Constraining sources of atmospheric trace constituents with Lagrangian particle dispersion modeling

by

Joshua Benmergui

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Earth Sciences

Waterloo, Ontario, Canada, 2013

© Joshua Benmergui 2013

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

This manuscript based thesis examines and advances methods for constraining sources of atmospheric trace constituents with a Lagrangian particle dispersion model. The method of Bayesian inversion is demonstrated, and a new method is introduced to a class of similar problems where established methods are not applicable. First, A new regression based methodology was developed and applied to observations of atmospheric methanesulfonic acid mass concentrations at Alert, Nunavut. The methodology was used to compare the importance of phytoplankton blooms vs. the ice-free ocean as sources of the dimethylsulfide precursor, and to compare the importance of bromine monoxide vs. hydroxyl as agents oxidizing dimethylsulfide to methanesulfonic acid. These issues are relevant to the application of methanesulfonic acid concentrations in ice cores to determine historic sea ice properties. The analysis indicated that source regions to Alert during the spring are primarily ice-free ocean with a significant contribution from ice edge blooms, and during the summer to be dominated by the ice-free ocean. The model also indicated that oxidation of DMS by BrO was the dominant source of MSA in the spring, while DMS oxidation by OH was the dominant source in the summer. Secondly, Bayesian inversion was applied to observations of atmospheric elemental carbon mass concentrations at Tsinghua University in Beijing, China. The analysis provided evidence that current bottom-up elemental carbon emissions estimates in northern China are likely underpredicted. Global chemical transport models show ubiquitous underestimates of the atmospheric burden of elemental carbon, especially near large sources of emissions. Northern China is among the regions with the most intensive elemental carbon emissions in the world, and an underestimate of emissions in this region may be partially responsible for the global chemical transport model underestimates.

Acknowledgements

I am forever indebted to John C. Lin, my mentor, who patiently developed me from a student to a scientist. I would like to thank Andrew Platt, who brought me to the Arctic, ignighting my passion for atmospheric science, Kevin Anderson, Marjorie Shepherd, and everyone else who has worked to make the N. B. A. Trivet Global Atmosphere Watch Lab in Alert, Nunavut such an extraordinary place. I thank the brilliant scientists who I was lucky enough to collaborate with on the publications included in this thesis; John C. Lin, Sangeeta Sharma, Deyong Wen, Lin Huang, Qiang Zhang, Sicong Kang, and Myung Gwang Kim. I would like to thank those who contributed resources, data, thoughts, and observations to this work. These include Desiree Toom-Sauntry, Xin Yang, Dylan Millet, Lu Hu, Fumo Yang, Fenqui Duan, Kebin He, Helen Macintyre, The United States Geological Survey, The European Centre for Medium-Range Weather Forecasts, the National Oceanic & Atmospheric Administration/Office of Oceanic and Atmospheric Reasearch/Earth System Reseach Laboratory Pysical Science Division, The National Snow and Ice Data Center, The Shared Hierarchical Academic Research Computing Network, and Compute Canada. I thank Environment Canada for funding my research. Lasty I thank Tonya Burgers, who makes life and science so much more enjoyable.

Dedication

To Patrick Castle, without him I would not know science.

Table of Contents

List of Tables xv

List of Figures xvii

1 Introduction 1

1.1 A Brief History of Lagrangian Particle Dispersion Models 2

1.2 Thesis Outlook 6

2 Quantitative attribution of processes affecting atmospheric chemical concentrations by combining a time-reversed Lagrangian particle dispersion model and a regression approach 7

Overview 9

2.1 Introduction 10

2.2 Application of Framework: Methanesulfonic Acid in the Arctic 13

2.3 Methodology 15

2.4 Fluxes 16

2.5 Deposition 18

2.6 Reactions and Initial Concentrations 19

2.7 Model Synthesis 21

2.8 Interpretation of Regression Results 23

2.9 Results of MSA Analysis 24

2.10 Discussion 32

3 Inverse analysis of elemental carbon aerosol in northern China: Evidence of an underprediction in bottom-up emissions 35

Overview . . . 36

3.1 Introduction . . . 37

3.2 Inverse Modeling Framework . . . 40

3.3 A Priori Emissions Estimates . . . 42

3.4 Measurements . . . 45

3.5 Simulating Elemental Carbon Concentrations . . . 46

3.5.1 Background Concentrations . . . 47

3.5.2 Contribution of Emissions . . . 49

3.5.3 Hydrophobic-to-Hydrophilic Conversion . . . 49

3.5.4 Deposition . . . 49

3.6 Uncertainty Characterization . . . 51

3.6.1 Measurements . . . 51

3.6.2 Aggregation . . . 51

3.6.3 Transport . . . 52

3.6.4 Planetary Boundary Layer . . . 54

3.6.5 Deposition . . . 56

3.6.6 Background . . . 56

3.6.7 Particle Number . . . 57

3.6.8 Construction of the $\mathbf{S}_\epsilon$ Matrix . . . 58

3.7 Results . . . 59

3.8 Discussion . . . 68

3.9 Conclusions . . . 70

4 Conclusions 73

Appendices **75**

Appendix A Description of Algorithms Used for Elemental Carbon Removal Processes **75**

A.1 Hydrophobic-to-Hydrophilic Conversion 75

A.1.1 Hydrophobic-to-Hydrophilic Conversion by Condensation of Sulphate and Nitrate 75

A.1.2 Hydrophobic-to-Hydrophilic Conversion by Coagulation 76

A.1.3 Hydrophobic-to-Hydrophilic Conversion by Oxidation 77

A.1.4 Consolidation of Hydrophobic-to-Hydrophilic Conversion Time Scale 78

A.2 Dry Deposition 79

A.3 Wet Deposition 83

A.3.1 Convective Updraft 83

A.3.2 Rainout 83

A.3.3 Washout 85

Copyright Permissions **87**

References **89**

List of Tables

2.1 Explanations of Symbols Used in Equations. 20

2.2 Regression Coefficients and Wald Tests of Significance for methanesulfonic acid Regression Model for May and August. "Ice-Free" is the variable for ice-free ocean dimethylsulfide emissions, "Bloom" is the variable for phytoplankton bloom dimethylsulfide emissions, "OH" is the variable for oxidation by hydroxyl, and "BrO" is the variable for oxidation by bromine monoxide. $\star$ signifies a regression variable that is significant at a p-value of $p \leq 0.05$. 24

2.3 Akaike Information Criterion For a Series of Nested MSA Regression Models, for May and August. Variables in this table are described fully in text. $\star$ signifies the statistically optimal model for each month 25

A.1 Date correspondence for seasonal categories from Zhang et al., 2001 [143]. 81

A.2 Land Use Categories: Translating MCD12C1 to Zhang et al., 2001 [143]. . 82

List of Figures

2.1 Schematic diagram of the data analysis framework, and the relation to other Lagrangian Particle Dispersion Model (LPDM) source attribution frameworks with respect to the level of prior scientific understanding of emissions and transformation processes. 11

2.2 Weekly climatology of methanesulfonic acid at Alert ($ng\,m^{-3}$), Nunavut, Canada during the years 1980–2009, from measurements by Sharma et al., 2012 [107]. 14

2.3 Residual (observed - predicted ($ng\,m^{-3}$)) plot for methanesulfonic acid regression model. 26

2.4 Annual climatology of residuals (observed - predicted ($ng\,m^{-3}$)) for the methanesulfonic acid regression model during the years 1981–2006. 27

2.5 Model predicted and measured methanesulfonic acid ($ng\,m^{-3}$) at Alert for 2006. 27

2.6 Climatology of methansulfonic acid ($ng\,m^{-3}$) at Alert, Nunavut, Canada as predicted by regression model during the years 1981–2006. 28

2.7 Stacked bar plot climatology of methanesulfonic acid ($ng\,m^{-3}$) at Alert, Nunavut, Canada during 1981–2006, as predicted by regression model, showing partitioning between contributions from hydroxyl (OH) and bromine monoxide (BrO) Oxidation. 29

2.8 Stacked bar plot climatology of ($ng\,m^{-3}$) at Alert, Nunavut, Canada during 1981–2006, as predicted by regression model, showing partitioning between contributions from dimethylsulfide emissions from ice-edge blooms and the ice-free ocean. 29

2.9 Mean spatial contributions of methanesulfonic acid ($ng\,m^{-3}$) to Alert, Nunavut, Canada in May and August 1981–2006, as predicted by regression model. . 31

3.1 Total black carbon emissions in the northern Chinese domain shown in Figure 3.2, as predicted by the bottom-up inventory of Zhang et al., 2009 [145]. 42

3.2 Spatial distribution of mean bottom-up inventory estimate of black carbon emissions in northern China During the period 1 January 2006 through 31 December 2008, with a resolution of 0.1° × 0.1°. Magnitude of emissions and spatial distribution to a 0.5° × 0.5° scale were taken from Zhang et al., 2009 [145]. Distribution at scales less than 0.5° × 0.5° were inferred from Lei et al., 2011 [60]. 43

3.3 Reduced Pearson χ^2 statistic comparing a posteriori STILT model predictions of elemental carbon concentrations at Tsinghua University (40.00°N, 116.33°E) to concentrations measured on weekly integrated $PM_{2.5}$ filter samples with the EnCan-total-900 thermal method, as a function of a priori uncertainty decorrelation length scale. The ideal χ^2 statistic is equal to 1. The decorrelation length scale where the χ^2 statistic is closest to 1 (50 km) is assumed correct. 44

3.4 Measurements of elemental carbon aerosol concentration from weekly integrated $PM_{2.5}$ filter samples collected at Tsinghua University during the period 1 January 2006 through 31 December 2008. Concentrations were measured by the Stable Isotope Research Lab at Environment Canada with the EnCan-total-900 thermal method. 45

3.5 Qualitative diagram of elemental carbon aerosol transport model implemented with the STILT Lagrangian particle dispersion model. The size of black dots represent a relative magnitude of elemental carbon concentration. + and − indicate the addition or subtraction of elemental concentration at each model step, and multiple + or − signs indicate a change in concentration that is typically large relative to other changes in the same model step. 48

3.6 Mean footprint for a receptor located at Tsinhghua University during the period 1 January 2006 through 31 December 2008. Calculated with STILT using meteorological fields from GDAS1. 50

3.7 Locations of RAOB stations that provided data for the analysis of transport uncertainty in a STILT model of elemental carbon concentration at Tsinghua University (40.00°N, 116.33°E). 53

3.8 A comparison between winds predicted by STILT (an interpolation GDAS1 reanalysis), and observations from RAOB for all stations shown in Figure 3.7, during the period 1 January 2006 through 31 December 2008. All available data below a height of 15000 m above ground level was used. Regression slopes were calculated with an ordinary least squares method. a) Zonal wind, b) Meridional wind. 53

3.9 A comparison between Planetary Boundary Layer (PBL) heights predicted by STILT, and by a bulk Richardson number analysis of RAOB radiosonde profiles for all stations shown in Figure 3.7, during the period 1 January 2006 through 31 December 2008. The RAOB PBL top corresponds with a critical bulk Richardson number of 0.25. STILT-predicted PBL heights are discretized to model vertical levels. The regression was calculated with a standard major axis method. 55

3.10 Plots demonstrating the effect of using a finite number of particles to simulate transport in a STILT model of elemental carbon aerosol concentrations at Tsinghua University (40.00°N, 116.33°E). a) Root-mean-square error for models computed with a varying number of particles compared against a model computed with 3000 particles. b) Mean bias for model computed with a varying number of particles compared against a model computed with 3000 particles. 57

3.11 Time series of standard deviation of error in a STILT model of elemental carbon concentration at Tsinghua University (40.00°N,116.33°E). Individual sources of error variance were calculated with a variety of numerical experiments. 58

3.12 STILT model estimate of the mean influence of elemental carbon emissions to Tsinghua University for the period 1 January 2006 through 31 December 2008, using bottom-up inventory (a priori) estimates of emissions. 59

3.13 Concentrations of elemental carbon aerosol at Tsinghua University (40.00°N, 116.33°E) as predicted by a STILT model using bottom-up inventory (a priori) estimates of emissions (red), and as measured on weekly integrated $PM_{2.5}$ filter samples with the EnCan-total-900 thermal method (blue). . . 60

3.14 A comparison between STILT model predictions of elemental carbon concentrations at Tsinghua University (40.00°N, 116.33°E) made using bottom-up inventory (a priori) estimates of emissions to concentrations measured on weekly integrated $PM_{2.5}$ filter samples with the EnCan-total-900 thermal method. A standard major axis regression fit is included. 60

3.15 The percentage of background concentrations in STILT model predicted elemental carbon at Tsinghua University (40.00°N, 116.33°E). Model predictions were made using bottom-up inventory (a priori) estimates of emissions. 61

3.16 Violin plot of a posteriori scaling factors for elemental carbon emissions in the northern Chinese domain shown in Figure 3.2, calculated using a Bayesian inversion analysis of a STILT model of the concentration at Tsinghua University (40.00°N, 116.33°E). The width of each element is representative of the number of emissions grid cells as a function of the scaling factor. . . 62

3.17 Mean spatial distribution of a posteriori scaling factors for elemental carbon emissions in nortnern China for the period 1 January 2006 through 31 December 2008, calculated using a Bayesian inversion analysis of a STILT model of the concentration at Tsinghua University (40.00°N, 116.33°E). . 63

3.18 Mean spatial distribution of a posteriori uncertainty of elemental carbon emissions in nortnern China for the period 1 January 2006 through 31 December 2008, calculated using a Bayesian inversion analysis of a STILT model of the concentration at Tsinghua University (40.00°N, 116.33°E). . 63

3.19 Concentrations of elemental carbon aerosol at Tsinghua University (40.00°N, 116.33°E) as predicted by a STILT model using a posteriori estimates of emissions calculated with a Bayesian inversion analysis (red), and as measured on weekly integrated $PM_{2.5}$ filter samples with the EnCan-total-900 thermal method (blue). 64

3.20 A comparison between STILT model predictions of elemental carbon concentrations at Tsinghua University (40.00°N, 116.33°E) made using a posteriori estimates of emissions calculated with a Bayesian inversion analysis to concentrations measured on weekly integrated $PM_{2.5}$ filter samples with the EnCan-total-900 thermal method. A standard major axis regression fit is included. 64

3.21 Violin plot of monthly climatology a posteriori scaling factors for elemental carbon emissions in the northern Chinese domain shown in Figure 3.2, calculated using a Bayesian inversion analysis of a STILT model of the concentration at Tsinghua University (40.00°N, 116.33°E). The width of each element is representative of the number of emissions grid cells as a function of the scaling factor. 65

3.22 Mean spatial distribution of monthly climatology a posteriori scaling factors for elemental carbon emissions in nortnern China for the period 1 January 2006 through 31 December 2008, calculated using a Bayesian inversion analysis of a STILT model of the concentration at Tsinghua University (40.00°N, 116.33°E). 66

3.23 Mean spatial distribution of monthly climatology a posteriori uncertainty of elemental carbon emissions in nortnern China for the period 1 January 2006 through 31 December 2008, calculated using a Bayesian inversion analysis of a STILT model of the concentration at Tsinghua University (40.00°N, 116.33°E). 66

3.24 Concentrations of elemental carbon aerosol at Tsinghua University (40.00°N, 116.33°E) as predicted by a STILT model using monthly climatology a posteriori estimates of emissions calculated with a Bayesian inversion analysis (red), and as measured on weekly integrated $PM_{2.5}$ filter samples with the EnCan-total-900 thermal method (blue). 67

3.25 A comparison between STILT model predictions of elemental carbon concentrations at Tsinghua University (40.00°N, 116.33°E) made using monthly climatology a posteriori estimates of emissions calculated with a Bayesian inversion analysis to concentrations measured on weekly integrated $PM_{2.5}$ filter samples with the EnCan-total-900 thermal method. A standard major axis regression fit is included. 67

Chapter 1

Introduction

Lagrangian particle dispersion modeling is a method of computing an ensemble of air parcel trajectories through a turbulent flow. This method is most commonly applied to problems relating to the transport of atmospheric trace constituents. In this thesis, we will explore and expand upon methodologies that apply the Stochastic Time-Inverted Lagrangian Transport (STILT) model, a Lagrangian Particle Dispersion Model (LPDM). This introductory chapter introduces the LPDM in a historical perspective and prepares for the advancements in application that are presented in this thesis.

1.1 A Brief History of Lagrangian Particle Dispersion Models

This section contains a brief description of the history of Lagrangian models of the atmosphere, following the works of Thomson and Wilson, 2012 [122], and Rodean, 1996 [95].

There are two frames of reference in which fluids can been modeled; the Eulerian and Lagrangian frames. In the Eulerian frame, properties of a fluid are specified as a function of time, t, and a position, (x, y, z), relative to an observer fixed to a point on the Earth's surface. In the Lagrangian frame the observer "rides" the fluid element as it moves along its trajectory. Co-ordinates in the Lagrangian frame are time, t, relative to an initial time, t_0, and the position of the fluid element, (x, y, z), relative to its position at t_0.

The first significant work leading to the development of Lagrangian models of the atmosphere was Langevin's 1908 [58] equation describing Brownian motion of a single particle as a stochastic differential equation in the Lagrangian frame of reference. This is the so-called "Langevin Equation":

$$\frac{d\mathbf{v}}{dt} = -a\mathbf{v} + b\boldsymbol{\eta}\left(t\right). \tag{1.1}$$

Where $\mathbf{v}$ is the velocity of the particle, a is a coefficient of damping due to viscous drag, and the product of b and $\boldsymbol{\eta}\left(t\right)$ form a random acceleration with a Gaussian probability distribution, due to molecular collisions.

The Eulerian equivalent to the Langevin equation (Eqn. 1.1) is the "Fokker-Planck equation", which predicts the evolution of the probability density of the locations of fluid elements, $P \equiv P\left(\mathbf{x}, \mathbf{u}, t\right)$ [122]:

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial x_i}\left(u_i P\right) - \frac{\partial}{\partial u_i}\left(a_i P\right) + \frac{1}{2}\frac{\partial^2}{\partial u_j \partial u_j}\left(b^2 P\right). \tag{1.2}$$

Where $\mathbf{x}$ is the velocity of a fluid element, $\mathbf{u}$ is the velocity of a fluid element, a_i is a (possibly anisotropic) coefficient of damping, b is an empirical constant describing the rate of diffusion, and i is an index for direction (1, 2, or 3).

The direct application of the Lagrangian frame of reference to problems of atmospheric transport began with the work of Taylor in 1921 [119]. Taylor derived analytical equations describing the dispersion of a plume of pollutant through the atmosphere under stationary, homogeneous turbulence. Taylor's solution describes the evolution of a plume as:

$$\frac{d\sigma_z^2}{dt} = 2\int_0^t \overline{w\left(t'\right)w\left(t'+\xi\right)}d\xi. \tag{1.3}$$

Where σ_z^2 is the variance of vertical displacement, w is the vertical component of air parcel velocity, t' is an integration dummy variable for time, and ξ is a time scale over which the autocorrelation of velocity of an air parcel is quantified.

The argument of the integral on the right side of Equation 1.3 contains an autocorrelation in time, and so the equation can be recast as:

$$\frac{d\sigma_z^2}{dt} = 2\sigma_w^2\int_0^t R_{ww}\left(\xi\right)d\xi. \tag{1.4}$$

Where σ_w^2 is the variance of the vertical component of air parcel velocity, and R_{ww} is the "Lagrangian autocorrelation function".

In the limit of short time of travel from the source, Equation 1.4 yields root-mean-square displacements that are proportional to the travel time. In the limit of long time, the root-mean-square displacements are proportional to the square root of travel time. These limits are of fundamental importance and provide a test for the behaviour of all contaminant transport models. The application of Taylor's work to dispersion in the complex environment of the real atmosphere was limited by the idealization to homogenous and stationary turbulence that was necessary to derive an analytical solution. It was not until the era of computer modeling that realistic problems with atmospheric conditions and turbulence varying in time and space could be addressed.

The first step towards creating a more practical analog to the work of Taylor came in 1959, when Obukhov [79] suggested that the Fokker-Planck equation could be applied to model evolution of fluid parcels undergoing turbulent diffusion. As work progressed, the focus shifted from the Fokker-Planck equation (Eqn. 1.2) to the Langevin equation (Eqn. 1.1). By the early 1970's numerical simulations of complex environments with inhomogeneous turbulence had begun to take place.

Early numerical simulations based on the Langevin equation (Eqn. 1.1) resulted in particles that drifted toward the direction of negative gradients in turbulent kinetic energy. This led to particles collecting around the local minima of turbulence, and models predicting an increase of entropy in violation of the second law of thermodynamics. To correct this, Wilson, 1981 [137, 138, 139] proposed adding an ad hoc corrective acceleration (the "drift correction") to the Langevin equation (Eqn. 1.1).

Thomson, 1987 [121] evaluated general forms of the Langevin and Fokker-Planck equations (Eqns. 1.1 and 1.2) to establish design criteria for stochastic Lagrangian models of turbulent diffusion. This work showed that the well-mixed criterion — that an initially uniform concentration field remains uniform in the absence of sources and sinks — served as the superlative criteria for selection of a model, and that random forcing terms must have Gaussian probability distributions. After the publication of this paper the foundations of Lagrangian stochastic models were sufficiently established to produce well behaved and practically applicable models.

The first popularized Lagrangian model was the HYbrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model [23]. HYSPLIT has most often been used as a mean trajectory model, ignoring the effects of turbulence. It is also capable of being run as puff dispersion model, where the modeled air parcel grows in dimension in accordance with Taylor's solution (Eqn. 1.4), as an LPDM, or as a hybrid of the two. HYSPLIT remains the most popular Lagrangian model today, largely due to its well established community and its web interface, which is accessible to a broad range of users. HYSPLIT is used for emergency preparedness in the United States and other nations (e.g., Draxler et al., 1994 [24]), and has also been used extensively for identifying source regions of atmospheric trace constituents (e.g., Wang et al., 2011 [130])

Another popular LPDM is FLEXPART [113, 112]. FLEXPART was originally developed for emergency preparedness, but has been applied to a wide variety of applications. FLEXPART has the capability to be run as a Langevin equation (Eqn. 1.1) model, or (with lower computational cost for emergency situations) as a random walk dispersion model. It includes native schemes for deposition and radioactive decay, which have made it popular for the estimation of effects of radioactive releases (e.g., Arnold et al., 2010 [4]).

STILT was introduced by Lin et al., 2003 [68], and provides numerous innovations that allow for the development of advanced methods of constraining the strength of sources of atmospheric trace constituents. One innovation is the "footprint", a function that computes the sensitivity of tracer concentrations at a receptor to emissions in the planetary boundary layer above a source region. Also, the propagation of transport and boundary layer height errors has been incorporated in STILT to improve the treatment of model uncertainty [67, 32].

Since the development of these and the many other Lagrangian transport models the field has flourished, and the number of published works has grown dramatically [64]. A significant community of specialists in Lagrangian modeling has evolved. The community met in Grindelwald, Switzerland in October 2011 for the Chapman Conference on Lagrangian Modeling of the Atmosphere. This conference resulted in the publication of the monograph "Lagrangian Modeling of the Atmosphere" [66], a collection of reviews, application studies, and theoretical explorations of methodology that provide the state of the science today.

1.2 Thesis Outlook

This is a manuscript based thesis, comprised of two manuscripts. The first has been published in the American Geophysical Union's Geophysical Monograph Series book "Lagrangian Modeling of the Atmosphere". It has been reproduced here with the expressed written consent of the American Geophysical Union. The second has been prepared for submission to The Journal of Geophysical Research-Atmospheres.

Both of these works deal with the application of STILT to the problem of using surface based ambient air monitoring sites as top-down constraints on sources of atmospheric trace constituents. When approaching these types of problems we strive to produce predictive models, and solve for a best estimate of the emissions field by Bayesian inversion.

This is not possible for all of the atmospheric constituents that we may be interested in, as gaps in the understanding of chemical and physical processes involving the constituent can prevent a sufficient model from being constructed. Simpler methods are available that leverage variations in a time series of observations to predict potential source regions. These methods take the lack of information to an unnecessary extreme, assuming that there is no scientific understanding of sources and transformations. In the first manuscript of this thesis, work to improve this situation is presented. A method was derived that writes a full model for the constituent, replaces poorly understood terms with a variable, then solves for this variable with a linear regression. The resulting model can be used to predict source regions, but can also be used to test hypotheses about the regression variables. In many cases, there are alternate chemical pathways for transformations, and physically different sources of the constituent. The method was designed to test the significance of these alternate processes, and quantify their contributions to the observed concentration.

The second manuscript presents a problem where a model amenable to Bayesian inversion can be developed. Results from this study demonstrate the strength of Bayesian inversion by providing evidence of an underprediction bias in a bottom-up emissions inventory, and providing constrained emissions with a reduction of uncertainty.

Chapter 2

Quantitative attribution of processes affecting atmospheric chemical concentrations by combining a time-reversed Lagrangian particle dispersion model and a regression approach

Reproduced by permission of American Geophysical Union.

Benmergui, J., Sharma, S., Wen, D. and Lin, J. C., Quantitative Attribution of Processes Affecting Atmospheric Chemical Concentrations by Combining a Time-Reversed Lagrangian Particle Dispersion Model and a Regression Approach, in Lagrangian Modeling of the Atmosphere, (eds J. Lin, D. Brunner, C. Gerbig, A. Stohl, A. Luhar and P. Webley), DOI: 10.1029/2012GM001254, 2013. Copyright [2013] American Geophysical Union.

This chapter is extracted from a paper [8] that was published in the American Geophysical Union's Geophysical Monograph Series book "Lagrangian Modeling of the Atmosphere", released 16 April 2013. Co-authors included Sangeeta Sharma, Deyong Wen, and John C. Lin. Sangeeta Sharma provided the motivation for the modeling experiment. Measurements and previous modeling work led her to questions about the cycle of methanesulfonic acid in the Arctic. Conversations between her and myself led to ideas about how the STILT model could be applied to answer these questions. She provided input at all stages of work to ensure a relevant methodology and adequate treatment of chemistry. She wrote the third paragraph in section 2.2, detailing the measurement of methanesulfonic acid at Alert. Deyong Wen contributed R and BASH scripts that facilitated the treatment of deposition in the model. He also helped in the acquisition of MOZART model output that was used for fields of hydroxyl concentration. John C. Lin contributed greatly to the development of the methodology, provided leadership, and guided the writing of the manuscript. All authors contributed to the editorial process of the paper.

Overview

A framework is presented for analyzing time series of atmospheric constituent measurements with respect to hypothesized processes by combining a Lagrangian particle dispersion model and a linear regression. This study illustrates the method by using the example of methanesulfonic acid (MSA) measured at Alert, Canada (82.45°N, 62.52°W). The model is used to compare the seasonal dependence of MSA concentrations at Alert on (1) the chemical source through oxidation of dimethylsulfide (DMS) by bromine monoxide (BrO) versus hydroxyl (OH) and (2) emissions of DMS from ice-edge phytoplankton blooms versus ice-free ocean. A linear regression model is fitted to the measurements, relating them to sources, transformations, and deposition in the upwind source region, as predicted by trajectories from a backward in time Lagrangian particle dispersion model. The model includes explicitly modeled processes as well as terms to be solved through linear regression, with each term comprised of a variable and the associated regression parameter. Processes that are modeled as linear regression terms are ones that can be carried out by multiple mechanisms — e.g., transformation from a precursor by different chemical mechanisms. Regression parameters for various mechanisms for the same process are compared to determine which are most significant. Predicted concentrations from the regression model are partitioned into contributions from mechanisms being contrasted for a given process. The model is used to determine source regions of the target constituent. Model predictions of MSA at Alert were calculated between January 1981 and December 2006. The model showed a net under-prediction, but clearly indicated the source regions during the spring to be primarily ice-free ocean with a significant contribution from ice edge blooms, and during the summer to be dominated by the ice-free ocean. The model also indicated that oxidation of DMS by BrO was the dominant source of MSA in the spring, while DMS oxidation by OH was the dominant source in the summer.

2.1 Introduction

Lagrangian Particle Dispersion Models (LPDM's) have been used to analyze concentration time series of atmospheric constituents since the 1980's (e.g., [139], [102], [124]). Reviews of the mathematical and practical basis of LPDM's can be found in [95], [136], [65], as well as many papers in this Monograph [38, 64, 122, 135].

