Atmos. Chem. Phys., 25, 3049–3107, 2025
https://doi.org/10.5194/acp-25-3049-2025
© Author(s) 2025. This work is distributed under
the Creative Commons Attribution 4.0 License.

Atmospheric Chemistry and Physics
Open Access
EGU

Research article

Critical load exceedances for North America and Europe using an ensemble of models and an investigation of causes of environmental impact estimate variability: an AQMEII4 study

Paul A. Makar[1], Philip Cheung[1], Christian Hogrefe[2], Ayodeji Akingunola[1], Ummugulsum Alyuz[3], Jesse O. Bash[2], Michael D. Bell[4], Roberto Bellasio[5], Roberto Bianconi[5], Tim Butler[6], Hazel Cathcart[1], Olivia E. Clifton[7,8], Alma Hodzic[9], Ioannis Kioutsioukis[10], Richard Kranenburg[11], Aurelia Lupascu[6,12], Jason A. Lynch[13], Kester Momoh[3], Juan L. Perez-Camanyo[14], Jonathan Pleim[2], Young-Hee Ryu[15], Roberto San Jose[14], Donna Schwede[2,☆], Thomas Scheuschner[16], Mark W. Shephard[1], Ranjeet S. Sokhi[3], and Stefano Galmarini[17]

[1]Environment and Climate Change Canada, Toronto, Canada
[2]Office of Research and Development (ORD), U.S. Environmental Protection Agency (EPA), Research Triangle Park, NC, USA
[3]Centre for Climate Change Research (C3R), University of Hertfordshire, Hatfield, UK
[4]Air Resources Division, National Park Service, Lakewood, CO, USA
[5]Enviroware srl, Concorezzo, Monza and Brianza, Italy
[6]Research Institute Sustainability – Helmholtz Centre Potsdam (RIFS Potsdam), Potsdam, Germany
[7]Goddard Institute for Space Studies, Earth Sciences Division, National Aeronautics and Space Administration, New York, NY, USA
[8]Center for Climate Systems Research, Columbia University, New York, NY, USA
[9]National Center for Atmospheric Research (NCAR), Boulder, CO, USA
[10]Department of Physics, University of Patras, Patras, Greece
[11]Netherlands Organisation for Applied Scientific Research (TNO), Utrecht, the Netherlands
[12]European Centre for Medium-Range Weather Forecasts (ECMWF), Bonn, Germany
[13]Office of Air and Radiation (OAR), U.S. Environmental Protection Agency (EPA), Washington, DC, USA
[14]Department of Computer Languages and Systems and Software Engineering, Polytechnic University of Madrid (UPM), Madrid, Spain
[15]Department of Atmospheric Sciences, Yonsei University, Seoul, South Korea
[16]Coordination Centre for Effects (CCE), Federal Environment Agency, Dessau, Germany
[17]Joint Research Centre (JRC), European Commission, Ispra, Italy
☆retired

Correspondence: Paul A. Makar (paul.makar@ec.gc.ca)

Received: 16 July 2024 – Discussion started: 30 July 2024
Revised: 19 December 2024 – Accepted: 14 January 2025 – Published: 14 March 2025

Abstract. Exceedances of critical loads for deposition of sulfur (S) and nitrogen (N) in different ecosystems were estimated using European and North American ensembles of air quality models, under the Air Quality Model Evaluation International Initiative Phase 4 (AQMEII4), to identify where the risk of ecosystem harm is expected to occur based on model deposition estimates. The ensembles were driven by common emissions and lateral boundary condition inputs. Model output was regridded to common North American and European 0.125° resolution domains, which were then used to calculate critical load exceedances. Targeted deposition diagnostics

Published by Copernicus Publications on behalf of the European Geosciences Union.

implemented in AQMEII4 allowed for an unprecedented level of post-simulation analysis to be carried out and facilitated the identification of specific causes of model-to-model variability in critical load exceedance estimates.

Datasets for North American critical loads for acidity for forest soil water and aquatic ecosystems were created for this analysis. These were combined with the ensemble deposition predictions to show a substantial decrease in the area and number of locations in exceedance between 2010 and 2016 (forest soils: 13.2 % to 6.1 %; aquatic ecosystems: 21.2 % to 11.4 %). All models agreed regarding the direction of the ensemble exceedance change between 2010 and 2016. The North American ensemble also predicted a decrease in both the severity and total area in exceedance between the years 2010 and 2016 for eutrophication-impacted ecosystems in the USA (sensitive epiphytic lichen: 81.5 % to 75.8 %). The exceedances for herbaceous-community richness also decreased between 2010 and 2016, from 13.9 % to 3.9 %. The uncertainty associated with the North American eutrophication results is high; there were sharp differences between the models in predictions of both total N deposition and the change in N deposition and hence in the predicted eutrophication exceedances between the 2 years. The European ensemble was used to predict relatively static exceedances of critical loads with respect to acidification (4.48 % to 4.32 % from 2009 to 2010), while eutrophication exceedance increased slightly (60.2 % to 62.2 %).

While most models showed the same changes in critical load exceedances as the ensemble between the 2 years, the spatial extent and magnitude of exceedances varied significantly between the models. The reasons for this variation were examined in detail by first ranking the relative contribution of different sources of sulfur and nitrogen deposition in terms of deposited mass and model-to-model variability in that deposited mass, followed by their analysis using AQMEII4 diagnostics, along with evaluation of the most recent literature.

All models in both the North American and European ensembles had net annual negative biases with respect to the observed wet deposition of sulfate, nitrate, and ammonium. Diagnostics and recent literature suggest that this bias may stem from insufficient cloud scavenging of aerosols and gases and may be improved through the incorporation of multiphase hydrometeor scavenging within the modelling frameworks. The inability of North American models to predict the timing of the seasonal peak in wet ammonium ion deposition (observed maximum was in April, while all models predicted a June maximum) may also relate to the need for multiphase hydrometeor scavenging (absence of snow scavenging in all models employed here). High variability in the relative importance of particulate sulfate, nitrate, and ammonium deposition fluxes between models was linked to the use of updated particle dry-deposition parameterizations in some models. However, recent literature and the further development of some of the models within the ensemble suggest these particulate biases may also be ameliorated via the incorporation of multiphase hydrometeor scavenging. Annual sulfur and nitrogen deposition prediction variability was linked to SO_2 and HNO_3 dry-deposition parameterizations, and diagnostic analysis showed that the cuticle and soil deposition pathways dominate the deposition mass flux of these species. Further work improving parameterizations for these deposition pathways should reduce variability in model acidifying-gas deposition estimates. The absence of base cation chemistry in some models was shown to be a major factor in positive biases in fine-mode particulate ammonium and particle nitrate concentrations. Models employing ammonia bidirectional fluxes had both the largest- and the smallest-magnitude biases, depending on the model and bidirectional flux algorithm employed. A careful analysis of bidirectional flux models suggests that those with poor NH_3 performance may underestimate the extent of NH_3 emission fluxes from forested areas.

Model–measurement fusion in the form of a simple bias correction was applied to the 2016 critical loads. This generally reduced variability between models. However, the bias correction exercise illustrated the need for observations which close the sulfur and nitrogen budgets in carrying out model–measurement fusion. Chemical transformations between different forms of sulfur and nitrogen in the atmosphere sometimes result in compensating biases in the resulting total sulfur and nitrogen deposition flux fields. If model–measurement fusion is only applied to some but not all of the fields contributing to the total deposition of sulfur or nitrogen, the corrections may result in greater variability between models or less accurate results for an ensemble of models, for those cases where an unobserved or unused observed component contributes significantly to predicted total deposition.

Based on these results, an increased process-research focus is therefore recommended for the following model processes and for observations which may assist in model evaluation and improvement: multiphase hydrometeor scavenging combined with updated particle dry-deposition, cuticle, and soil deposition pathway algorithms for acidifying gases, base cation chemistry and emissions, and NH_3 bidirectional fluxes. Comparisons with satellite observations suggest that oceanic NH_3 emission sources should be included in regional chemical transport models. The choice of a land use database employed within any given model was shown to significantly influence deposition totals in several instances, and employing a common land use database across chemical transport models and critical load calculations is recommended for future work.

1 Introduction

The concept of a critical load (CL) was first proposed as a means for evaluating the ecosystem impacts of the deposition of sulfur and nitrogen in response to the Convention on Long-Range Transboundary Air Pollution (CLRTAP), an international agreement on the mitigation and control of acidifying pollution, which entered into force in 1983 (CLRTAP, 2023). The convention provided some of the initial impetus for the development of comprehensive air quality models. The models provide a means of estimating the deposition fluxes of sulfur- and nitrogen-containing chemicals of anthropogenic origin, which may then be used to estimate the corresponding ecosystem impacts. Critical load exceedance estimates are the broadly accepted methodology for estimating the potential for ecosystem harm related to acidification and eutrophication. A critical load in this context was defined (Nilsson and Grennfelt, 1988) as "A quantitative estimate of an exposure to one or more pollutants below which significant harmful effects on specified sensitive elements of the environment do not occur, according to present knowledge". This definition is parsed in detail for readers unfamiliar with the critical load concept, in the Supplement.

The creation of critical loads for acidification and the calculation of their exceedances are based on the concept of the chemical charge balance steady state within soil water or aquatic ecosystems. The fluxes of anions and cations entering or leaving an ecosystem are used to determine whether an excess cation flux is available to the ecosystem, which could balance anion fluxes associated with acidifying deposition. Anion fluxes added to the system from anthropogenic sources include forms of deposited sulfur and nitrogen noted above. The S-containing forms of deposition (S_{dep}) are assumed to rapidly oxidize and are treated within critical load calculations as the sulfate ion. Every mole of deposited sulfur is assumed to be associated with two negative charges as the sulfate ion, SO_4^{2-}(aq) (aqueous); hence the deposition flux is tracked as charge equivalents per hectare per year, eq. $ha^{-1}\,yr^{-1}$. N-containing forms of deposition (N_{dep}) are assumed to rapidly oxidize and are treated as the nitrate ion – every mole of deposited nitrogen (including those of ammonia and ammonium) is assumed to be associated with one negative charge of nitrate ion deposition, NO_3^-(aq). Base cations and their deposition (Ca^{2+}, Mg^{2+}, K^+, and Na^+) are included in critical load calculations (collectively, BC_{dep}) and may incorporate anthropogenic base cation fluxes. The anthropogenic deposition fluxes in the ecosystem from the atmosphere are used in calculations of critical load exceedances. The critical loads themselves include estimates of natural atmospheric fluxes as well as other terms for fluxes of anions and cations. For example, in the steady-state or simple mass balance (SMB) model often used to define surface water critical loads for terrestrial ecosystems (Sverdrup and de Vries, 1994), BC_{dep} includes the release of soil base cations due to weathering, non-marine chloride deposition, the harvesting of base cations and/or nitrogen-containing biomass, denitrification, nitrogen immobilization in the rooting zone, the runoff volume, and a critical value of the non-sodium base cation to the aluminum ion ratio. Aquatic-ecosystem critical loads with respect to acidity are usually calculated using the steady-state water chemistry (SSWC) or the first-order acidity balance (FAB) methodologies (Henriksen and Posch, 2001; CLRTAP, 2023; de Vries et al., 2015) or other similar approaches (McDonnell et al., 2014). The SSWC model makes use of the difference between an estimate of the sea-salt-corrected pre-acidification concentration of base cations in the surface water and a specified biological indicator species' acid-neutralizing capacity limit above which no significant damage is expected to occur. The FAB methodology assumes the runoff fluxes at a lake outlet are charge-balanced and relates these runoff terms to fluxes of ions entering the lake and dimensionless retention factors and to terms for nitrogen immobilization, nitrogen growth uptake into vegetation, denitrification, atmospheric deposition, and weathering. An overview of the above methods for critical load (CL) estimation and how they are used in estimating exceedances may be found in CLRTAP (2023), Makar et al. (2018), and the references therein.

Critical loads of nutrient nitrogen and their exceedances are used to address the issue of the influx of airborne nitrogen resulting in changes in soil-based processes, plant growth, and inter-species relationships. Nitrogen-containing gases and aerosol components may be directly toxic to sensitive individual plant and animal species, while the accumulation of nitrogen (increased nitrogen availability) may also change species composition or relative abundance. Soil-mediated effects of acidification may include eutrophication, and species may have increased susceptibility to secondary stressors such as drought, frost, pathogens, or herbivores (CLRTAP, 2023). Critical loads of the eutrophication processes associated with nutrient nitrogen in terrestrial ecosystems may also make use of a version of the SMB model. This critical load model balances the input fluxes of all forms of nitrogen deposition plus biological fixation and soil nitrogen adsorption against ecosystem nitrogen losses (immobilization in soil organic matter, removal via harvesting of vegetation and animals, fluxes to the atmosphere (denitrification), erosion, combustion, ammonia volatilization, and leaching below the root zone). Biological fixation, soil adsorption, combustion, erosion, and ammonium leaching are usually considered negligible, and denitrification is assumed to be linearly dependent on the net input of nitrogen, leading to critical loads of nutrient nitrogen dependent only on immobilization, harvesting removal, the acceptable limit for nitrogen leaching of sensitive plant or animal species (nitrogen in soil water), and an ecosystem-dependent denitri-

fication fraction (CLRTAP, 2023). The acceptable limits for nitrogen concentrations in soil can range from 6.5 down to $0.2\,\mathrm{mg\,N\,L^{-1}}$, depending on the vegetation type (CLRTAP, 2023). A further means of estimating eutrophication is via comparison of measured nitrogen deposition with observed ecosystem damage over a large number of sites (Geiser et al., 2019; Simkin et al., 2016). Exceedances for eutrophication in this case may be estimated as the differences between the estimated nitrogen deposition and the observation-based critical load.

As noted in the Supplement, critical load exceedance calculations are carried out on an ongoing basis due to the ongoing cycle of chemical transport model (CTM) process improvement. The results of our analyses should thus be considered a "snapshot" of the state of both CTM science and critical load (CL) knowledge at the time the simulations and critical load data collection took place (2021). CTMs numerically integrate the system of time-dependent differential equations describing the rates of change in chemical species in the atmosphere in order to predict the changes in chemical concentrations and deposition over time. This is usually done by breaking the net differential equation for the rates of change into component processes (e.g. advection, diffusion, gas-phase chemistry, inorganic particle chemistry, dry deposition, particle microphysics treating the nucleation, condensation of gases, coagulation of particles, cloud processing of gases and aerosols including wet deposition), with the processes being solved in sequence to determine the future state of the atmosphere (Marchuk, 1990). However, there is usually not a complete scientific consensus on the best numerical methods to carry out the time stepping for each of these processes, and the level of detail in process representation in the models may also vary considerably, depending at times on external constraints such as the processing time available for CTM simulations. The individual processes are usually evaluated based on laboratory or other process-specific data wherever possible, but the selection of a specific process representation within a CTM is often based on comparisons of the output of an entire CTM relative to surface or satellite monitoring data. This latter approach may allow for the compensation errors in the process representation to take place (cf. Makar et al., 2014; Hyder et al., 2018; Huang et al., 2021; Vizuete et al., 2022). These considerations may contribute to the resulting variability in deposition estimates from the different modelling frameworks. The work conducted here uses analysis of new model diagnostic outputs added for AQMEII4 to attempt to determine the key causes of these model deposition estimate differences.

The ongoing reevaluation and improvement of CTMs is aided by ensemble model comparisons, where models driven by the same lateral boundary and emission inputs are cross-compared and evaluated against observations. The Air Quality Model Evaluation International Initiative (AQMEII) has comprised model CTM ensemble evaluation studies, to date in four phases. The initial phase of AQMEII utilized largely offline regional models used for research and public policy support to simulate a common year, 2006, with common emission inputs, in both North America (NA) and Europe (EU), with 22 modelling groups participating (Galmarini et al., 2012). Subsequent phases of AQMEII examined specific issues within the CTM community: AQMEII2 had as its focus the evaluation of both weather and air quality predictions for fully coupled, online air quality models, where the particulate matter generated by the models on any given time step feeds back into the coupled models' weather forecast radiative transfer and cloud formation processes (Galmarini et al., 2015). AQMEII3 addressed questions of hemispheric transport of air pollutants – the relative contributions of local versus long-range transport to predicted pollutant concentrations, and their impacts on ecosystem and human health (Galmarini et al., 2017).

The variety in underlying scientific theory encapsulated within CTMs and their process representation implies the need for cross-comparison of critical load exceedance predictions from a variety of models. As part of AQMEII3, 14 air quality models were used to calculate oxidized sulfur and oxidized and reduced nitrogen deposition, and hence critical load exceedances in Europe (Vivanco et al., 2018). This comparison revealed a high degree of variability in simulated wet- and dry-deposition fluxes. The models with the best performance relative to observations were used to provide ensemble critical loads – a "reduced ensemble" in that not all models submitting output for the study were used in generating ensemble critical loads. However, even within this reduced ensemble, local variations of over a factor of 4 in both sulfur and nitrogen deposition could be seen between the ensemble members, and the predicted percent area in exceedance for sensitive ecosystems varied by more than a factor of 2 for the best performing models (Vivanco et al., 2018). These results highlighted the large range of model-dependent variability possible in critical load exceedance estimates – but the causes for that variability, and how it might be reduced, were not investigated to any significant extent.

The study protocols of AQMEII Phase 4 (AQMEII4) were designed partly in response to the large variation in model sulfur and nitrogen deposition estimates noted in Vivanco et al. (2018), Solazzo et al. (2018) and Hogrefe et al. (2020). AQMEII4 protocols were also motivated by a similarly large variation in simulated ozone deposition velocities (Hardacre et al., 2015; Wu et al., 2018), and renewed emphasis on the importance of specific ozone deposition pathways (Clifton et al., 2017, 2020a, b).

AQMEII4 has two main activities: a regional model intercomparison with enhanced diagnostics for gas-phase dry deposition (Galmarini et al., 2021) and an observation-driven single-point model intercomparison study for ozone dry deposition at sites with ozone flux records (Clifton et al., 2023). The current work continues the regional model intercomparison driven by common boundary conditions, with a focus here on critical load exceedances for acidity and eutrophi-

cation, and the use of additional diagnostics to determine the underlying causes for the model-to-model variability in these exceedance estimates.

As described later in our analysis, two processes account for much of the variability in CTM predictions of the total deposition of sulfur and nitrogen (S_{dep} and N_{dep}): particle dry deposition and the scavenging of particles by depositing hydrometeors. We note that following the construction and application of the model versions applied in AQMEII4, new parameterizations for particle dry deposition became available. Emerson et al. (2020) compiled multiple particle dry-deposition velocity observations and compared these to the predictions of the commonly used Zhang et al. (2001) algorithm. Relative to these observations, the Zhang et al. (2001) algorithm tended to overestimate deposition velocity on vegetated surfaces at smaller particle sizes (<0.4 µm diameter), while it underestimated the deposition velocity for particles between 1 and 10 µm. Several papers prior to 2019 noted that the relationship between particle size and deposition velocity did not "capture observed relationships between particle deposition velocities and particle size, especially around the accumulation mode" (Clifton et al., 2024). Emerson et al. (2020) also noted a substantial overestimate of the Zhang et al. (2001) particle deposition velocity over water surfaces relative to observations. Emerson et al. (2020) proposed a modified version of the Zhang et al. (2001) algorithm, demonstrating a better fit to the ensemble of deposition velocity observations. The differences between the two parameterizations were substantial, with decreases in particle deposition velocities in the submicrometre range of 1 to 2 orders of magnitude relative to Zhang et al. (2001) across multiple land use types and increases over vegetated surfaces of up to an order of magnitude for particle diameters from 1 to 10 µm. The decrease in submicrometre deposition velocities might be expected to result in increases in air concentrations of Aitken to mid-accumulation mode particles and decreases in those of mid-accumulation mode to coarse-mode particles. Ryu and Min (2022) applied the Emerson et al. (2020) parameterization to the WRF-Chem model and found that $PM_{2.5}$ positive biases increased in magnitude, while PM_{10} negative biases were partially offset with the use of the new algorithm. Pleim et al. (2022) also re-examined aerosol dry-deposition velocities in the context of the CMAQ model, noting an increase in accumulation mode dry-deposition velocities of almost an order of magnitude in forested areas, an overall reduction in $PM_{2.5}$ concentrations, and an improvement in $PM_{2.5}$ prediction accuracy. The latter work does not necessarily contradict the Emerson et al. (2020) results, which imply possible increases in PM mass within the Aitken and accumulation modes. The increase in the removal of mass between the mid-accumulation mode to larger sizes may dominate over the particle deposition velocity decreases between the Aitken to mid-accumulation mode noted in the observations collected by Emerson et al. (2020).

Studies using sectional aerosol size representations have recently found that improved aerosol deposition velocity algorithms need to be combined with improved wet hydrometeor scavenging to result in net improvements of regional model performance. Ryu and Min (2022) found that the best overall WRF-Chem performance resulted from a combination of updates (when the new dry-deposition algorithm was combined with updates for cloud scavenging employing cloud fractions for rainout and a revised parameterization for below-cloud scavenging incorporating separate terms for rain and snow removal rates). Ghahreman et al. (2024), in updating the cloud scavenging parameterization of the GEM-MACH model, noted differences in rain and snow below-cloud scavenging rates of up to 2 orders of magnitude between the previously applied, temperature-based parameterization of Slinn (1984) and the newly implemented parameterization of multiphase scavenging (from both the underlying meteorological model and the empirical scavenging parameterization of Wang et al., 2014). Differences in scavenging rates were found to be strongly dependent on temperature, aerosol size, and the precipitation rate. The revised parameterizations resulted in an overall improvement in performance for wet SO_4^{2-} deposition, where the Emerson et al. (2020) algorithm was employed for the particle dry-deposition simulation in all the model runs.

A large part of the model-to-model variability and uncertainty resides in the two processes above, as demonstrated in our analysis. We next describe our methodology (including an overview of the two AQMEII4 model domains; descriptions of the construction of the critical load data employed herein; and descriptions of the models, their inputs, and boundary conditions). Our analysis follows, first presenting estimates of critical load exceedances for two different simulation years in each domain and then the exceedances estimated using ensembles of model deposition predictions. The bulk of the analysis then examines individual contributions of different sulfur and nitrogen species to their total deposition for each model and for the ensemble. The causes of the differences between the models are determined through process analysis. Our concluding section includes research recommendations based on the analysis in order to improve the performance of individual models and to reduce the variability between their estimates of critical load exceedances.

2 Methodology

2.1 Critical load data

Six critical load (CL) datasets were used in conjunction with our ensembles of CTM deposition estimates. North American CL datasets included terrestrial-ecosystem (forest-ecosystem) acidity critical loads for the continent, aquatic-ecosystem acidity critical loads combining data from Canada and the USA, and USA-specific eutrophication critical loads for sensitive-epiphytic-lichen species and herbaceous-

plant species. European CL datasets combined CL information from multiple countries for terrestrial- and aquatic-ecosystem acidity and terrestrial-ecosystem eutrophication. A brief summary of the six CL datasets used in this work is provided here; full descriptions of the methodology used to create the CL data are provided in Sect. S1.0 in the Supplement.

North American terrestrial ecosystem CL estimates were generated using the simple mass balance model (Sverdrup and Warfvinge, 1990; Sverdrup and de Vries, 1994), employing data from several studies within the USA and Canada (McNulty et al., 2007, 2013; Duarte et al., 2013; Phelan et al., 2014, 2016; Sullivan et al., 2011; Sullivan et al., 2012; Cathcart et al., 2025). Table S1 (Supplement) provides methodological information for these studies, such as the horizontal spatial resolution, the dataset extent, plant-species-specific ratio values of the critical base cation to aluminum soil water, the approaches used to estimate soil base cation weather rates, losses of (non-sodium) base cations from the ecosystem through uptake via harvesting or grazing, and whether nitrogen uptake via harvesting/grazing was included in the calculation of nitrogen minimum critical loads.

The North American aquatic-ecosystem acidity critical load dataset constructed here combined individual datasets from Canada and the USA, as follows.

Environment and Climate Change Canada (ECCC) data corresponding to the subset of 2997 lake surveys which reside within the North American AQMEII4 common grid were used in conjunction with the Steady-State Water Chemistry (SSWC) critical load model (Sverdrup et al., 1990) as described in Aherne and Jeffries (2015). SSWC is in widespread use for the aquatic-ecosystem CL (Posch et al., 2001; Cathcart et al., 2016; Henriksen et al., 2002; Jeffries et al., 2010; Scott et al., 2010; Whitfield et al., 2006; Williston et al., 2016; Dupont et al., 2005; Miller, 2012). CL calculations for Canada followed a hierarchy based on the available information for individual lakes; for example catchment runoff rates were determined by isotope mass balance estimates in preference to a GIS map-based approach using regional datasets, and when dissolved organic carbon estimates were available, an organic-acid-adjusted limiting value of the acid-neutralizing capacity was used to include the influence of organic acids in the lake in preference to a fixed value of $40\,\mu eq.\,L^{-1}$. Only sulfur deposition was used to determine exceedance, since the SSWC model does not consider non-acidifying nitrogen.

Aquatic-ecosystem critical loads for the USA were taken from the National Critical Loads Database version 3.2.1 (NCLDv3.2.1; Lynch et al., 2022), which contains both the critical load data used here and supporting information. A total of 21 667 critical loads were used for 14 334 unique lakes and streams across the USA (a combination of different methods for determining the critical loads were included in the US values, sometimes resulting in more than one CL estimate for the same waterbody). Most US aquatic critical loads (78 %) were determined using the SSWC model (Lynch et al., 2022; Scheffe et al., 2014; Dupont et al., 2005; Miller 2012; VDEC, 2003, 2004, 2012) and site-specific catchment runoff rates (US EPA, 2023). The remaining 22 % of US aquatic critical loads were determined by a dynamic modelling approach (Sullivan et al., 2005; Fakhraei et al., 2014; Lawrence et al., 2015) and a combination of dynamic modelling with a regionalization approach (McDonnell et al., 2012, 2014; Sullivan et al., 2012; and McDonnell et al., 2021). Organic-acid-adjusted values for limiting acid-neutralizing capacity were not used in generating these US aquatic CL with respect to acidity datasets, and an average critical load value was used for these waterbodies for which overlapping CL estimates were available. A more detailed description of the US aquatic critical loads used here can be found in Lynch et al. (2022).

North American critical loads for eutrophication were estimated using critical load exceedance (CLE) for two ecosystem types, sensitive-epiphytic-lichen species and herbaceous-species richness.

The CL for sensitive-epiphytic-lichen species richness made use of 9000 community surveys across the USA from 1990–2012 (Geiser et al., 2019), where a 90 % quantile regression was used to model relationships between deposition levels and observed species richness in order to estimate critical loads, and a −20 % decline in species richness was used to determine the critical load. These methods resulted in a single critical load of $3.1\,kg\,N\,ha^{-1}\,yr^{-1}$ for sensitive epiphytic lichen, which was applied to all broadleaf, conifer, or mixed-forest land cover types.

The CL for US herbaceous-species richness made use of data developed using over 14 000 vegetation survey plots across nitrogen deposition gradients (Simkin et al., 2016). An observation-based approach using median quantile regressions for the herbaceous-species richness response to deposition was employed to generate critical loads with respect to nitrogen deposition linked to various atmospheric and soil conditions. Separate CL models were developed for open and closed canopies. The resulting CL of N for open canopy systems ranged from 6.2 to $12.3\,kg\,N\,ha^{-1}\,yr^{-1}$, and the CLs of N for closed canopy systems ranged from 6.1 to $23.7\,kg\,N\,ha^{-1}\,yr^{-1}$.

Two EU CL datasets were employed for the EU AQMEII4 domain, for acidification and eutrophication of terrestrial ecosystems, respectively. The critical load database and the exceedance calculations for Europe were provided by the Coordination Centre for Effects (CCE) under the United Nations Economic Commission for Europe Convention on Long-range Transboundary Air Pollution (UNECE LRTAP convention), hosted by the Umweltbundesamt (UBA) in Germany, which develops and maintains the European critical loads database (Geupel et al., 2022). The most recent database available was used here, and while dependent on the country, all CL estimates made use of the simple mass balance model (Sverdrup and de Vries, 1994; CLRTAP, 2023;

Geupel et al., 2022), with gap filling using the CCE background database (Reinds et al., 2021). Critical loads for EU eutrophication ($CL_{nut}N$) were also based on the SMB method applied to nitrogen deposition, and two different methodologies were used to determine the accepted nitrogen leaching. Dependent on the country, empirical values were sometimes used as upper and lower boundaries for the SMB modelling results in order to avoid rather extreme results in ecosystems where the SMB model predicts very high or very low eutrophication CL values (Bobbink et al., 2022). The resulting EU CLEs were summarized as the share of the receptor area with critical load exceedance (bar charts) and the magnitude of the exceedance within each analysis grid cell (maps). The exceedance in a grid cell is defined as the so-called "average accumulated exceedance" (AAE), which is calculated as the area-weighted average of the exceedances of the critical loads of all ecosystems in this grid cell.

2.2 AQMEII4 overview description

The setup of the AQMEII4 regional model comparison is described in detail in Galmarini et al. (2021); a brief overview is provided here. The models within this analysis are a snapshot of the development of regional chemical transport models as of the time simulations were completed (2021).

Model simulations were carried out for the years 2010 and 2016 for North America and 2009 and 2010 for the European region. North American years were chosen due to policy relevance, with a significant change in SO_2 emission controls enacted between the 2 years. The European years were chosen due to a large difference in meteorology between the years 2009 and 2010, the latter being a year with unusually high summer temperatures in eastern Europe and on the western side of the Russian Federation (Barriopedro et al., 2011) leading to increased European forest fire activity and emissions (Schmuck et al., 2012). The July 2009 and July 2010 temperature and precipitation anomalies relative to the base year period of 1961 to 1990 are shown in Supplement Fig. S2 (NCDC, 2024)). The precipitation anomalies in July of each year are less significantly different than the temperature anomalies; similarly, the differences between the annual average temperature and precipitation anomalies between the 2 years are less significant than the July values. In the analysis which follows, the differences in the simulated deposition and critical load exceedances for the European region between the 2 years is shown to be relatively minor, implying that forest fire emissions contributed a relatively small proportion of sulfur and nitrogen deposition in 2010 and that the summer temperature anomalies in 2010 did not result in significant perturbations to total sulfur and nitrogen deposition.

Simulations were carried out by making use of the individual models' grid projection and resolution. Mass-conserving interpolation (for concentrations and fluxes) and nearest-neighbour interpolation (for diagnostics) were then used to map these "native-grid" outputs to corresponding North American and European AQMEII4 grids. The latter have $0.125° \times 0.125°$ resolution (North America: 23.5 to 58.5° N and 130 to 59.5° W; Europe: 25 to 70° N and 30° W to 60° E). Values extracted from the AQMEII4 grid locations were used for comparison to observations. Models made use of their own meteorological drivers or online meteorological components for meteorological field predictions. Models shared common inputs for emissions and chemical lateral boundary conditions. The latter provide a uniform chemical forcing and prevent input variations not associated with the models themselves from influencing simulation results.

North American anthropogenic emissions were generated using emission modelling platforms which included the anthropogenic inventories, temporal and spatial allocation from the county or state/province level to native model grids for each of the 2 model years, and adjustments for specific inventories by year. Emission processing was carried out by the United States Environmental Protection Agency (EPA) for the Carbon Bond 6 (revision 3; CB6r3) and Statewide Air Pollution Research Center 07 (SAPRC07) chemical mechanisms (Yarwood et al., 2010; Carter, 2010) and by Environment and Climate Change Canada for the Acid Deposition and Oxidant Model version II (ADOM-II; Stockwell and Lurmann, 1989). Note that while none of the modelling groups made use of the SAPRC07 mechanism itself within their simulations, this mechanism was sometimes used as a starting point for lumping individual models' volatile organic compound (VOC) species, due to the greater level of detail available within the SAPRC07 speciation. European anthropogenic emissions were prepared for the participating models' chemical mechanisms by the Netherlands Organization for Applied Scientific Research (TNO) as part of the Monitoring Atmospheric Composition and Climate, part 3 (MACC-III), project (Kuenen et al., 2014), with individual groups using their own emission data for the portion of their native model grids extending beyond the range of MACC-III emission grid if necessary.

North American forest fire emissions were generated by combining the US emission modelling platform values with Canadian data for 2010, while both US and Canadian data were based on the 2016 emission modelling platform estimates. These forest fire emissions included the criteria of air contaminant emission mass, heat flux, and acres burned. Fire plume rise calculations were carried out by individual modelling groups, typically based on large stack plume rise formulae (Briggs, 1984). European forest fire emissions were provided by the Finnish Meteorological Institute using eight layers from 50 to 6200 m. Both North American and European forest fire emissions were chemically disaggregated by the participating modelling groups and mapped onto the nearest grid cell with respect to their native model grids.

Lightning NO emissions were also prescribed in both domains, based on GEIA monthly climatology values (Price et al., 1997), diurnally disaggregated following Blakeslee et

al. (2014) and allocated vertically following Ott et al. (2010) by individual modelling groups.

Chemical lateral boundary conditions for simulations in both Europe and North America were taken from 3-hourly, 0.75° × 0.75°, 54-vertical-level ECMWF CAMS EAC4 reanalysis products (Inness et al., 2019), interpolated by participants to their own vertical and horizontal grid structures and chemically disaggregated to their own chemical speciation.

2.3 Common model diagnostics

The AQMEII4 protocol for ensemble participants included the reporting of gas-phase species' aerodynamic, bulk surface, stomatal, mesophyll, quasi-laminar sub-layer and within-canopy buoyant resistances (when present in the reporting model). Effective conductances (Paulot et al., 2018; Clifton et al., 2020a, b) and effective fluxes (Galmarini et al., 2021) were also reported. These latter two diagnostic terms provide the relative contribution of the four main pathways associated with gas-phase deposition to the deposition velocity and the deposition flux, respectively. The four main pathways include soil, the lower canopy, leaf cuticles, and stomata. Note that not all models specify a separate lower-canopy pathway (the conductance associated with this pathway tends to be relatively small, providing justification for its absence). Effective fluxes are of particular interest to critical load exceedance analysis, since they provide information on the charge equivalents deposited to different component surface types. Effective fluxes include the impact of other processes in addition to deposition on the concentrations and hence on the net flux of the deposited gases, via the net flux term (F). For example, the soil, lower canopy, cuticle, and stomatal effective fluxes in the Wesely (1989) dry-deposition parameterization are given by

$$\mathrm{DFLX_{SOIL}} = \left(\frac{(r_{ac}+r_{gs})^{-1}}{(r_s+r_m)^{-1}+(r_{lu})^{-1}+(r_{dc}+r_{cl})^{-1}+(r_{ac}+r_{gs})^{-1}} \right) F, \quad (1)$$

$$\mathrm{DFLX_{LCAN}} = \left(\frac{(r_{dc}+r_{cl})^{-1}}{(r_s+r_m)^{-1}+(r_{lu})^{-1}+(r_{dc}+r_{cl})^{-1}+(r_{ac}+r_{gs})^{-1}} \right) F, \quad (2)$$

$$\mathrm{DFLX_{CUT}} = \left(\frac{(r_{lu})^{-1}}{(r_s+r_m)^{-1}+(r_{lu})^{-1}+(r_{dc}+r_{cl})^{-1}+(r_{ac}+r_{gs})^{-1}} \right) F, \quad (3)$$

$$\mathrm{DFLX_{stom}} = \left(\frac{(r_s+r_m)^{-1}}{(r_s+r_m)^{-1}+(r_{lu})^{-1}+(r_{dc}+r_{cl})^{-1}+(r_{ac}+r_{gs})^{-1}} \right) F, \quad (4)$$

where F is the net flux to the surface and the r terms are resistances associated with different pathways of gas mass transfer to the four surface components (r_{ac}: aerodynamic mass transfer within the canopy, dependent on canopy height and density; r_{gs}: the soil and leaf litter resistance; r_{dc}: canopy buoyant convection resistance; r_{cl}: resistance associated with leaves, twigs, bark, and other exposed surface in the lower canopy; r_{lu}: resistance of leaf cuticles in healthy vegetation and other outer surfaces; r_s: leaf stomata; r_m: leaf mesophyll). The effective conductances can be generated from similar formulae, with the F term in Eqs. (1) through (4) being replaced by the deposition velocity of the gas V_d. Note that the formulae for individual models vary from the Wesely (1989) example shown above; see Galmarini et al. (2021) for details of the formulae for each of the gas-phase deposition algorithms used in the AQMEII4 regional model ensembles analyzed here.

2.4 Model parameterization descriptions

The models CMAQ-M3Dry, CMAQ-STAGE, WRF-Chem (IASS, Institute for Advanced Sustainability Studies), GEM-MACH (Base), GEM-MACH (Zhang), GEM-MACH (Ops, operational forecast), WRF-Chem (UPM, Technical University of Madrid), and WRF-Chem (UCAR, University Corporation for Atmospheric Research) provided simulations for AQMEII4, interpolated to the common the North American domain. The models WRF-Chem (IASS), LOTOS-EUROS (TNO), WRF-Chem (UPM), and CMAQ (Hertfordshire) provided simulations for AQMEII4, interpolated to the common European domain. Some of the modelling frameworks were repeated, but process implementation details were varied in order to examine the relative impact of these differences. We describe each of these models according to the starting framework (CMAQ, GEM-MACH, WRF-Chem, LOTOS-EUROS) below.

2.4.1 CMAQ-M3Dry, CMAQ-STAGE, CMAQ (Hertfordshire): WRF–CMAQ implementations

These three models make use of the WRF–CMAQ offline modelling framework (CMAQ v5.3.2; US EPA, 2020), with the North American implementations (CMAQ-M3Dry, CMAQ-STAGE) employing 12 km cell resolution and the EU implementation employing 10 km cell resolution (Lambert conformal conic projection, 459 × 299 and 500 × 681 grid cells, respectively). The CMAQ implementations employed 35 model layers with the lowest-layer thickness of ∼ 20 m. Both NA models operate in an offline configuration using the same driving weather forecast model output (NA: WRF4.1.1, EU: WRF 4.2.1; Skamarock et al., 2019). All three CMAQ model implementations use the same gas-phase chemical mechanism (Carbon Bond 6; Luecken et al., 2019), a modal aerosol size distribution representation with three modes (Binkowski and Roselle, 2003), aerosol microphysics through the AERO7 module (Appel et al., 2021; Binkowski and Shankar, 1995; Vehkamaki et al., 2002), and thermodynamic equilibrium partitioning for semivolatile inorganic species between gas- and aerosol-phase species (involving the components K^+–Ca^{2+}–Mg^{2+}–NH_4^+–Na^+–SO_4^{2-}–NO_3^-–Cl^-–H_2O) using the ISORROPIA II algorithm (Fountoukis and Nenes, 2007). Organic aerosol formation and monoterpene oxidation are modelled as described in AERO7 (Appel et al., 2021; Xu et al., 2018).

For all three model implementations, the impact scavenging of aerosols by cloud droplets is carried out for the Aitken mode particles, while accumulation mode and coarse-mode particles may form cloud condensation nuclei, resulting in their scavenging via cloud droplet nucleation (Binkowski and Roselle, 2003; Chaumerliac, 1984; Fahey et al., 2017). Aerosol scavenging in the Aitken mode is carried out as a simple exponential decay for the number, surface area, and mass concentration assuming a cloud droplet settling velocity based on Pruppacher and Klett (1978) and an assumed cloud droplet size distribution. Only Aitken mode particles (roughly 0.01 to 0.1 µm diameter) are impact-scavenged, for which only cloud liquid water is included as a scavenging hydrometeor. The wet deposition of all aqueous species is represented as a first-order loss rate based on the precipitation rate and total liquid water content (Fahey et al., 2017). The number of cloud droplets is parameterized following Bower and Choularton (1992) from the cloud liquid water content provided by the meteorological model.

The three CMAQ implementations differ in the algorithms employed for aerosol- and gas-phase dry-deposition algorithms.

The aerosol dry-deposition methodology of CMAQ-M3Dry was based on Binkowski and Shankar (1995), with updates as described in Venkatram and Pleim (1999), Giorgi (1986), and subsequent corrections to include the effect of mode width in the Stokes number (reducing previous large overpredictions in coarse-mode deposition velocities). Further modifications included changes to the Stokes number for vegetated surfaces, a modification of the impaction term, the scaling of diffusion layer resistance by the leaf area index (LAI) for the vegetated fraction of each grid cell, and improved mass conservation for the process of gravitational settling (Appel et al., 2021).

