

Regional Adjustment of Emission Strengths Via Four Dimensional Data Assimilation

Sun-Kyoung Park¹ and Armistead G. Russell²

¹School of Business Administration/Green Tech MBA, Hanyang Cyber University, Seoul, Korea

²School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, Georgia, U. S. A.

(Manuscript received 20 July 2012; revised 5 November 2012; accepted 20 November 2012)

© The Korean Meteorological Society and Springer 2013

Abstract: The Four-Dimensional Data Assimilation was performed to evaluate source emission strengths over the United States. The US-EPA Models-3 system (CMAQ/MM5/SMOKE) and ridge regression are used as the forward and inverse models, respectively. The continental US is divided into six regions, and data assimilation is performed for each region in July 2001 and January 2002. In addition, two separate scaling factors are calculated for weekdays and weekends. Results show that base emissions for CO and SO₂ sources are relatively accurate. Base emissions for PEC source are overestimated 100%, but those for POA source are underestimated up to 70% when compared with the adjusted emissions. Emissions for NH₃, NO_x, and PMFINE sources are relatively accurate in July 2001, but those in January 2002 are around 100% higher than the adjusted emissions. Base VOC emissions in July 2001 are similar to the adjusted emissions but those in January 2002 are underestimated up to 70% when compared with the adjusted emissions. Though the emission adjustment itself improves the overall air quality model performance, a better improvement is expected with the modification of speciation profiles and temporal allocations in the Models-3 system, as well.

Key words: Emission adjustment, US-EPA Models-3 system, ridge regression, four dimensional data Assimilation

1. Introduction

Photochemical air quality models are essential tools for the investigation of air quality management. Confidence of these models is bolstered when the pollutant concentrations simulated by the models compare well to observations. Accuracy of modeling results relies on model inputs, and one of the most uncertain inputs to air quality models includes the emission inventory (Placet *et al.*, 2000; Sawyer *et al.*, 2000; Mannschreck *et al.*, 2002; Gilliland *et al.*, 2003; Hogrefe *et al.*, 2003; Vautard *et al.*, 2003; Abdel-Aziz and Frey 2004; Taghavi *et al.*, 2005). A direct comparison of '1999 national emission inventory (NET99)' projected to 2001 with '2001 EPA platform emission inventory' also indicated that the emission inventory is one of the most uncertain factors in air quality modeling (Park *et al.*, 2013). Though a nonlinear response of pollutant

concentrations to emissions complicated the air quality model formulation, uncertainties in the responses of pollutant concentrations to various emissions controls were still less than the uncertainties in emissions inputs (Di *et al.*, 2010). In addition, a trans-boundary portion of air pollution also added more difficulty in accurate emission estimation (Park *et al.*, 2004).

Four Dimensional Data Assimilation (FDDA) (Mendoza-Dominguez and Russell, 2000) was developed to refine and evaluate emission inventories. The FDDA method links formal direct sensitivity analysis of three dimensional air quality models with inverse modeling, and incorporates observational data of multiple species. FDDA can estimate the emission strength of various source categories in an inventory simultaneously. The FDDA approach is iterative in nature as the sensitivity changes with emission strengths. Thus, the modeling procedure let us examine the changes of the sensitivity of pollutant concentrations to emission changes in the air quality model.

This method has been assessed by applying the perturbation to the base emission inventories to generate predictions that serve as observations (or pseudo-observations) in the assimilation process, and has successfully identified the predefined perturbation applied (Mendoza-Dominguez and Russell, 2000). This method was also applied to calibrate the 2004 elemental carbon (EC) emissions in the continental United States. The re-estimate of the total emissions was 13% higher than a priori 2004 emissions inventory and the adjusted emissions had an ~10% reduction in annual average fractional error based on 24 h EC observations (Hu *et al.*, 2009b). However, a separate analysis also showed that the results of the method could be affected by the choice of the observational networks (Hu *et al.*, 2009a).

Fundamental assumptions of FDDA are that the major source for the discrepancies between the simulated and observed concentrations of pollutants is the emission inventory. These limitations have been discussed in the previous studies (Mendoza-Dominguez and Russell, 2000; 2001a, b). The spatial allocation of the emissions is often less important to overall performance than the total mass of emissions in regional air quality modeling since the emission inventory is usually prepared with the spatial scale smaller than the grid size (Mendoza-Dominguez and Russell, 2001b; Park *et al.*, 2013). Though it may significantly affect the overall model performance particularly for

Corresponding Author: Sun-Kyoung Park, School of Business Administration/Green Tech MBA, Hanyang Cyber University, Cyber 2, Room 614, Wangsimriro 222 (Hangdangdong 17), Seoul 133-791, Korea.
E-mail: helena@hycu.ac.kr

relatively small scale modeling, for example urban-scale air quality modeling, if the spatial scale of the emission inventory prepared is larger than the grid size.

Previously, FDFA was applied to estimate adjustments in the emissions of gas-phase primary species and O₃ precursors during August 9-10, 1992 over the Atlanta, Georgia Metropolitan area (Mendoza-Dominguez and Russell, 2001b). In that study, the California Institute of Technology (CIT) airshed model (Harley *et al.*, 1993) and ridge regressions (Frank and Friedman, 1993) were used as the forward and the inverse models, respectively. This method was also applied to derive emission adjustments of domain-wide sources of fine organic aerosol and gas-phase species (e.g., NO_x, volatile organic compounds: VOCs, CO, SO₂, and NH₃) over the eastern United States during July 9-19, 1995 and May 22-29, 1995 (Mendoza-Dominguez and Russell, 2001a). The study employed the Urban-to-Regional multiscale Model (URM) (Odman and Russell, 1991) as a forward model.

The present study assessed the feasibility of using the FDFA approach to estimate regional emission strength adjustment over the continental United States. Comparing with the previous studies that were limited to the Atlanta, Georgia metropolitan area or the eastern United States, this study extended the spatial domain over the continental United States. Also, regionally different emission scaling factors were estimated to capture spatial variations of systematic biases in emissions. Furthermore, modeling was performed for two complete months (i.e., July 2001 and January 2002). This ensures that the estimated emission adjustment better reflects systematic bias because emission-scaling factors calculated are less dependent on the short-period abnormal meteorology or modeling system deficiencies.

2. Model description

The four dimensional data assimilation (FDFA) was performed using the US-EPA Models-3 system and ridge regression as the forward and inverse models, respectively. Models-3 is composed of the Community Multiscale Air Quality (CMAQ) model for air quality modeling (Byun and Ching, 1999), the Sparse Matrix Operator Kernel Emissions (SMOKE) for emission processing (US-EPA, 2012b), and the PSU/NCAR's 5th generation Mesoscale Model v5 (MM5) for meteorological modeling (PSU/NCAR, 2012). Meteorological fields were evaluated with the Barnes objective analysis scheme (Koch *et al.*, 1983) using the U.S. and Canada surface hourly data (UCAR, 2012), developed by the Meteorological Development Laboratory (MDL), formally the Techniques Development Laboratory (TDL), of NOAA's National Weather Service. Mean errors (MEs) in temperature, specific humidity, and wind speed over the continental United States were 1.7°C / 2.1°C (July 2001 / January 2002), 1.8 g kg⁻¹ / 0.5 g kg⁻¹, and 1.3 m sec⁻¹ / 1.4 m sec⁻¹, respectively. These values are within the benchmarks for the meteorological model evaluation (Emery and Tai, 2001).

The emission inventory for the continental United States is based on the 1999 National Emission Inventory (NEI 99) (US-EPA, 2012c), projected to 2001 and 2002 using EGAS 4.0 (US-EPA, 2012j). The continuous emissions monitoring data were also used for SO₂ (US-EPA, 2012g). The 1999 Mexico national emission inventory (US-EPA, 2012d) and the 2000 Canadian emission inventory (US-EPA, 2012e) were also used. The biogenic emissions were prepared using the Biogenic Emissions Inventory System version 3 (BEIS3). In order to prepare the meteorological fields using MM5, Pleim-Xiu land surface model, Pleim Chang Planetary boundary layer, simple ice microphysics, Kain-Fritsch cumulus scheme, and rapid radiative transfer model were used (PSU/NCAR, 2012).

Air quality modeling was conducted using CMAQ over the continental United States using the unified Regional Planning Organization (RPO) national grid with a 36 km resolution. The grid dimension is 147 by 111 with the origin of [-2628 km, -1980 km] (Fig. 1). The air quality model domain was larger than the continental United States, in which FDFA was conducted. The two episodes of modeling were from July 1 to 31, 2001 and from January 1 to 30, 2002, and the first 24 hours in each episode were used as ramp-up days. The chemical mechanism used for CMAQ was SAPRC-99 (Byun and Ching, 1999; Carter, 2004). The Modified Euler Backward Iterative method (MEBI) was used as a chemistry solver. The regional acid deposition model (RADM) was for cloud, and AERO3 was used for aerosol dynamics. Deposition velocities of aerosols was calculated using AERO_DEPV2, and horizontal and vertical advection was computed based on piecewise parabolic method (PPM). The initial and boundary conditions were generated using ICON and BCON processors, respectively (Byun and Ching, 1999).

The modeling period includes July 2001 and January 2002, which correspond to the coordinated intensive Eastern Supersite Program (ESP 01/02) period, in which additional observations are available. Observations were used for model evaluation as well as for ridge regression. Monitoring data of gaseous species were obtained from Ozone and Photochemical

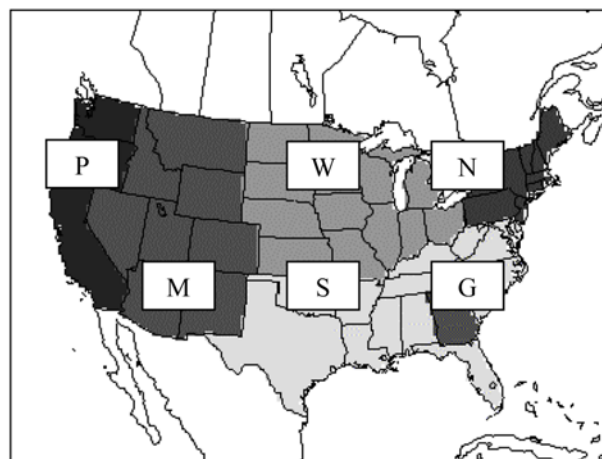


Fig. 1. Six regions defined for the FDFA approach.