Several methodologies have been developed that use LPDM's to analyze trajectories from a receptor site. These methods can be contrasted by the required level of prior scientific understanding about sources and atmospheric transformations (i.e., reactions with other constituents, deposition) of the constituent. If sources are well constrained and transformations are well understood, then a predictive model of the concentration of the constituent can be formulated (e.g., [132]). If a prior estimate of source strength can be constructed, and transformations and atmospheric transport are well understood, then Bayesian inversion modeling approaches [29, 94, 33, 52] use the concentration of a constituent at a point to provide top down constraints on the estimates of source strength. However, if knowledge of sources and transformations is insufficient for use in a Bayesian inversion, then the "potential source contribution function" [6] can be calculated to elucidate regions in the source field that are likely contributing to variations in measured concentrations. There exists a gap in this spectrum between the Bayesian inversion and potential source contribution function. If there are hypotheses about which sources are important contributors to a time series of an atmospheric constituent, but insufficient knowledge of these sources is available to provide an a priori estimate of fluxes for use in Bayesian inversion, then a framework to test the hypotheses would be useful. This paper develops such a modeling framework (Figure 2.1) by combining Lagrangian atmospheric modeling and a traditional method of hypothesis testing used in the statistics of linear models.

This method is intended to deduce key processes contributing to the observed concentration variations of a chemical species. Many earlier studies have compared measurements of one compound to another to make inferences about chemistry (examples include [103, 89], and [31]). These methods can provide important insights, but they neglect the fact that concentrations of an atmospheric constituent are mostly dependent on the history of the air mass and the processes during that history, prior to arrival at the receptor. Laboratory experiments have also been used to make inferences about atmospheric processes (cf., [20], [7], and [22]). These experiments can provide valuable information, such as rate constants for reactions; however, they do not account for the complexity of the real atmosphere. In our work, the concentrations of a constituent at a measurement site are correlated with the history of atmospheric conditions in the airmass through the use of an LPDM.

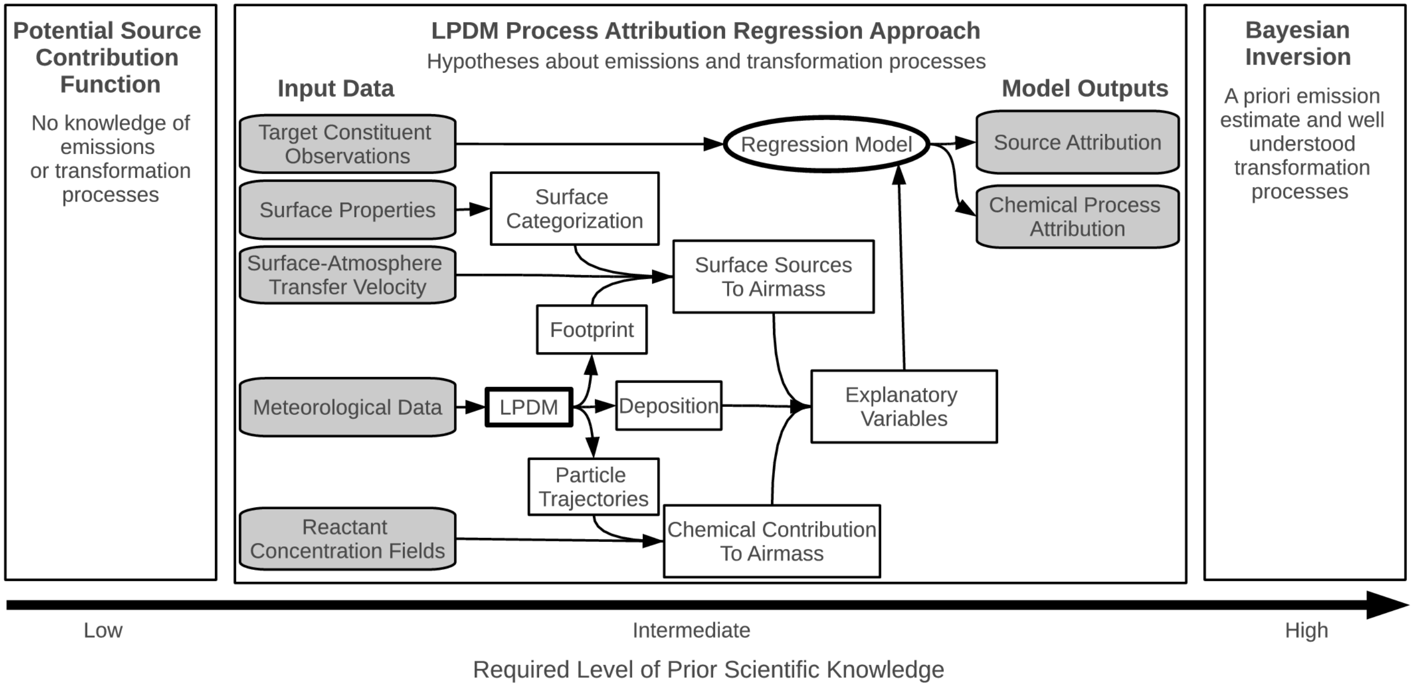

Potential Source Contribution Function
No knowledge of emissions or transformation processes
LPDM Process Attribution Regression Approach
Hypotheses about emissions and transformation processes
Input Data
Model Outputs
Target Constituent Observations
Surface Properties
Surface-Atmosphere Transfer Velocity
Meteorological Data
Reactant Concentration Fields
Surface Categorization
Footprint
LPDM
Deposition
Particle Trajectories
Surface Sources To Airmass
Chemical Contribution To Airmass
Explanatory Variables
Regression Model
Source Attribution
Chemical Process Attribution
Bayesian Inversion
A priori emission estimate and well understood transformation processes
Low
Intermediate
High
Required Level of Prior Scientific Knowledge

Figure 2.1: Schematic diagram of the data analysis framework, and the relation to other Lagrangian Particle Dispersion Model (LPDM) source attribution frameworks with respect to the level of prior scientific understanding of emissions and transformation processes.

In our regression method, a complete model for the concentration of a single atmospheric constituent at a point is constructed, and terms in the model where large uncertainties exist are replaced with linear regression terms. This allows measurements at the receptor site to be used within a regression to determine the strength and significance of association for the substituted terms. The regression formulation is useful when there are multiple mechanisms for individual processes affecting the constituent that, without constraints from the presence of sources and reactants in the upwind region, may lead to equal time series of concentrations. The different mechanisms can be compared to determine the most relevant in different atmospheric situations.

To test the significance of different mechanisms to the various processes and assess the validity of model results, the following analyses can be carried out: estimating the size of the regression coefficients, testing of statistical significance, computing confidence intervals, applying the Akaike Information Criterion (AIC) [1, 2, 3], and analyzing residuals. A statistically optimal model can be selected from nested subsets of the model with all parameters hypothesized to be important. Once the optimal model is achieved, the source regions can be estimated, and the concentrations can be partitioned into contributions from various mechanisms for the process of interest. This enables the investigator to determine the relative contributions from each mechanism to the concentration at the receptor site.

There have been previous studies that correlated atmospheric constituent time series with properties of airmass trajectories. Simpson et al., 2007 [108] correlated the concentrations of bromine monoxide (BrO) in Barrow, Alaska, USA with the amount of time that a single particle trajectory spent over surfaces of first year sea ice and "potential frost flower fields". Longhetto et al., 1997 [74] correlated carbon dioxide (CO_2) concentrations with average curvature, altitude, and potential temperature of clusters of homogeneous trajectories. In our paper, a methodology is developed that can be used to conduct these types of studies with a state of the art LPDM, that takes into account the atmospheric transport, which determines the sensitivity of atmospheric concentrations to surface emissions.

The method developed in this paper is applied to address the importance of ice-edge phytoplankton blooms to the concentration of methanesulfonic acid (CH_3SO_3H a.k.a. MSA) in the High Arctic and the relative importance of oxidation of dimethylsulfide ($(CH_3)_2S$ a.k.a. DMS) by hydroxyl (OH) and BrO. A regression model for MSA concentrations at Alert is calculated for the period between January 1981 and December 2006.

2.2 Application of Framework: Methanesulfonic Acid in the Arctic

MSA is a biogenic sulfur aerosol and stable oxidation product of DMS [41]. DMS is emitted from the ocean by phytoplankton metabolism and decay. The branching ratio of MSA from DMS oxidation is temperature dependent, but is on the order of 10% in the atmosphere [5]. Although phytoplanktons live in all surface ocean waters, there are blooms of very large populations along the edge of melting sea ice [98]. This is believed to be due to the influx of fresh water, nutrients, and exposure to solar radiation as the ice melts. Phytoplankton blooms along the edge of melting sea ice have been shown to be nearly ubiquitous (occurring in 77-89% of receding ice regions) by Perrette et al., 2011 [81].

Atmospheric MSA in the Arctic is studied as a proxy for DMS production at lower latitudes and its transport to the Arctic because it is the most abundant stable oxidation product of DMS that has no other sources [99]. Sulfate (${SO_4}^{2-}$) is the most abundant oxidation product, but other sources exist, and so sulfate concentration cannot be used as a proxy for DMS concentration. Significant scientific interest exists in studying DMS over the past few decades because its production was hypothesized to act as a negative climate feedback through increasing cloud albedo [15]. Warmer climate has been shown to increase DMS production, but the hypothesis of a significant negative feedback has recently been rejected by a wide ranging body of evidence [86]. MSA concentrations in ice cores continue to be used as a proxy for sea ice extent, atmospheric patterns and biological activity in the oceans (e.g., [109]). Variations of MSA concentrations in ice cores have been shown to be associated with sea surface temperature and sea ice extent [80], and the rate of glaciation [40, 100]. Recently, an increase in the MSA concentrations at three Arctic locations since 2000, in relation to declining sea ice extent in the Arctic, has been investigated by Sharma et al., 2012 [107]. This work expressed that further modeling exercises are required to examine MSA source region contributions and transport mechanisms to the Arctic.

Concentrations of MSA have been measured at Alert, Nunavut, Canada (82.45°N, 62.52°W) since 1980 by Li et al., 1993 [63], and Sharma et al., 2012 [107]. Weekly integrated samples of aerosol particles were collected on 20×25 cm^2 Whatman 41D filters with no size cut; integrated air volumes were approximately 16,000 m^3. The filters were cut into eight equal strips, and one strip was used for MSA analysis. Samples were analyzed for MSA by ion chromatography [63]. Details of the sampling inlet, handling procedures and chemical analyses can be found in Li et al., 1993 [63]. The detection limit for MSA measurements is 0.15 $ng\, m^{-3}$. The weekly climatology of MSA concentrations from 1980 to 2009 at Alert is shown in Figure 2.2.

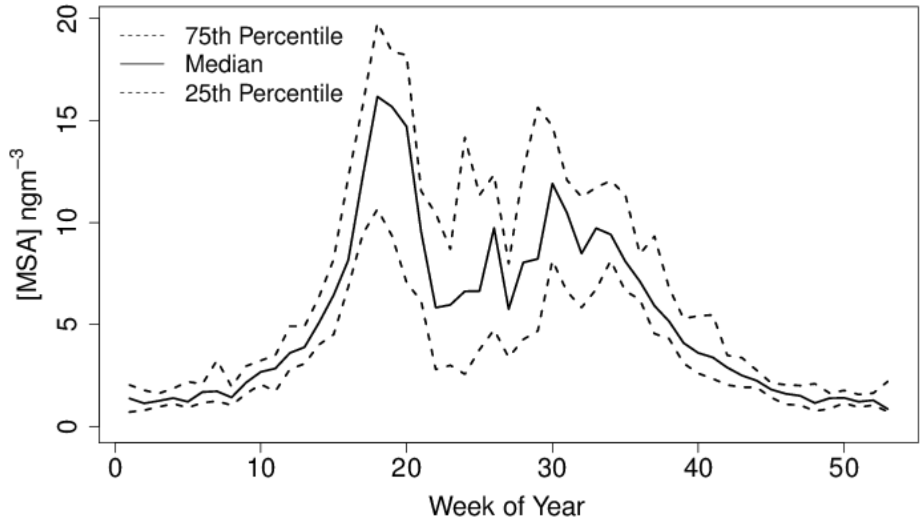


Figure 2.2: Weekly climatology of methanesulfonic acid at Alert ($ng\,m^{-3}$), Nunavut, Canada during the years 1980–2009, from measurements by Sharma et al., 2012 [107].

There are two clear peaks in the MSA concentration seasonal cycle. The peaks are hypothesized to occur due to variations of oxidants that convert DMS into MSA, and the increase in near-field sources of DMS, related to the retreat of sea ice in the summertime. The first peak is thought to be caused by an increase in BrO concentrations in the arctic springtime [78, 107] (the so called “bromine explosion” [133]). The second peak is believed to occur because sources of DMS are in closer proximity to Alert in the summer. Melting sea ice allows for phytoplankton growth in the waters that become ice-free, and bloom development at the ice-edge could contribute a significant amount of DMS to the atmosphere. The sea ice minimum usually occurs in September [13, 76].

2.3 Methodology

A general form of the framework used in this work is given in Equation 2.1. This is a model for the concentration of an atmospheric constituent, X, at a receptor site, from an LPDM trajectory with N particles.

$$X = \frac{1}{N} \sum_{n=1}^{N} \left[X_{init,n} + \Delta X_{flux,n} + \Delta X_{prec\,init,n} + \Delta X_{prec\,flux,n} \right]. \tag{2.1}$$

There are four contributions to the receptor concentration: $X_{init,n}$, $\Delta X_{flux,n}$, $\Delta X_{prec\,init,n}$ and $\Delta X_{prec\,flux,n}$. The concentration of X at the end of the trajectory (the "background") is advected to the receptor. The contribution from the background may be diminished due to deposition and chemical transformations. The resulting contribution is X_{init}. There may be surface fluxes of the constituent along the trajectory ΔX_{flux}. The portion of constituent contributed by these fluxes also may experience deposition and transformation along their journey to the receptor site. Furthermore, many atmospheric constituents are formed by secondary production where other "precursor" species react to form the constituent. It is necessary to include the initial concentrations and fluxes of these precursor species along the trajectory. The precursor species will also undergo transformations and deposition in a similar fashion as the target constituent, but these transformations produce the constituent, and thereby contribute to the receptor concentration. The initial and flux contributions of the precursor are $\Delta X_{prec\,init}$ and $\Delta X_{prec\,flux}$, respectively.

In order to implement this model, each of the four contributions to the concentration at the receptor site must be modeled. An LPDM is used to calculate the path of air parcels arriving at the receptor. In this study the Stochastic Time-Inverted Lagrangian Transport (STILT) model [68] was used to generate 10 day back trajectories with 1000 particles, driven by National Centers For Environmental Prediction / The National Center for Atmospheric Research (NCEP/NCAR) Reanalysis meteorological data [50]. Required inputs to the model are: fluxes of precursors and the target constituent, concentrations of reactants involved in transformation processes, initial concentrations of the modeled constituent, and deposition of all species.

2.4 Fluxes

The flux of a constituent into an LPDM air parcel is dependent on the emission rate at a given upwind location and time, and the sensitivity of the concentrations at the receptor site to these upwind emissions. Emission rates are often an external input in the model. They may be supplied by an inventory or calculated using a parameterization. Some LPDM's, including the STILT model adopted in this study, can calculate the footprint – i.e., the sensitivity of the measurements to fluxes at the surface of the Earth, with units of $ppm/\left(\mu mol\, m^{-2}\, s^{-1}\right)$ [68]. The enhancement of a constituent's concentration in the trajectory (in units of ppm) is then the product of the footprint and the emission rate (in units of $\mu mol\, m^{-2}\, s^{-1}$).

The emission rate provides an opportunity to introduce a regression variable. It might be hypothesized that the emission of the constituent or a precursor is dependent on some characteristic of the Earth's surface. A map of this surface characteristic can then be categorized, and the amount of contact between the trajectory and the categorizations can be used as a variable in the linear regression model. In many cases the emission is dependent on a transfer velocity for surface-atmosphere interactions. The transfer velocity can be calculated by parameterizations such as the "piston velocity" formulation [70, 131] over the ocean and models that include plant stomatal conductance such as the Simple Biosphere Model [106] over land.

The concentration increment added to the trajectory during a time step i is written $\alpha_s W_i F_{i,s}$. $F_{i,s}$ is the proportion of footprint falling in the s^{th} land surface categorization at time step i, in units of $ppm/\left(\mu mol\, m^{-2}\, s^{-1}\right)$. W_i is the value of the transfer velocity at time step i, in units of $m\, s^{-1}$. The regression parameter for the surface categorization s is α_s, which represents an effective concentration for the emission, in units of $\mu mol\, m^{-3}$.

The only atmospheric source of MSA is oxidation of DMS. There are no direct emissions of MSA from the Earth's surface. Therefore, only emissions of the precursor DMS concentration need to be calculated. The emission of DMS from the Earth's surface can be categorized into three categories: (1) ice-edge blooms, (2) the ice-free ocean, and (3) non-emitting surface (i.e., land, sea ice). This is a great simplification as it removes the heterogeneity of the emission strength of the open ocean and among various blooms, but allows for tests of significance of ice-edge blooms against the ice-free ocean.

In order to determine the regions of ice-free ocean and ice-edge bloom, sea ice concentrations derived from the Nimbus-7 Scanning Multichannel Microwave Radiometer (SMMR), the Defense Meteorological Satellite Program (DMSP) -F8, -F11 and -F13 Special Sensor Microwave/Imagers (SSM/Is) [13, 76] were used. This product is a daily percentage sea ice cover in the Northern Hemisphere, on a 25x25km spatial resolution. In this work, "ice-free ocean" is defined to be ocean that contains less than 10% surface coverage by sea ice. Blooms are defined by following the example of Perrette et al., 2011 [81]. An ice-edge bloom grid cell is defined as one in which there was greater than 10% ice some time within the last 10 days and that the ice has been present for at least 10 days, and where there is less than 10% ice for all of the next 10 days. This ensures that any grid cell counted as a bloom is indeed recently melted ice that had survived long enough to produce the conditions necessary for a bloom. The definition used was qualitatively found to prevent the advection of ice through a grid cell from being misrepresented as the retreating ice-edge.

The emissions in this model are oceanic, and are thus scaled by an ocean-atmosphere exchange parameterization. The piston velocity formulation of Liss and Merlivat, 1986 [70] is adopted, with sea surface temperatures from National Oceanic and Atmospheric Administration (NOAA) Optimum Interpolation [91] and surface winds from NCEP/NCAR Reanalysis [50]. This parameterization relates the rate of emission of any constituent to the wind speed, sea surface temperature and the concentration of the constituent in the surface ocean relative to the atmosphere. The flux of DMS into the atmosphere for each of the two categories is then calculated as the product of the regression parameter (α_{bloom} or $\alpha_{ice\,free}$), the STILT footprint ($F_{i,bloom}$ or $F_{i,bloom}$), and the piston velocity (W_i).

2.5 Deposition

All contributions to the concentration at the receptor site must be decreased by the fraction of the initial contribution that undergoes deposition. The rate of deposition for a given species can often be calculated by the LPDM. This feature is included in the HYbrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model [23]. HYSPLIT can be used to calculate an exponential deposition rate υ. The amount of a constituent X remaining from an initial value X_0 after a time Δt is $X = X_0 e^{-\upsilon \Delta t}$. For any contribution to the concentration at the receptor site, the deposition must be calculated as a fraction of the initial concentration surviving transport from the time of emission to the time that the trajectory meets the receptor site. When deposition is included, the concentration of a constituent contributed by fluxes from surface s at time step I, with previous time steps i of length Δt_i is $\alpha_s W_i F_{i,s} e^{\sum_{i=1}^{I} -\upsilon_i \Delta t_i}$.

Since STILT was based originally on the HYSPLIT model, it contains many core features of HYSPLIT. For the MSA example, the deposition rate was calculated with a HYSPLIT subroutine within STILT. MSA exists in both marine and continental air masses with particulate diameters ranging between 0.1 and $10\mu m$ [87]. Wyslouzil et al., 1991 [140] showed that the peak in the particle number diameter distribution varied with nucleation rate, but was always on the order of 1 μm. A single number is required by the HYSPLIT subroutine of STILT, so a value of $1\mu m$ was assumed. The dry deposition velocity V_{dry} can be estimated as a function of the mass mean diameter D as $V_{dry} = 0.388(D)^{0.76}$ [77]. Kerminen et al., 1997 [54] measured the mass median diameter of MSA in the Arctic atmosphere to be $0.28\mu m$, giving a deposition velocity of $1.47 \times 10^{-4} m\, s^{-1}$. The deposition scheme resulted in an atmospheric lifetime of MSA that ranged from 12 hours in the winter to 48 hours in the summer, with a mean of 24 hours. This calculated lifetime is too short when compared against the globally averaged lifetime of 7 days for MSA as predicted by Chin et al., 2000 [16], and so a scale correction was implemented to extend the mean lifetime of MSA to 7 days.

2.6 Reactions and Initial Concentrations

Chemical transformations provide a second opportunity to introduce a regression parameter. This regression requires the rate of any reactions in the form of a second order rate constant. Let C_p be the concentration of the p^{th} of P precursors (i.e., DMS in the MSA example), $C_{m,i}$ be the concentration of the m^{th} out of M reactant species that transforms the precursor into the target constituent (i.e., OH or BrO in the MSA example) at time step i. Let β_m be the rate constant for the reaction, multiplied by the branching ratio to give a constant for the production rate of X from C_p. Then,

$$\frac{dX_{prec}}{dt} = \sum_{m=1}^{M} \beta_m C_p C_{m,i}. \tag{2.2}$$

Equation 2.2, along with the initial condition $X_{prec}\,(t=0) = 0$ gives the concentration of X contributed from the precursor. For fluxes of the precursor at the time step i, this differential equation has the solution:

$$X_{prec} = C_p \sum_{i=1}^{I} \sum_{m=1}^{M} \beta_m C_{m,i} \Delta t_i. \tag{2.3}$$

Similarly, the remaining concentration of the target constituent, X_{rem}, from the portion X_I released at time step I, after reactions with the mth of M reacting compounds C_m is:

$$X_{rem} = X_I \left(1 - \sum_{i=1}^{I} \sum_{m=1}^{M} -\beta_m C_{m,i} \Delta t_i\right). \tag{2.4}$$

Chemical removals of the target constituent created from the precursor are not taken into account in this model, since the target constituent is assumed to be stable.

The concentrations of the target constituent and any precursors may be added at the end of the trajectories (the initial concentrations). A three-dimensional global chemical transport model such as the Goddard Earth Observing System Chemical Model (GEOS-Chem) [9] can be used to generate the required concentration fields. If the atmospheric lifetime of the constituent with respect to chemical removal, wet and dry deposition less than the duration of the trajectory, then the background concentration can be neglected.

For the MSA example, oxidation of DMS by BrO and OH that partially results in MSA was modeled. OH concentrations were taken from the Model for OZone And Related chemical Tracers (MOZART) [44], and BrO concentrations were taken from the parallel Toulouse Off-line Model of Chemistry And Transport (p-TOMCAT) [59, 141]. Both of these products provide global time-varying three dimensional fields of concentration.

Table 2.1: Explanations of Symbols Used in Equations.

Symbol	Explanation
X	Concentration of the constituent.
X_{init}	Contribution to X from initial concentration of X.
ΔX_{flux}	Contribution to X from fluxes of the constituent.
ΔX_{prec}	Contribution to X from initial concentration of precursors.
$\Delta X_{prec\,flux}$	Contribution to X from fluxes of precursors.
$\sum_{n=1}^{N}$	The sum across the N LPDM particles of index n.
$\sum_{i=1}^{I}$	The sum across I LPDM time steps of index i.
$\sum_{j=1}^{i}$	The sum across time steps j less than or equal to the time step i.
$\prod_{i=1}^{I}$	The product across I time steps of index i.
$\sum_{m=1}^{M}$	The sum across M reactant species of index m.
$\sum_{s=1}^{S}$	The sum across S surface categorizations of index s.
$\sum_{p=1}^{P}$	The sum across P precursor species of index p.
Δt_i	The length of the i^{th} time step.
υ_i	The total (wet + dry) exponential rate of deposition at time step i.
$C_{p,i}$	The concentration of the p^{th} precursor species at time step i.
$C_{m,i}$	The concentration of the m^{th} reactant species at time step i.
W_i	The scaling from the surface-atmosphere parameterization at time step i.
$F_{i,s}$	The footprint residing in surface categorization s at time step i.
α_s	The regression parameter for the effect of surface categorization s.
β_m	The regression parameter for the effect of the m^{th} reactant species.

2.7 Model Synthesis

Combining the results of Sections 2.3 to 2.6, a general model with the form of Equation 2.1 can be created. The model for each contribution is shown in Equations 2.5 to 2.8. Each contribution is written in the form (*precursor or constituent concentration*) × (*chemical reactions*) × (*deposition*). Table 2.1 provides a summary of the explanations for each symbol used in these equations.

$$X_{init} = X_{t_I}\left(\sum_{i=1}^{I}\sum_{m=1}^{M} -\beta_m C_{m,i}\Delta t_i\right)\left(\prod_{i=1}^{I} e^{-v_i \Delta t_i}\right). \tag{2.5}$$

$$\Delta X_{flux} = \sum_{i=1}^{I}\left[\left(\sum_{s=1}^{S}\alpha_s W_i F_{i,s}\right)\left(\sum_{j=1}^{i}\sum_{m=1}^{M} -\beta_m C_{m,j}\Delta t_j\right)\left(\prod_{j=1}^{i} e^{-v_j \Delta t_j}\right)\right]. \tag{2.6}$$

$$\Delta X_{prec\,init} = \sum_{p=1}^{P}\left[C_{p,I}\left(\sum_{j=1}^{i}\sum_{m=1}^{M} \beta_m C_{m,j}\Delta t_j\right)\left(\prod_{i=1}^{I} e^{-v_i \Delta t_i}\right)\right]. \tag{2.7}$$

$$\Delta X_{prec\,flux} = \sum_{i=1}^{I}\left[\left(\sum_{p=1}^{P}\sum_{s=1}^{S}\alpha_{s,p} W_i F_{i,s,p}\right)\left(\sum_{j=1}^{i}\sum_{m=1}^{M} \beta_m C_{m,j}\Delta t_j\right)\left(\prod_{j=1}^{i} e^{-v_j \Delta t_j}\right)\right]. \tag{2.8}$$

This model becomes linear with respect to the regression parameters when the variables are multiplied out and the appropriate substitutions are made. Anywhere a term is the product of two regression parameters α and β, a new regression parameter can be substituted and solved. If only one process affects X (i.e. chemical reactions or surface fluxes, but not both) then this is not necessary. When a new regression parameter is substituted in place of a product the α and β can no longer be solved independently.

When applying Equations 2.5 to 2.8 to the specific example of MSA, the following Equations (numbered 2.9 to 2.13; see below) are derived. The initial concentrations of MSA and DMS can be assumed to be zero, since the sum of the lifetimes are shorter than the trajectory duration (10 days), and the direct surface flux of MSA is zero for both bloom and ice-free ocean.

$$X_{init} = 0. \tag{2.9}$$

$$\Delta X_{prec\,init} = 0. \tag{2.10}$$

$$\Delta X_{flux} = 0. \tag{2.11}$$

DMS is emitted from blooms and the ice-free ocean and then oxidized to MSA by OH or BrO. The following equation gives the contribution from the precursor flux. The only precursor species is DMS, so for brevity the subscript p for precursor type is dropped.

$$\begin{aligned}\Delta X_{prec\,flux} =& \sum_{i=1}^{I} \Bigg[(\alpha_{bloom} W_i F_{i,bloom} + \alpha_{ice\,free} W_i F_{i,ice\,free}) \\ &\times \sum_{j=1}^{i} (\beta_{OH} C_{OH,j} \Delta t_j + \beta_{BrO} C_{BrO,j} \Delta t_j) \left(\prod_{j=1}^{i} e^{-v_j \Delta t_j} \right) \Bigg] .\end{aligned} \tag{2.12}$$

Writing this into the general model for MSA concentration:

$$\begin{aligned}MSA =& \frac{1}{1000} \sum_{n=1}^{1000} \sum_{i=1}^{I} \Bigg[(\alpha_{bloom} W_i F_{i,bloom,n} + \alpha_{icefree} W_i F_{i,icefree,n}) \\ &\times \sum_{j=1}^{i} (\beta_{OH} C_{OH,j,n} \Delta t_j + \beta_{BrO} C_{BrO,j,n} \Delta t_j) \left(\prod_{j=1}^{i} e^{-v_{j,n} \Delta t_j} \right) \Bigg] .\end{aligned} \tag{2.13}$$

The regression parameters that are solved in this model are $\alpha_{IceFree} \times \beta_{OH}$, $\alpha_{Bloom} \times \beta_{OH}$, $\alpha_{IceFree} \times \beta_{BrO}$ and $\alpha_{Bloom} \times \beta_{BrO}$. The individual α's and β's cannot be solved independently.

