# An Analysis of CMAQ Gas Phase Dry Deposition over North America Through Grid-Scale and Land-Use Specific Diagnostics in the Context of AQMEII4

Christian Hogrefe[1], Jesse O. Bash[1], Jonathan E. Pleim[1], Donna B. Schwede[1,*], Robert C. Gilliam[1], Kristen M. Foley[1], K. Wyat Appel[1], and Rohit Mathur[1]

[1]Center for Environmental Measurement and Modeling, US Environmental Protection Agency, 109 T.W. Alexander Dr., P.O. Box 12055, RTP, NC 27711, USA
[*]retired

*Correspondence to*: Christian Hogrefe (hogrefe.christian@epa.gov)

**Abstract.** The fourth phase of the Air Quality Model Evaluation International Initiative (AQMEII4) is conducting a diagnostic intercomparison and evaluation of deposition simulated by regional-scale air quality models over North America and Europe. In this study, we analyze annual AQMEII4 simulations performed with the Community Multiscale Air Quality Model (CMAQ) version 5.3.1 over North America. These simulations were configured with both the M3Dry and Surface Tiled Aerosol and Gas Exchange (STAGE) dry deposition schemes available in CMAQ. A comparison of observed and modeled concentrations and wet deposition fluxes shows that the AQMEII4 CMAQ simulations perform similarly to other contemporary regional-scale modeling studies. During summer, M3Dry has higher ozone ($O_3$) deposition velocities ($V_d$) and lower mixing ratios than STAGE for much of the eastern U.S. while the reverse is the case over eastern Canada and along the West Coast. In contrast, during winter STAGE has higher $O_3$ $V_d$ and lower mixing ratios than M3Dry over most of the southern half of the modeling domain while the reverse is the case for much of the northern U.S. and southern Canada. Analysis of the diagnostic variables defined for the AQMEII4 project, i.e. grid-scale and land-use (LU) specific effective conductances and deposition fluxes for the major dry deposition pathways, reveals generally higher summertime stomatal and wintertime cuticular grid-scale effective conductances for M3Dry and generally higher soil grid-scale effective conductances (for both vegetated and bare soil) for STAGE in both summer and winter. On a domain-wide basis, the stomatal grid-scale effective conductances account for about half of the total $O_3$ $V_d$ during daytime hours in summer for both schemes. Employing LU-specific diagnostics, results show that daytime $V_d$ varies by a factor of 2 between LU categories. Furthermore, M3Dry vs. STAGE differences are most pronounced for the stomatal and vegetated soil pathway for the forest LU categories, with M3Dry estimating larger effective conductances for the stomatal pathway and STAGE estimating larger effective conductances for the vegetated soil pathway for these LU categories. Annual domain total $O_3$ deposition fluxes differ only slightly between M3Dry (74.4 Tg/year) and STAGE (76.2 Tg/yr), but pathway-specific fluxes to individual LU types can vary more substantially on both annual and seasonal scales which would affect estimates of $O_3$ damages to sensitive vegetation. A comparison of two simulations differing only in their LU classification scheme shows that the differences in LU cause seasonal mean $O_3$ mixing ratio differences on

the order of 1 ppb across large portions of the domain, with the differences generally largest during summer and in areas characterized by the largest differences in the fractional coverages of the forest, planted/cultivated, and grassland LU categories. These differences are generally smaller than the M3Dry vs. STAGE differences outside the summer season but have a similar magnitude during summer. Results indicate that the deposition impacts of LU differences are caused both by differences in the fractional coverages and spatial distributions of different LU categories as well as the characterization of these categories through variables like surface roughness and vegetation fraction in look-up tables used in the land-surface model and deposition schemes. Overall, the analyses and results presented in this study illustrate how the diagnostic grid-scale and LU-specific dry deposition variables adopted for AQMEII4 can provide insights into similarities and differences between the CMAQ M3Dry and STAGE dry deposition schemes that affect simulated pollutant budgets and ecosystem impacts from atmospheric pollution.

## 1 Introduction

Over the past four decades, grid-based chemical transport models have been used to study air pollution on urban to global scales (McRae and Seinfeld, 1983; Chang et al., 1987; Russell et al., 1988; Harley et al., 1993; Hass et al., 1993; Scheffe and Morris, 1993; Kumar et al., 1994; Jacobson et al., 1996; Chang et al., 1997; Kasibhatla and Chameides, 2000; Sistla et al., 2001; Bey et al., 2001; Grell et al., 2005; Byun and Schere, 2006; Gaydos et al., 2007; Mathur et al., 2017). In these models, the removal of gases and aerosols from the atmosphere through wet and dry deposition is one of the key processes of the simulated pollutant budgets. While the representation of gas and aerosol dry deposition in many of these models is derived from the resistance framework introduced in Wesely and Hicks (1977) and Wesely (1989), its specific implementation can differ between models (Hardacre et al., 2015; Galmarini et al., 2021) and its use to represent aerosol dry deposition is an area of active model development (Saylor et al., 2019; Emerson et al., 2020; Pleim et al., 2022; Alapaty et al., 2022, Cheng et al., 2022). Likewise, the calculation of wet deposition fluxes in many models follows similar approaches as those used during initial acid deposition modeling (Chang et al., 1987; Irving and Smith, 1991; Hass et al., 1993) but differences exist in how models represent microphysics, precipitation, and aerosols.

Intercomparisons of air quality models can play an important role in assessing how such different process representations can impact simulated pollutant concentrations, model performance, and the use of models for planning applications. One such model intercomparison activity, the Air Quality Model Evaluation International Initiative (AQMEII), was launched in 2009 (Rao et al., 2011) and since then has organized several activities focused on evaluating regional-scale air quality models used for research and regulatory applications over North America and Europe. As discussed in Galmarini et al. (2021), its fourth phase (AQMEII4) employs both grid and box modeling techniques for a diagnostic intercomparison and evaluation of simulated deposition with a specific focus on dry deposition of gaseous species. The grid model component of AQMEII4 is based on eight groups performing annual simulations for two years over North America and Europe and collecting detailed

dry deposition diagnostics for a range of trace gases. The Community Multiscale Air Quality (CMAQ) model (Byun and Schere, 2006) has been a part of all AQMEII activities performed to date and its performance in previous activities has been documented in both detailed comparisons of CMAQ simulations to observations (Appel et al., 2012; Hogrefe et al., 2015; Hogrefe et al., 2018) and comparisons to other modeling systems participating in AQMEII (Solazzo et al., 2012a,b; Im et al., 2015a,b; Solazzo et al., 2017).

This present study is conducted in the context of AQMEII4 and has three main objectives. The first objective is to evaluate the CMAQ simulations contributed to AQMEII4 by comparing simulated pollutant fields and wet deposition fluxes to observations while leveraging results from a recent extensive CMAQ evaluation study (Appel et al., 2021) that used the same model version as the AQMEII4 simulations but differed in terms of several input fields and configuration options. The second objective is to use the AQMEII4 diagnostics introduced in Galmarini et al. (2021) and Clifton et al. (2023) to diagnostically compare the two dry deposition schemes implemented in CMAQ and used in the AQMEII4 simulations as discussed in Section 2, i.e. the M3Dry (Pleim et al., 1984; Pleim and Ran, 2011) and Surface Tiled Aerosol and Gaseous Exchange (STAGE) (Appel et al., 2021, Galmarini et al., 2021; Walker et al., 2023) schemes. The third objective is to quantify the impacts of differences in the representation of land use (LU) in the meteorological and air quality model on concentrations and fluxes and compare them to the impacts of different dry deposition schemes.

Section 2 provides an overview of the modeling system utilized in this study, the sensitivity simulations performed and analyzed, details on the configuration of the M3Dry and STAGE dry deposition schemes for these AQMEII4 simulations, and describes the observational datasets used for model evaluation. Section 3.1 presents results from the model performance evaluation, Section 3.2 presents results of a diagnostic gas phase dry deposition comparison between M3Dry and STAGE on both the grid scale and for specific LU types, and Section 3.3 analyzes the sensitivity of estimated dry deposition to the underlying LU classification scheme. Results are summarized and discussed in Section 4.

## 2 Description of model simulations and observational database

### 2.1 Base case model configuration

The 2010 and 2016 base case CMAQ simulations analyzed in this study closely follow the configuration of the "CMAQ531_WRF411_M3Dry_BiDi" and "CMAQ531_WRF411_STAGE_BiDi" 2016 CMAQv5.3.1 simulations analyzed in Appel et al. (2021). Key aspects of that configuration as well as deviations in the current study are summarized below.

#### 2.1.1. Meteorological modeling

As in the Appel et al. (2021) study, meteorological fields were generated with the Weather Research and Forecast (WRF) model version 4.1.1. WRFv4.1.1 configuration options in common with Appel et al. (2021) include the Rapid Radiation Transfer Model Global (RRTMG) for long- and short-wave radiation (Iacono et al., 2008), the Morrison microphysics scheme

(Morrison et al., 2005), and the Kain-Fritsch (KF) cumulus parametrization scheme (Kain, 2004). Furthermore, both Appel et al. (2021) and this study used the Pleim-Xiu land-surface model (PX-LSM; (Pleim and Xiu, 1995; Xiu and Pleim, 2001; Pleim and Gilliam, 2009) and Asymmetric Convective Mixing 2 planetary boundary layer (PBL) model (Pleim, 2007a,b) with the Pleim surface layer scheme (Pleim, 2006). WRF data assimilation followed Appel et al. (2021) and is described in greater detail in Gilliam et al. (2021), and soil temperature and moisture nudging followed Pleim and Gilliam (2009) and Pleim and Xiu (2003). In contrast to Appel et al. (2021), the WRF simulations used in this study obtained sea surface temperature from the North America Model (NAM) reanalysis data set (Mesinger et al., 2006) instead of the Group for High Resolution Sea Surface Temperature (GHRSST), did not include lightning assimilation (Heath et al., 2016), and classified LU with the 20-category Moderate Resolution Imaging Spectroradiometer (MODIS) satellite derived LU classification scheme instead of the 40-category National Land Cover Dataset (NLCD) (Dewitz and U.S. Geological Survey, 2021; Yang et al., 2018) LU data set. In a further contrast to Appel et al. (2021), the WRF PX LSM in this study was configured to obtain Leaf Area Index (LAI) and areal fraction covered by vegetation (VEGF) from the PX LSM MODIS LU-scheme lookup table values rather than directly ingesting MODIS satellite-derived inputs interpolated from monthly data as described in Ran et al. (2016). Finally, it should be noted that after performing the two 2010 CMAQ simulations analyzed in this study, it was discovered that the WRF fields for several time periods in September, October, and November 2010 were affected by inconsistencies in the WRF input file preparation that affected simulated precipitation. These time periods were excluded from the model performance evaluation.

#### 2.1.2 Emissions

The anthropogenic emissions for 2010 and 2016 were harmonized for all AQMEII4 modeling groups and are described in Galmarini et al. (2021). For 2016, these AQMEII4 emissions were based on an earlier version of emission inventories compared to the 2016 emissions described in Appel et al. (2021). For lightning NO emissions, the CMAQ simulations performed for this study used the GEIA monthly climatological data (Price et al., 1997) as described in Galmarini et al. (2021) for consistency with other AQMEII4 modeling groups while Appel et al. (2021) used lightning NO estimated from hourly year-specific lightning flash data from the National Lightning Detection Network (NLDN). Both this study and Appel et al. (2021) estimated biogenic VOC and soil NO emissions with the CMAQ inline Biogenic Emission Inventory System (BEIS) option with the same underlying LU dataset and emission factors.

#### 2.1.3 Boundary conditions

Lateral chemical boundary conditions for both 2010 and 2016 for all AQMEII4 model simulations, including the CMAQ simulations analyzed in this study, were obtained from the Copernicus Atmospheric Monitoring Service (CAMS) EAC4 reanalysis product (Inness et al., 2019) as described in Galmarini et al. (2021). This differs from the 2016 CMAQ simulations

analyzed in Appel et al. (2021) which used hemispheric CMAQ (Mathur et al., 2017) simulations to generate boundary conditions for the regional-scale modeling domain over North America.

#### 2.1.4 Air quality modeling

The base version of CMAQ used in this study is 5.3.1 (U.S. Environmental Protection Agency, 2019) and matches that used for the "CMAQ531_WRF411_M3Dry_BiDi" and "CMAQ531_WRF411_STAGE_BiDi" simulations in Appel et al. (2021). All CMAQ simulations were performed on the same 12 km modeling domain with 35 vertical layers covering the conterminous U.S, southern Canada, and northern Mexico that was used in Appel et al. (2021). Science configuration options include the cb6r3 chemical mechanism (Luecken et al., 2019), the aero7 aerosol module (Pye et al., 2017, 2019; Qin et al., 2021; Appel et al., 2021), and the bi-directional treatment of $NH_3$ fluxes, all matching the configuration options used in Appel et al. (2021). As in Appel et al. (2021), the fertilizer and soil $NH_3$ information required for the bi-directional treatment of $NH_3$ fluxes were generated by the Environmental Policy Integrated Climate (EPIC; Williams, 1995) model through the Java-based Fertilization Emission Scenario Tool for CMAQ (FEST-C; Ran et al., 2019), though it should be noted that the 2010 EPIC fields used in this study suffered from an EPIC configuration error that resulted in an unrealistically large allocation of annual fertilizer application to the beginning of the year. Dry deposition was simulated with both science options available in CMAQv5.3.1, i.e. the M3Dry and STAGE schemes. The application of both schemes to this study and modifications for the STAGE dry deposition option relative to the CMAQv5.3.1 base version are described in Section 2.3.

### 2.2 CMAQ Sensitivity Simulations

Besides the 2010 and 2016 CMAQ M3Dry and STAGE base case simulations described in Section 2.1, additional CMAQ sensitivity simulations were performed for 2016 to quantify the impacts of some of the differences relative to the Appel et al. (2021) CMAQ configuration and to gain further diagnostic understanding on the choice of model inputs on modeled deposition. Table 1 provides a listing of all base case and sensitivity simulations, their acronyms used in the following analyses and discussions, and the input data sets and/or configuration options differentiating them. As discussed further in Section 2.3, STAGE_REF_2016 is designed to quantify the impact of modifying the CMAQ STAGE code for AQMEII4 relative to the unmodified STAGE code in CMAQv5.3.1. M3DRY_NLCD40_2016 is designed to study the impact of using a different LU classification scheme in both WRF and CMAQ on simulated concentrations, deposition fields, and deposition diagnostics. M3DRY_HCMAQ_2016 can be used to assess the impact of using chemical boundary conditions from CAMS compared to using boundary conditions from H-CMAQ as in Appel et al. (2021) while M3DRY_APPEL_EMIS_2016 can be used to quantify the impacts of the different anthropogenic and wildland fire emissions used in this study vs. Appel et al. (2021). Finally, the set of M3DRY_LTGNO_BASE_2016 and M3DRY_LTGNO_NLDN_2016 simulations can be used to study the impacts of different lightning NO emission representations (GEIA vs. NLDN) on simulated concentration and deposition fields. It should be noted that these two simulations used the 2021 WRFv4.1.1 fields from Appel et al. (2021) rather than those

used for the base case simulations described in Section 2.1.1. Because these sensitivity simulations were performed for 2016, the focus of the analysis in Section 3 will be on that year.

| | Year | WRF Lightning Assimilation | WRF Land Use | CMAQ Land Use | CMAQ Lightning NO Emissions | CMAQ Boundary Conditions | Anthropogenic Emissions |
|---|---|---|---|---|---|---|---|
| **Base Case Simulations** | | | | | | | |
| M3DRY_2010 | 2010 | N | MODIS | MODIS | GEIA | CAMS | Galmarini et al (2021) |
| STAGE_2010 | 2010 | N | MODIS | AQMEII4 | GEIA | CAMS | Galmarini et al (2021) |
| M3DRY_2016 | 2016 | N | MODIS | MODIS | GEIA | CAMS | Galmarini et al (2021) |
| STAGE_2016 | 2016 | N | MODIS | AQMEII4 | GEIA | CAMS | Galmarini et al (2021) |
| **2016 Sensitivity Simulations** | | | | | | | |
| STAGE_REF_2016 | 2016 | N | MODIS | MODIS | GEIA | CAMS | Galmarini et al (2021) |
| M3DRY_NLCD40_2016 | 2016 | N | NLCD40 | NLCD40 | GEIA | CAMS | Galmarini et al (2021) |
| M3DRY_HCMAQ_2016 | 2016 | N | NLCD40 | NLCD40 | GEIA | H-CMAQ | Galmarini et al (2021) |
| M3DRY_APPEL_EMIS_2016 | 2016 | N | MODIS | MODIS | GEIA | CAMS | Appel et al. (2021) |
| M3DRY_LTNGNO_BASE_2016 | 2016 | Y | NLCD40 | NLCD40 | GEIA | CAMS | Galmarini et al (2021) |
| M3DRY_LTNGNO_NLDN_2016 | 2016 | Y | NLCD40 | NLCD40 | NLDN | CAMS | Galmarini et al (2021) |

**Table 1.** Configurations of the AQMEII4 CMAQ simulations performed for this study.

## 2.3 Application of M3Dry and STAGE for AQMEII4

A key component of the AQMEII4 activity is to compute and analyze dry deposition pathways and component resistances on both a grid-scale and LU-specific basis (Galmarini et al., 2021). This section describes how this diagnostic information is generated from the CMAQ simulations contributed to the AQMEII4 activity and analyzed in this study. A schematic representation of both M3Dry and STAGE and equations for the computations of the AQMEII4 dry deposition diagnostic variables can be found in Appendix B of Galmarini et al. (2021). As described in Appel et al. (2021), both M3Dry and STAGE originate from earlier versions of the M3Dry scheme which has a long history in CMAQ and other chemical transport models (Pleim et al., 1984). However, both algorithms follow different approaches in terms of their consideration of sub-grid scale variations in LU and the calculation of some component resistances.

### 2.3.1 M3Dry

The M3Dry dry deposition calculations performed in CMAQ are designed to maintain maximum consistency with the flux calculations performed in the WRF PX LSM. Specifically, the M3Dry calculations are performed on a grid scale basis. Sub-grid scale variations in LU are accounted for by computing relevant grid-scale parameters like roughness length ($z_0$), vegetation fraction (VEGF), and leaf area index (LAI) as LU-weighted averages from LU-specific lookup table values. Grid scale aerodynamic resistance ($R_a$), stomatal resistance ($R_s$), friction velocity ($u_*$), VEGF, and LAI computed in the WRF PX LSM are directly used in the CMAQ M3Dry deposition calculations. To compute the LU-specific and dry deposition diagnostic variables required for AQMEII4 (Galmarini et al., 2021), a post-processing tool was developed to estimate these variables by performing M3Dry calculations separately for each LU category encountered in a grid cell. While these LU-specific post-processor calculations used the same formulations and parameter values as the grid-scale deposition calculations performed in CMAQ M3Dry, the fact that WRF PX LSM uses parameter values weighted by LU fraction for the calculation of $R_a$, $R_s$, and other relevant variables means that the LU-weighted averages of the LU-specific post-processor estimates for deposition velocity ($V_d$) and effective conductance may slightly deviate from the grid-scale CMAQ M3Dry calculations and should therefore be viewed as an approximation. However, the LU-specific dry deposition fluxes computed by the post-processor were normalized by the grid cell values so that the LU-weighted flux sums equal the total grid cell fluxes. The CMAQ M3Dry calculations and post-processor estimates of LU-specific and aggregated diagnostic variables were performed using the native 20 category MODIS LU scheme (or 40 category NLCD LU scheme for the M3DRY_NLCD40_2016 sensitivity simulation) that was also used in the WRF simulations. Aggregation to the 16 category AQMEII4 LU scheme used in our analysis (Galmarini et al., 2021) was performed through mapping and LU-weighted averaging of equivalent categories (Table S1).

### 2.3.2 STAGE

The STAGE dry deposition option was first introduced in CMAQv5.3 (Appel et al., 2021). It unifies bi-directional and uni-directional deposition schemes following the resistance model frameworks of Massad et al. (2010) and Nemitz et al. (2001).

In contrast to M3Dry, STAGE computes individual resistances, $V_d$, and deposition fluxes for each LU category present in a grid cell and then aggregates these calculations to the grid scale value for use by the CMAQ surface exchange module. Therefore, some of the deposition diagnostics required for AQMEII4 (LU-specific and grid scale $V_d$ and fluxes) were readily available from the standard version of STAGE in CMAQv5.3.1. Code modifications were made to output the desired component resistances and conductances that were computed but not output in the standard version of STAGE. Moreover, the STAGE code used in this study was also modified to perform all deposition calculations directly on the 16 AQMEII4 LU categories (Galmarini et al., 2021) rather than the 20 MODIS categories used in the WRF PX LSM calculations. This was accomplished by applying the same mapping used in the M3Dry post-processing and shown in Table S1 and also defining LU-specific lookup table values for required parameters like $z_0$, VEGF, and LAI for each of the 16 AQMEII4 categories in the modified STAGE code. As a result of deriving grid-scale deposition-related variables from LU-specific calculations and using a more aggregated LU classification scheme with a separate set of lookup table values, some of these grid-scale variables such as $R_a$, $R_s$, $u_*$, VEGF, and LAI may differ from the corresponding values used in the WRF PX LSM flux calculations, creating a potential inconsistency in the treatment of surface-air exchange processes simulated in WRF PX LSM and CMAQ STAGE. It should be noted that in the most recent version of CMAQ released in October 2022 (U.S. Environmental Protection Agency, 2022), STAGE was updated to normalize LU-specific calculations for $R_a$, $R_s$, and $u_*$ such that their aggregated grid-scale values match the grid scale values obtained from the LSM of the driving meteorological model.

