JGR Atmospheres

Key Points:
- Modeled estimates of the wet deposition across the United States are improved by fusion with gridded precipitation data and observed wet deposition
- The fused wet deposition estimates evaluate well against observed values and maintain the spatial heterogeneity in the original model output
- Temporal trends in observed wet deposition of nitrate, ammonium, and sulfate from 2002 to 2012 are well captured by the adjusted model output

Supporting Information:
- Supporting Information S1

Correspondence to:
K. M. Foley,
foley.kristen@epa.gov

Author Contributions:
Conceptualization: Donna B. Schwede, Jesse O. Bash, Joseph P. Pinto, Robin L. Dennis
Methodology: Donna B. Schwede, Robin L. Dennis
Resources: Robin L. Dennis
Supervision: Robin L. Dennis
Writing – review & editing: Donna B. Schwede, Jesse O. Bash, Joseph P. Pinto, Robin L. Dennis

A Measurement-Model Fusion Approach for Improved Wet Deposition Maps and Trends

Yuqiang Zhang[1] (iD)**, Kristen M. Foley[2]** (iD)**, Donna B. Schwede[2], Jesse O. Bash[2]** (iD)**, Joseph P. Pinto[3], and Robin L. Dennis[2,4]**

[1]Oak Ridge Institute for Science and Education (ORISE), U.S. Environmental Protection Agency, Research Triangle Park, NC, USA, [2]U.S. Environmental Protection Agency, Research Triangle Park, NC, USA, [3]Department of Environmental Sciences and Engineering, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, [4]Retired

Abstract Air quality models provide spatial fields of wet deposition (WD) and dry deposition that explicitly account for the transport and transformation of emissions from thousands of sources. However, many sources of uncertainty in the air quality model including errors in emissions and meteorological inputs (particularly precipitation) and incomplete descriptions of the chemical and physical processes governing deposition can lead to bias and error in the simulation of WD. We present an approach to bias correct Community Multiscale Air Quality model output over the contiguous United States using observation-based gridded precipitation data generated by the Parameter-elevation Regressions on Independent Slopes Model and WD observations at the National Atmospheric Deposition Program National Trends Network sites. A cross-validation analysis shows that the adjusted annual accumulated WD for NO_3^-, NH_4^+, and SO_4^{2-} from 2002 to 2012 has less bias and higher correlation with observed values than the base model output without adjustment. Temporal trends in observed WD are captured well by the adjusted model simulations across the entire contiguous United States. Consistent with previous trend analyses, WD NO_3^- and SO_4^{2-} are shown to decrease during this period in the eastern half of the United States, particularly in the Northeast, while remaining nearly constant in the West. Trends in WD of NH_4^+ are more spatially and temporally heterogeneous, with some positive trends in the Great Plains and Central Valley of CA and slightly negative trends in the south.

Plain Language Summary We use observed wet deposition data and observation-based interpolated precipitation data to improve the Community Multiscale Air Quality model's performance in simulating wet deposition in the United States from 2002 to 2012 for nitrate, ammonium, and sulfate. Performance of the bias-corrected model output is improved, with lower bias and higher correlation with observations compared to base model output without adjustment. Community Multiscale Air Quality bias-corrected deposition fields can be used to more accurately estimate deposition for comparison to critical loads (i.e., the level of a pollutant below which no harmful ecological effect occurs) at locations lacking observations. Model-estimated trends during the 11-year time series after the bias correction are better able to capture observed trends compared to the raw model output. Model-based methods for estimating trends in wet deposition are important for accountability studies used to quantify the impact of emission control measures on ambient concentrations and wet deposition.

1. Introduction

The ultimate fate of atmospheric sulfur ($SO_x = SO_2 + SO_4^{2-}$) and reactive nitrogen compounds ($NO_y = NO + NO_2$ + other oxidized forms of N excluding N_2O and $NH_x = NH_3 + NH_4^+$) is removal by wet scavenging and dry deposition, which in turn lead to a variety of environmental effects including changes in net primary production (LeBauer & Treseder, 2008), acidification (Driscoll et al., 2003), eutrophication (Bergström & Jansson, 2006; Bouwman et al., 2002), and other nutrient-loading effects (Bowman et al., 2008). Long-term wet deposition (WD) and dry deposition (DD) data are needed to assess the changes and trends in total deposition (TD = WD + DD), as well as their effect on ecosystems. In the United States, the National Atmospheric Deposition Program/National Trends Network (NADP/NTN; https://nadp.slh.wisc.edu/ntn) provides the nation's primary WD data. These observed WD values are often used to estimate past and current loads of acidic (S + N) and nutrient (N) deposition on sensitive ecosystems for critical loads exceedance studies in the United States (Nanus et al., 2003; Sickles & Shadwick, 2007, 2015). Using

NADP/NTN data from 1985 to 2012, Du et al. (2014) found a positive trend in ammonium WD at the national scale and a significant negative trend in nitrate WD, especially in the Northeast. Sickles and Shadwick (2015) found a significant decrease of wet and DD of both S and N in the east from 1990 to 2009, using deposition data from 34 paired dry and wet monitoring sites.

More accurate, highly spatially resolved fields of deposition are needed for regional and national scale assessment of nitrogen and sulfur budgets and comparison of deposition loads of N + S with critical loads (i.e., the level of a pollutant below which no harmful ecological effect occurs) for different ecosystems and ecological endpoints. Previous studies have used different spatial interpolation methods, such as inverse distance weighting or kriging, to estimate the spatial distribution of N deposition (Grimm & Lynch, 1991, 2004; Jia et al., 2014; Li et al., 2016; Qu et al., 2017; Schwede & Lear, 2014) and temporal trends. Because deposition monitoring data are relatively sparse (on the order of 230–250 NADP/NTN sites across the United States, depending on the year) interpolation methods can miss important emission sources and geographic features that impact deposition unless additional covariate information is incorporated (e.g., regression kriging). Air quality model simulations offer an alternative to observation-only interpolation methods and have the advantage of incorporating detailed emissions information on hundreds of sources while also explicitly accounting for geographic features in the parameterization of physical and chemical processes.

Air quality model simulations have been extensively used to determine the distribution (Dennis et al., 2013; Dentener et al., 2006; Gu et al., 2015; Mathur & Dennis, 2003; Zhang et al., 2012) and temporal trends (Lamarque et al., 2013; Zhang et al., 2018) of nitrogen and sulfur WD and to assess critical load exceedances (Ellis et al., 2013; Sun et al., 2017; Williams et al., 2017), as they can provide more spatially complete coverage of TD. However, model simulations also have bias for estimating WD, especially at coarse resolution (Appel et al., 2011; Lamarque et al., 2013; Simon et al., 2012; Zhao et al., 2017). Zhang et al. (2018) showed that the long-term model simulation from coupled Weather Research and Forecasting (WRF)-Community Multiscale Air Quality Model (CMAQ) at 36 km underestimated WD of both nitrogen and sulfur in the United States from 1990 to 2010, as well as their negative trends by more than 20%. The biased model simulation of WD could lead to false negative or false positive deposition exceedances in critical loads analyses (Williams et al., 2017).

Measurement-model fusion methods have been widely used to improve air quality model output for ambient concentrations (e.g., Berrocal et al., 2010a, 2010b; Berrocal et al., 2012; Robichaud et al., 2016; Robichaud & Ménard, 2014), with more limited application to WD data (e.g., Makar et al., 2018; Sahu et al., 2010). In 2017 the World Meteorology Organization hosted a 3-day workshop on measurement-model fusion for global total atmospheric deposition. The resulting published report summarizes existing products and plans for developing improved maps of wet, dry, and total atmospheric deposition to "enable research into biogeochemical cycles and assessments of ecosystems and human health effects" (World Meteorological Organization (WMO), 2017).

One of the modeling products included in the WMO report was the annual TD maps for the contiguous United States (CONUS) developed by the NADP Total Deposition Science Committee (TDEP; Schwede & Lear, 2014). TDEP maps for DD are based on a measurement-model fusion method using CMAQ output and observations from the Clean Air Status and Trends Network, while spatial maps of annual WD are based on NADP/NTN measurements and Parameter-elevation Regressions on Independent Slopes Model (PRISM) precipitation data. In this study, we evaluate a measurement-model fusion method to combine CMAQ-simulated WD of major nitrogen (NO_3^- and NH_4^+) and sulfur (SO_4^{2-}) species with the NADP/NTN measurements and PRISM precipitation data used by TDEP. Bias-adjusted CMAQ WD fields are created for 2002–2012 and evaluated against NADP/NTN measurements through a cross-validation analysis. The bias-adjusted spatial fields reduce bias and improve correlation with observed values, without smoothing out the spatial heterogeneity of the original model output. We assess the long-term trends in WD of these species in the United States during the same period and find that estimated trends in bias-adjusted model WD evaluate well against observed trends. Although TD of both nitrate and sulfate can be dominated by DD in the United States (Zhang et al., 2018), improvements to DD maps are beyond the scope of the current work due to the lack of measured DD data currently available.

