https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Critical Load Exceedances for North America and Europe using an Ensemble of Models and an Investigation of Causes for Environmental Impact Estimate Variability: An AQMEII4 Study

Paul A. Makar[1], Philip Cheung[1], Christian Hogrefe[2], Ayodeji Akingunola[1], Ummugulsum Alyuz-Ozdemir[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], Iannis Kioutsioukis[10], Richard Kranenburg[11], Aura Lupascu[6,12], Jason A. Lynch[13], Kester Momoh[3], Juan Luis Perez-Camanyo[14], Jonathan Pleim[2], Young-Hee Ryu[15], Roberto San Jose[14], Donna Schwede[2,16], Thomas Scheuschner[17], Mark Shephard[1], Ranjeet S. Sokhi[3], Stefano Galmarini[18]

[1] Environment and Climate Change Canada, Toronto, Canada
[2] ORD, U.S. EPA, Research Triangle Park, NC, USA
[3] Centre for Climate Change Research (C3R), U. Hertfordshire, UK
[4] Air Resources Division, US National Park Service, USA
[5] Enviroware srl, Concorezzo, MB, Italy
[6] Research Institute for Sustainability – Helmholtz Centre Potsdam, Germany
[7] NASA Goddard Institute for Space Studies, New York, NY, USA
[8] Center for Climate Systems Research, Columbia University, New York, NY, USA
[9] NCAR, Boulder, CO, USA
[10] Department of Physics, University of Patras, Patras, Greece
[11] NOASR, Utrecht, the Netherlands
[12] ECMWF, Bonn, Germany
[13] OAR, US EPA, USA
[14] Tech. U. of Madrid (UPM), Madrid, Spain
[15] Yonsei University, Seoul, South Korea
[16] (Retired)
[17] CCE/Federal Environment Agency, Germany
[18] JRC, European Commission, Ispra, Italy

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

Abstract

Exceedances of critical loads for deposition of sulphur (S) and nitrogen (N) to different ecosystems were estimated using European and North American ensembles of air quality models, under Phase 4 of the Air Quality Model Evaluation International Initiative (AQMEII4), to identify where 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 Europe 0.125° resolution domains, which were then used to calculate critical load exceedances. New, targeted deposition diagnostics implemented in AQMEII4 allowed 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.

New datasets for North American critical loads for acidity for forest soil water and aquatic ecosystems 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 in the direction of the ensemble exceedance change

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

between 2010 and 2016. The North American ensemble also predicted a decrease in both 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 both predictions of total N deposition and the change in N deposition, and hence in the predicted eutrophication exceedances between the two 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 two 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 sulphur 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 observed wet deposition of sulphate, 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 sulphate, 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 further development of some of the models within the ensemble suggests these particulate biases may also be ameliorated via the incorporation of multiphase hydrometeor scavenging. Annual sulphur 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 emissions fluxes from forested areas.

Based on these results, an increased process-research focus is therefore recommended for the following model processes and on 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 emissions sources should be included in regional chemical transport models. The choice of 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

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Introduction

The concept of a Critical load (CL) was first proposed as a means for evaluating the ecosystem impacts of the deposition of sulphur and nitrogen in response to the Convention on Long-Range Transboundary Air Pollution (CLRTAP), an international agreement for 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 sulphur- 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 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 Supplemental Information (SI).

The creation of critical loads for acidification, and the calculation of their exceedances is based on the concept of 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 sulphur 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 sulphate ion. Every mole of deposited sulphur is assumed to be associated with two negative charges as the sulphate ion, $SO_4^{2-}(aq)$, 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 to 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 model (SMB) often used to define surface water critical loads for terrestrial ecosystems (Sverdrup and DeVries, 1994), BC_{dep} includes the release of soil base cations due to weathering, non-marine chloride deposition, harvesting of base cation and/or nitrogen-containing biomass, denitrification, nitrogen immobilization in the rooting zone, run-off volume, and a critical value of the non-sodium base cation to 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 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, 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)

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

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 for 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, a sensitive plant or animal species acceptable limit for nitrogen leaching (nitrogen in soil water), and an ecosystem-dependent denitrification fraction (CLRTAP, 2023). The acceptable limits for nitrogen concentrations in soil can range from 6.5 down to 0.2 mg N l^{-1}, depending on 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. CTMs numerically integrate the system of time-dependent differential equations describing the rates of change of 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 often the selection of a specific process representation within a CTM is often based on comparisons of the output of entire CTM relative to surface or satellite monitoring data. This latter approach may allow compensating errors in process representation to take place (c.f. 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, through process analysis, attempts to determine the key causes of these model deposition estimate differences.

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

The ongoing reevaluation and improvement of CTMs is aided by ensemble model comparisons, where models driven by the same lateral boundary and emissions 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 off-line regional models used for research and public policy support to simulate a common year, 2006, with common emissions inputs, in both North America and Europe, with 22 modelling groups participating (Galmarini *et al.,* 2012). Subsequent phases of AQMEII examined specific issues within the CTM community: AQMEII-2 had as its focus the evaluation of both weather and air-quality predictions for fully coupled, on-line air-quality models, where the particulate matter generated by the models on any given timestep feeds back into the coupled models' weather forecast radiative transfer and cloud formation processes (Galmarini *et al.,* 2015). AQMEII-3 addressed questions of hemispheric transport of air pollutants – the relative contributions of local versus long-range transport towards 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 AQMEII-3, 14 air-quality models were used to calculate oxidized sulphur and oxidized and reduced nitrogen deposition, and hence EU critical load exceedances (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 – however, even within this reduced ensemble, local variations of over a factor of four in both sulphur 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 two 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 four (AQMEII4) were designed partly in response to the large variation in model sulphur 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; Zhiyong 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 eutrophication, 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 sulphur and nitrogen (S_{dep} and N_{dep}): particle dry deposition and the scavenging of particles by depositing hydrometeors. We note that subsequent to 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 underestimating the deposition velocity for

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

particles between 1 and 10 μm). The accumulation mode of atmospheric particles tend to poorly capture the relationship between particle deposition velocity and particle size in 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 sub-μm range of one to two 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 sub-μm 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 PM2.5 positive biases increased in magnitude, while PM10 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 PM2.5 concentrations, and an improvement in PM2.5 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 two orders of magnitude between the previously applied, temperature-based parameterization 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 above two processes, 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 the exceedances estimated using ensembles of model deposition predictions. The bulk of the analysis then examines individual contributions of different sulphur and nitrogen species towards 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

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

analysis in order to improve the performance of individual models, and to reduce the variability between their estimates of critical load exceedances.

Methodology

1.0 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 (forest) ecosystem acidity critical loads for the continent, aquatic ecosystem acidity critical loads combining data from Canada and the USA, and USA-specific sensitive epiphytic lichen species and herbaceous plant species eutrophication critical loads. European CL datasets combined CL information from multiple countries for terrestrial and aquatic ecosystem acidity and terrestrial ecosystem eutrophication. Each CL dataset is described in this section.

1.1 North American Forest Soil Critical Loads of Acidity using the Steady-State Mass Balance Model

Forest soil critical loads maps were assembled from several studies within the U.S. and Canada (Figure 1 and Table 1). Critical loads were (in all but one study) calculated using the Steady-State (or Simple) Mass Balance (SMB) model (Sverdrup & Warfvinge, 1990; Sverdrup & De Vries, 1994) which has simple input parameter requirements and assumes the ecosystem is at long-term equilibrium. The SMB model defines the critical load as a line connecting three points in (S_{dep}, N_{dep}) space, $CL_{max}S$ (the maximum sulphur critical load), $CL_{max}N$ (the maximum nitrogen critical load) and the $CL_{min}N$ (the minimum nitrogen critical load). The regions above the (S_{dep}, N_{dep}) line connecting the points ($CL_{max}S,0$), ($CL_{max}S, CL_{min}N$) and ($0, CL_{max}N$) are said to be in exceedance of the critical load (see Figure 1). $CL_{max}S$ is determined by alkaline inputs to the ecosystem such as base cation deposition (BC_{dep}) and base cation weathering (BC_w) minus acidic inputs (chloride deposition, Cl_{dep}), losses through (non-sodium) base cation uptake through harvesting or grazing (BC_u) (Equation 1), and the critical leaching of the acid neutralizing capacity ($ANC_{le,crit}$, Equation 2).

$$CL_{max}S = BC_{dep} + BC_w - Cl_{dep} - BC_u - ANC_{le,crit} \quad (1)$$

$$ANC_{le,crit} = -Q^{2/3} \cdot \left(1.5 \cdot \frac{Bc_{dep} + Bc_w - BC_u}{K_{gibb} \cdot (Bc/Al)_{crit}}\right) \quad (2)$$

The Acid Neutralizing Capacity refers to the soil's ability to neutralize input fluxes of acidifying ions through the release of cations from the soil into the soil water. The addition of these neutralizing ions to soil water is a process known as leaching. However, the removal of base cations from soil water may also result in damage to plants via reductions in root growth, stem growth and crops, with the extent of damage dependent on the plant species. The plant-species-specific critical base cation to aluminum soil water ratio in equation (2), $(Bc/Al)_{crit}$, is linked to corresponding precent reductions of plant growth. If a larger percent reduction is deemed acceptable, the value of $(B_c/Al)_{crit}$ will be smaller, the magnitude of $ANC_{le,crit}$ will be larger, and the value of $CL_{max}S$ will be larger, and larger amounts of deposition will be required to exceed the critical load. Conversely, if a smaller impact is deemed acceptable, the value of $(B_c/Al)_{crit}$ will be larger, the magnitude of $ANC_{le,crit}$ will be smaller, the value of $CL_{max}S$ will be smaller, and smaller amounts of deposition will be required to exceed the critical load. Examples of $(Bc/Al)_{crit}$ values for different tree types and ground vegetation may be found in CLRTAP (2023), Chapter V, Table V.8). The critical base cation to aluminum ratio, $(B_c/Al)_{crit}$ (multiplied by the gibbsite equilibrium constant K_{gibb}) is thus the chemical criterion usually used to define the acceptable level of potential damage to biota, specifically via the definition of $ANC_{le,crit,}$ which includes the effect of soil runoff (Q).

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

The $CL_{min}N$ represents the long-term removal of N from the ecosystem as defined by nitrogen immobilization (N_i) and uptake (N_u) (Equation 3). The $CL_{max}N$ value is determined using $CL_{min}N$ and $CL_{max}S$, which is divided by unity minus the denitrification fraction (f_{de}) (Equation 4). Deposition points of S_{dep} and N_{dep} which fall outside (above) the critical load exceedance line defined by $Cl_{min}N$, $CL_{max}N$, and $CL_{max}S$ are considered to be in *exceedance of their critical loads* (see Figure 1, Regions 1 through 4). Note that these critical loads may be specific to a political jurisdiction, and hence caution should be applied when considering the critical loads and exceedance maps where there are cross-border discontinuities in data sources, parameterization and methodology, and resolution.

$$CL_{min}N = N_i + N_u \quad (3)$$

$$CL_{max}N = CL_{min}N + \left(\frac{CL_{max}S}{(1-f_{de})}\right) \quad (4)$$

Figure 1 illustrates the manner in which critical loads with respect to acidity are calculated using the SMB methodology. Based on the sulphur and nitrogen deposition amounts (S_{dep}, N_{dep}), the Region in which exceedance is occurring is first defined. The amount of exceedance is defined as the shortest possible path (in eq of deposition) to the shaded "no-exceedance" Region 0 of Figure 1, bordered by the line described above. Deposition amounts which fall above the critical load function defined by Region 0 are considered to be in exceedance of their critical loads. The shape of the critical load function is defined by $CL_{max}S$, $CL_{min}N$ and $CL_{max}N$, which in turn are functions of the ecosystems and at-risk species under consideration.

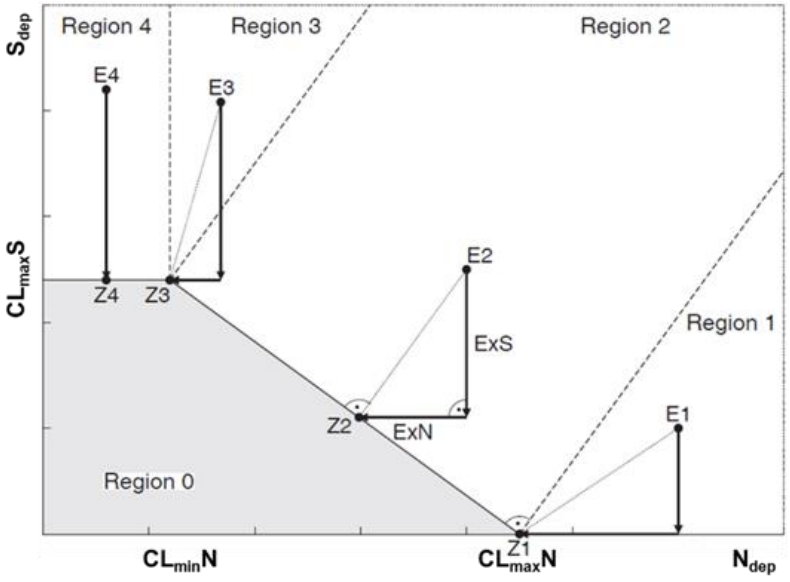


Figure 1. SMB Critical Load Function for acidification, showing exceedance regions 1 through 4 and "below exceedance" region 0. Deposition in exceedance of critical loads correspond to regions 1 through 4, while the grey region encompasses deposition below critical loads. The change in sulphur and nitrogen deposition required to bring a given ecosystem in exceedance to below exceedance is described by ExS, ExN, and the amount in exceedance is the dotted line linking E_i to Z_i. After CLRTAP, 2023, Figure 7.3

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Table 1: Data sources, model types and major parameters for North American forest soil critical loads maps. A database of maps within the U.S.A was provided in National Atmospheric Deposition Program (NADP, 2022). Table adapted from Lynch *et al.* (2022).

Source	Model	Resolution	Extent	Chemical criteria	BC_w approach	Uptake
(McNulty *et al.*, 2007, 2013)	SMB	1 km^2	U.S.A-wide	Bc/Al, Coniferous forest: 1, deciduous forest: 10	Clay correlation - substrate method	Bc_u, N_u
(Duarte *et al.*, 2011, 2013)	SMB	5 km^2	New England	Bc/Al = 10	Clay correlation - substrate method	Bc_u, N_u
(Phelan *et al.*, 2014; data corrected 2016)	SMB	1 m^2	Pennsylvania	Bc/Al =10	PROFILE	Bc_u, N_u
(Sullivan, 2011; Sullivan *et al.*, 2012)	MAGIC	Watershed	Virginia and New York	Bc/Al, Ca/Al = 1 and 10, Bsat = 5 and 10	MAGIC	Bc_u
Cathcart *et al.* (in prep.)	SMB	250 m x 250 m	Canada-wide	Bc/Al = site specific	Soil texture approximation	Bc_u, N_u

1.2 North American Aquatic Ecosystems Acidity Critical Loads

The North American Aquatic Ecosystem acidity critical load dataset constructed here combined individual datasets from the Canada and the USA.

1.2.1 Canadian Aquatic Ecosystem Data

Environment and Climate Change Canada data corresponding to the subset of 2,997 lake surveys which reside within the common AQMEII4 North American 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). The SSWC model has been widely used in regional lake critical load assessments across Europe (e.g. Posch *et al.*, 2001), Canada (e.g. Cathcart *et al.*, 2016; Henriksen *et al.*, 2002; Jeffries *et al.*, 2010; Scott *et al.*, 2010; Whitfield *et al.*, 2006; Williston *et al.*, 2016), and the United States (e.g. Dupont *et al.*, 2005; Miller, 2011). Briefly, the critical load exceedance is defined as the difference between the total sulphur deposition S_{dep} and the acidity critical load value CL(A). The latter is determined from the non-marine, pre-acidification base cation flux ($[BC^*]_0$) minus the Acid Neutralizing Capacity limit (ANC_{limit})for protecting aquatic biota from damage, scaled by the catchment runoff (Q):

$$CL(A) = Q([BC^*]_0 - ANC_{limit}) \quad (5)$$

Where available, a site-specific modelled isotope mass balance estimate of Q (Gibson *et al.*, 2010) was used (n=684) in preference to a Q value derived from a GIS-modelled map approach using regional datasets (Reinds *et al.*, 2015). When Dissolved Organic Carbon (DOC, mgC L^{-1}) values were available (n=2,875) the organic acid adjusted ANC_{limit} ($[ANC]_{oaa}$) was used to include the influence of organic acids in the lake as 1/3 the charge density (m, here set to 10.2 μeq mgC^{-1}) (Lydersen *et al.*, 2004; Hruska *et al.*, 2001),

$$[ANC]_{oaa} = [ANC]_{limit} - \frac{m}{3}DOC \quad (6)$$

Where the lake acid neutralizing capacity $[ANC]_{limit}$ is defined as the excess equivalents of cations – anions in lakewater:

$$[ANC]_{limit} = BC_{le} + NH_{4le} - SO_{4le} - NO_{3le} - Cl_{le} \quad (7)$$

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

BC_{le}, NH_{4le}, SO_{4le}, NO_{3le}, Cl_{le} are the charge equivalents (μeq L^{-1}) of ionic base cations, ammonium, sulphate, nitrate, and chloride in lakewater.

For lakes lacking DOC samples, an ANC_{limit} of 40 μeq L^{-1} was chosen as a conservative value, previously used in regional Canadian assessments (e.g. Henriksen *et al.,* 2002), and based on the response of brown trout (Lien *et al.,* 1996). Since the SSWC model does not consider non-acidifying nitrogen, only sulphur was used to determine exceedance (i.e. exceedance is defined as the total S deposition minus the critical load of Equation (5)).

1.2.2 USA Aquatic Ecosystem Data

Aquatic 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 USA values, sometimes resulting in more than one CL estimate for the same water body). Most critical loads (78%) were determined using the SSWC model as described above and by equations 5 and 7 (Lynch *et al.,* 2022; Scheffe *et al.,* 2014; Dupont *et al.,* 2005, Miller 2011, VDEC (2003, 2004, 2012)). Site-specific catchment Q estimates for these values were based on 30-year Normals that are included as a catchment parameter in the National Hydrography Dataset Plus (NHD+2, US EPA, 2023). The other 22% of critical loads were determined by a dynamic modelling approach (e.g., MAGIC and PnET-BGC models) (Sullivan *et al.,* 2005; Fakhraei *et al.,* 2014; Lawrence *et al.,* 2015) and a combination of dynamic modeling with a regionalization approach (e.g. hurdle/regional regression modeling) to determine the critical load across the landscape (McDonnell *et al.,* 2012, 2014; Sullivan *et al.,* 2012; and McDonnell *et al.,* 2021). Site-specific catchment Q estimates were also used; these were based on the specific research project. An ANC_{limit} of 50 μeq L^{-1} was used for the Eastern USA, with the exception of streams in the Adirondacks Mountain, NY, which used 20 μeq L^{-1} (McDonnell *et al.* 2021) and 20 μeq L^{-1} for the western USA. Organic acid-adjusted ANC_{limit} values were not used in generating the USA CL(A) datasets. In many cases, multiple studies estimated CL(A) for the same lake or stream, leading to multiple CL(A) estimates for a single water body. An average critical load value was therefore used for these waterbodies with more than one critical load. A more detailed description of the USA aquatic critical loads used here can be found in Lynch *et al.,* (2022).

1.3 USA Sensitive Epiphytic Lichen

Critical loads for sensitive epiphytic lichen species richness made use of 9,000 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. Here, Geiser *et al.* (2019) sets a -20% decline in species richness (their "Low ecological risk" critical load) as the level of ecosystem damage that can occur before the loss of species impacts the presence of plentiful forage, nesting materials or insect habitat; hence determining the critical load. The models show that there is a consistent relative response of lichen communities across climates, which results in a single critical load of 3.1 kg-N ha^{-1} yr^{-1} for sensitive epiphytic lichen, which can be applied across all ecosystems in which the lichen can be found. This value was applied to all broadleaf, conifer, or mixed forest landcover types as designated by the National Land Cover Database (NLCD, Dewitz 2021). The original 30m resolution NLCD dataset was aggregated to a 240m resolution grid including all cells with greater than 10% forest cover. Exceedances of the above critical load were calculated for each 240m resolution cell based on the annual deposition of the overlapping 0.125° resolution AQMEII4 CTM model cell.

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

1.4 USA Herbaceous Plants

The USA herbaceous plants dataset uses the critical load of total nitrogen for a decline in herbaceous species community richness, 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 herbaceous species richness response to deposition was employed, to generate critical loads with respect to nitrogen deposition linked to various atmospheric and soil conditions. A model was developed for open canopy ecosystems where the critical load varies with observed soil pH, precipitation, and mean temperature. A second model was developed for closed canopy ecosystems where the critical load varies with observed soil pH alone. The pant level critical loads were mapped across the continental U.S. using land cover from the NLCD. Open canopy systems were defined as the combination of the NLCD grassland and shrubland landcover types, while closed canopy ecosystems were defined as the combination of the NLCD's broadleaf, conifer, or mixed forest landcover classes. The resulting critical loads were aggregated to a 240m grid including all cells with greater than 10% cover. Using the United States Department of Agriculture gridded National Soil Survey Geographic Database (gNATSGO) soil pH dataset (https://www.nrcs.usda.gov/resources/data-and-reports/gridded-national-soil-survey-geographic-database-gnatsgo, last access July 12, 2024), and PRISM temperature and precipitation models (Daly *et al.*, 2008), the 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 $ha^{-1}yr^{-1}$. The two datasets were then merged into a single CL raster using the minimum CL when cells overlapped. Exceedances of the resulting critical loads for nitrogen deposition were then generated using the annual deposition of the overlapping 0.125° resolution AQMEII4 CTM model cell.

1.5 EU: Acidification of Terrestrial Ecosystems

The critical load database and the exceedance calculation 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 was also used within the review process of the Gothenburg protocol. It typically contains critical load values for acidification and eutrophication, and has two different components. The first component is data delivered by the member countries of the International Cooperative Programme on Modelling and Mapping. This data is collected within an officiated "Call for Data" (CfD) process within the framework of the Working Group on Effects (WGE). The most recent CfD was finalized in the year 2021. The methods used to determine acidification loads are country-dependent, but all make use of the Simple Mass Balance as described above (Sverdrup & De Vries, 1994; CLRTAP, 2023). The country-specific detailed methods and participating countries may be found in Geupel *et al.* (2022). If countries do not deliver their own CL data, the CCE fills these data gaps with its own background database (Reinds *et al.*, 2021).

The decision of the chemical criterion used to define exceedance (e.g., critical aluminium concentration, critical pH, and critical base saturation) and the chosen critical limit value is usually country-specific. The background CCE database makes use of a fixed value based on a critical pH value of 4.2.

1.6 EU: Eutrophication of Terrestrial Ecosystems

Critical loads for EU eutrophication ($CL_{nut}N$) are also based on the SMB method applied to nitrogen deposition – (Equation 8). Generally, the methods to derive the parameters of this equation are similar for national datasets and the CCE dataset (e.g. the estimation of the nitrogen uptake (N_u) is linked to growth potential of the vegetation, the fraction of the nitrogen which is denitrified (f_{de}) is connected to the soil

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

type). One major difference occurs when it comes to the derivation of the accepted nitrogen leaching ($N_{le(acc)}$) term. There are two ways to estimate the $N_{le(acc)}$. One way is to simply assign how much nitrogen is allowed to leave the ecosystem based on observations. Another way is to calculate the $N_{le(acc)}$ by using the amount of soil runoff (Q) and multiply it with a critical limit for nitrogen concentration. The latter limits can be linked to negative effects for the related ecosystems (such as fine root damage). The choice of the values for the critical limit for nitrogen is one of the main sources for differences in the modelled EC SMB eutrophication CL (see also CLRTAP, 2023). Another main source for differences in the CL values between countries is the integration of so-called empirical critical loads. These empirical values can be 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. Empirical CL were updated recently and are well documented in Bobbink *et al.* (2022).

$$CL_{nut}N = N_i + N_u + \left(\frac{N_{le\ (acc)}}{1-f_{de}}\right) \quad (8)$$

The CL exceedance was calculated for every available critical load value in the integrated CL database of the CCE (about 4 million EU data points) and later aggregated on the basis of the AQMEII4 deposition grid cells. The resulting EU CLE are 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. The units for critical loads and their exceedances are equivalents per hectare and year, making S and N deposition comparable on their impacts, which is important for acidity CLs.

2.0 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 regional chemical transport model development as of the time simulations were completed (2021).

Model simulations were carried out for the years 2009 and 2010 for the European region, and 2010 and 2016 for North America. North American years were chosen due to policy relevance, with a significant change in SO_2 emissions controls enacted between the two years. European years were chosen due to a large difference in meteorology between 2009 and 2010, hence allowing the effects of potential meteorological on deposition to be estimated. 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° x 0.125° resolution (North America: 23.5° N to 58.5° N, 130° W to 59.5° W; Europe 25° N to 70° N, 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 on-line 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 simulations results.

North American anthropogenic emissions were generated using emissions modelling platforms which included the anthropogenic inventories, temporal and spatial allocation from county or state/province level to native model grids, for each of the two model years, as well as adjustments for specific inventories by year. Emissions processing was carried out by the United States Environmental Protection Agency for the Carbon Bond 6 (revision 3; CB6r3)) and Statewide Air Pollution Research Center -07

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

(SAPRC07) chemical mechanisms (Yarwood *et al.*, 2010; Carter, 2010), and by Environment and Climate Change Canada for the Acid Deposition and Oxidant Mechanism version II (ADOM-II; Stockwell et al., 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' 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.,* 2015), with individual groups using their own emissions data for the portion of their native model grids extending beyond the range of MACC-III emissions grid if necessary.

North American forest fire emissions were generated by combining the US emissions modelling platform values with Canadian data for 2010, while both USA and Canadian data were based on the 2016 emissions modelling platform estimates. These forest fire emissions included criteria air contaminant emission mass, heat flux, and acres burned. Fire plume rise calculations were carried out by individual modeling groups, typically based on large stack plume rise formulae (Briggs, 1971, 1972). European forest fire emissions were provided by the Finnish Meteorological Institute using eight layers from 50 to 6200m. Both North American and European forest fire emissions were chemically disaggregated by the participating modelling groups and mapped on a nearest grid cell basis 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 both EU and NA simulations were taken from 3 hourly, 0.75° x 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.

2.1 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.,* 2020) 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 towards 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 criticalload 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:

$$DFLX_{SOIL} = \left(\frac{(r_{ac}+r_{gs})^{-1}}{(r_s+r_m)^{-1}+(r_{lu})^{-1}+(r_{dc}+r_{cl})^{-1}+(r_{ac}+r_{gs})^{-1}} \right) F \quad (9)$$

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

$$DFLX_{LCAN} = \left(\frac{(r_{dc}+r_{cl})^{-1}}{(r_s+r_m)^{-1}+(r_{lu})^{-1}+(r_{dc}+r_{cl})^{-1}+(r_{ac}+r_{gs})^{-1}} \right) F \quad (10)$$

$$DFLX_{CUT} = \left(\frac{(r_{lu})^{-1}}{(r_s+r_m)^{-1}+(r_{lu})^{-1}+(r_{dc}+r_{cl})^{-1}+(r_{ac}+r_{gs})^{-1}} \right) F \quad (11)$$

$$DFLX_{stom} = \left(\frac{(r_s+r_m)^{-1}}{(r_s+r_m)^{-1}+(r_{lu})^{-1}+(r_{dc}+r_{cl})^{-1}+(r_{ac}+r_{gs})^{-1}} \right) F \quad (12)$$

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 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 equations (9) through (12) being replaced by the deposition velocity of the gas V_d. Note that the formulae for individual models vary from the Wesely (1984) example shown above; see Galmarini *et al.* (2021) for details on the formulae for each of the gas-phase deposition algorithms used in the AQMEII4 regional model ensembles analyzed here.