### 2.4 Observational data

For the performance evaluation presented in Section 3.1 and the supplemental material, the base case CMAQ simulations were compared against observations obtained from the U.S. Environmental Protection Agency's Air Quality System (AQS; https://www.epa.gov/aqs) database, the Canadian National Air Pollution Surveillance (NAPS) program (https://donnees-data.ec.gc.ca/data/air/monitor/national-air-pollution-surveillance-naps-program/), and the National Atmospheric Deposition Program's National Trend Network (NADP NTN; https://nadp.slh.wisc.edu/networks/national-trends-network/). Specifically, model evaluation was performed for hourly nitrogen oxides ($NO_x$) and sulfur dioxide ($SO_2$), daily maximum 8-hr average $O_3$ (MDA8 $O_3$), and daily $PM_{2.5}$ sulfate ($SO_4^{2-}$), nitrate ($NO_3^-$), organic carbon (OC), elemental carbon (EC), and total $PM_{2.5}$ mass from AQS, MDA8 $O_3$ and hourly $NO_x$, $SO_2$, and total $PM_{2.5}$ mass from NAPS, and weekly integrated precipitation and wet deposition of $SO_4^{2-}$, $NO_3^-$, and ammonium ($NH_4^+$) from NADP NTN. The number of 2016 (2010) AQS monitors with available observations was 425 (374) for $NO_x$, 464 (446) for $SO_2$, 1323 (1278) for MDA8 $O_3$, 1926 (2006) for $PM_{2.5}$ mass, 318 (398) for $PM_{2.5}$ $SO_4^{2-}$, 312 (393) for $PM_{2.5}$ $NO_3^-$, 297 (177) for $PM_{2.5}$ OC, and 297 (177) for $PM_{2.5}$ EC. The number of 2016 (2010) NAPS monitors with available observations was 116 (100) for $NO_x$, 98 (95) for $SO_2$, 178 (168) for MDA8 $O_3$, and 172 (159) for $PM_{2.5}$ mass. In addition, there were 259 NADP NTN monitors that measured precipitation and $SO_4^{2-}$, $NO_3^-$, and $NH_4^+$ wet deposition in 2016 and 237 monitors that did so in 2010. All model values were matched in time and space against the available 2010 and 2016 observations using the Atmospheric Model Evaluation Tool (AMET; Appel et al., 2011) version 1.4.

## 3 Results

### 3.1 Model performance evaluation

Comparisons of modeled and observed MDA8 $O_3$, $SO_2$, $NO_x$, $PM_{2.5}$, $SO_4^{2-}$, $NO_3^-$, OC, and EC concentrations at AQS monitors, MDA8 $O_3$, $SO_2$, $NO_x$, and $PM_{2.5}$ at NAPS monitors, and precipitation and wet deposition of $SO_4^{2-}$, $NO_3^-$, and $NH_4$ at NADP NTN monitors are presented in Figures 1 - 2 and Tables 2 – 5. This section summarizes the performance of the M3DRY_2016, STAGE_2016, M3DRY_2010, and STAGE_2010 base case simulations. To provide context for these results, a comparison to the model performance of 2016 CMAQv5.3.1 simulations from a recent comprehensive evaluation study (Appel et al., 2021) and the differences in model configurations driving differences in model performance can be found in the supplemental material (Figures S1 – S7 and Tables S2 – S3). Even though the diagnostic analyses presented in subsequent sections of this manuscript are focused on 2016, this section documents model performance results for both 2010 and 2016 because results from the 2010 simulations will be included in forthcoming AQMEII4 analyses.

Table 2 shows key model performance metrics for the M3DRY_2016 and STAGE_2016 base case simulations for the gas phase and aerosol species listed above. These performance metrics are the observed and modeled mean values and standard deviations, the normalized and absolute mean bias (NMB and MB, respectively), the root mean square error (RMSE), and the correlation coefficient. The metrics shown in this table were computed across all stations and available observation – model pairs for the entire year. Corresponding time series of monthly mean observed and modeled values are shown in the left column of Figure 1.

For MDA8 $O_3$, both simulations show positive biases of about 3.5 ppb (AQS sites) and 6 ppb (NAPS sites), RMSE values of about 8 ppb (AQS sites) and 9 ppb (NAPS sites), and correlation coefficients of about 0.8 (AQS sites) and 0.74 (NAPS sites). The differences between performance statistics for the M3Dry and STAGE simulations are small but that is partially due to the seasonally and spatially varying nature of differences between these schemes that will be discussed in Section 3.2. For $NO_x$, the simulations show substantial underpredictions that are more pronounced at the AQS than NAPS sites and correlation coefficients of about 0.5. The time series indicate that the simulations deviate most substantially from observations during winter, and differences between M3Dry and STAGE are again much smaller than differences between model values and observations. For $SO_2$, both the M3Dry and STAGE simulations show negative biases at AQS sites and positive biases at NAPS sites, pointing to potential differences in monitor locations relative to sources between the two networks. Correlations for $SO_2$ are the lowest for any of the species considered in this analysis. For $PM_{2.5}$ mass, the simulations show a positive bias that is more pronounced at NAPS than AQS monitors and correlation coefficients of about 0.4. M3DRY_2016 shows slightly lower concentrations, MB and RMSE than STAGE_2016 for both networks. For $PM_{2.5}$ species, the simulations are biased high for all species except $NO_3^-$ with the highest normalized mean bias for OC. A comparison of the spatial patterns of MDA8 $O_3$ and $PM_{2.5}$ biases in Figure S3 shows that the positive bias for MDA8 $O_3$ is widespread throughout the modeling domain while

the positive bias for $PM_{2.5}$ is most pronounced in the eastern portion of the modeling domain and some portions of the West Coast with smaller overestimations and some underestimations in other regions.

| Species | Simulation | Observed Mean | Model Mean | Observed σ | Model σ | NMB | MB | RMSE | Correlation |
|---|---|---|---|---|---|---|---|---|---|
| AQS MDA8 $O_3$ (ppb) | M3DRY_BASE_2016 | 41.93 | 45.42 | 12.00 | 9.78 | 8.34 | 3.50 | 7.91 | 0.81 |
| | STAGE_BASE_2016 | | 45.24 | | 10.10 | 7.89 | 3.31 | 7.91 | 0.80 |
| NAPS MDA8 $O_3$ (ppb) | M3DRY_BASE_2016 | 32.08 | 37.95 | 10.40 | 8.86 | 18.30 | 5.87 | 9.24 | 0.74 |
| | STAGE_BASE_2016 | | 37.94 | | 9.04 | 18.30 | 5.86 | 9.25 | 0.74 |
| AQS $NO_X$ (ppb) | M3DRY_BASE_2016 | 13.36 | 8.41 | 21.20 | 10.80 | -37.10 | -4.95 | 18.60 | 0.53 |
| | STAGE_BASE_2016 | | 8.42 | | 10.80 | -37.00 | -4.94 | 18.60 | 0.53 |
| NAPS $NO_X$ (ppb) | M3DRY_BASE_2016 | 10.40 | 8.62 | 14.90 | 11.10 | -17.10 | -1.78 | 13.60 | 0.50 |
| | STAGE_BASE_2016 | | 8.62 | | 11.10 | -17.10 | -1.78 | 13.60 | 0.50 |
| AQS $SO_2$ (ppb) | M3DRY_BASE_2016 | 0.89 | 0.79 | 2.82 | 1.10 | -11.20 | -0.10 | 2.86 | 0.16 |
| | STAGE_BASE_2016 | | 0.80 | | 1.09 | -10.10 | -0.09 | 2.86 | 0.16 |
| NAPS $SO_2$ (ppb) | M3DRY_BASE_2016 | 0.73 | 1.02 | 4.43 | 3.14 | 39.70 | 0.29 | 5.08 | 0.13 |
| | STAGE_BASE_2016 | | 1.07 | | 3.24 | 46.80 | 0.34 | 5.14 | 0.13 |
| AQS Total $PM_{2.5}$ ($\mu g/m^3$) | M3DRY_BASE_2016 | 7.57 | 7.90 | 5.16 | 6.87 | 4.39 | 0.33 | 6.64 | 0.42 |
| | STAGE_BASE_2016 | | 8.47 | | 7.16 | 11.90 | 0.90 | 6.86 | 0.43 |
| NAPS Total $PM_{2.5}$ ($\mu g/m^3$) | M3DRY_BASE_2016 | 5.80 | 8.62 | 5.50 | 10.50 | 48.60 | 2.82 | 10.40 | 0.35 |
| | STAGE_BASE_2016 | | 9.03 | | 10.90 | 55.60 | 3.23 | 10.80 | 0.36 |
| AQS $SO_4^{2-}$ ($\mu g/m^3$) | M3DRY_BASE_2016 | 0.78 | 0.79 | 0.72 | 0.61 | 1.07 | 0.01 | 0.55 | 0.67 |
| | STAGE_BASE_2016 | | 0.83 | | 0.64 | 6.60 | 0.05 | 0.56 | 0.67 |
| AQS $NO_3^-$ ($\mu g/m^3$) | M3DRY_BASE_2016 | 0.59 | 0.46 | 1.24 | 0.97 | -21.30 | -0.13 | 1.00 | 0.62 |
| | STAGE_BASE_2016 | | 0.53 | | 1.07 | -8.89 | -0.05 | 1.00 | 0.63 |
| AQS OC ($\mu g/m^3$) | M3DRY_BASE_2016 | 1.31 | 1.83 | 1.56 | 1.68 | 39.90 | 0.52 | 1.71 | 0.50 |
| | STAGE_BASE_2016 | | 1.98 | | 1.77 | 51.90 | 0.68 | 1.80 | 0.51 |
| AQS EC ($\mu g/m^3$) | M3DRY_BASE_2016 | 0.31 | 0.39 | 0.40 | 0.46 | 23.50 | 0.07 | 0.40 | 0.61 |
| | STAGE_BASE_2016 | | 0.39 | | 0.47 | 25.20 | 0.08 | 0.40 | 0.61 |

**Table 2.** Model performance statistics for all daily maximum 8-hr average $O_3$ (MDA8 $O_3$), hourly $NO_x$ and $SO_2$, and 24-hr average total and speciated ($SO_4^{2-}$, $NO_3^-$, organic carbon (OC) and elemental carbon (EC)) $PM_{2.5}$ mass samples collected at AQS and NAPS monitors in 2016. The standard deviation over all samples is denoted as σ while NMB, MB, and RMSE represent the percentage normalized mean bias, mean bias, and root mean square error computed over all samples.

| Species | Simulation | Observed Mean | Model Mean | Observed σ | Model σ | NMB | MB | RMSE | Correlation | Observed 2016 – 2010 | Model 2016 – 2010 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| AQS MDA8 $O_3$ (ppb) | M3DRY_BASE_2010 | 43.87 | 48.55 | 13.50 | 11.50 | 10.70 | 4.68 | 9.41 | 0.80 | -1.95 | -3.13 |
| | STAGE_BASE_2010 | | 48.40 | | 11.80 | 10.30 | 4.53 | 9.43 | 0.79 | -1.95 | -3.16 |
| NAPS MDA8 $O_3$ (ppb) | M3DRY_BASE_2010 | 34.08 | 38.63 | 10.90 | 9.96 | 13.30 | 4.55 | 8.57 | 0.76 | 7.84 | 6.79 |
| | STAGE_BASE_2010 | | 38.59 | | 10.10 | 13.20 | 4.51 | 8.59 | 0.76 | 7.84 | 6.64 |
| AQS $NO_X$ (ppb) | M3DRY_BASE_2010 | 13.99 | 11.15 | 23.40 | 15.20 | -20.30 | -2.85 | 20.9 | 0.49 | -0.63 | -2.74 |
| | STAGE_BASE_2010 | | 11.15 | | 15.20 | -20.30 | -2.84 | 20.9 | 0.49 | -0.63 | -2.73 |
| NAPS $NO_X$ (ppb) | M3DRY_BASE_2010 | 13.36 | 13.26 | 20.10 | 23.70 | -0.74 | -0.10 | 21.7 | 0.52 | 0.01 | -4.85 |
| | STAGE_BASE_2010 | | 13.23 | | 23.60 | -0.91 | -0.12 | 21.7 | 0.52 | 0.01 | -4.81 |
| AQS $SO_2$ (ppb) | M3DRY_BASE_2010 | 2.05 | 2.01 | 5.23 | 2.57 | -2.14 | -0.04 | 5.22 | 0.25 | -1.16 | -1.22 |
| | STAGE_BASE_2010 | | 2.05 | | 2.60 | -0.01 | 0.00 | 5.23 | 0.25 | -1.16 | -1.25 |
| NAPS $SO_2$ (ppb) | M3DRY_BASE_2010 | 1.45 | 1.67 | 6.44 | 2.61 | 14.70 | 0.21 | 6.7 | 0.10 | -0.57 | -0.88 |
| | STAGE_BASE_2010 | | 1.72 | | 2.62 | 18.50 | 0.27 | 6.71 | 0.10 | -0.57 | -0.92 |
| AQS Total $PM_{2.5}$ ($\mu g/m^3$) | M3DRY_BASE_2010 | 9.51 | 10.33 | 6.38 | 7.36 | 8.66 | 0.82 | 5.88 | 0.65 | -1.94 | -2.43 |
| | STAGE_BASE_2010 | | 11.10 | | 7.93 | 16.80 | 1.60 | 6.32 | 0.65 | -1.94 | -2.63 |
| NAPS Total $PM_{2.5}$ ($\mu g/m^3$) | M3DRY_BASE_2010 | 6.81 | 9.21 | 10.20 | 12.30 | 35.30 | 2.40 | 11.9 | 0.47 | 0.76 | -1.31 |
| | STAGE_BASE_2010 | | 9.61 | | 12.70 | 41.20 | 2.80 | 12.3 | 0.47 | 0.76 | -1.14 |
| AQS $SO_4^{2-}$ ($\mu g/m^3$) | M3DRY_BASE_2010 | 1.52 | 1.43 | 1.50 | 1.28 | -6.23 | -0.09 | 0.951 | 0.78 | -0.74 | -0.64 |
| | STAGE_BASE_2010 | | 1.52 | | 1.37 | -0.58 | -0.01 | 0.974 | 0.77 | -0.74 | -0.68 |
| AQS $NO_3^-$ ($\mu g/m^3$) | M3DRY_BASE_2010 | 0.92 | 0.81 | 1.81 | 1.66 | -11.20 | -0.10 | 1.32 | 0.72 | -0.33 | -0.35 |
| | STAGE_BASE_2010 | | 0.96 | | 1.96 | 4.83 | 0.04 | 1.42 | 0.72 | -0.33 | -0.43 |
| AQS OC ($\mu g/m^3$) | M3DRY_BASE_2010 | 1.49 | 2.16 | 1.54 | 1.89 | 44.30 | 0.66 | 1.77 | 0.55 | -0.19 | -0.33 |
| | STAGE_BASE_2010 | | 2.34 | | 2.03 | 56.30 | 0.84 | 1.92 | 0.56 | -0.19 | -0.35 |
| AQS EC ($\mu g/m^3$) | M3DRY_BASE_2010 | 0.42 | 0.60 | 0.52 | 0.66 | 44.20 | 0.19 | 0.537 | 0.66 | -0.11 | -0.22 |
| | STAGE_BASE_2010 | | 0.61 | | 0.67 | 46.20 | 0.19 | 0.543 | 0.66 | -0.11 | -0.22 |

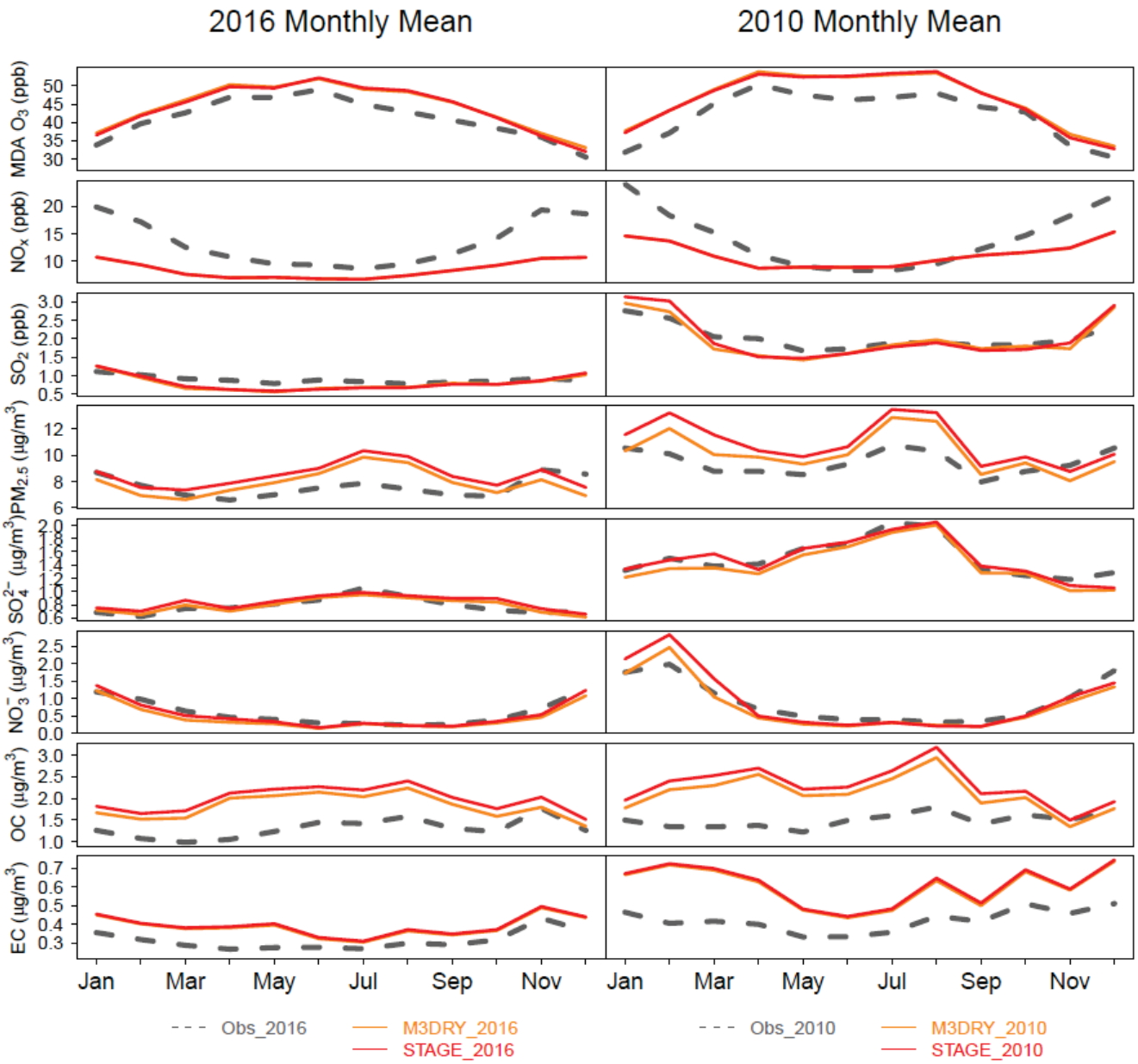
**Table 3.** Model performance statistics for all daily maximum 8-hr average $O_3$ (MDA8 $O_3$), hourly $NO_x$ and $SO_2$, and 24-hr average total and speciated ($SO_4^{2-}$, $NO_3^-$, organic carbon (OC) and elemental carbon (EC)) $PM_{2.5}$ mass samples collected at AQS and NAPS monitors in 2010. The standard deviation over all samples is denoted as σ while NMB, MB, and RMSE represent the percentage normalized mean bias, mean bias, and root mean square error computed over all samples. The last two columns show differences in average observed and modeled values between 2016 and 2010.

| Species | Simulation | Observed Mean | Model Mean | Observed σ | Model σ | NMB | MB | RMSE | Correlation |
|---|---|---|---|---|---|---|---|---|---|
| Precipitation (mm) | M3DRY_BASE_2016 | 19.89 | 19.52 | 26.40 | 25.00 | -1.84 | -0.37 | 19.30 | 0.72 |
| | STAGE_BASE_2016 | | 19.52 | | 25.00 | -1.84 | -0.37 | 19.30 | 0.72 |
| $SO_4^{2-}$ (kg/ha) | M3DRY_BASE_2016 | 0.088 | 0.078 | 0.130 | 0.130 | -11.600 | -0.010 | 0.115 | 0.61 |
| | STAGE_BASE_2016 | | 0.080 | | 0.131 | -9.050 | -0.008 | 0.115 | 0.61 |
| $NO_3^-$ (kg/ha) | M3DRY_BASE_2016 | 0.118 | 0.101 | 0.150 | 0.125 | -14.200 | -0.017 | 0.126 | 0.6 |
| | STAGE_BASE_2016 | | 0.103 | | 0.127 | -12.700 | -0.015 | 0.127 | 0.6 |
| $NH_4^+$ (kg/ha) | M3DRY_BASE_2016 | 0.055 | 0.027 | 0.085 | 0.046 | -50.200 | -0.027 | 0.079 | 0.49 |
| | STAGE_BASE_2016 | | 0.031 | | 0.051 | -43.000 | -0.024 | 0.076 | 0.52 |

**Table 4.** Model performance statistics for all weekly total precipitation and $SO_4^{2-}$, $NO_3^-$, and $NH_4^+$ wet deposition samples collected at NADP monitors in 2016. The standard deviation over all samples is denoted as σ while NMB, MB, and RMSE represent the percentage normalized mean bias, mean bias, and root mean square error computed over all samples.

| Species | Simulation | Observed Mean | Model Mean | Observed σ | Model σ | NMB | MB | RMSE | Correlation | Observed 2016 – 2010 | Model 2016 – 2010 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Precipitation (mm) | M3DRY_BASE_2010 | 19.47 | 17.45 | 26.50 | 24.00 | -10.40 | -2.02 | 19.70 | 0.71 | 0.42 | 2.07 |
| | STAGE_BASE_2010 | | 17.45 | | 24.00 | -10.40 | -2.02 | 19.70 | 0.71 | 0.42 | 2.07 |
| $SO_4^{2-}$ (kg/ha) | M3DRY_BASE_2010 | 0.128 | 0.107 | 0.201 | 0.162 | -16.300 | -0.021 | 0.171 | 0.58 | -0.040 | -0.030 |
| | STAGE_BASE_2010 | | 0.112 | | 0.168 | -12.900 | -0.017 | 0.172 | 0.58 | -0.04 | -0.03 |
| $NO_3^-$ (kg/ha) | M3DRY_BASE_2010 | 0.121 | 0.108 | 0.153 | 0.144 | -10.300 | -0.012 | 0.133 | 0.60 | 0.00 | -0.01 |
| | STAGE_BASE_2010 | | 0.111 | | 0.147 | -8.240 | -0.010 | 0.135 | 0.60 | -0.003 | -0.008 |
| $NH_4^+$ (kg/ha) | M3DRY_BASE_2010 | 0.043 | 0.032 | 0.072 | 0.056 | -25.700 | -0.011 | 0.064 | 0.53 | 0.01 | 0.00 |
| | STAGE_BASE_2010 | | 0.034 | | 0.058 | -20.800 | -0.009 | 0.063 | 0.55 | 0.01 | 0.00 |

**Table 5.** Model performance statistics for all weekly total precipitation and $SO_4^{2-}$, $NO_3^-$, and $NH_4^+$ wet deposition samples collected at NADP NTN monitors in 2010. The standard deviation over all samples is denoted as σ while NMB, MB, and RMSE represent the percentage normalized mean bias, mean bias, and root mean square error computed over all samples. The last two columns show differences in average observed and modeled weekly total values between 2016 and 2010.