The aerosol dry-deposition methodology of CMAQ-STAGE and CMAQ (Hertfordshire) followed that of CMAQ-M3Dry but made use of Slinn (1982) and Zhang et al. (2001) for impaction on vegetated surfaces and Giorgi (1986) for water and soil surfaces, with the resulting deposition velocities for smooth and vegetated surfaces weighted by the area of vegetated surface (Appel et al., 2021).

The gas-phase dry-deposition algorithms and diagnostic equations of CMAQ-M3Dry, CMAQ-STAGE, and CMAQ (Hertfordshire) are described in detail elsewhere (Table B2 of Galmarini et al., 2021; other implementation details in Hogrefe et al., 2023). The algorithms follow the original approach of Wesely (1989) but with separate resistance branches for the vegetated and non-vegetated fractions, dry versus wet fractions, and snow-covered versus non-snow-covered fractions.

Bidirectional fluxes of ammonia were found in the analysis which follows to be a major source of model-to-model variability and hence will be described here in more detail.

CMAQ-M3Dry simulated bidirectional fluxes of ammonia by first calculating soil ammonia concentrations using the Environmental Policy Integrated Climate (EPIC) agricultural ecosystem model (Williams, 1995; Ran et al., 2018) prior to the CTM simulations being carried out. Typically, the EPIC model simulation requires a model spin-up period of 25 years or more and requires a prior simulation of N deposition as input information. The soil NH_3 concentrations from this coupled system were then used as inputs for the AQMEII4 run (Pleim et al., 2019). While all dry-deposition diagnostics reported to AQMEII4 for CMAQ-M3Dry were computed making use of a post-processor, the post-processing did not include the generation of bidirectional flux calculations, and hence diagnostics such as the net compensation point concentration and the ground compensation point calculation were not provided from CMAQ-M3Dry for AQMEII4.

CMAQ-STAGE (Massad et al., 2010; Bash et al., 2013) also simulated bidirectional fluxes following Williams (1995), using a previous coupled EPIC simulation only for initial conditions, porting the methodology and information on daily fertilization and nitrification from EPIC into the CMAQ-STAGE framework while estimating evasion and deposition locally within the chemical transport model. This methodology, which operates on a land-use-specific basis and then aggregates to a grid cell basis, allowed for an additional AQMEII4 diagnostic to be incorporated into the CMAQ-STAGE simulations. This allows for greater consistency between the CTM and the resulting soil NH_3 calculations (and allows for the output of all of the diagnostics as specified under the AQMEII4 protocol; see Hogrefe et al., 2023). However, these calculations do not include other terms in EPIC dealing with N fixation, mineralization, denitrification, runoff, percolation, and plant uptake and hence will diverge from the EPIC-simulated soil ammonia concentrations due to the differences in evasion and deposition parameterizations between CMAQ-STAGE and EPIC.

2.4.2 NA WRF-Chem (IASS), EU WRF-Chem (IASS); NA WRF-Chem (UPM), EU WRF-Chem (UPM); NA WRF-Chem (UCAR): WRF-Chem implementations

All three of these models made use of the WRF-Chem chemical transport modelling framework (Grell et al., 2005), employing a 12 km Lambert conformal conic projection (400×360 grid cells in the European domain, 480×290 grid cells in the North American domain), two-way coupling between air quality and meteorology, a sectional aerosol size distribution representation (four bins), aerosol microphysics and chemistry via the MOSAIC model (Zaveri et al., 2008), organic aerosol formation following Knote et al. (2014, 2015), cloud microphysics following Morrison et al. (2009), the Noah-Multiparameterization (Noah-MP) land surface model (Niu et al., 2011), the Rapid Radiative Transfer Model (RRTM) for radiative transfer calculations (Iacono et al., 2008), biogenic emissions using the MEGAN model

(Guenther et al., 2006; Wiedinmyer et al., 2007), and the Fast-J algorithm for photolysis rate calculation (Chapman et al., 2009). All three code versions also make use of the Wesely (1989) parameterization for gas dry deposition and the Binkowski and Shankar (1995) approach for aerosol deposition. However, WRF-Chem has a large variety of configurations available for other model processes, allowing for the impact of those configurations on deposition results to be studied under AQMEII4. The differences between the model configurations are summarized in Table 1. It should also be noted that WRF-Chem is an online modelling framework; differences in the model parameterizations can influence the meteorological predictions through the aerosol direct and indirect effects, and consequently the meteorology generated by the implementations may also differ.

Not all of the WRF-Chem model implementations were able to report all of the information required to calculate exceedances: the WRF-Chem (IASS) implementation did not report all of the species contributing to S_{dep} and N_{dep} totals and also did not report several diagnostics requested under the AQMEII4 protocol. Consequently, the WRF-Chem (IASS) results were not included in ensemble deposition generation, and the model ensembles are referred to hereafter as "reduced ensembles". Our analysis is therefore based on these reduced ensembles, though WRF-Chem (IASS) values for deposition totals have been provided when available in figures and tables for comparison purposes.

2.4.3 LOTOS-EUROS (TNO): LOTOS-EUROS

LOTOS-EUROS (TNO) used in the AQMEII4 EU simulations is an open-source 3D chemistry transport model used extensively for air quality forecasts and scenarios for European domains (Timmermans et al., 2022; Manders et al., 2017). Gas dry-deposition fluxes made use of the approach based on Wesely (1989) (DEPosition of Acidifying Compounds, DEPAC; Van Zanten et al., 2010). Particle dry deposition was carried out using the approach of Zhang et al. (2001). Wet deposition followed the droplet saturation approach, and cloud chemistry with sulfate formation was dependent on cloud liquid water and droplet pH (Banzhaf et al., 2012). The dry deposition of ammonia makes use of a bidirectional flux approach (Wichink Kruit et al., 2012). Gas-phase chemistry was carried out using a modified form of the CBM-IV scheme (Gery et al., 1989; Whitten et al., 1980). N_2O_5 hydrolysis was included following Schaap et al. (2004), and inorganic thermodynamic particle chemistry was solved using the ISORROPIA II module (Fountoukis and Nenes, 2007). The model operated using 12 layers in the vertical in a hybrid coordinate system, with the near-surface layer having a thickness of ~ 20 m and a model top of approximately 8 km. The simulations carried out here made use of a 20×20 km grid cell size over Europe. Driving meteorology for the model was from 3-hourly ECMWF short-term forecasts. Land use data for the model comes from the CORINE Land Cover 2000 database (EEA, 2000, 2007).

2.4.4 GEM-MACH (Base), GEM-MACH (Zhang), GEM-MACH (Ops): GEM-MACH

All three of these NA models are variations on the Environment and Climate Change Canada GEM-MACH model. The first two configurations (GEM-MACH (Base), GEM-MACH (Zhang)) are based on the "research" version of the model, which has more detailed physical parameterizations, whereas GEM-MACH (Ops) is based on the "operational forecast" configuration, where more simplified parameterizations have been employed in order to reduce processing time for operational air quality forecast simulations. Common elements across all three implementations include a horizontal grid cell size of 0.09° in a rotated latitude–longitude domain (~ 10 km), 83 model levels, biogenic VOCs from BEIS versions 3.09 and 3.13 (Vukovich and Pierce, 2002; Stroud et al., 2010), a sectional aerosol size distribution (12 bins; Gong et al., 2003), the ADOM-II gas-phase mechanism (Stockwell and Lurmann, 1989), a modified Odum approach for secondary organic aerosol (SOA) formation (Stroud et al., 2018), and an inorganic aerosol chemistry module solving the thermodynamic equilibrium for the SO_4^{2-}–NO_3^-–NH_4^+–H_2O system (Makar et al., 2003). The GEM-MACH implementations also all make use of the GEM weather forecast model v4.9.8 for driving meteorology (Côté et al., 1998; Girard et al., 2014), with the ISBA land surface scheme (Belair et al., 2003a, b) and the Canadian Climate Center for Modeling and Analysis Radiative transfer algorithm 2 (CCCMA Rad2; Li and Barker, 2005). As was the case for the WRF-Chem implementations described above, GEM-MACH has several optional process representations used in operational forecast versus research versions of the model; hence the relative importance of model configurations versus deposition parameterizations may be studied. The differences between the configurations are summarized in Table 2.

Collectively, the differences between GEM-MACH (Base) and GEM-MACH (Zhang) provide an estimate of the relative importance of the gas-phase deposition parameterization for simulation results, while comparisons between GEM-MACH (Base or Zhang) and GEM-MACH (Ops) show the relative impact of the combination of ammonia bidirectional fluxes and the suite of more complex physical parameterizations used in the former model configurations compared to the operational framework.

2.5 Bias-corrected critical load exceedance estimates

As will be discussed in Sect. 3.2, model results were evaluated using the available data for North America and Europe (see Sect. S7.0 in the Supplement for species contributing significantly to total S and N deposition). Critical load exceedances were calculated, making use of the total sulfur

Table 1. AQMEII4 WRF-Chem configuration differences. PBL: planetary boundary layer.

Parameterization	WRF-Chem (IASS)	WRF-Chem (UPM)	WRF-Chem (UCAR)
WRF-Chem version number	3.9.1	4.0.3	4.1.2
Wet deposition	Convective: Grell and Devenyi (2002) Grid scale: Neu and Prather (2012) for gases and Chapman et al. (2009) for aerosols	Grid scale wet deposition: Easter et al. (2004)	Below cloud: Slinn (1984) In cloud: Easter et al. (2004)
Land use–land cover classification	Europe: CORINE, 33 classes North America: USGS-24 (Anderson et al., 1976), 24 classes	USGS-24 (Anderson et al., 1976), 24 classes	Modified IGBP MODIS NOAH, 21 classes including oceans and inland water, Friedl et al. (2010)
Cumulus cloud parameterization	Grell and Devenyi (2002)	Grell and Devenyi (2002)	Grell and Freitas (2014)
Windblown dust	Online, Shao et al. (2011)	MOSAIC, Zaveri et al. (2008)	GOCART, with AFWA modifications, Gong et al. (1997), Ginoux et al. (2001)
Gas-phase chemistry mechanism	MOZART, Emmons et al. (2010)	CMBZ, Zaveri and Peters (1999)	MOZART, Emmons et al. (2010)
Vertical resolution	38 levels up to 50 hPa	35 vertical levels	41 vertical levels
PBL scheme	Mellor–Yamada–Janjić, Janjić (2001)	Yonsei University (YSU), Hong et al. (2006), Hong (2010)	Mellor–Yamada level 2.5, Nakanishi and Niino (2006)

and total nitrogen deposition for each model in the ensemble, for 2009 and 2010 for Europe and for 2010 and 2016 for North America. In order to make a rough estimate of the impacts of model biases on the resulting exceedance estimates, a third set of exceedances were calculated for each model and each domain, for the year 2010 for Europe and 2016 for North America. For this last group, the ratio of the observed to model mean values at the observation station locations for individual species was used as scaling factors for the model annual deposition flux estimates prior to summation to total sulfur and total nitrogen deposition. Specifically, for North America, the ratio of the observed to measured mean concentrations of SO_2, NO_2, $PM_{2.5}$ sulfate, and $PM_{2.5}$ ammonium and Ammonia Monitoring Network (AMoN) NH_3 data were used to scale the corresponding dry flux variables, and the corresponding ratios for the wet deposition of sulfate, nitrate, and ammonium ions were used to scale the wet-deposition fluxes. Fewer observation data were available for Europe than North America: the ratios of observed to modelled SO_2 and NO_2 gas concentration mean values were used to scale the corresponding dry fluxes, and ratios of observed to modelled wet-deposition fluxes for sulfate, nitrate, and ammonium were used to scale the modelled wet-deposition fluxes.

We note that this approach makes simplifying assumptions. The corrections are inherently dependent on the assumption that the monitoring data are sufficiently representative of the model domain for the correction to be meaningful across the domain. While dry-deposition fluxes will be proportional to the concentrations in the lowest model layer, allowing an overall mean bias correction, we are also making the assumption that the bias ratios for $PM_{2.5}$ particulate matter will apply for larger particle sizes as well (note that size-resolved particulate fluxes were not reported under the AQMEII4 protocol). This form of bias correction is also the simplest possible means of model–measurement fusion; more complex methods appear in the literature. These methodologies for example may make use of a combination of observed wet and adjusted model dry deposition (Schwede and Lear, 2014), inverse distance weighting from observation stations (Rubin et al., 2023), and adjustment of modelled wet-deposition fluxes by the ratio of observed to simulated precipitation and by kriged observed wet deposition to model-predicted ratios (Zhang et al., 2019). An overview of model–measurement fusion approaches including advanced forms of data assimilation may be found in Fu et al. (2022). The methodology used here provides a first-order estimate of

Table 2. AQMEII4 GEM-MACH configuration differences.

Parameterization	GEM-MACH (Base)	GEM-MACH (Zhang)	GEM-MACH (Ops)
Gas dry deposition	Makar et al. (2018)	Zhang et al. (2003)	Makar et al. (2018)
Ammonia bidirectional fluxes	Zhang et al. (2010)	As in GEM-MACH (Base)	None
Particulate matter dry deposition	1D semi-Lagrangian mass transfer (Makar et al., 2018), using Emerson et al. (2020) correction to Zhang et al. (2001) coefficients	As in GEM-MACH (Base)	Zhang et al. (2001), applied as the flux lower boundary condition in the diffusion equation
Vertical resolution	83 levels plus 3 additional levels for forest canopy processes (Makar et al., 2017)	As in GEM-MACH (Base)	83 levels
Meteorological-model cloud parameterization	P3 explicit hydrometeor scheme (Morrison and Milbrandt, 2015; Milbrandt and Morrison, 2016).	As in GEM-MACH (Base)	Convective: Kain–Fritsch convective parameterization (Kain and Fritsch, 1990; Kain, 2004) Stratiform: Sundqvist et al. (1989)
Land use land cover classification	GEM-MACH 15 land use scheme (Makar et al., 2018), aggregated from 26 Zhang et al. (2002, 2003) land use categories	Zhang et al. (2002, 2003), 26 land use categories	As in GEM-MACH (Base)
Leaf area index data source	Satellite-derived (Zhang et al., 2020)	As in GEM-MACH (Base)	BEIS-based (Vukovich and Pierce, 2002; Stroud et al., 2010)
Seasonality for emissions	Based on satellite LAI (Zhang et al., 2020)	As in GEM-MACH (Base)	Fixed function of latitude and Julian day
Major point source plume rise algorithm	Akingunola et al. (2018)	As in GEM-MACH (Base)	Briggs (1984)
Gas-phase chemistry solver	KPP2.1 (Sandu and Sander, 2006)	As in GEM-MACH (Base)	Young and Boris (1977)
Vehicle-induced turbulence	Makar et al. (2021)	As in GEM-MACH (Base)	None
Forest canopy shading and turbulence	Makar et al. (2017)	As in GEM-MACH (Base)	None
CH_4 as a chemically active tracer	Yes	As in GEM-MACH (Base)	No
Aerosol direct- and indirect-effect feedback	Yes (Makar et al., 2015a, b)	As in GEM-MACH (Base)	No
Floor (minimum) PBL height imposed	No	As in GEM-MACH (Base)	Yes (100 m)
Area source emission treatment	Flux lower boundary condition on diffusion equation	As in GEM-MACH (Base)	Mass injection into the two lowest model layers
Advection mass conservation	ILMC, 3 sweeps (Sørensen et al., 2013)	As in GEM-MACH (Base)	ILMC, 2 sweeps, followed by a Bermejo and Conde (2002) global mass correction

the impact of model biases with respect to observations of critical load exceedances.

3 Results

3.1 Critical load exceedances

3.1.1 Europe: acidification

Critical load exceedances for acidification for each of the four models for Europe (EU) are shown in Fig. 1 for 2010, Fig. S3 (Supplement) for 2009, and Fig. S9 (Supplement) for bias-corrected 2010. Figure 2 shows the reduced-ensemble values for 2009 and 2010 (a, b), the bias-corrected value for 2010 (c), and AQMEII4 common-domain total bar charts for all models and the reduced ensemble (d).

The EU exceedances for acidity are similar between the 2 years (compare Figs. 1 and S3 and reduced-ensemble values for each year in Fig. 2). However, differences between models within a given year are larger (especially in an absolute sense; WRF-Chem (IASS) of <0.4 % in exceedance, WRF-Chem (UPM) of ∼6.5 %). Low WRF-Chem (IASS) exceedance levels are in part due to unreported deposition data (see Sect. 2.2.2); the reduced-ensemble maps in Fig. 2 show the ensemble average for LOTOS-EUROS (TNO), WRF-Chem (UPM), and CMAQ (Hertfordshire). The EU reduced ensemble shows the greatest extent of exceedance in the Netherlands along the Netherlands–Belgium border, northwestern Germany, southern Norway, and along the border between Poland and Germany (Fig. 2a, b). Individual models in Fig. 1 show additional acidity "hotspots" that may appear in one model and not in another (e.g. LOTOS-EUROS (TNO): near Lucerne and Bonn; WRF-Chem (UPM): westernmost Switzerland, south-central Germany, and Belgrade; CMAQ (Hertfordshire): southwestern Switzerland, south-central Germany, and southwestern Romania). Bias correction for the reduced ensemble for the 2010 data resulted in substantial increases in predicted exceedances (compare the last two columns of Fig. 2d and compare Fig. 1 to Fig. S9). However, we note that the European data did not include speciated particulate matter, and hence bias correction was not possible for part of the sulfur budget; much smaller impacts were noted for bias correction in North America where particulate sulfate data were available.

The percent area of the EU acidification CLE over the region for which CL data were available, for the reduced ensemble, was 4.48 % (range of 2.37 % to 6.85 %) in 2009 and 4.32 % (2.06 to 6.52 %) in 2010. Average reduced-ensemble accumulated exceedance for EU acidity was 13.8 (9.7 to 27.1) eq. ha^{-1} yr^{-1} in 2009 and 12.6 (7.8 to 23.7) eq. ha^{-1} yr^{-1} in 2010. The quoted range is from the highest and lowest members in the three-member reduced ensemble.

3.1.2 Europe: eutrophication

Critical load exceedances for eutrophication for each of the four EU models are shown in Fig. 3 for 2010, in Fig. S4 (Supplement) for 2009, and with bias-corrected deposition fields for 2010 in Fig. S10 (Supplement). Figure 4 shows the reduced-ensemble values for 2009 and 2010 (a, b), the bias-corrected values for 2010 (c), and the AQMEII4 common-domain summaries for all models and the ensembles (d).

As for EU acidity CLEs, the eutrophication CLEs are very similar between the 2 model years (compare Figs. 3 and S4 and the values for each year in Fig. 4). The spatial distribution of the greatest levels of exceedance also varies more strongly between models. All members in the three-member reduced ensemble identify the Po River valley as reaching the greatest level of exceedance, but LOTOS-EUROS (TNO) also shows high levels of exceedance from Benelux to northern Germany and in the Barcelona area, while WRF-Chem (UPM) shows high levels of exceedance of >800 eq. ha^{-1} yr^{-1} in multiple hotspots throughout the region. The relative impact of bias correction was smaller than for acidification in terms of the total area in exceedance, but the magnitude of exceedances increased significantly (e.g. larger proportion of red to black areas in Fig. 4c than Fig. 4b, comparing the last two columns of Fig. 4d and comparing Fig. 4 to Fig. S10). Again, the higher levels of exceedance predicted for Europe may reflect the impact of the lack of particulate sulfate and particulate nitrate data for bias correction purposes.

The percentage of the area in exceedance for eutrophication is much higher than that of acidification (reduced-ensemble CLE of 60.2 % (47.3 % to 73.3 %) in 2009 and 62.2 % (51.2 % to 74.4 %) in 2010). The average accumulated exceedance was 156.9 (89.4 % to 265.5) eq. ha^{-1} yr^{-1} in 2009 and 161.4 (109.4 to 261.8) eq. ha^{-1} yr^{-1} in 2010 (in Fig. 4 the range is from the lowest and highest members in the three-member reduced ensemble).

3.1.3 North America: forest-ecosystem simple mass balance critical load

Critical load exceedances with respect to the forest soil acidity for North America (NA) for the years 2016 and 2010 are shown in Figs. 5 and S5, respectively; the bias-corrected 2016 maps are in Fig. S11, and the reduced-ensemble maps for both years, as well as the domain summaries, including bias-corrected values for 2016, are shown in Fig. 6.

Unlike the EU domain comparison, the NA CLEs depicted in Fig. 5 show a large difference in the extent of regions in exceedance for the different models. While all models with the exception of WRF-Chem (IASS) identified the regions to the south and west of the Great Lakes, the US East Coast, and Florida as being in exceedance, the magnitude of the exceedances varied greatly between the models, with the GEM-MACH models (Fig. 5d–f) showing

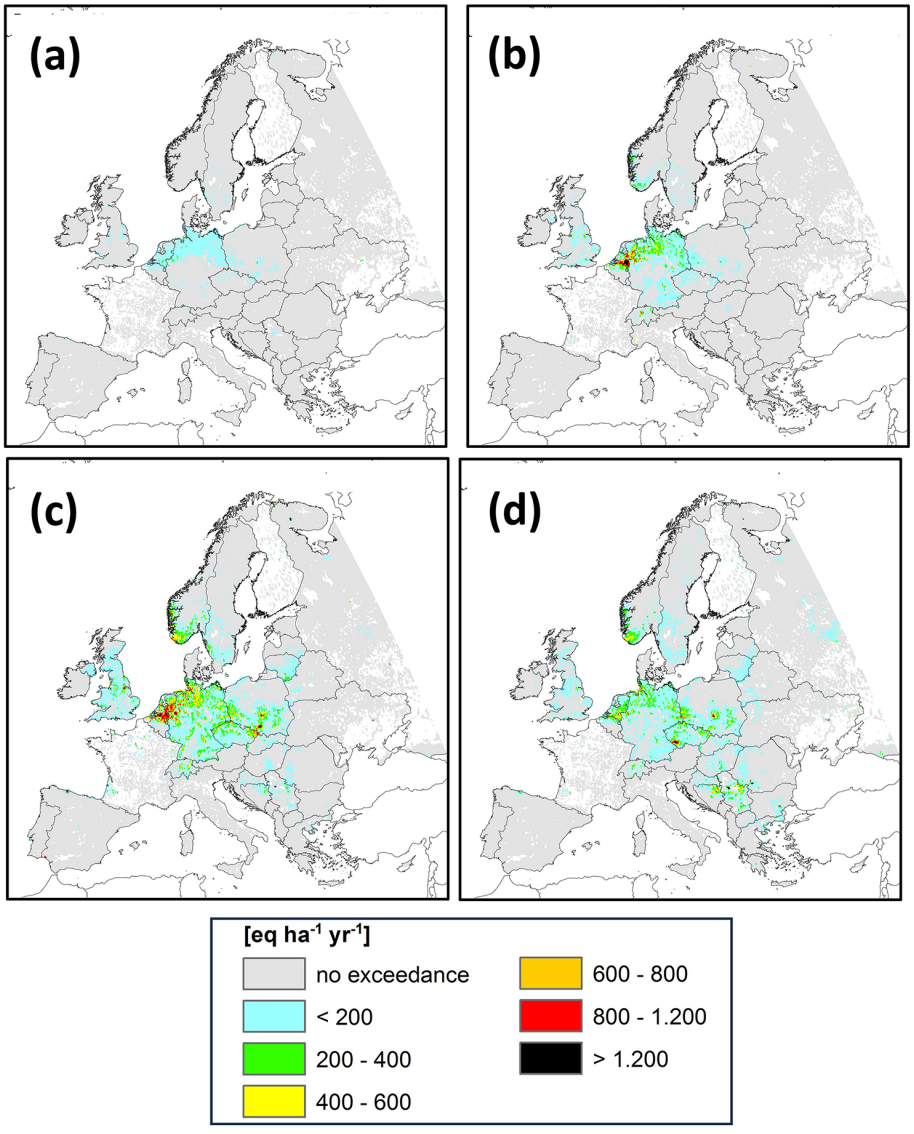

Figure 1. CLEs for acidity, EU AQMEII4 common domain, 2010, eq. ha^{-1} yr^{-1}. **(a)** WRF-Chem (IASS), **(b)** LOTOS-EUROS (TNO), **(c)** WRF-Chem (UPM), **(d)** CMAQ (Hertfordshire). Grey areas indicate regions for which critical load data are available but are not in exceedance of critical loads. Coloured areas indicate exceedance regions.

large regions with exceedances above 800 eq. ha^{-1} yr^{-1}, followed by, in descending order, WRF-Chem (UPM), CMAQ-M3Dry, CMAQ-STAGE, WRF-Chem (UCAR), and WRF-Chem (IASS).

The summary reduced-ensemble CLE values (Fig. 6) show the improvement in CLEs between the years 2010 and 2016, which occurred in response to the legislated reduction in SO_2 emissions during this time period. The summary chart (Fig. 6c) however shows that the magnitude of the response to the SO_2 reduction was model dependent: the change between 2010 and 2016 was the greatest for GEM-MACH (Base) in an absolute sense and the greatest for WRF-Chem (UCAR) in a relative sense. Similarly, the average accumulated exceedance (right-hand vertical axis and black diamonds, Fig. 6c) showed decreases in exceedance between 2010 and 2016 for all models, but the extent of these decreases differed, with WRF-Chem (UCAR) showing the smallest decrease in AAE from 2010 to 2016, followed in increasing order of the magnitude of change by CMAQ-STAGE, CMAQ-M3Dry WRF-Chem (UPM), GEM-MACH (Ops), GEM-MACH (Base), and GEM-MACH (Zhang).

The effect of bias correction was less pronounced than in Europe and, in general, reduced the variability between model results. Note that unlike the European case, North American observation data used for bias correction included corrections for particulate sulfate air concentrations, allowing for a greater degree of closure for the sulfur mass deposited. Comparing Figs. 5 and S11 it can be seen that the

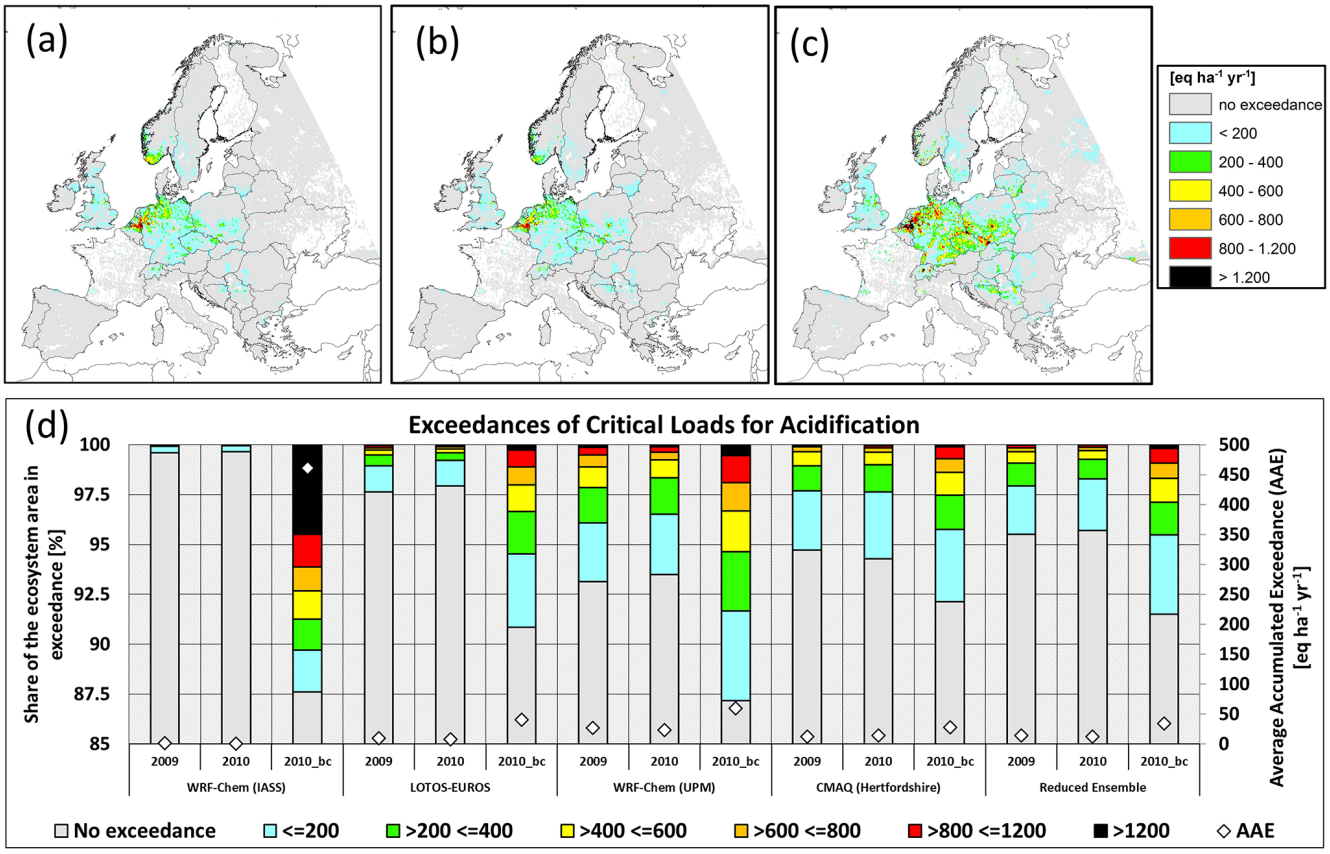

Figure 2. Summary CLEs for acidity, EU AQMEII4 common domain, eq. ha^{-1} yr^{-1}. **(a, b)** Spatial distribution of CLEs for the reduced ensemble for the years 2009 and 2010, respectively. **(c)** Spatial distribution of CLE for the bias-corrected reduced ensemble for the year 2010. **(d)** Percentage of ecosystems for which CL data are available that are in exceedance by model and year (left axis and colour bar) and average accumulated exceedance (eq. ha^{-1} yr^{-1}) (right axis and black diamonds). Reduced-ensemble CLEs for the years 2009, 2010, and bias-corrected 2010 (2010_bc) appear on the right side of panel **(d)**.

bias correction has increased exceedances for the CMAQ and WRF-Chem simulations and decreased exceedances for the GEM-MACH simulations, reducing the variability between the models. The extent to which model-to-model variability has been reduced following bias correction is also apparent in Fig. 6d (bias correction exceedance bars are closer in size across models compared to before bias correction). The net result is bias correction being a slight increase in the area of exceedance in the reduced ensemble, comparing the two right-hand bars of Fig. 6d.

The percentage of the NA forested area in exceedance for acidification for the reduced ensemble was 13.2 % (2.8 % to 22.2 %) in 2010 and 6.1 % (1.0 % to 12.9 %) in 2016. The ensemble thus shows a considerable improvement in exceedances with respect to acidification between the 2 years.

3.1.4 North America: aquatic-ecosystem critical-load exceedances with respect to acidity)

Exceedances with respect to the North American aquatic-ecosystem CL dataset for the years 2016 and 2010 are shown in Figs. 7 and S6, respectively; the bias-corrected maps for each model for 2016 are in Fig. S12; and the reduced-ensemble maps for both years and domain summaries, including bias correction, are shown in Fig. 8.

Comparison of Figs. 5 and 7 shows a similarity in the CLE response of the individual models between forest soil and aquatic ecosystems, with the GEM-MACH models predicting the highest number and magnitude of exceedances, followed by WRF-Chem (UPM), WRF-Chem (UCAR), and the two CMAQ implementations. Figure 8a and b show the expected decrease in the reduced-ensemble CLE between 2010 and 2016, as well as the higher levels of exceedance associated with the GEM-MACH and WRF-CHEM (UPM) models, followed in descending order by the two CMAQ implementations and WRF-CHEM (UCAR) (Fig. 8c).

The impact of bias correction on the North American aquatic-ecosystem critical load exceedances was relatively minimal for the models included in the reduced ensemble: differences between Figs. 7 and S12 are difficult to distinguish, and Fig. 8d shows slight increases in the exceedances for CMAQ and WRF-Chem simulations, slight increases in

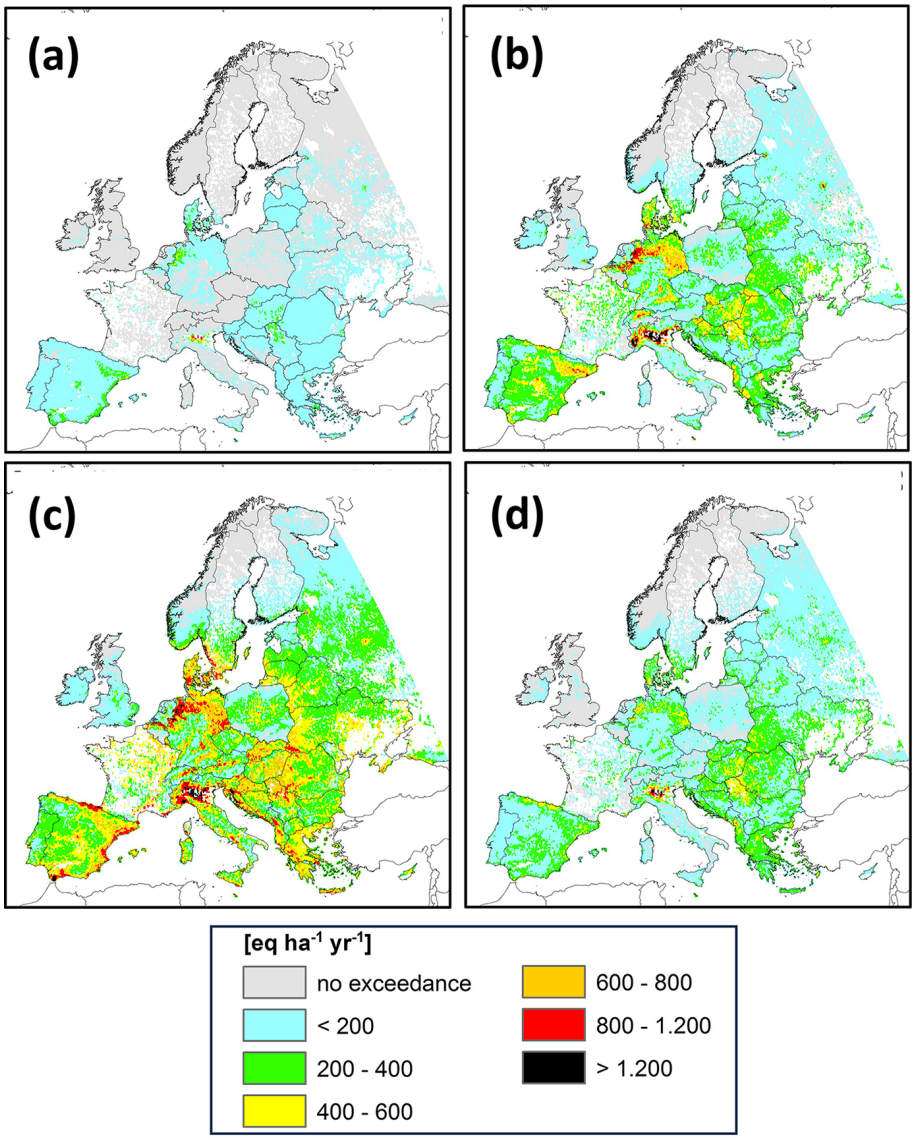

Figure 3. CLEs for eutrophication, EU AQMEII4 common domain, 2010, eq. $ha^{-1}\,yr^{-1}$. **(a)** WRF-Chem (IASS), **(b)** LOTOS-EUROS (TNO), **(c)** WRF-Chem (UPM), **(d)** CMAQ (Hertfordshire). Grey areas indicate regions for which critical load data are available but are not in exceedance of critical loads. Coloured areas indicate exceedance regions.

GEM-MACH simulations, and a very small change in the reduced-ensemble levels of exceedance.

The percentage of the NA aquatic ecosystems in exceedance for the reduced ensemble was 21.2 % (12.8 % to 28.9 %) in 2010 and 11.4 % (7.3 % to 15.8 %) in 2016. The reduced ensemble thus shows a considerable improvement in exceedances with respect to the exceedance of aquatic critical loads between the 2 years, again by almost a factor of 2.

3.1.5 US N deposition to lichen

Exceedances with respect to the US CL of N for a 20 % decline in the sensitive-epiphytic-lichen species richness (221 eq. N $ha^{-1}\,yr^{-1}$) dataset for the years 2016 and 2010 are shown in Figs. 9 and S7, respectively; bias-corrected 2016 values are in Fig. S13; and the reduced-ensemble maps for both years and domain summaries, including bias-corrected 2016 values, are shown in Fig. 10.

The overall pattern of exceedances and their magnitude across models (Fig. 9) is similar to that of the forest soil exceedances (Fig. 5), with the largest magnitudes in the northeastern continental USA and in North Carolina, though the lichen exceedances are more continuous across the region than for forest soil water acidity-impacted ecosystems. GEM-MACH (Base), GEM-MACH (Zhang), and GEM-MACH (Ops) have maximum exceedances usually between 800 and 1200 eq. $ha^{-1}\,yr^{-1}$, and the exceedances predicted

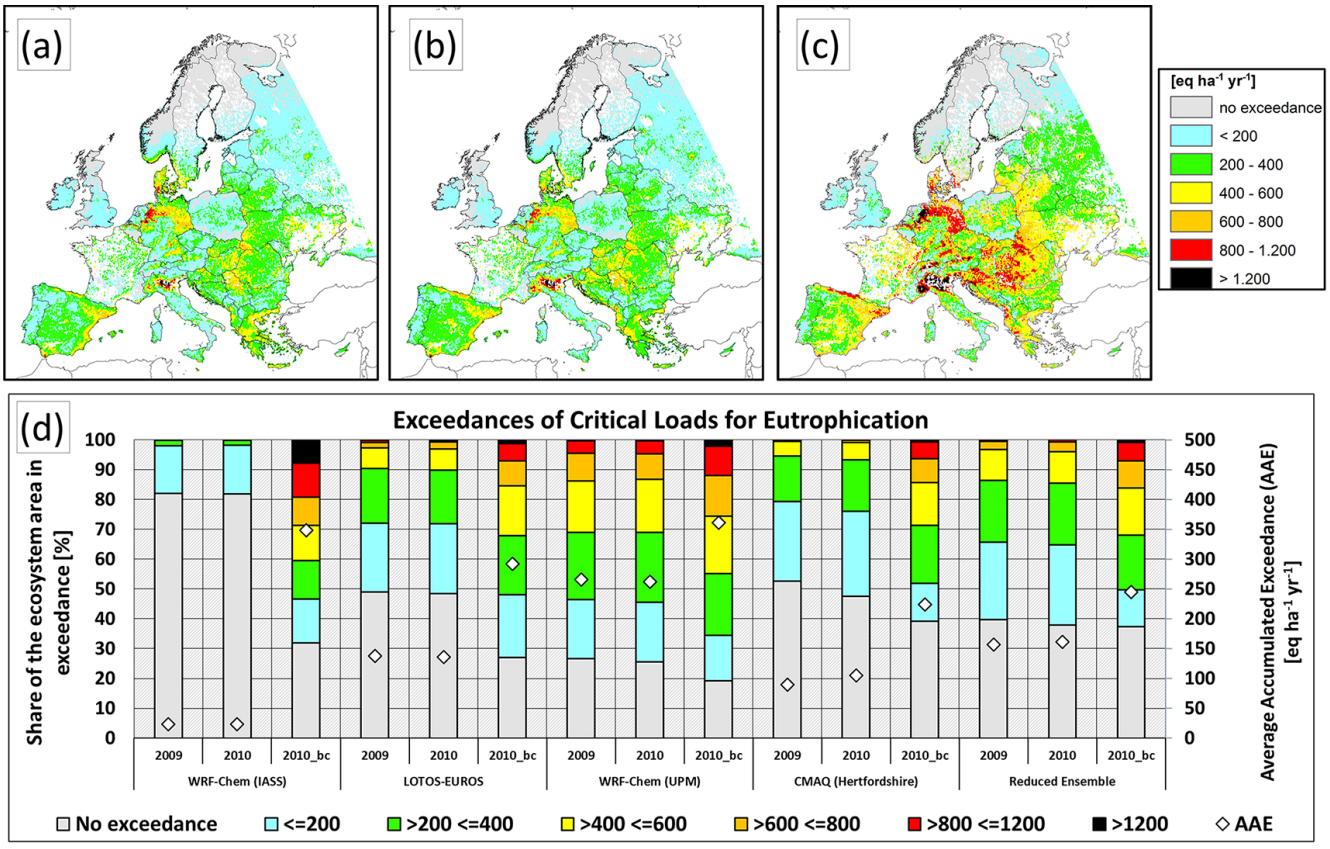

Figure 4. Summary CLEs for eutrophication, EU AQMEII4 common domain, eq. ha^{-1} yr^{-1}. **(a, b)** Spatial distribution of CLEs for the reduced ensemble for the years 2009 and 2010, respectively. **(c)** Spatial distributions of CLEs for the bias-corrected reduced ensemble for 2010. **(d)** Percentage of ecosystems for which CL data are available that are in exceedance by model and year (left axis and colour bar) and average accumulated exceedance (eq. ha^{-1} yr^{-1}) (right axis and black diamonds). Reduced-ensemble CLEs for the years 2009, 2010, and bias-corrected 2010 (2010_bc) appear on the right side of panel **(d)**.

by other models are less than 800 eq. ha^{-1} yr^{-1}, aside from a North Carolina exceedance hotspot which is predicted by all models. The overall reduced-ensemble magnitude of exceedances decreased significantly between 2010 and 2016 (Fig. 10a, b; fewer black and red regions in the more recent year). The reduced-ensemble total area in exceedance has decreased slightly (columns labelled "Reduced ensemble" in Fig. 10d). All models show a decreasing levels of exceedance between the 2 years and a slightly decreasing total area of exceedance. The magnitude of exceedance differs significantly between the models, with the highest-magnitude exceedances predicted by the GEM-MACH group of models, followed by WRF-Chem (UPM).