2.2 Model Parameterization Descriptions

2.2.1 CMAQ-M3Dry, *CMAQ-STAGE, CMAQ (Hertfordshire) – WRF-CMAQ Implementations*

These three models make use of the WRF-CMAQ off-line 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 10km cell resolution (Lambert Conformal Conic projection, 459x299 and 500x681 grid cells, respectively). The CMAQ implementations employed 35 model layers with the lowest layer thickness of ~20m. Both NA models operate in an off-line 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.,* 2018)), 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 phases 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 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 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. 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

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

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.

CMAQ-M3Dry's aerosol dry deposition methodology 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 Stoke's number for vegetated surfaces, modification of the impaction term, scaling of diffusion layer resistance by LAI for the vegetated fraction of each grid cell, and improved mass conservation for the process of gravitational settling (Appel *et al.,* 2021).

CMAQ-STAGE and CMAQ (Hertfordshire)'s aerosol dry deposition methodology followed that of CMAQ-M3Dry, but made use of Slinn (1982) and Zhang *et al.* (2001) for impaction on vegetated surfaces, and Georgi (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 (Galmarini *et al.,* 2021, Table B2, with other implementation details in Hogrefe *et al.,* 2023). The algorithms follow the original approach of Wesely *et al.* (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, 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 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 additional AQMEII4 diagnostic to be incorporated into the CMAQ-STAGE simulations. This allows a 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

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

concentrations due to the differences in evasion and deposition parameterizations between CMAQ-STAGE and EPIC.

2.2.2 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 12km Lambert Conformal Conic projection (400x360 grid cells in the European domain, 480x290 grid cells in the North American domain), 2-way coupling between air-quality and meteorology, a sectional aerosol size distribution representation (4 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 land surface model (Noah-MP, Niu *et al.,* 2011), the Rapid Radiative Transfer Model for radiative transfer calculations (RRTM, Iacono *et al.,* 2008), biogenic emissions using the MEGAN model (Guenther *et al.,* 2006, Wiedenmyer *et al.,* 2007), and the FAST-J algorithm for photolysis rate calculation (Fast *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 the impact of those configurations on deposition results to be studied under AQMEII4. The differences between the model configurations are summarized in Table 2. It should also be noted that WRF-Chem is an on-line 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.

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Table 2. AQMEII4 WRF-Chem Configuration Differences

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 : via Grell and Devenyi (2002); grid-scale following Neu and Prather (2012) for gases, Chapman *et al.* (2009) for aerosols	Grid scale wet deposition following 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 classes, (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	On-line, 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–Janjic, Janic (2001)	Yonsei University (YSU) Hong *et al.* (2006), Hong (2010)	Mellor-Yamada Nakahasi Niino, level 2.5 Nakanishi and Niino (2006)

2.2.3 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 Wesely (1989)-based approach (DEPosition of Acidifying Compounds; DEPAC, Van Zanten *et al.,* 2010). Particle dry deposition was carried out using the approach of Zhang (2001). Wet deposition followed the droplet saturation approach, and cloud chemistry with sulphate formation 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 ~20m and a model top of approximately 8 km. The simulations carried out here made use of a 20x20km 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 Corine2000 Land Cover database (EEA, 2000, 2007).

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

2.2.4 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 (~10km), 83 model levels, biogenic VOCs from BEIS3.09, 3.1.3 (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 *et al.,* 1989), a modified Odum approach for 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 CCMA Rad2 radiative transfer algorithm (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 3 Key differences between the models include:

(1) Similar to CMAQ-M3Dry and CMAQ-STAGE above, the only difference between GEM-MACH (Base) and GEM-MACH (Zhang) is the gas-phase dry deposition scheme employed (GEM-MACH (Base): Makar *et al.,* 2018; GEM-MACH (Zhang): Zhang *et al.,* 2003), though both models employ the Zhang *et al.* (2010) parameterization for ammonia bi-directional fluxes;

(2) GEM-MACH (Base) and GEM-MACH (Zhang) make use of the Emerson *et al.* (2020) dry deposition velocity correction to the approach of Zhang *et al.,* 2001) as well as semi-Lagrangian advection rather than a diffusion equation lower flux boundary condition for particle settling and deposition;

(3) GEM-MACH (Base) and GEM-MACH (Zhang) employ aerosol direct and indirect effect feedbacks between meteorology and chemistry, while GEM-MACH (Ops) does not. This requires the use of an explicit hydrometeor scheme in the former two model configurations and resulting in different meteorology between GEM-MACH (Base), GEM-MACH (Zhang) and GEM-MACH (Ops);

(4) GEM-MACH (Base) and GEM-MACH (Ops) make use of 15 land use categories aggregated from the 26 land use categories employed in GEM-MACH (Ops);

(5) Leaf Area Index values and seasonality for deposition for GEM-MACH (Base) and GEM-MACH (Zhang) are based on satellite retrieval data rather than the BEIS-based approach used in GEM-MACH (Ops) - the latter uses fixed LAI values for 232 land-use types and area-weights them to determine grid cell LAI;

(6) GEM-MACH (Base) and GEM-MACH (Zhang) make use of six additional physical parameterizations not present in GEM-MACH (Ops) (see Table 3).

Differences between GEM-MACH (Base) and GEM-MACH (Zhang) thus provide an estimate of the relative importance of the gas-phase deposition parameterization towards simulation results, while comparisons between GEM-MACH (Base or Zhang) and GEM-MACH (Ops) show the relative impact of

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

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.

Table 3. 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	1-D 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 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 Zhang *et al.,* (2002, 2003) 26 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 Pierces, 2002)
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 plumerise 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 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 (100m)

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Area source emissions treatment	Flux lower boundary condition on diffusion equation	As in GEM-MACH (Base)	Mass injection into two lowest model layers
Advection mass conservation	ILMC, 3 sweeps (Sorensen *et al.*, 2013)	As in GEM-MACH (Base)	ILMC, 2 sweeps, followed by Bermejo-Conde (2002) global mass correction

3.0 Results

3.1 Critical Load Exceedances

3.1.1 Europe, Acidification

Critical load exceedances for acidification for each of the four European (EU) models are shown in Figure 2 for 2010 and in Figure S1 (SI) for 2009. Figure 3 shows the reduced ensemble values for 2009 and 2010 (a,b) as well as common AQMEII4 domain total bar charts for all models and the ensembles (c).

The EU exceedances for acidity are similar between the two years (compare Figures 2 and S1, and values for each year in Figure 3). However, differences between models within a given year are larger (especially in an absolute sense; WRF-Chem (IASS) <0.4% in exceedance, WRF-Chem (UPM): ~6.5%). Low WRF-Chem (IASS) exceedance levels are in part due to unreported deposition data (see section 2.2.2); the reduced ensemble maps in Figure 3 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, north-western Germany, southern Norway, and along the border between Poland and Germany (Figure 3(a,b)). Individual models in Figure 2 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): south-west Switzerland, south-central Germany, and south-west Romania).

The percent area of EU acidification CLE over the region for which CL data was available, for the reduced ensemble, was 4.48% (range 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 3-member reduced ensemble.

3.1.2 Europe, Eutrophication

Critical load exceedances for eutrophication for each of the four EU models are shown in Figure 4 for 2010 and in Figure S2 (SI) for 2009. Figure 5 shows the reduced ensemble values for 2009 and 2010 (a,b) as well as common AQMEII4 domain summaries for all models and the ensembles (c).

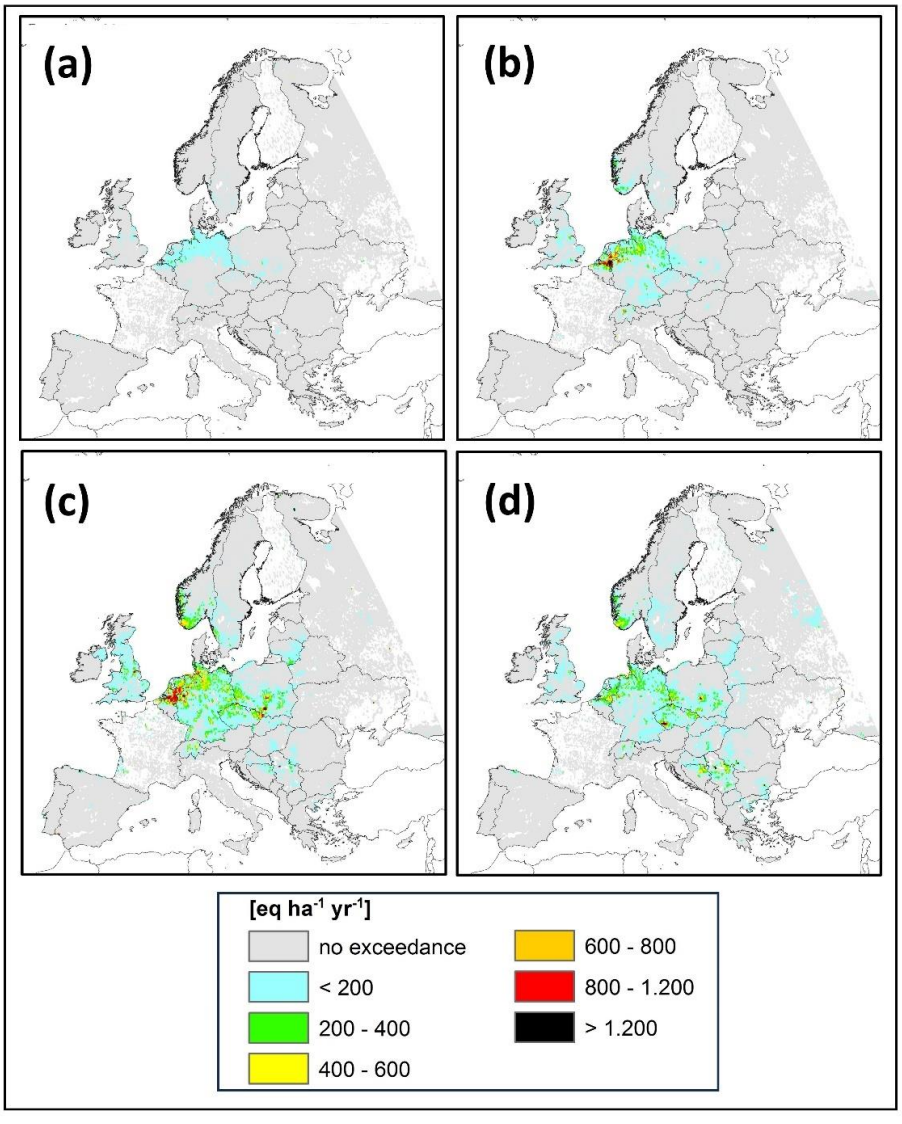
As for EU Acidity CLE's, the Eutrophication CLE's are very similar between the two model years (compare Figures 4 and S2, and the values for each year in Figure 5). The spatial distribution of the greatest levels of exceedance also varies more strongly between models. All members in the 3-member reduced ensemble identify the Po river valley as reaching the greatest level of exceedance, but LOTOS-EUROS (TNO) also shows high levels of exceedance in Benelux to northern Germany and in the

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Barcelona area, while WRF-Chem (UPM) shows high levels of exceedance > 800 eq ha^{-1} yr^{-1} in multiple hotspots throughout the region.

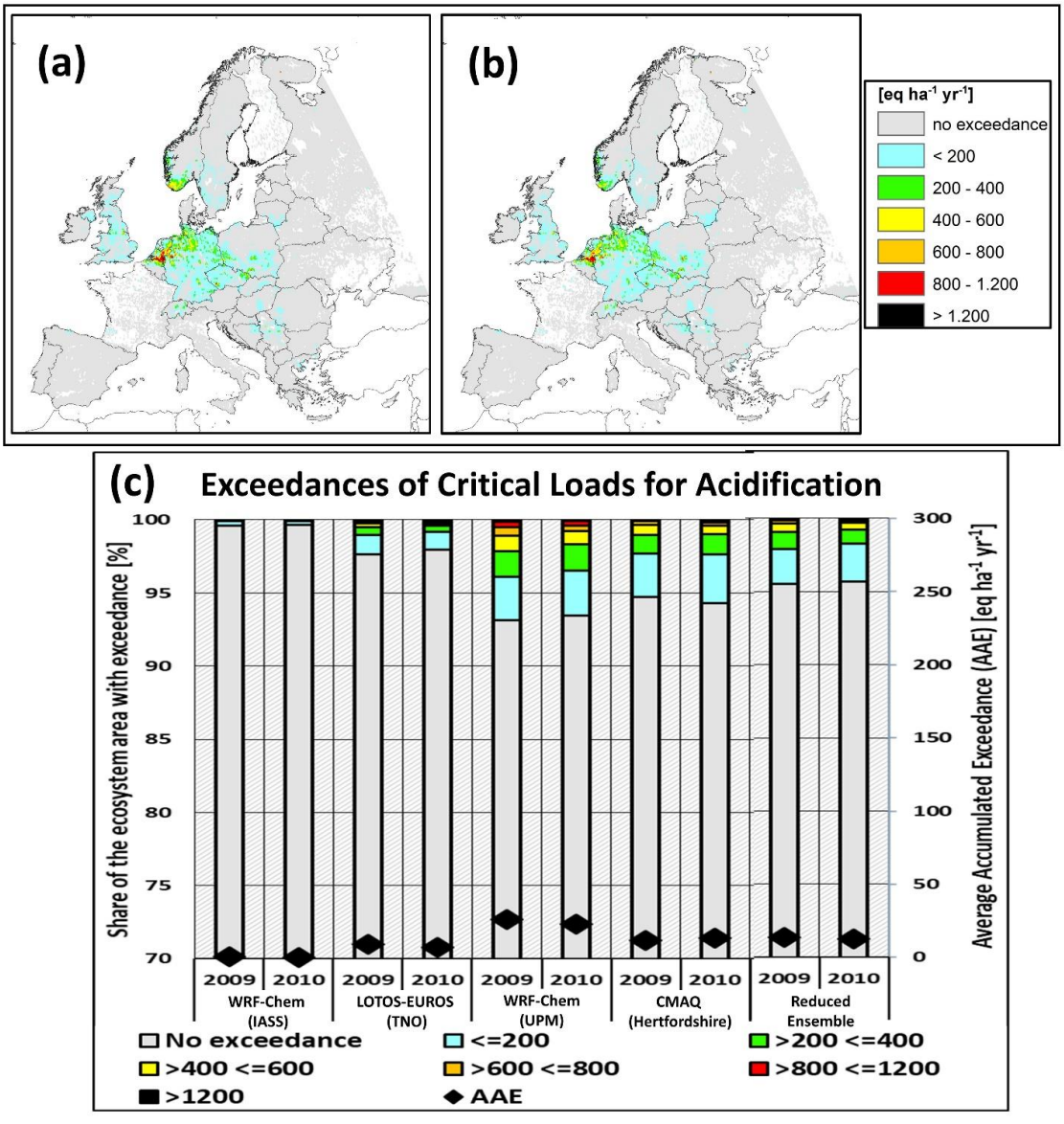
The percentage of the area in exceedance for eutrophication is much higher than that of acidification (reduced ensemble CLE 60.2% (47.3 to 73.3%) in 2009, and 62.2% 51.2 to 74.4%) in 2010). The average accumulated exceedance was 156.9 (89.4 to 265.5/) eq ha^{-1} yr^{-1} in 2009 and 161.4 (109.4 to 261.8) eq ha^{-1} yr^{-1} in 2010 (Figure 5, the range is from lowest and highest members in the 3-member reduced ensemble).

Figure 2. 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.



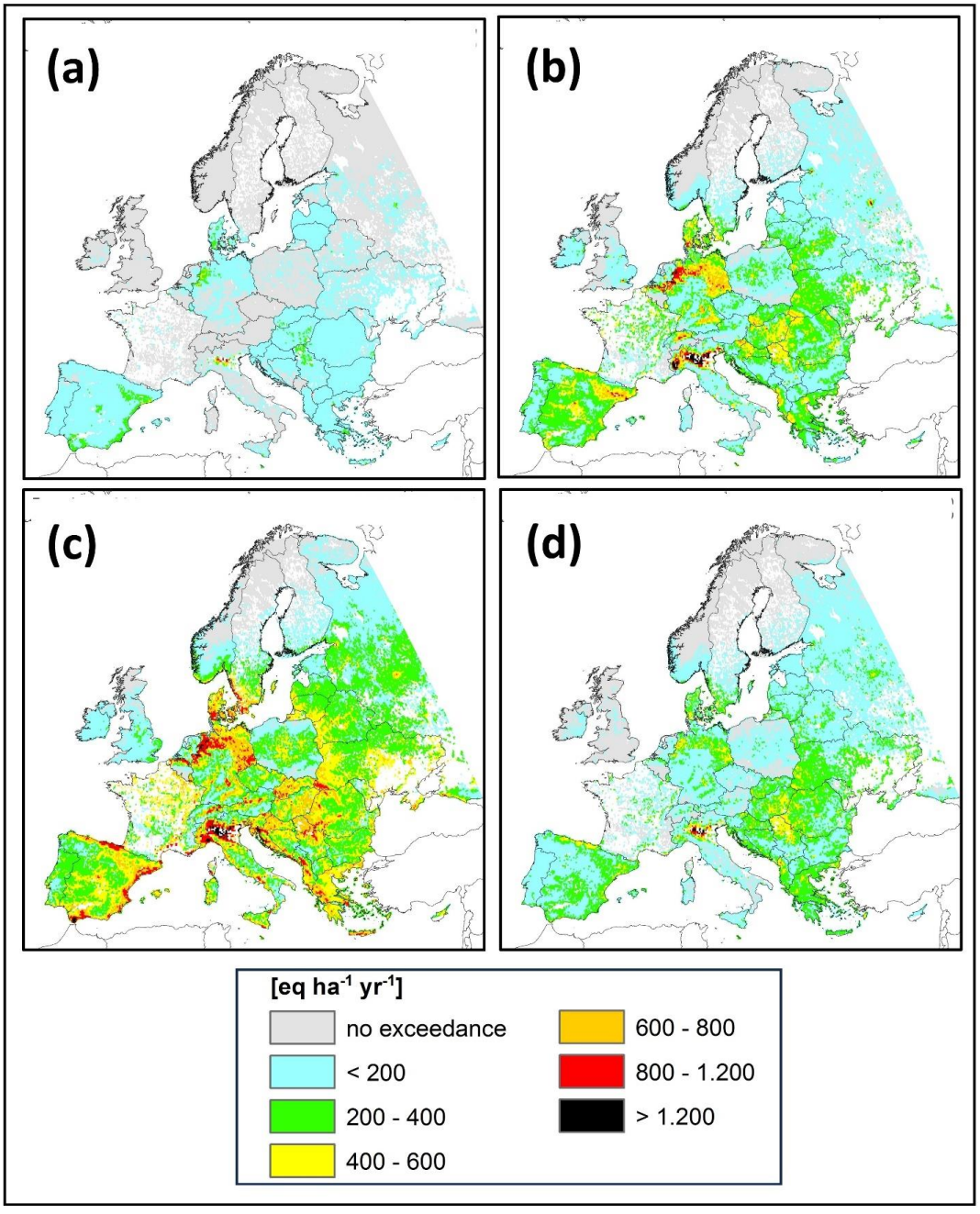
https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Figure 3. 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) 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 diamond symbols).



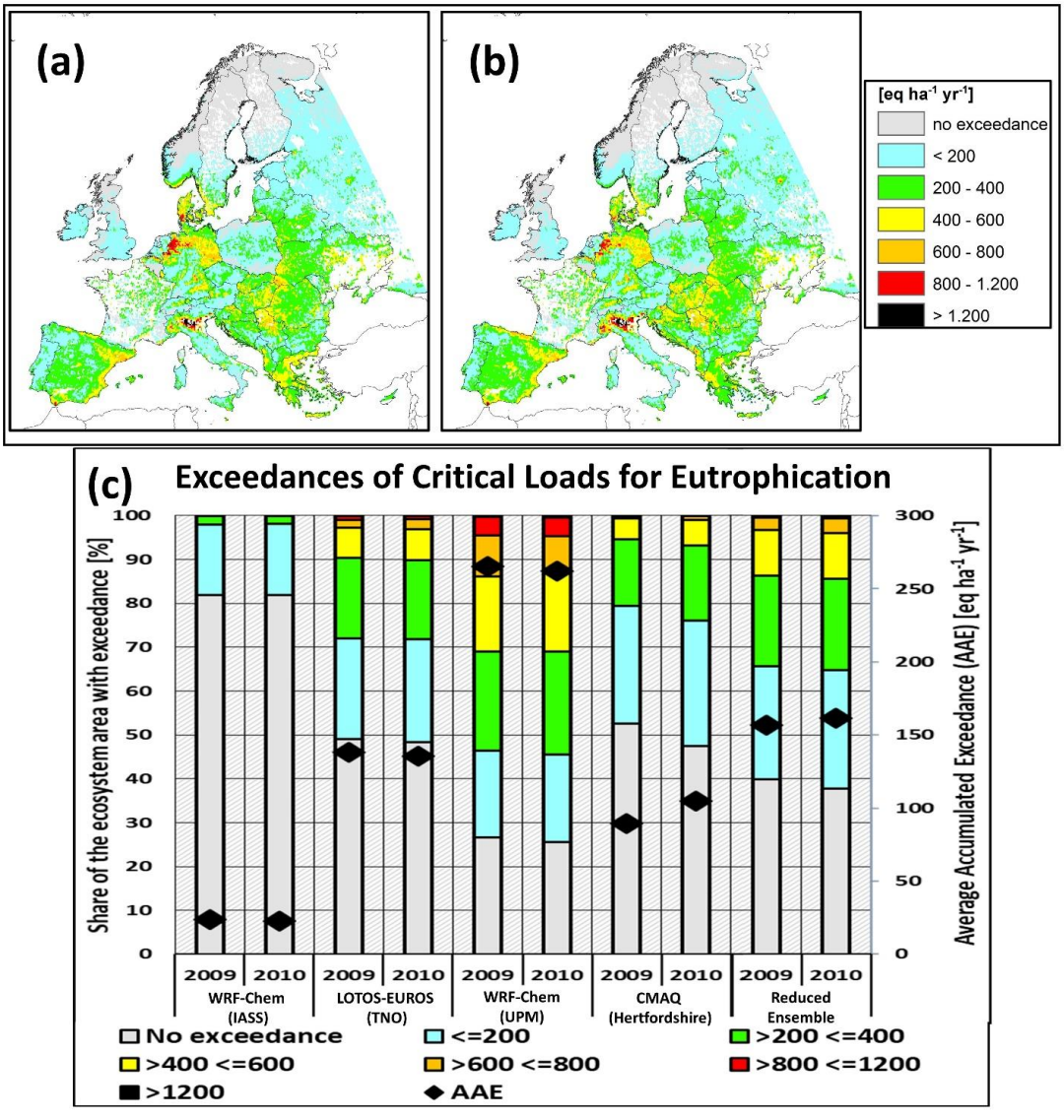
https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Figure 4. 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.



https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Figure 5. 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) 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 diamond symbols).



https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

3.1.3 North America, Forest Ecosystems Simple Mass Balance Critical Load

Critical load exceedances with respect to the North American (NA) forest soil acidity for the years 2016 and 2010 are shown in Figures 6 and S3, respectively, with the reduced ensemble maps for both years, and the domain summaries shown in Figure 7.

Unlike the EU domain comparison, the NA CLEs depicted in Figure 6 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 U.S. east coast, and Florida as being in exceedance, the magnitude of the exceedances varied greatly between the models, with the GEM-MACH models (Figure 6(d-f)) showing 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 (Figure 7) 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 (Figure 7 (c)) 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, Figure 7(c)) 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 byCMAQ-STAGE, CMAQ-M3Dry WRF-Chem (UPM), GEM-MACH-Ops, GEM-MACH-Base, and GEM-MACH-Zhang.

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 two years.

3.1.4 North America, Aquatic Ecosystems CL(A)

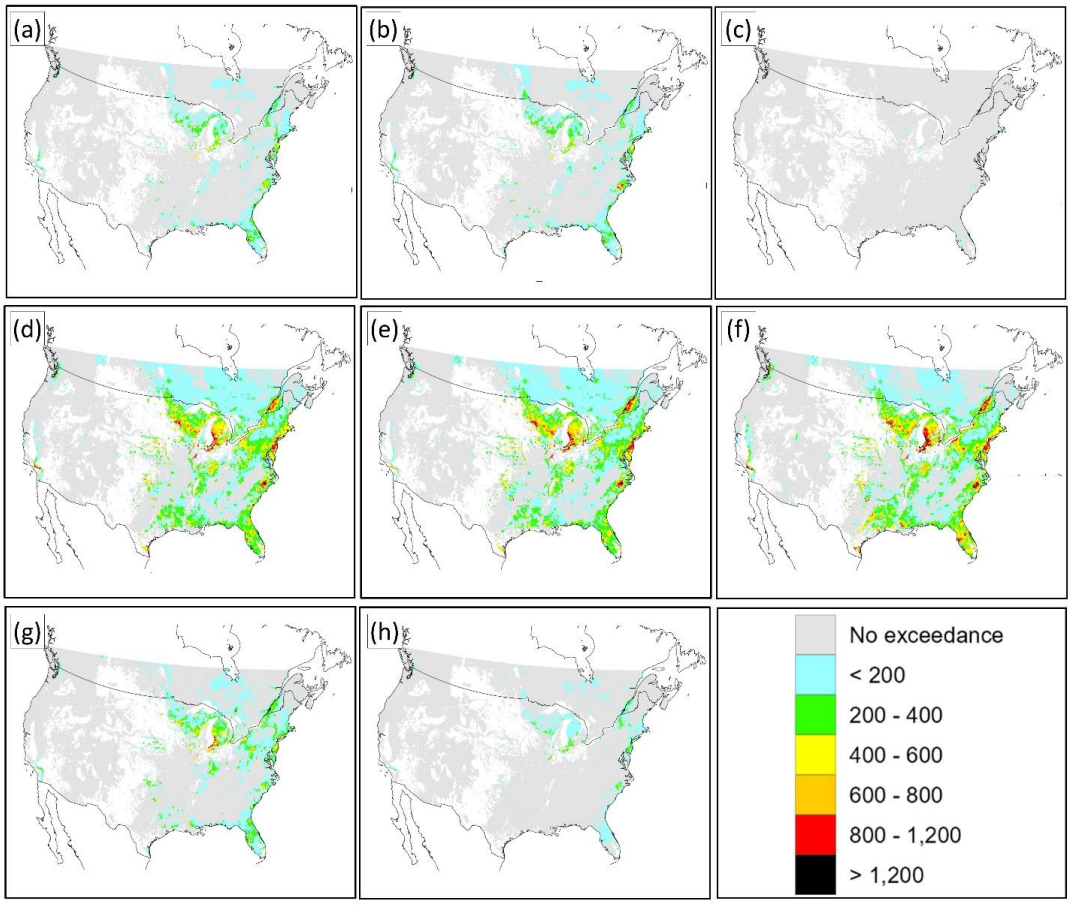
Exceedances with respect to the North American aquatic ecosystem CL dataset for the years 2016 and 2010 are shown in Figures 8 and S4, respectively, with the reduced ensemble maps for both years and domain summaries shown in Figure 9.