285 **Figure 1:** Monthly mean observed and modeled concentrations at AQS sites for MDA $O_3$, $SO_2$, $NO_x$, and total and speciated $PM_{2.5}$.

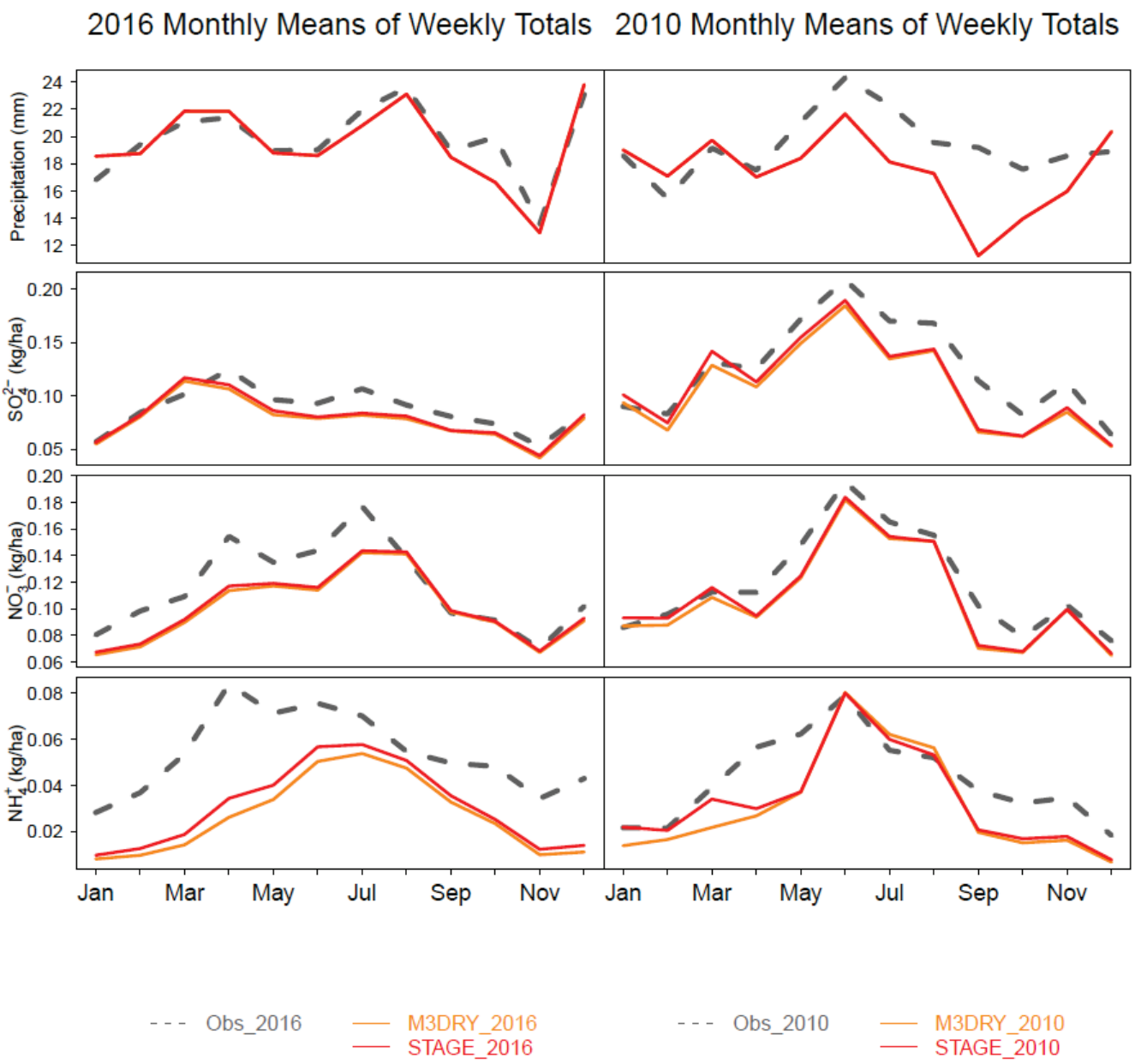



**Figure 2:** Monthly mean observed and modeled precipitation and wet deposition at NADP NTN sites.

290

Table 3 and the right column of Figure 1 show the corresponding results for the 2010 AQMEII4 simulations M3DRY_2010 and STAGE_2010. The results show that the sign of the NMB and MB for 2010 is the same as that for the 2016 results for MDA8 $O_3$, $NO_x$, $SO_2$, $PM_{2.5}$ mass, OC and EC. Bias results for $SO_4^{2-}$ and $NO_3^-$ show greater differences between the years, suggesting that the effects of the substantial reductions in $SO_2$ and $NO_x$ emissions (Foley et al., 2023) between 2010 and 2016 on these aerosol species may not be captured perfectly. The observed decrease in concentrations between 2010 and 2016 is captured by both simulations for all pollutants and is likely caused in large parts by substantial decreases in emissions (Foley et al., 2023) although differences in meteorological conditions between these years may also have played a role. The magnitude of the decrease is overestimated for $O_3$, $NO_x$, and $PM_{2.5}$ mass, though it is important to note that the monitors at which the statistics are calculated differ both between pollutants and between years. The differences in model performance between M3DRY_2010 and STAGE_2010 are similar to those between M3DRY_2016 and STAGE_2016.

Tables 4-5 and Figure 2 show model performance results for weekly precipitation and wet deposition at NADP NTN monitors in 2016 and 2010. Consistent with the negative precipitation bias, simulated wet deposition flux biases also are negative for all pollutants. This confirms that model performance for precipitation is a key driver for wet deposition model performance. The NMB results presented here fall within the range of retrospective long-term simulations over North America (Zhang et al., 2019). The results for 2010 are similar to those for 2016, again with negative biases for all variables. A comparison between 2010 and 2016 shows that the model captured the sign of the observed changes for precipitation and wet deposition fluxes, with wetter conditions in 2016 but mostly similar or slightly lower wet deposition fluxes, likely due to significant reductions in $NO_x$ and $SO_2$ emissions.

The results presented in this section demonstrate that the AQMEII4 CMAQ simulations perform similarly to other comparable regional-scale modeling studies (Emery et al., 2017; Kelly et al., 2019; Simon et al. 2012; Appel et al., 2021). The results presented in the supplemental material show that the choice of the CMAQ dry deposition scheme (M3Dry vs. STAGE) has a smaller impact on aggregated model performance metrics than the sensitivity of CMAQ results to model input data sets and boundary conditions that represent the large-scale chemical environment. However, it is important to note that M3Dry and STAGE share many structural similarities (see Figures B2 and B3 in Galmarini et al., 2021) and that the similarity in model evaluation results therefore does not imply that uncertainty in process-level representation of dry deposition is not a potentially important factor causing differences between model output and observations. An analysis of point model simulations at eight ozone flux measurement sites performed with all dry deposition schemes participating in AQMEII4 shows that differences in seasonal cycles of $V_d$ and deposition pathways between M3Dry and STAGE are generally smaller than differences relative to other schemes. In addition, the following sections demonstrate that the choice of M3Dry vs. STAGE in CMAQ can have more pronounced impacts for specific seasons, regions, and deposition pathways than its impact on these domain-wide model performance results.

### 3.2 Diagnostic gas phase dry deposition comparison M3Dry vs. STAGE

The following sections use the diagnostic variables generated for AQMEII4 to gain insights into the processes causing differences between the M3Dry and STAGE CMAQ simulations. This analysis starts with a comparison of grid-scale quantities, then proceeds to comparisons performed for the specific LU types defined for AQMEII4 (Galmarini et al., 2021), and finally compares and discusses different approaches for handling sub-grid LU variations and LU aggregation in M3Dry and STAGE. To avoid repetition, all analyses in these sections focus on the CMAQ AQMEII4 simulations performed for 2016 because the differences between the M3Dry and STAGE CMAQ simulations for 2010 were very similar to those for 2016 and because the sensitivity simulation quantifying the impacts of using a different LU classification scheme was performed for 2016.

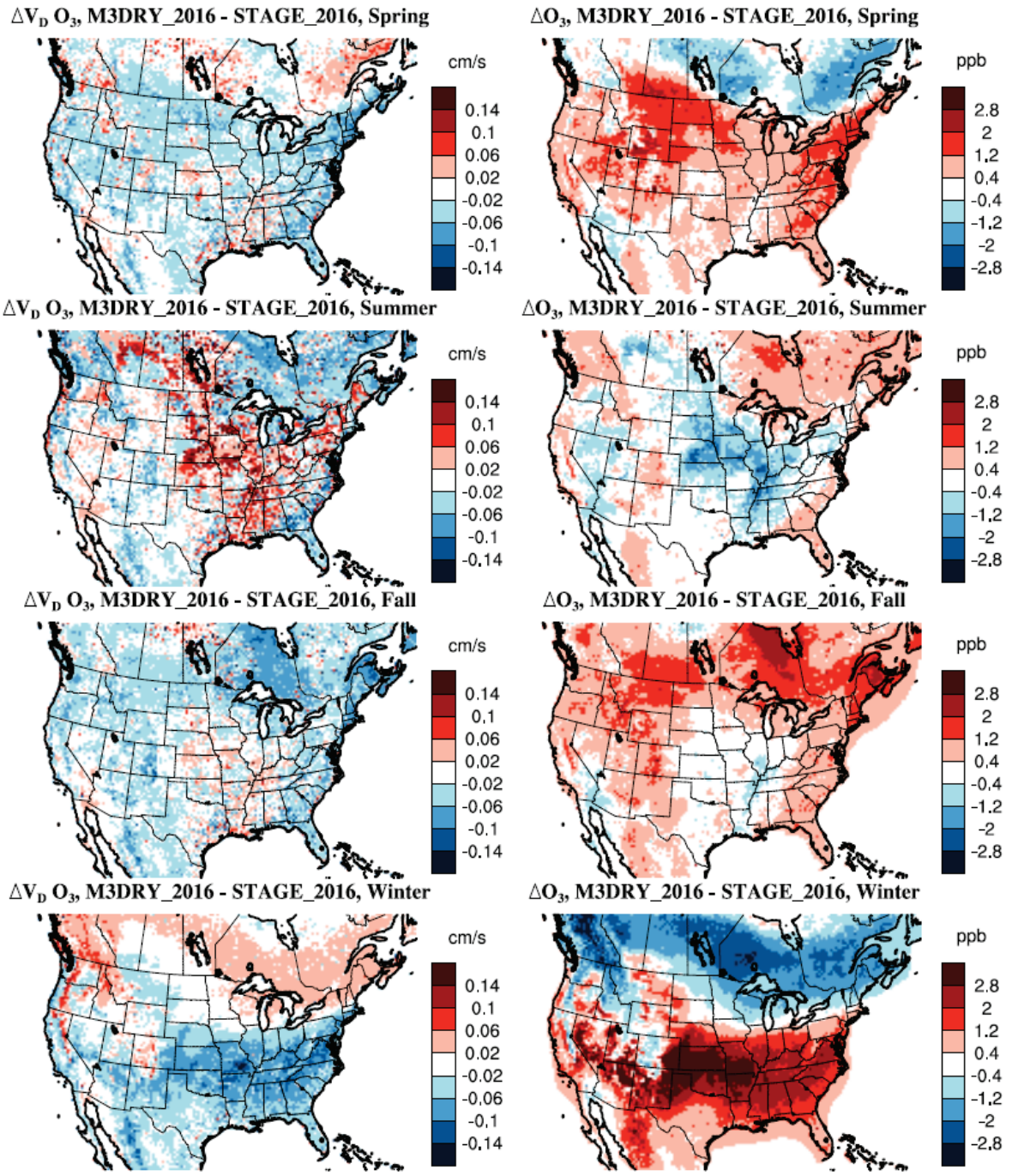
#### 3.2.1 Grid Scale Dry Deposition Diagnostics and Fluxes

Figure 3 shows differences in seasonal mean $O_3$ $V_d$ and mixing ratios between M3DRY_2016 and STAGE_2016. It can be seen that the spatial patterns of differences in $V_d$ and $O_3$ mixing ratios are closely linked, with areas of positive (negative) $V_d$ differences between M3Dry and STAGE generally corresponding to areas of negative (positive) mixing ratio differences. During summer, M3Dry has higher $V_d$ and lower mixing ratios than STAGE for much of the eastern U.S. while the reverse is the case over eastern Canada and along the West Coast. In contrast, during winter STAGE has higher $V_d$ and lower mixing ratios than M3Dry over most of the southern half of the modeling domain while the reverse is the case for much of the northern U.S. and southern Canada. The differences in seasonal mean mixing ratios reach 2-3 ppb in a number of locations, indicating that the effects of different dry deposition schemes can be more pronounced locally than in the spatially aggregated metrics presented in Section 3.1

To diagnose reasons for the differences in $V_d$ between M3Dry and STAGE, we apply the concept of effective conductances (Paulot et al., 2018; Clifton et al., 2020b) as adapted for AQMEII4 (Galmarini et al., 2021). Given that the sum of the effective conductances equals $V_d$, this allows an attribution of $V_d$ to distinct pathways controlled by different processes. The four pathways defined in Galmarini et al. (2021) and based on the original Wesely (1989) scheme are stomatal, cuticular, lower canopy, and soil. Neither M3Dry nor STAGE include a deposition pathway to the lower canopy (see Figures B2 and B3 in Galmarini et al., 2021) but both distinguish between deposition to bare vs. vegetated soil with the latter including an additional in-canopy convective resistance term. Therefore, we analyzed effective conductances for the stomatal, cuticular, vegetated soil, and bare soil pathways. We also note that the effective conductances analyzed here represent grid-scale values calculated from LU-weighted averages of the corresponding LU-specific diagnostic variables requested for AQMEII4.

Seasonal average effective conductance maps for each pathway for M3Dry and STAGE as well as their differences are shown in Figure 4 for summer and in Figure 5 for winter. Average diurnal cycles for summer and winter computed over all non-water grid cells are shown in Figure 6 (absolute effective conductances) and Figure 7 (percentage contribution to $V_d$). Figures 4-5

show that the absolute and relative magnitude of the different pathways varies both spatially and seasonally for both M3Dry and STAGE, with a generally greater importance of the stomatal and cuticular pathways in the more vegetated eastern and northern portions of the modeling domain especially during summer and a greater importance of the deposition to soil, especially bare soil, in the southwestern portion of the modeling domain and during winter. The comparison between M3Dry and STAGE shows generally higher summertime stomatal and wintertime cuticular effective conductances for M3Dry and generally higher soil effective conductances (both vegetated and bare) for STAGE in both summer and winter. The diurnal cycles in Figures 6-7 confirm the seasonal variations and differences between M3Dry and STAGE shown in Figures 4-5 and also illustrate the strong diurnal variation of several pathways, especially the stomatal pathway. In a relative sense, the stomatal effective conductance accounts for about half of the total $V_d$ during daytime hours in summer, though these diurnal cycles represent an average over the entire domain so the contribution would be expected to be higher over the eastern portion of the modeling domain and lower over the southwestern portion of the modeling domain.




365 **Figure 3:** Differences in seasonal mean $O_3$ deposition velocities and mixing ratios M3DRY_2016 minus STAGE_2016.

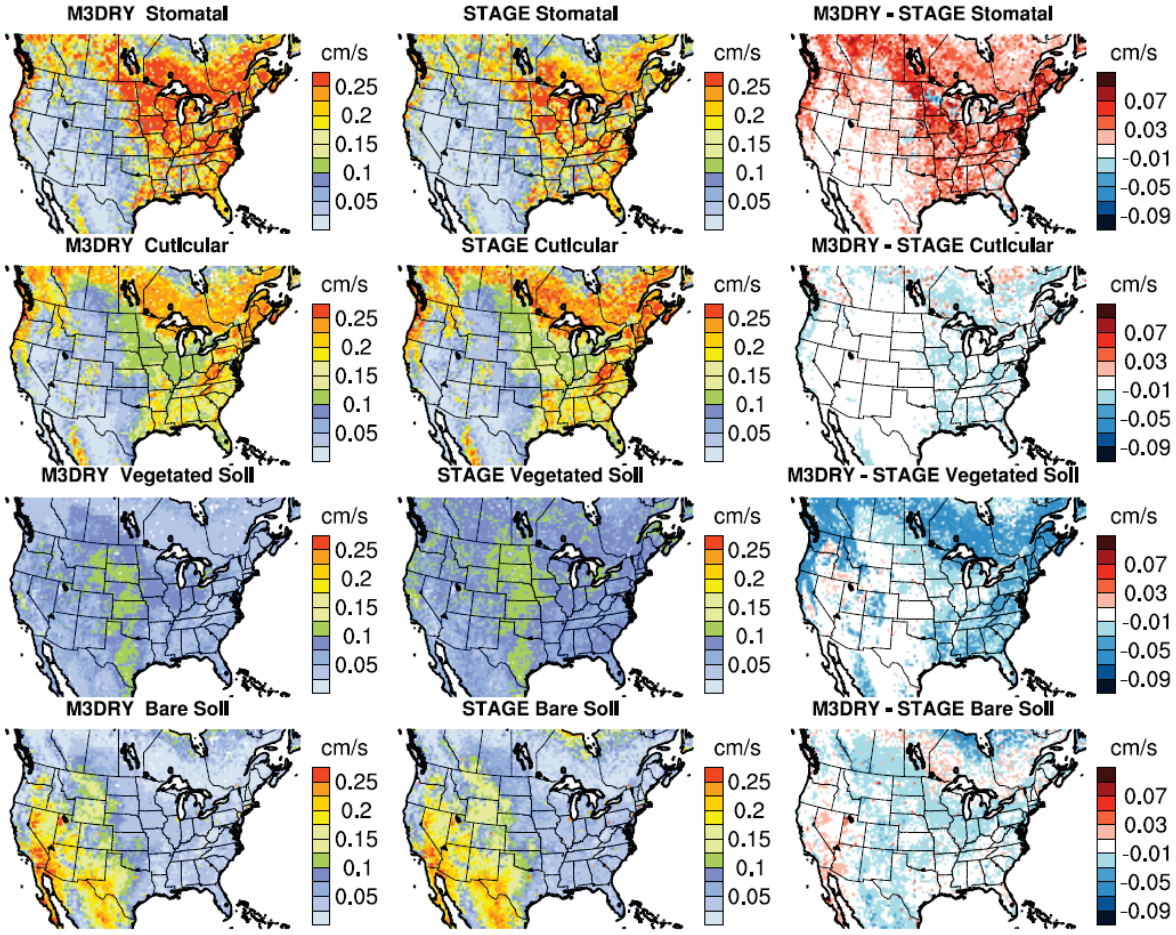



**Figure 4.** Grid-scale $O_3$ summer mean effective conductance maps for the stomatal, cuticular, vegetated soil, and bare soil pathways for M3DRY_2016, STAGE_2016, and M3DRY_2016 minus STAGE_2016.

Comparing the effective conductance difference maps in the right columns of Figures 4-5 to the summer and winter $V_d$ difference maps in Figure 3 shows that the higher summer M3Dry $V_d$ over the eastern U.S. are largely due to a larger stomatal conductance. While the M3Dry stomatal conductance is also moderately larger than for STAGE over eastern Canada, this is counteracted by a substantially larger STAGE soil effective conductance, leading to a net negative difference between M3Dry and STAGE $V_d$ values over that region. Similarly, the split between larger winter STAGE $V_d$ over the southern portion of the domain and larger M3Dry $V_d$ over much of the Canadian portion of the modeling domain is the result of substantially larger STAGE soil effective conductances with only small differences in stomatal and cuticular effective conductances over the southern portion of the domain and substantially larger cuticular effective conductances in M3Dry with only small differences in soil effective conductances over the Canadian portion of the modeling domain.