2.8 Interpretation of Regression Results

A synthesis of the methods used for analyzing linear regression models can be found in McCullagh et al., 1989 [75]. Relevant analyses are the estimation of coefficients, test of statistical significance, the computation of confidence intervals and residual analysis. When deciding which regression parameters are necessary, and which might be left out of the model, Akaike's Information Criterion (AIC) [1, 2, 3] can be employed. AIC is defined as $2k - 2ln(L)$, where k is the number of regression parameters and L is the maximized value of the likelihood function for the regression model. AIC analysis is conducted by removing one parameter from a model and comparing the AIC value. If removing the parameter from the model lowers the AIC, then the model is better parameterized by removing that regression term. A statistically optimum model has the lowest possible AIC. AIC is a relative measure among nested models; a single score is meaningless by itself.

Once the regression parameters are estimated, a predicted value for the constituent's concentration at the receptor site can be calculated using the regression model. These predicted concentrations should be compared to the measured concentration to assess whether the model captures the magnitudes and variability of the measured time series. A plot of residuals versus the fitted values of the model illustrates whether or not the model is well parameterized, and if the assumption that the model is linear is appropriate. The residuals should be normally distributed with a mean of 0 and standard deviation of 1, and should show no trend with the fitted values.

The regression parameters can be used to partition the concentration that arrives at the receptor site into contributions from different land surfaces or different chemical mechanisms. Each contribution to the predicted concentration can be calculated independently and then the results compared to each other. For instance, to partition the chemical mechanisms, set $\beta_m = 0$ for all but a single m, then calculate the predicted concentration. Repeat this for each m and the result will be a partitioning of concentration changes due to different m's. This analysis determines which processes are most important in controlling the constituent concentration at the receptor.

The estimated regression parameters can be used for the estimation of source regions. The footprint of the trajectory can be scaled by the amount of the constituent that is associated with that trajectory. The footprint can be separated into segments created during discrete time intervals within the duration of the trajectory. Each of these segments is scaled by the amount of deposition and the transformation processes that create and destroy the constituent. The scaled footprint represents a spatial plot of the contributions to the predicted concentrations.

2.9 Results of MSA Analysis

A time series of the model inputs was constructed, and then separated by the month of the year in order to account for seasonal variability in the productivity of the sources and temperature dependence of reactions. A linear regression was fitted and analyzed for each month using the R software package [21]. The model was calculated for the period January 1981 to December 2006.

Table 2.2 shows a summary of the linear regression coefficients for May and August, the months when there are peaks in the concentration observed at Alert (Fig. 2.2). Table 2.3 shows the AIC analysis for May and August. In May, all parameters except for BrO oxidized emissions from the open ocean were significant ($p \leq 0.05$). The AIC analysis showed that the statistically optimal model for predicting MSA concentration in May excluded only BrO oxidized emissions from the open ocean. In August, both parameters involving the ice-free ocean were significant, and both parameters involving blooms were insignificant ($p > 0.05$). The AIC analysis showed that the statistically optimal model for predicting MSA concentration in August included emissions from the ice-free ocean, oxidized by OH or BrO.

Table 2.2: Regression Coefficients and Wald Tests of Significance for methanesulfonic acid Regression Model for May and August. "Ice-Free" is the variable for ice-free ocean dimethylsulfide emissions, "Bloom" is the variable for phytoplankton bloom dimethylsulfide emissions, "OH" is the variable for oxidation by hydroxyl, and "BrO" is the variable for oxidation by bromine monoxide. $\star$ signifies a regression variable that is significant at a p-value of $p \leq 0.05$.

Factor	Estimate	Std. Error	t value	$Pr\,(Z(0,1) > \lvert t \rvert)$
May				
Ice-Free, OH	5.614×10^{9}	1.117×10^{9}	5.027	$2.11 \times 10^{-6}\star$
Bloom, OH	-1.245×10^{11}	5.232×10^{10}	-2.380	$0.0192\star$
Ice-Free, BrO	5.252×10^{-5}	1.628×10^{-3}	0.032	0.9743
Bloom, BrO	2.045	4.807×10^{-1}	4.255	$4.61 \times 10^{-5}\star$
August				
Ice-Free, OH	2.701×10^{9}	5.120×10^{8}	5.275	$7.38 \times 10^{7}\star$
Bloom, OH	7.115×10^{10}	4.459×10^{10}	1.596	0.1136
Ice-Free, BrO	2.630×10^{-3}	1.165×10^{-3}	2.258	$0.0260\star$
Bloom, BrO	-2.184×10^{-2}	1.150×10^{-1}	-0.190	0.8498

Table 2.3: Akaike Information Criterion For a Series of Nested MSA Regression Models, for May and August. Variables in this table are described fully in text. $\star$ signifies the statistically optimal model for each month

Model	AIC
May	
$\alpha_{IceFree} \times \beta_{OH}$, $\alpha_{Bloom} \times \beta_{OH}$, $\alpha_{IceFree} \times \beta_{BrO}$, $\alpha_{Bloom} \times \beta_{BrO}$	838
$\alpha_{IceFree} \times \beta_{OH}$, $\alpha_{IceFree} \times \beta_{BrO}$, $\alpha_{Bloom} \times \beta_{BrO}$	841.7
$\alpha_{IceFree} \times \beta_{OH}$, $\alpha_{Bloom} \times \beta_{OH}$, $\alpha_{IceFree} \times \beta_{BrO}$,	853.3
$\alpha_{IceFree} \times \beta_{OH}$, $\alpha_{Bloom} \times \beta_{OH}$, $\alpha_{Bloom} \times \beta_{BrO}$	836$\star$
$\alpha_{Bloom} \times \beta_{OH}$, $\alpha_{IceFree} \times \beta_{BrO}$, $\alpha_{Bloom} \times \beta_{BrO}$	859.4
$\alpha_{Bloom} \times \beta_{OH}$, $\alpha_{Bloom} \times \beta_{BrO}$	861.8
$\alpha_{IceFree} \times \beta_{OH}$, $\alpha_{IceFree} \times \beta_{BrO}$	851.4
$\alpha_{IceFree} \times \beta_{OH}$, $\alpha_{Bloom} \times \beta_{BrO}$	840.1
$\alpha_{IceFree} \times \beta_{BrO}$, $\alpha_{Bloom} \times \beta_{BrO}$	857.6
August	
$\alpha_{IceFree} \times \beta_{OH}$, $\alpha_{Bloom} \times \beta_{OH}$, $\alpha_{IceFree} \times \beta_{BrO}$, $\alpha_{Bloom} \times \beta_{BrO}$	762.2
$\alpha_{IceFree} \times \beta_{OH}$, $\alpha_{IceFree} \times \beta_{BrO}$, $\alpha_{Bloom} \times \beta_{BrO}$	762.8
$\alpha_{IceFree} \times \beta_{OH}$, $\alpha_{Bloom} \times \beta_{OH}$, $\alpha_{IceFree} \times \beta_{BrO}$,	760.3$\star$
$\alpha_{IceFree} \times \beta_{OH}$, $\alpha_{Bloom} \times \beta_{OH}$, $\alpha_{Bloom} \times \beta_{BrO}$	765.4
$\alpha_{Bloom} \times \beta_{OH}$, $\alpha_{IceFree} \times \beta_{BrO}$, $\alpha_{Bloom} \times \beta_{BrO}$	785.8
$\alpha_{Bloom} \times \beta_{OH}$, $\alpha_{Bloom} \times \beta_{BrO}$	789.2
$\alpha_{IceFree} \times \beta_{OH}$, $\alpha_{IceFree} \times \beta_{BrO}$	760.9
$\alpha_{IceFree} \times \beta_{OH}$, $\alpha_{Bloom} \times \beta_{BrO}$	765.7
$\alpha_{IceFree} \times \beta_{BrO}$, $\alpha_{Bloom} \times \beta_{BrO}$	796.8

Residuals (observed - predicted) of the regression model are plotted in Figure 2.3. The residuals of this model have a standard deviation of 6.17, and a mean of 3.31, suggesting that the model has an under-prediction bias. The residual plot has a slope of -0.44 and an intercept of 4.47 ($R^2 = 0.10$). A climatology of the residuals is plotted in Figure 2.4. The climatology of the residuals shows that the under-prediction occurs mostly at the peaks in MSA concentration. To help understand this result, the measured and model predicted MSA concentrations at Alert for the year 2006 were plotted (Figure 2.5). Peaks in the measured concentration do not always correspond in time to peaks in the predicted MSA concentration. We speculate that this is a result of transport error created by errors in the wind fields that drive the LPDM. The paucity of observations and presence of extreme conditions in the Arctic can produce erroneous values in the wind fields of reanalysis data

[27]. Other sources of error to the model that could have contributed to the lack of fit include: (1) uncertainty in the input concentration fields of reactant species, (2) failure of assumptions about the form of chemical processes, (3) error in the parameterization of atmosphere-sea exchange and uncertainty in the wind fields and sea surface temperatures used to calculate it, (4) uncertainty and aggregation error in the sea ice concentration fields, (5) error in the parameterization used to define blooms and failure of the assumption that blooms are ubiquitous at the receding ice-edge, (6) heterogeneity in the production of DMS in the open ocean and among blooms and (7) small amounts of MSA produced from continental water bodies and waters in the ice-pack where ice concentration is greater than 10% that were assumed negligible. Quantification of these errors is beyond the scope of this paper.

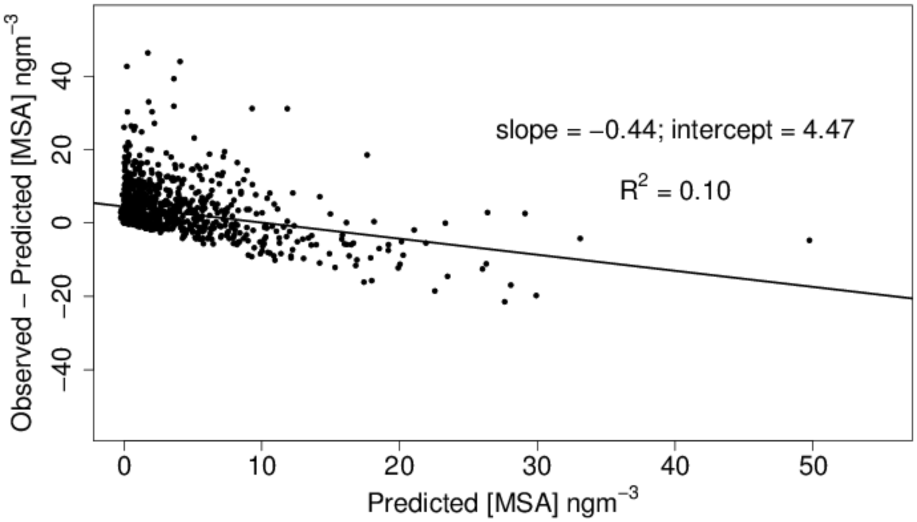


Figure 2.3: Residual (observed - predicted ($ng\,m^{-3}$)) plot for methanesulfonic acid regression model.

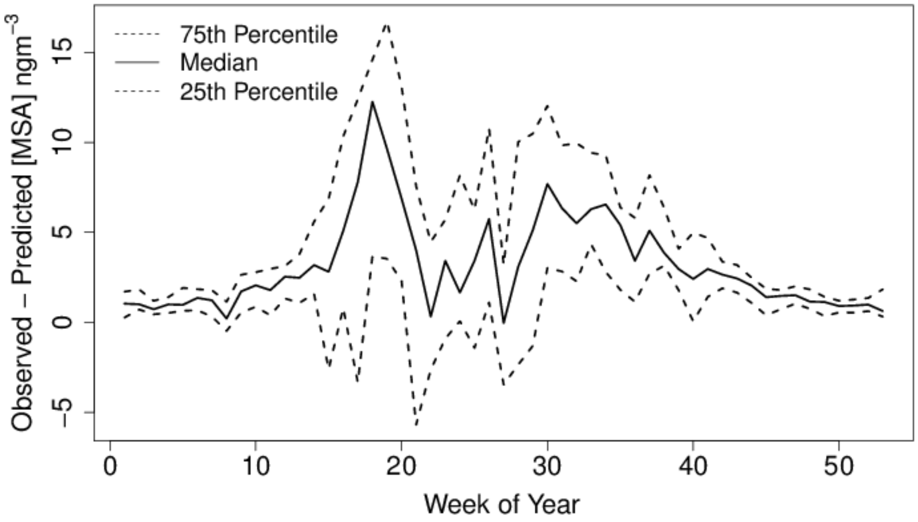


Figure 2.4: Annual climatology of residuals (observed - predicted ($ng\,m^{-3}$)) for the methanesulfonic acid regression model during the years 1981–2006.

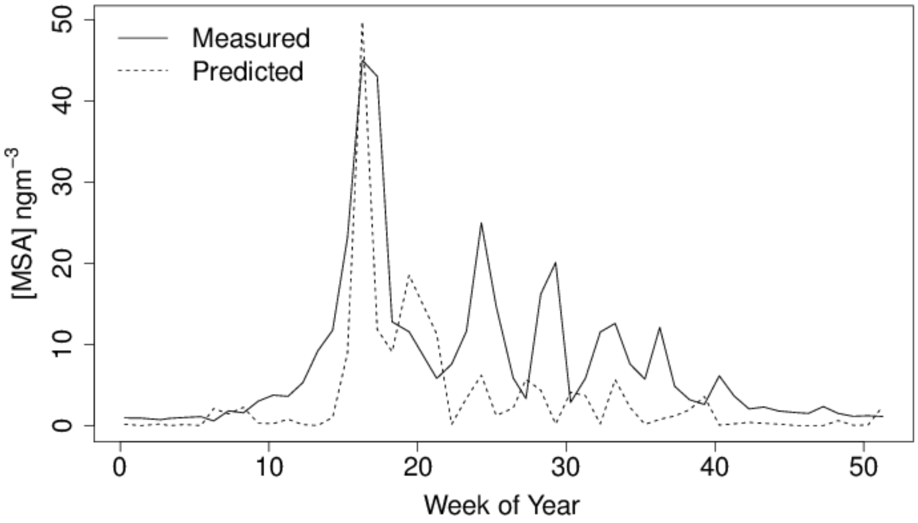


Figure 2.5: Model predicted and measured methanesulfonic acid ($ng\,m^{-3}$) at Alert for 2006.

The weekly climatology of MSA at Alert, as predicted by the model (Figure 2.6) shows that the model roughly replicates the double-peak pattern of the measured climatology. Since the model reproduces the seasonal pattern of the MSA concentration time series, inferences can be made about the relative strength of the sources.

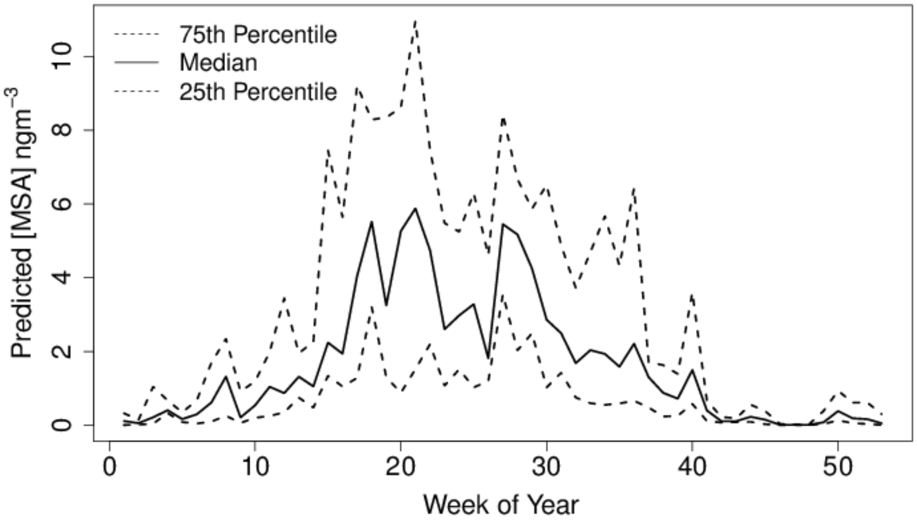


Figure 2.6: Climatology of methansulfonic acid ($ng\,m^{-3}$) at Alert, Nunavut, Canada as predicted by regression model during the years 1981–2006.

The predicted climatology of MSA was partitioned into mean contributions from ice-edge blooms versus ice-free ocean, and by oxidation by BrO versus OH. Figure 2.7 shows the partitioning of the predicted MSA into contribution from oxidation of DMS by BrO versus OH. The mean proportion of model predicted MSA that was created by oxidation by OH was 76%, while that created by oxidation by BrO was 24%. Oxidation by BrO dominated during the spring, with a maximum of 95% of the model predicted MSA being oxidized by BrO. This is evidence that the extra strength of the first peak might be caused by the extra oxidation by BrO, as hypothesized. Figure 2.8 shows the partition of the mean predicted concentration into ice-free ocean and ice-edge blooms. The mean proportion of model predicted MSA that was created from DMS emitted from the ice-free ocean was 94%, while that created by DMS emitted from blooms was only 6%. Emissions from blooms became more important during the spring, with a maximum of 37% of the model predicted MSA being created from DMS emitted from blooms.

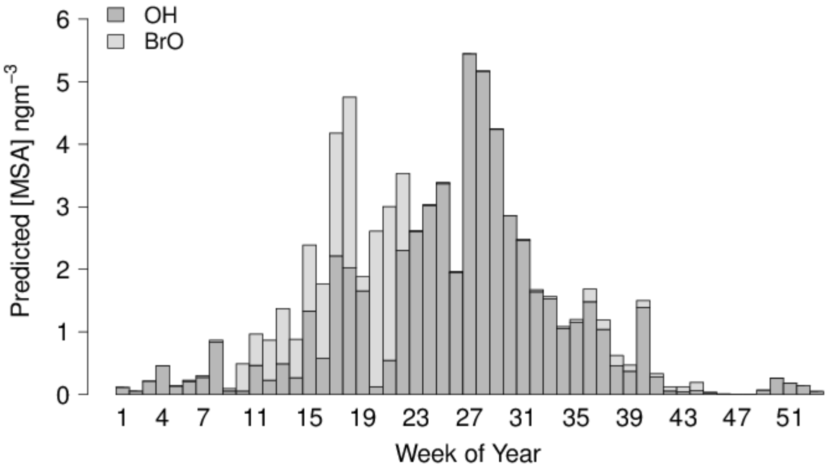


Figure 2.7: Stacked bar plot climatology of methanesulfonic acid ($ng\,m^{-3}$) at Alert, Nunavut, Canada during 1981–2006, as predicted by regression model, showing partitioning between contributions from hydroxyl (OH) and bromine monoxide (BrO) Oxidation.

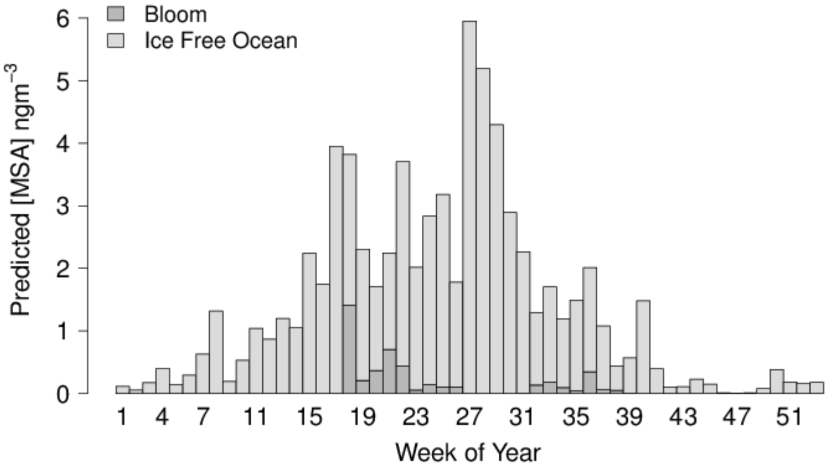
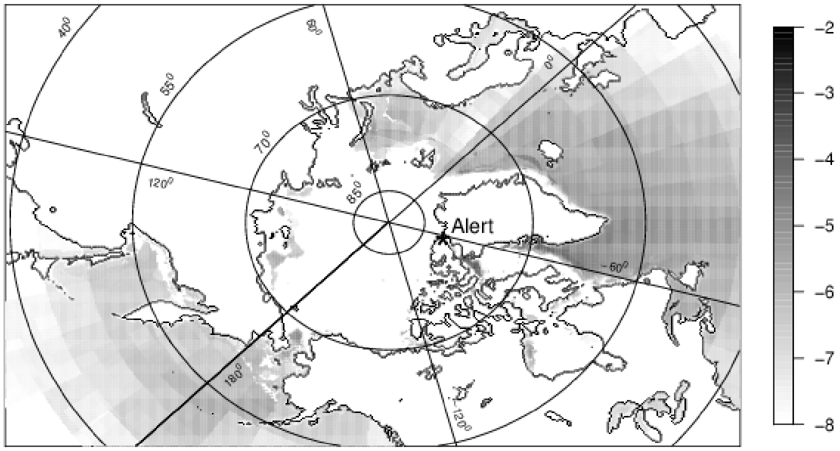


Figure 2.8: Stacked bar plot climatology of ($ng\,m^{-3}$) at Alert, Nunavut, Canada during 1981–2006, as predicted by regression model, showing partitioning between contributions from dimethylsulfide emissions from ice-edge blooms and the ice-free ocean.

Figure 2.9 shows the mean across all model years (1981–2006) of the distribution of regions contributing MSA to Alert for the months of May and August; the months where there are peaks in the concentration of MSA at Alert. The STILT footprint was separated into 12 hour time intervals, and the averages for each regression variable in the model are calculated for each time interval. The model was then reconstructed in each time interval to yield the spatial distribution of sources. The seasonal differences are clearly evident, as in August there are indeed sources that are much closer to Alert, especially in the Davis Strait (located between Greenland and the Canadian Archipelago), and in the many channels within the Canadian Archipelago. In May there are several regions at the edge of the sea ice extent that show strong sources. These are regions where the emissions are elevated by the presence of ice-edge phytoplankton blooms.

Mean Spatial MSA Contributions to Alert in May [$\log_{10}(ngm^{-3}/km^2)$]


40°
55°
60°
70°
85°
0°
120°
Alert
-60°
180°
-120°
-2
-3
-4
-5
-6
-7
-8

Mean Spatial MSA Contributions to Alert in August [$\log_{10}(ngm^{-3}/km^2)$]

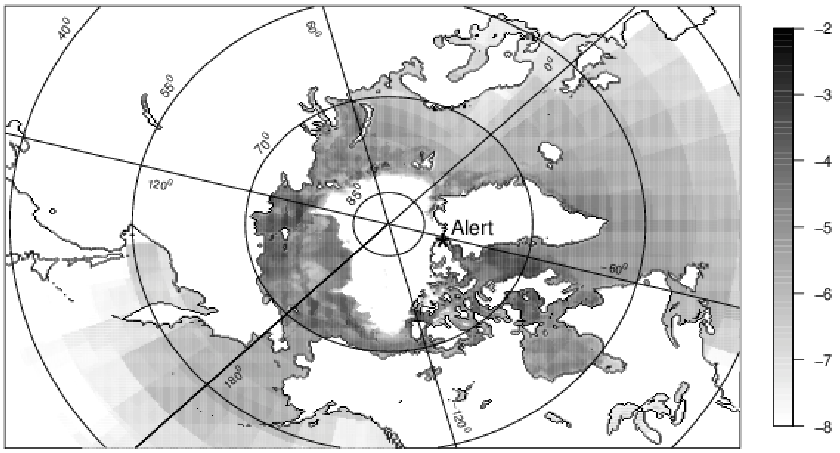

40°
55°
60°
70°
85°
0°
120°
Alert
-60°
180°
-120°
-2
-3
-4
-5
-6
-7
-8

Figure 2.9: Mean spatial contributions of methanesulfonic acid ($ng\,m^{-3}$) to Alert, Nunavut, Canada in May and August 1981–2006, as predicted by regression model.

2.10 Discussion

The use of a regression type model based on a time-reversed Lagrangian particle dispersion model has been demonstrated to have utility in making inferences about processes affecting atmospheric chemical concentrations at a measurement site (Fig. 2.1). The method was driven by two motivations: (1) to improve upon knowledge of the chemical transformations of a constituent when there is an intermediate level of scientific understanding to that required by the potential source contribution function method and Bayesian Inversion, and (2) to combine observations of the real atmosphere with knowledge of airmass history to determine the processes important to the concentration at the measurement site.

This method was applied to concentrations of MSA measured in the High Arctic at Alert, Nunavut, Canada. The linear regression model reproduced the seasonal cycle of MSA in Alert, but had a significant under-prediction bias. The reproduction of the seasonal cycle allowed inferences to be made about the relative strength of the sources. The relative importance of DMS emissions from ice-edge phytoplankton blooms and the ice-free ocean, as well as their oxidation by OH and BrO, to concentrations of MSA at Alert were compared. It was seen that the ice-free ocean was the dominant source of MSA, representing 94% of the model predicted MSA at Alert. Ice-edge blooms were important for the spring peak in MSA concentration, contributing up to 37% to the median climatology of MSA contributions for a single week. It was also seen that MSA was created primarily by the oxidation of DMS by OH, accounting for 76% of the model predicted MSA at Alert. During the peak in the spring BrO was the dominant oxidant, and was responsible for up to 95% of the MSA for a single week in the median contribution climatology. Spatial plots of MSA sources contributing to concentrations at Alert were created that illustrated the strong sources associated with blooms at the ice-edge in May, and a dramatic increase in near-field open ocean sources in August.

Acknowledgments

We gratefully acknowledge funding from Environment Canada for supporting J. Benmergui. This work was made possible by the facilities of the Shared Hierarchical Academic Research Computing Network (SHARCNET: www.sharcnet.ca) and Compute Canada. We would like to thank D. Toom-Sauntry for analyzing the MSA samples, technicians, operators and staff of CFS Alert for their support in maintaining the aerosol program. NOAA OI SST V2 data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at http://www.esrl.noaa.gov/psd/. NCEP Reanalysis data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at http://www.esrl.noaa.gov/psd/. Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data provided by The National Snow and Ice Data Center, Boulder, Colorado, USA, from their Web site at http://nsidc.org/data/nsidc-0051.html. Concentrations of OH from MOZART, Version 3 were kindly provided by The European Centre for Medium-Range Weather Forecasts. Concentrations of BrO from p-TOMCAT were kindly provided by Xin Yang, Department of Chemistry, University of Cambridge, UK.

Chapter 3

Inverse analysis of elemental carbon aerosol in northern China: Evidence of an underprediction in bottom-up emissions

This chapter is adapted from a paper that is to be submitted to the Journal of Geophysical Research-Atmospheres. Co-authors on the paper include John C. Lin, Lin Huang, Qiang Zhang, Sicong Kang, and Myung Gwang Kim. Lin Huang provided the scientific motivation for the study. Samples of elemental carbon that were collected at Tsinghua University were analyzed in her lab at Environment Canada. These measurements were well suited to be used in an inversion framework to constrain the very uncertain emissions of black carbon in northern China. Qiang Zhang and Sicong Kang provided bottom-up emissions inventories that served as a priori estimates. Myung Gwang Kim provided assistance and R scripts that aided in the evaluation of transport and boundary layer height errors. John C. Lin provided leadership and training in the use of STILT and Bayesian inversion, and aided in the editorial process.

Overview

Atmospheric concentrations of elemental carbon aerosol were measured on weekly integrated filter samples collected at Tsinghua University in Beijing (40.00°N, 116.33°E) from 1 January 2006 through 31 December 2008, and used as top-down constraints on a bottom-up emissions inventory with a Bayesian inversion technique. Model predicted elemental carbon concentrations were calculated using a spatially resolved emission inventory, air parcel trajectories from the Stochastic Time-Inverted Lagrangian Transport (STILT) model, background concentrations from the Goddard Earth Observing System Chemical Transport Model (GEOS-Chem), and modeled removal processes. Scaling factors for the magnitude of emissions were derived, and an uncertainty analysis was performed. Results showed that emissions of elemental carbon in northern China were likely underpredicted by the bottom-up inventory. The severity of the underprediction increased during the winter months, coinciding with the increased use of coal for residential heating, suggesting that the portion of the inventory attributable to residential coal use was underestimated. An apparent interference by emissions outside of the domain obfuscated results for the early summer emissions scaling, resulting in poorly constrained emissions during this period.