Bias correction values varied between the models, with CMAQ exceedances increasing slightly, GEM-MACH exceedances decreasing slightly, WRF-Chem exceedances increasing, and a slight increase in the overall extent and magnitude of the reduced-ensemble exceedances in the last two columns of Fig. 10d. The similarity in the spatial distribution of exceedances is greater across models following bias correction (compare Fig. 9 with Fig. S13 in the Supplement).

The percentage of the NA sensitive-epiphytic-lichen ecosystems in exceedance for the reduced ensemble was 81.5 % (69.3 % to 95.0 %) in 2010 and 75.8 % (63.7 % to 90.7 %) in 2016.

3.1.6 US N deposition to herbaceous plants

Exceedances with respect to the US CL of N for a decline in the herbaceous-species richness (436 to 1693 eq. N ha^{-1} yr^{-1}) dataset for the years 2016 and 2010 are shown in Figs. 11 and S8, respectively; bias-corrected exceedances for 2016 appear in Fig. S14 (Supplement); and the reduced-ensemble maps for both years and domain summaries, including bias correction for 2016, are shown in Fig. 12.

The spatial distribution of the regions of the highest exceedance shares some common features with that of sensitive epiphytic lichen (compare Fig. 11 with Fig. 9), such as maximum exceedances in the northeastern USA, in North Carolina, and extending along a region north of Texas. However, both the magnitude and extent of exceedance is much

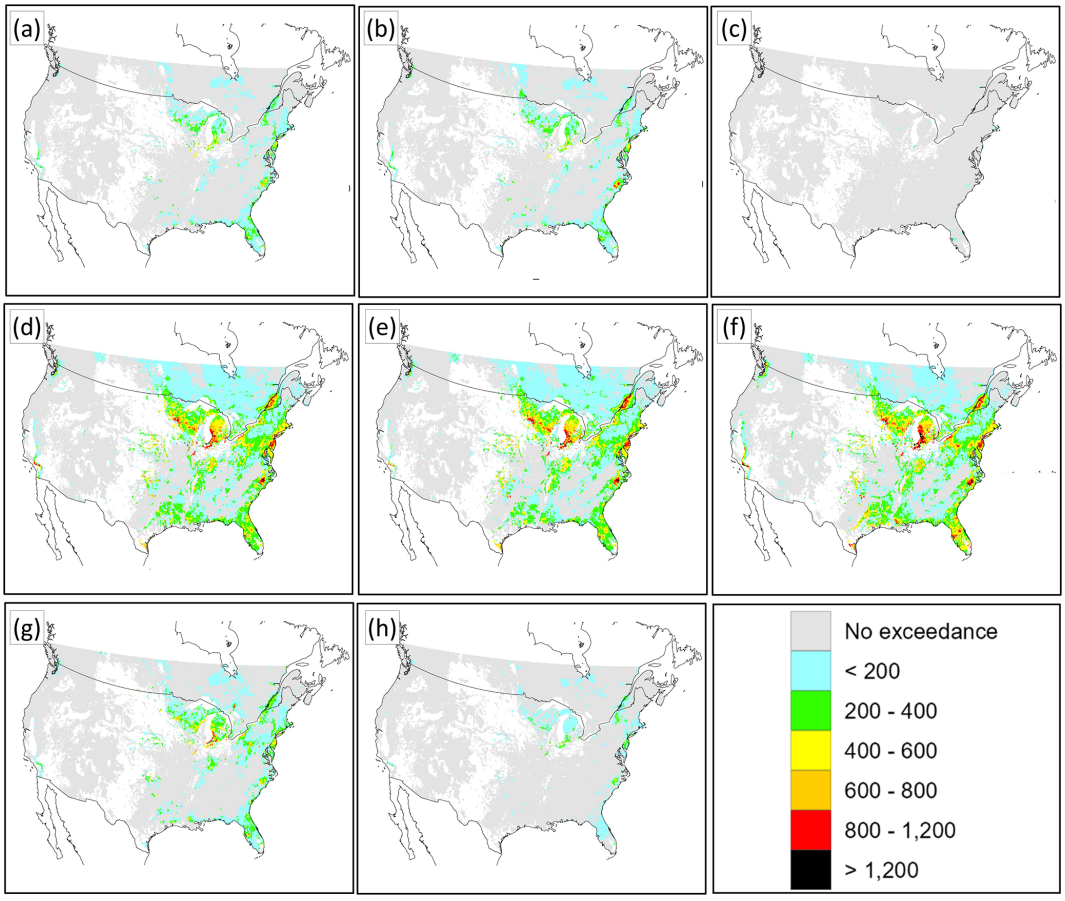

Figure 5. CLEs for forest soil acidification, NA AQMEII4 common domain, 2016, eq. ha^{-1} yr^{-1}. **(a)** CMAQ-M3Dry, **(b)** CMAQ-STAGE, **(c)** WRF-Chem (IASS), **(d)** GEM-MACH (Base), **(e)** GEM-MACH (Zhang), **(f)** GEM-MACH (Ops), **(g)** WRF-Chem (UPM), **(h)** WRF-Chem (UCAR). Grey areas indicate regions for which critical load data are available but are not in exceedance of critical loads. Coloured areas indicate exceedance regions.

more varied for herbaceous-species richness than for lichen species richness, with the GEM-MACH suite of models (Figs. 11d–f and 12d) predicting the highest exceedance levels and up to 18.4 % of the area in exceedance in 2016, with the CMAQ implementations varying between 0.6 % and 0.8 % and WRF-Chem (UCAR) predicting 0.1 %.

The impacts of bias correction may be more easily distinguished for herbaceous-species richness critical load exceedances compared to some of the other exceedance estimates (compare Figs. 11 and S14), with the CMAQ and WRF-Chem exceedances increasing and the GEM-MACH exceedances decreasing. The overall impact was a slight increase in the area and extent of the ensemble average exceedance (Fig. 12d).

The percentage of the NA herbaceous-plant ecosystems in exceedance for the reduced ensemble was 13.9 % (0.4 % to 39.5 %) in 2010 and 3.9 % (0.1 % to 18.4 %) in 2016, with the higher exceedance levels in the range resulting from the GEM-MACH suite of models. Reduced-ensemble herbaceous-species richness exceedances have decreased considerably between the 2 years in all models.

3.1.7 Critical load exceedances: key results

The percent exceedance for the reduced ensemble and ranges from the reduced ensembles for the ecosystems examined here are summarized in Table 3. The values suggest acidification in Europe will happen over a smaller region than eutrophication at 2009–2010 emission levels, with a slight decrease in acidification and a slight increase in eutrophication between the 2 years. About 60 % of EU ecosystems would be subject to eutrophication at some point in the future at 2009–2010 emission levels. One striking difference between the different model estimates of CLE is in the magnitude of exceedances (as opposed to the total area in exceedance). WRF-Chem (UPM) for example in Figs. 1 and 3 predicts more severe levels of exceedance across Europe than the other models. The North American results suggest that reductions in SO_2 and NO_x emissions between 2010 and 2016

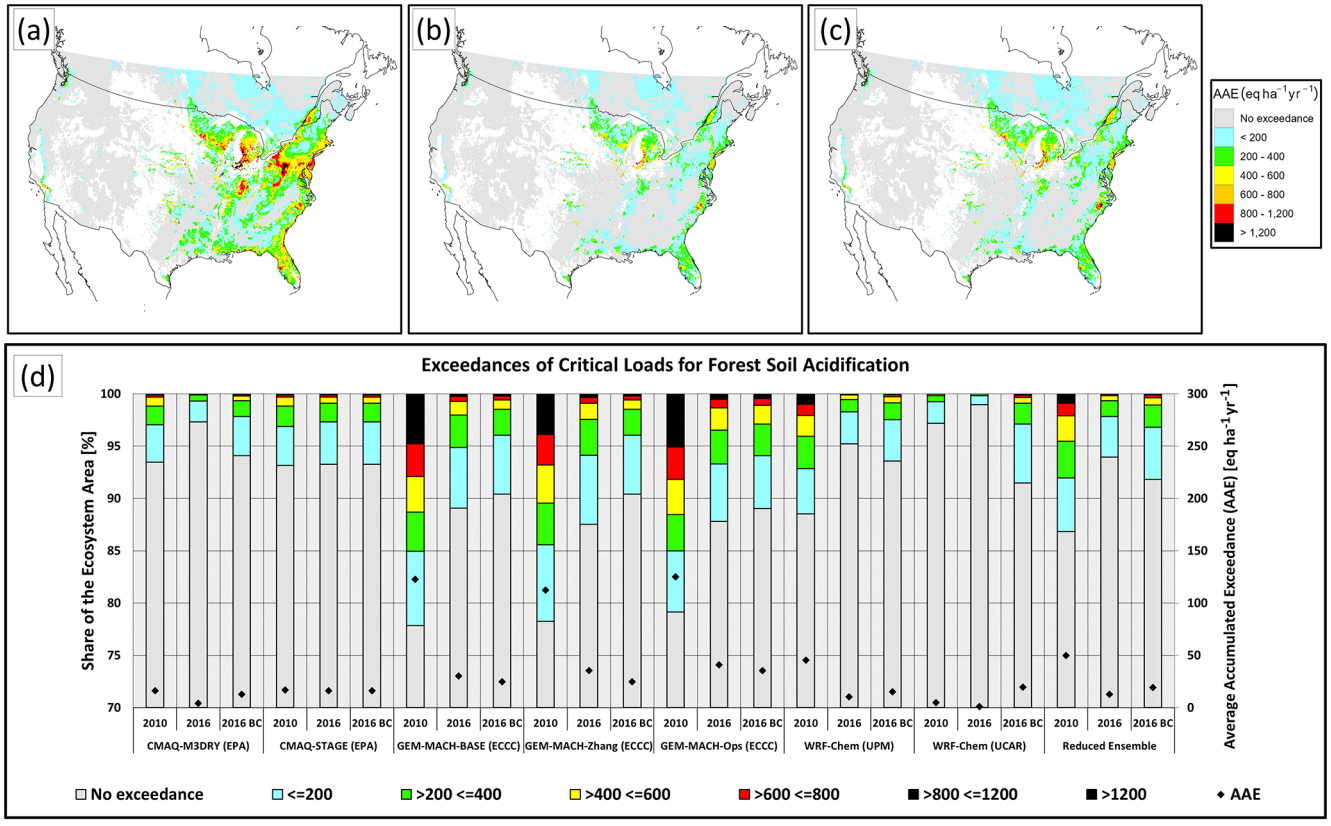

Figure 6. Summary CLEs for forest soil acidification, NA AQMEII4 common domain, eq. ha^{-1} yr^{-1}. **(a, b)** Spatial distribution of CLEs for the reduced ensemble for the years 2010 and 2016, respectively. **(c)** Spatial distribution of CLEs for the reduced ensemble for the year 2016. **(d)** Percentage of ecosystems for which CL data are available that are in exceedance by model and year (left axis and colour bar) and average accumulated exceedance (eq. ha^{-1} yr^{-1}) (right axis and black diamonds). Reduced-ensemble CLEs for the years 2010, 2016, and bias-corrected 2016 (2016 BC) appear on the right side of panel **(d)**.

resulted in a substantial reduction in the number of forest soil and aquatic-ecosystem acidification exceedances (by nearly a factor of 2). The impacts of nitrogen deposition on herbaceous species also improved (by nearly a factor of 3), while impacts of nitrogen deposition on sensitive lichen had a more modest improvement (from 81.5 % to 75.8 % in exceedance). The magnitude and spatial extent of these eutrophication exceedances were highly dependent on the model and on the variations in the representation of sub-processes within each model, used for predictions. Understanding the large range of model predictions is one of the main aims of the current work. The next section discusses the underlying causes driving the model-to-model differences, using the AQMEII4 deposition diagnostics.

3.2 Analysis of model deposition predictions

3.2.1 Causes of S deposition variability in North American domain simulations

The AQMEII4 common-grid average and percent contribution of each depositing species to total S deposition in 2016 are given in Table 4. The averages and standard deviation for the reduced ensemble show that wet deposition of the sum of the sulfate and bisulfite ions ($SO_4^{(2-)}$ and $HSO_3^{(-)}$) contributes more to total S deposition than particulate sulfate dry deposition, which is in turn contributes more than SO_2(g) (gas) dry deposition. However, the model-to-model variability is also large, particularly for the contribution of particulate sulfate, which varies by nearly 2 orders of magnitude between GEM-MACH (Base), GEM-MACH (Zhang), and GEM-MACH (Ops) and WRF-Chem (UPM). The contributions to the average reduced-ensemble total S deposition are 62.0 ± 19.3, 44.8 ± 39.0, and 28.8 ± 9.9 eq. ha^{-1} yr^{-1} for wet, particle dry, and gas dry deposition, respectively ($\pm$ ranges in Table 4 are the standard deviation of the component). The greatest cause of model *variability* in absolute total deposition is associated with the contribution of particulate sulfate dry deposition, followed by sulfur wet deposition and then gaseous SO_2 dry deposition.

The spatial distributions of the two largest components of the total S deposition variability (wet S and dry particle S) are shown in Fig. 13. The WRF-Chem (IASS) val-

Table 3. Summary of reduced-ensemble percent exceedance mean values and their range in the EU and NA domains, along with total S deposition and total N deposition predicted by the ensemble. All models used the same starting inventories for emissions.

EU ecosystem	Year 2009 percent exceedance (lower to upper bound)	Year 2010 percent exceedance (lower to upper bound)	Total S deposition, 2010 (eq. ha^{-1} yr^{-1})	Total N deposition, 2010 (eq. ha^{-1} yr^{-1})
Acidification	4.48 (2.37 to 6.85)	4.32 (2.06 to 6.52)	158.4 (81.5 to 221.6)	376.5 (304.8 to 481.9)
Eutrophication	60.2 (47.3 to 73.3)	62.2 (51.2 to 74.4)		
NA ecosystem	Year 2010 percent exceedance (lower to upper bound)	Year 2016 percent exceedance (lower to upper bound)	Total S deposition, 2016 (eq. ha^{-1} yr^{-1})	Total N deposition, 2016 (eq. ha^{-1} yr^{-1})
Forest soil acidification	13.2 (2.8 to 22.2)	6.1 (1.0 to 12.9)	135.6 (56.1 to 193.4)	321.7 (182.4 to 430)
Lake ecosystems	21.2 (12.8 to 28.9)	11.4 (7.3 to 15.8)		
US N deposition lichen	81.5 (69.3 to 95.0)	75.8 (63.7 to 90.7)		
US N deposition herbaceous	13.9 (0.4 to 39.5)	3.9 (0.1 to 18.4)		

Table 4. Average S deposition contributions in the NA AQMEII4 common-grid area (eq. ha^{-1} yr^{-1}) and percent contribution to average total S deposition, 2016. n/d: no data submitted or insufficient data to calculate a percentage.

	Average deposition (eq. ha^{-1} yr^{-1})				Percent of total S deposition		
Model number	$SO_4^{(2-)} + HSO_3^{(-)}$ wet deposition	Particle sulfate dry deposition	SO_2(g) dry deposition	Total S deposition	$SO_4^{(2-)} + HSO_3^{(-)}$ wet deposition	Particle sulfate dry deposition	SO_2(g) dry deposition
CMAQ-M3Dry	79.0	19.0	24.9	122.9	64.3	15.4	20.2
CMAQ-STAGE	79.2	21.0	23.3	123.4	64.2	17.0	18.8
WRF-Chem (IASS)	0.9	n/d	26.7	n/d	n/d	n/d	n/d
GEM-MACH (Base)	52.4	90.7	23.0	166.1	31.5	54.6	13.9
GEM-MACH (Zhang)	51.4	88.8	25.1	165.3	31.1	53.7	15.2
GEM-MACH (Ops)	81.3	88.2	23.9	193.4	42.0	45.6	12.4
WRF-Chem (UPM)	66.3	2.8	52.8	121.9	54.4	2.3	43.3
WRF-Chem (UCAR)	24.4	3.0	28.7	56.1	43.5	5.3	51.2
Reduced-ensemble average	62.0	44.8	28.8	135.6	45.7	33.0	21.2
Reduced-ensemble standard deviation	19.3	39.0	9.9	41.3	13.0	21.2	14.5

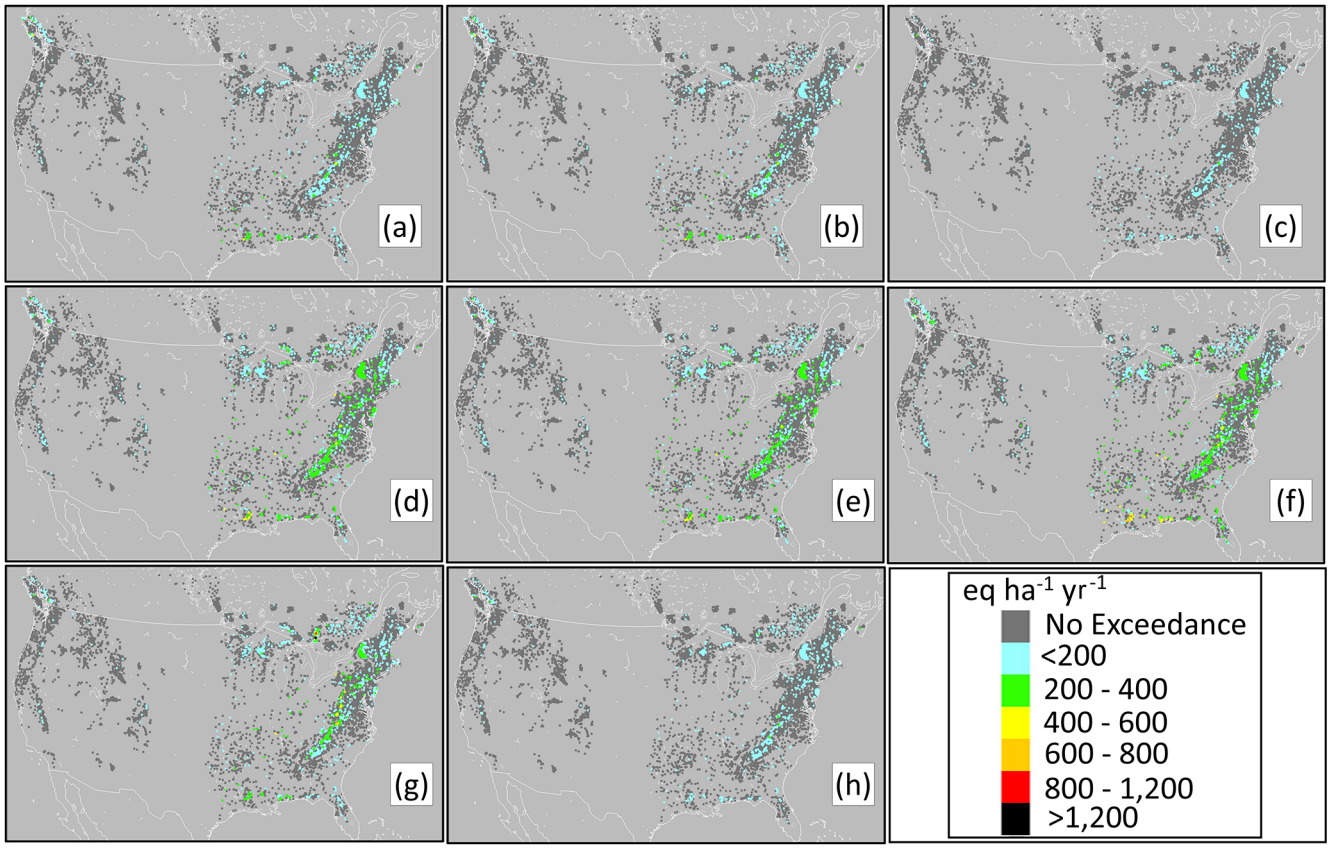

Figure 7. CLEs for aquatic ecosystems, NA AQMEII4 common domain, 2016, eq. ha^{-1} yr^{-1}. Panels arranged by model as in Fig. 6; individual sites are shown as pixels. Dark-grey pixels indicate regions for which critical load data were available but were not in exceedance of critical loads. Coloured areas indicate exceedance regions; overplotting in precedence by the extent of exceedance was carried out for overlapping pixels. Areas of no CL data are shown in lighter grey.

ues did not represent the expected sources of S deposition over the continent, and some deposition fields such as the total particulate sulfate dry deposition were not submitted. The S wet-deposition maps are qualitatively similar between the other models (note that the colour scale is logarithmic), with WRF-Chem (UCAR) having the lowest values (Fig. 13a). As shown in Table 4, the greatest degree of variability between the different modelling platforms is in the particle deposition fluxes (Fig. 13b). This variability extends over orders of magnitude. WRF-Chem (UPM) and WRF-Chem (UCAR) predict the lowest deposition fluxes of dry particulate sulfate over both land and ocean. CMAQ-STAGE and CMAQ-M3Dry predict higher values over parts of the ocean but relatively low values over land. GEM-MACH (Base), GEM-MACH (Zhang), and GEM-MACH (Ops) have the highest particulate sulfate dry-deposition fluxes, roughly equivalent to the wet-deposition fluxes.

We next evaluate each of the models' predictions against North American network observations for concentrations of SO_2 and particulate sulfate and sulfur wet deposition for the year 2016. The monitoring network databases employed included the U.S. Environmental Protection Agency Air Quality System (AQS; https://www.epa.gov/aqs, last access: 7 July 2024), the National Atmospheric Deposition Program National Trends Network (NADP NTN; https://nadp.slh.wisc.edu/networks/national-trends-network/, last access: 7 July 2024), the Canadian National Air Pollution Surveillance (NAPS) program (https://www.canada.ca/en/environment-climate-change/services/air-pollution/monitoring-networks-data/national-air-pollution-program.html, last access: 7 July 2024), and the Canadian National Atmospheric Chemistry database (https://www.canada.ca/en/environment-climate-change/services/air-pollution/monitoring-networks-data/national-atmospheric-chemistry-database.html, last access: 7 July 2024).

The NA models' monthly average values of hourly near-surface SO_2(g) concentrations and daily $PM_{2.5}$ sulfate concentrations are compared to observations in Fig. 14. The monthly averages of daily (Canadian Air and Precipitation Monitoring Network, CAPMoN) and weekly (NADP) S wet deposition are shown in Fig. 15. Model–observation evaluation statistics are compared in Table S2 (Supplement). Sta-

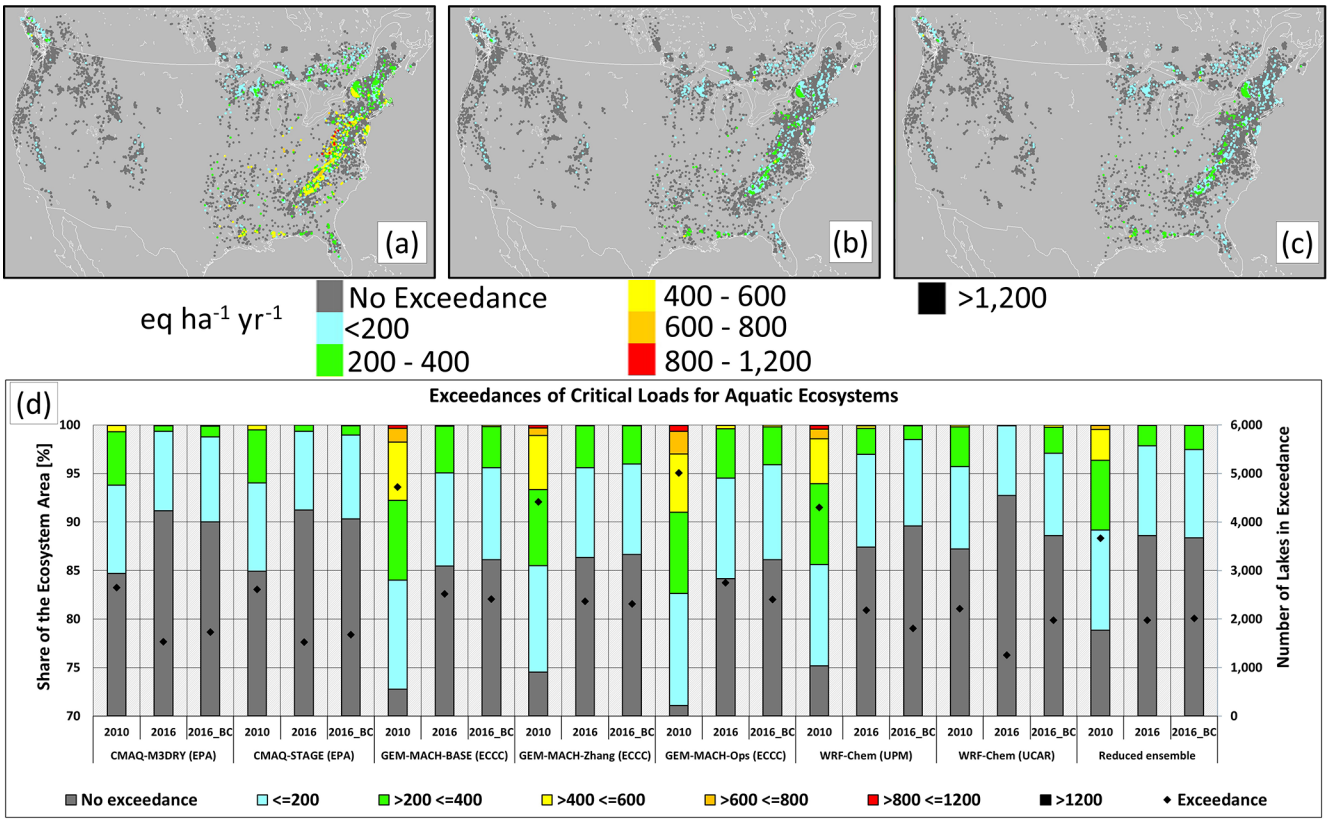

Figure 8. Summary CLEs for aquatic ecosystems, NA AQMEII4 common domain. **(a, b)** Spatial distribution of CLEs for the reduced ensemble for the years 2010 and 2016, respectively. **(c)** Spatial distribution of CLEs for the bias-corrected reduced ensemble for the year 2016. **(d)** Percentage of lakes for which CL data are available that are in exceedance by model and year (left axis and colour bar) and number of lakes in exceedance (right axis and black diamonds). Reduced-ensemble CLEs for the years 2010, 2016, and bias-corrected 2016 (2016_BC) appear on the right side of panel **(d)**.

tion locations for the observations are shown in Supplement Figs. S15, S16, and S17.

Table S2 shows that CMAQ-M3Dry and CMAQ-STAGE had the best values for most metrics, for the concentrations of SO_2 and $PM_{2.5}$ sulfate, and for daily sulfur wet deposition. CMAQ-M3Dry, CMAQ-STAGE, and WRF-Chem (IASS) had predominantly negative biases, and all other models had positive biases. The same tendency can be seen in Fig. 14a, where CMAQ-M3Dry and CMAQ-STAGE negative biases can be seen to occur in the warmer months, whereas WRF-Chem (IASS) has negative biases in the spring. Despite these differences, the net contribution of SO_2 dry-deposition flux to total sulfur deposition on an *annual* basis is relatively similar across the models (Table 4), with the standard deviation being relatively small, mostly driven by the SO_2 deposition flux for WRF-Chem (UPM) being higher than for the other models.

Particle sulfate (Fig. 14b and Table S2) values were also closest to monthly observed values for CMAQ-M3Dry and CMAQ-STAGE, while they were biased negative for WRF-Chem (IASS) and biased positive for the remaining models. The evaluation of total S wet deposition (Fig. 15a, Table S2) showed that all models with the exception of GEM-MACH (Ops) had negative biases relative to the Canadian daily S wet-deposition observations. Weekly S wet-deposition biases are also negative for most models (Table S2, Fig. 15b), with only GEM-MACH (Ops) having a positive bias in the ensemble.

Factors aside from emissions which affect the SO_2 concentrations within the models are the loss processes of gas oxidation, uptake into hydrometeor water (and subsequent in-cloud oxidation), and dry deposition. Both the gas oxidation and hydrometeor uptake pathways may lead to particulate sulfate formation (through nucleation/condensation of sulfuric acid into particles and through evaporation of hydrometeors). An underestimate of chemical conversion of SO_2 within hydrometeors may thus be expected to result in underestimates of particulate sulfate and in sulfate ion wet deposition. However, Table S2 shows relatively little bias for $PM_{2.5}$ sulfate relative to observations for CMAQ-M3Dry and CMAQ-STAGE and positive biases for the GEM-MACH models and WRF-Chem (UPM); these positive biases in predicted particulate sulfate would argue against an insufficient conversion of SO_2 to particulate sulfate in the latter group

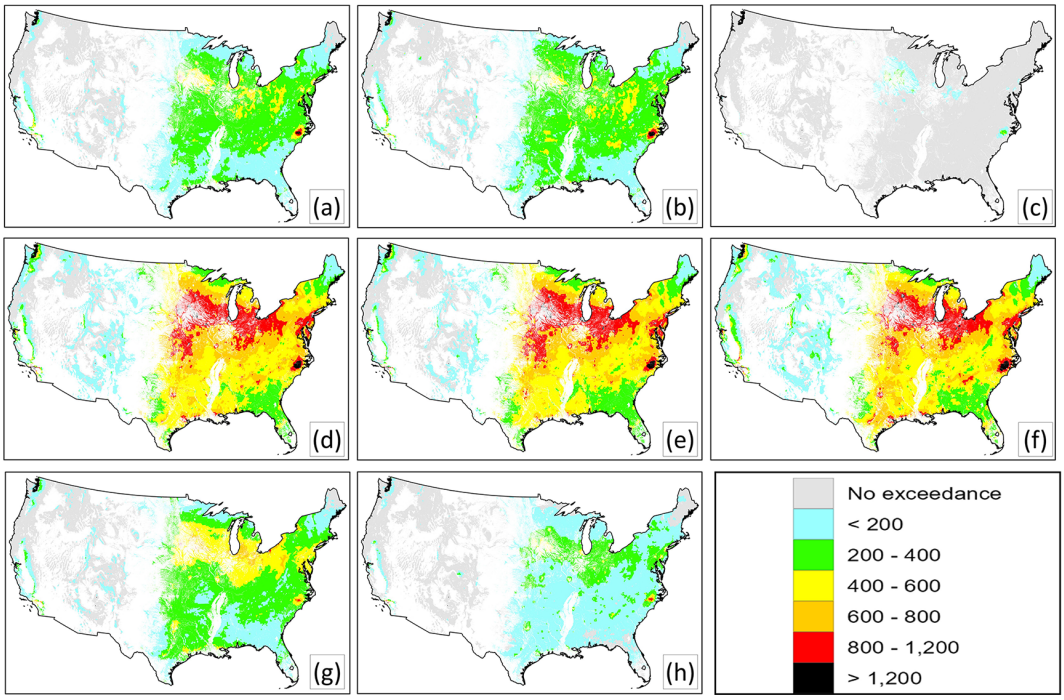

Figure 9. CLEs for sensitive-epiphytic-lichen species, NA AQMEII4 common domain, 2016, eq. $ha^{-1}\,yr^{-1}$. Panels arranged by model as in Fig. 6. Light-grey areas indicate regions for which critical load data were available but were not in exceedance of critical loads. Coloured areas indicate exceedance regions.

of models. Rather, the general tendency of negative biases in sulfur wet deposition may indicate insufficient hydrometeor scavenging and subsequent aqueous-phase oxidation of aerosols across all models. We also note that the mean bias of SO_2 concentrations for GEM-MACH (Ops) is more positive than those of GEM-MACH (Base) and GEM-MACH (Zhang), while the particulate sulfate bias was lower and the sulfate wet-deposition bias was higher. GEM-MACH (Ops) makes use of an operational weather forecast for cloud fields, while GEM-MACH (Base) and GEM-MACH (Zhang) make use of an explicit cloud microphysics scheme, which allows for weather–air quality feedbacks to be simulated but tends to underestimate the cloud amounts when used at lower resolution such as the 10 km grid cell size used in the simulations for these three models in this study. The differences between GEM-MACH (Base), GEM-MACH (Zhang), and GEM-MACH (Ops) may thus reflect weaker scavenging of aerosols into clouds in the Base and Zhang implementations.

GEM-MACH (Base), GEM-MACH (Zhang), and WRF-Chem (UCAR) have the most positive biases for particulate sulfate. As noted above, GEM-MACH (Base) and GEM-MACH (Zhang) share a common framework, and unlike other models in the ensemble, they also share an implementation of the updated particle deposition parameters of Emerson et al. (2020). The Emerson et al. (2020) implementation makes use of extensive measurement data and, compared to earlier parameterizations such as Zhang et al. (2001), results in decreased dry-deposition velocities for submicrometre particles and increased dry-deposition velocities for particles larger than 0.2 to 0.8 µm diameter, depending on land use type. The increased $PM_{2.5}$ SO_4 values in GEM-MACH (Base) and GEM-MACH (Zhang) in Fig. 14b may thus reflect decreases in the deposition removal flux in the submicrometre portion of the bins in this 12-bin sectional model framework. WRF-Chem (UPM) and WRF-Chem (UCAR) are also both sectional models making use of a common modelling framework, with WRF-Chem (UPM) being a slightly earlier release than WRF-Chem (UCAR). Neither model made use of the Emerson et al. (2020) update at the time the AQMEII4 simulations took place. However, this option was later examined for the WRF-Chem (UCAR) configuration by Ryu and Min (2022), who found that the Emerson et al. (2020) dry-deposition parameterization, applied following the runs carried out here, resulted in an increase in the positive $PM_{2.5}$ bias from $+4.5$ to $+6.7\,\mu g\,m^{-3}$ and a shift towards less negative biases in PM_{10} from -19.7 to $-1.77\,\mu g\,m^{-3}$, similar to the biases in particulate sulfate and ammonium observed in Fig. 14b between GEM-MACH (Base), GEM-MACH (Zhang), and GEM-MACH (Ops). Ryu and Min (2022) further found that the additional update of replacing the default Slinn (1984) aerosol cloud scavenging parameterization with the Wang et al. (2014) parameterization offset the increase in $PM_{2.5}$ SO_4 biases associated with the new particle dry-deposition scheme, illustrating the extent

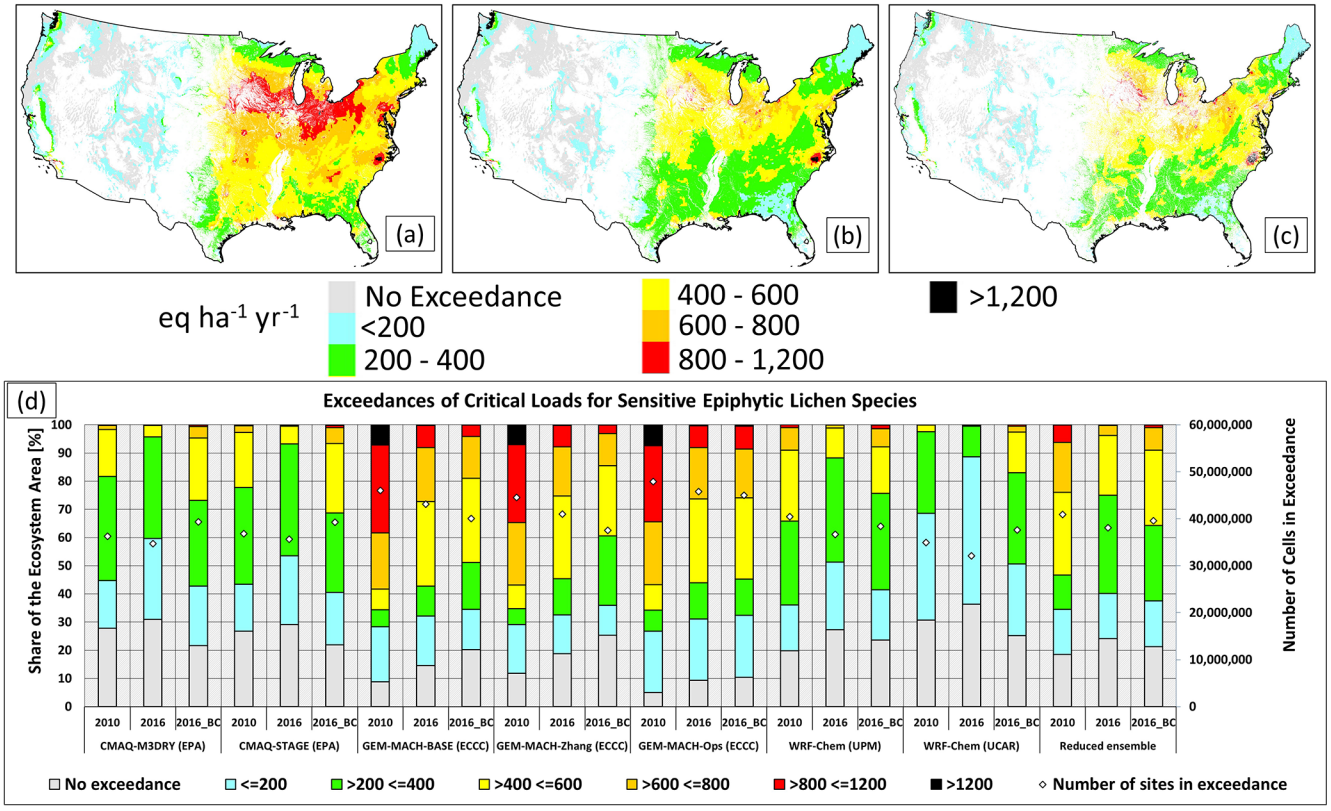

Figure 10. Summary CLEs, sensitive-epiphytic-lichen species, NA AQMEII4 common domain, eq. ha^{-1} yr^{-1}. **(a, b)** Spatial distribution of CLEs for the reduced ensemble for the years 2010 and 2016, respectively. **(c)** Spatial distribution of CLEs for the bias-corrected reduced ensemble for the year 2016. **(d)** Percentage of sensitive-epiphytic-lichen ecosystems for which CL data are available that are also are in exceedance by model and year (left axis and colour bar) and number of sites in exceedance (right axis and white diamonds). Reduced-ensemble CLEs for the years 2010, 2016, and bias-corrected 2016 (2016_BC) appear on the right side of panel **(d)**.

to which combinations of parameterizations are sometimes needed to improve model performance. More recent versions of GEM-MACH also make use of multiphase hydrometeor partitioning, with and without the Wang et al. (2014) semi-empirical scavenging scheme, with a significant increase in the uptake of particulate sulfate depending on precipitation rate and improvements in the wet sulfate performance relative to previous model versions (Ghahreman et al., 2024). Implementation of both updated particle dry-deposition velocities and wet-scavenging methodology have thus resulted in reduced biases for these fields, for several of the models examined here, in work following the simulations for AQMEII4.

With regards to sulfur wet deposition, Fig. 15a and Table S2 show a tendency of most models towards negative biases for total *daily* S wet deposition. However, this negative bias is much less pronounced or even positive in comparison to the *weekly* S wet-deposition data. Other metrics of model performance differed sharply between the two wet-deposition observation datasets for some metrics, with the weekly wet SO_4^{2-} deposition data comparison having higher mean gross error (MGE), normalized mean gross error (NMGE), and root mean square error (RMSE) values than the daily wet SO_4^{2-} deposition data comparison. The overall tendency of the performance was similar for both datasets, with the CMAQ models having the best scores for metrics other than mean bias. We note that the daily and weekly NA wet-deposition values correspond to monitoring networks in two different locations (see Fig. S15a). The daily values are from the Canadian CAPMoN network (stations in the AQMEII4 common domain are located mostly in south-eastern Canada), while the weekly data from the US NADP network are distributed throughout the USA. The differences in model performance may thus reflect regional differences in predicted meteorological and/or emission fields.