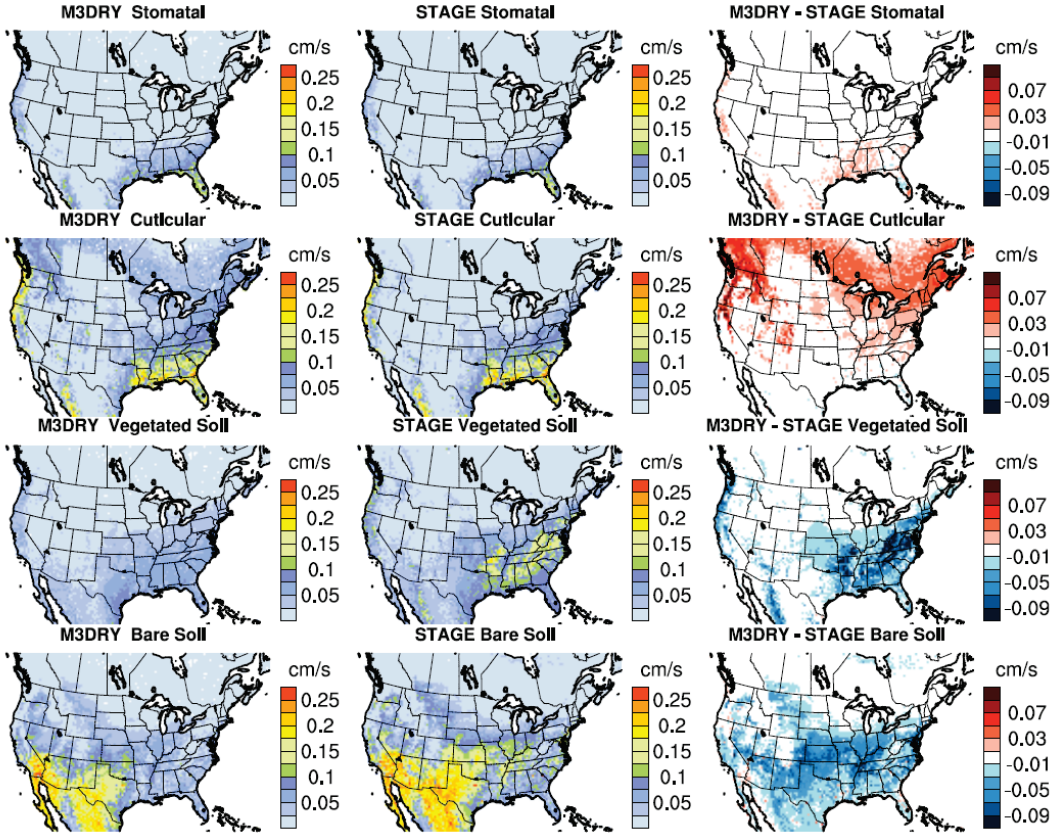



**Figure 5.** As in Figure 4 but for winter mean effective conductances.

In addition to the effective conductances, AQMEII4 also requests participating models to save and submit key component resistances common to all schemes (see Galmarini et al., 2021 Table 4 for a listing of all requested component resistances and Tables B2 and B3 for their definition in the M3Dry and STAGE deposition diagrams in Figures B2 and B3). Like for the effective conductances, these diagnostic variables are calculated for each LU category and grid scale values are calculated as LU-weighted averages from LU-specific values. Figure 8 shows average summer and winter diurnal cycles of six inverse component resistances averaged over all non-water grid cells for both M3Dry and STAGE. These component resistances are the stomatal ($R_s$), cuticular ($R_{cut}$), in-canopy convective ($R_{dc}$), quasi-laminar sublayer ($R_b$ for M3Dry and $R_{can,qlsb}$ and $R_{gnd,qlsb}$ for STAGE), and aerodynamic ($R_a$) resistances and they are plotted as inverse resistances for easier comparison to $V_d$ and effective conductances. As seen in Galmarini et al. (2021) Figures B2 and B3, $R_s$, $R_{cut}$, and $R_{dc}$ are pathway specific and parallel to each other, while they are also serial with $R_a$ and the quasi-laminar sublayer resistance. In M3Dry, the quasi-laminar sublayer resistance is pathway independent while in STAGE it differs between the canopy (cuticular and stomatal) and ground (vegetated and bare soil) pathways. Note that the y-axis range differs across different resistances to better highlight seasonal variations and differences between M3Dry and STAGE for a given resistance.

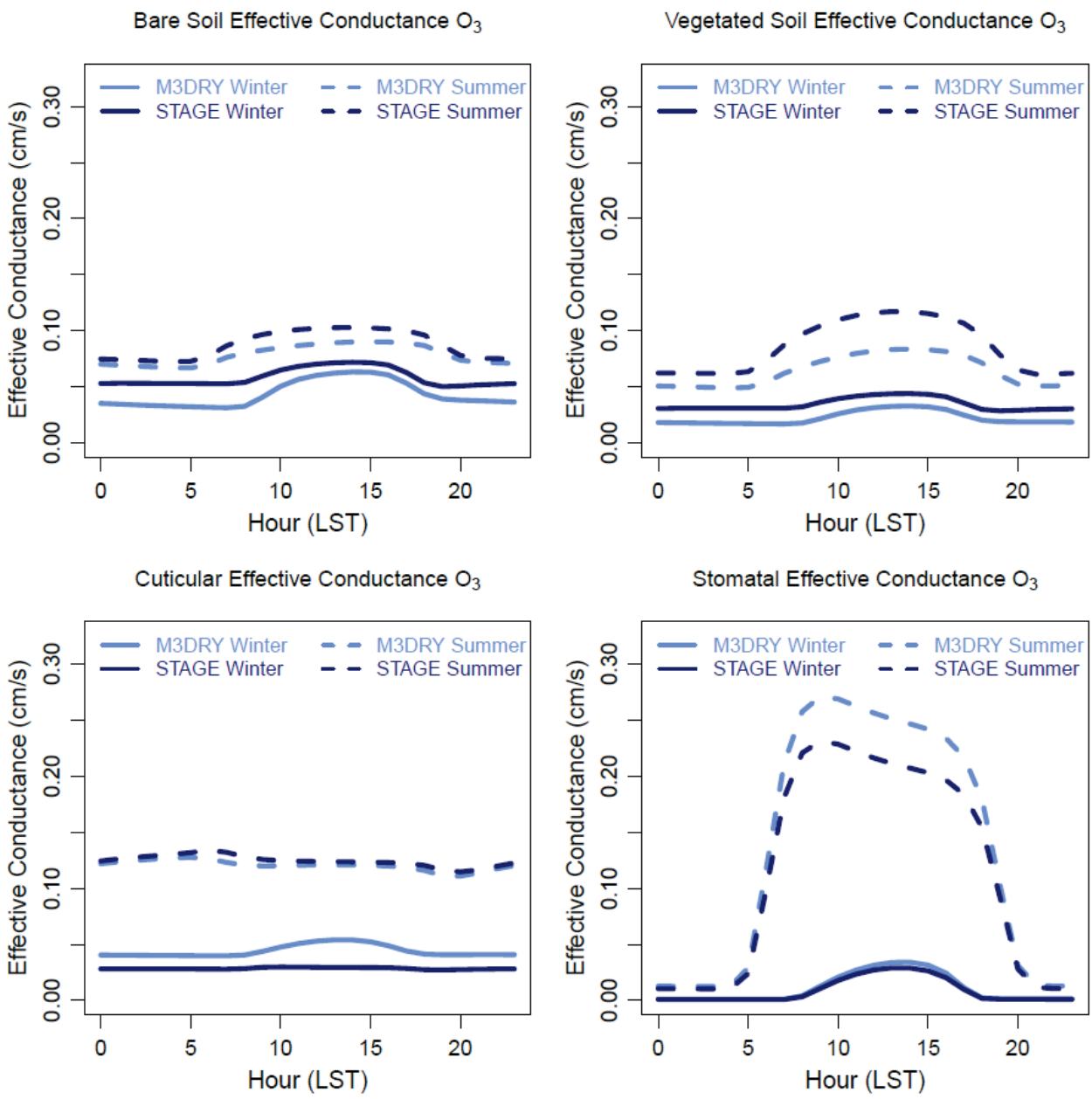



**Figure 6:** Summer and winter domain average diurnal cycles of grid-scale $O_3$ effective conductances for the stomatal, cuticular, vegetated soil, and bare soil pathways for M3DRY_2016 and STAGE_2016.

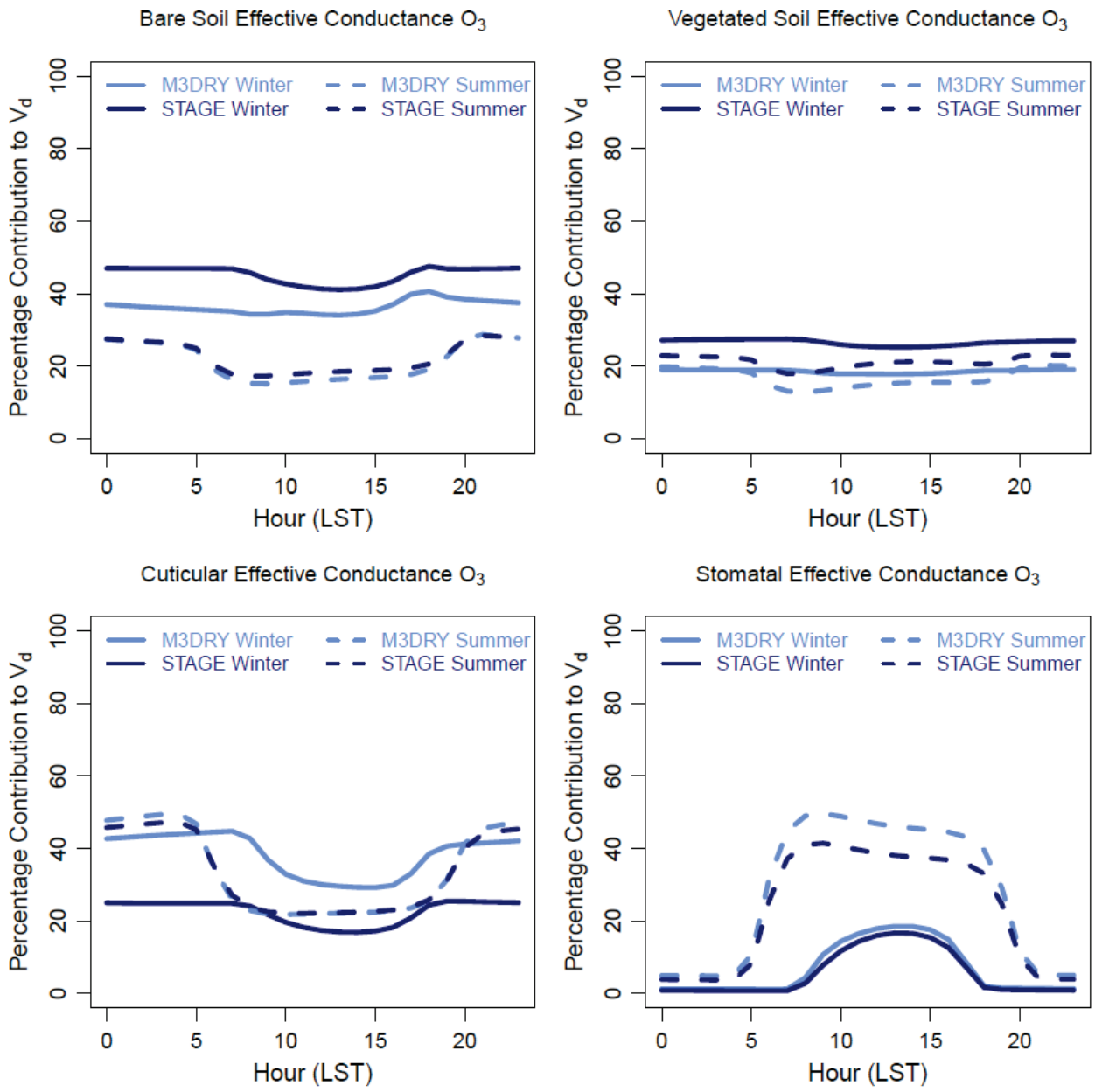
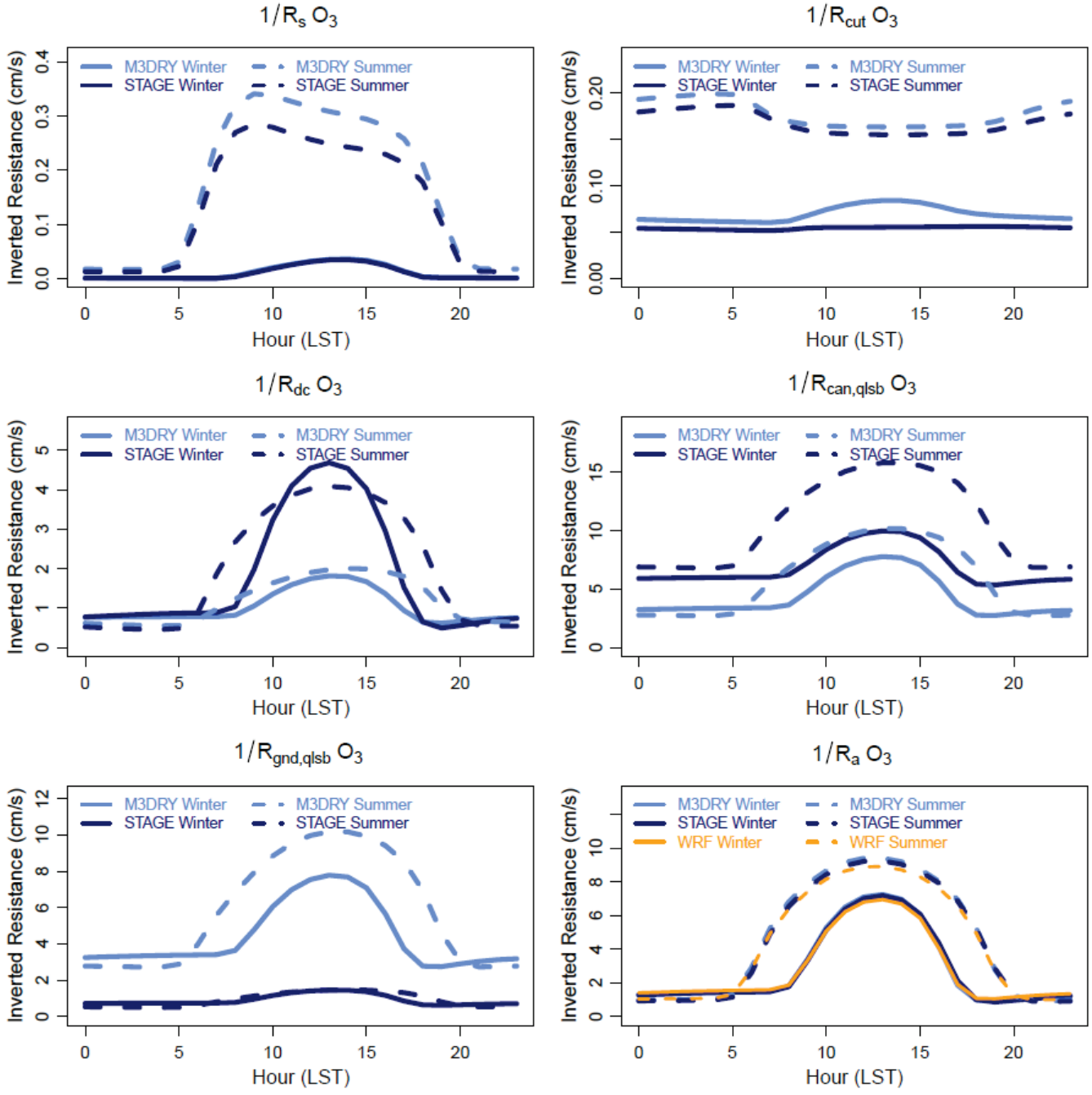



**Figure 7:** As in Figure 6, but showing the percentage contribution of the effective conductance for each pathway to the total deposition velocity.




405 **Figure 8:** Summer and winter domain average diurnal cycles of grid-scale $O_3$ component resistances for M3DRY_2016 and STAGE_2016. Note the different y-axis ranges for different inverted resistances. The panel showing results for the inverted aerodynamic resistance $1/R_a$ also includes the values computed in the WRF PX LSM.

Consistent with the effective conductances shown above, values for summertime $1/R_s$ and wintertime $1/R_{cut}$ differ between M3Dry and STAGE, with the higher values of these inverted resistances in M3Dry causing higher effective conductances for these pathways. Inverted $R_{dc}$ is substantially higher in STAGE than M3Dry. While soil resistance was not requested and saved as a separate term in AQMEII4 as its definition differs across schemes, the higher inverted STAGE $R_{dc}$ values likely had an effect on the higher STAGE effective conductance values for the vegetated soil pathway. For the quasi-laminar sublayer resistances, inverted $R_{can,qlsb}$ for STAGE is higher than the pathway-independent inverted $R_b$ for M3Dry which in turn is higher than the inverted $R_{gnd,qlsb}$ for STAGE. However, these resistances are typically small compared to the other resistances (i.e. their inverted values are larger) so that they generally only have a small impact on effective conductances and overall $V_d$. The inverted aerodynamic resistance $R_a$ is very similar between M3Dry and STAGE for both summer and winter. As discussed in Section 2.3.1, all LU-specific diagnostic deposition variables for M3Dry are estimated through a post-processor while the M3Dry deposition calculations within CMAQ used grid-scale $R_a$ calculated by the WRF PX LSM to maintain maximum consistency between the representation of land-surface exchange processes in WRF and CMAQ. Therefore, the panel for inverted $R_a$ also shows the WRF-based grid-scale value used in the M3Dry CMAQ deposition calculations. This value is very similar to the post-processor based estimate for M3Dry. The different approaches taken by M3Dry and STAGE to handle sub-grid variability in LU and their impacts on deposition calculations is discussed further in Section 3.2.3.

While this section so far has focused on applying the AQMEII4 diagnostics to analyze M3Dry vs. STAGE differences in $O_3$ deposition, these diagnostics are also being calculated for $SO_2$, $NO_2$, NO, $HNO_3$, $NH_3$, PAN, $HNO_4$, $N_2O_5$, organic nitrates, $H_2O_2$, and HCHO (Galmarini et al., 2021). Moreover, the effective conductances and total $V_d$ can also be used to calculate *effective fluxes*, i.e., apportion total dry deposition fluxes to specific pathways (Galmarini et al., 2021). Figure 9 shows annual domain-average (excl. water cells) $V_d$ for M3Dry and STAGE for $O_3$, $H_2O_2$, HCHO, $SO_2$, and oxidized nitrogen species as sum of the effective conductances for the four pathways while Figure 10 shows the corresponding annual total domain-wide (excl. water cells) dry deposition fluxes as sum of the effective fluxes. The $V_d$ and deposition flux results for $O_3$ reflect the larger contributions from the vegetated and bare soil pathways for STAGE and the larger contributions for the stomatal and cuticular pathways for M3Dry. Annual domain total $O_3$ deposition fluxes differ only slightly between M3Dry and STAGE due to the seasonal and spatial variation of the $V_d$ differences shown in Figure 3. In contrast to $O_3$, most other pollutants show larger effective conductances and effective fluxes for the bare soil and sum of bare and vegetated soil pathways for M3Dry than STAGE. The cuticular effective flux is larger for M3Dry than STAGE for HCHO, $SO_2$, $HNO_4$, and organic nitrates, smaller for M3Dry than STAGE for $H_2O_2$ and $HNO_3$, and similar between M3Dry and STAGE for other species. Stomatal effective fluxes are small for all species except $O_3$, HCHO, and $NO_2$ for both M3Dry and STAGE. Total dry deposition fluxes differ the most between M3Dry and STAGE for HCHO and organic nitrate, though again it should be noted that these fluxes represent annual domain totals and larger differences likely exist at sub-seasonal scales for different regions.

