



EPA Public Access

Author manuscript

Atmos Environ (1994). Author manuscript; available in PMC 2020 January 01.

About author manuscripts

Submit a manuscript

Published in final edited form as:

Atmos Environ (1994). 2019 ; 214: 1–116872. doi:10.1016/j.atmosenv.2019.116872.

Assessing PM_{2.5} Model Performance for the Conterminous U.S. with Comparison to Model Performance Statistics from 2007–2015

James T. Kelly¹, Shannon N. Koplitz¹, Kirk R. Baker¹, Amara L. Holder², Havala O.T. Pye², Benjamin N. Murphy², Jesse O. Bash², Barron H. Henderson¹, Norm Possiel¹, Heather Simon¹, Alison M. Eyth¹, Carey Jang¹, Sharon Phillips¹, Brian Timin¹

¹Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711, USA

²Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711, USA

Abstract

Previous studies have proposed that model performance statistics from earlier photochemical grid model (PGM) applications can be used to benchmark performance in new PGM applications. A challenge in implementing this approach is that limited information is available on consistently calculated model performance statistics that vary spatially and temporally over the U.S. Here, a consistent set of model performance statistics are calculated by year, season, region, and monitoring network for PM_{2.5} and its major components using simulations from versions 4.7.1–5.2.1 of the Community Multiscale Air Quality (CMAQ) model for years 2007–2015. The multi-year set of statistics is then used to provide quantitative context for model performance results from the 2015 simulation. Model performance for PM_{2.5} organic carbon in the 2015 simulation ranked high (i.e., favorable performance) in the multi-year dataset, due to factors including recent improvements in biogenic secondary organic aerosol and atmospheric mixing parameterizations in CMAQ. Model performance statistics for the Northwest region in 2015 ranked low (i.e., unfavorable performance) for many species in comparison to the 2007–2015 dataset. This finding motivated additional investigation that suggests a need for improved speciation of wildfire PM_{2.5} emissions and modeling of boundary layer dynamics near water bodies. Several limitations were identified in the approach of benchmarking new model performance results with previous results. Since performance statistics vary widely by region and season, a simple set of national

Correspondence to, James T Kelly, Air Quality Modeling Group | Air Quality Assessment Division, Office of Air Quality Planning & Standards | Office of Air & Radiation, US Environmental Protection Agency, 109 TW Alexander Drive | Mail Drop: C439-01, Research Triangle Park, NC 27711, Tel: (919) 541-0886 | Fax: (919) 541-0044 | kelly.james@epa.gov.

Disclaimer

The views in this manuscript are those of the authors alone and do not necessarily reflect the policy of the U.S. Environmental Protection Agency.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

The following are the Supplementary data to this article:

performance benchmarks (e.g., one or two targets per species and statistic) as proposed previously are inadequate to assess model performance throughout the U.S. Also, trends in model performance statistics for sulfate over the 2007 to 2015 period suggest that model performance for earlier years may not be a useful reference for assessing model performance for recent years in some cases. Comparisons of results from the 2015 base case with results from five sensitivity simulations demonstrated the importance of parameterizations of NH₃ surface exchange, organic aerosol volatility and production, and emissions of crustal cations for predicting PM_{2.5} species concentrations.

Keywords

CMAQ; model evaluation; crustal cations; ammonia; organic aerosol

1. Introduction

Photochemical grid models (PGMs) simulate concentrations of trace gases and particles in the atmosphere using numerical representations of the major physical and chemical production and loss processes. Since PGMs are based on mechanistic parameterizations, they are believed to have suitable predictive capability to be used in a wide range of assessments including human health and welfare risk analyses (USEPA, 2009, 2014b), policy cost-benefit assessments (USEPA, 2012c), air quality forecasting (Lee et al., 2017), and air quality management (SJVAPCD, 2018). The predictive capability of PGMs is established in part during model development by deriving process parameterizations from first principles, using fundamental laboratory and other scientific datasets, and evaluating model developments against field study measurements designed to isolate processes of interest. PGM predictions are also assessed by comparison with routine observations according to operational, diagnostic, dynamic, and probabilistic performance evaluation techniques (Dennis et al., 2010).

Operational model evaluation uses statistical and graphical comparisons to assess the overall agreement of model predictions and observations from routine monitoring networks. Operational evaluation statistics are often used in judging the appropriateness of modeling for a given application, and several studies have recommended approaches and best practices for operational evaluation. Boylan and Russell (2006) recommended goals and criteria for model performance in terms of mean fractional bias (MFB) and mean fractional error (MFE) statistics based on performance in earlier modeling studies for particulate matter (PM) and visibility impairment. The Boylan and Russell (2006) goals and criteria relax under low concentration conditions to accommodate a range of environments, but do not vary by region or season. Simon et al. (2012) compiled model performance statistics for multiple pollutants from 69 peer-reviewed publications during 2006–2012 and provided recommendations on a minimum set of statistics that should be reported in studies of regulatory relevance. The set includes absolute and normalized bias and error statistics along with information to assess performance for correlation and variability. Recently, Emery et al. (2017) provided recommendations on statistics and benchmarks to assess PGM performance based on information from 31 of the 69 studies compiled by Simon et al. (2012) in combination with

results from seven additional studies published between 2012 and 2015. For operational evaluations of $PM_{2.5}$ and several $PM_{2.5}$ components, Emery et al. (2017) recommended model performance goals and criteria for mean bias (MB), normalized mean bias (NMB), and Pearson correlation coefficient (r). The Emery et al. (2017) benchmarks do not vary with region, season, or pollutant concentration.

The previous studies identified several challenges in assessing the state of operational model performance using statistics available in the peer-reviewed literature. Limitations include inconsistency in the statistics reported, inconsistency in the temporal and spatial scales of aggregation, relatively large influence on multi-study compilations of a minority of studies that report many statistics, and possible publication bias in the model performance literature. Artifacts in measurements of total and speciated $PM_{2.5}$ (e.g., El-Sayed et al., 2016; Kim et al., 2015; Pye et al., 2018; Solomon et al., 2014) also complicate interpretation of model performance statistics. Another issue in compiling information across studies is inconsistency in the type of modeling performed. For instance, some applications are based on an optimal model configuration developed for a specific region and period. For routine annual modeling of the conterminous U.S., computational considerations and differences in performance in different regions and seasons limit the ability to optimize model configuration options and grid resolution. As a result, typical model performance statistics likely differ for routine national modeling compared to modeling tailored to a specific region and period (e.g., Murphy et al., 2017). In summary, limited information is available in the peer-reviewed literature to provide quantitative context for interpreting spatially and temporally varying operational model performance statistics for national simulations over the U.S.

In the current study, $PM_{2.5}$ is simulated over the conterminous U.S. during 2015 with the Community Multiscale Air Quality (CMAQ) model using 12-km horizontal resolution. Operational model performance statistics are calculated and compared with a consistent set of performance statistics developed for the years 2007–2015 from national 12-km modeling of the U.S. with varying versions of CMAQ. The 2015 model predictions are also compared with results of five sensitivity simulations to help interpret model performance and examine the influence of alternative model configurations. Information on model performance statistics for 2007–2015 developed in this study are provided in the supporting information (Tables S1–S9 and a supplementary file) for use in assessing 12-km modeling studies of the conterminous U.S.

2. Methods

2.1 Air Quality Modeling

The 2015 base case simulation was based on CMAQ version 5.2.1 (www.epa.gov/cmaq; <https://doi.org/10.5281/zenodo.1212601>) (Appel et al., 2018) on a domain covering the conterminous U.S. with 12-km horizontal resolution and 35 vertical layers as part of a recent study (Kelly et al., 2019). Gas-phase chemistry was represented with the Carbon Bond 2006 mechanism (CB6r3; Emery et al., 2015), inorganic aerosol thermodynamics were based on ISORROPIA II (Fountoukis and Nenes, 2007; Nenes et al., 1998), primary organic aerosol (POA) was modeled as non-volatile (Appel et al., 2017; Simon and Bhawe, 2012), and

secondary organic aerosol (SOA) from volatile organic compounds was based on Pye et al. (2017). Chemical boundary conditions (BCs) were developed from a CMAQ simulation on a larger domain that used BCs based on a hemispheric CMAQ simulation (Mathur et al., 2017). U.S. anthropogenic emissions were based on version 2 of the 2014 national emission inventory (NEI) (USEPA, 2019a). Day-specific satellite-based fire activity data and fuel-specific emissions were used to generate wild, prescribed (Baker et al., 2016, 2018), and cropland fire (Pouliot et al., 2017; Zhou et al., 2018) emissions for 2015. Electric generation unit emissions were based on continuous emission monitoring data from 2015. Mobile source emissions were simulated for 2014 and 2016 with MOVES2014a (www.epa.gov/moves) and were interpolated to 2015. Emissions of biogenic compounds (Bash et al., 2016), windblown dust (Foroutan et al., 2017), and sea-spray aerosol (Gantt et al., 2015; Kelly et al., 2010) were simulated online using 2015 meteorology. NH₃ surface-exchange was simulated using an updated version of the CMAQv5.2.1 bidirectional exchange parameterization (Bash et al., 2013; Pleim et al., 2013). Specifically, the resistance parameterization of Pleim et al. (2013) was replaced with that of Massad et al. (2010), and the maximum amount of NH₄⁺ in soil-water solution was estimated using the sorption model of Venterea et al. (2015) rather than a fixed fraction of total NH₄⁺. Meteorological fields for CMAQ modeling were based on version 3.8 of the Weather Research and Forecasting (WRF) model (Skamarock et al., 2008).

Five sensitivity simulations were conducted with CMAQ to examine the influence of emissions and model configuration options on PM_{2.5} predictions in the 2015 base case. First, a simulation with wild and prescribed (but not cropland) fire emissions set to zero (“no.fire” case) was conducted to understand the influence of modeled fires on predictions. Second, a simulation with gridded emissions of crustal PM_{2.5} components set to zero (“no.crustal” case) was conducted to examine the influence of crustal cations on nitrate predictions via their effects on inorganic aerosol thermodynamics. Crustal cation emissions from sea spray and windblown dust are calculated during CMAQ execution and are included in all simulations. Third, a simulation was conducted using the default version of the NH₃ bidirectional exchange parameterization in CMAQv5.2.1 and updated versions of the emissions inventory and BCs that became available during the study (“nei.bc.nh3” case). National total emissions for NO_x, SO₂, and primary EC and OC in the nei.bc.nh3 case were within 1% of the emissions in the base case. Fourth, a simulation without the bidirectional surface-exchange parameterization for NH₃ (“no.bidi” case) was conducted to understand the influence of this model option on performance. Finally, a simulation was conducted where the organic aerosol treatment of Murphy et al. (2017) was used (“pc.soa” case) instead of the non-volatile POA treatment. The Murphy et al. (2017) parameterization treats POA as semi-volatile and produces SOA from an additional species that is emitted in proportion to POA emissions to approximate potential missing SOA production from combustion sources.

As part of previous studies, CMAQ simulations for the conterminous U.S. were conducted for years 2007–2014. Consistent with the 2015 base case, these simulations used 12-km horizontal resolution for annual simulations of air quality over the U.S. CMAQ versions ranged from 4.7.1 to 5.2 and WRF versions ranged from 3.1 to 3.8.1 for the 2007 to 2014 simulations. The differences in model configuration over the years introduced variability

into the model performance statistics, which is useful in providing a broader range of recent performance results for comparison with the new modeling. More details on the configuration of the 2007–2014 simulations are provided in Table 1 and references therein.

2.2 Model Performance Statistics

Model performance statistics were calculated consistently for all simulations using measurements of major PM_{2.5} components from Chemical Speciation Network (CSN) and Interagency Monitoring for the Protection of Visual Environments (IMPROVE) monitoring sites (Solomon et al., 2014). Comparisons were made between modeled and observed concentrations that were paired in space and time by averaging predictions to the 24-h sampling period of each measurement. Particle mass in the sub-2.5 μm diameter size range was calculated directly from the predicted particle size distributions. Nolte et al. (2015) reported that summation of particles mass in the Aitken and accumulation modes provides a similar estimate as the direct calculation used here. Since CSN sites tend to be in urban areas and IMPROVE sites tend to be in rural areas, model performance statistics are considered separately for the two networks. Statistics were calculated for PM_{2.5} components by season, year, network, and U.S. climate region (Karl and Koss, 1984; Fig. S1). Information on model performance for total PM_{2.5} is available in Tables S1 and S10 and previous work (Kelly et al., 2019) but is not discussed below for brevity.