Assessment Monitoring Stations (PAMS) program (US-EPA, 2012a). Those of speciated $PM_{2.5}$ data were obtained from the Assessment of Spatial Aerosol Composition in Atlanta (ASACA) project (Butler *et al.*, 2003), the Southeastern Aerosol Research and Characterization (SEARCH) study (Hansen *et al.*, 2003), the Environmental Protection Agency's Speciated Trends Network (EPA-STN) databases (Park *et al.*, 2006), and the Interagency Monitoring of Protected Visual Environments (IMPROVE) network (Ames and Malm, 2001).

The FDDA approach adjusts emissions by incorporating the surface observations into the three-dimensional air quality model. The difference between observations and simulated concentrations, along with the sensitivity of pollutant concentrations to emissions, are used to estimate how much emissions from a specific source should be altered to optimize model performance (Mendoza-Dominguez and Russell, 2000; 2001a, b; Hu *et al.*, 2009a, b). Detailed descriptions of FDDA is available elsewhere (Mendoza-Dominguez and Russell, 2000;

2001a, b), and summary is provided below.

Let O_{ik} be the k^{th} observation (in space and time) of the i^{th} species, and $P_{ik}^{<adjusted>}$, be the adjusted prediction at the same location and time with the measured concentrations. Then, the weighted error is defined as:

$$e_{ik} = w_{ik} (O_{ik} - P_{ik}^{<adjusted>}) \quad (1)$$

where w_{ik} is a weighting factor that accounts for measurement uncertainties, and other properties of the measurements. w_{ik} is computed as:

$$w_{ik} = w_{ik}^{<n>} \cdot w_{ik}^{<c>} \quad (2)$$

where $w_{ik}^{<n>}$ is defined as the inverse of the total (valid) number of observations for species i , i.e., $1/N_i$. This allows each species to have equal weight in the solution. $w_{ik}^{<c>}$ is defined as the inverse of the variance of the k^{th} observation (space and

Table 1. Observed species assimilated and the coefficient of variation used for each species in the ridge regression module.

Species	Coefficient of Variation (%)	No. of Observations											
		Pacific (P)		Mountain (M)		midWest (W)		Northeast (N)		Southeast (S)		Gorgia (G)	
		July	January	July	January	July	January	July	January	July	January	July	January
		2001	2002	2001	2002	2991	2002	2001	2002	2001	2002	2001	2002
O ₃	30	444	71	501	104	1330	0	615	0	1407	13	91	0
CO	50	778	1202	472	920	644	811	625	805	1004	1202	37	25
NO ₂	45	1040	1350	308	379	502	588	682	870	1039	1408	50	65
NO _x	45	901	1212	145	124	289	300	579	718	536	818	50	65
NO _y	45	0	0	0	0	19	0	50	51	231	49	37	24
SO ₂	40	222	301	152	215	1080	1636	893	1333	1050	1367	100	91
Formaldehyde (HCHO)	50	34	41	7	0	73	52	25	31	30	20	0	0
Acetaldehyde (CCHO)	50	34	41	7	0	63	42	28	31	29	21	0	0
Acetone	50	6	0	7	0	34	66	30	19	27	21	0	0
Ethylene	50	14	19	6	0	73	6	111	21	186	191	15	10
Isoprene	50	14	19	6	0	67	3	117	18	166	144	25	1
Ethane (ALK1)	50	14	19	6	0	84	6	130	21	186	206	26	10
ALK2	50	14	13	6	0	84	3	128	21	186	121	26	7
ALK3	50	14	19	6	0	84	6	123	21	186	206	26	10
ALK4	50	14	17	6	0	84	3	124	21	177	164	26	10
ALK5	50	14	17	6	0	73	6	109	21	185	154	23	10
ARO1	50	14	17	6	0	74	6	128	21	184	197	25	10
ARO2	50	2	2	0	0	53	3	129	21	71	75	8	8
OLE1	50	14	19	0	0	50	6	103	21	178	201	22	10
OLE2	50	14	19	6	0	51	2	109	21	159	197	6	10
TNMOC	50	77	139	6	0	84	30	162	66	144	136	26	8
SO ₄ ²⁻ PM _{2.5}	30	132	198	173	252	104	172	147	194	269	378	48	64
NO ₃ ⁻ PM _{2.5}	50	132	174	173	241	102	171	147	194	193	331	48	64
NH ₄ ⁺ PM _{2.5}	30	27	44	21	40	58	111	85	112	197	271	46	55
OC PM _{2.5}	40	132	193	175	251	105	174	151	193	274	376	54	63
EC PM _{2.5}	40	132	198	176	252	105	174	151	193	273	376	53	59
Crustal PM _{2.5}	60	19	40	16	25	54	86	82	84	110	183	23	39
Total PM _{2.5}	30	727	1029	607	714	1358	1733	848	1124	1909	2387	168	206

time) for species i , i.e., $1/(\sigma_{ik})^2$. The more accurate observations will have greater weight than the inaccurate observations. The standard deviation of each species is calculated by multiplying the coefficient of variation by the average concentration of each species.

The number of observations and the coefficient of variation of twenty-eight species used in the assimilation process are illustrated in Table 1. The coefficient of variation for each species is selected following the previous study (Mendoza-Dominguez and Russell, 2001a). The coefficient of variation multiplied by the average concentration of each species becomes the standard deviation of each species, which is used to calculate the observation weight ($w_{ik}^{<cv>}$) described previously.

The value of the adjusted simulated concentration (i.e., $P_{ik}^{<adjusted>}$) is approximated by the linear combination of the product of the sensitivity coefficients and the perturbed

emission strengths:

$$P_{ik}^{<adjusted>} = P_{ik} + \sum_{j=1}^J \delta_{ij} s_{ik,j} m_j \quad (3)$$

where P_{ik} is the predicted concentration at base emission, $s_{ik,j}$ is the sensitivity of the i^{th} species to emissions from the j^{th} source at the k^{th} location, δ_{ij} is a weighting factor applied to the sensitivity coefficients, J is the total number of sources involved, and m_j is the emission adjustment scaling factor of the j^{th} source from its base level. The emission adjustment scaling factor (m_j) is the unknown that should be calculated.

A new weighting factor for the sensitivity coefficients (δ_{ij}) is introduced in an ad hoc fashion to compensate for a lack of emission constraints in the penalty function (W_m), which will be shown in Eq. (6). δ_{ij} used in this study are given in Table 2

Table 2. The weighting factors for the sensitivity coefficients (δ_{ij}).

Species	Sources*									
	CO (T)	NH ₃ (T)	NO _x (AMN)	NO _x (P)	SO ₂ (T)	VOC (AMNP)	VOC (B)	POA (T)	PEC (T)	PMFINE (T)
O ₃	0.5	0.5	1	1	0.5	1	1	0.5	0	0
CO	1	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0	0
NO ₂	0	0	0.5	0.5	0	0	0	0	0	0
NO _x	0.5	0.5	1	1	0.5	1	1	0.5	0	0
NO _y	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0	0
SO ₂	0.5	0.5	0.5	0.5	1	0.5	0.5	0.5	0	0
NMOC	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
Ethylene	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
Isoprene	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
HCHO	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
CCHO	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
Acetone	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
ALK1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
ALK2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
ALK3	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
ALK4	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
ALK5	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
OLE1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
OLE2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
ARO1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
ARO2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
Total PM _{2.5}	0	0.5	0.1	0.1	0.5	0	0	0.5	0	0
SO ₄ ²⁻ -PM _{2.5}	0	0.5	0.1	0.1	1	0	0	0.5	0	0
NO ₃ ⁻ -PM _{2.5}	0	0.5	0	0	0.5	0	0	0.5	0	0
NH ₄ ⁺ -PM _{2.5}	0	1	0.1	0.1	0.5	0	0	0.5	0	0
OC PM _{2.5}	0	0.5	0.1	0.1	0.5	0	0	1	0	0
EC PM _{2.5}	0	0.5	0.1	0.1	0.5	0	0	0.5	1	0
Crustal PM _{2.5}	0	0.5	0.1	0.1	0.5	0	0	0.5	0	1

* T (total): area, biogenic, mobile, nonroad and point sources

* A: area source

* B: biogenic source

* M: mobile source

* N: nonroad source

* P: point source

(Mendoza-Dominguez and Russell, 2001a). Please note that the weighting factor (w_{ik}) introduced in Eqs. (1) and (2) is to account for measurement uncertainties and the number of observation.

Sensitivity coefficients (s_{ikj}) can be calculated by direct sensitivity methods such as DDM-3D (Dunker 1981; Yang *et al.*, 1997) or by Brute Force (BF) (i.e., applying the model once with, then without the target sources) (Park *et al.*, 2013), the latter of which was applied in this study. Ten target emission sources used to calculate the sensitivity coefficients are total carbon monoxide (CO; T), total ammonium (NH₃; T), area, on-road mobile, and nonroad mobile source nitrogen oxides (NO_x; AMN), point source nitrogen oxides (NO_x; P), total sulfur dioxide (SO₂; T), anthropogenic volatile organic carbon (VOC; AMNP), biogenic volatile organic carbon (VOC; B), total primary organic aerosol (POA; T), total primary elemental carbon (PEC; T), and total crustal element (PMFINE; T). Thus, the air quality model for each region in the 1st iteration was run eleven times to calculate the sensitivity coefficients for ten emission sources, one with base emissions, and the other ten times with emissions without each target source. Since the modeling domain was divided into six regions, and iterations were conducted, the eleven times of modeling steps were repeated for each region in every iteration.