Comparison of Figures 6 and 8 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 9 (a,b) shows the expected decrease of the reduced ensemble's 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) (Figure 9 (c)).

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 exceedance of aquatic critical loads between the two years, again by almost a factor of two.

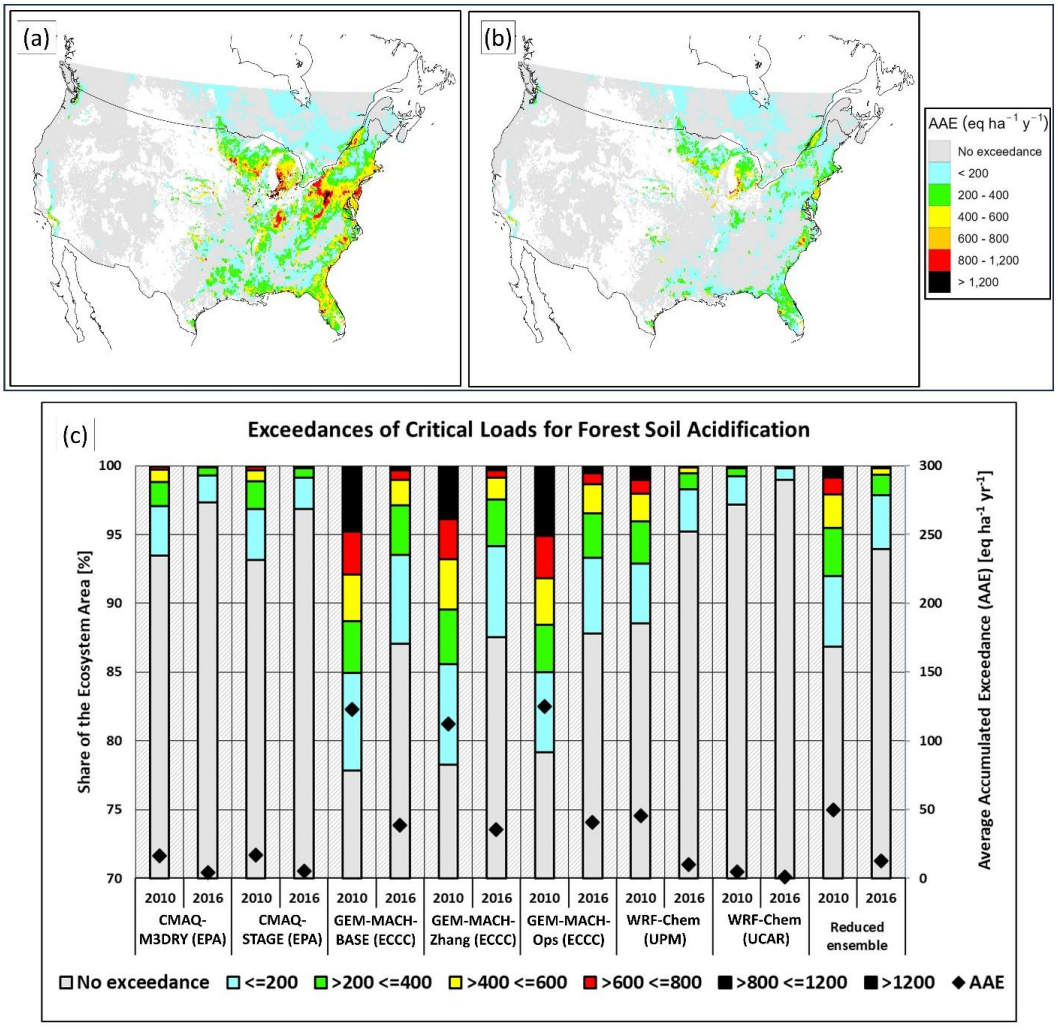
https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Figure 6. CLEs for Forest Soils, 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.



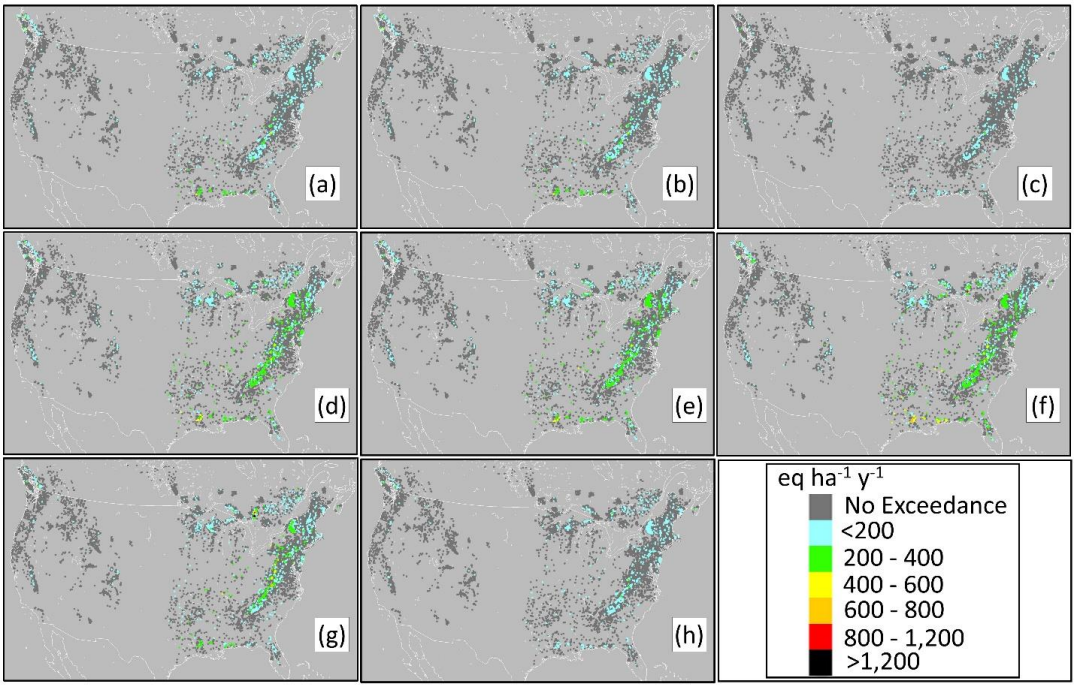
https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Figure 7. Summary CLEs for Forest Soils, 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) 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 diamond symbols).



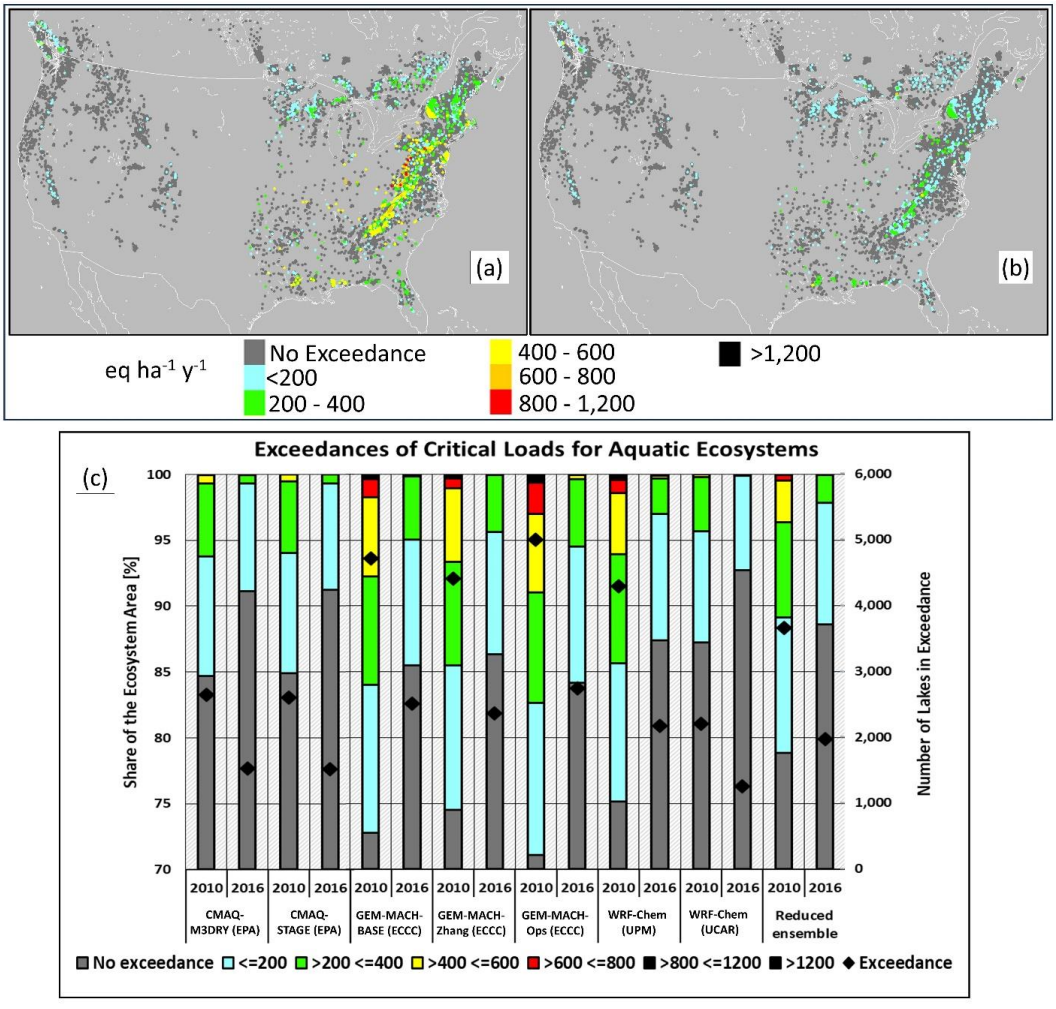
https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Figure 8. CLEs for Aquatic Ecosystems, NA AQMEII4 common domain, 2016, eq ha^{-1} yr^{-1}. Panels arranged by Model as in Figure 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.



https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Figure 9. 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) 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 diamond symbols).



https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

3.1.5 U.S. N Deposition to Lichen

Exceedances with respect to the USA CL of N for a 20% decline in sensitive epiphytic lichen species richness (221 eq-N ha^{-1} yr^{-1}) dataset for the years 2016 and 2010 are shown in Figures 10 and S5, respectively, with the reduced ensemble maps for both years and domain summaries shown in Figure 11.

The overall pattern of exceedances and their magnitude across models (Figure 10) is similar to that of the Forest Soil exceedances (Figure 6), with the largest magnitudes in the north-eastern 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 1,200 eq ha^{-1} yr^{-1}, and the exceedances predicted 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 reduced ensemble overall magnitude of exceedances decreased significantly between 2010 and 2016 (Figure 11(a,b), less black and red regions in the more recent year). The reduced ensemble total area in exceedance has decreased slightly (Figure 11(c), "reduced ensemble" columns). All models show a decreasing levels of exceedance between the two years, and slightly decreasing total area of exceedance. The magnitude of exceedances differs significantly between the models, with the highest magnitude exceedances predicted by the GEM-MACH group of models, followed by WRF-Chem (UPM).

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 U.S. N Deposition to Herbaceous Plants

Exceedances with respect to the USA CL of N for a decline in herbaceous species richness (436 to 1693 eq-N ha^{-1} yr^{-1}) dataset for the years 2016 and 2010 are shown in Figures 12 and S6, respectively, with the reduced ensemble maps for both years and domain summaries shown in Figure 13.

The spatial distribution of the regions of highest exceedance shares some common features with that of sensitive epiphytic lichen (compare Figure 12 with Figure 10), such as maximum exceedances in NE USA, North Carolina, and extending along a region north of Texas. However, both the magnitude and extent of exceedance is much more varied for herbaceous species richness than for lichen species richness, with the GEM-MACH suite of models (Figure 12 d-f and Figure 13c) predicting the highest exceedance levels and up to 18.4% of the area in exceedance in 2016, the CMAQ implementations varying between 0.6% and 0.8%, and WRF-Chem (UCAR) predicting 0.1%.

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 two 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 4. The values suggest acidification in EU will happen over a smaller region than eutrophication at 2009/2010 emissions levels, with a slight decrease in acidification and a slight increase in eutrophication between the two years. About 60% of EU ecosystems would be subject to eutrophication at some point in the future at 2010/2009 emissions levels. One

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

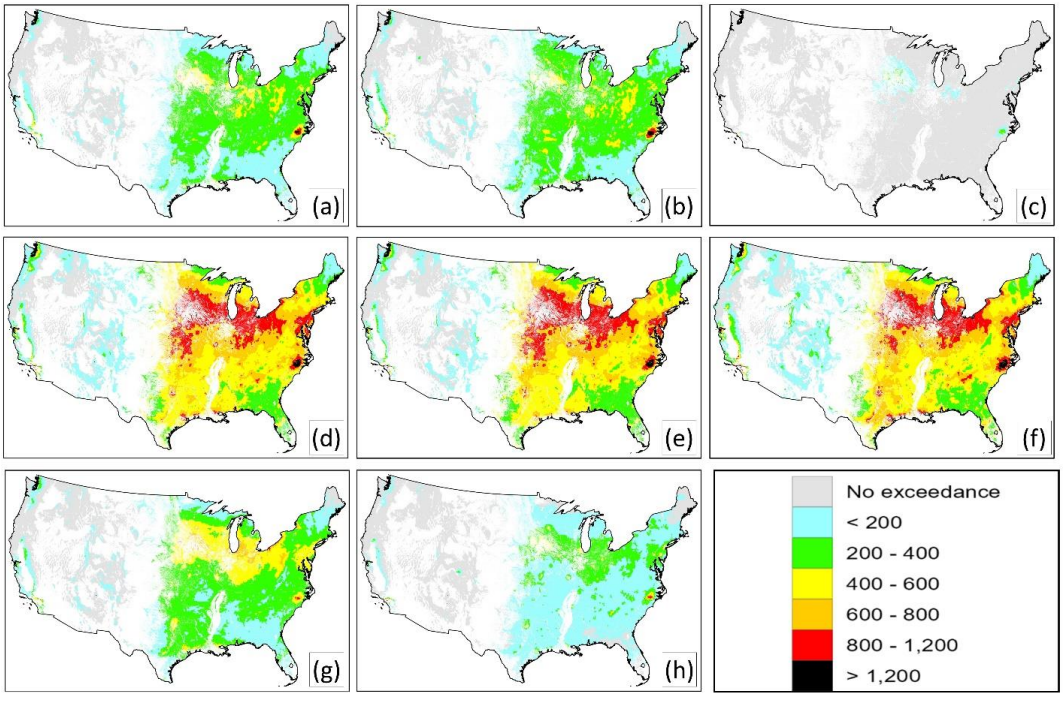
striking difference between the different model estimates of CLE is in the magnitude of exceedances (as opposed to the total area in exceedance). WRF-Chem (UPM) for example in Figures 2 and 4 predicts more severe levels of exceedance across Europe than the other models. The North America results suggest that reductions in SO_2 and NOx emissions between 2010 and 2016 resulted in a substantial reduction in the number of forest soil and aquatic ecosystem acidification exceedances (by nearly a factor of two). The impacts of nitrogen deposition on herbaceous species also improved (by nearly a factor of three), while impacts of nitrogen deposition on sensitive lichen had 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.

Table 4. Summary of reduced ensemble percent exceedance mean values and their range in EU and NA domains.

EU Ecosystem	Year 2009 Percent Exceedance (lower to upper bound)	Year 2010 Percent Exceedance (lower to upper bound)
Acidification	4.48 (2.37 to 6.85)	4.32 (2.06 to 6.52)
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)
Forest Soils Acidification	13.2 (2.8 to 22.2)	6.1 (1.0 to 12.9)
Lake Ecosystems	21.2 (12.8 to 28.9)	11.4 (7.3 to 15.8)
USA N Deposition Lichen	81.5 (69.3 to 95.0)	75.8 (63.7 to 90.7)
USA N Deposition Herbaceous	13.9 (0.4 to 39.5)	3.9 (0.1 to 18.4)

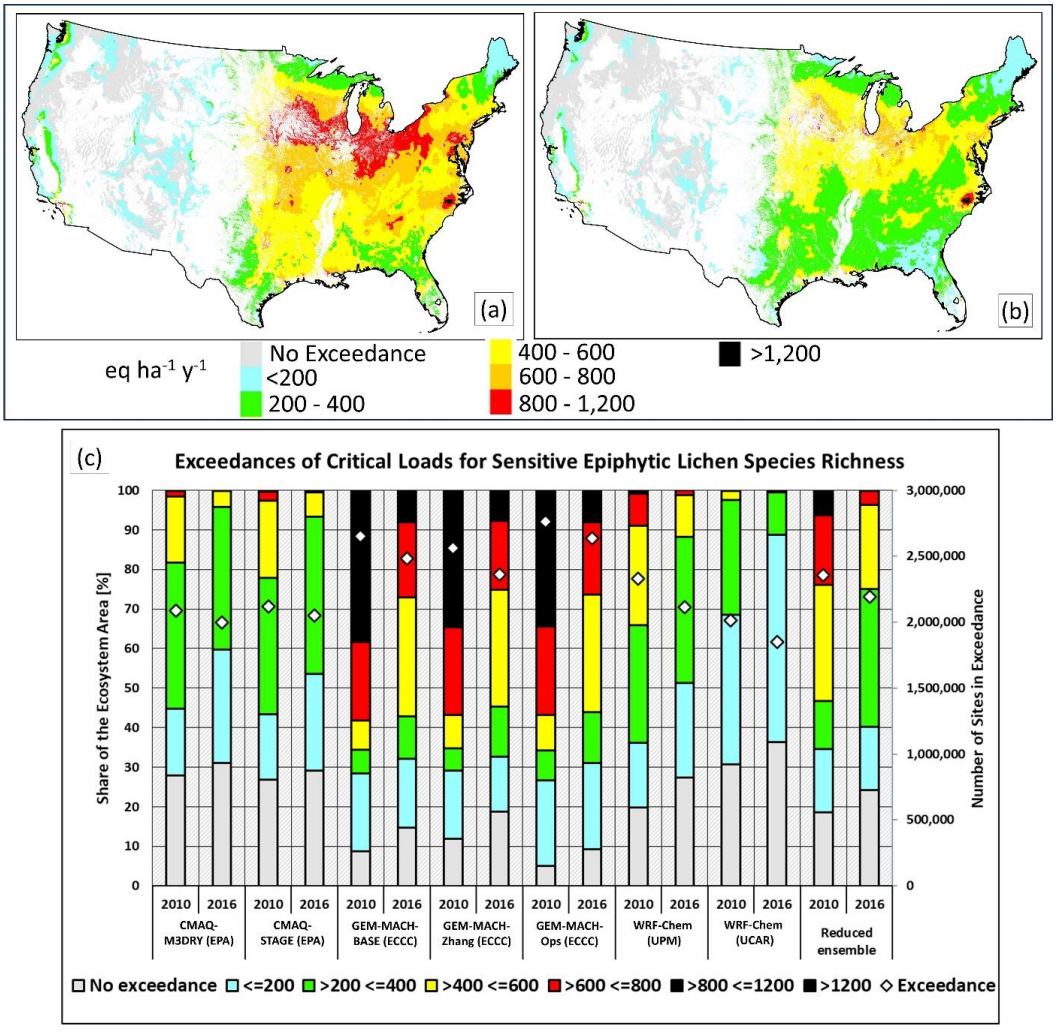
https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Figure 10. CLEs for Sensitive Epiphytic Lichen Species, NAAQMEII4 common domain, 2016, eq ha^{-1} yr^{-1}. Panels arranged by model as in Figure 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.



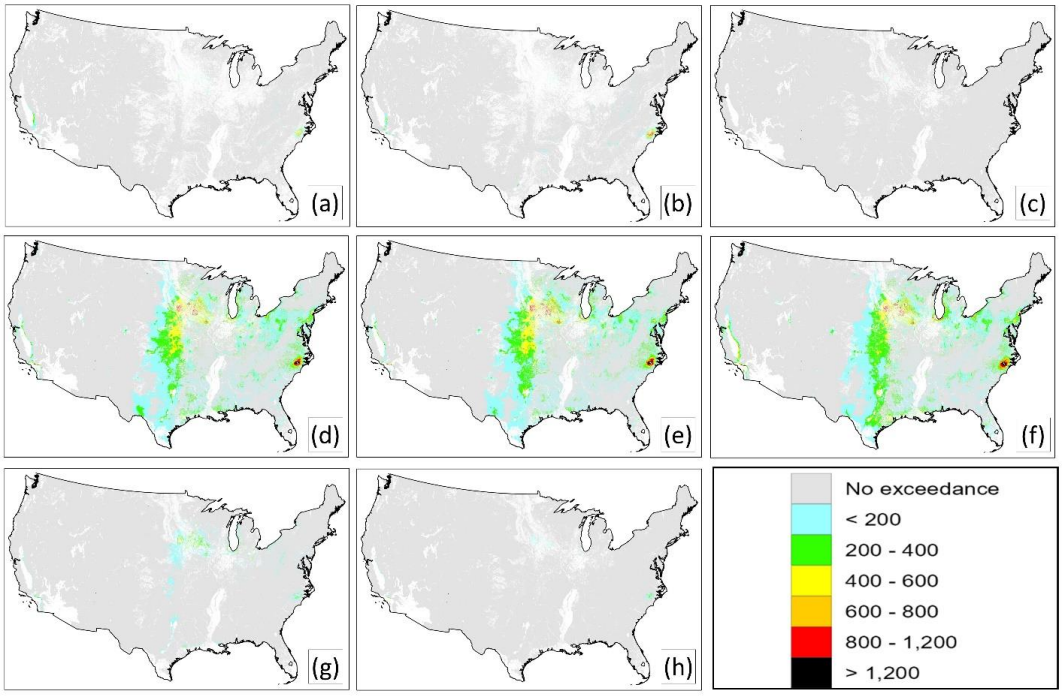
https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Figure 11. 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) 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 diamond symbols).



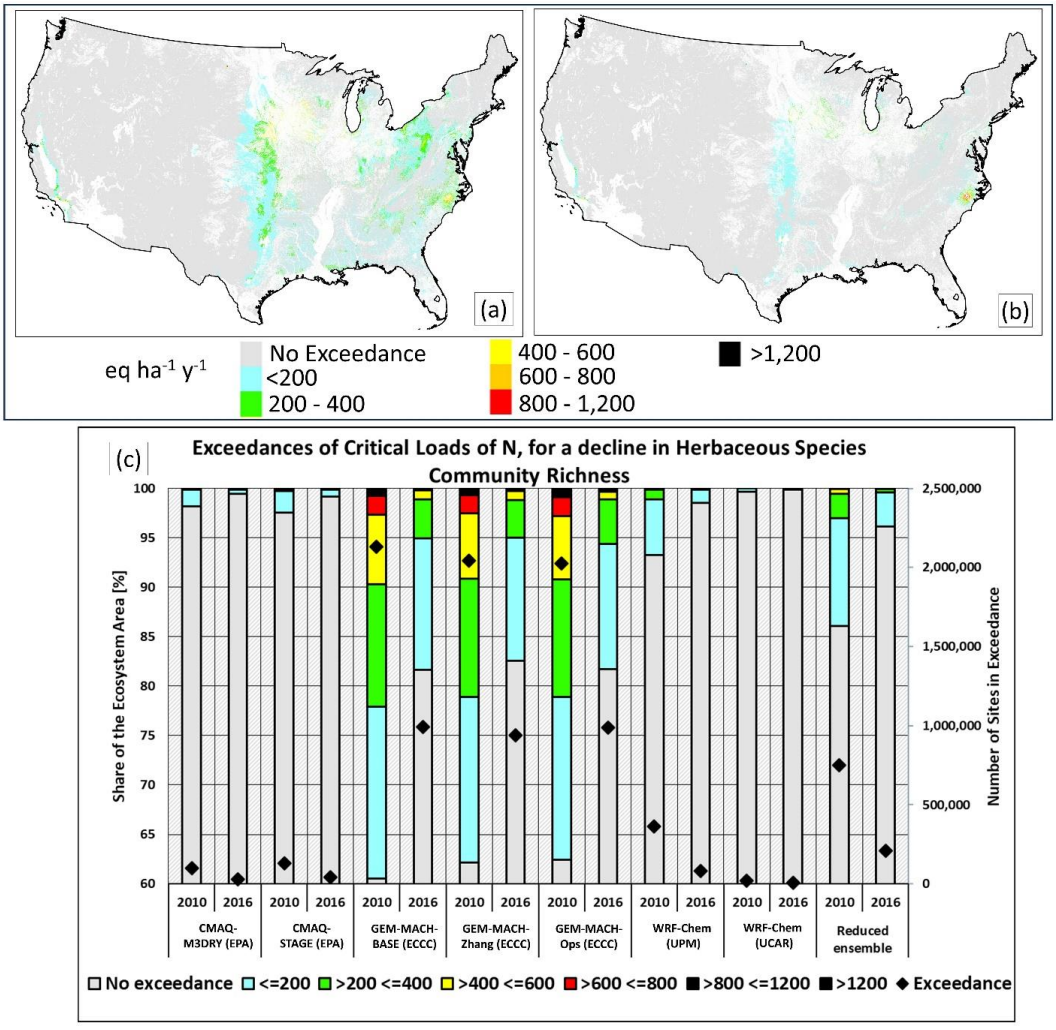
https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Figure 12. CLEs for a decline in Herbaceous Species Community Richness, NA common domain, 2016, eq $ha^{-1}yr^{-1}$. Panels arranged by model as in Figure 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.



https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Figure 13. Summary CLEs for a decline in Herbaceous Species Community Richness, AQMEII4 NA 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) 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 diamond symbols).



https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

3.2 Analysis of Model Deposition Predictions

3.2.1 Causes of S Deposition Variability in North America Domain Simulations

The AQMEII4 common grid average, and percent contribution of each depositing species towards total S deposition in 2016, are given in Table 5. The averages and standard deviation for the reduced ensemble show that wet deposition of the sum of the sulphate and bisulphite ions ($SO_4^{(2-)}$ and $HSO_3^{(-)}$) contributes more to total S deposition than particulate sulphate dry deposition, which is in turn contributes more than SO_2 (g) dry deposition. However, the model-to-model variability is also large, particularly for the contribution of particulate sulphate, which varies by nearly two orders of magnitude between GEM-MACH (Base, Zhang 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 (± ranges in Table 5 are the standard deviation of the component). The greatest cause of model *variability* in absolute total deposition is associated with the contribution of particulate sulphate dry deposition, followed by sulphur wet deposition and then gaseous SO_2 dry deposition.

Table 5. Average S deposition contributions in common AQMEII4 NA 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 percentage.