The following statistics are considered below (see Table S11 for definitions): NMB (%), MB ($\mu\text{g m}^{-3}$), root-mean-square error (RMSE) ($\mu\text{g m}^{-3}$), and Pearson r . Normalized mean error (NME) (%), MFB, and MFE values are also provided in the supporting information. This set of statistics is consistent with recommendations of Simon et al. (2012) and Emery et al. (2017). Absolute values of NMB and MB were used in ranking performance for the 2015 base case against the multi-year set of statistics because NMB and MB can have positive and negative values.

3. Results

3.1 Overview

In this section, model performance statistics for PM_{2.5} sulfate, nitrate, organic carbon (OC), and elemental carbon (EC) are discussed. PM_{2.5} ammonium is not discussed because of measurement uncertainties (e.g., Yu et al., 2006) and its strong correlation with sulfate and nitrate. Due to the large number of possible comparisons (e.g., species, seasons, regions, networks, statistics), key features of model performance are considered below, and the supporting information is used to provide additional details. The median and range of the model performance statistics for simulations of 2007–2015 are provided in Tables S1–S9, and performance statistics for the 2015 base case are provided in Tables S10 and S12–S15. The full table of performance statistics for the individual annual simulations are available in a supporting file. Since distinct performance issues were observed for the Northwest, performance for the Northwest is discussed in more detail following the discussion of performance for the species.

3.2 Sulfate

Model performance statistics for PM_{2.5} sulfate at CSN sites for the 2015 base case simulation are illustrated in Fig. 1 by season and U.S. climate region. Values of the sulfate performance statistics at CSN and IMPROVE sites are provided in Table S12. NMB for sulfate at CSN sites is generally within $\pm 20\%$ (Fig. 1), with a notable exception of the Northwest region, where NMB is greater than 80% for most seasons. MB is generally within $\pm 0.2 \mu\text{g m}^{-3}$ and has a similar spatial and seasonal dependence as NMB. The highest RMSE values are in the Northeast and Ohio Valley regions due to the relatively high absolute sulfate concentrations. For instance, the mean observed sulfate concentrations at CSN sites in the Northeast ($1.45 \mu\text{g m}^{-3}$) and Ohio Valley ($1.81 \mu\text{g m}^{-3}$) are several times that in the Northwest ($0.52 \mu\text{g m}^{-3}$). The correlation coefficient for sulfate predictions in the Northeast and Ohio Valley is relatively low in winter compared with other seasons, which could be due to challenges in simulating oxidation mechanisms in winter (Shah et al., 2018). Correlation coefficients are also generally lower in the western U.S., where concentrations are relatively low. The annual correlation coefficient for predictions at CSN sites ranged from 0.31 (Southwest) to 0.79 (Upper Midwest) (Table S12).

Model performance statistics for sulfate at CSN sites for the 2015 base case are compared with statistics for 2007–2015 in Fig. 2. The percentiles give the relative level of performance for the base case compared with the multi-year set of statistics. Since the 2015 base case is included in the multi-year set, the model performance percentiles for the base case fall between 0% (lowest rank) and 100% (highest rank). Sulfate performance for the 2015 base case compares favorably with previous modeling in terms of NMB, MB, and RMSE, with exceptions of the Southwest in Spring and Northwest in all seasons. The rank of the correlation coefficient for the 2015 sulfate predictions is relatively low compared with that of NMB, MB, and RMSE. A challenge in comparing sulfate performance statistics across years is that model performance for sulfate appears to have been influenced by the substantial decreases in ambient sulfate in the U.S. during the 2007 to 2015 period. For instance, negative correlation exists between year and RMSE ($r: 0.70$), $r(r: 0.74)$, and observed concentration ($r: 0.57$) based on annual values at IMPROVE and CSN sites over 2007–2015 (Figs. S2 and S3). The relatively low rank for r performance for the 2015 base case for regions in the eastern U.S. is consistent with the trend of decreasing r over 2007–2015. The trend is associated with reductions in SO₂ emissions from electric generation units that have decreased the summertime sulfate peaks (i.e., signal-to-noise ratio) in the eastern U.S. (Chan et al., 2018; gispub.epa.gov/neireport/2014/). Therefore, although performance statistics from modeling of earlier years may be helpful in providing context for a new model case, performance for earlier years may set inappropriate standards to judge new modeling in cases where the underlying atmospheric conditions have changed substantially between the modeling periods.

A prominent feature of Fig. 1, Fig. 2 is the relatively poor performance for sulfate in the Northwest in the 2015 base case simulation. Predicting sulfate in the Northwest is relatively challenging because the concentrations are typically low and representing source-receptor relationships is difficult due to the complex terrain. In Fig. 3, NMB for sulfate at CSN and IMPROVE sites is compared for the base case simulation and the nei.bc.nh3 and no.fire

simulations. Updates to the emission inventory and BCs in the nei.bc.nh3 case improved sulfate performance statistics in the Northwest and Southwest. These improvements are due in part to emission regulations for shipping sources (USEPA, 2019b) that were better represented in the nei.bc.nh3 case than the base case. Recent reductions in SO₂ emissions in China (e.g., Krotkov et al., 2016; van der A et al., 2017; Zheng et al., 2018) were also better represented in the nei.bc.nh3 case and contributed to improved sulfate performance in the western U.S. Annual NMB for sulfate improved from 108% for the base case to 96% for the no.fire case at IMPROVE sites in the Northwest. High bias in sulfate predictions was evident on days where the model estimated a high fire contribution to the concentrations (e.g., Fig. S4). These results suggest that the model may overestimate sulfate from wildfires in the west. Multiple factors could have contributed to model performance issues for sulfate from wildfires including problems with plume rise and transport, excessive mixing of aloft plumes to the surface, and emissions issues (e.g., overestimates in SO₂ emissions, primary PM_{2.5} emissions, or the fraction of primary PM_{2.5} emissions speciated to sulfate). Laboratory measurements suggest that the modeled percent of primary PM_{2.5} emissions speciated as sulfate (0.33%) is not too high, because much larger values have been reported for fuels typical of the western U.S. such as needle leaf trees (0.68%) and chaparral (1.72%) (McMeeking et al., 2009). In previous modeling of fires in 2011 and 2013 (Baker et al., 2016, 2018), regional transport of wildfire plumes was captured reasonably well by the modeling system and sulfate predictions were relatively unbiased on days with wildfire impacts. Modeling wildfires in the 2015 base case could be relatively challenging because the 2015 fire season in the Pacific Northwest was the most severe in modern history by some metrics (USDA, 2016).

3.3 Nitrate

Model performance statistics for PM_{2.5} nitrate at CSN sites in the 2015 base case simulation are shown in Fig. 4 by season and U.S. climate region. Values of the nitrate performance statistics at CSN and IMPROVE sites are provided in Table S13. In the eastern U.S., NMB for nitrate at CSN sites is generally within $\pm 40\%$, with exceptions such as the Southeast in Fall (NMB: 98%) and Winter (NMB: 77%). Seasonal average modeled NO_y concentrations were within 14% of measured values in all seasons at Southeastern Aerosol Research Characterization (SEARCH) sites, and so overpredictions of nitrate in the Southeast do not appear to be due to overpredictions in the total oxides of nitrogen. In the western U.S., nitrate is generally biased low with NMBs between about -20 and -60% , except for the Northwest in Fall (62%), Spring (NMB: 107%), and Summer (NMB: 145%). Underpredictions in the West and Southwest in winter appear more pronounced when viewed in terms of MB (Fig. 4b) than NMB (Fig. 4a). These relatively large negative MBs are driven by underpredictions of nitrate in mountain valleys during ammonium nitrate episodes associated with strong meteorological temperature inversions in winter (e.g., Chen et al., 2012; Chow et al., 2006; Franchin et al., 2018). Modeling stagnant meteorology in complex terrain often requires finer grid resolution than is currently possible in national-scale modeling (e.g., $<1-4$ km; Crosman and Horel, 2017), and wintertime nitrate episodes have been simulated reasonably well in the western U.S. for higher-resolution CMAQ simulations (e.g., Chen et al., 2014; Kelly et al., 2018). Since model performance can differ for PGM simulations at different grid resolutions (e.g., Zakoura and Pandis, 2018),

comparisons of performance statistics among simulations to assess operational model performance are most meaningful when a consistent grid resolution is used. The correlation coefficients for nitrate predictions at CSN sites are greater than 0.6 in much of the U.S., with exceptions in cases where nitrate concentrations are low (e.g., during summer and in the Southwest) and in the Northwest (Fig. 4d).

Model performance statistics for nitrate at CSN sites for the 2015 base case are compared with statistics for 2007–2015 in Fig. 5. In general, NMB, MB, and RMSE statistics for nitrate predictions from the 2015 base simulation rank high compared with the full 2007–2015 dataset, with exceptions of the Northern Rockies and Plains in Winter, Upper Midwest in Spring, and the Northwest. The high NMBs for nitrate predictions in the Southeast in Fall and Winter (Fig. 4a) are not anomalous compared with the full 2007–2015 dataset (Fig. 5). A persistent high bias in $PM_{2.5}$ nitrate predictions occurred in the Southeast in Fall and Winter during 2007–2015 (NMB: 32–119%, Table S4), although nitrate is typically a small fraction of $PM_{2.5}$ in the Southeast (e.g., about 6% on average for CSN sites in 2015). Correlation coefficients for nitrate in the 2015 base case tend to compare less favorably to the multi-year dataset than do NMB, MB, and RMSE. However, correlation coefficients for nitrate predictions in the 2015 base case were high in some cases (e.g., r : 0.63 for Ohio Valley in Fall; Table S13) where performance ranked in the lowest category (0–20%) in comparison to the 2007–2015 dataset. This behavior illustrates that relatively weak performance in new modeling compared with previous modeling does not necessarily imply a model performance issue. The annual correlation coefficient for nitrate predictions over all CSN sites in the 2015 base case (0.71) was the highest in the 2007–2015 dataset (0.62–0.71; Fig. S6d).

Recent studies have identified that over-predictions of crustal cations can influence model performance for nitrate by affecting particle pH and gas-particle partitioning of total nitrate (i.e., $NO_3^- + HNO_3$) (Pye et al., 2018; Shah et al., 2018; Vasilakos et al., 2018). Crustal cations were biased high in the 2015 base case over much of the U.S. (Fig. S7). These biases may derive in part from issues with the nonpoint and fugitive dust emission sectors considering the magnitude and spatial correspondence of the emissions (Fig. S8) and predicted concentrations (Fig. S9). Also, crustal cation concentrations were underpredicted in the Southwest in Summer when windblown dust is relatively active (Fig. S7). Dust concentrations in the northern U.S. were influenced by transport from Canada in the model (Fig. S9), but sources of dust in Canada are unlikely to explain overpredictions of crustal cations in regions far from the border, such as the Southeast. In Fig. 6, NMB for nitrate at CSN and IMPROVE sites for the 2015 base case is compared with values for sensitivity simulations. For the no.crustal simulation, nitrate concentrations and NMBs are lower than for the base case due to increases in particle acidity associated with lower concentrations of water-soluble crustal cations in the model (i.e., Ca^{2+} , Mg^{2+} , and K^+). The increases in particle acidity reduce the fraction of total nitrate in the particle phase in the no.crustal simulation compared with the base case simulation (Fig. S10). In areas of the country with overpredictions of nitrate (e.g., Southeast and Northwest), the reductions in crustal cation emissions improved model bias (e.g., annual NMB improved from 47 to 22% at CSN sites in the Southeast). In the West and Southwest, the nitrate NMB was slightly worse in the

no.crustal simulation than the base case consistent with underpredictions of crustal cation concentrations in the Southwest in the 2015 base case.