2. Methodology

2.1. Model Inputs and Outputs

The model simulations used in this analysis were conducted using CMAQ (Byun & Schere, 2006) version 5.0.2 (Appel et al., 2013; doi:10.5281/zenodo.1079898), which includes bidirectional NH_3 air-surface exchange (Bash et al., 2013; Pleim et al., 2013) and the Carbon Bond version 2005 (CB05) chemical mechanism (Yarwood et al., 2005) updated with toluene reactions (CB05-TU; Whitten et al., 2010). The simulations were performed over the CONUS from 1 January 2002 to 31 December 2012 at 12-km × 12-km horizontal resolution with 35 vertical layers from the surface to the top of the free troposphere and a nominal height of 19 m for layer one. Meteorological inputs were developed using the WRF model version 3.4 using the Kain-Fritsch trigger for convective precipitation (Ma & Tan, 2009), the Pleim-Xiu land-surface model (Gilliam & Pleim, 2010), and four-dimensional data assimilation (Gilliam et al., 2012). From 2002 to 2005, the 2001 National Land Cover Database((NLCD) land cover data were used as an input to the WRF simulations; remaining years were based on 2006 NLCD land cover data. The land use classification methods and changes in the land use over time were documented and evaluated in Homer et al. (2015). Output from WRF was processed to create inputs for CMAQ using the Meteorology-Chemistry Interface Processor (Otte & Pleim, 2010) version 4.1.3. Emission inputs were based on 2002, 2005, 2008, and 2011 National Emission Inventory data (NEI; https://www.epa.gov/air-emissions-inventories, accessed April 26, 2018) using the Sparse Matrix Operator Kernel Emissions (Houyoux et al., 2000) processing system version 3.1 and included year-specific continuous emission monitoring system data for large combustion and industrial processes (mainly electric generating units). Year-specific mobile emissions were derived from U.S. Environmental Protection Agency (EPA) Motor Vehicle Emission Simulator (www.epa.gov/otaq/models/moves, accessed 26 April 2018) version 2010b, and the Biogenic Emissions Inventory System version 3.14 was used for inline biogenic emissions (Carlton & Baker, 2011). NEI data on annual total ammonia emissions from confined animal feeding operations were based on national estimates from the Carnegie Mellon University model (Davidson et al., 2004) and state submitted information. These annual totals were temporally allocated using the Sparse Matrix Operator Kernel Emissions programme with the temporalization from Zhu et al. (2015). Ammonia emissions from fertilizer were estimated using CMAQ with bidirectional exchange; in this parameterization, the net flux is broken into emissions and deposition components, the former making use of a compensation point concentration with zero ambient ammonia, the latter from assuming zero stomatal and soil compensation point concentrations (Bash et al., 2013). Emission inputs also included inline NO produced from lightning using year-specific flash rate data from the National Lightning Detection Network (Allen et al., 2012). For simulation years 2005 to 2012 chemical boundary conditions for CMAQ were based on year-specific simulations of GEOS-Chem version 8–03-02 (http://wiki.seas.harvard.edu/geos-chem/index.php/Main_Page, accessed 26 April 2018) using version 8-02-01 chemistry, GEOS-5 meteorology and ICOADS shipping emissions (Henderson et al., 2014). The CB05 aerosol chemical mechanism used in CMAQ has more aerosol species than that in GEOS-Chem, and so speciation remapping between these two models was done following Henderson et al. (2014). GEOS-Chem simulations were not available for 2002–2004; chemical boundary conditions for these years were based on the monthly median values from the 2005 simulation. We expect the influence of these boundary conditions to be relatively small, especially compared to other uncertainties in the modeling system such as uncertainties in the emissions inventories in the early 2000s. Additional information on model inputs is provided in the supporting information.

2.2. Observational Data Sets

Model predicted WD for nitrate (NO_3^-), ammonium (NH_4^+), and sulfate (SO_4^{2-}) were compared to observed WD values from NADP/NTN (National Atmospheric Deposition Program) (NRSP-3) (2019). Modeled and observed values were paired in space by matching the observations to the model values from the grid cell that contains the monitor. See Appel et al. (2011), section 2.4, for a description of how the chemical species in the CMAQ model were configured to match the measured WD values. There were 282 NADP/NTN sites that had data for at least 1 year of the study period. The number of sites with available data in each year ranged from 231 to 252. In addition, NADP/NTN provides four completeness criteria (CC) to ensure the reported values represent a valid annual accumulated measurement. About 20% of the sites in the network in any given year are dropped through application of these CC. As a result, only 68 sites met all four criteria for the full 11-year time series, complicating spatial analysis of temporal trends. Here we
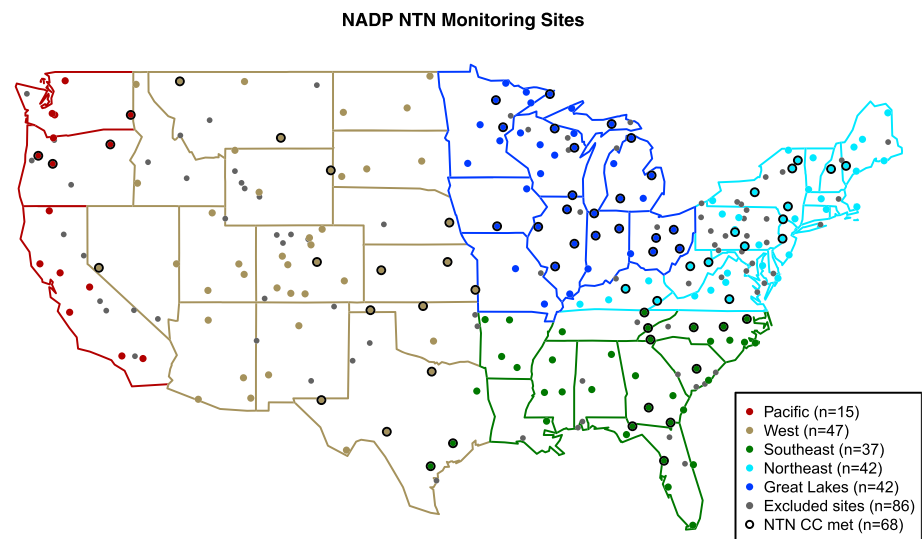
Figure 1. Location of 183 NADP/NTN monitors used in the analysis. Color-coded regions are used in Figure 4 and Figures S3, S4, and S7 provided in the supporting information. The black-boarded colored circles (NTN CC met) indicate sites that met the full NADP/NTN completeness criteria for the entire 11-year time series. NADP = National Atmospheric Deposition Program; NTN = National Trends Network; CC = completeness criteria.

have relaxed the CC and included any site that had 10 or 11 years of data that met the revised criteria, resulting in a final sample size of 183 sites. Figure 1 shows the location of the 183 sites used in the analysis and highlights which of these sites met the full NADP/NTN criteria. A full description of the CC used in this analysis is included in the supporting information.

The observation-based gridded precipitation fields generated by the PRISM (Daly et al., 2008; PRISM Climate Group, Oregon State University, http://prism.oregonstate.edu, created 10 September 2015) were used to evaluate model-predicted accumulated precipitation from the WRF meteorological model. Annual accumulated PRISM precipitation data at 4-km resolution were regridded to the CMAQ 12-km resolution grid for the entire time series, which were then used to scale the CMAQ WD fields, as described in section 2.3. The precipitation estimates from PRISM have been used to improve maps of WD provided by NADP/NTN and TDEP in complex terrain regions in the United States, as PRISM incorporates point observation, elevation, and climatic factors when estimating precipitation (Latysh & Wetherbee, 2012; Schwede & Lear, 2014). The gridded PRISM precipitation estimates are also more commensurable with the WRF grid cell volume precipitation values compared to the precipitation measurements at NADP/NTN sites which represent rainfall totals at a single point location.

2.3. Description of the Measurement-Model Fusion Approach to Combine CMAQ Output and Observations

The relationship between WD and precipitation is affected by the frequency, duration, and intensity of the rainfall as well as the ambient concentration. While this relationship is typically nonlinear on hourly and daily time scales, on an annual accumulation basis the relationship can become more linear. For example, Figure 2 shows the modeled and observed relationship between annual accumulated precipitation and WD of sulfate at the 42 sites in the Great Lakes subregion. At locations where the modeled and observed relationships are similar, the model/observed errors in annual accumulated WD can also be highly correlated with model/observed errors in precipitation as shown in Figure 3.

The variability in the log of the ratio of model/observed annual accumulated precipitation explains at least 40% of the variability in the log of the ratio of annual accumulated modeled/observed WD at 94, 78, and 104 of the observation stations for nitrate, ammonium, and sulfate, respectively (i.e., 52%, 43%, and 57% of the stations). For about a quarter of the locations in the eastern half of the United States, these R^2 values are on the order of 60–90%, particularly in the Great Lakes subregion (Figures 3a and 3c). While high correlations do not hold for all stations, we have made use of the calculated linear relationships to recalculate the modeled WD using PRISM rainfall volume. Specifically, the modeled WD for grid cell s and year t
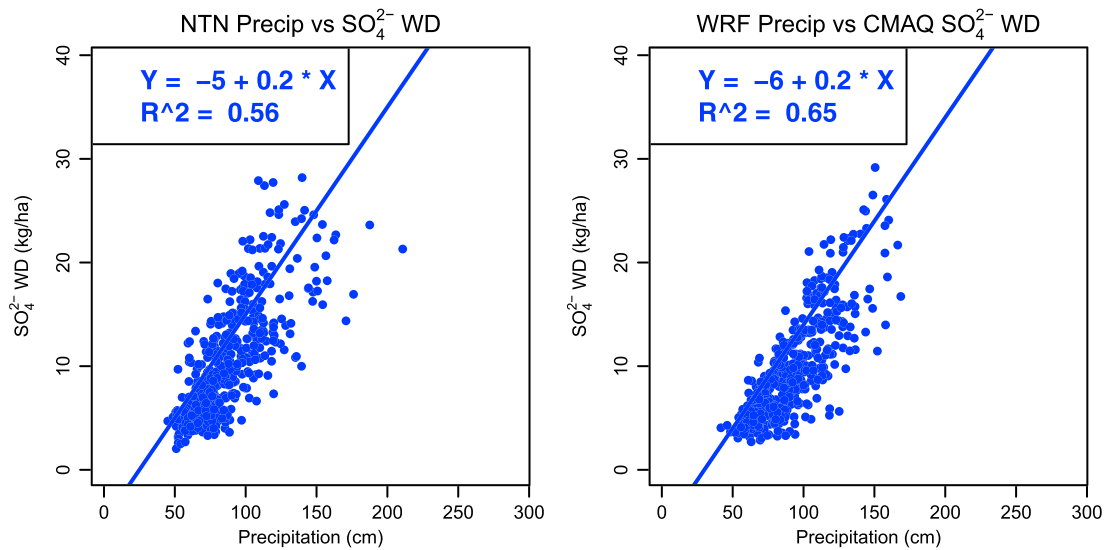
Figure 2. Strength of linear relationship between annual accumulated precipitation and sulfate wet deposition at NTN sites in the Great Lakes subregion based on observed data (left) and modeled data (right). NTN = National Trends Network.

