (Fig 3.8e).

The NO$_x$ emission in each inversion region is recalculated after applying adjustments to each emission sector, and model performance is evaluated by the ground
and aircraft measurements. The scaling factors in each region now are less and different than those generated by the region-based inversion, ranging from 0.86 in NE TX to 1.17 in DFW. The modeled bias and error in simulating OMI NO2 are all decreased in five urban areas. Within the inversion region, the overall modeled bias and error are reduced from -0.11 to -0.04 and from 0.17 to 0.14, respectively using inverted NOx emissions (Table 3.3). The 50% cut in the ARNR sector helps to improve the model performance in simulating ground-level NO2 and O3 which already have been overestimated using a priori NOx emissions. The inverted NOx emissions decrease modeled bias and error in all five urban areas and overall decrease bias by 0.25 and 0.04, and error by 0.13 and 0.04 in simulating ground-level NO2 and O3, respectively (Table 3.4 and Table 3.6). The model performance is also improved compared against P-3 measurements, with 0.09 reduction in both modeled bias and error for NO2 simulation and 0.16 reduction in bias and 0.11 reduction in error for NOy simulation (Table 3.5). The scaled-down ground NOx emissions lead to 2-5 ppb lower modeled 8-h (10:00-18:00LT) ground O3 and make O3 formation chemistry less sensitive to the VOC emissions, with reduction of 1-3 ppb sensitivity coefficients over the inversion region. The O3 sensitivity to NOx emissions also decreases by approximately 1-2 ppb over most of the inversion region; however, the O3 formation chemistry in the urban cores of the DFW, HGB, and Austin and San Antonio regions shifts toward being more NOx-limited, leading to 1-3 ppb increase of O3 sensitivity to NOx emissions (Fig. 3.9).

Although the inversion improves the model performance, the sensitivity analysis (Appendix B sect. 3) shows that the aviation and ARNR sectors are relatively responsive to the emission uncertainty values and offset each other (Fig. B2), indicating the DKF
inversion may not be capable of fully distinguishing these two emission sectors. Therefore, the aviation source is then merged with ARNR and the DKF inversion is re-conducted on five emission sectors: area, nonroad, and aviation (ARNRAV), mobile, biogenic, lightning, and non-EGU points (Case II). In case II, the inversion results are more stable and insensitive to the emission uncertainties in each emission sector (Fig. B2). However, the inversion tends to scale up all three source categories in the ARNRAV sector together by 50% to compensate for the rural NO2 gap, which leads to increases of NOx emissions in both ground area and nonroad sources. The inversion reduces mobile and biogenic NOx emissions by 12% and 16%, respectively. The adjustments for the lightning and non-EGU points sectors are still less than 4% (Table 3.2). On the region basis, the inversion tends to increase NOx emissions in all regions with increments ranging from 1% in the Austin and San Antonio region to 18% in the NE TX region; thus increases the modeled NO2 VCD by 7% on average toward the OMI measurement. The inversed NO2 VCD in this case is very similar to that from the region-based inversion (Fig. 3.8f). The model performance of simulating OMI NO2 VCD is improved and similar to the results from case I (Table 3.3). However, unlike case I, no improvements are found in simulating ground measured NO2 and O3 and P-3-measured NO2 and NOy using the inverted NOx emissions in case II (Table 3.4-3.6). Because the ground NOx emissions are increased in this case, the inversion impacts the O3 simulations in the opposite direction than in case I. The modeled 8-h ground O3 increases by around 2ppb and becomes more sensitive to both NOx and VOC emissions over most of the inversion region; however, the O3 formation chemistry shifts toward being more VOC-limited in DFW and HGB (Fig. 3.9).
Figure 3-9. Monthly 8-h (10:00-18:00LT) averaged ground O₃ concentrations (top), O₃ sensitivity to NOₓ (middle), and O₃ sensitivity to VOC (bottom) for the a priori case (left), and differences between a posteriori and a priori for the sector-based DKF inversions case I (middle) and case II (right).
Table 3-6. Evaluation of CAMx modeled O₃ using hourly AQS ground-measured O₃.

<table>
<thead>
<tr>
<th>Source region</th>
<th>Priori</th>
<th>Posteriori: region-based inversion</th>
<th>Posteriori: sector-based inversion I</th>
<th>Posteriori: sector-based inversion II</th>
<th>Sector-I inversed NOₓ emissions &amp; GOES photolysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R²</td>
<td>NMB</td>
<td>NME</td>
<td>R²</td>
<td>NMB</td>
</tr>
<tr>
<td>HGB</td>
<td>0.46</td>
<td>0.68</td>
<td>0.75</td>
<td>0.47</td>
<td>0.67</td>
</tr>
<tr>
<td>DFW</td>
<td>0.64</td>
<td>0.21</td>
<td>0.32</td>
<td>0.64</td>
<td>0.23</td>
</tr>
<tr>
<td>BPA</td>
<td>0.47</td>
<td>0.61</td>
<td>0.70</td>
<td>0.47</td>
<td>0.59</td>
</tr>
<tr>
<td>NE Texas</td>
<td>0.49</td>
<td>0.36</td>
<td>0.43</td>
<td>0.49</td>
<td>0.38</td>
</tr>
<tr>
<td>Austin and San Antonio</td>
<td>0.52</td>
<td>0.40</td>
<td>0.46</td>
<td>0.52</td>
<td>0.40</td>
</tr>
<tr>
<td>Overall</td>
<td>0.50</td>
<td>0.42</td>
<td>0.50</td>
<td>0.51</td>
<td>0.42</td>
</tr>
</tbody>
</table>

a. Compare to all ground sites

Figure 3-10. Monthly 8-h (10:00-18:00LT) averaged differences in modeled (a) ground O₃ concentrations, (b) O₃ sensitivity to NOₓ, and (c) O₃ sensitivity to VOC resulting from use of both satellite-derived photolysis rates and NOₓ emissions in place of a priori data.
3.4. Conclusions

Satellite-derived photolysis rates and NO\textsubscript{x} emissions are both applied to a Texas SIP modeling episode to investigate the capabilities of using satellite data to enhance state-level O\textsubscript{3} regulatory modeling. Results show that the ground-level O\textsubscript{3} simulations are improved with reductions of modeled bias from 0.42 to 0.37 and modeled error from 0.50 to 0.45 by using GOES-derived photolysis rates and sector-based DKF (case I) with OMI NO\textsubscript{2} inverted NO\textsubscript{x} emission inventory (Table 3.6). The GOES-derived photolysis rates and OMI-constrained NO\textsubscript{x} emissions decrease monthly averaged 8-h O\textsubscript{3} concentrations by 2-5ppb over the entire inversion region and turn O\textsubscript{3} formation chemistry toward being less sensitive to NO\textsubscript{x} and VOC emissions over most of inversion areas, while being more NO\textsubscript{x} sensitive in the two O\textsubscript{3} nonattainment areas, DFW and HGB (Fig. 3.10).