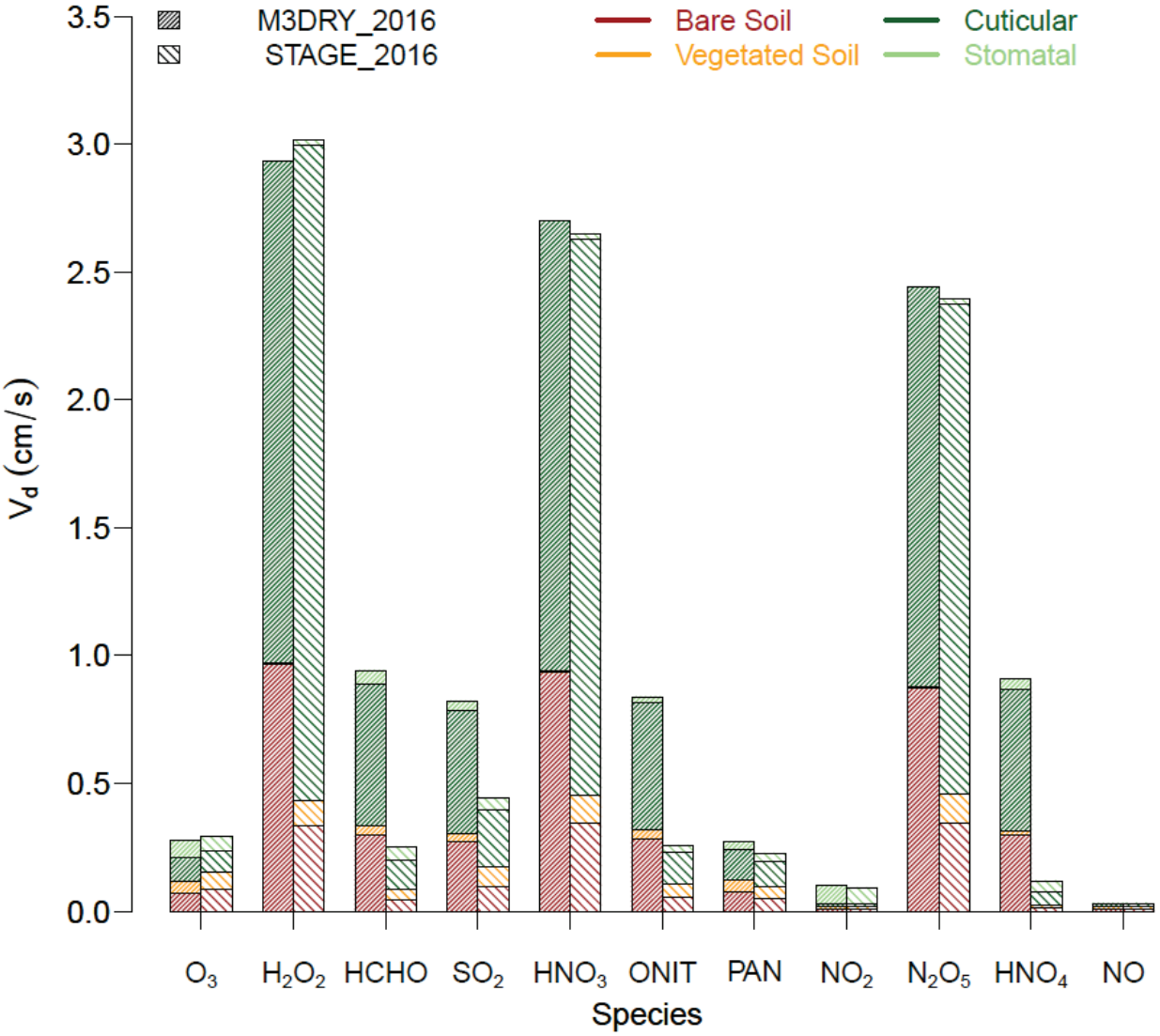



**Figure 9.** Grid-scale annual average domain-wide (excl. water cells) effective conductances for $O_3$, $H_2O_2$, HCHO, $SO_2$, and oxidized nitrogen species.

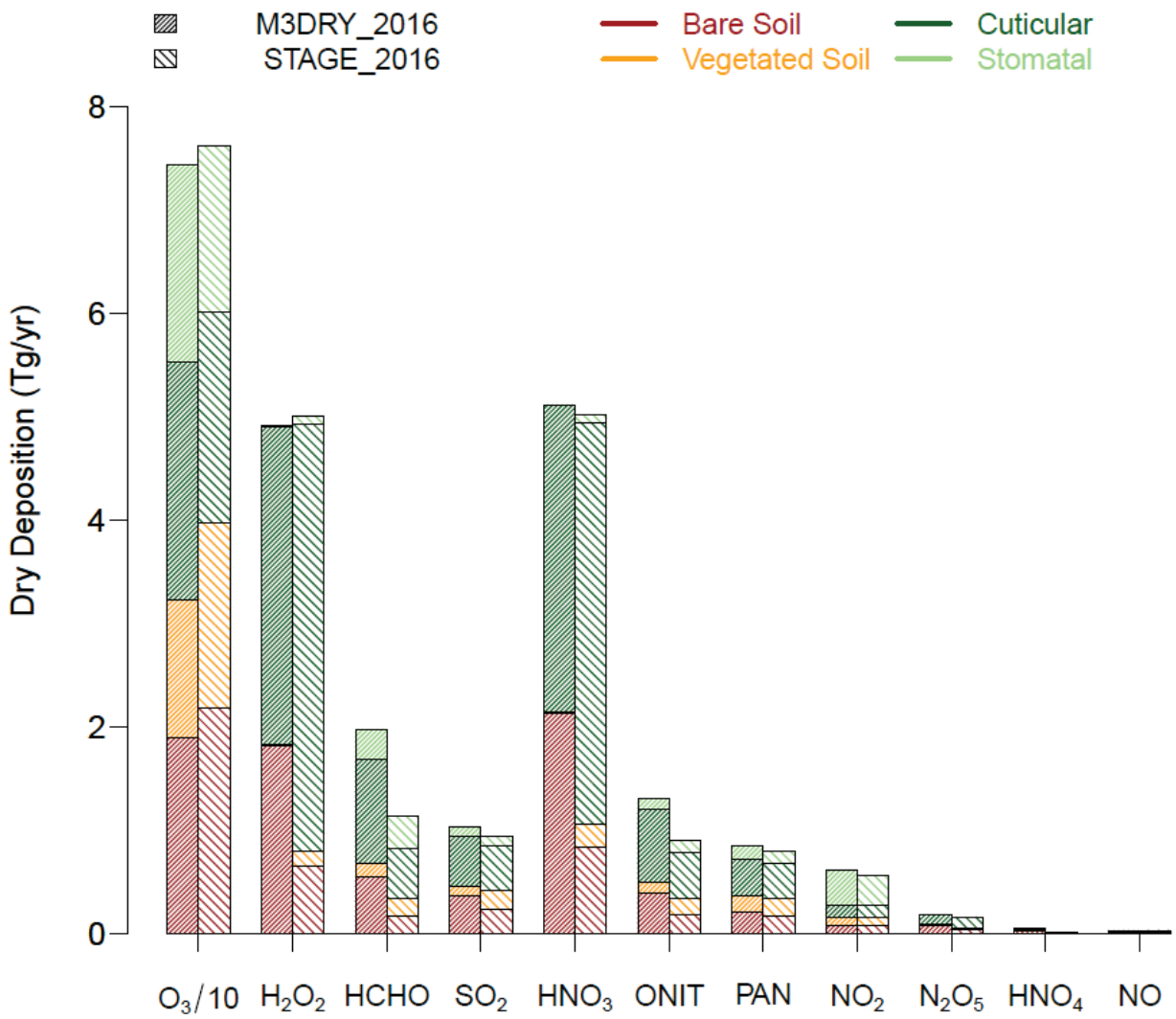



**Figure 10.** Grid-scale annual total domain-wide (excl. water cells) pathway-specific dry deposition fluxes ("effective fluxes") for $O_3$, $H_2O_2$, HCHO, $SO_2$, and oxidized nitrogen species. Ozone dry deposition values are divided by a factor 10 to use the same y-axis as for the other pollutants.

### 3.2.2 LU-Specific Dry Deposition Diagnostics and Fluxes

In this section, we utilize the LU-specific dry deposition diagnostics generated during AQMEII4 to provide further insights into the grid-scale comparisons presented above. These LU-specific diagnostics were generated for the 16 common LU types defined in Galmarini et al. (2021). Figure S8 depicts spatial maps of the fractional coverage for each of these 16 categories, aggregated through post-processing from the native 20 MODIS LU categories used in the M3DRY_2016 simulations (see

Section 2.3.1). The aggregation of the MODIS LU categories to the AQMEII4 LU categories (performed during post-processing for M3Dry and prior to the deposition calculations in STAGE as discussed in Section 2.3.2) is documented in Table S1. Note that none of the 20 MODIS LU categories correspond to the AQMEII4 herbaceous category. Figure S8 shows the prevalence of evergreen needleleaf, mixed, and deciduous forest in the northern, northeastern, and southeastern portions of the modeling domain, planted/cultivated LU in the north-central portion of the modeling domain, grassland in a belt stretching from Texas to Montana, and shrubland in the southwestern portion of the modeling domain. Figure S9 shows bar charts of the fraction of the modeling domain covered by each LU category. Separate bar charts are shown for M3DRY_2016 and STAGE_2016. While both started with the fractional coverages of the MODIS LU categories in the 12 km modeling domain, they differ in their treatment of grid cell with partial water coverage. Consistent with the approach taken in the WRF PX LSM, M3DRY_2016 treats cells with more than 10% water coverage as either all land or all water. For grid cells with water fractions between 10% and 50%, the water fraction is reset to zero and the non-water categories are renormalized to 100%. For grid cells with water fractions exceeding 50%, the water fraction is reset to 100% and the fractions for non-water categories are set to zero. No renormalization is performed for grid cells with water fraction coverage below 10%. The rationale for this approach is that meteorological flux calculations are distinctly different between land and water and both WRF PX LSM and M3Dry are designed to retain that distinctiveness imposed by the underlying spatial discretization of the modeling domain. On the other hand, STAGE does not apply any special treatment to grid cells with partial water coverage. As a result of these different approaches, Figure S9 shows that a slightly larger fraction of the modeling domain is covered by water for STAGE_2016 compared to M3DRY_2016, with correspondingly slightly smaller coverages for non-water categories.

First, we investigate summer and winter average diurnal cycles of LU-specific component resistances in the context of the grid-scale resistance shown in Figure 8 and discussed in Section 3.2.1. Figure S10 shows that $R_s$ conductances are higher for M3Dry than STAGE during summer for all forest types, agricultural land, savanna, and tundra, i.e. common LU types within the modeling domain (Figures S8 and S9). Only wetlands have higher $R_s$ conductance for STAGE than M3dry, with the remaining LU categories having similar $R_s$ conductances for both schemes. This is consistent with the grid scale domain-average summer $R_s$ conductance being higher for M3Dry than STAGE (Figure 8). Only small differences are seen in winter for both the grid-scale (Figure 8) and LU-specific (Figure S10) stomatal conductances. The higher domain-average grid-scale $1/R_{cut}$ conductance for M3Dry during winter (Figure 8) is driven by the higher cuticular conductance values for evergreen needleleaf forest, mixed forest, and agricultural land Figure S11), particularly since these LU types cover a substantial portion of the domain (Figures S8-S9). Figure S12 shows that the higher domain-average grid-scale in-canopy convective conductance ($1/R_{dc}$) for STAGE during both summer and winter (Figure 8) is present for almost all LU categories with the exception of snow and ice which (along with barren) uses a placeholder value to represent the non-existent canopy and which is present in very few grid cells in the domain (Figure S9). LU-specific $R_{can,qlsb}$ conductances ($1/R_{can,\,qlsb}$) are higher for STAGE than M3Dry for almost all LU categories in both summer and winter (Figure S13) while the opposite is the case for $R_{gnd,qlsb}$ (Figure S14). This demonstrates that the same behavior noted for the grid-scale $R_{can,qlsb}$ and $R_{gnd,qlsb}$ conductances (Figure 8) was not driven

by differences in their representation by M3Dry vs. STAGE for select LU categories only but rather resulted from consistent differences in representation across all LU categories. LU-specific $R_a$ conductances ($1/R_a$) are generally very similar between STAGE and M3Dry with the exception of urban and tundra for which M3Dry is higher than STAGE and wetlands for which STAGE is higher than M3Dry (Figure S15). This general similarity between M3Dry and STAGE is consistent with the net grid scale domain-average $R_a$ conductances shown in Figure 8. Differences between grid-scale and LU-weighted aggregated LU-specific $R_a$ are investigated further in the next section.

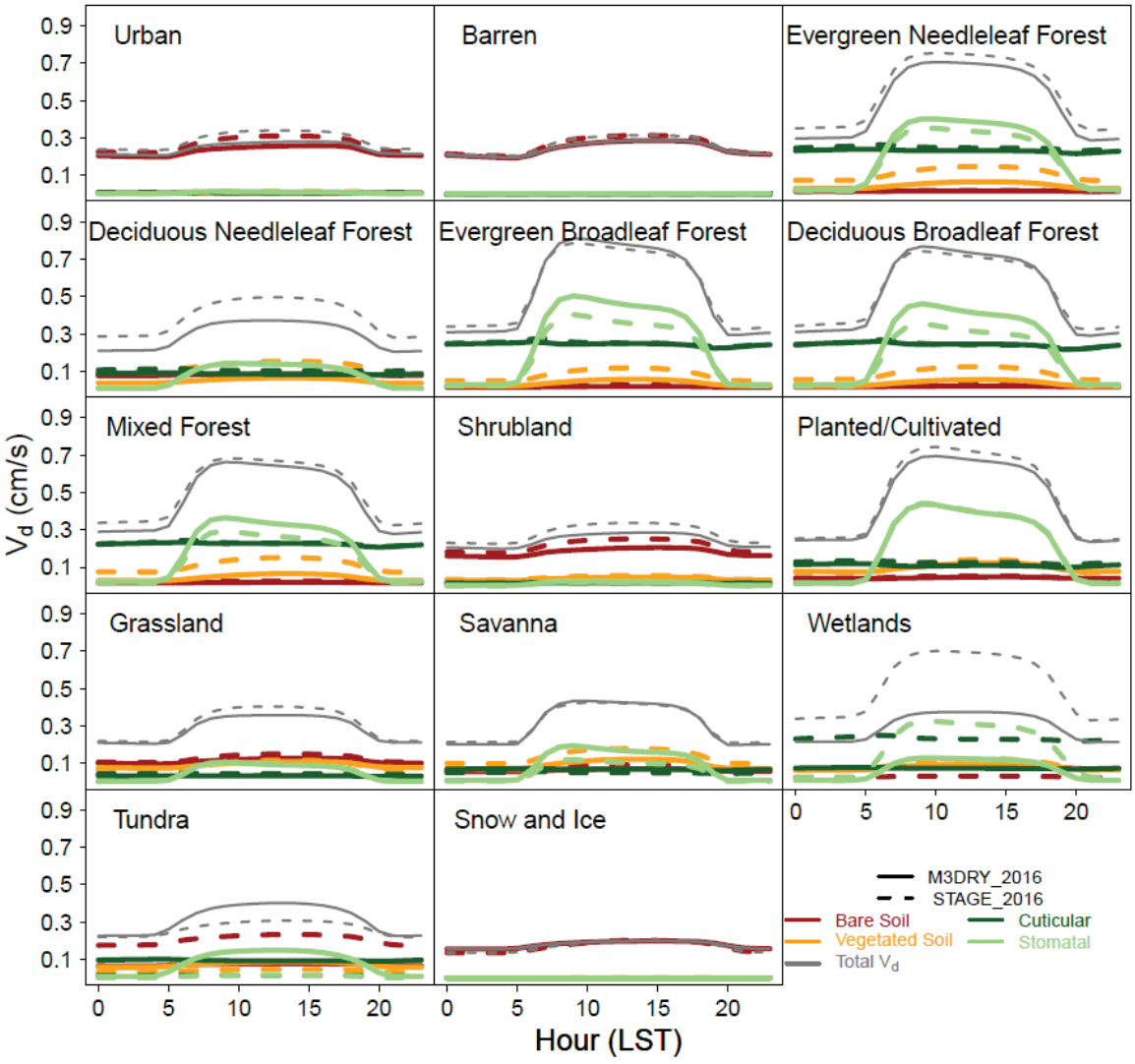



**Figure 11.** Summer domain average diurnal cycles of LU-specific $O_3$ effective conductances and total deposition velocities for M3DRY_2016 and STAGE_2016.

Next, we investigate the effects of the differences in the LU-specific component resistances discussed above on LU-specific effective conductances and effective fluxes, i.e. pathway-specific $V_d$ and dry deposition fluxes. Figure 11 depicts summer average diurnals of LU-specific effective conductances and total $V_d$ for $O_3$ for both M3Dry and STAGE, averaged over all grid cells with a non-zero fractional coverage for a given LU category. As expected, these diurnal cycles illustrate that both total $V_d$ and the relative importance of the different effective conductance pathways varies between LU categories for both M3Dry and STAGE as a result of different underlying surface characteristics like VEGF, vegetation type, etc. for the different LU categories. Total $V_d$ during daytime varies by a factor of 2 between the LU categories with the lowest values (urban, barren, and snow and ice) and those with the highest values (evergreen needleleaf and broadleaf forest, deciduous forest). Effective conductances for the bare soil pathway dominate for the urban, barren, and shrubland LU categories while conductances for the cuticular and stomatal pathways dominate for the forest LU categories, especially during daytime. Results also show that the M3Dry vs. STAGE differences are most pronounced for the stomatal and vegetated soil pathway for the forest LU categories, with M3Dry estimating larger effective conductances for the stomatal pathway and STAGE estimating larger effective conductances for the vegetated soil pathway for these LU categories. Higher effective conductances for the bare soil pathway in STAGE are particularly noticeable for the urban, shrubland, and tundra LU categories. When considering how these differences in LU-specific effective conductances impact the grid-scale summertime effective conductances shown in Figures 4 - 5, the abundance and spatial distribution of each LU category (Figures S8 and S9) needs to be taken into account. For example, given the abundance of the evergreen needleleaf forest, deciduous broadleaf forest, and mixed forest categories over the eastern and northern portions of the modeling domain, the higher M3Dry stomatal conductances and lower M3Dry vegetated soil conductances for these LU categories shown in Figure 11 can explain the spatial pattern of the corresponding M3Dry vs. STAGE grid-scale differences for these two pathways shown in Figures 4 - 5.

The analysis above provided diagnostic insights into the M3Dry vs. STAGE differences in grid-scale effective conductances shown in Figures 4 – 5 and discussed in Section 3.2.1. Analogously, Figure 12 shows annual and summer domain-total LU-specific effective fluxes for $O_3$ to provide further insights into the corresponding grid-scale effective fluxes shown in the left two bars of Figure 10 and also discussed in Section 3.2.1. For annual total deposition, the overall slightly larger grid-scale $O_3$ deposition flux for STAGE (Figure 10) is present for almost all LU categories, with only a few categories (e.g. evergreen needleleaf and broadleaf forest) showing no discernible differences between M3Dry and STAGE. Likewise, the generally larger contribution of the stomatal and cuticular deposition pathways in M3Dry and bare and vegetated soil in STAGE are present for most LU categories, but the differences in the magnitude of the individual deposition pathways are especially pronounced for the evergreen needleleaf forest, mixed forest, and agricultural LU categories. Considering effective fluxes for summertime only, the greater importance of the cuticular and stomatal pathways during this season for LU categories most strongly affected by seasonal variations in LAI (deciduous broadleaf forest and planted/cultivated, see Table 6), along with the greater importance of these pathways in M3Dry compared to STAGE, yield greater overall estimated deposition to these LU categories for M3Dry compared to STAGE. In contrast, for several other LU categories (e.g. evergreen needleleaf and mixed

forest, shrubland, and grassland), STAGE still estimates higher deposition even during summertime, largely due to its higher estimated deposition to vegetated and bare soil and the dominance of these pathways for several of these LU categories (shrubland and grassland). Overall, Figure 12 demonstrates that even though overall annual total $O_3$ deposition fluxes estimated by M3Dry and STAGE are fairly similar (Figure 10), pathway-specific fluxes to individual LU types can vary more substantially on both annual and seasonal scales which might affect estimates of $O_3$ damages to sensitive vegetation.