3.1 Introduction

Black carbon has a major effect on the climate system through direct absorption of solar radiation, influence on cloud properties, and the reduction of snow/ice albedo. Bond et al., 2013 [10] recently estimated the net global mean climatic forcing of industrial era black carbon in the year 2000 as +1.1 $W\,m^{-2}$ with 90% uncertainty bounds of (+0.17, +2.1). Estimates from previous studies ranged from +0.4 to +1.4 $W\,m^{-2}$ [42, 47, 17, 101, 88], and the Intergovernmental Panel on Climate Change, 2007 [110] estimated the total industrial era forcing as +0.44 $W\,m^{-2}$ with 90% uncertainty bounds of (+0.09, +0.79). The estimate of Bond et al., 2013 [10] places black carbon as the second most important agent of global anthropogenic climate forcing, behind only carbon dioxide. With an atmospheric lifetime on the order of days to weeks [73, 111, 88], elemental carbon aerosol emission reductions would likely be quick to yield reductions in the anthropogenic climate forcing [62].

Elevated black carbon concentrations are also known to cause adverse health effects in humans. Short term elevations in atmospheric black carbon concentrations have been shown to increase rates of cardiovascular and all-cause mortality, and cardiopulmonary hospital admissions [48]. Black carbon is not itself toxic or carcinogenic, but can increase the bioavailability of carcinogenic compounds by carrying and depositing them into the respiratory system [126, 118].

China's State Council has recently approved its first national standard for limiting the amount of fine particulates in the atmosphere [144]. A major component of the urban fine particulate aerosol is black carbon. Efforts to reduce particulate pollution are likely to reduce black carbon as well, both improving the health of the local population and reducing the anthropogenic radiative forcing. A thorough understanding of existing emission sources is necessary in order to make targeted emissions reductions. Black carbon emissions from northern China have been identified as some of the strongest in the world [11].

In this study we used the Stochastic Time-Inverted Lagrangian Transport (STILT) model [68], a Lagrangian particle dispersion model, and elemental carbon measurements from weekly integrated fine aerosol at Tsinghua University, Beijing, from 1 January 2006 through 31 December 2008 in a Bayesian inversion analysis to constrain bottom-up inventory estimates of black carbon emissions in northern China [145].

The terms "elemental carbon" and "black carbon" are often used synonymously in the literature on carbonaceous aerosols. There is an important distinction, as each is operationally defined by different measurement methodologies. Elemental carbon is defined by its thermal refractory property and measured as a mass concentration. Black carbon is defined by an optical property, i.e., measured as light absorption, and related to a mass concentration by a mass attenuation coefficient. The relation between the mass concentration of black carbon and elemental carbon concentration in a given sample are generally close to one-to-one, except in the case when there are significant concentrations of dark organic carbon [90]. Also, methods are not standardized, and so differences exist between the results of measurements of elemental carbon or black carbon carried out at different sites.

Measurements utilized in this study were of elemental carbon, but many of the model inputs were reported in the literature as relating to black carbon. We assumed that model inputs relating to black carbon are valid for the prediction of elemental carbon. In most cases this is exactly true, as black and elemental carbon are used synonymously in the literature describing the model inputs. In the model description we followed the source literature in describing parameters as relating to black or elemental carbon.

The global black carbon modeling literature reports a ubiquitous underestimate in model predicted surface concentrations in East Asia [57]. This may be partially explained by an inadequate representation of the effects of internal mixing on the optical properties of atmospheric aerosol, but a major component of the underestimate must be due to underprediction of the atmospheric burden of black carbon by either underpredicting emissions or overpredicting removal processes [10]. Since measurements of elemental carbon were used in our study, we avoided having to consider any optical properties of the aerosol and were left to focus on the component relating to the atmospheric burden. Furthermore, the site of measurement is in close proximity to sources, limiting the effect of biases in the modeling of removal processes.

Bottom-up inventories of black/elemental carbon emissions in China have reported large uncertainties, with ranges as large as (-100%, +500%) [117, 11, 145, 60, 147]. A major source of uncertainty is difficulty in quantifying emission factors from the use of coal for residential cooking and heating. The emission factor for coal is highly dependent on both the coal type and preparation. Bituminous coal has a much higher emission factor than anthracite coal, and raw coal stones have a much higher emission factor than coal processed into briquettes. The emission factor for raw, bituminous coal can be as large as 500 times the emission factors for anthracite coal briquettes. The fractions of coal types and preparations are poorly constrained, leading to the large uncertainties in the emission factors. In the region that we analyzed emissions were dominated by residential coal use, comprising 55% of the bottom-up emissions estimate [145]. Differences between the bottom-up emissions estimates and the inversion retrieval were likely caused by biases in the bottom-up estimates of residential coal emissions.

Top-down constraints have been applied to black carbon concentrations in the vicinity of Beijing in the past. Wang et al., 2011 [130] examined the ratio of black carbon to the trace gasses CO, CO_2, NO_y, and SO_2 from April to October of 2010. By analyzing measurements with air parcel trajectories from the HYbrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model they found that residential coal use must either have a smaller total share of black carbon emissions than predicted in the Zhang et al., 2009 [145] inventory, or the total emissions must have a lower than predicted BC/CO emission ratio. They suggested that the industrial and transportation sectors may be the primary source of black carbon to Beijing, in contrast to the predictions of the bottom-up inventory. The time frame that was studied in this work did not include the winter residential heating season, which limited the ability of the study to make inferences on the impact of residential coal. Also, the use of a single particle trajectory model may have over-simplified transport processes and led to a poor constraints on potential source regions.

Hakami et al., 2005 [39] conducted a four-dimensional variational data assimilation study to constrain emissions of black carbon during the Asian Pacific Regional Aerosol Characterization Experiment. By applying the adjoint version of the STEM-2k1 model with a bottom-up inventory from Streets et al., 2003 [117], they found model assimilated emissions of black carbon in northeastern China that were 74% higher than a priori values. This study used direct mass concentration observations of black carbon from several sites, and with a high time resolution. This study only analyzed the month of April 2001, and so extrapolations to analyze the annual emissions budgets are not possible.

Schutgens et al., 2012 [104] applied a fixed-lag Kalman smoother to constrain aerosol emissions in the MIROC-SPRINTARS global aerosol model, with top-down constraints from aerosol optical thickness (AOT) measurements from the MODIS Terra AOT over oceans and AERONET AOT over land. They found that global emissions of carbonaceous aerosol are underpredicted by 64%. This work is dependent on the relationship between the mass concentration of black carbon and the AOT, which is poorly understood and has been an ongoing research goal of the black carbon modeling community [10].

The data and model that were available for the current study permitted an analysis of a region that is important to the global burden of elemental carbon. They included both the heating and non-heating seasons, allowing for the interpretation of differences between the seasons with relation to differences in fuel usage. The use of elemental carbon measurements allowed for a direct analysis of the atmospheric burden, without having to consider the relationship between mass concentrations and aerosol optical properties. The use of the STILT model in our study allowed us to attribute variations in concentrations directly to emissions. The present study therefore provided an inversion with a temporal coverage and spatial targeting in a manner that previous studies have not achieved.

3.2 Inverse Modeling Framework

Bayesian inversion frameworks have been applied to Lagrangian particle dispersion models many times in the past by a variety of authors, (e.g., [69, 28, 114, 147, 120, 53, 66]). Rodgers, 2000 [96] provides a detailed review of the subject. Here we present only the aspects relevant for the application to the problem at hand.

The methodology is based on the linear relationship between the observation vector $\mathbf{y}$ and the vector of quantities $\boldsymbol{\lambda}$ to be inverted for:

$$\mathbf{y} = \mathbf{K}\boldsymbol{\lambda} + \mathbf{b} + \boldsymbol{\epsilon}. \tag{3.1}$$

In this study, the vector $\mathbf{y}$ represents measured concentrations of elemental carbon. The vector $\boldsymbol{\lambda}$ represents scaling factors applied to emissions within the modeling domain. Solving for an optimized estimate of $\boldsymbol{\lambda}$, $\hat{\boldsymbol{\lambda}}$, was the goal of this study. The vector $\mathbf{b}$ is the background concentration — the concentration of elemental carbon at the site of measurement that originated from outside of the spatial or temporal domain of the model. The matrix $\mathbf{K}$ represents the STILT model predicted influence of emissions within the modeling domain, accounting for the effects of transport and deposition. Each column of $\mathbf{K}$ corresponds to a grid cell in the emissions data, and each row corresponds to a time represented by an element of $\mathbf{y}$. The vector $\boldsymbol{\epsilon}$ is the uncertainty associated with the prediction of $\mathbf{y}$ by the model, including uncertainties in the measurements. The uncertainty covariance matrix associated with $\boldsymbol{\epsilon}$ is denoted S_ϵ.

To a close approximation, the model responds linearly to uniform scaling of the emissions. There is one deviation from linearity; an increase in emissions will increase the coagulation of particulates, increasing the sensitivity to wet deposition (see Section A.1.2). We can write $\mathbf{K} \approx \mathbf{K_E}\mathbf{diag}\left(\boldsymbol{\phi}\right)$, where $\boldsymbol{\phi}$ is a vector of emission strength and $\mathbf{K}_E$ is a matrix representing transport and removal processes.

$\boldsymbol{\lambda_a}$, the a priori value of $\boldsymbol{\lambda}$, is a vector with all entries equal to 1. This represents no scaling of the a priori emissions estimates. There is an associated uncertainty covariance matrix, $\boldsymbol{S_a}$.

A posteriori values of $\boldsymbol{\lambda}$ are given by:

$$\hat{\boldsymbol{\lambda}} = \boldsymbol{\lambda_a} + \mathbf{G}\left(\mathbf{y} - \mathbf{b} - \mathbf{K}\boldsymbol{\lambda_a}\right). \tag{3.2}$$

Where $\mathbf{G}$ is the gain matrix, given by:

$$\mathbf{G} = \boldsymbol{S_a}\mathbf{K}^T\left(\mathbf{K}\boldsymbol{\sigma_\epsilon}\mathbf{K}^T + \boldsymbol{S_a}\right)^{-1}. \tag{3.3}$$

And the a posteriori uncertainty covariance matrix is:

$$\hat{\boldsymbol{S}} = \left(\mathbf{K}^T\boldsymbol{S_\epsilon}^{-1}\mathbf{K} + \boldsymbol{S_a}^{-1}\right)^{-1}. \tag{3.4}$$

3.3 A Priori Emissions Estimates

A Priori emissions estimates were taken from Zhang et al., 2009 [145]. These are bottom-up estimates, spatially resolved at 0.5° × 0.5° and temporally resolved at one month.

The measurement site at Tsinghua University was in close proximity to an area of high emissions intensity, making modeled elemental carbon concentrations particularly sensitive to potential errors associated with coarse resolution [51, 34]. To reduce this error, emissions in each $0.5° \times 0.5°$ grid cell were distributed according to $0.1° \times 0.1°$ coal consumption data from Lei et al., 2011 [60]. Magnitudes of emissions from Lei et al., 2011 [60] were not used directly as a priori emissions estimates because the dataset contained only fuel usage data, and did not relate the fuel usage to emissions of elemental carbon. A time series of the magnitude of total emissions in the domain is shown in Figure 3.1, and a map showing the spatial distribution of the mean emissions for the period is shown in Figure 3.2.

Zhang et al., 2009 [145] reported uncertainties in the bottom-up black carbon emissions for all of China as having a standard deviation of ±104%. This was used in our study as the uncertainty in individual elements of of the a priori emissions estimate. Uncertainty covariance was calculated with a spatial decorrelation length scale, D, so that the uncertainty covariance between gridcells i and j, separated at their centers by a distance d would be $\mathbf{S}_{\mathbf{a}i,j} = \sqrt{var(i)var(j)}\exp(-d/D)$. D was found by following the suggestion of Gerbig et al., 2006 [33]; by calculating the a posteriori model predicted elemental carbon concentrations using a range of estimates of D, and choosing the value where the reduced Pearson χ^2 statistic approached 1. This resulted in a value of $D = 50\,km$ (see Fig. 3.3).

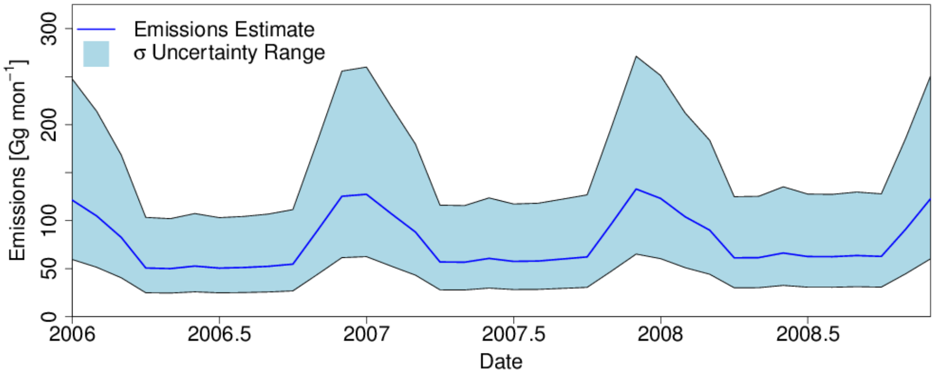


Figure 3.1: Total black carbon emissions in the northern Chinese domain shown in Figure 3.2, as predicted by the bottom-up inventory of Zhang et al., 2009 [145].

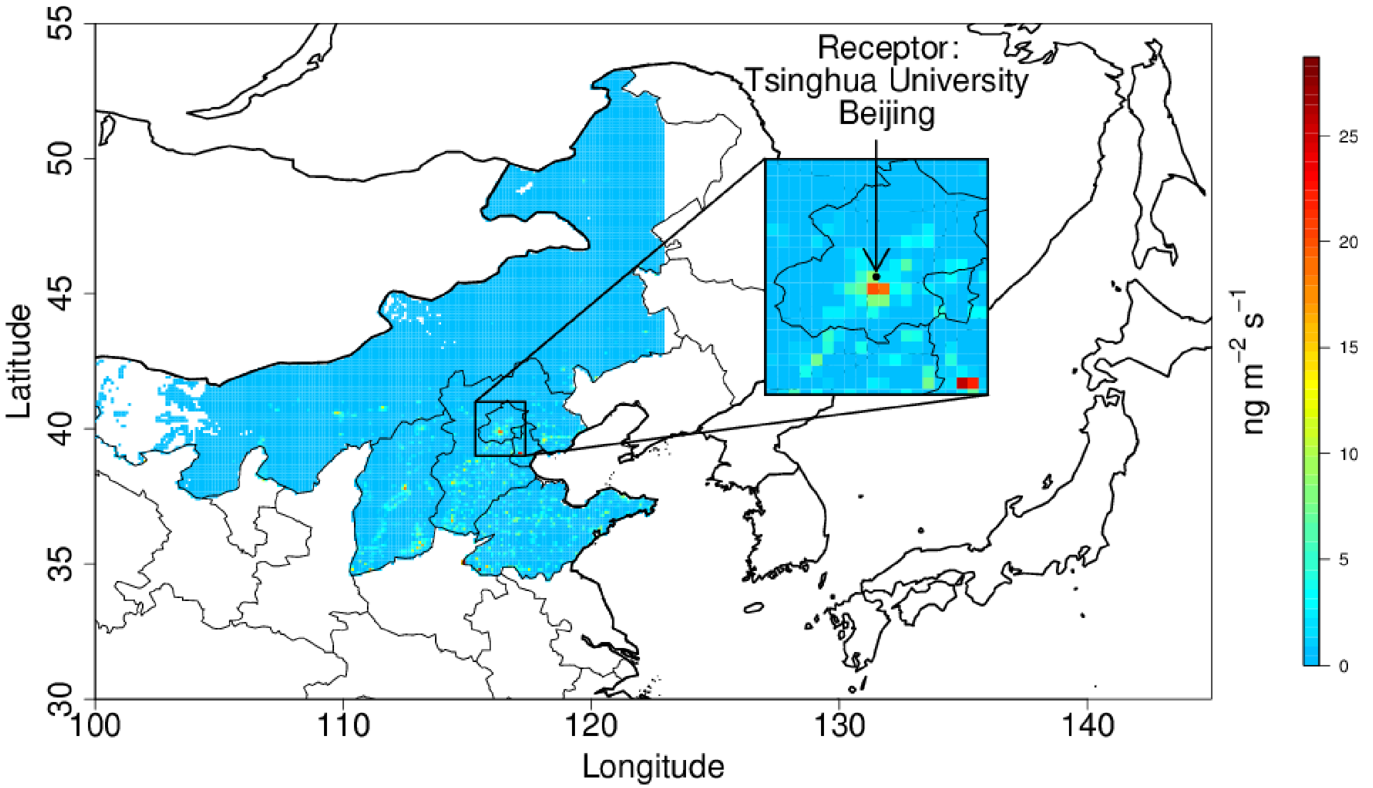

Receptor:
Tsinghua University
Beijing
Latitude
55
50
45
40
35
30
100
110
120
130
140
Longitude
ng m^{-2} s^{-1}
25
20
15
10
5
0

Figure 3.2: Spatial distribution of mean bottom-up inventory estimate of black carbon emissions in northern China During the period 1 January 2006 through 31 December 2008, with a resolution of 0.1° × 0.1°. Magnitude of emissions and spatial distribution to a 0.5° × 0.5° scale were taken from Zhang et al., 2009 [145]. Distribution at scales less than 0.5° × 0.5° were inferred from Lei et al., 2011 [60].

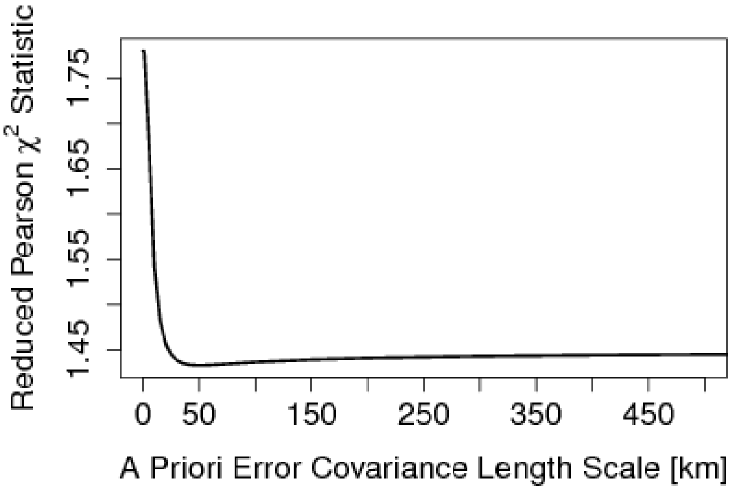


Figure 3.3: Reduced Pearson χ^2 statistic comparing a posteriori STILT model predictions of elemental carbon concentrations at Tsinghua University (40.00°N, 116.33°E) to concentrations measured on weekly integrated $PM_{2.5}$ filter samples with the EnCan-total-900 thermal method, as a function of a priori uncertainty decorrelation length scale. The ideal χ^2 statistic is equal to 1. The decorrelation length scale where the χ^2 statistic is closest to 1 (50 km) is assumed correct.

3.4 Measurements

Samples of $PM_{2.5}$ were collected on weekly integrated quartz fiber filters with a 4.25 cm radius and 77.56 $L\,min^{-1}$ flow rate at Tsinghua University in Beijing (40.00°N, 116.33°E) from 1 January 2006 through 31 December 2008. The filter samples were shipped to the Stable Isotope Research Lab at Environment Canada's Atmospheric Science and Technology Directorate in Toronto, where they were analyzed for elemental carbon concentrations with the EnCan-total-900 thermal method [14], on a Sunset Laboratories thermal-optical semi continuous OC/EC analyzer. A time series of the measurements are shown in Figure 3.4.

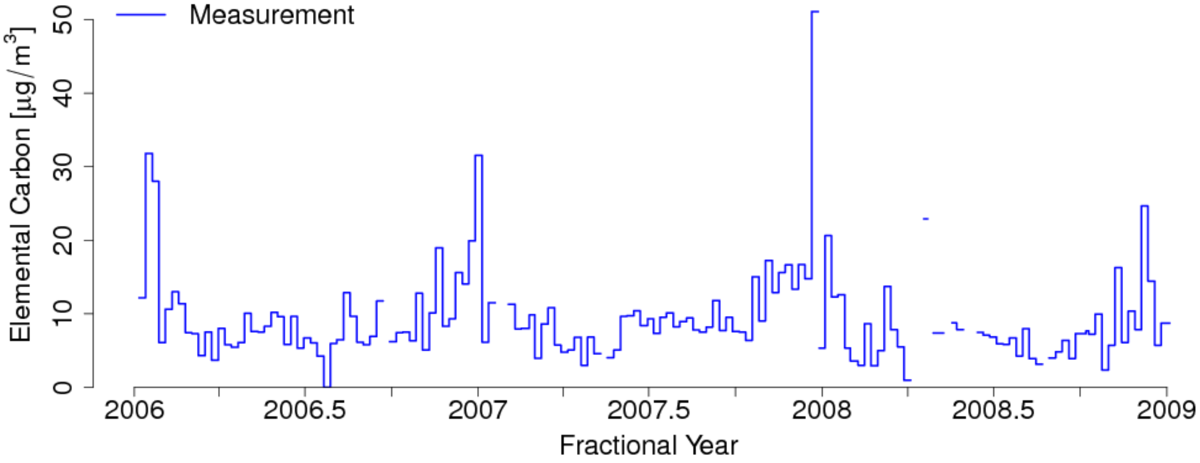


Figure 3.4: Measurements of elemental carbon aerosol concentration from weekly integrated $PM_{2.5}$ filter samples collected at Tsinghua University during the period 1 January 2006 through 31 December 2008. Concentrations were measured by the Stable Isotope Research Lab at Environment Canada with the EnCan-total-900 thermal method.

3.5 Simulating Elemental Carbon Concentrations

The model used to simulate elemental carbon concentrations in the inversion framework is given by Equation 3.1. This is a model relating the concentration of elemental carbon measured at Tsinghua University (the "receptor", or "measurement site") to emissions and deposition on a regional scale, and the background concentration. This section describes the algorithm used to calculate model estimates of elemental carbon concentrations. Each component of the model is described in Sections 3.5.1 through 3.5.4, with the details of calculation contained in the appendices.

Ensembles of 500 STILT air parcel trajectories were traced backwards in time for a duration of 5 days, initialized at Tsinghua University in Beijing (40.00°N, 116.33°E, 7 *m* above ground level) twice daily (noon and midnight, local time) from 1 January 2006 through 31 December 2008. The model was driven by meteorological fields from the Global Data Assimilation System (GDAS1), which are resolved at a $1^\circ \times 1^\circ$ spatial resolution and represent instantaneous conditions at 3 hour time intervals. The trajectory of any particle which left the domain of emissions (regions of specified emissions in Fig. 3.2) prior to 5 days was truncated. On average 99.77% of the particles in an ensemble exited the domain before 5 days had elapsed. 27.24% of the ensembles had at least one particle remain within the domain for the entire 5 day modeling period, and 11.03% of ensembles had at least 10% of their particles remain in the domain for the entire 5 day modeling period. The mean total transit time of a particle was 1.34 days. A model predicted concentration was calculated for each particle individually, and the mean of all 500 particles in the ensemble was taken as the model predicted concentration at the receptor.

Pure elemental carbon is hydrophobic, but the processes of condensation, coagulation and oxidation can lead to elemental carbon that is internally mixed with other constituents, and vulnerable to wet deposition [71]. To account for this, separate models were calculated for the concentration of "hydrophobic" and "hydrophilic" elemental carbon, along with a conversion of hydrophobic elemental carbon into hydrophilic elemental carbon. It was assumed that hydrophobic-to-hydrophilic conversion does not affect the measured concentration of elemental carbon directly, since compounds in the internal mixed aerosol do not cause significant interference in EC concentrations from the EnCan-total-900 method.

STILT outputs air parcel trajectory data at dynamically determined time intervals ranging in duration from 20 to 60 *min*. The concentrations of both hydrophobic and hydrophilic elemental carbon were calculated on one minute time steps nested in each of the STILT time steps, starting from the point where the trajectory was truncated, and moving forward in time. Changes in the concentrations were due to emission, hydrophobic-to-hydrophilic conversion, and deposition. Concentrations were calculated as a molar mixing ratio until arrival at the measurement site where they were converted to a mass concentration. The model is unconditionally stable for all time step durations, and the error due to time discretization was assumed to be negligible.

During each 1 minute time step the concentration of hydrophilic and hydrophobic elemental carbon was calculated as follows. A graphical representation of the algorithm is displayed in Figure 3.5.

1. The mixing ratios of hydrophilic and hydrophobic elemental carbon were initialized to the final value from the previous time step. If the current time step was the first time step for this particle's trajectory, then the mixing ratios were initialized to the respective background concentrations.

2. The influence of elemental carbon emissions was added to the mixing ratios in the trajectory.

3. A fraction of the hydrophobic elemental carbon was converted to hydrophilic elemental carbon, according to a time scale determined with a hydrophobic-to-hydrophilic conversion scheme.

4. Removal by dry deposition was applied to the hydrophobic elemental carbon, and removal by both dry and wet deposition were applied to hydrophilic elemental carbon.

5. The resulting mixing ratio was passed to the next time step in the trajectory. If the current time step corresponded to arrival at the receptor site, the mixing ratio was converted to a mass concentration.

6. The mean concentration for the trajectory ensemble was taken as the modeled concentration.

3.5.1 Background Concentrations

Background concentrations of hydrophobic and hydrophilic black carbon were provided by personal communication with Dylan Millet and Lu Hu at the University of Minnesota, who calculated them using the Goddard Earth Observing System Chemical Transport Model (GEOS-Chem) [9, 128] on a 2.5° × 2° grid with 47 vertical levels. Monthly mean concentrations were calculated for the year 2010, and applied to all years under the assumption that the inter-annual variation is insignificant compared to the model uncertainty.

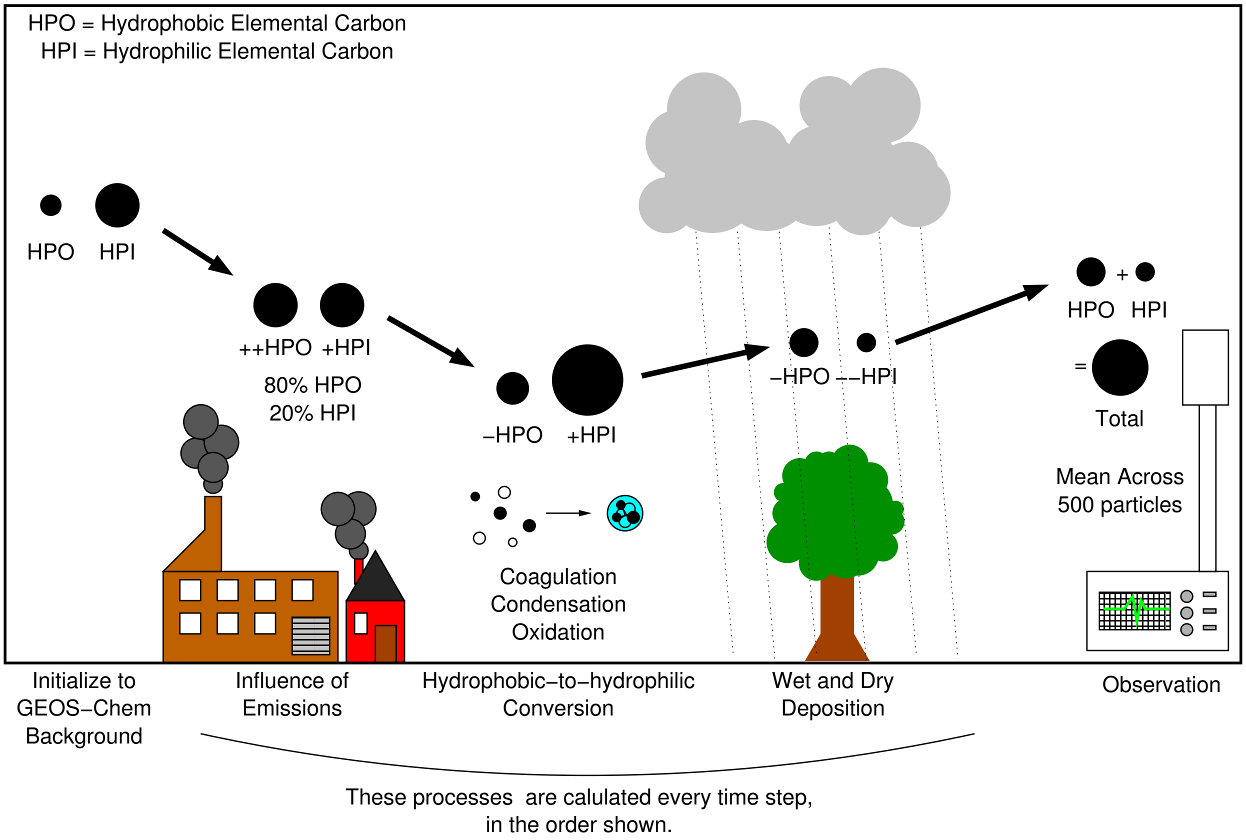

HPO = Hydrophobic Elemental Carbon
HPI = Hydrophilic Elemental Carbon
HPO
HPI
++HPO
+HPI
80% HPO
20% HPI
−HPO
+HPI
Coagulation
Condensation
Oxidation
−HPO
−−HPI
HPO
+
HPI
=
Total
Mean Across
500 particles
Initialize to
GEOS−Chem
Background
Influence of
Emissions
Hydrophobic−to−hydrophilic
Conversion
Wet and Dry
Deposition
Observation
These processes are calulated every time step,
in the order shown.

