Atmosphere **2012**, *3*, 400-418; doi:10.3390/atmos3030400

OPEN ACCESS

atmosphere

ISSN 2073-4433

www.mdpi.com/journal/atmosphere

Article

Evaluation of the Variability in Chemical Transport Model Performance for Deposition and Ambient Concentrations of Nitrogen and Sulfur Compounds

Bonyoung Koo [1,*], Piti Piyachaturawat [1], Ralph Morris [1] and Eladio Knipping [2]

[1] ENVIRON International Corporation, Novato, CA 94998, USA;
E-Mails: ppiyachaturawat@environcorp.com (P.P.); rmorris@environcorp.com (R.M.)

[2] Electric Power Research Institute, Washington, DC 20036, USA; E-Mail: eknippin@epri.com

* Author to whom correspondence should be addressed; E-Mail: bkoo@environcorp.com;
Tel.: +1-415-899-0700; Fax: +1-415-899-0707.

Received: 14 June 2012; in revised form: 6 July 2012 / Accepted: 26 July 2012 / Published: 10 August 2012

Abstract: Air quality models are increasingly used to develop estimates of dry and wet deposition of sulfate and nitrate in watersheds (because of lack of measurements) in an effort to determine the acidifying deposition load into the aquatic systems. These models need to be rigorously evaluated to ensure that one can rely on the modeled quantities instead of the measured quantities. In the United State (U.S.), these models have been proposed for use in establishing national standards based on modeled quantities. The U.S. Environmental Protection Agency (EPA) is considering aquatic acidification as the main ecological endpoint of concern in determining the secondary national ambient air quality standards for nitrogen oxides and sulfur oxides. Acidification is tied to depositions of sulfur and nitrogen, which are linked to ambient concentrations of the elements. As EPA proposes to use a chemical transport model in linking deposition to ambient concentration, it is important to investigate how the currently used chemical transport models perform in predicting depositions and ambient concentrations of relevant chemical species and quantify the variability in their estimates. In this study, several annual simulations by multiple chemical transport models for the entire continental U.S. domain are evaluated against available measurement data for depositions and ambient concentrations of sulfur oxides and reactive nitrogen species. The model performance results vary by evaluation time-scale and geographical region. Evaluation of annualized quantities (annual average ambient concentrations and annual total depositions) suppresses the large variances shown in the evaluation using the observation's native shorter-term time-scales (e.g., weekly). In

addition, there is a large degree of bias and error (especially for deposition fluxes) in the modeling results that brings to question the suitability of using air quality models to provide estimates of deposition loads. Variability in the ratio of deposition to ambient concentration, so-called the Transference Ratio that EPA has proposed to use in linking deposition to ambient concentration, is also examined. Our study shows that the Transference Ratios as well as total reduced nitrogen deposition, another modeled parameter EPA proposed to use in the process of determining the new secondary standard, vary considerably by geographical region and by model simulation.

Keywords: sulfur oxides; nitrogen oxides; deposition; air quality standard; chemical transport model; transference ratio

---

1. Introduction

The current secondary (welfare-based) National Ambient Air Quality Standard (NAAQS) for oxides of nitrogen (NO_X) is 0.053 ppm (53 ppb or 100 μg/m^3), calculated as the annual arithmetic average of the 1-hour NO_2 concentrations. The standard, which was selected to provide protection to the public welfare against acute injury to vegetation from direct exposure and resulting phytoxicity, was originally set to identical to the primary standard (to protect public health) in 1971. While the U.S. Environmental Protection Agency (EPA) has recently strengthened the primary standard by establishing a new 1-hour standard at a level of 100 ppb based on the 3-year average of the 98th percentile of the yearly distribution of 1-hour daily maximum concentrations, the secondary standard remains as it was set in 1971. The current secondary standard for oxides of sulfur (SO_X) was also set in 1971 and uses SO_2 as the atmospheric indicator. The standard of 0.5 ppm, based on a 3-hour average SO_2 concentration, is not to be exceeded more than once per year.

EPA has recently conducted a review of the secondary NAAQS for NO_X and SO_X and argued that there is now sufficient information available to support the concept that the largest impact to public welfare from SO_X and NO_X are the ecological effects associated with the deposition of nitrogen and sulfur compounds to terrestrial and aquatic environments [1–4]. Based on the findings from this review, EPA has concluded that the current secondary NO_X and SO_X standards are not adequate to provide sufficient protection against adverse ecological effects associated with deposition of oxides of nitrogen and sulfur to sensitive ecosystems. Three issues were identified as main concerns of the current structure of the standards:

- Ecological effects related to deposition are usually associated with depositional loads occurring over periods of months to years, which differ significantly from the short-term concentrations used by the current standards;
- The current standards, which use NO_2 and SO_2 as indicators, address only a fraction of total atmospheric oxides of nitrogen and sulfur because they do not capture all relevant chemical species of nitrogen and sulfur that contribute to deposition; and

- Under the current standards, NO_2 and SO_2 are assessed individually even though they play a joint role in acidifying aquatic ecosystems.

Addressing these issues, EPA proposed the following as the elements of the new secondary NAAQS:

- *Indicator*: The concentration of all the major oxides of sulfur (SO_X = SO_2 + particulate SO_4) and nitrogen (NO_Y = NO + NO_2 + HNO_3 + PAN + 2 N_2O_5 + HONO + NO_3 + organic nitrates + particulate NO_3);
- *Form*: The Aquatic Acidification Index (AAI) designed to be an ecologically relevant form of the standard that determines the allowable levels of ambient NO_Y and SO_X;
- *Averaging Time*: An annual averaging time based on the average of each year over a consecutive 3 to 5 year period; and
- *Level*: A target chronic Acid Neutralizing Capacity (ANC) value in range of 20 to 70 μeq/L.

The Clean Air Scientific Advisory Committee (CASAC) reviewed EPA's assessment and agreed that aquatic acidification should be the focus for developing a secondary NO_X and SO_X NAAQS because of the quantity and quality of the ecological effects data [5,6].

The complexities in the responses of ecosystems to deposition of atmospheric NO_Y and SO_X require a more sophisticated form of the standard than the current structure. The conceptual design of the new secondary NAAQS proposed by EPA consists of three main components: (1) linkage between ecological indicators and ecological effects; (2) linkage between an ecological indicator and atmospheric deposition; and (3) linkage between deposition and ambient air indicators. With respect to linking ecological indicator to adverse effects of aquatic acidification, EPA proposed to use ANC which is the most widely used indicator of acid sensitivity and has been found in various studies to be the best single indicator of the biological response and health of aquatic communities in acid sensitive systems. To link atmospheric deposition to the ecological indicator, ANC, EPA proposed to use ecosystem acidification models that quantify the relationship between deposition of nitrogen and sulfur and the resulting ANC in surface waters based on an ecosystem's inherent generation of ANC and ability to neutralize nitrogen deposition through biological and physical processes. Finally, the linkage between deposition and ambient concentrations of NO_Y and SO_X introduces a new quantity termed as Transference Ratio (T) which is defined as the ratio of total wet and dry deposition to concentration:

$$\mathrm{TSO_X} = \frac{\mathrm{Dep(SO_X)}}{\mathrm{[SO_X]}} \tag{1}$$

$$\mathrm{TNO_Y} = \frac{\mathrm{Dep(NO_Y)}}{\mathrm{[NO_Y]}} \tag{2}$$

where Dep(SO_X or NO_Y) is the annual total wet and dry deposition of SO_X or NO_Y, and [SO_X or NO_Y] is the annual average concentration of ambient SO_X or NO_Y. Due to lack of appropriate measurement data, EPA proposed to use EPA's Community Multiscale Air Quality (CMAQ) modeling system [7] to calculate the Transference Ratios.

EPA defines the AAI in terms of four ecological and atmospheric factors and ambient air indicators (NO_Y and SO_X) as follows:

$$\mathrm{AAI} = \mathrm{F1} - \mathrm{F2} - \mathrm{F3[NO_Y]} - \mathrm{F4[SO_X]} \tag{3}$$

where F1 represents the ecosystem's natural neutralizing capability; F2 represents acidifying depositions associated with reduced nitrogen (NH_X) based on CMAQ model simulations; F3 and F4 are the Transference Ratios to convert ambient NO_Y and SO_X concentrations to nitrogen and sulfur deposition, respectively, which is also based on CMAQ modeling. The value of F1 would be based on a representative runoff rate as well as a representative critical load (the amount of acidifying atmospheric deposition of nitrogen and sulfur beyond which a target ANC is not reached) for the ecosystem associated with a single national target ANC level. The values of F1–F4 will vary spatially and thus an appropriate regionalization scheme is necessary for spatial aggregation (or averaging). More details on each of the above F factors are given elsewhere [4].