The receptor model is written using Eqs. (1) and (3) as:

$$w_{ik} d_{ik} = w_{ik} \sum_{j=1}^J \delta_{ij} s_{ikj} m_j + e_{ik} \quad (4)$$

where d_{ik} ($= O_{ik} - P_{ik}$) is the difference between observed and modeled concentrations. Eq. (4) can be expressed in matrix notation:

$$\mathbf{W}_c \mathbf{d} = \mathbf{W}_c \mathbf{G} \mathbf{m} + \mathbf{e} \quad (5)$$

where \mathbf{W}_c is a ($I \times K$) diagonal matrix of observation weights (the w_{ik} 's are the elements of the diagonal), \mathbf{d} ($d_{ik} = O_{ik} - P_{ik}$) and \mathbf{e} (error term) are ($I \times K$) length vectors, \mathbf{G} ($g_{ikj} = \delta_{ij} s_{ikj}$) is the ($I \times K \times J$) matrix of weighted sensitivity coefficients, and \mathbf{m} is a J -length vector. I , K , and J are the total number of species, sources, and monitoring stations, respectively.

The unknown variable, \mathbf{m} , the amount of emissions to be adjusted is selected by minimizing the following:

$$\Gamma = \mathbf{e}^T \mathbf{W}_c \mathbf{e} + \mathbf{m}^T \mathbf{W}_m \mathbf{m} \quad (6)$$

The penalty function (\mathbf{W}_m) is introduced to bind the emission adjustment within the physically meaningful ranges. The adjusted emissions should lie inside a range defined by uncertainty limits of the base emission estimates. Uncertainties in emission estimates are often represented by a log-normal distribution. Assuming that the base inventory is located at the median of the distribution, a Confidence Interval (CI) of the emissions estimates ranges from $1/\sigma_g^n$ to σ_g^n (σ_g is the geometric standard deviation), where n depends on the CI selected. Here, n was selected as 3.3 for CI to be 99.9% (Mendoza-

Dominguez and Russell, 2000). The upper limits of the CI of emissions, i.e., $\sigma_g^{3.3}$, are 3.73 for VOC and POA, and 2.25 for all other sources (Mendoza-Dominguez and Russell, 2001a).

\mathbf{W}_m is a ($J \times J$) diagonal matrix, which has parameters w_{jj} (positive constants) in the diagonal, i.e., $\mathbf{W}_m = \lambda \mathbf{I}$, where λ is known as the ridge parameter and \mathbf{I} is the identity matrix. The method to calculate the ridge parameter λ can be found in elsewhere (Marquardt, 1963; Hoerl and Kennard, 1970; Frank and Friedman, 1993; Aldrin, 1997; Mendoza-Dominguez and Russell, 2000).

The resulting \mathbf{m} after minimizing the Eq. (6) is as follows:

$$\mathbf{m} = (\mathbf{G}^T \mathbf{W}_c \mathbf{G} + \mathbf{W}_m)^{-1} \mathbf{G}^T \mathbf{W}_c \mathbf{d} \quad (7)$$

The base emissions multiplied by the emission adjustment scaling factor (\mathbf{m}) from Eq. (7) are the adjusted emissions of the first iteration. The FDDA approach is iterative in nature as the sensitivity changes with emission strengths. The sensitivity coefficients multiplied by the weighting factor of the sensitivity coefficient (\mathbf{G}) and the difference between modeled and observed concentrations (\mathbf{d}) are recalculated based on the adjusted emissions from the first iteration. \mathbf{G} is again calculated by Brute Force described previously. \mathbf{m} of the second iteration is calculated using Eq. (7) with \mathbf{G} and \mathbf{d} obtained using the adjusted emissions from the first iteration. In this way, \mathbf{m} is iteratively recalculated.

The iteration is conducted to have a stable response in the assimilation process, and the following expression is used to terminate the iteration process (Hoerl and Kennard, 1970):

$$\frac{\lambda^t - \lambda^{t-1}}{\lambda^{t-1}} \leq 20 \left\{ \frac{\text{Tr}[(\mathbf{G}^T \mathbf{W}_c \mathbf{G})^{-1}]}{J} \right\}^{-1.3} \quad (\text{for } t > 1) \quad (8)$$

where λ^t is the λ (ridge parameter) at t th iteration, and $\text{Tr}[\mathbf{M}]$ denotes the trace of the square matrix \mathbf{M} . After five iterations, Eq. (8) was satisfied to terminate the iteration in this study.

The error bound (1σ) for each scaling factor is obtained from the square root of the diagonal elements of the variance-covariance matrix (\mathbf{V}) in the emission adjustment estimates.

$$\mathbf{V} = (\mathbf{G}^T \mathbf{W}_c \mathbf{G} + \mathbf{W}_m)^{-1} (\mathbf{G}^T \mathbf{W}_c \mathbf{G}) (\mathbf{G}^T \mathbf{W}_c \mathbf{G} + \mathbf{W}_m) \quad (9)$$

Since iteration is conducted to obtain the ultimate estimates for vector \mathbf{m} , the values of \mathbf{G} and \mathbf{W}_m from the last iteration are used to compute \mathbf{V} (Menke, 1989). For the nonlinear problem, σ does not necessarily represent a standard deviation from a typical Gaussian distribution.

The robustness of this method was tested in the previous study (Mendoza-Dominguez and Russell, 2000). A predefined perturbation was applied to the base emissions inventory to generate predictions that serve as observations in the assimilation process. Therefore, a known solution exist in test scenarios. Five test scenarios, each one with different mix of observed species and emission sources, were investigated. In addition, several combinations of the weighting factors were

evaluated for its impact on convergence and stability of the emission adjustments. Results showed that the FDDA method could adjust perturbed emissions close to their known original values in these test scenarios. The weighting factors helped to accelerate the convergence, but they have little impact on the final results

3. Results

The continental United States is divided into six regions based on the U.S. Census Bureau (Census, 2012): Pacific (P), Mountain (M), midWest (W), Southeast (S), Northeast (N), and Georgia (G) (Fig. 1). Emission adjustment scaling factors are calculated separately for those six regions. In addition to the different scaling factor for each region, scaling factors were calculated for weekdays (Monday through Friday) and weekends (Saturday and Sunday) to account for the day-of-week variation. Since emissions strength continuously varies, hourly emissions assimilation would give more useful information than daily assimilation. However, observed $PM_{2.5}$ species data were only available on daily basis (24-hour average) though measurements of other species were collected hourly, so the FDDA was performed on daily basis.

Emissions were adjusted separately for each region as well as for weekdays and weekends (Fig. 2 and Table 3). The error bounds (expressed as 1σ) for scaling factors are also represented (Table 3). CO emissions are relatively accurate with

little difference in emission scaling factors between weekdays and weekends. However, scaling factors in the Southeast region in July 2001 and the Mountain region in January 2002 are significantly different from 1.0. The scaling factor of CO emissions in the State of Georgia on weekends in January 2002 is also deviated from 1.0. The relatively small errors in the CO emissions are consistent with the previous studies (Petron *et al.*, 2004; Muller and Stavrou, 2005; Stavrou and Muller, 2006). However, a detailed comparison with the previous studies was not done because the spatial resolutions of the previous studies differed from those of this work.

NH_3 emissions are slightly overestimated in July 2001, but they are overestimated 100% in January 2002. The formation of NH_3NO_3 from NO_x and NH_3 favors low temperature, the correct estimation of NH_3 improves the predicted $NO_3-PM_{2.5}$ concentrations especially in winter (Gilliland *et al.*, 2003). Emission scaling factor for NO_x from point sources and NO_x from area/mobile/nonroad sources had similar trends. The base NO_x emissions were overestimated around 120% for Georgia and Mountain regions in July 2001. However, NO_x emissions were underestimated in Southeast and Pacific regions in July 2001. Previous studies revealed that the NO_x in the Southeastern region was underestimated (Mendoza-Dominguez and Russell, 2001a). NO_x emissions were around 100% overestimated in January 2002 except in the Pacific region.

SO_2 emissions were relatively accurate because the continuous emissions monitoring data were used for modeling

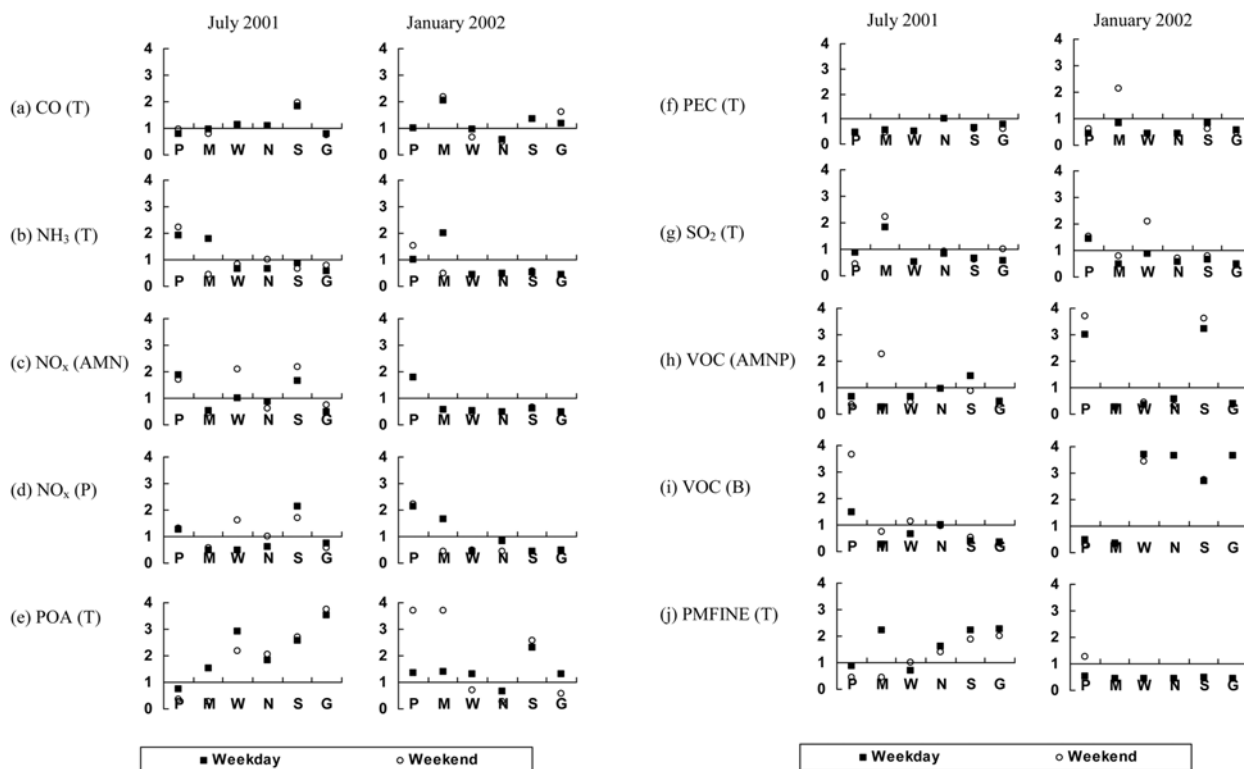


Fig. 2. Emission adjustment scaling factor for July 2001 and January 2002. The scaling factor of one indicates that the adjustment is not needed. Two and 0.5 indicate the adjustment is needed twice of the base level and a half of the base level, respectively. Some scaling factors on weekday are overlapped with those on weekends.