The effect of removing the bidirectional NH₃ treatment (no.bidi case) on nitrate NMB was often directionally similar but smaller than for removing the gridded crustal cation emissions, with some exceptions such as the Ohio Valley, South, and Southwest at IMPROVE sites (Fig. 6). Turning off the bidirectional surface-exchange parameterization led to relatively near-source NH₃ deposition and lower NH₃ concentrations. Lower NH₃ concentration, in turn, led to greater partitioning of total nitrate to HNO₃ (Fig. S10), which deposits rapidly. Model predictions for NH₃ were closer to observations from U.S. monitoring networks in the base case than in the no.bidi case, although model predictions were biased low in both cases (Table S16, Figs. S11 and S12). For the nei.bc.nh3 simulation, nitrate NMB was higher than in the base case indicating degraded performance in the Northeast, Southeast, and Northwest. The greater nitrate NMBs are associated with the greater NH₃ concentrations in the nei.bc.nh3 case, which used the default version of the CMAQv5.2.1 bidirectional exchange parameterization, compared with the base case, which used the updated bidirectional exchange parameterization (Figs. S11 and S12). The higher NH₃ levels in the nei.bc.nh3 case increased the fraction of total nitrate in the particle phase compared with the base case (Fig. S10). Despite the degraded performance for nitrate in the nei.bc.nh3 case, NH₃ predictions for the nei.bc.nh3 case were closer to observations in Winter and Spring at SEARCH and Atmospheric Ammonia Network (AMoN; Butler et al., 2016) sites than were the base case predictions.

For the no.fire simulation, model bias for nitrate is similar to that in the base case with some exceptions such as the Northwest, where bias is lower and performance is improved (e.g., NMB improved from 76 to 48% at IMPROVE sites in the Northwest; Fig. 6). The wildfire sector emits significant amounts of NH₃ (Bray et al., 2016) and NO_x that can contribute to nitrate formation. In previous studies (Baker et al., 2018; Cai et al., 2016), the modeling system simulated plume-top NO_x concentrations reasonably well and produced little HNO₃ in wildfire plumes, which suggests that the wildfire-driven nitrate overpredictions in the Northwest are due to factors other than NO_x emissions and chemistry. Laboratory testing of fuels common to this region (McMeeking et al., 2009) suggest the nitrate fraction of primary PM_{2.5} emissions from wildfires (0.85–1.02%) is similar to the value used here (1.07%).

Another challenge in simulating nitrate is representing nitrate production due to heterogeneous hydrolysis of N₂O₅ (Chang et al., 2011). Nitrate production from this pathway is highly variable and difficult to parameterize in PGMs (Jaeglé et al., 2018; McDuffie et al., 2018). McDuffie et al. (2018) reported that the Davis et al. (2008) parameterization used in CMAQ produced higher N₂O₅ uptake coefficients than were estimated from aircraft measurements in 2015. Future work is planned to better constrain estimates of N₂O₅ uptake using field study data.

3.4 Organic carbon

Model performance statistics for PM_{2.5} OC at CSN sites for the 2015 base case are shown in Fig. 7 by season and U.S. climate region. Values of the OC performance statistics at CSN and IMPROVE sites are provided in Table S14. For half of the region-season cases in Fig. 7,

NMB is within $\pm 20\%$, but underpredictions of -23 to -50% occur in six of the nine regions in summer. Summertime OC underpredictions could be due in part to too little SOA production, although predictions are relatively unbiased in the Southeast in summer (NMB: 7.6%), where biogenic SOA concentrations are high. In contrast to the underpredictions for summer, OC predictions were biased high (3–76%) in seven of the nine regions in winter. Emissions associated with home heating (e.g., wood combustion) and prescribed burning in the Southeast are relatively important in winter (Odman et al., 2018; Watson et al., 2015) and lead to model OC overestimates under non-volatile POA assumptions (Murphy et al., 2017). Challenges in simulating prescribed burning along the Gulf of Mexico contributed to $PM_{2.5}$ overpredictions in winter there (Fig. S17). For instance, the predicted OC concentration at Breton Island, LA on 14 February was $250 \mu\text{g m}^{-3}$ in the base case (and $2.2 \mu\text{g m}^{-3}$ in the no.fire case) when the observed concentration was less than $5 \mu\text{g m}^{-3}$ (Fig. S18). RMSE for OC is generally between 1 and $2 \mu\text{g m}^{-3}$ in the eastern U.S. and tends to be higher in the western U.S., especially in the Northwest region. The correlation coefficient for OC is frequently between 0.4 and 0.6, but is between 0.6 and 0.8 in the West in all seasons and six of the nine regions in Fall. Overall, model performance for the 2015 simulation ranked high in the full 2007–2015 dataset (Fig. 8), with a few exceptions (e.g., Northern Rockies and Plains in Fall and Summer). The relatively good performance for OC could be due to a range of factors including recent improvements in the biogenic SOA parameterizations (Pye et al., 2013, 2017) and atmospheric mixing (USEPA, 2015b) in CMAQ.

NMB values for OC at CSN and IMPROVE sites for the 2015 base case are compared with values for the sensitivity simulations in Fig. 9. The lower NMB values for the no.fire simulation than the base case indicate generally worse model performance for the simulation without wild and prescribed fire emissions. This behavior is most evident at IMPROVE sites in regions of the western U.S. where fires were most prevalent (Fig. S19). Consistent with previous studies (Baker et al, 2016, 2018), model performance for OC is mixed within regions and across seasons (Fig. S20) on days with modeled fire influence. The model had some skill in predicting OC associated with large fires for sites in the Northwest (Flathead and Glacier), but underpredictions (Fresno) and overpredictions (Redwood) are evident at other sites (e.g., see Fig. S21).

For the pc.soa simulation, NMB for OC is lower at CSN sites and higher at IMPROVE sites compared with the base case. This behavior is consistent with the organic aerosol parameterization used in the pc.soa simulation in which organic aerosol may locally decrease due to POA evaporation, but generally increase in source regions and downwind due to additional SOA production relative to the base case (Murphy et al., 2017). The Murphy et al. (2017) treatment tends to increase OC during photochemically active daytime conditions (Fig. S22) and was previously shown to substantially reduce negative bias in at urban sites in California during summer when evaluated with hourly observations. OC model performance statistics were generally similar for the base case and the pc.soa case, although NMB performance degraded moderately in some regions in the pc.soa simulation (e.g., at IMPROVE sites in the east) (Fig. 9). These results should be interpreted cautiously, however, because differences in OC measurements from collocated IMPROVE and CSN sites have been previously reported (e.g., Kim et al., 2015; Weakley et al., 2016). NMB for

OC was noticeably better in the Northwest for the pc.soa case than the base case, although this behavior may be related to issues with simulating boundary layer mixing (see section 3.5). The influence on OC NMB of updates to the emission inventory, BCs, and NH₃ surface-exchange parameterization in the nei.bc.nh3 case was moderate and mixed overall. Correlation coefficients for OC improved significantly in the Upper Midwest and several other regions in the nei.bc.nh3 case compared with the base case (Fig. S23).

3.5 Elemental carbon

Model performance statistics for PM_{2.5} EC at CSN sites in the 2015 base case simulation are shown in Fig. 10 by season and U.S. climate region. Values of the EC performance statistics at CSN and IMPROVE sites are provided in Table S15. Modeled EC is biased high for all regions and seasons compared with CSN (and IMPROVE, Table S15) observations. NMB values are between 20 and 60% for 23 of the 36 region-season cases in Fig. 10a. Despite some large NMBs, MB values are generally less than 0.2 μg m⁻³ due to the low concentrations of EC (e.g, Fig. S15), with an exception of the Northwest region for which MBs are greater than 0.8 μg m⁻³ (see section 3.5 below). Annual correlation coefficients for EC predictions at CSN sites ranged from 0.39 (Northwest) to 0.65 (Northern Rockies and Plains). Model performance for EC in the base case often ranks high in comparison to statistics for the 2007–2015 dataset (Fig. 11), with exceptions of the Northern Rockies and Plains and Northwest regions, which experienced unusually high wildfire activity in 2015. PM_{2.5} EC predictions were biased high on median throughout the 2007–2015 period (Table S9).

NMB for PM_{2.5} EC predictions in the 2015 base case is compared with NMB for the PM_{2.5} EC in the sensitivity simulations in Fig. 12. The bias in EC predictions was smaller in the no.fire simulation than the base case, and the biggest improvements in the no.fire simulation compared with the base case occurred at IMPROVE sites in the western U.S. For instance, NMB improved from 90 to 7% in the West, and 120 to -12% in the Northern Rockies and Plains. On days where the model predicted substantial influence of fire emissions on concentrations at monitoring sites, EC concentrations were generally overestimated in the base case (Fig. S4) consistent with previous studies (Baker et al., 2016, 2018). Issues with primary PM_{2.5} emissions from the wildfire sector or the fraction of PM_{2.5} emissions speciated to EC may contribute to the overpredictions. In the nei.bc.nh3 case, model performance for EC improved slightly compared with the base case.

3.6 Model Performance in the Northwest

Consistently higher MB and NMB for the predictions in Northwest compared to the other regions are evident in Fig. 1, Fig. 4, Fig. 7, Fig. 10, with NMB exceeding 60% for most seasons across all species examined. This pattern is particularly pronounced for EC, where NMB exceeds 100% across the whole year. The greater high biases for EC in the Northwest are likely due in part to assumptions about the speciation of PM_{2.5} emissions from wildland fires and challenges in simulating boundary layer dynamics near the Puget Sound and other coastal areas.

The NMB for EC improved dramatically in the Northwest in the no.fire simulation, reducing from 309% in the base case to 149% in the no.fire case at IMPROVE sites. The speciation profile for wildland fire PM_{2.5} emissions, which was based on profiles estimated in two studies (Chow et al., 2004; Watson et al., 1996) according to Reff et al. (2009), allocates 9.5% of total PM_{2.5} from all wildfire emissions as EC. The Chow et al. (2004) and Watson et al. (1996) speciation profiles were developed by combining measurements across several burn experiments, some of which included measurements from pile burns of dried vegetative clippings and fence posts in addition to natural biomass. A more recent laboratory study also suggests that EC may contribute a relatively large percentage of PM_{2.5} emissions for fuels in the southeastern (7.5%) and southwestern (e.g., 20% for chaparral) U.S. (Hosseini et al., 2013). However, separate laboratory measurements for fuels common in the western U.S. suggest that EC makes up a smaller percentage of total PM_{2.5} emissions (e.g., 1.4% for montane and 4.3% for chaparral ecosystems) (McMeeking et al., 2009) than the 9.5% value used in our modeling. Additionally, recent field measurements from forest fires in the southeastern U.S. (Aurell et al., 2015) suggest that the composite profile developed by Reff et al. (2009) that was used in developing the NEI may have overestimated the contribution of EC to emitted PM_{2.5} mass by a factor of four (Fig. 13). Updated speciated contributions for wildland fire EC are scheduled for implementation in future U.S. EPA emission inventories and should correct some of the relatively high biases in EC performance in the Northwest and the other western fire-prone regions observed in this assessment.

In addition to issues with speciation profiles for fire emissions, overpredictions for EC and other primary species at some sites in the Northwest appear to be due to meteorological factors. For instance, EC was strongly overestimated in the no.fire simulation at the Puget Sound IMPROVE site (PUGE1) in Seattle, which is west of the Cascade Mountain Range and upwind from many of the major fire areas in eastern Washington, Oregon, and Idaho. Unusually high concentrations of EC (>20 µg m⁻³) and other species (e.g., soil and OC) were predicted at this site during periods where the modeled boundary layer height was extremely low (<50 m) (e.g., Figs. S24 and S25). Simulating atmospheric mixed layers is challenging near the shoreline (McNider et al., 2018), and modeled boundary layer heights may have been underestimated at times due to the low water temperature of the Puget Sound. Other monitoring sites in Washington and Oregon show similar patterns of elevated EC and soil concentrations during periods with low simulated boundary layer heights (e.g., Fig. S26) suggesting that this issue is not confined to the site in Seattle. Further evaluation of modeled boundary layer dynamics in this region would be of interest. Also, although this section emphasizes the role of fire emissions and boundary layer dynamics, other factors likely contributed to the high biases for EC predictions in the Northwest and other regions.