$(WD_{mod}(s,t))$ was converted to precipitation concentration by dividing the modeled precipitation $(Precip_{mod}(s,t))$ and then converted back to deposition by multiplying by the gridded PRISM precipitation $(Precip_{obs}(s,t))$ for that location and year

$$WD_{precip-adj}(s, t) = \frac{WD_{mod}(s, t)}{Precip_{mod}(s, t)} \times Precip_{obs}(s, t) \tag{2.1}$$

where $WD_{precip-adj}(s,t)$ is the precipitation adjusted WD at grid cell s and year t.

While this precipitation adjustment (precip-adj) increased the correlation between observed and modeled WD, model bias remained due to (1) cases where the WD does not scale with changes in precipitation in a one-to-one linear fashion and (2) bias in emission inputs (including missing emissions sources) and/or errors in other model processes effecting deposition. A second bias-adjustment (bias-adjust) was applied to the precip-adj WD $(WD_{precip-adj}(s,t))$, based on interpolating the ratio of the precip-adj WD and the NADP/NTN WD observations using a kriging model. A cross-validation analysis was used to select the interpolation approach used in the bias adjustment. Cross-validation results are provided in section 3.2 with supporting figures in the supporting information. The cross-validation analysis showed that the model bias at a single NADP/NTN site can be influenced by very local emission sources and precipitation events and is not necessarily representative of the bias at surrounding grid cells. Specifically, the cross-validation analysis showed that kriging the bias extended the influence of sites with very high bias to other locations, degrading the effectiveness of the bias-adjustment. To account for this, a moving window averaging approach was applied prior to kriging to reduce the influence of these very site-specific biases. At each site, the median of the ratio, $WD_{precip-adj}/WD_{obs}$, of all sites within a 300-km radius was calculated, dampening any large bias
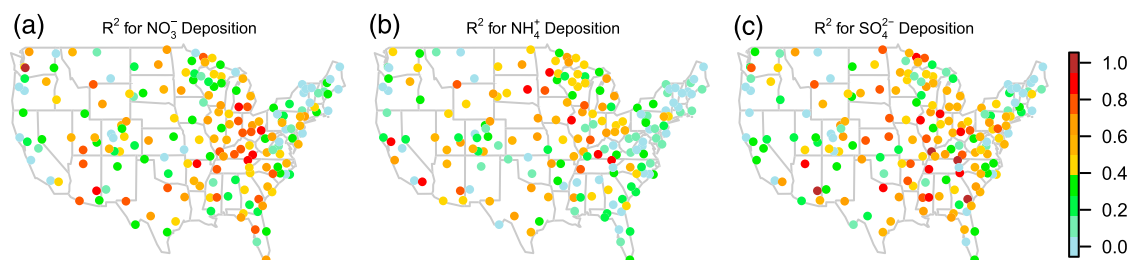

Figure 3. Coefficient of determination (i.e., Pearson correlation squared or R^2) between the natural log of model/observed annual wet deposition, (a) for NO_3^-, (b) for NH_4^+, and (c) for SO_4^{2-}, for 2002–2012. A seasonal version of this plot is included in the supporting information (Figure S1).

values to create a more regional estimate of bias. The new median biases were then interpolated using universal kriging with a linear trend in the spatial coordinates and an exponential covariance structure (Cressie, 1993). Note that since the bias metric was a ratio, the averaging and kriging steps were done on a log scale and then back transformed. Both the moving average calculation and the universal kriging were done using the gstat library in the R statistical software package (Pebesma, 2004) The cross-validation analysis was used to determine the size of the moving window (300 km) and the choice of the type of kriging model.

The result of the spatial interpolation steps above is a set of estimated spatial bias fields for each year, $\widehat{b}(s, t)$. The final adjusted WD for a given grid cell and year t was calculated based on multiplying the $WD_{\text{precip-adj}}(s, t)$ and the inverse of the estimated bias at that location:

$$WD_{\text{adj}}(s, t) = \frac{1}{\widehat{b}(s, t)} \times WD_{\text{precip-adj}}(s, t) \qquad (2.2)$$

An example of the spatial bias field and the map of the final adjustment factor for WD of nitrate in 2012 is provided in the supporting information (Figure S5).

A second cross-validation analysis was performed to compare this precip-adjusted and bias-adjusted approach (Method 1) with three other measurement-model fusion methods. First, the individual precipitation and bias adjustments were tested separately. Method 2 includes the precipitation adjustment (precip-adj) of CMAQ WD using PRISM based on equation (2.1) with no additional bias adjustment. Method 3 uses the bias-adj CMAQ model outputs using NADP/NTN WD data with no initial PRISM adjustment. A fourth method (Method 4) that does not use CMAQ output was also compared. In this case, WD measurements from NADP/NTN were divided by the regridded 12-km resolution PRISM precipitation for the grid cell containing the monitor to create a precipitation-weighted wet concentration value. The precipitation-weighted wet concentrations for each ion were kriged using the same kriging approach as Method 1 (i.e., universal kriging with a linear trend and exponential covariance) to produce a spatial field (units = mg/L) and then multiplied by the full PRISM precipitation grid (units converted to millimeter per year) to estimate WD at every grid cell (units = kg/ha) (Latysh & Wetherbee, 2012; Nanus et al., 2003). Results of the cross-validation analysis comparing the different methods are provided in section 3.2.

Note that Method 4 differs from the TDEP approach for combining PRISM and WD observations. TDEP uses inverse distance weighting to interpolate precipitation-weighted concentrations from NADP/NTN and AIRMoN and adjusts the PRISM data based on the precipitation amounts measured at NADP/NTN monitoring network sites (http://nadp.slh.wisc.edu/committees/tdep/tdepmaps/; accessed 26 April 2018). It was not possible to incorporate this method into the cross-validation study applied in this paper; however, differences between the TDEP maps and the methods applied here are discussed in section 3.3.

2.4. Description of Trend Analysis

The temporal trends in modeled and observed WD were estimated using the Theil-Sen slope estimator (Sen, 1968; Theil, 1950). Theil-Sen slopes were calculated using the mblm library in the R statistical software package. The Theil-Sen estimator is a nonparametric method for estimating the slope through a set of points, in this case the annual accumulated deposition for each year at a given grid cell or NADP/NTN location. This data set has a very small sample size ($n = 10$ or 11) and can exhibit large interannual variability in some locations driven by year-to-year changes in precipitation. The Theil-Sen estimator was used because it is insensitive to outliers and offers a more robust approach compared to simple linear regression. The Kendall rank test was used to test if the linear trend in the observed time series was statistically significant.

3. Results

3.1. Evaluation of Model Estimated WD

Figure 4 provides the normalized mean bias (NMB) of annual precipitation (top left), WD of sulfate (top right), nitrate (bottom left), and ammonium (bottom right) based on WRF precipitation and CMAQ WD estimates compared to NADP/NTN observations from 2002 to 2012. The time series of NMB was calculated for each of the five subregions to compare how the bias differs across space. Observed precipitation totals in the intermountain West are very low yet are consistently overestimated by the model by 20–40%. Precipitation is also overestimated every year in Pacific Coast states, with the greatest overestimation in 2008 (NMB = 35%).
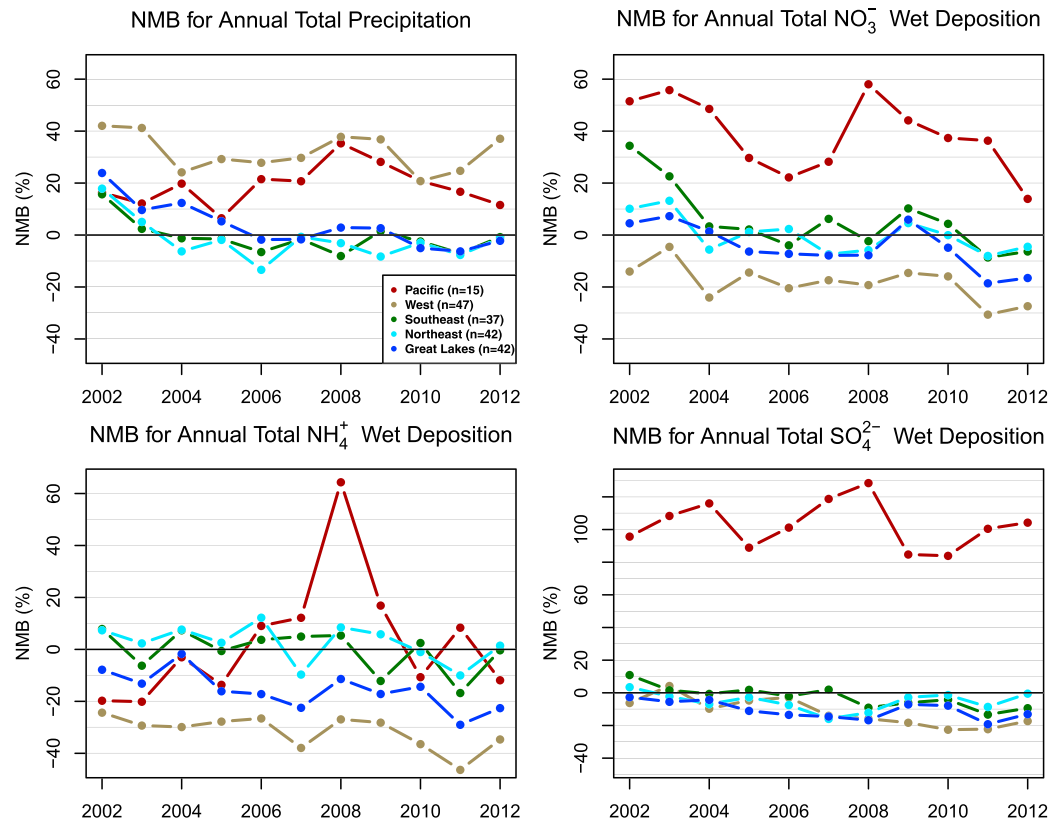
Figure 4. Time series of normalized mean bias (NMB; %) of annual Weather Research and Forecasting precipitation and annual Community Multiscale Air Quality Model wet deposition for 2002 to 2012 based on model sums compared to National Atmospheric Deposition Program/National Trends Network observations. NMB values were calculated for each of the five subregions shown in Figure 1. NMB = normalized mean bias.