Table 3. Emission adjustment scaling factors and the error bound (1σ) in each region on weekdays and weekends. (a) is for July 2001, and (b) is for January 2002. A scaling factor of one indicates that the adjusted emissions are the same as the base emissions. Two and 0.5 indicate that the adjusted emissions are twice and a half of the base emissions, respectively.

(a) July 2001						
(a)-1. Weekdays (Monday through Friday)						
	Pacific (P)	Mountain (M)	midWest (W)	Northeast (N)	Southeast (S)	Georgia (G)
CO	0.78 ± 1.01	0.96 ± 1.14	1.14 ± 1.09	1.11 ± 0.00	1.82 ± 0.29	0.76 ± 0.79
NH3	1.90 ± 0.00	1.78 ± 0.02	0.65 ± 0.00	0.65 ± 0.00	0.87 ± 0.00	0.58 ± 0.00
NOXA	1.86 ± 0.09	0.50 ± 0.00	0.98 ± 0.03	0.85 ± 0.00	1.65 ± 0.07	0.47 ± 0.00
NOXP	1.28 ± 0.02	0.49 ± 0.07	0.46 ± 0.00	0.60 ± 0.00	2.11 ± 0.00	0.75 ± 0.01
PEC	0.47 ± 0.00	0.57 ± 0.01	0.50 ± 0.00	1.00 ± 0.00	0.65 ± 0.00	0.78 ± 0.00
PMFI	0.86 ± 0.01	2.21 ± 0.01	0.69 ± 0.01	1.62 ± 0.00	2.23 ± 0.01	2.24 ± 0.00
POA	0.73 ± 0.01	1.51 ± 0.01	2.90 ± 0.00	1.84 ± 0.00	2.59 ± 0.00	3.50 ± 0.00
SO2	0.85 ± 0.01	1.82 ± 0.01	0.51 ± 0.00	0.81 ± 0.00	0.64 ± 0.00	0.57 ± 0.00
VOCA	0.67 ± 0.09	0.28 ± 0.01	0.64 ± 0.08	0.97 ± 0.00	1.44 ± 0.49	0.48 ± 0.27
VOCB	1.48 ± 1.68	0.28 ± 0.00	0.63 ± 0.30	0.98 ± 0.00	0.39 ± 0.57	0.35 ± 0.32
(a)-2. Weekends (Saturday and Sunday)						
	Pacific (P)	Mountain (M)	midWest (W)	Northeast (N)	Southeast (S)	Georgia (G)
CO	0.97 ± 1.20	0.77 ± 0.98	1.14 ± 1.22	1.09 ± 0.00	1.97 ± 0.19	0.73 ± 0.76
NH3	2.20 ± 0.01	0.45 ± 0.00	0.82 ± 0.00	1.00 ± 0.00	0.65 ± 0.00	0.77 ± 0.00
NOXA	1.69 ± 0.01	0.45 ± 0.00	2.11 ± 0.00	0.61 ± 0.00	2.18 ± 0.12	0.74 ± 0.01
NOXP	1.31 ± 0.01	0.59 ± 0.00	1.60 ± 0.02	0.99 ± 0.00	1.70 ± 0.01	0.57 ± 0.01
PEC	0.45 ± 0.00	0.47 ± 0.01	0.48 ± 0.00	1.00 ± 0.00	0.58 ± 0.00	0.58 ± 0.00
PMFI	0.45 ± 0.00	0.45 ± 0.01	1.01 ± 0.01	1.40 ± 0.00	1.87 ± 0.01	2.00 ± 0.00
POA	0.33 ± 0.02	0.27 ± 0.00	2.17 ± 0.00	2.06 ± 0.00	2.70 ± 0.00	3.72 ± 0.00
SO2	0.45 ± 0.01	2.22 ± 0.01	0.54 ± 0.00	0.91 ± 0.00	0.62 ± 0.00	0.98 ± 0.00
VOCA	0.34 ± 0.22	2.28 ± 0.14	0.42 ± 0.47	0.97 ± 0.00	0.88 ± 0.47	0.34 ± 0.03
VOCB	3.65 ± 0.00	0.74 ± 0.29	1.15 ± 0.24	0.98 ± 0.00	0.51 ± 0.48	0.31 ± 0.07
(b) January 2002						
(b)-1. Weekdays (Monday through Friday)						
	Pacific (P)	Mountain (M)	midWest (W)	Northeast (N)	Southeast (S)	Georgia (G)
CO	0.98 ± 1.20	2.02 ± 0.24	0.98 ± 7.97	0.57 ± 0.53	1.35 ± 0.86	1.18 ± 1.03
NH3	1.00 ± 0.00	2.01 ± 0.01	0.45 ± 0.01	0.46 ± 0.00	0.51 ± 0.00	0.44 ± 0.00
NOXA	1.79 ± 0.01	0.56 ± 0.07	0.50 ± 0.15	0.50 ± 0.01	0.63 ± 0.01	0.49 ± 0.01
NOXP	2.12 ± 0.00	1.67 ± 0.03	0.45 ± 0.38	0.85 ± 0.01	0.44 ± 0.00	0.46 ± 0.00
PEC	0.45 ± 0.00	0.81 ± 0.00	0.45 ± 0.00	0.45 ± 0.00	0.81 ± 0.00	0.56 ± 0.00
PMFI	0.52 ± 0.00	0.45 ± 0.00	0.45 ± 0.00	0.45 ± 0.00	0.46 ± 0.00	0.45 ± 0.00
POA	1.36 ± 0.00	1.38 ± 0.01	1.32 ± 0.00	0.63 ± 0.00	2.31 ± 0.00	1.29 ± 0.00
SO2	1.43 ± 0.02	0.48 ± 0.02	0.88 ± 0.01	0.55 ± 0.00	0.65 ± 0.00	0.47 ± 0.00
VOCA	3.00 ± 1.32	0.27 ± 0.00	0.35 ± 17.2	0.56 ± 0.06	3.20 ± 0.84	0.37 ± 0.03
VOCB	0.49 ± 0.08	0.37 ± 0.02	3.70 ± #####	3.64 ± 0.01	2.72 ± 0.04	3.66 ± 0.00
(b)-2. Weekends (Saturday and Sunday)						
	Pacific (P)	Mountain (M)	midWest (W)	Northeast (N)	Southeast (S)	Georgia (G)
CO	1.00 ± 1.23	2.19 ± 0.02	0.64 ± 0.74	0.48 ± 0.26	1.34 ± 0.72	1.60 ± 0.73
NH3	1.53 ± 0.01	0.46 ± 0.01	0.45 ± 0.01	0.45 ± 0.01	0.56 ± 0.00	0.45 ± 0.00
NOXA	1.76 ± 0.01	0.56 ± 0.04	0.45 ± 0.00	0.48 ± 0.01	0.63 ± 0.01	0.48 ± 0.00
NOXP	2.20 ± 0.00	0.45 ± 0.04	0.48 ± 0.00	0.45 ± 0.00	0.45 ± 0.00	0.45 ± 0.00
PEC	0.61 ± 0.00	2.13 ± 0.01	0.45 ± 0.00	0.45 ± 0.00	0.61 ± 0.00	0.47 ± 0.00
PMFI	1.28 ± 0.00	0.44 ± 0.00	0.45 ± 0.00	0.44 ± 0.00	0.45 ± 0.00	0.45 ± 0.00
POA	3.71 ± 0.00	3.71 ± 0.00	0.69 ± 0.00	0.27 ± 0.00	2.55 ± 0.00	0.57 ± 0.00
SO2	1.52 ± 0.04	0.80 ± 0.02	2.09 ± 0.01	0.71 ± 0.00	0.78 ± 0.00	0.46 ± 0.00
VOCA	3.70 ± 0.00	0.27 ± 0.00	0.42 ± 0.67	0.28 ± 0.01	3.62 ± 0.05	0.34 ± 0.10
VOCB	0.29 ± 0.01	0.29 ± 0.00	3.45 ± 0.00	3.63 ± 0.00	2.75 ± 0.28	3.67 ± 0.00

(US-EPA, 2012g). Emissions of all other sources were modeled using the annual emissions (1999 National Emission Inventory: NEI 99) (US-EPA, 2012c). The annual emissions were allocated for each month, day of week, and hour of day according to the temporal allocation file in SMOKE (US-EPA, 2012h). Thus, the accuracy of the hourly or daily emissions is limited by the accuracy of temporal allocation as well as that of the annual emissions. The temporal allocation used in SMOKE does not always reflect the actual temporal variation. It is usually smoother than the real variations (Unal *et al.*, 2005; Park *et al.*, 2013). Therefore, the different model performance between July 2001 and January 2002 for sources other than SO₂ is also partly attributed from the temporal allocation used in SMOKE.