As described above, determination of the AAI value (and establishment of the new secondary NAAQS) relies heavily on use of chemical transport models. Therefore, it is important to investigate how the currently used chemical transport models such as CMAQ and the Comprehensive Air-quality Model with Extensions (CAMx) [8] perform in predicting depositions and ambient concentrations of relevant chemical species and quantify the variability in their estimates. The CMAQ modeling system has been periodically given updates, and for each new release EPA performed a comprehensive operational evaluation mostly focusing on eastern U.S. [9–13]. CAMx model performance has also been investigated in various U.S. regions [14–17]. While providing detailed analysis on the model performance, conventional model evaluation studies typically focus on single model application.

The Transference Ratios, TSO_X and TNO_Y, are calculated from model simulations and are key parameters in determining AAI. Another key parameter in the AAI formulation is total reduced nitrogen loading (LNH_X, *i.e.*, the term F2 in the AAI equation) that is also based on model simulations. EPA suggested that these parameters are sufficiently stable across time and space based on the CMAQ modeling results over the Adirondacks and Shenandoah case study areas [4]. We examined variability in these parameters over various regions in contiguous U.S. using several annual model simulation datasets.

2. Modeling Databases

Several existing chemical transport model simulation outputs were selected based on two conditions: (1) since the new secondary NAAQS proposed by EPA is based on annualized depositions and ambient concentrations, the modeling period should cover at least a year; (2) both the concentration and deposition outputs should be available. Since many of the historical model runs were performed to address concentration issues (e.g., ozone, $PM_{2.5}$ and regional haze SIPs), the deposition outputs were often not archived. Table 1 summarizes the modeling datasets included in this study. The compiled datasets show variations in the model, emission/meteorological modeling year and grid resolution used. Different models may employ different algorithms for the same atmospheric process. For example, CAMx adopted a dry deposition model based on the approaches of Wesely [18] and Slinn and Slinn [19] while CMAQ uses the M3 dry deposition model by Pleim *et al*. [20,21]. Details on the science algorithms used in the modes can be found in the references given below Table 1 and references therein. Figure 1 shows modeling domains used in these datasets. Each of the VISTAS and UBAQS modeling includes nested grid simulation with 12-km resolution (while covering a different region) in addition to the 36-km continental U.S. (CONUS) modeling domain. The FCAQTF modeling

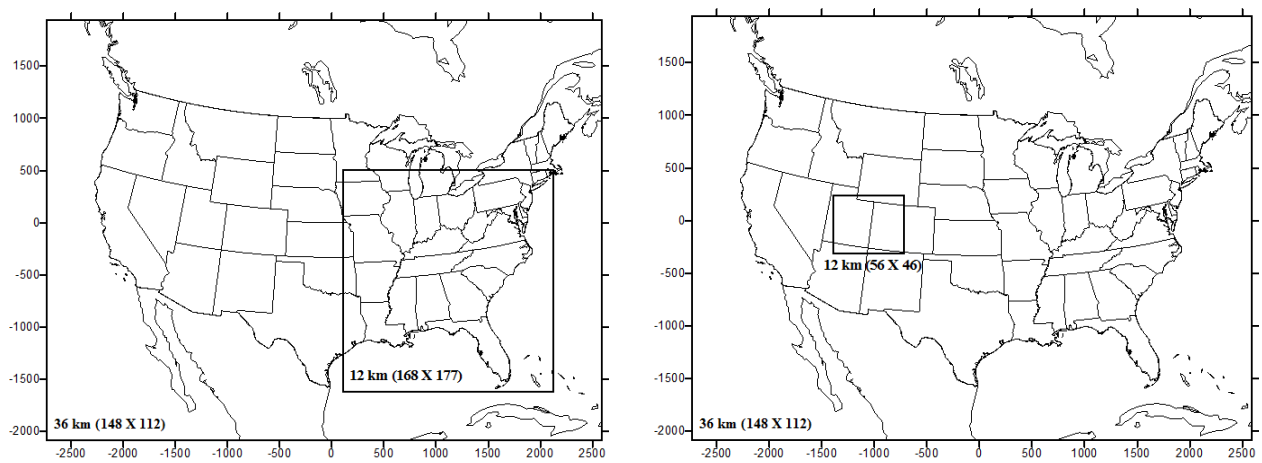
domain consists of 12-km and 4-km nested grids covering southwestern U.S. and the Four Corners region, respectively, as well as the common 36-km CONUS domain.

Table 1. Annual chemical transport model datasets compiled and used for this study.

Study	Model (Version)	Emiss/Met Year	Domain (Grid Resolution)		
			Grid #1	Grid #2	Grid #3
VISTAS [1]	CMAQ (V4.5.1_SOAmods)	2002/2002 2009/2002 2012/2002 2018/2002	CONUS (36 km) [6]	SE US (12 km) [7]	-
UBAQS [2]	CMAQ (V4.6)	2005/2005 2006/2006 2012/2005 2012/2006	CONUS (36 km)	UT-CO (12 km) [8]	-
EPRI [3]	CMAQ (AMSTERDAM)	2002/2002	-	E US (12 km) [9]	-
General [4]	CAMx (V5.21)	2005/2005 2006/2006	CONUS (36 km)	-	-
FCAQTF [5]	CAMx (V4.51)	2018/2005	CONUS (36 km)	SW US (12 km) [10]	4Corners (4 km) [11]

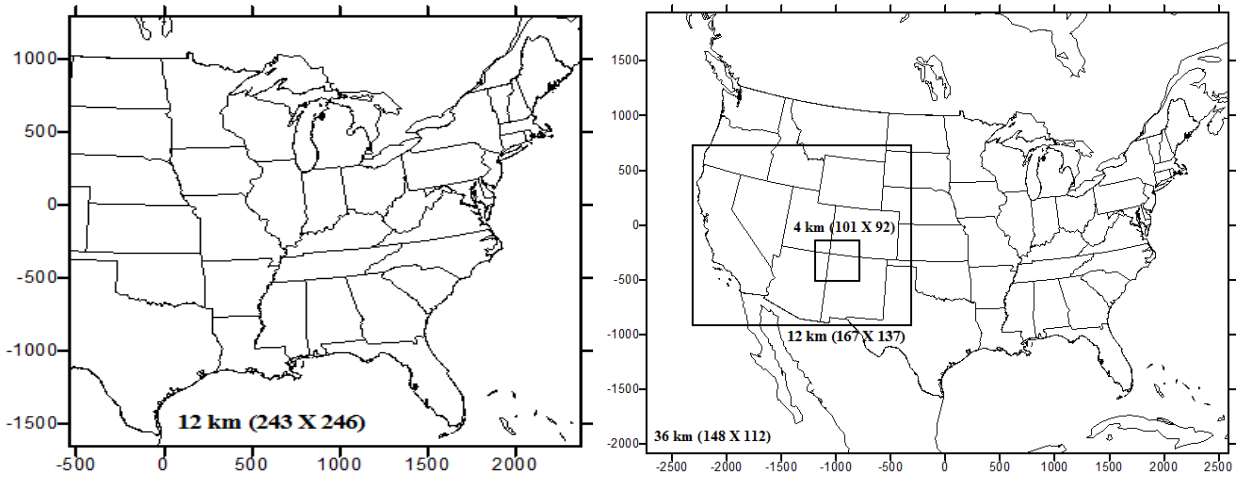
[1] Visibility Improvements State and Tribal Association of the Southeast (VISTAS) regional haze SIP modeling study [22]; [2] Uinta Basin Air Quality Study (UBAQS) [23]; [3] Electric Power Research Institute (EPRI) project for development and application of the Advanced Modeling System for Transport, Emissions, Reactions and Deposition of Atmospheric Matter (AMSTERDAM) [24]; [4] CAMx modeling over the 36-km continental U.S. domain for various applications including Denver SIP modeling study [25]; [5] Four Corners Air Quality Task Force (FCAQTF) modeling study for the Four Corners region [26]; [6] Continental U.S. 36-km modeling domain (see Figure 1(a)); [7] Southeastern U.S. 12-km modeling domain (see Figure 1(a)); [8] Utah-Colorado 12-km modeling domain (see Figure 1(b)); [9] Eastern U.S. 12-km modeling domain (see Figure 1(c)); [10] Southwestern U.S. 12-km modeling domain (see Figure 1(d)); [11] Four Corners 4-km modeling domain (see Figure 1(d)).