**Figure 12.** Land-use specific domain-wide (excl. water cells) pathway-specific $O_3$ dry deposition fluxes ("effective fluxes"), a) annual and b) summer

### 3.2.3 Impact of different approaches for handling subgrid LU variations and LU aggregation in M3Dry and STAGE

Both WRF PX LSM (M3Dry) and STAGE rely on LU-specific parameters prescribed in look-up tables to account for subgrid-scale LU variations when performing grid-scale calculations. As discussed in Sections 2.3.1 and 2.3.2, the WRF PX LSM bases grid-scale calculations on LU-weighted aggregated parameters from such LU-specific look-up table values and CMAQ M3Dry then directly uses these calculations in its grid-scale deposition calculations, maintaining consistency with WRF PX LSM. On the other hand, CMAQ STAGE and the M3Dry post-processor developed for AQMEII4 perform LU-specific deposition calculations using the look-up table values for each LU type, and in the case of CMAQ STAGE then use these LU-specific deposition calculations to calculate grid-scale deposition, maintaining consistency between LU-specific and grid-scale deposition but potentially differing from WRF PX LSM for variables like $R_a$, Rs, and $u_*$ (note that starting with CMAQv5.4 released in October 2022, STAGE normalizes $R_a$, Rs, and $u_*$ to match the aggregated grid-scale values to the grid-scale values calculated in the LSM of the driving meteorological model). To provide further context for these differences, Table 6 documents several of these LU-specific look-up table parameters used in deposition calculations, i.e. LAI, VEGF, and $z_0$. Values are shown for the 20 LU category MODIS configuration of WRF PX LSM and M3Dry, the 16 LU category AQMEII4 configuration of STAGE used in this study, and the 20 LU category MODIS configuration in the unmodified version of STAGE used for sensitivity simulation STAGE_REF_2016 discussed below. The boldfaced table entries indicate instances where parameter values differed from WRF PX LSM (and M3Dry) for the STAGE AQMEII4 and/or STAGE MODIS configuration. For example, STAGE prescribes lower $z_0$ for the urban and tundra categories for both the AQMEII4 and MODIS configurations and higher $z_0$ for wetlands for the AQMEII4 configuration. These differences in $z_0$ are consistent with the $1./R_a$ differences for these three LU categories shown in Figure S15. The higher wetlands $z_0$ specified in STAGE results in lower $R_a$ and higher $1./R_a$ compared to M3Dry while the reverse is true for the tundra and urban categories.

| MODIS LU | AQMEII4 LU | LAI (Seasonal Minimum) | | | LAI (Seasonal Maximum) | | | VEGF (Seasonal Minimum) | | | VEGF (Seasonal Maximum) | | | $Z_0$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | PX LSM / M3DRY | STAGE | STAGE REF | PX LSM / M3DRY | STAGE | STAGE REF | PX LSM / M3DRY | STAGE | STAGE REF | PX LSM / M3DRY | STAGE | STAGE REF | PX LSM / M3DRY | STAGE | STAGE REF |
| Evergreen Needleleaf Forest | Evergreen Needleleaf Forest | 3.5 | 3.5 | 3.5 | 5.5 | 5.5 | 5.5 | 93 | 93 | 93 | 93 | 93 | 93 | 100 | 100 | 100 |
| Evergreen Broadleaf Forest | Evergreen Broadleaf Forest | 3.5 | 3.5 | 3.5 | 6 | 6 | 6 | 92 | 92 | 92 | 92 | 92 | 92 | 90 | 90 | 90 |
| Deciduous Needleleaf Forest | Deciduous Needleleaf Forest | 1.5 | 1.5 | 1.5 | 3 | 3 | 3 | 60 | 60 | 60 | 60 | 60 | 60 | 100 | 100 | 100 |
| Deciduous Broadleaf Forest | Deciduous Broadleaf Forest | 1.5 | **2** | **2** | 6 | 6 | 6 | 91 | 91 | 91 | 91 | 91 | 91 | 100 | 100 | 100 |
| Mixed Forest | Mixed Forest | 2.2 | 2.5 | 2.5 | 5.5 | 5.5 | 5.5 | 92 | 92 | 92 | 92 | 92 | 92 | 100 | 100 | 100 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Closed Shrublands | Shrublands | 1 | **1.3** | 1 | 1.5 | 1.5 | 1.5 | 20 | **10** | 20 | 40 | **20** | 40 | 15 | 15 | 15 |
| Open Shrublands | | 1.3 | | 1.3 | 1.5 | | 1.5 | 17 | | **10** | 40 | | **20** | 15 | | 15 |
| Woody Savanna | Savanna | 2 | **1.5** | 2 | 2.3 | **1.5** | 2.3 | 60 | **40** | 60 | 70 | 70 | 70 | 25 | **15** | 25 |
| Savanna | | 0.5 | | **1.5** | 1.5 | | 1.5 | 40 | | 40 | 70 | | 70 | 15 | | 15 |
| Grasslands | Grassland | 0.5 | **1.5** | **1.5** | 1.5 | **2** | 1.5 | 20 | 20 | 20 | 50 | 50 | 50 | 7 | 7 | 7 |
| Permanent Wetlands | Wetlands | 1 | **2.2** | **2** | 2.5 | **5.5** | 2.5 | 35 | **50** | 35 | 65 | **92** | 65 | 20 | **55** | 20 |
| Croplands | Planted / Cultivated | 0.7 | **1.5** | **1.5** | 3.5 | **4** | 3.5 | 20 | **10** | 20 | 90 | **90** | 90 | 10 | 10 | 10 |
| Crop/Natural Mosaic | | 1 | | **1.5** | 3.5 | | 3.5 | 40 | | 40 | 80 | | 80 | 30 | | 30 |
| Urban and Built-up | Developed/ Urban | 0.5 | **1.7** | **2** | 2 | **2.5** | 2 | 5 | 5 | 5 | 5 | 5 | 5 | 80 | **60** | 80 |
| Snow and Ice | Snow/Ice | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 1.2 | 1.2 | 1.2 |
| Sparsely Vegetated | Barren | 0.1 | 0.1 | 0.1 | 0.2 | 0.2 | 0.2 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 5 | 5 | 5 |
| IGBP water | Water | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.1 | 0.1 | 0.1 |
| Wooded Tundra | Tundra | 2 | **1** | 2 | 3.4 | **1** | 3.4 | 50 | **10** | 50 | 70 | **15** | 70 | 30 | **10** | 30 |
| Mixed Tundra | | 1 | | 1 | 2.4 | | 2.4 | 20 | | 20 | 40 | | 40 | 15 | | 15 |
| Barren Tundra | | 0.1 | | 0.1 | 1.4 | | 1.4 | 5 | | 5 | 20 | | 20 | 10 | | 10 |

**Table 6.** LU-specific look-up table parameter values for seasonal minimum and maximum Leaf Area Index (LAI), seasonal minimum and maximum vegetation fraction (VEGF), and surface roughness ($z_0$) used in the dry deposition calculations. See the beginning of Section 3.2.2 for details on the differences between the "PX LSM / M3DRY", "STAGE", and "STAGE REF" columns for each variable. The boldfaced table entries indicate instances where parameter values differed between the WRF PX LSM (also used in M3DRY) and the STAGE AQMEII4 and/or STAGE reference configuration.

Figures S16 and S17 compare grid-scale vs LU-weighted aggregates of LU-specific values for $u_*$ and $R_a$ for both M3Dry and STAGE. For M3Dry, the sum of the LU-weighted aggregated $u_*$ estimates matches the grid-scale values within 10%, with generally larger values for the aggregated estimates. For STAGE_2016 and STAGE_REF_2016, the aggregated values also typically fall within 10% of the grid-scale values, but differences can be both positive and negative. For $R_a$, differences between the LU-weighted aggregated and grid-scale values can reach 25% for all cases examined here, with generally higher values for the LU-weighted aggregated $R_a$ values which is consistent with Figure 8.

Figure 13 compares seasonal mean $O_3$ Vd and mixing ratio differences between STAGE_2016 and STAGE_REF_2016 to assess the impact of performing STAGE deposition calculations on the more aggregated 16 category AQMEII4 LU classification scheme (see section 2.3.2) rather than the 20 category MODIS LU classification scheme used in the WRF PX LSM. Results show that the use of the AQMEII4 LU classification scheme can cause seasonal mean $O_3$ Vd increases of 0.02 – 0.06 cm/s and corresponding season mean $O_3$ mixing ratio decreases of 0.5 – 1 ppb mostly over the eastern and northern portions of the modeling domain. The location of these changes suggests that they are likely at least partially due to the

collapsing of two MODIS agricultural LU categories (croplands and crop/natural mosaic) with different look-up table values to a single AQMEII4 agricultural LU category (see Table 6 and Figure S8).

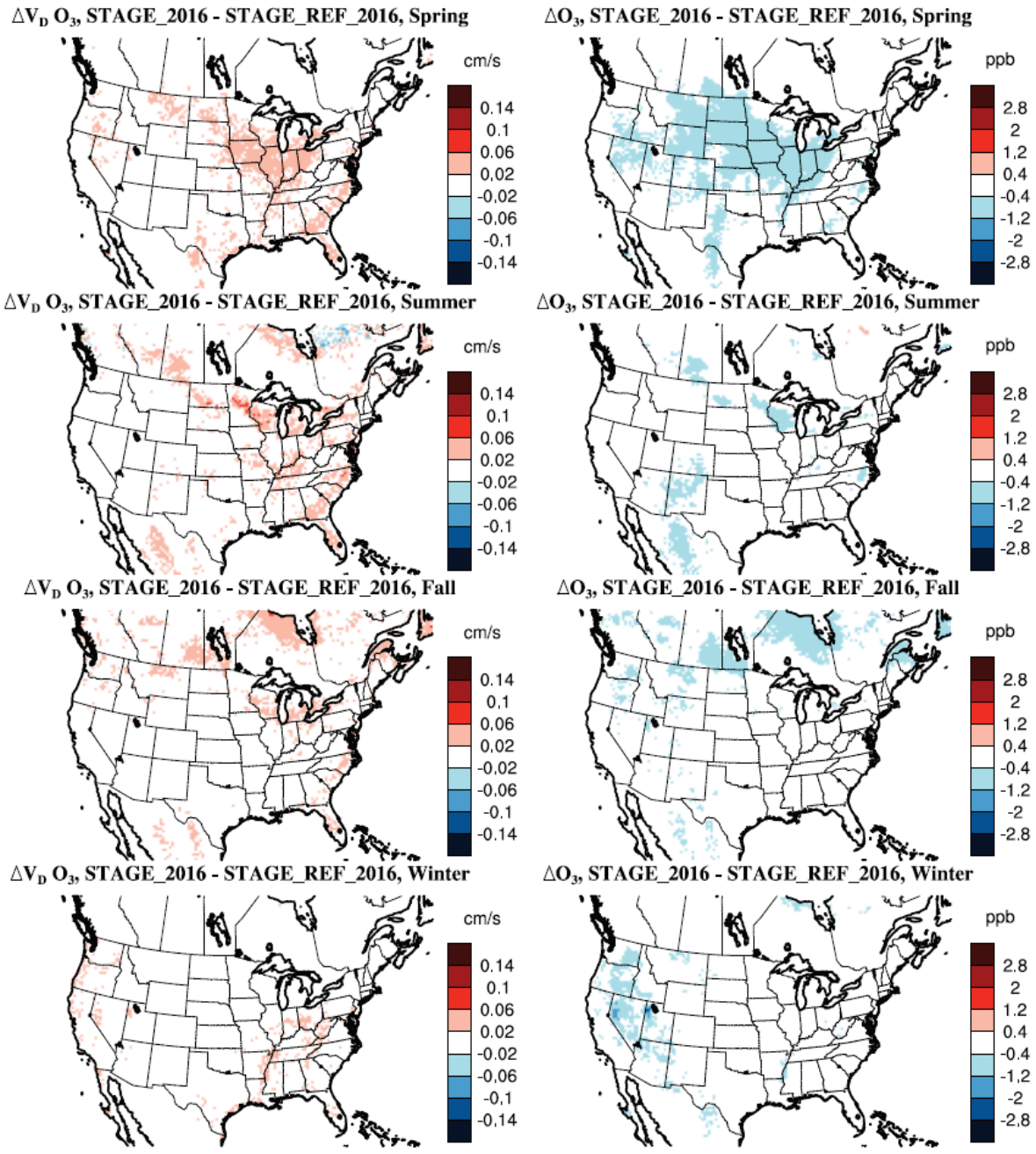



**Figure 13.** Differences in seasonal mean $O_3$ deposition velocities and mixing ratios STAGE_2016 minus STAGE_2016_REF.

The results presented in this section showed that differences in LU-specific look-up table values between different deposition schemes and/or between the WRF PX LSM calculations and the deposition scheme calculations within CMAQ as well as the aggregation of the MODIS LU categories to the AQMEII4 LU categories in the STAGE deposition calculations can introduce differences in the estimated $V_d$ and resulting mixing ratios. In the next section, we investigate the effects of using a different underlying LU classification scheme in both the WRF PX LSM calculations and the CMAQ M3Dry calculations.

### 3.3 Impacts of Land-Use Classification Scheme

In this section, we compare the effects of replacing the MODIS dataset to represent LU in both WRFv4.1.1 and CMAQ M3Dry with the National Land Cover Dataset (NLCD) (Dewitz and U.S. Geological Survey, 2021; Yang et al., 2018). Similar comparisons between the U.S. Geological Survey (USGS) and NLCD LU datasets in WRF have previously been presented by Mallard et al. (2018). In contrast to the changes discussed in the previous section, the differences between MODIS and NLCD are due to different underlying remote sensing data, their spatial resolution, and different approaches for their classification into distinct LU categories. Specifically, the MODIS LU categorization scheme as implemented in WRFv4.1.1 uses 20 categories following the IGBP land cover classification scheme (Loveland et al. 1999), has a native underlying spatial resolution of 500m, is based on a 2001 – 2010 climatology of MODIS satellite data, and features no geophysical boundaries in LU classification categories between the U.S., Canada, and Mexico. In contrast, the NLCD LU characterization scheme as implemented in WRFv4.1.1 uses 40 categories and is hereafter referred to as NLCD40. The first 20 of these categories mirror those of the MODIS scheme and are used for areas outside of the U.S. where NLCD data is not available. The remaining categories are used over the U.S. and follow a modified Anderson land cover classification scheme (Homer et al., 2015, Yang et al., 2021), are based on a native underlying spatial resolution of 30m, and are derived from NLCD satellite data for the year 2011 (Homer et al., 2015). The combination of the NLCD and MODIS satellite data for use in WRF across North America is described by Ran et al. (2010).

Table S1 contains a listing of the 20 MODIS and 40 NLCD40 LU categories used in WRFv4.1.1 and CMAQ M3Dry and their mapping to the aggregated 16 AQMEII4 LU categories (Galmarini et al., 2021) used for analysis purposes. As noted in Section 2.3.1, the CMAQ M3Dry calculations and post-processor estimates of LU-specific and aggregated diagnostic variables were performed using native LU categories from the MODIS and NLCD40 schemes. Aggregation to the 16 category AQMEII4 LU scheme was then performed through LU-weighted averaging of equivalent categories using the mappings from Table S1. None of the MODIS or NLCD40 LU categories correspond to the AQMEII4 herbaceous category. Figure 14 shows the domain-wide fractional coverage of the AQMEII4 LU categories aggregated from the MODIS and NLCD40 categories while Figure S18 shows maps of differences for each of these 16 LU categories between the MODIS and NLCD40 configurations. The barchart in Figure 14 shows that MODIS has overall higher fractional coverage than NLCD40 for the evergreen needleleaf forest, mixed forest, planted/cultivated, and grassland categories and lower fractional coverage for the urban, shrubland, savanna, and wetlands categories. However, the maps in Figure S18 show a more nuanced picture, with both positive and negative

differences in the forest categories. For example, MODIS generally has higher evergreen needleleaf forest fractions than NLCD40 over Canada and lower fractions over the U.S. with the exception of the West Coast. The higher planted/cultivated fractions in MODIS are most pronounced over the central and northern U.S. while the higher grassland fractions are most pronounced over the Western U.S. MODIS urban and wetland fractions are lower than NLCD40 at most grid cells although the magnitude of the difference varies. For shrublands and savanna, MODIS fractions are generally lower than NLCD40 fractions but exceptions exist in the Southeastern U.S. and portions of Canada.

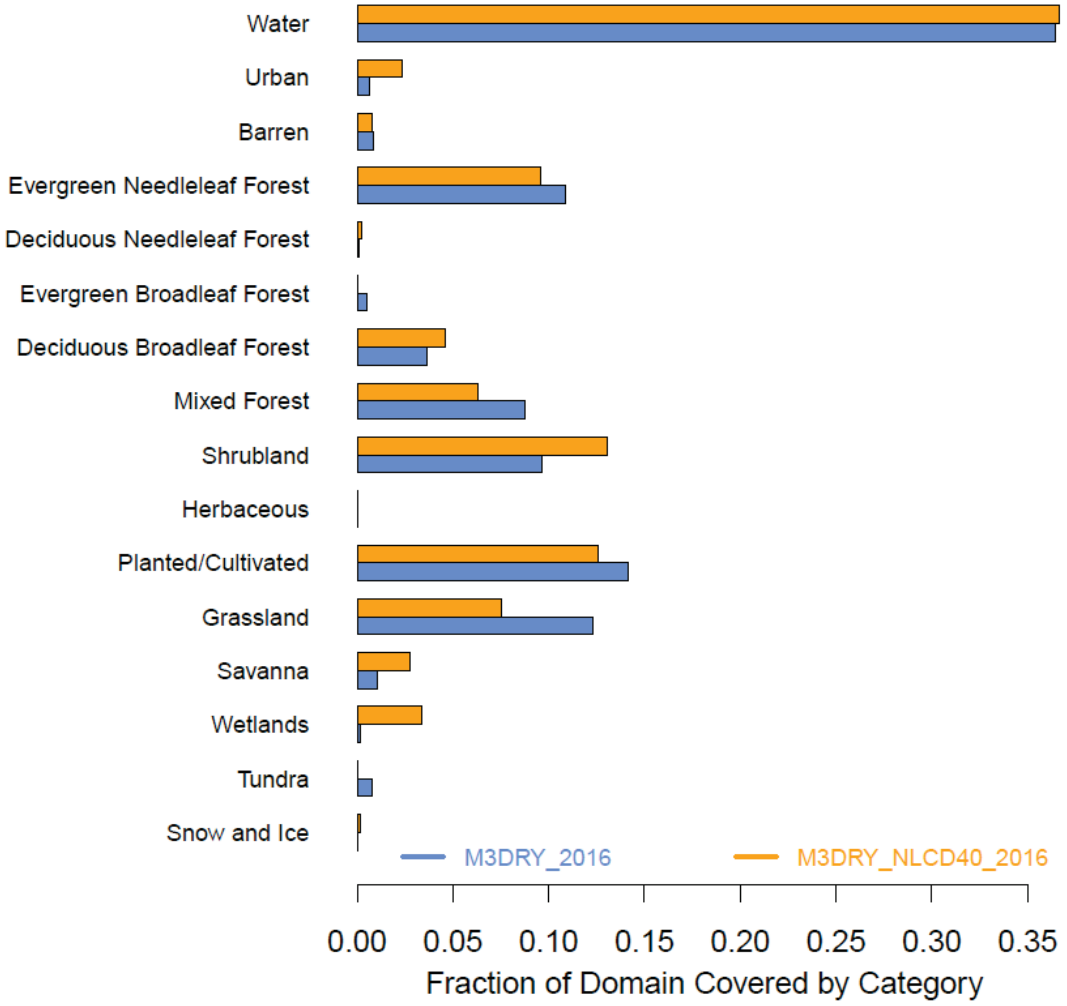



**Figure 14:** Domainwide fractional coverage of the 16 AQMEII4 land-use categories (Galmarini et al., 2021) in the M3DRY_2016 (using MODIS LU in WRF PX LSM) and M3DRY_NLCD40_2016 (using NLCD40 LU in WRF PX LSM) simulations. None of the 20 MODIS land-use categories correspond to the AQMEII4 herbaceous category.

Figure 15 shows differences in seasonal mean $O_3$ $V_d$ and mixing ratios between M3DRY_2016 and M3DRY_NLCD40_2016 in columns 1 and 3 to quantify the effects of using different LU schemes on these variables. Results show that the M3DRY_2016 simulation using MODIS LU tends to have higher $O_3$ $V_d$ (and correspondingly lower mixing ratios) than the M3DRY_NLCD40_2016 simulation using NLCD40 LU for most seasons and regions. The differences in LU cause seasonal mean $O_3$ mixing ratio differences on the order of 1 ppb across large portions of the domain, with the differences generally largest during summer and in areas characterized by the largest differences in the fractional coverages of the forest, planted/cultivated, and grassland LU categories. Columns 2 and 4 of Figure 15 also show corresponding differences in seasonal mean $O_3$ $V_d$ and mixing ratios between M3DRY_2016 and STAGE_2016 (the same data already depicted in Figure 3) to provide a side-by-side contrast of the LU effect on these variables with the dry deposition scheme effect. While the M3Dry vs. STAGE differences generally are larger than the MODIS vs. NLCD40 differences outside the summer season, the magnitude of both effects is comparable during summer.

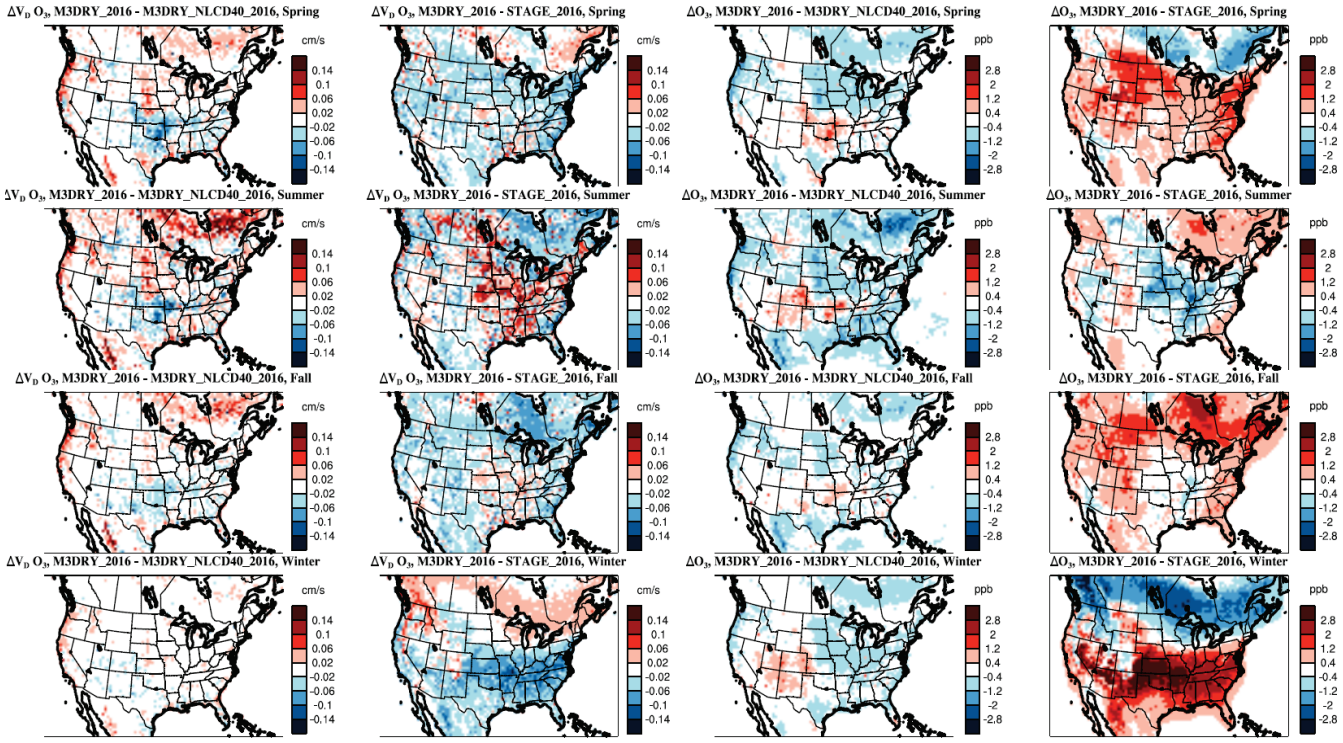



**Figure 15.** Differences in seasonal mean $O_3$ deposition velocities (columns 1 and 2) and mixing ratios (columns 3 and 4) M3DRY_2016 minus M3DRY_NLCD40_2016 (columns 1 and 3) and M3DRY_2016 minus STAGE_2016 (columns 2 and 4)

To assess the impacts of the LU-induced differences in $V_d$ and mixing concentrations on effective fluxes and total dry deposition not only for $O_3$ but also other species, Figure 16 shows annual total domain-wide pathway-specific dry deposition effective fluxes over all non-water grid cells for $O_3$, H2O2, HCHO, SO2, and oxidized nitrogen species for M3DRY_2016 and M3DRY_NLCD40_2016. Note that like in equivalent Figure 10 comparing M3DRY_2016 and STAGE_2016, $O_3$ dry deposition values are divided by a factor 10 to use the same y-axis as for the other pollutants. Overall, the use of MODIS vs. NLCD40 results in only very minor differences in domain-total dry deposition for all species analyzed here. However, the small domain total annual differences of these grid-scale deposition fluxes likely mask larger differences existing in different regions and seasons such as those shown for $O_3$ $V_d$ and mixing ratios in Figure 15 as well as differences existing for LU-specific fluxes examined below. Moreover, there is a tendency for slightly larger effective fluxes through the bare soil pathway and slightly lower effective fluxes through the cuticular and stomatal for the M3DRY_NLCD40_2016 simulation, indicating that the choice of LU datasets can result in different estimates of dry deposition fluxes through specific pathways.