Applying GOES-retrieved cloud coverage and transmissivity reduce the modeled photolysis rates over most of the domain, leading to less photochemical activity and O\textsubscript{3} production and shifting O\textsubscript{3} formation chemistry toward being less sensitive to NO\textsubscript{x} emissions, except in the DFW region where modeled photolysis rates are increased by the GOES retrieval, leading to impacts in the opposite direction. In comparing with the AQS ground measurements, the GOES-derived photolysis rates improve the ground-level O\textsubscript{3} simulations but not the NO\textsubscript{2} simulations, indicating other model errors may dominate the accuracy of model performance in simulating ground-level NO\textsubscript{2}. Future work could extend use of GOES-retrieved clouds to further correct model dynamics and aqueous phase chemistry and investigate their impacts on the O\textsubscript{3} modeling.
The DKF inversion approach has been successfully incorporated with CAMx-DDM model and was conducted on both region-based and sector-based NOx emissions. The controlled pseudodata test is conducted on the sector-based DKF inversion and shows the inversion accurately captures the known perturbation in each NOx emission sector and adjusts back to its original conditions. In addition to implementing lightning and aviation NOx emissions in the upper troposphere and doubling soil NOx emissions from the ground, the NOx lifetime is increased by reducing 25% the reaction rate constant of the reaction OH + NO2. The upper tropospheric NO2 underestimation is further eliminated by adding a 40ppt homogenous NO2 layer in the model top. On the other hand, the high resolution OMI retrieval with a priori profile from the nested GEOS-Chem simulation further enhances NO2 in the urban areas and reduces NO2 in rural. However, the comparison still shows that the OMI has higher NO2 VCD than CAMx in the rural areas, by around 2×10^{14} molecules/cm^2. It is not clear that the discrepancy between OMI and CAMx in rural areas is caused by uncertainties in NOx emission inventory or errors in OMI retrieval and other model uncertainties. The OMI NO2 retrieval can be further improved by using the finer resolution terrain and albedo data (Russell et al., 2011) and via findings from the recent DISCOVER-AQ Houston measurement campaign (Crawford et al., 2014). The accuracy of CAMx modeled NO2 VCD can benefit from further improving the modeled chemical and transport processes (ENVIRON 2013), such as updating NOx recycling process to increase NOx lifetime, or adding cross-tropopause transport process to allow more stratospheric NO2 penetrate to upper troposphere. This may obtain better spatial distribution of modeled NO2 rather than adding a homogeneous layer at top to compensate the model deficiency.
The region-based DKF inversion still over scales NO\textsubscript{x} emissions in urban areas to compensate for the rural NO\textsubscript{2} differences because the NO\textsubscript{2} VCD gap in rural areas is not eliminated, leading to 10-50\% increase of NO\textsubscript{x} emissions at most regions and worsening the ground-level O\textsubscript{3} simulations; however, the scaling factors generated in this study are much more moderate than those that were found in Tang et al. (2013). The sector-based DKF inversion (case I) takes the aviation source to compensate the NO\textsubscript{2} gap in the rural area, probably because its relatively uniform emission pattern over the rural area mostly matches the NO\textsubscript{2} discrepancy distributions, leading to appropriate adjustments in the ground emissions and improving both ground-level NO\textsubscript{2} and O\textsubscript{3} simulations; however, the aviation source is unrealistically adjusted by applying a suggested factor of 4 to its base value, and the adjustments offset the area and nonroad sector with varying emission uncertainties in the sensitivity analysis. Although merging the aviation source into the area and nonroad emission sector makes the inversion (case II) more stable, the large scaling factor for the aviation sector is now shared with area and nonroad emissions, leading to area and nonroad NO\textsubscript{x} emissions being scaled up by 50\%. Thus, the model performance in ground-level NO\textsubscript{2} and O\textsubscript{3} simulations is deteriorated and is even worse than the results generated from the region-based inversion. The lightning NO\textsubscript{x} emissions seem to be well estimated and require almost no adjustment suggested by the inversion from both sector-based cases. However, it may also indicate that the OMI retrieved NO\textsubscript{2} is insensitive to the lightning source, most probably due to the NO/NO\textsubscript{2} partitioning in the upper troposphere and the clear-sky cloud screening criterion used in the OMI retrieval. The NO\textsubscript{2} discrepancy between OMI and CAMx drives the DKF inversion and is assumed to be mostly contributed by the uncertainties in the NO\textsubscript{x} emission inventory.
However, findings from this study indicate that if the uncertainty in the a priori NO\textsubscript{x} emissions is low, errors in the satellite retrieval and model itself cannot be neglected, making the inversion less capable of reducing the uncertainties in the bottom-up NO\textsubscript{x} emission inventory.

The sector-based DKF inversion applies a single scaling factor to each emission sector, assuming that the spatial distribution of NO\textsubscript{x} emissions in each sector is accurately estimated in the bottom-up NO\textsubscript{x} emission inventory. However, this assumption may be not true. For example, TCEQ recently developed a single-day aviation emission inventory using the Advanced Emission Model (AEM3) for the new Rider 8 modeling domain, which has more accurate flight profile and distributes emissions more broadly in the vertical direction, leading to the spatial pattern of NO\textsubscript{x} emissions somewhat different than that obtained from EDGAR (ENVIRON 2013). In addition, the newly developed Berkeley-Dalhousie Soil NO\textsubscript{x} Parameterization (BDSNP) scheme (Hudman et al., 2012) recently was implemented into the CMAQ model to estimate soil NO\textsubscript{x} emissions, showing large spatial and temporal differences compared to those estimated by the YL95 scheme over eastern Texas. All these changes described above in the a priori NO\textsubscript{x} emission inventory may have significant impact on the inversion results.

The direct scaling inversion (Appendix B sect. 4) using PAMS measured VOCs improves the model performance in simulating five chosen VOC species and indicates the TCEQ VOC emission inventory used in HGB SIP modeling is now much better than the previous reported emissions with values off by an order of magnitude. However, the inverted VOC emissions have insignificant impact on the ground-level NO\textsubscript{2} and O\textsubscript{3}
simulations, probably because of the limited spatial coverage of the PAMS measurement sites and most VOC-saturated conditions in the inversion region. Future work could explore the capabilities of using satellite-observed formaldehyde data to constrain the Texas isoprene or even other anthropogenic VOC emissions (Defour et al., 2009; Curci et al., 2010).

The statistical results show that although the modeled bias and error are reduced, OMI-constrained NOx emissions barely improve the spatial and temporal correlations (R²) with ground-measured NO₂ and O₃, indicating that either applying the scaling factors generated at the OMI passing time is unable to reduce the emission uncertainty at each hour or the current OMI resolution is insufficient to capture the spatial distributions of the NOx emission pattern. This limitation could be solved by using measurements from the geostationary satellite which has the temporal resolution down to an hour and the spatial resolution down to 4km×4km. These data will be available in the NASA’s new mission called Tropospheric Emission: Monitoring of Pollution (TEMPO) (Streets et al., 2013).
Chapter 4

Conclusions and Future Research

This doctoral research demonstrates the applicability of using satellite-derived photolysis rates and a top-down NOx emission inventory to enhance Texas O3 abatement modeling. The GOES-retrieved cloud fractions and transmissivity were implemented into CAMx to reproduce broadband photolysis rates by overwriting the model standard procedure. The DKF inversion was incorporated successfully with CAMx-DDM and conducted on constraining both region-based and sector-based NOx emissions. Final results show that the model performance in simulating ground-level O3 is improved by applying satellite-derived photolysis rates and NOx emissions either separately or together, and that using satellite-derived photolysis rates and NOx emissions decrease modeled 8-h O3 concentrations and make O3 more sensitive to NOx emissions in the O3 nonattainment areas.
4.1. Conclusions

Studies show that WRF and MM5 which are used to provide meteorological fields to the regional air quality models underpredict clouds and thus make more solar irradiance reach the ground. Hence, using GOES-observed clouds corrects that issue and reduces the solar irradiance, making modeled photolysis rates decrease. In this study, generally, GOES-derived photolysis rates decrease modeled photochemical activity and O₃ production and shift O₃ formation chemistry toward being less sensitive to NOₓ emissions over most of the domain. Using GOES-derived photolysis rates improve the model performance in simulating ground-level O₃ and change the modeled O₃ concentrations up to 80ppb. However, no improvements are found in the ground-level NO₂ simulations.

The pseudodata analysis is first performed to test the capability of the DKF inversions to capture the known perturbations in the NOₓ emission inventory, and the results show that the DKF inversions accurately adjust perturbed NOₓ emissions in either designated emission regions or sectors back to their base values. Comparing the CAMx modeled NO₂ to the OMI observations, a NO₂ gap is found over the rural areas in the inversion region, and this limits the capability of the DKF inversion to correctly adjust the bottom-up NOₓ emission inventory. It is still not clear if this gap is caused by the uncertainties in the NOₓ emission inventory or by other model uncertainties and satellite retrieval error.

In the first study, missing lightning and aviation NOₓ emissions and extra soil NOₓ emissions are first added to the base case NOₓ emission inventory to reduce the NO₂
gap in the rural areas. However, it narrowed but did not fully close the gap. Therefore, the region-based DKF inversion with OMI NO$_2$ over scales the NO$_x$ emissions in the urban areas in order to compensate for the gap in the rural areas and deteriorates the ground-level NO$_2$ and O$_3$ simulations. An alternative DS method is performed to adjust the NO$_x$ emissions by comparing the OMI and CAMx NO$_2$ cell by cell, which features more spatial heterogeneous adjustments and avoids the limitation of only applying a single scaling factor to the entire region in the region-based DKF inversion. However, the nonlinear influence of NO$_x$ on its own lifetime makes the DS method fail to adjust emissions correctly and to improve the model performance. The corrected AQS ground NO$_2$ with mostly urban measurements are then used in the region-based DKF inversion to eliminate the influence of the model underprediction errors in the upper troposphere and in the rural areas. However, the inversion results are unreliable because of model shortcomings in simulating PBL heights in the early morning and late afternoon. It also fails to improve the ground O$_3$ modeling. Thus, we switch our focus back on using OMI NO$_2$ to adjust the NO$_x$ emissions and try to alleviate the other known uncertainties that could deteriorate the inversion accuracy.

In the second study, the NO$_2$ gap is further reduced by using high resolution SP2 OMI NO$_2$ product, increasing the NO$_x$ lifetime, and adding a 40ppt homogeneous layer in CAMx before performing the DKF inversion; however, the gap between the model and satellite observations in rural areas still exists. Hence, the region-based DKF inversion still overly scales up NO$_x$ emissions in the urban areas to compensate the rural NO$_2$ gap, while the adjustments are much more moderate than that in the first study. An alternative DKF inversion of constraining NO$_x$ emissions in different emission sectors is
then conducted on six emission sectors, area and nonroad, mobile, biogenic, lightning, aviation, and non-EGU points, to have more spatial adjustments rather than applying a single scaling factor to the whole region. The sector-based DKF inversion uses the relatively homogeneous distributed aviation source to remedy the NO\textsubscript{2} gap in the rural areas, leading to a 50\% reduction in the ground area and nonroad emissions and improvements in both ground-level NO\textsubscript{2} and O\textsubscript{3} simulations; however, the factor of 4 scaling-up for the aviation source is unrealistic, and the inversion system seems to be somewhat ill-posed. The DKF inversion is re-conducted on five emission sectors with the aviation source merged with the area and nonroad sector, in order to have more stable inversion results; however, in this case, the DKF inversion takes the merged area and nonroad and aviation sector together to compensate for the rural NO\textsubscript{2} gap, leading to a 50\% increase in the ground area and nonroad NO\textsubscript{x} emissions and worsening the model performance in simulating ground-level NO\textsubscript{2} and O\textsubscript{3}. Although the aviation source is inappropriately adjusted and the inversion results are dependent on the chosen emission uncertainties, the OMI-constrained NO\textsubscript{x} emissions via the sector-based DKF inversion finally enhance the ground O\textsubscript{3} modeling by reducing model bias from 0.42 to 0.38, and model error from 0.50 to 0.46.