One possible cause for the negative biases in wet deposition common to most models could be underestimates in the amount of model-predicted precipitation, which in turn would reduce the wet flux. The net precipitation totals converted to liquid water for the eight NA models and observations are shown in Fig. S18, for both daily (CAPMoN) and weekly (NADP) monthly averages. While the monthly averages of daily precipitation (Fig. S18b) suggest a tendency towards negative biases in the summer months for some mod-

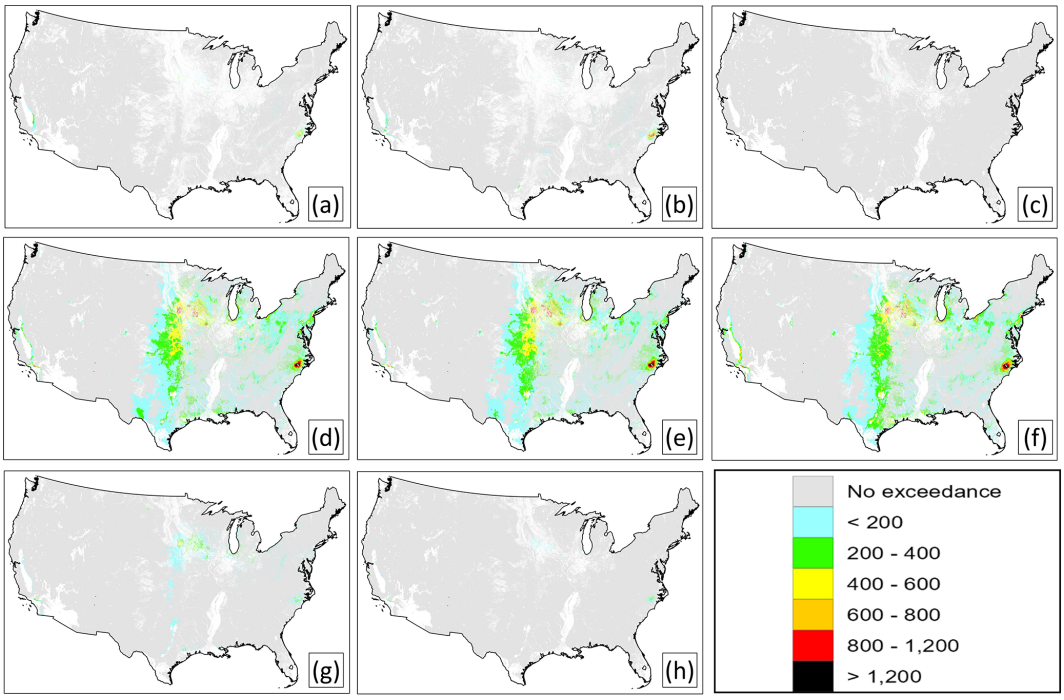

Figure 11. CLEs for a decline in herbaceous-species community richness, NA common domain, 2016, eq. $ha^{-1}\,yr^{-1}$. Panels arranged by model as in Fig. 6. Light-grey areas indicate regions for which critical load data were available but were not in exceedance of critical loads. Coloured areas indicate exceedance regions.

els, the time series of the precipitation biases does not follow that of the sulfate wet-deposition biases; for example, the difference relative to wet sulfate observations in Fig. 15a remains relatively constant for CMAQ-M3Dry and CMAQ-STAGE, while the predicted precipitation difference relative to observations for the same models in Fig. S18a shows more negative biases in the summer than wintertime. Model total precipitation biases thus do not appear to be a major contributing factor to the sulfur flux biases found in this work.

We also note the potential for the lower-magnitude biases in the daily wet SO_4^{2-} evaluation, compared to the weekly evaluation, to be the result of the respective regions represented by the two monitoring networks. Figure S16a shows that the daily data are derived from a smaller geographic area than the weekly data; hence regional performance differences may be affecting the two evaluation results.

Summary: North American S deposition variability

Sulfur deposition results from a complex balance between SO_2 oxidation, particulate sulfate formation, scavenging, and the release of particles within clouds, in addition to the processes governing deposition of each of the components. The largest contributing pathways to North American sulfur deposition, in descending order of importance, were wet deposition ($SO_4^{2-} + HSO_3^-$), particulate sulfate dry deposition, and dry SO_2(g) deposition in the reduced ensemble of model runs. The largest contributors to model-to-model variability in sulfur deposition, in descending order of importance, were particulate sulfate dry deposition, wet deposition ($SO_4^{2-} + HSO_3^-$), and dry SO_2(g) deposition.

CMAQ-M3Dry, CMAQ-STAGE, and GEM-MACH (Ops) had both the highest levels of wet deposition and also the best scores relative to wet-deposition observations. Models with higher $PM_{2.5}$ sulfate positive biases relative to observations also had stronger negative biases for sulfate wet deposition, indicating that the magnitude of particle scavenging in hydrometeors may play a role in both biases in the models. Comparisons between GEM-MACH (Base), GEM-MACH (Zhang), and GEM-MACH (Ops) provide some evidence for this effect. WRF-Chem (UPM) and WRF-Chem (UCAR) have very low particulate sulfate deposition fluxes relative to the other models and substantial positive biases in $PM_{2.5}$ sulfate and negative biases in sulfate wet deposition, relative to observations, likely related to insufficient wet scavenging of sulfate particles in hydrometeors (Ryu and Min, 2022)

3.2.2 Causes of N deposition variability in North American domain simulations

The common-grid spatial average and percent contribution of each of the species contributing to total annual N deposition for 2016 are given in Table 5. The columns in the table are arranged in descending order from left to right regarding contribution to the reduced-ensemble total nitrogen de-

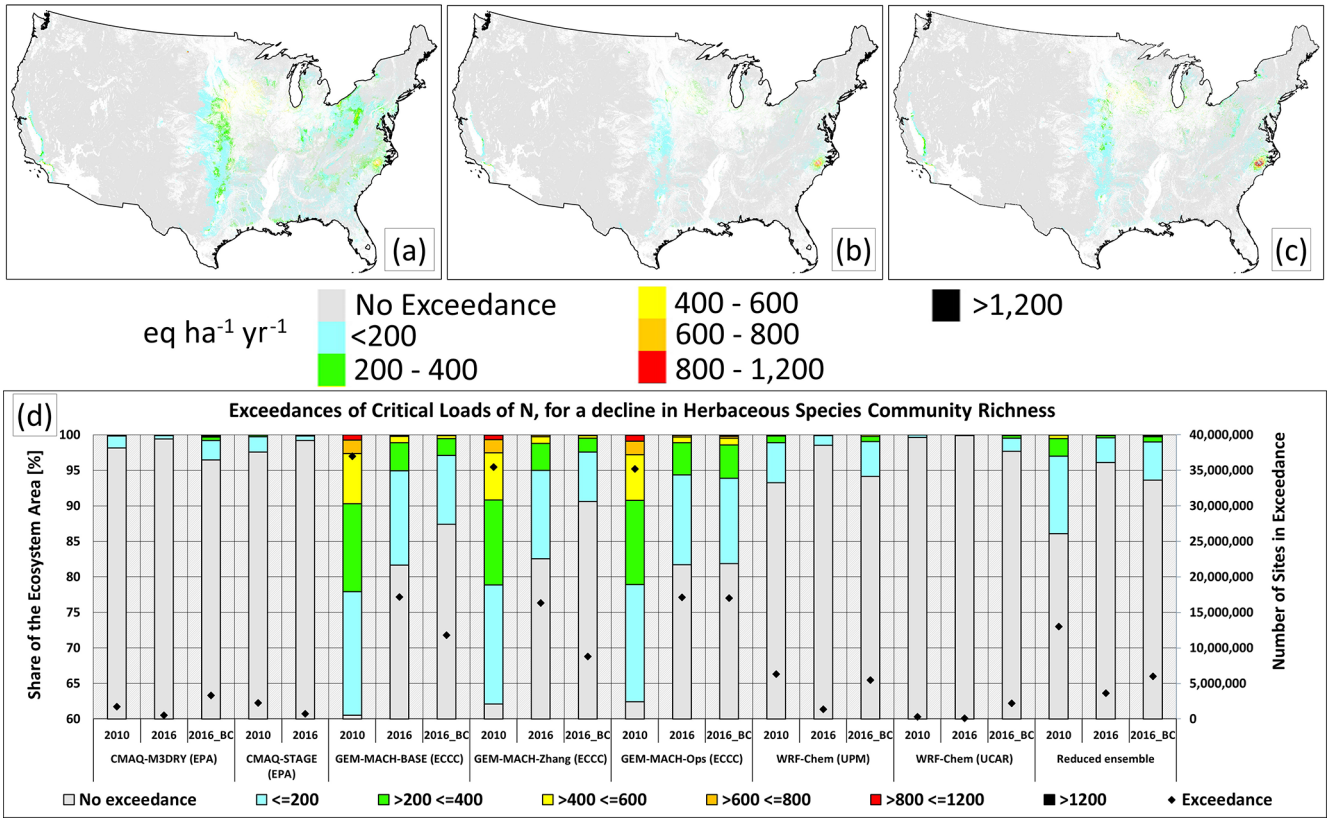

Figure 12. Summary CLEs for a decline in herbaceous-species community richness, NA AQMEII4 common domain, eq. ha^{-1} yr^{-1}. **(a, b)** Spatial distribution of CLEs for the reduced ensemble for the years 2010 and 2016, respectively. **(c)** Spatial distribution of CLEs for the bias-corrected reduced ensemble for the year 2016. **(d)** Percentage of herbaceous-species communities for which CL data are available that are also are in exceedance by model and year (left axis and colour bar) and number of sites in exceedance (right axis and white diamonds). Reduced-ensemble CLEs for the years 2010, 2016, and bias-corrected 2016 (2016 BC) appear on the right side of panel **(d)**.

position for each contributing chemical (row labelled "Red. ens. avg"). The impact of variability on the model deposition from each component for each model is once again shown as the standard deviation across the models used for the reduced ensemble (row labelled "Red. ens. SD"). From the row representing the standard deviation, it can be seen that the variation (standard deviation) between models for the contributions to total N deposition are driven, in descending order, by particle ammonium (DAM column, where the standard deviation for particle ammonium deposition is larger than the reduced-ensemble mean value), followed by wet ammonium ion (WNH4), wet nitrate ion (WNO3), dry HNO_3 (DHNO3), dry particle nitrate (DNI), dry NO_2 (DNO2), and dry ammonia gas (DNH3), with the remaining species contributing a small percentage of the total variability. Both the particle ammonium and wet ammonium variability between the models is largely driven by the GEM-MACH group of models, which have average dry particle ammonium and wet ammonium fluxes which are 17.4 and 1.76 times higher than the other models, respectively.

We next evaluate the models' nitrogen performance using the available concentration and wet-deposition flux data to determine the impact of the parameterization differences on model performance and hence identify which components in which models might be improved.

Dry deposition of particle ammonium

The largest source of variability between North American models' total N predictions resides in the dry particle ammonium deposition fluxes, with Table 5 showing that the standard deviation of this deposition flux across models was essentially as large as the reduced-ensemble average. Particle ammonium dry deposition contributes a disproportionately high contribution to total N *variability* across the North American ensemble, despite the magnitude of the ensemble average particle ammonium dry-deposition flux being less than the deposition of wet ammonium ion, dry nitric acid gas, or wet nitrate ion.

Figure 16 compares the monthly average $PM_{2.5}$ ammonium concentrations with observations (station locations appear in Fig. S15b), and Table S3 provides detailed statistics. From the latter, CMAQ-M3Dry and CMAQ-STAGE have the best overall performance for particulate ammonium

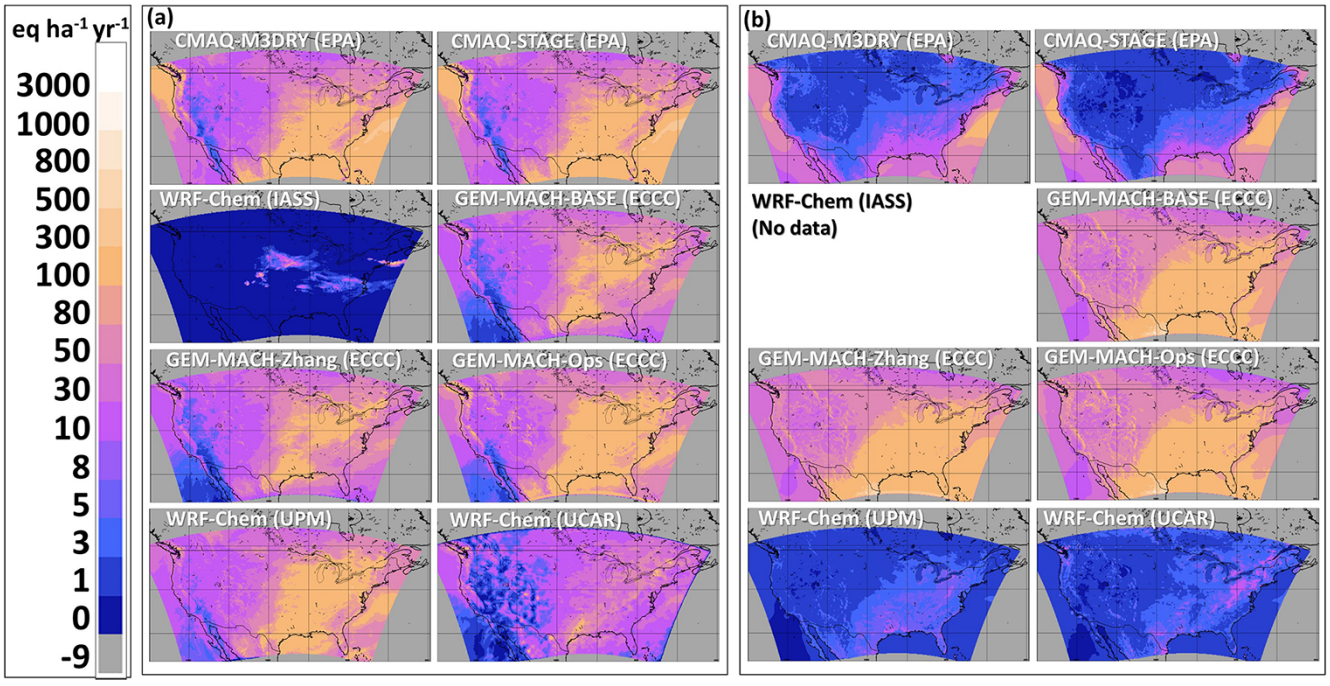

Figure 13. The 2016 total annual deposition flux (eq. ha^{-1} yr^{-1}) of **(a)** wet S and **(b)** dry particulate sulfate. Note that regions outside the AQMEII4 common domain have been assigned an outside-domain mask value of −9.

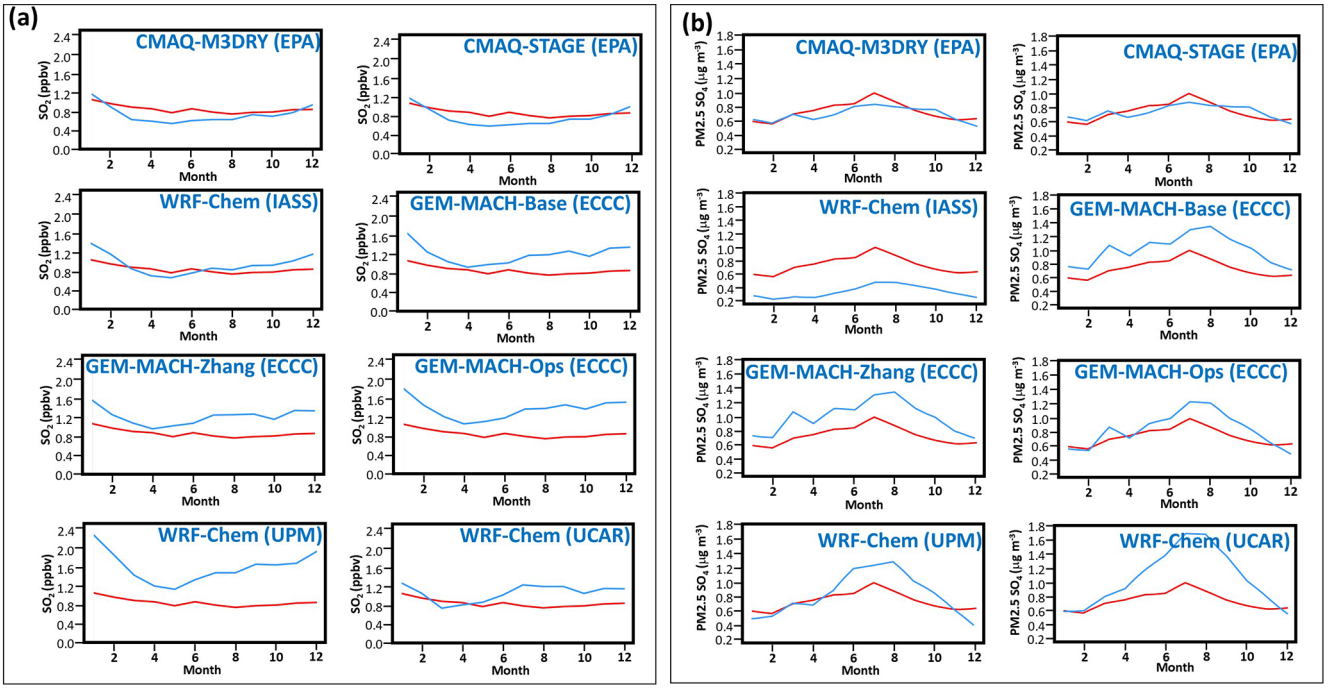

Figure 14. Comparison of model (blue line) and observed (red line) monthly average surface concentrations of **(a)** hourly SO_2 (ppbv) and **(b)** daily $PM_{2.5}$ sulfate (μg m^{-3}) for the year 2016 (AQS, NAPS data).

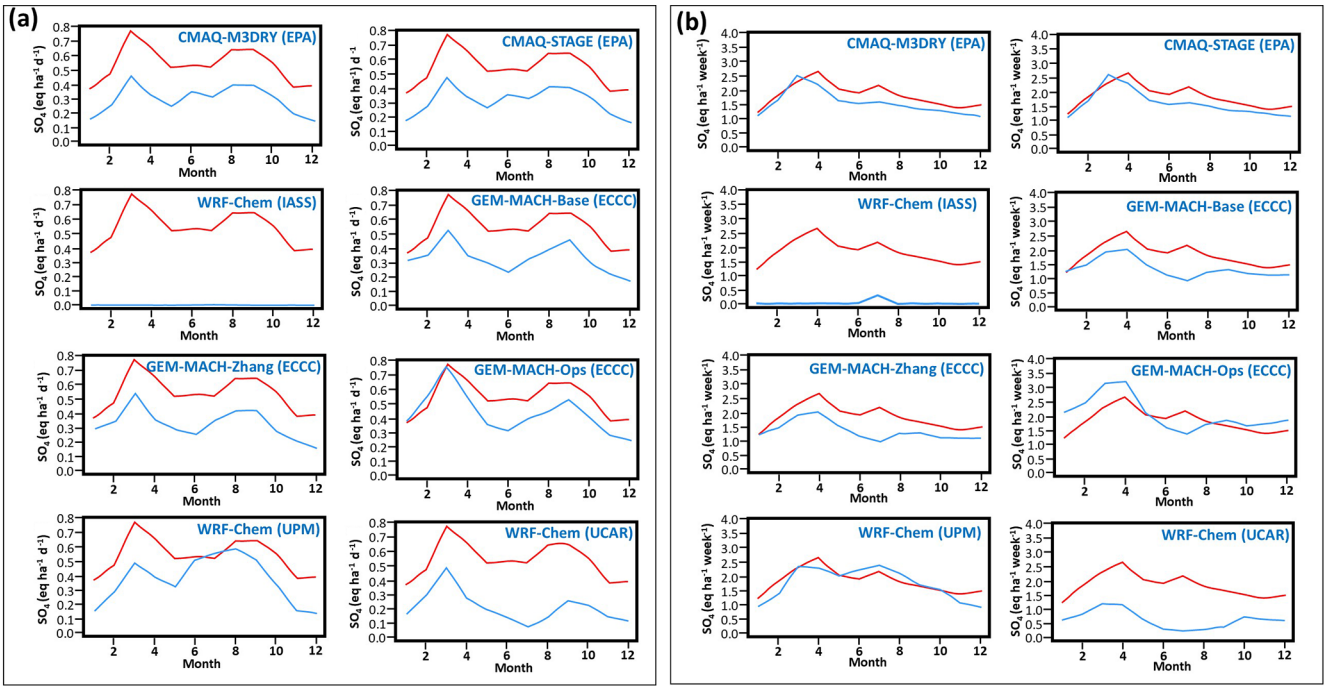

Figure 15. Comparison of model (blue line) and observed (red line) monthly average values of sulfur wet deposition for **(a)** daily CAPMoN data (eq. ha^{-1} d^{-1}) and **(b)** weekly NADP data (eq. ha^{-1} per week) for the year 2016.

and GEM-MACH (Base), GEM-MACH (Zhang) and GEM-MACH (Ops) have the worst performance by the statistical measures used here. This latter group of models also have the largest magnitude of positive biases relative to observed $PM_{2.5}$ ammonium concentrations, while the CMAQ implementations have negative biases and the remaining models have smaller-magnitude positive biases. Figure 16 shows that CMAQ-M3Dry, CMAQ-STAGE, WRF-Chem (IASS), and to a lesser extent WRF-Chem (UPM) have a greater seasonal variability in model particle ammonium (blue line) than observed (red line), with the difference between summer and winter (months 1 and 12 versus months 5 through 9) being higher in the models than in observations.

The GEM-MACH contributions to model N variability in critical load exceedances are thus linked to poor model performance for $PM_{2.5}$ ammonium. This poor performance is likely due to two factors, which can be deduced from comparing the process representations implemented in the models (Sect. 2.2).

The first factor, which differentiates GEM-MACH (Base), GEM-MACH (Zhang), and GEM-MACH (Ops) from the other ensemble members relates to how inorganic aerosol thermodynamic partitioning chemistry has been implemented: while the process representation in the models of the ensemble is derived from the ISORROPIA module (Nenes et al., 1998; Fountoukis and Nenes, 2007), the GEM-MACH implementations in AQMEII4 employ a partial speciation of SO_4^{2-}, NH_4^+, and NO_3^- (Makar et al., 2003) and do not in-

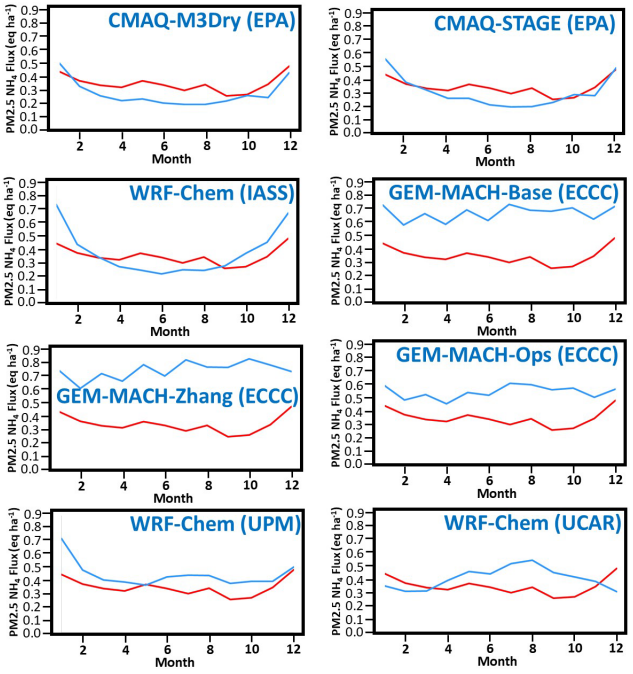

Figure 16. $PM_{2.5}$ ammonium compared to observations, North American model ensemble for the simulation year 2016. Red line: monthly observed average. Blue line: monthly model average.

Table 5. Contributions of N species to total deposition (eq. ha^{-1} yr^{-1}) and percent of total N deposited, over the NA AQMEII4 common grid, arranged in descending order of importance to the reduced-ensemble average. WNH4: wet deposition of NH_4^+(aq). DHNO3: dry deposition of HNO_3(g). WNO3: wet deposition of NO_3^-(aq). DAM: dry deposition of particulate ammonium. DNH3: dry deposition of NH_3(g). DNI: dry deposition of particulate nitrate. DNO2: dry deposition of NO_2(g). DPAN: dry deposition of peroxyacetyl nitrate gas. DRN3: dry deposition of gaseous organic nitrate gases. DN2O5: dry deposition of N_2O_5(g). DHNO4: dry deposition of pernitric acid gas. DNO: dry deposition of NO(g). WRF-Chem (IASS) did not report dry particle fluxes. The GEM-MACH models and WRF-CHEM (UPM) do not include dry deposition of N_2O_5(g), and the GEM-MACH models do not dry deposit HNO_4(g). n/d: no data.

Average (eq. ha^{-1} yr^{-1})										
	Model									
Species	CMAQ-M3Dry	CMAQ-STAGE	WRF-Chem (IASS)	GEM-MACH (Base)	GEM-MACH (Zhang)	GEM-MACH (Ops)	WRF-Chem (UPM)	WRF-CHEM (UCAR)	Red. ens. avg	Red. ens. SD
WNH4	51	60.4	0.2	129	129	114.2	64.3	29.4	82.5	37.7
DHNO3	52.5	51.9	0	66.9	56.2	62.4	75.1	46.8	58.8	9.1
WNO3	65.6	66.9	0.2	45	51.3	71.9	73.1	33.6	58.2	14
DAM	8.5	8.4	n/d	98.5	100.7	82.6	2.7	2	43.3	44.2
DNH3	33.2	29.5	36.3	26.9	26.6	40	40.3	47.2	34.8	7.3
DNI	18.3	18.9	n/d	26.8	32.7	19	7.6	7.1	18.6	8.6
DNO2	7.9	7.3	7.7	23.8	21.9	26.7	10.9	10.8	15.6	7.6
DPAN	4.9	4.7	2	7.7	7.4	10	2.7	2	5.6	2.7
DRN3	6.6	4.9	0.4	1.8	2.4	3.1	0.7	3.1	3.2	1.8
DN2O5	1.2	1.1	2.2	n/d	n/d	n/d	n/d	n/d	1.2	0.1
DHNO4	0.4	0.1	0	n/d	n/d	n/d	0.8	0.4	0.3	0.1
DNO	0.5	0.5	0	0.1	1.2	0.2	0	0	0.4	0.4
Total N	250.7	254.7	49	426.5	429.4	430	278.2	182.4	321.7	96.5
Percent contribution										
	Model									
Species	CMAQ-M3Dry	CMAQ-STAGE	WRF-Chem (IASS)	GEM-MACH (Base)	GEM-MACH (Zhang)	GEM-MACH (Ops)	WRF-Chem (UPM)	WRF-CHEM (UCAR)	Red. ens. avg	Red. ens. SD
WNH4	20.4	23.7	0.4	30.2	30	26.5	23.1	16.1	25.6	4.7
DHNO3	21	20.4	0	15.7	13.1	14.5	27	25.7	18.3	5
WNO3	26.2	26.3	0.3	10.6	11.9	16.7	26.3	18.4	18.1	6.4
DAM	3.4	3.3	n/d	23.1	23.5	19.2	1	1.1	13.5	9.9
DNH3	13.2	11.6	74.2	6.3	6.2	9.3	14.5	25.9	10.8	7.6
DNI	7.3	7.4	n/d	6.3	7.6	4.4	2.7	3.9	5.8	1.8
DNO2	3.2	2.9	15.8	5.6	5.1	6.2	3.9	5.9	4.9	1.3
DPAN	1.9	1.9	4.1	1.8	1.7	2.3	1	1.1	1.7	0.5
DRN3	2.6	1.9	0.7	0.4	0.6	0.7	0.2	1.7	1	0.8
DN2O5	0.5	0.4	4.4	n/d	n/d	n/d	n/d	n/d	0.4	0
DHNO4	0.2	0	0	n/d	n/d	n/d	0.3	0.2	0.1	0.1
DNO	0.2	0.2	0.1	0	0.3	0	0	0	0.1	0.1
WNH4	20.4	23.7	0.4	30.2	30	26.5	23.1	16.1	25.6	4.7

clude the reactions involving particulate base cations (Ca^{2+}, Mg^{2+}, Na^+, K^+). The other models in the ensemble do include these additional reactions. In the absence of base cation chemistry, the formation of particle ammonium will be controlled by the availability of ammonia gas in excess of that required to charge balance particulate sulfate, as well as by the availability of nitric acid gas. In the presence of base cations, nitric acid gas will preferentially associate with base cations rather than ammonia, leaving less HNO_3 available for particle ammonium nitrate formation. Several observational studies have shown that when base cations are present, their peak mass occurs in the coarse particle size mode (>2.5 µm diameter), where they will have higher deposition velocities (e.g. inland, agricultural dust sources (Makar et al., 1998); ocean sources of sea-salt (Anlauf et al., 2006)). Base cation inorganic heterogeneous chemistry thus provides a competing pathway for the uptake of nitrate into particles and, when present, will also reduce the amount of NH_3 that may be taken up by particles, especially in the fine mode. The positive bias of $PM_{2.5}$ ammonium in Fig. 16 for GEM-MACH relative to the other models likely represents the impact of simplified inorganic aerosol chemistry.

The second factor influencing the GEM-MACH models' positive particulate ammonium biases may be reflected in the biases for GEM-MACH (Base) and GEM-MACH (Zhang), which are higher by 50 % to a factor of 2 than that of GEM-MACH (Ops), respectively; that is, an additional source of bias resides in the former two model implementations that is not present in the latter implementation. The likely source of this additional bias is the use of Emerson et al. (2020) particle deposition velocities in these implementations, in the absence of enhanced wet scavenging of aerosols, as discussed above for $PM_{2.5}$ sulfate and described in Ryu and Min (2022) and Ghahreman et al. (2024). Ryu and Min (2022) showed that the use of the updated particle deposition velocity as per Emerson et al. (2020), when implemented in the absence of

concurrent multiphase wet-scavenging updates, led to positive biases in $PM^{2.5}$ concentrations in the WRF-Chem model.

We note that the manner in which inorganic heterogeneous chemistry is simulated also differs between the models. CMAQ-M3Dry and CMAQ-STAGE calculate local equilibrium concentrations at different modes of the size distribution, and WRF-Chem (UPM) and WRF-Chem (UCAR) also calculate the equilibrium with respect to specific size bins, while GEM-MACH (Base), GEM-MACH (Zhang), and GEM-MACH (Ops) carry out a single bulk calculation across all size bins. The use of a bulk calculation is a third simplification for the latter group of models and may also affect the particulate ammonium performance of these models.

The spatial distribution of $PM_{2.5}$ ammonia biases was examined in Fig. 17, for the month of July 2016 (July was chosen due to the expectation that bidirectional fluxes would have a higher impact in the summer months). The region with the highest positive biases (dark-red circles, Fig. 17) are in the same station locations for all models, in the agricultural region to south of the Great Lakes. Positive $PM_{2.5}$ ammonium mean biases also occur near urban regions in the western USA (Seattle–Tacoma, Yakima, Portland, Sacramento, San Jose, Boise, Butte, Helena, Denver, Boulder, and Albuquerque) and at one eastern site, Miami. A re-examination of ammonia gas deposition and emission parameters and primary particle ammonium emission inventories is recommended for these locations, given that they are likely having a large impact on model performance statistics. The CMAQ models and WRF-Chem (IASS) have negative to minimal biases along the coastlines and southwestern USA (regions of sea-spray NaCl and windblown base cation containing dust, respectively), while WRF-Chem (UPM) and WRF-Chem (UCAR) have small negative to positive biases in these regions and the GEM-MACH models are uniformly biased positive in these regions. This provides support for the possibility that the GEM-MACH positive bias in particulate ammonium concentrations is due to missing particulate base cation chemistry; the regions where particulate base cations would be expected to contribute significantly to total particulate mass are also the regions where the GEM-MACH models have positive biases, and the biases in the other model biases are not as significant.

Wet deposition of ammonium and nitrate ions

Wet deposition of ammonium ion is the largest contributor to the North America reduced-ensemble N_{dep} and the second largest contributor to model-to-model variability in N deposition (Table 5). Wet deposition of nitrate ion is the third largest contributor to both the NA ensemble total N deposition and model-to-model variability in N deposition. Time series of the monthly averages of observed and modelled daily (CAPMoN) and weekly (NADP) NH_4^+ wet-deposition fluxes are shown in Fig. 18. The monthly mean of modelled daily values (Fig. 18a) are generally biased negative, with the exceptions of the months of July and August for GEM-MACH (Base) and GEM-MACH (Zhang). The observed maximum in NH_4^+ wet deposition occurs in April (Fig. 18a, red line, month 4); this seasonal variation is captured only by GEM-MACH (Ops) and WRF-Chem (UCAR), with the other models predicting peak deposition in June through August. The monthly average of the weekly NH_4^+ wet-deposition fluxes (Fig. 18b) shows a similar pattern, with the observed values (red lines, Fig. 18b) peaking in April and all of the models except for WRF-Chem (UCAR) peaking in June. As was the case for sulfate wet deposition, the observed seasonal variation is apparently not connected with biases in precipitation predictions (see Fig. S18a and b and the Supplement), with the possible exception of WRF-Chem (UCAR), for which total precipitation is biased substantially negative throughout the year.

As noted above, the models taking part in this ensemble did not make use of multiphase hydrometeor scavenging in precipitation. The maximum NH_4^+ wet-deposition negative bias in April featuring for several models may reflect the absence of this level of detail in hydrometeor scavenging, with the absence of snow scavenging potentially impacting early-spring deposition. We note that the weekly and daily monitoring networks cover different geographical regions, hence the differences in model performance relative to the two observation datasets (compare the CAPMoN and NADP station locations in yellow and green circles in Fig. S15a, respectively).

The mean biases in average daily and weekly NH_4^+ wet deposition for the month of April are shown in Fig. 19. WRF-Chem (IASS), CMAQ-M3Dry, and CMAQ-STAGE have predominantly negative biases throughout the region; WRF-Chem (UCAR) and WRF-Chem (UPM) have a few stations with more positive biases; and the GEM-MACH models have both positive and negative biases throughout the domain. Insight into the differences in model performance can be gained through reviewing the manner in which each model parameterizes aerosol activation and scavenging:

1. GEM-MACH (Base), GEM-MACH (Zhang), GEM-MACH (Ops), WRF-Chem (UPM), and WRF-Chem (UCAR) make use of the aerosol activation scheme of Abdul-Razzak and Ghan (2000) and the Slinn (1984) approach to aerosol scavenging.

2. In GEM-MACH (Ops), the aerosol activation and aerosol-scavenging schemes are decoupled from meteorological feedbacks, while GEM-MACH (Base), GEM-MACH (Zhang), WRF-Chem (UPM), and WRF-Chem (UCAR) are "aerosol-aware" and full feedback models incorporating parameterizations for the aerosol direct and indirect effects. The latter will result in cloud formation from model-produced aerosols acting as cloud-condensation nuclei; clouds are more likely to form where aerosol concentrations are high (and thus more likely to scavenge aerosols below the clouds as well),

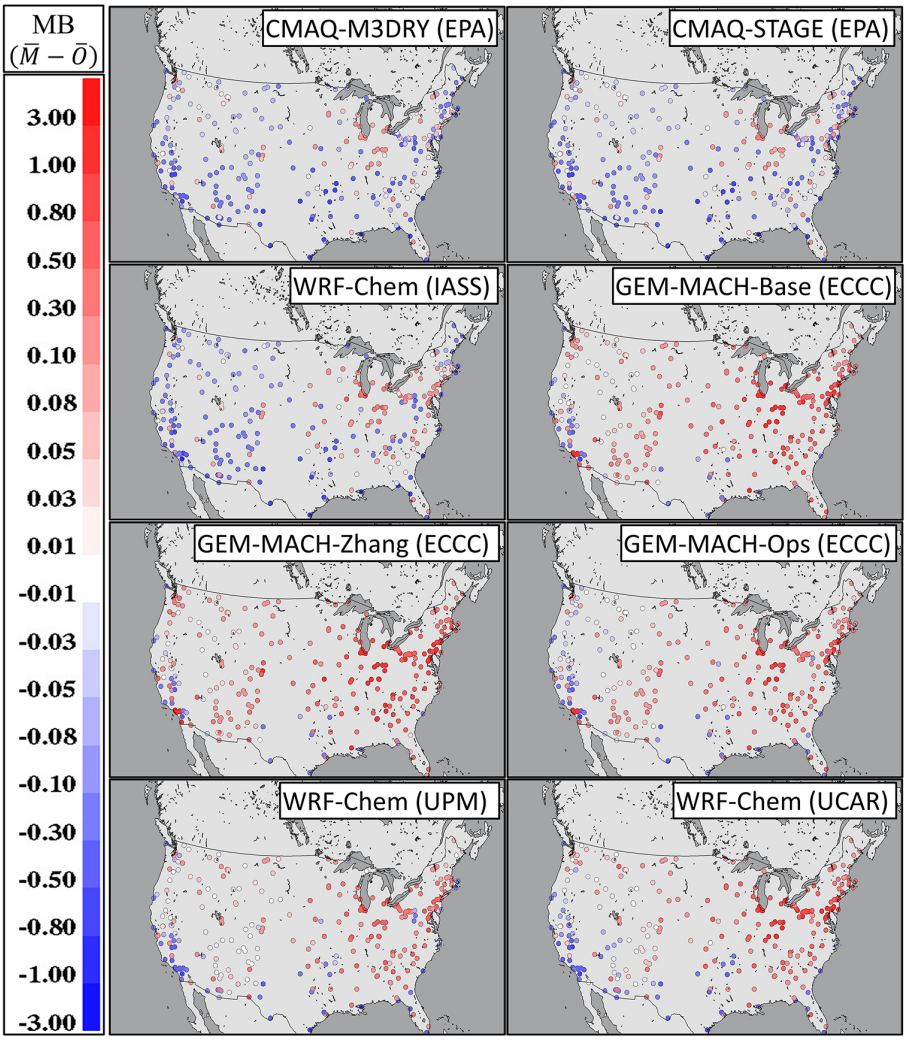

Figure 17. Mean bias (MB $= \overline{M} - \overline{O}$, where $\overline{M}$ is the model mean and $\overline{O}$ is the observed mean), $PM_{2.5}$ NH_4, July 2016, by station (µg m^{-3}). Negative values are given in blue; positive biases are given in red. Note that the colour scale is logarithmic.

compared to offline models. Very high aerosol concentrations may also reduce cloud droplet size and cloud-to-precipitation conversion, potentially making clouds more persistent, while reducing precipitation.

3. WRF-Chem (IASS) also makes use of aerosol direct- and indirect-effect feedbacks but employs the approach of Chapman et al. (2009) for aerosol scavenging.

4. CMAQ-M3Dry and CMAQ-STAGE are offline models (no feedbacks between aerosols, cloud formation, and radiative transfer takes place), where interstitial and nucleation aerosol scavenging by cloud droplets is modelled following Binkowski and Roselle (2003), and the wet-deposition rate is a simple parameterization dependent on the cloud total liquid water content, cloud thickness, and cloud precipitation rate (Fahey et al., 2017).