Figure 1. Modeling domains for (**a**) the Visibility Improvements State and Tribal Association of the Southeast (VISTAS) regional haze State Implementation Plan (SIP) modeling study; (**b**) the Uinta Basin Air Quality Study (UBAQS) modeling study; (**c**) the Electric Power Research Institute (EPRI) modeling study; and (**d**) the Four Corners Air Quality Task Force (FCAQTF) modeling study.



(**a**) VISTAS modeling domains (**b**) UBAQS modeling domains

Figure 1. *Cont.*



(**c**) EPRI modeling domain (**d**) FCAQTF modeling domains

3. Model Performance Evaluation

As mentioned earlier, the proposed new standard relies extensively on modeling results rather than monitoring data, therefore, a rigorous evaluation of model performance is crucial. While EPA has conducted performance evaluation for their CMAQ modeling over the continental U.S. modeling domain (the 2002 annual simulation with CMAQ v4.6 and the 2002 through 2005 simulations with CMAQ v4.7) as part of their review of the secondary NAAQS for NO_X and SO_X [3,4], they focused on statistics based on annualized quantities and/or averaged over large regions. For this study, we performed more comprehensive model evaluation using several annual CMAQ and CAMx simulations for the 36-km CONUS modeling domain, and also examined how different statistics affect the evaluation results.

Several ambient and deposition monitoring networks were used for the model performance evaluation. Ambient concentrations and dry deposition of gaseous and particulate sulfur and nitrogen compounds (SO_2, HNO_3, NH_3, particulate sulfate and nitrate) were evaluated against measurements from the Clean Air Status and Trends Network (CASTNET). Note that CASTNET reports dry deposition fluxes estimated by an inferential method which calculates the fluxes as the product of a locally-measured atmospheric concentration and a modeled deposition velocity [27]. The deposition velocity is calculated using the Multi-Layer Model (MLM) where the vegetated canopy is discretized into multiple layers and stomatal and boundary layer resistances are calculated for each of these layers [28]. Wet deposition of total (gaseous and particulate) sulfate, nitrate and ammonium were evaluated against measurements from the National Trends Network (NTN) of the National Atmospheric Deposition Program (NADP). Standard sampling interval for both CASTNET and NADP is one week (weekly average concentrations and weekly total deposition). Modeled NO_Y concentrations were evaluated against hourly NO_Y measurements from the Air Quality System (AQS) database. Additionally, hourly NO_Y measurements at eight monitoring stations in the southeastern states (Alabama, Florida, Georgia and Mississippi) from the Southeastern Aerosol Research and Characterization (SEARCH) study were also included in the evaluation.

Table 2 summarizes performance of the selected model simulations using both annualized quantities (annual average for concentrations and annual total for deposition) and those with the monitoring network's native sampling intervals (one week for the CASTNET and NADP monitors; one hour for the AQS and SEARCH monitors). Additional model performance metrics are given in the Supplementary Material (see Tables S1–S3). In general, model performance for ambient concentrations is relatively better than that for depositions with dry deposition performance being especially poor. As noted earlier, CASTNET's dry deposition data is not "true" measurements, but estimates based on an "inferential model" involving measured air concentrations coupled with species- and location-dependent deposition velocities that reflect local land use and meteorological conditions at each monitoring site. Therefore, the performance evaluation results for dry depositions should be taken with caution. However, the large discrepancies (120–230% for gas species and 70–800% for PM species in MNE with native sampling frequency) between CMAQ-/CAMx-predicted dry depositions and CASTNET estimates may indicate significant uncertainties in dry deposition modeling. Spatial variability of deposition surface is one of the main factors that contribute to these discrepancies. Deposition velocities estimated at a monitoring site may not be adequately represented with relatively large size of model's computational grid cell which would contain complex mix of various land use types. Since the deposition flux is expressed as the product of deposition velocity and ambient concentration, the uncertainty in concentration measurements will be propagated directly into the calculated fluxes. Previous field studies indicated 5–30% of biases between different measurement methods for several gas and PM species [27]. The models also exhibit particularly poor performance for wet depositions of total ammonium (NH_3 and particulate NH_4). As EPA has noted in their assessment [4], modeling ammonia deposition is a difficult task due to complexity of ammonia deposition processes (e.g., bi-directional flux of ammonia) and relatively large uncertainty in the ammonia emission inventory. EPA has attempted to improve CMAQ model performance for depositions, for example, by adjusting the CMAQ model output based on precipitation data generated by the Parameter-elevation Regressions on Independent Slopes Model (PRISM) or by incorporating a bi-directional ammonia flux algorithm in CMAQ [4]. However, these corrections should be evaluated in a more comprehensive and thorough way before being utilized in a regulatory framework.

Table 2. Model performance statistics [1] for various sulfur or nitrogen compounds using annualized quantities and those with native sampling frequency [2].

		VISTAS 2002		**UBAQS 2005** [3]		**UBAQS 2006**		**General 2005**		**General 2006**	
		MNB	**MNE**	**MNB**	**MNE**	**MNB**	**MNE**	**MNB**	**MNE**	**MNB**	**MNE**
Concentrations (CASTNET)											
SO_2	Native	33	55	23	54	47	68	67	80	96	106
	Annual	27	38	9.4	28	30	42	43	52	69	76
SO_4	Native	−16	30	−22	33	−16	31	13	45	20	45
	Annual	−22	22	−30	30	−23	24	−3.2	19	7.1	26
HNO_3	Native	36	63	−6.0	50	6.2	51	8.6	48	20	51
	Annual	20	35	−12	31	−1.3	29	−3.5	25	6.8	27
NO_3	Native	58	146	171	229	142	199	92	169	87	158
	Annual	35	97	116	127	100	112	60	91	47	77

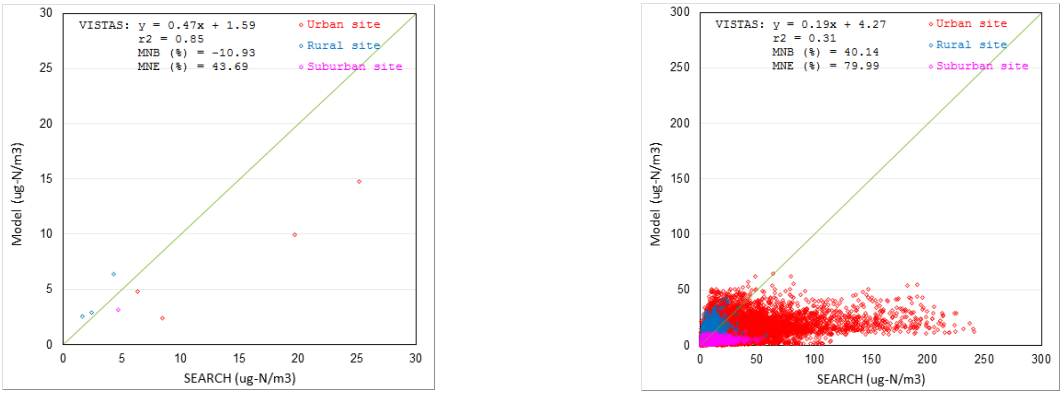
Table 2. *Cont.*

		VISTAS 2002		UBAQS 2005 [3]		UBAQS 2006		General 2005		General 2006	
		MNB	MNE	MNB	MNE	MNB	MNE	MNB	MNE	MNB	MNE
Dry deposition (CASTNET)											
SO_2	Native	148	163	176	187	205	216	196	204	224	233
	Annual	104	109	134	136	148	154	141	143	155	161
SO_4	Native	−11	105	257	279	301	321	132	151	140	157
	Annual	−58	66	112	118	123	130	45	57	55	68
HNO_3	Native	131	145	100	125	105	127	219	227	220	228
	Annual	75	81	68	75	76	83	141	142	154	157
NO_3	Native	−28	116	739	801	652	714	342	405	280	344
	Annual	−57	71	326	335	308	316	147	170	114	140
NH_4	Native	−58	74	116	156	127	161	69	106	77	107
	Annual	−72	72	35	55	49	66	12	36	22	41
Wet deposition (NADP)											
Total SO_4	Native	88	147	138	186	97	147	71	148	40	123
	Annual	48	74	75	99	27	50	33	73	6.1	45
Total NO_3	Native	13	91	41	107	30	102	−25	81	−33	79
	Annual	−4.8	52	44	88	−15	38	−3.6	99	−51	57
Total NH_4	Native	52	131	318	371	223	277	153	246	59	157
	Annual	3.5	66	26	64	11	45	−24	61	−36	51
Concentrations (AQS)											
NO_Y	Native	76	101	-	-	56	88	60	90	56	88
	Annual	41	54	-	-	29	53	44	61	29	51
Concentrations (SEARCH)											
NO_Y	Native	40	80	-	-	37	75	38	81	42	83
	Annual	−11	44	-	-	−7.4	42	−7.7	47	−3.5	47