4. Summary and Conclusions

Comparing model performance statistics for new modeling with consistently calculated statistics from previous modeling studies can be useful in assessing model performance (Boylan and Russell, 2006; Emery et al., 2017; Simon et al., 2012; USEPA, 2018b). However, limited information on performance statistics from previously published studies is available to provide quantitative context for annual 12-km modeling of the conterminous U.S. Here, a consistent set of model performance statistics by season, year, network, and

U.S. climate region are developed for PM_{2.5} and its major components (i.e., sulfate, nitrate, OC, and EC) using 12-km CMAQ simulations for 2007 to 2015. The multi-year set of statistics are then used to interpret model performance for the 2015 base simulation.

Several insights were provided by comparison of model performance statistics for the 2015 base case simulation with statistics for the full 2007–2015 dataset. Performance statistics for OC in the base case generally ranked high compared with performance statistics for the 2007–2015 dataset and built confidence in the 2015 simulation results. The relatively good performance for OC in 2015 could be due to improvements in emissions modeling as well as the modeling of atmospheric mixing (USEPA, 2015b) and biogenic SOA (Pye et al., 2013, 2017) in recent versions of CMAQ. Model performance for sulfate and nitrate in the 2015 base case also ranked high in general (excluding the Northwest) for NMB, MB, and RMSE compared with the 2007–2015 dataset. Comparison of 2015 model performance statistics with results for previous years helped identify relatively weak model skill in the Northwest in 2015. Additional investigation of performance in the Northwest indicated that the fraction of wildfire PM_{2.5} emissions speciated as EC is likely too high and the simulated boundary layer height is frequently too low at coastal urban sites in the Northwest in the 2015 simulation. Model performance for sulfate in the Northwest was also found to improve with updates to shipping emissions and boundary conditions that better reflected recent reductions in SO₂ emissions in Asia.

Limitations in comparing performance statistics for the 2015 base case with results of earlier modeling also emerged from our study. First, performance statistics were found to vary widely by region and season due to the spatially and temporally varying nature of the underlying processes. This behavior suggests that a simple set of nationally representative statistical benchmarks (e.g., one or two national benchmarks per statistic and species) as has been proposed previously would not be adequate to assess model results throughout the U.S. Second, trends in model performance due to trends in ambient air quality can compromise the value of using performance statistics from previous modeling in assessing the appropriateness of new modeling. For example, recent decreases in sulfate concentration in the eastern U.S. appear to have contributed to improvements in MB and RMSE and degradation in r for CMAQ predictions over the 2007–2015 period. Third, relatively high rank of performance for new modeling compared with previous modeling can occur even when model skill is weak in the new modeling, and vice-versa. For example, NMB for nitrate at CSN sites in the Southeast in Winter appeared high in the 2015 base case (NMB: 77%), but this value ranked reasonably high (i.e., 40–60% category) in the 2007–2015 dataset. Conversely, good correlation for nitrate (e.g., r : 0.63 for Ohio Valley in Fall) in the 2015 base case ranked in the lowest category (0–20%) compared with the full 2007–2015 period. These limitations should be considered when using previous model performance results to provide context for new modeling. Summary statistics for the full 2007–2015 dataset are provided in Tables S1–S9 and statistics for individual cases are available in a supplementary file.

In addition to providing a database of model performance statistics and illustrating their use, sensitivity simulations were used to examine model performance for the 2015 base case. Improved performance for sulfate, nitrate, and EC at IMPROVE sites in the western U.S. in

a simulation with wild and prescribed fire emissions removed suggested that sulfur and EC (and possibly NH₃) emissions may be overestimated from the wildfire sector. Fire-related OC performance was mixed but removing all wild and prescribed fire emissions resulted in unreasonable concentrations at monitors in the West and Southeast, consistent with previous studies (Baker et al., 2016, 2018). A simulation with reduced crustal cation emissions demonstrated that these species have an important influence on nitrate concentrations, and crustal cation concentrations were frequently overestimated in the base case. Simulations with different treatments of NH₃ surface exchange in CMAQ demonstrated that NH₃ and nitrate concentrations are sensitive to NH₃ resistance and bidirectional flux parameterizations. These parameterizations are active areas of development in CMAQ. A simulation using a semi-volatile POA treatment and SOA formation from potentially missing combustion sources yielded higher OC concentrations at IMPROVE sites and lower OC concentrations at CSN sites. This behavior generally had a small impact on performance statistics, with some degradation in annual NMB at IMPROVE sites in the east. However, interpreting OC performance across CSN and IMPROVE networks is challenging due to differences in OC measurements at collocated CSN and IMPROVE sites identified in recent studies (e.g., Kim et al., 2015; Weakley et al., 2016). Additional work is warranted on understanding the volatility distribution of POA emissions and the magnitude of SOA precursor emissions, particularly from wild and prescribed fires and residential wood combustion. CMAQ predictions for wild and prescribed fires may need to be supplemented with other relevant information to gain a more comprehensive understanding of fire influence in cases where such predictions are of importance (e.g., extreme events in rural and remote areas).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

- Appel KW, Napelenok S, Hogrefe C, Pouliot G, Foley KM, Roselle SJ, Pleim JE, Bash J, Pye HOT, Heath N, Murphy B, Mathur R, 2018 Overview and Evaluation of the Community Multiscale Air Quality (CMAQ) Modeling System Version 5.2, in: Mensink C, Kallos G (Eds.), *Air Pollution Modeling and its Application XXV*. Springer International Publishing, Cham, pp. 69–73.
- Appel KW, Napelenok SL, Foley KM, Pye HOT, Hogrefe C, Luecken DJ, Bash JO, Roselle SJ, Pleim JE, Foroutan H, Hutzell WT, Pouliot GA, Sarwar G, Fahey KM, Gantt B, Gilliam RC, Heath NK, Kang DW, Mathur R, Schwede DB, Spero TL, Wong DC, Young JO, 2017 Description and evaluation of the Community Multiscale Air Quality (CMAQ) modeling system version 5.1. *Geoscientific Model Development* 10, 1703–1732. [PubMed: 30147852]
- Aurell J, Gullett BK, Tabor D, 2015 Emissions from southeastern U.S. Grasslands and pine savannas: comparison of aerial and ground field measurements with laboratory burns. *Atmos. Environ* 111, 170–178.
- Baker KR, Woody MC, Tonnesen GS, Hutzell W, Pye HOT, Beaver MR, Pouliot G, Pierce T, 2016 Contribution of regional-scale fire events to ozone and PM_{2.5} air quality estimated by photochemical modeling approaches. *Atmospheric Environment* 140, 539–554.
- Baker KR, Woody MC, Valin L, Szykman J, Yates EL, Iraci LT, Choi HD, Soja AJ, Koplitz SN, Zhou L, Campuzano-Jost P, Jimenez JL, Hair JW, 2018 Photochemical model evaluation of 2013 California wild fire air quality impacts using surface, aircraft, and satellite data. *Science of The Total Environment* 637–638, 1137–1149.

- Bash JO, Baker KR, Beaver MR, 2016 Evaluation of improved land use and canopy representation in BEIS v3.61 with biogenic VOC measurements in California. *Geosci. Model Dev* 9, 2191–2207.
- Bash JO, Cooter EJ, Dennis RL, Walker JT, Pleim JE, 2013 Evaluation of a regional air-quality model with bidirectional NH₃ exchange coupled to an agroecosystem model. *Biogeosciences* 10, 1635–1645.
- Boylan JW, Russell AG, 2006 PM and light extinction model performance metrics, goals, and criteria for three-dimensional air quality models. *Atmospheric Environment* 40, 4946–4959.
- Bray C, Battye W, Aneja V, Tong DQ, Lee P, Tang Y, 2016 Impact of Wildfires on Atmospheric Ammonia Concentrations in the US: Coupling Satellite and Ground Based Measurements. *Electronic Conference on Atmospheric Sciences*, DOI: 10.3390/ecas2016-B001 , 10.3390/ecas2016-B001 <https://sciforum.net/manuscripts/3406/manuscript.pdf>, <https://sciforum.net/manuscripts/3406/manuscript.pdf>
- Butler T, Vermeylen F, Lehmann CM, Likens GE, Puchalski M, 2016 Increasing ammonia concentration trends in large regions of the USA derived from the NADP/AMoN network. *Atmospheric Environment* 146, 132–140.
- Cai C, Kulkarni S, Zhao Z, Kaduwela AP, Avise JC, DaMassa JA, Singh HB, Weinheimer AJ, Cohen RC, Diskin GS, Wennberg P, Dibb JE, Huey G, Wisthaler A, Jimenez JL, Cubison MJ, 2016 Simulating reactive nitrogen, carbon monoxide, and ozone in California during ARCTAS-CARB 2008 with high wildfire activity. *Atmospheric Environment* 128, 28–44.
- Chan EAW, Gantt B, McDow S, 2018 The reduction of summer sulfate and switch from summertime to wintertime PM_{2.5} concentration maxima in the United States. *Atmospheric Environment* 175, 25–32.
- Chang WL, Bhawe PV, Brown SS, Riemer N, Stutz J, Dabdub D, 2011 Heterogeneous Atmospheric Chemistry, Ambient Measurements, and Model Calculations of N₂O₅: A Review. *Aerosol Science and Technology* 45, 665–695.
- Chen J, Lu J, Avise JC, DaMassa JA, Kleeman MJ, Kaduwela AP, 2014 Seasonal modeling of PM_{2.5} in California's San Joaquin Valley. *Atmospheric Environment* 92, 182–190.
- Chen LWA, Watson JG, Chow JC, Green MC, Inouye D, Dick K, 2012 Wintertime particulate pollution episodes in an urban valley of the Western US: a case study. *Atmos. Chem. Phys* 12, 10051–10064.
- Chow JC, Chen L-WA, Watson JG, Lowenthal DH, Magliano KA, Turkiewicz K, Lehrman DE, 2006 PM_{2.5} chemical composition and spatiotemporal variability during the California Regional PM₁₀/PM_{2.5} Air Quality Study (CRPAQS). *Journal of Geophysical Research: Atmospheres* 111.
- Chow JC, Watson JG, Kuhns H, Etyemezian V, Lowenthal DH, Crow D, Kohl SD, Engelbrecht JP, Green MC, 2004 Source profiles for industrial, mobile, and area sources in the Big Bend Regional Aerosol Visibility and Observational study. *Chemosphere* 54, 185–208. [PubMed: 14559270]
- Crosman ET, Horel JD, 2017 Large-eddy simulations of a Salt Lake Valley cold-air pool. *Atmospheric Research* 193, 10–25.
- Davis JM, Bhawe PV, Foley KM, 2008 Parameterization of N₂O₅ reaction probabilities on the surface of particles containing ammonium, sulfate, and nitrate. *Atmos. Chem. Phys* 8, 5295–5311.
- Dennis R, Fox T, Fuentes M, Gilliland A, Hanna S, Hogrefe C, Irwin J, Rao ST, Scheffe R, Schere K, Steyn D, Venkatram A, 2010 A framework for evaluating regional-scale numerical photochemical modeling systems. *Environmental Fluid Mechanics* 10, 471–489. [PubMed: 21461126]
- El-Sayed MMH, Amenumey D, Hennigan CJ, 2016 Drying-Induced Evaporation of Secondary Organic Aerosol during Summer. *Environmental Science & Technology* 50, 3626–3633. [PubMed: 26910726]
- Emery C, Jung J, Koo B, Yarwood G, 2015 Improvements to CAMx snow cover treatments and Carbon Bond chemical mechanism for winter ozone Final Report, prepared for Utah Department of Environmental Quality, Salt Lake City, UT Prepared by Ramboll Environ, Novato, CA, 8 2015. .
- Emery C, Liu Z, Russell AG, Odman MT, Yarwood G, Kumar N, 2017 Recommendations on statistics and benchmarks to assess photochemical model performance. *Journal of the Air & Waste Management Association* 67, 582–598. [PubMed: 27960634]