Figure 3.5: Qualitative diagram of elemental carbon aerosol transport model implemented with the STILT Lagrangian particle dispersion model. The size of black dots represent a relative magnitude of elemental carbon concentration. + and − indicate the addition or subtraction of elemental concentration at each model step, and multiple + or − signs indicate a change in concentration that is typically large relative to other changes in the same model step.

3.5.2 Contribution of Emissions

Emissions contribute to the elemental carbon concentration along a particle's trajectory when the particle height drops below $z^\star$; one half the planetary boundary layer height. The surface flux due to emissions is vertically diluted in a grid volume defined by the grid of emissions in the horizontal and $z^\star$ in the vertical [68]. This assumes that there is rapid mixing in the planetary boundary layer, and that the sources of emissions are below the height $z^\star$. The spatio-temporal distribution of sensitivity to emissions are then defined as a "footprint", in units of $ppm/\,(\mu mol\, m^{-2}\, s^{-1})$. The enhancement in concentration of elemental carbon during the time step (in units of ppm) is then the product of the footprint and the emission rate (in units of $\mu mol\, m^{-2}\, s^{-1}$). A plot of the mean footprint during the study period is shown in Figure 3.6.

3.5.3 Hydrophobic-to-Hydrophilic Conversion

Following several published studies [73, 56, 18, 19], it was assumed that 80% of freshly emitted elemental carbon is in a hydrophobic form. In order to calculate hydrophilic-to-hydrophobic conversion, the condensation and coagulation scheme of Riemer et al., 2005 [93], and the oxidation algorithm of Tsigaridis and Kanakidou, 2003 [123] were adapted. Details of the implementation are given in Appendix A.1.

3.5.4 Deposition

Dry and wet deposition were calculated independently as timescales of exponential decay, and then converted into fractions of the elemental carbon concentration that is removed during a one minute time step. Dry deposition was estimated using an adapted version of the size-segregated and land use dependent scheme of Zhang et al., 2001 [143]. Details of the calculations are given in Appendix A.2. Wet deposition was estimated by adapting the GEOS-Chem wet deposition scheme described by Jacob et al., 2000 [46] and Liu et al., 2001 [72]. Details of these calculations are given in Appendix A.3.

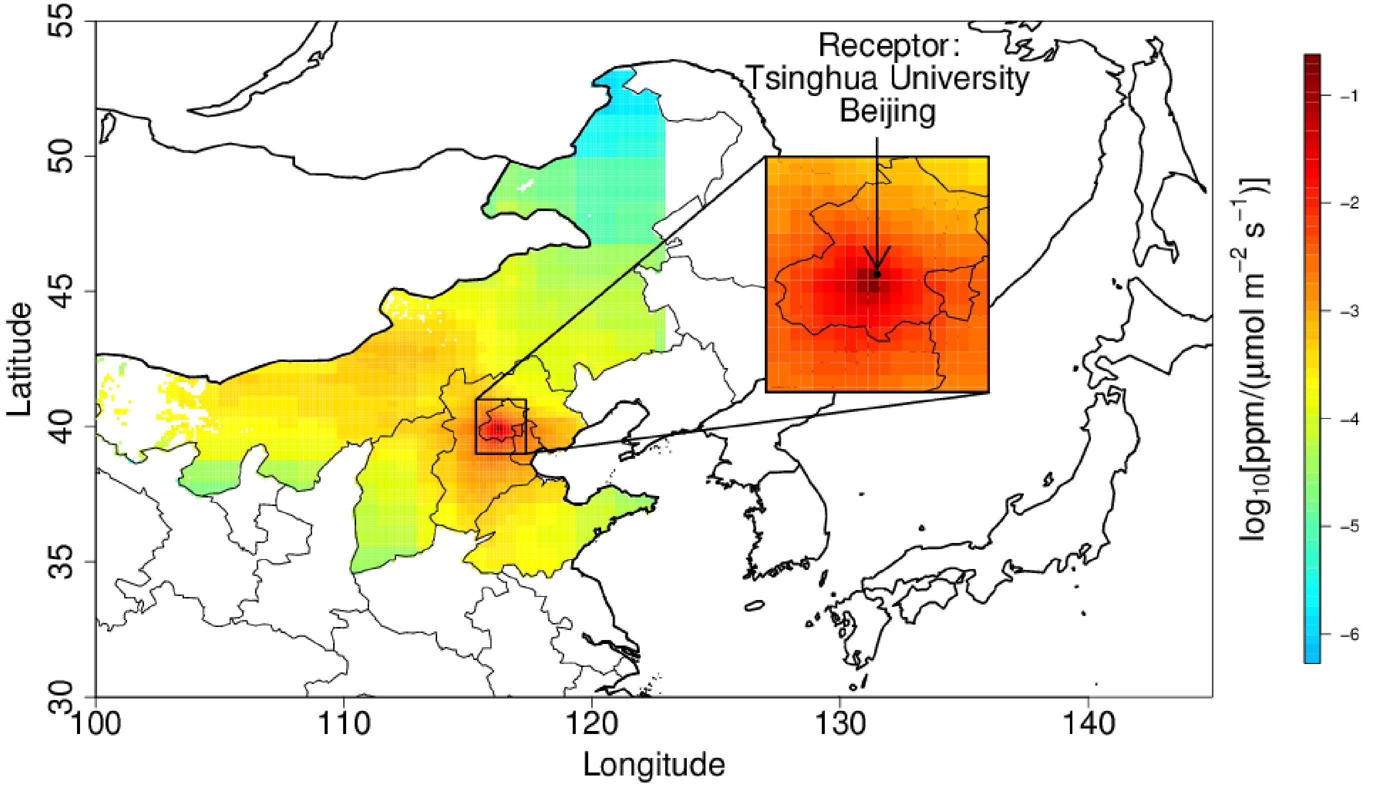

Receptor:
Tsinghua University
Beijing
Latitude
55
50
45
40
35
30
100
110
120
130
140
Longitude
log$_{10}$[ppm/(μmol m^{-2} s^{-1})]
−1
−2
−3
−4
−5
−6

Figure 3.6: Mean footprint for a receptor located at Tsinhghua University during the period 1 January 2006 through 31 December 2008. Calculated with STILT using meteorological fields from GDAS1.

3.6 Uncertainty Characterization

In order to properly apply a Bayesian inversion methodology, it is imperative to have a good handle on model uncertainty. In this section we describe the calculation of the contribution to model uncertainty for all known major sources of error in the model.

3.6.1 Measurements

The uncertainty in measurements was reported by Chan et al., 2010 [14] as a standard deviation of 10% of the measured concentration. This was assumed to be a random error, and uncorrelated with any of the other sources of error.

3.6.2 Aggregation

In reality, anthropogenic elemental carbon is emitted from from point-like sources. The emissions inventory used in this work has a resolution of $0.1° \times 0.1°$, introducing a spatial emissions aggregation error to the model.

Zhao et al., 2009 [146] evaluated the spatial aggregation error of methane emissions resolved at $1° \times 1°$. The emission inventory investigated in their work included emissions from landfill, which were fully resolved. By aggregating this subset of emissions to a $1° \times 1°$ grid, and evaluating the difference in influence on the concentration at the receptor they were able to obtain an estimate of the spatial aggregation error. They obtained a value of 5%.

In the present study, we did not have access to fully resolved emissions. We therefore could not explicitly calculate the aggregation error in such a straightforward manner. An alternative method was implemented, where emissions were successively aggregated and the model was calculated with each resolution. A third degree multiple polynomial regression relating the predicted concentration to the latitudinal and longitudinal resolution was calculated for every point in time during the study period. A reference concentration was estimated by extrapolating the regression to zero grid spacing in latitude and longitude. The error due to aggregation was taken as the difference between the reference concentration and the regression estimate of the $0.1° \times 0.1°$ resolution concentration.

The mean standard deviation of the errors in modeled elemental carbon concentrations was 1.45 $\mu g\, m^{-3}$ with a mean bias of +10.7% of the predicted concentration. The bias was corrected for in the model by scaling the effect of emissions.

3.6.3 Transport

Uncertainty in transport pathways represents a large source of uncertainty in any atmospheric transport model. This was especially true for the present study because of the proximity of the receptor to large emission sources. Small variations in transport could cause an air parcel to deviate between regions of highly variable emissions. Uncertainty due to transport error was evaluated by adopting the method of Lin and Gerbig, 2005 [67], where an added stochastic velocity component representing uncertainty in wind fields is propagated through the model. Uncertainty in model predictions made with a trajectory ensemble is then equivalent to the difference in spread of concentrations among particles caused by the added stochastic component.

Observations of wind profiles from the Universal RAwinsonde OBservation program (RAOB) for 32 stations in China (see Figure 3.7) from 1 January 2006 through 31 December 2008 were downloaded from the National Oceanic and Atmospheric Administration/Earth System Research Laboratory (NOAA/ESRL) Radiosonde Database. These observations were compared to estimates of winds interpolated from GDAS1 to the sub-grid scale by STILT.

The residual error (STILT estimate - RAOB observation) was calculated at each station, each available elevation, and each available point in time. A standard deviation of 4.3 $m\,s^{-1}$ was found. Error covariance statistics in space and time were calculated by fitting exponential variogram models to the data using the geoR package for the R Project for Statistical Computing [92]. Figure 3.8 shows a comparison between the winds predicted by STILT and observations from RAOB, for all available RAOB data points between ground level and 15000 m.

Error decorrelation time scales were calculated by fitting an exponential variogram in time to wind errors for each of the stations. These fittings resulted in mean value of 157 min for the error decorrelation timescale. Vertical decorrelation of horizontal wind errors were calculated independently for each station at each point in time, and independently for winds blowing in the zonal and meridional directions. The error decorrelation was found to be nearly isotropic, with mean values of 1103.1 m for zonal winds and 1066.1 m for meridional winds. The arithmetic mean of 1084.6 m was applied to the trajectories. For horizontal decorrelation of horizontal winds the error decorrelation length scales were calculated across all stations using all height levels. A mean of 329.2 km was found for zonal winds and 337.1 km for meridional winds, with no dependence on height. The arithmetic mean of 333.1 km was applied to the trajectories.

The mean standard deviation of errors in modeled elemental carbon concentrations due to transport uncertainty was 1.71 $\mu g\,m^{-3}$.

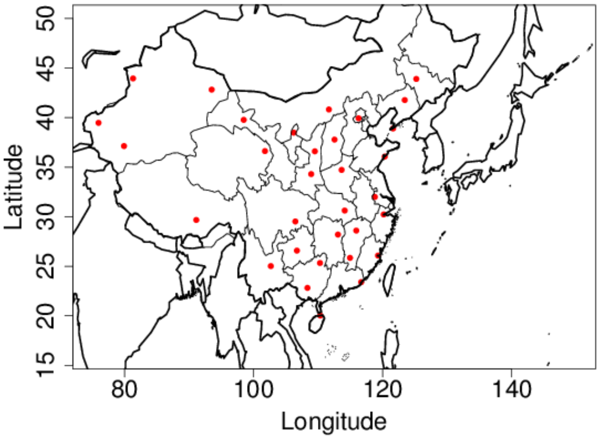


Figure 3.7: Locations of RAOB stations that provided data for the analysis of transport uncertainty in a STILT model of elemental carbon concentration at Tsinghua University (40.00°N, 116.33°E).

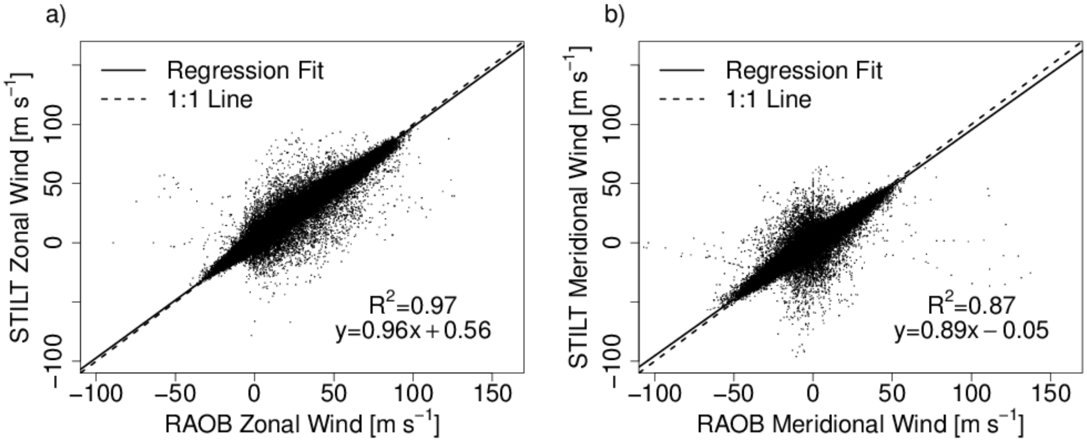


Figure 3.8: A comparison between winds predicted by STILT (an interpolation GDAS1 reanalysis), and observations from RAOB for all stations shown in Figure 3.7, during the period 1 January 2006 through 31 December 2008. All available data below a height of 15000 m above ground level was used. Regression slopes were calculated with an ordinary least squares method. a) Zonal wind, b) Meridional wind.

3.6.4 Planetary Boundary Layer

STILT computes an estimate of the planetary boundary layer (PBL) height. The PBL height affects the model predicted concentration by dictating the height of emissions influence, the volume over which emissions are diluted, and the height where boundary layer turbulent diffusion affects transport.

Errors in the estimation of the PBL height were calculated by comparing the STILT-predicted PBL height to observations estimated by applying the bulk Richardson number method to RAOB radiosonde profiles. The bulk Richardson number was taken as:

$$RiB = \frac{g \Delta \theta_v \Delta z}{\theta_v \left[(\Delta U)^2 + (\Delta V)^2 \right]}. \tag{3.5}$$

Where $g = 9.8\, m\, sec^{-2}$ is the gravitational acceleration, θ_v is virtual potential temperature, z is height above ground level where RiB is evaluated, U is the zonal wind velocity component, V is the meridional wind velocity component, and Δ denotes a difference in a quantity between the height where RiB is evaluated and the lowest RAOB level (i.e. $\Delta x = x(z) - x(lowest\, level)$).

The top of the PBL was assumed to be where the bulk Richardson number falls below a critical value of 0.25. STILT-predicted PBL heights are discretized to model vertical levels. A comparison of the STILT-predicted PBL heights with the bulk Richardson Number derived RAOB PBL heights is shown in Figure 3.9.

The effect of PBL errors on the model was quantified in a manner similar to the treatment of transport errors in Section 3.6.3. This method is described in Gerbig et al., 2008 [32]. The standard deviation of relative error in the planetary boundary layer height was 93.8%, the horizontal decorrelation length scale was 573.5 km, and the decorrelation time scale was 31.1 hr.

The mean standard deviation of errors in predicted elemental carbon concentrations due to PBL errors was 1.35 $\mu g\, m^{-3}$.

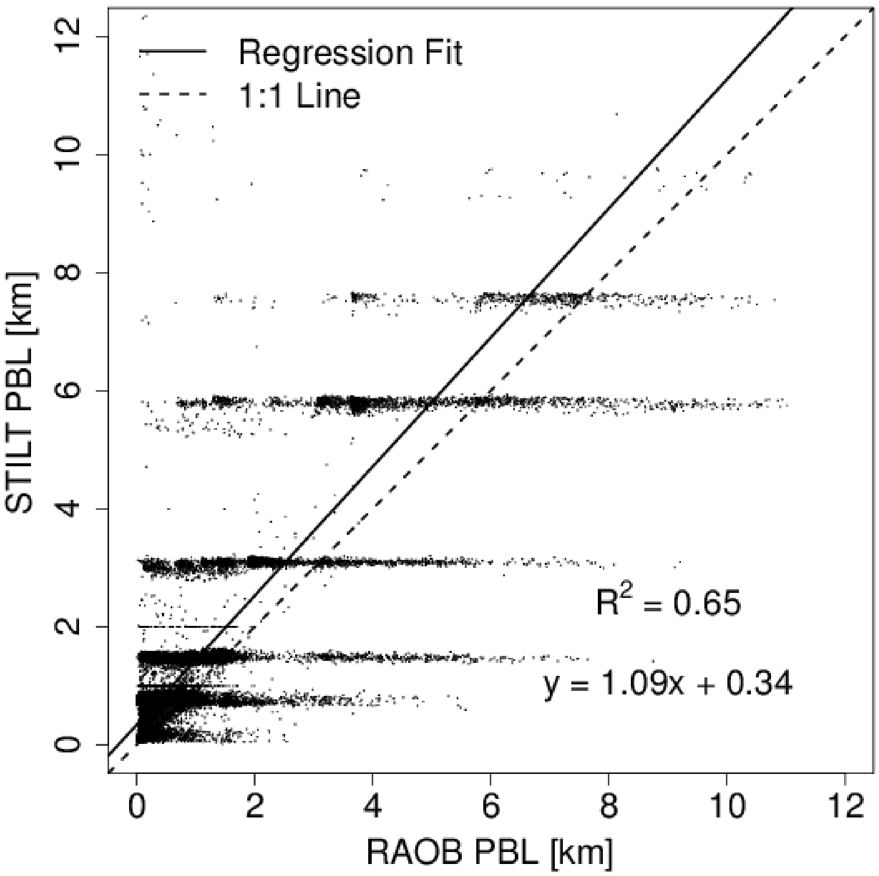


Figure 3.9: A comparison between Planetary Boundary Layer (PBL) heights predicted by STILT, and by a bulk Richardson number analysis of RAOB radiosonde profiles for all stations shown in Figure 3.7, during the period 1 January 2006 through 31 December 2008. The RAOB PBL top corresponds with a critical bulk Richardson number of 0.25. STILT-predicted PBL heights are discretized to model vertical levels. The regression was calculated with a standard major axis method.

3.6.5 Deposition

There were no available measurements with the spatio-temporal coverage to quantify the uncertainty in the deposition scheme applied in this work. To obtain a conservative estimate the model was run in the absence of any deposition, and the uncertainty was taken as 100% of the difference in predicted concentration. Since the greatest emissions occur close to the receptor site, deposition has a relatively small effect on the contribution from regional emissions, but is important for its effect on background concentrations.

The mean standard deviation of errors in modeled elemental carbon concentrations due to deposition uncertainty was 0.24 $\mu g\, m^{-3}$.

3.6.6 Background

Background concentrations were taken from a GEOS-Chem output that has no published uncertainty analysis. Since global black carbon models have been shown to underpredict surface level concentrations in Asia by a mean factor of 2, but overpredict concentrations aloft [57], we made a conservative assumption that the standard deviation of uncertainty in the background concentrations input to the model was 200%. A scaling was not applied to correct for the net underestimate because airmass trajectories are often aloft when they reach the background. Biased input parameters can introduce biases in the results of a Bayesian inversion, and so it was necessary to treat periods of larger background concentrations with caution during interpretation.

The mean standard deviation of errors in the model predicted elemental carbon concentrations due to uncertainty in the background concentrations was 1.91 $\mu g\, m^{-3}$.

3.6.7 Particle Number

A STILT particle represents a parcel of air that includes enough molecules to have macroscopic properties, but few enough such that it can be represented by a point in space. Close to an infinite number of particles (of order 10^{23}—i.e., Avogadro's number) would be required in order to achieve a perfect representation of the dispersion of molecules [66]. This is impossible in practise, as the number of particles used is limited by computational resources.

To quantify particle number error, and to choose the most appropriate number of particles for model computation, the model was computed for every fifth date in the study period using a number of particles ranging between 1 and 3000. Plots of the bias and root-mean-square error of models with respect to 3000 particle trajectory ensembles (the largest number of particles that was practical to use) are shown in Figure 3.10. At 500 particles the plots both approach an asymptote. 500 particle ensembles were therefore used for all other analyses. The particle error was estimated as a normal random error with standard deviation equal to the difference in concentration predicted with 500 particles and 3000 particles. This resulted in an estimate of the mean standard deviation of errors in modeled elemental carbon of 0.08 $\mu g\, m^{-3}$.

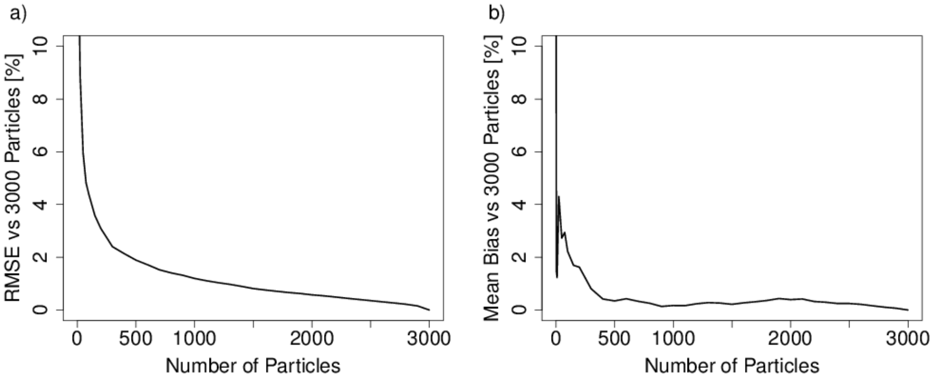


Figure 3.10: Plots demonstrating the effect of using a finite number of particles to simulate transport in a STILT model of elemental carbon aerosol concentrations at Tsinghua University (40.00°N, 116.33°E). a) Root-mean-square error for models computed with a varying number of particles compared against a model computed with 3000 particles. b) Mean bias for model computed with a varying number of particles compared against a model computed with 3000 particles.

3.6.8 Construction of the $\mathbf{S}_\epsilon$ Matrix

A time series of total model uncertainty, along with each of the components of uncertainty, is shown in Figure 3.11. It was assumed that the individual components contributing to the error were uncorrelated.

The model uncertainty covariance matrix was calculated with a temporal decorrelation length scale, $\tau_{\epsilon\, cov}$, so that the uncertainty covariance of model estimated elemenatal carbon concentrations separated by a time t would be $\mathbf{S}_{\epsilon i,j} = \sqrt{var(i)var(j)}\exp\left(-t/\tau_{\epsilon\, cov}\right)$. $\tau_{\epsilon\, cov}$ was found by fitting the experimental autocorrelation function to an exponential decay function. This yielded a decorrelation time scale of 10.1 days.

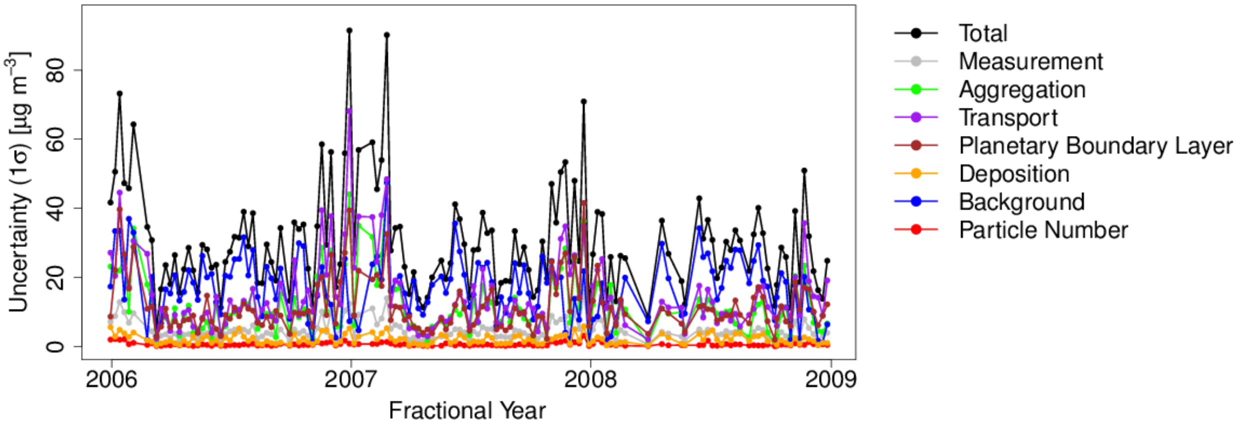


Figure 3.11: Time series of standard deviation of error in a STILT model of elemental carbon concentration at Tsinghua University (40.00°N,116.33°E). Individual sources of error variance were calculated with a variety of numerical experiments.

3.7 Results

Figure 3.12 shows the mean value of the model predicted influence of emissions — the mean row of the **K** matrix. Figure 3.13 shows measured concentrations of elemental carbon at Tsinghua University, along with the a priori model predicted concentration. A regression analysis is shown in Figure 3.14. The a priori model underpredicted the concentrations, with a standard major axis linear regression slope of 0.32, intercept of 0.90, and a squared coefficient of correlation of 0.43.

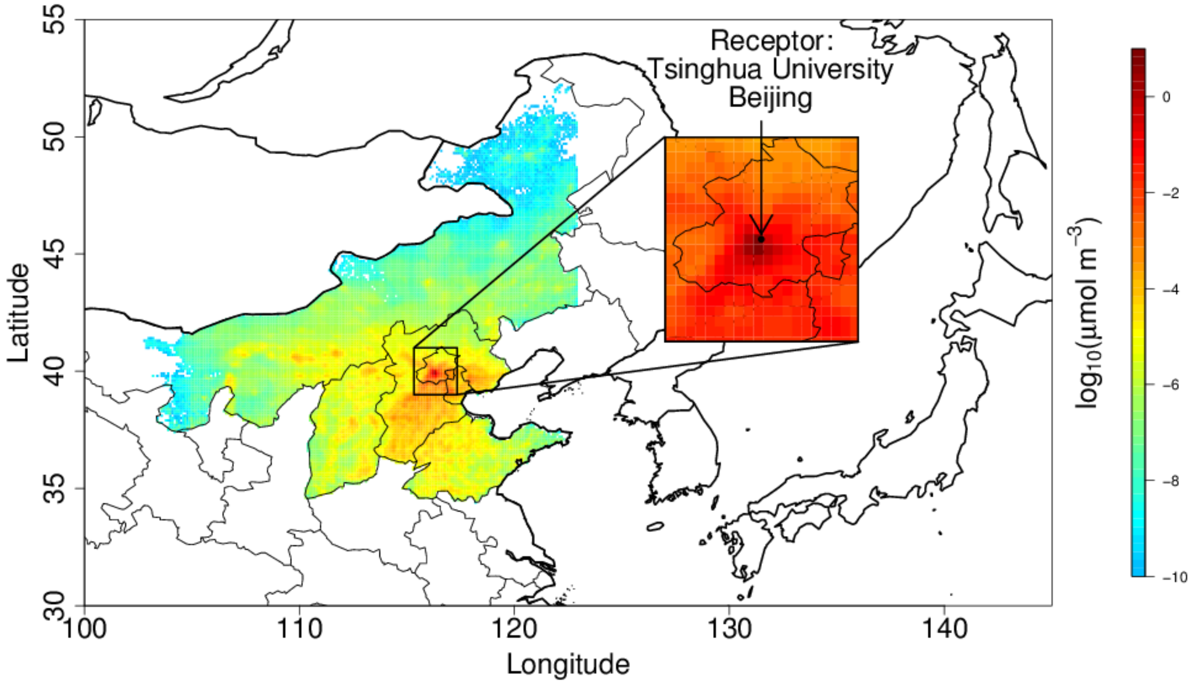

Receptor:
Tsinghua University
Beijing
Latitude
Longitude
55
50
45
40
35
30
100
110
120
130
140
$\log_{10}(\mu\text{mol m}^{-3})$
0
−2
−4
−6
−8
−10

Figure 3.12: STILT model estimate of the mean influence of elemental carbon emissions to Tsinghua University for the period 1 January 2006 through 31 December 2008, using bottom-up inventory (a priori) estimates of emissions.