[1] Two types of performance statistics, mean normalized bias (MNB) and mean normalized error (MNE), are shown here: $\text{MNB (\%)} = \frac{1}{N}\sum\frac{\text{Model}-\text{Observation}}{\text{Observation}} \times 100$ $\text{MNE (\%)} = \frac{1}{N}\sum\frac{|\text{Model}-\text{Observation}|}{\text{Observation}} \times 100$ where N is the number of pairs of modeled and observed quantities matched in time and space; [2] Weekly average concentrations or weekly total depositions for the CASTNET and NADP measurements; hourly average concentrations for the AQS and SEARCH measurements; [3] NO_Y is not evaluated for the UBAQS 2005 modeling dataset because results of some NO_Y species were not archived.

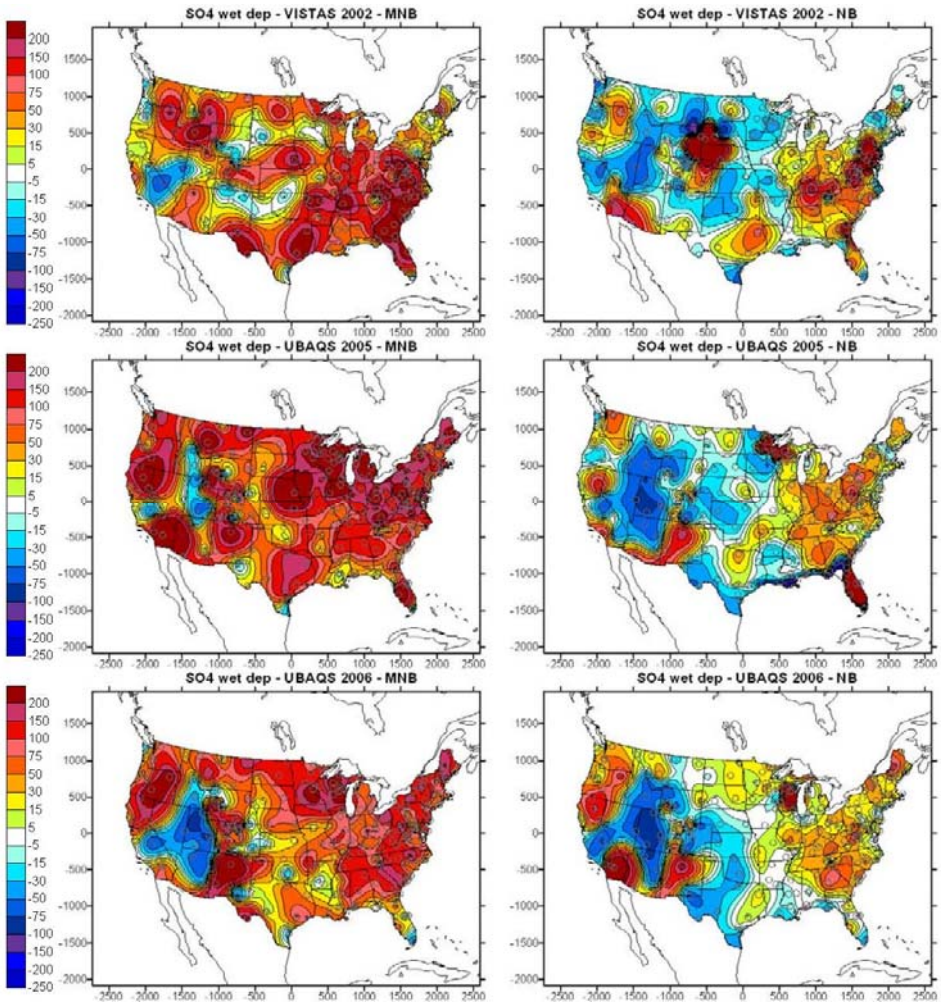
In most cases, the biases and errors with weekly or hourly quantities are poorer than those with annualized quantities (see Table 2) as using annualized quantities is subject to compensation of errors due to temporal averaging/aggregation. The effect of temporal averaging is obvious in Figure 2 where the hourly average scatter plot clearly shows model's failure to reproduce high NO_Y concentrations observed at urban and suburban sites (scatter plots for individual sites are shown in Figure S1) while the annual average scatter plot suggests more acceptable performance.

Figure 2. Scatter plots of (**a**) Annual average NO_Y concentrations and (**b**) Hourly average NO_Y concentrations: VISTAS 2002 annual Community Multiscale Air Quality (CMAQ) simulation *vs*. Southeastern Aerosol Research and Characterization (SEARCH) monitoring data.



(**a**) Annual average NO_Y conc. (μg-N/m^3) (**b**) Hourly average NO_Y conc. (μg-N/m^3)

Figure 3. Spatial maps of mean normalized bias (MNB) for modeled SO_4 wet depositions using weekly total (**left**) and annual total (**right**) deposition values from the VISTAS 2002 (**top**) and UBAQS 2005 (**middle**) and 2006 (**bottom**) CMAQ simulations. Contours are created using Kriging interpolation of values at NADP stations (circle markers).



Spatial aggregation can also affect results of model performance evaluation. Averaging across large regions (e.g., over continental U.S.) may hide spatial variations in model performance because areas of overestimation and underestimation can compensate each other. Spatial variation of the model performance is clearly demonstrated in Figure 3 which shows both overestimation and underestimation of SO_4 wet deposition by the model depending on location of the monitoring site. Factors contributing to the spatial variation in model performance include variations in geographical topography and land use/land cover. Temporal averaging in the performance statistics also affects this spatial variation. Model evaluation for weekly total wet deposition of SO_4 (left panels of Figure 3) shows that the model overestimates in most areas while exhibiting underestimations in certain regions in western and central U.S. while using annualized quantities (right panels of Figure 3) lessens level of the model overestimations in eastern U.S. and even changes sign of the biases in certain areas in central U.S. (from overestimation to underestimation). Note that since conservative estimation is more desirable in regulatory applications, model overestimation is considered more acceptable than underestimation.

4. Variability in Transference Ratios and Reduced Nitrogen Loading

To examine variability in TSO_X, TNO_Y and LNH_X, we calculated these parameters using compiled set of annual chemical transport model simulations (listed in Table 1) over various regions in U.S. We first focused our analysis on the two case study areas (Adirondacks and Shenandoah; see Figure 4) that have been extensively studied by EPA for their review of the secondary SO_X and NO_X NAAQS. Variability in TSO_X, TNO_Y and LNH_X is displayed by utilizing a "box and whisker" plot in which gray boxes represent 50% of the distribution around the median and any data points beyond whiskers are regarded as outliers. While EPA presented the inverse of TSO_X and TNO_Y in their variability analysis, we directly plotted these quantities as they were defined. We focused our analysis on the data range within the top and bottom whisker ends excluding outliers. Figure 5 presents variability in the TSO_X, TNO_Y, and LNH_X values across the grid cells within each case study area calculated by all the model simulations considered in this study. Both TSO_X and TNO_Y vary considerably. At Adirondacks, the TSO_X values range from 0.56 to 2.3 cm/s (varying by a factor of ~4) and TNO_Y from 0.24 to 1.5 cm/s (varying by a factor of ~6). TSO_X and TNO_Y at Shenandoah show similar variability: The TSO_X values spans from 0.35 to 1.3 cm/s (by a factor of ~4) and TNO_Y from 0.22 to 0.96 (by a factor of ~4). LNH_X changes its value by factors of 4 and 5 at Adirondacks and Shenandoah, respectively.

Figure 4. Regions used for analysis of variability in TSO_X, TNO_Y and LNH_X. The Adirondack region includes 18 grid cells in a 36-km modeling grid and 164 cells in a 12-km grid. The Shenandoah region has 23 cells in a 36-km grid and 207 cells in a 12-km grid.