One major assumption of using inverse modeling with satellite NO\textsubscript{2} to adjust NO\textsubscript{x} emissions is that the differences in satellite-observed and modeled NO\textsubscript{2} VCD are mostly attributed to the uncertainties in the NO\textsubscript{x} emission inventory. However, findings from this study indicate that other uncertainties in the satellite retrieval or model itself may invalidate the above assumption and make inverse modeling less capable of creating a reasonable top-down NO\textsubscript{x} emission inventory. The rural NO\textsubscript{2} gap between satellite data
and model output may come from uncertainties in the model chemistry and mixing processes or from errors in the OMI retrieval. It makes the region-based DKF inversions over scale the urban NO\textsubscript{x} emissions and the sector-based DKF inversions over scale the aviation NO\textsubscript{x} emissions. On the other hand, the appropriate estimated spatial patterns of different emission sources in the bottom-up emission inventory are essential to have reliable inversion results, especially for the sector-based DKF inversion. For example, the recently developed aviation NO\textsubscript{x} emission using the Advanced Emission Model (AEM3) and the soil NO\textsubscript{x} emission using the Berkeley-Dalhousie Soil NO\textsubscript{x} Parameterization (BDSNP) scheme show different spatial distribution of NO\textsubscript{x} emissions compared to those in the a priori case and may have significant impact on the inversion results. Although limitations still exist in the DKF inversions, the reductions in the area and nonroad sources suggested by the sector-based DKF inversion improve the ground-level NO\textsubscript{2} and O\textsubscript{3} simulations. It could be used as an indicator for TCEQ in improving the bottom-up NO\textsubscript{x} emission inventory. However, we think the top-down NO\textsubscript{x} emissions derived by the satellite observations should for now remain a complement to the bottom-up NO\textsubscript{x} emission inventory instead of replacing it in the Texas SIP modeling.

The VOC emissions in the Texas SIP emission inventory also are adjusted based on the PAMS VOC measurements. Although the model performance is improved in simulating these adjusted VOC species, the inverted VOC emissions have insignificant impact on the ground-level NO\textsubscript{2} and O\textsubscript{3} simulations, most probably due to the scarce VOC measurement sites and VOC-abundant conditions for the O\textsubscript{3} formation.
4.2. Future research

Based on the findings in this research, a few things could be done in the future to further enhance the O$_3$ abatement modeling in Texas using satellite measurements. First of all, the GOES-observed clouds have been used only to correct photolysis rates in this study, which creates the inconsistency between photolysis rates and other cloud features. For example, the difference in cloud occurrences also can affect the modeled PBL height and vertical mixing and the chemical reactions happening in the cloud. Hence, future work can extend the use of GOES-observed clouds to correct model dynamics and aqueous-phase chemistry.

In addition, the accuracy of top-down NO$_x$ emission constrained by space-based NO$_2$ measurements can be improved by further reducing the OMI and CAMx uncertainties or by using high resolution geostationary satellite. The OMI retrieval could benefit from using finer resolution terrain and albedo data and measured NO$_2$ vertical profile from the DISCOVER-AQ measurements. For CAMx modeling, a few things need to be improved, as suggested in ENVIRON (2013), such as NO$_x$ recycling chemistry, inter-tropopause NO$_2$ exchange processes, and vertical mixing processes in the upper troposphere. The geostationary satellite used in NASA’s new mission called Tropospheric Emission: Monitoring of Pollution (TEMPO) (Streets et al., 2013) is capable of measuring NO$_2$ with temporal resolution down to an hour and the spatial resolution down to 4km$\times$4km, which can provide more information for the NO$_x$ inversions. For example, the higher spatial resolution satellite measurement dataset would provide more observation points to better capture the spatial pattern of NO$_x$ emissions.
and would make the inversion results more reliable statistically. Meanwhile, the higher temporal resolution satellite measurement dataset would enable the DKF inversions to constrain the NO\textsubscript{x} emissions hourly, in order to better capture the temporal pattern of NO\textsubscript{x} emissions.

Furthermore, the impacts of using satellite-derived photolysis rates and NO\textsubscript{x} emissions on emission control strategies need to be investigated. For example, using GOES-photolysis rates makes O\textsubscript{3} more sensitive to NO\textsubscript{x} emissions in the two O\textsubscript{3} nonattainment areas, indicating current NO\textsubscript{x} reduction plan may actually decrease O\textsubscript{3} concentrations to a greater extent in the future. The OMI-constrained NO\textsubscript{x} emissions indicate area and nonroad sources in the current bottom-up NO\textsubscript{x} emission inventory are overestimated. Hence, if the current NO\textsubscript{x} reduction strategies are on the percentage basis of each emission category, then the reduction may be insufficient to bring down enough O\textsubscript{3}. However, if the NO\textsubscript{x} reduction plan is on the actual quantity basis, then O\textsubscript{3} concentrations may attain the standard earlier.

Although applying DKF inversions to adjust the NO\textsubscript{x} emissions in Texas regulatory modeling shows limitations, the region- and sector-based DKF inversion methods with the CAMx-DDM model are well established, and can be applied to adjust NO\textsubscript{x} emissions in other regions or extended to adjust other species in the emission inventory using space-based observations, such as formaldehyde or aerosol optical depth (AOD).
4.2.1. Constrain VOC emissions using satellite-retrieved formaldehyde

Formaldehyde (HCHO) is a very important VOC species for tropospheric O₃ formation and is an intermediate compound from oxidations of various VOC species, such as methane, biogenic isoprene and terpene, and anthropogenically emitted VOCs (Atkinson, 1994; Dufour et al., 2009; Curci et al., 2010). Recent studies have used SCIAMACHY (Dufour et al., 2009) and OMI (Curci et al., 2010) HCHO observations combined with inversions to constrain VOC emissions over Europe. Since DDM can track the spatial and temporal relationships between HCHO concentrations and VOC emissions, DKF inversions can be performed to constrain Texas VOC emissions using satellite-retrieved HCHO data in the future. However, due to the measurement limitations and retrieval errors, current space-based HCHO observations only are capable of constraining isoprene emissions.

4.2.2. Constrain aerosol emissions using satellite-retrieved AOD

AOD indicates the quantity of light extinction by aerosol through a vertical column of atmosphere, relying on aerosol mass concentration, mass extinction efficiency, relative humidity, and aerosol vertical distributions (Kaufman and Fraser, 1983; Seinfeld and Pandis, 2006). The Multiple Imaging SepectroRadiometer (MISR) on the EOS-Terra satellite (Diner et al., 1998) and the Moderate Resolution Imaging SpectroRadiometer (MODIS) on EOS-Terra/Aqua satellites (Remer et al., 2005) are capable of measuring fine mode AOD from space, which roughly corresponds to atmospheric PM₂.₅ concentrations. The relationship between satellite-derived AOD and PM₂.₅ concentrations can be determined quantitatively either by linear regression models (Chu et al., 2002;
Wang and Christopher, 2003) or more complex models (Liu et al., 2005; 2007). Inverse modeling has been performed to constrain black carbon, particulate organic matter, sea salt, and desert dust emissions in global models using satellite-derived AOD recently (Dubovik et al., 2008; Huneeus et al., 2012). DDM is capable of tracking the relationships between PM$_{2.5}$ concentrations and emissions. Hence, the DKF inversion could be extended to constrain PM emissions in regional models using space-based AOD measurements.