There is less model bias throughout the eastern half of the United States, particularly for the latter half of the time series when the NMB is within ±10%. For most years, the NMB for both NO_3^- and SO_4^{2-} WD is less than 10% in the Northeast, Great Lakes and Southeast, with a slight tendency toward more underestimation of SO_4^{2-} WD. In the Pacific Coast states NO_3^- and SO_4^{2-} WD are overestimated with SO_4^{2-} WD NMB values exceeding 100% in seven of the 10 years. The year with the largest NMB for the Pacific region for all three WD species is 2008, corresponding to the peak in the precipitation bias. The modeled ammonium WD is comparable to observations in the Northeast and Southeast regions (NMB within ±15%) but is underestimated in the Great Lakes and Pacific regions. The underprediction of ammonium WD in the central United States may be the result of missing emissions from large concentrated animal feeding operations (Warner et al., 2017; Zhang et al., 2018). The overestimation in Pacific Coast states is seen in previous studies (Appel et al., 2011; Zhang et al., 2018) and could be caused by complex terrain effects. Besides the errors in modeled precipitation, other errors could also lead to the bias in modeled WD: errors in atmospheric concentrations due to misspecification of precursor emission levels, imperfect representation of the chemical and physical processes related to the nitrogen and sulfur cycles, the underestimation of deposition velocity of ammonia gas, and errors in the bidirectional fluxes (Bash et al., 2013; Pleim et al., 2013).

Table 1 provides performance statistics for the simulated annual and seasonal totals for WRF precipitation and CMAQ WD of nitrate, ammonium, and sulfate for the entire time series. The model consistently overestimates precipitation except during fall, with NMB for the annual total of 5.3%. R^2 between modeled and observed seasonal precipitation is lowest in summer (0.47) and highest in winter (0.65). Correlation for the annual accumulated precipitation is higher, with an R^2 value of 0.67. The model overestimates nitrate WD in winter (8.6%) and fall (12.4%) and underestimates it in spring (−8.2%) and summer (−6.7%) with NMB of the annual accumulation values of −11.0%. R^2 for seasonal nitrate WD between the model and measurements is highest in spring (0.73) and lowest in summer (0.54), with a higher annual R^2 of 0.76. The model

Table 1
Performance Metrics for Annual and Seasonal Accumulated WRF Precipitation Predictions and CMAQv5.0.2 Wet Deposition Predictions for Sulfate, Nitrate, and Ammonium Including Observation Mean (MEAN OBS), Model Mean (MEAN MOD), Mean Bias (MB), Normalized Mean Bias (NMB), Root-Mean-Square Error (RMSE), Mean Error (ME), Normalized Mean Error (NME), and Pearson Correlation Squared (R^2)

Seasons	Variables	N	MEAN OBS	MEAN MOD	MB	NMB (%)	RMSE	ME	NME (%)	R^2
	Precipitation	1,964	94.41	99.4	4.99	5.3	26.81	19.82	21.9	0.67
	NO_3^-	1,964	7.53	7.43	−0.09	−1.2	2.19	1.60	21.2	0.76
Annual	NH_4^+	1,964	2.46	2.19	−0.27	−11.0	0.94	0.68	27.6	0.60
	SO_4^{2-}	1,964	8.87	8.47	−0.40	−4.5	2.76	1.92	21.7	0.81
	Precipitation	1,631	20.94	22.37	1.43	6.8	9.69	7.02	33.5	0.65
	NO_3^-	1,631	1.45	1.58	0.13	8.6	0.77	0.55	37.8	0.58
Winter	NH_4^+	1,631	0.32	0.24	−0.07	−23.4	0.20	.012	37.9	0.46
	SO_4^{2-}	1,631	1.51	1.23	−0.28	−18.6	0.85	0.57	37.4	0.57
	Precipitation	1,801	24.35	29.06	2.71	11.1	9.61	6.99	28.7	0.61
	NO_3^-	1,801	2.20	2.02	−0.18	−8.2	0.77	0.57	26.0	0.73
Spring	NH_4^+	1,801	0.85	0.63	−0.22	−26.2	0.46	0.31	36.9	0.55
	SO_4^{2-}	1,801	2.63	2.53	−0.10	−3.7	1.09	0.74	28.3	0.71
	Precipitation	1,802	27.24	29.15	1.91	7.0	13.35	9.34	34.3	0.47
	NO_3^-	1,802	2.59	2.42	−0.17	−6.7	1.04	0.77	29.7	0.54
Summer	NH_4^+	1,802	0.87	0.98	0.11	12.6	0.49	0.36	41.3	0.41
	SO_4^{2-}	1,802	3.14	3.20	0.06	2.0	1.51	0.99	31.5	0.69
	precipitation	1,853	23.77	22.11	−1.65	−7.0	8.87	6.38	26.9	0.64
	NO_3^-	1,853	1.44	1.62	0.18	12.4	0.71	0.48	33.5	0.65
Fall	NH_4^+	1,853	0.46	0.39	−0.07	−14.9	0.25	0.17	36.5	0.50
	SO_4^{2-}	1,853	1.83	1.73	−0.10	−5.4	0.79	0.54	29.6	0.69

Note. The definition of the evaluation metrics can be found in the supporting information. Statistics are based on the modeled and observed data from 183 NADP/NTN sites. Note that not all 183 sites have a valid observation for every season of every year (i.e., $N < 183$ sites×11 years). Modeled and observed values were paired in space by matching the observations to the model values from the grid cell that contains the monitor. The units are centimeter for precipitation, and kilogram/hectare for wet deposition. WRF = Weather Research and Forecasting; cmaq = Community Multiscale Air Quality Model.

generally underestimates ammonium WD for all seasons except for summer (NMB = 12.6%), with annual NMB of −11.0%. R^2 for seasonal accumulated ammonium WD is below 0.5 except for spring (0.55), with an annual R^2 of 0.60. For sulfate WD, the model only overestimates in summer (2.0%), with annual NMB of −4.5%. The R^2 is comparable for all seasons (from 0.69 to 0.71), with slightly lower values in winter (0.57). In general, the model had better performance in simulating the WD of nitrate and sulfate than that of ammonium. As an additional reference, Table S1 in the supporting information provides the NMB for the precipitation and WD of nitrate, ammonium, and sulfate, based on years 2002–2006 compared to the analogous statistics from the evaluation of CMAQv4.7 WD estimates published in Appel et al. (2011).

3.2. Cross-Validation Comparison of Five Methods for Creating WD Maps

Two leave-one-out, cross-validation analyses were conducted to evaluate different methods for combining NADP/NTN, PRISM, and model data to create WD maps. The cross validation was conducted as follows. For each year and each monitoring location, the annual WD from that site and any site within 48 km (four grid cells) was withheld from the estimation of the bias adjustment field for that year. This approach was used to evaluate how well the methods perform in areas without a NADP/NTN measurement. The cross-validation prediction for the location and year of the withheld observation was then compared to the original NADP/NTN measurement.

3.2.1. Smoothing Radius

The first cross-validation comparison was used to select the radius of the moving window averaging used to smooth the bias in the precipitation-adjusted model values. The first row of Figure S2 shows the NMB, NME, and R^2 values for WD of nitrate, ammonium, and sulfate for a radius of 100–700 km compared to not performing the smoothing step prior to creating bias maps via kriging. Results for a radius of 300–600 km are very similar. A radius of 300 km was selected because, on aggregate, this distance produced the best cross-validation statistics for WD of sulfate and nitrate. For ammonium, the moving window smoothing increases the bias but decreases the error and improves the R^2. The second row of Figure S2 shows a similar comparison across seven spatial interpolation methods available in the gstat library. Inverse distance weight (IDW) was compared to ordinary kriging (constant mean) and universal kriging (linear trend in site coordinates).