- Foroutan H, Young J, Napelenok S, Ran L, Appel KW, Gilliam RC, Pleim JE, 2017 Development and evaluation of a physics-based windblown dust emission scheme implemented in the CMAQ modeling system. *Journal of Advances in Modeling Earth Systems* 9, 585–608. [PubMed: 30245776]
- Fountoukis C, Nenes A, 2007 ISORROPIA II: a computationally efficient thermodynamic equilibrium model for K⁺, Ca²⁺, Mg²⁺, NH₄⁺, Na⁺, SO₄²⁻, NO₃⁻, Cl⁻, H₂O aerosols. *Atmos. Chem. Phys* 7, 4639–4659.
- Franchin A, Fibiger DL, Goldberger L, McDuffie EE, Moravek A, Womack CC, Crosman ET, Docherty KS, Dube WP, Hoch SW, Lee BH, Long R, Murphy JG, Thornton JA, Brown SS, Baasandorj M, Middlebrook AM, 2018 Airborne and ground-based observations of ammonium-nitrate-dominated aerosols in a shallow boundary layer during intense winter pollution episodes in northern Utah. *Atmos. Chem. Phys* 18, 17259–17276.
- Gantt B, Kelly JT, Bash JO, 2015 Updating sea spray aerosol emissions in the Community Multiscale Air Quality (CMAQ) model version 5.0.2. *Geosci. Model Dev* 8, 3733–3746.
- Hays MD, Geron CD, Linna KJ, Smith ND, Schauer JJ, 2002 Speciation of Gas-Phase and Fine Particle Emissions from Burning of Foliar Fuels. *Environmental Science & Technology* 36, 2281–2295. [PubMed: 12075778]
- Hosseini S, Urbanski SP, Dixit P, Qi L, Burling IR, Yokelson RJ, Johnson TJ, Shrivastava M, Jung HS, Weise DR, Miller JW, Cocker DR III, 2013 Laboratory characterization of PM emissions from combustion of wildland biomass fuels. *Journal of Geophysical Research: Atmospheres* 118, 9914–9929.
- Jaeglé L, Shah V, Thornton JA, Lopez-Hilfiker FD, Lee BH, McDuffie EE, Fibiger D, Brown SS, Veres P, Sparks TL, Ebben CJ, Wooldridge PJ, Kenagy HS, Cohen RC, Weinheimer AJ, Campos TL, Montzka DD, Digangi JP, Wolfe GM, Hanisco T, Schroder JC, Campuzano-Jost P, Day DA, Jimenez JL, Sullivan AP, Guo H, Weber RJ, 2018 Nitrogen Oxides Emissions, Chemistry, Deposition, and Export Over the Northeast United States During the WINTER Aircraft Campaign. *Journal of Geophysical Research: Atmospheres* 123, 3123–3139.
- Karl TR, Koss WJ, 1984 Regional and National Monthly, Seasonal, and Annual Temperature Weighted by Area, 1895–1983 Historical Climatology Series 4–3, National Climatic Data Center, Asheville, NC, 38 pp. <https://www.ncdc.noaa.gov/monitoring-references/maps/us-climate-regions.php>
- Kelly JT, Bhawe PV, Nolte CG, Shankar U, Foley KM, 2010 Simulating emission and chemical evolution of coarse sea-salt particles in the Community Multiscale Air Quality (CMAQ) model. *Geoscientific Model Development* 3, 257–273.
- Kelly JT, Jang CJ, Timin B, Gantt B, Reff A, Zhu Y, Long S, Hanna A, 2019 A system for developing and projecting PM_{2.5} spatial fields to correspond to just meeting National Ambient Air Quality Standards. *Atmospheric Environment: X*, 100019.
- Kelly JT, Parworth CL, Zhang Q, Miller DJ, Sun K, Zondlo MA, Baker KR, Wisthaler A, Nowak JB, Pusede SE, Cohen RC, Weinheimer AJ, Beyersdorf AJ, Tonnesen GS, Bash JO, Valin LC, Crawford JH, Fried A, Walega JG, 2018 Modeling NH₄NO₃ Over the San Joaquin Valley During the 2013 DISCOVER-AQ Campaign. *Journal of Geophysical Research: Atmospheres* 123, 4727–4745.
- Kim PS, Jacob DJ, Fisher JA, Travis K, Yu K, Zhu L, Yantosca RM, Sulprizio MP, Jimenez JL, Campuzano-Jost P, Froyd KD, Liao J, Hair JW, Fenn MA, Butler CF, Wagner NL, Gordon TD, Welti A, Wennberg PO, Crouse JD, St. Clair JM, Teng AP, Millet DB, Schwarz JP, Markovic MZ, Perring AE, 2015 Sources, seasonality, and trends of southeast US aerosol: an integrated analysis of surface, aircraft, and satellite observations with the GEOS-Chem chemical transport model. *Atmos. Chem. Phys* 15, 10411–10433.
- Krotkov NA, McLinden CA, Li C, Lamsal LN, Celarier EA, Marchenko SV, Swartz WH, Bucsela EJ, Joiner J, Duncan BN, Boersma KF, Veefkind JP, Levelt PF, Fioletov VE, Dickerson RR, He H, Lu Z, Streets DG, 2016 Aura OMI observations of regional SO₂ and NO₂ pollution changes from 2005 to 2015. *Atmos. Chem. Phys* 16, 4605–4629.
- Lee P, McQueen J, Stajner I, Huang J, Pan L, Tong D, Kim H, Tang Y, Kondragunta S, Ruminski M, Lu S, Rogers E, Saylor R, Shafran P, Huang H-C, Gorline J, Upadhyay S, Artz R, 2017 NAQFC Developmental Forecast Guidance for Fine Particulate Matter (PM_{2.5}). *Weather and Forecasting* 32, 343–360.

- Massad RS, Nemitz E, Sutton MA, 2010 Review and parameterisation of bi-directional ammonia exchange between vegetation and the atmosphere. *Atmos. Chem. Phys* 10, 10359–10386.
- Mathur R, Xing J, Gilliam R, Sarwar G, Hogrefe C, Pleim J, Pouliot G, Roselle S, Spero TL, Wong DC, Young J, 2017 Extending the Community Multiscale Air Quality (CMAQ) modeling system to hemispheric scales: overview of process considerations and initial applications. *Atmos. Chem. Phys* 17, 12449–12474. [PubMed: 29681922]
- McDuffie EE, Fibiger DL, Dubé WP, Lopez-Hilfiker F, Lee BH, Thornton JA, Shah V, Jaeglé L, Guo H, Weber RJ, Michael Reeves J, Weinheimer AJ, Schroder JC, Campuzano-Jost P, Jimenez JL, Dibb JE, Veres P, Ebben C, Sparks TL, Wooldridge PJ, Cohen RC, Hornbrook RS, Apel EC, Campos T, Hall SR, Ullmann K, Brown SS, 2018 Heterogeneous N₂O₅ Uptake During Winter: Aircraft Measurements During the 2015 WINTER Campaign and Critical Evaluation of Current Parameterizations. *Journal of Geophysical Research: Atmospheres* 123, 4345–4372.
- McMeeking GR, Kreidenweis SM, Baker S, Carrico CM, Chow JC, Collett JL Jr., Hao WM, Holden AS, Kirchstetter TW, Malm WC, Moosmüller H, Sullivan AP, Wold CE, 2009 Emissions of trace gases and aerosols during the open combustion of biomass in the laboratory. *Journal of Geophysical Research: Atmospheres* 114.
- McNider RT, Pour-Biazar A, Doty K, White A, Wu Y, Qin M, Hu Y, Odman T, Cleary P, Knipping E, Dornblaser B, Lee P, Hain C, McKeen S, 2018 Examination of the Physical Atmosphere in the Great Lakes Region and Its Potential Impact on Air Quality—Overwater Stability and Satellite Assimilation. *Journal of Applied Meteorology and Climatology* 57, 2789–2816.
- Murphy BN, Woody MC, Jimenez JL, Carlton AMG, Hayes PL, Liu S, Ng NL, Russell LM, Setyan A, Xu L, Young J, Zaveri RA, Zhang Q, Pye HOT, 2017 Semivolatile POA and parameterized total combustion SOA in CMAQv5.2: impacts on source strength and partitioning. *Atmos. Chem. Phys* 17, 11107–11133.
- Nenes A, Pandis SN, Pilinis C, 1998 ISORROPIA: A New Thermodynamic Equilibrium Model for Multiphase Multicomponent Inorganic Aerosols. *Aquatic Geochemistry* 4, 123–152.
- Nolte CG, Appel KW, Kelly JT, Bhawe PV, Fahey KM, Collett JL Jr, Zhang L, Young JO, 2015 Evaluation of the Community Multiscale Air Quality (CMAQ) model v5.0 against size-resolved measurements of inorganic particle composition across sites in North America. *Geosci. Model Dev* 8, 2877–2892.
- Odman MT, Huang R, Pophale AA, Sakhpara RD, Hu Y, Russell AG, Chang ME, 2018 Forecasting the Impacts of Prescribed Fires for Dynamic Air Quality Management. *Atmosphere* 9, 220.
- Pleim JE, Bash JO, Walker JT, Cooter EJ, 2013 Development and evaluation of an ammonia bidirectional flux parameterization for air quality models. *Journal of Geophysical Research-Atmospheres* 118, 3794–3806.
- Pouliot G, Rao V, McCarty JL, Soja A, 2017 Development of the crop residue and rangeland burning in the 2014 National Emissions Inventory using information from multiple sources. *Journal of the Air & Waste Management Association* 67, 613–622. [PubMed: 27964698]
- Pye HOT, Murphy BN, Xu L, Ng NL, Carlton AG, Guo H, Weber R, Vasilakos P, Appel KW, Budisulistiorini SH, Surratt JD, Nenes A, Hu W, Jimenez JL, Isaacman-VanWertz G, Misztal PK, Goldstein AH, 2017 On the implications of aerosol liquid water and phase separation for organic aerosol mass. *Atmos. Chem. Phys* 17, 343–369. [PubMed: 30147709]
- Pye HOT, Pinder RW, Piletic IR, Xie Y, Capps SL, Lin Y-H, Surratt JD, Zhang Z, Gold A, Luecken DJ, Hutzell WT, Jaoui M, Offenberg JH, Kleindienst TE, Lewandowski M, Edney EO, 2013 Epoxide Pathways Improve Model Predictions of Isoprene Markers and Reveal Key Role of Acidity in Aerosol Formation. *Environmental Science & Technology* 47, 11056–11064. [PubMed: 24024583]
- Pye HOT, Zuend A, Fry JL, Isaacman-VanWertz G, Capps SL, Appel KW, Foroutan H, Xu L, Ng NL, Goldstein AH, 2018 Coupling of organic and inorganic aerosol systems and the effect on gas–particle partitioning in the southeastern US. *Atmos. Chem. Phys* 18, 357–370. [PubMed: 29963078]
- Reff A, Bhawe PV, Simon H, Pace TG, Pouliot GA, Mobley JD, Houyoux M, 2009 Emissions Inventory of PM_{2.5} Trace Elements across the United States. *Environmental Science & Technology* 43, 5790–5796. [PubMed: 19731678]