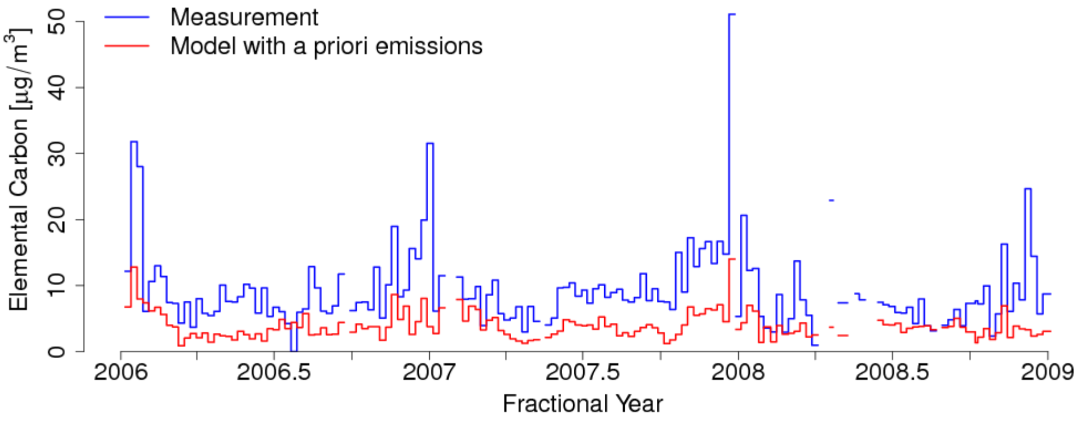


Figure 3.13: Concentrations of elemental carbon aerosol at Tsinghua University (40.00°N, 116.33°E) as predicted by a STILT model using bottom-up inventory (a priori) estimates of emissions (red), and as measured on weekly integrated $PM_{2.5}$ filter samples with the EnCan-total-900 thermal method (blue).

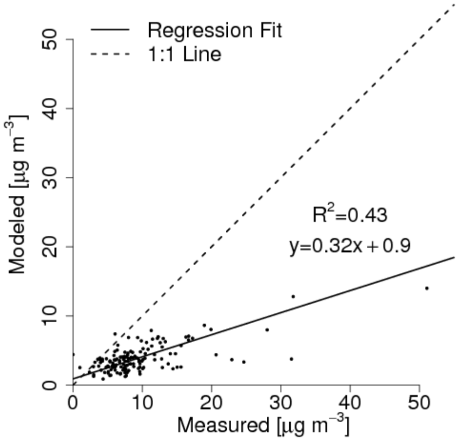


Figure 3.14: A comparison between STILT model predictions of elemental carbon concentrations at Tsinghua University (40.00°N, 116.33°E) made using bottom-up inventory (a priori) estimates of emissions to concentrations measured on weekly integrated $PM_{2.5}$ filter samples with the EnCan-total-900 thermal method. A standard major axis regression fit is included.

Due to the high uncertainty and possible bias in the background concentrations, there was concern about periods where they made up a large portion of the total modeled concentration of elemental carbon. In order to asses this, Figure 3.15 shows the percentage of the a priori modeled elemental carbon concentration that was a result of the background concentrations. If the background concentrations are biased low, then the scaling factors will be biased high during periods when the background makes up a significant portion of the model.

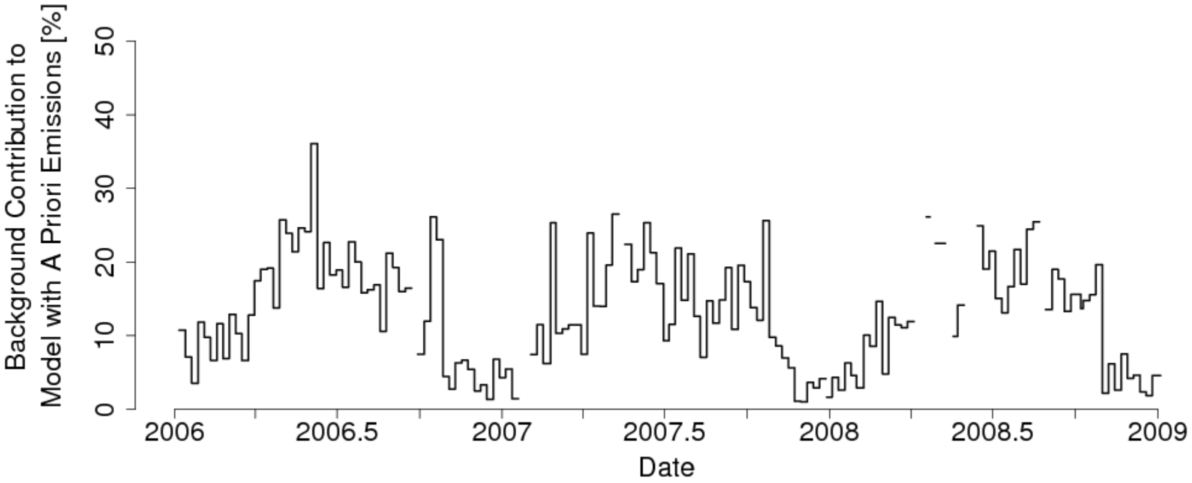


Figure 3.15: The percentage of background concentrations in STILT model predicted elemental carbon at Tsinghua University (40.00°N, 116.33°E). Model predictions were made using bottom-up inventory (a priori) estimates of emissions.

The inverse analysis of Section 3.2 was applied to calculate scaling factors for emissions that optimize the model. A violin plot of the resulting scaling factors is shown in Figure 3.16. This plot shows the kernel density smoothed distribution of scaling factors, where the width of each element is representative of the number of grid cells with a particular scaling factor. Figure 3.17 shows a plot of the mean spatial distribution of scaling factors. Figure 3.18 shows a map of the mean a posteriori uncertainty variance.

Figure 3.19 shows model predicted concentrations of elemental carbon at Tsinghua University, created using a posteriori scaled emissions, along with the measured concentrations. A standard major axis regression analysis was computed, and is shown in Figure 3.20. A portion of the underprediction persisted as a result of the inclusion of a priori emissions in the inversion framework. The regression slope was 0.68, the intercept was 1.46, and the squared coefficient of correlation was 0.60. This represents a marked improvement from the a priori model, as the model is both closer to one-to-one and explains a greater portion of the variability.

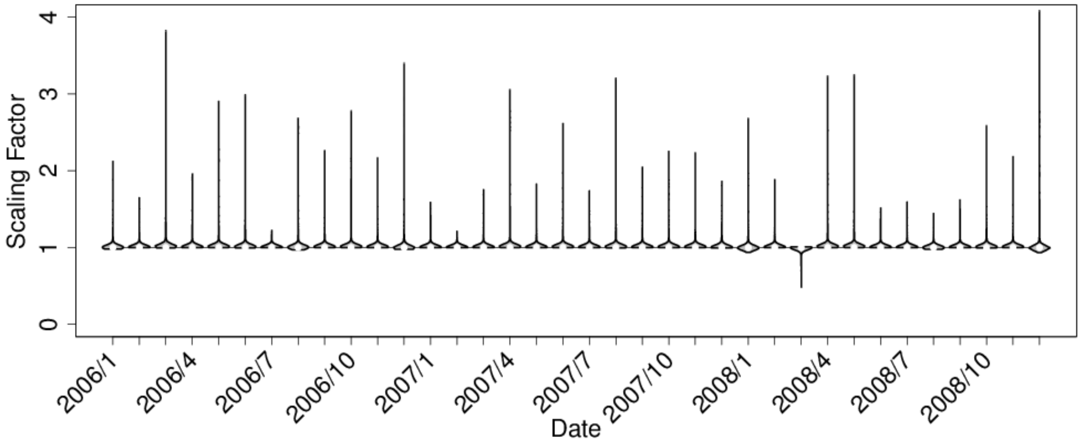


Figure 3.16: Violin plot of a posteriori scaling factors for elemental carbon emissions in the northern Chinese domain shown in Figure 3.2, calculated using a Bayesian inversion analysis of a STILT model of the concentration at Tsinghua University (40.00°N, 116.33°E). The width of each element is representative of the number of emissions grid cells as a function of the scaling factor.

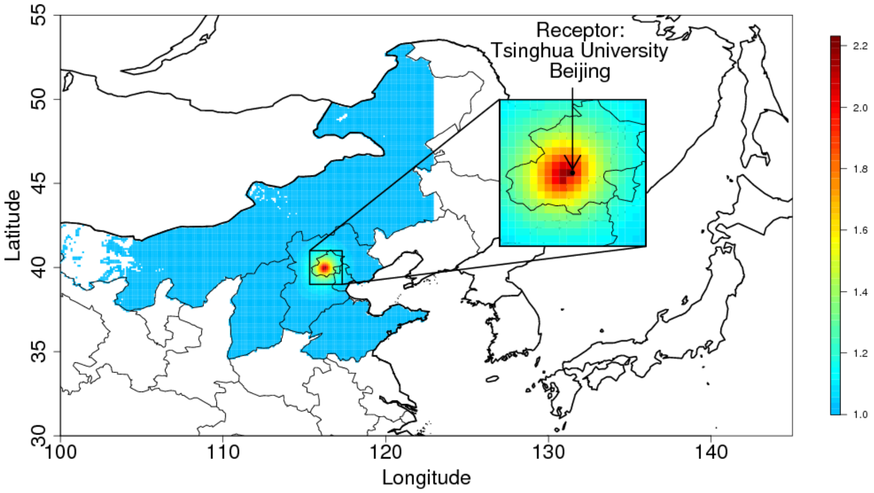

Receptor:
Tsinghua University
Beijing
Latitude
Longitude
30
35
40
45
50
55
100
110
120
130
140
1.0
1.2
1.4
1.6
1.8
2.0
2.2

Figure 3.17: Mean spatial distribution of a posteriori scaling factors for elemental carbon emissions in nortnern China for the period 1 January 2006 through 31 December 2008, calculated using a Bayesian inversion analysis of a STILT model of the concentration at Tsinghua University (40.00°N, 116.33°E).

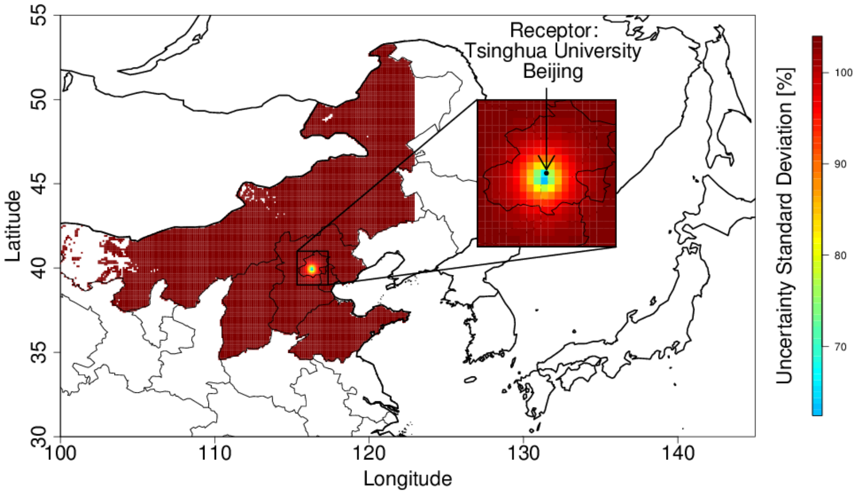

Receptor:
Tsinghua University
Beijing
Latitude
Longitude
30
35
40
45
50
55
100
110
120
130
140
Uncertainty Standard Deviation [%]
70
80
90
100

Figure 3.18: Mean spatial distribution of a posteriori uncertainty of elemental carbon emissions in nortnern China for the period 1 January 2006 through 31 December 2008, calculated using a Bayesian inversion analysis of a STILT model of the concentration at Tsinghua University (40.00°N, 116.33°E).

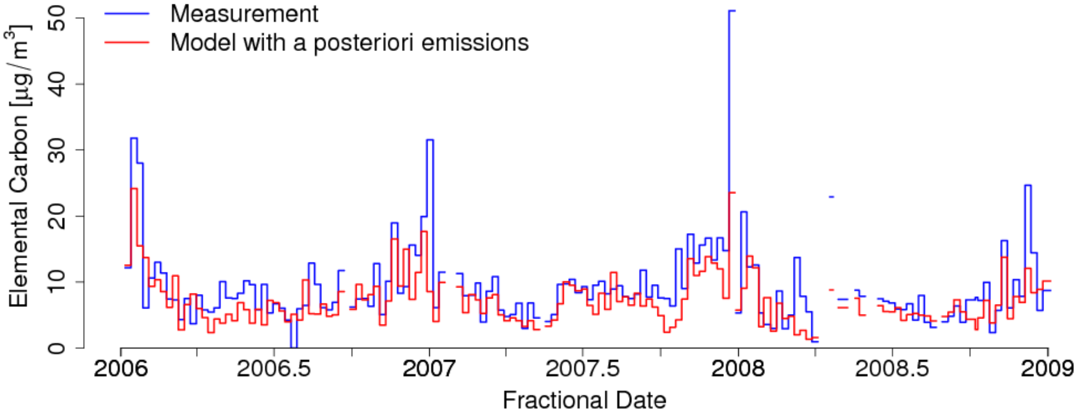


Figure 3.19: Concentrations of elemental carbon aerosol at Tsinghua University (40.00°N, 116.33°E) as predicted by a STILT model using a posteriori estimates of emissions calculated with a Bayesian inversion analysis (red), and as measured on weekly integrated $PM_{2.5}$ filter samples with the EnCan-total-900 thermal method (blue).

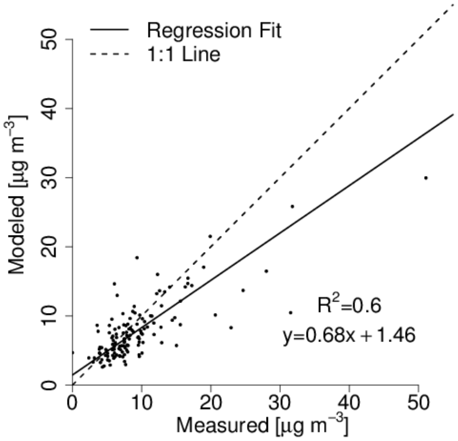


Figure 3.20: A comparison between STILT model predictions of elemental carbon concentrations at Tsinghua University (40.00°N, 116.33°E) made using a posteriori estimates of emissions calculated with a Bayesian inversion analysis to concentrations measured on weekly integrated $PM_{2.5}$ filter samples with the EnCan-total-900 thermal method. A standard major axis regression fit is included.

A second set of scaling factors were calculated to elucidate seasonal patterns in the scaling. Scaling factors were calculated for each month of the year, using data across all years. A violin plot of these "monthly climatology" scaling factors are shown in Figure 3.21. Figure 3.22 shows a plot of the mean spatial distribution of scaling factors. Figure 3.23 shows a map of the mean a posteriori uncertainty variance.

Figure 3.24 shows model predicted concentrations of elemental carbon at Tsinghua University, created using the monthly climatology a posteriori scaled emissions, along with the measured concentrations. A standard major axis regression was computed, and is shown in Figure 3.25. The regression slope was 0.64, the intercept was 1.94, and the squared coefficient of correlation is 0.52. While the relationship is weaker and the slope is further from unity than in the case with scaling factors for each of the 36 months of the model period, the difference is small considering the much smaller number of scaling factors.

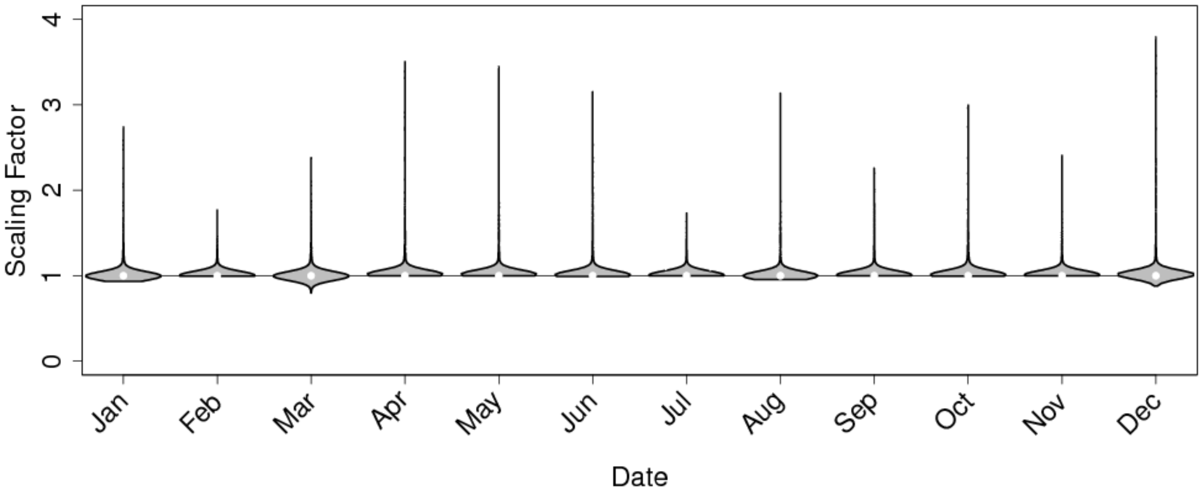


Figure 3.21: Violin plot of monthly climatology a posteriori scaling factors for elemental carbon emissions in the northern Chinese domain shown in Figure 3.2, calculated using a Bayesian inversion analysis of a STILT model of the concentration at Tsinghua University (40.00°N, 116.33°E). The width of each element is representative of the number of emissions grid cells as a function of the scaling factor.

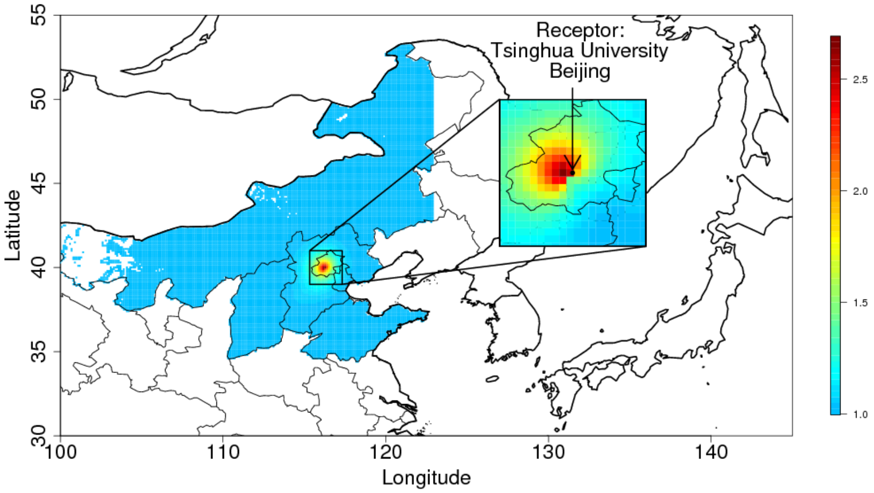

Receptor:
Tsinghua University
Beijing
Latitude
Longitude
55
50
45
40
35
30
100
110
120
130
140
2.5
2.0
1.5
1.0

Figure 3.22: Mean spatial distribution of monthly climatology a posteriori scaling factors for elemental carbon emissions in nortnern China for the period 1 January 2006 through 31 December 2008, calculated using a Bayesian inversion analysis of a STILT model of the concentration at Tsinghua University (40.00°N, 116.33°E).

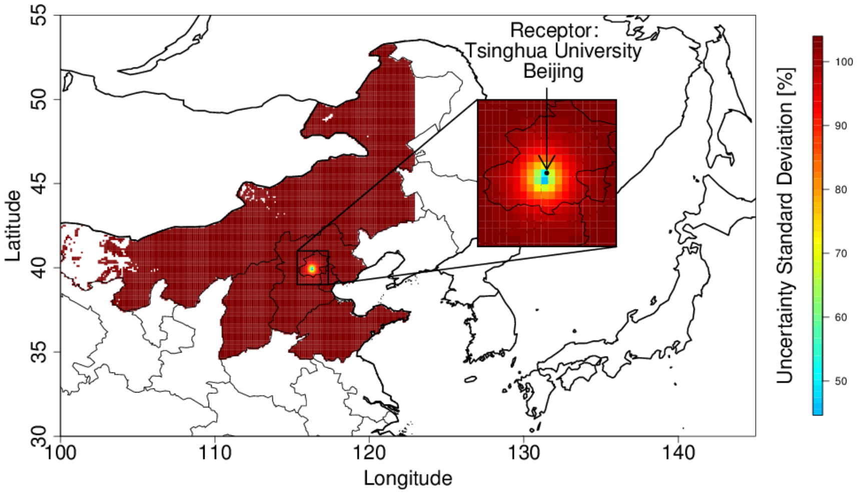

Receptor:
Tsinghua University
Beijing
Latitude
Longitude
55
50
45
40
35
30
100
110
120
130
140
Uncertainty Standard Deviation [%]
100
90
80
70
60
50

Figure 3.23: Mean spatial distribution of monthly climatology a posteriori uncertainty of elemental carbon emissions in nortnern China for the period 1 January 2006 through 31 December 2008, calculated using a Bayesian inversion analysis of a STILT model of the concentration at Tsinghua University (40.00°N, 116.33°E).

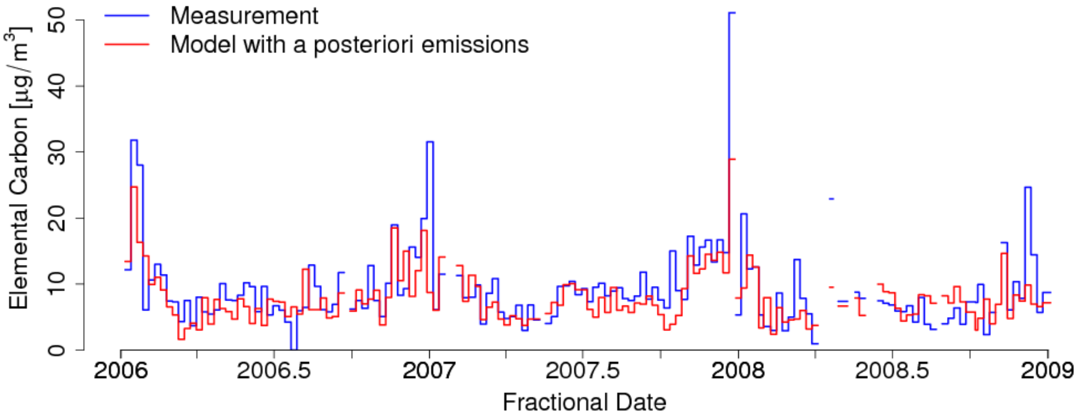


Figure 3.24: Concentrations of elemental carbon aerosol at Tsinghua University (40.00°N, 116.33°E) as predicted by a STILT model using monthly climatology a posteriori estimates of emissions calculated with a Bayesian inversion analysis (red), and as measured on weekly integrated $PM_{2.5}$ filter samples with the EnCan-total-900 thermal method (blue).

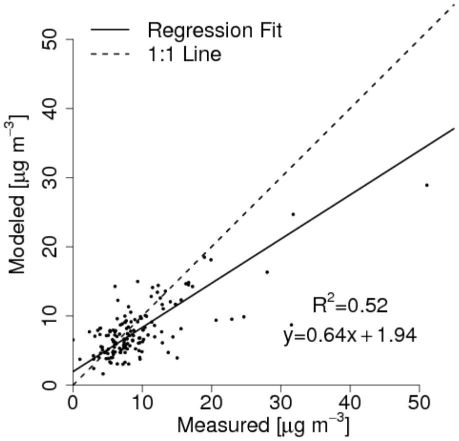


Figure 3.25: A comparison between STILT model predictions of elemental carbon concentrations at Tsinghua University (40.00°N, 116.33°E) made using monthly climatology a posteriori estimates of emissions calculated with a Bayesian inversion analysis to concentrations measured on weekly integrated $PM_{2.5}$ filter samples with the EnCan-total-900 thermal method. A standard major axis regression fit is included.

3.8 Discussion

The results of the Bayesian inversion analysis presented in Section 3.7 demonstrated that bottom-up emissions were likely underpredicted at nearly all times during the study. This is a significant underestimation in a region that includes some of the most intensive elemental carbon emissions on Earth [11]. This result could help to explain the underestimation in the atmospheric burden of black carbon in global chemical transport models [10].

There was a bimodal seasonal cycle in a posteriori scaling factors. One mode peaked in the wintertime, when residential heating by coal use increases in the domain of emissions, while other sources of emissions remain fairly constant. The emission factor for residential coal has been cited as a large uncertainty in bottom-up emissions inventories, as the relative use of anthracite versus bituminous coal, and raw coal versus briquettes have very strong influences on the emission factor, but are poorly constrained [145].

A second peak in the seasonal pattern of emissions scaling factors occurred in the early summer, when background concentrations were largest. If the GEOS-Chem model underpredicts the atmospheric burden, then an artifact will appear in the model and drive a bias in the scaling factors. This source of error was reflected in the conservative assumption of 200% uncertainty in the background concentration, but a bias would have affected the position of the center of the a posteriori emissions distribution.

There were two possible sources for the increased background concentrations observed during the early summer. Air parcel trajectories had a tendency to originate from the south during the summer, where larger anthropogenic emissions lay just outside of the domain of emissions. If GEOS-Chem used underestimated emissions, then emissions south of the domain could have increased the burden of elemental carbon at the boundary of the domain. Also, Wiedinmyer et al., 2011 [134] showed that forest fire emissions in the tropical and subtropical Northern Hemisphere peak between February and May, coinciding with the period of increased background concentrations. These emissions had the potential to be advected to the receptor, and if they were underestimated in GEOS-Chem, then they could have driven a high bias in the scaling factors. It is thus probable that background concentrations were underestimated during the summer peak, and the emission scaling was overestimated.

We speculate that the true emission scaling has a single mode that peaks in the wintertime, but is considerably above 1 for at all times. The use of residential coal has been identified as a target for particulate emissions reductions in the regions around Beijing. We suspect that residential coal likely made up a larger portion of elemental carbon emissions than had been previously believed. This means that future efforts to replace coal as a household fuel may provide an even greater efficacy than previously expected.

Wang et al., 2012 [129] predicted black carbon emissions for the period 2008–2050. They predicted that black carbon emissions changes could range from decreases of up to 53% to increases of up to 12%. Included in this prediction was an estimated 45% decrease in residential emissions. If residential emissions are initially underestimated by more than the rest of the inventory, these predicted reductions could be greater than anticipated.

Future research needs to better constrain the proportion of the inventory relating to residential coal use. Methods could include analyses similar to this study, but utilizing carbon isotope ratio measurements of elemental carbon. The contribution of various fuel types to the bottom-up inventory, along with their isotopic signatures, could be used in the model presented in this paper to predict the isotopic signature of elemental carbon samples. The model could then be inverted, providing top-down constraints on relative rates of fuel use. The method of Huang et al., 2006 [45] could provide measurements suitable for such a study.

3.9 Conclusions

In this work we have provided evidence that bottom-up emissions inventories of elemental carbon emissions in northern China are underestimated. We speculate that an apparent interference by emissions outside of the domain obfuscated results for the early summer emissions scaling. Seasonal patterns in the magnitude of the underestimate suggest that the portion of the inventory relating to the use of coal for residential heating and cooking may be underestimated by a greater degree than the rest of the inventory.

Acknowledgments

We gratefully acknowledge funding from Environment Canada for supporting J. Benmergui. This work was made possible by the facilities of the Shared Hierarchical Academic Research Computing Network (SHARCNET: www.sharcnet.ca) and Compute Canada. We thank the United States Geological Survey for providing the Land Cover Type Modeling Grid product MCD12C1. We would like to thank Dylan Millet and Lu Hu from the University of Minnesota for supplying GEOS-Chem model output that were used as boundary conditions. Special thanks to Helen Macintyre of the Massachusetts Institute of Technology for pointing out an apparent forest fire signal in the monthly climatology a posteriori scaling factors. We thank James Sloan for his many comments.

Chapter 4

Conclusions

In this thesis we have examined methods that constrain sources of atmospheric constituents and processes affecting their concentration. The Bayesian inversion method worked well, providing effective and informative constraints on the sources of elemental carbon in the important region of northern China. A new method was derived and applied to methanesulfonic acid concentrations in Alert, Nunavut. This method yielded information that would not have been available with previously existing methods. It proved to be particularly powerful for testing hypothesis about which processes dominate the variation of concentrations. Information such as this would be especially powerful for defining a direction that research should take to improve future modeling efforts. Statistical tests called into question the quality of the model estimates. The method may work better in less remote regions where sources are nearer to the receptor, and driving meteorological fields are of higher quality.