The Slinn (1984) aerosol-scavenging approach makes use of different observation-based aerosol collection efficiency formulae for rain and snow, respectively, where temperature dependence in the collection efficiency such as at 0 °C may be used to distinguish between liquid and solid hydrometeor collection efficiencies. Following the AQMEII4 simulations carried out here, parameterizations that utilize multiphase precipitation data with multiple hydrometeor classes, such as those of Wang et al. (2014), have been tested within the modelling framework of GEM-MACH (Ghahreman et al., 2024). Similarly, Ryu and Min (2022) describe the impact of multiphase hydrometeor scavenging as implemented in the WRF-Chem modelling framework. These tests resulted in significant improvements in particulate concentrations and wet deposition compared to previous implementations employing the approach of Slinn (1984). The approach for scavenging in Binkowski and Roselle (2003) assumes scavenging only

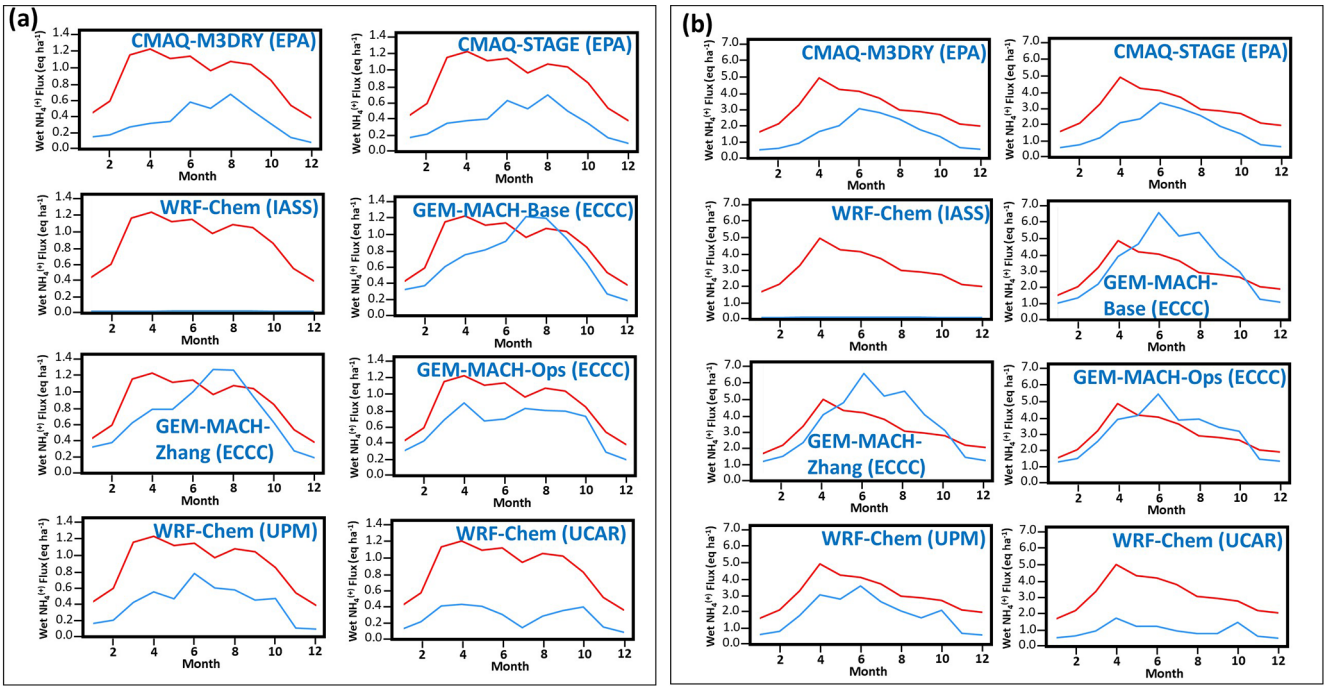

Figure 18. Time series of monthly average observed (red line) and modelled (blue line) ammonium wet-deposition fluxes for **(a)** daily CAPMoN data (eq. ha^{-1} d^{-1}) and **(b)** weekly NADP data (eq. ha^{-1} per week).

occurs in cloud droplets; snow scavenging is not considered. However, snow scavenging at higher precipitation rates is known to be 1 to 2 orders of magnitude more efficient than scavenging by rain. Hence the use of the Slinn (1984) parameterization instead of multiphase hydrometeor scavenging and the Wang et al. (2014) parameterization in GEM-MACH, as well as the omission of multiphase hydrometeor scavenging in CMAQ, may account for the springtime bias in all models noted here.

The causes of the differences in wet deposition of NH_4 between WRF-Chem (IASS), WRF-Chem (UPM), and WRF-Chem (UCAR) may be the use of the Chapman et al. (2009) wet-scavenging approach in the first model and the implementation of the Abdul-Razzak and Ghan (2000) and the Slinn (1984) approaches in the latter two models. All three models make use of the Morrison two-moment cloud microphysics scheme (Morrison et al., 2009), though WRF-Chem (IASS) and WRF-Chem (UPM) differ from WRF-Chem (UCAR) in the parameterization of convective clouds (See Table 1). Differences in aerosol-scavenging implementations may account for some of the differences in ammonium wet deposition between these models, as may the manner in which convective clouds identify cloud condensation nuclei from aerosol size distribution and speciation within their convective parameterizations.

Wet nitrate ion deposition is the third largest source of N deposition in the North American ensemble as well as the third largest source of model-to-model variability (Table 5). CMAQ-M3Dry, CMAQ-STAGE, and GEM-MACH (Ops) have the best performance scores for nitrate wet deposition (Table S3 in the Supplement). GEM-MACH (Base) and GEM-MACH (Zhang) have larger-magnitude and more negative biases than GEM-MACH (Ops), despite all three models making use of the same modelling framework. The only difference between GEM-MACH (Base) and GEM-MACH (Zhang) is the gas-phase dry-deposition algorithm employed (see Table 2). The increase in the wet-deposition negative bias magnitude going from GEM-MACH (Zhang) to GEM-MACH (Base) in Table S3 (from -0.19 to -0.26 eq. ha^{-1} d^{-1} for daily CAPMoN data and from -0.41 to -0.64 eq. ha^{-1} d^{-1} for weekly NADP data) is therefore attributable to gas-phase deposition differences. This is also reflected in the HNO_3 dry-deposition flux for the two models in Table 5, with the deposition flux for GEM-MACH (Base) at 66.9 eq. ha^{-1} yr^{-1} being 19 % higher than the GEM-MACH (Zhang) value of 56.2 eq. ha^{-1} yr^{-1}.

The remainder of the difference in nitrate wet-deposition bias between GEM-MACH (Base) and GEM-MACH (Zhang) and GEM-MACH (Ops) must be due to other factors in the model configuration as described in Table 2. Based on the $PM_{2.5}$ sulfate and $PM_{2.5}$ nitrate evaluations (Tables S2, S3), as well as the work of Ghahreman et al. (2024) and Ryu and Min (2022), we believe that the cause of the additional wet nitrate negative bias resides in the use of the new particle deposition velocity algorithm in the absence of a simultaneous update in the wet-deposition algorithm to make use of

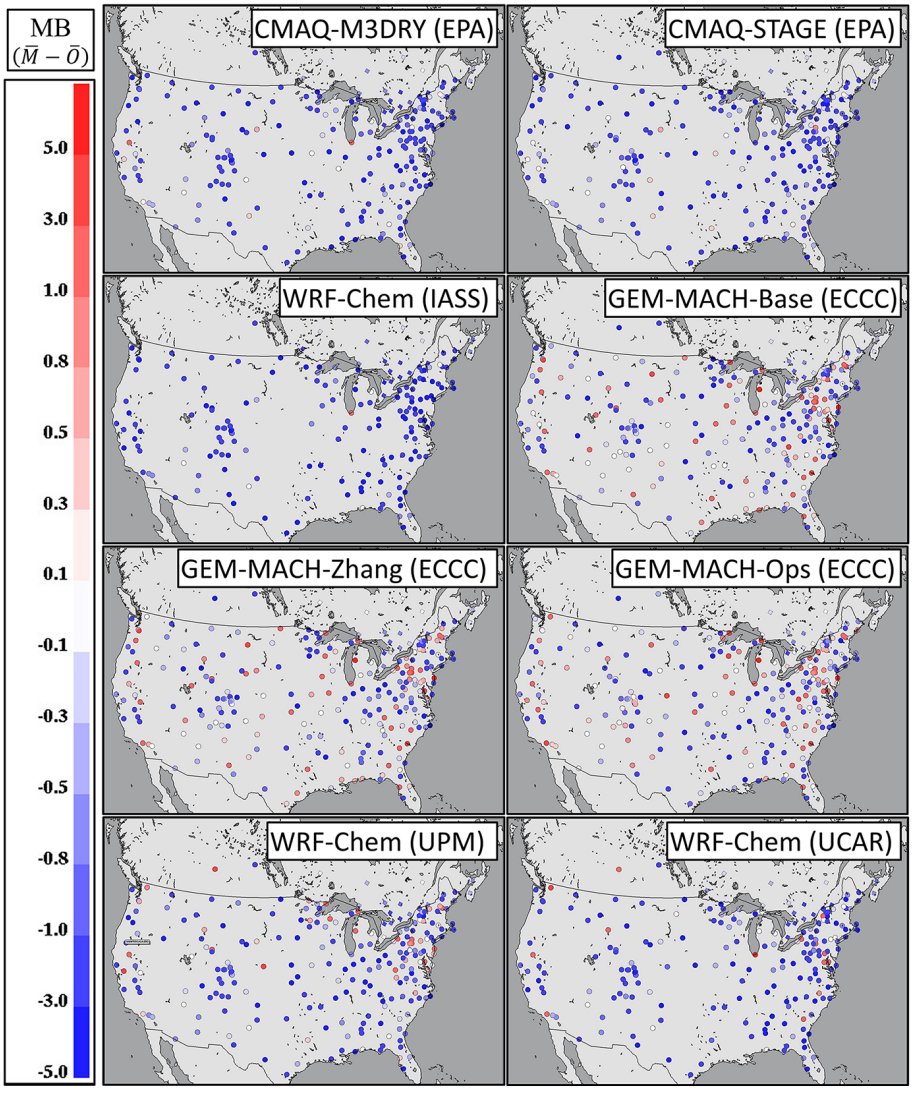

Figure 19. Model mean biases in ammonium wet deposition for the month of April 2016, North America (eq. ha^{-1} yr^{-1}). Daily station values of the mean bias (CAPMoN network) are shown as diamonds; weekly station values (NADP network) are shown as circles. Positive biases are shown in red; negative biases are shown in blue. Note that the colour scale intervals are logarithmic.

multiphase hydrometeor scavenging of aerosols. For example, the particulate matter scavenging coefficients for snow are 1 to 2 orders of magnitude more efficient than for rain – including snow scavenging (which may occur at higher elevations even in the summer) will lead to greater uptake of particles (Ghahreman et al., 2024). The Emerson et al. (2020) parameterization will lead to less particle deposition in submicrometre particle sizes (and hence would otherwise increase $PM_{2.5}$ concentrations – the increased scavenging associated with multiphase hydrometeors will offset this effect).

Dry deposition of HNO_3

Dry deposition of HNO_3 is the second largest source of N_{dep} in the reduced ensemble and the fourth largest source of model-to-model variability.

The spatial variation in the annual sum of the effective deposition fluxes for HNO_3 dry deposition are shown in Figs. S19, S20, S21, and S22, representing the mass of HNO_3 transferred to the surface via the cuticle, soil, stomatal, and lower-canopy pathways, respectively, and are summarized as common-grid totals in Fig. 20. Effective fluxes build on the concept of effective conductance: the product of the hourly deposition flux with the ratio of specific pathway conductance to total deposition velocity, for each of the four pathways (Galmarini et al., 2021). The figures thus de-

pict the contributions of each pathway to the HNO_3 dry-deposition mass flux for each model[1]. Effective fluxes incorporate changes in the flux resulting from changes in chemical concentration associated with factors in addition to deposition. However, comparison of the effective-flux values of Fig. 20 to effective conductances (not shown) has a similar pattern, implying that the deposition velocity is the dominating factor in the HNO_3 deposition flux. The HNO_3 mass flux is dominated by the cuticle pathway (Figs. S19, 20), followed by the soil pathway (Figs. S20, 20). All models show a similar pattern in HNO_3 annual cuticle flux (largest fluxes in the southeastern USA, lowest fluxes over the western mountain ranges and the Canadian boreal forest), though the magnitudes of the fluxes vary, with WRF-Chem (UPM) having the highest flux and GEM-MACH (Zhang) showing much lower fluxes for specific land use types over the western mountains compared to the other models.

The HNO_3 dry-deposition velocity parameterizations in the GEM-MACH models depend in part on deposition pathway parameterizations employing functions of the ozone and sulfur dioxide pathway values (Makar et al., 2018; Zhang et al., 2003). Other recent AQMEII4 work for ozone dry deposition using an observation-driven single-point modelling framework (Clifton et al., 2023) found that the ozone deposition velocity for GEM-MACH (Base) has positive biases in the summer months (average across eight sites of +73 %) and negative biases in the winter months (eight-site average of −33 %), while GEM-MACH (Zhang) has smaller summer biases (+3 %) and high winter biases (+50 %). This is consistent with the increase in HNO_3 dry-deposition flux going from GEM-MACH (Zhang) to GEM-MACH (Base), though HNO also deposits via dissociation (sulfur dioxide pathway); not all of the observed effects can be attributed to the use of O_3 as a proxy in part of the deposition algorithm. A portion of the increase in the negative bias in nitrate wet deposition going from GEM-MACH (Zhang) to GEM-MACH (Base) is thus the result of the higher HNO_3 dry-deposition removal of the available nitrate which would otherwise be taken up into clouds.

NH_3 and the role of bidirectional flux algorithms

NH_3 deposition fluxes were the fifth largest driver of ensemble nitrogen deposition and the seventh largest driver of N_{dep} variability in North America. Two different observation datasets for the year 2016 were used to evaluate model NH_3 concentration performance and cross-track infrared sounding (CrIS) satellite retrievals of NH_3 (see the Supplement for retrieval procedure and references) and AMoN (Chen et al., 2014; NADP, 2025b) surface monitoring network observations (see Supplement Fig. S16b for AMoN measurement locations). The two datasets evaluate model NH_3 performance in different ways. The CrIS observations (and model values extracted for evaluation) correspond to the specific time of day of the satellite overpass, for the polar-orbiting platform upon which the CrIS instrument is based. The evaluation against CrIS data is thus a measure of the model performance at early afternoon local time. The AMoN observations in contrast are 2-week integrated average concentrations; the AMoN comparison evaluates average model performance on this integrated timescale and hence includes into that average diurnal variations in NH_3 concentrations not available in the CrIS observations.

The evaluation of the models' NH_3 against CrIS observations at overpass time is shown in Table S4 (Supplement) and Fig. 21. The general trend for the models is one of negative biases in NH_3 concentrations. CMAQ-M3Dry and CMAQ-STAGE have the largest negative NH_3 biases, lowest number of model values within a factor of 2 of the observations (FAC2), highest MGE, lowest R, lowest coefficient of efficiency (COE), and lowest index of agreement (IOA) scores in Table S4. This suggests that the magnitude of the fluxes and/or the balance between positive (downward; deposition) and negative (upward; emission) fluxes for CMAQ-M3Dry and CMAQ-STAGE are the cause of the models' relatively poor performance for NH_3. GEM-MACH (Base) and GEM-MACH (Zhang) have the smallest (and positive) biases compared to the other models, and these two models as well as WRF-Chem (UPM) and WRF-Chem (UCAR) have the best overall scores for NH_3 against satellite data.

The satellite data comparison of Fig. 21 also shows some significant differences between observed ammonia and all models' predicted ammonia, particularly over waterbodies (oceans, Great Lakes), with observed NH_3 in the range 1–3 ppbv in the Atlantic and near Baja California, while the models all show NH_3 over the oceans always below 0.3 to 0.5 ppbv and decreasing with increasing distance from the shoreline. All models reach 0.0–0.01 ppbv at the greatest distances from the shoreline, while the satellite observations are above 0.5 ppbv (lower detection limit of ∼ 0.3 ppbv) throughout the AQMEII4 common domain.

NH_3 emissions from natural sources have been a source of ongoing interest in the global modelling community due to its properties as a greenhouse gas. Paulot et al. (2015) reviewed estimates of global oceanic NH_3 emissions, with a range of 7–23 $Tg\,N\,yr^{-1}$ and their own estimate being lower at 2.5 $Tg\,N\,yr^{-1}$. Their estimated maps of NH_3 emissions showed relatively lower values on the western shoreline of North America (Pacific coast) than on eastern shoreline (Atlantic coast) and high emissions in three out of the four oceanic NH_3 flux models tested, in the Gulf of Mexico and along the Gulf Stream between North America and Europe (their Fig. 3). Subsequent simulations of oceanic outgassing (Paulot et al., 2020) showed oceanic out-

[1] Note that the CMAQ-M3Dry and CMAQ-STAGE models incorporate the lower-canopy pathway into the soil pathway; the lower-canopy effects are not *absent* in these models but form part of the soil pathway, and hence they are reported here as part of the soil pathway.

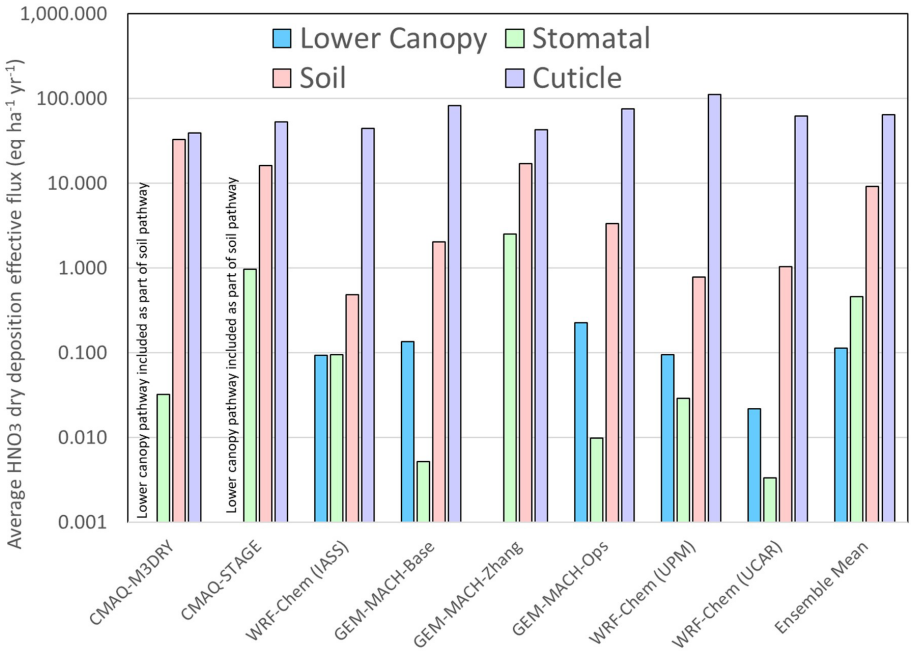

Figure 20. Averages of flux pathway contributions to HNO_3 dry deposition, NA AQMEII4 common grid, 2016 (eq. ha^{-1} yr^{-1}).

gassing in the Gulf of Mexico in excess of 0.03 g N m^{-2} yr^{-1} (17.6 eq. ha^{-1} yr^{-1}) and between 0.01 and 0.02 g N m^{-2} yr^{-1} (5.9 to 11.8 eq. ha^{-1} yr^{-1}) in the Gulf Stream. The oceanic emission model of Paulot et al. (2020) would be relatively straightforward to implement in a regional modelling context; our work suggests that a considerable deficit in oceanic NH_3 may be occurring in the current regional air quality models.

The evaluation of the models' NH_3 against biweekly surface observations at the AMoN sites is shown in Table S5 (Supplement), where biweekly values have been used to create annual averages from both model and observed values at observation sites. GEM-MACH (Base) and GEM-MACH (Zhang) once again have the lowest-magnitude (and positive) biases relative to observations; CMAQ-M3Dry and CMAQ-STAGE have the most negative biases, though CMAQ-STAGE has the best correlation coefficient score; and WRF-Chem (UPM) has the best scores overall aside from the mean bias and correlation coefficient.

Figure 22 shows the contributions to total N deposition flux from the dry deposition of NH_3(g) and the difference in overall deposition patterns between the models employing bidirectional NH_3 flux parameterizations (CMAQ-M3Dry, CMAQ-STAGE, GEM-MACH (Base), and GEM-MACH (Zhang)) and the models which do not employ such a parameterization (WRF-Chem (IASS), GEM-MACH (Ops), WRF-Chem (UPM), WRF-Chem (UCAR)). The models utilizing bidirectional fluxes have large regions where the net downward flux is given as zero in the panels of Fig. 22 (dark-blue regions, CMAQ-M3Dry, CMAQ-STAGE, GEM-MACH (Base), GEM-MACH (Zhang) models); these are locations where the annual total NH_3 flux is *upward*, with net *emissions* of NH_3 when summed over the course of the year. The size of these regions differs between CMAQ-M3Dry and CMAQ-STAGE, indicating differences in the bidirectional flux parameterizations between these models. GEM-MACH (Base) and GEM-MACH (Zhang) also use a bidirectional flux parameterization, which differs from those of CMAQ-M3Dry and CMAQ-STAGE, and consequently have relatively similar patterns of net NH_3 dry deposition versus emissions. Differences in land use data as well as country-specific differences in the level of detail utilized in the bidirectional flux schemes also result in differences between the two modelling platforms (e.g. border of the northwestern USA and southwestern Canada shows up as a sharp contrast in the CMAQ models, with NH_3 fluxes that utilize information from EPIC over the USA and less detailed information outside the USA, while this differences is much less pronounced in the GEM-MACH models).

The AQMEII4 diagnostics for NH_3 deposition provide further insight into the causes of the differences between the models employing NH_3 bidirectional fluxes. The most generic formula for NH_3 bidirectional fluxes is

$$F_T = \frac{c_a - c_c}{r_{sum}}, \tag{5}$$

where F_T is the net flux, c_a is the atmospheric concentration of ammonia gas, and r_{sum} is a sum of resistances associated with turbulent eddies and molecular diffusion of gaseous NH_3 across the reference height of air and the vegetation canopy. c_c is the canopy compensation point concen-

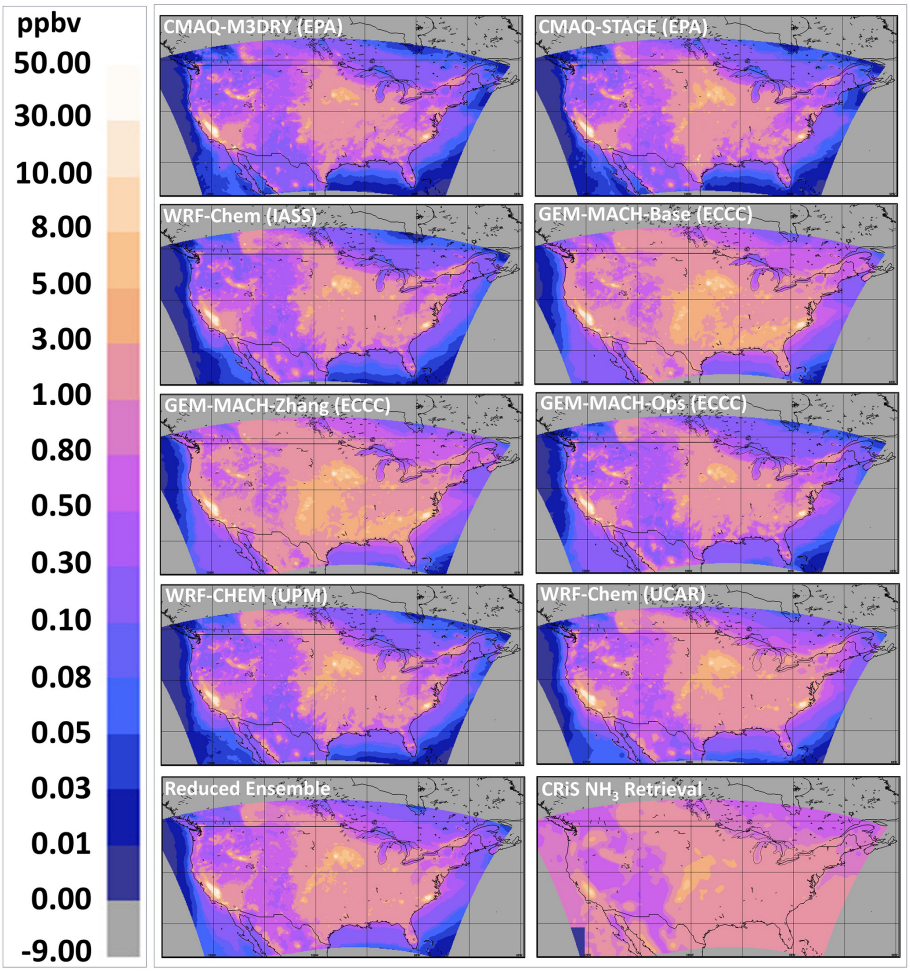

Figure 21. Comparison of annual average surface NH_3 concentrations at CrIS overpass times (participating models and the reduced ensemble) and the corresponding CrIS observed average NH_3 at overpass times. Note that regions outside the AQMEII4 common domain have been assigned an outside-domain mask value of −9.

trations of ammonia gas at the top of the canopy and may be expressed as a function of the atmospheric concentration as well as compensation point concentrations near stomata and the ground (c_s, c_g) and of the aerodynamic resistance of ammonia gas (r_a). As can be seen from Eq. (5), if the atmospheric concentration is greater than the compensation point concentration, the flux will be positive (downward). If the atmospheric concentration is less than the compensation point concentration, the flux will be negative (upward). Galmarini et al. (2021, Appendix C) give the detailed formulae for the terms in Eq. (5), for the bidirectional flux models participating in AQMEII4. A comparison of r_a, r_{sum}, c_a, c_c, c_g, and c_s may thus provide insight into the differences between the predicted NH_3 dry-deposition fluxes for the models employing bidirectional flux parameterizations for the North American AQMEII4 ensemble. These terms were reported by AQMEII4 participants as the diurnal median (50th percentile) at each hour (UT) within each month. The median values for 16:00 UT (noon EDT) for July 2016 are shown in Fig. 23. It is important to note that the median values for a given hour (UT) may correspond to different days within a given month. For example, the median values of r_{sum} and r_a at 16:00 UT in July may not occur on the same day, and hence the median value of r_{sum} will not necessarily be greater than the median value of r_a, as might be expected from the equations governing the resistances as given in Appendix C of Galmarini et al. (2021). Also, not all models were able to report all variables (as noted above, for CMAQ-M3Dry, the net and ground compensation point concentrations were calculated offline of the model simulation and could not be included as AQMEII4 diagnostic parameters). However, substantial differences between the panels of Fig. 23 provide a useful indication of the relative importance of different pathways in the participating models.

From Fig. 23, we note the following:

1. The median aerodynamic resistance r_a for July 2016 at 16:00 UT is similar for all four models (Fig. 23a); con-

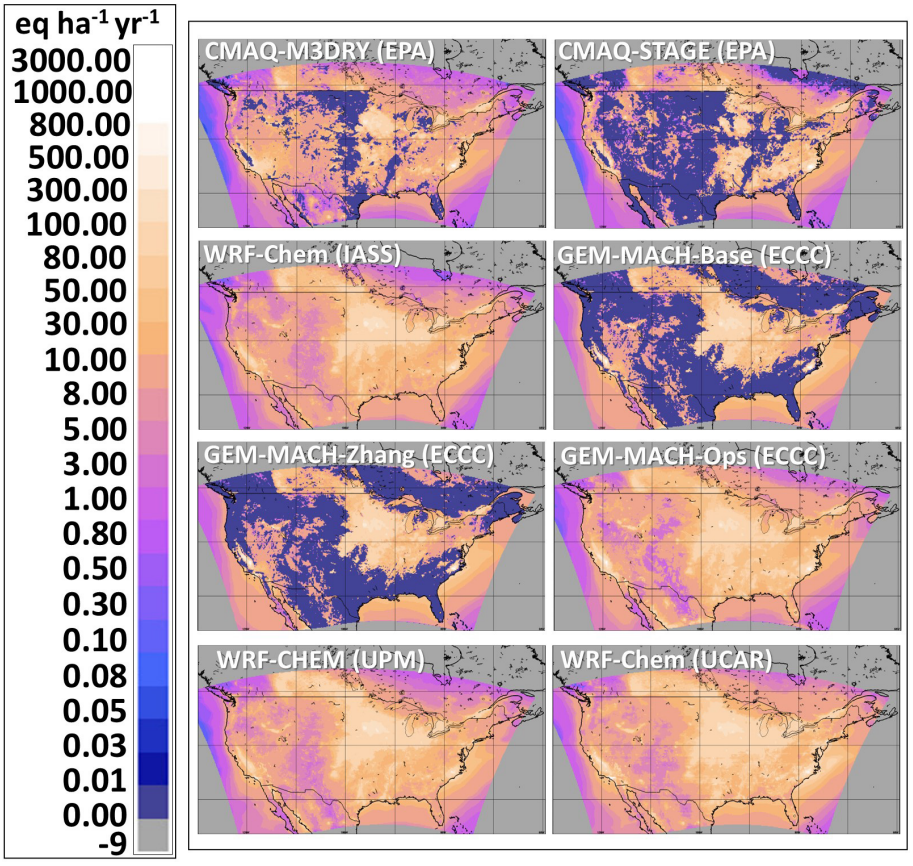

Figure 22. The 2016 N dry-deposition fluxes (eq. ha^{-1} yr^{-1}) for NH_3(g) (eq. ha^{-1} yr^{-1}). Note that regions outside the AQMEII4 common domain have been assigned an outside-domain mask value of −9.

sequently, differences in r_a are unlikely to be the cause of the model flux differences.

2. The median r_{sum} values (Fig. 23b) for CMAQ-M3Dry for July 2016 at 16:00 UT is considerably smaller than for other models; at least some relatively high fluxes for CMAQ-M3Dry are due to these smaller r_{sum} values (which, appearing in the denominator for Eq. 5, will increase the magnitude of the fluxes).

3. The median r_{sum} values for CMAQ-STAGE over land for July 2016 at 16:00 UT are equal to those for r_a for this model. This is expected ($r_{sum} = r_a$ for this model; Galmarini et al., 2021); other terms influence the magnitude and direction of the fluxes.

4. The median values of the air concentrations of NH_3 c_a (Fig. 23c) for July 2016 at 16:00 UT are lower for CMAQ-M3Dry and CMAQ-STAGE than for GEM-MACH (Base) and GEM-MACH (Zhang), as might be expected from the above-mentioned bias calculations relative to CrIS and AMoN data.

5. The median net compensation point concentration c_c (Fig. 23d) for CMAQ-STAGE for July 2016 at 16:00 UT is an order of magnitude smaller than for GEM-MACH (Base) and GEM-MACH (Zhang). From Eq. (5), this likely drives much of the large NH_3 flux for this model and its negative bias values; smaller c_c values will result in larger positive (downward) net fluxes F_T.

6. Some of the locations where CMAQ-STAGE's median ground compensation point concentration (c_g) for July 2016 at 16:00 UT maximized are where GEM-MACH (Base) and GEM-MACH (Zhang) have zero to near-zero ground compensation point values (Fig. 23e; e.g. Rocky Mountains, north-central US agricultural region; dark-blue areas in the GEM-MACH results compared to much lighter values in the CMAQ-STAGE results). The larger CMAQ-STAGE c_g values (local values were up to 1×10^4 ppbv for this model), if dominant, would be expected to result in larger c_c values in Eq. (5) (see Galmarini et al., 2021) and hence a tendency towards smaller downward fluxes. This is not the case from the above analysis (DNH3 values in Table 5 for CMAQ-STAGE are greater than those of the GEM-MACH models, and CMAQ-STAGE NH_3 concentrations have more negative biases than the two GEM-

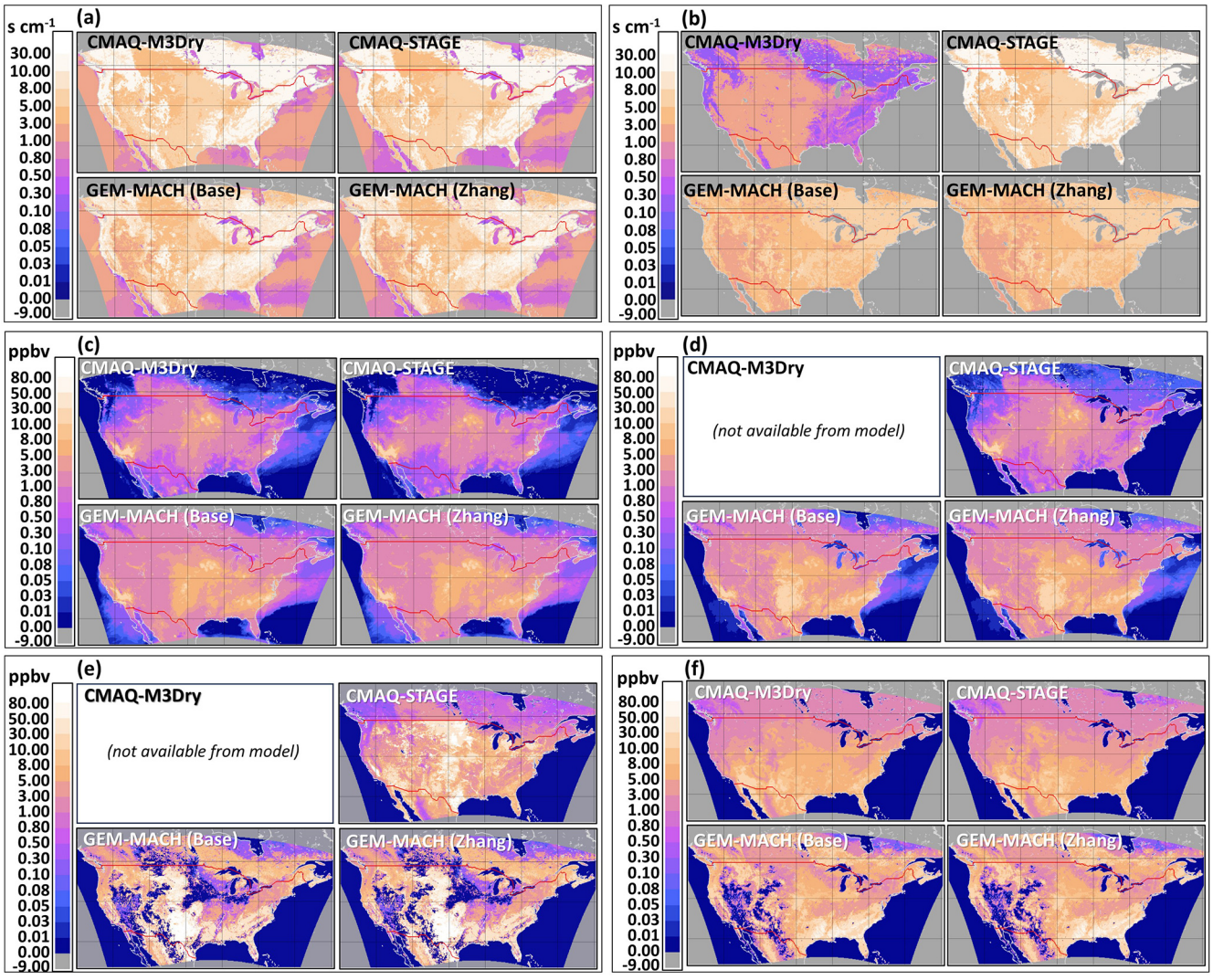

Figure 23. The 2016 spatial distribution from July 2016 at 16:00 UT with median N values for key bidirectional flux diagnostic variables. **(a)** Aerodynamic resistance (r_a; s cm^{-1}). **(b)** Sum resistance (r_{sum}; s cm^{-1}). **(c)** Air concentration of NH_3 (c_a; ppbv). **(d)** Net compensation point concentration (c_c; ppbv). **(e)** Ground compensation point concentration (c_g; ppbv). **(f)** Stomatal compensation point concentration (c_s; ppbv). Note that regions outside the AQMEII4 common domain have been assigned an outside-domain mask value of −9.

MACH models), suggesting that the ground pathway is not the main term affecting the differences in model NH_3 dry-deposition fluxes.

7. For much of the AQMEII4 common domain (aside from the southwestern USA), CMAQ-M3Dry and CMAQ-STAGE have lower median stomatal compensation point concentrations for July 2016 at 16:00 UT than either GEM-MACH (Base) or GEM-MACH (Zhang) (Fig. 23f). This in turn implies that the difference in model dry-deposition fluxes is via the stomatal pathway.

The main factors resulting in higher-magnitude downward fluxes in CMAQ-M3Dry and CMAQ-STAGE relative to GEM-MACH (Base) and GEM-MACH (Zhang) are thus lower net compensation point concentrations (CMAQ-STAGE), lower stomatal compensation point concentrations (CMAQ-M3Dry, CMAQ-STAGE), and lower r_{sum} values (CMAQ-M3Dry).

All four bidirectional flux models calculate fluxes on specific land use types within each grid cell and use some form of land use fraction weighting to generate the values of the key parameters in the bidirectional flux equations. The native land use types used by each modelling platform were converted to a common set of 16 AQMEII4 land use types (see Galmarini et al., 2021). We investigated the CMAQ and GEM-MACH spatial and temporal patterns of ammonia bidirectional fluxes in the context of the AQMEII4 land use types, along with the relationship to the highest regions of nitrogen CLE. This is shown in Figs. 24 and 25, where Fig. 24a

and b are the sum of AQMEII4 land use types 11 and 12 (i.e. the sum of the "planted/cultivated" and "grassland" land use types) used in CMAQ and GEM-MACH, respectively. Figure 24c and d are the sum of AQMEII4 land use fractions for land use types 6, 7, 8, and 13 (evergreen broadleaf forest, deciduous broadleaf forest, mixed forest, and savanna, respectively), for CMAQ and GEM-MACH, respectively. We note that these forested areas are the ecosystems of interest for many of the CLE values calculated earlier in this work. The land use summations of Fig. 25 are also worth noting in the context of the typical timing of the direction of NH_3 fluxes during the course of a day. Figure 25 shows an example of the diurnal behaviour of the NH_3 bidirectional fluxes for the CMAQ and GEM-MACH models, at (a) 15:00 CDT and (b) 07:00 CDT. Midafternoon fluxes (Fig. 26a) tended to be largely negative (upward; emissions; blue colours). However, the spatial location of the fluxes differs between the models even within a given model framework. CMAQ-M3Dry predicts afternoon emissions (blue colours) largely restricted to the combined grassland and agricultural land use types, with deposition (red colours) to the forested areas in southeastern Canada and the southeastern USA. CMAQ-STAGE predicts midafternoon emissions throughout western North America, though with a pattern of deposition similar to CMAQ-M3Dry in southeastern Canada and the southeastern USA. The GEM-MACH bidirectional fluxes in the afternoon are mostly negative (emissions; blue). All three models show midafternoon NH_3 deposition in the north-central USA, corresponding to a known region of high NH_3 concentrations (Fig. 21; CrIS NH_3 retrieval maximum). In contrast, early-morning fluxes (Fig. 25b) predicted by both CMAQ implementations are largely positive (downward; deposition; red colours), across all land use types, while GEM-MACH predicts deposition in agricultural areas and emissions further downwind in southeastern Canada and the southeastern USA.

The generic diurnal sign changes in the direction of the ammonia flux across all four models is easily explained in reference to Eq. (5): in midafternoon (Fig. 25a), both the height of the planetary boundary layer and the magnitude of thermal coefficients of diffusivity are relatively high, reducing the ambient air concentration of ammonia gas (c_a in Eq. 5), resulting in negative fluxes (emissions; blue colours). In the early morning (Fig. 25b), both the boundary layer height and the magnitude of thermal coefficients of diffusivity are lower, hence increasing the ambient air concentrations of ammonia gas, resulting in more positive fluxes and prevalent deposition. However, the different bidirectional flux models show differences in diurnal behaviour by land use type. CMAQ-M3Dry and CMAQ-STAGE show a diurnal pattern of afternoon emissions from agricultural and grassland areas, deposition in forested regions downwind, and early-morning deposition irrespective of land use type. GEM-MACH shows stronger afternoon emissions regardless of land use type and morning lower-magnitude emissions in forested areas and deposition only in agricultural areas and the western USA.

We note that Table S4 measures model performance specifically at the satellite overpass time in the afternoon, i.e. at close to the time shown in Fig. 25a, and that the performance of CMAQ-M3Dry and CMAQ-STAGE is lower than the other models at this time, while the differences between the models aside from the magnitude of the bias is less pronounced in the integrated surface observations of Table S4. This analysis thus suggests that the CMAQ negative biases may be reduced in magnitude by re-examining the factors contributing to compensation point concentrations in forested areas in the day; c_c values (Eq. 5) are probably too low in these regions at these times, leading to excessive positive (downward) fluxes. That is, the analysis suggests that the CMAQ negative NH_3 biases may be the result of excessive deposition and/or insufficient emissions, in forested areas, in both the daytime and early morning, with the effect most noticeable in the afternoon. The bulk of the differences likely resides in the stomatal deposition pathway. Conversely, we note that the GEM-MACH bidirectional flux algorithm overestimates midafternoon ammonia in the southeastern USA relative to satellite observations (Fig. 21), indicating that compensation point concentrations may be overestimated in this region.