- Shah V, Jaeglé L, Thornton JA, Lopez-Hilfiker FD, Lee BH, Schroder JC, Campuzano-Jost P, Jimenez JL, Guo H, Sullivan AP, Weber RJ, Green JR, Fiddler MN, Bililign S, Campos TL, Stell M, Weinheimer AJ, Montzka DD, Brown SS, 2018 Chemical feedbacks weaken the wintertime response of particulate sulfate and nitrate to emissions reductions over the eastern United States. *Proceedings of the National Academy of Sciences* 115, 8110–8115.
- Simon H, Baker KR, Phillips S, 2012 Compilation and interpretation of photochemical model performance statistics published between 2006 and 2012. *Atmospheric Environment* 61, 124–139.
- Simon H, Bhawe PV, 2012 Simulating the Degree of Oxidation in Atmospheric Organic Particles. *Environmental Science & Technology* 46, 331–339. [PubMed: 22107341]
- SJVAPCD, 2018 San Joaquin Valley Air Pollution Control District, 2018 Plan for the 1997, 2006, and 2012 PM_{2.5} Standards, Available: <http://valleyair.org/pmplans/documents/2018/pm-plan-adopted/2018-Plan-for-the-1997-2006-and-2012-PM2.5-Standards.pdf>.
- Skamarock WC, Klemp JB, Dudhia J, Gill DO, Barker DM, Duda MG, Huang X, Wang W, Powers JG, 2008 A description of the Advanced Research WRF version 3. NCAR Technical Note NCAR/TN-475+STR.
- Solomon PA, Crumpler D, Flanagan JB, Jayanty RKM, Rickman EE, McDade CE, 2014 US National PM_{2.5} Chemical Speciation Monitoring Networks-CSN and IMPROVE: Description of networks. *Journal of the Air & Waste Management Association* 64, 1410–1438. [PubMed: 25562937]
- USDA, 2016 U.S. Forest Service, Department of Agriculture, Narrative Timeline of the Pacific Northwest 2015 Fire Season, 1–281. Available: https://wfmrda.nwcg.gov/docs/_Reference_Materials/2015_Timeline_PNW_Season_FINAL.pdf.
- USEPA, 2009 Risk and Exposure Assessment for Review of the Secondary National Ambient Air Quality Standards for Oxides of Nitrogen and Oxides of Sulfur, U.S. EPA, Office of Air Quality Planning and Standards Research Triangle Park, NC EPA-452/R-09-008a, Available: <https://www3.epa.gov/ttn/naaqs/standards/no2so2sec/data/NOxSOxREASep2009MainContent.pdf>.
- USEPA, 2012a Hierarchical Bayesian Model (HBM)-Derived Estimates of Air Quality for 2007: Annual Report. Office of Research and Development, Office of Air and Radiation, Research Triangle Park, NC EPA/600/R-12/538 Available: <https://www.epa.gov/hesc/rsig-related-downloadable-data-files>.
- USEPA, 2012b Hierarchical Bayesian Model (HBM)-Derived Estimates of Air Quality for 2008: Annual Report. Office of Research and Development, Office of Air and Radiation, Research Triangle Park, NC EPA/600/R-12/048 Available: <https://www.epa.gov/hesc/rsig-related-downloadable-data-files>.
- USEPA, 2012c Regulatory Impact Analysis for the Final Revisions to the National Ambient Air Quality Standards for Particulate Matter. U.S. EPA, Office of Air Quality Planning and Standards, Research Triangle Park, NC EPA-452/R-12-005. Available: <https://www3.epa.gov/ttnecas1/regdata/RIAs/finalria.pdf>.
- USEPA, 2013 Bayesian Space-Time Downscaling Fusion Model (Downscaler) -Derived Estimates of Air Quality for 2009. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC EPA-454/R-13-003 Available: <https://www.epa.gov/hesc/rsig-related-downloadable-data-files>.
- USEPA, 2014a Bayesian Space-Time Downscaling Fusion Model (Downscaler) -Derived Estimates of Air Quality for 2010. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC EPA-454/S-14-001 Available: <https://www.epa.gov/hesc/rsig-related-downloadable-data-files>.
- USEPA, 2014b Health Risk and Exposure Assessment for Ozone. Final Report. U.S. EPA, Office of Air Quality Planning and Standards, Research Triangle Park, NC EPA-452/R-14-004a. Available: <https://www3.epa.gov/ttn/naaqs/standards/ozone/data/20140829healthrea.pdf>.
- USEPA, 2015a Bayesian Space-Time Downscaling Fusion Model (Downscaler) -Derived Estimates of Air Quality for 2011. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC EPA-454/S-15-001 Available: <https://www.epa.gov/hesc/rsig-related-downloadable-data-files>.
- USEPA, 2015b Release notes for CMAQv5.1 - november 2015. [https://www.airqualitymodeling.org/index.php/CMAQ_version_5.1_\(November_2015_releaseTechnical_Documentation\)](https://www.airqualitymodeling.org/index.php/CMAQ_version_5.1_(November_2015_releaseTechnical_Documentation)).

- USEPA, 2016 Bayesian Space-Time Downscaling Fusion Model (Downscaler) -Derived Estimates of Air Quality for 2012. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC EPA-450/R-16-001 Available: <https://www.epa.gov/hesc/rsig-related-downloadable-data-files>.
- USEPA, 2017 Bayesian Space-Time Downscaling Fusion Model (Downscaler) -Derived Estimates of Air Quality for 2013. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC EPA-450/R-17-001. Available: <https://www.epa.gov/hesc/rsig-related-downloadable-data-files>.
- USEPA, 2018a Bayesian Space-Time Downscaling Fusion Model (Downscaler) -Derived Estimates of Air Quality for 2014. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC EPA-454/R-18-008. Available: <https://www.epa.gov/hesc/rsig-related-downloadable-data-files>.
- USEPA, 2018b Modeling Guidance for Demonstrating Attainment of Air Quality Goals for Ozone, PM_{2.5}, and Regional Haze. U.S. EPA, Office of Air Quality Planning and Standards, Research Triangle Park, NC EPA 454/R-18-009. Available: https://www3.epa.gov/ttn/scram/guidance/guide/O3-PM-RH-Modeling_Guidance-2018.pdf.
- USEPA, 2019a Technical support document (TSD) - preparation of emissions inventories for the version 7.1 2015 emissions modeling platform. Available: <https://www.epa.gov/air-emissions-modeling/2015v71-alpha-platform>.
- USEPA, 2019b website (accessed March 2019): www.epa.gov/enforcement/act-preventpollution-ships-apps-enforcement-case-resolutions.
- van der A RJ, Mijling B, Ding J, Koukouli ME, Liu F, Li Q, Mao H, Theys N, 2017 Cleaning up the air: effectiveness of air quality policy for SO₂ and NO_x emissions in China. *Atmos. Chem. Phys* 17, 1775–1789.
- Vasilakos P, Russell A, Weber R, Nenes A, 2018 Understanding nitrate formation in a world with less sulfate. *Atmos. Chem. Phys* 18, 12765–12775.
- Venterea RT, Clough TJ, Coulter JA, Breuillin-Sessoms F, Wang P, Sadowsky MJ, 2015 Ammonium sorption and ammonia inhibition of nitrite-oxidizing bacteria explain contrasting soil N₂O production. *Scientific Reports* 5, 12153. [PubMed: 26179972]
- Watson JG, Blumenthal DL, Chow JC, Cahill CF, Richards LW, Dietrich D, Morris R, Houck JE, Dickson RJ, Andersen SR, 1996 Mt. Zirkel Wilderness Area reasonable attribution study of visibility impairment, Vol. II: Results of data analysis and modeling.
- Watson JG, Chow JC, Lowenthal DH, Antony Chen LW, Shaw S, Edgerton ES, Blanchard CL, 2015 PM_{2.5} source apportionment with organic markers in the Southeastern Aerosol Research and Characterization (SEARCH) study. *Journal of the Air & Waste Management Association* 65, 1104–1118. [PubMed: 26102211]
- Weakley AT, Takahama S, Dillner AM, 2016 Ambient aerosol composition by infrared spectroscopy and partial least-squares in the chemical speciation network: Organic carbon with functional group identification. *Aerosol Science and Technology* 50, 1096–1114.
- Yu X-Y, Lee T, Ayres B, Kreidenweis SM, Malm W, Collett JL, 2006 Loss of fine particle ammonium from denuded nylon filters. *Atmospheric Environment* 40, 4797–4807.
- Zakoura M, Pandis SN, 2018 Overprediction of aerosol nitrate by chemical transport models: The role of grid resolution. *Atmospheric Environment* 187, 390–400.
- Zhou L, Baker KR, Napelenok SL, Pouliot G, Elleman R, O'Neill SM, Urbanski SP, Wong DC, 2018 Modeling crop residue burning experiments to evaluate smoke emissions and plume transport. *Science of The Total Environment* 627, 523–533. [PubMed: 29426175]

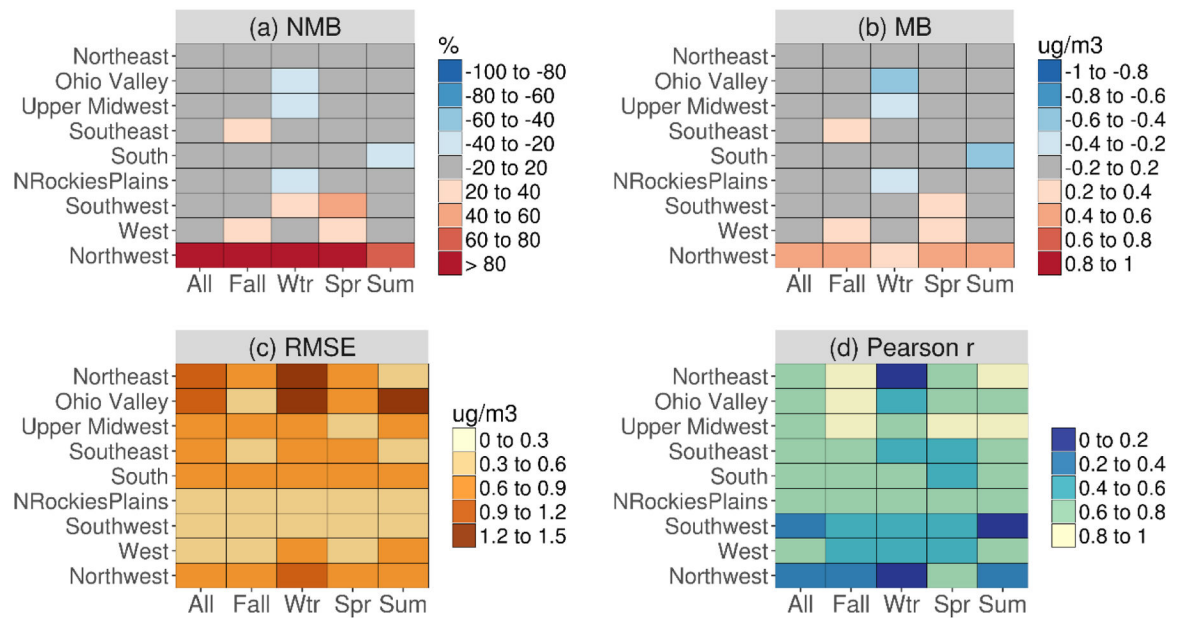


Figure 1. PM_{2.5} sulfate performance statistics for the 2015 base case at CSN sites by season and U.S. climate region.

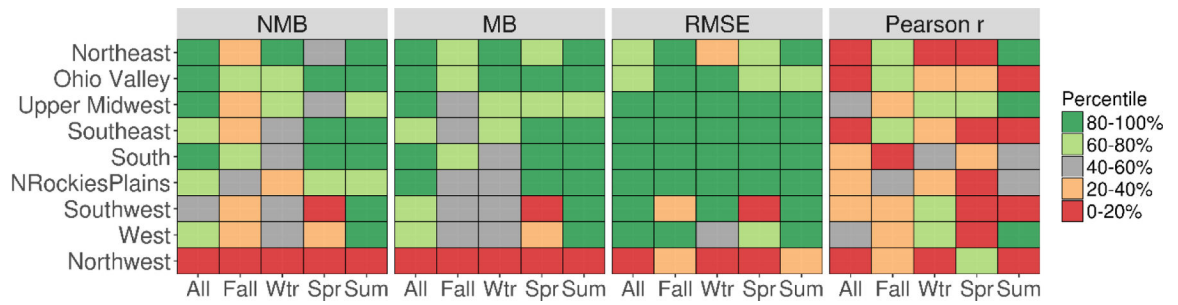


Figure 2. Percentile rank of PM_{2.5} sulfate performance statistics for the 2015 simulation relative to performance statistics for 2007–2015. Scale ranges from low rank (0–20%) to high rank (80–100%).