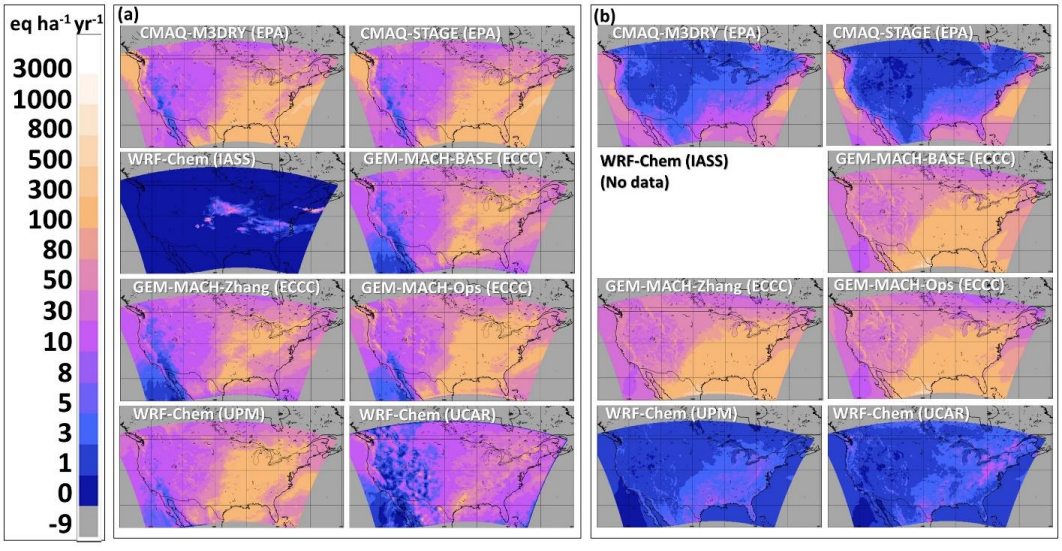
	Average Deposition (eq ha^{-1} yr^{-1})				Percent of total S deposition		
Model Number	$SO_4^{(2-)}$ + $HSO_3^{(-)}$ Wet Deposition	Particle Sulphate Dry Deposition	SO_2(g) Dry Deposition	Total S Deposition	$SO_4^{(2-)}$ + $HSO_3^{(-)}$ Wet Deposition	Particle Sulphate Dry Deposition	SO_2(g) Dry Deposition
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	nd	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

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

The spatial distributions of the two largest components of the total S deposition variability (wet S and dry particle S) are shown in Figure 14. The WRF-Chem (IASS) values did not represent the expected sources of S deposition over the continent and some deposition fields such as the total particulate sulphate dry deposition were not submitted. The wet S deposition maps are qualitatively similar between the other models (note that the colour scale is logarithmic), with WRF-Chem (UCAR) having the lowest values (Figure 14(a)). As shown in Table 5, the greatest degree of variability between the different modelling platforms is in the particle deposition fluxes (Figure 14(b)). This variability extends over orders of magnitude. WRF-Chem (UPM) and WRF-Chem (UCAR) predict the lowest deposition fluxes of dry particulate sulphate 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 sulphate 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 sulphate, and wet sulphur deposition for the year 2016. The monitoring network databases employed included the U.S. Environmental Protection Agency's Air Quality System (AQS; https://www.epa.gov/aqs , last access: 7 July 2024), the National Atmospheric Deposition Program's National Trend 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).

Figure 14. 2016 total annual deposition flux (eq ha^{-1} yr^{-1}) of (a) wet S, and (b) dry particulate sulphate.



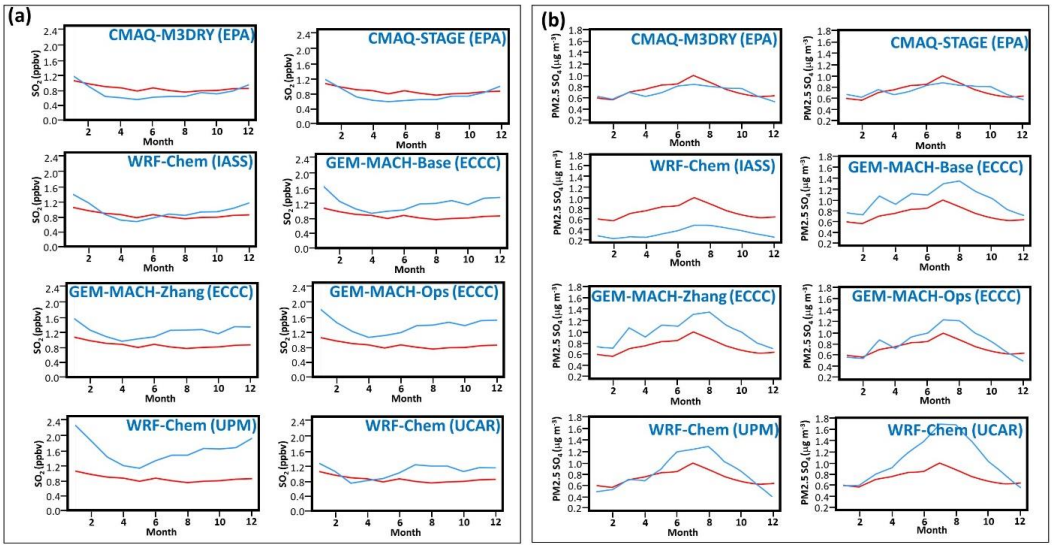
https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

The NA models' monthly average values of hourly near-surface SO_2 (g) concentrations and daily particulate sulphate concentrations are compared to observations in Figure 15. The monthly averages of daily (CAPMoN) and weekly (NADP) wet S deposition are shown in Figure 16. Model-observation evaluation statistics are compared in Table 6. Station locations for the observations are shown in SI Figures S7, S8, and S9.

Table 6 shows that CMAQ-M3Dry and CMAQ-STAGE had the best values for most metrics, for the concentrations of SO_2 and PM2.5 sulphate, and daily wet sulphur deposition. The 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 Figure 15(a), where CMAQ-M3Dry and CMAQ-STAGE negative biases can be seen to occur in the warmer months, WRF-Chem (IASS) negative biases in the spring. Despite these differences, the net contribution of SO_2 dry deposition flux towards total sulphur deposition on an *annual* basis is relatively similar across the models (Table 5), 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.

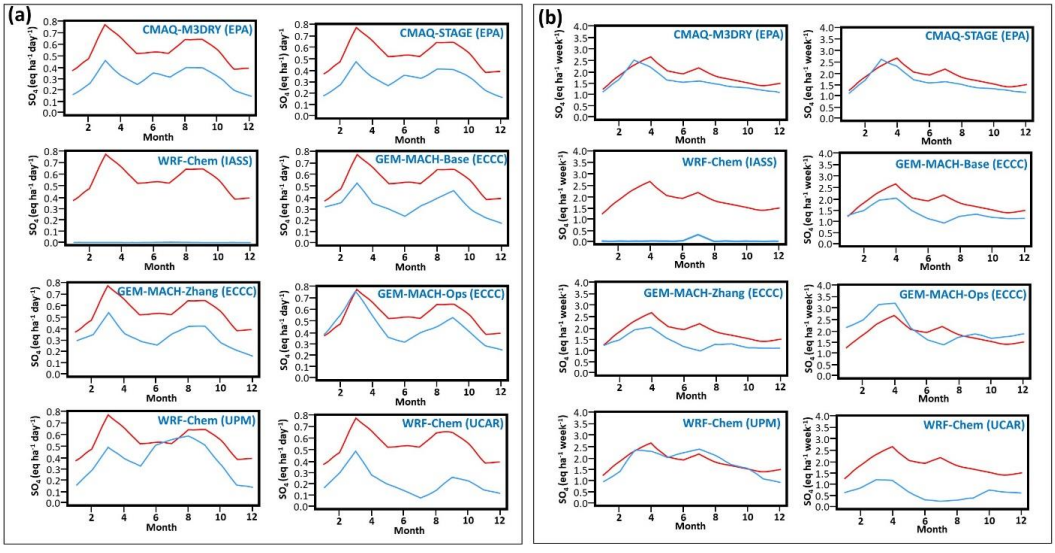
Particle sulphate (Figure 15(b), and Table 6) values were also closest to monthly observed values for CMAQ-M3Dry and CMAQ-STAGE, while being biased negative for WRF-Chem (IASS) and biased positive for the remaining models. The evaluation of total S wet deposition (Figure 16(a), Table 6) showed that all models with the exception of GEM-MACH (Ops) had negative biases relative to the Canadian daily wet S deposition observations. Weekly wet S deposition biases are also negative for most models (Table 6, Figure 16(b)), with only GEM-MACH-Ops having a positive bias in the ensemble.

Figure 15. Comparison of model (blue line) and observed (red line) monthly average surface concentrations of (a) hourly SO_2 (ppbv)) and (b) daily PM2.5 sulphate (ug m^{-3}), for the year 2016 (AQS, NAPS data).



https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Figure 16. Comparison of model (blue line) and observed (red line) monthly average values of wet sulphur deposition for (a) daily CAPMoN data (eq ha^{-1} day^{-1}), (b) weekly NADP data (eq ha^{-1} $week^{-1}$), for the year 2016.



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 sulphate formation (through nucleation/condensation of sulphuric 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 sulphate and in sulphate ion wet deposition. However, Table 16 shows relatively little bias for PM2.5 sulphate 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 sulphate would argue against an insufficient conversion of SO_2 to particulate sulphate in the latter group of models. Rather, the general tendency of negative biases in wet sulphur 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 sulphate bias was lower, and the wet sulphate 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 weather/air quality feedbacks to be simulated, but tends to underestimate the cloud amounts when used at lower resolution such as the 10km 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.

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Table 6. Model Performance Metrics for SO_2, PM2.5 SO_4, Wet deposition of S, AQMEII4 North American domain, 2016. Bold-face letters show the highest scoring model.

Hourly SO_2 (units ppbv where applicable)								
Performance Measure	CMAQ-M3Dry	CMAQ-STAGE	WRF-Chem (IASS)	GEM-MACH (Base)	GEM-MACH (Zhang)	GEM-MACH (Ops)	WRF-Chem (UPM)	WRF-Chem (UCAR)
FAC2	0.27	**0.28**	0.26	**0.28**	**0.28**	**0.28**	0.26	0.29
MB	-0.18	-0.17	**-0.03**	0.11	0.14	0.24	0.61	0.17
MGE	**0.91**	**0.91**	1.02	1.08	1.09	1.17	1.43	1.09
NMGE	**1.02**	**1.02**	1.15	1.21	1.22	1.32	1.60	1.22
RMSE	**3.14**	**3.14**	3.29	3.33	3.34	3.51	3.75	3.21
R	**0.15**	**0.15**	0.12	0.14	0.14	0.13	0.13	0.13
COE	**0.04**	0.03	-0.08	-0.14	-0.16	-0.24	-0.51	-0.15
IOA	**0.52**	**0.52**	0.46	0.43	0.42	0.38	0.25	0.43
PM2.5 SO_4 (units μg m^{-3}, where applicable)								
FAC2	**0.77**	0.76	0.33	0.65	0.66	0.63	0.67	0.59
MB	-0.04	**0.00**	-0.41	0.28	0.26	0.10	0.10	0.32
MGE	**0.31**	0.32	0.45	0.50	0.50	0.46	0.43	0.55
NMGE	**0.43**	**0.43**	0.60	0.68	0.67	0.62	0.58	0.75
RMSE	**0.89**	**0.89**	1.00	1.10	1.09	1.06	1.00	1.12
R	0.45	**0.46**	0.40	0.40	0.40	0.38	0.39	0.40
COE	**0.37**	0.36	0.10	-0.02	0.00	0.07	0.13	-0.12
IOA	**0.68**	**0.68**	0.55	0.49	0.50	0.54	0.57	0.44
Daily Total Wet S Deposition (units eq ha^{-1} d^{-1}, where applicable)								
FAC2	0.35	0.36	0.00	0.40	0.40	**0.41**	0.39	0.19
MB	-0.19	-0.17	-0.57	-0.07	-0.08	0.09	**-0.06**	-0.31
MGE	**0.37**	**0.37**	0.57	0.42	0.42	0.48	0.45	0.46
NMGE	**0.65**	**0.65**	1.00	0.74	0.74	0.85	0.79	0.81
RMSE	**0.71**	**0.71**	1.02	0.81	0.81	0.88	0.90	0.89
R	**0.61**	**0.61**	0.06	0.52	0.52	0.54	0.47	0.44
COE	**0.31**	**0.31**	-0.06	0.21	0.22	0.10	0.16	0.14
IOA	**0.65**	**0.65**	0.47	0.60	0.61	0.55	0.58	0.57
Weekly Total Wet S Deposition (units eq ha^{-1} $week^{-1}$, where applicable)								
FAC2	0.46	**0.47**	0.00	0.41	0.41	0.41	0.45	0.21
MB	-0.21	-0.17	-1.78	-0.41	-0.42	0.30	**-0.03**	-1.18
MGE	**1.12**	**1.12**	1.81	1.18	1.18	1.40	1.28	1.38
NMGE	**0.62**	**0.62**	1.00	0.65	0.66	0.78	0.71	0.76
RMSE	**2.30**	**2.30**	3.26	2.30	2.30	2.54	2.48	2.64
R	**0.63**	**0.63**	0.03	0.55	0.55	0.57	0.53	0.46
COE	**0.34**	**0.34**	-0.07	0.30	0.30	0.17	0.24	0.18
IOA	**0.67**	**0.67**	0.46	0.65	0.65	0.58	0.62	0.59

GEM-MACH (Base), GEM-MACH (Zhang) and WRF-Chem (UCAR) have the most positive biases for particulate sulphate. 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) makes use of extensive measurement data, and compared to earlier parameterizations such as Zhang *et al.* (2001), results in decreased dry deposition velocities for sub-micrometer particles and increased dry deposition velocities for particles larger than 0.2 to 0.8 um, depending on land use type. The increased PM2.5 SO_4 values in GEM-MACH (Base) and GEM-MACH (Zhang) in Figure 15(b) may thus reflect decreases in the deposition removal flux in the sub-micrometer portion of the bins in these 12-bin sectional model

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

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 subsequent to the runs carried out here, resulted in an increase in the positive PM2.5 bias from +4.5 to +6.7 ug m^{-3} and a shift towards less negative biases in PM10, from -19.7 to -1.77 ug m^{-3}, similar to the biases in particulate sulphate and ammonium observed in Figure 15(b) 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 by the Wang (2014) parameterization offset the increase in PM2.5 SO4 biases associated with the new particle dry deposition scheme, illustrating the extent 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 (2014) semi-empirical scavenging scheme, with a significant increase in the uptake of particulate sulphate depending on precipitation rate, and improvements in the wet sulphate 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 subsequent to the simulations for AQMEII4.

With regards to wet sulphur deposition, Figure 16(a) and Table 6 shows a tendency of most models towards negative biases for total *daily* wet S deposition. However, this negative bias is much less pronounced or even positive in comparison to the *weekly* wet S 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 MGE, NMGE, and 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 Figure SI7(a)). The daily values are from the Canadian CAPMoN network (stations in the common AQMEII4 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 emissions 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 Figure SI10, for both daily (CAPMoN) and weekly(NADP) monthly averages. While the monthly averages of daily precipitation (Figure SI10(b)) suggest a tendency towards negative biases in the summer months for some models, the time series of the precipitation biases does not follow that of the wet sulphate deposition biases (for example, the difference relative to wet sulphate observations in Figure 06(a) remains relatively constant for CMAQ-M3Dry and CMAQ-STAGE, while the predicted precipitation difference relative to observations for the same models in Figure SI10(a) 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 sulphur flux biases found in this work T.

We also note the potential for the lower magnitude biases in the daily wet wet SO_4^{2-} evaluation, compared to the weekly evaluation, to be the result of the region represented by the two monitoring

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

networks. Figure SI7 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

Sulphur deposition results from a complex balance between SO_2 oxidation, particulate sulphate formation, scavenging and release of particles within clouds, in addition to the processes governing deposition of each of the components. The largest contributing pathways to North American sulphur deposition, in descending order of importance, were wet deposition ($SO_4^{2-} + HSO_3^-$), particulate sulphate dry deposition, and dry SO_2(g) deposition in the reduced ensemble of mnodel runs. The largest contributors to model-to-model variability in sulphur deposition, in descending order of importance, were particulate sulphate 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 PM2.5 sulphate positive biases relative to observations also had stronger negative biases for wet sulphate deposition, indicating that the magnitude of particle scavenging into 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 sulphate deposition fluxes relative to the other models, and substantial positive biases in PM2.5 sulphate and negative biases in wet sulphate deposition, relative to observations, likely related to insufficient wet scavenging of sulphate particles into hydrometeors (Ryu and Min, 2022)

3.2.2 Causes of N Deposition Variability in North America 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 7. The columns in the Table are arranged in descending order from left to right of contribution to the reduced ensemble total nitrogen deposition for each contributing chemical ("Red. Ens. Avg" row). 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 ("Red. Ens. Std. Dev" row). From the standard deviation row, it can be seen that the variation (standard deviation) between models for the contributions towards 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 (DHNO$_3$), dry particle nitrate (DNI), dry NO_2 (DNO2), 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 respectively 17.4x and 1.76x higher than the other models

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.

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Table 7. Contributions of N species towards total deposition (eq ha^{-1} yr^{-1}) and percent of total N deposited, over the common AQMEII4 NA 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. DNO_2: dry deposition of NO_2(g). DPAN: dry deposition of peroxyactylnitrate 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).

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. Std Dev
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	nd	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	nd	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	nd	nd	nd	nd	nd	1.2	0.1
DHNO4	0.4	0.1	0	nd	nd	nd	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. Std Dev
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	nd	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	nd	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	nd	nd	nd	nd	nd	0.4	0
DHNO4	0.2	0	0	nd	nd	nd	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

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Dry deposition of particle ammonium

The largest source of variability between North America models' total N predictions resides in the dry particle ammonium deposition fluxes, with Table 7 showing that the standard deviation of this deposition flux across models was essentially as large as the reduced ensemble average. Particle dry ammonium 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 17 compares the monthly average PM2.5 ammonium concentrations with observations (station locations appear in Figure S7(b)), and Table 8 provides detailed statistics. From the latter, CMAQ-M3Dry and CMAQ-STAGE have the best overall performance for particulate ammonium, 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 PM2.5 ammonium concentrations, while the CMAQ implementations have the negative biases, and the remaining models have smaller magnitude positive biases. Figure 17 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 PM2.5 ammonium. This poor performance is likely due to two factors, which can be deduced from comparing the process representations implemented in the models (section 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 all this process representation in the models of the ensemble is derived from the ISORROPIA module (Nenes *et al.,* 1998; Fountoukis *et al.*, 2007), the GEM-MACH implementations in AQMEII-4 employ a partial speciation of SO_4^{2-}, NH_4^+ and NO_3^- (Makar *et al.,* 2003), and do not include 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 sulphate, 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 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 PM2.5 ammonium in Figure 17 for GEM-MACH relative to the other models likely represents the impact of simplified inorganic aerosol chemistry.

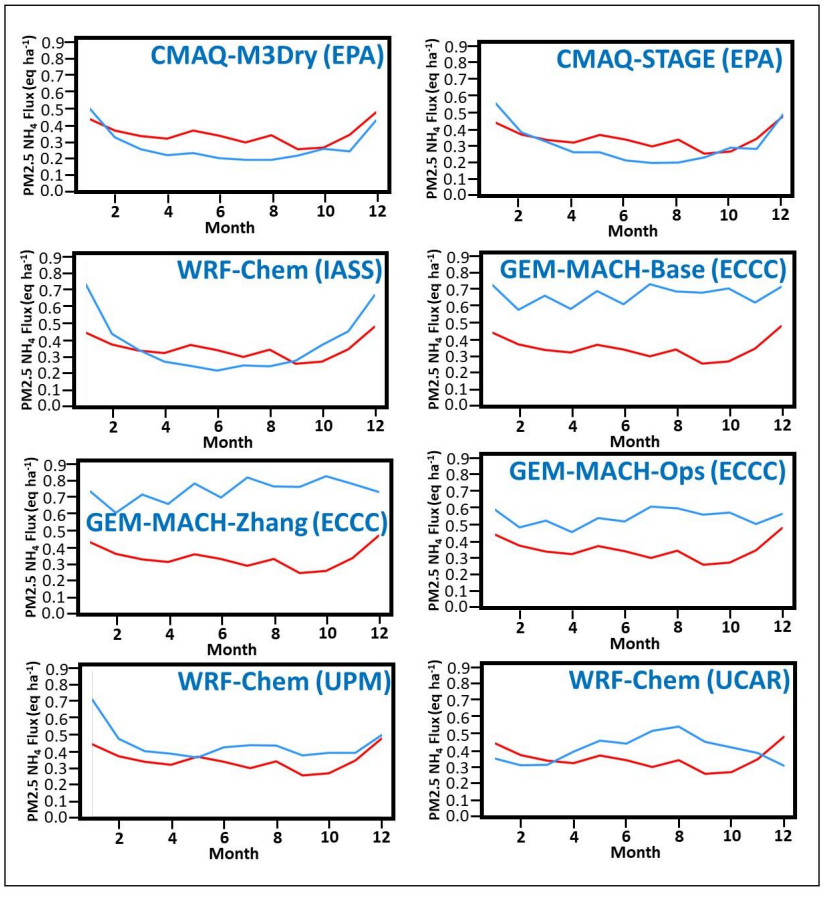
The second factor influencing the GEM-MACH models positive particulate ammonium biases may be reflected in the biases for GEM-MACH (Base) and GEM-MACH (Zhang), which are 50% to a factor of two, respectively, higher than that of GEM-MACH (Ops): that is, an additional source of bias resides in the former two model implementations that is not present in the latter implementation. The likely source of this additional bias is their use of Emerson *et al.* (2020) particle deposition velocities in these

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

implementations, in the absence of enhanced wet scavenging of aerosols, as discussed above for PM2.5 sulphate, 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 PM2.5 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.

Figure 17. PM2.5 Ammonium compared to observations, North American Model Ensemble, 2016. Red line: monthly observed average. Blue line: monthly model average.



https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

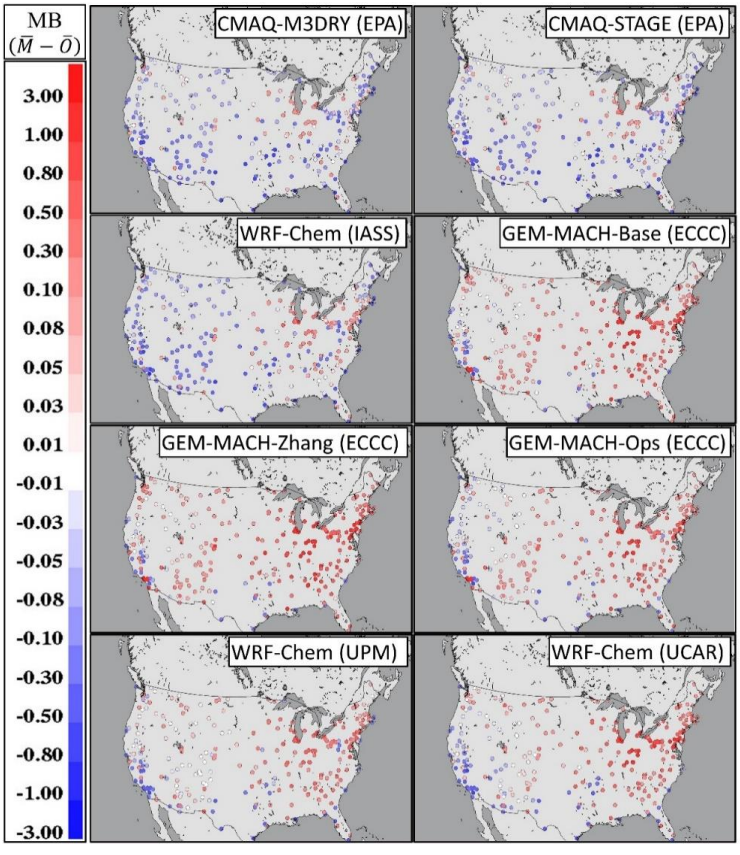
Table 8. Model Performance Metrics for PM2.5 ammonium, wet deposition of ammonium ion, wet deposition of nitrate ion, AQMEII4 North American domain, 2016. Bold-face letters show the highest scoring model.

PM2.5 NH_4 (units μg m^{-3}, where applicable)								
Performance Measure	CMAQ-M3Dry	CMAQ-STAGE	WRF-Chem (IASS)	GEM-MACH (Base)	GEM-MACH (Zhang)	GEM-MACH (Ops)	WRF-Chem (UPM)	WRF-Chem (UCAR)
FAC2	0.48	**0.49**	0.31	0.45	0.42	0.46	0.51	0.46
MB	-0.07	-0.04	**0.03**	0.32	0.41	0.20	0.10	0.06
MGE	**0.23**	0.24	0.33	0.45	0.52	0.38	0.31	0.31
NMGE	**0.68**	0.70	0.96	1.31	1.53	1.10	0.91	0.91
RMSE	**0.59**	0.60	0.75	0.81	0.93	0.75	0.69	0.69
R	**0.37**	**0.37**	0.30	0.33	0.32	0.32	0.30	0.23
COE	**0.19**	0.17	-0.13	-0.55	-0.80	-0.30	-0.08	-0.08
IOA	**0.60**	0.58	0.43	0.23	0.10	0.35	0.46	0.46
Daily Total Wet NH_4 Deposition (units eq ha^{-1} d^{-1}, where applicable)								
FAC2	0.26	0.29	0.00	0.39	0.38	**0.43**	0.28	0.14
MB	-0.49	-0.44	-0.94	-0.01	**0.00**	-0.10	-0.39	-0.59
MGE	0.67	**0.65**	0.94	0.76	0.78	0.68	0.71	0.80
NMGE	0.72	**0.69**	1.00	0.81	0.83	0.73	0.76	0.86
RMSE	1.46	**1.43**	1.90	1.66	1.71	1.45	1.54	1.73
R	0.55	0.57	0.26	0.52	0.51	**0.59**	0.49	0.37
COE	0.32	**0.34**	0.05	0.23	0.21	0.31	0.28	0.19
IOA	0.66	**0.67**	0.53	0.61	0.61	0.65	0.64	0.59
Weekly Total Wet NH_4 Deposition (units eq ha^{-1} $week^{-1}$, where applicable)								
FAC2	0.28	0.33	0.00	0.41	0.42	**0.44**	0.31	0.14
MB	-1.51	-1.29	-2.97	0.39	0.38	**0.08**	-1.19	-2.18
MGE	2.13	**2.03**	2.97	2.46	2.44	2.18	2.12	2.43
NMGE	0.72	**0.68**	1.00	0.82	0.82	0.73	0.71	0.82
RMSE	4.29	**4.13**	5.49	5.06	5.02	4.42	4.25	4.78
R	0.50	0.53	0.29	0.51	0.51	**0.54**	0.50	0.40
COE	0.25	**0.28**	-0.05	0.13	0.14	0.23	0.25	0.14
IOA	0.62	**0.64**	0.47	0.57	0.57	0.62	0.63	0.57
Daily Total Wet NO_3 Deposition (units eq ha^{-1} d^{-1}, where applicable)								
FAC2	0.39	0.39	0.00	0.38	0.39	**0.49**	0.43	0.28
MB	-0.18	-0.16	-0.68	-0.26	-0.19	-0.07	**-0.05**	-0.34
MGE	**0.44**	**0.44**	0.68	0.45	0.46	**0.44**	0.48	0.52
NMGE	0.65	0.65	1.00	0.66	0.68	**0.64**	0.71	0.76
RMSE	**0.80**	**0.80**	1.16	0.84	0.85	0.83	0.89	0.97
R	0.61	**0.62**	0.22	0.56	0.56	0.59	0.55	0.44
COE	0.28	0.28	-0.11	0.27	0.25	**0.29**	0.22	0.15
IOA	**0.64**	**0.64**	0.45	0.63	0.63	**0.64**	0.61	0.58
Weekly Total Wet NO_3 Deposition (units eq ha^{-1} $week^{-1}$, where applicable)								
FAC2	**0.50**	**0.50**	0.00	0.42	0.45	0.49	0.43	0.33
MB	-0.10	**-0.06**	-1.86	-0.64	-0.41	**0.06**	0.10	-0.87
MGE	**1.09**	**1.09**	1.86	1.12	1.12	1.17	1.34	1.26
NMGE	**0.58**	0.59	1.00	0.60	0.60	0.63	0.72	0.68
RMSE	**1.86**	1.88	2.93	1.96	1.95	1.93	2.23	2.19
R	**0.65**	**0.65**	0.35	0.58	0.58	0.60	0.53	0.48
COE	**0.32**	**0.32**	-0.16	0.30	0.30	0.27	0.16	0.21
IOA	**0.66**	**0.66**	0.42	0.65	0.65	0.64	0.58	0.61

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

The spatial distribution of PM2.5 ammonia biases was examined in Figure 18, 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, Figure 18) are in the same station locations for all models, in the agricultural region to south of the Great Lakes. Positive PM2.5 ammonium MB also occur near urban regions in 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 emissions parameters and primary particle ammonium emissions inventories are 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 SW USA (regions of sea-spray NaCl and wind-blown 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 to 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.