4.3. Closing Remarks

To sum up, this research uses GOES observations to correct cloud underprediction errors and to readjust photolysis rates, and uses OMI observations combined with a well established DKF inversion method to identify and reduce uncertainties in the bottom-up NO$_x$ emission inventory for Texas O$_3$ attainment modeling. Results show both the promise and limitations of using satellite measurements to enhance O$_3$ regulatory modeling. The methods established in this study can be applied for the other regions. Although the satellite-derived top-down NO$_x$ emission inventory is recommended as a complement to the current bottom-up NO$_x$ emission inventory in Texas SIP modeling, it could be more useful in air quality modeling for places with immature bottom-up NO$_x$ emissions or remote areas with isolated emission sources.
References


Buzcu, B and Fraser, M. P.: Source identification and apportionment of volatile organic


Nam, J., Kimura, Y., Vizuete, W., Murphy, C., and Allen, D. T.: Modeling the impacts of
emission events on ozone formation in Houston, Texas. Atmos. Environ., 40, 5329–
5341, 2006.
sensitivity analysis for particulate matter (DDM-3D/PM). Atmos. Environ., 40,
Napelenok, S.L., Cohan, D.S, Odman, M.T., and Tonse, S.: Extension and evaluation of
sensitivity analysis capabilities in a photochemical model. Environ. Modell. Softw.,
evaluating spatially-resolved NOx emissions using Kalman filter inversion, direct
sensitivities, and space-based NO2 observations. Atmos. Chem. Phys., 8, 5603-5614,
2008.
National Research Council.: Models in environmental regulatory decision making.
National Research Council.: Air Quality Management in the United States.
Parrish, D. D., Allen, D. T., Bates, T. S., Estes, M., Fehsenfeld, F. C., Feingold, G.,
Ferrare, R., Hardey, R. M., Meagher, J. F., Nielsen-Gammon, J. W., Pierce, R. B.,
Ryserson, T. B., Seinfeld, J. H., and Williams, E. J.: Overview of the second Texas
Air Quality Study (TexAQS II) and the Gulf of Mexico Atmosphere Composition
and Climate Study (GoMACCS). J. Geophys. Res., 114, D00F13, doi:
Pinder, R.W., Gilliam, R.C., Appel, K.W., Napelenok, S.L., Foley, K.M., and Gilliland,
A.B.: Efficient probabilistic estimates of surface ozone concentration using an
ensemble of model configurations and direct sensitivity calculations. Environ. Sci.
Pison, I., Menut, L., and Bergametti, G.: Inverse modeling of surface NOx anthropogenic
emission fluxes in the Paris area during the Air Pollution Over Paris Region
2007.
Pleim, J.E.: A combined local and nonlocal closure model for the atmospheric boundary
layer. Part II: Application and evaluation in a mesoscale meteorological model. J.
Pour-Biazar, A., McNider, R. T., Roselle, S. J., Suggs, R., Jedlovec, G., Byun, D. W.,
Kim, S., Lin, C. J., Ho, T. C., Haines, S., Domblasser, B., and Cameron, R.:,
Correcting photolysis rates on the basis of satellite observed clouds. J. Geophys.
Prinn, R. G.: Measurement equation for trace chemicals in fluids and solution of its
inverse, in Inverse Methods in Global Biogeochemical Cycles, vol. 114, edited by
Kasibhatla, P., Heimann, M., Rayner, P., Mahowald, N., Prinn, R. G., and Hartley,
Quélo, D., Mallet, V., and Sportisse, B.: Inverse modeling of NOx emissions at regional
scale over northern France: preliminary investigation of the second-order sensitivity.


TCEQ.: Dallas-Fort Worth Attainment Demonstration SIP Revision for the 1997 Eight-Hour Ozone Standard Non-attainment Area, Austin, TX, 2011.


Appendix A

Supporting Information for Chapter 2

A1. Meteorological parameters evaluations

The MM5 modeled hourly temperature, wind speed, and wind direction are evaluated with measured data from 34 ground monitoring sites over the 12km CAMx domain for both modeling episodes (Table A1). The model simulates the temperature and wind speed well, showing that the mean bias error (MBE) is less than 0.5K and the root mean square error (RMSE) is less than 2K for the temperature, and the MBE and RMSE are around 2 m/s for the wind speed, which are similar to evaluation results from the study done by Kim et al. (2011). The simulated wind direction shows slightly weak performance in terms of RMSE. However, according to TCEQ (2010 and 2011), the large discrepancy between observed and modeled wind directions mostly occurred with very low wind speed. Thus, it only has slight influence on the CAMx modeling.

Modeled PBL heights are evaluated with available measurement data from two sites at Huntsville airport, Huntsville, TX (30.75°N, 95.58°W) and Jefferson airport, Port
Arthur, TX (29.94°N, 94.00°W) for the June episode (Fig. A1), and from three sites at Huntsville airport, Jefferson airport, and LaPorte airport, La Porte, TX (29.67°N, 95.06°W) for the August-September episode (Fig. A2). The model overpredicts the daytime PBL height for the June episode with exceptions at 8am and 7pm in the Huntsville site and 8am at Jefferson site (Fig. A1), while the August-September episode tends to underpredict PBL heights at Huntsville and Jefferson sites, but it shows a good agreement with measurement at LaPorte site (Fig. A2). The overprediction of PBL heights in the June episode, but underprediction of that in the August-September episode is probably caused by using different vertical mixing schemes in the MM5 modeling. The June episode uses ACM2 vertical mixing scheme to simulate PBL heights which tends to have stronger mixing (Pleim 2007).

![Figure A1. Temporal variations of monthly averaged modeled and measured PBL heights at Huntsville airport (30.75°N, 95.58°W) and Jefferson airport (29.94°N, 94.00°W).]
Figure A2. Temporal variations of monthly averaged modeled and measured PBL heights at Huntsville airport (30.75°N, 95.58°W), LaPorte airport (29.67°N, 95.06°W), and Jefferson airport (29.94°N, 94.00°W).

Table A1. Evaluation of MM5 in simulating hourly temperature, wind speed and wind direction from 34 ground monitoring sites for both June and Aug-Sep episodes.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>June 3 to July 1, 2006</th>
<th>August 16 to September 15, 2006</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T (K)</td>
<td>Wind Speed (m/s)</td>
</tr>
<tr>
<td>MBE(^a)</td>
<td>0.22</td>
<td>1.76</td>
</tr>
<tr>
<td>RMSE(^b)</td>
<td>1.58</td>
<td>2.13</td>
</tr>
</tbody>
</table>

\(^a\) Mean bias error  
\(^b\) Root mean square error
A2. Sensitivity analysis of the DKF inversions

To evaluate the responsiveness of the DKF inversion to the uncertainties in the a priori emission inventory and the observations, sensitivity tests by varying uncertainty values in the error covariance matrices for both OMI-based and ground-based DKF inversions in the actual inversion case are performed to investigate their impacts on the scaling factors generated by the inversion for each inversion region (Fig. A3). In the sensitivity tests for the emission error covariance matrix, the observation uncertainties were fixed to 30% for OMI and 15% for the AQS ground monitors (Fig. A3 top), but the emission uncertainties were varied from 50% to 200%. In contrast, the sensitivity tests for the observation error covariance matrix (Fig. A3 bottom) varied the observation uncertainties from 10% to 40% for OMI and from 10% to 25% for the ground measurements, but the emission uncertainties were fixed to 200%. The results from the sensitivity tests for the actual inversion cases are very similar to the pseudodata test: the adjustments for the posteriors are insensitive to the emission error covariance matrix and slightly responsive to the assumed observation errors, indicating the inversion results are robust.
Figure A3. Sensitivity analysis of Kalman filter inversion by changing emission uncertainties (top), and observation uncertainties (bottom) using OMI NO2 (left) and AQS ground NO2 (right).

A3. Direct Scaling (DS) inversion

A3.1 Method

The DS method applies the ratio between satellite NO2 observations and modeled NO2 concentrations to scale the bottom-up NOx emissions in each grid cell:
\[ E_t = E_b \times \frac{\Omega_s}{\Omega_m} \]  

where \( E_t \) is the top-down NO\(_x\) emission rate, \( E_b \) is the bottom-up NO\(_x\) emission rate, and \( \Omega_s \) and \( \Omega_m \) are the satellite and modeled NO\(_2\) column densities, respectively.

This method was developed in a global model with coarse grid resolution and assumes that the NO\(_2\) concentration in each model grid will not be affected by the NO\(_x\) emitted from surrounding grids. However, in a regional model with relatively small grid size, this assumption may fail, generating a spatial smearing error when NO\(_x\) lifetime is longer than the horizontal transport time (Martin et al., 2003; Boersma et al., 2008; Lamsal et al., 2010; Turner et al., 2012). Martin et al. (2003) indicated that the spatial smearing error can be neglected if the grid length is greater than 100km. Therefore, smoothing kernels (Toenges-Schuller et al., 2006; Boersma et al., 2008; Lamsal et al., 2010) need to be applied in order to alleviate the spatial smearing error in CAMx by accounting for the emissions from adjacent grid cells in developing the top-down NO\(_x\) emissions. The smoothing kernel is defined as

\[ K = \frac{1}{k + 8} \begin{pmatrix} 1 & 1 & 1 \\ 1 & k & 1 \\ 1 & 1 & 1 \end{pmatrix} \]  

where \( k \) is a smoothing parameter that is determined by applying the smoothing kernel \( K \) to each grid cell in the bottom-up NO\(_x\) emission inventory with different \( k \) values until the correlation between smoothed bottom-up NO\(_x\) emissions and corresponding CAMx modeled NO\(_2\) column density reaches a maximum. The smoothing kernel \( K \) is then applied to Eq. (A1) to form Eq. (A3),
\[ E'_{i,j} = \frac{\sum_{n=-1}^{1} \sum_{m=-1}^{1} K_{l,n} E^b_{i+l,j+n}}{\sum_{n=-1}^{1} \sum_{m=-1}^{1} \Omega_{l,n} E^b_{i+l,j+n}} \times E^b_{i,j} \]  

(A3)

where \( i \) and \( j \) represent column and row in horizontal model grids.