Table 2
Cross-Validation Statistics for the WD From CMAQv5.0.2 Output Only and Four Measurement-Model Fusion Methods (Precip- Adj and Bias-Adj (Method 1), Only Precipitation-Adjustment (Precip-Adj, Method 2), Only Bias-Adjustment (Bias-Adj, Method 3), and NTN × PRISM (Method 4)

Species	Fusion method	Mean Obs (kg/ha)	Mean mod (kg/ha)	MB (kg/ha)	NMB (%)	RMSE (kg/ha)	ME (kg/ha)	NME (%)	R^2
NO_3^-	CMAQ outputs	7.53	7.43	−0.09	−1.2	2.19	1.60	21.2	0.76
	Method 1 (precip-adj and bias-adj)	7.53	7.53	**0.00**	**0.0**	1.53	1.10	14.6	**0.86**
	Method 2 (precip-adj)	7.53	7.47	−0.05	−0.7	1.95	1.48	19.6	0.81
	Method 3 (bias-adj)	7.53	7.63	0.10	1.4	1.92	1.36	18.1	0.79
	Method 4 (NTN × PRISM)	7.53	7.54	**0.00**	0.1	**1.48**	**1.09**	**14.5**	**0.86**
NH_4^+	CMAQ outputs	2.46	2.19	−0.27	−11.0	0.94	0.68	27.6	0.60
	Method 1 (precip-and bias-adj)	2.46	2.39	−0.07	−2.9	**0.69**	**0.48**	**19.6**	**0.76**
	Method 2 (precip-adj)	2.46	2.21	−0.25	−10.2	0.87	0.65	26.3	0.66
	Method 3 (bias-adj)	2.46	2.41	**−0.05**	**−1.8**	0.79	0.57	23.0	0.70
	Method 4 (NTN × PRISM)	2.46	2.37	−0.09	−3.6	0.82	0.58	23.5	0.67
SO_4^{2-}	CMAQ outputs	8.87	8.47	−0.40	−4.5	2.76	1.92	21.7	0.81
	Method 1 (precip-and bias-adj)	8.87	8.84	−0.03	−0.3	**1.98**	**1.35**	**15.3**	**0.90**
	Method 2 (precip-adj)	8.88	8.56	−0.31	−3.4	2.37	1.66	18.7	0.86
	Method 3 (bias-adj)	8.87	8.97	0.10	1.1	2.57	1.72	19.4	0.84
	Method 4 (NTN × PRISM)	8.87	8.86	**−0.01**	**−0.1**	2.08	1.39	15.7	0.89

Note. Summary statistics are based on $N = 1,964$ observations and include observation mean (MEAN OBS), model mean (MEAN MOD), mean bias (MB), normalized mean bias (NMB), root mean square error (RMSE), mean error (ME), normalized mean error (NME), and Pearson correlation squared (R^2). Evaluation metrics with the best performance values are shown in bold. WD = wet deposition; NTN = National Trends Network; PRISM = Parameter-elevation Regressions on Independent Slopes Model.

Five different spatial covariance models were also compared: spherical and Matérn with smoothness parameter 0.5, 1.0, 1.5, and 2.0. The bias and error are reduced for all three species by using a kriging approach as opposed to the IDW. The statistics for the different kriging approaches are very similar, with some worsening of performance with the increased smoothness parameters. Universal kriging was selected with a Matérn covariance model with smoothness parameter 0.5, which is equivalent to what is referred to as an exponential covariance model.

3.2.2. Comparison of Five Methods for Creating WD Maps

A second cross validation was used to compare the raw CMAQ output and four methods for combining different sources of model and observed data. Table 2 provides evaluation summary statistics based on 1,964 observations, which include the 10 or 11 years of data from the 183 sites used in the study. Using only a precip-adj (Method 2) or only a multiplicative deposition bias-adjustment (Method 3) to adjust the modeled WD decreases the negative bias and increases the correlation with observations, compared to the unadjusted model values. Method 1, which combines both adjustments, provides additional improvement over using only one adjustment (i.e., even lower bias and higher R^2 compared to Methods 2 and 3; see Table 2). The precip-adj and bias-adj results show very good agreement with observations, with NMB within 3% and R^2 values between 0.77 and 0.90 (Figure S4). This is a clear improvement over the raw model output, especially for the Pacific, West, and Great Lakes regions (Figures S3 and S4) where the model performance for WD is usually poor due to terrain effects (Appel et al., 2011; Zhang et al., 2018). This result illustrates the value in incorporating data for both precipitation and WD in the bias-adjusting of the model-simulated WD. By taking advantage of the improved WD estimations from Method 1 in this study, Williams et al. (2017) found reduced bias in informing critical load exceedance calculations in the Pacific Northwest.

Using NADP/NTN wet concentration and PRISM data alone (NTN × PRISM) in Method 4 provides estimates for WD of sulfate and nitrate that have similar correlation and bias compared to the approach of precip-adj and bias-adj, which utilizes CMAQ output. For ammonium, Method 1 offers a higher correlation (R^2 of 0.77 compared to 0.67) as well as slightly lower bias (NMB of −2.9% compared to 3.6%) compared to Method 4, which relies on observations only. Additional differences in the methods that are not reflected in the cross-validation are evident in the final spatial maps of WD shown in the next section.

3.3. Spatial Maps of WD

Figure 5 shows maps of annual accumulated WD of nitrate in 2012 based on four different approaches. TDEP in the top left and the NTN × PRISM approach (Method 4) in the bottom left are observation-based
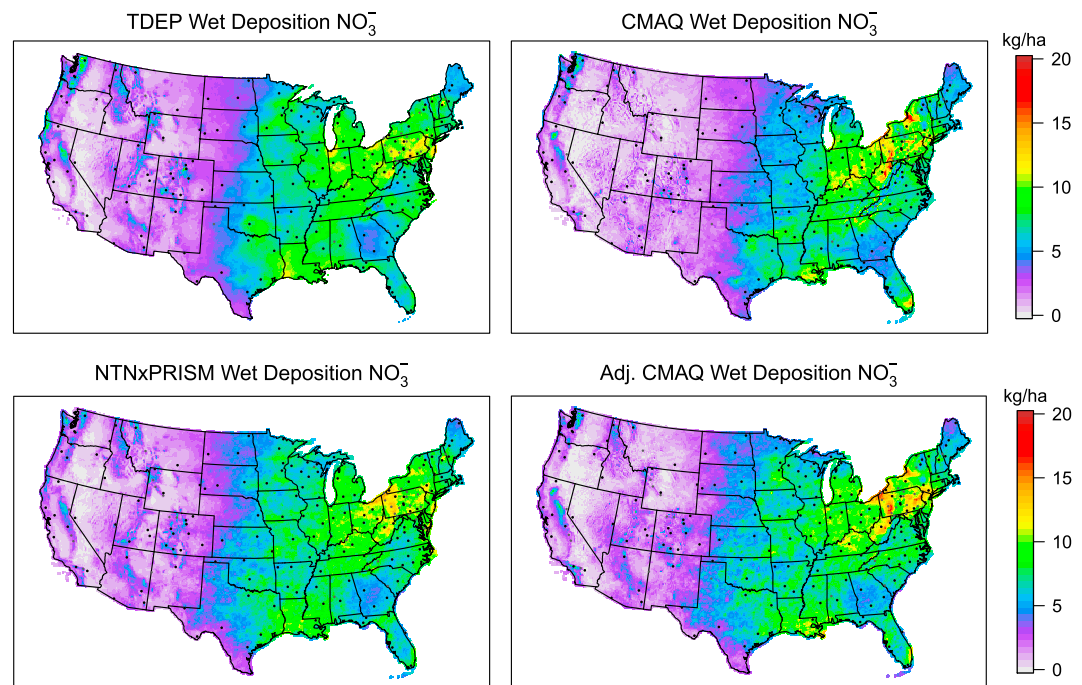
Figure 5. Annual wet deposition of nitrate (kg/ha) in 2012 based on TDEP (top left), CMAQ (top right), NTN × PRISM (Method 4; bottom left) and precip-adjusted and bias-adjusted CMAQ (Method 2; bottom right). National Atmospheric Deposition Program/National Trends Network monitor locations with annual data for 2012 are shown on all the maps. Sites that were dropped from the trend analysis in section 3.4 because they did not have a sufficiently long time series of measurements are shown in gray. TDEP = NADP Total Deposition; CMAQ = Community Multiscale Air Quality Model; NTN = National Trends Network; PRISM = Parameter-elevation Regressions on Independent Slopes Model.

approaches. The base model CMAQ and precip-adjusted and bias-adjusted CMAQ maps (Method 2) are shown in the top right and bottom right, respectively. Maps of the bias adjustment applied to the base model output to produce the final map of WD of nitrate are shown in supporting information Figure S5. Maps of precip-adjusted and bias-adjusted CMAQ WD of sulfate and ammonium for 2012 are provided in Figure S6.

The TDEP map in Figure 5 is more spatially smooth than the other approaches, with some "bullseye" spatial gradients common with IDW interpolation. The NTN × PRISM map has less of this bullseye effect but is missing some of the hot spots in the Northeast that are captured by the model. The base model output underestimates the WD throughout much of the domain west of the Mississippi River. The spatial gradients in the precip- adjusted and bias-adjusted map in this region are more consistent with the observation-only based maps without smoothing out the spatial heterogeneity of the original model output. These differences are not necessarily reflected in the evaluation in section 3.2 because the cross validation is limited by the spatial coverage of the observation network. The model-based approaches reflect spatial heterogeneity in local emissions sources that cannot be captured by the measurement network.