Emission scaling factors of PEC carbon were less than 1.0 in most cases suggesting the overestimation of the PEC emissions. The low scaling factor of PEC reflects the overestimation of the elemental carbon concentrations because the weighting factor for sensitivity coefficient (δ_{ij}) for PEC is positive only for elemental carbon species (Table 2). Results may vary depending on the year of analysis as well as monitoring networks chosen (Hu *et al.*, 2009a). POA emissions were overestimated in most regions. This result is partly due to the underestimation of the secondary organic carbon concentrations.

PMFINE emissions were underestimated in July 2001, but overestimated in January 2002. These results are also directly related with the performance of the crustal PM_{2.5} concentrations because δ_{ij} for PMFINE is positive only for crustal PM_{2.5}. Studies showed that the modeling of crustal PM_{2.5} has difficulty due to the uncertainty in fugitive dust emissions. Zero to 100% of the fugitive dust emissions are emitted very near ground level where it can immediately be re-deposited on the ground or trapped in vegetation depending on the regions (Pace, 2003). Here, 75% of the fugitive dust emissions are assumed

that they are not mixed in the modeling grid.

Before assimilation, sulfate (SO₄²⁻PM_{2.5}) concentrations in July 2001 and nitrate (NO₃⁻PM_{2.5}) concentrations in January 2002 were overestimated, whereas organic carbon concentrations (OC PM_{2.5}) in July 2001 were underestimated for the midWestern, Northeastern, Southeastern, and Georgia regions (Fig. 3). After assimilation, sulfate concentrations in July 2001 and nitrate concentrations in January 2002 decreased, and organic carbon concentrations in July 2001 increased (Fig. 3).

The biogenic VOC emissions were underestimated in the Pacific region in July 2001, and those in the midWestern, the Northeastern, and the Southeastern regions and the State of Georgia in January 2002 (Fig. 2). Previous studies indicated that the emissions of isoprene estimated using BEIS3 were lower than the measured value in the southeastern United States (Warneke *et al.*, 2010). In addition, isoprene emissions in the United States from BEIS3 were lower than those estimated using the Model of Emissions of Gases and Aerosols from Nature (MEGANv2.02) (Sakulyanotvittaya *et al.*, 2008).

The anthropogenic VOC emissions were overestimated in most cases, consistent with the previous study (Chang *et al.*, 1996). VOC emission adjustments involved fifteen VOC species, O₃ and other species (Table 2). ALK1 and ALK2 are underestimated in most regions. Other VOC species concentrations were higher or lower than the measured values depending on the regions and on the episode (Table 4 and Fig. 4). These different model performances of multiple VOC species complicated the VOC emission adjustment.

The emission adjustment in this work was performed for the emissions inventory, so the adjustment is for the sum of the VOC emissions, not for the individual VOC emission sources (e.g., isoprene, etc.). Thus, after the sum of VOC emissions are adjusted, the emissions of individual VOC emissions sources

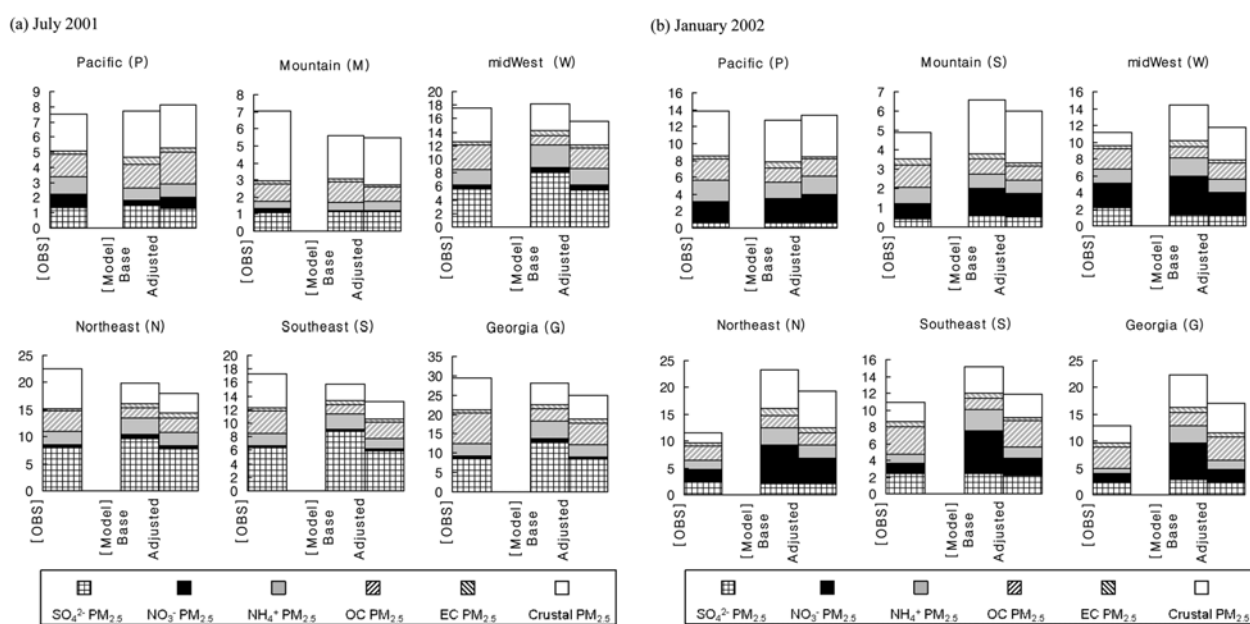


Fig. 3. Average observed and modeled concentrations of PM_{2.5} species in $\mu\text{g m}^{-3}$. (a) is for July 2001, and (b) is for January 2002.

Table 4. Average observed and modeled concentrations, and mean biases and errors of O₃, CO, NO₂, NO_x, NO_y, SO₂, and TNMOC in ppb (a) is for July 2001 and (b) is for January 2002.

(a) July 2001								
Region	Species	OBS conc.	MODEL conc.		MB		ME	
			(Base)	(Adjusted)	(Base)	(Adjusted)	(Base)	(Adjusted)
Pacific	O ₃	50.25	40.92	39.57	-9.3	-11.0	13.5	18.4
	CO	452.53	327.87	294.81	-124.7	-166.6	267.3	262.9
	NO ₂	11.48	12.25	18.32	0.8	6.3	7.2	11.7
	NO _x	16.64	15.93	28.20	-0.7	10.8	10.6	19.7
	NO _y	NA	NA	NA	NA	NA	NA	NA
	SO ₂	2.44	1.17	0.90	-1.3	-1.6	1.5	1.8
	TNMOC	56.18	23.84	25.74	-32.3	-31.2	40.9	39.8
Mountain	O ₃	46.40	45.59	44.25	-0.8	-2.7	6.3	9.5
	CO	565.50	404.70	361.28	-160.8	-215.8	318.0	326.1
	NO ₂	11.94	11.39	6.70	-0.6	-5.5	7.5	7.3
	NO _x	16.50	9.63	4.83	-6.9	-11.3	14.0	12.9
	NO _y	NA	NA	NA	NA	NA	NA	NA
	SO ₂	6.03	1.15	2.23	-4.9	-5.3	4.9	6.1
	TNMOC	47.79	139.85	45.23	92.1	-2.6	99.7	42.7
Midwest	O ₃	49.60	63.67	58.45	14.1	9.0	16.4	14.5
	CO	569.43	424.80	433.52	-144.6	-127.3	286.1	293.3
	NO ₂	14.88	15.19	17.30	0.3	2.4	7.1	8.2
	NO _x	20.73	19.45	26.62	-1.3	4.7	10.2	15.4
	NO _y	11.19	17.83	19.28	6.6	8.4	11.9	13.3
	SO ₂	5.85	4.51	2.33	-1.3	-3.1	4.2	3.8
	TNMOC	54.45	43.20	26.60	-11.3	-24.5	48.4	40.3
Northeast	O ₃	51.70	62.90	62.26	11.2	10.6	14.7	14.1
	CO	544.89	401.19	411.82	-143.7	-126.4	268.9	269.9
	NO ₂	15.36	15.04	12.24	-0.3	-2.6	8.3	7.6
	NO _x	20.59	17.32	14.04	-3.3	-6.3	11.3	11.3
	NO _y	9.46	9.34	7.99	-0.1	-1.5	3.6	3.5
	SO ₂	5.56	6.51	4.91	1.0	-0.4	4.2	3.7
	TNMOC	25.94	39.58	36.13	13.6	10.8	28.6	26.3
South	O ₃	47.66	61.15	54.98	13.5	7.3	15.9	15.4
	CO	523.24	296.49	391.41	-226.7	-129.0	270.4	255.5
	NO ₂	10.32	10.94	17.16	0.6	6.9	6.7	10.7
	NO _x	12.12	13.12	25.23	1.0	13.2	7.4	15.9
	NO _y	9.34	9.84	15.38	0.5	6.0	5.5	9.0
	SO ₂	5.35	5.39	3.34	0.0	-2.0	3.7	3.4
	TNMOC	33.81	35.26	25.63	1.5	-7.1	28.8	24.7
Georgia	O ₃	47.55	68.71	62.25	21.2	14.2	21.6	19.3
	CO	460.21	472.31	397.62	12.1	-40.8	216.4	186.3
	NO ₂	9.57	12.54	7.81	3.0	-1.4	5.8	5.5
	NO _x	15.64	13.35	8.34	-2.3	-6.3	10.9	11.1
	NO _y	25.00	24.92	15.61	-0.1	-7.9	15.3	14.5
	SO ₂	1.90	3.40	2.09	1.5	0.1	2.1	1.3
	TNMOC	24.16	54.89	22.62	30.7	0.0	38.3	15.0

were recalculated based on the VOC speciation profile (US-EPA, 2012i). Therefore, even though the sum of VOC emissions is properly adjusted, the adjusted VOC species concen-

trations could be incorrectly estimated if the VOC speciation profile is not correct. The relative contribution of fourteen VOC species to the sum of them indicates that the contribution

Table 4. (Continued) Average observed and modeled concentrations, and mean biases and errors of O₃, CO, NO₂, NO_x, NO_y, SO₂, and TNMOC in ppb (a) is for July 2001 and (b) is for January 2002.