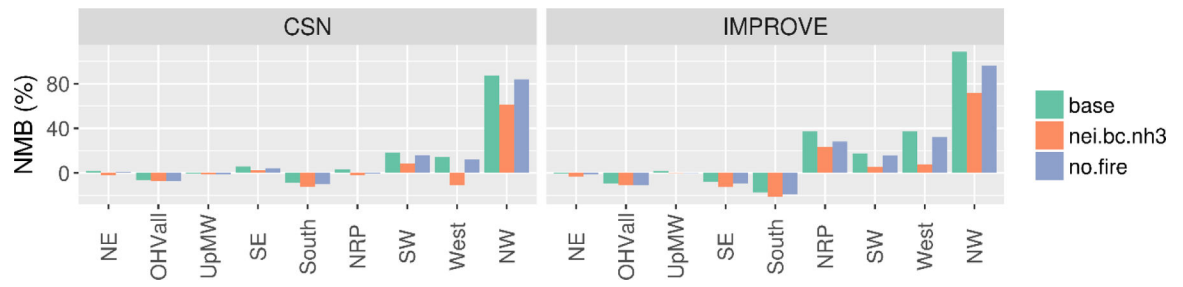


Figure 3. Comparison of annual NMB for PM_{2.5} sulfate in the 2015 base simulation with NMB for the no.fire and nei.bc.nh3 cases at CSN and IMPROVE sites.

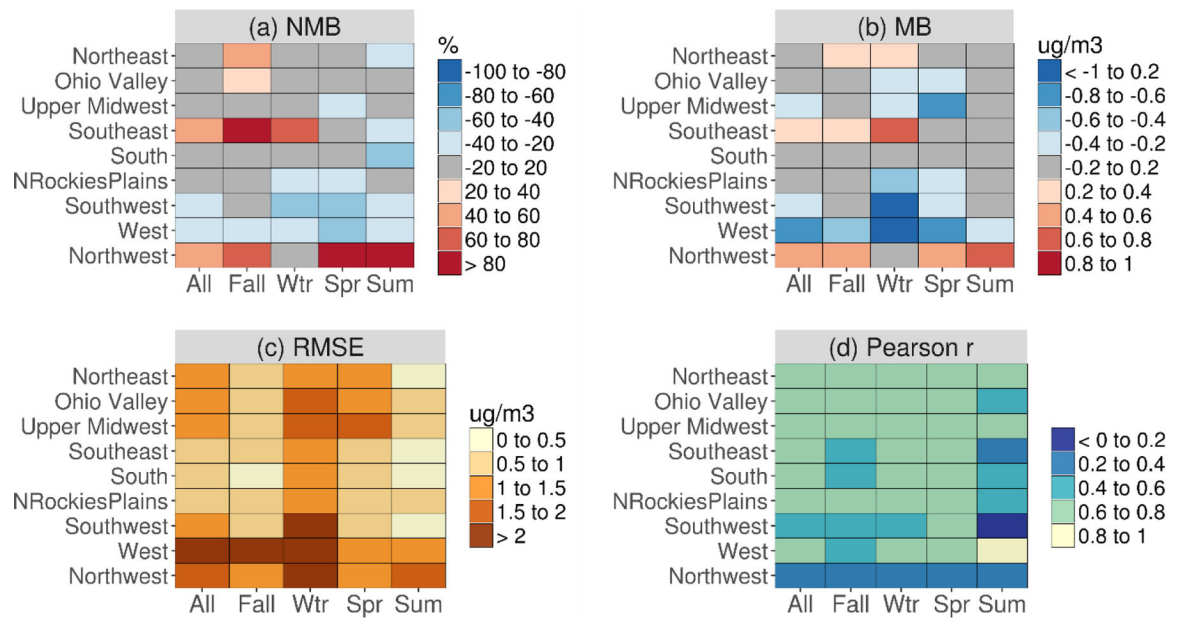


Figure 4. PM_{2.5} nitrate performance statistics for the 2015 base case at CSN sites by season and U.S. climate region.

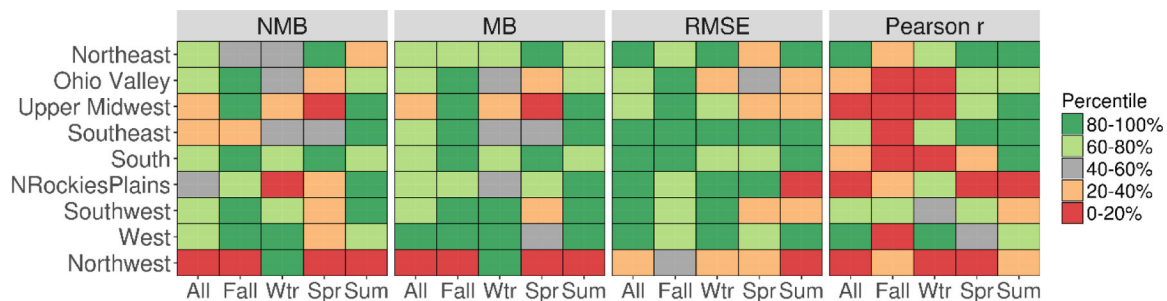


Figure 5. Percentile rank of PM_{2.5} nitrate performance statistics for the 2015 simulation relative to performance statistics for 2007–2015. Scale ranges from low rank (0–20%) to high rank (80–100%).

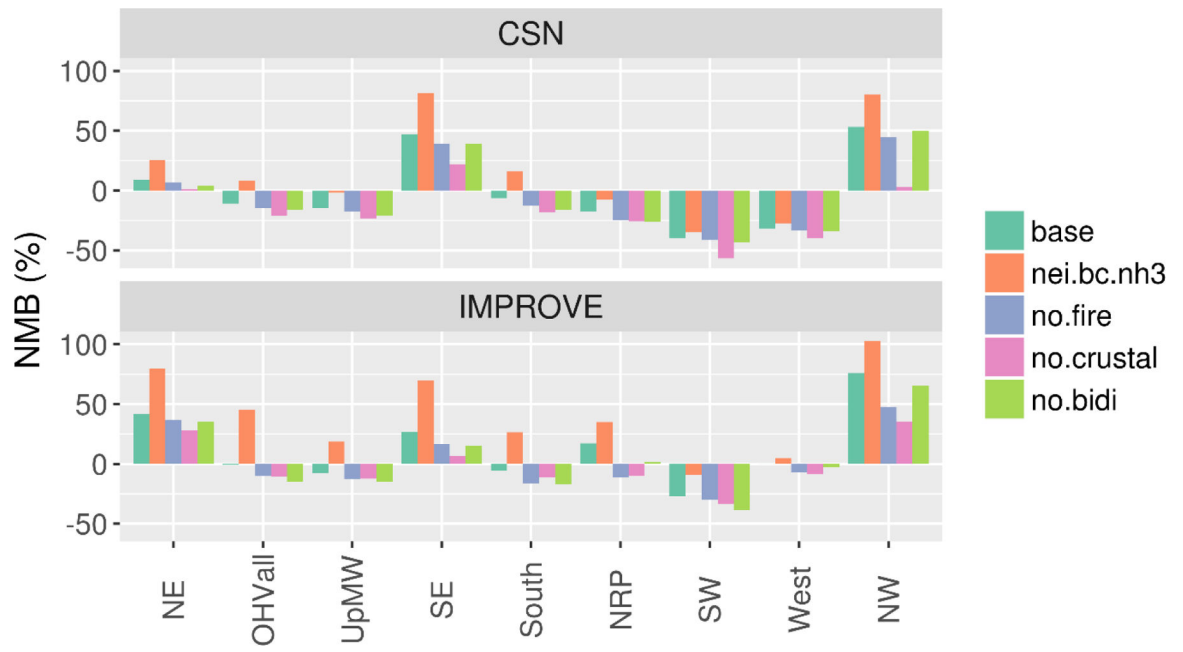


Figure 6. Comparison of annual NMB for PM_{2.5} nitrate in the 2015 base simulation with NMB for the sensitivity modeling cases at CSN and IMPROVE sites.

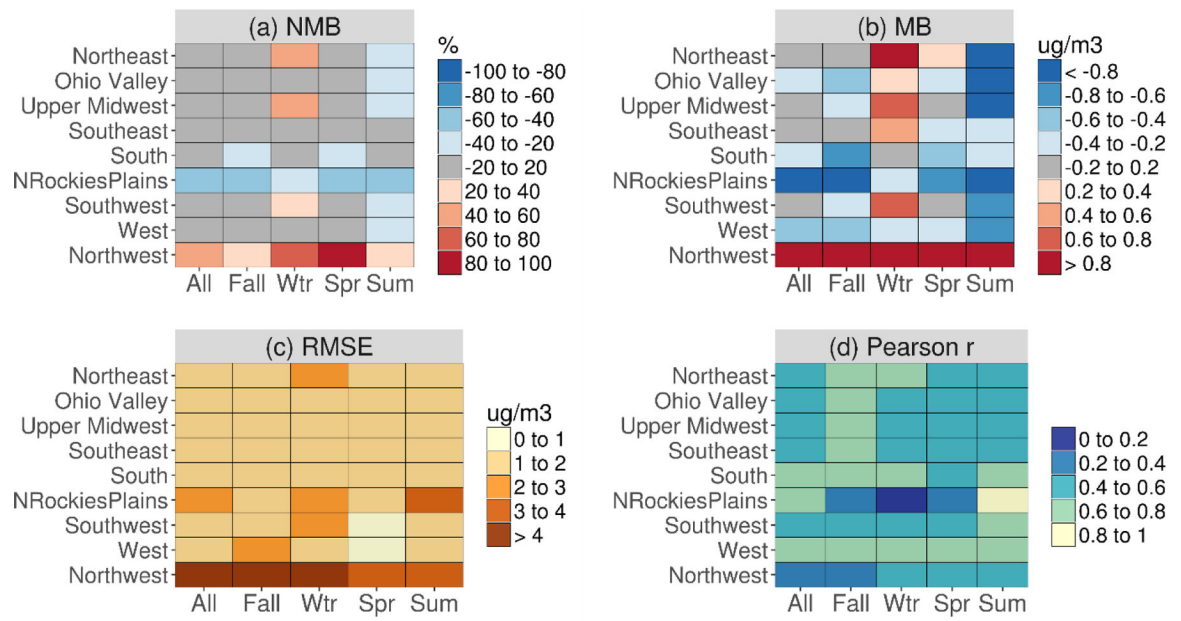


Figure 7. PM_{2.5} OC performance statistics for the 2015 base case at CSN sites by season and U.S. climate region.

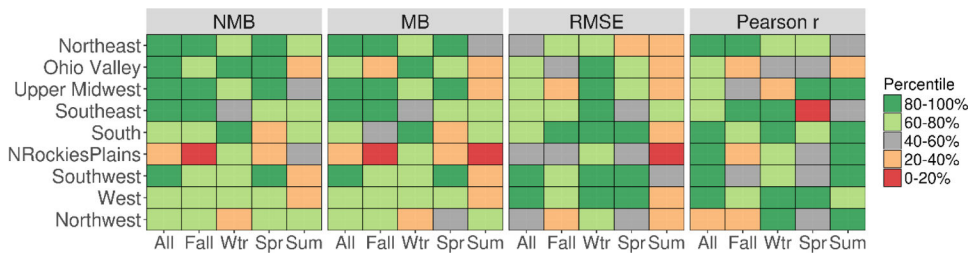


Figure 8. Percentile rank of PM_{2.5} OC performance statistics for the 2015 simulation relative to performance statistics for 2007–2015. Scale ranges from low rank (0–20%) to high rank (80–100%).