3.4. Evaluation of Model Predicted Trends

A trend analysis was used to assess the CMAQv5.0.2 system′s ability to predict changes in WD during 2002 to 2012 with and without the precipitation and bias adjustments. During this 11-year period, emission reduction programs such as the EPA′s Acid Rain Program, Nitrogen Oxides State Implementation Plan Call, and Clear Air Interstate Rule led to a substantial decrease in emissions of nitrogen oxides and sulfur dioxide (e.g., Gilliland et al., 2008; Hand et al., 2014). The impact of such emission reductions on WD can be seen in Figure 6 which shows the Theil-Sen slope estimator for the annual WD of nitrate, ammonium, and sulfate. The trends shown in Figure 6 are based on raw CMAQ output (left column) and adjusted CMAQ values (right column, Method 1). Trends based on the NADP/NTN measurements are overlaid on these spatial fields and show that there has been a steady decrease in WD of nitrate and sulfate from 2002 to 2012 in the eastern half of the United States, particularly in the Northeast where the average decrease across the
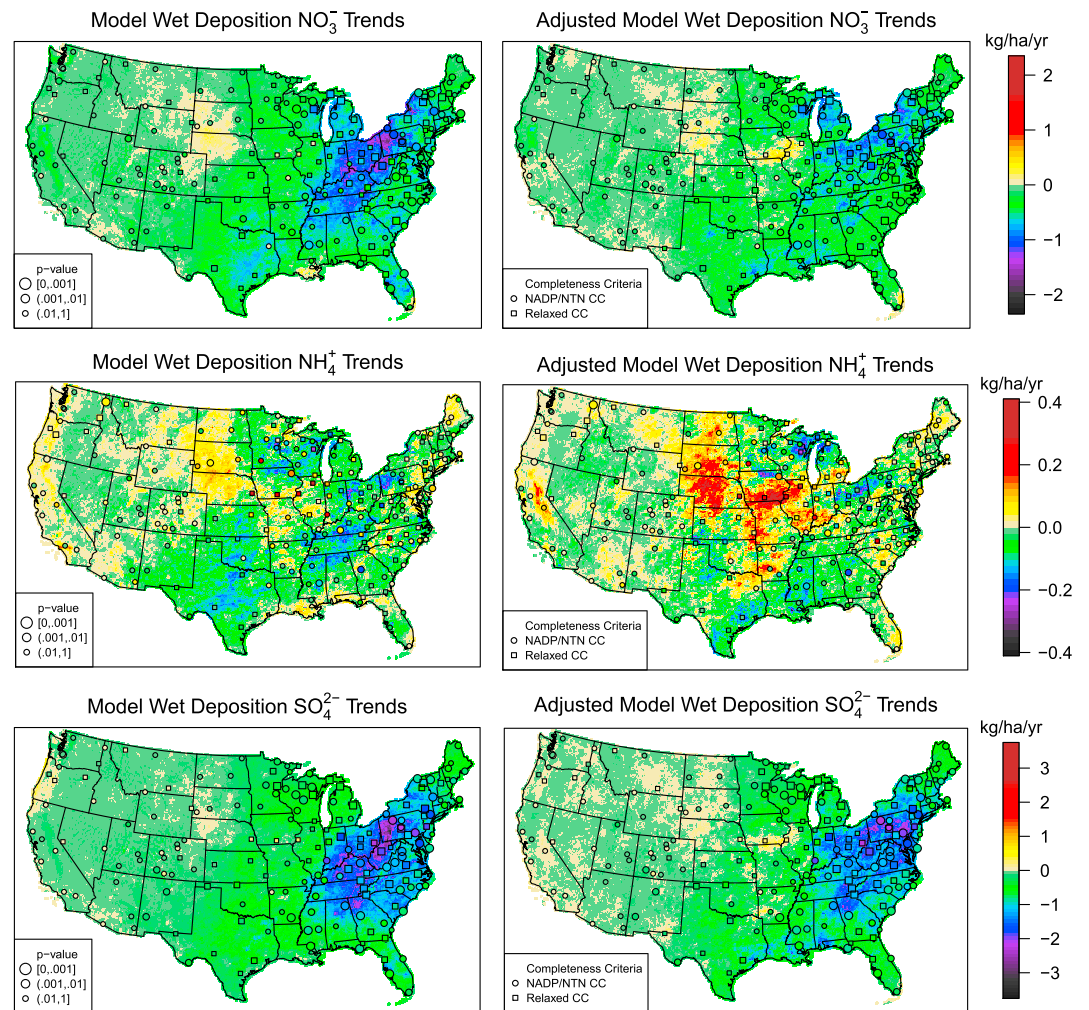
Figure 6. Theil-Sen trends (i.e., slopes) for the 2002–2012 time series of annual wet deposition at each NADP/NTN location and each model grid cell based on raw Community Multiscale Air Quality Model output (left column) and precip-and bias-adjusted Community Multiscale Air Quality Model output (right column) for wet deposition of nitrate (top row), ammonium (middle row) and sulfate (bottom row). Note the change in scale for ammonium. Trends at the NADP/NTN locations that met the full NADP/NTN completeness criteria are shown as colored circles with the remaining sites used in the analysis shown as colored squares. The size of the circle/square is based on the p-value of a nonparametric Kendall rank test. A larger symbol indicates a statistically significant linear trend in the time series. NADP/NTN = National Atmospheric Deposition Program/National Trends Network; CC = completeness criteria.

entire 11 years is 30% for nitrate and 50% for sulfate. Figure S7 provides scatterplots of the observed versus modeled and adjusted modeled trends at the 68 NADP/NTN locations that met the full CC. Use of the raw CMAQ output led to an overestimation of this observed negative trend in the Great Lakes and Northeast regions and parts of states on the Pacific Coast. The trends in WD of nitrate based on the adjusted model output are more consistent with the spatial patterns in the observed data, with the root-mean-square error (RMSE) of the estimated trends decreasing from 0.35 to 0.24 kg/ha/year in the Northeast and from 0.22 to 0.18 in the Great Lakes (Figure S7). Trends in WD of nitrate and sulfate are also improved in the Southeast with the RMSE for nitrate decreasing from 0.29 to 0.10 kg/ha/year and the RMSE for sulfate decreasing from 0.29 to 0.17 kg/ha/year. This improvement illustrates the value of adjusting the model output to account for errors in modeled precipitation and model bias due to emissions or other processes impacting WD. Note that the 68 NADP/NTN locations that met the full CC are plotted in Figure 5 as circles. The sites that were added with the relaxed CC (plotted as squares) help fill in spatial information, while not changing the overall conclusions drawn from the model predicted trends.

Trends in WD of ammonium from the adjusted model outputs are more spatially and temporally heterogeneous, with essentially no trend throughout much of the CONUS except for some slightly positive trends in parts of the Great Plains and Central Valley of California. (Note the change in scale in Figure 6 for ammonium compared to the maps for nitrate and sulfate.) Spatial features in ammonium are better captured by the adjusted model time series compared to the raw model output, with the RMSE of the trends decreasing in the West, Southeast, and Great Lakes and remaining constant in the other regions (Figure S7).

4. Summary and Conclusions

Using observational data from NADP/NTN, we evaluated an 11-year time series of CMAQv5.0.2 simulations to assess the model's performance in estimating annual WD. The model underestimates the annual WD of NO_3^-, NH_4^+, and SO_4^{2-}, with NMB of −1.2%, −11.0%, and −4.5%, respectively. To further improve the model's performance in simulating annual WD, we developed an approach to bias correct CMAQ model outputs using observation-based gridded PRISM precipitation and the WD measurements at the NADP/NTN sites. Cross-validation analysis shows that the new bias-adjusted model outputs have much lower bias and higher R^2 compared to the unadjusted model values. The NMB values for the annual WD of NO_3^-, NH_4^+, and SO_4^{2-} from the adjusted model outputs are 0.0%, −2.9%, and 0.3%, respectively. Spatial maps of the adjusted WD reflect spatial heterogeneity in local emissions and precipitation compared to the observation-based approaches considered here that tended to be more spatially smooth. The improved measurement-model fused WD estimates can be used to develop more precise cause-effect relationships between atmospheric deposition and sensitive aquatic and terrestrial ecosystems, particularly at high elevations that are impacted by orographic effects on precipitation.

The temporal trends in annual WD of NO_3^-, NH_4^+, and SO_4^{2-} are also captured by the model. Statistically significant negative trends are found for WD of NO_3^- and SO_4^{2-}, especially in the eastern United States using the precipitation- adjusted and bias-adjusted model outputs, while this trend is underestimated using the raw model outputs. The trends based on the precipitation- adjusted and bias-adjusted model outputs show much greater consistency with the spatial patterns in the observed data, illustrating the merits of our new methods. The use of air quality modeling for estimating trends offers the ability to perform accountability studies where simulations are used to quantify the impact of the control measures on ambient concentrations and WD. Air quality modeling also allows for assessment of alternative emission strategies, changes in land use and meteorology. Future simulations are typically run using base-year meteorology, allowing for the same precipitation adjustment to be applied. The appropriateness of applying the second bias adjustment to future or alternative simulations will depend on the scenario being modeled.

Accurate quantification of WD and DD is essential for assessing critical loads exceedances and effects on ecosystems. Improvements to WD have been demonstrated in the current work. Measurement-model fusion can be used to correct biases in model output while maintaining the spatial detail that the model provides. As described in the GAW report (WMO, 2017), measurement-model fusion approaches for deposition have already been adopted by multiple groups. TDEP is widely used in the United States and uses a measurement-model fusion approach for the calculation of DD. Inclusion of measurement-model fusion for WD has been identified by the NADP TDEP committee as a next step for U.S. TD maps. The analysis presented here can inform the selection of the fusion method for WD. Note that the incommensurability between volume-average model output values and point measurements can impact evaluation results and interpretation. This issue is present in model evaluation and data fusion methods for all modeled pollutants and has been discussed in the literature most frequently for ambient concentration of ozone and $PM_{2.5}$ (e.g., Gelfand et al. (2001); Fuentes and Raftery (2005); Swall and Foley (2009)). Addressing the issue of incommensurability in WD estimates is complicated by the fact that the subgrid variability in precipitation can be quite different from the subgrid variability in concentrations, and these differences will vary across the country. For example, a WD measurement in the Rocky Mountains might be different from a 12-km × 12-km average due to orographic effects while a measurement site in the Ohio River Valley may not be representative because it is upwind of a large source that does influence the 12-km × 12-km volume average. The NADP/NTN network is not dense enough to evaluate how much subgrid variability in WD can be expected within a 12-km × 12-km grid cell and how this subgrid variability changes across the country. Based on our current and prior analyses, we believe that the issue of incommensurability does not

preclude the use of the NADP/NTN measurements to provide information on broad regional biases in the model due to errors in emissions, meteorology, chemistry, and boundary conditions. Future work will include a comparison of alternative fusion methods such as the optimal interpolation methods developed by Robichaud et al. for the GEM-MACH air quality model (WMO, 2017). Improvements to deposition modeling within the CMAQ system can also reduce the need for bias correction of model outputs. Finally, the analysis methods developed here focused on the estimation of annual total WD. An assessment of data fusion methods for seasonal totals is needed to evaluate the effectiveness of these adjustment methods across different seasons.