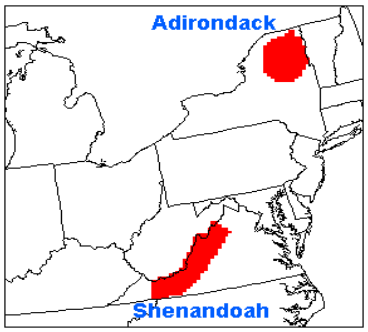
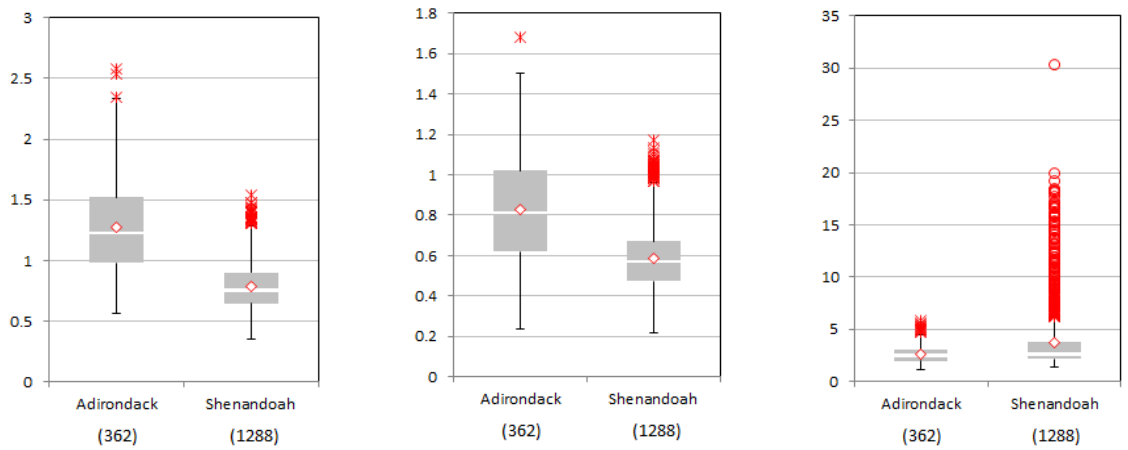


Figure 5. Box and whisker plots for variability of (**a**) TSO_X; (**b**) TNO_Y; and (**c**) LNH_X across all the grid cells at each case study area and all the model simulations. The number in parentheses represents the number of each sample data. The inner quartiles of each data sample are represented by a stack of two gray boxes, separated at the median by a thin line. The height of the gray boxes together makes up the interquartile range (IQR). The range of data falling within 1.5 IQRs of the median is represented by whiskers. Any outliers that fall between 1.5 and 3 IQRs from the median are represented by asterisk markers, while any outliers falling beyond 3 IQRs from the median are represented by circular markers. The average of each sample is represented by a diamond marker.

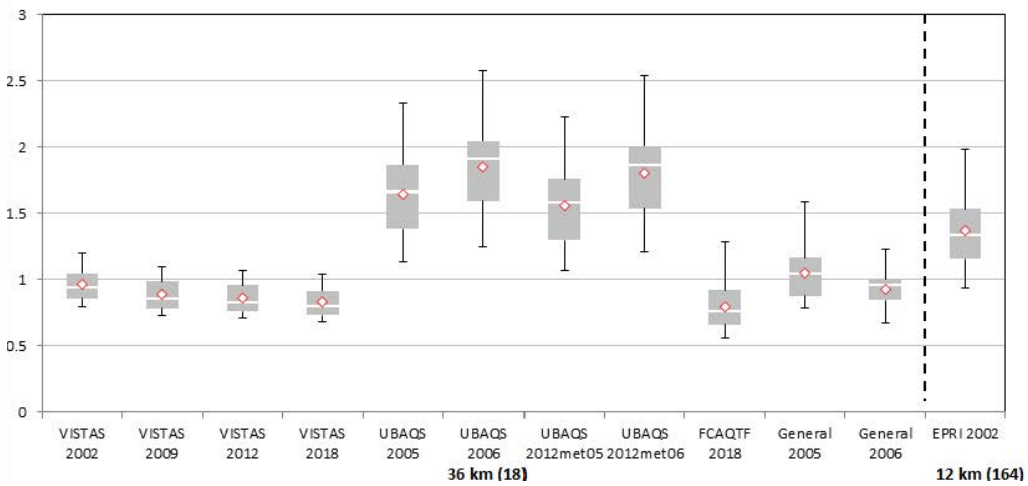


(**a**) TSO_X (cm/s) (**b**) TNO_Y (cm/s) (**c**) LNH_X (kg-N/ha/year)

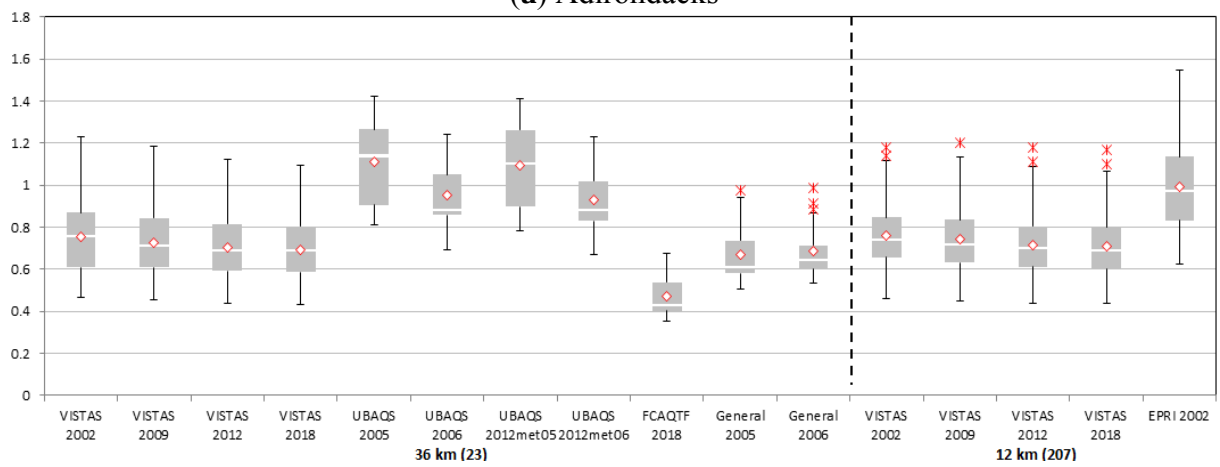
Figures 6 through 8 show spatial variability in TSO_X, TNO_Y, and LNH_X at each of the two case study areas for each annual model simulation. Comparing the model simulations of the same meteorological modeling year but with different emission scenarios (VISTAS 2002 *vs*. 2009 *vs*. 2012 *vs*. 2018; UBAQS 2005 *vs*. 2012met05; UBAQS 2006 *vs*. 2012met06) shows that emission changes have minor impacts on the TSO_X and TNO_Y variability while emission reductions slightly increase the LNH_X variability. The impact of different meteorological modeling years on the TSO_X and TNO_Y is not clear. The mean TSO_X slightly increased from 2005 to 2006 of the UBAQS simulations at Adirondacks, but slightly decreased at Shenandoah. Choice of chemical transport model appears to have bigger impacts. The CMAQ V4.6 (UBAQS) model shows higher TSO_X than the CAMx V5.21 (General) at both Adirondacks and Shenandoah. TNO_Y is slightly higher with CMAQ V4.6 than CAMx at Adirondacks, but slightly lower at Shenandoah. CMAQ-AMSTERDAM (EPRI) tends to estimate higher TSO_X and TNO_Y with greater variability than the CMAQ V4.5.1_SOAmods code used for VISTAS. Comparing VISTAS 2002 with UBAQS 2005 or 2006 shows combined impact of using different meteorological modeling years (2002 *vs*. 2005 or 2006) and different models (CMAQ V4.5.1_SOAmods *vs*. CMAQ V4.6), and TSO_X, TNO_Y, and LNH_X from the VISTAS simulation are generally lower than those from the UBAQS simulation. The simulations with a fine grid (12 km) show higher variability in TNO_Y and slightly lower variability in TSO_X and LNH_X than those with a coarse grid (36 km) at the Shenandoah case study areas (see the VISTAS 36 km *vs*. 12 km results). In addition to the two case study areas, we also looked at variability in these parameters at several Class I areas across the U.S.: Olympic NP (OLYM1); Mount Rainier NP (MORA1); Yosemite NP (YOSE1);

Grand Canyon NP (GRCA2); Glacier NP (GLAC1); Yellowstone NP (YELL2); Rocky Mountain NP (ROMO1); Mesa Verde NP (MEVE1); Great Smoky Mountains NP (GRSM1). Unlike the Adirondacks and Shenandoah areas discussed above, these Class I areas are each represented as a single grid cell (at the corresponding IMPROVE monitoring sites), therefore, only the variability across the model simulations are shown in Figure 9. Again, TSO_X, TNO_Y, and LNH_X show significant variability. TSO_X spans from 0.23 to 1.8 cm/s with higher variability at Mount Rainier NP, Yellowstone NP, and Rocky Mountain NP. The TNO_Y values range from 0.23 to 0.90 with higher variability at Yosemite NP, Yellowstone NP, and Rocky Mountain NP. The LNH_X values show similar variability. Overall, areas in western U.S. exhibit greater variability in these parameters.