(b) January 2002								
January 2002								
Region	Species	OBS conc.	MODEL conc.		MB		ME	
			(Base)	(Adjusted)	(Base)	(Adjusted)	(Base)	(Adjusted)
Pacific	O ₃	43.10	29.59	28.74	-13.5	-14.5	13.9	14.8
	CO	909.53	564.40	560.41	-345.1	-358.2	539.0	543.9
	NO ₂	20.83	20.80	24.41	0.0	3.4	10.4	12.9
	NO _x	49.05	32.21	46.45	-16.8	-3.8	30.4	36.6
	NO _y	NA	NA	NA	NA	NA	NA	NA
	SO ₂	2.52	2.18	2.81	-0.3	0.2	1.9	2.2
	TNMOC	118.13	31.47	60.60	-86.7	-54.1	90.3	85.2
Mountain	O ₃	44.41	36.43	37.09	-8.0	-7.4	9.3	8.5
	CO	1074.17	440.47	647.55	-633.7	-440.6	667.8	627.3
	NO ₂	24.18	16.42	8.55	-7.8	-15.5	11.1	16.3
	NO _x	63.63	16.52	7.29	-47.1	-57.4	52.4	58.6
	NO _y	NA	NA	NA	NA	NA	NA	NA
	SO ₂	4.97	1.50	1.08	-3.5	-4.1	3.7	4.1
	TNMOC	NA	NA	NA	NA	NA	NA	NA
Midwest	O ₃	NA	NA	NA	NA	NA	NA	NA
	CO	626.68	501.55	486.48	-125.1	-125.0	325.2	309.3
	NO ₂	15.38	14.27	8.63	-1.1	-6.4	4.7	6.9
	NO _x	32.19	28.25	12.89	-3.9	-18.2	17.2	19.4
	NO _y	NA	NA	NA	NA	NA	NA	NA
	SO ₂	5.65	4.82	4.10	-0.8	-1.6	3.6	3.6
	TNMOC	107.01	18.63	12.76	-88.4	-94.3	88.9	94.3
Northeast	O ₃	NA	NA	NA	NA	NA	NA	NA
	CO	716.64	742.53	574.50	25.9	-119.2	392.3	333.3
	NO ₂	22.64	21.77	14.79	-0.9	-6.9	6.9	8.7
	NO _x	44.94	39.60	20.42	-5.3	-21.6	25.6	24.6
	NO _y	28.89	18.63	9.77	-10.3	-17.3	18.3	18.6
	SO ₂	8.36	10.76	8.14	2.4	0.0	5.3	4.4
	TNMOC	58.22	65.38	47.33	7.2	-10.1	31.5	25.2
South	O ₃	44.65	24.07	32.30	-20.6	-12.4	21.3	13.2
	CO	754.76	435.28	557.39	-319.5	-192.7	414.7	403.8
	NO ₂	14.81	13.25	8.71	-1.6	-6.0	6.0	7.6
	NO _x	28.87	19.39	11.03	-9.5	-17.6	16.0	19.0
	NO _y	11.24	10.41	5.74	-0.8	-4.9	5.7	7.4
	SO ₂	6.23	6.25	4.42	0.0	-1.7	3.9	3.6
	TNMOC	53.46	18.25	38.66	-35.2	-13.7	35.8	29.0
Georgia	O ₃	NA	NA	NA	NA	NA	NA	NA
	CO	833.38	628.18	677.84	-205.2	-149.8	313.2	313.9
	NO ₂	17.19	16.65	8.54	-0.5	-9.1	7.0	9.7
	NO _x	42.22	24.72	9.78	-17.5	-34.0	26.8	34.8
	NO _y	50.29	31.48	13.72	-18.8	-39.1	32.2	39.8
	SO ₂	2.73	3.97	2.52	1.2	-0.6	2.2	1.8
	TNMOC	28.09	21.97	16.45	-6.1	-11.6	11.0	12.3

of individual VOC species were quite different between observed and measured concentrations (compare first and second bar graphs in Fig. 5). Even after the assimilation, the

contribution of individual VOC species did not change much (compare second and third bar graphs in Fig. 5) indicating that the adjustment of VOC speciation profile also needed to better

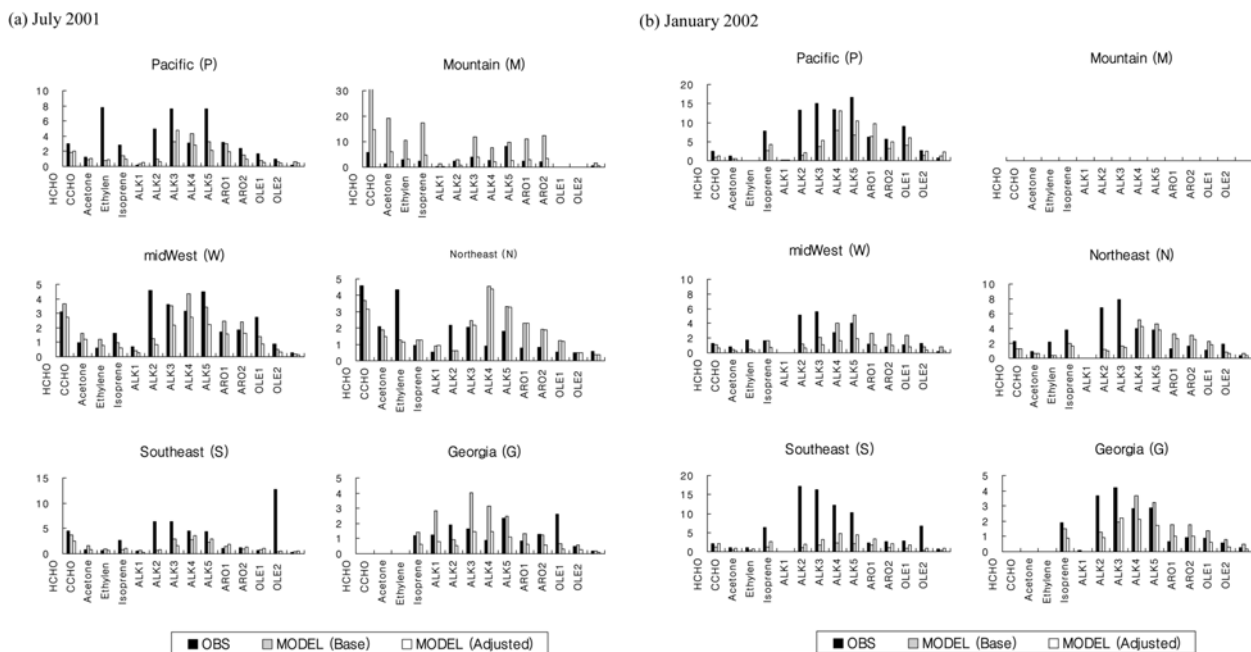


Fig. 4. Average observed and modeled concentrations of VOC species [ppb]. (a) is for July 2001, and (b) is for January 2002. Observed VOC spec concentrations were not available in the Mountain region (M) in January 2002.

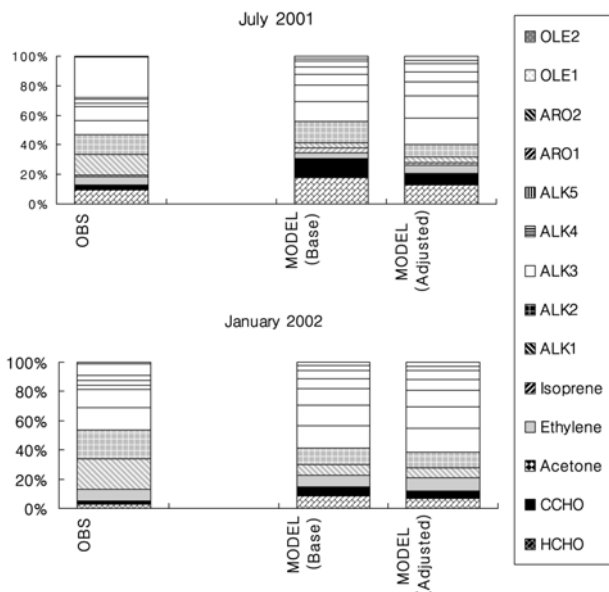


Fig. 5. The relative contribution of HCHO, CCHO, acetone, ethylene, isoprene, ALK1, ALK2, ALK3, ALK4, ALK5, ARO1, ARO2, OLE1, OLE2 to the sum of those species in the Southeastern region. (a) is for July 2001, and (b) is for January 2002.

improve the model performance.

The performance of the model using base level emissions was compared with that using adjusted emissions (Fig. 6). The mean errors decreased in the Mountain, the midWestern, and the Northeastern regions, and the State of Georgia. However, the mean errors rather increased in the Pacific and Southeastern regions partly due to the uncertainty in the VOC

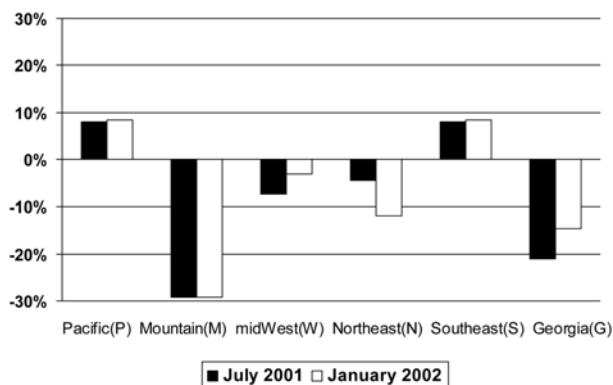


Fig. 6. Normalized change* of Mean Errors (MEs) between base emissions and adjusted emissions.

*Normalized change of MEs (Park *et al.*, 2006): $(ME_{adjusted\ emissions} - ME_{base\ emissions}) / ME_{base\ emissions} \times 100(\%)$

speciation profile.