While NH_3 fluxes are only the fifth largest source of N deposition in the North American reduced ensemble, we also note that the manner in which NH_3 bidirectional fluxes are treated in the context of critical load exceedance calculations may be open to interpretation. Exceedances with respect to critical loads are calculated with respect to annual *total* deposition of N and S, but what constitutes total N deposition in the context of bidirectional fluxes is less clear. Here, we have taken the approach of assuming that negative fluxes (emissions) of NH_3 during the course of a year constitute a loss of N from the ecosystem but that NH_3 contained within the ecosystem cannot be converted to other forms of N. Consequently, the approach taken here was to sum the hourly NH_3 fluxes (positive downward and negative upward) for the year simulated, with only those grid cells with net positive summations (i.e. net annual deposition fluxes) adding to total N deposition. However, other interpretations are possible. For example, only the positive contributions on an hourly basis could be accumulated, and any losses of N from the same ecosystems associated with NH_3 emissions could be ignored/excluded from the N balance of the ecosystem. A third interpretation would be to assume that deposited NH_3 within the ecosystem may be converted to other forms of N, and hence the net NH_3 flux (which may be positive or negative in different parts of the region simulated) is added to N_{dep}, with N_{dep} being set to zero only when the NH_3 emission flux exceeds the deposition flux of all other forms of N. Here, we have taken the first of these approaches. We note that the second approach would lead to higher estimates of total N_{dep} than generated here, while the third approach would result

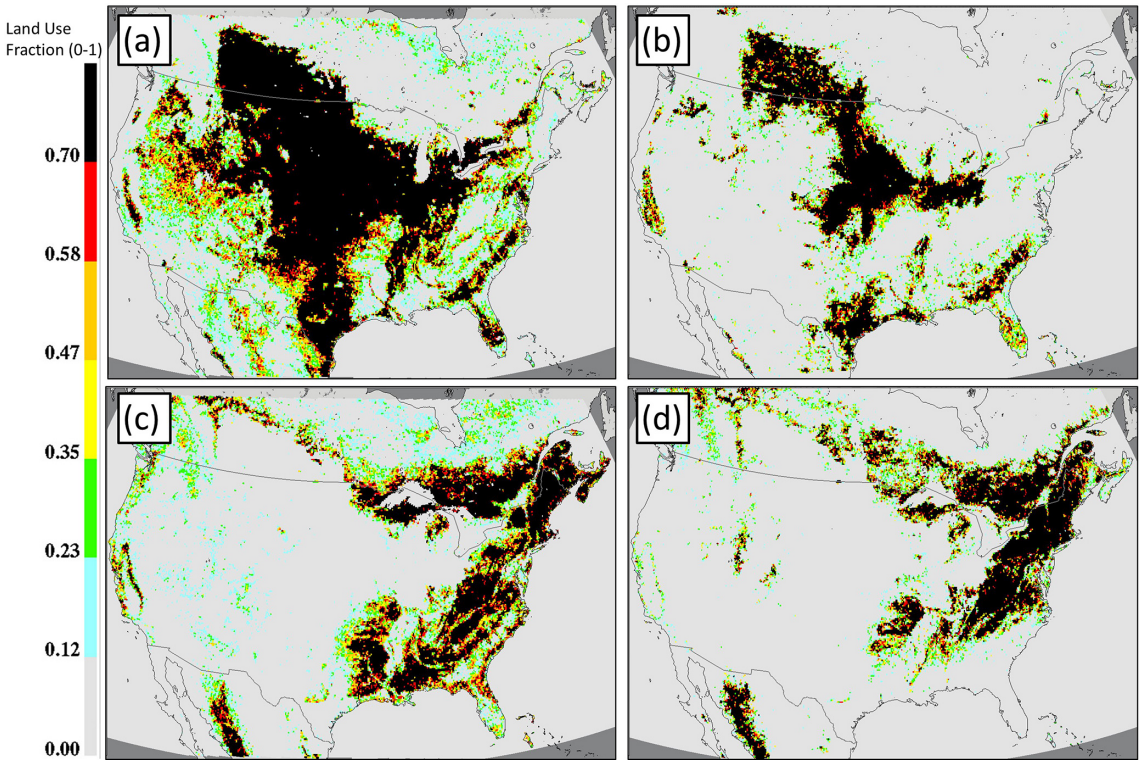

Figure 24. Comparison of AQMEII4 land use type fractions with locations of the highest CLE for forest ecosystems: CMAQ versus GEM-MACH. **(a, b)** Grid cell fractional area composed of the sum of AQMEII4 land use types 11 and 12 (planted/cultivated and grassland) for **(a)** CMAQ-M3Dry and CMAQ-STAGE and **(b)** GEM-MACH (Base) and GEM-MACH (Zhang). **(c, d)** Grid cell fractional area composed of the sum of AQMEII4 land use types 6, 7, 8, and 13 (evergreen broadleaf forest, deciduous broadleaf forest, mixed forest, and savanna) for **(c)** CMAQ-M3Dry and CMAQ-STAGE and **(d)** GEM-MACH (Base) and GEM-MACH (Zhang).

in lower estimates of total N_{dep}. Although NH_3 is the fifth largest contributor to total N_{dep} across North America, these differences in approach may affect critical load exceedance estimates in regions of high NH_3 fluxes.

3.2.3 Causes of S deposition variability in European domain simulations

The relative contributions of the different sources of S deposition in the EU AQMEII4 common domain for the year 2010 are shown in Table 6 and Fig. 26.

The European ensemble contributions to total S deposition contrasted with those in North America; both the contribution to total S deposition and the magnitude of variability between the models follow the same descending order of importance: SO_2 dry deposition followed by $(SO_4^{2-} + HSO_3^-)$ wet deposition, followed by particulate sulfate dry deposition (see Table 6). The relatively higher importance of SO_2 dry deposition to total sulfur deposition, compared to North America, may reflect a denser spatial distribution of SO_2 emissions in the EU domain compared to the North American domain, as well as higher EU emissions in 2010 compared to the NA 2016 year focused on here for model variability analysis. Another potential cause of differences between the two domains may reflect differences in the quality of the emission data (and emission reporting requirements) between the two jurisdictions. SO_2 emissions are largely from industrial stacks in both locations. In North America, regulations require that facility operators for large stack sources report their emissions and stack parameters making use of continuous emission monitoring, on an hourly basis (USA) or as annual reports (Canada). Plume rise algorithms may then be used to distribute the emissions in the vertical within air quality models. In the EU, stack sources are reported as annual totals without stack parameters which could be used for more accurate plume rise estimates (e.g. volume flow rates, effluent temperatures); the lack of these more detailed data necessitates approximations (either making use of "typical" plume rise rates or treating stack sources as surface emissions without plume rise). The larger variation in SO_2 performance in the simulations may thus reflect differences in the level of detail available within SO_2 emission inventories in the two regions.

European observation data for model evaluation were taken from the European Monitoring and Evaluation Programme (EMEP; https://www.emep.int/, last accessed 11 July 2024) and the European air quality database

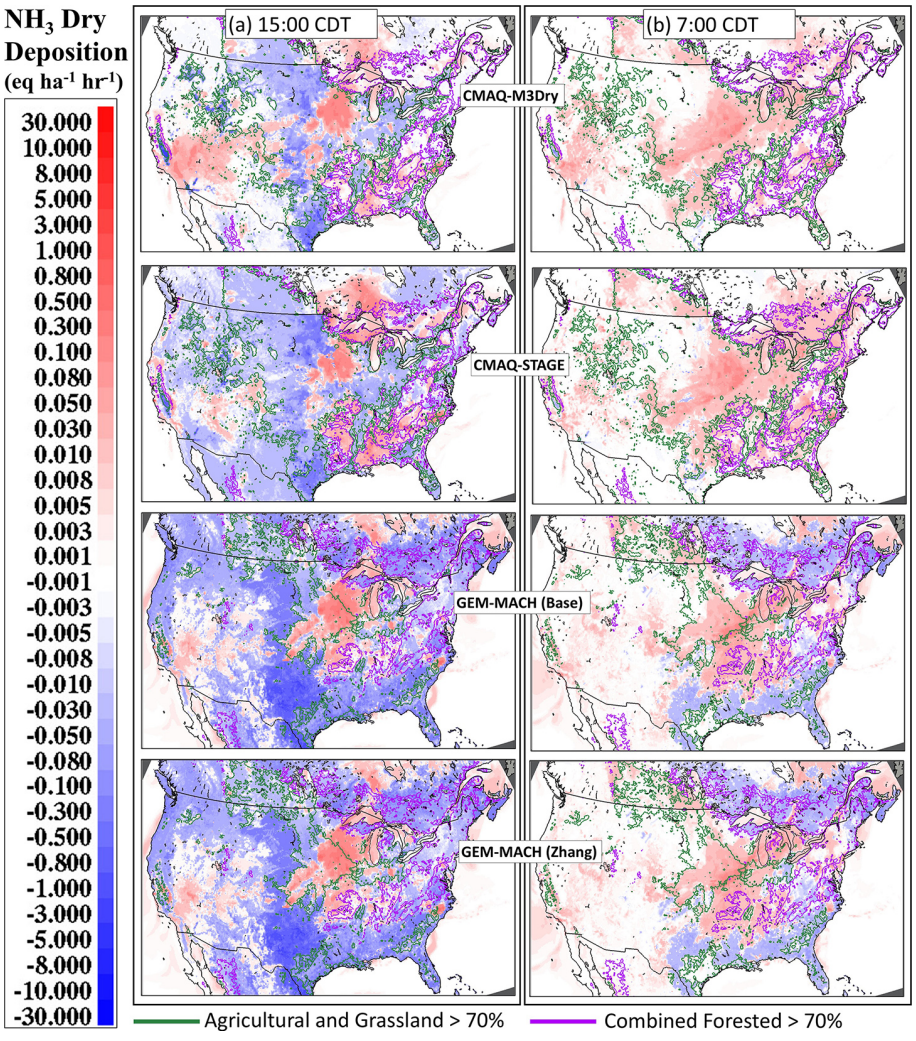

Figure 25. NH_3(g) flux (eq. ha^{-1} h^{-1}) on **(a)** 4 August 2016 at 15:00 CDT and **(b)** 5 August 2016 at 07:00 CDT. Blue lines in the CMAQ and GEM-MACH model (horizontal row) panels enclose areas which are predominantly agricultural and grassland, whereas red lines enclose areas which are predominantly evergreen broadleaf forest, deciduous broadleaf forest, mixed forest, and savanna in each model's respective land use databases (see Fig. 24). Blue-shaded regions indicate negative (upward; emissions) NH_3 fluxes; red-shaded regions indicate positive (downward; deposition) NH_3 fluxes. Green lines show the boundary of regions where combined agricultural and grassland land use types comprise greater than 70 % of land cover. Purple lines show the boundary of regions where combined forest land use types comprise greater than 70 % of land cover.

(AirBase; https://discomap.eea.europa.eu/map/fme/AirQualityExportAirBase.htm, last access: 3 March 2025.).

Dry deposition of SO_2

The model SO_2 performance metrics relative to observations at stations closer to urban centers (AirBase network) and more broadly distributed over the EU region (EMEP network), as well as comparisons to wet ($SO_4^{2-} + HSO_3^-$) deposition (EMEP wet-deposition network), are shown in Table S6 (Supplement). Observation station locations are shown in Fig. S17a. WRF-Chem (IASS) had the best SO_2 performance relative to both networks for most statistics, with the exceptions of a slightly smaller FAC2 score compared to other models for both AirBase and EMEP, and the largest negative bias for SO_2 relative to AirBase observations. The proximity of AirBase station locations to SO_2 sources can also be seen in Fig. 27, where the AirBase monthly concentration y axis (Fig. 27a) is almost twice that of the EMEP monthly concentration y axis (Fig. 27b). Observed SO_2 close to sources (Fig. 27a, red lines) shows a strong seasonal variability, with concentrations in the winter being a factor of 2 higher than in summer, likely showing the effect of increased winter stability on plume rise. This tendency is greatly reduced at regional stations (Fig. 27b, red lines). LOTOS-EUROS (TNO) matches the near-source SO_2 time series the most closely, while CMAQ (Hertfordshire)

Table 6. Average S deposition contributions in the EU AQMEII4 common grid area (eq. ha^{-1} yr^{-1}) and percent contribution to average total S deposition, 2010. n/d: no data. nr: not reported.

Model number	Average deposition (eq. ha^{-1} yr^{-1})				Percent of total S deposition		
	SO_2(g) dry deposition	$SO_4^{(2-)} + HSO_3^{(-)}$ wet deposition	Particle sulfate dry deposition	Total S deposition	SO_2(g) dry deposition	$SO_4^{(2-)} + HSO_3^{(-)}$ wet deposition	Particle sulfate dry deposition
WRF-Chem (IASS)	92.1	42.1	nr	134.2	68.6	31.4	n/d
LOTOS-EUROS (TNO)	38.3	37.9	5.4	81.5	47.0	46.4	6.6
WRF-Chem (UPM)	105.6	63.2	3.2	172.0	61.4	36.7	1.9
CMAQ (Hertfordshire)	125.7	75.9	20.1	221.6	56.7	34.3	9.0
Reduced-ensemble average	89.9	59.0	9.5	158.4	56.7	37.2	6.0
Reduced-ensemble standard deviation	37.3	15.8	7.5	58.0	23.6	10.0	4.7

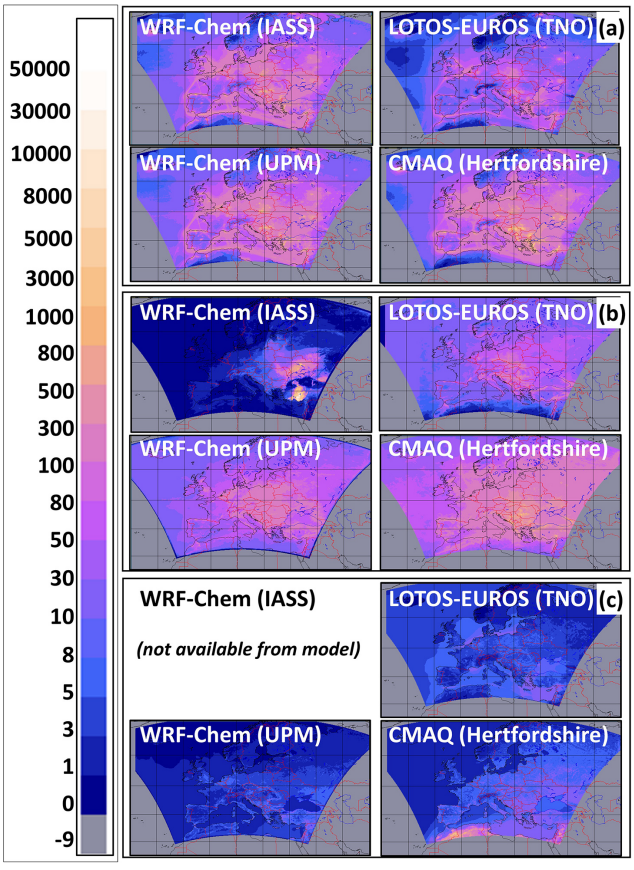

Figure 26. Spatial distribution and magnitude of contributions to annual S deposition, EU AQMEII4 common domain, 2010 (eq. ha^{-1} yr^{-1}). **(a)** SO_2(g) dry deposition. **(b)** Total S wet deposition. **(c)** Particle sulfate dry deposition. Note that regions outside the AQMEII4 common domain have been assigned an outside-domain mask value of −9.

overestimates the impact of seasonal variability (Fig. 27a). At regional stations, LOTOS-EUROS (TNO) and CMAQ (Hertfordshire) overestimate seasonal variation, while WRF-Chem (IASS) most closely matches observations. At least some of the variation in the simulated SO_2 performance relative to observations and hence in SO_2 deposition fluxes and critical load exceedance estimates is due to some models overestimating the seasonal variation in SO_2 at regional locations further from cities. This may reflect differences in atmospheric stability, the seasonal response of the deposition algorithms, or the manner in which plume rise is simulated between the models.

WRF-Chem (IASS) has the best overall performance for SO_2; while this model's mean bias is the most negative for observation sites close to the sources (AirBase comparison), the remaining statistics are the best of the ensemble, and the model bias performance is also better than the other models as the distance from the sources increases (EMEP comparison). The large negative biases in WRF-Chem (IASS) model values may indicate an overestimate of SO_2 deposition, though other model processes may also play a role.

Wet deposition of sulfur

As was the case for most models in the North American domain, all EU domain models underestimated wet deposition relative to observations (note negative biases in Table S6 and the monthly time series comparison versus observations in Fig. 27c). CMAQ (Hertfordshire) outperforms the other models relative to observations, though we note that the sulfur wet-deposition bias for this model is nevertheless −0.39 eq. ha^{-1} yr^{-1}, with a correlation coefficient of 0.15. In contrast to the North American sulfur wet-deposition comparison time series (Fig. 15, Table S2), the European wet-deposition observations do not show a springtime peak in values, rather a seasonality centered around the month of

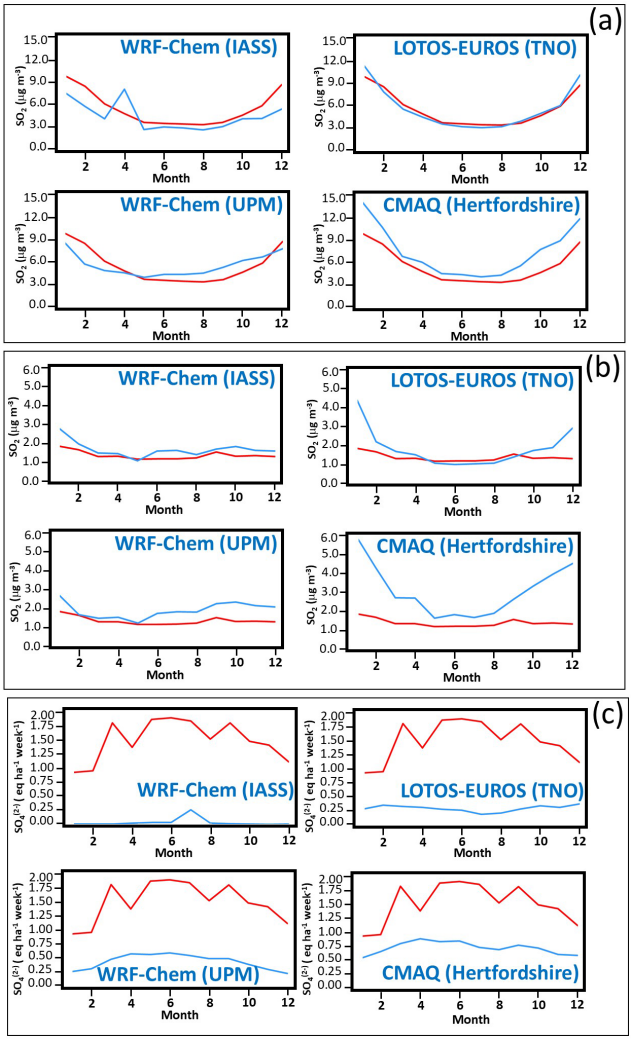

Figure 27. Comparison of observed and modelled S, EU AQMEII4 common domain, 2010. **(a)** AirBase SO_2 ($\mu g\,m^{-3}$). **(b)** EMEP SO_2 ($\mu g\,m^{-3}$). **(c)** Wet flux of total S deposition (eq. ha^{-1} per week). Red: observations. Blue: model.

June, with higher values extending from March to September.

None of the EU models made use of updated particle dry-deposition velocities available in more recent literature; as a result, the relative contribution of particle dry deposition to EU model-to-model variability is small. Speciated PM observations were not available for comparison to model predictions in the EU region.

Returning to the spatial distribution of the relative contributions of the three forms of sulfur deposition for the year 2010 shown in Fig. 26, CMAQ (Hertfordshire), with the highest SO_2 deposition flux (Fig. 26a; see also Tables 6, S6), also has the most positive SO_2 concentration mean bias. With increasing distance from the sources, the SO_2 loss or conversion processes of all four models are likely underestimated (EMEP SO_2 biases are positive for all models, Table S6). In contrast, all models have significant negative biases in sulfur wet deposition (Table S6); hence at least one reason for this underestimate may be insufficient conversion of SO_2 to ionic sulfate and bisulfite in simulated cloud water, through uptake of SO_2 and scavenging of particulate sulfate. The wet deposition of sulfur in WRF-Chem (IASS) in particular seems anomalously low (Figs. 26c, 27b), with much of Europe having little to no sulfate wet deposition in this model. A comparison of the relative differences in the deposition pathway strength for the models may help shed light on the causes of SO_2 deposition flux variability between the models. However, no effective fluxes were reported by LOTOS-EUROS (TNO). Figures S23 and S24 show the spatial distribution of the summed annual effective fluxes for the reporting models, with the results in the EU AQMEII4 common domain summarized in Fig. 28.

Despite having the highest average SO_2 deposition flux (Table 6), CMAQ (Hertfordshire) also has the highest positive biases for SO_2 ambient concentrations (Table S6). From Figs. S23, S24, and 28, the CMAQ (Hertfordshire) positive biases may be the result of spatial variations in deposition, specifically, to low contributions to the cuticle effective fluxes in northern Europe for this model (Fig. S23a). Despite these relatively low values, the SO_2 net dry-deposition flux for this model (Table 6) is higher than that of the other models, implying that the low northern EU fluxes are being offset by higher values elsewhere (e.g. via the soil flux; compare soil and cuticle values in Fig. 28). We note that the effective-flux analysis is restricted to grid cells that do not have water as a dominant land use type (a maximum of 1 % water land fraction was used as an exclusion criterion); for grid cells held in common (mostly land), the CMAQ (Hertfordshire) the cuticle effective-flux pathway specifically is lower than that of the other models, while the differences are less noticeable for the other terms, as reflected by the summary values in Fig. 28. Other than northern Europe, CMAQ (Hertfordshire) has higher soil fluxes than WRF-Chem (IASS). Similar to AQMEII4 analyses for ozone (Hogrefe et al., 2025), the relative importance of the different pathways for total deposition varies between the models. For example, WRF-Chem (IASS), with the best overall performance for SO_2 concentrations aside from bias and factor of 2, has flux contributions in descending order of importance: cuticle, stomatal, soil, and lower canopy. For CMAQ (Hertfordshire), with relatively poor performance and high positive biases (Table S6), the flux contributions in descending order of importance are soil, cuticle, and stomatal (with the lower canopy being incorporated as part of soil flux, for this model), and the cuticle pathway contributes less to deposition in northern Europe than the other models.

3.2.4 Causes of N deposition variability in European domain simulations

The EU AQMEII4 common-domain relative contributions for each model's deposited species to total nitrogen deposition and its variability are shown in Table 7. The contributions to total N deposition for the reduced ensemble, in descending order of importance, were wet NO_3^-, dry HNO_3, wet NH_4^+, dry NH_3, dry particulate nitrate, dry NO_2, and dry particle ammonium, with relatively small contributions from the other depositing N species. The spatial distributions of the four largest contributions to total N deposition are shown in Fig. 29. The largest contributions to model-to-model variability, in descending order, were wet NO_3^-, dry HNO_3, dry NH_3, wet NH_4^+, and dry NO_2, with smaller contributions to variability from the other species.

Wet-deposition fluxes of NO_3^- and NH_4^+ and the ground-level concentration of NO_2 are evaluated in Table S7 (Supplement); monthly average time series comparisons of wet deposition to the observations are provided in Fig. 30. From Fig. 29, WRF-Chem (IASS) predicted much lower-magnitude NO_3^- wet-deposition and NH_4^+ wet-deposition fluxes than the other three models, and from Table S7, these result in larger negative biases and poor overall performance relative to observations for WRF-Chem (IASS) in comparison to the other models. LOTOS-EUROS (TNO) had the best overall performance for NH_4^+ and NO_3^- wet-deposition fluxes. However, similar to the case for S wet deposition, all models have significant negative biases for both nitrogen ion wet fluxes, as can be seen from Table S7 and Fig. 30. LOTOS-EUROS (TNO) has the best performance for statistics relating to the spatial and temporal distribution of wet deposition, while WRF-Chem (UPM) has the lowest bias for NO_3^- wet deposition. A common feature of the AQMEII4 ensemble of models for both EU and NA domains are these negative biases for wet deposition of both sulfate and nitrogen species. Also, we note that the observed NH_4^+ wet-deposition (Fig. 30b, red line) peaks in June, while the model values (blue lines) peak earlier, in March. This in contrast to the North American NH_4^+ comparison (Fig. 18), where observed peaks occur in April and model peaks occur in June.

Dry deposition of HNO_3 was the second largest source of modelled EU nitrogen deposition variability. The spatial distribution of the relative contributions of the four pathways to the mass flux of HNO_3 is shown in Figs. S25 and S26 and is summarized for the entire grid in Fig. 31. There is more heterogeneity between the EU models regarding the relative importance of the HNO_3 deposition pathways than was observed for the North American simulations (compare Figs. 20 and 31). In the North American simulations, the cuticle deposition pathway also dominated for all models, followed by the soil pathways. In the EU simulations, the reported soil pathway for WRF-Chem (UPM) was several orders of magnitude smaller than the same pathway for CMAQ (Hertfordshire). The cuticle pathway dominated for WRF-Chem (IASS) (not shown) and CMAQ (Hertfordshire). The stomatal pathway magnitude is less than the cuticle pathway for the EU models but greater in general than for the North American models, where the stomatal pathway had a smaller contribution to HNO_3 dry deposition than the lower-canopy pathway.

Observations of 2010 HNO_3(g), NH_3(g), and dry particle nitrate were not available for comparison to the model predictions. However, observations of the NO_2 concentrations, the sixth largest contributor to total N deposition and the fifth largest contributor to model-to-model variability, were available at near-source AirBase and regionally distributed EMEP stations (Table S7). Aside from having the second largest-magnitude mean bias, LOTOS-EUROS (TNO) had the best performance for NO_2 relative to stations positioned close to emission sources (AirBase), while WRF-Chem (IASS) and CMAQ (Hertfordshire) had the best performance for NO_2 for stations distributed more widely across the region (EMEP).

4 Conclusions

We have used the AQMEII4 North American and European ensembles to calculate net sulfur and nitrogen deposition from individual models and a reduced ensemble of all models. These deposition estimates were used to calculate exceedances of critical loads for these two regions, using several critical load datasets. An in-depth analysis of the causes of model-to-model variability followed, using diagnostics designed for AQMEII4. We therefore subdivide these conclusions by the domain simulated, as well as the critical load exceedance and causes of model variability, within each domain.

4.1 North America: critical load exceedances

All simulations showed a decrease in the size of the area of exceedance and the severity of exceedances with respect to acidification of forest ecosystems and aquatic-ecosystem acidity between the years 2010 and 2016. The total area in exceedance for sensitive-epiphytic-lichen species richness improved slightly, but the severity of exceedance was greatly reduced. Given that the lichen community has a dose–response relationship with increasing deposition, this indicates reduced harm to forest health, even when the CL is still in exceedance. CLEs for herbaceous-species community richness had substantial improvements in the total area of exceedance and severity of exceedance. The amount of exceedance in any given year and the extent of reduction between the 2 years varied considerably between the models. Any individual model provided a similar direction of the change between the 2 years; the range of estimates suggests the utility of model ensembles where possible in estimating critical load exceedances, as well as model–measurement fusion, when sufficient S and N species data are available.

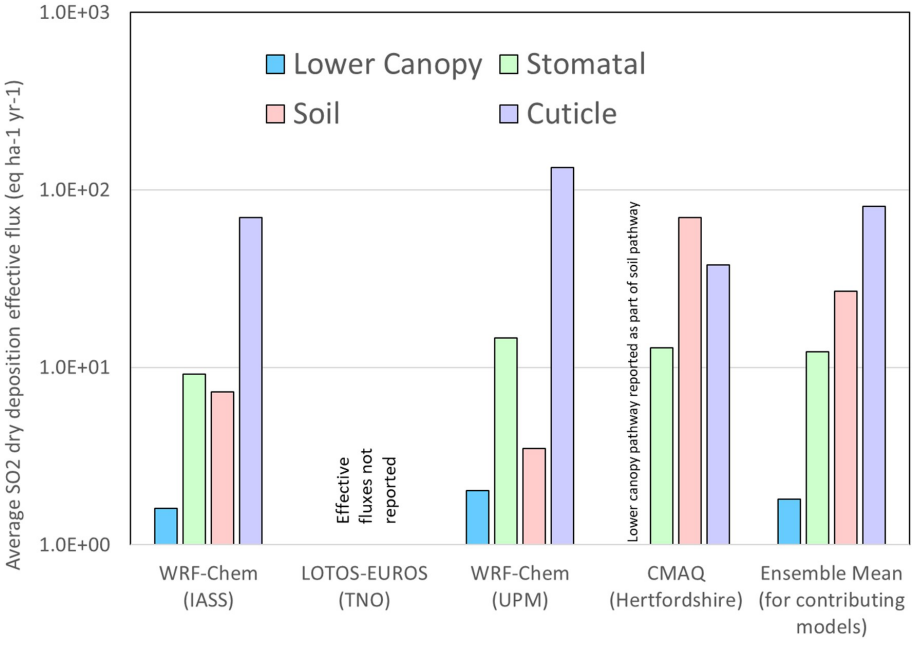

Figure 28. Averages of effective-flux pathway contributions to SO_2 dry deposition, EU AQMEII4 common grid, 2010 (eq. ha^{-1} yr^{-1}).

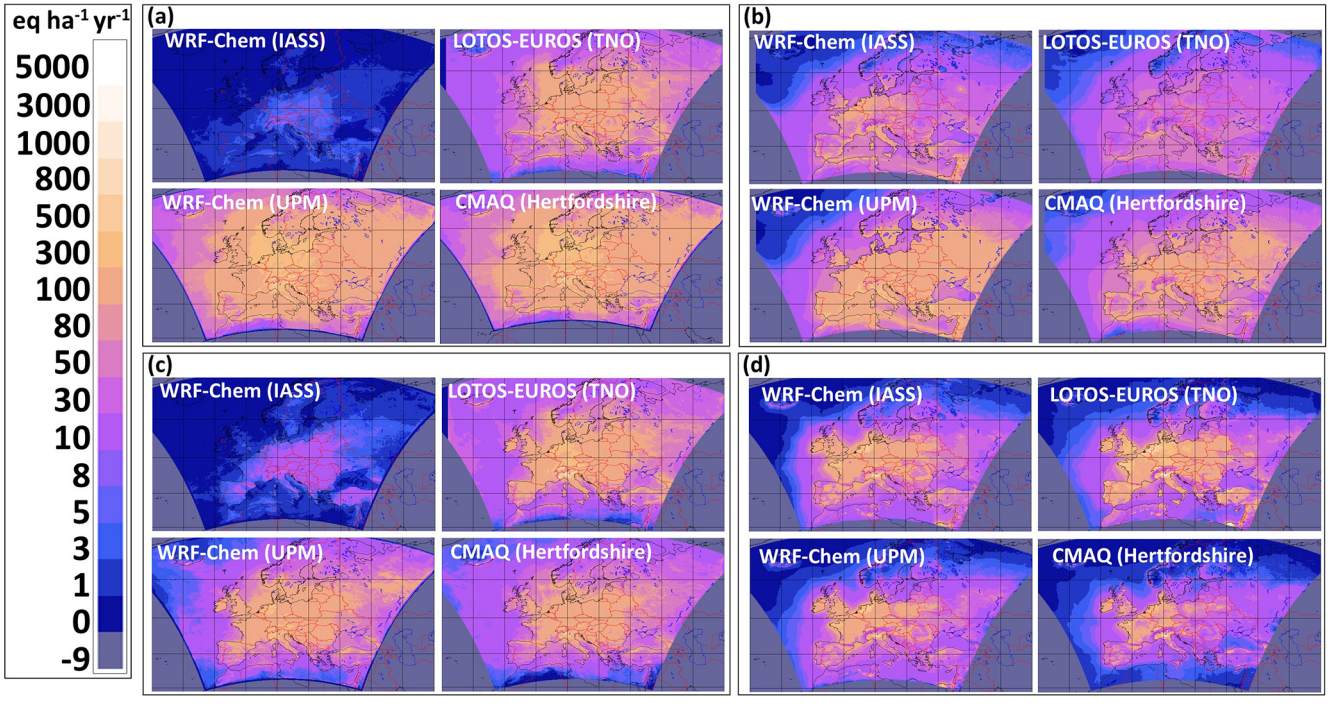

Figure 29. Spatial distribution of contributions of **(a)** wet nitrate ion deposition, **(b)** dry gaseous HNO_3 deposition, **(c)** wet ammonium ion deposition, and **(d)** dry gaseous ammonia deposition to total N deposition in the EU AQMEII4 common domain, 2010 (eq. ha^{-1} yr^{-1}). Note that regions outside the AQMEII4 common domain have been assigned an outside-domain mask value of −9.

Table 7. Contributions of N species to total deposition (eq. ha^{-1} yr^{-1}) and percent of total N deposited, EU AQMEII4 common grid, 2010, arranged in descending order of importance to the reduced-ensemble average. DNH3: dry deposition of NH_3(g). WNH4: wet deposition of NH_4^+(aq). DHNO3: dry deposition of HNO_3(g). WNO3: wet deposition of NO_3^-(aq). DAM: dry deposition of particulate ammonium. DNI: dry deposition of particulate nitrate. DNO2: dry deposition of NO_2(g). DPAN: dry deposition of peroxyacetyl nitrate gas. DRN3: dry deposition of organic nitrate gases. DN2O5: dry deposition of N_2O_5(g). DHNO4: dry deposition of pernitric acid gas. DNO: dry deposition of NO(g). n/d: no data. nr: not reported. ndd: no dry deposition.

Average (eq. ha^{-1} yr^{-1})						
	Model					
Species	WRF-Chem (IASS)	LOTOS-EUROS	WRF-Chem (UPM)	CMAQ (Hertfordshire)	Red. ens. avg	Red. ens. SD
WNO3	1.8	77.8	174.8	96.2	116.2	42
DHNO3	50.2	38.4	120.5	78.6	79.2	33.5
WNH4	4.3	90.3	74.6	64.1	76.3	10.8
DNH3	60.5	76.8	47.9	29.6	51.5	19.4
DNI	nr	18.2	25.9	13.5	19.2	5.1
DNO2	11.6	23.6	27.5	6.3	19.2	9.2
DAM	nr	14.2	6.2	6.6	9	3.7
DPAN	2.3	ndd	2.7	5.2	4	1.2
DN2O5	5.3	1.2	ndd	1	1.1	0.1
DRN3	0.3	ndd	0.6	3.2	1.9	1.3
DHNO4	1.4	ndd	0.9	0.2	0.5	0.4
DNO	0.1	2	0.2	0.4	0.9	0.8
Total N	137.6	342.7	481.9	304.8	376.5	76.1
Percent contribution						
	Model					
Species	WRF-Chem (IASS)	LOTOS-EUROS	WRF-Chem (UPM)	CMAQ (Hertfordshire)	Red. ens. avg	Red. ens. SD
WNO3	1.3	22.7	36.3	31.5	30.9	5.6
DHNO3	36.5	11.2	25	25.8	21	6.7
WNH4	3.1	26.4	15.5	21	20.3	4.4
DNH3	43.9	22.4	9.9	9.7	13.7	5.9
DNI	nr	5.3	5.4	4.4	5.1	0.4
DNO2	8.4	6.9	5.7	2.1	5.1	2.1
DAM	nr	4.1	1.3	2.2	2.4	1.2
DPAN	1.7	n/d	0.6	1.7	1.1	0.6
DN2O5	3.8	0.3	n/d	0.3	0.3	0
DRN3	0.2	n/d	0.1	1.1	0.5	0.5
DHNO4	1	n/d	0.2	0.1	0.1	0.1
DNO	0	0.6	0	0.1	0.2	0.2

4.2 North America: causes of model S deposition variability

The total mass of North American sulfur deposition is as follows, in decreasing order of importance: wet deposition of S ($SO_4^{2-} + HSO_3^-$), dry deposition of particulate sulfate, and dry deposition of SO_2. Dry deposition of particulate sulfate contributed the most to model-to-model variability in total sulfur deposition, followed by wet deposition and dry SO_2 deposition. The models with the highest S wet-deposition levels had the best performance relative to monitoring network observations (CMAQ-M3Dry, CMAQ-STAGE, GEM-MACH (Ops)), though all models' S wet deposition was biased low relative to observations. A subgroup of models (GEM-MACH (Base), GEM-MACH (Zhang), GEM-MACH (Ops)) had the highest positive biases in observed $PM_{2.5}$ sulfate concentrations relative to monitoring network observations, contributing to the model-to-model variability. Recent work by Ryu and Min (2022) and Ghahreman et al. (2024) suggests that model negative biases for wet deposition may be improved through incorporation of multiphase hydrometeor scavenging, and this may also reduce positive biases in particulate mass resulting from the implementation of the Emerson et al. (2020) particle dry-deposition al-

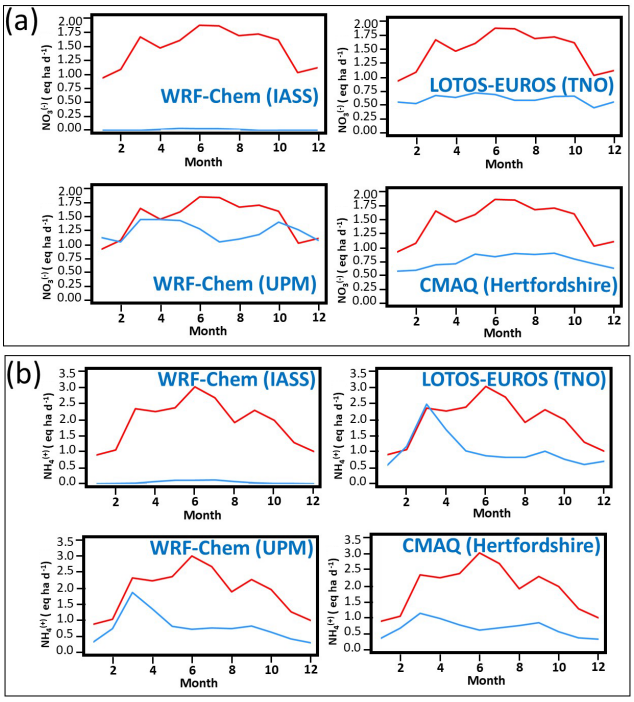

Figure 30. Monthly average comparison of nitrogen wet deposition, EU AQMEII4 common grid, 2010. **(a)** Average flux of NO_3^-(aq). **(b)** Average flux of NH_4^+(aq) (eq. ha^{-1} d^{-1}).

gorithm (GEM-MACH (Base) and GEM-MACH (Zhang)). Most North American reduced-ensemble models were in relatively good agreement with regards to their predictions for the total dry-deposition flux of SO_2(g).

4.3 North America: causes of N deposition variability

The largest contributors to the average total nitrogen *deposition fluxes* across North America in 2016 were wet ammonium ion, dry HNO_3, wet nitrate ion, dry particle ammonium, dry ammonia gas, dry particle nitrate, and dry NO_2, with relatively minor contributions from the other depositing gases. The largest contributors to the average total N deposition flux *variability* across models in descending order of importance were the deposition of dry particulate ammonium, wet ammonium ion, wet nitrate ion, dry nitric acid, dry particle nitrate, dry NO_2, and dry NH_3.

The first and second contributions to model-to-model variability between the members of the North American reduced ensemble were due to the three GEM-MACH implementations (Base, Zhang, and Ops) all having much higher particle ammonium dry-deposition and ammonium ion wet-deposition fluxes, zero to positive biases in ammonium ion wet deposition relative to observations during the summer, and the largest positive biases for $PM_{2.5}$ ammonium concentrations relative to observations, as a result of the simplified sulfate–ammonium–nitrate–water inorganic aerosol thermodynamics algorithm they employed. The positive biases in fine-mode particle ammonium concentrations and positive biases in ammonium ion wet deposition for this subgroup of models are likely caused by the absence of base cations as an alternative sink of nitric acid in addition to ammonium nitrate formation. Updates to these model implementations making use of a new, highly efficient solver for inorganic heterogeneous chemistry which includes the base cation reactions (Miller et al., 2024) should reduce these positive biases. The absence of multiphase hydrometeor scavenging of particle mass may also play a role in the particle ammonium positive biases for these models and in the negative biases across all North American models for ammonium wet deposition and nitrate wet deposition (Ghahreman et al., 2024).