Figure 18. Mean Biases, PM2.5 NH_4, July, 2016, by station (μg m^{-3}). Negative values given in blue, positive biases given in red. Note that colour scale is logarithmic.

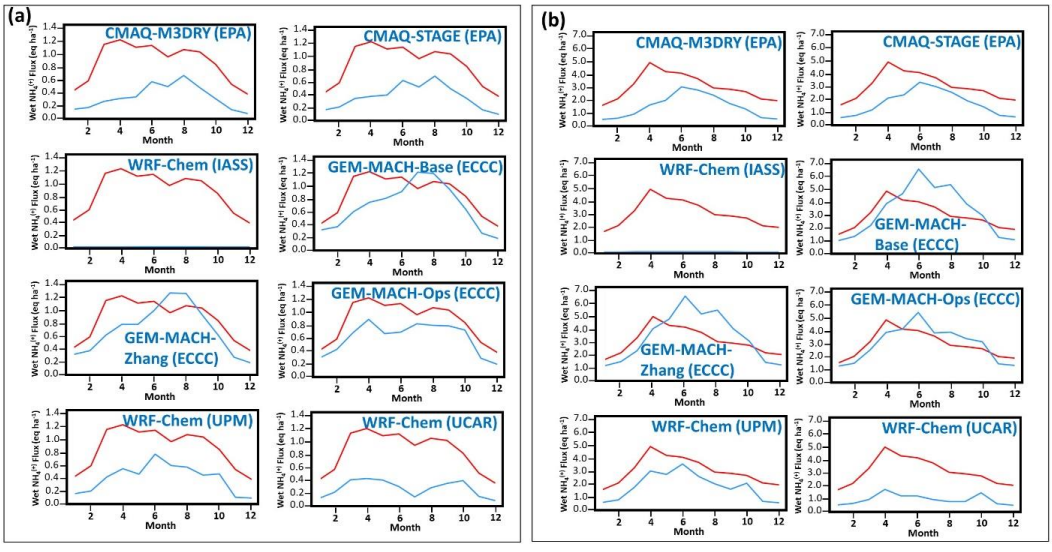


https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

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 7). 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) wet NH_4^+ deposition fluxes are shown in Figure 19. The monthly mean of modelled daily values (Figure 19(a)) 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 (Figure 19(a), 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 between June through August. The monthly average of the weekly wet NH_4^+ deposition fluxes (Figure 19(b)) shows a similar pattern, with the observed values (red lines, Figure 19(b)) peaking in April, and all of the models except for WRF-Chem (UCAR) peaking in June. As was the case for wet sulphate deposition, the observed seasonal variation is apparently not connected with biases in precipitation predictions (see Figure S10(a,b), supplemental information), with the possible exception of WRF-Chem (UCAR), for which total precipitation is biased substantially negative throughout the year.

Figure 19. Time series of monthly average observed (red line) and modelled (blue line) wet ammonium deposition fluxes, for (a) Daily CAPMoN data (eq ha^{-1} day^{-1}), and (b) Weekly NADP data (eq ha^{-1} week^{-1}).



As noted above, the models taking part in this ensemble did not make use of multiphase hydrometeor scavenging in precipitation. The maximum wet NH_4^+ 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 respectively, Figure S7(a).

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

The mean biases in average daily and weekly wet NH_4^+ deposition for the month of April are shown in Figure 20. 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 scavenging schemes are decoupled from meteorological feedbacks, while GEM-MACH (Base), GEM-MACH (Zhang), WRF-Chem (UPM) and WRF-Chem (UCAR) are "aerosol-aware"/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), 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 off-line 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 a 0 C may be used to distinguish between liquid and solid hydrometeor collection efficiencies. Subsequent to the AQMEII-4 simulations carried out here, parameterizations that utilize multiphase precipitation data with multiple hydrometeor classes, such as that of Wang *et al.* (2014), have been tested within the modelling framework of GEM-MACH (Ghahreman *et al.,* 2024).(. Similarly, Ryu and Min (2022) describes 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 occurs to cloud droplets; snow scavenging is not considered. However, snow scavenging at higher precipitation rates is known to be one to two 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, (2014) parameterization in GEM-MACH, and the omission of multiphase hydrometeor scavenging in CMAQ, may account for the springtime bias in all models noted here.

The causes for the differences in wet deposition of NH_4 between WRF-Chem (IASS), WRF-Chem (UPM) and WRF-Chem (UCAR) may result from the use of the Chapman *et al.* (2009) wet scavenging approach in the first model, and the implementation of 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 and (Morrison *et al.,* 2009), though WRF-Chem (IASS and WRF-Chem (UPM)

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

differ from WRF-Chem (UCAR) in the parameterization of convective clouds (See Table 2). Differences in aerosol scavenging implementations may account for some of the differences in wet ammonium 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 7). CMAQ-M3Dry, CMAQ-STAGE and GEM-MACH (Ops) have the best performance scores for wet nitrate deposition (Table 8). 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 3). The increase in wet deposition negative bias magnitude going from GEM-MACH (Zhang) to GEM-MACH (Base) in Table 8 (from -0.19 to -0.26 eq ha^{-1} d^{-1} for daily CAPMoN data, and from -0.41 to -0.64 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 7, 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 wet nitrate deposition bias between (GEM-MACH (Base, Zhang) and GEM-MACH (Ops) must be due to other factors in the model configuration as described in Table 3. Based on the PM2.5 sulphate and PM2.5 nitrate evaluations (Table 6, Table 8), 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 multiphase hydrometeor scavenging of aerosols. For example, the particulate matter scavenging coefficients for snow are one to two 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 sub-micrometer particle sizes (and hence would otherwise increase PM2.5 concentrations – the increased scavenging associated with multiphase hydrometeors will offset this effect.

Dry Deposition of HNO_3

Dry deposition of HNO_3 is the 2nd largest source of N_{dep} in the reduced ensemble, and the 4th largest source of model-to-model variability.

The spatial variation of the annual sum of the effective deposition fluxes for HNO_3 dry deposition are shown in Figure S11, Figure S12, Figure S13 and Figure S14, 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 Figure 21. 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 depict the contributions of each pathway towards 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 Figure 21 to effective

[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 are reported here as part of the soil pathway.

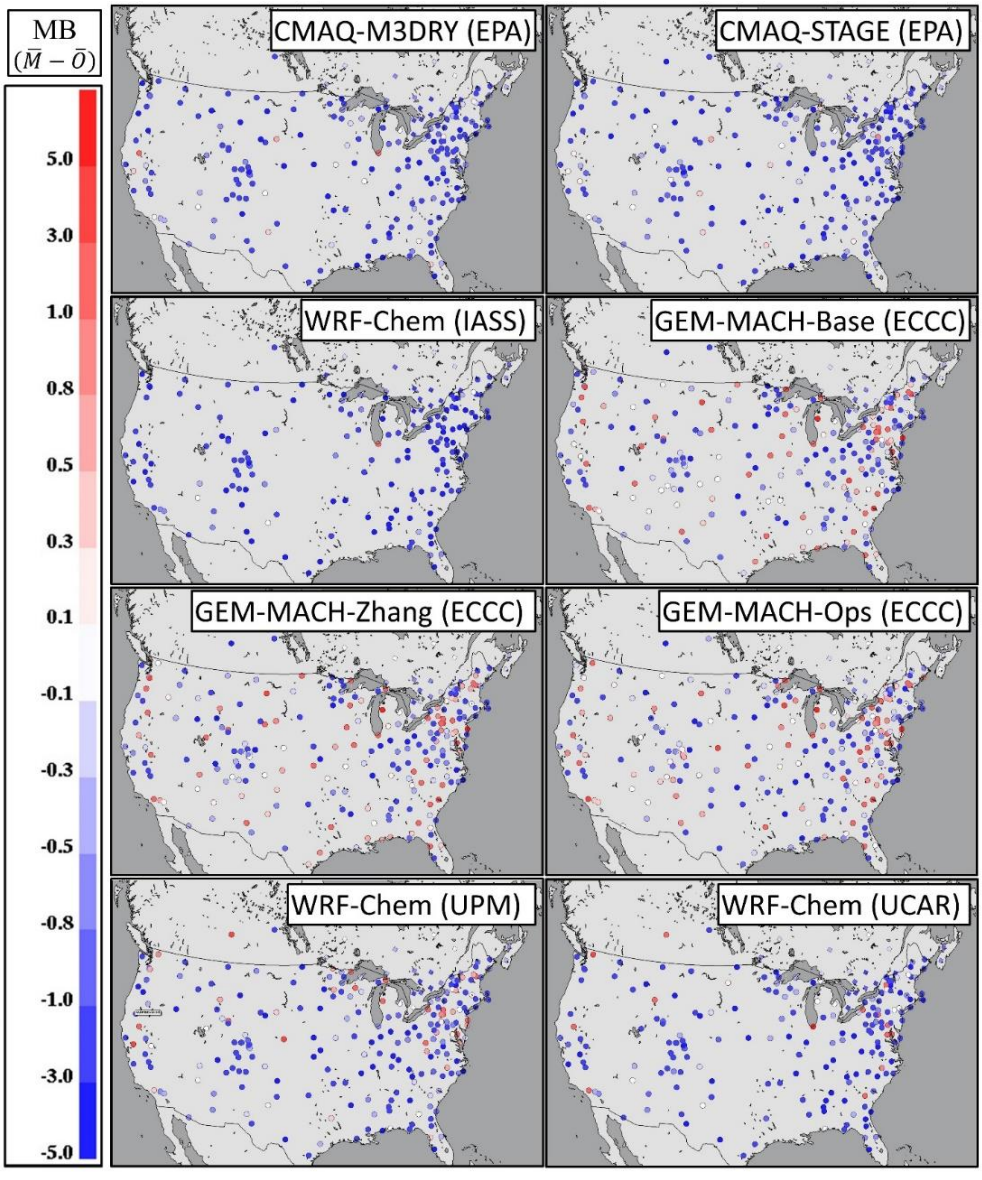
https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

conductances (not shown) has a similar pattern, implying that the deposition velocity is the dominating factor in the HNO3 deposition flux. The HNO_3 mass flux is dominated by the cuticle pathway (Figures S11, 21), followed by the soil pathway (Figures S12, 21). All models show a similar pattern in HNO_3 annual cuticle flux (largest fluxes in the south-eastern 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, 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 depends in part on deposition pathway parameterizations employing functions of the ozone and sulphur 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 modeling framework (Clifton *et al.,* 2023) found that the ozone deposition velocity for GEM-MACH (Base) has positive biases in the summer months (average across 8 sites +73%), negative in the winter months (8 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 dry HNO_3 deposition flux going from GEM-MACH (Zhang) to GEM-MACH (Base) though HNO also deposits via dissociation (sulphur dioxide pathway); not all of the observed effects can be attributed to the use of O3 as a proxy in part of the deposition algorithm.. A portion of the increase in the negative bias in wet nitrate deposition going from GEM-MACH (Zhang) to GEM-MACH (Base) is thus the result of higher HNO_3 dry deposition removal of the available nitrate which would otherwise be taken up into clouds.

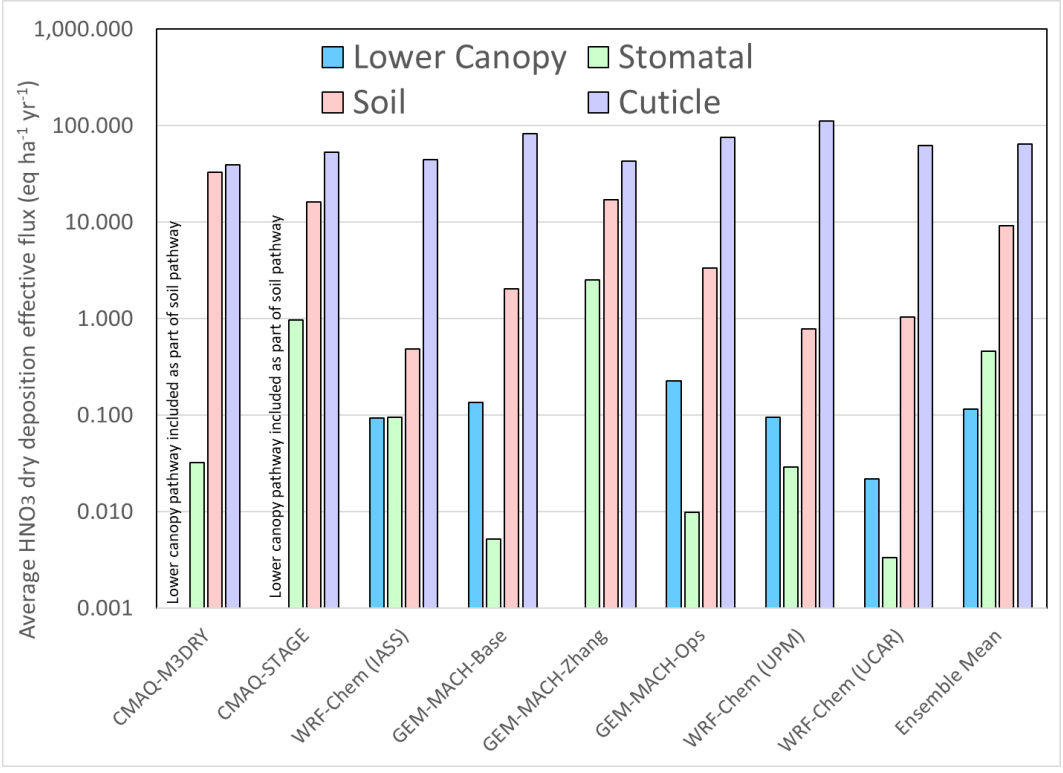
https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Figure 20. Model mean biases in wet ammonium deposition for the month of April, 2016, North America (eq ha^{-1} yr^{-1}). Daily station values of the mean bias (CAPMoN network) shown as diamond symbols, weekly station values (NADP network) as circles. Positive biases shown in red, negative biases shown in blue; note that colour scale intervals are logarithmic.



https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Figure 21. Averages of flux pathway contributions to HNO_3 dry deposition, AQMEII4 common NA grid, 2016 (eq ha^{-1} yr^{-1}).



NH_3 and the role of bidirectional flux algorithms

NH_3 deposition fluxes were the fifth largest driver of ensemble nitrogen deposition, and the 7th 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, Cross-track Infrared Sounding (CrIS) satellite retrievals of NH_3 (see SI for retrieval procedure and references) and AMoN (Chen *et al.,* 2014; AMoN, 2024) surface monitoring network observations (see SI Figure S8(b) 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 two-week integrated average concentrations; the AMoN comparison evaluates average model performance on this integrated time scale, 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 9 and Figure 22. 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 FAC2, highest MGE, lowest R, lowest COE and lowest IOA scores in Table 9. This suggests that the magnitude of the fluxes and/or the balance between positive (downward; deposition) and negative (upward; emission) fluxes for CMAQ-

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

M3Dry and CMAQ-STAGE are the cause of the model's relatively poor performance for NH_3. GEM-MACH (Base) and GEM-MACH (Zhang) have the smallest (and positive) baises 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 Figure 22 also shows some significant differences between observed ammonia and all models' predicted ammonia, particularly over water bodies (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 ~0.3 ppbv) throughout the common AQMEII4 domain.

NH_3 emissions from natural sources has 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 Figure 4). Subsequent simulations of oceanic outgassing (Paulot *et al.*, 2020) showed oceanic outgassing 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 emissions 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.

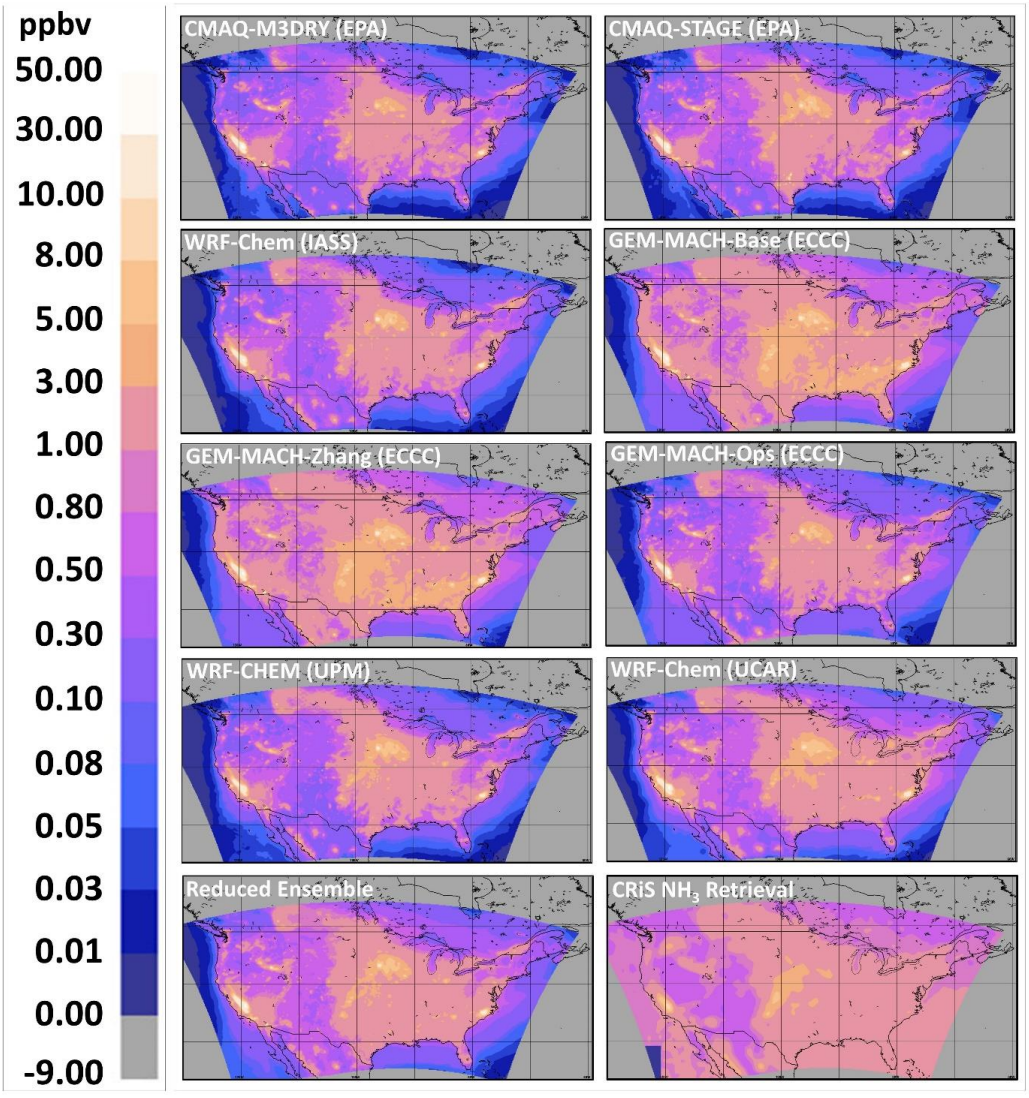
The evaluation of the models' NH_3 against biweekly surface observations at the AMoN sites is shown in Table 10, where biweekly values have been used to create annual averages from both model and observed values at observation sites. GEM-MACH (Base) and GEM-MACH (Zhang) once again have the lowest magnitude (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 mean bias and correlation coefficient.

Table 9. Evaluation of model predictions of NH_3 against retrieved CrIS NH_3 concentrations at overpass time, AQMEII4 common NA grid, 2016. Units ppbv where required.

Evaluation Metric	CMAQ-M3Dry	CMAQ-STAGE	GEM-MACH (Base)	GEM-MACH (Zhang)	GEM-MACH (Ops)	WRF-Chem (UPM)	WRF-Chem (UCAR)
FAC2	0.28	0.38	**0.68**	**0.68**	0.40	0.38	0.58
MB	-0.68	-0.57	**0.09**	**0.09**	-0.54	-0.54	-0.27
MGE	0.83	0.76	0.63	0.63	0.72	0.72	**0.61**
NMGE	0.64	0.58	0.48	0.48	0.55	0.56	**0.47**
RMSE	1.16	1.03	1.07	1.06	1.00	**0.94**	1.00
R	0.66	0.72	0.77	**0.78**	0.70	0.76	0.74
COE	-0.63	-0.50	-0.24	-0.24	-0.41	-0.43	**-0.21**
IOA	0.18	0.25	0.38	0.38	0.29	0.29	**0.40**

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Figure 22. Comparison of annual average surface NH_3 concentrations at CrIS overpass times, participating models, reduced ensemble, and corresponding CrIS observed average NH_3 at overpass time.



https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

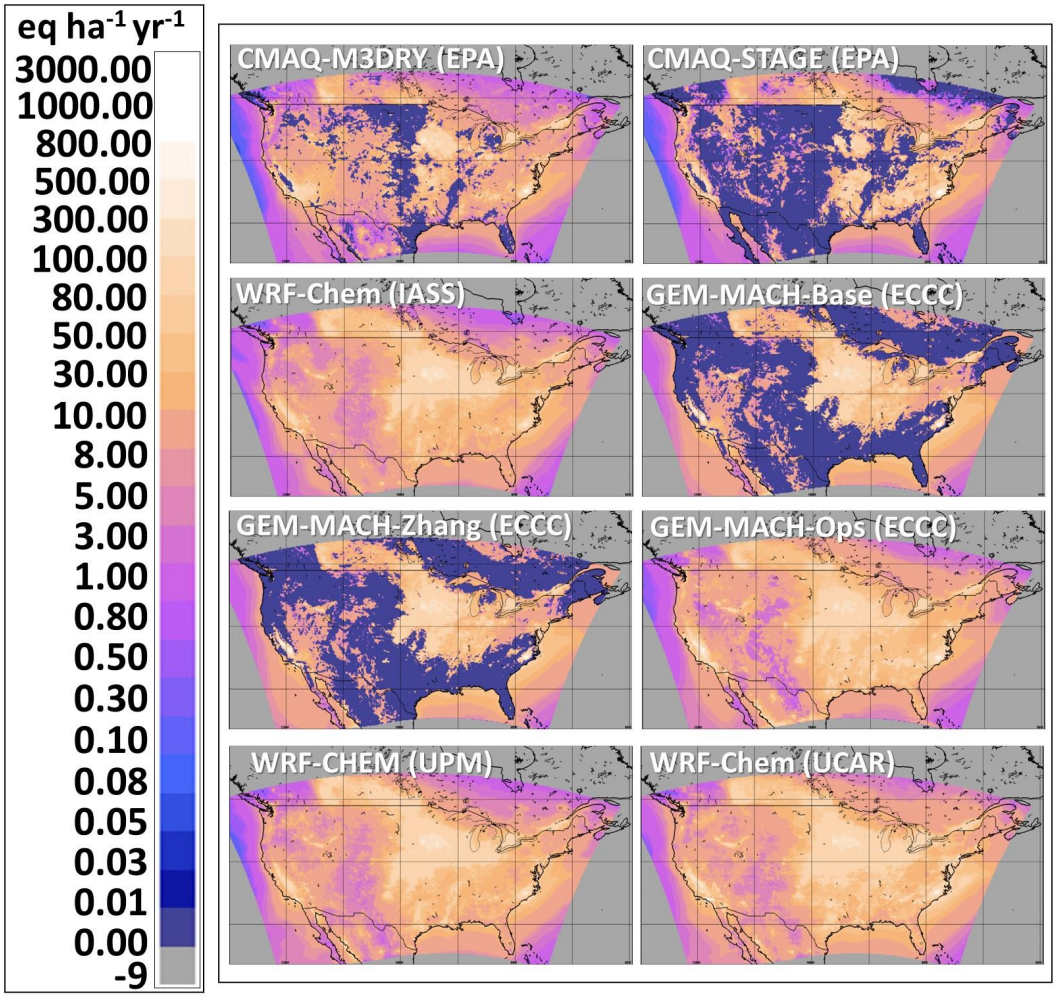
Table 10. Evaluation of model predictions of NH_3 against annual average AMoN biweekly NH_3 concentrations model-observation pairs, 2016. Units ppbv where required.

Evaluation Metric	CMAQ-M3Dry	CMAQ-STAGE	GEM-MACH (Base)	GEM-MACH (Zhang)	GEM-MACH (Ops)	WRF-Chem (UPM)	WRF-Chem (UCAR)
FAC2	0.66	0.62	0.67	0.67	0.72	**0.76**	0.66
MB	-0.82	-0.88	**0.09**	**0.02**	-0.80	-0.61	0.27
MGE	1.24	1.12	1.21	1.18	1.12	**1.08**	1.28
NMGE	0.60	0.54	0.59	0.57	0.54	**0.52**	0.62
RMSE	2.71	2.53	2.72	2.72	2.65	**2.57**	2.95
R	0.37	**0.45**	0.39	0.39	0.39	0.40	0.38
COE	0.21	0.29	0.23	0.25	0.29	**0.32**	0.19
IOA	0.61	0.65	0.61	0.62	0.64	**0.66**	0.59

Figure 23 shows the contributions to total N deposition flux from 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 Figure 23 (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*; 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 details utilized in the bidirectional flux schemes also are resulting in differences between the two modelling platforms (e.g. the north-western USA/south-western Canada border shows up as a sharp contrast in the CMAQ models NH_3 fluxes that utilize information from EPIC over the US and less detailed information outside the US while this differences is much less pronounced in the GEM-MACH models).