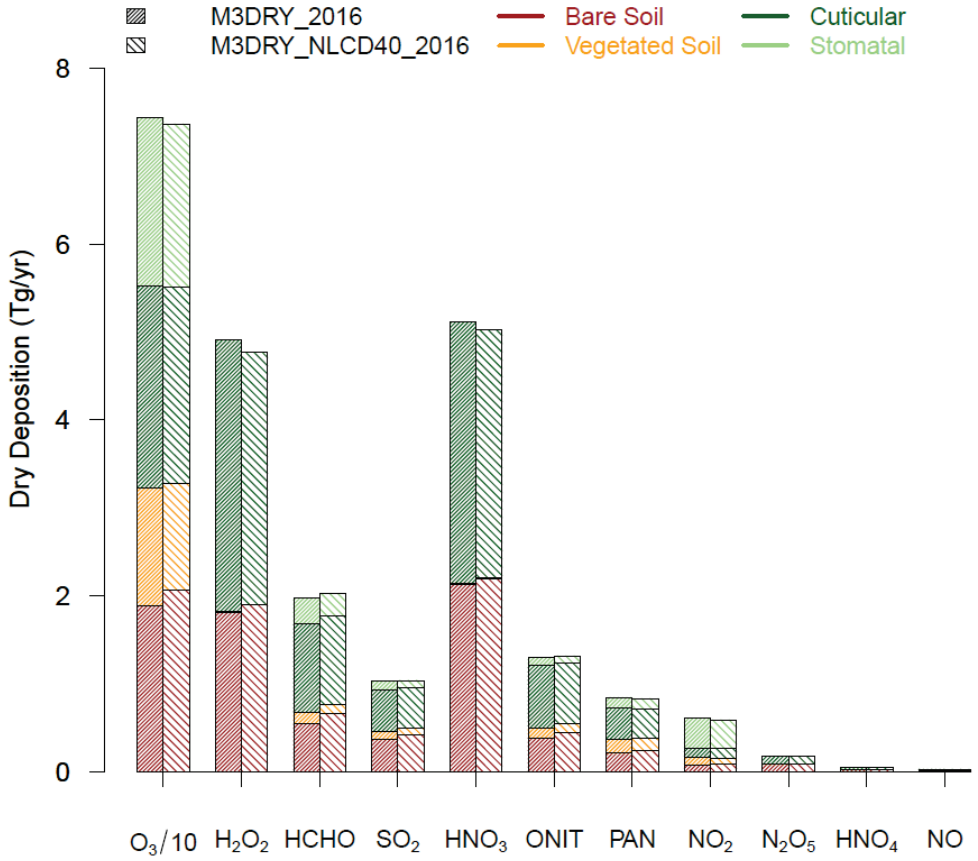



**Figure 16:** Grid-scale annual total domain-wide (excl. water cells) pathway-specific dry deposition fluxes ("effective fluxes") for $O_3$, $H_2O_2$, HCHO, $SO_2$, and oxidized nitrogen species for M3DRY_2016 and M3DRY_NLCD40_2016. $O_3$ dry deposition values are divided by a factor 10 to use the same y-axis as for the other pollutants. (compare to Figure 10)

Figure 17 further disaggregates the grid-scale effective fluxes shown in Figure 16 by LU category using $O_3$ as an example. Consistent with the differences in fractional coverage for the different LU categories between MODIS and NLCD40 shown in Figure 14, domain-total $O_3$ deposition fluxes to the evergreen needleleaf forest, mixed forest, planted/cultivated, and grassland

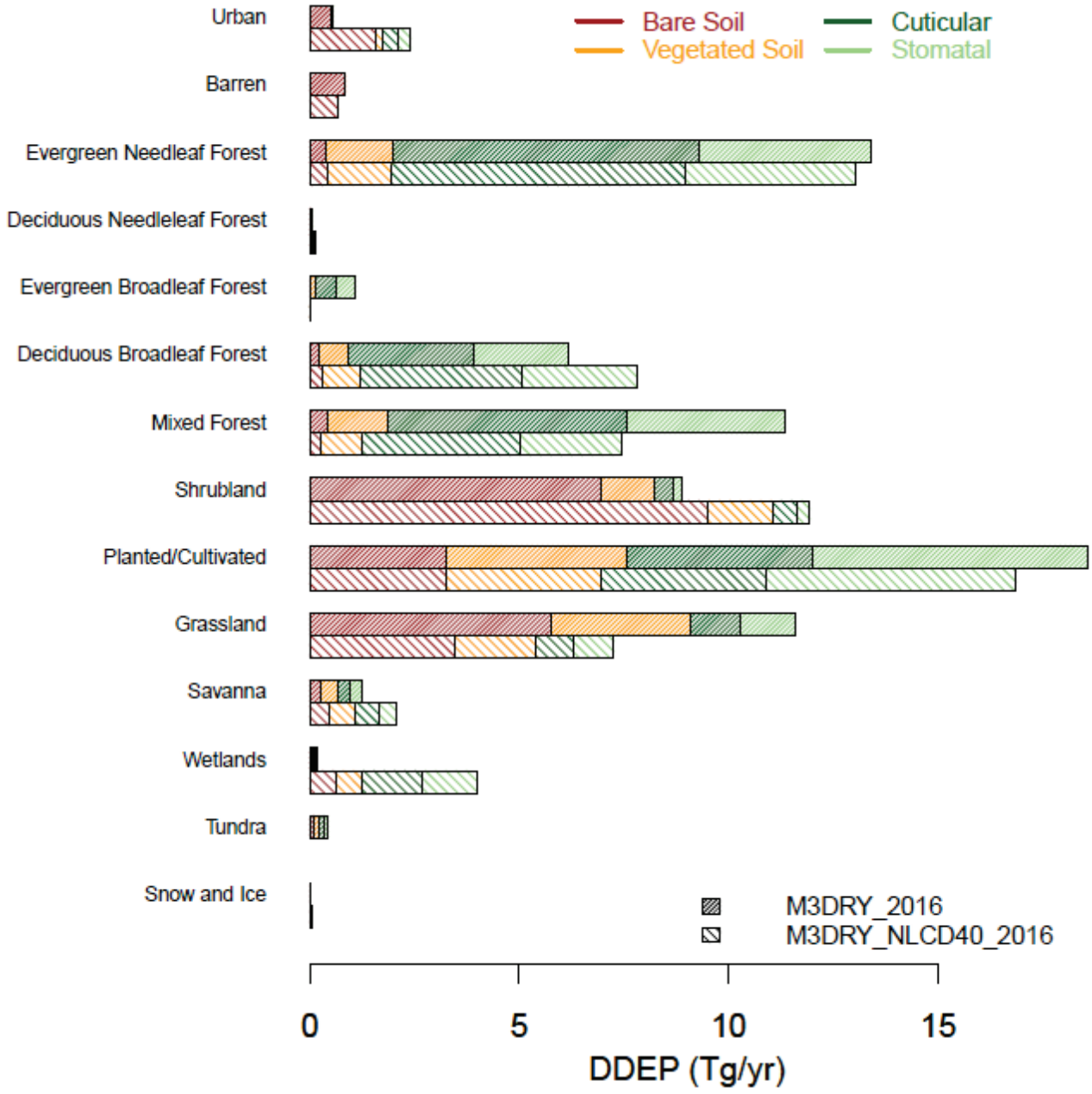



Figure 17. Land-use specific annual total domain-wide (excl. water cells) pathway-specific $O_3$ dry deposition fluxes ("effective fluxes") for M3DRY_2016 and M3DRY_NLCD40_2016. (compare to Figure 12)

categories are higher for M3DRY_2016 than M3DRY_NLCD40_2016 while the opposite is the case for the urban, deciduous broadleaf forest, shrubland, savanna, and wetlands categories. For example, the fraction of the modeling domain classified as urban is 1% for MODIS and 2.4% for NLCD40, resulting in higher deposition estimates for this category in the M3DRY_NLCD40_2016 simulations. Moreover, because of the higher underlying spatial resolution of the NLCD satellite data and the inclusion of lower density developed areas characterized by more vegetation in the urban LU category, this category has higher estimated effective fluxes through the stomatal and cuticular pathways in the M3DRY_NLCD40_2016 simulated compared to the M3DRY_2016 simulations. Overall, this demonstrates that the deposition impacts of LU differences are caused both by differences in the fractional coverages and spatial distributions of different LU categories as well as their characterization in terms of variables like VEGF, $z_0$, etc. in WRF PX LSM look-up tables which are partially tied to the spatial resolution of the underlying satellite datasets.

## 4 Summary and Discussion

The model evaluation results presented in Section 3.1 demonstrate that the AQMEII4 CMAQ simulations perform similarly to other comparable regional-scale modeling studies (Emery et al., 2017; Kelly et al., 2019; Simon et al., 2012; Appel et al., 2021). The analysis of several sensitivity simulations presented in the supplemental material indicates that the choice of lateral boundary conditions was the largest driver of differences in mean model concentrations and biases compared to the corresponding CMAQv5.3.1 simulations analyzed in Appel et al. (2021). Moreover, the results also indicate that while the choice of the M3Dry vs. STAGE dry deposition option can impact CMAQ performance, these impacts tend to be smaller than those caused by choices of model input data sets and particularly boundary conditions to represent the large-scale chemical environment.

The analysis of $O_3$ $V_d$ and mixing ratios showed that during summer, M3Dry has higher $V_d$ and lower mixing ratios than STAGE for much of the eastern U.S. while the reverse is the case over eastern Canada and along the West Coast. In contrast, during winter STAGE has higher $V_d$ and lower mixing ratios than M3Dry over most of the southern half of the modeling domain while the reverse is the case for much of the northern U.S. and southern Canada. The differences in seasonal mean mixing ratios reach 2-3 ppb in a number of locations, indicating that the effects of different dry deposition schemes can be more pronounced locally than in the spatially aggregated model evaluation metrics presented in Section 3.1. Absolute differences in seasonal mean $O_3$ $V_d$ are on the order of 0.05 – 0.1 cm/s for many locations, While these differences tend to be smaller than the range of model differences reported in intercomparison studies performed at flux measurement sites (e.g. Wu et al., 2018 and Clifton et al., 2023) and for global models (Hardacre et al., 2015), their magnitude nevertheless represents a variation of about 10-30% of CMAQ simulated seasonal mean $V_d$ with the highest relative differences generally occuring during winter.

When using grid-scale effective conductances to further analyze the differences in $O_3$ $V_d$, the comparison between M3Dry and STAGE showed generally higher summertime stomatal and wintertime cuticular effective conductances for M3Dry and generally higher soil effective conductances (both vegetated and bare) for STAGE in both summer and winter. On a domain-wide basis, the stomatal effective conductance accounted for about half of the total $O_3$ $V_d$ during daytime hours in summer for both schemes, though regional variations in these contributions exist due to variations in vegetation coverage. Examining grid-scale component resistances for $O_3$ shows that values for summertime $1/R_s$ and wintertime $1/R_{cut}$ differ between M3Dry and STAGE, with the higher values of these inverted resistances in M3Dry causing higher effective conductances for these pathways compared to STAGE. Extending the concept of effective conductances to effective fluxes, annual domain total $O_3$ dry deposition flux results show the larger contributions from the vegetated and bare soil pathways for STAGE and the larger contributions from the stomatal and cuticular pathways for M3Dry.

In contrast to $O_3$, most other pollutants show larger effective conductances and effective fluxes for the bare soil and sum of bare and vegetated soil pathways for M3Dry than STAGE. The cuticular effective flux is larger for M3Dry than STAGE for HCHO, $SO_2$, $HNO_4$, and organic nitrates, smaller for M3Dry than STAGE for $H_2O_2$ and $HNO_3$, and similar between M3Dry and STAGE for other species. Stomatal effective fluxes are small for all species except $O_3$, HCHO, and $SO_2$ for both M3Dry and STAGE. Domain-wide annual total dry deposition fluxes differ the most between M3Dry and STAGE for HCHO and organic nitrate.

Extending the analysis of grid-scale dry deposition diagnostics further to specific LU categories, results show that total $V_d$ during daytime varies by a factor of 2 between the LU categories with the lowest values (urban, barren, and snow and ice) and those with the highest values (evergreen needleleaf and broadleaf forest, deciduous forest). Effective conductances for the bare soil pathway dominate for the urban, barren, and shrubland LU categories while conductances for the cuticular and stomatal pathways dominate for the forest LU categories, especially during daytime. Results also show that the M3Dry vs. STAGE differences are most pronounced for the stomatal and vegetated soil pathway for the forest LU categories, with M3Dry estimating larger effective conductances for the stomatal pathway and STAGE estimating larger effective conductances for the vegetated soil pathway for these LU categories. Higher effective conductances for the bare soil pathway in STAGE are particularly noticeable for the urban, shrubland, and tundra LU categories. For annual total deposition, the overall slightly larger grid-scale $O_3$ deposition flux for STAGE is present for almost all LU categories. Considering effective fluxes for summertime only, the greater importance of the cuticular and stomatal pathways during this season for LU categories most strongly affected by seasonal variations in LAI, along with the greater importance of these pathways in M3Dry compared to STAGE, yield greater overall estimated deposition to these LU categories for M3Dry compared to STAGE. Additional analysis showed that minor differences in LU-specific look-up table values between different deposition schemes as well as the aggregation of the MODIS LU categories to the AQMEII4 LU categories in the STAGE deposition calculations can also contribute to differences in the estimated $V_d$ and resulting mixing ratios. Overall, the analysis of LU-specific diagnostic variables for both the entire year and summer only revealed that even though annual total $O_3$ deposition fluxes estimated by

M3Dry and STAGE are fairly similar, pathway-specific fluxes to individual LU types can vary more substantially on both annual and seasonal scales which is likely to affect estimates of $O_3$ damages to sensitive vegetation.

A comparison of two simulations differing only in their LU classification scheme (MODIS vs. NLCD40) showed that the differences in LU cause seasonal mean $O_3$ mixing ratio differences on the order of 1 ppb across large portions of the domain, with the differences generally largest during summer and in areas characterized by the largest differences in the fractional coverages of the forest, planted/cultivated, and grassland LU categories. These differences are generally smaller than the M3Dry vs. STAGE differences outside the summer season but have a similar magnitude during summer. When considering LU-specific effective fluxes for both simulations, results show that domain-total $O_3$ deposition fluxes to the evergreen needleleaf forest, mixed forest, planted/cultivated, and grassland categories are higher for the simulation using MODIS LU than the simulation using NLCD40 LU while the opposite is the case for the urban, deciduous broadleaf forest, shrubland, savanna, and wetlands categories. Moreover, because of the higher underlying spatial resolution of the NLCD satellite data and the inclusion of lower density developed areas characterized by more vegetation in the urban LU category, this category has higher estimated effective fluxes through the stomatal and cuticular pathways in the M3DRY_NLCD40_2016 simulated compared to the M3DRY_2016 simulations. Results indicate that the deposition impacts of LU differences are caused both by differences in the fractional coverages and spatial distributions of different LU categories as well as their characterization in terms of variables like VEGF, $z_0$, etc. in the look-up tables used in the LSM and deposition scheme.

Refining the representation of dry deposition in regional modeling systems is an area of ongoing research, leveraging new insights from observational datasets, field scale modeling, and model intercomparisons such as those performed in AQMEII4 (Galmarini et al., 2021; Clifton et al., 2023). For example, both the M3Dry and STAGE deposition schemes in CMAQ have updated the representation of aerosol dry deposition in the recent release of CMAQv5.4 in October 2022 (U.S. EPA, 2022; Pleim et al., 2022) while STAGE has also been revised to normalize $R_a$, $R_s$, and $u_*$ to grid-scale values from the LSM used in the driving meteorological model. The analyses and results presented in this study serve as an example of how the diagnostic grid-scale and LU-specific dry deposition variables adopted for AQMEII4 can provide insights into a key model process affecting simulated pollutant budgets and ecosystem impacts of atmospheric pollution. Initial analysis of results from all AQMEII4 grid model simulations show that the differences in simulated $O_3$ $V_d$, deposition pathways, and deposition fluxes between the CMAQ M3Dry and STAGE simulations analyzed in this study tend to be smaller than the differences relative to other AQMEII4 grid models. It is expected that the forthcoming diagnostic analyses of results from all AQMEII4 grid models as well as a potentially more wide-spread adoption of such diagnostic variables in other modeling studies will serve the modeling community in their future development efforts. Additionally, it is hoped that the ongoing point intercomparison of the M3Dry and STAGE schemes implemented in CMAQ with schemes implemented in other models (Clifton et al., 2023), along with planned point model simulations to quantify the sensitivity of specific resistances and conductances in different schemes towards individual meteorological, soil, and biophysical forcing variables can help guide future model development efforts.

## Code availability

The CMAQ version 5.3.1 (https://doi.org/10.5281/zenodo.3585898) code is available from the CMAQ GitHub site (https://github.com/USEPA/CMAQ). The AMET code is available from the AMET GitHub site (https://github.com/USEPA/AMET)

## Data availability

All data used in this work are available upon request from the authors. Please contact the corresponding author to request any data related to this work.

## Author contributions

CH led the development of this manuscript, performed all CMAQ simulations, and conducted most of the analyses presented in Section 3. JOB and JEP developed code to adapt the CMAQ STAGE and M3Dry dry deposition schemes to provide diagnostic variables for AQMEII4 through directly modifying CMAQ (STAGE) or creating a post-processor (M3Dry). RCG performed the WRF simulations used as inputs to all CMAQ simulations. CH, DBS and RM helped conceptualize both the AQMEII4 model intercomparison framework and the CMAQ model simulations and analyses presented in this study. KMF and KWA contributed to the model evaluation analysis presented in Section 3.1. All authors contributed to the writing and editing of the manuscript.

## Competing Interests

The authors declare no competing interests.

## 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.

## Special issue statement

This manuscript is part of the special issue "AQMEII4: a detailed assessment of atmospheric deposition processes from point models to regional-scale models".

## Acknowledgments

We gratefully acknowledge members of the AQMEII4 steering committee (Olivia Clifton, Paul Makar, Christopher Holmes, Lisa Emberson, Johannes Flemming, Sam Silva, Johannes Bieser, Jason Ducker, and Martijn Schaap) and AQMEII4 co-chair Stefano Galmarini for facilitating the analysis described in this manuscript by designing and coordinating regional-scale air quality model simulations that provide diagnostic insights into modeled dry deposition.

## References

Alapaty, K., Cheng, B., Bash, J., Munger, J. W., Walker, J. T., and Arunachalam, S.: Dry deposition methods based on turbulence kinetic energy: 1. Evaluation of various resistances and sensitivity studies using a single-point model. Journal of Geophysical Research: Atmospheres, 127, e2022JD036631. https://doi.org/10.1029/2022JD036631, 2022

Appel, K. W., Gilliam, R. C., Davis, N., Zubrow, A., and Howard, S. C.: Overview of the Atmospheric Model Evaluation Tool (AMET) v1.1 for evaluating meteorological and air quality models, Environ. Modell. Softw., 26, 4, 434-443, doi.org/10.1016/j.envsoft.2010.09.007, 2011.

Appel, K.W., Chemel, C., Roselle, S.J., Francis, X.V., Hu R.-M., Sokhi, R.S., Rao, S.T., and Galmarini, S. Examination of the Community Multiscale Air Quality (CMAQ) Model Performance over the North American and European Domains. Atmos. Environ., 53(June):142-155, 2012

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

Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., Li, Q., Liu, H. Y., Mickley, L. J., and Schultz, M. G.: Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, J. Geophys. Res., 106(D19), 23073– 23095, doi:10.1029/2001JD000807, 2001

Byun, D. W. and Schere, K. L.: Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling system, Appl. Mech. Rev., 59, 51–77, 2006.