The overall mean fractional bias (MFB) of the modeled concentrations from the base emissions were compared with that from the adjusted emissions (Fig. 7). A scatter plot was drawn with MFBs from the base emissions of each species of each region to X axis, and those from the adjusted emissions to Y axis. The slope of the first order regression line was less than one suggesting that the FDDA approach helped to minimize the systematic bias of the emissions. However, the overall mean fractional error (MFE) did not improve, indicating that there are other sources of error (e.g., the speciation profile of VOC), which cannot be reduced by emission adjustments alone (Fig. 7).

It is also interesting to compare the emission adjustment through the FDDA with the relatively recent emission inven-

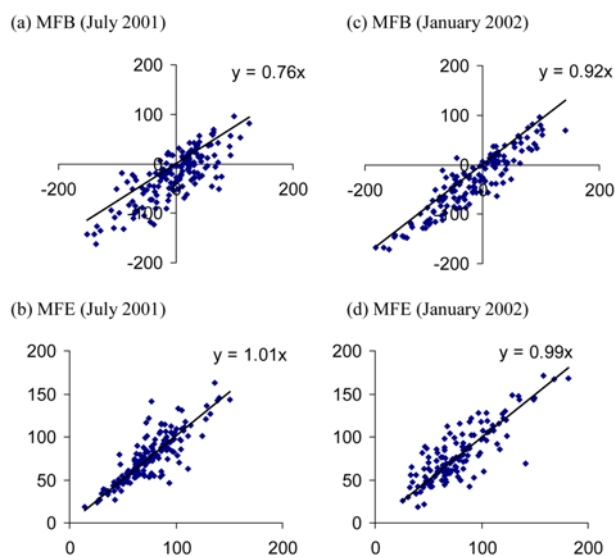


Fig. 7. Comparison of the air quality model performance. X-axis is the model performance with base emissions, and y-axis is that with adjusted emissions. The slope less than 1.0 indicates that the model error (or bias) decreases, so the model performance improves. The equations for MFB and MFE is available elsewhere (Boylan *et al.*, 2006).

tory, which is prepared using the bottom-up approach. Here, NEI2001 and NEI2002 prepared by the bottom-up approach (US-EPA, 2012k, 2012l) are compared with the adjusted emissions from the FDDA. Since NEI99 projected to 2001 are the original emissions used as the starting guess for 2001, the ratio of ‘NEI2001’ to ‘NEI99 projected to 2001’ becomes the emission adjustment through the bottom-up approach.

The ratio of ‘NEI2001’ to ‘NEI99 projected to 2001’ for each source is illustrated in Fig. 8a and that of ‘NEI 2002’ to ‘NEI99 projected to 2002’ is represented in Fig. 8b. The ratio calculated in Fig. 8 is based on the annual emissions, while the emission adjustment scaling factor represented in Fig. 2 is based on the monthly emissions. Thus, please note that the exact comparison of emission adjustment of the FDDA with the bottom-up should not be made based on Figs. 2 and 8. Results show that, in general, the emission adjustment from the bottom-up approach is relatively mild compared with that from the FDDA (Fig. 8). However, an obvious decrease of the adjusted emissions from the bottom-up approach for $PM_{2.5}$ in January 2002 (Fig. 8) is consistent with the results from the FDDA for PEC and PMFINE in January 2002 (Fig. 2).

4. Concluding remarks

Emission strengths biases were identified using FDDA. Regionally separate scaling factors were calculated to account for the regionally different biases. Further, a separate scaling factor is calculated for weekdays and weekends to capture the day-of-week variation in the emission bias. This study considered the adjustment of the strength for ten emission sources: CO (total), NH_3 (total), SO_2 (total), NO_x (area/mobile/nonroad),

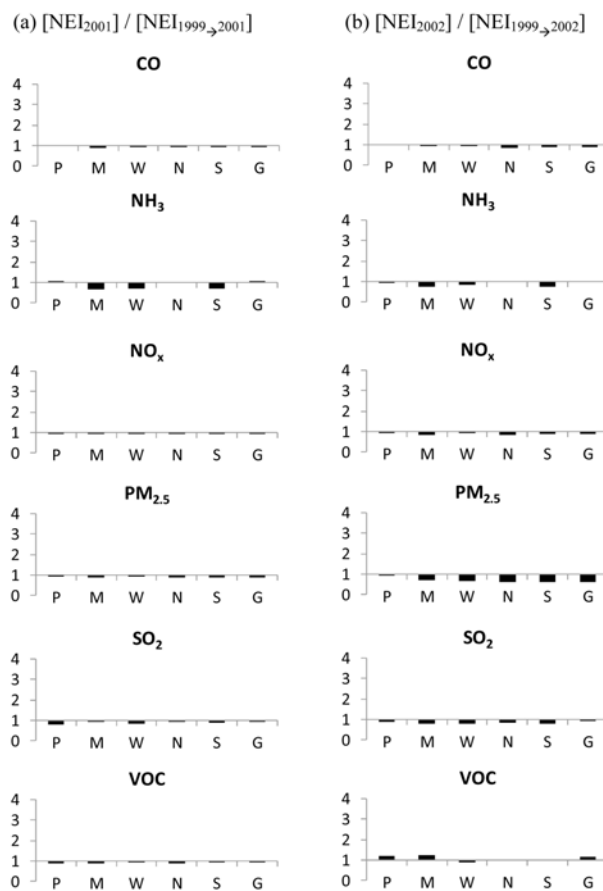


Fig. 8. (a) is the ratio of ‘NEI2001’ to ‘NEI1999 projected to 2001’, and (b) is the ratio of ‘NEI2002’ to ‘NEI1999 projected to 2002’ for each region.

NO_x (point), VOC (area/mobile/nonroad/point), VOC (biogenic), POA (total), and PMFINE (total).

Results showed that base emissions for CO and SO_2 sources were estimated reasonably well, while emissions for NH_3 , NO_x , PEC and PMFINE, POA, and VOC require more significant revision. After assimilation, NH_3 , NO_x , PEC, and PMFINE emissions decreased in overall. The adjustment of POA and VOC emissions was significantly different among regions. The different model performance between VOC species complicated the VOC emission assimilation. This suggests that a modification of speciation profile of VOC is needed. The difference in the emission adjustment between weekdays and weekends was minor in most cases. The model performance was relatively better in July 2001 compared with that in January 2002. Those differences were partly due to the temporal allocation used in SMOKE. The emission adjustment improved the air quality model performance. However, this method could not improve the VOC speciation profile or temporal allocation of the emissions since the assimilation was performed on the annual emissions inventory. Further application of the FDDA to the speciated and temporally allocated emissions could improve the speciation profile or temporal allocation used to process emissions in SMOKE as well.

The importance of the emission adjustment has been discussed in previous studies. When NO_x and VOC emissions are adjusted, the ratio of the VOC to NO_x can change from 2 (NO_x scavenges O_3) to 11 (O_3 production is favored by NO_x) (Chang *et al.*, 1996). Thus, the emission control strategy needed to be modified with the changes of the sensitivity of O_3 to NO_x emissions after the adjustment (Mendoza-Dominguez and Russell, 2001b). Also, the relative importance of the contributed mass to $\text{PM}_{2.5}$ changes as the emission adjustment is performed. For example, if the PEC emissions are adjusted negatively and POA emissions are adjusted positively, the resulting mass contributions from diesel exhaust (one of the large source of PEC emissions) would decrease, whereas those from meat cooking combustion (one of the large source of POA emissions) would increase. Therefore, accurate emissions increase the reliability of the air quality model, and thus enhance the confidence of further applications.

Acknowledgements. This research was supported by the U.S. Environmental Protection Agency under Agreements RD-82897602, RD83107601, RD83096001, and RD83479901.