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Figure 23. 2016 N dry deposition fluxes (eq ha^{-1} yr^{-1}) for NH_3(g) (eq ha^{-1} yr^{-1})



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:

$$F_T = \frac{c_a - c_c}{r_{sum}} \quad (13)$$

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 is the canopy compensation point concentrations 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 equation (13), if the atmospheric

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

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) gives the detailed formulae for the terms in equation (13), 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 in the between the predicted NH_3 dry deposition fluxes for the models employing bidirectional flux parameterizations for the AQMEII4 North American ensemble. These terms were reported by AQMEII4 participants as the diurnal median (50th percentile) at each UT hour within each month. The median values for 16UT (noon EDT) for July 2016 are shown in Figure 24. It is important to note that the median values for a given UT hour may correspond to different days within a given month. For example, the median values of r_{sum} and r_a at 16 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 off-line of the model simulation, and could not be included as AQMEII4 diagnostic parameters). However, substantial differences between the panels of Figure 24 provide a useful indication of relative importance of different pathways in the participating models.

From Figure 24, we note:

(1) The 2016 July, 16 UT median aerodynamic resistance r_a is similar for all four models (Figure 24(a)) – consequently, differences in r_a are unlikely to be the cause of the model flux differences.
(2) The 2016 July, 16 UT median r_{sum} values (Figure 24(b)) for CMAQ-M3Dry 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 equation (13), will increase the magnitude of the fluxes). *et al.*
(3) The 2016 July, 16 UT median r_{sum} values for CMAQ-STAGE over land 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 2016 July, 16 UT median values of the air concentrations of NH_3, c_a (Figure 24(c)) 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 2016 July, 16 UT median net compensation point concentration c_c (Figure 24(d)) for CMAQ-STAGE is an order of magnitude smaller than for GEM-MACH (Base) and GEM-MACH (Zhang). From equation (13), 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 2016 July, 16 UT median ground compensation point concentration (c_g) has maximized are where GEM-MACH (Base) and GEM-MACH (Zhang) have zero to near-zero ground compensation point values (Figure 24(e) – e.g. Rocky mountains, north-central USA 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 1E4 ppbv for this model), if dominant, would be expected to result in larger c_c values in equation (13) (see Galmarini *et al.* 2021) and hence a tendency towards smaller downward fluxes. This is not the case from the above analysis (DNH_3 values in Table 7 for CMAQ-STAGE are greater than those of the GEM-MACH models, and CMAQ-STAGE NH3 concentrations have more negative biases than the two GEM-MACH models),

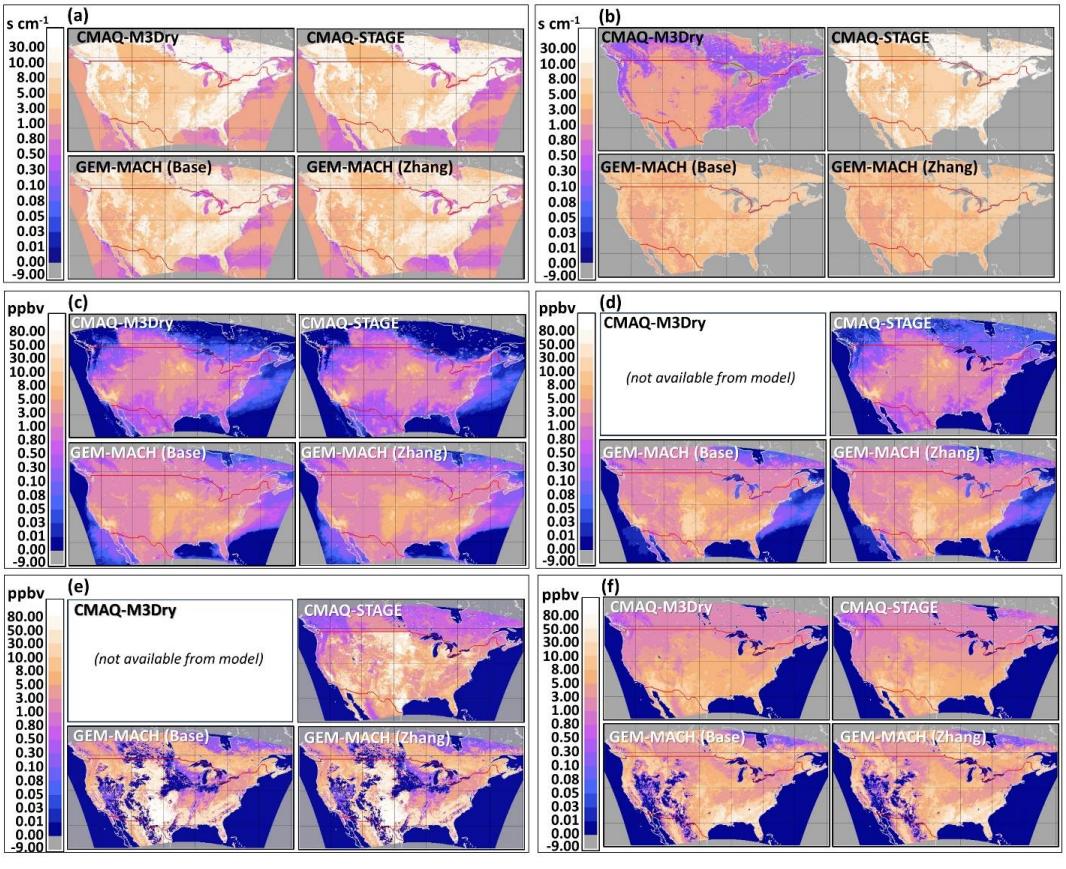
https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

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 SW USA), CMAQ-M3Dry and CMAQ-STAGE have lower 2016 July, 16 UT median stomatal compensation point concentrations than either GEM-MACH (Base) or GEM-MACH (Zhang) (Figure 24(f)). 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).

Figure 24. 2016 Spatial distribution of 2016 July, 16 UT median n values for key bidirectional flux diagnostic variables. (a) Aerodynamic resistance (s cm^{-1}), r_a. (b) Sum resistance (s cm^{-1}), r_{sum}. (c) Air Concentration of NH_3 (ppbv), c_a. (d) Net compensation point concentration (ppbv), c_c. (e) Ground compensation point concentration (ppbv), c_g. (f) Stomatal compensation point concentration (ppbv), c_s.



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

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

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 Figures 25 and 26, where Figure 25 panels (a and b) are the sum of AQMEII4 land use types 11 and 12 (i.e. the sum of "planted/cultivated" and "grassland" land use types) used in CMAQ and GEM-MACH respectively. Figure 25 panels (c and) 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 Figure 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 26 shows an example of this diurnal behaviour of the NH_3 bidirectional fluxes for the CMAQ and GEM-MACH models, at (a) 15:00 CDT and (b) 7:00 CDT. Mid-afternoon fluxes (Figure 26(a)) 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 south-east Canada and south-east USA. CMAQ-STAGE predicts mid-afternoon emissions throughout western North America, though a similar pattern of deposition as CMAQ-M3Dry in south-east Canada and south-east USA. The GEM-MACH bidirectional fluxes in 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 (Figure 22, CrIS NH_3 retrieval maximum). In contrast, early morning fluxes (Figure 26(b)) 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 south-east Canada and south-east USA.

The generic diurnal sign changes in the direction of the ammonia flux across all four models is easily explained with reference to equation (13): in mid-afternoon (Figure 26(a)), 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 eqn 13), resulting in negative fluxes (emissions; blue colours). In the early morning (Figure 26(b)), 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, and 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.

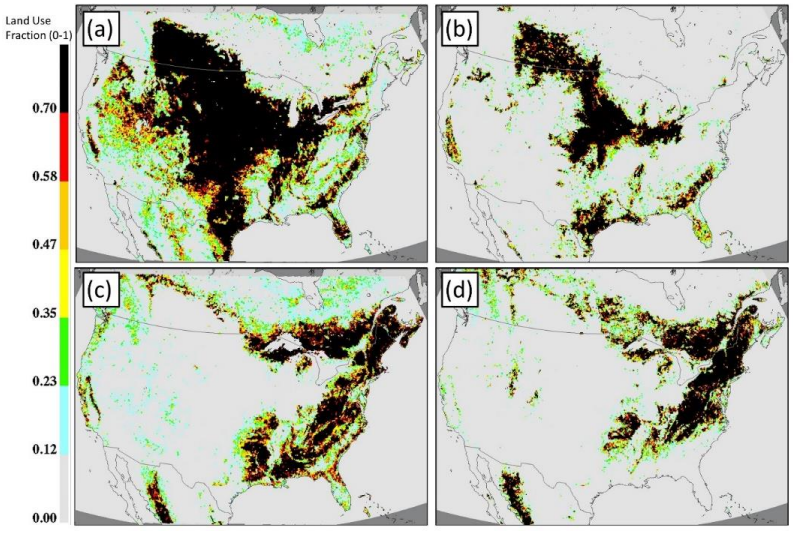
We note that Table 9 measures model performance specifically at satellite overpass time in the afternoon – i.e. at close to the time shown in Figure 26(a), and that the performance of CMAQ-M3Dry and CMAQ-STAGE is lower than the other models at this time, while the differences between the models aside from magnitude of the bias is less pronounced in the integrated surface observations of Table 9. 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 (eqn. 13) 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

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

pathway. Conversely, we note that the GEM-MACH bidirectional flux algorithm is overestimating midafternoon ammonia in the SE USA relative to satellite observations (Figure 22), indicating that compensation point concentrations may be overestimated in this region.

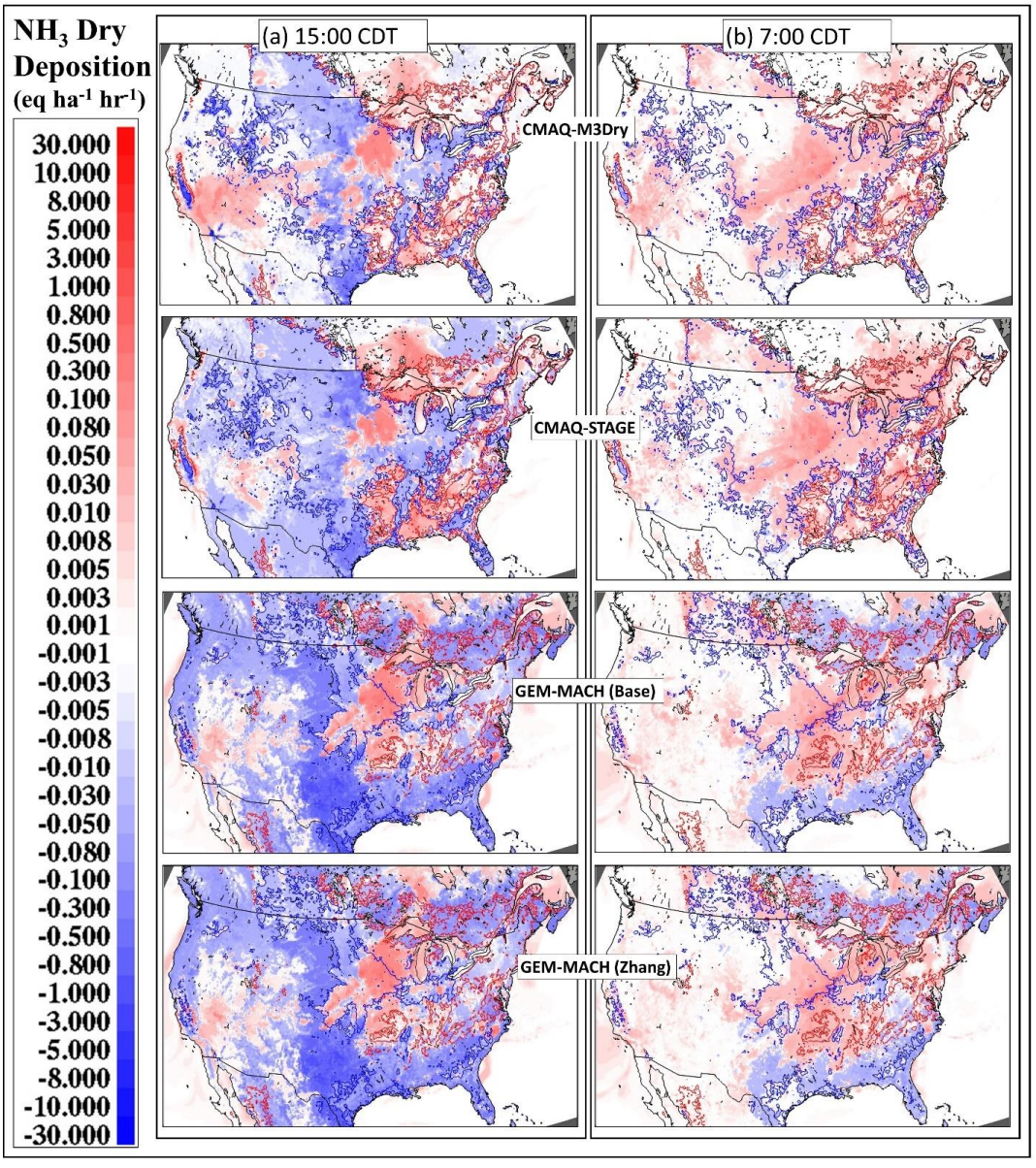
While NH_3 fluxes are only the 5th largest source of N deposition in the North American reduced ensemble, we also note that the manner in which NH_3 bidirectional fluxes are treated in the context of critical load exceedance calculations may be open to interpretation. Exceedances with respect to critical 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 towards 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 emissions 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 in lower estimates of total N_{dep}. Although NH_3 is the 5th 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.

Figure 25. Comparison of AQMEII4 land use type fractions with locations of highest CLE for forest ecosystems, CMAQ versus GEM-MACH. Upper row: grid cell fractional area composed of sum of AQMEII4 land use types 11+12 (planted/cultivated and grassland), for: (a) CMAQ-M3Dry and CMAQ-STAGE, (b) GEM-MACH (Base) and GEM-MACH (Zhang). Lower row: grid cell fractional area composed of sum of AQMEII4 land use types 6+7+8+13 (evergreen broadleaf forest, deciduous broadleaf forest, mixed forest, and savanna), for (c) CMAQ-M3Dry and CMAQ-STAGE, (d) GEM-MACH (Base) and GEM-MACH (Zhang).



https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Figure 26. NH_3(g) flux (eq ha^{-1} hr^{-1}) at (a) 15:00 CDT August 4, 2016 and (b) 7:00 CDT August 5, 2016. Blue lines in the CMAQ and GEM-MACH models (horizontal row) panels encloses areas which are predominantly agricultural and grassland, red line encloses areas which are predominantly evergreen broadleaf forest, deciduous broadleaf forest, mixed forest and savanna, in each model's respective land use databases (se Figure 25). Blue shaded regions indicate negative (upward; emissions) NH_3 fluxes, red shaded regions indicate positive (downward; deposition) NH_3 fluxes.



https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

3.2.3 Causes of S Deposition Variability in European Domain Simulations

The relative contributions of the different sources of S deposition in the AQMEII4 EU common domain for the year 2010 are shown in Table 11 and Figure 27.

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 wet (SO_4^{2-} + HSO_3^-) deposition, followed by particulate sulphate dry deposition (see Table 11). The relatively higher importance of SO_2 dry deposition towards total sulphur 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 emissions data (and emissions 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 Emissions 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 this 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 emissions 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 July 11, 2024), and the European Air Quality Database (AIRBASE; https://data.europa.eu/data/datasets/data_airbase-the-european-air-quality-database-1?locale=en , last accessed July 11, 2024).

Table 11. Average S deposition contributions in common AQMEII4 EU grid area (eq ha^{-1} yr^{-1}) and percent contribution to average total S deposition, 2010.

	Average Deposition (eq ha^{-1} yr^{-1})				Percent of total S deposition		
Model Number	SO_2(g) Dry Deposition	$SO_4^{(2-)}$ + $HSO_3^{(-)}$ Wet Deposition	Particle Sulphate Dry Deposition	Total S Deposition	SO_2(g) Dry Deposition	$SO_4^{(2-)}$ + $HSO_3^{(-)}$ Wet Deposition	Particle Sulphate Dry Deposition
WRF-Chem (IASS)	92.1	42.1	n.r.	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

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Dry deposition of SO_2

The model SO_2 performance 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 towet ($SO_4^{2-} + HSO_3^-$) deposition (EMEP wet deposition network), are shown in Table 12. Observation station locations are shown in Figure S9. 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 Figure 28, where the AIRBASE monthly concentration y-axis (Figure 29(a)) is almost twice that of the EMEP monthly concentration y-axis (Figure 28(b)). Observed SO_2 close to sources (Figure 28(a), red lines) shows a strong seasonal variability, with concentrations in the winter being a factor of two higher than HNO3summer than summer, likely showing the effect of increased winter stability on plume rise. This tendency is greatly reduced at regional stations (Figure 28(b), red lines). LOTOS-EUROS (TNO) matches the near-source SO_2 time series the most closely, while CMAQ (Hertfordshire) overestimates the impact of seasonal variability (Figure 28(a)). 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 simulated SO_2 performance relative to observations and hence in SO2 deposition fluxes and critical load exceedance estimates is due to some models overestimating the seasonal variation in SO2 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 Sulphur

As was the case for most models on the North American domain, all EU domain models underestimated wet deposition relative to observations (note negative biases in Table 12 and monthly time series comparison versus observations in Figure 28(c)). CMAQ (Hertfordshire) outperforms the other models relative to observations, though we note that the wet sulphur 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 wet sulphur deposition comparison time series (Figure 16, Table 6), the European wet deposition observations do not show a spring-time peak in values, rather a seasonality centered around the month of June, with higher values extending from March to September.

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Table 12. Model performance statistics for EU domain SO_2 concentrations and total wet S deposition, μg m^{-3} and eq ha^{-1} yr^{-1}, respectively.

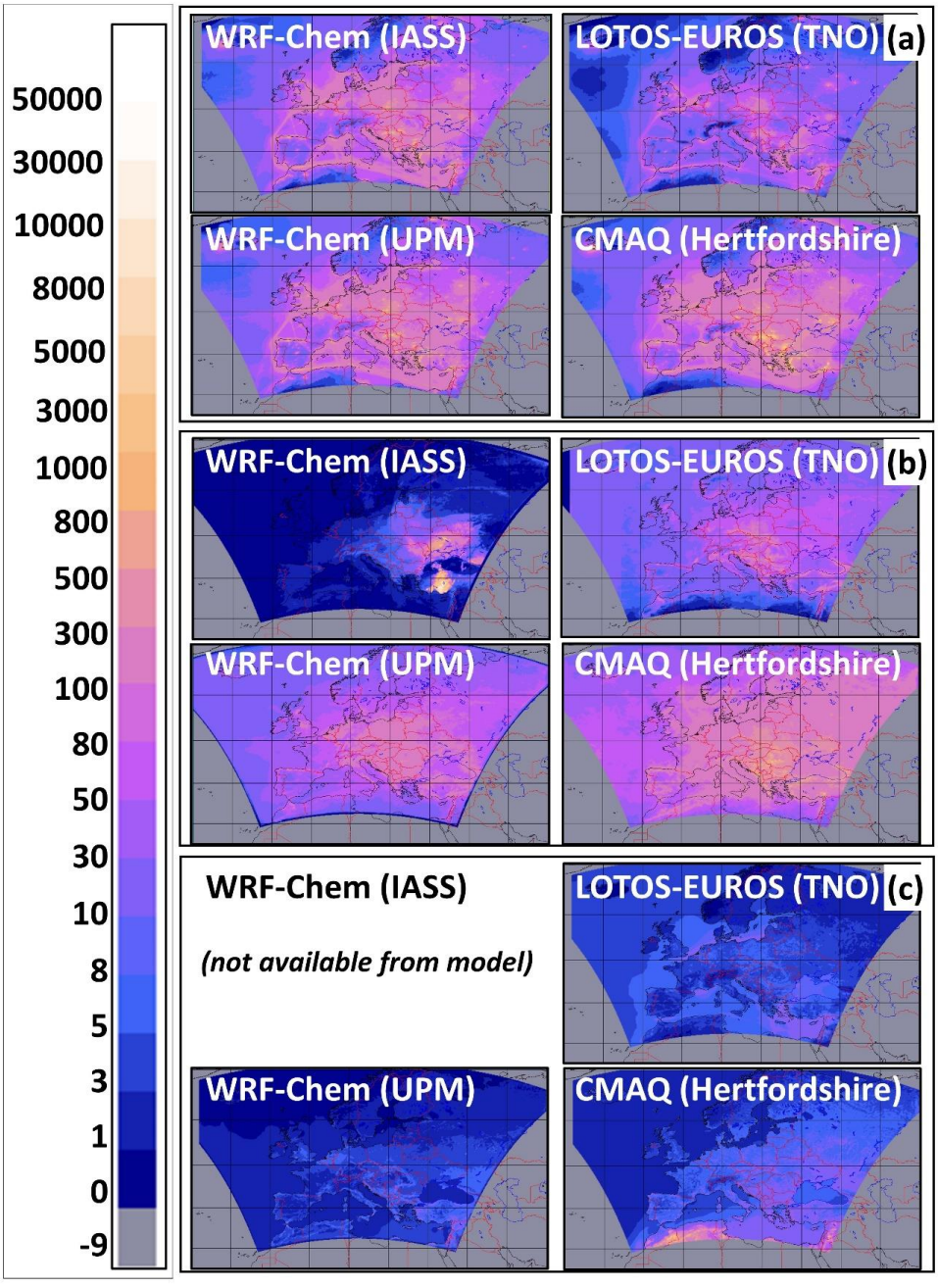
	SO_2 (Airbase)					SO_2 (EMEP)			
	WRF-Chem (IASS)	LOTOS-EUROS (TNO)	WRF-Chem (UPM)	CMAQ (Hertfordshire)		WRF-Chem (IASS)	LOTOS-EUROS (TNO)	WRF-Chem (UPM)	CMAQ (Hertfordshire)
FAC2	0.35	0.36	**0.38**	0.35		0.35	**0.36**	0.34	0.29
MB	-1.42	**0.04**	0.06	1.89		**0.32**	0.48	0.58	1.76
MGE	**4.60**	5.32	5.29	6.35		**1.48**	1.66	1.63	2.57
NMGE	**0.85**	0.98	0.97	1.17		**1.07**	1.20	1.18	1.87
RMSE	**14.47**	15.64	15.27	17.60		**2.92**	3.58	2.98	5.80
R	**0.28**	0.26	0.24	0.26		**0.38**	0.33	0.34	0.35
COE	**0.12**	-0.01	-0.01	-0.21		**-0.08**	-0.21	-0.19	-0.88
IOA	**0.56**	0.49	0.50	0.40		**0.46**	0.40	0.40	0.06
	Total Wet S deposition								
	WRF-Chem (IASS)	LOTOS-EUROS (TNO)	WRF-Chem (UPM)	CMAQ (Hertfordshire)					
FAC2	0.00	0.19	0.28	**0.31**					
MB	-1.51	-1.22	-1.08	**-0.39**					
MGE	1.53	1.34	**1.29**	1.42					
NMGE	1.00	0.87	**0.84**	0.92					
RMSE	6.61	6.50	6.48	**6.46**					
R	0.02	0.11	0.11	**0.15**					
COE	0.04	0.16	**0.19**	0.11					
IOA	0.52	0.58	**0.60**	0.56					

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 towards EU model-to-model variability is small. Speciated PM observations were not available for comparison to model predictions in the EU region.

The spatial distribution of the relative contributions of the three forms of sulphur deposition for the year 2010 is shown in Figure 27. CMAQ (Hertfordshire), with the highest SO_2 deposition flux (Figure 27(a), see also Table 11, Table 12) 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 12). In contrast, all models have significant negative biases in wet sulphur deposition (Table 12), hence at least one reason for this underestimate may be insufficient conversion of SO_2 to ionic sulphate and bisulphite in simulated cloud water, through uptake of SO_2 and scavenging of particulate sulphate. The wet deposition of sulphur in WRF-Chem (IASS) in particular seems anomalously low (Figure 27(c), Figure 28 (b)), with much of Europe having little to no wet sulphate deposition in this model.

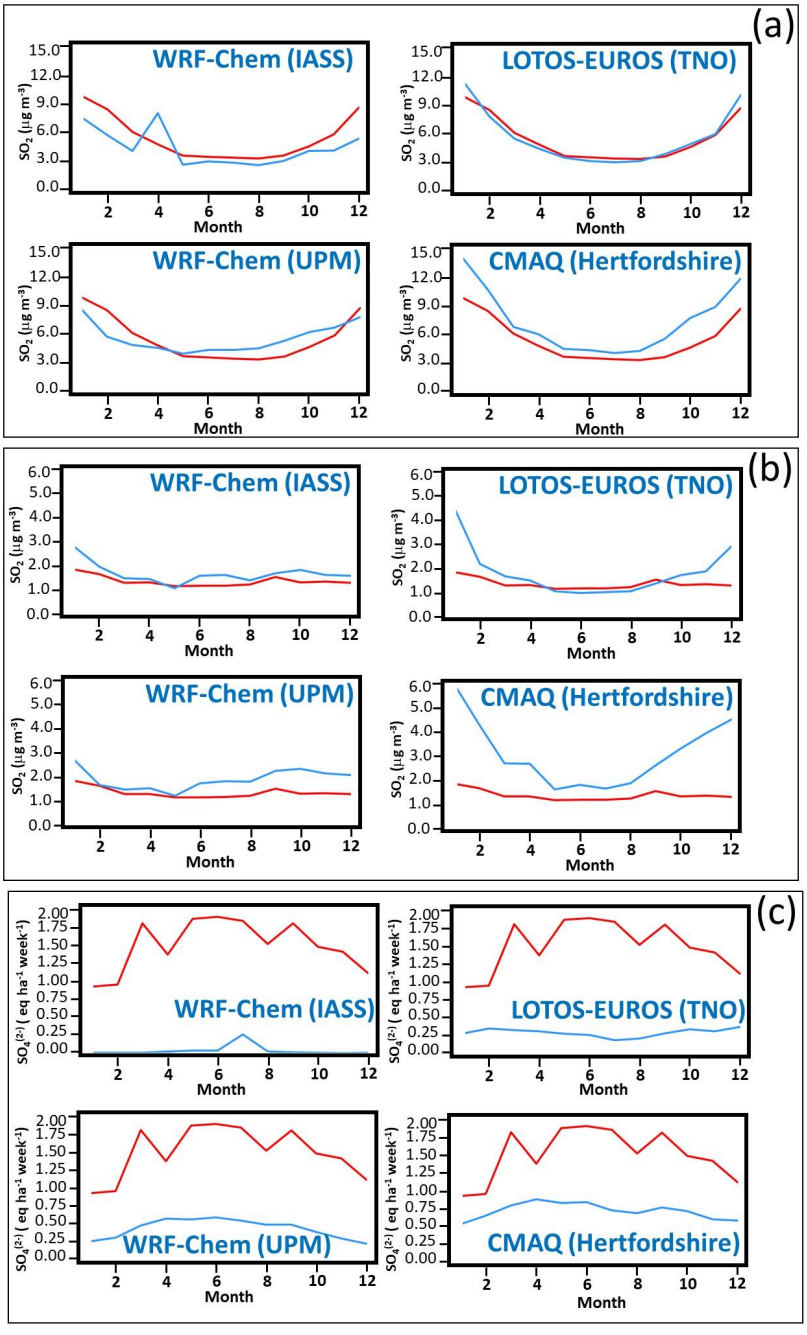
https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Figure 27. Spatial distribution and magnitude of contributions to annual S deposition, AQMEII4 common EU domain, 2010 (eq ha^{-1} yr^{-1}). (a) SO_2(g) dry deposition. (b) Total wet S deposition. (c) Particle sulphate dry deposition.



https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

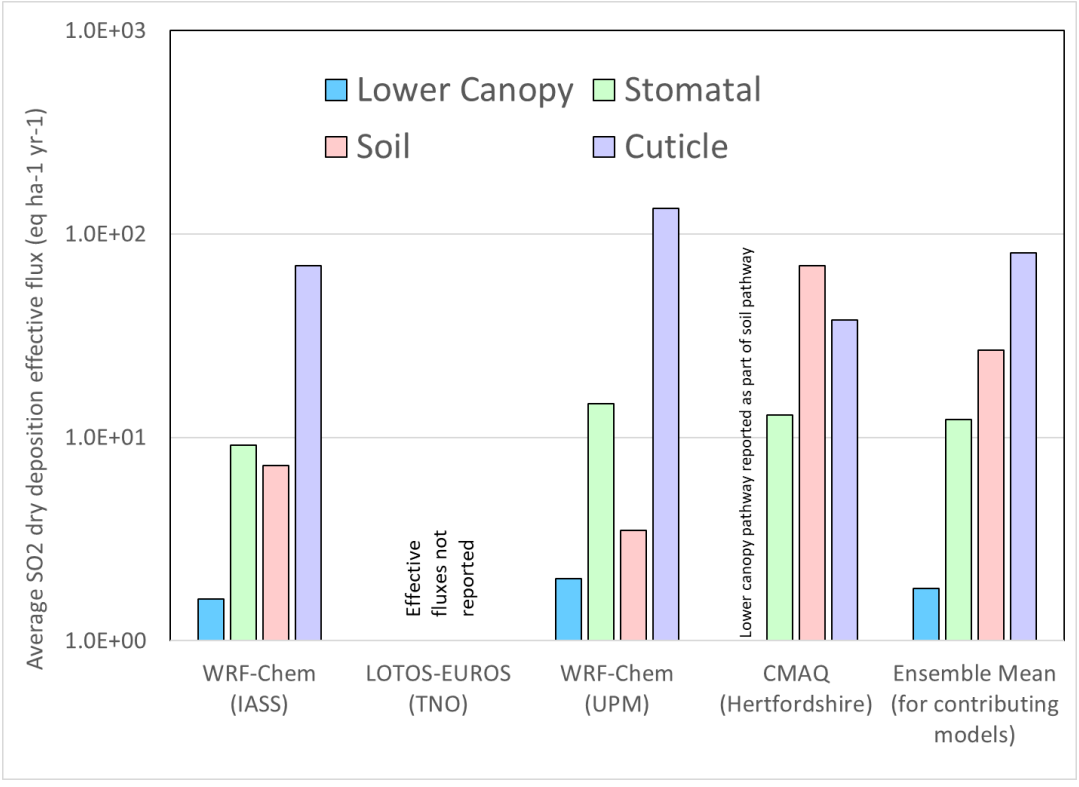
Figure 28. Comparison of observed and modelled S, AQMEII4 EU common domain, 2010. (a) AIRBASE SO_2 (ug m^{-3}). (b) EMEP SO_2 (ug m^{-3}). (c) Wet flux of total S deposition (eq ha^{-1} $week^{-1}$). Red: observations. Blue: model.



https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

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 S15 and S16 show the spatial distribution of the summed annual effective fluxes for the reporting models, with the results in the common AQMEII4 EU domain summarized in Figure 29.