Chang, J. S., Brost, R. A., Isaksen, I. S. A., Madronich, S., Middleton, P., Stockwell,W. R., and Walcek, C. J.: A three-dimensional Eulerian acid deposition model: Physical concepts and formulation, J. Geophys. Res., 92, 14681–14700, 1987.

Cheng, B., Alapaty, K., Shu, Q., & Arunachalam, S. Dry deposition methods based on turbulence kinetic energy: 2. Extension to particle deposition using a single-point model. Journal of Geophysical Research: Atmospheres, 127, e2022JD037803. https://doi.org/10.1029/2022JD037803, 2022.

Clifton, O. E., Fiore, A. M., Massman, W. J., Baublitz, C. B., Coyle, M., Emberson, L., et al.: Dry deposition of ozone over land: processes, measurement, and modeling, Rev. Geophys., 58, https://doi.org/10.1029/2019RG000670, 2020a.

Clifton, O. E., Paulot, F., Fiore, A. M., Horowitz, L. W., Correa, G., 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., & Weng, E.: Influence of dynamic ozone dry deposition on ozone pollution. Journal of Geophysical Research: Atmospheres. https://doi.org/10.1029/2020JD032398, 2020b

Clifton, O. E., Schwede, D., Hogrefe, C., Bash, J. O., Bland, S., Cheung, P., Coyle, M., Emberson, L., Flemming, J., Fredj, E., Galmarini, S., Ganzeveld, L., Gazetas, O., Goded, I., Holmes, C. D., Horváth, L., Huijnen, V., Li, Q., Makar, P. A., Mammarella, I., Manca, G., Munger, J. W., Pérez-Camanyo, J. L., Pleim, J., Ran, L., San Jose, R., Silva, S. J., Staebler, R., Sun, S., Tai, A. P. K., Tas, E., Vesala, T., Weidinger, T., Wu, Z., and Zhang, L.: A single-point modeling approach for the intercomparison and evaluation of ozone dry deposition across chemical transport models (Activity 2 of AQMEII4), EGUsphere [preprint], https://doi.org/10.5194/egusphere-2023-465, 2023.

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

Emerson, E. W., Hodshire, A. L., DeBolt, H. M., Bilsback, K. R., Pierce, J. R., McMeeking, G. R., & Farmer, D. K.: Revisiting particle dry deposition and its role in radiative effect estimates. Proceedings of the National Academy of Sciences of the United States of America, 117(42), 26076– 26082. https://doi.org/10.1073/pnas.2014761117, 2020

Emery, C., L. Zhen, A., Russell, G., Odman, M. T., Yarwood, G., and Kumar, N.: Recommendations on statistics and benchmarks to assess photochemical model performance, J. Air Waste Manage., 67, 582–598, https://doi.org/10.1080/10962247.2016.1265027, 2017.

Foley, K.M., Pouliot, G.A., Eyth, A., Aldridge, M.F., Allen, C., Appel, K.W., Bash, J.O., Beardsley, M., Beidler, J., Choi, D., Farkas, C., Gilliam, R.C., Godfrey, J., Henderson, B.H., Hogrefe. C., Koplitz, S.N., Mason, R., Mathur, R., Misenis, C., Possiel, N., Pye, H.O.T., Reynolds, L., Roark, M., Roberts, S., Schwede, D.B., Seltzer, K.M., Sonntag, D., Talgo, K., Toro,

C., Vukovich, J., and Xing, J., 2022. 2002-2017 Anthropogenic Emissions Data for Air Quality Modeling over the United States, Data in Brief, 47, 109022, https://doi.org/10.1016/j.dib.2023.109022, 2023

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

Gilliam, R. C., Herwehe, J. A., Bullock, O. R., Pleim, J. E., Ran, L., Campbell, P. C., and Foroutan, H.: Establishing the suitability of the model for prediction across scales for global retrospective air quality modeling. Journal of Geophysical Research: Atmospheres, 126, e2020JD033588. https://doi.org/10.1029/2020JD033588, 2021

Gaydos, T., Pinder, R., Koo, B., Fahey, K.,Yarwood, G., and Pandis, S.N.: Development and application of a three-dimensional Chemical Transport Model, PMCAMx, Atmos. Environ., 41, 2594– 2611, 2007

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W., and Eder, B.: Fully-coupled online chemistry within the WRF model. Atmos. Environ., 39, 6957-6975. doi:10.1016/j.atmosenv.2005.04.027, 2005

Harley, R. A., Russell, A. G., McRae, G. J., Cass, G., and Seinfeld, J. H.: Photochemical modeling of the Southern California air quality study, Environ. Sci. Technol., 27, 378–388, doi:10.1021/es00039a019, 1993.

Hass, H., Ebel, A., Feldmann, H., Jakobs, H.J., and Memmesheimer, M.: Evaluation studies with a regional chemical transport model (EURAD) using air quality data from the EMEP monitoring network, Atmos. Environ., 27, 867-887, https://doi.org/10.1016/0960-1686(93)90007-L, 1993

Heath, N., Pleim, J., Gilliam, R. and Kang, D.: A simple lightning assimilation technique for improving retrospective WRF simulations. J. Adv. Model Earth Syst., 8, doi:10.1002/2016MS000735, 2016

Hogrefe, C., Pouliot, G., Wong, D., Torian, A., Roselle, S.J., Pleim, J., and Mathur, R.: Annual application and evaluation of the online coupled WRF–CMAQ system over North America under AQMEII phase 2, Atmos. Environ., 115, 683-694, https://doi.org/10.1016/j.atmosenv.2014.12.034, 2015

Hogrefe, C., Liu, P., Pouliot, G., Mathur, R., Roselle, S., Flemming, J., Lin, M., and Park, R. J.: Impacts of different characterizations of large-scale background on simulated regional-scale ozone over the continental United States, Atmos. Chem. Phys., 18, 3839–3864, https://doi.org/10.5194/acp-18-3839-2018, 2018

Homer, C., Dewitz, J., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N., Wickham, J., and Megown, K.: Completion of the 2011 National Land Cover Database for the Conterminous United States – Representing a Decade of Land

Cover Change Information. PHOTOGRAMMETRIC ENGINEERING AND REMOTE SENSING. American Society for Photogrammetry and Remote Sensing, Bethesda, MD, 81(0):345-354, 2015

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

Im, U., Bianconi, R., Solazzo, E., Kioutsioukis, I., Badia, A., Balzarini, A., Baro, R., Belassio, R., Brunner, D., Chemel, C., Curci, G., Flemming, J., Forkel, R., Giordano, L., Jimenez-Guerrero, P., Hirtl, M., Hodzic, A., Honzak, L., Jorba, O., Knote, C., Kuenen, J. J. P., Makar, P. A., Manders-Groot, A., Neal, L., Perez, J. L., Piravano, G., Pouliot, G., San Jose, R., Savage, N., Schroder, W., Sokhi, R. S., Syrakov, D., Torian, A., Werhahn, K., Wolke, R., Yahya, K., Zabkar, R., Zhang, Y., Zhang, J., Hogrefe, C., and Galmarini, S.: Evaluation of operational online-coupled regional air quality models over Europe and North America in the context of AQMEII phase 2. Part I: Ozone, Atmos. Environ., 115, 404–420, 2015a.

Im, U., Bianconi, R., Solazzo, E., Kioutsioukis, I., Badia, A., Balzarini, A., Baro, R., Bellasio, R., Brunner, D., Chemel, C., Curci, G., Denier van der Gon, H. A. C., Flemming, J., Forkel, R., Giordano, L., Jimenez-Guerrero, P., Hirtl, M., Hodzic, A., Honzak, L., Jorba, O., Knote, C., Makar, P. A., Manders-Groot, A., Neal, L., Perez, J. L., Pirovano, G., Pouliot, G., San Jose, R., Savage, N., Schroder, W., Sokhi, R.S., Syrakov, D., Torian, A., Tuccella, P., Werhahn, K., Wolke, R., Yahya, K., Zabkar, R., Zhang, Y., Zhang, J., Hogrefe, C., and Galmarini, S.: Evaluation of operational online-coupled regional air quality models over Europe and North America in the context of AQMEII phase 2. Part II: Particulate Matter, Atmos. Environ., 115, 421–441, 2015b.

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

Irving, P. M. and Smith, E.: Acid deposition: State of science and technology Summary report of the US National Acid Precipitation Assessment Program, US Govt. Printing Office, United States, ISBN 0-16-035925-2, 274 pp., 1991.

Jacobson, M.Z., Lu, R., Turco, R.P., Toon, O.B.: Development and application of a new air pollution modeling system-part I: Gas-phase simulations, Atmos. Environ., 30, 12, 1939-1963, https://doi.org/10.1016/1352-2310(95)00139-5, 1996

Kain, J. S.: The Kain-Fritsch convective parameterization: An update, J. Appl. Meteor., 43, 170-181, doi.org/10.1175/1520-0450(2004)043%3C0170:TKCPAU%3E2.0.CO;2, 2004.

Kasibhatla, P., Chameides, W.L.: Seasonal modeling of regional ozone pollution in the eastern United States. Geophysical Research Letters 27, 1415–1418, 2000

Kelly, J.T, Koplitz, S.N., Baker, K.R., Holder, A.L., Pye, H.O.T., Murphy, B.N., Bash, J.O., Henderson, B.H., Possiel, N.C., Simon, H., Eyth, A.M., Jang, C., Phillips, S., and Timin, B.: Assessing PM2.5 model performance for the conterminous U.S. with comparison to model performance statistics from 2007 - 2015, Atmos. Environ., 214, 116872, https://doi.org/10.1016/j.atmosenv.2019.116872, 2019

Kumar, N., Odman, M.T., Russell, A.G.: Multiscale air quality modeling: application to Southern California. Journal of Geophysical Research 99, 5385–5397, 1994

Loveland, T.R.; Zhu, Z.L.; Ohlen, D.O.; Brown, J.F.; Reed, B.C.; Yang, L.M. An analysis of the IGBP global land-cover characterization process. Photogramm. Eng. Remote Sens. 65, 1021–1032, 1999

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

Mallard, M. S., Spero, T. L., and Taylor, S. M.: Examining WRF's sensitivity to contemporary land use datasets across the contiguous U.S. using dynamical downscaling. Journal of Meteorology and Climatology, 57(11), 2561– 2583. https://doi.org/10.1175/JAMC-D-17-0328.1, 2018

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

Mathur, R., Xing, J., Gilliam, R., Sarwar, G., Hogrefe, C., Pleim, J., Pouliot, G., Roselle, S., Spero, T. L., Wong, D. C., and Young, J.: Extending the Community Multiscale Air Quality (CMAQ) modeling system to hemispheric scales: overview of process considerations and initial applications, Atmos. Chem. Phys., 17, 12449–12474, doi.org/10.5194/acp-17-12449-2017, 2017.

McRae, G. J. and Seinfeld, J. H.: Development of a second-generation mathematical model for urban air pollution – II. Evaluation of model performance, Atmos. Environ., 17, 501–522, doi:10.1016/0004-6981(83)90124-5, 1983.

Morrison, H., Curry, J. A., and Khvorostyanov, V. I.: A new double-moment microphysics parameterization for application in cloud and climate models. Part I: Description, J. Atmos. Sci., 62, 1665–1677, doi.org/10.1175/JAS3446.1, 2005.

Nemitz, E., Milford, C., and Sutton, M. A.: A two-layer canopy compensation point model for describing bi-directional biosphere-atmosphere exchange of ammonia, Q. J. Roy. Meteor. Soc., 127, 815–833, doi.org/10.1002/qj.49712757306, 2001.

Paulot, F., Malyshev, S., Nguyen, T., Crounse, J. D., Shevliakova, E., and Horowitz, L. W.: Representing sub-grid scale variations in nitrogen deposition associated with land use in a global Earth system model: implications for 7present and future nitrogen deposition fluxes over North America. Atmospheric Chemistry and Physics, 18(24), 17963-17978, 2018

Pleim, J., Venkatram, A., Yamartino, R.: ADOM/TADAP Model Development Program: The Dry Deposition Module; Ontario Ministry of the Environment, 4, 1984.

Pleim, J. E., and Xiu, A.: Development and testing of a surface flux and planetary boundary layer model for application in mesoscale models. J. Appl. Meteor., 34 , 16–32, 1995

Pleim, J. E., and Xiu, A.: Development of a land surface model. Part II: Data assimilation. J. Appl. Meteor., 42 , 1811–1822, 2003.

Pleim, J. E.: A simple, efficient solution of flux-profile relationships in the atmospheric surface layer, J. Appl. Meteor. Climatol., 45, 341 – 347, doi.org/10.1175/JAM2339.1, 2006.

Pleim, J. E.: A combined local and nonlocal closure model for the atmospheric boundary layer. Part I: model description and testing, J. Appl. Meteor. Clim., 46, 1383-1395, doi.org/10.1175/JAM2539.1, 2007a.

Pleim, J. E.: A combined local and nonlocal closure model for the atmospheric boundary layer. Part II: application and evaluation in a mesoscale meteorological model, J. Appl. Meteor. Clim., 46, 1396–1409, doi.org/10.1175/JAM2534.1, 2007b.

Pleim J. E. and Gilliam, R. C.: An indirect data assimilation scheme for deep soil temperature in the Pleim-Xiu land surface model, J. Appl. Meteor. Clim., 48, 1362-1376, doi.org/10.1175/2009JAMC2053.1, 2009.

Pleim J.E. and Ran L.: Surface Flux Modeling for Air Quality Applications. Atmosphere. 2(3):271-302. https://doi.org/10.3390/atmos2030271, 2011

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

Price, C. G., Penner, J. E., and Prather, M. J.: NOx from lightning, Part I: Global distribution based on lightning physics, J. Geophys. Res., 102, 5229–5241, 1997.

Pye, H. O. T., Murphy, B. N., Xu, L., Ng, N. L., Carlton, A. G., Guo, H., Weber, R., Vasilakos, P., Appel, K. W., Budisulistiorini, S. H., Surratt, J. D., Nenes, A., Hu, W., Jimenez, J. L., Isaacman-VanWertz, G., Misztal, P. K., and Goldstein, A. H.: On the implications of aerosol liquid water and phase separation for organic aerosol mass, Atmos. Chem. Phys., 17, 343–369, https://doi.org/10.5194/acp-17-343-2017, 2017.

Pye, H. O. T., D'Ambro, E. L., Lee, B. H., Schobesberger, S., Takeuchi, M., Zhao, Y., Lopez-Hilfiker, F., Liu, J., Shilling, J. E., Xing, J., Mathur, R., Middlebrook, A. M., Liao, J., Welti, A., Graus, M., Warneke, C., de Gouw, J. A., Holloway, J. S., Ryerson, T. B., Pollack, I. B., and Thornton, J. A.: Anthropogenic enhancements to production of highly oxygenated molecules from autoxidation, P. Natl. Acad. Sci. USA, 116, 6641–6646, https://doi.org/10.1073/pnas.1810774116, 2019.

Qin, M., Murphy, B. N., Isaacs, K. K., McDonald, B. C., Lu, Q., McKeen, S. A., Koval, L., Robinson, A. L., Efstathious, C., Allen, C., and Pye, H. O. T.: Criteria pollutant impacts of volatile chemical products informed by near-field modeling, Nat Sustain, 4, 129–137, https://doi.org/10.1038/s41893-020-00614-1, 2021

Ran, L., Pleim, J., and Gilliam, R.: Impact of high resolution land-use data in meteorology and air quality modeling systems, in Air Pollution Modeling and its Applications XX, edited by D. G. Steyn and S. T. Rao, chap. 1, C (section 1.1): 3– 7, Springer, Netherlands, 2010

Ran, L., Pleim, J., Gilliam, R., Binkowski, F. S., Hogrefe, C., and Band, L.: Improved meteorology from an updated WRF/CMAQ modeling system with MODIS vegetation and albedo, J. Geophys. Res. Atmos., 121, 2393– 2415, doi:10.1002/2015JD024406, 2016

Ran, L., Yuan, Y., Cooter, E., Benson, V., Yang, D., Pleim, J., Wang, R. and Williams, J.: An integrated agriculture, atmosphere, and hydrology modeling system for ecosystem assessments. Journal of Advances in Modeling Earth Systems, 11 (12), 4645-4668, doi.org/10.1029/2019MS001708, 2019

Rao, S. T., Galmarini, S., and Puckett, K.: Air quality model evaluation international initiative (AQMEII), B. Am. Meteorol. Soc., 92, 23–30, https://doi.org/10.1175/2010BAMS3069.1, 2011.

Russell, A.G., McCue, K.F., and Cass, G.R.: Mathematical modeling of the formation of nitrogen-containing air pollutants-I. Evaluation of an Eulerian photochemical model, Environ. Sci. Technol. 22:263, 1988

Scheffe, R.D. and Morris, R.E.: A Review of the Development and Application of the Urban Airshed Model; Atmos. Environ. 27B, 23-39, 1993

Saylor, R. D., Baker, B. D., Lee, P., Tong, D., Pan, L., and Hicks, B. B.: The particle dry deposition component of total deposition from air quality models: Right, wrong or uncertain? Tellus B: Chemical and Physical Meteorology, 71(1), 1550324. https://doi.org/10.1080/16000889.2018.1550324, 2019

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

Sistla, G., Hao, W., Ku, J.-Y., Kallos, G., Zhang, K., Mao, H., and Rao, S. T.: An operational evaluation of two regional-scale ozone air quality modeling systems over the eastern United States, B. Am. Meteor. Soc., 82, 945–964, 2001.

Solazzo, E., Bianconi, R., Vautard, R., Appel, K. W., Moran, M. D., Hogrefe, C., Bessagnet, B., Brandt, J., Christensen, J. H., Chemel, C., Coll, I., van der Gon, H. D., Ferreira, J., Forkel, R., Francis, X. V., Grell, G., Grossi, P., Hansen, A. B., Jericevic, A., Kraljevic, L., Miranda, A. I., Nopmongcol, U., Pirovano, G., Prank, M., Riccio, A., Sartelet, K. N., Schaap, M., Silver, J. D., Sokhi, R. S., Vira, J., Werhahn, J., Wolke, R., Yarwood, G., Zhang, J., Rao, S. T., and Galmarini, S.: Model evaluation and ensemble modelling and for surface-level ozone in Europe and North America, Atmos. Environ., 53, 60–74, 2012a.

Solazzo, E., Bianconi, R., Pirovano, G., Matthias, V., Vautard, R., Moran, M. D., Appel, K. W., Bessagnet, B., Brandt, J., Christensen, J. H., Chemel, C., Coll, I., Ferreira, J., Forkel, R., Francis, X. V., Grell, G., Grossi, P., Hansen, A. B., Hogrefe, C., Miranda, A. I., Nopmongcol, U., Prank, M., Sartelet, K. N., Schaap, M., Silver, J. D., Sokhi, R. S., Vira, J., Werhahn, J.,

Wolke, R., Yarwood, G., Zhang, J., Rao, S. T., and Galmarini, S.: Operational model evaluation for particulate matter in Europe and North America, Atmos. Environ., 53, 75–92, 2012b.

Solazzo, E., Hogrefe, C., Colette, A., Garcia-Vivanco, M., and Galmarini, S.: Advanced error diagnostics of the CMAQ and Chimere modelling systems within the AQMEII3 model evaluation framework, Atmos. Chem. Phys., 17, 10435–10465, https://doi.org/10.5194/acp-17-10435-2017, 2017

U.S. Environmental Protection Agency: CMAQ (Version 5.3.1) [Software]. Available from doi:10.5281/zenodo.3585898, 2019

U.S. Environmental Protection Agency: CMAQ (Version 5.4) [Software]. Available from doi: 10.5281/zenodo.7218076, 2022

Wesely, M.L, and Hicks, B.B.: Some Factors that Affect the Deposition Rates of Sulfur Dioxide and Similar Gases on Vegetation, Journal of the Air Pollution Control Association, 27:11, 1110-1116, DOI: 10.1080/00022470.1977.10470534, 1977

Walker, J. T., Chen, X., Wu, Z., Schwede, D., Daly, R., Djurkovic, A., Oishi, A. C., Edgerton, E., Bash, J., Knoepp, J., Puchalski, M., Iiames, J., and Miniat, C. F.: Atmospheric deposition of reactive nitrogen to a deciduous forest in the southern Appalachian Mountains, Biogeosciences, 20, 971–995, https://doi.org/10.5194/bg-20-971-2023, 2023.

Wesely, M.L.: Parameterization of surface resistances to gaseous dry deposition in regional-scale numerical models, Atmospheric Environment, 23, 1293-1304, 1989.

Williams J. R.: The EPIC model, Computer models in watershed hydrology, Water Resources Publications, Highlands Ranch, 909–1000, 1995.

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 Syst., 10, 1571–1586, https://doi.org/10.1029/2017MS001231, 2018.

Xiu, A. and Pleim, J.E.: Development of a land surface Model. Part I: Application in a mesoscale meteorological model. J. Appl. Meteor., 40 , 192–209, 2001.

Yang, L., Jin, S., Danielson, P., Homer, C., Gass, L., Case, A., Costello, C., Dewitz, J., Fry, J., Funk, M., Grannemann, B., Rigge, M. and G. Xian: A New Generation of the United States National Land Cover Database: Requirements, Research Priorities, Design, and Implementation Strategies, ISPRS Journal of Photogrammetry and Remote Sensing, 146, pp.108-123, 2018

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