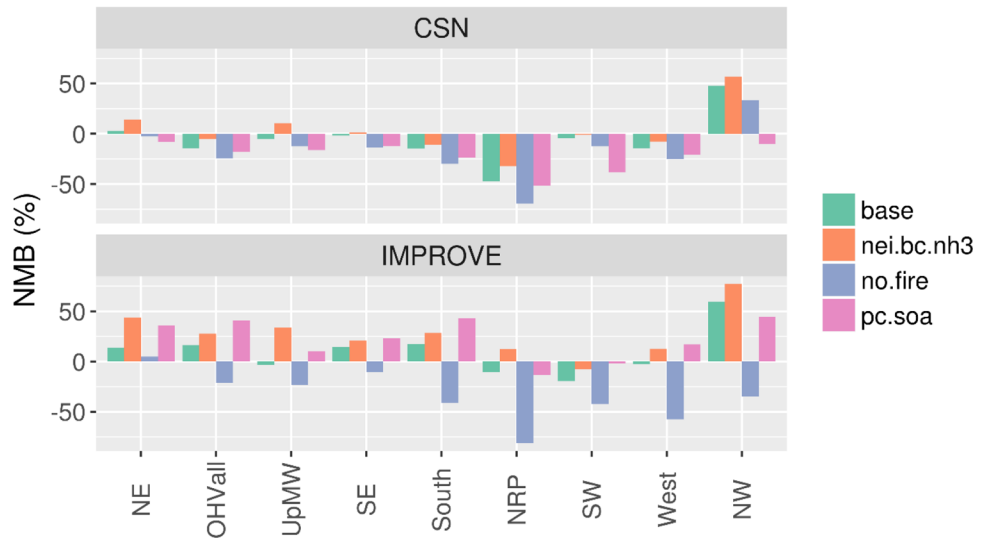


Figure 9. Comparison of annual NMB for PM_{2.5} OC in the 2015 base simulation with NMB for the no.fire, pc.soa, and nei.bc.nh3 cases at CSN and IMPROVE sites.

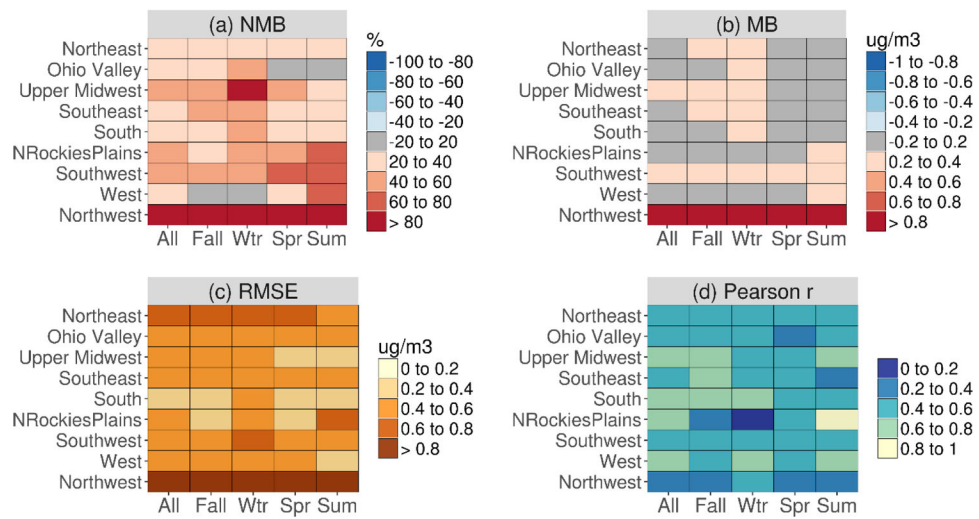


Figure 10. PM_{2.5} EC performance statistics for the 2015 base case at CSN sites by season and U.S. climate region.

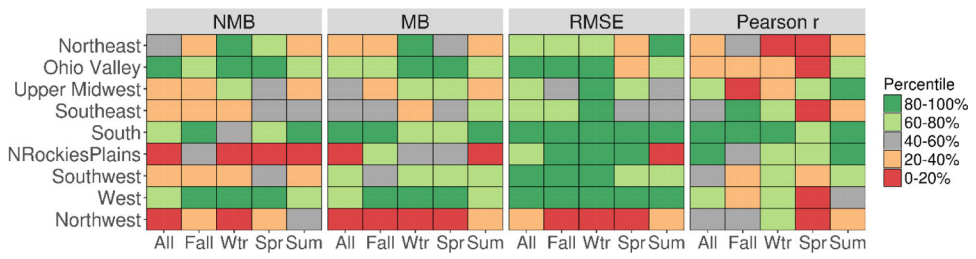


Figure 11. Percentile rank of PM_{2.5} EC performance statistics for the 2015 simulation relative to performance statistics for 2007–2015. Scale ranges from low rank (0–20%) to high rank (80–100%).

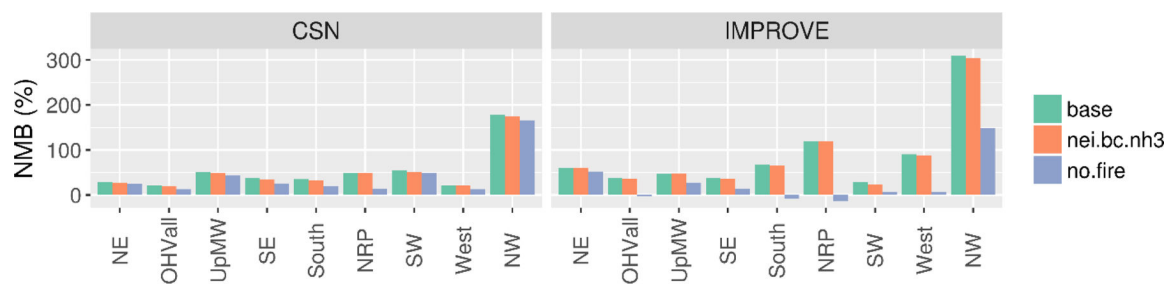


Figure 12. Comparison of annual NMB for PM_{2.5} EC in the 2015 base simulation with NMB for the no.fire and nei.bc.nh3 cases at CSN and IMPROVE sites.

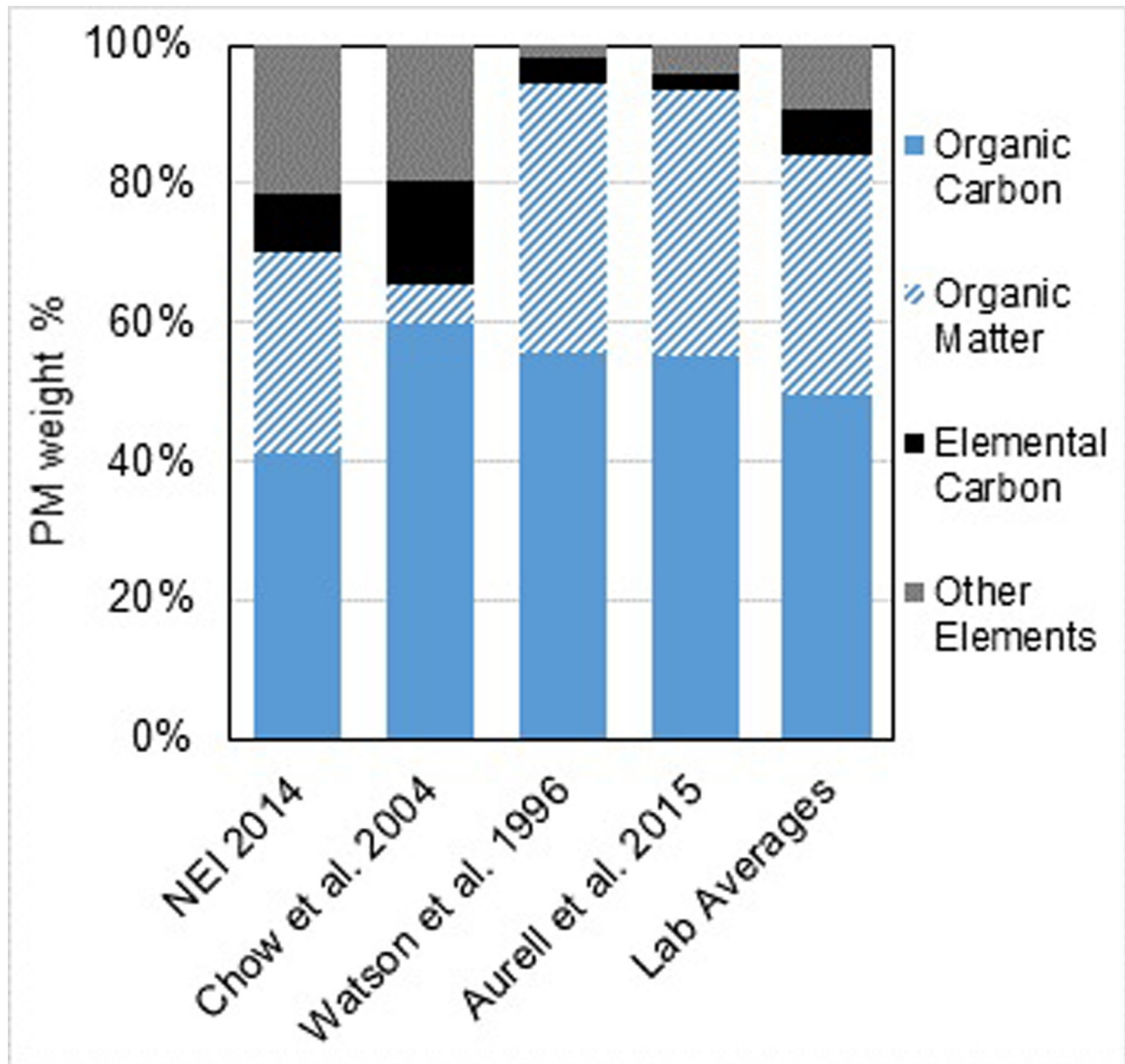


Figure 13. Comparison of wildfire PM_{2.5} speciation profiles. The assumed profile for the 2015 modeling platform is labeled “NEI 2014”. The “Lab Averages” profile represents the mean profile across several additional lab studies (Hays et al., 2002; Hosseini et al., 2013; McMeeking et al., 2009).

Table 1.

Model configuration for 2007–2015 simulations.

Year	Case name	Air Quality Model	Meteorological Model	Reference
2007	2007aq_07c_N5ao_inline	CMAQv4.7.1	WRFv3.1	USEPA (2012a)
2008	2008aa_08c_N5ao_inline	CMAQv4.7.1	WRFv3.1	USEPA (2012b)
2009	2009ef2_v5_09d_N5ao_inline	CMAQv4.7.1	WRFv3.2	USEPA (2013)
2010	2010ef_v5_10f_N5ao_inline	CMAQv4.7.1	WRFv3.4	USEPA (2014a)
2011	2011ef_v6_11g_ltngNO_bidi_25L	CMAQv5.0.2	WRFv3.4	USEPA (2015a)
2012	2012eh_cb05v2_v6_12g	CMAQv5.0.2	WRFv3.6.1	USEPA (2016)
2013	2013ej_v6_13i	CMAQv5.1	WRFv3.7.1	USEPA (2017)
2014	2014 fb_cb6r3_ae6nvPOA_aq	CMAQv5.2	WRFv3.8.1	USEPA (2018a)
2015	2015fd_cb6_15j (“base case”)	CMAQv5.2.1	WRFv3.8	This study
2015	2015fd_cb6_15j_noptf (“no.fire”)	CMAQv5.2.1	WRFv3.8	This study
2015	2015fd_cb6_15j_0cr (“no.crustal”)	CMAQv5.2.1	WRFv3.8	This study
2015	2015fe_cb6_15j (“nei.bc.nh3”)	CMAQv5.2.1	WRFv3.8	This study
2015	2015fd_cb6_15j_nobidi (“no.bidi”)	CMAQv5.2.1	WRFv3.8	This study
2015	2015fd_cb6_15j_pcSOA (“pc.soa”)	CMAQv5.2.1	WRFv3.8	This study