### A3.2 Results

The DS inversion method was performed with OMI NO\(_2\) column densities to create top-down NO\(_x\) emissions for the 12-km modeling domain. The monthly averaged (June 3 to July 1, and August 16 to September 15) NO\(_2\) column densities at 1-2pm were used to calculate the ratio of OMI to CAMx (Eq. A1). The first three modeling days were discarded for both modeling episodes to avoid the influence of initial conditions. The monthly 24-h averaged NO\(_x\) emissions and modeled NO\(_2\) column densities were used to determine the value of the smoothing parameter, \( k \). In this case, \( k \) equals 2.0 for both episodes, indicating large influence of NO\(_x\) emissions transported from surrounding grid cells.

Results (Table A2) show the DS inversion scales up the NO\(_x\) emissions in all seven regions, leading to higher estimates of modeled NO\(_2\) column densities (Fig. A4) in most of the domain. However, especially in urban areas, the simulated NO\(_2\) column densities with inverted NO\(_x\) emissions overshoot those observed by OMI. This indicates that the ability of NO\(_x\) to influence its own lifetime via changes in OH radical concentrations results in significant nonlinearity between NO\(_2\) concentration and NO\(_x\) emission that are neglected by the DS method. Use of inverted NO\(_x\) emissions does reduce bias and error in simulating OMI observed column densities, but \( R^2 \) decreases
(Table A3), indicating no improvement in the spatial distribution. The comparisons with AQS ground NO2 (Table A4) and O3 (Table A5) measurements indicate that the inverted NOx emissions actually deteriorate both simulations of ground-level NO2 and O3, with bias and error increasing by 0.7 against measured NO2, and with bias and error increasing by 0.1 against measured O3. Similar results are shown in evaluating model performance against P-3 NO2 and NOy measurements (Table A6). The DS inversion increases bias and error by approximately 0.3 and 0.2, respectively, in simulating P-3 NO2 and by around 0.7 and 0.6, respectively, in simulating P-3 NOy.

Table A2. Scaling factors for each region from DS inversions.

<table>
<thead>
<tr>
<th>Source Region</th>
<th>June 3 to July 1, 2006</th>
<th>August 16 to September 15, 2006</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Base NOx emission</td>
<td>Priori NOx emission*</td>
</tr>
<tr>
<td></td>
<td>(tons/day)</td>
<td>(tons/day)</td>
</tr>
<tr>
<td>HGB</td>
<td>374</td>
<td>455</td>
</tr>
<tr>
<td>DFW</td>
<td>335</td>
<td>435</td>
</tr>
<tr>
<td>BPA</td>
<td>81</td>
<td>97</td>
</tr>
<tr>
<td>NE Texas</td>
<td>141</td>
<td>164</td>
</tr>
<tr>
<td>Austin and San Antonio</td>
<td>252</td>
<td>319</td>
</tr>
<tr>
<td>N rural</td>
<td>522</td>
<td>823</td>
</tr>
<tr>
<td>S rural</td>
<td>472</td>
<td>728</td>
</tr>
</tbody>
</table>

* Adds lightning and aircraft NOx and doubled soil NOx emissions to the base case.
Table A3. Performance of CAMx in simulating OMI-observed NO$_2$ column densities.

<table>
<thead>
<tr>
<th>Statistical Parameters</th>
<th>June 3 to July 1, 2006</th>
<th>August 16 to September 15, 2006</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Base case</td>
<td>Priori$^c$</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.62</td>
<td>0.61</td>
</tr>
<tr>
<td>NMB$^a$</td>
<td>-0.47</td>
<td>-0.30</td>
</tr>
<tr>
<td>NME$^b$</td>
<td>0.48</td>
<td>0.32</td>
</tr>
</tbody>
</table>

a. Normalized mean bias  
b. Normalized mean error  
c. Adds lightning and aircraft NO$_x$ and doubled soil NO$_x$ emissions to the base case

Figure A4. Monthly averaged tropospheric NO$_2$ vertical columns at 1-2pm from OMI observations (left) and from CAMx simulations using a priori NO$_x$ emissions
(middle) and DS method inverted NO\textsubscript{x} emissions for both June (top) and August-September episode.

Table A4. Performance of CAMx in simulating AQS hourly ground-level NO\textsubscript{2}.

<table>
<thead>
<tr>
<th>Statistical Parameters</th>
<th>June 3 to July 1, 2006</th>
<th>August 16 to September 15, 2006</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Base case</td>
<td>Priori</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.56</td>
<td>0.56</td>
</tr>
<tr>
<td>NMB</td>
<td>0.89</td>
<td>0.98</td>
</tr>
<tr>
<td>NME</td>
<td>1.01</td>
<td>1.09</td>
</tr>
</tbody>
</table>

Table A5. Performance of CAMx in simulating AQS hourly ground-level O\textsubscript{3}.

<table>
<thead>
<tr>
<th>Statistical Parameters</th>
<th>June 3 to July 1, 2006</th>
<th>August 16 to September 15, 2006</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Priori</td>
<td>Posteriori OMI-based DS inversion</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.61</td>
<td>0.61</td>
</tr>
<tr>
<td>NMB</td>
<td>0.01</td>
<td>0.12</td>
</tr>
<tr>
<td>NME</td>
<td>0.29</td>
<td>0.37</td>
</tr>
</tbody>
</table>

Table A6. Performance of CAMx in simulating P-3 aircraft-observed NO\textsubscript{2} and NO\textsubscript{y}.

<table>
<thead>
<tr>
<th>Statistical Parameters</th>
<th>( \text{NO}_2 )\textsuperscript{a}</th>
<th>( \text{NO}_y )\textsuperscript{a}</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Base case</td>
<td>Priori</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.23</td>
<td>0.23</td>
</tr>
<tr>
<td>NMB</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>NME</td>
<td>0.99</td>
<td>0.99</td>
</tr>
</tbody>
</table>

\textsuperscript{a}. Comparison available for only four days (August 31, September 11, September 13, and September 15, 2006).
Appendix B

Supporting Information for Chapter 3

B1. CAMx modeled profile-based OMI retrieval

The OMI-retrieved tropospheric NO$_2$ vertical column density (VCD) used in this study is calculated via Eq. (B1) (Bucsela et al., 2013),

\[
V_{c(GEOSChem)} = \frac{S_{c(OMI)}}{AMF_{GEOSChem}}
\]  

(B1)

where $S_{c(OMI)}$ is the OMI tropospheric NO$_2$ slant column density, $AMF$ stands for the air mass factor which is computed based on a priori GEOS-Chem modeled profile and scattering weights ($SW$) calculated by the TOMRAD model (Bucsela et al., 2013), and $V_{c(GEOSChem)}$ is the GEOS-Chem modeled profile-based OMI tropospheric NO$_2$ VCD. A satellite NO$_2$ retrieval error analysis study (Boersma et al., 2004) shows that the estimated a priori profile from global models may contribute approximately 10% uncertainty to the AMF calculations and propagate that uncertainty to the retrieved NO$_2$ VCD. Therefore, when OMI VCD is compared to any modeled VCD, OMI averaging kernels ($AKs$) (Eskes and Boersma, 2003), calculated in Eq. (B2), are recommended to be
applied to the modeled VCDs via Eq. (B3), in order to remove the influence from the a priori profile used in the OMI retrievals.

\[ AK_i = \frac{SW_i}{AMF_{\text{GEOSChem}}} \]  

(B2)

\[ C_{\text{NO}_2}^{\text{predicted}} = \sum (AK_i \times CAMx_{\text{vci}}) = \sum (\frac{SW_i}{AMF_{\text{GEOSChem}}} \times CAMx_{\text{vci}}) = \frac{\sum (SW_i \times CAMx_{\text{vci}})}{AMF_{\text{GEOSChem}}} \]  

(B3)

In Eq. (B3), \( CAMx_{\text{vci}} \) represents the CAMx modeled NO\(_2\) VCD at each model layer \( i \), and \( CAMx_{\text{vctot}} \) is the CAMx modeled total tropospheric VCD. The AMF which contains the a priori GEOS-Chem modeled profile is now merged with the CAMx modeled VCD.