Figure 29. Averages of effective flux pathway contributions to SO_2 dry deposition, AQMEII4 common EU grid, 2010 (eq ha^{-1} yr^{-1}).



Despite having the highest average SO_2 deposition flux (Table 11), CMAQ (Hertfordshire) also has the highest positive biases for SO_2 ambient concentrations (Table 12). From Figures S15, S16 and 29, 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 (Figure S15(a)). Despite these relatively low values, the SO_2 net dry deposition flux for this model (Table 11) is higher than that of the other models, implying that the low northern EU fluxes are being offset by higher values elsewhere (eg. via the soil flux, compare soil and cuticle values in Figure 29). 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 Figure 29. Other than Northern Europe, CMAQ (Hertfordshire) has higher soil fluxes than WRF-Chem (IASS). Similar to AQMEII4 analyses for ozone (Hogrefe *et al.,* 2024, under preparation), the

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

relative importance of the different pathways towards 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 12), the flux contributions in descending order of importance are soil, cuticle, and stomatal (with 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 common AQMEII4 EU domain relative contributions for each model's deposited species towards total nitrogen deposition and its variability are shown in Table 13. The contributions towards 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 Figure 30. 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 towards 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 14; monthly average time series comparisons wet deposition to the observations are provided in Figure 31. From Figure 30, WRF-Chem (IASS) predicted much lower magnitude wet NO_3^- and wet NH_4^+ deposition fluxes than the other three models, and from Table 14, 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 wet S deposition, all models have significant negative biases for both nitrogen ion wet fluxes, as can be seen from Table 14 and Figure 31. 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 wet NO_3^- 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 sulphate and nitrogen species. Also, we note that the observed wet NH_4^+ deposition (Figure 31(b), red line) peaks in June, while the model values (blue lines) peak earlier, in March. This in in contrast to the North American NH_4^+ comparison (Figure 19), where observed peaks occur in April and model peaks occur in June.

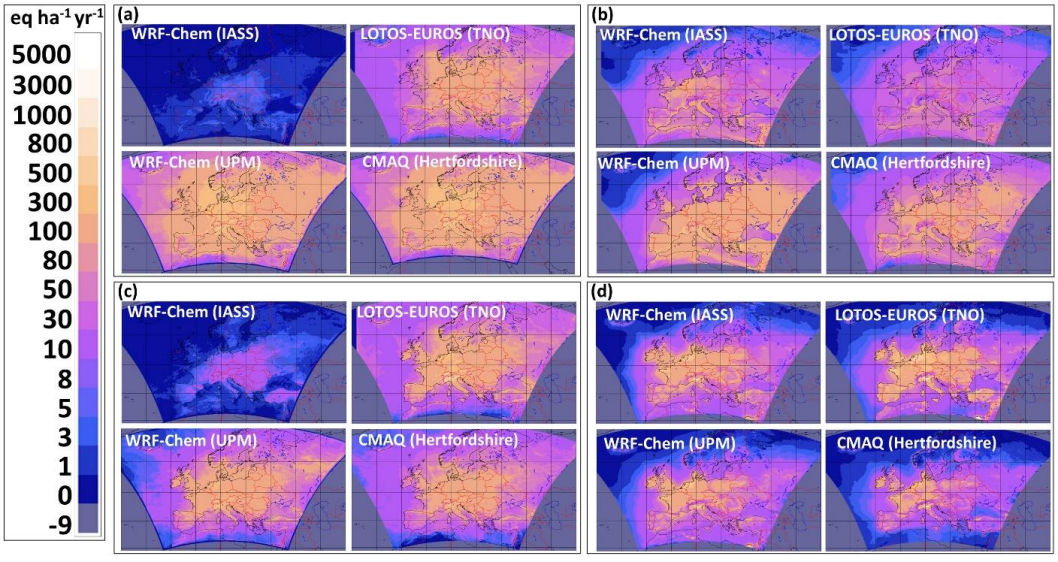
https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Table 13. Contributions of N species towards total deposition (eq ha^{-1} yr^{-1} and percent of total N deposited, common AQMEII4 EU 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. DNO_2: dry deposition of NO_2(g). DPAN: dry deposition of peroxyactylnitrate 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). 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. Std Dev
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. Std Dev
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	nd	0.6	1.7	1.1	0.6
DN2O5	3.8	0.3	nd	0.3	0.3	0
DRN3	0.2	nd	0.1	1.1	0.5	0.5
DHNO4	1	nd	0.2	0.1	0.1	0.1
DNO	0	0.6	0	0.1	0.2	0.2

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Figure 30. Spatial distribution of contributions of (a) wet nitrate ion deposition, (b) dry gaseous HNO3 deposition, (c) wet ammonium ion deposition, and (d) dry gaseous ammonia deposition towards total N deposition in the common AQMEII4 EU domain, 2010 (eq ha^{-1} yr^{-1}).



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 towards the mass flux of HNO_3 is shown in Figures S17 and S18 and are summarized for the entire grid in Figure 32. 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 Figures 21 and 32). 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) is was several orders of magnitude smaller than the same pathway for CMAQ (Hertfordshire). The cuticle pathway dominates 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 6th largest contributor to total N deposition and the 5th largest contributor to model-to-model variability, were available at near-source AIRBASE and regionally distributed EMEP stations (Table 14). Aside from having the 2nd largest magnitude mean bias, LOTOS-EUROS (TNO) had the best performance for NO_2 relative to stations positioned close to emissions 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).

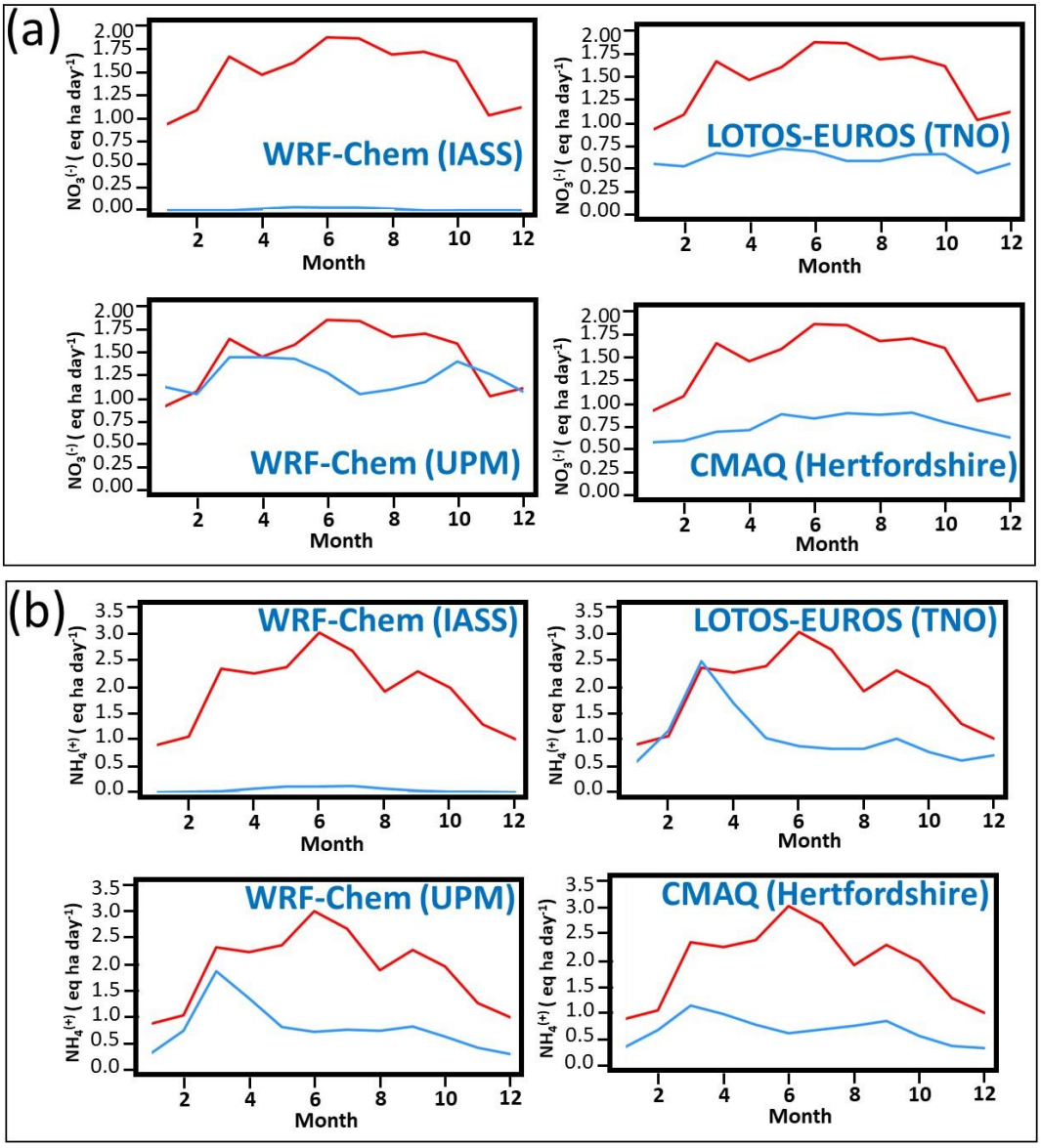
https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Table 14. Model performance statistics for wet deposition of nitrate and ammonium ions, and ground level concentrations of NO_2, AQMEII4 EU domain, 2010

	Wet NO_3^- deposition (eq ha^{-1} yr^{-1})					Wet NH_4^+ deposition (eq ha^{-1} yr^{-1})			
	WRF-Chem (IASS)	LOTOS-EUROS (TNO)	WRF-Chem (UPM)	CMAQ (Hertfordshire)		WRF-Chem (IASS)	LOTOS-EUROS (TNO)	WRF-Chem (UPM)	CMAQ (Hertfordshire)
FAC2	0.00	0.32	**0.35**	0.31	FAC2	0.02	**0.32**	0.28	0.24
MB	-1.38	-0.75	**-0.04**	-0.58	MB	-1.80	**-0.80**	-1.01	-1.13
MGE	1.38	**1.04**	1.33	1.11	MGE	1.81	**1.52**	1.55	1.53
NMGE	0.99	**0.75**	0.96	0.80	NMGE	0.98	**0.82**	0.84	0.83
RMSE	2.66	**2.19**	2.53	2.25	RMSE	3.83	**3.37**	3.45	3.42
R	0.16	**0.43**	0.36	0.38	R	0.18	**0.33**	0.32	**0.33**
COE	-0.10	**0.17**	-0.06	0.11	COE	0.00	**0.15**	0.14	**0.15**
IOA	0.45	**0.59**	0.47	0.56	IOA	0.50	**0.58**	0.57	**0.58**
	AIRBASE NO_2 concentrations (µg m^{-3})					EMEP NO_2 concentrations (µg m^{-3})			
	WRF-Chem (IASS)	LOTOS-EUROS (TNO)	WRF-Chem (UPM)	CMAQ (Hertfordshire)		WRF-Chem (IASS)	LOTOS-EUROS (TNO)	WRF-Chem (UPM)	CMAQ (Hertfordshire)
FAC2	0.45	**0.56**	0.55	0.35	FAC2	**0.57**	0.53	0.39	0.50
MB	-10.00	-5.68	**2.38**	-12.40	MB	**0.36**	2.35	9.54	-2.02
MGE	12.67	**11.22**	13.61	13.84	MGE	5.01	6.18	11.49	**4.90**
NMGE	0.60	**0.53**	0.65	0.66	NMGE	0.57	0.70	1.31	**0.56**
RMSE	19.25	**16.76**	19.19	20.41	RMSE	**8.17**	10.01	17.28	8.29
R	0.49	**0.56**	0.47	0.50	R	**0.71**	0.64	0.61	0.67
COE	0.10	**0.20**	0.03	0.02	COE	0.31	0.15	-0.59	**0.32**
IOA	0.55	**0.60**	0.52	0.51	IOA	0.65	0.57	0.21	**0.66**

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

1769 Figure 31. Monthly average comparison of wet nitrogen deposition, AQMEII4 common EU grid, 2010. (a) Average
1770 flux of NO_3^-(aq). (b) Average flux of NH_4^+(aq). (eq ha^{-1} day^{-1})



1771

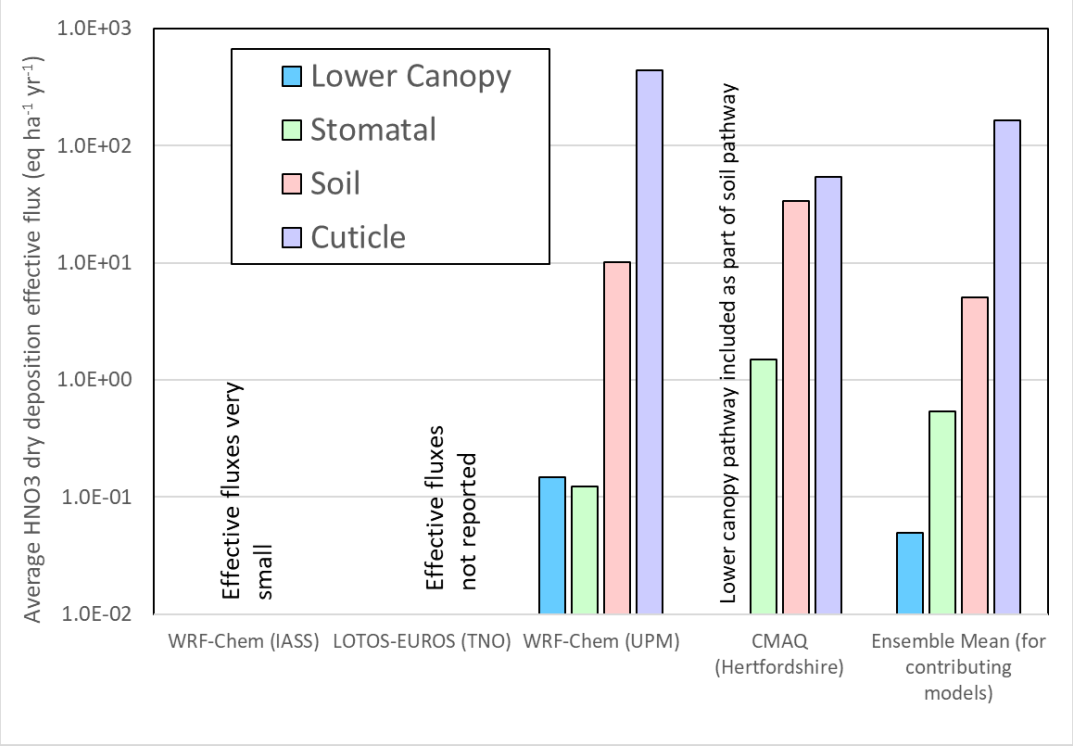
1772

1773

1774

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Figure 32. Averages of flux pathway contributions to HNO_3 dry deposition, AQMEII4 common EU grid, 2010 (eq ha^{-1} yr^{-1}).



Conclusions

We have used the AQMEII4 North American and European ensembles to calculate net Sulphur 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, and the critical load exceedance and causes of model variability, within each domain.

North America, Critical Load Exceedances

All simulations showed a decrease in the size of the area in exceedance and the severity of exceedances with respect to acidification of forest ecosystems between the years 2010 and 2016. The percentage of North American forest area in exceedance of acidification critical loads was 13.2% (range 2.8% to 22.2%) in 2010, and 6.1% (1.0% to 12.9%) in 2016. Similarly, the percent exceedance with respect to acidity for aquatic ecosystems and the number of water bodies in exceedance decreased between 2010 and 2016 from 21.2% (12.8 to 28.9%) in 2010 to 11.4% (7.3% to 15.8%) in 2016). The decrease in SO_2 emissions between these two years, and the resulting decreases in S deposition for all models, as well as decreases in N deposition for all models, drive these reductions in potential ecosystem damage.

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

For sensitive epiphytic lichen species richness, all models also showed an improvement in the exceedances between 2010 and 2016. The improvement in the *total area* in exceedance predicted by most models was relatively small, 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. The reduced ensemble predicted a decrease in both severity and total area in exceedance from 2010 to 2016 from 81.5% (range 69.3 to 95%) in 2010 to 75.8% (range 63.7% to 90.7%) in 2016.

For herbaceous species community richness, most models showed an improvement, often substantial, in both the total area in exceedance and the severity of exceedances between 2010 and 2016 (13.9%, range 0.4 to 39.5%, in 2010 to 3.9%, range 0.1% to 18.4% in 2016).

All models and the reduced ensemble thus showed improvements in the extent of potential ecosystem damage due to acidifying and eutrophying deposition between 2010 and 2016, in accordance with legislated emissions reductions policies coming into force between the two years. However, the amount of exceedance in any given year and the extent of reduction between the two years varied considerably between the models. Any individual model provides a similar direction of the change between the two years; the range of estimates suggests the utility of model ensembles where possible in estimating critical load exceedances.

North America, Causes of Model S Deposition Variability

The total mass of North American Sulphur deposition followed, in decreasing order of importance, wet deposition of S ($SO_4^{2-} + HSO_3^-$), dry deposition of particulate sulphate, and dry deposition of SO_2. Dry deposition of particulate sulphate contributed the most to model-to-model variability in total Sulphur deposition, followed by wet deposition, and dry SO_2 deposition. The models with the highest wet S deposition levels had the best performance relative to monitoring network observations (CMAQ-M3Dry, CMAQ-STAGE, GEM-MACH (Ops)), though all models' wet S 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 PM2.5 sulphate 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 algorithm (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).

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 reduced North American ensemble were due to the three GEM-MACH implementations (Base, Zhang, and Ops) all having much higher dry particle ammonium and wet ammonium ion deposition fluxes than the other

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

models. These models also had zero to positive biases in wet ammonium ion deposition relative to observations during the summer (all other models had negative biases throughout the year), and the largest positive biases for PM2.5 ammonium concentrations relative to observations. These three models employed an inorganic aerosol thermodynamics algorithm solving only the sulphate-ammonium-nitrate-water system, while the other models in the ensemble employed a solver which incorporated base cation chemistry. The presence of base cations is known to be a limiting factor on the formation of particle ammonium nitrate, with the available nitric acid preferentially partitioning to the base cations over forming particle ammonium nitrate. The positive biases in fine mode particle ammonium concentrations and positive biases in wet ammonium ion deposition for this subgroup of models are thus likely caused by the absence of this alternative sink of nitric acid, leaving only the particle ammonium formation pathway. This in turn led to higher predicted concentrations of particle ammonium, subsequent particle ammonium scavenging in clouds, and greater wet ammonium deposition in these models. 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, with improved performance associated with the incorporation of this more detailed wet scavenging approach (Ghahreman *et al.* , 2024). We note that all North American models had negative biases in wet ammonium and wet nitrate deposition, indicating insufficient uptake of gas and particle precursors into cloud and precipitation water, possibly due to the absence of multiphase hydrometeor scavenging in all models participating in the ensemble.

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. The cuticle deposition pathway followed by the soil pathway were shown to dominate the HNO_3 mass flux, usually by more than an order of magnitude. While effective fluxes include potential changes due to other terms such as chemical production, cuticle values also dominate for the effective conductances – deposition velocity is a key term affecting HNO3 concentrations near the surface.

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 CMAQ-M3Dry and CMAQ-STAGE had the most negative biases in NH_3. These models, employ two different bidirectional flux algorithms for NH_3 emissions and deposition. The GEM-MACH (Base) and GEM-MACH (Zhang) models (both of which employed a common bidirectional flux algorithm) had the smallest magnitude NH_3 biases. A detailed analysis of the magnitude and direction of the bidirectional flux models 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 for these differences.

Europe, Critical Load Exceedances

The AQMEII4 ensemble for Europe predicted similar exceedances with respect to acidity and eutrophication in 2009 and 2010, with the 3-member reduced ensemble showing slightly reduced exceedance levels for acidity, and slightly increased exceedance levels for eutrophication, in 2010. The EU percent reduced ensemble acidification exceedance and its range was 4.48% (range 2.37% to 6.85%) in 2009, and 4.32% (range 2.06% to 6.52%) in 2010. EU eutrophication exceedance areas were higher, at 60.2% (range 47.3% to 73.3%) in 2009, and 62.2% (range 51.2% to 74) in 2010 .

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

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, as well as the order of importance of different forms of Sulphur towards total deposition differed from the North American ensemble.

Europe, Causes of Model S Deposition Variability

The common domain average reduced ensemble sulphur dry deposition contributions and their variability followed the same order (SO_2: 56.7%; range 47.0% to 61.4%, Wet S: 37.2%; range 34.3% to 46.4%, dry particulate sulphate: 6.0%; range 1.9 to 9.0%). WRF-Chem (IASS) had the best overall performance relative to observations for SO_2 concentrations, while CMAQ (Hertfordshire) had the best performance for wet S 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 America ensemble, all models had negative biases for wet S deposition. CMAQ (Hertfordshire), with the best overall wet S deposition performance, nevertheless had a 2010 bias of -0.39 eq ha^{-1} yr^{-1}. This may be compared to the monthly observed values which ranged from 1.0 to 1.95 eq ha^{-1} yr^{-1}). As in North America, the manner in which cloud scavenging of particulate sulphate 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.

EU SO_2 deposition pathways were investigated with AQMEII4 diagnostics; the soil and cuticle pathways dominated, though their relative importance varied between reporting models. The stomatal pathway was relatively unimportant (e.g., the cuticle pathway flux was approximately an order of magnitude higher than the stomatal pathway for the three models reporting these data). 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 of 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.

Europe, Causes of Model N Deposition Variability

The relative contributions towards total N deposition and the range in the EU domain were: wet nitrate ion (30.9%; range 22.7% to 31.5%), dry HNO_3 (21.0%; range 11.2 to 25.8%), wet ammonium ion (20.3%; range 15.5% to 26.4%), dry ammonia gas (13.7%; range 9.7% to 22.4%), dry particle nitrate (5.1%; range 4.4% to 5.4%), and dry NO_2 (5.1%; range 2.1% to 6.9%). 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 dry NH_3 deposition was the 4th 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. While monitoring network and satellite NH_3 observations for the time periods simulated were not

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

available for Europe, the latter data source should be used in simulations in subsequent years for evaluation of European model NH_3 performance.

LOTOS-EUROS (TNO) had the best overall performance for wet nitrate deposition, wet ammonium deposition, and near source NO_2 concentrations compared to the other models. However, wet nitrate and ammonium deposition had substantial negative biases for all models, in common with the North American models. The seasonality of wet N deposition was poorly simulated, with most models failing to predict the observed summertime maximum of wet ammonium 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.

Two models for the EU comparison (WRF-Chem (UPM) and CMAQ (Hertfordshire)) reported effective flux diagnostics for all four HNO_3 dry deposition effective flux pathways: these models showed a similar result to North America, with cuticle and soil pathways being approximately an order of magnitude higher than the stomatal and (when present) the lower canopy pathways.

We also note that the importance of the details of the underlying land-use database may be seen in the HNO_3 deposition flux diagnostics (Figures S17 and S18). The effective flux values are with respect to land (vegetated) regions (i.e. the relative importance of stomatal, cuticle, lower canopy and soil pathways is only relevant on land). Here, a >1% water was used as an exclusion criterion to allow a comparison between grid cells that were "mostly land". In these Figures, the regions which do not meet this exclusion criterion are represented as grey areas. Some models clearly have a greater fraction of inland water (possibly at low land use fractions) than others in boreal forest areas, from the relative proportion of grey "no land" grid cells in these Figures (NA: (CMAQ-M3Dry, CMAQ-STAGE, GEM-MACH (Base), GEM-MACH (Zhang), GEM-MACH (Ops)) have more grey regions than (WRF-Chem (IASS), WRF-Chem (UPM), WRF-Chem (UCAR)); EU: CMAQ (Hertfordshire) has more grey regions than the other three models). 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 a companion paper (Hogrefe *et al.,* 2024, ACPD, in preparation).