Dry deposition of nitric acid was the second largest contributor to total nitrogen deposition fluxes in North America and the fourth largest contributor to model-to-model variability, with cuticle and the soil pathway dominating the HNO_3 mass flux, usually by more than an order of magnitude.

Comparisons of model-predicted 2016 concentrations of NH_3(g) to both CrIS satellite-based observations (in the afternoon, at overpass time) and ground-based AMON monitoring network values (biweekly averages) showed that the details of the implementation of ammonia bidirectional flux algorithms have a large impact on model NH_3 performance, with CMAQ-M3Dry and CMAQ-STAGE having the most negative NH_3 biases in NH_3 and GEM-MACH (Base) and GEM-MACH (Zhang) having the smallest-magnitude NH_3 biases. A detailed analysis of the magnitude and direction of these models employing bidirectional flux algorithms showed a common diurnal behaviour of daytime emissions from agricultural and grassland areas and deposition in downwind forested areas and nighttime deposition in all regions. However, the GEM-MACH models predicted low-magnitude net emissions from forested areas downwind of agricultural areas in the early morning, while the CMAQ models predicted net deposition at all locations. Differences in the relative magnitudes of compensation point concentrations and the strength of the daytime stomatal deposition pathway were shown to be the cause of these differences.

4.4 Europe: critical load exceedances

The AQMEII4 ensemble for Europe predicted similar exceedances with respect to acidity and eutrophication in 2009 and 2010, with the three-member reduced ensemble showing slightly reduced exceedance levels for acidity and slightly increased exceedance levels for eutrophication, in 2010. We note that the models used made use of inorganic aerosol thermodynamics algorithms which included reactions of base cations, and none made use of more recent updates to the particle dry-deposition parameterization (Emerson et al., 2020; Pleim et al., 2022). Consequently, the magnitude of differences between the models varied from the North American models, and the order of importance of different forms of

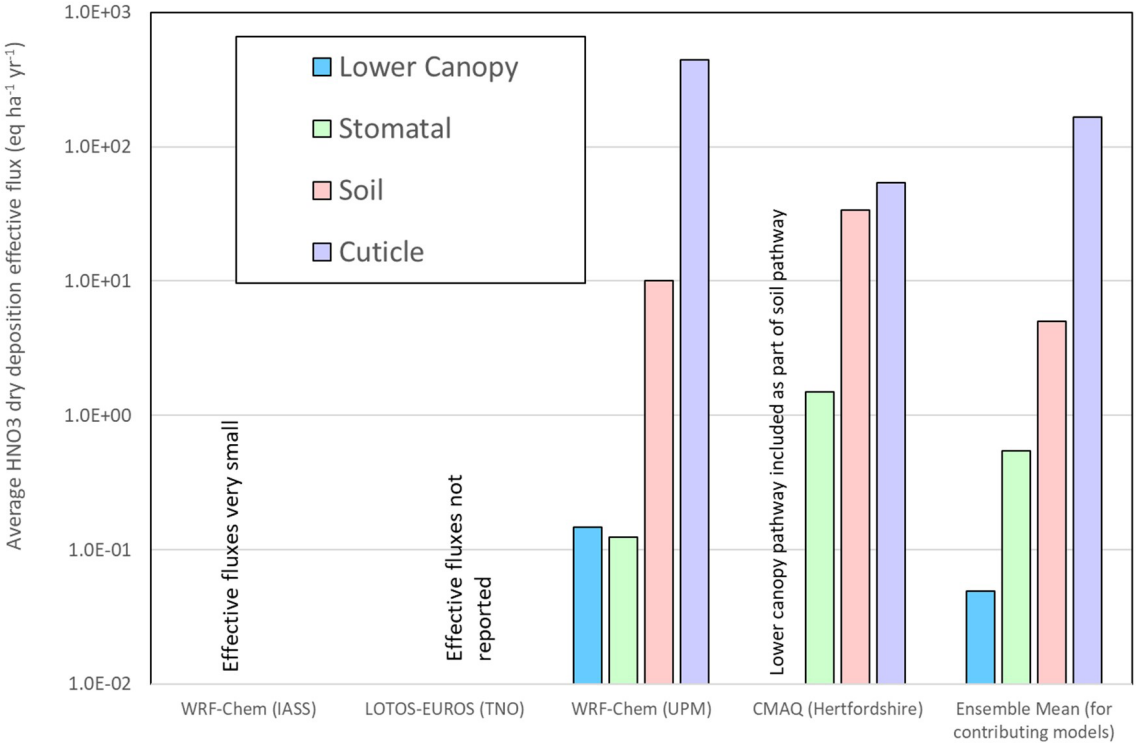

Figure 31. Averages of flux pathway contributions to HNO_3 dry deposition, EU AQMEII4 common grid, 2010 (eq. ha^{-1} yr^{-1}).

sulfur to total deposition differed from the North American ensemble.

4.5 Europe: causes of model S deposition variability

The common-domain average reduced-ensemble sulfur dry-deposition contributions and their variability followed the same decreasing order of importance (SO_2, wet S, dry particulate sulfate). WRF-Chem (IASS) had the best overall performance relative to observations for SO_2 concentrations, while CMAQ (Hertfordshire) had the best performance for S wet deposition. LOTOS-EUROS (TNO) and CMAQ (Hertfordshire) tended to overestimate regional SO_2 seasonality, with much higher concentrations in winter than summer compared to observations in the EMEP SO_2 network. Near-source observations (AirBase network) had higher winter than summer values, though this seasonal variation was largely absent in the observations for stations more representative of regional conditions (EMEP). The positive biases in modelled regional SO_2 concentrations for LOTOS-EUROS (TNO) and CMAQ (Hertfordshire) (the latter relative to both EMEP and AirBase stations) may reflect differences in plume rise distribution between the models or in their driving meteorology's vertical stability (e.g. the modelled wintertime atmosphere may be more stable than is observed, for these models). As was the case in the North American ensemble, all models had negative biases for S wet deposition. As in North America, the manner in which cloud scavenging of particulate sulfate and SO_2 was implemented in these models may be the cause of the wet-deposition negative biases. Unlike North America, speciated PM measurements were unavailable for model evaluation and bias correction.

EU SO_2 deposition pathways were investigated with AQMEII4 diagnostics; the soil and cuticle pathways dominated, and the stomatal pathway was relatively unimportant. This order of importance may reflect diurnal and seasonal SO_2 concentration variations. SO_2 concentrations are more likely to be high under more stable atmospheric conditions (these inhibit the rise in buoyant SO_2 plumes from large stack sources); these conditions are more likely to occur more frequently at night and in the winter, when the influence of the stomatal pathway is at its minimum.

4.6 Europe: causes of model N deposition variability

The relative contributions to total N deposition and the range in the EU domain were the following, in decreasing order of importance: wet nitrate ion, dry HNO_3, wet ammonium ion, dry ammonia gas, dry particle nitrate, and dry NO_2. The variations in the N deposition values between models were smaller than in North America, likely due to the use of base-cation-inclusive inorganic aerosol thermodynamic algorithms in all models and the use of older implementations of wet scavenging and particle dry deposition than in the North American models. We note that NH_3 dry deposition

was the fourth largest contributor to European N deposition model-to-model variability, with the model employing a bidirectional flux algorithm (LOTOS-EUROS) having the highest NH_3 deposition. Satellite-based NH_3 data were unavailable for Europe for the years simulated but are recommended for simulation evaluation in more recent years.

LOTOS-EUROS (TNO) had the best overall performance for nitrate wet deposition, ammonium wet deposition, and near-source NO_2 concentrations compared to the other models. However, all EU models had substantial negative biases in nitrate and ammonium wet deposition, which is in common with the North American models. The seasonality of N wet deposition was poorly simulated, with most models failing to predict the observed summertime maximum of ammonium wet deposition. Given that this negative bias has its maximum in the summer, when agricultural NH_3 emissions are also likely to maximize, evaluation in more recent years of NH_3 predictions against satellite data is recommended.

In accordance with the NA ensemble, those EU models which reported effective-flux diagnostics for all four HNO_3 dry-deposition effective-flux pathways showed the cuticle and soil pathways dominating. The details of an individual land use database may be seen in the HNO_3 deposition flux diagnostics (Figs. S25 and S26), with differences in the amount of inland water being apparent. Furthermore, we note that the land use databases employed in critical load exceedance calculations may *also* differ from those used in individual models. Such mismatches are another source of uncertainty in the estimation the critical load exceedances for the dry-deposition portions of total S and N deposition. The effect of land use type classifications on model deposition fluxes for ozone will be examined in more detail in an upcoming paper (Hogrefe et al., 2025).

4.7 Impact of bias correction as a simple form of model–measurement fusion

A simple form of model–measurement fusion (bias correction) was applied to each of the models' species contributing to total sulfur and nitrogen deposition, for those component species for which observations were available, and corresponding bias-corrected critical load estimates were generated. This sometimes resulted in substantial decreases in model-to-model variability in the CLEs generated, indicating that model–measurement fusion will decrease model-to-model variability, as well as improved CLE estimates, provided sufficient data are available for the main contributors to total sulfur and total nitrogen deposition. In the case of Europe, the application of bias correction *increased* CLE variability for acidification, likely due to the lack of particulate sulfate observations in Europe for the years simulated. The substantial contrast to North American bias-corrected values suggests that the bias corrections for individual species contributing to total sulfur deposition may offset each other (e.g. positive biases in particle sulfate may be offset by negative biases in wet deposition). In the absence of speciated particle observation data in Europe, this compensating effect could not be captured using bias correction, and hence the European CLE variability increased with bias correction.

An important implication of the bias correction exercise conducted here is the need for observation data which close the sulfur and nitrogen deposition budgets to the greatest extent possible when carrying out model–measurement fusion. The biases with respect to observations for sulfur species may reflect inaccuracies in the transformation of one species to another, for example. If model–measurement fusion is applied to only some of the species contributing to sulfur deposition, the resulting total sulfur deposition field and exceedance estimates may be *less* accurate than the original model fields. Similarly, we note that the observations available here did not include particle nitrate or nitric acid data – and hence the impacts of model–measurement fusion on total nitrogen deposition may potentially lead to *less* accurate estimates than the original model values.

Recommendations: air quality modelling needs identified by the analysis.

Our analysis suggests that model biases and model-to-model variability may be reduced through targeted research into specific model process components. These include the following:

- multiphase hydrometeor scavenging of gases and aerosols into clouds to reduce the magnitude of wet deposition and particle concentration biases
- incorporation of improved particle deposition velocity algorithms (e.g. Emerson et al., 2020) – but only in combination with multiphase wet scavenging (Ryu and Min, 2022; Ghahreman et al., 2024)
- incorporation of base cation inorganic chemistry (if not already present) (Fountoukis and Nenes, 2007; Miller et al., 2024) and improved base cation emission inventory development
- NH_3 bidirectional fluxes evaluated using satellite data, with particular reference to improving compensation point estimates for forested areas
- land use type database harmonization across models and between models and critical load databases.

Code availability. It should be noted that the regional model code used in this work was the current version for each model as of 2021, but some of the models are no longer under active development, while others have publicly available code. Below we list the best methods for obtaining copies of the code at the time of writing.

- CMAQ-M3Dry is based on the standard CMAQv5.3.2 code, available via Zenodo link: https://doi.org/10.5281/zenodo.4081737 (US EPA, 2020). CMAQ-STAGE is a custom version updating from CMAQ v5.3.2, available upon email request to Christian Hogrefe (hogrefe.christian@epa.gov).

- WRF-Chem: as of October 2024, WRF-Chem is no longer being developed by NOAA GSL (see https://www2.acom.ucar.edu/wrf-chem, Pfister and Barth, 2024). Readers interested in obtaining copies of the specific code versions used here should contact the co-authors who contributed the following model simulations. WRF-Chem (IASS) (WRF-Chem v3.9.1): Aurelia Lupascu (aura.lupascu@ecmwf.int). WRF-Chem (UPM) (WRF-Chem v4.0.3): Roberto San Jose (roberto@fi.upm.es). WRF-Chem (UCAR) (WRF-Chem v4.1.2): Young-Hee Ryu (yhryu@yonsei.ac.kr, younghee.ryu.ncar@gmail.com) and Alma Hodzic (alma@ucar.edu).
- LOTOS-EUROS is an open source version of the LOTOS-EUROS code based on LOTOS- EUROSv2.3; see https://airqualitymodeling.tno.nl/lotos-euros/open-source-version/ (TNO, 2025).
- GEM-MACH: the chemistry code of the versions of GEM-MACH used in this work can be provided by email request to Paul Makar (paul.makar@ec.gc.ca).

Data availability. Observation data used in this study for model evaluation are publicly available at the following monitoring network data links. In North America, the U.S. Environmental Protection Agency's Air Quality System (AQS; https://www.epa.gov/aqs, US EPA, 2025), National Atmospheric Deposition Program's National Trend Network (NADP NTN; https://nadp.slh.wisc.edu/networks/national-trends-network/, NADP, 2025a), National Atmospheric Deposition Program's Ammonia Monitoring Network (AMON; https://nadp2.slh.wisc.edu/data/AMoN/, NADP, 2025b), Canadian National Air Pollution Surveillance program (NAPS; https://www.canada.ca/en/environment-climate-change/services/air-pollution/monitoring-networks-data/national-air-pollution-program.html, Government of Canada, 2025a), and Canadian National atmospheric chemistry database (https://www.canada.ca/en/environment-climate-change/services/air-pollution/monitoring-networks-data/national-atmospheric-chemistry-database.html, Government of Canada, 2025b). In Europe, the European Monitoring and Evaluation Programme (EMEP; https://www.emep.int/, EMEP, 2025) and European Air Quality Database (AIRBASE; https://discomap.eea.europa.eu/map/fme/AirQualityExportAirBase.htm, AIRBASE, 2025). Satellite ammonia retrieval products used in this study were constructed at Environment and Climate Change Canada. For retrieval data, contact Mark W. Shephard (mark.shephard@ec.gc.ca).

Critical load information used in this study and information on its processing may be obtained by email request to the individual co-authors contributing this information. For North American terrestrial ecosystem critical loads, contact Hazel Cathcart (hazel.cathcart@ec.gc.ca, hazel.cathcart@ec.gc.ca). For North American aquatic ecosystem critical loads contact Jason A. Lynch, Hazel Cathcart, and (lynch.jason@epa.gov, lynch.jason@epa.gov, hazel.cathcart@ec.gc.ca, hazel.cathcart@ec.gc.ca). For North American critical loads for eutrophication, contact Michael D. Bell (michael_d_bell@nps.gov, michael_d_bell@nps.gov). For European critical loads, contact Thomas Scheuschner (thomas.scheuschner@uba.de, thomas.scheuschner@uba.de).

Model output used in this study constitutes very large datasets (Tb), and their access requires software and documentation for their interpretation and for processing from the storage format into netCDF format. For current information and access to the model output for AQMEII4, contact the AQMEII4 co-leads Stefano Galmarini (stefano.galmarini@ec.europa.eu) and Christian Hogrefe (hogrefe.christian@epa.gov).

Supplement. The supplement related to this article is available online at https://doi.org/10.5194/acp-25-3049-2025-supplement.

Author contributions. PAM: study design and analysis, manuscript writing, GEM-MACH simulations, generation of figures and tables. PC: study analysis support, generation of figures and tables. CH: coordination of modelling team, CMAQ-M3Dry and CMAQ-STAGE simulations, manuscript writing, analysis checking and verification. AA: GEM-MACH simulations. UA: WRF-Chem (IASS) simulations, comments on manuscript. JOB: CMAQ-STAGE (EPA) simulations, comments on manuscript. MDB: critical load exceedance generation from model output, US critical loads for lichen- and herbaceous-community richness. RBe: ENSEMBLE system for submission of model output, coordination of model output library. RBi: ENSEMBLE system for submission of model output, coordination of model output library. TB: WRF-Chem (IASS) simulations. HC: North American critical load exceedance generation for aquatic and forest ecosystems, comments on manuscript. OEC: comments on manuscript. AH: WRF-Chem (UCAR) simulations, comments on manuscript. IK: comments on manuscript, discussions on observation data. RK: LOTOS-EUROS simulations. AL: WRF-Chem (IASS) simulations, comments on paper. JAL: US aquatic-ecosystem critical loads, contributions to North American critical load exceedances. KM: WRF-Chem (IASS) simulations. JLPC: WRF-Chem (UPM) simulations. JP: CMAQ-M3Dry simulations. YHR: WRF-Chem (UCAR) simulations, comments on manuscript. RSJ: WRF-Chem (UPM) simulations, reanalysis of WRF-Chem output. DS: discussions on initial AQMEII4 work, including the work described in this manuscript. TS: European critical load exceedance analysis, design of common format for critical load exceedance bar charts, comments on manuscript. RSS: CMAQ (Hertfordshire) simulations, comments on manuscript. SG: ENSEMBLE model output submission system coordination, co-chairing regular meetings at which the manuscript was discussed. PAM, CH, OEC, DS, SG: AQMEII4 steering committee coordination, manuscript discussion.

Competing interests. At least one of the (co-)authors is a member of the editorial board of *Atmospheric Chemistry and Physics*. The peer-review process was guided by an independent editor, and the authors also have no other competing interests to declare.

Special issue statement. This article is part of the special issue "AQMEII-4: A detailed assessment of atmospheric deposition processes from point to the regional-scale models". It is not associated with a conference.

Disclaimer. The views expressed in this article are those of the authors and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency.

Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Acknowledgements. We gratefully acknowledge the members of the AQMEII4 steering committee who were not co-authors on the current work (Christopher Holmes, Lisa Emberson, Johannes Flemming, Sam Silva, Johannes Bieser, Jason Ducker, and Martijn Schaap) for facilitating the analysis described in this paper by designing and coordinating regional-scale air quality model simulations that provide diagnostic insights into modeled dry deposition. The first author would also like to acknowledge the assistance of Junhua Zhang of ECCC for emissions processing of the GEM-MACH simulations and Amanda Cole of ECCC and Julian Aherne of Trent University for initial discussions on the critical loads part of the project.

Review statement. This paper was edited by Joshua Fu and reviewed by two anonymous referees.

References

Abdul-Razzak, H. and Ghan, S. J.: A parameterization of aerosol activation: 2. Multiple aerosol types, J. Geophys. Res.-Atmos., 105, 6837–6844, https://doi.org/10.1029/1999JD901161, 2000.

Aherne, J. and Jeffries, D.: Critical Load Assessments and Dynamic Applications for Lakes in North America, In W. de Vries, J.-P. Hettelingh, and Posch, M. (Eds.): Critical Loads and Dynamic Risk Assessments: Nitrogen, Acidity and Metals in Terrestrial and Aquatic Ecosystems, Springer Netherlands, 485–503, https://doi.org/10.1007/978-94-017-9508-1_19, 2015.

AIRBASE: Download of historic Airbase air quality data (2000–2012), Discover Map Series, AIRBASE [data set], https://discomap.eea.europa.eu/map/fme/AirQualityExportAirBase.htm, last access: 6 March 2025.

Akingunola, A., Makar, P. A., Zhang, J., Darlington, A., Li, S.-M., Gordon, M., Moran, M. D., and Zheng, Q.: A chemical transport model study of plume-rise and particle size distribution for the Athabasca oil sands, Atmos. Chem. Phys., 18, 8667–8688, https://doi.org/10.5194/acp-18-8667-2018, 2018.

Anderson, J., Hardy, E., Roach, J., and Witmer, R.: A land use and land cover classification system for use with remote sensor data. U.S. Geological Survey, Geological survey professional paper 96, https://pubs.usgs.gov/publication/pp964 (last access: 3 March 2025), 1976.

Anlauf, K., Li, S.-M., Leaitch, R., Brook, J., Hayden, K., Toom-Sauntry, D., and Wiebe, A.: Ionic composition and size characteristics of particles in the Lower Fraser Valley: Pacific 2001 field study, Atmos. Environ., 40, 2662–2675, https://doi.org/10.1016/j.atmosenv.2005.12.027, 2006.

Appel, K. W., Bash, J. O., Fahey, K. M., Foley, K. M., Gilliam, R. C., Hogrefe, C., Hutzell, W. T., Kang, D., Mathur, R., Murphy, B. N., Napelenok, S. L., Nolte, C. G., Pleim, J. E., Pouliot, G. A., Pye, H. O. T., Ran, L., Roselle, S. J., Sarwar, G., Schwede, D. B., Sidi, F. I., Spero, T. L., and Wong, D. C.: The Community Multiscale Air Quality (CMAQ) model versions 5.3 and 5.3.1: system updates and evaluation, Geosci. Model Dev., 14, 2867–2897, https://doi.org/10.5194/gmd-14-2867-2021, 2021.

Banzhaf, S. , Schaap,M., Kerschbaumer, A., Reimer, E., Stern, R., Van Der Swaluw, E., and Builtjes, P.: Implementation and evaluation of pH-dependent cloud chemistry and wetdeposition in the chemical transport model REM-Calgrid, Atmos. Environ., 49, 378–390, https://doi.org/10.1016/j.atmosenv.2011.10.069, 2012.

Barriopedro, D., Fischer, E. M., Luterbacher, J., Trigo, R. M., and Garcia-Herrera, R.: The hot summer of 2010: redrawing the temperature record map of Europe, Science, 332, 220–224, https://doi.org/10.1126/science.1201224, 2011.

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

Belair, S., Crevier, L.-P., Mailhot, J., Bilodeau, B., and Delage, Y.: Operational implementation of the ISBA land surface scheme in the Canadian regional weather forecast model. Part I: warm season results, J. Hydrometeorol., 4, 352–370, https://doi.org/10.1175/1525-7541(2003)4<352:OIOTIL>2.0.CO;2, 2003a.

Belair, S., Brown, R., Mailhot, J., Bilodeau, B., and Crevier, L.-P.: Operational implementation of the ISBA land surface scheme in the Canadian regional weather forecast model. Part II: cold season results, J. Hydrometeorol., 4, 371–386, https://doi.org/10.1175/1525-7541(2003)4<371:OIOTIL>2.0.CO;2, 2003b.

Bermejo, R. and Conde, J.: A conservative quasi-monotone semi-Lagrangian scheme, Mon. Weather Rev., 130, 423–430, https://doi.org/10.1175/1520-0493(2002)130<0423:ACQMSL>2.0.CO;2, 2002.

Binkowski, F. S. and Roselle, S. J.: Models-3 Community Multiscale Air Quality (CMAQ) model aerosol component 1. Model description, J. Geophys. Res., 108, 4183, https://doi.org/10.1029/2001JD001409, 2003.

Binkowski, F. S. and Shankar, U.: The Regional Particulate Matter Model: 1. Model description and preliminary results, J. Geophys. Res., 100, 26191–26209, https://doi.org/10.1029/95JD02093, 1995.

Blakeslee, R. J., Mach, D. M., Bateman, M. G., and Bailey, J. C.: Seasonal variations in the lightning diurnal cycle and implications for the global electric circuit, Atmos. Res., 135–136, 228–243, https://doi.org/10.1016/j.atmosres.2012.09.023, 2014.

Bobbink, R., Loran, C., and Tomassen, H.: Review and revision of empirical critical loads of nitrogen for Europe, Dessau-Roßlau, UBA TEXTE 02/2022, https://www.umweltbundesamt.de/en/publikationen/review-revision-of-empirical-critical-loads-of (last access: 3 March 2025), 2022.

Bower, K. N. and Choularton, T. W.: A parameterisation of the effective radius of ice free clouds for use in global climate mod-

els, Atmos. Res., 27, 305–339, https://doi.org/10.1016/0169-8095(92)90038-C, 1992.

Briggs, G. A.: Plume Rise and Buoyancy Effects, in Atmospheric Science and Power Production, Chap. 8, edited by: Randerson, D., Office of Health and Environmental Research and Office of Energy Research, United States Department of Energy, Washington, D.C., 372–366, https://www.osti.gov/servlets/purl/6503687 (last access: 10 March 2025), 1984.

Carter, W. P. L.: Development of the SAPRC-07 chemical mechanism, Atmos. Environ., 44, 5324-=5335, https://doi.org/10.1016/j.atmosenv.2010.01.026, 2010.

Cathcart, H., Aherne, J., Jeffries, D. S., and Scott, K. A.: Critical loads of acidity for 90,000 lakes in northern Saskatchewan: A novel approach for mapping regional sensitivity to acidic deposition, Atmos. Environ., 146, 290–299, https://doi.org/10.1016/j.atmosenv.2016.08.048, 2016.

Cathcart, H., Aherne, J., Moran, M. D., Savic-Jovcic, V., Makar, P. A., and Cole, A.: Estimates of critical loads and exceedances of acidity and nutrient nitrogen for mineral soils in Canada for 2014–2016 average annual sulfur and nitrogen atmospheric deposition, Biogeosciences, 22, 535–554, https://doi.org/10.5194/bg-22-535-2025, 2025.

Chapman, E. G., Gustafson Jr., W. I., Easter, R. C., Barnard, J. C., Ghan, S. J., Pekour, M. S., and Fast, J. D.: Coupling aerosol-cloud-radiative processes in the WRF-Chem model: Investigating the radiative impact of elevated point sources, Atmos. Chem. Phys., 9, 945–964, https://doi.org/10.5194/acp-9-945-2009, 2009.

Chaumerliac, N.: Evaluation des Termes de Captation Dynamique dans un Modele Tridimensionel à Mesoechelle de Lessivage de L'Atmo-sphere, PhD thesis, Univ. de Clermont II, U.E.R. de Rech. Sci. et Tech., 1984.

Chen, X., Day, D., Schichtel, B., Malm, W., Matzoll, A.K., Mjica., J., McDade, C. E., Hardison, E. D., Hardison, D. L., Walters, S., Van De Water, M., and Collett Jr., J. L.: Seasonal ambient ammonia and ammonium concentrations in a pilot IMPROVE NHx monitoring network in the western United States, Atmos. Environ., 91, 118–126, 2014.

Clifton, O. E., Fiore, A. M., Munger, J. W., Malyshev, S., Horowitz, L. W., Shevliakova, E., Paulot, F., Murray, L. T., and Griffin, K. L.: Interannual variability in ozone removal by a temperate deciduous forest, Geophys. Res. Lett., 44, 542–552, https://doi.org/10.1002/2016GL070923, 2017.

Clifton, O. E., Fiore, A. M., Massman, W. J., Baublitz, C. B., Coyle, M., Emberson, L., Fares, S., Farmer, D. K., Gentine, P., Gerosa, G, Guenther, A. B., Helmig, D., Lombardozzi, D. L., Munger, J. W., Patton, E. G., Pusede, S. E., Schwede, D. B., Silva, S. J., Sörgel, M., Steiner, A. L., and Tai, A. P. K.: Dry deposition of ozone over land: processes, measurement, and modeling, Rev. Geophys., 58, e2019RG000670, https://doi.org/10.1029/2019RG000670, 2020a.

Clifton, O. E., Paulot, F., Fiore, A. M., Horowitz, L. W., Correa, G., Baublitz, C. B., Fares, S., Goded, I., Goldstein, A. H., Gruening, C., Hogg, A. J., Loubet, B., Mammarella, I., Munger, J. W., Neil, L., Stella, P., Uddling, J., Vesala T., and Weng, E.: Influence of dynamic ozone dry deposition on ozone pollution, J. Geophys. Res.-Atmos., 125, e2020JD032398, https://doi.org/10.1029/2020JD032398, 2020b.

Clifton, O. E., Schwede, D., Hogrefe, C., Bash, J. O., Bland, S., Cheung, P., Coyle, M., Emberson, L., Flemming, J., Fredj, E., Galmarini, S., Ganzeveld, L., Gazetas, O., Goded, I., Holmes, C. D., Horváth, L., Huijnen, V., Li, Q., Makar, P. A., Mammarella, I., Manca, G., Munger, J. W., Pérez-Camanyo, J. L., Pleim, J., Ran, L., San Jose, R., Silva, S. J., Staebler, R., Sun, S., Tai, A. P. K., Tas, E., Vesala, T., Weidinger, T., Wu, Z., and Zhang, L.: A single-point modeling approach for the intercomparison and evaluation of ozone dry deposition across chemical transport models (Activity 2 of AQMEII4), Atmos. Chem. Phys., 23, 9911–9961, https://doi.org/10.5194/acp-23-9911-2023, 2023.

Clifton, O. E., Bauer, S. E., Tsigaridis, K., Aleinov, I., Cowan, T. G., Faluvegi, G., and Kelley, M.: Influence of more mechanistic representation of particle dry deposition on 1850-2000 changes in global aerosol burdens and radiative forcing, J. Adv. Model. Earth Sy., 16, e2023MS003952, https://doi.org/10.1029/2023MS003952, 2024.

Convention on Long-Range Transboundary Air Pollution (CLRTAP): Manual on Methodologies and Criteria for Modelling and Mapping Critical Loads and Levels and Air Pollution Effects, Risks, and Trends. Dessau-Roßlau, UBA TEXTE 109/2023, https://www.umweltbundesamt.de/en/publikationen/manual-on-methodologies-criteria-for-modelling-0 (last access: 3 March 2025), 2023.

Côté, J., Gravel, S., Méthot, A., Patoine, A., Roch, M., and Staniforth, A.: The operational CMC/MRB global environmental multiscale (GEM) model. Part 1: design considerations and formulation, Mon. Weather Rev., 126, 1373–1395, https://doi.org/10.1175/1520-0493(1998)126<1373:TOCMGE>2.0.CO;2, 1998.

de Vries, W., Hettelingh, J.-P., and Posch, M.: Critical Loads and Dynamic Risk Assessments: Nitrogen, Acidity and Metals in Terrestrial and Aquatic Ecosystems, Springer, Dordrecht, the Netherlands, 647 pp., https://link.springer.com/book/10.1007/978-94-017-9508-1 (last access: 3 March 2025), 2015.

Duarte, N., Pardo, L. H., and Robin-Abbott, M. J.: Susceptibility of forests in the northeastern USA to nitrogen and sulfur deposition: critical load exceedance and forest health, Water Air Soil Pollution, 22, 1355, https://link.springer.com/article/10.1007/s11270-012-1355-6 (last access: 3 March 2025), 2013.

Dupont, J., Clair, T. A., Gagnon, C., Jeffries, D. S., Kahl, J. S., Nelson, S. J., and Peckenham, J. M.: Estimation of Critical Loads of Acidity for Lakes in Northeastern United States and Eastern Canada, Environ. Monit. Assess., 109, 275–292, https://doi.org/10.1007/s10661-005-6286-x, 2005.

Easter, R. C., Ghan, S. J., Zhang, Y., Saylor, R. D., Chapman, E. G., Laulainen, N. S., Abdul-Razzak, H., Leung, L. R., Bian, X., and Zaveri, R. A.: MIRAGE: Model description and evaluation of aerosols and trace gases, J. Geophys. Res., 109, D20210, https://doi.org/10.1029/2004JD004571, 2004.

Emerson, E. W., Hodshire, A. L., DeBolt, H. M., Bilsmack, K. R., Pierce, J. R., McMeeking, G. R., and Farmber, D. K.: Revisiting particle dry deposition and its role in radiative effect estimates, P. Natl. Acad. Sci. USA, 117, 26076–26082, https://doi.org/10.1073/pnas.2014761117, 2020.

Emmons, L. K., Walters, S., Hess, P. G., Lamarque, J.-F., Pfister, G. G., Fillmore, D., Granier, C., Guenther, A., Kinnison, D., Laepple, T., Orlando, J., Tie, X., Tyndall, G., Wiedinmyer, C., Baughcum, S. L., and Kloster, S.: Description and eval-

uation of the Model for Ozone and Related chemical Tracers, version 4 (MOZART-4), Geosci. Model Dev., 3, 43–67, https://doi.org/10.5194/gmd-3-43-2010, 2010.

European Environment Agency (EEA): CORINE Land Cover 2000, https://trac.osgeo.org/geonetwork/raw-attachment/ticket/650/GeoNetwork-chrome-simple.pdf (last access: 26 December 2023), 2000.

European Environment Agency (EEA): CLC2006 technical guidelines. EEA Technical report 17/2007. ISBN 978-92-9167-968-3, https://www.eea.europa.eu/ds_resolveuid/6ee7e1406e694f6adacf6cd349aff89a (last access: 3 March 2025), 2007.

European Monitoring and Evaluation Program (EMEP): EMEP [data set], https://www.emep.int/, last access: 6 March 2025.

Fahey, K. M., Carlton, A. G., Pye, H. O. T., Baek, J., Hutzell, W. T., Stanier, C. O., Baker, K. R., Appel, K. W., Jaoui, M., and Offenberg, J. H.: A framework for expanding aqueous chemistry in the Community Multiscale Air Quality (CMAQ) model version 5.1, Geosci. Model Dev., 10, 1587–1605, https://doi.org/10.5194/gmd-10-1587-2017, 2017.

Fakhraei, H. A., Driscoll, C. T., Selvendiran, P., DePinto, J. V., Bloomfield, J., Quinn, S., and Rowell, H. C.: Development of a total maximum daily load (TMDL) for acid-impaired lakes in the Adirondack region of New York, Atmos. Environ., 95, 277–287, https://doi.org/10.1016/j.atmosenv.2014.06.039, 2014.

Fountoukis, C. and Nenes, A.: ISORROPIA II: a computationally efficient thermodynamic equilibrium model for K^+–Ca^{2+}–Mg^{2+}–NH_4^+–Na^+–SO_4^{2-}––NO_3^-––Cl^-–H_2O aerosols, Atmos. Chem. Phys., 7, 4639–4659, https://doi.org/10.5194/acp-7-4639-2007, 2007.

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., and Huang, X.: MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets, Remote Sens. Environ., 114, 168–182, https://doi.org/10.1016/j.rse.2009.08.016, 2010.

Fu, J. S., Carmichael, G. R., Dentener, F., Aas, W., Andersson, C., Barrie, L. A., Cole, A., Galy-Lacaux, C., Geddes, J., Itahashi, S., Kanakidou, M., Labrador, L., Paulot, F., Schwede, D., Tan, J., and Vet., R.: Improving estimates of sulphur, nitrogen and ozone total deposition through multi-model and measurement-model fusion approaches, Env. Sci. Tech., 56, 2134–2142, https://pubs.acs.org/doi/10.1021/acs.est.1c05929, 2022.

Galmarini, S., Hogrefe, C., Brunner, D. Makar, P., and Baklanov, A.: Preface, Atmos. Environ., 115, 340–344, https://doi.org/10.1016/j.atmosenv.2015.06.009, 2015.

Galmarini, S., Koffi, B., Solazzo, E., Keating, T., Hogrefe, C., Schulz, M., Benedictow, A., Griesfeller, J. J., Janssens-Maenhout, G., Carmichael, G., Fu, J., and Dentener, F.: Technical note: Coordination and harmonization of the multi-scale, multi-model activities HTAP2, AQMEII3, and MICS-Asia3: simulations, emission inventories, boundary conditions, and model output formats, Atmos. Chem. Phys., 17, 1543–1555, https://doi.org/10.5194/acp-17-1543-2017, 2017.

Galmarini, S., Rao, S. T., and Steyn, D. G.: Preface, Atmos. Environ., 53, 1–3, https://doi.org/10.1016/j.atmosenv.2012.03.001, 2012.

Galmarini, S., Makar, P., Clifton, O. E., Hogrefe, C., Bash, J. O., Bellasio, R., Bianconi, R., Bieser, J., Butler, T., Ducker, J., Flemming, J., Hodzic, A., Holmes, C. D., Kioutsioukis, I., Kranenburg, R., Lupascu, A., Perez-Camanyo, J. L., Pleim, J., Ryu, Y.-H., San Jose, R., Schwede, D., Silva, S., and Wolke, R.: Technical note: AQMEII4 Activity 1: evaluation of wet and dry deposition schemes as an integral part of regional-scale air quality models, Atmos. Chem. Phys., 21, 15663–15697, https://doi.org/10.5194/acp-21-15663-2021, 2021.

Geiser, L. H., Nelson, P. R., Jovan, S. E., Root, H. T., and Clark, C. M.: Assessing Ecological Risks from Atmospheric Deposition of Nitrogen and Sulfur to US Forests Using Epiphytic Macrolichens, Diversity, 11, 87, https://doi.org/10.3390/d11060087, 2019.

Gery, M. W., Whitten, G. Z., Killus, J. P., and Dodge, M.C.: A photochemical kinetics mechanism for urban and regional scale computer modeling, J. Geophys. Res., 94, 12925–12956, https://doi.org/10.1029/JD094iD10p12925, 1989.

Geupel, M., Loran, C., Scheuschner, T., and Wohlgemuth, L.: CCE Status Report, Dessau-Roßlau, UBA TEXTE 135/2022, https://www.umweltbundesamt.de/en/publikationen/cce-status-report-2022 (last access: 21 December 2023), 2022.

Ghahreman, R., Gong, W., Makar, P. A., Lupu, A., Cole, A., Banwait, K., Lee, C., and Akingunola, A.: Modeling below-cloud scavenging of size-resolved particles in GEM-MACHv3.1, Geosci. Model Dev., 17, 685–707, https://doi.org/10.5194/gmd-17-685-2024, 2024.

Ginoux, P., Chin, M., Tegen, I., Prospero, J. M., Holben, B., Dubovik, O., and Lin, S.-J.: Sources and distributions of dust aerosols simulated with the GOCART model, J. Geophys. Res.-Atmos., 106, 20255–20273, https://doi.org/10.1029/2000JD000053, 2001.

Giorgi, F.: A particle dry-deposition parameterization scheme for use in tracer transport models, J. Geophys. Res.-Atmos., 91, 9794–9806, https://doi.org/10.1029/JD091iD09p09794, 1986.

Girard, C., Plante, A., Desgagne, M., McTaggart-Cowan, R., Cote, J., Charron, M., Gravel, S., Lee, V., Patoine, A., Qaddouri, A., Roch, M., Spacek, L., Tanguay, M., Vaillancourt, P. A., and Zadra, A.: Staggered vertical discretization of the Canadian Environmental Multiscale (GEM) model using a coordinate of the log-hydrostatic-pressure type, Mon. Weather Rev., 142, 1183–1196, https://doi.org/10.1175/MWR-D-13-00255.1, 2014.

Gong, S. L., Barrie, L. A., and Blanchet, J.-P.: Modeling sea-salt aerosols in the atmosphere: 1. Model development, J. Geophys. Res.-Atmos., 102, 3805–3818, https://doi.org/10.1029/96JD02953, 1997.

Gong, S. L., Barrie, L. A., Blanchet, J.-P., von Salzen, K., Lohmann, U., Lesins, G., Spacek, L., Zhang, L. M., Girard, E., Lin, H., Leaitch, R., Leighton, H., Chylek, P., and Huang, P.: Canadian Aerosol Module: a size-segregated simulation of atmospheric aerosol processes for climate and air quality models. 1. Module development, J. Geophys. Res., 108, 4007, https://doi.org/10.1029/2001JD002002, 2003.

Government of Canada: National Air Pollution Surveillance Program, Government of Canada [data set], https://www.canada.ca/en/environment-climate-change/services/air-pollution/monitoring-networks-data/national-air-pollution-program.html, last access: 6 March 2025a.

Government of Canada: National atmospheric chemistry database and analysis facility, Government of Canada [data set] https://www.canada.ca/en/environment-climate-change/services/air-pollution/monitoring-networks-data/

national-atmospheric-chemistry-database.html, last access: 6 March 2025b.

Grell, G. A. and Devenyi, D.: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques, Geophys. Res. Lett., 29, 38-1–38-4, https://doi.org/10.1029/2002GL015311, 2002.

Grell, G. A. and Freitas, S. R.: A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling, Atmos. Chem. Phys., 14, 5233–5250, https://doi.org/10.5194/acp-14-5233-2014, 2014.

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W .C., and Eder, B.: Fully coupled "online" chemistry in the WRF model, Atmos. Environ., 39, 6957–6976, https://doi.org/10.1016/j.atmosenv.2005.04.027, 2005.

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), Atmos. Chem. Phys., 6, 3181–3210, https://doi.org/10.5194/acp-6-3181-2006, 2006.

Hardacre, C., Wild, O., and Emberson, L.: An evaluation of ozone dry deposition in global scale chemistry climate models, Atmos. Chem. Phys., 15, 6419–6436, https://doi.org/10.5194/acp-15-6419-2015, 2015.

Henriksen, A. and Posch, M.: Steady-state models for calculating critical loads of acidity for surface waters, Water Air Soil Poll., 1, 375–398, https://doi.org/10.1023/A:1011523720461, 2001.