The way of removing the a priori GEOS-Chem modeled profile via applying AKs is carried out by generating the CAMx modeled profile-based \( AMF_{\text{CAMx}} \) as shown in Eq. (B4),

\[ AMF_{\text{CAMx}} = \sum (SW_i \times \frac{CAMx_{\text{vci}}}{CAMx_{\text{vctot}}}) = \frac{\sum (SW_i \times CAMx_{\text{vci}})}{CAMx_{\text{vctot}}} \]  

(B4)

using \( AMF_{\text{CAMx}} \) to replace \( AMF_{\text{GEOSChem}} \) in Eq. (B1) and then creating a CAMx modeled profile-based OMI tropospheric NO\(_2\) VCD \( (V_{c(CAMx)}) \). However, this procedure can only be realized in the inversion process by comparing the AKs applied CAMx VCD \( (C_{\text{NO}_2}^{\text{predicted}}) \) and original OMI retrieved VCD \( (V_{c(GEOSChem)}) \).
The numerator in Eq. (B3) can be replaced by the $AMF_{\text{CAMx}}$ generated in Eq. (B4) to form Eq. (B5).

\[
C_{\text{NO}_2}^{\text{predicted}} = \text{CAMx}_{\text{vctot}} \times \frac{AMF_{\text{CAMx}}}{AMF_{\text{GEOSChem}}} \tag{B5}
\]

When applying $C_{\text{NO}_2}^{\text{predicted}}$ to the direct scaling method (Martin et al., 2003; Tang et al., 2013) in Eq. (B6),

\[
E_t = E_b \times \frac{V_{v(\text{GEOSChem})}}{C_{\text{NO}_2}^{\text{predicted}}} = E_b \times \frac{S_{v(\text{OMI})}}{AMF_{\text{GEOSChem}}} \times \frac{AMF_{\text{GEOSChem}}}{AMF_{\text{CAMx}}} = E_b \times \frac{V_{v(\text{CAMx})}}{\text{CAMx}_{\text{vctot}}} = E_b \times \frac{V_{v(\text{CAMx})}}{\text{CAMx}_{\text{vctot}}} = E_b \times \frac{S_{v(\text{OMI})}}{AMF_{\text{CAMx}}} \tag{B6}
\]

the $AMF_{\text{GEOSChem}}$ is canceled, and $V_{v(\text{CAMx})}$ is formed through $AMF_{\text{CAMx}}$ to compare with the CAMx modeled VCD directly.

When applying OMI AKs to the CAMx modeled NO$_2$ and its sensitivity VCDs in the DKF method (Tang et al., 2013) as shown in Eq. (B7),

\[
\hat{x}_{\text{NO}_2} = x_{\text{NO}_2}^{*} + P_{\text{NO}_2}^{*} \times (AMF_{\text{CAMx}}^{\text{AMF}}_{\text{GEOSChem}})S_{w}^{T} \times ((AMF_{\text{CAMx}}^{\text{AMF}}_{\text{GEOSChem}})_{w}^{2} S_{w}^{T} P_{\text{NO}_2}^{*} S_{w}^{T} + (\frac{S_{v(\text{OMI})}}{AMF_{\text{GEOSChem}}})^{2})^{-1} \times ((\frac{S_{v(\text{OMI})}}{AMF_{\text{GEOSChem}}}) - (\text{CAMx}_{\text{vctot}} \times \frac{AMF_{\text{CAMx}}}{AMF_{\text{GEOSChem}}}) - (\frac{AMF_{\text{CAMx}}}{AMF_{\text{GEOSChem}}})S_{w}^{*} x_{\text{NO}_2}^{*}) \tag{B7}
\]

where $\varepsilon_{\text{out}}$ is the OMI measurement uncertainty, Eq. (B8) is derived.
\[ \hat{x}_{\text{NO}_x} = x_{\text{NO}_x}^* + P_{\text{NO}_x}^* \times S_w^T \times (S_w^* P_{\text{NO}_x}^* S_w^T + \left( \frac{S_{\text{OMI}}^*}{AMF_{\text{GEOSChem}}} \right)^2 I)^{-1} \times \left( \frac{AMF_{\text{GEOSChem}}}{AMF_{\text{CAMx}}} S_{\text{OMI}}^* - \left( \frac{AMF_{\text{CAMx}}}{AMF_{\text{GEOSChem}}} \right) S_w^* x_{\text{NO}_x}^* \right) \]  

(B8)

and further transformed to Eq. (B9),

\[ \hat{x}_{\text{NO}_x} = x_{\text{NO}_x}^* + P_{\text{NO}_x}^* \times S_w^T \times (S_w^* P_{\text{NO}_x}^* S_w^T + (V_{c(CAMx)}^* e_{\text{OMI}}^*)^2)^{-1} \times \left( (V_{c(CAMx)}^* - CAMx_{\text{nox}} - S_w^* x_{\text{NO}_x}^*) \right) \]  

(B9)

where all \( AMF_{\text{GEOSChem}} \) are removed, and the original \( V_{c(GEOSChem)} \) becomes \( V_{c(CAMx)} \).

There is an alternative way to create \( V_{c(CAMx)} \) instead of applying OMI AKs, which is to use the CAMx modeled profile directly in the OMI retrieval process. In this case, the error of interpolating AKs values into the CAMx layer can be avoided, and the CAMx profile-based OMI retrieval can be calculated and viewed directly. In this study, we have created a CAMx profile-based OMI product that uses a CAMx profile in the retrieval process for the AMF calculation and planned to use this new OMI retrieval product at the beginning for the inversion study. However, we find that the CAMx profile-based OMI overestimates NO2 VCD by approximately 30% compared to the original OMI retrieval using a GEOS-Chem profile (Fig. B1 right). We further compare the monthly averaged 13:00-14:00LT CAMx NO2 profile to the GEOS-Chem NO2 profile over the 12km domain (Fig. B1 left) and find that the CAMx profile shows much higher amounts of NO2 in the boundary layer but lower amounts of NO2 in the upper troposphere. This may reduce the AMF values (Eq. B4) because instrument sensitivity related SW is much higher in the upper troposphere than in the boundary layer and thus increases the total retrieval quantity. Unfortunately, there are no corresponding
measurement data available to validate the CAMx and GEOS-Chem profiles in Fig. (B1), but similar bias has been found in the CAMx modeled NO$_2$ profile compared to the DC-8 and P-3 aircraft NO$_2$ measurements (Fig. 3.8). Using the CAMx profile here may introduce more errors to the OMI retrieval and inversions; hence, we do not recommend to either apply AK to the CAMx modeled VCD or to use the CAMx profile-based OMI in this study unless the CAMx profile is validated.

Figure B1. Comparisons between GEOS-Chem and CAMx modeled NO$_2$ vertical profiles (left) and corresponded OMI retrievals (right). Filled circles represent observations under clear sky condition (cloud fraction $<$0.5), and open circles are all observations.
B2. Impact of increased NOx lifetime and artificial layer on modeled NO2 VCD

The NASA OMI high resolution product used in this study shows reduced NO2 in the rural areas, while enhanced NO2 in the urban, compared to the NASA standard retrieval, version 2 (Tang et al., 2013); however, it still shows more smeared-out pattern than the CAMx modeled NO2 VCD (Fig. B3a). The CAMx simulations with the a priori NOx emission inventory created in Tang et al. (2013) shows larger NO2 VCD in the cities, while lower NO2 VCD in the rural places than OMI (Fig. B3b). Reducing the reaction rate constant of the reaction OH + NO2 by 25% in the CB05 chemical mechanism increases the NOx lifetime, makes more NOx transport to the rural, and enhances around 3% NO2 VCD on average in the inversion region, but the impact is small (Fig. B3c). Implementing 40ppt NO2 homogeneously into the model top layer adds about $1.6 \times 10^{14}$ molecules/cm$^2$ NO2 densities to each model grid and increases approximately 8% NO2 VCD in the inversion region, further alleviating the NO2 gap between OMI and CAMx in the rural areas (Fig. B3d).
Figure B2. Sensitivities of the DKF inversions to the uncertainties in emissions (left) and in OMI observations (right) in region-based inversion (top), sector-based inversion case I (middle), and sector-based inversion case II (bottom).
**B3. Sensitivity of DKF inversion to error covariance matrices**

The sensitivities of the DKF inversion-generated scaling factors to the uncertainties in the emission and observation error covariance matrices are tested for both region-based and sector-based DKF inversions to evaluate the robustness of the inversion results (Fig. B2). The OMI observation uncertainties are fixed to 30% in the sensitivity tests for the emission error covariance matrix, while the emission uncertainties are varied from 50% to 100% (Fig. B2 left). In contrast, the OMI observation uncertainties are varied from 10% to 50% in the sensitivity tests for the observation error covariance matrix, while the emission uncertainties in each sector are fixed to 100% (Fig. B2 right).