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:

Multiphase hydrometeor scavenging of gases and aerosols into clouds: Wet deposition was usually the largest or second largest source of total S and N deposition in both ensembles for Europe and North America, however, every model in the two ensembles had negative biases for wet deposition. The portion of critical load exceedances associated with wet deposition has been underestimated, due to this model bias. Improvements in the process representation of cloud processing of S- and N-containing particles and gases is thus a priority for research. Recent papers indicate that multiphase hydrometeor scavenging implementations may reduce this bias.

Incorporation of improved particle deposition velocity algorithms – but only in combination with multiphase wet scavenging. Most of the models and model simulations employed in the AQMEII4 ensemble did not make use of the more recent algorithms for particle dry deposition based on an exhaustive review of available deposition observations in a variety of land use types (Emerson *et al.*, 2020). However, the models which *did* incorporate these algorithms (GEM-MACH (Base), GEM-MACH

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

(Zhang)) had substantially increased positive biases in fine mode particulate matter concentrations. An algorithm known to result in a better fit to observed *deposition velocities* thus resulted in worse *regional model* performance – suggesting that other factors and compensating errors in the modelling platforms may be influencing performance in the model implementations examined here.

Recent work by Ryu and Min (2022) and Ghahreman *et al.* (2024) suggest that the above two issues (positive particle concentration biases and negative wet deposition flux biases) are related, and that a *combination* of particle dry deposition updates and multiphase hydrometeor wet scavenging may be required to generate a net improvement in model performance. Ryu and Min (2022) demonstrated using the WRF-Chem model that when the dry deposition updates are implemented alone, a similar effect on particulate concentrations was seen as for GEM-MACH in the current study – an increase in positive PM2.5 biases relative to observations. They also showed that the implementation of multiphase scavenging alone also decreased WRF-Chem performance, while simultaneously implementing both particle dry deposition velocity updates and multiphase scavenging resulted in a net improvement in performance of that model. Ghahreman *et al.* (2024) added multiphase hydrometeor scavenging to a GEM-MACH implementation incorporating the Emerson *et al.* (2020) particle dry deposition velocity updates, and showed significant improvements in sulphate fluxes relative to earlier model implementations. The need for a combined approach for these process updates is supported by the current work.

Base cation inorganic chemistry and improved base cation emissions inventory development. A striking cause of differences between NA GEM-MACH (Base), GEM-MACH (Zhang), and GEM-MACH (Ops) in comparison to the other models were their positive biases for fine model particle nitrate and ammonium concentrations relative to observations. Unlike the other models in both ensembles, these models employed an inorganic aerosol thermodynamics algorithm which did not include base cation chemistry, and hence some of the excess fine mode particle nitrate and ammonium concentrations and fluxes may be due to the absence of coarse mode formation of base cation nitrates, allowing excessive fine mode particle ammonium and nitrate formation. Updating these models to include base cation chemistry using more recent (Fountoukis and Nenes, 2007) and highly computationally efficient/accurate algorithms now incorporated into GEM-MACH and GEOS-Chem (Miller *et al.,* 2024) should reduce these biases and overall ensemble variability.

NH_3 bidirectional fluxes. The comparison to both surface and satellite observations for the North American ensemble models showed the potential of these algorithms to improve nitrogen deposition performance. However, the comparison also suggested that the compensation point concentrations and stomatal pathway parameterization choices for forested areas in the models compared here may have led to reduced performance in these CMAQ implementations. The balance between NH_3 emissions and deposition for this group of models in forested regions may overestimate the relative importance of deposition fluxes; a reexamination of forest stomatal pathway algorithms and compensation point concentrations for forested regions may improve the performance of this group of models. We also note that ambiguities exist in the methodology for accounting for the influence of bidirectional NH_3 fluxes towards net nitrogen deposition – this is a task for resolution by both the atmospheric modelling and critical load data communities. Furthermore, comparisons of North America regional model simulations of NH_3 with satellite data suggest that oceanic NH_3 emissions, which are included in global chemical transport modelling platforms, should also be considered for regional modelling platforms, due to negative biases observed over oceans in the simulations taking place here.

Land use type harmonization. We note here that the databases used for land use classification may differ between models, and between the land use databases used for air-quality model simulations and those

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

used for critical load data generation; these factors influenced the simulated NH_3 fluxes shown above, and are the subject of a separate AQMEII4 paper analyzing model O_3 deposition performance (Hogrefe *et al.* , 2024, in preparation). Future work centered on harmonizing the land use classification data used for models, and between models and those used in critical load estimates, is therefore recommended.

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.

Competing Interests: At least one of the co-authors is a member of the editorial board of Atmospheric Chemistry and Physics

Author Contributions: PAM: study design and analysis, manuscript writing, GEM-MACH simulations, generation of figures, tables. PC: study analysis support, generation of figures, tables. CH: coordination of modelling team, CMAQ-M3Dry, 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. RB: ENSEMBLE system for submission of model output, coordination of model output library. RB: ENSEMBLE system for submission of model output, coordination of model output library. T. Butler: 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 the 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.

References

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

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

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, 2018. https://doi.org/10.5194/acp-18-8667-2018

AMoN, 2024: https://nadp2.slh.wisc.edu/data/AMoN/, with data downloaded from https://nadp2.slh.wisc.edu/dataLib/AMoN/csv/all-rep.csv, and metadata available at https://nadp.slh.wisc.edu/wp-content/uploads/2022/08/amon_metadata_v2.pdf

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

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, 1976. https://pubs.usgs.gov/publication/pp964

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

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. CMAQ (Hertfordshire)ev., 14, 2867–2897, 202. https://doi.org/10.5194/gmd-14-2867-2021

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

Bash, J. O., Cooter, E. J., Dennis, R. L., Walker, J. T., & Pleim, J. E., Evaluation of a regional air-quality model with bidirectional NH3exchange coupled to an agroecosystem model.Biogeosciences,10(3), 1635–1645, 2013. https://doi.org/10.5194/bg-10-1635-2013

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

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

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

Binkowski F.S., and Roselle, S.J., Models-3 Community Multiscale Air Quality(CMAQ) model aerosol component1. Model description, J. Geophys. Res. 108, No. D6, 4183, 18pp, 2003. https://doi.org/10.1029/2001JD001409

Binkowski, F.S., and Shankar, U., The Regional Particulate Matter Model: 1. Model description and preliminary results, J. Geophys. Res., 100, no D12, 26,191-26,209, 1995. https://doi.org/10.1029/95JD02093

Bobbink, R., Loran, C., Tomassen, H. (2022). 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

Bower, K.N. and T.W. Choularton, A parameterisation of the effective radius of ice free clouds for use in global climate models, Atmospheric Research, 27, 305-339, 1992. https://doi.org/10.1016/0169-8095(92)90038-C

Briggs, G. A.: Plume rise and buoyancy effects, atmospheric sciences and power production, in: DOE/TIC-27601 (DE84005177), edited by: Randerson, D., TN, Technical Information Center, U.S. Dept. of Energy, Oak Ridge, USA,327–366, 1984

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

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

Chapman, E. G., Gustafson Jr., W. I., Easter, R. C., Bamard, J. C., Ghan, S. J., Pekour, M. S., and Fast, J. D., Coupling aerosol-cloud-radiative processes in the WRF-Chem model: Investigating the

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

radiative impact of elevated point sources, Atmos. Chem. Phys., 9, 945–964, 2009. https://doi.org/10.5194/acp-9-945-2009

Chaumerliac, N., Evaluation des Termes de Captation Dynamique dans un Modele Tridimensionel a` Mesoechelle de Lessivage de L'Atmo-sphere, 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., Collett, J.L. Jr., Seasonal ambient ammonia and ammonium concentrations in a pilot IMPROVE NHx monitoring network in the western United States, Atm. Env., 91, 118-126, 2104.

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, 62pp, , 2020a. https://doi.org/10.1029/2019RG000670

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, 2017. https://doi.org/10.1002/2016GL070923

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, 2020b. https://doi.org/10.1029/2020JD032398

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, 2023. https://acp.copernicus.org/articles/23/9911/2023/

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. Mod. Earth Systems, 16, e2023MS003952, 19 pp., 2024. https://doi.org/10.1029/2023MS003952

CLRTAP, 2023: UNECE CLRTAP (2023). 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

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, 1998. https://doi.org/10.1175/1520-0493(1998)126<1373:TOCMGE>2.0.CO;2

Daly, C., Halbleib, M., Smith, J.I., Gibson, W.P., Doggett, M.K., Taylor, G.H., Curtis, J., Pasteris, P.P., Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States, In.t J. Climatol., 2031-2064, 2008. https://rmets.onlinelibrary.wiley.com/doi/10.1002/joc.1688

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, Netherlands, 647 pp., 2015. https://link.springer.com/book/10.1007/978-94-017-9508-1

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Dewitz, J., National Land Cover Database (NLCD) 2019 Products (ver. 2.0, June 2021). U.S. Geological Survey. 2021 https://doi.org/10.5066/P9KZCM54

Dupont, J., Clair, T. A., Gagnon, C., Jeffries, D. S., Kahl, J. S., Nelson, S. J., & Peckenham, J. M., Estimation of Critical Loads of Acidity for Lakes in Northeastern United States and Eastern Canada. *Environmental Monitoring and Assessment*, *109*(1–3), 275–292, 2005. https://doi.org/10.1007/s10661-005-6286-x

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., Zaveri, R.A., 2004. MIRAGE: Model description and evaluation of aerosols and trace gases. Journal of Geophysical Research, 109, 46 pp., 2004. https://doi.org/10.1029/2004JD004571

EEA 2000: CORINE Land Cover 2000, https://trac.osgeo.org/geonetwork/raw-attachment/ticket/650/GeoNetwork-chrome-simple.pdf, Last accessed December 26, 2023.

EEA: CLC2006 technical guidelines. EEA Technical report 17/2007. ISBN 978-92-9167-968-3, https://www.eea.europa.eu/ds_resolveuid/6ee7e1406e694f6adacf6cd349aff89a , 2007.

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, Proc. Nat. Acad. Sci., 117, 26076–26082, 2020. www.pnas.org/cgi/doi/10.1073/pnas.2014761117

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

Fahey, K.M., Carlton, A.G., Pye, H.O.T., Baeka, 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. CMAQ (Hertfordshire)ev., 10, 1587–1605, 2017. https://doi.org/10.5194/gmd-10-1587-2017

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, Atm. Env., 95, 277-287, 2014. https://doi.org/10.1016/j.atmosenv.2014.06.039

Fast, J. D.: Coupling aerosol-cloud-radiative processes in the WRF-Chem model: Investigating the radiativeimpact of elevated point sources, Atmos. Chem. Phys., 9, 945–964, 2009. https://dor.org/10.5194/acp-9-945-2009

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

Fountoukis,C. and Nenes A., ISORROPIAII: 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, 2007. https://doi.org/10.5194/acp-7-4639-2007

Galmarini, S., Hogrefe, C., Brunner, D. Makar, P., Baklanov, A., Preface, Atm. Env., 115, 340-344, 2015. https://doi.org/10.1016/j.atmosenv.2015.06.009

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, 2017. https://doi.org/10.5194/acp-17-1543-2017

Galmarini, S., Makar, P., Clifton, O.E., Hogrefe, C., Bash, J.O., Bellasio, R., Bianconi, R., Beiser, J., Butler, T., Ducker, J., Flemming, J., Hodzic, A., Holmes, C.D., Kioutsioukis, I., Kranenberg, R., Lupascu, A., Perez-Cmanyo, 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

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

as an integral part of regional-scale air quality models, Atmos. Chem. Phys., 21, 15663–15697, 2021. https://doi.org/10.5194/acp-21-15663-2021

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

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(6): 87, 2019. https://doi.org/10.3390/d11060087

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, 1989. https://doi.org/10.1029/JD094iD10p12925

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

Ghahreman, R., Gong, W., Makar, P.A., Lupu, A., Cole, A., Banwait, K., Lee, C., Akingunola, A., Modelling below-cloud scavenging of size resolved particles in GEM-MACHv3.1, Geo. CMAQ (Hertfordshire)ev., 17, 685-707, 2024. https://gmd.copernicus.org/articles/17/685/2024/ .

Gibson, J. J., Birks, J. S., Jeffries, D. S., Kumar, S., Scott, K. A., Aherne, J., & Shaw, P. D., Site-specific estimates of water yield applied in regional acid sensitivity surveys across western Canada. *Journal of Limnology*, *69*(1s), 67, 2010. https://doi.org/10.4081/jlimnol.2010.s1.67

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

Ginoux, P., Chin, M., Tegen, I., Prospero, J. M., Holben, B., Dubovik, O., & Lin, S.-J., Sources and distributions of dust aerosols simulated with the GOCART model. Journal of Geophysical Research: Atmospheres, 106(D17), 20255–20273. https://doi.org/10.1029/2000JD000053. 2001.

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, 2014. https://doi.org/10.1175/MWR-D-13-00255.1

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, 2003. https://doi.org/10.1029/2001JD002002

Gong, S. L., Barrie, L. A., & Blanchet, J.-P., Modeling sea-salt aerosols in the atmosphere: 1. Model development. Journal of Geophysical Research: Atmospheres, 102(D3), 3805–3818, 1997. https://doi.org/10.1029/96JD02953

Grell GA, SE Peckham, R Schmitz, and SA McKeen, G Frost, WC Skamarock, and B Eder, Fully coupled 'online' chemistry in the WRF model. Atm. Env., 39, 6957-6976, 2005. https://doi.org/10.1016/j.atmosenv.2005.04.027

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, 2002. https://doi.org/10.1029/2002GL015311

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, 2014. https://doi.org/10.5194/acp-14-5233-2014

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, 2006. https://doi.org/10.5194/acp-6- 3181-2006

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

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, 2015. https://doi.org/10.5194/acp15-6419-2015

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

Henriksen, A., Dillon, P. J., & 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*, 9., 2002. https://doi.org/10.1139/f02-092

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, 2023. https://doi.org/10.5194/acp-23-8119-2023

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, 2020. https://doi.org/10.1007/978-3-030-22055-6_48

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

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

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: PM2.5 and chemical species, Atmos. Chem. Phys., 21, 2725-2743, 2021. https://doi.org/10.5194/acp-21-2725-2021

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 (2018). https://doi.org/10.1038/s41467-018-05634-2

Hruska, J., Köhler, S., Laudon, H., Bishop, K., Comparison of acid/base character of organic acids in boreal zone of Sweden and mountainous regions in the Czech Republic. Water Air Soil Pollut (S4-2), p. 100 (in: *Acid Raing 2000, Proceedings from the 6th International Conference on Acid Deposition, Kenichi Satake, Ed*.), 2001. https://link.springer.com/book/10.1007/978-94-007-0810-5

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

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, 2019. https://doi.org/10.5194/acp-19-3515-2019

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., 2001. http://www.emc.ncep.noaa.gov/officenotes/newernotes/on437.pdf

Jeffries, D. S., Semkin, R. G., Gibson, J. J., & Wong, I., Recently surveyed lakes in northern Manitoba and Saskatchewan, Canada: Characteristics and critical loads of acidity. *Journal of Limnology*, *69*(1s), 45, 2010. https://doi.org/10.4081/jlimnol.2010.s1.45

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

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

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, 1990. https://doi.org/10.1175/1520-0469(1990)047<2784:AODEPM>2.0.CO;2

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

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. Atmospheric Chemistry and Physics, 14(12), 6213– 6239, 2014. https://doi.org/10.5194/acp-14-6213-2014

Lawrence, G. B., T. J. Sullivan, D. A. Burns, S. W. Bailey, B. J. Cosby, M. Dovciak, H. A. Ewing, T. C. McDonnell, R. Minocha, R. Riemann, J. Quant, K. C. Rice, J. Siemion, and K. Weathers, 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, 2015. https://irma.nps.gov/DataStore/Reference/Profile/2223220

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

Lien, L., Raddum, G. G., Fjellheim, A., & Henriksen, A., A critical limit for acid neutralizing capacity in Norwegian surface waters, based on new analyses of fish and invertebrate responses. Science of The Total Environment, 177(1–3), 173–193, 1996. https://doi.org/10.1016/0048-9697(95)04894-4

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

Lydersen, E., Larssen, T., & Fjeld, E., The influence of total organic carbon (TOC) on the relationship between acid neutralizing capacity (ANC) and fish status in Norwegian lakes. *Science of The Total Environment*, *326*(1), 63–69, 2004. https://doi.org/10.1016/j.scitotenv.2003.12.005

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., 2022. https://nadp.slh.wisc.edu/filelib/claddb/DB_Version/Documentation/NCLD_Documentation_v321.pdf

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, 2018. https://doi.org/10.5194/acp-18-9897-2018

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, 2021. https://doi.org/10.5194/acp-21-12291-2021

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

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,

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

R., Tuccella, P., Werhahn, J., Zhang, J., Galmarini, S., Feedbacks between air pollution and weather, part 2: Effects on chemistry, Atm. Env., 115, 499-526, 2015. https://doi.org/10.1016/j.atmosenv.2014.10.021

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, Atm. Env., 115, 442-469, 2015. https://doi.org/10.1016/j.atmosenv.2014.12.003

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

Makar, P.A., Vouchet, V.S., 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, Atm. Env, 37, 2279-2294, 2003. https://doi.org/10.1016/S1352-2310(03)00074-8

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. Atm., 103, 13095-13110, 1998. https://doi.org/10.1029/98JD00978

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. CMAQ (Hertfordshire)ev., 10, 4145–4173, 2017. https://doi.org/10.5194/gmd-10-4145-2017

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

Massad, R. S., Nemitz, E., & Sutton, M. A., Review and parameterization of bi-directional ammonia exchange between vegetationand the atmosphere.Atmospheric Chemistry and Physics,10(21), 10,359–10,386, 2010. https://doi.org/10.5194/acp-10-10359-2010

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

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. Environmental Pollution, 162: 338-344, 2012. https://doi.org/10.1016/j.envpol.2011.11.025.

McDonnell, T.D., 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. Journal of Environmental Management 146 (2014) 407-419, 2014. http://dx.doi.org/10.1016/j.jenvman.2014.07.019

McNulty, S. G., Cohen, E. C., & Myers, J. A. M., Climate change impacts on forest soil critical acid loads and exceedances at a national scale. In: Potter, Kevin M.; Conkling, Barbara L., Eds. Forest Health Monitoring: National Status, Trends, and Analysis 2010. Gen. Tech. Rep. SRS-GTR-176. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station. 95-108., 176, 95–108, 2013.

McNulty, S. G., Cohen, E. C., Moore Myers, J. A., Sullivan, T. J., & Li, H., Estimates of critical acid loads and exceedances for forest soils across the conterminous United States. Environmental Pollution, 149(3), 281–292, 2017. https://doi.org/10.1016/j.envpol.2007.05.025

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

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, 2016. https://doi.org/10.1175/JAS-D-15-0204.1

Miller, E., *Steady-state critical loads and exceedances for terrestrial and aquatic eocsystems in the northeastern United States*. Technical report, National Park Service, Air Resources Division, 2011.

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

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, 2015. https://doi.org/10.1175/JAS-D-14-0065.1

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, 2009. https://doi.org/10.1175/2008MWR2556.1

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, 2006. https://doi.org/10.1007/s10546-005-9030-8

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

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, 2012. https://doi.org/10.5194/acp-12-3289-2012

Nilsson, J., and Grennfelt, P., Critical loads for sulphur and nitrogen, in: Report from a workshop held at Skokloster, Sweden 19–24 March 1988, J. Nilsson, Ed., Miljorapport, Volume 15 of Nordic Council of Ministers-Publications-Nord, 418pp, 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., 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, 2011. https://doi.org/10.1029/2010JD015139

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 NOx and its vertical distribution calculated from three-dimensional cloud-scale chemical transport model simulations, J. Geophys. Res. Atm., 115, D04301, 2010. https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2009JD011880

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

Paulot, F., Malyshev, S., Nguyen, T., Crounse, J. D., Shevliakova, E., & 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. *Atmospheric Chemistry and Physics*, *18*(24), 17963-17978, 2018. https://doi.org/10.5194/acp-18-17963-2018

Paulot, F., Stock, C., John, J.G., Zadeh, N., Horowitz, L.W., Ocean Ammonia Outgassing: Modulation by CO2 and Anthropogenic Nitrogen Deposition, J. Adv. Mod. Earth Sys., 12, e2019MS002026, 2020. https://doi.org/10.1029/2019MS002026

Pleim, J. E., Ran, L., Appel, W.,Shephard, M. W., & Cady-Pereira, K., New bidirectional ammonia flux model in an air quality model coupled with an agricultural model, Journal of Advances in Modeling Earth Systems,11,2934–2957, 2019. https://doi.org/10.1029/2019MS0017282934

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Pleim, J. E., Ran, L., Saylor, R. D., Willison, J., & Binkowski, F. S. (2022). A new aerosol dry deposition model for air quality and climate modeling. Journal of Advances in Modeling Earth Systems, 14, e2022MS003050. https://doi.org/10.1029/2022MS003050

Posch, M., de Smet, P. A., Hettelingh, J.-P., & 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 accessed December 22, 2023.

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

Ran, L., E. Cooter, D. Yang, V. Benson, Y. Yuan, A. Hanna, AND V. Garcia. 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, 2018. https://www.cmascenter.org/fest-c/documentation/1.4/FESTC_v1_4_UserManual.pdf

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

Reinds, G., Posch, M., Aheme, J., Forsius, M., Assessment of Critical Loads of Sulphur and Nitrogen and Their Exceedances for Terrestrial Ecosystems in the Northern Hemisphere. In: de Vries, W., Hettelingh, JP., Posch, M. (eds) Critical Loads and Dynamic Risk Assessments. Environmental Pollution, vol 25. Springer, Dordrecht, 2015. https://doi.org/10.1007/978-94-017-9508-1_15

Russell, A.G., McRae, G.J., and Cass, G.R., The dynamics of nitric acid production and the fate of nitrogen oxides, Atm. Env.,(1967) 19893-903, 1985.

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

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, 2006. https://doi.org/10.5194/acp-6-187-2006

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

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 Pollution, 225:1838, 2014. https://doi.org/10.1007/s11270-013-1838-0

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, 2011. https://doi.org/10.1029/2010JD014527

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, Proc. Nat. Acad. Sci., 113(15), 4086-4091, 2016. https://doi.org/10.1073/pnas.1515241113

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

Slinn, W. G. N., Precipitation Scavenging, in Atmospheric Science and Power Production, CH. 11, edited by: 680 Randerson, D., Tech. Inf. Cent., Off. of Sci. and Techn. Inf., Dep. of Energy, Washington DC, USA, 466–532, 1984.

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

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

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, 2018. https://doi.org/10.1016/j.scitotenv.2018.03.317

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

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. 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, 2018. https://doi.org/10.5194/acp-18-13531-2018

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, 2010.

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 Pollution, 223:5759-5771, 2012. https://doi.org/10.1007/s11270-012-1312-4.

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 RESOURCES RESEARCH, VOL. 41, W01021, 2005. https://doi.org/10.1029/2004WR003414

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

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

Sverdrup, H., and Warfvinge, P., The role of weathering and forestry in determining the acidity of Lakes in Sweden. Water, Air, and Soil Pollution, 52(1), 71–78, 1990. https://doi.org/10.1007/BF00283115

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

Timmermans, R., van Pinxteren, D. , Kranenburg, R., Hendriks, C., Fomba, K.W., Herrmann, H., Schaap, M., Evaluation of modelled LOTOS-EUROS with observational based PM10 source attribution, Atm. Env.: X, 14, 100173, 2022. https://doi.org/10.1016/j.aeaoa.2022.100173

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

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

Van Zanten, M.C., Sauter, F.J., Wichink Kruit, R.J., Van Jaarsveld, J.A., Van Pul, W.A.J., Description of the DEPAC module: dry deposition modelling with DEPAC GCN2010 RIVM Rep., 2010. https://rivm.openrepository.com/handle/10029/256555

VDEC, 2003: Vermont Department of Environmental Conservation (VDEC), (2003, 2004, 2012). TOTAL MAXIMUM DAILY LOADS: Acid Impaired Lakes. Watershed Management Division, 103 South Main Street, Building 10 North, Waterbury, VT 05671-0408.

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

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(D22),4622, 2002. https://doi.org/10.1029/2002JD002184

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

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, 2018. https://doi.org/10.5194/acp-18-10199-2018

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 Man., 72, 969-984, 2002. https://doi.org/10.1080/10962247.2022.2064004

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 , 2002. https://www.epa.gov/sites/default/files/2015-10/documents/vukovich.pdf

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

Wesely, M. L., Parameterization of surface resistances to gaseous dry deposition in regional-scale numerical models. Atmospheric Environment (1967), 23(6), 1293– 1304, 1989. https://doi.org/10.1016/0004-6981(89)90153-4

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

Whitfield, C. J., Aherne, J., Watmough, S. A., Dillon, P. J., & Clair, T. A., Recovery from acidification in Nova Scotia: Temporal trends and critical loads for 20 headwater lakes. *Canadian Journal of Fisheries and Aquatic Sciences*, *63*(7), 1504–1514, 2016. https://doi.org/10.1139/f06-053

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

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

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

Williams, J. R., The EPIC model. In V. P. Singh (Ed.),Computer models in watershed hydrology(Chapter 25, (pp. 909–1000), 1995. Littleton, CO: Water Resources Publications.

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

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.

https://doi.org/10.5194/egusphere-2024-2226
Preprint. Discussion started: 30 July 2024

Xu, L., Pye, H. O. T., He, J., Chen, Y. L., Murphy, B. N., 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, 2018. https://doi.org/10.5194/acp-18-12613-2018

Yarwood, G., Jung, J., Whitten, G., Heo, G., J, M., and M, E.: Updates to the Carbon Bond Mechanism for Version 6 (CB6), in: 9th Annual CMAS Conference, Chapel Hill, NC, 11–13 October 2010, pp. 1–4, 2010. https://cmascenter.org/conference/2010/abstracts/emery_updates_carbon_2010.pdf 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., Easter, R. C., Fast, J. D., and Peters, L. K.: Model for Simulating Aerosol Interactions and Chemistry (MOSAIC), Journal of Geophysical Research, 113, D13 204, 2008. https://doi.org/10.1029/2007JD008782

Zaveri, R.A., and L.K. Peters, A new lumped structure photochemical mechanism for large-scale applications, J. Geophys. Res., 104, 30,387 - 30,415, 1999. https://doi.org/10.1029/1999JD900876

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 CMOS Conference, 2020, https://www.cmascenter.org/conference/2020/slides/ZhangJ_MODIS_LAI_CMAS_2020.pdf, last accessed December 19, 2023.

Zhang, L., Brook, J. R., and Vet, R.: A revised parameterization for gaseous dry deposition in air-quality models, Atmos. Chem. Phys, 3(6), 2067–2082, 2003. https://doi.org/10.5194/acp-3-2067-2003

Zhang, L., Gong, S., Padro, J., & Barrie, L., A size-segregated particle dry deposition scheme for an atmospheric aerosol module. Atmos-pheric Environment, 35(3), 549–560, 2001. https://doi.org/10.1016/S1352-2310(00)00326-5

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, 2001. https://doi.org/10.1016/S1352-2310(00)00326-5

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(3), 537–560, 2002. https://doi.org/10.1016/S1352-2310(01)00447-2, 2002.

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, 2010. https://doi.org/10.1029/2009JD013589