Figure 6. Box and whisker plots for variability in TSO_X (cm/s) at (**a**) Adirondacks and (**b**) Shenandoah across all the model simulations. The number in parentheses represents the number of each sample data. How to read a box and whisker plot is explained in the Figure 5 caption.

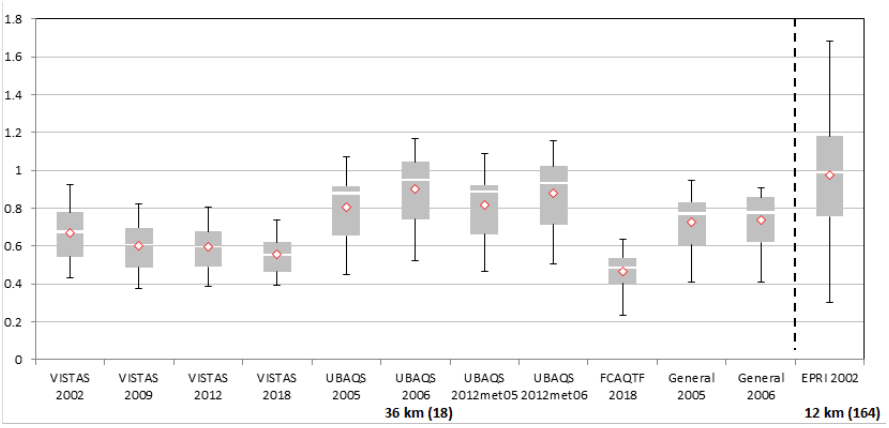


(**a**) Adirondacks

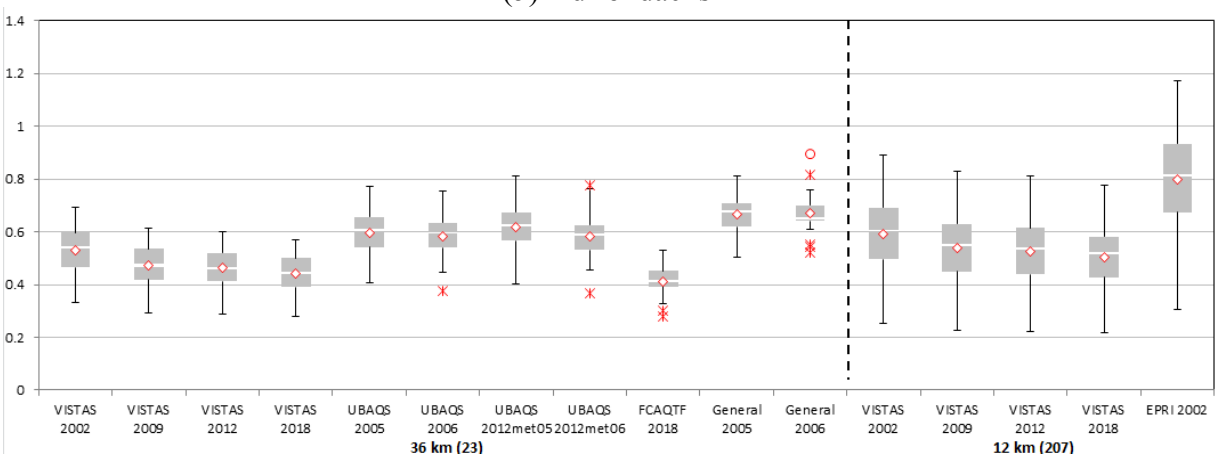


(**b**) Shenandoah

Figure 7. Box and whisker plots for variability in TNO_Y (cm/s) at (**a**) Adirondacks and (**b**) Shenandoah across all the model simulations. The number in parentheses represents the number of each sample data. How to read a box and whisker plot is explained in the Figure 5 caption.

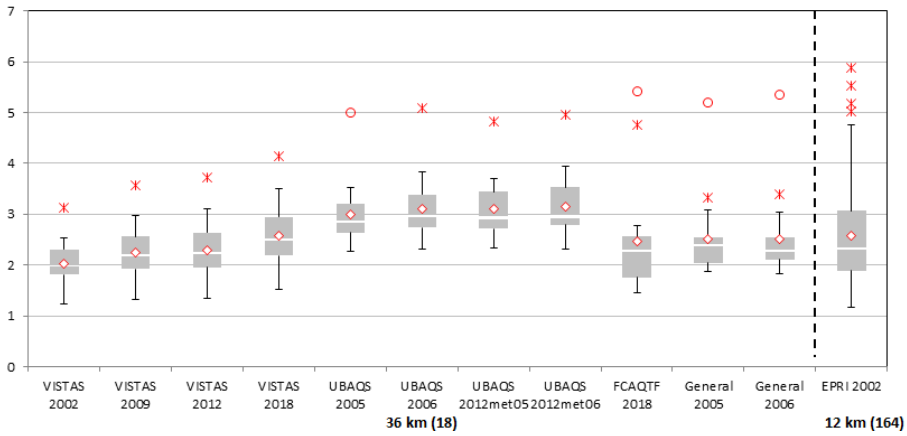


(**a**) Adirondacks



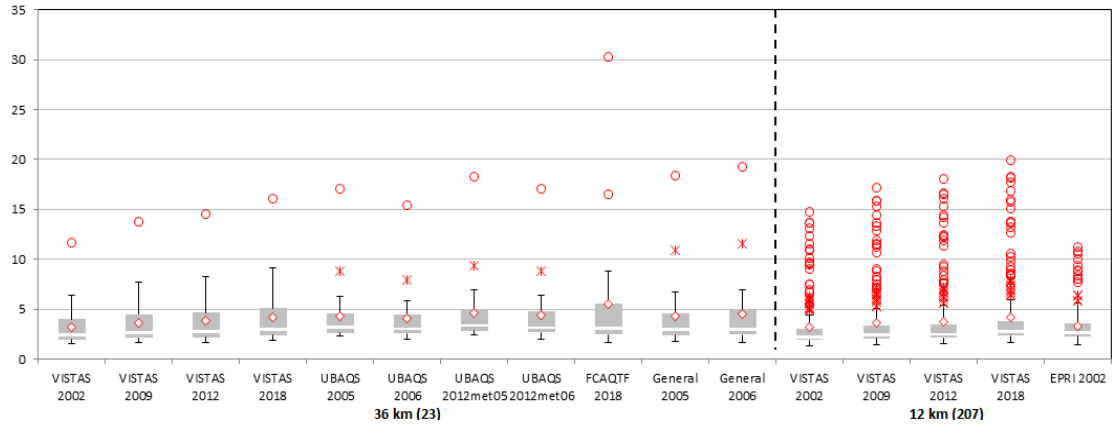
(**b**) Shenandoah

Figure 8. Box and whisker plots for variability in LNH_X (kg-N/ha/year) at (**a**) Adirondacks and (**b**) Shenandoah across all the model simulations. The number in parentheses represents the number of each sample data. How to read a box and whisker plot is explained in the Figure 5 caption.