In the region-based inversion, the emission uncertainties have insignificant impact on the inversion results. The inversion seems to be relatively responsive to the lower observation uncertainties, but results become more stable when the uncertainties are over 30% (Fig. B2 top). In the sector-based inversion, the scaling factors decrease when uncertainties in the observations increase, but the inversion results are less sensitive to the emission uncertainties. However, an exception occurs in the sector-based DKF inversion case I, where the adjustments in the aviation sector are relatively more sensitive to the emission uncertainty, ranging from 3.9 to 4.6 when emission uncertainty increases from 50% to 100%. It seems to offset against area and nonroad sector which the scaling factors reduce from 0.6 to 0.5 (Fig. B2 middle). However, the inversion becomes insensitive to the emission uncertainties in the sector-based DKF inversion case II when merging aviation into the area and nonroad sector (Fig. B2 bottom), indicating the DKF inversion in case II is more stable and less responsive to the uncertainty matrices than that in case I.
Figure B3. Monthly averaged (16 August to 15 September) tropospheric NO₂ VCDs at 13:00-14:00LT from (a) OMI, (b) simulations using NOₓ emissions from Tang et al., (2013), (c) simulations with the lower rate constant of the reaction OH+NO₂ from (b), and (d) simulations with added 40ppt NO₂ layer from (c).
B4. Top-down VOC emissions

Five VOC species, ethylene (ETH), ethane (ETHA), isoprene (ISOP), toluene (TOL), and xylene (XYL) are chosen to conduct the inversion in this study because of their explicit model outputs and sufficient measurement data. ETH, ISOP, TOL, and XYL are defined as highly reactive VOC (HRVOC) by TCEQ for regulatory purposes, due to their high reactivity with OH and propensity for contributing to rapid O₃ formation (Thomas et al., 2008). Although ETHA is not a HRVOC, the high concentrations in urban environments make it also play very important role in forming O₃ (Katzenstein et al., 2003; Buzcu and Fraser, 2006).

Table B1. Emission rates of five VOC species for six emission sectors in the inversion region.

<table>
<thead>
<tr>
<th>VOCs</th>
<th>Area (tons/day)</th>
<th>Mobile (tons/day)</th>
<th>Non-road (tons/day)</th>
<th>Biogenic (tons/day)</th>
<th>No-EGU points (tons/day)</th>
<th>EGU points (tons/day)</th>
<th>Total (tons/day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ETH</td>
<td>19.2 (11.5%)</td>
<td>14.9 (8.9%)</td>
<td>11.1 (6.6%)</td>
<td>104.8 (62.6%)</td>
<td>17.2 (10.3%)</td>
<td>0.1 (0.06%)</td>
<td>167.3</td>
</tr>
<tr>
<td>ETHA</td>
<td>232.4 (82.3%)</td>
<td>0 (0%)</td>
<td>5 (1.8%)</td>
<td>22.5 (8.0%)</td>
<td>20.4 (7.2%)</td>
<td>2.1 (0.7%)</td>
<td>282.4</td>
</tr>
<tr>
<td>ISOP</td>
<td>0.4 (0.002%)</td>
<td>0.8 (0.005%)</td>
<td>0.5 (0.003%)</td>
<td>15835.8 (99.9%)</td>
<td>0.2 (0.001%)</td>
<td>0 (0%)</td>
<td>15837.9</td>
</tr>
<tr>
<td>TOL</td>
<td>53.3 (48.9%)</td>
<td>24.5 (22.5%)</td>
<td>25.1 (23.1%)</td>
<td>0 (0%)</td>
<td>5.3 (4.9%)</td>
<td>0.7 (0.6%)</td>
<td>108.9</td>
</tr>
<tr>
<td>XYL</td>
<td>116.7 (58.3%)</td>
<td>38.2 (19.1%)</td>
<td>39.7 (19.8%)</td>
<td>0 (0%)</td>
<td>3.3 (1.6%)</td>
<td>2.2 (1.1%)</td>
<td>200.1</td>
</tr>
</tbody>
</table>

Note: percentage indicates the apportionment of each emission sector to the regional total.
B4.1 Base case VOC emission inventory

The base case VOC emission inventory for the HGB SIP modeling from 13 August to 15 September 2006 was developed by TCEQ (Table B1). The non-EGU point source VOC emissions were from the State of Texas Air Reporting System (STARS) database, a special inventory containing reported hourly VOC emissions from 15 August to 15 September targeting a specific list of non-EGU points and from Tank Landing Loss surveys of hourly landing loss VOC emissions. The EGU point source VOC emissions were from the EPA Acid Rain database (ARD) with the emissions calculated based on VOC: NO\textsubscript{x} ratios. The VOC emissions from motor vehicles were generated by the Motor Vehicle Emission Simulator 2010a (MOVES2010a) model for the on-road vehicles and the Texas NONROAD (TexN) model for the off-road vehicles. The VOC emissions from the other non-road and area sources were from the Texas Air Emissions Repository (TexAER) database (TCEQ 2010). The Global Biosphere Emissions and Interactions System model, version 3.1 (GloBEIS3.1) was used for developing biogenic VOC emissions (Yarwood et al., 1999). Four HRVOC species emissions, ethylene, propylene, 1,3-butadiene, and butenes were further corrected using the Potential Source Contribution Function (PSCF) technique with Automatic Gas Chromatographs (Auto-GC) measured data in the HGB area (TCEQ 2010).

For the five chosen VOC species, ETH and ISOP emissions are mostly contributed by the biogenic source around 60% and 99%, respectively, while TOL and XYL are entirely anthropogenic, originating mostly from area emissions. Area sources
also dominate emissions of ETHA, which does not appear in the on-road mobile source. EGUs emissions are minor contributors to all five VOC species (Table B1).

**B4.2 VOC observations**

The U.S. EPA Photochemical Assessment Monitoring Stations (PAMS) VOC measurement data ([http://www.epa.gov/ttn/airs/airsaq/s/](http://www.epa.gov/ttn/airs/airsaq/s/)) are used here to adjust emissions for the five chosen VOC species. All five VOC species were measured by the gas chromatographs-flame ionization detector (GC-FID) with 1-hr resolution for the entire modeling period from 13 August to 15 September 2006 in the unit of ppmC (U.S. EPA 1998). Measurements are available only for a total of 11 PAMS monitoring sites in the inversion region: 2 in DFW, 3 in BPA and 6 in HGB (Fig 3.1). The measurement data are first converted into the unit of ppb for each VOC species, and then averaged monthly over all monitoring sites in each region and compared to the corresponding modeled data.

The NOAA P-3 aircraft measured VOC data ([http://www.esrl.noaa.gov/csd/tropchem/2006TexAQS/](http://www.esrl.noaa.gov/csd/tropchem/2006TexAQS/)) are further used for evaluating the model performance in simulating aloft VOCs. Only four chosen VOC species, ETH, ISOP, TOL, and XYL are measured by P-3. ETH is measured using Laser Photoacoustic Spectroscopy (LPAS) with 20s resolution (de Gouw et al., 2009), and ISOP, TOL, and XYL are measured using Proton Transfer Reaction Mass Spectrometer (PTRMS) with 15s resolution (de Gouw et al., 2003). The P-3 measured ISOP, TOL, and XYL are available on 4 days (31 August, 11 September, 13 September, and 15 September 2006), while measured ETH is only available on 3 days (31 August, 13 September, and 15
September 2006) during our modeling period. The P-3 measured VOC data are averaged hourly and compared with the hourly modeled data at corresponding grid cells.

### B4.3 Results

Since all modeled ETH, ETHA, ISOP, TOL, and XYL are from the primary emissions, a direct scaling (DS) inversion method that adjusts VOC emissions based on the ratios between modeled VOC and PAMS measured VOC is applied here. The inversion is conducted on a regional basis, which means the scaling factor calculated from the measurement data in one region only applies to adjust the emissions in that region. Therefore, due to the availability of observations, the five chosen VOC species emissions are adjusted in only three regions, DFW, HGB, and BPA.

The scaling factors generated from the inversions vary significantly in different regions (Table B2) and show that the HRVOC emissions in the 2006 TCEQ emission inventory for HGB SIP modeling are much better than the reported uncertainty of an order of magnitude (Ryerson et al., 2003; Parrish et al., 2009) but still much higher than the uncertainty in NOx emissions. The ETHA emissions require the largest adjustments in all three regions with scaling factors ranging from 3.14 to 4.63. The inversion scales down ETH emissions in the HGB and DFW regions by only 10%, but in BPA, it requires a scaling factor of 3.33. The mostly biogenic source contributed ISOP emission only requires 4% scale-up adjustment in HGB, but relatively larger scale-down adjustments ranging from 30-50% in DFW and BPA. The anthropogenic source contributed TOL emissions require scale-up adjustments in all three regions by scaling factors ranging
from 1.32 to 2.22. The XYL emissions are well estimated in the base case emission inventory for the HGB region, but require scale-down by approximately 70% in DFW and scale-up around 50% in BPA.

The temporal variations of the five VOC species (Fig.B4) show that the discrepancies between observed VOCs and the a priori modeled VOCs are significantly reduced by using the a posteriori emissions. The inverted ETHA emission improves modeled $R^2$ and reduces model bias and error by 0.5 and 0.1, respectively (Table B3). The inverted ETH shows increased $R^2$ and 0.13 reduced bias, but no improvement in the model error against ground measurement (Table B3); however, it shows 0.4 reductions in both model bias and error against P-3 measured data (Table B4). The inverted ISOP emissions reduce approximately 20% bias and error in ground ISOP simulation (Table B3), but no improvements are found compared against aircraft measurement (Table B4). The model bias in the inverted TOL is reduced by approximately 0.4 (Table B3) compared against PAMS and 0.13 compared against P-3 (Table B4), while the model error has not been improved. The inverted XYL shows increased $R^2$ and around 0.2 reduced model bias and error compared to ground measurement (Table B3) and 0.02 reduced model bias and error compared to aircraft measurement (Table B4). However, no improvements are found in the model performance of simulating ground-level NO$_2$ (Table B5), and there is a slight decreasing, around 0.01, of model bias and error in ground-level O$_3$ simulations using the inverted VOC emissions (Table B6).
Figure B4. Comparisons of monthly averaged daily variation between observed (black) and modeled VOC species using the a priori (red) and the a posteriori (blue) VOC emission inventory over all monitoring sites.