Appendix A

Description of Algorithms Used for Elemental Carbon Removal Processes

A.1 Hydrophobic-to-Hydrophilic Conversion

A.1.1 Hydrophobic-to-Hydrophilic Conversion by Condensation of Sulphate and Nitrate

Riemer et al., 2005 [93] computed the conversion timescales due to condensation of sulphate and nitrate onto carbonaceous aerosols, and of coagulation of carbonaceous aerosols onto other hydrophilic aerosols. They found that condensation was the dominant process during the daytime, and that coagulation became dominant during the night. We followed their suggestion of using a time scale for conversion by condensation during the daytime of 2 h for aerosols above 250 m above ground level and 8 h for those below 250 m. Daytime was defined as any time with a non-zero downward short-wave radiation flux. We applied an infinite time scale (no conversion) for conversion by condensation during the nighttime.

A.1.2 Hydrophobic-to-Hydrophilic Conversion by Coagulation

An adapted version of the algorithm of Riemer et al., 2005 [93] was implemented for hydrophobic-to-hydrophilic conversion by coagulation. The timescale for conversion is given by:

$$\tau_{coag} = \frac{A}{N} + \frac{B}{N^2}. \tag{A.1}$$

Where N is the number density of all aerosols in coagulating mixture, in cm^{-3}, $A = 6 \times 10^4 \, h \, cm^{-3}$ and $B = 3 \times 10^8 \, h \, cm^{-6}$ are empirical constants.

For the number density of all aerosols, it was assumed that the mixture of aerosols was constant in composition, and that the modeled elemental carbon concentration was representative of the mixture. The mean contribution of elemental carbon to total mass of PM2.5 in Beijing is about 10% [25], and so the total mass concentration of aerosol in the mixture, M, was estimated as 10 times the modeled elemental carbon.

Including the concentration of elemental carbon in this calculation introduced a non-linearity to Equation 3.1. An increase in the emissions rate will increase the concentration of elemental carbon, increasing the rate of coagulation, and thus the sensitivity to wet deposition. The deviation from linearity was assumed to be negligible. This was the only deviation from linearity with respect to emissions in **K**.

Conversion from mass concentration to number concentrations was calculated following Lesins and Lohmann, 2005 [61] and Croft, 2005 [19], assuming a log-normal size distribution with a mode radius $r_a = 0.069 \, \mu m$, and a geometric standard deviation of $\sigma_g = 1.538$. The number concentration of particles in the mixture was then:

$$N_a = \frac{\rho M}{\rho_i} \left(\frac{\pi}{6} 2 r_a^3 exp \left(\frac{9}{2} \right) \ln^2 \sigma_g \right)^{-1}. \tag{A.2}$$

Where it was assumed that the collective mass density of the aerosol mixture was $\rho_i = 1.5 \, g \, cm^{-3}$, based on the work of Pitz et al., 2003 [85]. ρ is the density of air.

A.1.3 Hydrophobic-to-Hydrophilic Conversion by Oxidation

The oxidation algorithm of Tsigaridis and Kanakidou, 2003 [123] calculates a timescale for hydrophobic-to-hydrophilic conversion by oxidation, dependent on concentrations of ozone and water vapor at the locations of the particles in time and space as:

$$\tau_{ox} = \frac{1 + K_{O_3}[O_3] + K_{H_2O}[H_2O]}{K_\infty K_{O_3}[O_3]}. \tag{A.3}$$

Where $K_\infty = 0.015\,s^{-1}$ is the pseudo-first-order decay rate coefficient in limit of high ozone concentration, and the reaction rates K_{O_3} and K_{H_2O} are:

$$K_i = \frac{S_{0_i}\omega_i\tau_i}{4SS_s}. \tag{A.4}$$

Where i denotes the species H_2O or O_3, $S_{0,H_2O} = 0.4 \times 10^{-3}$ and $S_{0,O_3} = 3.3 \times 10^{-3}$ are dimensionless sticking coefficients for water vapor and ozone, respectively, $\tau_{H_2O} = 3 \times 10^{-3}\,s$ and $\tau_{O_3} = 5\,s$ are the residence times of water vapor and ozone on a surface site, and $[SS]_s = 5.7 \times 10^{14}\,sites\,cm^{-2}$ is the surface concentration of adsorption sites.

ω_i are the mean thermal velocities, given by:

$$\omega_i = \left(\frac{8k_bT}{M_i\pi}\right)^{\frac{1}{2}}. \tag{A.5}$$

Where $k_b = 1.38 \times 10^{-23}\,J\,K^{-1}$ is Boltzmann's constant, T is the ambient temperature, $M_{H_2O} = 2.99 \times 10^{-26}\,kg$ is the mass of a single water vapour molecule, and $M_{O_3} = 7.97 \times 10^{-26}\,kg$ is the mass of a single ozone molecule.

Ozone concentrations were taken from the Model of Ozone And Related Tracers, version 4 [26], with National Centers for Environmental Protection, National Center for Atmospheric Research Reanalysis meteorology (MOZART-4/NCEP). The model output was provided on a 2.8° × 2.8° grid with 28 vertical levels and a 6 hour temporal resolution. Water vapor concentrations and temperature were taken directly from the STILT output, which interpolates them from the GDAS1 meteorology fields.

A.1.4 Consolidation of Hydrophobic-to-Hydrophilic Conversion Time Scale

It was assumed that the hydrophobic-to-hydrophilic conversion from each mechanism acts independently. If the initial concentration of hydrophobic elemental carbon is $M_{PO,i}$, then the mass, M_{conv}, converted to hydrophilic elemental carbon after a time step Δt is:

$$\begin{aligned} M_{conv} &= M_{PO,i}\left[1 - exp\left(\frac{-\Delta t}{\tau_{ox}}\right) exp\left(\frac{-\Delta t}{\tau_{cond}}\right) exp\left(\frac{-\Delta t}{\tau_{coag}}\right)\right] \\ &= M_{PO,i}\left[1 - exp\left(-\Delta t\frac{\tau_{ox}\tau_{cond} + \tau_{ox}\tau_{coag} + \tau_{cond}\tau_{coag}}{\tau_{ox}\tau_{cond}\tau_{coag}}\right)\right]. \end{aligned} \tag{A.6}$$

And so the effective time constant for hydrophobic-to-hydrophilic conversion is:

$$\tau = \frac{\tau_{ox}\tau_{cond}\tau_{coag}}{\tau_{ox}\tau_{cond} + \tau_{ox}\tau_{coag} + \tau_{cond}\tau_{coag}}. \tag{A.7}$$

A.2 Dry Deposition

The size-segregated and land use dependent dry deposition scheme of Zhang et al., 2001 [143] was adapted for use in this work. The dry deposition velocity, V_{dry} was calculated with a gravitational settling velocity V_g, and aerodynamic and surface resistances R_a and R_s as:

$$V_{dry} = V_g + \frac{1}{R_a + R_s}. \tag{A.8}$$

Where the gravitational settling velocity is given by:

$$V_g = \frac{\rho D^2 g C_c}{18\eta}. \tag{A.9}$$

Where D is the diameter of the aerosol, $g = 9.8\, m\, sec^{-2}$ is the acceleration due to gravity, $\eta = 1.81 \times 10^{-5}\, kg\, m^{-1}\, s^{-1}$ is the dynamic viscosity of air.

C_c is the Cunningham slip correction, given by:

$$C_c = 1 + 2\frac{\lambda_{mfp}}{D}\left[1.257 + 0.4exp\left(\frac{-0.55D}{\lambda_{mfp}}\right)\right]. \tag{A.10}$$

λ_{mfp} is the mean free path of air molecules, given by:

$$\lambda_{mfp} = \lambda_{stp}\left(\frac{\rho_{stp}}{\rho}\right). \tag{A.11}$$

Where $\lambda_{stp} = 6.53 \times 10^{-8}\, m$ and $\rho_{stp} = 1.2754\, kg\, m^{-3}$ are the mean free path of air molecules and mass density of air at standard temperature and pressure.

R_a is is the aerodynamic resistance:

$$R_a = \frac{ln\left(\frac{z}{Z_0}\right) - \psi_h}{\kappa u_\star}. \tag{A.12}$$

Where z is the height of the STILT particle above the ground, $u_\star$ is the friction velocity, $\kappa = 0.4$ is the von Karman constant, and the roughness length Z_0 is taken from the land use category. ψ_h is the Businger profile estimate of a dimensionless temperature gradient [12].

R_s is the vegetative resistance:

$$R_s = \frac{1}{\epsilon_0 u_\star R_1 \left(E_B + E_{IM} + E_{IN}\right)}. \tag{A.13}$$

Where $\epsilon_0 = 3$ is [143]'s empirical constant for surface resistance.

The efficiency of surface collection by Brownian diffusion, E_B, impaction E_{IM}, and interception E_{IN} are given by:

$$E_B = \left(\frac{nu}{\frac{k_b T}{3\eta\pi D}}\right)^{-\gamma_{LUC}}. \tag{A.14}$$

$$E_{IM} = \left(\frac{St}{\alpha_{LUC} + St}\right)^2. \tag{A.15}$$

$$E_{IN} = 0.5\left(\frac{D}{A_{LUC}}\right)^2. \tag{A.16}$$

With Stokes number St_{veg} for vegetated surfaces and St_{smo} for smooth surfaces are given by:

$$St_{veg} = \frac{V_g u_\star}{g A_{LUC}}. \tag{A.17}$$

$$St_{smo} = \frac{V_g u_\star^2}{\nu}. \tag{A.18}$$

and R_1 is given by:

$$R_1 = exp\left(-St^{\frac{1}{2}}\right). \tag{A.19}$$

No specific size distribution data were available, and so a typical accumulation mode size distribution was assumed, with size bins of 0.1(8.3% of particulate mass distribution), 0.5(16.7%), 1(25%), 1.5(25%), 2(16.7%) and 2.5(8.3%) μm. Densities of the particles were taken to have bins of 1(16.7%), 3(33.3%), 5(33.3%) and 7(16.7%) $g\,cm^{-3}$.

Land use categorizations from the IGBP global vegetation classification scheme in the Land Cover Type Climate Modeling Grid product (MCD12C1 Type 1) were provided by the United States Geological Survey Land Processes Distributed Active Archive Center for use in the dry deposition scheme. The theoretical basis of the algorithm is described in Strahler et al., 1999 [116].

The land use categories of the MCD12C1 product were translated into the categories of Zhang et al., 2001 [143] as described in Table A.2. The seasonal categories used in Zhang et al., 2001 [143] were applied to date ranges that roughly correspond to the seasonal climate patterns. The correspondence between dates and seasonal categories are shown in Table A.1.

Table A.1: Date correspondence for seasonal categories from Zhang et al., 2001 [143].

Date Range	Zhang et al., 2001 [143] category
May-Aug	Midsummer with lush vegetation.
Sep-Oct	Autumn with cropland that has not been harvested.
Nov	Late autumn after frost, no snow.
Dec-Feb	Winter, snow on ground and subfreezing.
Mar-Apr	Transitional spring with partially green short annuals.

Table A.2: Land Use Categories: Translating MCD12C1 to Zhang et al., 2001 [143].

Zhang	MCD12C1	% of Domain		
		2006	2007	2008
Inland water	Water	0	0	0
Evergreen needleleaf trees	Evergreen needleleaf forest	2.16	2.47	2.63
Evergreen broadleaf trees	Evergreen broadleaf forest Woody savannas	1.51	1.64	1.47
Deciduous needleleaf trees	Deciduous needleleaf forest	3.78	3.46	3.76
Deciduous broadleaf trees	Deciduous broadleaf forest	0.23	0.23	0.15
Mixed broadleleaf/needleleaf trees	Mixed forest	7.42	7.87	7.73
Shrubs and interrupted woodlands	Closed shrubland Open shrubland Savannas	3.32	2.85	2.50
Grass	Grasslands	42.90	42.92	42.98
Wet land with plants	Permanent wetlands	0.01	0.01	0.02
Crops, mixed farming	Croplands Cropland/Natural vegetation mosaic	22.52	22.04	22.55
Urban	Urban and built-up	0.32	0.34	0.34
Ice cap and glacier	Snow and ice	0	0	0
Desert	Barren or sparsely vegetated	15.79	16.15	15.85

A.3 Wet Deposition

The GEOS-Chem wet deposition scheme described by Jacob et al., 2000 [46] and Liu et al., 2001 [72] was adapted for application to this work. This parameterization includes losses due to wet convective updrafts, in cloud scavenging (rainout), and below cloud scavenging (washout). It was assumed that only hydrophilic elemental carbon was susceptible to wet deposition.

A.3.1 Convective Updraft

As air is lifted through a wet convective updraft, a fraction of the hydrophilic elemental carbon is removed by the conversion of cloud condensate to precipitation. The fraction of elemental carbon, $F_{updraft}$, removed during a single time step is given by:

$$F_{updraft} = 1 - exp\left(-\beta_{updraft}\Delta t\right). \tag{A.20}$$

Where the rate constant, $\beta_{updraft}$, is given by:

$$\beta_{updraft} = k\frac{w}{w_{updraft}}. \tag{A.21}$$

Where w is the vertical velocity, and $k = 5 \times 10^{-3}\,sec^{-1}$, is the rate constant for conversion of cloud condensate to precipitation, relative to the reference velocity, $w_{updraft} = 10\,m\,sec^{-1}$.

The GDAS1 meteorological data used to drive the STILT model is of insufficient horizontal resolution to capture convective updrafts. Vertical velocities in GDAS1 are typically on the order of $0.1\,m\,s^{-1}$. This results in an underprediction of deposition due to wet convective updrafts, and may result in slightly overpredicted concentrations of elemental carbon at the receptor.

A.3.2 Rainout

The fraction of elemental carbon removed by rainout during a time step is given by:

$$F_{rainout} = f\left[1 - exp\left(-\beta_{rainout}\Delta t\right)\right]. \tag{A.22}$$

The rate constant for removal by rainout, $\beta_{rainout}$, was assumed to be constant and equal to k.

In the GEOS-Chem deposition scheme f is the fraction of a grid cell that is experiencing precipitation. STILT particle locations are resolved at points, so f was adapted to be a probability of a STILT particle experiencing precipitation. STILT outputs the total precipitation rate P for a GDAS1 grid cell, along with precipitation rates for convective and stratiform rain, P_{strat} and P_{conv}. The contributions of each type of precipitation to the probability of precipitation at the particle point were calculated separately:

$$f = \frac{P_{strat}}{P} f_{strat} + \frac{P_{conv}}{P} f_{conv}. \tag{A.23}$$

Where f_{strat} is the fractional area coverage of stratiform rain, given by:

$$f_{strat} = \frac{CQ}{k_{strat}\left(L + W\right)}. \tag{A.24}$$

Where C is the fractional area of cloud cover in the grid box containing a particle, and $L + W$ is the condensed water content (liquid + ice) within the precipitating cloud, taken as a constant $1.5 \times 10^{-6}\, m^3\, water\, m^{-3}\, air$.

Q is the rate of precipitation formation, estimated as in Giorgi, 1986 [36]:

$$Q = C\beta_{cl} LWC. \tag{A.25}$$

Where LWC is the in-cloud liquid water content for a precipitating cloud, taken as a constant $0.0004\, kg\, m^{-3}$, and β_{cl} is the frequency of conversion of cloud water to rainwater, in units of sec^{-1}.

The rate constant for rainout loss in stratiform clouds takes the form:

$$k_{strat} = k_{min} + \frac{Q}{L + W}. \tag{A.26}$$

with a minimum value of $k_{min} = 10^{-4}\, sec^{-1}$.

f_{conv}, the fractional area coverage of convective rain, is given by:

$$f_{conv} = \frac{0.3CQ}{Q + 0.3k\left(L + W\right)}. \tag{A.27}$$

A.3.3 Washout

The fraction of elemental carbon removed by washout during a time step is given by:

$$F_{washout} = f \left[1 - exp \left(-\beta_{washout} \Delta t \right) \right]. \tag{A.28}$$

Where the rate constant for removal by washout is given by:

$$\beta_{washout} = \frac{k_{washout} P}{C}. \tag{A.29}$$

Where $k_{washout} = 0.01\, m^{-1}$ is a constant relating precipitation rate to washout rate.

The total fraction of elemental carbon aerosol removed by wet deposition during a time step is then:

$$F_{wet} = F_{updraft} + F_{rainout} + F_{washout}. \tag{A.30}$$

Copyright Permissions

"Quantitative attribution of processes affecting atmospheric chemical concentrations by combining a time-reversed Lagrangian particle dispersion model and a regression approach" was reproduced by permission of American Geophysical Union. The written permission was given via email from Michael Connolly (MConnolly@agu.org) and Maria Lindgren (MLindgren@agu.org) of the American Geophysical Union, as follows:

> We are pleased to grant permission for the use of the material requested for inclusion in your thesis. The following non-exclusive rights are granted to AGU authors:
>
> - All proprietary rights other than copyright (such as patent rights).
> - The right to present the material orally.
> - The right to reproduce figures, tables, and extracts, appropriately cited.
> - The right to make hard paper copies of all or part of the paper for classroom use.
> - The right to deny subsequent commercial use of the paper.
>
> Further reproduction or distribution is not permitted beyond that stipulated. The copyright credit line should appear on the first page of the article or book chapter. The following must also be included, "Reproduced by permission of American Geophysical Union" To ensure that credit is given to the original source(s) and that authors receive full credit through appropriate citation to their papers, we recommend that the full bibliographic reference be cited in the reference list. The standard credit line for journal articles is: "Author(s), title of work, publication title, volume number, issue number, citation number (or page number(s) prior to 2002), year. Copyright [year] American Geophysical Union."

If an article was placed in the public domain, in which case the words “Not subject to U.S. copyright” appear on the bottom of the first page or screen of the article, please substitute “published” for the word “copyright” in the credit line mentioned above.

References

[1] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, AC-19(6):716–723, 1974.

[2] H. Akaike. Bayesian statistics. In J. M. Bernardo, editor, *Bayesian Statistics*, chapter Likelihood and the Bayes procedure, pages 143–166. University Press, Valencia, 1980.

[3] H. Akaike. Factor analysis and AIC. *Psychometrika*, 52(3):317–332, 1987.

[4] D. Arnold, A. Vargas, A. Vermeulen, B. Verheggen, and P. Siebert. Analysis of radon origin by backward atmospheric transport modeling. *Atmos. Environ.*, 44(4):494–502, 2010.

[5] C. Arsene, I. Barnes, and K. H. Becker. FT-IR product study of the photo-oxidation of dimethyl sulfide: Temperature and o_2 partial pressure dependence. *Phys. Chem. Chem. Phys.*, 1:5463–5470, 1999.

[6] L. L. Ashbaugh, W. C. Malm, and W. Z. Sadeh. A residence time probability analysis of sulfer concentrations at Grand Canyon National Park. *Atmos. Environ.*, 19(8):1263–1270, 1985.

[7] R. Atkinson, D. L. Baulch, R. A. Cox, J. N. Crowley, R. F. Hampson, R. G. Hynes, M. E. Jenkin, M. J. Rossi, and J. Troe. Evaluated kinetic and photochemical data for atmospheric chemistry: Volume i - gas phase reactions of o_x, ho_x, no_x and so_x species. *Atmos. Chem. Phys.*, 4(6):1461–1738, 2004.

[8] J. Benmergui, S. Sharma, D. Wen, and J. C. Lin. Quantitative attribution of processes affecting atmospheric chemical concentrations by combining a time-reversed Lagrangian particle dispersion model and a regression approach. *Geophys. Monogr. Ser.*, Lagrangian Modeling of the Atmosphere:251–264, 2012.

[9] I. Bey, D. J. Jacob, R. M. Yantosca, J. A. Logan, B. D. Field, A. M. Fiore, Q. Li, H. Y. Liu, L. J. Mickley, and M. G. Schultz. Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. *J. Geophys. Res.*, 106(D19):23073–23096, 2001.

[10] T. C. Bond, S. J. Doherty, D. W. Fahey, P. M. Forster, T. Berntsen, B. J. DeAngelo, M. G. Flanner, S. Ghan, B. Krcher, D. Koch, S. Kinne, Y. Kondo, P. K. Quinn, M. C. Sarofim, M. G. Schultz, M. Schulz, C. Venkataraman, H. Zhang, S. Zhang, N. Bellouin, S. K. Guttikunda, P. K. Hopke, M. Z. Jacobson, J. W. Kaiser, Z. Klimont, U. Lohmann, J. P. Schwarz, D. Shindell, T. Storelvmo, S. G. Warren, and C. S. Zender. Bounding the role of black carbon in the climate system: A scientific assessment. *J. Geophys. Res.*, 2013.

[11] T. C. Bond, D. G. Streets, K. F. Yarber, S. M. Nelson, J.-H. Woo, and Z. Kilmont. A technology-based global inventory of black and organic carbon emissions from combustion. *J. Geophys. Res.*, 109(D14), 2004.

[12] J. A. Businger, J. C. Wyngaard, Y. Izumi, and E. F. Bradley. Flux-profile relationship in the atmospheric surface layer. *J. Atmos. Sci.*, 28(2):181–189, 1971.

[13] D. Cavalieri, C. Parkinson, P. Gloersen, and H. J. Zwally. Sea ice concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS passive microwave data, [January 1980 - December 2007]. Version 1.0, 1996, updated yearly.

[14] T. W. Chan, L. Huang, W. R. Leaitch, S. Sharma, J. R. Brook, J. G. Slowik, J. P. D. Abbatt, P. C. Brickell, J. Liggio, S.-M. Li, and H. Moosmüller. Observations of OM/OC and specific attenuation coefficients (SAC) in ambient fine PM at a rural site in central Ontario, Canada. *Atmos. Chem. Phys.*, 10(5):2392–2411, 2010.

[15] R. J. Charlson, J. E. Lovelock, M. O. Andreae, and S. G. Warren. Oceanic phytoplankton, atmospheric sulphur, cloud albedo and climate. *Nature*, 326:655–661, 1987.

[16] M. Chin, R. B. Rood, S.-J. Lin, J.-F. Muller, and A. M. Thompson. Atmospheric sulfur cycle simulated in the global model GOCART: Model description and global properties. *J. Geophys. Res.*, 105(D20):24671–24687, 2000.

[17] S. H. Chung and J. H. Seinfeld. Global distribution and climate forcing of carbonaceous aerosols. *J. Geophys. Res.*, 107(D19):4407, 2002.

[18] W. F. Cooke, V. Ramaswamy, and P. Kasibhatla. A general circulation model study of the global carbonaceous aerosol distribution. *J. Geophys. Res.*, 107(D16), 2002.

[19] B. Croft, U. Lohmann, and K. von Salzen. Black carbon ageing in the Canadian Centre for Climate modelling and analysis atmospheric general circulation model. *Atmos. Chem. Phys.*, 5(7):1931–1949, 2005.

[20] W. B. DeMore. Experimental and estimated rate constants for the reactions of hydroxyl radicals with several halocarbons. *J. Phys. Chem.*, 100:5813–5820, 1996.

[21] R Development Core Team. *R: A language and environment for statistical computing, reference index version 2.10.0.* R Foundation for Statistical Computing, Vienna, Austria, 2005.

[22] S. Dhanya, K. K. Pushpa, and P. D. Naik. Measurement of kinetic parameters of gas-phase reactions relevant to atmospheric chemistry. *Current Sci.*, 102(3):452–459, 2012.

[23] R. R. Draxler and G. D. Hess. An overview of the HYSPLIT 4 modeling system of trajectories, dispersion, and deposition. *Aust. Meteor. Mag.*, 47:295–308, 1998.

[24] R. R. Draxler, J. T. McQueen, and B. J. B. Stunder. An evaluation of air pollutant exposures due to the 1991 Kuwait oil fires using a Lagrangian model. *Atmos. Environ.*, 28:2197–2210, 1994.

[25] F. K. Duan, K. B. He, Y. L. Ma, F. M. Yang, X. C. Yu, S. H. Cadle, T. Chan, and P. A. Mulawa. Concentration and chemical characteristics of $PM_{2.5}$ in Beijing, China: 20012002. *Sci. Tot. Env.*, 335:264–275, 2006.

[26] L. K. Emmons, S. Walters, P. G. Hess, J.-F. Lamarque, G. G. Pfister, D. Fillmore, C. Granier, A. Guenther, D. Kinnison, T. Laepple, J. Orlando, X. Tie, G. Tyndall, C. Wiedinmyer, S. L. Baughcum, and S. Kloster. Description and evaluation of the Model for Ozone and Related chemical Tracers, version 4 (MOZART-4). *Geosci. Model Dev.*, 3:43–67, 2010.

[27] X. Fan, J. E. Walsh, and J. Krieger. A one year experimental Arctic reanalysis and comparisons with ERA-40 and NCEP/NCAR reanalyses. *Geophys. Res. Lett.*, 35(L19811), 2008.

[28] T. K. Flesch, J. D. Wilson, L. A. Harper, R. W. Todd, and N. A. Cole. Determining ammonia emissions from a cattle feedlot with an inverse dispersion technique. *Agr. Forest Meteorol.*, 144:139–155, 2007.

[29] T. K. Flesch, J. D. Wilson, and E. Yee. Backward-time Lagrangian stochastic dispersion models and their application to estimate gaseous emissions. *J. Appl. Meteorol.*, 34:1320–1332, 1995.

[30] H. Florig, G. Sun, and G. Song. Evolution of particulate regulation in China—prospects and challenges of of exposure-based control. *Chemosphere*, 49(9):1163–1174, 2002.

[31] C. J. Gaston, H. Furutani, S. A. Guazzotti, K. R. Coffee, T. S. Bates, P. K. Quinn, L. I. Aluwihare, B. G. Mitchell, and K. A. Prather. Unique ocean-derived particles serve as a proxy for changes in ocean chemistry. *J. Geophys. Res.*, 116(D18), 2011.

[32] C. Gerbig, S. Körner, and J. C. Lin. Vertical mixing in atmospheric tracer transport models: error characterization and propagation. *Atmos. Chem. Phys.*, 8(3):591–602, 2008.

[33] C. Gerbig, J. C. Lin, J. W. Munger, and S. C. Wofsy. What can tracer observations in the continental boundary layer tell us about surface-atmosphere fluxes? *Atmos. Chem. Phys.*, 6(2):539–554, 2006.

[34] C. Gerbig, J. C. Lin, S. C. Wofsy, B. C. Daube, A. E. Andrews, B. B. Stephens, P. S. Bakwin, and C. A. Grainger. Towards constraining regional scale fluxes of CO_2 with atmospheric observations over a continent: 1. observed spatial variability from airborne platforms. *J. Geophys. Res.*, 108(D24), 2003.

[35] C. Gerbig, J. C. Lin, S. C. Wofsy, B. C. Daube, A. E. Andrews, B. B. Stephens, P. S. Bakwin, and C. A. Grainger. Towards constraining regional scale fluxes of CO_2 with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor oriented framework. *J. Geophys. Res.*, 108(D24), 2003.

[36] F. Giorgi and L. Chameides. Rainout lifetimes of highly soluble aerosols and gases as inferred from simulations with a general circulation model. *J. Geophys. Res.*, 91(D13):14367–14376, 1986.

[37] M. Gosselin, M. Levasseur, P. A. Wheeler, R. A. Horner, and B. C. Booth. New measurements of phytoplankton and ice algal production in the Arctic Ocean. *Deep-Sea Res. II*, 44(8):1623–1644, 1997.

[38] P. Haertel. A Lagrangian method for simulating geophysical fluids. *Geophys. Monogr. Ser.*, Lagrangian Modeling of the Atmosphere:85–98, 2012.

[39] A. Hakami, D. K. Hanze, J. H. Seinfeld, T. Chai, Y. Tang, G. R. Carmichael, and A. Sandu. Adjoint inverse modeling of black carbon during the Asian Pacific Regional Aerosol Characterization Experiment. *J. Geophys. Res.*, 110:D14301, 2005.

[40] M. E. Hansson and E. S. Saltzman. The first Greenland ice core record of methanesulfonate and sulfate over a full glacial cycle. *Geophys. Res. Lett.*, 20(12):1163–1166, 1993.

[41] S. Hatakeyama, M. Okuda, and H. Akimoto. Formation of sulfur dioxide and methanesulfonic acid in the photooxidation of dimethyl sulfide in the air. *Geophysical Res. Lett.*, 9(5):583–586, 1982.