Henriksen, A., Dillon, P. J., and Aherne, J., Critical loads of acidity for surface waters in south-central Ontario, Canada: Regional application of the Steady-State Water Chemistry (SSWC) model, Can. J. Fish. Aquat. Sci, 59, 1287–1295, https://doi.org/10.1139/f02-092, 2002.

Hogrefe, C., Galmarini, S., Solazzo, E., Bianconi, R., Bellasio, R., Liu, P., and Mathur, R.: Continental-Scale Analysis of Atmospheric Deposition Over North America and Europe Using the AQMEII Database, in: Air Pollution Modeling and its Application XXVI, ITM 2018, edited by: Mensink, C., Gong, W., and Hakami, A., Springer Proceedings in Complexity, Springer, Cham, https://doi.org/10.1007/978-3-030-22055-6_48, 2020.

Hogrefe, C., Bash, J. O., Pleim, J. E., Schwede, D. B., Gilliam, R. C., Foley, K. M., Appel, K. W., and Mathur, R.: An analysis of CMAQ gas-phase dry deposition over North America through grid-scale and land-use-specific diagnostics in the context of AQMEII4, Atmos. Chem. Phys., 23, 8119–8147, https://doi.org/10.5194/acp-23-8119-2023, 2023.

Hogrefe, C., Galmarini, S., Makar, P. A., Kioutsioukis, I., Clifton, O. E., Alyuz, U., Bash, J. O., Bellasio, R., Bianconi, R., Butler, T., Cheung, P., Hodzic, A., Kranenburg, R., Lupascu, A., Momoh, K., Perez-Camanyo, J. L., Pleim, J. E., Ryu, Y.-H., San Jose, R., Schaap, M., Schwede, D. B., and Sokhi, R.: A Diagnostic Intercomparison of Modeled Ozone Dry Deposition Over North America and Europe Using AQMEII4 Regional-Scale Simulations, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2025-225, 2025.

Hong, S. Y.: A new stable boundary-layer mixing scheme and its impact on the simulated East Asian summer monsoon, Q. J. Roy. Meteor. Soc., 136, 1481–1496, https://doi.org/10.1002/qj.665, 2010.

Hong, S. Y., Noh, Y., and Dudhia, J.: A new vertical diffusion package with an explicit treatment of entrainment processes, Mon. Weather Rev., 134, 2318–2341, https://doi.org/10.1175/MWR3199.1, 2006.

Huang, L., Zhu, Y., Zhai, H., Xue, S., Zhu, T., Shao, Y., Liu, Z., Emery, C., Yarwood, G., Wang, Y., Fu, J., Zhang, K., and Li, L.: Recommendations on benchmarks for numerical air quality model applications in China – Part 1: $PM_{2.5}$ and chemical species, Atmos. Chem. Phys., 21, 2725–2743, https://doi.org/10.5194/acp-21-2725-2021, 2021.

Hyder, P., Edwards, J. M., Allan., R. P., Hewitt, H. T., Bracegirdle, T. J., Gregory, J. M., Wood, R. A., Meijers, A. J. S., Mulcahy, J., Field, P., Furtado, K., Bodas-Salcedo, A., Williams, K. D., Copesy, D., Josey, S. A., Liu, C., Robverts, C. D., Sanchez, C., Ridley, J., Thrope, L., Hardiman, S. C., Mayer, M., Berry, D. I., and Belcher, S. E.: Critical Southern Ocean climate model biases traced to atmospheric model cloud errors, Nat. Commun., 9, 3625, https://doi.org/10.1038/s41467-018-05634-2, 2018.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long–lived greenhouse gases: Calculations with the AER radiative transfer models, J. Geophys. Res., 113, D13103, https://doi.org/10.1029/2008JD009944, 2008.

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M., and Suttie, M.: The CAMS reanalysis of atmospheric composition, Atmos. Chem. Phys., 19, 3515–3556, https://doi.org/10.5194/acp-19-3515-2019, 2019.

Janjić, Z. I.: Nonsingular Implementation of the Mellor–Yamada Level 2.5 Scheme in the NCEP Meso model, Technical Report, National Centers for Environmental Prediction, Office Note No. 437, 61 pp., http://www.emc.ncep.noaa.gov/officenotes/newernotes/on437.pdf (last access: 3 March 2025), 2001.

Jeffries, D. S., Semkin, R. G., Gibson, J. J., and Wong, I.: Recently surveyed lakes in northern Manitoba and Saskatchewan, Canada: Characteristics and critical loads of acidity, J. Limnol., 69, 45, https://doi.org/10.4081/jlimnol.2010.s1.45, 2010.

Kain, J. S.: The Kain-Fritsch convective parameterization: an update, J. Appl. Meteorol., 43, 170–181, https://doi.org/10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2, 2004.

Kain, J. S. and Fritsch, J. M.: A one-dimensional entraining/detraining plume model and its application in convective parameterizations, J. Atmos. Sci., 47, 2784–2802, https://doi.org/10.1175/1520-0469(1990)047<2784:AODEPM>2.0.CO;2, 1990.

Knote, C., Hodzic, A., Jimenez, J. L., Volkamer, R., Orlando, J. J., Baidar, S., Brioude, J., Fast, J., Gentner, D. R., Goldstein, A. H., Hayes, P. L., Knighton, W. B., Oetjen, H., Setyan, A., Stark, H., Thalman, R., Tyndall, G., Washenfelder, R., Waxman, E., and Zhang, Q.: Simulation of semi-explicit mechanisms of SOA formation from glyoxal in aerosol in a 3-D model, Atmos. Chem. Phys., 14, 6213–6239, https://doi.org/10.5194/acp-14-6213-2014, 2014.

Knote, C., Hodzic, A., and Jimenez, J. L.: The effect of dry and wet deposition of condensable vapors on secondary organic aerosols concentrations over the continental US, Atmos. Chem. Phys., 15, 1–18, https://doi.org/10.5194/acp-15-1-2015, 2015.

Kuenen, J. J. P., Visschedijk, A. J. H., Jozwicka, M., and Denier van der Gon, H. A. C.: TNO-MACC_II emission inventory; a multi-year (2003–2009) consistent high-resolution European emission inventory for air quality modelling, Atmos. Chem. Phys., 14, 10963–10976, https://doi.org/10.5194/acp-14-10963-2014, 2014.

Lawrence, G. B., Sullivan, T. J., Burns, D. A., Bailey, S. W., Cosby, B. J., Dovciak, M., Ewing, H. A., McDonnell, T. C., Minocha, R., Riemann, R., Quant, J., Rice, K. C., Siemion, J., and Weathers, K.: Acidic deposition along the Appalachian Trail corridor and its effects on acid-sensitive terrestrial and aquatic resources: Results of the Appalachian Trail MEGA-transect atmospheric deposition effects study. Natural Resource Report NPS/NRSS/ARD/NRR–2015/996. National Park Service, Fort Collins, Colorado, https://irma.nps.gov/DataStore/Reference/Profile/2223220 (last access: 3 March 2025), 2015.

Li, J. and Barker, H. W.: A radiation algorithm with correlated k-distribution. Part I: local thermal equilibrium, J. Atmos. Sci., 62, 286–309, https://doi.org/10.1175/JAS-3396.1, 2005.

Luecken, D. J., Yarwood, G., and Hutzell, W. H.: Multipollutant of ozone, reactive nitrogen and HAPs across the continental US with CMAQ-CB6, Atmos. Environ., 201, 62–72, https://doi.org/10.1016/j.atmosenv.2018.11.060, 2019.

Lynch, J. A., Phelan, J., Pardo, L. H., McDonnell, T. C., Clark, C. M., and Bell, M. D.: Detailed Documentation of the National Critical Load Database (NCLD) for U.S. Critical Loads of Sulfur and Nitrogen, version 3.2.1, National Atmospheric Deposition Program, Wisconsin State Laboratory of Hygiene, Madison, WI, https://nadp.slh.wisc.edu/filelib/claddb/DB_Version/ (last access: 3 March 2025), 2022.

Makar, P. A., Wiebe, H. A., Staebler, R. M., Li, S. M., and Anlauf, K.: Measurement and modeling of particle nitrate formation, J. Geophys. Res.-Atmos., 103, 13095–13110, https://doi.org/10.1029/98JD00978, 1998.

Makar, P. A., Vouchet, V. S., and Nenes, A.: Inorganic chemistry calculations using HETV – a vectorized solver for the SO_4^{2-}–NO_3^-–NH_4^+ system based on the ISORROPIA algorithms, Atmos. Environ., 37, 2279–2294, https://doi.org/10.1016/S1352-2310(03)00074-8, 2003.

Makar, P. A., Nissen, R., Teakles, A., Zhang, J., Zheng, Q., Moran, M. D., Yau, H., and diCenzo, C.: Turbulent transport, emissions and the role of compensating errors in chemical transport models, Geosci. Model Dev., 7, 1001–1024, https://doi.org/10.5194/gmd-7-1001-2014, 2014.

Makar, P. A., Gong, W., Milbrandt, J., Hogrefe, C., Zhang, Y., Curci, G., Žabkar, R., Im, U. Balzarini, A., Baro, R., Bianconi, R., Cheung, P., Forkel, R., Gravel, S., Hirtl, M., Honzak, L., Hou, A., Jiménez-Guerrero, P., Langer, M., Moran, M. D., Pabla, B., Pérez, J. L., Pirovano, G., San José, R., Tuccella, P., Werhahn, J., Zhang, J., and Galmarini, S.: Feedbacks between air pollution and weather, Part 1: Effects on weather, Atmos. Environ., 115, 442-469, https://doi.org/10.1016/j.atmosenv.2014.12.003, 2015a.

Makar, P. A., Gong, W., Hogrefe, C., Zhang, Y., Curci, G., Žabkar, R., Milbrandt, J., Im, U., Balzarini, A., Baró, R. , Bianconi, R., Cheung, P., Forkel, R., Gravel, S., Hirtl, M., Honzak, L., Hou, A., Jiménez-Guerrero, P., Langer, M., Moran, M.D., Pabla, B., Pérez, J.L., Pirovano, G., San José, R., Tuccella, P., Werhahn, J., Zhang, J., and Galmarini, S.: Feedbacks between air pollution and weather, part 2: Effects on chemistry, Atmos. Environ., 115, 499–526, https://doi.org/10.1016/j.atmosenv.2014.10.021, 2015b.

Makar, P. A., Staebler, R. M., Akingunola, A., Zhang, J., McLinden, C., Kharol, S. K., Pabla, B., Cheung, P., and Zheng, Q.: The effects of forest canopy shading and turbulence on boundary layer ozone, Nat. Commun., 8, 15243, https://doi.org/10.1038/ncomms15243, 2017.

Makar, P. A., Akingunola, A., Aherne, J., Cole, A. S., Aklilu, Y.-A., Zhang, J., Wong, I., Hayden, K., Li, S.-M., Kirk, J., Scott, K., Moran, M. D., Robichaud, A., Cathcart, H., Baratzedah, P., Pabla, B., Cheung, P., Zheng, Q., and Jeffries, D. S.: Estimates of exceedances of critical loads for acidifying deposition in Alberta and Saskatchewan, Atmos. Chem. Phys., 18, 9897–9927, https://doi.org/10.5194/acp-18-9897-2018, 2018.

Makar, P. A., Stroud, C., Akingunola, A., Zhang, J., Ren, S., Cheung, P., and Zheng, Q.: Vehicle-induced turbulence and atmospheric pollution, Atmos. Chem. Phys., 21, 12291–12316, https://doi.org/10.5194/acp-21-12291-2021, 2021.

Manders, A. M. M., Builtjes, P. J. H., Curier, L., Denier van der Gon, H. A. C., Hendriks, C., Jonkers, S., Kranenburg, R., Kuenen, J. J. P., Segers, A. J., Timmermans, R. M. A., Visschedijk, A. J. H., Wichink Kruit, R. J., van Pul, W. A. J., Sauter, F. J., van der Swaluw, E., Swart, D. P. J., Douros, J., Eskes, H., van Meijgaard, E., van Ulft, B., van Velthoven, P., Banzhaf, S., Mues, A. C., Stern, R., Fu, G., Lu, S., Heemink, A., van Velzen, N., and Schaap, M.: Curriculum vitae of the LOTOS–EUROS (v2.0) chemistry transport model, Geosci. Model Dev., 10, 4145–4173, https://doi.org/10.5194/gmd-10-4145-2017, 2017.

Marchuk, G.I.: Splitting and alternating direction methods, Handbook of Numerical Analysis, 1, 197–462, 1990.

Massad, R.-S., Nemitz, E., and Sutton, M. A.: Review and parameterisation of bi-directional ammonia exchange between vegetation and the atmosphere, Atmos. Chem. Phys., 10, 10359–10386, https://doi.org/10.5194/acp-10-10359-2010, 2010.

McDonnell, T. C., Cosby, B. J., and Sullivan, T. J.: Regionalization of soil base cation weathering for evaluating stream water acidification in the Appalachian Mountains, USA, Environ. Pollut., 162, 338–344, https://doi.org/10.1016/j.envpol.2011.11.025, 2012.

McDonnell, T. C., Sullivan, T. J., Hessburg, P. F., Reynolds, K. M., Povak, N. A., Cosby, B. J., Jackson, W., and Salter, R. B.: Steady-state sulfur critical loads and exceedances for protection of aquatic ecosystems in the U.S. southern Appalachian Mountains, J. Environ. Manage., 146, 407–419, https://doi.org/10.1016/j.jenvman.2014.07.019, 2014.

McDonnell, T. C., Driscoll, C. T., Sullivan, T. J., Burns, D. A., Baldigo, B. P., Shao, S., and Lawrence, G. B.: Regional target loads of atmospheric nitrogen and sulfur deposition for the protection of stream and watershed soil resources of the Adirondack Mountains, USA, Environ. Pollut., 281, 117110, https://doi.org/10.1016/j.envpol.2021.117110, 2021.

McNulty, S. G., Cohen, E. C., Moore Myers, J. A., Sullivan, T. J., and Li, H.: Estimates of critical acid loads and exceedances for forest soils across the conterminous United States, Environ. Pollut., 149, 281–292, https://doi.org/10.1016/j.envpol.2007.05.025, 2007.

McNulty, S. G., Cohen, E. C., and Myers, J. A. M.: Climate change impacts on forest soil critical acid loads and exceedances at

a national scale, in: Forest Health Monitoring: National Status, Trends, and Analysis 2010, edited by: Potter, K. M. and Conkling, B. L., Gen. Tech. Rep. SRS-GTR-176, Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station, 176, 95–108, https://research.fs.usda.gov/treesearch/43980 (last access: 7 March 2025), 2013.

Milbrandt, J. A. and Morrison, H.: Parameterization of Cloud Microphysics Based on the Prediction of Bulk Ice Particle Properties. Part III: Introduction of Multiple Free Categories, J. Atmos. Sci., 73, 975–995, https://doi.org/10.1175/JAS-D-15-0204.1, 2016.

Miller, E.: Steady-state critical loads and exceedances for terrestrial and aquatic ecosystems in the northeastern United States. Technical report, National Park Service, Air Resources Division, https://www-f.nescaum.org/documents/steady-state-critical-loads-and-exceedance-for-terrestrial-and-aquatic-ecosystems-in-the-northeastern-united-states (last access: 3 March 2025), 2012.

Miller, S. J., Makar, P. A., and Lee, C. J.: HETerogeneous vectorized or Parallel (HETPv1.0): an updated inorganic heterogeneous chemistry solver for the metastable-state NH_4^+–Na^+–Ca^{2+}–K^+–Mg^{2+}–SO_4^{2}––NO_3^-–Cl^-–H_2O system based on ISORROPIA II, Geosci. Model Dev., 17, 2197–2219, https://doi.org/10.5194/gmd-17-2197-2024, 2024.

Morrison, H. and Milbrandt, J. A.: Parameterization of Cloud Microphysics Based on the Prediction of Bulk Ice Particle Properties. Part I: Scheme Description and Idealized Tests, J. Atmos. Sci., 72, 287–311, https://doi.org/10.1175/JAS-D-14-0065.1, 2015.

Morrison, H., Thompson, G., and Tatarskii, V.: Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: comparison of one- and two-moment schemes, Mon. Weather Rev., 137, 991–1007, https://doi.org/10.1175/2008MWR2556.1, 2009.

Nakanishi, M. and Niino, H.: An Improved Mellor–Yamada Level-3 Model: Its Numerical Stability and Application to a Regional Prediction of Advection Fog, Bound.-Lay. Meteorol., 119, 397–407, https://doi.org/10.1007/s10546-005-9030-8, 2006.

National Atmospheric Deposition Program (NADP): National Trends Network, National Atmospheric Deposition Program [data set], https://nadp.slh.wisc.edu/networks/national-trends-network/, last access: 6 March 2025a.

National Atmospheric Deposition Program (NADP): AMoN Data Retrieval page, National Atmospheric Deposition Program [data set], https://nadp2.slh.wisc.edu/data/AMoN/, last access: 6 March 2025b.

NCDC: National Centers for Environmental Information, Gridded Maps, https://www.ncei.noaa.gov/access/monitoring/ghcn-gridded-products/maps/, last access: 19 November 2024.

Nenes, A., Pandis, S. N., and Pilinis, C.: Isorropia: A new thermodynamic equilibrium model for multiphase multicomponent inorganic aerosols, Aquat. Geochem., 4, 123–152, https://doi.org/10.1023/A:1009604003981, 1998.

Neu, J. L. and Prather, M. J.: Toward a more physical representation of precipitation scavenging in global chemistry models: cloud overlap and ice physics and their impact on tropospheric ozone, Atmos. Chem. Phys., 12, 3289–3310, https://doi.org/10.5194/acp-12-3289-2012, 2012.

Nilsson, J. and Grennfelt, P. (Eds.): Reprint of the workshop report on critical loads for sulphur and nitrogen, UN-ECE and the Nordic Council of Ministers, 16 pp., https://www.umweltbundesamt.de/sites/default/files/medien/4292/dokumente/nilssongrennfelt_1988.pdf (last access: 7 March 2025), 1988.

Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Ek. M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., and Xia, Y.: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements, J. Geophys. Res., 116, D12109, https://doi.org/10.1029/2010JD015139, 2011.

Ott, L. E., Pickering, K. E., Stenchikov, G. L., Allen, D. J., DeCaria, A. J., Ridley, B., Line, R.-F., Lang, S., and Tao, W.-K.: Production of lightning NO_x and its vertical distribution calculated from three-dimensional cloud-scale chemical transport model simulations, J. Geophys. Res.-Atmos., 115, D04301, https://doi.org/10.1029/2009JD011880, 2010.

Paulot, F., Jacob., D. J., Johnson, M. T., Bell, T. G., Baker, A. R., Keene, W. C., Lima, I. D., Doney, S. C., and Stock, C. A.: Global oceanic emission of ammonia: Constraints from seawater and atmospheric observations, Global Geochem. Cy., 29, 1165–1178, https://doi.org/10.1002/2015GB005106, 2015.

Paulot, F., Malyshev, S., Nguyen, T., Crounse, J. D., Shevliakova, E., and Horowitz, L. W.: Representing sub-grid scale variations in nitrogen deposition associated with land use in a global Earth system model: implications for present and future nitrogen deposition fluxes over North America, Atmos. Chem. Phys., 18, 17963–17978, https://doi.org/10.5194/acp-18-17963-2018, 2018.

Paulot, F., Stock, C., John, J. G., Zadeh, N., and Horowitz, L. W.: Ocean Ammonia Outgassing: Modulation by CO_2 and Anthropogenic Nitrogen Deposition, J. Adv. Model. Earth Sy., 12, e2019MS002026, https://doi.org/10.1029/2019MS002026, 2020.

Pfister, G. and Barth, M.: WRF-Chem status announcement, https://www2.acom.ucar.edu/wrf-chem (last access: 6 March 2025), 2024.

Phelan, J., Belazid, S., Kurz, D., Guthrie, S., Cajka, J., Sverdrup, H., and Waite, R.: Estimation of soil base cation weathering rates with the PROFILE model to determine critical loads of acidity for forested ecosystems in Pennsylvania, USA: pilot application of a potential national methodology, Water Air Soil Poll., 225, 2109–2128, https://doi.org/10.1007/s11270-014-2109-4, 2014.

Phelan, J., Balzid, S., Jones, P., Cajka, J., Buckley, J., and Clark, C.: Assessing the effects of climate change and air pollution on soil properties and plant diversity in sugar maple – beech – yellow birch hardwood forests in the northeastern United States: model simulations from 1900 to 2100, Water Air Soil Pollut., 22, 30 , https://doi.org/10.1007/s11270-016-2762-x, 2016.

Pleim, J. E., Ran, L., Appel, W., Shephard, M. W., and Cady-Pereira, K.: New bidirectional ammonia flux model in an air quality model coupled with an agricultural model, J. Adv. Model. Earth Sy., 11, 2934–2957, https://doi.org/10.1029/2019MS001728, 2019.

Pleim, J. E., Ran, L., Saylor, R. D., Willison, J., and Binkowski, F. S.: A new aerosol dry deposition model for air quality and climate modeling, J. Adv. Model. Earth Sy., 14, e2022MS003050, https://doi.org/10.1029/2022MS003050, 2022.

Posch, M., de Smet, P. A., Hettelingh, J.-P., and Downing, R. J.: Modelling and mapping of critical thresholds in Europe: Status Report 2001, Citeseer, https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=a585c6eff5f0c93938e4ef21874d4dc9425fd282 (last access: 22 December 2023), 2001.

Price, C., Penner, J., and Prather, M.: NO_x from lightning: 1. global distribution based on lightning physics, J. Geophys. Res.-Atmos., 102, 529–5941, https://doi.org/10.1029/96JD03504, 1997.

Pruppacher, H. R. and Klett, J. D.: Growth of Cloud Drops by Collision and Coalescence, in: Microphysics of Clouds and Precipitation, Springer, Dordrecht, https://doi.org/10.1007/978-94-009-9905-3_15, 1978.

Ran, L., Cooter, E., Yang, D., Benson, V., Yuan, Y., Hanna, A., and Garcia, V.: User's Guide for the Fertilizer Emission Scenario Tool for CMAQ (FEST-C) Version 1.4. U.S. EPA Office of Research and Development, Washington, DC, https://www.cmascenter.org/fest-c/documentation/1.4/FESTC_v1_4_UserManual.pdf (last access: 3 March 2025), 2018.

Reinds, G. J., Thomas D., Posch M., and Slootweg J.: Critical loads for eutrophication and acidification for European terrestrial ecosystems, Final report, Dessau-Roßlau, https://pure.iiasa.ac.at/id/eprint/17341/ (last access: 3 March 2025), 2021.

Rubin, H. J., Fu, J. S., Dentener, F., Li, R., Huang, K., and Fu, H.: Global nitrogen and sulfur deposition mapping using a measurement–model fusion approach, Atmos. Chem. Phys., 23, 7091–7102, https://doi.org/10.5194/acp-23-7091-2023, 2023.

Ryu, Y.-H. and Min, S.-K.: Improving wet and dry deposition of aerosols in WRF-Chem: Updates to below-cloud scavenging and coarse-particle dry deposition, J. Adv. Model. Earth Sy., 14, e2021MS002792, https://doi.org/10.1029/2021MS002792, 2022.

Sandu, A. and Sander, R.: Technical note: Simulating chemical systems in Fortran90 and Matlab with the Kinetic PreProcessor KPP-2.1, Atmos. Chem. Phys., 6, 187–195, https://doi.org/10.5194/acp-6-187-2006, 2006.

Schaap, M., van Loon, M., ten Brink, H. M., Dentener, F. J., and Builtjes, P. J. H.: Secondary inorganic aerosol simulations for Europe with special attention to nitrate, Atmos. Chem. Phys., 4, 857–874, https://doi.org/10.5194/acp-4-857-2004, 2004.

Scheffe, R. D., Lynch, J. A., Reff, A., Hubbell, B., Greaver, T. L., and Smith, J. T.: The Aquatic Acidification Index: A New Regulatory Metric Linking Atmospheric and Biogeochemical Models to Assess Potential Aquatic Ecosystem Recovery, Water Air Soil Pollut., 225, 1838, https://doi.org/10.1007/s11270-013-1838-0, 2014.

Schmuck, G., San-Miguel-Ayanz, J., Camia, A., Durrant, T., Boca, R., Whitmore, C. J., Liberta, G., Corti, P., and Ernst, S.: Forest fires in Europe, Middle East and North Africa, 2011, Joint Research Centre, European Commission, Ispra, Italy, https://publications.jrc.ec.europa.eu/repository/handle/JRC74152 (last access: 7 March 2025), 2012.

Schwede, D. B. and Lear, G. G.: A novel hybrid approach for estimating total deposition in the United States, Atmos. Environ., 92, 207–220, https://doi.org/10.1016/j.atmosenv.2014.04.008, 2014.

Scott, K., Wissel, B., Gibson, J., and Birks, S.: Chemical characteristics and acid sensitivity of boreal headwater lakes in northwest Saskatchewan, J. Limnol., 69, 33–44, https://doi.org/10.4081/jlimnol.2010.s1.33, 2010.

Shao, Y., Ishizuka, M., Mikami, M., and Leys, J. F.: Parameterization of size-resolved dust emission and validation with measurements, J. Geophys. Res., 116, D08203, https://doi.org/10.1029/2010JD014527, 2011.

Simkin, S. M., Allen, E. B., Bowman, W. D., Clark, C. M., Belnap, J., Brooks, M. L., Cade, B. S., Collins, S. L., Geiser, L. H., Gilliam, F. S., Jovan, S. E., Pardo, L. H., Schulz, B. K., Stevens, C. I., Suding, K. N., Throop, H. L., and Waller, D. M.: Conditional vulnerability of plant diversity to atmospheric nitrogen deposition across the United States, P. Natl. Acad. Sci. USA, 113, 4086-4091, https://doi.org/10.1073/pnas.1515241113, 2016.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Wang, W., Powers, J. G., Duda, M. G., Barker, D. M., and Huang, X.-Y.: A Description of the Advanced Research WRF Version 4, NCAR Tech. Note NCAR/TN-556+STR, 145 pp., https://doi.org/10.5065/1dfh-6p97, 2019.

Slinn, W. G. N.: Predictions for particle deposition to vegetative canopies, Atmos. Environ., 16, 1785–1794, https://doi.org/10.1016/0004-6981(82)90271-2, 1982.

Slinn, W. G. N.: Precipitation Scavenging, in Atmospheric Science and Power Production, Chap. 11, edited by: Randerson, D., Office of Health and Environmental Research and Office of Energy Research, United States Department of Energy, Washington, D.C., 466–532, https://www.osti.gov/servlets/purl/6503687 (last access: 10 March 2025), 1984.

Solazzo, E., Riccio, A., Van Dingenen, R., Valentini, L., and Galmarini, S.: Evaluation and uncertainty estimation of the impact of air quality modelling on crop yields and premature deaths using a multi-model ensemble, Sci. Total Environ., 633, 1437–1452, https://doi.org/10.1016/j.scitotenv.2018.03.317, 2018.

Sørensen, B., Kaas, E., and Korsholm, U. S.: A mass-conserving and multi-tracer efficient transport scheme in the online integrated Enviro-HIRLAM model, Geosci. Model Dev., 6, 1029–1042, https://doi.org/10.5194/gmd-6-1029-2013, 2013.

Stockwell, W. R. and Lurmann, F. W.: Intercomparison of the ADOM and RADM gas-phase chemical mechanisms, Electric Power Institute Topical Report, Electric Power Institute, Palo Alto, California, 323 pp., 1989.

Stroud, C., Moran, M., Makar, P., Gong, W., Gong, S., Mourneau, G., Bouchet, V. Dann, T., Wang, D., and Huang, L.: Impact of Updates to BEIS v3 Boreal Forest Emissions on Canadian Air Quality Forecasts, 2nd International Workshop on Air Quality Forecasting Research, Quebec, Canada, 16 November 2010.

Stroud, C. A., Makar, P. A., Zhang, J., Moran, M. D., Akingunola, A., Li, S.-M., Leithead, A., Hayden, K., and Siu, M.: Improving air quality model predictions of organic species using measurement-derived organic gaseous and particle emissions in a petrochemical-dominated region, Atmos. Chem. Phys., 18, 13531–13545, https://doi.org/10.5194/acp-18-13531-2018, 2018.

Sullivan, T. J, Cosby, B. J. Tonnessen, K. A., and Clow, D. W.: Surface water acidification responses and critical loads of sulfur and nitrogen deposition in Loch Vale watershed, Colorado, Water Resour. Res., 41, W01021, https://doi.org/10.1029/2004WR003414, 2005.

Sullivan, T. J., Cosby, B. J., Driscoll, C. T., McDonnell, T. C., and Herlihy, A. T.: Target loads of atmospheric sulfur deposition to

protect terrestrial resources in the Adirondack Mountains, New York against biological impacts caused by soil acidification, J. Environ. Stud. Sci. 1, 301–314, https://link.springer.com/article/10.1007/s13412-011-0062-8 (last access: 3 March 2025), 2011.

Sullivan, T. J., Cosby, B. J., McDonnell, T. C., Porter, E. M., Blett, T., Haeuber, R., Huber, C. M., and Lynch, J.: Critical loads of acidity to protect and restore acid-sensitive streams in Virginia and West Virginia, Water Air Soil Pollut., 223, 5759–5771, https://doi.org/10.1007/s11270-012-1312-4, 2012.

Sundqvist, H., Berge, E., and Kristjansson, J. E.: Condensation and cloud parameterization studies with a mesoscale numerical weather prediction model, Mon. Weather Rev., 117, 1641–1657, https://doi.org/10.1175/1520-0493(1989)117<1641:CACPSW>2.0.CO;2, 1989.

Sverdrup, H. and de Vries, W.: Calculating critical loads for acidity with the simple mass balance method, Water Air Soil Poll., 72, 143–162, https://doi.org/10.1007/BF01257121, 1994.

Sverdrup, H. and Warfvinge, P.: The role of weathering and forestry in determining the acidity of Lakes in Sweden, Water Air Soil Pollut., 52, 71–78, https://doi.org/10.1007/BF00283115, 1990.

Sverdrup, H., De Vries, W., and Henriksen, A.: Mapping critical loads (Miljörapport 14), Nordic Council of Ministers, 124 pp., 1990.

Timmermans, R., van Pinxteren, D., Kranenburg, R., Hendriks, C., Fomba, K. W., Herrmann, H., and Schaap, M.: Evaluation of modelled LOTOS-EUROS with observational based PM_{10} source attribution, Atmos. Environ. X, 14, 100173, https://doi.org/10.1016/j.aeaoa.2022.100173, 2022.

TNO: LOTOS-EUROS Open Source Version, TNO [code], https://airqualitymodeling.tno.nl/lotos-euros/open-source-version/, last access: 6 March 2025.

US EPA: US EPA Office of Research and Development, CMAQ (5.3.2). Zenodo [code], https://doi.org/10.5281/zenodo.4081737, 2020.

US EPA: NHDPlus (National Hydrography Dataset Plus), https://www.epa.gov/waterdata/nhdplus-national-hydrography-dataset-plus, last access: 26 December 2023.

US EPA: Air Quality System (AQS), United States Environmental Protection Agency [data set], https://www.epa.gov/aqs, last access: 6 March 2025.

Van Zanten, M. C., Sauter, F. J., Wichink Kruit, R. J., Van Jaarsveld, J. A., and Van Pul, W. A. J.: Description of the DEPAC module: dry deposition modelling with DEPAC_GCN2010 RIVM Rep., https://rivm.openrepository.com/handle/10029/256555 (last access: 3 March 2025), 2010.

Vehkamaki, H., Kulmala, M., Napari, I., Lehtinen, K. E. J., Timmreck, C., Noppel, M., and Laaksonen, A.: An improved parameterization for sulfuric acid – water nucleation rates for tropospheric and stratospheric conditions, J. Geophys. Res., 107, 4622, https://doi.org/10.1029/2002JD002184, 2002.

Venkatram, A. and Pleim, J.: The electrical analogy does not apply to modeling dry deposition of particles, Atmos. Environ., 33, 3075–3076, https://doi.org/10.1016/S1352-2310(99)00094-1, 1999.

Vermont Department of Environmental Conservation (VDEC): Total Maximum Daily Loads: Acid Impaired Lakes, Watershed Management Division, 103 South Main Street, Building 10 North, Waterbury, VT 05671-0408, https://dec.vermont.gov/sites/dec/files/documents/WSMD_mapp_TMDL_2003_Acid.pdf (last access: 7 March 2025), 2003.

Vermont Department of Environmental Conservation (VDEC): Total Maximum Daily Loads: Acid Impaired Lakes, Watershed Management Division, 103 South Main Street, Building 10 North, Waterbury, VT 05671-0408, https://dec.vermont.gov/sites/dec/files/documents/WSMD_mapp_TMDL_2004_Acid.pdf (last access: 7 March 2025), 2004.

Vermont Department of Environmental Conservation (VDEC): Total Maximum Daily Loads: Acid Impaired Lakes, Watershed Management Division, 103 South Main Street, Building 10 North, Waterbury, VT 05671-0408, https://dec.vermont.gov/sites/dec/files/documents/WSMD_mapp_TMDL_2012_Acid.pdf (last access: 7 March 2025), 2012.

Vivanco, M. G., Theobald, M. R., García-Gómez, H., Garrido, J. L., Prank, M., Aas, W., Adani, M., Alyuz, U., Andersson, C., Bellasio, R., Bessagnet, B., Bianconi, R., Bieser, J., Brandt, J., Briganti, G., Cappelletti, A., Curci, G., Christensen, J. H., Colette, A., Couvidat, F., Cuvelier, C., D'Isidoro, M., Flemming, J., Fraser, A., Geels, C., Hansen, K. M., Hogrefe, C., Im, U., Jorba, O., Kitwiroon, N., Manders, A., Mircea, M., Otero, N., Pay, M.-T., Pozzoli, L., Solazzo, E., Tsyro, S., Unal, A., Wind, P., and Galmarini, S.: Modeled deposition of nitrogen and sulfur in Europe estimated by 14 air quality model systems: evaluation, effects of changes in emissions and implications for habitat protection, Atmos. Chem. Phys., 18, 10199–10218, https://doi.org/10.5194/acp-18-10199-2018, 2018.

Vizuete, W., Nielsen-Gammon, J., Dickey, J., Couzo, E., Blanchard, C., and Breitenbach, P.: Meteorological based parameters and ozone exceedances in Houston and other cities in Texas, J. Air Waste Manage., 72, 969–984, https://doi.org/10.1080/10962247.2022.2064004, 2022.

Vukovich, J. M. and Pierce, T.: The Implementation of BEIS3 within the SMOKE modeling framework, Proceedings of the 11th International Emissions Inventory Conference, Atlanta, Georgia, USA, 13–16 April 2015, https://www.epa.gov/sites/default/files/2015-10/documents/vukovich.pdf (last access: 3 March 2025), 2002.

Wang, X., Zhang, L., and Moran, M. D.: Development of a new semi-empirical parameterization for below-cloud scavenging of size-resolved aerosol particles by both rain and snow, Geosci. Model Dev., 7, 799–819, https://doi.org/10.5194/gmd-7-799-2014, 2014.

Wesely, M. L.: Parameterization of surface resistances to gaseous dry deposition in regional-scale numerical models, Atmos. Environ., 23, 1293–1304, https://doi.org/10.1016/j.atmosenv.2007.10.058, 1989.

Whitfield, C. J., Aherne, J., Watmough, S. A., Dillon, P. J., and Clair, T. A.: Recovery from acidification in Nova Scotia: Temporal trends and critical loads for 20 headwater lakes, Can. J. Fish. Aquat. Sci., 63, 1504–1514, https://doi.org/10.1139/f06-053, 2006.

Whitten, G., Hogo, H., and Killus, J.: The carbon bond mechanism for photochemical smog, Environ. Sci. Technol., 14, 14690–14700, 1980.

Wichink Kruit, R. J., Schaap, M., Sauter, F. J., van Zanten, M. C., and van Pul, W. A. J.: Modeling the distribution of ammonia across Europe including bi-directional surface–atmosphere exchange, Biogeosciences, 9, 5261–5277, https://doi.org/10.5194/bg-9-5261-2012, 2012.

Wiedinmyer, C., Sakulyanontvittaya, T., and Guenther, A.: MEGAN FORTRAN code v2.04 User Guide, https://www.acom.ucar.edu/webt/MEGAN/MEGANguideFORTRAN204.pdf (last access: 12 December 2023), 2007.

Williams, J. R.: The EPIC model, in: Computer models in watershed hydrology, Chap. 25, edited by: Singh, V. P., Water Resources Publications, Littleton, CO, 909–1000, ISBN-13: 978-1-887201-74-2, 1995.

Williston, P., Aherne, J., Watmough, S., Marmorek, D., Hall, A., de la Cueva Bueno, P., Murray, C., Henolson, A., and Laurence, J. A.: Critical levels and loads and the regulation of industrial emissions in northwest British Columbia, Canada, Atmos. Environ., 146, 311–323, https://doi.org/10.1016/j.atmosenv.2016.08.058, 2016.

Wu, Z., Schwede D. B., Vet R., Walker J. T., Shaw M., Staebler R., and Zhang L.: Evaluation and intercomparison of five North American dry deposition algorithms at a mixed forest site, J. Adv. Model. Earth Sy., 10, 1571–1586, https://doi.org/10.1029/2017MS001231, 2018.

Xu, L., Pye, H. O. T., He, J., Chen, Y., Murphy, B. N., and Ng, N. L.: Experimental and model estimates of the contributions from biogenic monoterpenes and sesquiterpenes to secondary organic aerosol in the southeastern United States, Atmos. Chem. Phys., 18, 12613–12637, https://doi.org/10.5194/acp-18-12613-2018, 2018.

Yarwood, G., Jung, J., Whitten, G., Heo, G., Mellberg, J., and Estes, M.: Updates to the Carbon Bond Mechanism for Version 6 (CB6), in: 9th Annual CMAS Conference, Chapel Hill, NC, 11–13 October 2010, https://cmascenter.org/conference/2010/abstracts/emery_updates_carbon_2010.pdf (last access: 3 March 2025), 2010.

Young, T. R. and Boris, J. P.: A numerical technique for solving stiff ordinary differential equations associated with the chemical kinetics of reactive-flow problems, J. Phys. Chem., 81, 2424–2427, 1977.

Zaveri, R. A. and Peters, L. K.: A new lumped structure photochemical mechanism for large-scale applications, J. Geophys. Res., 104, 30387–30415, https://doi.org/10.1029/1999JD900876, 1999.

Zaveri, R. A., Easter, R. C., Fast, J. D., and Peters, L. K.: Model for Simulating Aerosol Interactions and Chemistry (MOSAIC), J. Geophys. Res., 113, D13204, https://doi.org/10.1029/2007JD008782, 2008.

Zhang, J., Moran, M. D., Makar, P. A., and Kharol, S.: Examination of MODIS Leaf Area Index (LAI) Product for Air Quality Modelling, 19th Annual CMAS Conference, 26–30 October 2020, https://www.cmascenter.org/conference/2020/slides/ZhangJ_MODIS_LAI_CMAS_2020.pdf (last access: 19 December 2023), 2020.

Zhang, L., Gong, S., Padro, J., and Barrie, L.: A size-segregated particle dry deposition scheme for an atmospheric aerosol module, Atmos. Environ., 35, 549–560, https://doi.org/10.1016/S1352-2310(00)00326-5, 2001.

Zhang, L., Moran, M. D., Makar, P. A., Brook, J. R., and Gong, S.: Modelling gaseous dry deposition in AURAMS: a unified regional air-quality modelling system, Atmos. Environ., 36, 537–560, https://doi.org/10.1016/S1352-2310(01)00447-2, 2002.

Zhang, L., Brook, J. R., and Vet, R.: A revised parameterization for gaseous dry deposition in air-quality models, Atmos. Chem. Phys., 3, 2067–2082, https://doi.org/10.5194/acp-3-2067-2003, 2003.

Zhang, L., Wright, L. P., and Asman, W. A. H.: Bi-directional air-surface exchange of atmospheric ammonia: A review of measurements and a development of a big-leaf model for applications in regional-scale air-quality models, J. Geophys. Res., 115, D20310, https://doi.org/10.1029/2009JD013589, 2010.

Zhang, Y., Foley, K. M., Schwede, D. B., Bash, J. O., Pinto, J. P., and Dennis, R. L.: A measurement-model fusion approach for improved wet deposition maps and trends, J. Geophys. Res.-Atmos., 124, 4237–4251, https://doi.org/10.1029/2018JD029051, 2019.