Table B2. Direct scaling factors for VOC species in three inversion regions.

<table>
<thead>
<tr>
<th>Source Region</th>
<th>ETHA</th>
<th>ETH</th>
<th>ISOP</th>
<th>TOL</th>
<th>XYL</th>
<th>ETHA</th>
<th>ETH</th>
<th>ISOP</th>
<th>TOL</th>
<th>XYL</th>
</tr>
</thead>
<tbody>
<tr>
<td>HGB</td>
<td>52.7</td>
<td>26.4</td>
<td>635.5</td>
<td>23.9</td>
<td>42.1</td>
<td>3.45</td>
<td>0.92</td>
<td>1.04</td>
<td>1.71</td>
<td>0.98</td>
</tr>
<tr>
<td>DFW</td>
<td>14.3</td>
<td>11.5</td>
<td>780.5</td>
<td>20.6</td>
<td>45.1</td>
<td>4.63</td>
<td>0.90</td>
<td>0.71</td>
<td>1.32</td>
<td>0.33</td>
</tr>
<tr>
<td>BPA</td>
<td>27.6</td>
<td>7.1</td>
<td>282.2</td>
<td>5.7</td>
<td>6.9</td>
<td>3.14</td>
<td>3.33</td>
<td>0.50</td>
<td>2.22</td>
<td>1.47</td>
</tr>
</tbody>
</table>
Table B3. Evaluation of CAMx modeled VOCs using hourly PAMS-measured VOCs.

<table>
<thead>
<tr>
<th>Source Region</th>
<th>ETH</th>
<th>ETH</th>
<th>ISOP</th>
<th>TOL</th>
<th>XYL</th>
<th>ETH</th>
<th>ETH</th>
<th>ISOP</th>
<th>TOL</th>
<th>XYL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R²</td>
<td>NMB</td>
<td>NME</td>
<td>R²</td>
<td>NMB</td>
<td>NME</td>
<td>R²</td>
<td>NMB</td>
<td>NME</td>
<td></td>
</tr>
<tr>
<td>ETHA</td>
<td>0.12</td>
<td>-0.71</td>
<td>0.73</td>
<td>0.05</td>
<td>-0.20</td>
<td>0.80</td>
<td>0.04</td>
<td>1.04</td>
<td>0.63</td>
<td>1.04</td>
</tr>
<tr>
<td>ETH</td>
<td>0.09</td>
<td>0.32</td>
<td>0.63</td>
<td>0.07</td>
<td>0.24</td>
<td>0.90</td>
<td>0.13</td>
<td>0.90</td>
<td>0.61</td>
<td>0.90</td>
</tr>
<tr>
<td>ISOP</td>
<td>0.09</td>
<td>-0.41</td>
<td>-0.53</td>
<td>0.07</td>
<td>-0.22</td>
<td>-0.53</td>
<td>0.13</td>
<td>0.22</td>
<td>-0.53</td>
<td>0.13</td>
</tr>
<tr>
<td>TOL</td>
<td>0.13</td>
<td>-0.07</td>
<td>-0.07</td>
<td>0.04</td>
<td>0.05</td>
<td>0.04</td>
<td>0.10</td>
<td>0.05</td>
<td>0.04</td>
<td>0.10</td>
</tr>
<tr>
<td>XYL</td>
<td>0.12</td>
<td>-0.03</td>
<td>0.05</td>
<td>0.09</td>
<td>0.01</td>
<td>0.09</td>
<td>0.12</td>
<td>0.01</td>
<td>0.09</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Table B4. Evaluation of CAMx modeled VOCs using P-3 aircraft-measured VOCs\(^a\).

<table>
<thead>
<tr>
<th>Source Region</th>
<th>ETH(^b)</th>
<th>ISOP</th>
<th>TOL</th>
<th>XYL(^c)</th>
<th>ETH(^b)</th>
<th>ISOP</th>
<th>TOL</th>
<th>XYL(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NMB</td>
<td></td>
<td></td>
<td></td>
<td>NMB</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ETH</td>
<td>-0.63</td>
<td>-0.81</td>
<td>-0.60</td>
<td>-0.53</td>
<td>-0.81</td>
<td>-0.47</td>
<td>-0.51</td>
<td></td>
</tr>
<tr>
<td>ISOP</td>
<td>0.84</td>
<td>1.05</td>
<td>0.72</td>
<td>0.80</td>
<td>1.05</td>
<td>0.72</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>TOL</td>
<td>0.46</td>
<td>0.43</td>
<td>0.66</td>
<td>0.49</td>
<td>0.43</td>
<td>0.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>XYL(^c)</td>
<td>0.46</td>
<td>0.43</td>
<td>0.66</td>
<td>0.49</td>
<td>0.43</td>
<td>0.66</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. Comparison available for four days (31 August, 11 September, 13 September, and 15 September 2006).
b. Comparison only available for three days (31 August, 13 September, and 15 September 2006).
c. Compared with measured C-8 aromatics.

Table B5. Evaluation of CAMx modeled NO\(_2\) using hourly AQS ground-measured NO\(_2\).

<table>
<thead>
<tr>
<th>Source Region</th>
<th>R(^2)</th>
<th>NMB</th>
<th>NME</th>
<th>R(^2)</th>
<th>NMB</th>
<th>NME</th>
</tr>
</thead>
<tbody>
<tr>
<td>HGB</td>
<td>0.51</td>
<td>0.46</td>
<td>0.67</td>
<td>0.51</td>
<td>0.46</td>
<td>0.67</td>
</tr>
<tr>
<td>DFW</td>
<td>0.49</td>
<td>0.43</td>
<td>0.66</td>
<td>0.49</td>
<td>0.43</td>
<td>0.66</td>
</tr>
<tr>
<td>BPA</td>
<td>0.45</td>
<td>0.92</td>
<td>1.02</td>
<td>0.45</td>
<td>0.92</td>
<td>1.02</td>
</tr>
<tr>
<td>Overall</td>
<td>0.51</td>
<td>0.51</td>
<td>0.72</td>
<td>0.51</td>
<td>0.51</td>
<td>0.73</td>
</tr>
</tbody>
</table>
Table B6. Evaluation of CAMx modeled O₃ using hourly AQS ground-measured O₃.

<table>
<thead>
<tr>
<th>Source Region</th>
<th>Priori R²</th>
<th>Priori NMB</th>
<th>Priori NME</th>
<th>Posteriori R²</th>
<th>Posteriori NMB</th>
<th>Posteriori NME</th>
</tr>
</thead>
<tbody>
<tr>
<td>HGB</td>
<td>0.46</td>
<td>0.68</td>
<td>0.75</td>
<td>0.46</td>
<td>0.68</td>
<td>0.75</td>
</tr>
<tr>
<td>DFW</td>
<td>0.64</td>
<td>0.21</td>
<td>0.32</td>
<td>0.64</td>
<td>0.20</td>
<td>0.31</td>
</tr>
<tr>
<td>BPA</td>
<td>0.47</td>
<td>0.66</td>
<td>0.70</td>
<td>0.46</td>
<td>0.65</td>
<td>0.69</td>
</tr>
<tr>
<td>Overall</td>
<td>0.50</td>
<td>0.42</td>
<td>0.50</td>
<td>0.50</td>
<td>0.41</td>
<td>0.49</td>
</tr>
</tbody>
</table>
Appendix C

List of publications and conference presentations

C1. Journal publications

1. **Wei Tang, Daniel Cohan, Lok Lamsal, Arastoo Pour-Biazar, Xue Xiao, Wei Zhou.** Influence of satellite-derived photolysis rates and NOx emissions on Texas ozone modeling. (In preparation for *Atmospheric Chemistry and Physics*)

2. **Wei Tang, Daniel Cohan, Lok Lamsal, Xue Xiao, Wei Zhou.** Inverse modeling of Texas NOx emissions using space-based and ground-based NO2 observations. *Atmospheric Chemistry and Physics*, 2013, 13, 11005-11018.


C2. Conference presentations

1. **Wei Tang, Daniel Cohan.** Inverse modeling of Texas NOx emissions. 6th biannual AQAST meeting, Houston, TX. Jan, 2014.

2. Daniel Cohan, **Wei Tang, Ben Lash, Lun Li.** Characterization of nitrogen oxide emissions and trends. 5th biannual AQAST meeting, College Park, MD. June, 2013.

3. **Wei Tang, Daniel Cohan, Lok Lamsal, Arastoo Pour-Biazar.** Influence of uncertainties on the inverse modeling of Texas NOx emissions. 93th AMS conference, Austin, TX. January, 2013

4. **Wei Tang, Daniel Cohan, Lok Lamsal.** Kalman Filter inversion on Texas NOx emissions using space-based NO2 observations. 11th CMAS Conference, Chapel