[42] J. M. Haywood and V. Ramaswamy. Global sensitivity studies of the direct radiative forcing due to anthropogenic sulfate and black carbon aerosols. *J. Geophys. Res.*, 103(D6):6043–6058, 1998.

[43] J. D. Hegarty, R. Draxler, A. F. Stein, J. Brioude, J. Eluszkiewicz, M. Mountain, T. Nehrkorn, and A. E. Andrews. Lagrangian particle dispersion model intercomparison and evaluation utilizing measurements from controlled tracer release experiments. In *Lagrangian Particle Dispersion Model Intercomparison and Evaluation Utilizing Measurements from Controlled Tracer Release Experiments*, volume A53A-0130 of *AGU Fall Meeting*. American Geophysical Union, 2012.

[44] L. W. Horowitz, S. Walters, D. L. Mauzerall, L. K. Emmons, P. J. Rasch, C. Granier, X. Tie, J-F. Lamarque, M. G. Schultz, G. S. Tyndall, J. J. Orlando, and G. P. Brasseur. A global simulation of tropospheric ozone and related tracers: Description and evaluation of MOZART, version 2. *J. Geophys. Res.*, 108(D24), 2003.

[45] L. Huang, J. R. Brook, W. Zhang, S. M. Li, L. Graham, D. Ernst, A. Chivulescu, , and G. Lu. Stable isotope measurements of carbon fractions (OC/EC) in airborne particulate: A new dimension for source characterization and apportionment. *Atmos. Env.*, 40(15):2690–2705, 2006.

[46] D. J. Jacob, H. Liu, C. Mari, and R. M. Yantosca. Harvard wet deposition scheme for GMI.

[47] M. Z. Jacobson. Strong radiative heating due to the mixing state of black carbon in atmospheric aerosols. *Nature*, 409:695–697, 2001.

[48] N. A. H. Janssen, M. E. Gerlofs-Nijland, T. Lanki, R. O. Salonen, F. Cassee, G. Hoek, P. Fischer, B. Brunekreef, and M. Krzyzanowski. Health effects of black carbon. In

Rosemary Bohr, editor, *Health effects of black carbon.* World Health Organization, Copenhagen, Denmark, 2012.

[49] J. W. Kaiser, A. Heil, M. O. Andreae, A. Benedetti, N. Chubarova, L. Jones, J.-J. Morcrette, M. Razinger, M. G. Schultz, M. Suttie, and G. R. van der Werf. Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power. *Biogeosciences*, 9(1):527–554, 2012.

[50] E. Kalnay, M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woolen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Jonawiak, K. C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, R. Jenne, and D. Joseph. The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, 77:437–470, 1996.

[51] T. Kaminski, P. J. Rayner, M. Heimann, and I. G. Enting. On aggregation errors in atmospheric transport inversions. *J. Geophys. Res.*, 106(D5):4703–4715, 2001.

[52] A. Keats, E. Yee, and F-S. Lien. Bayesian inference for source determination with applications to a complex urban environment. *Atmos. Environ.*, 41(3):465–479, 2007.

[53] C. A. Keller, M. Hill, M. K. Vollmer, S. Henne, D. Brunner, S. Reimann, S. O'Doherty, J. Arduini, M. Maione, Z. Ferenczi, L. Haszpra, A. J. Manning, and T. Peter. European emissions of halogenated greenhouse gases inferred from atmospheric measurements. *Env. Sci. Tech.*, 46(1):217–225, 2012.

[54] V-M. Kerminen, M. Aurela, R. E. Hillamo, and A. Virkkula. Formation of particulate MSA: deductions from size distribution measurements in the Finnish Arctic. *Tellus B*, 49(2):159–171, 1997.

[55] P. K. Kitanidis. *Introduction to geostatistics: applications to hydrology.* Cambridge University Press, 1997.

[56] D. Koch. Transport and direct radiative forcing of carbonaceous and sulfate aerosols in the GISS GCM. *J. Geophys. Res.*, 106(D17):20311–20322, 2001.

[57] D. Koch, M. Schulz, S. Kinne, C. McNaughton, J. R. Spackman, Y. Balkanski, S. Bauer, T. Berntsen, T. C. Bond, O. Boucher, M. Chin, A. Clarke, N. De Luca, F. Dentener, T. Diehl, O. Dubovik, R. Easter, D. W. Fahey, J. Feichter, D. Fillmore, S. Freitag, S. Ghan, P. Ginoux, S. Gong, L. Horowitz, T. Iversen, A. Kirkevag, Z. Klimont, Y. Kondo, M. Krol, X. Liu, R. Miller, V. Montanaro, N. Moteki, G. Myhre, J. E. Penner, J. Perlwitz, G. Pitari, S. Reddy, L. Sahu, H. Sakamoto,

G. Schuster, J. P. Schwarz, . Seland, P. Stier, N. Takegawa, T. Takemura, C. Textor, J. A. van Aardenne, and Y. Zhao. Evaluation of black carbon estimations in global aerosol models. *Atmos. Chem. Phys.*, 9(22):9001–9026, 2009.

[58] P. Langevin. Sur la théorie du mouvement brownien. *C. R. Hebd. Seances Acad. Sci.*, 146:530–533, 1980.

[59] K. S. Law, P. H. Plantevin, D. E. Shallcross, H. J. Rogers, J. A. Pyle, C. Grouheland V. Thouret, and A. Marenco. Evaluation of modeled O_3 using Measurement of Ozone by Airbus In-Service Aircraft (MOZAIC) data. *J. Geophys. Res.*, 103(D19):25721–25737, 1998.

[60] Y. Lei, Q. Zhang, K. B. He, and D. G. Streets. Primary anthropogenic aerosol emission trends for China, 19902005. *Atmos. Chem. Phys.*, 11(3):931–954, 2011.

[61] G. Lesins and U. Lohmann. Using MODIS and AERONET to determine GCM aerosol size. *J. Atmos. Sci.*, 63(4):1338–1347, 2005.

[62] H. Levy, H. M. D. Schwarzkopf, L. Horowitz, V. Ramaswarmy, and K. L. Findell. Strong sensitivity of late 21st century climate to projected changes in short-lived air pollutants. *J. Geophys. Res.*, 113(D6), 2008.

[63] S. M. Li, L. A. Barrie, R. W. Talbot, R .C. Harriss, C. I. Davidson, and J.-L. Jaffrezo. Seasonal and geographic variations of methanesulfonic acid in the arctic troposphere. *Atmos. Environ. A: General Topics*, 27(17–18):3011–3024, 1993.

[64] J. C. Lin. Lagrangian modeling of the atmosphere: an introduction. *Geophys. Monogr. Ser.*, Lagrangian Modeling of the Atmosphere:1–11, 2012.

[65] J. C. Lin, D. Brunner, and C. Gerbig. Studying atmospheric transport through Lagrangian models. *Eos Trans. AGU*, 92(21):177–184, 2011.

[66] J. C. Lin, D. Brunner, C. Gerbig, A. Stohl, A. Luhar, and P. Webley, editors. *Lagrangian modeling of the atmosphere.* Number 200 in Geophys. Monogr. Ser. AGU, Washington, D. C., 2012.

[67] J. C. Lin and C. Gerbig. Accounting for the effect of transport errors on tracer inversions. *Geophys. Res. Lett.*, 32(1), 2005.

[68] J. C. Lin, C. Gerbig, S. C. Wofsy, A. E. Andrews, B. C. Daube, K. J. Davis, and C. A. Grainger. A near-field tool for simulating the upstream influence of atmospheric

observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model. *J. Geophys. Res.*, 108(D16), 2003.

[69] J. C. Lin, C. Gerbig, S. C. Wofsy, A. E. Andrews, B. C. Daube, C. A. Grainger, B. B. Stephens, P. S. Bakwin, and D. Y. Hollinger. Measuring fluxes of trace gases at regional scales by Lagrangian observations: Application to the CO_2 Budget and Rectification Airborne (COBRA) study. *109*, D15, 2004.

[70] P. S. Liss and L. Merlivat. Air-sea gas exchange rates: introduction and synthesis. In P. Buat-Ménard, editor, *The Role of Air-Sea Exchange in Geochemical Cycling*, volume 185 of *Nato Science Series C*, pages 113–127, Dordrecht, Holland, 1986. D.Reidel Publishing Company. Proceedings of the NATO Advanced Study Institute, Bombannes.

[71] D. Liu, J. Allen, J. Whitehead, D. Young, M. Flynn, H. Coe, G. McFiggans, Z. L. Fleming, and B. Bandy. Ambient black carbon particle properties controlled by mixing state and composition. *Atmos. Chem. Phys.*, 13(4):2015–2029, 2013.

[72] H. Liu, D. J. Jacob, I. Bey, and R. Yantosca. Constraints from $^($210)pb and $^($7)be on wet deposition and transport in a global three-dimensional chemical tracer model driven by assimilated meteorological fields. *J. Geophys. Res.*, 106:12,109–12,128, D11.

[73] U. Lohmann, J. Feichter, C. C. Chuang, and J. E. Penner. Prediction of the number of cloud droplets in the ECHAM GCM. *J. Geophys. Res.*, 104(D8):9169–9198, 1999.

[74] A. Longhetto, S. Ferrarese, C. Cassardo, C. Giraud, F. Apadula, P. Bacci, P. Bonelli, and A. Marzorati. Relationships between atmospheric circulation patterns and CO_2 greenhouse-gas concentration levels in the alpine troposphere. *Adv. Atmos. Sci.*, 14(3):309–322, 1997.

[75] P. McCullagh and J. A. Nelder. Generalized linear models. In D. R. Cox, D. V. Hinkley, N. Reid, D. B. Rubin, and B. W. Silverman, editors, *Generalized Linear Models*, Monographs on Statistics and Applied Probability. Chapman and Hall/CRC, Boca Raton, Florida, USA, second edition, 1989.

[76] W. Meier, F. Fetterer, K. Knowles, M. Savoie, and M. J. Brodzik. Sea ice concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS passive microwave data, [January 1980 - December 2007]. Version 1.0, 2006.

[77] J. B. Milford and C. L. Davidson. The sizes of particulate trace elements in the atmosphere – a review. *J. Air Pollut. Control Assoc.*, 35(12):1249–1260, 1985.

[78] S. Netcheva, D. Toom-Sauntry, S. Li, and J. W. Bottenheim. Is the spring maximum of MSA (Methane Sulfonic Acid) in the Arctic due to oxidation of DMS (DiMethyl Sulphide) by BrO? In *Is the spring maximum of MSA (Methane Sulfonic Acid) in the Arctic due to oxidation of DMS (DiMethyl Sulphide) by BrO?*, volume abstract A42B-07, 2007.

[79] A. M. Obukhov. Description of turbulence in terms of Lagrangian variables. *Adv. Geophys.*, 6:113–116, 1959.

[80] J. O'Dwyer, E. Isaksson, T. Vinje, T. Jauhiainen, J. Moore, V. Pohjola, R. Vaikmäe, and R. S. W. van de Wal. Methanesulfonic acid in a Svalbard ice core as an indicator of ocean climate. *Geophys. Res. Lett.*, 27(8), 2000.

[81] M. Perrette, A. Yool, G. D. Quartly, and E. E. Popova. Near-ubiquity of ice-edge blooms in the Arctic. *Biogeosci.*, 8:515–524, 2011.

[82] R. A. Pielke. *Mesoscale meteorological modeling.* Academic Press, San Diego, second edition, 2002.

[83] D. Pillai, C. Gerbig, R. Kretschmer, V. Beck, U. Karstens, B. Neininger, and M. Heimann. Comparing Lagrangian and Eulerian models for CO_2 transport — a step towards Bayesian inverse modeling using WRF/STILT-VPRM. *Atmospheric Chemistry and Physics Discussions*, 12(1):1267–1298, 2012.

[84] C. A. Pio, M. A. Cerqueira, L. M. Castro, and M. L. Salgueiro. Sulphur and nitrogen compounds in variable marine/continental air masses at the southwest European coast. *Atmos. Env.*, 30(18):3115–3127, 1996.

[85] M. Pitz, J. Cyrys, E. Karg, A. Wiedensohler, H.-E. Wichmann, and J. Heinrich. Variability of apparent particle density of an urban aerosol. *Env. Sci. Tech.*, 37(19):4336–4342, 2003.

[86] P. K. Quinn and T. S. Bates. The case against climate regulation via oceanic phytoplankton sulphur emissions. *Nature*, 480:51–56, 2011.

[87] P. K. Quinn, T. S. Bates, D. J. Coffman, T. L. Miller, J. E. Johnson, D. S. Covert, J. P. Putaud, C. Neususs, and T. Novakov. A comparison of aerosol chemical and optical properties from the 1st and 2nd Aerosol Characterization Experiments. *Tellus*, 52B:239–257, 2000.

[88] V. Ramanathan and G. Carmichael. Global and regional climate changes due to black carbon. *Nature Geoscience*, 1:221–227, 2008.

[89] K. A. Read, A. C. Lewis, S. Bauguitte, A. M. Rankin, R. A. Salmon, E. W. Wolff, A. Saiz-Lopez, W. J. Bloss, D. E. Heard, J. D. Lee, and J. M. C. Plane. DMS and MSA measurements in the Antarctic Boundary Layer: impact of BrO on MSA production. *Atmos. Chem. Phys.*, 8(11):2985–2997, 2008.

[90] P. Reisinger, A. Wonaschultz, R. Hitzenberger, A. Petzold, H. Bauer, and N. Jankowski. Intercomparison of measurement techniques for black or elemental carbon under urban background conditions in wintertime: influence of biomass combustion. *Environ. Sci. Tech.*, 42(3):884–889, 2008.

[91] R. W. Reynolds, N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang. An improved in situ and satellite SST analysis for climate. *J. Climate*, 15:1609–1625, 2002.

[92] P. J. Ribiero, Jr. and P. J. Diggle. geoR: A package for geostatistical analysis. *R-NEWS*, 1(2), 2005.

[93] N. Riemer, H. Vogel, and B. Vogel. Soot aging time scales in polluted regions during day and night. *Atmos. Chem. Phys.*, 4(7):1885–1893, 2005.

[94] L. Robertson and J. Langer. Source function estimate by means of variational data assimilation applied to the ETEX-I tracer experiment. *Atmos. Environ.*, 32(24):4219–4225, 1998.

[95] H. C. Rodean. Stochastic Lagrangian models of turbulent diffusion. In Donald R. Johnson, editor, *Stochastic Lagrangian models of turbulent diffusion*, volume 26 of *Meteorological Monographs*. American Meteorological Society, Boston, 1996.

[96] C. D. Rodgers. *Inverse Methods for Atmospheric Sounding: Theory and Practice*. World Scientific Publication Company Incorporated, 2000.

[97] R. B. Rood. Numerical advection algorithms and their role in atmospheric transport and chemistry models. *Rev. Geophys.*, 25(1):71–100, 1987.

[98] E. Sakshaug and H. Skjoldal. Life at the ice edge. *Ambio. Stockholm*, 18(1):60–67, 1989.

[99] E. S. Saltzman, D. L. Savoie, R. G. Zika, and J. M. Prospero. Methane sulfonic acid in the marine atmosphere. *J. Geophys. Res.*, 88(C15):10897–10902, 1983.

[100] E. S. Saltzman, P-Y. Whung, and P. A. Mayewski. Methanesulfonate in the Greenland ice sheet project 2 ice core. *J. Geophys. Res.*, 102(C12):26649–26658, 1997.

[101] M. Sato, J. Hansen, D. Koch, A. Lacis, R. Ruedy, O. Dubovik, B. Holben, M. Chin, and T. Novakov. Global atmospheric black carbon inferred from AERONET. *Proc. Natl. Acad. Sci. U.S.A.*, 100(11):6319–6324, 2003.

[102] B. L. Sawford. Lagrangian statistical simulation of concentration mean and fluctuation fields. *J. Clim. Appl. Meteorol.*, 24:1152–1166, 1985.

[103] W. H. Schroeder, K. G. Anlauf, L. A. Barrie, J. Y. Lu, and A. Steffen. Arctic springtime depletion of mercury. *Nature*, 394:331–332, 1998.

[104] N. Schutgens, M. Nakata, and T. Nakajima. Estimating aerosol emissions by assimilating remote sensing observations into a global transport model. *Remote Sensing*, 4(11):3528–3543, 2012.

[105] J. P. Schwartz, J. R. Spackman, R. S. Gao, L. A. Watts, P. Stier, M. Schultz, S. M. Davis, S. C. Wofsy, and D. W. Fahey. Global-scale black carbon profiles observed in the remote atmosphere and compared to models. *Geophys. Res. Lett.*, 37(L18812), 2010.

[106] P. J. Sellers, D. A. Randall, G. J. Collatz, J. A. Berry, C. B. Field, D. A. Dazlich, C. Zhang, G. D. Collelo, and L. Bounoua. A revised land surface parameterization (SiB2) for atmospheric GCMs. Part 1: Model formulation. *J. Climate*, 9:676–705, 1996.

[107] S. Sharma, E. Chan, M. Ishizawa, D. Toom-Sauntry, S. Gong, S.-M. Li, D. W. Tarasick, R. Leaitch, A.-L. Norman, P. K. Quinn, T. S. Bates, M. Levasseur, L. Barrie, and W. Maenhaut. Influence of transport and ocean ice extent on biogenic aerosol sulfur in the Arctic atmosphere. *J. Geophys. Res.*, 117(D12), 2012.

[108] W. R. Simpson, D. Carlson, G. Honninger, T. A. Douglas, M. Sturm, D. Perovich, and U. Platt. First-year sea-ice contact predicts bromine monoxide (BrO) levels at Barrow, Alaska better than potential frost flower contact. *Atmos. Chem. Phys.*, 7(3):621–627, 2007.

[109] S. B. Sneed, P. A. Mayewski, and D. A. Dixon. An emerging technique: multi-ice-core multi-parameter correlations with Antarctic sea-ice extent. *Ann. Glaciol.*, 52(57):347–354, 2011.

[110] S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor, and H.L. Miller, editors. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

[111] P. Stier, J. Feichter, S. Kinne, S. Kloster, E. Vignati, J. Wilson, L. Ganzeveld, I. Tegen, M. Werner, Y. Balkanski, M. Schulz, O. Boucher, A. Minikin, and A. Petzold. The aerosol-climate model ECHAM5-HAM. *Atmos. Chem. Phys.*, 5(4):1125–1156, 2005.

[112] A. Stohl, C. Foster, A. Frank, P. Seibert, and G. Wotawa. Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2. *Atmos. Chem. Phys.*, 5(9):2461–2474, 2005.

[113] A. Stohl, M. Hittenberger, and G. Wotawa. Validation of the Lagrangian particle dispersion model FLEXPART against large-scale tracer experiment data. *Atmos. Environ.*, 32(24):4245–4264, 1998.

[114] A. Stohl, P. Seibert, J. Arduini, S. Eckhardt, P. Fraser, B. R. Greally, C. Lunder, M. Maione, Muhle J, S. O'Doherty, R. G. Prinn, S. Reimann, T. Saito, N. Schmidbauer, P. G. Simmonds, M. K. Vollmer, R. F. Weiss, and Y. Yokouchi. An analytical inversion method for determining regional and global emissions of greenhouse gases: Sensitivity studies and application to halocarbons. *Atmos. Chem. Phys.*, 9(5):1597–1620, 2009.

[115] A. Stohl, P. Seibert, G. Wotawa, D. Arnold, J. F. Burkhart, S. Eckhardt, C. Tapia, A. Vargas, and T. J. Yasunari. Xenon-133 and caesium-137 releases into the atmosphere from the Fukushima Dai-ichi nuclear power plant: Determination of the source term, atmospheric dispersion, and deposition. *Atmos. Chem. Phys.*, 12(5):2313–2343, 2012.

[116] A. Strahler, D. Muchoney, J. Borak, M. Friedl, S. Gopal, E. Lambin, and A. Moody. MODIS land cover product Algorithm Theoretical Basis Document (atbd) version 5.0: MODIS land cover and land-cover change. Technical report, Boston University, 1999.

[117] D. G. Streets, T. C. Bond, G. R. Carmichael, S. D. Fernandes, Q. Fu, D. He, Z. Klimont, S. M. Nelson, N. Y. Tsai, M. Q. Wang, J.-H. Woo, and K. F. Yarber.

An inventory of gaseous and primary aerosol emissions in Asia in the year 2000. *J. Geophys. Res.*, 108(D21), 2003.

[118] B. Sundelin, A.-K. Eriksson, G. Lithner, and O. Gustafsson. Evaluation of the role of black carbon in attenuating bioaccumulation of polycyclic aromatic hydrocarbons from field-contaminated sediments. *Environ. Tox. Chem.*, 23(11):2611–2617, 2004.

[119] G. I. Taylor. Diffusion by continuous movements. *Proc. London Math. Soc.*, s2–20:196–212, 1921.

[120] R. L. Thompson, C. Gerbig, and C. Rodenbeck. Bayesian inversion estimate of N_2O emissions for western and central europe and the assessment of aggregation errors. *Atmos. Chem. Phys.*, 11(7):3443–3458, 2004.

[121] D. J. Thomson. Criteria for the selection of stochastic models of particle trajectories in turbulent flows. *J. Fluid Mech.*, 180:529–556, 1987.

[122] D. J. Thomson and J. D. Wilson. History of the Lagrangian stochastic model for turbulent dispersion. *Geophys. Monogr. Ser.*, Lagrangian Modeling of the Atmosphere:19–36, 2012.

[123] K. Tsigaridis and M. Kanakidou. Global modelling of secondary organic aerosol in the troposphere: a sensitivity analysis. *Atmos. Chem. Phys.*, 3(5):1849–1869, 2003.

[124] M. Uliasz and R. A. Pielke. Computer techniques in environmental studies iii. In P. Zannetti, editor, *Computer Techniques in Environmental Studies III*, Computational Mechanics Publications, chapter Receptor-oriented Lagrangian-Eulerian model of mesoscale air pollution dispersion, pages 57–68. Springer–Verlag, Berlin, 1990. Envirosoft 88-2nd International Conference, Porto Carras, Greece, September 1988.

[125] S. M. Uppala, P. W. Kllberg, A. J. Simmons, U. Andrae, V. Da Costa Bechtold, M. Fiorino, J. K. Gibson, J. Haseler, A. Hernandez, G. A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R. P. Allan, E. Andersson, K. Arpe, M. A. Balmaseda, A. C. M. Beljaars, L. Van De Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher, M. Fuentes, S. Hagemann, E. Hlm, B. J. Hoskins, L. Isaksen, P. A. E. M. Janssen, R. Jenne, A. P. Mcnally, J.-F. Mahfouf, J.-J. Morcrette, N. A. Rayner, R. W. Saunders, P. Simon, A. Sterl, K. E. Trenberth, A. Untch, D. Vasiljevic, P. Viterbo, and J. Woollen. The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, 131(612):2961–3012, 2005.

[126] S. Utsunomiya, K. A. Jensen, G. J. Keeler, and R. C. Ewing. Uraninite and fullerene in atmospheric particulates. *Environ. Sci. Technol.*, 36(23):4943–4947, 2002.

[127] R. von Glasow and P. J. Crutzen. Model study of multiphase DMS oxidation with a focus on halogens. *Atmos. Chem. Phys.*, 4(3):589–608, 2004.

[128] Q. Wang, D. J. Jacob, J. A. Fisher, J. Mao, E. M. Leibensperger, C. C. Carouge, P. Le Sager, Y. Kondo, J. L. Jimenez, M. J. Cubison, , and S. J. Doherty. Sources of carbonaceous aerosols and deposited black carbon in the Arctic in winter-spring: implications for radiative forcing. *Atmos. Chem. Phys.*, 11(23):12453–12473, 2011.

[129] R. Wang, S. Tao, W. Wang, J. Liu, H. Shen, G. Shen, B. Wang, X. Liu, W. Li, Y. Huang, Y. Zhang, Y. Lu, H. Chen, Y. Chen, C. Wang, D. Zhu, X. Wang, B. Li, W. Liu, and J. Ma. Black carbon emissions in China from 1949 to 2050. *Environ. Sci. Tech.*, 46(14):7595–7603, 2012.

[130] Y. Wang, X. Wang, Y. Kondo, M. Kajino, J. W. Munger, and J. Hao. Black carbon and its correlation with trace gasses at a rural site in Beijing: Top-down constraints from ambient measurements on bottom-up emissions. *J. Geophys. Res.*, 116(D24), 2011.

[131] R. Wanninkhof. Relationship between wind speed and gas exchange over the ocean. *J. Geophys. Res.*, 97(C5):7373–7382, 1992.

[132] D. Wen, J. C. Lin, D. B. Millet, A. F. Stein, and R. R. Draxler. A backward-time stochastic Lagrangian air quality model. *Atmos. Env.*, 54:373–386, 2012.

[133] P. Wennberg. Atmospheric chemistry: Bromine explosion. *Nature*, 397:299–301, 1999.

[134] C. Wiedinmyer, S. K. Akagi, R. J. Yokelson, L. K. Emmons, J. A. Al-Saadi, J. J. Orlando, and A. J. Soja. The Fire INventory from NCAR (FINN): a high resolution global model to estimate the emissions from open burning. *Geosci. Model. Dev.*, 4(3):635–641, 2011.

[135] J. D. Wilson, T. K. Flesch, and B. P. Crenna. Estimating surface-air gas fluxes by inverse dispersion using a backward Lagrangian stochastic trajectory model. *Geophys. Monogr. Ser.*, Lagrangian Modeling of the Atmosphere:149–161, 2012.

[136] J. D. Wilson and B. L. Sawford. Review of Lagrangian stochastic models for trajectories in the turbulent atmosphere. *Boundary Layer Meteorol.*, 78(1–2):191–210, 1996.

[137] J. D. Wilson, G. W. Thurtell, and G. E. Kidd. Numerical simulation of particle trajectories in inhomogeneous turbulence, I: Systems with constant turbulent velocity scale. *Boundary Layer Meteorol.*, 21(3):295–313, 1981.

[138] J. D. Wilson, G. W. Thurtell, and G. E. Kidd. Numerical simulation of particle trajectories in inhomogeneous turbulence, II: Systems with variable turbulent velocity scale. *Boundary Layer Meteorol.*, 21(4):423–441, 1981.

[139] J. D. Wilson, G. W. Thurtell, and G. E. Kidd. Numerical simulation of particle trajectories in inhomogeneous turbulence, III: Comparison of predictions with experimental data for the atmospheric surface layer. *Boundary Layer Meteorol.*, 21(4):443–463, 1981.

[140] B. E. Wyslouzil, J. H. Seinfeld, and R. C. Flagan. Binary nucleation in acid-water systems. 1. Methanesulfonic acid-water. *J. Chem. Phys.*, 94(10):6827–6841, 1991.

[141] X. Yang, R. A. Cox, N. J. Warwick, J. A. Pyle, G. D. Carver, F. M. O'Connor, and N. H. Savage. Tropospheric bromine chemistry and its impacts on ozone: A model study. *J. Geophys. Res.*, 110(D23), 2005.

[142] E. Yee, F-S. Lien, A. Keats, and R. D'Amours. Bayesian inversion of concentration data: Source reconstruction in the adjoint representation of atmospheric diffusion. *J. Wind Eng. and Indust. Aerodynamics*, 96(10–11):1805–1816, 2008.

[143] L. Zhang, S. Gong, J. Prado, and L. Barrie. A size-segregated particle dry deposition scheme for an atmospheric aerosol module. *Atmos. Env.*, 35(3):549–560, 2001.

[144] Q. Zhang, K. He, and H. Huo. Cleaning China's air. *Nature*, 484:161–162, 2012.

[145] Q. Zhang, D. G. Streets, G. R. Charmichael, K. B. He, H. Huo, A. Kannari, Z. Klimont, I. S. Park, S.Reddy, J. S. Fu, D. Chen, L. Duan, Y. Lei, L. T. Wang, , and Z. L. Yao. Asian emissions in 2006 for the NASA INTEX-B mission. *Atmos. Chem. Phys.*, 9(14):5131–5153, 2009.

[146] C. Zhao, A. E. Andrews, L. Bianco, J. Eluszkiewicz, A. Hirsch, C. MacDonald, T. Nehrkorn, and M. L. Fischer. Atmospheric inverse estimates of methane emissions from Central California. *J. Geophys. Res.*, 114(D16), 2009.

[147] Y. Zhao, C. P. Nielsen, Y. Lei, M. B. McElroy, and J. Hao. Quantifying the uncertainties of a bottom-up emission inventory of anthropogenic atmospheric pollutants in China. *Atmos. Chem. Phys.*, 11(5):2295–2308, 2011.