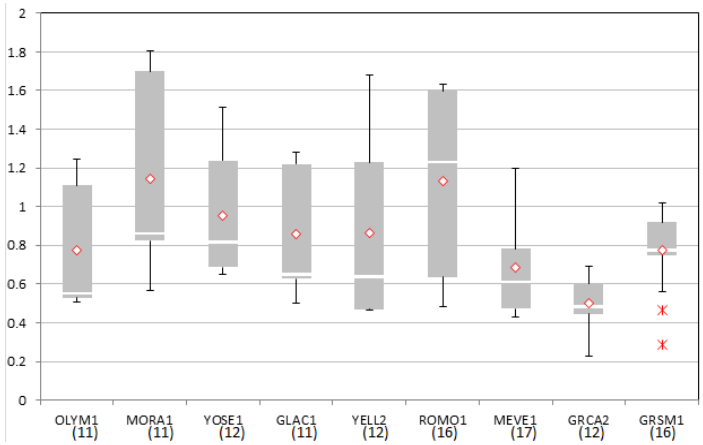
(**a**) Adirondacks

Figure 8. *Cont.*

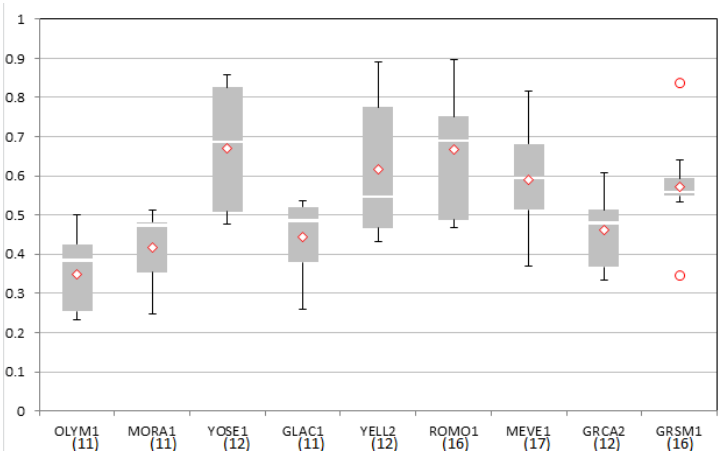


(**b**) Shenandoah

Figure 9. Box and whisker plots for variability in (**a**) TSO_X; (**b**) TNO_Y and (**c**) LNH_X across all the model simulations at each IMPROVE site. The number in parentheses represents the number of each sample data. How to read a box and whisker plot is explained in the Figure 5 caption.

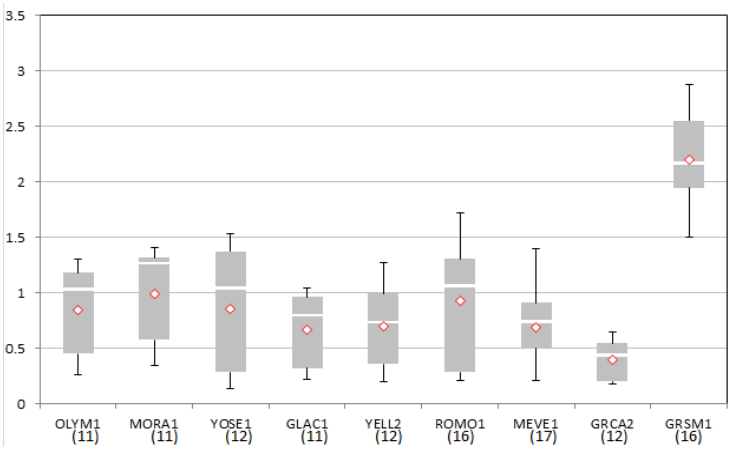


(**a**) TSO_X (cm/s)



(**b**) TNO_Y (cm/s)

Figure 9. *Cont.*



(**c**) LNH_X (kg-N/ha/year)

5. Conclusions

EPA's recent proposal for new secondary NAAQS of SO_X and NO_Y rely heavily on chemical transport modeling simulations. Therefore, it is crucial to adequately evaluate model's ability to accurately predict ambient concentrations and deposition of the relevant species. The rigor required when using models to establish a standard (and to determine non-attainment of those standards) should be much higher than when the models are used in a relative sense for implementation of those standards, such as development of control strategies. To supplement EPA's model performance evaluation of the CMAQ model, we conducted a rigorous and comprehensive model performance evaluation using several annual model simulations by multiple chemical transport models. The results show that there are significant uncertainty and variability in model performance of estimating concentrations and depositions of SO_X and NO_Y species. We also note that a complete model performance evaluation is problematic at the moment due to two issues. Firstly, routine measurements of total oxidized and reduced nitrogen is limited. The National Core (NCore) multi-pollutant monitoring program [29] routinely measures NO_Y and SO_X beginning in 2011, but the coverage is still not sufficient (83 sites at mainly urban locations). Secondly, significant uncertainties exist in measurement of dry deposition. As noted earlier, CASTNET dry deposition data are model-based estimates, not direct measurements. Uncertainties associated with the inferential measurements of dry deposition have been discussed elsewhere [30–32]. One of the main uncertainties is introduced by scaling up estimates of deposition velocities from a small area with relatively uniform surface characteristics to a large area consisting of various land use categories. Even with uniform surface condition, conventional inferential approaches cannot be reliably applicable to very stable atmospheric conditions that can occur near surface during nighttime. Evaluation studies where modeled deposition velocities were compared with direct measurements in relatively small areas showed that differences as large as ±30% are common [33].

The conceptual design of the new standard proposed by EPA includes many model-based parameters such as the Transference Ratios (TSO_X and TNO_Y) and total reduced nitrogen deposition (LNH_X) to link ambient air indicators (SO_X and NO_Y) to an ecological indicator (ANC). EPA argued that these parameters are sufficiently stable to be used for establishing the new standard [4]. However,

we found that, unlike EPA's analysis, there is a substantial spatial variability in these parameters across the U.S. as well as within a given ecosystem watershed area. It is also shown that the choice of the simulation year, the chemical transport model, and the model configuration (e.g., the grid resolution) can make a large difference in the calculated TSO_X, TNO_Y and LNH_X parameters. Such variability can potentially introduce significant uncertainty in the definition of the new secondary SO_X and NO_Y standards proposed by EPA. Given identified uncertainties and biases in the modeling results, these parameters should be based on observations, not model estimates. If the model can be shown to reproduce the observed values of these parameters then, and only then, can the model-based parameters be used for the regulatory purpose.

Acknowledgment

This work was funded by the Electric Power Research Institute.

References

1. Environmental Protection Agency (EPA). *Integrated Review Plan for the Secondary National Ambient Air Quality Standards for Nitrogen Dioxide and Sulfur Dioxide*; EPA-452/R-08-006; December 2007. Available online: http://www.epa.gov/ttnnaaqs/standards/no2so2sec/cr_pd.html (accessed on 29 December 2011).
2. Environmental Protection Agency (EPA). *Integrated Science Assessment for Oxides of Nitrogen and Sulfur Ecological Criteria (Final Report)*; EPA/600/R-08/082F; December 2008. Available online: http://www.epa.gov/ttnnaaqs/standards/no2so2sec/cr_isi.html (accessed on 29 December 2011).
3. Environmental Protection Agency (EPA). *Risk and Exposure Assessment for Review of the Secondary National Ambient Air Quality Standards for Oxides of Nitrogen and Oxides of Sulfur-Main Content—Final Report*; EPA-452/R-09-008a; September 2009. Available online: http://www.epa.gov/ttnnaaqs/standards/no2so2sec/cr_rea.html (accessed on 29 December 2011).
4. Environmental Protection Agency (EPA). *Policy Assessment for the Review of the Secondary National Ambient Air Quality Standards for Oxides of Nitrogen and Oxides of Sulfur*; EPA-452/R-11-005a; February 2011. Available online: http://www.epa.gov/ttnnaaqs/standards/no2so2sec/cr_pa.html (accessed on 29 December 2011).
5. Russell, A.; Samet, J.M. *Review of the Policy Assessment for the Review of the Secondary National Ambient Air Quality Standard for NO_X and SO_X: First Draft*; EPA-CASAC-10-014; March 2010. Available online: http://yosemite.epa.gov/sab/sabproduct.nsf/264cb1227d55e02c85257402007446a4/0fc13c821ee6181a85257473005ae1ec!OpenDocument&TableRow=2.3#2 (accessed on 29 December 2011).
6. Russell, A.; Samet, J.M. *Review of the Policy Assessment for the Review of the Secondary National Ambient Air Quality Standard for NO_X and SO_X: Second Draft*; EPA-CASAC-11-003; December 2010. Available online: http://yosemite.epa.gov/sab/sabproduct.nsf/264cb1227d55e02c85257402007446a4/0fc13c821ee6181a85257473005ae1ec!OpenDocument&TableRow=2.3#2 (accessed on 29 December 2011).