References

Allen, D. J., Pickering, K. E., Pinder, R. W., Henderson, B. H., Appel, K. W., & Prados, A. (2012). Impact of lightning-NO on eastern United States photochemistry during the summer of 2006 as determined using the CMAQ model. *Atmospheric Chemistry and Physics*, *12*(4), 1737–1758. https://doi.org/10.5194/acp-12-1737-2012

Appel, K. W., Foley, K. M., Bash, J. O., Pinder, R. W., Dennis, R. L., Allen, D. J., & Pickering, K. (2011). A multi-resolution assessment of the Community Multiscale Air Quality (CMAQ) model v4.7 wet deposition estimates for 2002-2006. *Geoscientific Model Development*, *4*(2), 357–371. https://doi.org/10.5194/gmd-4-357-2011

Appel, K. W., Pouliot, G. A., Simon, H., Sarwar, G., Pye, H. O. T., Napelenok, S. L., et al. (2013). Evaluation of dust and trace metal estimates from the Community Multiscale Air Quality (CMAQ) model version 5.0. *Geoscientific Model Development*, *6*(4), 883–899. https://doi.org/10.5194/gmd-6-883-2013

Bash, J. O., Cooter, E. J., Dennis, R. L., Walker, J. T., & Pleim, J. E. (2013). Evaluation of a regional air-quality model with bidirectional NH_3 exchange coupled to an agroecosystem model. *Biogeosciences*, *10*(3), 1635–1645. https://doi.org/10.5194/bg-10-1635-2013

Bergström, A. K., & Jansson, M. (2006). Atmospheric nitrogen deposition has caused nitrogen enrichment and eutrophication of lakes in the northern hemisphere. *Global Change Biology*, *12*(4), 635–643. https://doi.org/10.1111/j.1365-2486.2006.01129.x

Berrocal, V., Gelfand, A. E., & Holland, D. M. (2010a). A spatiotemporal downscaler for output from numerical models. *Journal of Agricultural, Biological, and Environmental Statistics*, *15*(2), 176–197. https://doi.org/10.1007/s13253-009-0004-z

Berrocal, V., Gelfand, A. E., & Holland, D. M. (2010b). A bivariate space-time downscaler under space and time misalignment. *Ann. Appl. Stat.*, *4*(4), 1942–1975. https://doi.org/10.1214/10-AOAS351

Berrocal, V., Gelfand, A. E., & Holland, D. M. (2012). Space-time fusion under error in computer model output: An application to modeling air quality. *Biometrics*, *68*(3), 837–848. https://doi.org/10.1111/j.1541-0420.2011.01725.x

Bouwman, A. F., Van Vuuren, D. P., Derwent, R. G., & Posch, M. (2002). A global analysis of acidification and eutrophication of terrestrial ecosystems. *Water, Air, and Soil Pollution*, *141*(1/4), 349–382. https://doi.org/10.1023/A:1021398008726

Bowman, W. D., Cleveland, C. C., Halada, Ĺ., Hreško, J., & Baron, J. S. (2008). Negative impact of nitrogen deposition on soil buffering capacity. *Nature Geoscience*, *1*(11), 767–770. https://doi.org/10.1038/ngeo339

Byun, D., & Schere, K. L. (2006). Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling system. *Applied Mechanics Reviews*, *59*(2), 51–77. https://doi.org/10.1115/1.2128636

Carlton, A. G., & Baker, K. R. (2011). Photochemical modeling of the ozark isoprene volcano: MEGAN, BEIS, and their impacts on air quality predictions. *Environmental Science and Technology*, *45*(10), 4438–4445. https://doi.org/10.1021/es200050x

Cressie, N. (1993). *Statistics for spatial data*. New York, NY: John Wiley & Sons, Inc. https://doi.org/10.1002/9781119115151.ch1

Daly, C., Halblerb, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., et al. (2008). Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International Journal of Climatology*, *28*(15), 2031–2064. https://doi.org/10.1002/joc.1688

Davidson, C., Adams, P., Strader, R., Pinder, R., Anderson, N., Goebes, M., & Ayers, J. (2004). The Environmental Institute, Carnegie Mellon University, CMU Ammonia Model v.3.6., at , accessed 25 April 2009.http://www.cmu.edu/ammonia/

Dennis, R. L., Schwede, D. B., Bash, J. O., Pleim, J. E., Walker, J. T., & Foley, K. M. (2013). Sensitivity of continental United States atmospheric budgets of oxidized and reduced nitrogen to dry deposition parametrizations. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *368*(1621), 20130124. https://doi.org/10.1098/rstb.2013.0124

Dentener, F., Drevet, J., Lamarque, J. F., Bey, I., Eickhout, B., Fiore, A. M., et al. (2006). Nitrogen and sulfur deposition on regional and global scales: A multimodel evaluation. *Global Biogeochemical Cycles*, *20*, GB4003. https://doi.org/10.1029/2005GB002672

Driscoll, C. T., Whitall, D., Aber, J., Castro, M., Cronan, C., Goodale, C. L., et al. (2003). Nitrogen pollution in the northeastern United States: Sources, effects, and management options. *Bioscience*, *53*(4), 357–374. https://doi.org/10.1641/0006-3568(2003)053

Du, E., de Vries, W., Galloway, J. N., Hu, X., & Fang, J. (2014). Changes in wet nitrogen deposition in the United States between 1985 and 2012. *Environmental Research Letters*, *9*(9), 95,004. https://doi.org/10.1088/1748-9326/9/9/095004

Ellis, R. A., Jacob, D. J., Sulprizio, M. P., Zhang, L., Holmes, C. D., Schichtel, B. A., et al. (2013). Present and future nitrogen deposition to national parks in the United States: Critical load exceedances. *Atmospheric Chemistry and Physics*, *13*(17), 9083–9095. https://doi.org/10.5194/acp-13-9083-2013

Fuentes, M., & Raftery, A. E. (2005). Model evaluation and spatial interpolation by Bayesian. *Biometrics*, 36–45.

Gelfand, A. E., Zhu, L., & Carlin, B. P. (2001). On the change of support problem for spatio-temporal data. *Biostatistics*, *2*(1), 31–45. https://doi.org/10.1093/biostatistics/2.1.31

Gilliam, R. C., Godowitch, J. M., & Rao, S. T. (2012). Improving the horizontal transport in the lower troposphere with four dimensional data assimilation. *Atmospheric Environment*, *53*, 186–201. https://doi.org/10.1016/j.atmosenv.2011.10.064

Gilliam, R. C., & Pleim, J. E. (2010). Performance assessment of new land surface and planetary boundary layer physics in the WRF-ARW. *Journal of Applied Meteorology and Climatology*, *49*(4), 760–774. https://doi.org/10.1175/2009JAMC2126.1

Gilliland, A. B., Hogrefe, C., Pinder, R. W., Godowitch, J. M., Foley, K. L., & Rao, S. T. (2008). Dynamic evaluation of regional air quality models: Assessing changes in O_3 stemming from changes in emissions and meteorology. *Atmospheric Environment*, *42*(20), 5110–5123. https://doi.org/10.1016/j.atmosenv.2008.02.018

Grimm, J. W., & Lynch, J. A. (1991). Statistical analysis of errors in estimating wet deposition using five surface estimation algorithms. *Atmospheric Environment*, *25*(2), 317–327. https://doi.org/10.1016/0960-1686(91)90303-O

Grimm, J. W., & Lynch, J. A. (2004). Enhanced wet deposition estimates using modeled precipitation inputs. *Environmental Monitoring and Assessment*, *90*(1–3), 243–268. https://doi.org/10.1023/B:EMAS.0000003592.56006.a0

Gu, B., Ju, X., Chang, J., Ge, Y., & Vitousek, P. M. (2015). Integrated reactive nitrogen budgets and future trends in China. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(28), 8792–8797. https://doi.org/10.1073/pnas.1510211112

Hand, J. L., Schichtel, B. A., Malm, W. C., Copeland, S., Molenar, J. V., Frank, N., & Pitchford, M. (2014). Widespread reductions in haze across the United States from the early 1990s through 2011. *Atmospheric Environment*, *94*, 671–679. https://doi.org/10.1016/j.atmosenv.2014.05.062

Henderson, B. H., Akhtar, F., Pye, H. O. T., Napelenok, S. L., & Hutzell, W. T. (2014). A database and tool for boundary conditions for regional air quality modeling: Description and evaluation. *Geoscientific Model Development*, *7*(1), 339–360. https://doi.org/10.5194/gmd-7-339-2014

Homer, C. G., Dewitz, J. A., Yang, L., Jin, S., Danielson, P., Xian, G., et al. (2015). Completion of the 2011 National Land Cover Database for the conterminous United States—Representing a decade of land cover change information. *Photogrammetric Engineering and Remote Sensing*, *81*(5), 345–354.