Edited by: Rokjin Park

REFERENCES

- Aldrin, M., 1997: Length modified ridge regression. *Comput. Stat. Data An.*, **25**, 377-398.
- Ames, R. B., and W. C. Malm, 2001: Chemical species' contribution to the upper extremes of aerosol fine mass. *Atmos. Environ.*, **35**, 5193-5204.
- Abdel-Aziz, A., and H. C. Frey, 2004: Propagation of uncertainty in hourly utility NO_x emissions through a photochemical grid air quality model: A case study for the Charlotte, NC, modeling domain. *Environ. Sci. Technol.*, **38**, 2153-2160.
- Boylan, J. W., M. Odman, J. Wilkinson, and A. G. Russell, 2006: Integrated Assessment Modeling of Atmospheric Pollutants in the Southern Appalachian Mountains: Part 2. Fine Particulate Matter and Visibility. *J. Air Waste Manage. Assoc.*, **56**, 12-22.
- Butler, A. J., M. S. Andrew, and A. G. Russell, 2003: Daily sampling of $\text{PM}_{2.5}$ in Atlanta: results of the first year of the assessment of spatial aerosol composition in Atlanta study. *J. Geophys. Res.-Atmos.*, **108**, D1, doi:10.1029/2002JD002234.
- Byun, D. W., and J. K. Ching, 1999: Science algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) modeling system. US-EPA/600/R-99/030
- Carter, W. P. L., 2004: Development of a Chemical Speciation Database and Software for Processing VOC Emissions for Air Quality Models. 13th International Emission Inventory Conference, "Working for Clean Air in Clearwater", Hilton Clearwater Beach Resort, Clearwater, Florida USA, June 8-10, 2004. [Available online at <http://www.epa.gov/ttnchie1/conference/ei13/modeling/carter.pdf>]
- Census, cited 2012. Census Regions and Divisions of the United States. [Available online at http://www.census.gov/geo/www/us_regdiv.pdf]
- Chang, M. E., D. E. Hartley, C. Cardelino, and W. L. Chang, 1996: Inverse modeling of biogenic isoprene emissions. *Geophys. Res. Lett.*, **23**, 3007-3010.
- Di, T., D. S., Cohan, N. Napelenok, M. S. Bergin, Y. T. Hu, M. Chang, A. G. Russell, 2010: Uncertainty analysis of ozone formation and response to emission controls using higher-order sensitivities. *J. Air Waste Manage. Assoc.*, **60**, 797-804.
- Dunker A. M., 1981: Efficient calculation of sensitivity coefficients for complex atmospheric models. *Atmos. Environ.*, **15**, 1155-1161.
- Emery, C., E. and Tai, 2001: Enhanced meteorological modeling and performance evaluation for two Texas ozone episodes. project report prepared for the Texas Natural Resources Conservation Commission. Prepared by ENVIRON, International Corporation, Novato, CA, U.S.A.
- Frank, I. E., and J. H. Friedman, 1993: A statistical view of some chemometrics regression tools. *Technometrics.*, **35**, 109-148.
- Gilliland, A. B., R. L. Dennis, S. J. Roselle, and T. E. Pierce, 2003: Seasonal NH_3 emission estimates for the eastern United States based on ammonium wet concentrations and an inverse modeling method. *J. Geophys. Res.-Atmos.*, **108**, doi:10.1029/2002JD003063.
- Hansen, D. A., E. S. Edgerton, B. E. Hartsell, J. J. Jansen, N. Kandasamy, G. M. Hidy, and C. L. Blanchard, 2003: The southeastern aerosol research and characterization study: Part 1 - Overview. *J. Air Waste Manage. Assoc.*, **53**, 1460-1471.
- Harley, R. A., A. G. Russell, G. J. McRae, G. R. Cass, and J. H. Seinfeld, 1993: Photochemical modeling of the Southern California air-quality study. *Environ. Sci. Technol.*, **27**, 378-388.
- Hoerl, A. E., and R. W. Kennard, 1970: Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics.*, **12**, 55-67.
- Hogrefe, C., G. Sistla, E. Zalewsky, W. Hoa, and J. Y. Ku, 2003: An assessment of the emissions inventory processing systems EMS-2001 and SMOKE in grid-based air quality models. *J. Air Waste Manage. Assoc.*, **53**, 1121-1129.
- Hu, Y., S. L. Napelenok, M. T. Odman, and A. G. Russell, 2009a: Sensitivity of inverse estimation of 2004 elemental carbon emissions inventory in the United States to the choice of observational networks. *Geophys. Res. Lett.*, **36**, L15806, doi:10.1029/2009GL039655.
- _____, M. T. Odman, and A. G. Russell, 2009b: Top-down analysis of the elemental carbon emissions inventory in the United States by inverse modeling using Community Multiscale Air Quality model with decoupled direct method (CMAQ-DDM). *J. Geophys. Res.-Atmos.*, **114**, D24302, doi:10.1029/2009JD011987.
- Koch, S.E., M. Desjardins, and P. J. Kocin, 1983: An Interactive Barnes objective map analysis scheme for use with satellite and conventional data. *J. Climate Appl. Meteor.*, **22**, 1487-1503.
- Mannschreck, K., D. Klemp, D. Kley, R. Friedrich, J. Kuhlwein, B. Wickert, P. Matuska, M. Habram, and F. Slemr, 2002: Evaluation of an emission inventory by comparisons of modeled and measured emission rates of individual HCs, CO and NO_x . *Atmos. Environ.*, **36**, S81-S94.
- Marquardt, D. W., 1963: An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Indust. Appl. Math.*, **11**, 431-441.
- Mendoza-Dominguez, A., and A. G. Russell, 2000: Iterative inverse modeling and direct sensitivity analysis of a photochemical air quality model. *Environ. Sci. Technol.*, **34**, 4974-4981.
- _____, and _____, 2001a: Emission strength validation using four-dimensional data assimilation: Application to primary aerosol and precursors to ozone and secondary aerosol. *J. Air Waste Manage. Assoc.*, **51**, 1538-1550.
- _____, and _____, 2001b: Estimation of emission adjustments from the application of four-dimensional data assimilation to photochemical air quality modeling. *Atmos. Environ.*, **35**, 2879-2894.
- Menke, W., 1989: *Geophysical data analysis, vol. 45: Discrete inverse theory*. Academic Press, 289 pp.
- Muller, J. F., T. Stavrou, 2005: Inversion of CO and NO_x emissions using the adjoint of the IMAGES model. *Atmos. Chem. Phys.*, **5**, 1157-1186.
- Odman, M. T., A. G. Russell, 1991: Multiscale modeling of pollutant transport and chemistry. *J. Geophys. Res.-Atmos.*, **96**, D4, doi:10.1029/91JD00387.
- Pace, T. G., 2003: A conceptual model to adjust fugitive dust emissions to

- account for near source particle removal in grid model applications. *Proc. of Joint Meeting of WRAP Emissions & Fugitive Dust Forums*, Las Vegas, NV, U.S.A.
- Park, R., D. J. Jacob, B. D. Field, R. M. Yantosca, and M. Chin, 2004: Natural and transboundary pollution influences on sulfate-nitrate-ammonium aerosols in the United States: Implication for policy. *J. Geophys. Res.-Atmos.*, **109**, D15204, doi:10.1029/2003JD004473.
- Park, S.-K., C. E. Cobb, K. Wade, J. Mulholland, Y. Hu, A. G. Russell, 2006: Uncertainty in air quality model evaluation from spatial variation. *Atmos. Environ.*, **40**, S563-S573.
- _____, A. Marmur, A. G. Russell, 2013: Environmental risk assessment: comparison of receptor and air quality models for Source apportionment. *Hum. Ecol. Risk Assess. in press*.
- Petron, G., C. Granier, B. Khattatov, V. Yudin, J. F. Lamarque, L. Emmons, J. Gille, D. P. Edwards, 2004: Monthly CO surface sources inventory based on the 2000-2001 MOPITT satellite data. *Geophys. Res. Lett.*, **31**, L21107, doi:10.1029/2004GL020560.
- Placet, M., C. O. Mann, R. O. Gilbert, M. J. Niefer, 2000: Emissions of ozone precursors from stationary sources: a critical review. *Atmos. Environ.*, **34**, 2183-2204.
- PSU/NCAR, cited 2012: PSU/NCAR Mesoscale Modeling System Tutorial Class Notes and User's Guide: MM5 Modeling System Version 3. Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research. [Available online at <http://www.mmm.ucar.edu/mm5/documents/tutorial-v3-notes.html>]
- Russell, A., R. Dennis, 2000: NARSTO critical review of photochemical models and modeling. *Atmos. Environ.*, **34**, 2283-2324.
- Sakulyanontvittaya, T., T. Duhl, C. Wiedinmyer, D. Helmig, S. Matsunaga, M. Potosnak, J. Milford, A. Guenther, 2008: Monoterpene and Sesquiterpene emission estimates for the United States. *Environ. Sci. Technol.*, **42**, 1623-1629.
- Sawyer, R. F., R. A. Harley, S. H. Cadle, J. M. Norbeck, R. Slott, H. A., Bravo, 2000: Mobile sources critical review: 1998 NARSTO assessment. *Atmos. Environ.*, **34**, 2161-2181
- Stavrakou, T., J. F. Muller, 2006: Grid-based versus big region approach for inverting CO emissions using Measurement of Pollution in the Troposphere (MOPITT) data. *J. Geophys. Res.-Atmos.*, **111**, D15304, doi:10.1029/2005JD006896.
- Taghavi, M., S. Cautenet, and J. Arteta, 2005: Impact of a highly detailed emission inventory on modeling accuracy. *Atmos. Res.*, **74**, 65-88.
- UCAR, cited 2012: TDL U.S. and Canada surface hourly observations. [Available online at <http://dss.ucar.edu/datasets/ds472.0/>]
- Unal A., Y. T. Hu, M. E. Chang, M. T. Odman, A. G. Russell, 2005: Airport related emissions and impacts on air quality: Application to the Atlanta International Airport. *Atmos. Environ.* **39**, 5787-5798
- US-EPA, cited 2012a: Ozone and Photochemical Assessment Monitoring Stations (PAMS) program. [Available online at <http://www.epa.gov/ttn/amtic/pamsmain.html>.]
- _____, cited 2012b: SMOKE User's manual. [Available online at <http://www.smoke-model.org>]
- _____, cited 2012c: Technology Transfer Network: Clearinghouse for Inventories & Emissions Factors. [Available online at <http://www.epa.gov/ttn/chief/>]
- _____, cited 2012d: North American Emissions Inventories - Mexico. [Available online at <http://www.epa.gov/ttn/chief/net/mexico.html>]
- _____, cited 2012e: North American Emissions Inventories - Canada. [Available online at <http://www.epa.gov/ttn/chief/net/canada.html>]
- _____, cited 2012f: Biogenic Emissions Inventory System (BEIS) Modeling. [Available online at <http://www.epa.gov/asmdner1/biogen.html>]
- _____, cited 2012g: Continuous Emission Monitoring - Information, Guidance, etc. [Available at <http://www.epa.gov/ttn/emc/cem.html>]
- _____, cited 2012h: Emissions Modeling Clearinghouse Temporal Allocation. [Available at <http://www.epa.gov/ttnchie1/emch/temporal/>]
- _____, cited 2012i: Emissions Modeling Clearinghouse Speciation. [Available at <http://www.epa.gov/ttnchie1/emch/speciation/>]
- _____, cited 2012j: EGAS Version 5.0. [Available at <http://www.epa.gov/ttnchie1/egas5.htm>]
- _____, cited 2012k: Criteria Pollutant Emissions Summary Files [Available at <http://www.epa.gov/ttn/chief/net/critsummary.html>]
- _____, cited 2012l: 2002 National Emissions Inventory Data & Documentation. [Available at <http://www.epa.gov/ttn/chief/net/2002inventory.html>]
- Vautard, R., and Coauthors, 2003: Paris emission inventory diagnostics from ESQUIF airborne measurements and a chemistry transport model. *J. Geophys. Res.-Atmos.*, **108**, D17, doi:10.1029/2002JD002797.
- Warneke C., and Coauthors, 2010: Biogenic emission measurement and inventories determination of biogenic emissions in the eastern United States and Texas and comparison with biogenic emission inventories. *J. Geophys. Res.-Atmos.*, **115**, D00F18, doi:10.1029/2009JD012445.
- Yang Y. J., J. G. Wilkinson, and A. G. Russell, 1997: Fast, direct sensitivity analysis of multidimensional photochemical models. *Environ. Sci. Technol.* **31**, 2859-2868.