7. Byun, D.W.; Ching, J.K.S. *Science Algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) Modeling System*; EPA-600/R-99-030; March 1999. Available online: http://www.epa.gov/asmdnerl/CMAQ/CMAQscienceDoc.html (accessed on 29 December 2011).
8. *CAMx. User's Guide—Comprehensive Air-Quality Model with Extensions*, Version 5.40; ENVIRON International Corporation: Novato, CA, USA, 2011. Available online: http://www.camx.com (accessed on 29 December 2011).
9. Appel, K.W.; Gilliland, A.B.; Sarwar, G.; Gilliam, R.C. Evaluation of the Community Multiscale Air Quality (CMAQ) model version 4.5: Sensitivities impacting model performance Part I—Ozone. *Atmos. Environ.* **2007**, *41*, 9603–9615.
10. Appel, K.W.; Bhave, P.V.; Gilliland, A.B.; Sarwar, G.; Roselle, S.J. Evaluation of the Community Multiscale Air Quality (CMAQ) model version 4.5: Sensitivities impacting model performance; Part II—Particulate matter. *Atmos. Environ.* **2008**, *42*, 6057–6066.
11. Appel, K.W.; Roselle, S.J.; Gilliam, R.C.; Pleim, J.E. Sensitivity of the community multiscale air quality (CMAQ) model v4.7 results for the eastern United States to MM5 and WRF meteorological drivers. *Geosci. Model Dev.* **2010**, *3*, 169–188.
12. Appel, K.W.; Foley, K.M.; Bash, J.O.; Pinder, R.W.; Dennis, R.L.; Allen, D.J.; Pickering, K. A multi-resolution assessment of the Community Multiscale Air Quality (CMAQ) model v4.7 wet deposition estimates for 2002–2006. *Geosci. Model Dev.* **2011**, *4*, 357–371.
13. Foley, K.M.; Roselle, S.J.; Appel, K.W.; Bhave, P.V.; Pleim, J.E.; Otte, T.L.; Mathur, R.; Sarwar, G.; Young, J.O.; Gilliam, R.C.; *et al*. Incremental testing of the Community Multiscale Air Quality (CMAQ) modeling system version 4.7. *Geosci. Model Dev.* **2010**, *3*, 205–226.
14. Baker, K.; Scheff, P. Photochemical model performance for $PM_{2.5}$ sulfate, nitrate, ammonium, and precursor species SO_2, HNO_3, and NH_3 at background monitor locations in the central and eastern United States. *Atmos. Environ.* **2007**, *41*, 6185–6195.
15. Baker, K.; Scheff, P. Assessing meteorological variable and process relationships to modeled $PM_{2.5}$ ammonium nitrate and ammonium sulfate in the central United States. *J. Appl. Meteorol. Climatol.* **2008**, *47*, 2395–2404.
16. Rodriguez, M.A.; Barna, M.G.; Moore, T. Regional impacts of oil and gas development on ozone formation in the western United States. *J. Air Waste Manag. Assoc.* **2009**, *59*, 1111–1118.
17. Rodriguez, M.A.; Barna, M.G.; Gebhart, K.A.; Hand, J.L.; Adelman, Z.E.; Schichtel, B.A.; Collett, J.L., Jr.; Malm, W.C. Modeling the fate of atmospheric reduced nitrogen during the Rocky Mountain Atmospheric Nitrogen and Sulfur Study (RoMANS): Performance evaluation and diagnosis using integrated processes rate analysis. *Atmos. Environ.* **2011**, *45*, 223–234.
18. Wesely, M.L. Parameterization of surface resistances to gaseous dry deposition in regional-scale numerical models. *Atmos. Environ.* **1989**, *23*, 1293–1304.
19. Slinn, S.A.; Slinn, W.G.N. Predictions for particle deposition on natural waters. *Atmos. Environ.* **1980**, *24*, 1013–1016.
20. Pleim, J.E.; Clarke, J.F.; Finkelstein, P.L.; Cooter, E.J.; Ellestad, T.G.; Xiu, A.; Angevine, W.M. Comparison of Measured and Modeled Surface Fluxes of Heat, Moisture and Chemical Dry Deposition. In *Air Pollution Modeling and Its Application*, XI ed.; Plenum Press: New York, NY, USA, 1996.

21. Pleim, J.E.; Xiu, A.; Finkelstein, P.L.; Clarke, J.F. Evaluation of a Coupled Land-Surface and Dry Deposition Model through Comparison to Field Measurements of Surface Heat, Moisture, and Ozone Fluxes. In *Proceedings of the 12th Symposium on Boundary Layers and Turbulence*, Vancouver, BC, Canada, 28 July–1 August 1997.
22. Morris, R.E.; Koo, B.; Piyachaturawat, P.; Stella, G.; McNally, D.; Loomis, C.; Chien, C.-J.; Tonnesen, G. *Technical Support Document for VISTAS Emissions and Air Quality Modeling to Support Regional Haze State Implementation Plans (Final Report)*; March 2009. Available online: http://www.metro4-sesarm.org/vistas/data/RHR/Modeling/Reports/VISTASII_TSD_FinalReport_3-09.pdf (accessed on 29 December 2011).
23. Morris, R.E.; Kemball-Cook, S.; Sakulyanonyvittaya, T.; Parker, L.; Shah, T.; Nopmomngcol, U.; Piyachaturawat, P. *Uinta Basin Air Quality Study (UBAQS) Technical Report (Final Report)*; June 2009. Available online: http://westernenergyalliance.org/wp-content/uploads/UBAQS_Final_Report_Jun30_2009.pdf (accessed on 29 December 2011).
24. Karamchandani, P.; Vijayaraghavan, K.; Bronson, R.; Chen, S.-Y.; Knipping, E.M. Development and application of a parallelized version of the Advanced Modeling System for Transport, Emissions, Reactions and Deposition of Atmospheric Matter (AMSTERDAM): 1. Model performance evaluation. *Atmos. Poll. Res.* **2010**, *1*, 260–270.
25. Morris, R.E.; Tai, E.; Sakulyanonyvittaya, T.; McNally, D.; Loomis, C. *Model Performance Evaluation for the June–July 2006 Ozone Episode for the Denver 8-Hour Ozone State Implementation Plan (Final Report)*; August 2008. Available online: http://www.colorado.gov/airquality/documents/deno308/Denver_2006MPE_DraftFinal_Aug29_2008.pdf (accessed on 29 December 2011).
26. Stoeckenius, T.E.; Emery, C.A.; Shah, T.P.; Johnson, J.R.; Parker, L.K.; Pollack, A.K. *Air Quality Modeling Study for the Four Corners Region (Final Report)*; June 2009. Available online: http://www.nmenv.state.nm.us/aqb/4C/Documents/FinalRepRev20090806.pdf (accessed on 29 December 2011).
27. Clarke, J.F.; Edgerton, E.S.; Martin, B.E. Dry deposition calculations for the clean air status and trends network. *Atmos. Environ.* **1997**, *31*, 3667–3678.
28. Meyers, T.P.; Finkelstein, P.; Clarke, J.; Ellestad, T.G.; Sims, P.F. A multilayer model for inferring dry deposition using standard meteorological measurements. *J. Geophys. Res.* **1998**, *103*, 22645–22661.
29. Environmental Protection Agency (EPA). *Ambient Air Monitoring Strategy for State, Local, and Tribal Air Agencies*; December 2008. Available online: http://www.epa.gov/ttn/amtic/monstratdoc.html (accessed on 29 December 2011).
30. Foken, T.; Dlugi, R.; Kramm, G. On the determination of dry deposition and emission of gaseous compounds at the biosphere-atmosphere interface. *Meteorol. Z.* **1995**, *4*, 91–118.
31. Kramm, G.; Dlugi, R. Modelling of the vertical fluxes of nitric acid, ammonia, and ammonium nitrate. *J. Atmos. Chem.* **1994**, *18*, 319–357.
32. Kramm, G.; Dlugi, R.; Dollard, G.J.; Foken, T.; Molders, N.; Muller, H.; Seiler, W.; Sievering, H. On the dry deposition of ozone and reactive nitrogen species. *Atmos. Environ.* **1995**, *29*, 3209–3231.

33. Wesely, M.L.; Hicks, B.B. A review of the current status of knowledge on dry deposition. *Atmos. Environ.* **2000**, *34*, 2261–2282.