Houyoux, M. R., Vukovich, J. M., Coats, C. J. Jr., Wheeler, N. J. M., & Kasibhatla, P. S. (2000). Emission inventory development and processing for the Seasonal Model for Regional Air Quality (SMRAQ) project. *Journal of Geophysical Research*, *105*(D7), 9079–9090. https://doi.org/10.1029/1999JD900975

Jia, Y., Yu, G., He, N., Zhan, X., Fang, H., Sheng, W., et al. (2014). Spatial and decadal variations in inorganic nitrogen wet deposition in China induced by human activity. *Scientific Reports*, *4*(1). https://doi.org/10.1038/srep03763

Lamarque, J. F., Dentener, F., McConnell, J., Ro, C. U., Shaw, M., Vet, R., et al. (2013). Multi-model mean nitrogen and sulfur deposition from the atmospheric chemistry and climate model intercomparison project (ACCMIP): Evaluation of historical and projected future changes. *Atmospheric Chemistry and Physics*, *13*(16), 7997–8018. https://doi.org/10.5194/acp-13-7997-2013

Latysh, N. E., & Wetherbee, G. A. (2012). Improved mapping of National Atmospheric Deposition Program wet-deposition in complex terrain using PRISM-gridded data sets. *Environmental Monitoring and Assessment*, *184*(2), 913–928. https://doi.org/10.1007/s10661-011-2009-7

LeBauer, D., & Treseder, K. (2008). Nitrogen limitation of net primary productivity. *Ecology*, *179*(2), 35–40. https://doi.org/10.1016/j.agee.2013.04.020

Li, Y., Schichtel, B. A., Walker, J. T., Schwede, D. B., Chen, X., Lehmann, C. M. B., et al. (2016). Increasing importance of deposition of reduced nitrogen in the United States. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(21), 5784–5879. https://doi.org/10.1073/pnas.1525736113

Ma, L. M., & Tan, Z. M. (2009). Improving the behavior of the cumulus parameterization for tropical cyclone prediction: Convection trigger. *Atmospheric Research*, *92*(2), 190–211. https://doi.org/10.1016/j.atmosres.2008.09.022

Makar, P. A., Akingunola, A., Aherne, J., Cole, A. S., Aklilu, Y.-A., Zhang, J., et al. (2018). Estimates of exceedances of critical loads for acidifying deposition in Alberta and Saskatchewan. *Atmospheric Chemistry and Physics*, *18*(13), 9897–9927. https://doi.org/10.5194/acp-18-9897-2018

Mathur, R., & Dennis, R. L. (2003). Seasonal and annual modeling of reduced nitrogen compounds over the eastern United States: Emissions, ambient levels, and deposition amounts. *Journal of Geophysical Research*, *108*(D15), 4481. https://doi.org/10.1029/2002JD002794

NADP/NTN (National Atmospheric Deposition Program) (NRSP-3) (2019). NADP Program Office, Wisconsin State Laboratory of Hygiene, Madison, WI.

Nanus, L., Campbell, D. H., Ingersoll, G. P., Clow, D. W., & Mast, M. A. (2003). Atmospheric deposition maps for the Rocky Mountains. *Atmospheric Environment*, *37*(35), 4881–4892. https://doi.org/10.1016/j.atmosenv.2003.08.024

Otte, T. L., & Pleim, J. E. (2010). The Meteorology-Chemistry Interface Processor (MCIP) for the CMAQ modeling system: Updates through MCIPv3.4.1. *Geoscientific Model Development*, *3*(1), 243–256. https://doi.org/10.5194/gmd-3-243-2010

Pebesma, E. J. (2004). Multivariable geostatistics in S: The gstat package. *Computers & Geosciences*, *30*(7), 683–691. https://doi.org/10.1016/j.cageo.2004.03.012

Pleim, J. E., Bash, E. O., Walker, J. T., & Cooter, E. J. (2013). Development and evaluation of an ammonia bidirectional flux parameterization for air quality models. *Journal of Geophysical Research: Atmospheres*, *118*, 3794–3806. https://doi.org/10.1002/jgrd.502622013

Qu, L., Xiao, H., Zheng, N., Zhang, Z., & Xu, Y. (2017). Comparison of four methods for spatial interpolation of estimated atmospheric nitrogen deposition in South China. *Environmental Science and Pollution Research*, *24*(3), 2578–2588. https://doi.org/10.1007/s11356-016-7995-0

Robichaud, A., & Ménard, R. (2014). Multi-year objective analyses of warm season ground-level ozone and $PM_{2.5}$ over North America using real-time observations and Canadian operational air quality models. *Atmospheric Chemistry and Physics*, *14*(4), 1769–1800. https://doi.org/10.5194/acp-14-1769-2014

Robichaud, A., Ménard, R., Zaïtseva, Y., & Anselmo, D. (2016). Multi-pollutant surface objective analyses and mapping of air quality health index over North America. *Air Quality, Atmosphere and Health*, *9*(7), 743–759. https://doi.org/10.1007/s11869-015-0385-9

Sahu, S., Gelfand, A. E., & Holland, D. M. (2010). Fusing point and areal level space–time data with application to wet deposition. *Applied Statistics Series C*, *59*(1), 77–103. https://doi.org/10.1111/j.1467-9876.2009.00685.x

Schwede, D. B., & Lear, G. G. (2014). A novel hybrid approach for estimating total deposition in the United States. *Atmospheric Environment*, *92*, 207–220. https://doi.org/10.1016/j.atmosenv.2014.04.008

Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's Tau. *Journal of the American Statistical Association*, *63*(324), 1379–1389. https://doi.org/10.1080/01621459.1968.10480934

Sickles, J. E. II, & Shadwick, D. S. (2007). Changes in air quality and atmospheric deposition in the eastern United States: 1990-2004. *Journal of Geophysical Research*, *112*, D17301. https://doi.org/10.1029/2006JD007843

Sickles, J. E. II, & Shadwick, D. S. (2015). Air quality and atmospheric deposition in the eastern US: 20 years of change. *Atmospheric Chemistry and Physics*, *15*(1), 173–197. https://doi.org/10.5194/acp-15-173-2015

Simon, H., Baker, K. R., & Phillips, S. (2012). Compilation and interpretation of photochemical model performance statistics published between 2006 and 2012. *Atmospheric Environment*, *61*, 124–139. https://doi.org/10.1016/j.atmosenv.2012.07.012

Sun, J., Fu, J. S., Lynch, J. A., Huang, K., & Gao, Y. (2017). Climate-driven exceedance of total (wet+dry) nitrogen (N)+sulfur (S) deposition to forest soil over the conterminous U.S. *Earth's Future*, *5*, 560–576. https://doi.org/10.1002/2017EF000588

Swall, J. L., & Foley, K. M. (2009). The impact of spatial correlation and incommensurability on model evaluation. *Atmospheric Environment*, *43*(6), 1204–1217. https://doi.org/10.1016/j.atmosenv.2008.10.057

Theil, H. (1950). A rank-invariant method of linear and polynomial regression analysis. I, II, III. *Proceedings of the Royal Netherlands Academy of Sciences*, *53*, 386–392. 521–525, 1397–1412

Warner, J. X., Dickerson, R. R., Wei, Z., Strow, L. L., Wang, Y., & Liang, Q. (2017). Increased atmospheric ammonia over the world's major agricultural areas detected from space. *Geophysical Research Letters*, *44*, 2875–2884. https://doi.org/10.1002/2016GL072305

Whitten, G. Z., Heo, G., Kimura, Y., McDonald-Buller, E., Allen, D. T., Carter, W. P. L., & Yarwood, G. (2010). A new condensed toluene mechanism for carbon bond: CB05-TU. *Atmospheric Environment*, *44*(40), 5346–5355. https://doi.org/10.1016/j.atmosenv.2009.12.029

Williams, J. J., Chung, S. H., Johansen, A. M., Lamb, B. K., Vaughan, J. K., & Beutel, M. (2017). Evaluation of atmospheric nitrogen deposition model performance in the context of U.S. critical load assessments. *Atmospheric Environment*, *150*, 244–255. https://doi.org/10.1016/j.atmosenv.2016.11.051

World Meteorological Organization (2017). Global Atmosphere Watch Workshop on Measurement-Model Fusion for Global Total Atmospheric Deposition (MMF-GTAD) (GAW Report No. 234), Geneva, Switzerland. https://library.wmo.int/doc_num.php?explnum_id=3563

Yarwood, G., Rao, S., Yocke, M., & Whitten, G. (2005). Updates to the carbon bond chemical mechanism: CB05. Final report to the US EPA, RT-0400675, available at: www.camx.com/files/cb05_final_report_120805.aspx, (accessed April 26, 2018).

Zhang, L., Jacob, D. J., Knipping, E. M., Kumar, N., Munger, J. W., Carouge, C. C., et al. (2012). Nitrogen deposition to the United States: Distribution, sources, and processes. *Atmospheric Chemistry and Physics*, *12*(10), 4539–4554. https://doi.org/10.5194/acp-12-4539-2012

Zhang, Y., Mathur, R., Bash, J. O., Hogrefe, C., Xing, J., & Roselle, S. J. (2018). Long-term trends in total inorganic nitrogen and sulfur deposition in the US from 1990 to 2010. *Atmospheric Chemistry and Physics*, *18*(12), 9091–9106. https://doi.org/10.5194/acp-18-9091-2018

Zhao, Y., Zhang, L., Chen, Y., Liu, X., Xu, W., Pan, Y., & Duan, L. (2017). Atmospheric nitrogen deposition to China: A model analysis on nitrogen budget and critical load exceedance. *Atmospheric Environment*, *153*, 32–40. https://doi.org/10.1016/j.atmosenv.2017.01.018

Zhu, L., Henze, D. K., Bash, J. O., Jeong, G.-R., Cady-Pereira, K. E., Shephard, M. W., et al. (2015). Global evaluation of ammonia bi-directional exchange and livestock diurnal variation schemes. *Atmospheric Chemistry and Physics*, *15*(22), 12,823–12,843. https://doi.org/10.5194/acp-15-12823-2015